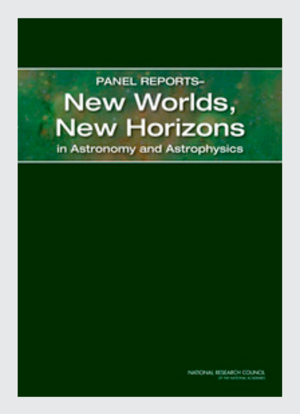


Panel Reports New Worlds, New Horizons in Astronomy and Astrophysics (2011)

DETAILS

578 pages | 7 x 10 | PAPERBACK
ISBN 978-0-309-15962-3 | DOI 10.17226/12982

CONTRIBUTORS

Science Frontiers Panels; Program Prioritization Panels; Committee for a Decadal Survey of Astronomy and Astrophysics; Board on Physics and Astronomy; Space Studies Board; Division on Engineering and Physical Sciences; National Research Council

SUGGESTED CITATION

National Research Council 2011. *Panel Reports New Worlds, New Horizons in Astronomy and Astrophysics*. Washington, DC: The National Academies Press. https://doi.org/10.17226/12982.

PANEL REPORTS—
New Worlds, New Horizons
in Astronomy and Astrophysics

Science Frontiers Panels
Program Prioritization Panels

Committee for a Decadal Survey of Astronomy and Astrophysics

Board on Physics and Astronomy
Space Studies Board

Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the panels responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by Contract NNX08AN97G between the National Academy of Sciences and the National Aeronautics and Space Administration, Contract AST-0743899 between the National Academy of Sciences and the National Science Foundation, and Contract DE-FG02-08ER41542 between the National Academy of Sciences and the U.S. Department of Energy. Support for this study was also provided by the Vesto Slipher Fund. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the organizations that provided support for the project.

International Standard Book Number-13: 978-0-309-15962-3
International Standard Book Number-10: 0-309-15962-8
Library of Congresss Control Number: 2011925026

Available in limited quantities from:

Board on Physics and Astronomy
500 Fifth Street, N.W.
Washington, DC 20001
bpa@nas.edu
http://www.nationalacademies.edu/bpa

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.

Printed in the United States of America

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

Copyright National Academy of Sciences. All rights reserved.

SCIENCE FRONTIERS PANELS

Panel on Cosmology and Fundamental Physics

DAVID N. SPERGEL, Princeton University, *Chair*
DAVID WEINBERG, Ohio State University, *Vice Chair*
RACHEL BEAN, Cornell University
NEIL CORNISH, Montana State University
JONATHAN FENG, University of California, Irvine
ALEX V. FILIPPENKO, University of California, Berkeley
WICK C. HAXTON, University of California, Berkeley
MARC P. KAMIONKOWSKI, California Institute of Technology
LISA RANDALL, Harvard University
EUN-SUK SEO, University of Maryland
DAVID TYTLER, University of California, San Diego
CLIFFORD M. WILL, Washington University

Panel on the Galactic Neighborhood

MICHAEL J. SHULL, University of Colorado, *Chair*
JULIANNE DALCANTON, University of Washington, *Vice Chair*
LEO BLITZ, University of California, Berkeley
BRUCE T. DRAINE, Princeton University
ROBERT FESEN, Dartmouth University
KARL GEBHARDT, University of Texas
JUNA KOLLMEIER, Observatories of the Carnegie Institution of Washington
CRYSTAL MARTIN, University of California, Santa Barbara
JASON TUMLINSON, Space Telescope Science Institute
DANIEL WANG, University of Massachusetts
DENNIS ZARITSKY, University of Arizona
STEPHEN E. ZEPF, Michigan State University

Panel on Galaxies Across Cosmic Time

C. MEGAN URRY, Yale University, *Chair*
MITCHELL C. BEGELMAN, University of Colorado, *Vice Chair*
NETA A. BAHCALL, Princeton University
ANDREW J. BAKER, Rutgers University
ROMEEL DAVÉ, University of Arizona
TIZIANA DI MATTEO, Carnegie Mellon University
HENRIC S.W. KRAWCZYNSKI, Washington University

JOSEPH MOHR, Ludwig Maximilian University
RICHARD F. MUSHOTZKY, NASA Goddard Space Flight Center
CHRISTOPHER S. REYNOLDS, University of Maryland
ALICE SHAPLEY, University of California, Los Angeles
TOMMASO TREU, University of California, Santa Barbara
JAQUELINE H. VAN GORKOM, Columbia University
ERIC M. WILCOTS, University of Wisconsin

Panel on Planetary Systems and Star Formation

LEE W. HARTMANN, University of Michigan, *Chair*
DAN M. WATSON, University of Rochester, *Vice Chair*
HECTOR ARCE, Yale University
CLAIRE CHANDLER, National Radio Astronomy Observatory
DAVID CHARBONNEAU, Harvard University
EUGENE CHIANG, University of California, Berkeley
SUZAN EDWARDS, Smith College
ERIC HERBST, Ohio State University
DAVID C. JEWITT, University of California, Los Angeles
JAMES P. LLOYD, Cornell University
EVE C. OSTRIKER, University of Maryland
DAVID J. STEVENSON, California Institute of Technology
JONATHAN C. TAN, University of Florida

Panel on Stars and Stellar Evolution

ROGER A. CHEVALIER, University of Virginia, *Chair*
ROBERT P. KIRSHNER, Harvard-Smithsonian Center for Astrophysics, *Vice Chair*
DEEPTO CHAKRABARTY, Massachusetts Institute of Technology
SUZANNE HAWLEY, University of Washington
JEFFREY R. KUHN, University of Hawaii
STANLEY OWOCKI, University of Delaware
MARC PINSONNEAULT, Ohio State University
ELIOT QUATAERT, University of California, Berkeley
SCOTT RANSOM, National Radio Astronomy Observatory
HENDRIK SCHATZ, Michigan State University
LEE ANNE WILLSON, Iowa State University
STANFORD E. WOOSLEY, University of California, Santa Cruz

Staff

DONALD C. SHAPERO, Director, Board on Physics and Astronomy (BPA)
MICHAEL H. MOLONEY, Astro2010 Study Director and Director, Space Studies Board (SSB)
BRANT L. SPONBERG, Associate Director, SSB (until December 2009)
ROBERT L. RIEMER, Senior Program Officer, BPA
DAVID B. LANG, Program Officer, BPA
CARMELA CHAMBERLAIN, Administrative Coordinator, SSB
CATHERINE A. GRUBER, Editor, SSB
CARYN J. KNUTSEN, Research Associate, BPA
LaVITA COATES-FOGLE, Senior Program Assistant, BPA (until October 2009)
BETH DOLAN, Financial Associate, BPA

PROGRAM PRIORITIZATION PANELS

Panel on Electromagnetic Observations from Space

ALAN DRESSLER, Observatories of the Carnegie Institution of Washington, *Chair*
MICHAEL BAY, Bay Engineering Innovations
ALAN P. BOSS, Carnegie Institution of Washington
MARK DEVLIN, University of Pennsylvania
MEGAN DONAHUE, Michigan State University
BRENNA FLAUGHER, Fermi National Accelerator Laboratory
TOM GREENE, NASA Ames Research Center
PURAGRA (RAJA) GUHATHAKURTA, University of California Observatories/Lick Observatory
MICHAEL G. HAUSER, Space Telescope Science Institute
HAROLD McALISTER, Georgia State University
PETER F. MICHELSON, Stanford University
BEN R. OPPENHEIMER, American Museum of Natural History
FRITS PAERELS, Columbia University
GEORGE H. RIEKE, Steward Observatory, University of Arizona
ADAM G. RIESS, Johns Hopkins University
PAUL L. SCHECHTER, Massachusetts Institute of Technology
TODD TRIPP, University of Massachusetts at Amherst

Panel on Optical and Infrared Astronomy from the Ground

PATRICK S. OSMER, Ohio State University, *Chair*
MICHAEL SKRUTSKIE, University of Virginia, *Vice Chair*
CHARLES BAILYN, Yale University
BETSY BARTON, University of California, Irvine
TODD A. BOROSON, National Optical Astronomy Observatory
DANIEL EISENSTEIN, University of Arizona
ANDREA M. GHEZ, University of California, Los Angeles
J. TODD HOEKSEMA, Stanford University
ROBERT P. KIRSHNER, Harvard-Smithsonian Center for Astrophysics
BRUCE MACINTOSH, Lawrence Livermore National Laboratory
PIERO MADAU, University of California, Santa Cruz
JOHN MONNIER, University of Michigan
I. NEILL REID, Space Telescope Science Institute
CHARLES E. WOODWARD, University of Minnesota

Panel on Particle Astrophysics and Gravitation

JACQUELINE N. HEWITT, Massachusetts Institute of Technology, *Chair*
ERIC G. ADELBERGER, University of Washington
ANDREAS ALBRECHT, University of California, Davis
ELENA APRILE, Columbia University
JONATHAN ARONS, University of California, Berkeley
BARRY C. BARISH, California Institute of Technology
JOAN CENTRELLA, NASA-Goddard Space Flight Center
DOUGLAS FINKBEINER, Harvard University
KATHRYN FLANAGAN, Space Telescope Science Institute
GABRIELA GONZALEZ, Louisiana State University
JAMES B. HARTLE, University of California, Santa Barbara
STEVEN M. KAHN, Stanford University
N. JEREMY KASDIN, Princeton University
TERESA MONTARULI, University of Wisconsin–Madison
ANGELA V. OLINTO, University of Chicago
RENE A. ONG, University of California, Los Angeles
HELEN R. QUINN, Stanford National Accelerator Laboratory (retired)

Panel on Radio, Millimeter, and Submillimeter Astronomy from the Ground

NEAL J. EVANS II, University of Texas, *Chair*
JAMES M. MORAN, Harvard University, *Vice Chair*
CRYSTAL BROGAN, National Radio Astronomy Observatory
AARON S. EVANS, University of Virginia
SARAH GIBSON, National Center for Atmospheric Research, High Altitude Observatory
JASON GLENN, University of Colorado at Boulder
NICKOLAY Y. GNEDIN, Fermi National Accelerator Laboratory
CORNELIA C. LANG, University of Iowa
MAURA McLAUGHLIN, West Virginia University
MIGUEL MORALES, University of Washington
LYMAN A. PAGE, JR., Princeton University
JEAN L. TURNER, University of California, Los Angeles
DAVID J. WILNER, Smithsonian Astrophysical Observatory

Staff

DONALD C. SHAPERO, Director, Board on Physics and Astronomy (BPA)
MICHAEL H. MOLONEY, Astro2010 Study Director and Director, Space Studies Board (SSB)
BRANT L. SPONBERG, Associate Director, SSB (until December 2009)
ROBERT L. RIEMER, Senior Program Officer, BPA
BRIAN DEWHURST, Program Officer, ASEB (until July 2009)
JAMES LANCASTER, Program Officer, BPA
CATHERINE A. GRUBER, Editor, SSB
CARYN J. KNUTSEN, Research Associate, BPA
CARMELA CHAMBERLAIN, Administrative Coordinator, SSB
LaVITA COATES-FOGLE, Senior Program Assistant, BPA (until October 2009)
BETH DOLAN, Financial Associate, BPA

COMMITTEE FOR A DECADAL SURVEY OF ASTRONOMY AND ASTROPHYSICS

ROGER D. BLANDFORD, Stanford University, *Chair*
MARTHA P. HAYNES, Cornell University, *Vice Chair*
JOHN P. HUCHRA,[1] Harvard University, *Vice Chair*
MARCIA J. RIEKE, University of Arizona, *Vice Chair*
LYNNE HILLENBRAND, California Institute of Technology, *Executive Officer*
STEVEN J. BATTEL, Battel Engineering
LARS BILDSTEN, University of California, Santa Barbara
JOHN E. CARLSTROM, University of Chicago
DEBRA M. ELMEGREEN, Vassar College
JOSHUA FRIEMAN, Fermi National Accelerator Laboratory
FIONA A. HARRISON, California Institute of Technology
TIMOTHY M. HECKMAN, Johns Hopkins University
ROBERT C. KENNICUTT, JR., University of Cambridge
JONATHAN I. LUNINE, University of Rome, Tor Vergata
CLAIRE E. MAX, University of California, Santa Cruz
DAN McCAMMON, University of Wisconsin
STEVEN M. RITZ, University of California, Santa Cruz
JURI TOOMRE, University of Colorado
SCOTT D. TREMAINE, Institute for Advanced Study
MICHAEL S. TURNER, University of Chicago
NEIL deGRASSE TYSON, Hayden Planetarium, American Museum of Natural History
PAUL A. VANDEN BOUT, National Radio Astronomy Observatory
A. THOMAS YOUNG, Lockheed Martin Corporation (retired)

Staff

DONALD C. SHAPERO, Director, Board on Physics and Astronomy (BPA)
MICHAEL H. MOLONEY, Astro2010 Study Director and Director, Space Studies Board (SSB)
BRANT L. SPONBERG, Senior Program Officer, BPA (until December 2009)
ROBERT L. RIEMER, Senior Program Officer, BPA
BRIAN D. DEWHURST, Program Officer, Aeronautics and Space Engineering Board (until July 2009)
JAMES C. LANCASTER, Program Officer, BPA

---

[1] This report is dedicated to John P. Huchra, who served as a vice chair for the Astro2010 decadal survey.

DAVID B. LANG, Program Officer, BPA
TERI THOROWGOOD, Administrative Coordinator, BPA (from November 2009)
CARMELA CHAMBERLAIN, Administrative Coordinator, SSB
CATHERINE A. GRUBER, Editor, SSB
CARYN J. KNUTSEN, Research Associate, BPA
LaVITA COATES-FOGLE, Senior Program Assistant, BPA (until October 2009)
BETH DOLAN, Financial Associate, BPA

BOARD ON PHYSICS AND ASTRONOMY

ADAM S. BURROWS, Princeton University, *Chair*
PHILIP H. BUCKSBAUM, Stanford University, *Vice Chair*
RICCARDO BETTI, University of Rochester
PATRICK L. COLESTOCK, Los Alamos National Laboratory (until June 30, 2010)
JAMES DRAKE, University of Maryland
JAMES EISENSTEIN, California Institute of Technology
DEBRA M. ELMEGREEN, Vassar College
PAUL A. FLEURY, Yale University
ANDREA M. GHEZ, University of California, Los Angeles (until June 30, 2010)
PETER F. GREEN, University of Michigan
LAURA H. GREENE, University of Illinois, Urbana-Champaign
MARTHA P. HAYNES, Cornell University
JOSEPH HEZIR, EOP Group, Inc.
MARC A. KASTNER, Massachusetts Institute of Technology (*Chair* until June 30, 2010)
MARK B. KETCHEN, IBM Thomas J. Watson Research Center
JOSEPH LYKKEN, Fermi National Accelerator Laboratory
PIERRE MEYSTRE, University of Arizona
HOMER A. NEAL, University of Michigan
MONICA OLVERA de la CRUZ, Northwestern University
JOSE N. ONUCHIC, University of California, San Diego
LISA RANDALL, Harvard University
CHARLES V. SHANK, Janelia Farm, Howard Hughes Medical Institute (until June 30, 2010)
MICHAEL S. TURNER, University of Chicago
MICHAEL C.F. WIESCHER, University of Notre Dame

Staff

DONALD C. SHAPERO, Director
MICHAEL H. MOLONEY, Associate Director (until March 30, 2010)

SPACE STUDIES BOARD

CHARLES F. KENNEL, Scripps Institution of Oceanography at the University of California, San Diego, *Chair*
A. THOMAS YOUNG, Lockheed Martin Corporation (retired), *Vice Chair*
DANIEL N. BAKER, University of Colorado, Boulder (until June 30, 2010)
STEVEN J. BATTEL, Battel Engineering
CHARLES L. BENNETT, Johns Hopkins University (until June 30, 2010)
YVONNE C. BRILL, Aerospace Consultant
ELIZABETH R. CANTWELL, Oak Ridge National Laboratory
ANDREW B. CHRISTENSEN, Dixie State College; Aerospace Corporation
ALAN DRESSLER, Observatories of the Carnegie Institution of Washington
JACK D. FELLOWS, University Corporation for Atmospheric Research
HEIDI B. HAMMEL, Space Science Institute
FIONA A. HARRISON, California Institute of Technology
ANTHONY C. JANETOS, Pacific Northwest National Laboratory
JOAN JOHNSON-FREESE, Naval War College
KLAUS KEIL, University of Hawaii (until June 30, 2010)
MOLLY K. MACAULEY, Resources for the Future
BERRIEN MOORE III, Climate Central (until June 30, 2010)
JOHN F. MUSTARD, Brown University
ROBERT T. PAPPALARDO, Jet Propulsion Laboratory, California Institute of Technology
JAMES PAWELCZYK, Pennsylvania State University
SOROOSH SOROOSHIAN, University of California, Irvine
DAVID N. SPERGEL, Princeton University
JOAN VERNIKOS, Thirdage LLC
JOSEPH F. VEVERKA, Cornell University (until June 30, 2010)
WARREN M. WASHINGTON, National Center for Atmospheric Research
CHARLES E. WOODWARD, University of Minnesota
THOMAS H. ZURBUCHEN, University of Michigan
ELLEN G. ZWEIBEL, University of Wisconsin (until June 30, 2010)

Staff

MICHAEL H. MOLONEY, Director (from April 2010)
RICHARD E. ROWBERG, Interim Director (from March 2009 through March 2010)
MARCIA S. SMITH, Director (through February 2009)
BRANT L. SPONBERG, Associate Director (until December 2009)

Preface

This volume contains the reports of the five Astro2010 Science Frontiers Panels (SFPs) and the four Astro2010 Program Prioritization Panels (PPPs). These panels were appointed by the National Research Council (NRC) to assist the Committee for a Decadal Survey of Astronomy and Astrophysics in surveying the field of space- and ground-based astronomy and astrophysics and identifying promising areas of research in the coming decade, taking into consideration both new and previously identified concepts. The tasks assigned to the SFPs and PPPs are outlined in Appendix A. This volume also reflects the work of six Infrastructure Study Groups (ISGs) whose members were appointed by the NRC's Division on Engineering and Physical Sciences to gather and analyze data on issues related to several broad topics. The results of the work of the three groups—the SFPs, PPPs, and ISGs—were advisory to the survey committee, whose task was to recommend priorities for the most important scientific and technical activities in astronomy and astrophysics for the decade 2010-2020.[1] The survey committee's recommendations are presented in a separate volume, *New Worlds, New Horizons in Astronomy and Astrophysics* (The National Academies Press, Washington, D.C., 2010).

The survey was conducted in two overlapping phases. In the first phase, the five SFPs worked to identify science themes that define the research frontiers for the 2010-2020 decade in five areas: Cosmology and Fundamental Physics, the Galactic Neighborhood, Galaxies Across Cosmic Time, Planetary Systems and Star

---

[1] In this context, "activities" include any project, telescope, facility, mission, or research program of sufficient scope to be identified separately in this report.

Formation, and Stars and Stellar Evolution. Drawing on the 324 white papers on science opportunities submitted to the NRC in response to an open call from the survey committee to the astronomy and astrophysics research community,[2] as well as on briefings received from federal agencies that provide support for the field, the SFPs strove to identify the scientific drivers of the field and the most promising opportunities for progress in research in the next decade, taking into consideration those areas where the technical means and the theoretical foundations are in place for major steps forward. The SFPs were instructed to avoid advocacy for prioritization of specific new missions, telescopes, and other research activities. They also worked ahead of and therefore independent of the PPPs. As delineated in Chapters 1 through 5 of this volume, the input of each of the SFPs to the survey committee was organized by four science questions ripe for answering and general areas with unusual discovery potential.

In the second phase of the survey, the PPPs were charged to develop a ranked program of research activities in four programmatic areas: Electromagnetic Observations from Space; Optical and Infrared Astronomy from the Ground; Particle Astrophysics and Gravitation; and Radio, Millimeter, and Submillimeter Astronomy from the Ground. In addition to the draft science questions and discovery areas received from the SFP chairs at a joint meeting held in May 2009, the PPPs also reviewed the more than 100 proposals for research activities presented by the astronomy and astrophysics community for consideration by the survey.[3] In addition the PPPs received briefings from federal agencies, project proponents, and other stakeholders at public sessions held in June 2009 at the summer meeting of the American Astronomical Society in Pasadena, California. In their final assembly of priorities the PPPs also took into account assessments of cost and schedule risk, and of the technical readiness of the research activities under consideration for prioritization, that were provided by an NRC-hired contractor, the Aerospace Corporation. As presented in Chapters 6 through 9 of this volume, each PPP report contains a proposed program of prioritized, balanced, and integrated research activities, reflecting the results of its in-depth study of the technical and programmatic issues and its consideration of the results of the independent technical evaluation and cost and schedule risk estimate. The survey committee received draft reports of the PPPs' input on proposed programs at its fourth committee meeting in October 2009.

The SFPs and the PPPs conducted their work independent of each other, although coordinating calls among the panel chairs were held frequently. No members of the panels served on the survey committee, but the panel chairs did attend

---

[2] The set of white papers submitted is available at http://sites.nationalacademies.org/BPA/BPA_050603.

[3] For more information see http://sites.nationalacademies.org/BPA/BPA_049855.

all but the final committee meeting, and liaisons from the committee attended panel meetings.

The six Infrastructure Study Groups that also provided input for the survey committee's consideration consisted of 71 volunteer consultants drawn for the most part from the astronomy and astrophysics community. These groups gathered and analyzed data on issues in six areas—Computation, Simulation, and Data Handling (including archiving of astronomical data); Demographics (encompassing astronomers and astrophysicists working in different environments and subfields); Facilities, Funding, and Programs (including infrastructure issues such as support for laboratory astrophysics and technology development and theory); International and Private Partnerships; Education and Public Outreach; and Astronomy and Public Policy (benefits to the nation that accrue from federal investment in astronomy and from the potential contributions that professional astronomers can make to research of societal importance, and mechanisms by which the astronomy community provides advice to the federal government)—to describe recent trends and past quantifiable impacts on research programs in astronomy and astrophysics. The ISGs provided preliminary reports to the survey committee and the PPPs at the May 2009 so-called jamboree meeting, and their final reports were completed by the fall of 2009.

It then became the task of the survey committee to integrate the inputs from the SFPs and the PPPs, along with that from the ISGs, into a recommended program for all of astronomy and astrophysics for the decade 2010-2020.

The five SFPs, four PPPs, and six ISGs were critical components of the survey, not only for the content and critical analysis they supplied but also because of the connections they provided to the astronomy and astrophysics community. Moreover the panels completed a Herculean set of tasks in an extraordinarily short time. As presented in this volume, the results of their efforts were essential to the deliberations of the survey committee, the success of whose work depended critically on the sequential and orderly flow of information from the SFPs to the PPPs, and then to the committee as provided for in the survey plan and structure.

In addition, the survey as a whole benefited immensely from the broader participation of the astronomy and astrophysics community, which, over the course of the study and in particular in the first half of 2009, undertook a massive effort to provide input to the survey process. Included were informal reports from 17 community town hall meetings, more than 20 unsolicited e-mails, and 90-plus notices of interest for project activities, in addition to more than 450 white papers on topics including science opportunities, the state of the profession and infrastructure, and opportunities in technology development, theory, computation, and laboratory astrophysics. Critical to the success of the nine panels' and six study groups' work, these inputs were also an early product of the survey in that the white papers and various reports were made available on the NRC Web pages. On behalf of the survey committee and the panels, sincere thanks are extended to the volunteers from the research community who gave so much of their time to formulate this backbone of information and data as input for the Astro2010 survey process.

The survey committee also acknowledges with heartfelt thanks the critical input represented by the material provided in this volume. The reports of the SFPs and the PPPs stand as a testament to the hard work done by the panels, and especially their chairs. The full value of this tremendous effort will be recognized through the decade to come. The survey committee and the entire field of astronomy and astrophysics owe a great deal of thanks to all those who dedicated their time and effort to the Astro2010 survey activities.

Roger D. Blandford, *Chair*
Committee for a Decadal Survey of
Astronomy and Astrophysics

Acknowledgment of Members of the Astro2010 Infrastructure Study Groups

The Committee for a Decadal Survey of Astronomy and Astrophysics acknowledges with gratitude the contributions of the members of the Astro2010 Infrastructure Study Groups, who gathered information on issues related to the broad topics listed below.

Computation, Simulation, and Data Handling: Robert Hanisch, Space Telescope Science Institute, *Co-Chair;* Lars Hernquist, Harvard University, *Co-Chair;* Thomas Abel, Stanford University; Keith Arnaud, NASA Goddard Space Flight Center; Tim Axelrod, LSST; Alyssa Goodman, Harvard-Smithsonian Center for Astrophysics; Kathryn Johnston, Columbia University; Andrey Kravtsov, University of Chicago; Kristen Larson, Western Washington University; Carol Lonsdale, National Radio Astronomy Observatory; Mordecai-Mark Mac Low, American Museum of Natural History; Michael Norman, University of California, San Diego; Richard Pogge, Ohio State University; and James Stone, Princeton University.

Demographics: James Ulvestad, National Radio Astronomy Observatory, *Chair;* Jack Gallimore, Bucknell University; Evalyn Gates, University of Chicago; Rachel Ivie, American Institute of Physics; Christine Jones, Harvard-Smithsonian Center for Astrophysics; Patricia Knezek, WIYN Consortium, Inc.; Travis Metcalfe, National Center for Atmospheric Research; Naveen Reddy, National Optical Astronomy Observatory; Joan Schmelz, University of Memphis; and Louis-Gregory Strolger, Western Kentucky University.

Facilities, Funding, and Programs: J. Craig Wheeler, University of Texas at Austin, *Chair;* Rebecca A. Bernstein, University of California, Santa Cruz; David Burrows, Pennsylvania State University; Webster Cash, University of Colorado; R. Paul Drake, University of Michigan; Jeremy Goodman, Princeton University; W. Miller Goss, National Radio Astronomy Observatory; Kate Kirby, Harvard-Smithsonian Center for Astrophysics; Anthony Mezzacappa, Oak Ridge National Laboratory; Robert Millis, Lowell Observatory; Catherine Pilachowski, Indiana University; Farid Salama, NASA Ames Research Center; and Ellen Zweibel, University of Wisconsin.

International and Private Partnerships: Robert L. Dickman, National Radio Astronomy Observatory, *Chair;* Michael Bolte, University of California, Santa Cruz; George Helou, California Institute of Technology; James Hesser, Herzberg Institute of Astrophysics; Wesley T. Huntress, Carnegie Institution of Washington; Richard Kelley, NASA Goddard Space Flight Center; Rolf-Peter Kudritzki, University of Hawaii; Eugene H. Levy, Rice University; Antonella Nota, Space Telescope Science Institute; and Brad Peterson, Ohio State University.

Education and Public Outreach: Lucy Fortson, Adler Planetarium, *Co-Chair;* Chris Impey, University of Arizona, *Co-Chair;* Carol Christian, Space Telescope Science Institute; Lynn Cominsky, Sonoma State University; Mary Dussault, Harvard-Smithsonian Center for Astrophysics; Richard Tresch Feinberg, Phillips Academy; Andrew Fraknoi, Foothill College; Pamela Gay, Southern Illinois University; Jeffrey Kirsch, Reuben H. Fleet Science Center; Robert Mathieu, University of Wisconsin; George Nelson, Western Washington University; Edward Prather, University of Arizona; Philip Sadler, Harvard-Smithsonian Center for Astrophysics; Keivan Stassun, Vanderbilt University; and Sidney Woolf, LSST.

Astronomy and Public Policy: Daniel F. Lester, University of Texas at Austin, *Chair;* Jack Burns, University of Colorado; Bruce Carney, University of North Carolina; Heidi Hammel, Space Science Institute; Noel W. Hinners, Lockheed (retired); John Leibacher, National Solar Observatory; J. Patrick Looney, Brookhaven National Laboratory; Melissa McGrath, NASA Marshall Space Flight Center; and Annelia Sargent, California Institute of Technology.

Acknowledgment of Reviewers

These panel reports have been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the Report Review Committee of the National Research Council (NRC). The purpose of this independent review is to provide candid and critical comments that will assist the institution in making the published reports as sound as possible and to ensure that the reports meet institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscripts remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of these reports:

Science Frontiers Panel Reports

Jonathan Bagger, Johns Hopkins University
Sarbani Basu, Yale University
Timothy Beers, Michigan State University
John H. Black, Chalmers University of Technology, Sweden
Joseph Burns, Cornell University
Len Cowie, University of Hawaii
Marc Davis, University of California, Berkeley
Henry Ferguson, Space Telescope Science Institute
Marla Geha, Yale University
Andrew Gould, Ohio State University
Craig Hogan, University of Chicago

Michael Jura, University of California, Los Angeles
Vicky Kalogera, Northwestern University
Gillian Knapp, Princeton University
Richard McCray, University of Colorado, Boulder
Christopher McKee, University of California, Berkeley
Ramesh Narayan, Harvard-Smithsonian Center for Astrophysics
Roman Rafikov, Princeton University
Michael Strauss, Princeton University
Ann Wehrle, Space Science Institute
Rainer Weiss, Massachusetts Institute of Technology (professor emeritus)
Bruce Winstein, University of Chicago
Mark Wyatt, University of Cambridge

Program Prioritization Panel Reports

Jonathan Bagger, Johns Hopkins University
James Barrowman, NASA (retired)
John H. Black, Chalmers University of Technology, Sweden
Darrel Emerson, National Radio Astronomy Observatory
Reinhard Genzel, Max Planck Institute for Extraterrestrial Physics
Ronald L. Gilliland, Space Telescope Science Institute
James E. Gunn, Princeton University Observatory
Craig Hogan, University of Chicago
Vicky Kalogera, Northwestern University
Richard McCray, University of Colorado, Boulder
Christopher McKee, University of California, Berkeley
Ramesh Narayan, Harvard-Smithsonian Center for Astrophysics
Anthony Readhead, California Institute of Technology
Anneila Sargent, California Institute of Technology
Michael Strauss, Princeton University
Edward L. Wright, University of California, Los Angeles

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the reports' conclusions or recommendations, nor did they see the final draft of the reports before their release. The review of the Science Frontiers Panel reports was overseen by Kenneth H. Keller, Johns Hopkins School of Advanced International Studies, and Bernard F. Burke, Massachusetts Institute of Technology. The review of the Program Prioritization Panel reports was overseen by Louis J. Lanzerotti, New Jersey Institute of Technology, and Bernard F. Burke, Massachusetts Institute of Technology. Appointed by the NRC, they were responsible for making certain that an independent examination of the reports was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of the reports rests entirely with the authoring panels and the institution.

Contents

PART I
REPORTS OF THE ASTRO2010 SCIENCE FRONTIERS PANELS

PART II
REPORTS OF THE ASTRO2010 PROGRAM PRIORITIZATION PANELS

APPENDIXES

Part I

Reports of the Astro2010 Science Frontiers Panels

1

Report of the Panel on Cosmology and Fundamental Physics

SUMMARY

Astronomical observations have become a vital tool for studying fundamental physics, and advances in fundamental physics are now essential for addressing the key problems in astronomy and cosmology.

The past 15 years have been a period of tremendous progress in cosmology and particle physics:

- There is now a simple cosmological model that fits a host of astronomical data. Fifteen years ago, cosmologists considered a wide range of possible models; their best estimates of the Hubble constant differed by nearly a factor of two, and estimates of the mass density of the universe differed by as much as a factor of five. Today, the Lambda cold dark matter model is remarkably successful in explaining current observations, and the key cosmological parameters in this model have been measured by multiple techniques to better than 10 percent.
- Measurements of the cosmic microwave background (CMB), supplemented by observations of large-scale structure (LSS), suggest that the very early universe underwent a period of accelerated expansion that is likely to be attributable to a period of cosmological "inflation." The inflationary model predicts that the universe is nearly flat and that the initial fluctuations were Gaussian, nearly scale-invariant, and adiabatic. Remarkably, all of these predictions have now been verified.
- The astronomical evidence for the existence of dark matter has been improving for more than 60 years. Within the past decade, measurements of acoustic peaks

in the CMB have confirmed the predictions of big bang nucleosynthesis (BBN) that the dark matter must be nonbaryonic. Gravitational lensing measurements have directly mapped its large-scale distribution, and the combination of lensing and X-ray measurements has severely challenged many of the modified-gravity alternatives to dark matter.

- Supernova data, along with other cosmological observations, imply that the expansion of the universe is accelerating. This surprising result suggests either a breakdown of general relativity on the scale of the observable universe or the existence of a novel form of "dark energy" that fills space, exerts repulsive gravity, and dominates the energy density of the cosmos.
- The discovery that neutrinos oscillate between their electron, muon, and tau flavors as they travel, and hence that they have mass, provides evidence for new physics beyond the standard model of particle physics. The effects of oscillations were seen in the first experiment to measure solar neutrinos, and the interpretation was confirmed by measurements of atmospheric neutrinos produced by cosmic rays and by new solar neutrino experiments with flavor sensitivity.
- In the past few years, a cutoff has been seen in the energy spectrum of ultrahigh-energy cosmic rays (UHECRs) consistent with that predicted to arise from interactions with the CMB. UHECRs have become a powerful tool for probing the active galactic nuclei (AGN), galaxy clusters, or radio sources responsible for accelerating such particles.

Looking forward to the coming decade, scientists anticipate further advances that build on these results.

The Astro2010 Science Frontiers Panel on Cosmology and Fundamental Physics was tasked to identify and articulate the scientific themes that will define the frontier in cosmology and fundamental physics (CFP) research in the 2010-2020 decade. The scope of this panel report encompasses cosmology and fundamental physics, including the early universe; the cosmic microwave background; linear probes of large-scale structure using galaxies, intergalactic gas, and gravitational lensing; the determination of cosmological parameters; dark matter; dark energy; tests of gravity; astronomical measurements of physical constants; and fundamental physics derived from astronomical messengers such as neutrinos, gamma rays, and ultrahigh-energy cosmic rays.

In response to its charge, the panel identified four central questions that are ripe for answering and one general area in which there is unusual discovery potential:

- How did the universe begin?
- Why is the universe accelerating?
- What is dark matter?
- What are the properties of neutrinos?
- *Discovery area:* Gravitational wave astronomy.

How Did the Universe Begin?

Did the universe undergo inflation, a rapid period of accelerating expansion within its first moments? If so, what drove this early acceleration, when exactly did it occur, and how did it end? When introduced in the early 1980s, the inflationary paradigm made a number of generic observational predictions: we live in a flat universe seeded by nearly scale-invariant, adiabatic, Gaussian scalar fluctuations. Over the past decade, cosmological observations have confirmed these predictions. Over the coming decade, it may be possible to detect the gravitational waves produced by inflation, and thereby infer the inflationary energy scale, through measurements of the polarization of the microwave background. It may also be possible to test the physics of inflation and distinguish among models by precisely measuring departures from the predictions of the simplest models.

Why Is the Universe Accelerating?

Is the acceleration of the universe the signature of a breakdown of general relativity on the largest scales, or is it due to dark energy? The current evidence for the acceleration of the universe rests primarily on measurements of the relationship between distance and redshift based on observations of supernovae, the CMB, and LSS. Improved distance measurements can test whether the distance-redshift relationship follows the form expected for vacuum energy or whether the dark energy evolves with redshift. Measurements of the growth rate of LSS provide an independent probe of the effects of dark energy. The combination of these measurements tests the validity of general relativity on large scales. The evidence for cosmic acceleration provides further motivation for improving tests of general relativity on laboratory, interplanetary, and cosmic scales, and for searching for variations in fundamental parameters.

What Is Dark Matter?

Astronomical observations imply that the dark matter is nonbaryonic. Particle theory suggests several viable dark matter candidates, including weakly interacting massive particles (WIMPs)[1] and axions. Over the coming decade, the combination of accelerator experiments at the Large Hadron Collider (LHC), direct and indirect dark matter searches, and astrophysical probes are poised to test these and other leading candidates and may identify the particles that constitute dark matter.

---

[1] WIMPs are hypothetical particles serving as one possible solution to the dark matter problem. These particles interact through the weak nuclear force and gravity, and possibly through other interactions no stronger than the weak force. Because they do not interact with electromagnetism, they cannot be seen directly, and because they do not interact with the strong nuclear force, they do not react strongly with atomic nuclei.

Successful detections would mark the dawn of dark matter astronomy: the use of measurements of dark matter particles or their annihilation products to probe the dynamics of the galaxy and the physics of structure formation.

What Are the Properties of Neutrinos?

What are the masses of the neutrinos? What are their mixing angles and couplings to ordinary matter? Is the universe lepton-number symmetric? Solar neutrino astronomy determined the Sun's central temperature to 1 percent and provided the first evidence for neutrino oscillations. Neutrino events from Supernova 1987A confirmed scientists' basic ideas about stellar core collapse and placed important new constraints on neutrino properties. Owing to rapid advances in neutrino-detection technologies, over the coming decade astronomers will be able to use neutrinos as precise probes of solar and supernova interiors and of ultrahigh-energy cosmic accelerators. Cosmological measurements of structure growth offer the most sensitive probe of the neutrino mass scale, with the potential to reach the 0.05-eV lower limit already set by oscillation experiments. The next generation of neutrino detectors could detect the cosmic background of neutrinos produced over the history of star formation and collapse. Ultrahigh-energy neutrino detectors will record the neutrino by-products of the interactions of UHECRs with CMB photons, the same interactions that degrade the energy of the charged particles and cause the high-energy cutoff. These experiments offer a unique probe of physics at and beyond the TeV scale. Improved measurements of light-element abundances might relieve the current tension between the predictions of BBN and observations, or they might amplify this tension and point the way to a revised model of neutrino physics or the early universe.

Discovery Area: Gravitational Wave Astronomy

With upcoming and prospective experiments about to open a new window on the universe, gravitational wave astronomy is an area of unusual discovery potential that may yield truly revolutionary results. Gravitational waves, on the verge of being detected, can be used both to study astrophysical objects of central importance to current astronomy and to perform precision tests of general relativity. The strongest known sources of gravitational waves involve extreme conditions—black holes and neutron stars (and especially the tight binary systems containing them), core-collapse supernovae, evolving cosmic strings, and early-universe fluctuations—and studies of these phenomena can advance the understanding of matter at high energy and density. General relativity predicts that gravitational waves propagate at the speed of light and produce a force pattern that is transverse and quadrupolar. Observations of black hole mergers with high signal-to-noise ratios will make possible extremely precise tests of many predictions of general relativity in the strong-field regime, such as whether black holes really exist and whether the warped space-time that surrounds them obeys the theorems developed by Hawking, Penrose, and others. And because merging black hole binaries act as "standard sirens," there is a well-understood relationship between their waveform and their intrinsic luminosity. If their optical counterparts can be detected, they will enable a novel approach to absolute distance measurements of high-redshift objects.

A worldwide network of terrestrial laser interferometric gravitational wave observatories is currently in operation, covering the 10- to 1,000-Hz frequency range. This network may soon detect neutron star–black hole mergers and stellar mass black hole–black hole mergers. Operating in the much lower nanohertz (10^{-9} Hz) frequency range are pulsar timing arrays. The low-frequency range, between 10^{-5} and 10^{-1} Hz, is believed to be rich in gravitational wave sources of strong interest for astronomy, cosmology, and fundamental physics. This portion of the gravitational wave spectrum can be accessed *only* from space. Space-based detections can achieve much higher precision measurements of black hole mergers and thus much stronger tests of general relativity.

Theoretical and Computational Activities

Theory and observation are so closely intertwined in investigations of cosmology and fundamental physics that it is often difficult to define the border between them. Many of the ideas that are central to the next decade's empirical investigations originated decades ago as theoretical speculations. Many of the tools now being used for these investigations grew out of theoretical studies that started long before the methods were technically feasible. Theory plays an important role in designing experiments, optimizing methods of signal extraction, and understanding and mitigating systematic errors. Theoretical advances often amplify the scientific return of a data set or experiment well beyond its initial design. More-speculative, exploratory theory may produce the breakthrough that leads to a natural explanation of observed phenomena or a prediction of extraordinary new phenomena. In all these areas, high-performance computing plays a critical and growing role. Robust development of a wide range of theoretical and computational activities is essential in order to reap the return from the large investments in observational facilities envisioned over the next decade.

Key Activities Identified by the Panel

The panel identified the following key activities as essential to realizing the scientific opportunities within the decade 2010-2020 (the list is unranked):

Inflation and Acceleration

- Measure the amplitude of the initial scalar fluctuations of both matter and space-time across all observationally accessible scales through measurements of CMB E-mode polarization, the LSS of galaxies, weak lensing of galaxies and the CMB, and fluctuations in the intergalactic medium.
- Search for ultra-long-wavelength gravitational waves through measurements of CMB B-mode polarization, achieving sensitivities to the tensor-scalar ratio at the level set by astronomical foregrounds. Detection of these gravitational waves would determine the energy scale of inflation.
- Search for isocurvature modes, non-Gaussian initial conditions, and other deviations from the fluctuations predicted by the simplest inflationary models.
- Measure the curvature of the universe to a precision of 10^{-4}, the limit set by horizon-scale fluctuations.
- Determine the history of cosmic acceleration by measuring the distance-redshift relation and Hubble parameter to sub-percent accuracy over a wide range of redshifts.
- Determine the history of structure growth by measuring the amplitude of matter clustering to sub-percent accuracy over a wide range of redshifts.
- Improve measurements that test the constancy of various physical constants and the validity of general relativity.

Dark Matter

- Probe both spin-independent and spin-dependent dark matter scattering cross sections with searches that explore much of the parameter space of WIMP candidates, through both underground experiments and searches for dark matter annihilating to neutrinos. Although a review of laboratory dark matter detection methods is outside the scope of this panel's charge, progress in this area (as well as progress at the Large Hadron Collider) is critical for determining the properties of dark matter. As noted in the NRC report *Revealing the Hidden Nature of Space and Time*,[2] the proposed International Linear Collider may turn out to be an essential tool for studying dark matter.
- Carry out indirect searches for dark matter that probe the annihilation cross sections of weakly interacting thermal relics. Identifying "smoking gun" signals is essential for detecting dark matter annihilation products above the astronomical backgrounds.
- Improve astrophysical constraints on the local dark matter density and

---

[2] National Research Council, *Revealing the Hidden Nature of Space and Time: Charting the Course for Elementary Particle Physics*, The National Academies Press, Washington, D.C., 2006.

structure on subgalactic scales to test the paradigm of cold, collisionless, and stable dark matter and to look for evidence for alternative dark matter candidates. These astronomical observations, particularly of dwarf galaxies, help to optimize dark matter search strategies and will be critical for determining the implications of dark matter signals for the particle properties of dark matter.

Neutrinos

- Develop the sensitivity to detect and study the ultrahigh-energy (UHE) neutrinos that can be expected if the cosmic-ray energy cutoff is due to protons annihilating into neutrinos and other particles. UHE neutrino fluxes above those expected from the Greisen-Zatsepin-Kuzmin (GZK) mechanism would be the signature of new acceleration processes.
- Measure the neutrino mass to a level of 0.05 eV, the lower limit implied by current neutrino mixing measurements, through its effects on the growth of structure.
- Enable precision measurement of the multiflavor neutrino “light curves” from a galactic supernova.
- Improve measurements of light-element abundances in combination with big bang nucleosynthesis theory to test neutrino properties and dark matter models.

Gravitational Waves

- Detect gravitational waves from mergers of neutron stars and stellar mass black holes.
- Detect gravitational waves from inspiral and mergers of supermassive black holes at cosmological distances.
- Achieve high signal-to-noise ratio measurements of black hole mergers to test general relativity in the strong-field, highly dynamical regime.
- Identify electromagnetic counterparts to gravitational wave sources.
- Open a radically new window on the universe, with the potential to reveal new phenomena in stellar-scale astrophysics, early-universe physics, or other unanticipated areas.

Theory

- Advance theoretical work that provides the foundation for empirical approaches, through the development of methods, design of experiments, calculation of systematic effects, and statistical analysis.
- Advance theoretical work that provides interpretation of empirical results in terms of underlying physical models.

- Push the frontiers of exploratory theory, which can lead to breakthrough ideas needed to address the deep mysteries of cosmology and fundamental physics.

INTRODUCTION

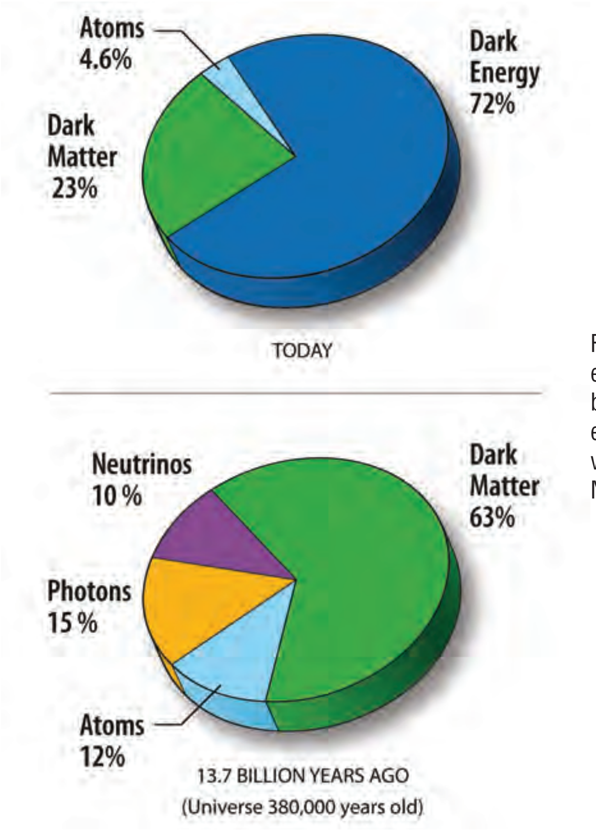
Since the dawn of modern science, advances in fundamental physics have elucidated the deepest mysteries of astronomy. As part of this symbiotic relationship, astronomical observations have stimulated new advances in fundamental physics. Kepler, Galileo, and Newton devised new theories of motion, force, and universal gravity to explain the wandering of the planets across the sky. Quantum mechanics enabled the understanding of stellar spectra and revealed that stars are made primarily of hydrogen and helium rather than of the oxygen, silicon, and iron that dominate Earth and meteorites. Advances in nuclear physics were essential for explaining the unknown energy source of stars. Today, astronomers confront new mysteries: dark matter, cosmic acceleration, and the origin of structure (Figure 1.1). Again, advances in fundamental physics are needed—and astronomy again offers a powerful laboratory for probing fundamental physics.

Over the past three decades, astronomers and physicists have made remarkable progress toward a detailed scientific theory of the cosmos, a "standard model" of cosmology that explains observations that probe an enormous range of time and distance. But this theory is still incomplete, and it relies on three key physical ideas that are at best partly understood: inflation, cold dark matter, and vacuum energy.

The inflation hypothesis, first proposed in the early 1980s, asserts that the universe grew by an enormous factor during its first moments. This accelerating expansion not only erases any pre-existing fluctuations but also generates a nearly scale-invariant spectrum of Gaussian fluctuations that leave an imprint on the CMB and grow to form galaxies and clusters of galaxies. Cold dark matter, composed of weakly interacting particles with low thermal velocities in the early universe, explains the dynamics of galaxies and clusters and allows consistency between CMB and LSS observations. Vacuum energy exerts repulsive gravity, driving the present-day acceleration of cosmic expansion (which is many, many orders of magnitude slower than the acceleration hypothesized to occur during inflation).

Despite its observational successes, this standard model is unsatisfying in several ways. Scientists do not know the physics that caused inflation to happen or to end, nor do they know for sure that inflation is the mechanism that created a large, radiation-filled universe and seeded it with fluctuations. There are several plausible ideas for what the dark matter might be, but it is not known which of them, if any, is correct.

By far the most surprising element of the model is the vacuum energy. While quantum physics does allow "empty" space to be filled with energy, the naively predicted value of this energy is 10^{120} times larger than the value allowed by ob-



TODAY

13.7 BILLION YEARS AGO
(Universe 380,000 years old)

FIGURE 1.1 The composition of the universe has evolved over the 13.7 billion years since the big bang. Today the universe is dominated by dark energy (*upper*), whereas after the big bang it was dominated by dark matter (*lower*). SOURCE: NASA/WMAP Science Team.

servations. It may be that the true vacuum energy is zero and a pervasive, previously unknown fundamental field, akin to the field that caused inflation in the early universe, is driving the present-day acceleration. Alternatively, the observed acceleration could be a sign that general relativity itself breaks down on the scale of the observable universe.

The most-studied hypothesized dark matter particles are in many ways analogous to neutrinos in that they interact with baryonic matter only by way of gravity and the weak interaction (although neutrinos are much, much less massive). Over the past four decades, and especially over the past decade, progress in neutrino physics has been driven principally by astronomical observations. Most notably, observations of solar neutrinos and of atmospheric neutrinos produced by cosmic rays have demonstrated that the three neutrino species in the standard model of particle physics have non-zero mass and that they oscillate from one form to another as they propagate through matter or empty space. Cosmological observations set the strongest upper limits on the neutrino mass; they show that the standard-model neutrinos cannot be the main form of dark matter, but it remains possible that a fourth, sterile neutrino species could constitute the dark matter.

These developments, and the successes and limitations of the current cosmological model, suggest the following four questions to guide research in cosmology and fundamental physics over the next decade:

- How did the universe begin?
- Why is the universe accelerating?
- What is dark matter?
- What are the properties of neutrinos?

The sections that follow elucidate these questions and describe the capabilities needed to answer them.

The panel also identified gravitational wave astronomy as an emerging science area with unusual discovery potential. Scientists expect the next decade to see the first direct detection of gravitational waves, the propagating ripples of space-time predicted by Einstein nearly a century ago. The strongest expected sources of gravitational waves are violent events such as mergers of black holes and neutron stars; gravitational wave measurements will provide unique insights into the physics of these events and allow powerful tests of general relativity in a completely new regime. More enticing still are the prospects for sources that have not yet been imagined or have only been speculated about, perhaps new classes of stellar implosions or collisions, or backgrounds of gravitational waves produced in the early universe. If the history of radio astronomy and X-ray astronomy is any guide, then the dawn of gravitational wave astronomy will fundamentally change our view of the cosmos and the objects that it contains. The panel discusses the extraordinary discovery potential of gravitational wave astronomy and the technical capabilities needed to realize it in the section below titled "CFP Discovery Area—Gravitational Wave Astronomy: Listening to the Universe."

Three themes connect the observational approaches to all of these questions:

- The first is the mapping of cosmological initial conditions over the widest possible dynamic range with measurements of CMB temperature and polarization fluctuations, observations of weak lensing, and optical and radio observations that use galaxies and intergalactic gas to map the distribution of matter at lower redshifts. The enormous increase in statistical precision and dynamic range now possible will enable new tests of models of inflation, precision measurements of the geometry of space, the determination of the masses of neutrinos by means of their cosmological effects, and tests of theories for the origin of cosmic acceleration.

Realizing these vast improvements in statistical power requires exquisitely careful control of systematic uncertainties, which often present the greatest challenge to these methods.

- The second theme is the opening of new windows that allow scientists to view astrophysical phenomena in radically new ways. Gravitational waves are the most dramatic example of such a new window, but they are not the only one. Dark matter searches hinge on great advances in the sensitivity and sky coverage of high-energy gamma-ray and cosmic-ray experiments. New facilities should achieve the first detections of ultrahigh-energy neutrinos. Searches for highly redshifted 21-cm radiation will provide the first three-dimensional maps of structure at the epoch of cosmic reionization, and advances in these techniques should eventually allow maps of cosmic initial conditions over unprecedented volumes.
- The third theme is the universe as a laboratory for fundamental physics. Studies of primordial fluctuations probe early-universe physics at energies that can never be achieved in terrestrial accelerators. The explanation of cosmic acceleration may radically reshape the understanding of gravity, the quantum vacuum, or both. Dark matter experiments provide windows on extensions of the standard model that beautifully complement the traditional tools of particle physics. Astrophysical measurements provide the most powerful and varied constraints on neutrino properties. Gravitational waves probe general relativity in the strong-field regime, a test that can be done only in the extreme environment near black holes.

CFP 1. HOW DID THE UNIVERSE BEGIN?

Although little is understood about the origin of the universe, cosmologists have made significant progress in studying its very early history. In the early 1980s, they theorized that during its first moments, the universe underwent a rapid period of accelerated expansion called inflation. During inflation, microscopic, causally connected regions expanded exponentially, driving the spatial curvature to nearly zero and producing a homogeneous universe. This inflationary paradigm not only explained many of the open questions in cosmology but also predicted that quantum fluctuations of both matter and space-time curvature create nearly scale-invariant, adiabatic, Gaussian random phase fluctuations. During the past decade, observations have shown impressive agreement with these predictions.

Although inflation is a successful paradigm, its underlying mechanism remains a mystery. Inflation may have something to do with grand unified theories that amalgamate the strong and electroweak interactions at an ultrahigh-energy scale. It may derive from a quantum-gravity theory such as string theory. Inflation may arise at some lower-energy phase transition, such as the breaking of the Peccei-Quinn symmetry posited to explain the lack of charge-parity violation in the strong interaction. It may be a consequence of the compactification of large extra dimensions or of a departure from general relativity at high densities. In any scenario, however, inflation is driven by some new physics at ultrahigh densities and energies well beyond those accessible in the laboratory.

Testing the Inflationary Paradigm and Identifying the Underlying Physics Model

A key aim of cosmological observation and theory in the coming decade will be to test the inflationary paradigm further and to identify the underlying physical model responsible for inflation; cosmologists desire a full description of the high-energy physics responsible for inflation.

Single-Field Slow-Roll Inflation

In the simplest scenario, inflation is driven by the displacement of a scalar field from the minimum of its potential. If the potential has the right shape, then the scalar field will roll slowly toward its minimum, and the vacuum energy associated with this displacement will drive the accelerated expansion. The single-field, slow-roll (SFSR) model makes a number of testable predictions: (1) a flat universe with a curvature scale much larger than the horizon, (2) nearly scale-invariant fluctuations in the spatial distribution of matter, (3) adiabatic fluctuations, (4) Gaussian fluctuations, (5) homogenous and isotropic fluctuations, and (6) a stochastic background of inflationary gravitational waves (IGWs) with a nearly scale-invariant spectrum. The IGW amplitude is proportional to the square root of the height of the scalar-field potential or, equivalently, to the energy density or expansion rate during inflation, whereas the deviations from scale invariance in matter and IGWs describe the potential's shape. The Wilkinson Microwave Anisotropy Probe (WMAP) and the current generation of ground- and balloon-based experiments have already tested the first five of the above predictions. The Planck spacecraft and the upcoming generation of CMB experiments will test these predictions to even higher precision and hold the promise of potentially detecting the IGW background.

Toward a Full Description of Physics During Inflation

It is quite surprising that single-field "toy models" have done so well in explaining such a large and precise body of data. Most theorists surmise that SFSR models are simply an approximation to more complex inflationary physics. The "true" model could be very different from SFSR models, with the differences including modifications to the kinetic-energy term of the scalar field, multiple fields driving inflation, models with features in the scalar-field potential, and alternative scenarios such as the ekpyrotic model that posits a collapsing phase prior to the big bang. These alternatives to the simplest SFSR model make new, measurable predictions including non-Gaussian correlations and non-adiabatic (isocurvature) density fluctuations. In addition, physical processes such as phase transitions at the end of inflation can produce topological defects (e.g., cosmic strings) or alter the IGW background.

The primary goal in the coming decade is to test precisely each of the predictions of SFSR inflation. If consistency with the predictions of SFSR inflation persists, then the range of allowable values of the scalar-field potential will need to be narrowed. If departures from the simplest predictions are found, this will provide insights into fundamental physics and into the first moments of the early universe.

Testing the Predictions of Single-Field Slow-Roll Inflation

Flatness

The expected contribution of curvature from inflation is determined by the amplitude of large-scale density fluctuations, $\Omega_k \approx 10^{-5}$ to 10^{-4}. The combination of Planck CMB measurements with baryon acoustic oscillations (BAOs), distance measures from high-redshift galaxies, and 21-cm line surveys can improve constraints on the curvature of the universe by two orders of magnitude from the current limits of $\Omega_k \approx 10^{-2}$. Such values may also be achievable by measuring weak lensing of the CMB with a high-angular-resolution, high-sensitivity CMB polarization experiment. Observations in the coming decade should allow the direct testing of this key inflationary prediction (Figure 1.2).

Scalar Power Spectrum

The mechanism by which structure was generated in the early universe is constrained by the linear power spectrum of fluctuations. On the largest observable spatial scales, the CMB provides the cleanest window into inflation. The Planck satellite should give an order-of-magnitude improvement in the measurement of the slope of the primordial power spectrum and its scale dependence, both measured to 10^{-3} within the early part of the decade. At angular scales smaller than those accessible with Planck, the kinematic Sunyaev-Zel'dovich (SZ) effect due to the scattering off of moving electrons at low redshift starts to dominate the temperature fluctuations, thus obscuring primordial-spectrum information.

Further improvements in power-spectrum characterization will come from a combination of CMB polarization maps and LSS surveys down to few-megaparsec (Mpc) scales. A variety of LSS observations (galaxy, weak gravitational lensing, Lyman-α, and 21-cm surveys) give a wealth of information about the power spec-

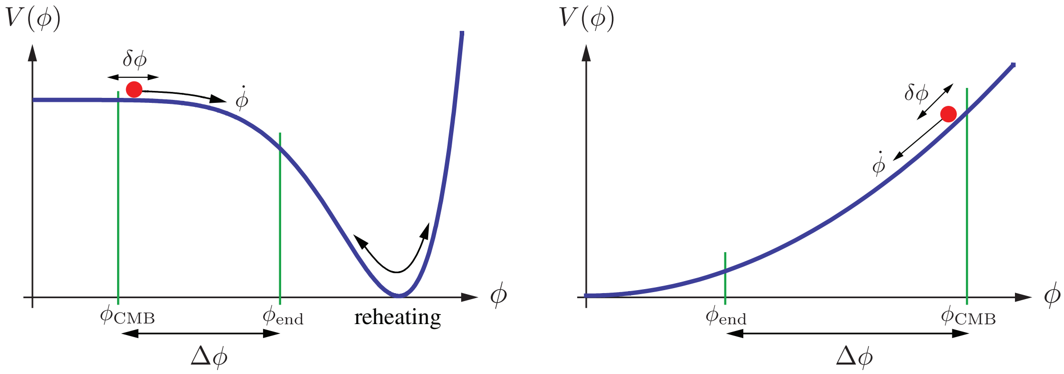


FIGURE 1.2 Cosmic microwave background (CMB) and large-scale structure observations can reveal key information about the shape and amplitude of the inflationary potential. Inflation occurs while the potential's gradient is small, as the scalar field slowly rolls down the potential. When the gradient of the field becomes too large, at ϕ_{end}, inflation ends. Quantum fluctuations in the scalar field, $\delta\phi$, are sensitive to the shape of the potential. They become frozen during inflation, and a snapshot at ϕ_{CMB} is imprinted in the temperature and polarization of the CMB and the distribution of galaxies and galaxy clusters. The amplitude of CMB B-mode polarization is sensitive to the change in ϕ ($\Delta\phi$) during inflation. Two examples of single-field slow-roll inflation potentials are shown: (*left*) small-field, $\Delta\phi < M_p$ and (*right*) large field, $\Delta\phi > M_p$. For large-field models this primordial signal could well be measurable. SOURCE: Reprinted with permission from D. Baumann, M.G. Jackson, P. Adshead, et al., *CMBPol Mission Concept Study: Probing Inflation with CMB Polarization*, AIP Conference, copyright 2009, American Institute of Physics.

trum of fluctuations in the early universe through a three-dimensional, rather than two-dimensional, picture of the matter distribution extending to smaller scales inaccessible to the CMB. At present, only 10^{-6} of the linear spatial modes in the observable universe have been measured.[3] Galaxy surveys over larger volumes could survey 10^8 to 10^9 modes out to redshift $z < 2$, characterizing the linear matter power spectrum with statistical uncertainties that approach the cosmic-variance limit.

The 21-cm signal is currently untested as a cosmological probe, but it offers the potential to measure the matter power spectrum from $z > 6$ down to Mpc (and possibly smaller) scales, with these modes in the linear regime at such redshifts. These observations have the potential to detect far more modes than are accessible in the CMB alone.

Observations of redshifted 21-cm emission at $z > 20$ will be extremely challenging; lower redshift observations of 21-cm emission, probing the epoch of reionization (EoR), will be an important first step. These observations help outline

---

[3] The number of independent modes in the visible universe is the volume of the observable universe divided by the volume of the smallest measurable scale. The number of modes is a measure of the amount of information about initial conditions accessible to experiments.

a seamless history of the universe from the primordial perturbations laid down by inflation to the diversity of stars, galaxies, and clusters that are seen in the universe today. The Panel on Galaxies Across Cosmic Time discusses these EoR observations in more detail in its report in Chapter 3 of this volume.

Gravitational Waves

Gravitational waves produce a distinctive B-mode signal in the CMB that could provide the most promising method for detecting the "smoking gun" of inflation. Different inflationary models predict different amplitudes for r, the ratio of gravitational wave fluctuations to scalar fluctuations. If the IGW background is detected, additional tests of the inflationary consistency relations can be obtained by measuring its spectrum.

Ground- and balloon-based experiments with high sensitivity but modest angular resolution could detect an IGW amplitude of $r \approx 0.01$. Measuring $r < 0.01$ will require higher signal-to-noise ratios and better angular resolution in both temperature and polarization in order to separate the primordial B modes from those induced by weak lensing. Because a large spectral range will likely be essential for removing galactic foregrounds, a dedicated space-based B-mode survey will be required to obtain the values $r \approx 10^{-4}$ to 10^{-3} required to access the full parameter space implied by grand unified theory (GUT)-scale inflation. Such an experiment would also provide a wealth of other extremely valuable information (including detailed measurements of the gravitational lensing of the CMB), which will constrain the small-scale matter power spectrum at intermediate redshifts. These measurements will be useful for further testing inflation, constraining neutrino masses, studying the effects of dark energy, and determining cosmological parameters with further precision.

Searching for Evidence of Deviations from SFSR Inflation

Non-Gaussianity

The gravitational evolution of primordial fluctuations introduces non-Gaussian correlations in the distribution of matter with a non-Gaussianity parameter f_{NL}[4] of order unity, dwarfing those predicted by SFSR inflation ($f_{\mathrm{NL}} \approx 10^{-2}$). Current limits constrain f_{NL} to be <100. Non-slow-roll inflationary models inspired by string theory, and some alternatives to inflation, however, can have amplified primordial non-Gaussianity that could be detectable. By increasing the number of small-scale modes measured, Planck should improve current constraints on f_{NL}

---

[4] f_{NL} is a measure of the skewness of the potential fluctuations.

by a factor of five, to $f_{NL} \approx 4$. Obtaining greater sensitivity will require a shift from the two-dimensional CMB map to three-dimensional correlations in LSS, where more modes are accessible. A 100 to 1,000 million galaxy survey would target $f_{NL} \approx 1$. If there is a detection of primordial non-Gaussianity, the detailed structure, obtained from the shape dependence of the three-point (and higher) correlation function, can discriminate among inflationary models.

Isocurvature Modes

Multifield inflation models can introduce significant non-adiabatic (isocurvature) fluctuations in the distribution of matter. The Planck satellite's improved measurements of temperature and polarization will increase sensitivity to these signatures of physics beyond the SFSR models. A post-Planck CMB-polarization map that is cosmic-variance-limited to multipole moments $l \approx 2{,}000$ would provide an additional order-of-magnitude improvement on both correlated and uncorrelated isocurvature modes.

The requirements for testing inflation overlap those for investigating cosmic acceleration. They are summarized together in Box 1.1.

CFP 2. WHY IS THE UNIVERSE ACCELERATING?

Cosmic acceleration is widely regarded as the most profound puzzle in fundamental physics today. Even the *least* exotic explanations imply an energetically dominant new component of the universe with extraordinary physical properties. Alternative proposals include a breakdown of general relativity on cosmological scales, perhaps tied to extra dimensions or to low-energy manifestations of quantum gravity. While measurements of the distance-redshift relation using Type Ia supernovae provide the most direct evidence for cosmic acceleration, there are now multiple lines of supporting evidence, including the measurements of CMB fluctuations, the integrated Sachs-Wolfe effect, the Hubble constant, BAO, and the growth rate of structure based on X-ray observations and lensing. Cosmic acceleration is a surprising but accepted piece of modern cosmology. The two top-level questions in this field are these:

1. Is this acceleration caused by a breakdown of general relativity or by a new form of energy?
2. If dark energy is causing the acceleration, is its energy density constant in space and time?

Cosmologists describe the evolution of dark energy in terms of its equation-of-state parameter, $w = P/(\rho c^2)$, where P and ρ are the pressure and mass density, respec-

BOX 1.1
Conclusions on Inflation and Acceleration by the Science Frontiers Panel on Cosmology and Fundamental Physics

Goals

- Measure the amplitude of the initial scalar fluctuations across all observationally accessible scales through measurements of cosmic microwave background (CMB) E-mode polarization, the large-scale structure (LSS) of galaxies, weak lensing of galaxies and the CMB, and fluctuations in the intergalactic medium.
- Search for ultra-long-wavelength gravitational waves through measurements of CMB B-mode polarization, achieving sensitivities to the tensor-scalar ratio at the level set by astronomical foregrounds.
- Search for isocurvature modes, non-Gaussian initial conditions, and other deviations from the fluctuations predicted by the simplest inflationary models.
- Measure the curvature of the universe to precision of 10^{-4}, the limit set by horizon-scale fluctuations.
- Determine the history of cosmic acceleration by measuring the distance-redshift relation and Hubble parameter to sub-percent accuracy over a wide range of redshifts.
- Determine the history of structure growth by measuring the amplitude of matter clustering to sub-percent accuracy over a wide range of redshifts.
- Improve measurements that test the constancy of various physical constants and the validity of general relativity.

Needed Capabilities

- CMB experiments that measure E-mode polarization out to the limits set by foregrounds.
- CMB experiments that measure B-mode polarization, both for tensor modes and for lensing. These should begin with ground-based and balloon experiments with a variety of techniques, which can probe to $T/S \approx 10^{-2}$, and eventually proceed to space-based experiments, which may reach $T/S \approx 10^{-3}$ or below.
- Supernova campaigns that provide high-quality observations of several thousand Type Ia supernovae, including rest-frame infrared photometry.
- Weak-lensing surveys with accurate shape measurements and photometric redshifts of approximately 10^9 galaxies.
- Spectroscopic redshift surveys of approximately 10^8 galaxies.
- Measurements of small-scale structure in the intergalactic medium, from the Lyman-α forest and from pre-reionization 21-cm radiation.
- Significant improvements in precision tests of general relativity and time-variation of fundamental "constants."

tively. The mass density scales with redshift as $\rho_{DE}(z) = \rho_{DE,0} \times \exp[3\int(1 + w(z))\, d\ln(1 + z)]$. Vacuum energy, the simplest and arguably best-motivated model for dark energy, is constant with time, and so for this model $w = -1$ at all z. Alternative forms of dark energy have different values of $w(z)$.

The main line of attack for addressing these questions is to improve measurements of the Hubble parameter $H(z)$, the distance-redshift relation $D(z)$, and the growth function $G(z)$ that describes the strength of matter clustering. In a spatially flat universe, $H^2(z)$ is proportional to the *total* energy density—the sum of matter, radiation, and dark energy—and $D(z)$ is given by an integral of $H^{-1}(z)$. General relativity predicts a specific relation between $H(z)$ and $G(z)$; modified gravity models could alter this relation, or they could make $G(z)$ dependent on spatial scale. At present, the $D(z)$ relation is measured with a precision of roughly 5 percent at $z \leq 0.8$, with much weaker constraints at higher redshift. The function $G(z)$ is known to about 5 percent at low redshift, while direct constraints on the rate of growth $d\ln G(z)/dz$ are approximately 25 percent. Current data are consistent with general relativity and $w = -1 \pm 0.2$ (with the exact central value and error bar depending on the adopted data sets and on the assessment of systematic uncertainties). Proposed observations could realistically achieve one to two orders-of-magnitude improvement in the precision on w, and significant measurements of its redshift history if it turns out not to be constant.

Methods for Measuring Distances and Structure Growth

Supernovae

Type Ia supernovae have so far played the most critical role in demonstrating the existence of cosmic acceleration and measuring its evolution. Of the techniques used thus far, this one is the best understood from a practical point of view—both its strengths and its potential limitations. The statistical precision of the method is high, as each well-observed supernova yields a distance estimate with roughly 10 percent uncertainty. The key systematic uncertainties are corrections for dust extinction, the accuracy of photometric calibrations across a wide range of redshift, and the possible cosmic evolution of the supernova population.

Baryon Acoustic Oscillations

The BAO technique measures $D(z)$ and $H(z)$ using a scale imprinted by pressure waves in the pre-recombination universe as a standard ruler (Figure 1.3). This scale (≈150 Mpc) can be calculated precisely using parameters that are well determined by CMB observations, and it can be measured in the clustering of galaxies, quasars, the Lyman-α forest, or 21-cm emission. The ultimate statistical limit for BAO measurements is set by cosmic variance—that is, by the finite volume of the observable universe. For a similar survey volume and redshift range, a spectroscopic survey (with redshift errors smaller than the typical galaxy-peculiar velocity) provides several times higher precision than a photometric survey, and it allows separate determination of $D(z)$ and $H(z)$, whereas a photometric survey yields only the former. Large photometric surveys (e.g., for weak lensing) will yield

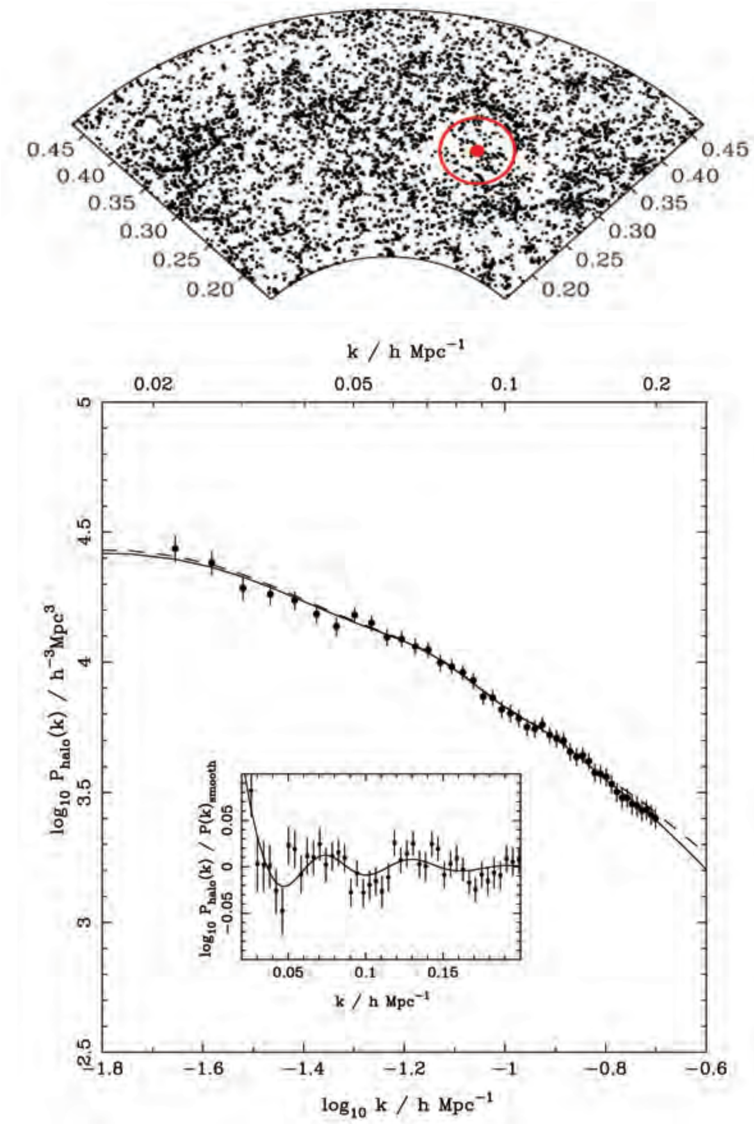


FIGURE 1.3 *Upper:* A slice of the Sloan Digital Sky Survey (SDSS) luminous red galaxy distribution out to redshift $z = 0.47$, with a superimposed "bulls-eye" showing the baryon acoustic oscilation (BAO) scale. *Lower:* The power spectrum measured from the final SDSS galaxy redshift survey, with the inset emphasizing the BAO signature of wiggles relative to the smooth power spectrum that would arise in a baryon-free universe. SOURCE: *Upper:* Daniel Eisenstein and Sloan Digital Sky Survey (SDSS) Collaboration, http://www.sdss.org. *Lower:* B.A. Reid, W.J. Percival, D.J. Eisenstein, L. Verde, D.N. Spergel, R.A. Skibba, N.A. Bahcall, T. Budavari, M. Fukugita, J.R. Gott, J.E. Gunn, et al., Cosmological constraints from the clustering of the Sloan Digital Sky Survey DR7 luminous red galaxies, *Monthly Notices of the Royal Astronomical Society* 404(1):60-85, copyright Royal Astronomical Society, 2010.

BAO measurements as a by-product, but spectroscopic surveys will be needed to harness the power of the BAO technique. There are possible systematic uncertainties from nonlinear bias of galaxies (or other tracers used), but current theory suggests that the systematic errors in the BAO technique will be smaller than the statistical errors, even for experiments that approach the cosmic variance limit.

Weak Lensing

Weak lensing depends on the amplitude of matter clustering, hence $G(z)$, and on the distance-redshift relation $D(z)$. A given weak-lensing data set can be analyzed with multiple statistical methods, including the power spectrum of cosmic shear in bins of photometric redshift (also known as tomography), higher-order measures such as the three-point correlation function, and galaxy-galaxy lensing. These multiple analyses can be used to increase the overall statistical precision, to carry out internal consistency checks for systematic errors, and to break degeneracies among cosmological parameters. The key systematic uncertainties for weak-lensing studies are as follows: the accuracy of shear measurements themselves, intrinsic galaxy alignments (which can mimic cosmic shear), the influence of baryons on small-scale theoretical predictions, and errors in photometric redshift distributions. Statistical errors depend mainly on the total number of galaxies with well-measured shapes.

Cluster Abundance

Measurements of the abundance of galaxy clusters as a function of redshift are sensitive to $G(z)$, which governs the evolution of the mass function, and to $D(z)$, which determines the volume element. In principle, cluster surveys can achieve high statistical sensitivity. The key challenge is that the mean relation between cluster observables and halo mass must be known to high accuracy, and the scatter and redshift evolution of these relations must be known to moderate accuracy. One can fit these relations to the cluster-count data themselves, but this reduces the statistical precision of the cosmological parameter measurements. X-ray, SZ, and optical galaxy selection are all viable approaches to assembling cluster catalogs. At present, weak lensing is the only observable whose fundamental physics is understood well enough to achieve the required, sub-percent-level mass calibration. While weak-lensing measurements are noisy for any individual cluster, they can precisely measure the mean mass profile of clusters binned by observable quantities.

Complementarity

These four methods are more powerful in combination than in isolation. Comparing the two pure distance methods, supernovae can achieve higher statistical precision than is possible with BAO at $z < 0.7$, and the control of systematic errors in supernova measurements is easier at low redshifts. BAO surveys are more effective at higher redshift because there is more available volume and because $H(z)$ is more sensitive to dark energy than is $D(z)$. Even though the dynamical effects of dark energy are smaller at high redshift, the sensitivity of a cosmic-variance-limited BAO survey to dark energy (more precisely, to a cosmological constant) is roughly flat over the redshift range $0.5 < z < 2.5$. High-redshift BAO measurements also yield precise constraints on curvature, breaking degeneracies among low-redshift measurements.

Weak lensing and cluster surveys provide independent measurements of $D(z)$, and they provide the $G(z)$ constraints needed to test modified gravity models. The galaxy redshift surveys designed for BAO studies can also measure $G(z)$ through redshift-space distortions, the apparent anisotropy of structure induced by galaxy-peculiar velocities. Recent theoretical work suggests that redshift-space distortions could be competitive with weak-lensing measurements of structure growth, but the systematic uncertainties of this method have not yet been explored.

Next-Generation Experiments for Distance and Structure Growth

The next decade should include aggressive observational programs that employ all of the methods described above. In addition to providing complementary information, the mix of methods allows cross-checks that will be crucial to drawing robust conclusions about cosmic acceleration.

Supernova Surveys

Supernova surveys require deep, high-cadence (frequent revisits to the same field), multiband imaging over areas of several square degrees to several tens of square degrees, with follow-up spectroscopy and excellent photometric calibration. Wider, shallower surveys are needed for low-redshift calibration samples. The top priority for supernova studies is to obtain high statistical precision and correspondingly tight control of systematic errors at redshifts $z \leq 0.7$, where their precision exceeds that of BAO. Systematic errors can be reduced by including observations in the rest-frame near-infrared (IR), where dust extinction is much lower and the range of luminosities is smaller than in the optical (even when the latter are corrected for light-curve duration). Improvements in photometric calibration systems are essential to supernova cosmology, and they will benefit many other areas of astronomy. Samples of several thousand supernovae can achieve statistical precision of approximately 0.01 mag in multiple redshift bins, and systematic uncertainties could possibly be brought to this level or below.

BAO Surveys

BAO surveys require spectroscopy over large co-moving volumes. At $z > 1$, samples of 10^8 to several times 10^8 galaxies are needed to reach the cosmic variance limit over a large fraction of the sky. Ground-based optical surveys can straightforwardly reach to redshifts $z \sim 1.3$, but accessing most of the co-moving volume in the range $1 < z < 2$ requires deep near-IR spectroscopy that can be done only from space. At higher redshifts, ground-based optical methods using emission-line galaxies and the Lyman-α forest become feasible, although it is not clear that these can approach the cosmic-variance limit. Intensity mapping of redshifted 21-cm emission is an emerging method that may allow efficient BAO measurements over a wide redshift range.

Weak Lensing and Cluster Surveys

Weak lensing and cluster surveys require high-resolution imaging with a well-characterized point-spread function (PSF) over wide areas (Figure 1.4). Ground-based surveys over 20,000 square degrees could achieve accurate shape measurements for approximately 10^9 galaxies. Higher-resolution space observations could increase the surface density of usable galaxies by a factor of several, and the more stable PSF obtainable in space could reduce systematic uncertainties in shape measurements. Achieving accurate photometric redshifts requires optical *and* near-IR photometry for all source galaxies, and spectroscopic samples of approximately 10^5 galaxies to the depth of the imaging surveys. Large-area X-ray and SZ surveys that overlap these imaging surveys may allow cleaner selection of clusters than is possible from the optical data alone. High signal-to-noise ratio X-ray and SZ measurements of smaller numbers of clusters help illuminate the underlying cluster physics and constrain the scatter of the mass-observable relations.

Space and Ground

Space observations could benefit all of these methods, in different ways. A common theme is the need for deep near-IR imaging and spectroscopy over wide areas, which is a unique space capability. A wide-field near-IR space telescope could provide the following:

1. Rest-frame near-IR photometry for supernovae at $z < 1$, coordinated with ground-based optical observations;
2. Galaxy redshifts over the huge co-moving volume available at $1 < z < 2$, complementing ground-based BAO surveys at other redshifts; and

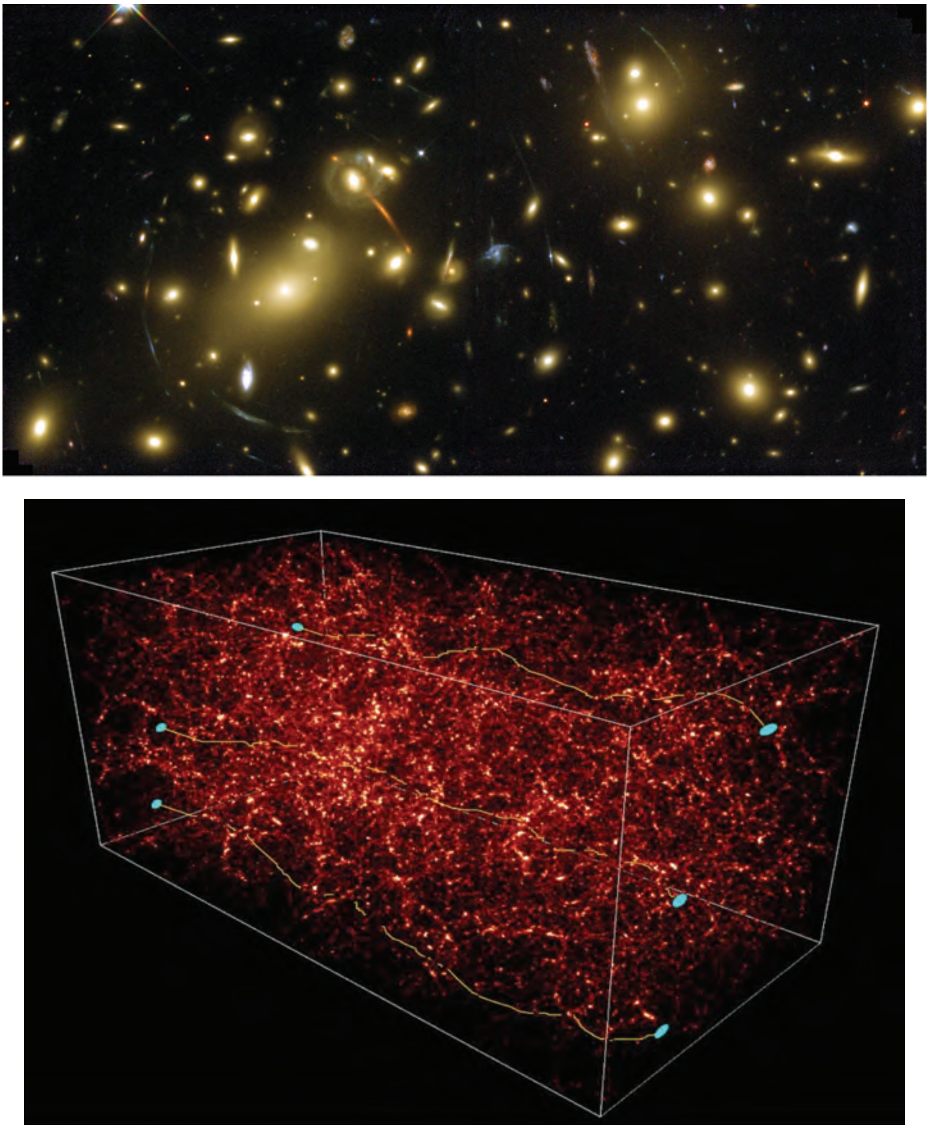


FIGURE 1.4 *Upper:* Gravitational lensing by the massive galaxy cluster Abell 2218, shown here in a Hubble Space Telescope image, stretches the images of background galaxies into arcs, elongated tangentially with respect to the center of the mass distribution. Far away from massive clusters, the image distortions are much weaker, but they can still be detected statistically because galaxies lensed by the same foreground mass will be stretched in the same direction. *Lower:* A simulation of light propagation through the distribution of dark matter in a large volume of the universe. The box shown spans a distance of about 1 billion light-years and is displayed so that brighter regions represent a higher density of dark matter. SOURCE: *Upper:* NASA, A. Fuchler, and the ERO Team (STSci, ST-ECF). *Lower:* Courtesy of S. Colombi and Y. Mellier (UPMC and CNRS, Institut d'Astrophysique de Paris).

3. IR photometry for the photometric redshifts of weak-lensing source galaxies, and IR spectroscopy of calibration samples. The combination of ground- and space-based shape determinations should reduce systematic uncertainties in the weak-lensing measurements.

Such a space mission would greatly enhance the precision of cosmic acceleration constraints and greatly reduce the associated systematic errors.

Large ground-based surveys could also benefit all of these methods. Large-area imaging surveys will provide optical photometry and independent shape measurements for weak-lensing and cluster studies. With proper choice of cadence, the same surveys can also identify supernovae at $z < 1$, where supernovae have been the most powerful tool for measuring distances. Ground-based optical spectroscopic surveys are already measuring BAO features at $z < 0.7$ and can push to higher-redshift windows. Measurements of redshifted 21-cm emission also have the potential of making powerful BAO measurements.

The combination of ground-based and space-based surveys should be able to improve measurements of both distance and growth rate of structure by an order of magnitude.

Precision Measurements and Cosmic Acceleration

Although there are many ideas about possible causes of cosmic acceleration, none of them is compelling. The current cosmological data are consistent with general relativity and a cosmological constant, which is a viable model for acceleration even though the magnitude of acceleration appears surprising. There are divergent opinions within this panel on the a priori likelihood of $w = -1$. If new measurements over the next decade improve the constraints on w by a factor of 10 and find no evidence for any deviation from general relativity, this would be important progress. However, it would be much less profound than detecting a signature of dynamical dark energy or a breakdown of general relativity.

Fortunately, the ground- and space-based experiments outlined in the previous subsection would generate rich data sets with broad scientific return, and achieving this return should be a strong consideration in designing observational programs, provided that the primary measurements are not impaired. In addition, low-cost but higher-risk experiments that explore a much wider range of new physics should be supported as a complement to the distance and structure-growth measurements.

Clues to the origin of acceleration could come from several directions, such as high-precision tests of general relativity and tests for time variation of fundamental "constants" such as the gravitational constant, the fine-structure parameter, the electron-to-proton mass ratio, and the speed of light. There are few clear targets for an "expected" level of deviations from conventional physics (and the same could be said for deviations from a cosmological constant). An appropriate metric for evaluating potential investigations might be the number of orders of magnitude of improvement that can be achieved for a given cost.

Panel Conclusions Regarding Inflation and Acceleration

As noted above, the requirements for testing inflation overlap those for investigating cosmic acceleration and so are presented together. The conclusions of the panel, including goals and needed capabilities in these areas, are summarized in Box 1.1.

CFP 3. WHAT IS DARK MATTER?

Dark matter is currently the flagship topic at the interface of cosmology and particle physics. It was first noticed in the 1930s that there must be more matter in galaxy clusters than the luminous matter could provide, and the evidence that accrued over the intervening decades showed that this dark matter must be non-baryonic. The existence of dark matter is now the strongest empirical evidence for physics beyond the standard model, and, as discussed below, the dark matter may well be tied to the physics of electroweak symmetry breaking, one of the central problems in particle physics today.

Over the past decade, galactic rotation curves, weak- and strong-lensing data, the hot gas in clusters, the Bullet Cluster and similar systems (see Figure 1.5), and measurements of large-scale structure and the Lyman-α forest have further constrained the distribution of dark matter and its properties. Dark matter must be long-lived and kinematically cold or warm (slow enough to seed structure formation); it must interact gravitationally but must not have strong interactions with itself or with baryonic matter. Dark matter is an essential ingredient in theories of structure formation, galaxy formation, and galactic dynamics. Efforts to detect and identify dark matter are therefore extremely important and, if successful, may one day lead to the field of "dark matter astronomy," in which the measured velocities of dark matter particles or spatial distribution of their annihilation products are used to probe the structure of the Milky Way and the physics of galaxy formation.

Although there is no definitive idea about what dark matter is, elementary particle theory provides a particularly compelling class of candidates—weakly interacting massive particles, or WIMPs—that are the target of a variety of experimental searches.

The WIMP Miracle

The origin of the symmetry breaking that provides particles with mass in the standard model of particle interactions remains a mystery, but every proposed

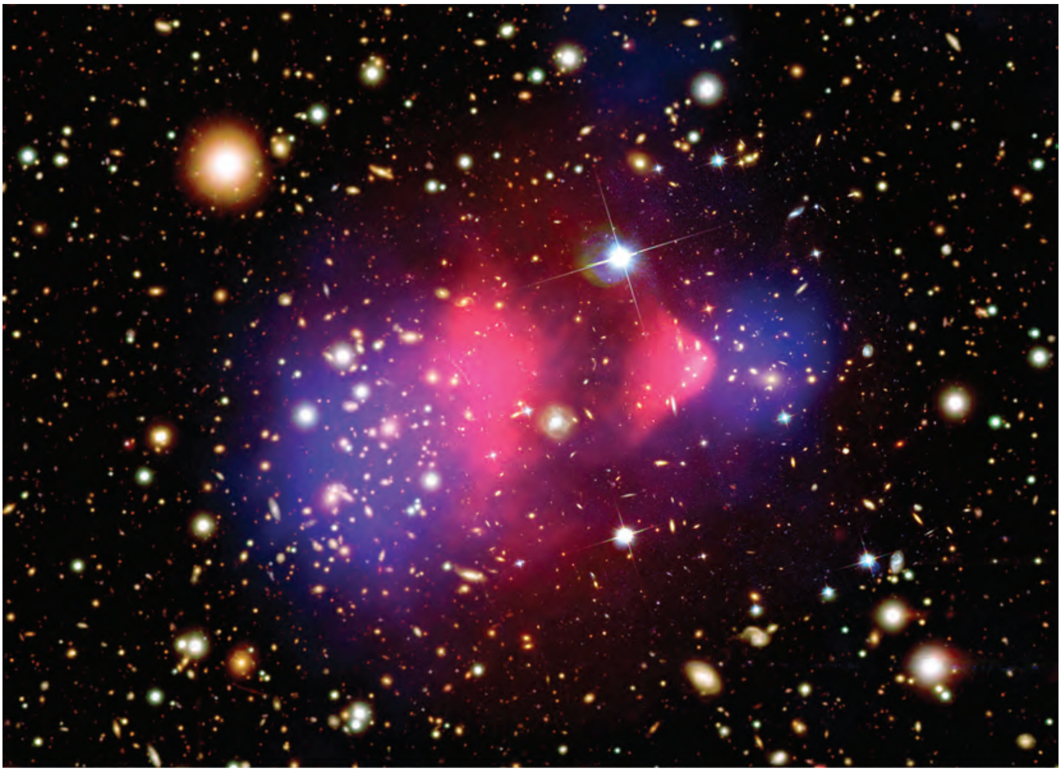


FIGURE 1.5 The Bullet Cluster displays particularly clear visual evidence for dark matter. In this collision of two clusters of galaxies, normal matter (hot gas, red, measured by X-ray emission) is distorted by the collision, but the rest of the cluster mass (dark matter, blue, measured by gravitational lensing) passes through unperturbed. SOURCE: *X-ray:* NASA/CXC/CfA/M. Markevitch et al. *Optical:* NASA/STScI; Magellan/University of Arizona/D. Clowe et al. *Lensing map:* NASA/STScI; ESO WFI; Magellan/University of Arizona/D. Clowe et al.

explanation requires new particles at the weak mass scale $m_W \sim 10$ GeV to a few TeV. Examples of the new weak-scale physics that could provide such particles include supersymmetry (in which every fermion in the standard model has a bosonic partner, and vice versa) and models with extra spatial dimensions. Such particles must have been in thermal equilibrium in the early universe until the $\chi\chi \rightarrow ff$ interactions (where χ is the WIMP and f a standard model particle) that convert them to ordinary particles became slower than the expansion timescale. If these particles are stable, then there will be relics from this early universe population. The relic density is inversely proportional to the annihilation cross section. Although there is no reason, a priori, to expect the electroweak scale to have anything to do with dark matter, a straightforward calculation shows that WIMP annihilation cross sections imply relic densities close to $\Omega_{DM} \sim 0.1$. Although there may be no connection between electroweak symmetry and dark matter, the remarkable coincidence between the mass scale required for electroweak symmetry breaking and the current dark matter density, the "WIMP miracle," provides strong motivation for a broad family of dark matter models. Such WIMPs would, moreover, be effectively collisionless and cold. Combined, these facts provide a very strong case that such particles are the dark matter.

Detection of Dark Matter

An important implication of the WIMP miracle is that if WIMPs make up the dark matter, then there are clear targets for experimental searches for WIMPs. As shown in Figure 1.6, the annihilation process $\chi\chi \rightarrow f\bar{f}$ that determines the relic density also implies that dark matter may be indirectly detected by searches for the products of dark matter annihilating now. This interaction also implies that dark matter may elastically scatter off standard model particles through $\chi f \rightarrow \chi f$ or be created at particle colliders through $f\bar{f} \rightarrow \chi\chi$.

Although such experimental efforts have been underway for a number of years, *the panel emphasizes that the conditions are now ripe for the detection of dark matter in the coming decade.* WIMPs may be produced directly at the Large Hadron Collider; direct searches for dark matter scattering in low-background experiments will dig deeply into the favored WIMP parameter space; and there are a variety of indirect astrophysical signatures of WIMPs that will be probed.

Dark Matter Searches at Particle Accelerators

Although accelerator searches are beyond the scope of the Astro2010 study, it is essential to note the synergy of the science goals of the Large Hadron Collider and the proposed International Linear Collider and those of dark matter searches.

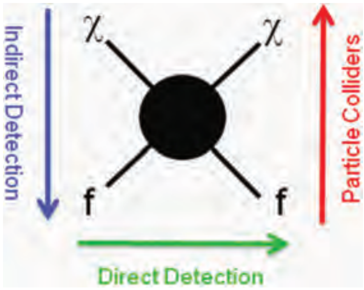


FIGURE 1.6 Three complementary approaches to searching for dark matter. Read from top to bottom, the diagram shows two dark matter particles χ annihilating to produce standard model particles *f*, which can be detected by indirect searches (e.g., for gamma rays or cosmic rays). Read from left to right, dark matter particles scatter off of standard model particles, producing signatures that are detectable in direct search experiments. Read from bottom to top, collisions of standard model particles can produce dark matter particles, interactions that can be detected in particle colliders like the LHC.

The LHC is very likely to identify the mechanism of electroweak symmetry breaking, thereby dramatically narrowing the range of possible dark matter candidates. The NRC's *Revealing the Hidden Nature of Space and Time*[5] discusses the essential role of particle accelerators in studying dark matter physics and emphasizes the link between the WIMP miracle and dark matter searches. Astronomical dark matter searches are essential to show that any particle produced in an accelerator is long-lived.

Detection of WIMPs in Underground Laboratories

The $\chi f \rightarrow \chi f$ reaction implies that WIMPs elastically scatter from atomic nuclei. Inelastic collisions are also possible if there are other dark particles with mass nearly identical to that of the dark matter. Direct-detection experiments use low-background detectors to search for the small recoil energy imparted to nuclei from such interactions. The WIMP-nucleus scattering may be either spin-independent (with the cross section depending on the nuclear mass) or spin-dependent (with the cross section depending on the nuclear spin).

Although the crossing symmetry illustrated in Figure 1.6 provides a relation between the relic abundance (determined by annihilation rates) and the elastic scattering cross section, particle physics models predict a range of elastic scattering cross sections over several orders of magnitude. The current bounds on spin-independent to WIMP-nucleon cross sections of $\sigma_{SI} \sim 10^{-42}$ to 10^{-43} cm^2, for WIMP masses in the range of 10 GeV to 1 TeV, are beginning to probe the parameter space of supersymmetric neutralinos and other WIMP candidates. In the next few years, cryogenic detectors and liquid noble gas detectors, located deep underground in extremely low background environments, are expected to be sensitive to $\sigma_{SI} \sim 10^{-44}$ to 10^{-45} cm^2. These experiments are exceptionally promising, both because they will improve present bounds so significantly and because the predictions for neutralinos in many models cluster around $\sigma_{SI} \sim 10^{-44}$ cm^2.

By the end of the coming decade, sensitivities extending to $\sigma_{SI} \sim 10^{-46}$ to 10^{-47} cm^2 should be possible with ton-scale detectors. Such experiments will probe even more deeply into WIMP parameter space or, if a signal is detected earlier, will provide detailed follow-up studies. Additional promising directions are low-threshold experiments extending sensitivities to lower WIMP masses, experiments that significantly improve limits on spin-dependent cross sections, and detectors that are sensitive to WIMP direction.

---

[5] National Research Council, *Revealing the Hidden Nature of Space and Time: Charting the Course for Elementary Particle Physics*, The National Academies Press, Washington, D.C., 2006.

Astronomical Detection of WIMPs

The reach of astronomical methods is potentially beyond that of direct detection and particle colliders, especially if the dark matter is very heavy, with mass approximately TeV or above, or if it couples dominantly to Higgs bosons or gauge bosons. In the coming decade a wide variety of experiments have the capacity to detect WIMP annihilation products by targeting a variety of final states.

Neutrinos

WIMPs from the galactic halo that pass through the Sun or Earth may scatter from nuclei. If they scatter to a velocity less than the escape velocity, they will be gravitationally captured. These WIMPs may then annihilate through the same processes as in the early universe, producing neutrinos, which may escape the Sun or Earth and be observed in terrestrial under-ice or underwater neutrino observatories. The Sun, typically the more promising target, is focused on here. The resulting neutrinos are distinctive because they come from the Sun but are far more energetic than the solar neutrinos produced by nuclear reactions.

It is emphasized that unlike all other indirect signals, the flux of neutrinos from the Sun is determined by scattering cross sections, and so these searches may be directly compared to direct-detection searches. This is because, given mild assumptions, the neutrino signal is determined by the WIMP capture rate, which is set by WIMP scattering cross sections. Given a particular WIMP candidate, it is straightforward to predict the neutrino flux, with no more astrophysical uncertainty than enters the prediction for the direct-detection rate.

For spin-dependent interactions, neutrino searches currently provide the leading constraints and are on the threshold of constraining regions of parameter space for neutralinos and other types of WIMP dark matter. For spin-independent interactions, neutrino searches are not competitive with direct searches for WIMP masses in the range of approximately 10 GeV to 1 TeV, but future water Čerenkov detectors and cubic-kilometer neutrino telescopes may be able to probe new territory at the low- and high-mass ends, respectively.

Gamma Rays

Gamma rays from dark matter annihilation in or near the center of the Milky Way Galaxy or in nearby dwarf galaxies may be detected by atmospheric Čerenkov telescopes or space-based observatories. The annihilation rates depend sensitively on the detailed dark matter distribution (the radial profile of the halo, substructure, etc.), which introduces considerable theoretical uncertainty.

The rates also depend on the annihilation cross section. For WIMPs, the correct thermal relic density is achieved if the thermally averaged annihilation cross section at freeze-out is $\langle\sigma_{\mathrm{ann}}v\rangle \sim 10^{-26}$ cm^3/s. This provides a rough benchmark to guide expectations. Furthermore, the annihilation rate $\langle\sigma_{\mathrm{ann}}v\rangle$ into observable decay products may differ significantly from the annihilation rate into all possible states in the early universe. At WIMP decoupling, $v \approx 0.3$ c, whereas the WIMP velocities in the galactic halo are a hundred times smaller. Clearly it is advantageous to achieve the lowest sensitivity limits possible to maximize the range of candidates that can be tested or discovered.

The resulting photons may be continuum photons from decays of an assortment of halo-annihilation products or monoenergetic photons from direct annihilation to gamma rays. The line signal is typically, although not always, significantly suppressed, but it is far more distinctive. Since photons point back to their source, directionality is a useful diagnostic for disentangling signal from background. In the case of detection, gamma-ray signals may be used to constrain dark matter distributions. Gamma rays from inverse Compton scattering in the inner galaxy may also constrain the distribution of electrons and positrons from WIMP halo annihilation.

Charged Particles

Indirect searches also target a variety of charged particles produced by annihilations in the galactic halo. As in the case of gamma rays, these rates depend sensitively on the distribution of dark matter and its annihilation cross sections. Charged-particle targets include the following:

- *Positrons from nearby halo annihilations.* Recent results have highlighted the need for a better understanding of astrophysical backgrounds. In some cases, however, direct annihilations $\chi\chi \rightarrow e^+e^-$ may yield a sharp feature in the energy spectrum that is not erased by propagation, and future searches may distinguish such signals from conventional astrophysical sources.
- *Antiprotons.* In general, antiproton backgrounds can obscure dark matter signals, but antiproton searches will nonetheless provide useful constraints on models for dark matter annihilation in the halo.
- *Antideuterons.* Because of the negligible background of antideuterons with kinetic energies below approximately 1 GeV, antideuterons can be a promising search target for dark matter that annihilates significantly to hadrons.

The dominant annihilation channel and the distinctiveness of the resulting signal depend on the dark matter candidate. In the absence of a single compelling candidate, it is of paramount importance to maintain a diverse program. However, given that charged particles do not point back to their source, and in view of the difficulty of distinguishing generic charged-particle signals from background, "smoking gun" signals that can be cleanly differentiated from background merit special attention. Precise measurements of various cosmic-ray elements over a wide energy range are also necessary to constrain cosmic-ray acceleration and propagation models and to determine the astrophysical background.

Alternative Candidates and Astrophysical Probes

WIMPs are not the only viable dark matter candidates. The axion is another interesting possibility. Motivated by the strong charge-parity problem in particle physics, the axion is extremely light and weakly coupled. There is no reason to expect the axion, unlike WIMPs, to have a relic density near Ω_{DM}. In simple models, however, the axion may be all or much of the dark matter for axion masses in the range approximately 1 to 100 meV. In the next decade, microwave cavity experiments are expected to probe much of this favored parameter range in canonical models, potentially establishing the existence of axion dark matter.

The WIMP and axion dark-matter candidates discussed so far are cold, collisionless, and stable. These are not required properties for dark matter, however, and there are viable candidates that are also motivated independently by particle physics but do not have one or more of these properties. Examples include sterile neutrinos; super-WIMPs (dark matter candidate particles with extremely small cross sections); dark matter candidates produced in decays during or after big bang nucleosynthesis; hidden-sector dark matter—that is, dark matter without standard model gauge interactions; and metastable dark matter with lifetime greater than the age of the universe.

These dark matter candidates may leave their imprint on astrophysical observables. For example, warm dark matter candidates, such as sterile neutrinos and super-WIMPs, may erase small-scale structure, with observable implications for galactic halo profiles and the abundance of dwarf galaxies. Hidden-sector dark matter may be significantly self-interacting and constrained by the Bullet Cluster and similar systems, or by the observation of elliptical halos. Dark matter that decays now may be detectable through cosmic rays or the diffuse photon spectrum. Studies in all of these areas are promising, as they may point to candidates beyond the standard cold dark matter paradigm.

Finally, as is evident from the discussion above, astrophysical inputs are also essential for interpreting WIMP searches. Direct searches and indirect searches for neutrinos are subject to uncertainties in the local dark matter density. Gamma-ray and charged-particle searches are subject to much larger uncertainties from halo profile parameters and structure on subgalactic scales. Precise astrophysical inputs will enhance what can be learned from these observations for many reasons—from setting the search strategies of atmospheric Čerenkov telescopes targeting dwarf galaxies to interpreting a confirmed signal in terms of particle physics parameters—and they will become essential given the expected progress in the coming decade.

For all of these reasons, astrophysical probes of the gravitational effects of dark matter are complementary to direct, indirect, and collider searches and may play an important role in constraining the nature of dark matter. In particular, theoretical, observational, and computational studies that improve constraints on small-scale structure and the local phase space distribution of dark matter are highly desirable.

Panel Conclusions Regarding Dark Matter

The conclusions of the panel with respect to dark matter are presented in Box 1.2. Included are goals and needed capabilities in this area.

CFP 4. WHAT ARE THE PROPERTIES OF NEUTRINOS?

Neutrinos are among the most abundant and most elusive particles in the universe. Created in enormous numbers by the big bang, they played an essential role in the synthesis of primordial light nuclei and limited the growth of galaxy-size density perturbations. Today, they are created in stars, supernovae, and cosmic-ray interactions, as well as on Earth in nuclear fission reactors, particle accelerators, and by Earth's natural radioactivity.

The new field of neutrino physics and astronomy has already discovered neutrino mass and neutrino oscillations. Soon after Raymond Davis, Jr., began his historic Homestake Mine experiment in 1968, it became clear that the flux of electron neutrinos coming from the Sun was well below the standard solar model prediction of John Bahcall and his collaborators. In fact, this experiment and follow-up results from other detectors established a pattern of solar neutrino fluxes that was incompatible with any plausible variation in that model. In 1998 neutrino oscillations, long suspected as a possible explanation for the missing solar neutrinos, were discovered by the Super-Kamiokande collaboration, which found a deficit in the muon neutrino flux produced when cosmic rays hit Earth's atmosphere. Then in 2001 and 2002, the Sudbury Neutrino Observatory showed that oscillations were also responsible for the missing solar neutrinos, directly measuring the flux of muon and tau neutrinos into which the solar electron neutrinos had oscillated. The demonstration of neutrino oscillations is a marvelous example of the use of astrophysics as a laboratory for fundamental physics.

In the coming decade, astrophysical and cosmological observations should resolve several more profound and so-far-unanswered questions about neutrino properties, including the sum of their absolute masses and the pattern of the individual masses. Astrophysical measurements are sensitive to values for the unknown

BOX 1.2
Conclusions Regarding Dark Matter by the Science Frontiers Panel on Cosmology and Fundamental Physics

Goals

- Probe both spin-independent and spin-dependent dark matter scattering cross sections with searches that explore much of the parameter space of weakly interacting massive particle (WIMP) candidates, both through underground experiments with a variety of techniques and through searches for dark matter annihilating to neutrinos. Although a review of laboratory dark matter detection methods is outside the panel's charge, progress in this area (as well as progress at the Large Hadron Collider) is critical for determining the properties of dark matter.
- Carry out indirect searches for dark matter that probe the annihilation cross sections of weakly interacting thermal relics. Identifying "smoking gun" signals is essential for detecting dark matter annihilation products above the astronomical backgrounds.
- Improve astrophysical constraints on the local dark matter density and on structure on subgalactic scales to test the paradigm of cold, collisionless, and stable dark matter and to look for evidence for alternative dark matter candidates. These astronomical observations, particularly of dwarf galaxies, help optimize dark matter search strategies and will be critical for determining the implications of dark matter signals for the particle properties of dark matter.

Needed Capabilities

- Direct-detection experiments that achieve sensitivities of 10^{-47} cm^2 or better for spin-independent WIMP-nucleon scattering cross sections.
- Indirect-detection experiments that target "smoking gun" signals: neutrinos from annihilation in the Sun, gamma-ray lines, antideuterons with kinetic energies of approximately GeV and below, and sharp features in the positron spectrum from $\chi\chi \rightarrow e^+e^-$.
- Measurements of gamma-ray and charged-particle spectra that sharpen constraints on astrophysical backgrounds to dark matter signals.
- Theoretical, observational, and computational studies that determine the local density of dark matter and structure on subgalactic scales.
- Axion searches that exhaustively scan the preferred masses and couplings of simple models.

mixing angle θ_{13} far below those that can be measured in other ways, and they may reveal the lepton number of the universe.

Neutrinos are also unique probes of the cosmos. Rapid advances in large and sophisticated neutrino detectors have created new opportunities for measuring the metallicity of the solar core, determining properties of neutron stars, and probing the most distant regions of the universe for nature's most energetic particle accelerators. These advances include massive ultra-clean low-energy detectors mounted a kilometer or more underground, cubic meter ice and water instruments sensitive to high-energy cosmic neutrinos, and radio and fluorescence detectors designed to measure ultrahigh-energy neutrinos.

Neutrino Mass

Neutrinos have a distinctive effect on cosmological evolution. Contributing approximately 10 percent of the energy density of the universe at the time of decoupling, they leave an imprint on the CMB. As neutrinos are light, they remain relativistic until late, and tend to erase the clustering of matter on small and intermediate scales. Thus, neutrinos produce a distinctive suppression of power in the matter (and hence galaxy) distribution.

The influence of neutrinos on cosmology is connected to a crucial property of neutrinos, the sum of the neutrino masses. Oscillation experiments, whether done with solar, atmospheric, or accelerator neutrinos, only measure the differences in the squares of the neutrino masses. Thus, the absolute scale of these masses is unknown. If that scale is as low as many theories suggest, it may be measurable only through cosmological observations, at least in the foreseeable future. Existing cosmological data and analyses place an upper bound of <0.7 eV on the sum of the masses of the three known neutrinos. Direct laboratory constraints on this mass sum (determined from the tritium beta-decay limit on the electron neutrino mass) have so far reached 6.6 eV, with improvements in the next decade possibly lowering this bound to 0.6 eV. But pushing beyond this level with terrestrial experiments could prove very difficult, requiring some fundamentally new idea.

Neutrino oscillations reveal that the mass could be as small as 0.05 eV. Cosmology is the only known tool for probing such tiny neutrino masses. Although neutrinos today may contribute only approximately 0.1 percent of the energy density of the universe, they still could suppress the power spectrum on galaxy scales by several percent, an effect that may be seen if measurements of the amplitude of the matter power spectrum are improved by an order of magnitude. Specifically, for a mass sum of 0.05 eV, the matter power spectrum is suppressed by 2.1 percent for wave numbers $k > 0.6$ (1/Mpc) at $z = 1.5$ and by 3.5 percent at $z = 0$. The suppression is scale-dependent, and the amplitude is predicted to be larger for larger masses. There are several promising approaches for measuring the small-scale matter power spectrum to this precision, including large galaxy redshift surveys, gravitational lensing, and Lyman-α absorption or 21-cm emission studies. For each method, there are challenges similar to those discussed in the subsection above titled “Precision Measurements and Cosmic Acceleration.” But the critical role of these measurements in both cosmology and neutrino astrophysics makes overcoming these challenges a very high priority.

A determination of the absolute scale of neutrino mass would be of fundamental importance to particle physics, potentially probing new physics phenomena far beyond the reach of accelerators. Models of neutrino mass and the existing results from solar and atmospheric neutrino oscillations suggest physics close to the GUT scale. There is a further exciting possibility: a demonstration that the sum of the neutrino masses is well below 0.1 eV would demonstrate that the mass pattern is the normal hierarchy, one of the two possibilities described in Figure 1.7.

Ultrahigh-Energy Cosmic Rays and Neutrinos

In 1966, Greisen, and separately Zatsepin and Kuzmin, noted that cosmic-ray protons or nuclei originating from distances greater than approximately 100 Mpc (depending on the composition) will react with the CMB photons, creating pions that decay into electrons and neutrinos, or photo-dissociating, degrading the energy and producing lower-energy secondary particles. Such interactions produce a cutoff in the cosmic-ray spectrum, as UHE protons and nuclei from distant sources will not be able to reach Earth. First the High Resolution Fly's Eye (HiRes) experiment in Utah and then, with better statistics in 2008, the Pierre Auger Observatory in Argentina appear to find the expected GZK cutoff (see Figure 1.8).

The results from these and other cosmic-ray observatories have stimulated great interest in UHE particle propagation. Some event simulators suggest that

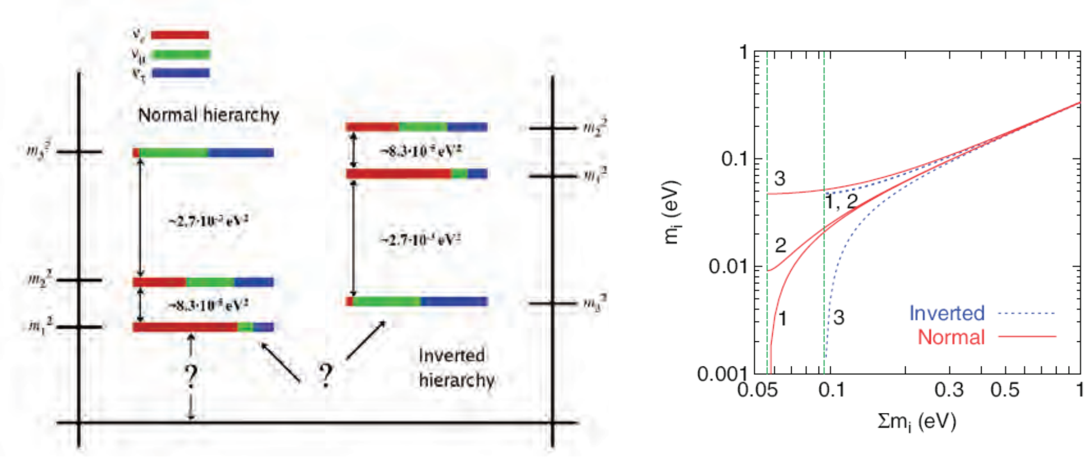


FIGURE 1.7 *Left:* The electron, muon, and tau neutrinos are each a mixture of three "mass states" m_1, m_2, and m_3. Scientists know two differences of the squares of the masses, indicated by the vertical arrows. Yet we do not know the individual masses (suggested by the left "?"). The ordering where m_3 is the most massive is the normal hierarchy on the left, while the inverted hierarchy is on the right. *Right:* Masses of the neutrino "mass states" 1, 2, and 3 as a function of the sum of the masses that we know are between 0.05 V and 0.7 eV. If measurements show that the sum of the masses is <0.1 eV, we will know that the hierarchy is normal and m_3 is near 0.05 eV. SOURCE: *Left:* F.B. Abdalla and S. Rawlings, Determining neutrino properties using future galaxy redshift surveys, *Monthly Notices of the Royal Astronomical Society* 381(4):1313-1328, copyright Royal Astronomical Society, 2007. *Right:* Reprinted from J. Lesgourgues and S. Pastor, Massive neutrinos and cosmology, *Physics Reports* 429(6):307-379, copyright 2006, with permission from Elsevier.

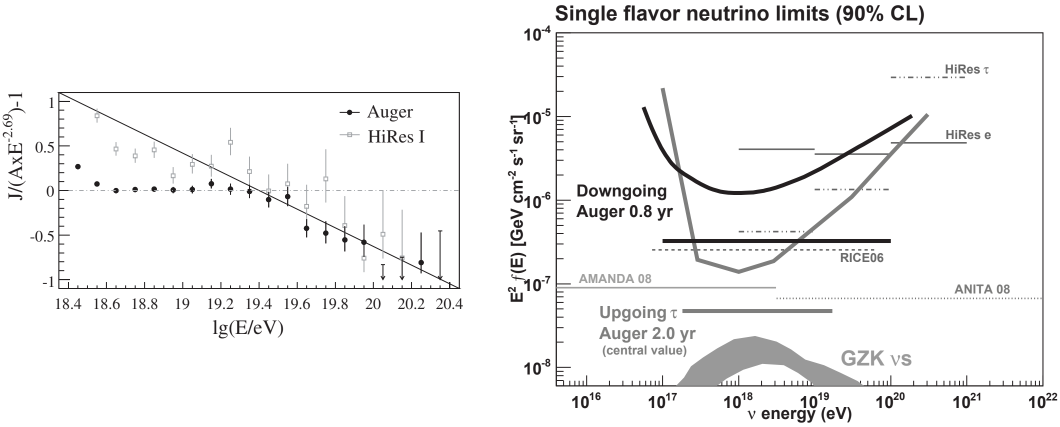


FIGURE 1.8 *Left:* The flux of ultrahigh-energy cosmic rays recorded by the High Resolution Fly's Eye (HiRes) experiment (atmospheric fluorescence detectors in Utah) and the Pierre Auger Observatory (water Čerenkov and atmospheric fluorescence detectors in Argentina), both showing the decline above 4×10^{19} eV—resembling the Greisen-Zatsepin-Kuzmin cutoff. *Right:* Limits on the flux of ultrahigh-energy neutrinos from HiRes, Auger, Radio Ice Čerenkov Experiment (RICE; Askaryan effect in ice), Antarctic Muon and Neutrino Detector (AMANDA), and IceCube (in South Pole ice), and Antarctic Impulsive Transient Antenna (ANITA; Askaryan effect from long-duration balloon flights over the South Pole). SOURCE: *Left:* Courtesy of the Pierre Auger Observatory. Figure reprinted with permission from J. Abraham, G. Snow, and the Pierre Auger Collaboration, Observation of the suppression of the flux of cosmic rays above 4×10^{19} eV, *Physical Review Letters* 101:061101, 2008, copyright 2008 by the American Physical Society, available at http://link.aps.org/doi/10.1103/PhysRevLett.101.061101. *Right:* Courtesy of the Pierre Auger Observatory.

events observed at lower energies (10^{15} to 10^{16} eV) are protons, while there is a trend toward heavy nuclei at intermediate energies that may persist in part to the GZK cutoff. This interpretation depends in part on the modeling of proton-nucleus and nucleus-nucleus reactions at ultrarelativistic energies that extend beyond the center-of-mass energies that can be tested in collisions at Brookhaven National Laboratory's Relativistic Heavy Ion Collider or at the European Organization for Nuclear Research (CERN). High-quality data at the GZK cutoff could be particularly helpful in reducing such uncertainties, as the location and even precise shape of this cutoff depend on composition: nuclei and nucleons of the same energy travel at different velocities and thus experience different interactions with the CMB. For this reason new UHE cosmic-ray observatories that can provide additional data on the GZK cutoff are very well motivated.

The detection of GZK neutrinos would complement observatories such as Pierre Auger, as the flux of neutrinos associated with near-GZK hadronic events is a function of the energy distribution and composition of the hadronic events.

Because neutrinos rarely interact with the CMB or other matter and fields and because they point back to their sources, neutrinos also provide a rare opportunity to determine what the ultimate energy limits of cosmological accelerators might be.

Novel techniques are being developed to detect UHE neutrinos. One possibility for extending measurements to and beyond GZK energies is the *Askaryan effect*, the coherent emission of a radio pulse by electrons swept along by the UHE neutrino shower front. The Antarctic Impulse Transient Array (ANITA) balloon-borne antenna has demonstrated the method over Antarctica in 2006-2009, though so far it has not reported a neutrino with GZK energies, while the Radio Ice Čerenkov Experiment (RICE) explored the effect with antennas in the ice. From the detection of the GZK cutoff in the cosmic-ray protons, one can estimate a baseline level for the UHE neutrino flux (see Figure 1.8). An exciting possibility would be the development of a radio antenna array in the Antarctic ice or elsewhere capable of detecting this flux. This might begin with a coverage of 100 km^2 and could be extended well beyond this scale potentially to detect many tens of GZK neutrinos a year. Scientists expect to learn significant astrophysics and physics from UHE neutrinos, including verification of the origin of the cutoff and determination of the origin and evolution with redshift of sources beyond the GZK cutoff.

It is fortunate, in the quest to understand cosmic rays and to exploit them as a probe of new astrophysical sources (and potentially new fundamental physics), that the field has multiple probes, including nucleons and nuclei, neutrinos, and gamma rays. There is a good understanding of the common mechanisms affecting the production and propagation of such "messengers," and thus multiple opportunities for constraining the properties of their astrophysical sources. Neutrinos, unique as probes of sources at cosmological distances and arbitrary energies, present an exciting frontier where further developments of instrumentation could be rewarded by major discoveries.

Supernova Neutrinos

The approximately 20 neutrinos detected from Supernova 1987A marked the dramatic opening of extrasolar neutrino astronomy. Today, deep-underground neutrino detectors are an order-of-magnitude larger than those operating in 1987, with much improved cleanliness and lower triggering thresholds. A galactic supernova 10 kpc from Earth would produce roughly 10,000 events in Super-Kamiokande's 50-kiloton volume of water and a similar event rate in NOvA, a 14-kiloton liquid scintillator detector under construction in Minnesota. On the order of 100,000 events might be seen in future megaton-scale experiments.

A precise measurement of the supernova neutrino "light curve" out to long times would provide a great deal of information about the supernova mechanism and fundamental neutrino properties. These neutrinos are thought to be essential to the explosion, transporting energy and driving the convection responsible for mantle ejection. They also create the explosive, neutron-rich stellar environments in which about half of our galaxy's heavy elements may have been synthesized—the "star stuff" that is crucial to the evolution of complex structures such as the planets of our solar system and the life that they sustain. The neutrino flux, originating from deep within the supernova, provides one of the best tests of the theoretical understanding of the core-collapse mechanism, long considered one of the major high-performance-computing challenges in theoretical astrophysics.

The neutrino flux provides a very accurate measurement of the gravitational energy released in core collapse, and it also marks the onset of core collapse, providing a "clock" against which gravitational wave and optical signals can be compared. Changes in the late-time neutrino cooling curve could signal the onset of phase transitions at supernuclear densities; a sudden termination would accompany black hole formation. The correlation between neutrino energies and their flavors is a powerful diagnostic of neutrino oscillations. In particular, supernova neutrinos are sensitive to values of the unknown mixing angle θ_{13} as small as 10^{-4}, two orders of magnitude beyond the goals of reactor and accelerator experiments currently planned. Exotic "matter effects" connected with oscillations in an intense neutrino background could produce distinctive changes in the neutrino flux that depend on the neutrino mass hierarchy (normal or inverted).

Bursts from individual supernovae are rare, but the flux of neutrinos from all past supernovae is continuous and potentially measurable. The detection of this flux would determine an integral over the redshifted neutrino emission of all core-collapse supernovae, from the time of the first stars until now. This would constrain the mass cut for core-collapse supernovae and potentially provide some information on the redshift evolution of these massive stars. Future megaton detectors could record dozens of events per year from this source, in the energy region above the solar neutrino end point.

Light-Element Abundances

Five light nuclei (H, D, ^{3}He, ^{4}He, and ^{7}Li) are fossils from the first few minutes of the big bang. Their abundances are direct probes of physical conditions in the very early universe. As CMB measurements now independently fix the baryon-to-photon ratio, big bang nucleosynthesis predictions for the primordial abundances are largely parameter-free in the context of standard-model physics, subject only to the uncertainties in the input parameters (e.g., the baryon-to-photon ratio and nuclear cross sections).

The agreement between the observed and predicted D/H ratio is a major pillar of the big bang model. Although there is fair agreement for the other nuclei, the precision of the measurements lags behind that of the theory, a situation that should be improved in order to more fully test understanding of the big bang. In detail, there have been persistent discrepancies, such as measurements of the abundance of ^{4}He that are typically below the modern prediction. There are plausible explanations for this discrepancy, including helium flux that was missed in absorption lines and inaccurate atomic data. These possibilities need to be thoroughly explored.

The observed abundance of ^{7}Li in old halo stars is constant to within measurement errors of 5 percent in stars with a wide range of metal abundances and masses, yet the amount of ^{7}Li is a factor of four below BBN predictions. An astrophysical explanation of this anomaly would have to produce this large and near-constant reduction. The exciting alternative is that the ^{4}He or ^{7}Li anomalies are the signature of some new physics beyond standard BBN.

Plausible sources of such new physics exist. Measurements of light-element abundances are sensitive to neutrino properties. The addition of extra neutrino species increases the universe's expansion rate, leading to more ^{4}He and less D. An excess of neutrinos over antineutrinos (a lepton number asymmetry) decreases the n/p ratio, leading to less ^{4}He and less D. New particles could be involved, such as sterile neutrinos with ultraweak interactions and/or dark matter that decays to change the abundances after BBN.

Improved measurements are required to test such possibilities. Much remains to be done with existing 10-m-class optical telescopes for deuterium, ^{7}Li, and especially ^{4}He, although in other cases observations already stretch the limits of such telescopes. This is especially true for ^{6}Li, which is more fragile than ^{7}Li and hence limits the amount of ^{7}Li that has been destroyed in stellar atmospheres. A telescope in the 30-m range with a high-resolution stable spectrograph, should lead to dramatic improvements in the measurements of the light nuclei. Improved accuracy is also needed for several nuclear reactions at low (MeV) energies and for the atomic rate coefficients that are used to determine the ^{4}He abundance in H II regions. The measurements of the abundance of ^{6}Li in halo stars require improved three-dimensional stellar atmosphere models, whereas interpretation of ^{7}Li measurements would benefit from stellar models that incorporate more physical treatments of turbulence.

Panel Conclusions Regarding Neutrinos

The conclusions of the panel with respect to neutrinos are presented in Box 1.3. Included are goals and needed capabilities in this area.

BOX 1.3
Conclusions Regarding Neutrinos by the Science Frontiers Panel on Cosmology and Fundamental Physics

Goals

- Develop the sensitivity to detect and study ultrahigh-energy (UHE) neutrinos expected if the cosmic-ray energy cutoff is due to protons annihilating into neutrinos and other particles. The detection of UHE neutrino fluxes above those expected from the Greisen-Zatsepin-Kuzmin (GZK) mechanism would be the signature of new acceleration processes.
- Measure the neutrino mass to a level of 0.05 eV, the lower limit implied by current neutrino mixing measurements, through its effects on the growth of structure.
- Enable precision measurement of the multiflavor neutrino "light curves" from a galactic supernova.
- Improve measurements of light-element abundances in combination with big bang nucleosynthesis theory to test neutrino properties and dark matter models.

Needed Capabilities

- Measurements of small-scale structure from dwarf-galaxy dynamics, gravitational lensing, and the Lyman-α forest.
- Precision measurements of the power-spectrum amplitude using a combination of cosmic microwave background lensing, weak lensing, the galaxy power spectrum, and measurements of neutral hydrogen fluctuations.
- Neutrino detectors that can measure the time, energy, and flavor distribution of neutrinos from a nearby supernova and detect the integrated supernova neutrino background.
- Radio-frequency experiments for UHE neutrinos, with sensitivity to detect the expected events associated with the proton GZK cutoff.
- Improved characterization of the energy spectrum and sources of cosmic rays near the GZK cutoff.
- Improved measurements of light-element abundances in stellar atmospheres and the interstellar and intergalactic medium, principally through high-resolution spectroscopy on approximately 30-m telescopes.

CFP DISCOVERY AREA—GRAVITATIONAL WAVE ASTRONOMY: LISTENING TO THE UNIVERSE

In the past century, our ability to view the universe expanded to encompass a vast sweep of the electromagnetic spectrum from gamma rays to radio waves, bringing with it the discovery of many unexpected phenomena. In the coming decade, some of the most exciting discoveries may come from opening a new observational window with the first direct detections of gravitational waves.

In the same way that the sense of hearing complements the sense of sight, gravitational wave observations complement and enrich what can be learned electromagnetically. Gravitational waves are produced by the bulk motions of matter, and they propagate essentially unabsorbed through even the densest material to convey information about the overall dynamics of the source. In contrast, electromagnetic waves tell only about the thermal and magnetic environment of the gas that surrounds a source, and they can be bent or absorbed along their propagation paths to telescopes.

However, the weak coupling to matter that allows gravitational waves to travel unimpeded also makes them very hard to detect: the merger of two stellar remnant black holes at 10 Mpc would bathe Earth with a peak energy flux exceeding 10 percent of the solar constant, but so little of this energy is captured that the mirrors of a kilometer-scale detector are displaced by less than 10 percent of the width of a proton. Compelling though indirect evidence for their existence can be seen in the orbital decay of binary pulsar systems, whose discovery earned Hulse and Taylor the 1993 Nobel Prize in physics. Efforts at direct detection initially employed massive bar detectors fitted with vibration sensors but are now focused on laser interferometers, which use interference of laser beams to detect the minute motions of mirrors suspended at the ends of kilometers-long evacuated cavities. A worldwide network of terrestrial interferometers is currently in operation, covering the frequency range 10 to 1,000 Hz with the sensitivity to detect relative displacements a thousand times smaller than the width of a proton. Operating in the much lower nanohertz (10^{-9} Hz) frequency range are pulsar-timing arrays, which seek to detect gravitational waves by the delays that the waves impart on the arrival times of pulses from radio pulsars. The low-frequency range, between 10^{-5} and 10^{-1} Hz, is believed to be rich in gravitational wave sources of strong interest for astronomy, cosmology, and fundamental physics. Because terrestrial sources of noise dominate at such low frequencies, this portion of the gravitational wave spectrum can be accessed *only* from space.

Gravitational Wave Astrophysics

All astrophysical objects emit gravitational radiation at some level, but the extreme stiffness of space-time implies that only systems that pack a large amount of rapidly moving material into a small volume will emit detectable signals. As a general rule, the more massive systems radiate deeper and louder signals. Accordingly, for example, high-frequency ground-based interferometers look for black hole mergers with total mass of 1 to 10^3 solar masses, whereas low-frequency space-based systems will look for mergers with total mass of 10^3 to 10^7 solar masses (Figure 1.9).

Among the anticipated results from ground-based detectors are measurements of the merger rates of binary systems containing two black holes or two neutron stars (or one of each), measurements of the deviations from spherically symmetric

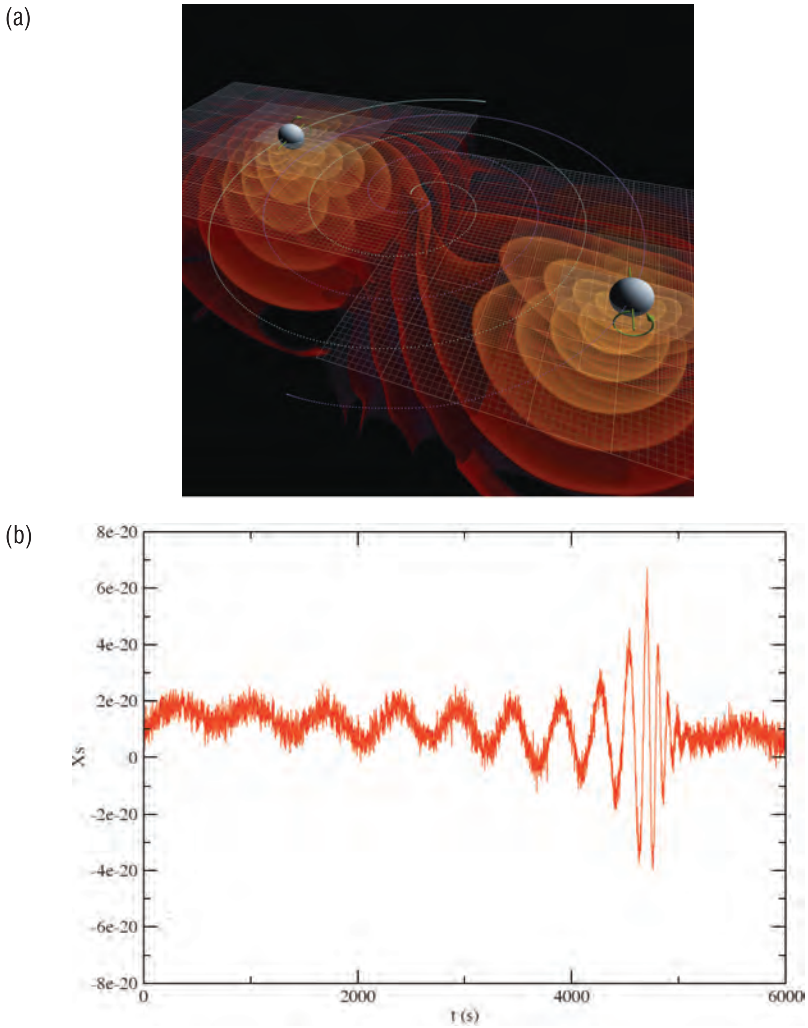


FIGURE 1.9 (a) Recent advances in numerical relativity make it possible to simulate accurately the ripples in the curvature of space-time stirred up by the merger of two black holes. The top image shows two spinning black holes that are spiraling together toward merger. The black hole horizons (gray surfaces) and spin directions (green arrows) are shown. A cross section of the adaptive mesh refinement computational grid is shown, with a cut-away view of the strong dynamical gravitational fields underneath. (b) The simulated response to the late inspiral and merger of two 10^5 $M_\odot$ black holes at $z = 15$ is superimposed on the galactic foreground from white-dwarf binaries and the instrument noise from a space detector. In contrast to the ground-based detectors, these signals will be clearly visible in the raw data, with signal-to-noise ratios in the hundreds of thousands. The noise in the plot is the unresolved gravitational radiation from the ordinary binary stellar systems in our galaxy.
SOURCE: (a) Image courtesy of Chris Henze, NASA Ames Research Center. (b) Figure reprinted with permission from J. Baker, S.T. McWilliams, J.R. van Meter, J. Centrella, D.-I. Choi, B.J. Kelly, and M. Koppitz, Binary black hole late inspiral: Simulations for gravitational wave observations, *Physical Review (Section) D: Particles and Fields* 75:124024, 2007, copyright 2007 by the American Physical Society.

or pure dipole signals[6] that can be supported by millisecond pulsars, constraints on the equation of state of matter at nuclear densities through observations of tidally distorted neutron star mergers, and exploration of the dynamics of nearby core-collapse supernovae. By mid-decade it should known if compact binary mergers are indeed responsible for short-hard gamma-ray bursts, and if gravitational wave emission is indeed balancing the accretion torque in low-mass X-ray binaries. Even the most pessimistic estimates of rates and strengths of signals predict a detection by the advanced ground-based interferometers scheduled to be operational by 2014. Beyond the anticipated discoveries, the opening of a new observational window will likely produce surprising, and perhaps even revolutionary, results.

Pulsar-timing arrays have the potential to detect a stochastic background from the slow inspiral of supermassive black hole binaries (10^8 to 10^{10} solar masses) in the frequency range 10^{-9} to 10^{-6} Hz, and perhaps to resolve a few of the brighter systems. With modest enhancements to the existing arrays, a positive detection is likely by the latter part of the decade. Such measurements will fix the merger rate of supermassive black holes and provide unique constraints on models of black hole growth.

In the low-frequency (10^{-5} to 10^{-1} Hz) portion of the spectrum, accessible only from space, there are likely to be so many sources that a background of waves from weaker systems will be a dominant source of noise. Above this noise it will still be possible to detect black hole mergers in the mass range 3×10^3 to 10^7 solar masses out to redshift 20 or greater. These observations will reveal the masses and spins of the black holes and will indicate the merger rate as a function of distance. For example, with typical sensitivity parameters for a space interferometer, including noise from foreground sources, for two $10^6\ M_\odot$ nonspinning black holes merging at $z = 10$, the total mass of the system could be measured to 0.1 percent, and at $z = 1$, to 0.001 percent.

This information can be used to test theories of how galaxies and black holes coevolve, and to determine the relative importance of gas accretion and mergers in massive black hole growth. If the basic ideas of massive black hole growth are qualitatively correct, tens to hundreds of events per year for inspirals at the high-mass end may be detectable. For inspirals at the low-mass end, the rates are highly uncertain. Observations of low-redshift systems can be used to confirm the existence of intermediate-mass (500 to 10^4 solar mass) black holes and to probe their properties.

Also visible out to $z \approx 2$ will be the capture of stellar remnants (mostly black holes and neutron stars) by supermassive black holes in galactic nuclei. Observations

---

[6] Only nonsymmetric distributions of mass emit gravitational waves. Spherically symmetric distributions of charge or mass emit neither electromagnetic nor gravitational waves. Distributions of charge with a dipolar symmetry emit electromagnetic waves; however, there are no dipolar gravitational waves.

of these "extreme mass ratio inspirals" (EMRIs) will allow for precision measurements of the mass and spin of the central black holes, and the rate of these mergers as a function of eccentricity can be used to constrain models of galactic cores.

The most common signals detected from space, however, will be from binary systems of white dwarfs in our own galaxy, and tens of thousands of these systems are expected to be individually resolved. The masses and positions of all short-period binaries (periods less than about 11 minutes) in our galaxy will be measured, providing unique constraints on population-synthesis models and providing a new window for the study of white-dwarf interiors through tidally induced oscillations. Joint electromagnetic/gravitational wave observations of white-dwarf binaries can be used to constrain the mass of the graviton and the polarization pattern of gravitational waves. Once again, opening up several decades of a new observational spectrum is fertile ground for surprises, from the exotic (for example, cosmic superstrings) to others that are wholly unanticipated.

Looking for New Physics with Gravitational Waves

The gravitational waves in the universe today preserve a record of all macroscopic mass-energy flows over the entire history of the universe. They can be used to probe aspects of new physics never before explored. The possibilities include first-order phase transitions leading to bubble nucleation and collision, the dynamics of extra spatial dimensions, inflationary reheating, and a writhing network of cosmic (super)strings. The radiation from these and other exotic processes occurring in the early universe when the temperature was 0.1 to 1,000 TeV will have been redshifted to the frequency range explored by a space-based instrument. This nice coincidence means that gravitational waves have the potential to explore weak-scale physics.

Binary black hole mergers will provide stringent tests of general relativity (Figure 1.10). These systems are "simple," consisting of pure space-time curvature, while their strong signals, even when emitted from cosmological distances, can dominate over noise in space-based measurements. Because the final inspiral and merger of two compact bodies are dominated by their mutual gravity, the orbit and gravitational wave signal will reflect strong-field, dynamical, curved space-time general relativity in its full glory. Detailed comparisons between the measured waveforms and theoretical waveforms calculated from combinations of analytical and numerical solutions of Einstein's equations will give a rich variety of tests of the theory in a regime that has hitherto been inaccessible to experiment or observation.

Also detectable will be "ringdown" waves, emitted by the distorted black hole produced by the merger as it settles down to a stationary state. These waves have discrete frequencies and damping rates that depend on the mass and spin of the hole. Thus, measurements of the ringdown will test whether geometry obeys the no-hair theorems predicted by general relativity.

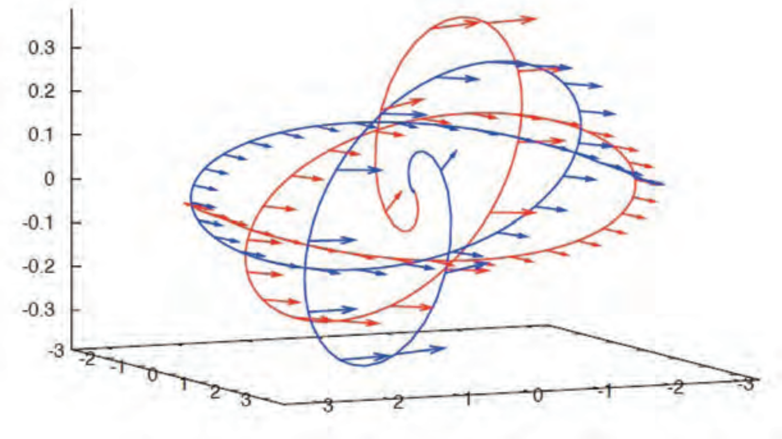


FIGURE 1.10 Simulations of black hole mergers suggest that spin may play a critical role in how black holes merge; depending on the magnitude and alignment of the spins, the mergers could be very rapid or could experience a momentary "hang-up," with significant consequences for the observed waveform. Shown in this figure are the final few orbits of two black holes (one in red, the other in blue), with arrows denoting their spin magnitude and direction. These complex spin-induced orbital effects are the consequences of "frame dragging," a fundamental prediction of Einstein's theory that has been probed in the solar system using Gravity Probe B, Laser Geodynamics (LAGEOS) satellites, and lunar laser ranging, and has been hinted at in observations of accretion onto neutron stars and black holes. Observing the effects of frame dragging in such an extreme environment would be a stunning test of general relativity. Furthermore, with spinning progenitors, the final black hole could experience substantial recoil resulting from the emission of linear momentum in the gravitational waves, large enough to eject it from the host galaxy. SOURCE: Reprinted with permission from M. Campanelli, C.O. Lousto, and Y. Zlochower, Spinning-black-hole binaries: The orbital hang-up, *Physical Review (Section) D: Particles and Fluids* 74:041501, 2006, copyright 2006 by the American Physical Society.

EMRIs may be observed by a space-based detector and can provide incredibly precise quantitative tests of the space-time geometry of black holes. Over the 10^4 to 10^5 eccentric, precessing orbits traced out by the smaller mass, the emitted waves encode details about the space-time structure of the larger hole with a variety of distinct signatures. In addition to providing determinations of the black hole's mass and angular momentum to fractions of a percent, the observations can also be used to test whether the space-time that encodes the waves is the unique Kerr geometry that general relativity predicts for rotating black holes.

Gravitational waves can also be used to test specific theories alternative to general relativity. If gravity propagates with a speed that depends on wavelength (for example if the putative "graviton" were massive), a strong constraint on its mass could be placed by searching for deviations in the phases of the arriving waves from the general-relativistic predictions. It may be possible to test whether the prediction of only two transverse quadrupolar modes is correct. If astronomers can detect the electromagnetic signature associated with the inspiral, then differences in arrival times would place powerful constraints on the graviton mass.

Binary black hole inspirals provide standard candles free of calibrations for measuring the distance to the source. For these "clean" systems, there is an expected relationship between the waveform shape and the luminosity of the source. A space-based detector could measure luminosity distances to a few percent at redshift 2, and to tens of percent at $z = 10$. At the same time, because of the changing orientation of the space array with respect to the source, its sky location could be determined to 10 arcminutes for massive inspirals at $z = 1$. This positional information can be used to search for electromagnetic counterparts, which can in turn be used to measure the redshift. The combination of several such measurements could give a dark energy bound that begins to be competitive with conventional dark energy approaches. Absent an electromagnetic counterpart, statistical associations of galaxies with EMRIs can be used to measure the Hubble constant to a few percent, and to place constraints on the acceleration rate.

Realizing the Discovery Potential

A number of steps are needed to realize the discovery potential of gravitational wave astronomy. The first is to complete the upgrade of the network of ground-based, high-frequency interferometers in the United States and Europe. If these arrays achieve their proposed sensitivities of roughly $4 \times 10^{-24}/(\mathrm{Hz})^{1/2}$ at 100 Hz, there will be a very good chance of discovery of gravitational waves from stellar remnant inspirals and mergers. Development of pulsar timing arrays with the capability of detecting and timing approximately 40 millisecond pulsars with 100-nanosecond accuracy should be completed. Initially, the primary need is increased time allocations at existing facilities, with advanced kilometer-scale detector arrays envisioned for the future.

For the low-frequency band between 10^{-5} and 10^{-1} Hz, a space-based detector is essential, with sensitivity capable of detecting massive black hole mergers to redshift 20 with a signal-to-noise ratio of at least 10. This requirement translates roughly to a strain sensitivity of $3 \times 10^{-21}/(\mathrm{Hz})^{1/2}$ in the millihertz range.

The science to be learned from gravitational wave detections will be greatly enhanced if observations of the same phenomena can be done in the electromagnetic (and possibly the neutrino) window. Because many gravitational wave sources are transient in nature (mergers, collapse), this will require wide-angle, high-cadence electromagnetic surveys, together with potentially rapid follow-up observations of sources with large telescopes. Observing electromagnetic counterparts will aid in source identification, sky localization, and redshift determination, and will make possible novel measurements of the distance-redshift relation. For massive black hole inspirals, a space detector could give weeks of warning of the final merger

BOX 1.4
Conclusions Relating to Gravitational Waves by the Science Frontiers Panel on Cosmology and Fundamental Physics

Goals

- Detect gravitational waves from mergers of neutron stars and stellar mass black holes.
- Detect gravitational waves from inspirals and mergers of supermassive black holes at cosmological distances.
- Achieve high signal-to-noise ratio measurements of black hole mergers to test general relativity in the strong-field, highly dynamical regime.
- Identify electromagnetic counterparts to gravitational-wave sources.
- Open a radically new window on the universe, with the potential to reveal new phenomena in stellar-scale astrophysics, early-universe physics, or other unanticipated areas.

Needed Capabilities

- A space-based gravitational wave interferometer probing the 10^{-5} to 10^{-1} Hz frequency range to the sensitivity limits imposed by astrophysical "foreground" noise from galactic binaries.
- Ground-based interferometers probing the 10- to 1,000-Hz range with the sensitivity to detect neutron star mergers at 300-Mpc distances.
- Pulsar-timing arrays probing the nanohertz frequency range with the sensitivity to detect the stochastic background from supermassive black hole binaries.
- Time-domain electromagnetic facilities with the sensitivity, speed, and flexibility needed to find the counterparts of gravitational wave sources.

event together with degree-accurate source positions, permitting narrower fields to be viewed.

Panel Conclusions Regarding Gravitational Waves

The conclusions of the panel with respect to gravitational waves are presented in Box 1.4. Included are goals and needed capabilities in this area.

THEORY AND SYNTHESIS

Theory is and has been essential to what we choose to observe and how we arrange to do so. Many of the ideas that are central to the next decade's empirical investigations—inflation, supersymmetric dark matter, neutrino oscillations, black holes, and gravitational waves—began life decades ago as theoretical speculations. Many of the tools that are being used for these investigations—CMB polarization, weak lensing, BAO observations, laser interferometers—grew out of theoretical studies that started long before the methods were technically feasible.

As these methods are moving toward implementation, theory plays an important role in designing experiments, optimizing methods of signal extraction, and understanding and mitigating systematic errors. Theory and observation are so closely intertwined in investigations of cosmology and fundamental physics that it is often difficult to define the border between them. Examples include predicting the signals and backgrounds for dark-matter-detection observations, calculating the impact of intrinsic galaxy alignments on weak-lensing measurements, and computing waveforms for template matching in gravitational wave searches. Theoretical advances often amplify the scientific return of a data set or experiment well beyond its initial design. Potentially powerful techniques that are subjects of active theoretical scrutiny include redshift-space distortions as a precision measure of structure growth, and scale-dependent galaxy bias as a sensitive probe of primordial non-Gaussianity. As data come in, theory assumes the pivotal role in tying them back to the underlying physics, whether it be computer models of core-collapse supernovae, phase transitions in the early universe, or extensions of the standard model of particle physics. Finally, more speculative, exploratory theory may produce the breakthrough that leads to a natural explanation of cosmic acceleration, a compelling physical mechanism for inflation, or the prediction of an extraordinary gravitational wave phenomenon yet to be observed.

The above examples illustrate four distinct modes of theoretical work that are essential to progress in the next decade:

1. *Before observation.* Development of new methods, identification of new observables, and statistical forecasting.
2. *During observation.* Design of experiments, calculation of systematic effects, and statistical analysis to optimize the use of the data.
3. *After observation.* Interpretation of empirical results in terms of underlying physical models.
4. *Exploratory theory at the frontiers of current knowledge.* Although often speculative and high risk, this mode of theoretical research can lead to breakthrough ideas that transform the field.

Advances in high-performance computing are driving rapid progress in many areas of cosmology and fundamental physics. Examples that are central to the themes of this report include numerical simulations of structure formation needed to interpret maps of the galaxy distribution or to predict signals of dark matter annihilation; computational studies of core collapse and thermonuclear supernovae; calculations of gravitational wave emission from mergers of spinning black holes; statistical analyses of large and complex data sets from CMB observations, LSS surveys, neutrino observations, and gravitational wave searches; and massive searches through high-dimensional parameter spaces to evaluate the statistical uncertainties from varied combinations of data sets. To exploit these advances requires both cutting-edge computer hardware and the software, personnel support, and the training of researchers needed to maximize its scientific reach. This support is needed at many levels, from the handful of ultrapowerful machines that enable the most ambitious calculations, through the larger and more varied tier of supercomputers available at national and state-supported centers, and on to high-performance clusters in individual research groups and the networks of workstations and laptops by which scientists access these facilities and examine the results of their computations.

Although the advances in computational theory are dramatic, it is often pencil-and-paper theory that leads to novel ideas or identifies the connections between seemingly disparate phenomena. The frontiers of cosmology today present grand theoretical challenges: rooting models of inflation in more fundamental descriptions of underlying physics; explaining the asymmetry between matter and antimatter and thus the origin of the particles that form Earth and the life on its surface; describing the interior structure of black holes and explaining their entropy in terms of quantum gravity; determining whether there are spatial dimensions beyond the three of everyday experience; explaining the surprising magnitude of cosmic acceleration and the seeming coincidence of the densities of baryons, dark matter, and dark energy; and determining whether our observable cosmos is a fully representative sample of the universe or one of many disparate bubbles in a much larger inflationary sea. Robust support for the full span of theoretical activities is essential in order to reap the return from large investments in observational facilities over the next decade, and also to ensure that the scientific opportunities in the 2020-2030 decade will be as exciting as those of today.

BIBLIOGRAPHY

Bahcall, J. 2004. Solar neutrinos: A popular account. *Encyclopedia of Physics.* Edition 3, arXiv:physics/0411190.

Bertone, G., D. Hooper, and J. Silk. 2005. Particle dark matter: Evidence, candidates and constraints. *Physics Reports* 405:279.

Blumer, J., R. Engel, and J.R. Horanfel. 2009. Cosmic rays from the knee to the highest energies. *Progress in Particle and Nuclear Physics* 63:292.

Camilleri, L., E. Lisi, and J. Wilkerson. 2008. Neutrino masses and mixings: Status and prospects. *Annual Review of Nuclear and Particle Science* 58:343-369.

Fields, B., and K. Olive. 2006. Big bang nucleosynthesis. *Nuclear Physics A* 777:208.

Frieman, J.A., M.S. Turner, and D. Huterer. 2008. Dark energy and the accelerating universe. *Annual Review of Astronomy and Astrophysics* 46:385-432.

Hannestad, S. 2006. Primordial neutrinos. *Annual Review of Nuclear and Particle Science* 56:137-161.

Hoffman, K.D. 2009. High energy neutrino telescopes. *New Journal of Physics* 11:055006.

Hughes, S.A. 2003. Listening to the universe with gravitational-wave astronomy. *Annals of Physics* 303:142.

Sathyaprakash, B.S., and B.F. Schutz. 2009. Physics, astrophysics and cosmology with gravitational waves. *Living Reviews in Relativity* 12:2. Available at http://www.livingreviews.org/lrr-2009-2.

Sommers, P., and S. Westerhoff. 2009. Cosmic ray astronomy. *New Journal of Physics* 11:055004.

Stecker, F.W., and S.T. Scully. 2009. Searching for new physics with ultra-high energy cosmic rays. *New Journal of Physics* 11:085003.

Steigman, G. 2009. Primordial nucleosynthesis after WMAP. In *Chemical Abundances in the Universe: Connecting First Stars to Planets*, Proceedings of the IAU, Volume 265, arXiv 0909.3270.

Uzan, J.-P. 2003. The fundamental constants and their variation: Observational status and theoretical motivations. *Reviews of Modern Physics* 75:403-455.

Uzan, J.-P. 2009. Fundamental constants and tests of general relativity—Theoretical and cosmological considerations. *Space Science Reviews* 148:249-265.

2

Report of the Panel on the Galactic Neighborhood

SUMMARY

The galactic neighborhood occupies a key role in our quest to understand the universe. Extending from the Milky Way and the Local Group out to redshifts $z \approx 0.1$, the galactic neighborhood contains galaxies of all morphological types, metallicities, masses, histories, environments, and star-formation rates. However, unlike galaxies seen at greater distances, those within the galactic neighborhood can be probed with parsec-scale resolution down to faint luminosities. The resulting sensitivity permits the dissection of galaxies into their individual components, reaching the scale of individual stars and gas clouds. Moreover, these constituents can be studied in their proper context and with full knowledge of their galactic environment, allowing one to connect the stars and gas to the larger structure within which each formed. Thus, only in the galactic neighborhood can galaxies be studied as the complex, interconnected systems that they truly are, governed by microphysical processes. Probing this complexity involves studying processes that connect galaxies to extended gaseous systems: the interstellar medium (ISM), circumgalactic medium (CGM), and intergalactic medium (IGM).

The detailed observations possible in the galactic neighborhood also make it the critical laboratory for constraining the physics that governs the assembly and evolution of galaxies and their components across cosmic time. Indeed, almost every field of astrophysics—from the evolution of stars to the structure of dark matter halos, from the formation of supermassive black holes to the flows of gas in and out of galaxies—benefits from the detailed physical constraints that are

possible to achieve only in the galactic neighborhood. Not surprisingly, these constraints have been woven into the modern theoretical framework for galaxy formation and evolution.

To appreciate the impact of the galactic neighborhood, first consider studies of the universe on the largest scales. The interpretation of observations of the most distant galaxies is built on a foundation of knowledge established in the galactic neighborhood, including knowledge about the evolution of stellar populations, the existence of dark matter, the scaling relations of supermassive black holes, the effects of feedback from supernovae, the importance of accretion, the relationship between star formation and gas density, and the stellar initial mass function, among many others. Likewise, the evidence for dark energy from high-redshift supernovae was predicated on years of characterization of the properties of supernovae in nearby galaxies, along with more mundane constraints on the properties of dust extinction and exhaustive calibrations of the local distance scale.

The impact of the galactic neighborhood has been equally significant on smaller scales. The galaxies of the Local Group offer millions of observationally accessible stars, assembled into systems with a common distance and foreground extinction. The resulting samples of stars, their ancestors, and their descendants (e.g., planetary nebulae, supernova remnants, variable stars, transients, supernovae, molecular clouds, H II regions, X-ray binaries, etc.) can be analyzed with fewer uncertainties than in the Milky Way, where unknown distances and reddenings present challenging obstacles to assembling large samples. Moreover, such samples span a wide range of environment and metallicity, adding these new dimensions to the understanding of the physics of stellar evolution and the interstellar medium. The galactic neighborhood is also the only region where one can study the smallest scales of galaxy formation, revealing the presence of galaxies whose masses are scarcely more than a globular cluster. This fact is particularly important for assessing processes of feedback from star formation to the ISM, CGM, and IGM.

In assessing the scientific potential of the galactic neighborhood over the coming decade, the Panel on the Galactic Neighborhood faced a difficult task, given that the galactic neighborhood is the arena within which the interaction of nearly all astrophysical systems can be witnessed. Thus, narrowing down the scientific potential to only four key questions involved both the exclusion of research areas and unavoidable overlap with the scientific realms covered by other Science Frontiers Panels participating in the National Research Council's (NRC's) Astronomy and Astrophysics (Astro2010) Survey. This panel chose to focus its questions on areas in which the constraints from the galactic neighborhood are most powerful and unique. As a result, the four science questions developed by the panel exploit the use of the galactic neighborhood as a venue for studying interconnected astrophysical systems, for constraining complex physical processes, and for probing small scales. The key science questions are as follows:

- *What are the flows of matter and energy in the circumgalactic medium?* This question concerns the understanding of the circumgalactic medium that is needed to understand the mass, energy, and chemical feedback cycle that appears to shape the growth of galaxies and the metal enrichment of the universe. In this report the panel identifies a program of detailed observations of the accretion and outflow processes in nearby galaxies that can inform the understanding of these processes at all epochs and mass scales.
- *What controls the mass-energy-chemical cycles within galaxies?* This question explores the rich system of gas and stellar physics that shapes, and is shaped by, the interstellar medium. The panel outlines multiwavelength and theoretical studies of gas, dust, and magnetic fields within galaxies. Such studies can unravel the complexities of the gaseous ecosystem, with a level of detail critical to isolating the relevant physics but that cannot be obtained outside the galactic neighborhood.
- *What is the fossil record of galaxy assembly from first stars to present?* This question focuses on probes of the fossil record of star formation, galaxy assembly, and the first stars. The panel identifies the value of surveys for resolved stars at high spatial resolution, with spectroscopic follow-up of stellar populations and metal-poor halo stars providing high-impact science unique to the galactic neighborhood. Furthermore, this fossil record promises to reveal the properties of galaxies at epochs where they cannot be seen directly.
- *What are the connections between dark and luminous matter?* This question addresses the use of the galactic neighborhood as a laboratory of fundamental physics. The local universe offers the opportunity to isolate the nearest and smallest dark matter halos and to study astrophysically "dark" systems at high spatial resolution. The panel discusses the many observational and theoretical advances that could be expected as a result of these unique capabilities.

The prospects for advances in the coming decade are especially exciting in these four areas, particularly if supported by a comprehensive program of theory and numerical calculation, together with laboratory astrophysical measurements or calculations. The sections that follow this Summary describe the unresolved scientific issues in more detail, highlighting specific observational and theoretical programs that offer significant opportunities for advancing scientific understanding. Also highlighted is the discovery potential of *time-domain astronomy* and *astrometry* for capitalizing on powerful new techniques and facilities that provide precise connections among stars, galaxies, and newly discovered transient events.

Highlights of Top Activities Identified by the Panel

To make significant progress in addressing the four science questions, the panel suggests a broad program of ground-based and space-based science, together with theoretical studies. In the highest overview, galactic neighborhood science uses the local universe as a laboratory for fundamental physics and astrophysics, galactic and dark matter structures, gas flows in and out of galaxies, and the fossil record of galaxy assembly. The astronomical goal is toward an understanding of how gas gets into galaxies, arranges itself to form stars, and returns to the galactic surroundings, reprocessed in the form of radiative, mechanical, and chemical "feedback."

The science goals discussed in this panel's report depend on the ability to trace the interconnected, multiphase nature of galaxies and their surroundings. This complexity naturally leads to a very broad range of desired observational and theoretical capabilities. Tables 2.1 through 2.4 at the end of this panel report summarize in some detail many of the possible capabilities that are mentioned in the sections following the Summary.

The panel recommends powerful new ultraviolet (UV) and X-ray missions for spectroscopic studies of these gaseous structures, chemical abundances, and flows. Studying processes within the galaxies requires capability at longer wavelengths (infrared [IR], submillimeter, millimeter, radio) to probe the processes that transform accreted gas into stars. Measuring the fossil record requires the identification of large numbers of stars through photometric and kinematic surveys and the subsequent study of their chemical content. Studies of star-formation histories through color-magnitude diagrams require both high spatial resolution on large optical and infrared (OIR) telescopes in space and high-resolution stellar spectroscopy on very large telescopes on Earth. Pursuing the connections between dark and luminous matter requires kinematic and abundance studies of dwarf galaxies and their stars, as well as of black holes that reside in many galactic nuclei, particularly in the Milky Way center. Progress in the areas of discovery potential identified by the panel can be made with new OIR and radio facilities that follow the transient universe and with powerful astrometric facilities.

GAN 1. WHAT ARE THE FLOWS OF MATTER AND ENERGY IN THE CIRCUMGALACTIC MEDIUM?

Observations over the past decade have revealed strong evidence for the infall and outflow of matter from galaxies. Less understood is the circulation of mass, energy, and chemical elements between galaxies and the IGM. This section discusses key issues of how gas gets into and out of galaxies—processes colloquially named "accretion" and "feedback," respectively. The major feedback processes include the deposition and transport of mass, momentum, energy, and heavy elements into the ISM, CGM, and IGM by stars and supermassive black holes in galactic nuclei. *The bulk of the energy and metals from these feedback channels is still not accounted for observationally.* The common theme of these observations is emission-line and absorption-line spectra at moderate to high resolution.

The current ideas of how gas gets into galaxies have emerged from cosmological simulations, with galactic mass as the key parameter. Galaxies with $M_{\text{halo}} \geq 10^{12}$ solar mass ($M_\odot$) accrete gas in the so-called hot mode after being heated to temperature (T) > 10^6 K by shocks near the virial radius. This shock does not develop in models of lower-mass galaxies, where gas flows along narrow filaments ("cold-mode" accretion), extending well inside the virial radius (Figure 2.1). This simple picture may provide a natural interpretation for the observed galaxy bimodality of stellar color and morphology, if one associates cold mode with blue, star-forming galaxies and hot mode with red, inactive galaxies. Firmly establishing such links to the stellar properties of galaxies requires observations that probe directly the gas-accretion modes.

Direct evidence for galactic outflows comes from the multispectral imaging of local galaxies (Figure 2.2) as well as the heavy elements detected throughout much of the IGM. Indirect evidence for gas flows comes from the observation that many galaxies contain lower fractions of baryons to dark matter than the primordial ratio, Ω_b/Ω_m. These missing baryons may reside in the IGM or in an extended multiphase CGM produced by interactions between galaxies and their intergalactic environment. Displacing these baryons to large spatial scales requires substantial energy deposition. Therefore, probing the content and circulation of matter, energy, and

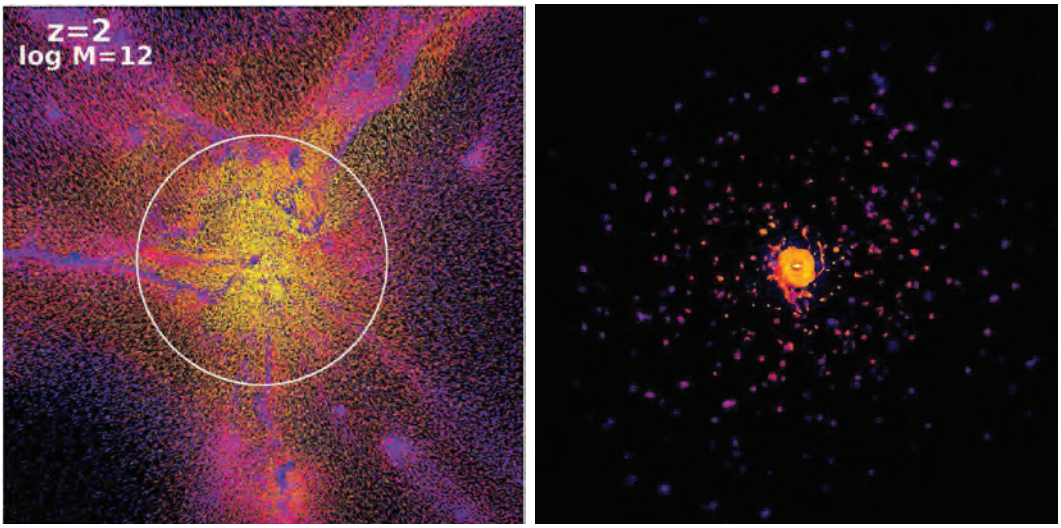


FIGURE 2.1 Milky Way-size halo (left) at z = 2 fed by cold filamentary streams of gas (purple 10^4 K; yellow 10^6 K) into the virial radius. Milky Way-size halo at z = 0.1 (right) with infalling dense gas. SOURCE: M.E. Putman et al., How do galaxies accrete gas and form stars?, Astro2010 science white paper, submitted 2008.

FIGURE 2.2 Messier 82 outflow—composite of Chandra, HST, and Spitzer images. X-ray data recorded by Chandra appears in blue; infrared light recorded by Spitzer appears in red; Hubble's observations of hydrogen emission appear in orange; bluest visible light appears in yellow-green. SOURCE: *X-ray:* Courtesy of NASA/CXC/Johns Hopkins University/D. Strickland. *Optical:* Courtesy of NASA/ESA/STScI/AURA/The Hubble Heritage Team. *IR:* Courtesy of NASA/JPL-Caltech/University of Arizona/C. Engelbracht.

heavy elements between galaxies and the CGM is central to an understanding of galaxies and the IGM.

The high spatial resolution and panchromatic approach possible for nearby galaxies make the galactic neighborhood the best location for establishing the content and circulation of the circumgalactic gas. Indeed, observations from radio to X-ray have established the presence of galactic winds in starburst galaxies and feedback from active galactic nuclei (AGN) in massive elliptical galaxies. Optical observations that probe the CGM of distant galaxies sample a limited temperature range with little spatial information. A comprehensive understanding of galaxies and the IGM depends on an ability to probe the content and circulation of matter, energy, and key elements. In the next decade, transformative gains in knowledge in this area will require two proven observational strategies: absorption-line tomography and spectral imaging. These techniques need to be applied to more targets at better velocity resolution so as to detect gas over a broad temperature range.

New observations at wavelengths most sensitive to these flows will help to remove uncertainties in simulations of galaxy formation and evolution, which currently lack the resolution required for the direct modeling of all physical processes. Physically motivated and empirically constrained recipes have been prescribed to model outflows in such simulations. But the amount of matter and energy in these outflows, the ultimate fate of the matter (escape or circulation), and their relation to the feedback process in more normal galaxies remain controversial. The coupling of the radiation, momentum, and kinetic energy from stars and supermassive black holes with the surrounding interstellar and circumgalactic gas is extremely difficult to model from first principles. Indeed, such models provide theoretical challenges for future simulations. Feedback mechanisms may enrich the IGM in heavy elements, create the mass-metallicity scaling relation for galaxies, suppress star formation in the smallest and largest galaxies, alter masses and shapes of galaxies over cosmic time, and connect bulge and black hole growth.

Distinguishing reality from plausibility will require well-articulated tests of the models. The observations must be sensitive to the physical and kinematic properties of the CGM, and they must be tested against higher-resolution simulations with additional physics.

What Is in the Circumgalactic Medium?

The circumgalactic medium is defined as the region around a galaxy that is strongly influenced by the galaxy's gravity and by chemical and mechanical feedback. What are the overall mass, energy, and metal contents of the CGM? What are its spatial, thermal, chemical, and kinematic structures? These questions can be addressed in many wavelength bands, but most sensitively through absorption-line tomography of the CGM around nearby galaxies. For the Milky Way, the bulk of the CGM likely resides in a hot phase with $T \sim 10^{6\text{-}7}$ K, which can affect infalling clouds. The most sensitive way to detect and characterize this hot medium is absorption-line spectroscopy in the soft X-ray and UV regimes. These wavelength regions complement one another by probing different temperature regimes and ionization states in the CGM.

X-ray absorption-line spectroscopy is a powerful new tool for solving the missing-baryon and missing-feedback problems of galaxies, particularly for the Milky Way. This technique has been demonstrated in recent studies of the global hot gas in and around the Milky Way, based on the Chandra and X-ray Multi-Mirror Mission (XMM)-Newton X-ray grating observations of a few brightest

AGNs and X-ray binaries, typically with very long exposures. X-ray absorption lines produced by ions such as O VII, O VIII, Ne IX, and Fe XVII directly probe their column densities, which are proportional to the mass of the gas. Although individual lines are not resolved with the existing grating instruments, the velocity dispersion of the hot gas can often be inferred from the relative line saturation (e.g., by comparing O VII, Kα, and Kβ). These studies are leading to the first global characterization of the spatial, thermal, chemical, and kinematic properties of the ISM in our galaxy.

Requirements for X-Ray Absorption-Line Spectroscopy

Tomography of the CGM will require a substantially improved line detection sensitivity ($S/N \propto \sqrt{RA}$, where $R = \lambda/\Delta\lambda$ and A are the spectral resolution and photon collecting area, respectively) in the 0.3- to 1-keV range, which encloses the key metal lines. A resolution of 100 km s^{-1} ($R = 3{,}000$) and effective area $A > 1{,}000$ cm^2 would provide 100 times improvement over Chandra and XMM-Newton, opening up the detection of metal lines from enriched outflows in the CGM and IGM. To conduct effective tomographic mapping, a few hundred sight lines are needed, which can be achieved by observing AGNs with fluxes greater than 10^{-11} ergs s^{-1} cm^{-2} for a reasonable exposure (for example, ~30 ks each), according to the source list from the Röntgen Satellite (R∝OSAT) All-Sky Survey. The required line detection sensitivity is a factor of approximately 15 higher than that offered by the Chandra and XMM-Newton grating instruments. This sensitivity will allow one to use essentially all low-mass X-ray binaries in our galaxy and relatively bright ones in Local Group galaxies. The desired velocity resolution is ~100 km s^{-1}, which will produce line centroiding at 10 to 20 km s^{-1} (for $S/N \geq 10$).

Tomography of the hot gas in and around the Milky Way will allow the decomposition of hot gaseous components of the Milky Way (galactic disk, bulge, halo) as well as of galaxies in the Local Group and local large-scale structure. Their global spatial, thermal, chemical, and kinematic structures can then be characterized. Absorption-line spectra toward the AGN sight lines can also be used to sample the CGM around intervening galaxies, at different impact parameters. With longer exposures, one can observe the fields of Local Group galaxies, and fainter AGN and binaries can be observed with longer exposures. The overall metal content and metal transport around galaxies can be inferred directly, allowing one to determine how these properties depend on the star-formation rate, mass, morphological type, and environment.

How Are Galaxies Fed?

Understanding Cool and Warm Gas in the Circumgalactic Medium

An important issue related to the first key question is how the CGM controls the gas supply for star formation in galaxies. To address this question, both cool ($T \leq 10^4$ K) and warm ($T \leq 10^5$ K) gas in the CGM must be understood. This gas has been probed in 21-cm emission and a few optical lines (Na I, Ca II). However, by far the most sensitive probe is spectroscopy of the UV resonance absorption lines of metal ions and H I. The UV-optical wavebands include numerous diagnostic absorption lines from a wide range of ionization states of common elements from H to Zn, which provide measurements of gas temperature, density, metallicity, and kinematics in cold or cooling gas within the hot CGM seen in the X-ray. A great deal has already been learned about the cool CGM around the Milky Way, based on existing observations, particularly those from the Far Ultraviolet Spectroscopic Explorer (FUSE) and the Hubble Space Telescope (HST) Space Telescope Imaging Spectrograph. In the next decade, the HST Cosmic Origins Spectrograph (COS) is expected to advance this technique significantly by applying it to potentially hundreds of galaxies in the galactic neighborhood. Obtaining these nearby external views will allow for a characterization of the projected distribution of cool gas around galaxies.

Requirements for Far-Ultraviolet Spectroscopy

The most desirable new facility for far-UV absorption-line spectroscopy of the CGM would be a follow-up of FUSE, but with a 10-fold increase in effective area. This capability will allow spectra to be taken of thousands of fainter AGN, as background targets for intervening galaxy halos. This facility will enable observations of O VI-bearing gas (not accessible in the COS pass-band) around many nearby galaxies, which are especially important for studying the interplay between hot and cool components of the CGM. Absorption studies can also probe the relationship between the hot CGM and infalling material, such as cool high-velocity gas clouds. Full three-dimensional information about the cool components of the CGM can be obtained from emission-line mapping: a wide-field spectrograph for Hα-emitting gas at approximately 10^4 K and 21-cm studies of H I clouds, using existing and upcoming optical and radio telescopes. Together with ever-improving numerical simulations (see, e.g., Figure 2.1), one can address questions such as the following: What is the origin of these clouds? How do they evolve in the CGM? Do they interact strongly with galactic disks?

How Does Galaxy Feedback Work?

Owing to the broad range of gas temperatures present in galactic flows, an accurate accounting of the outflowing material requires panchromatic, spectral-imaging observations (Figure 2.2). Such observations can identify the dominant physical mechanism driving the winds, as well as quantify the amount of mass, energy, and chemical feedback. Targeting galaxies out to redshift $z \sim 0.05$ would sample the full range of feedback activity, from dwarf and elliptical galaxies to extreme starbursts (star-formation rates $> 100\ M_{\odot}\ \mathrm{yr}^{-1}$). This census would form the foundation for building a physical understanding of galaxy feedback.

The energy-carrying phase of a galactic wind emits primarily at X-ray frequencies, where observations have been limited to a small number of galaxies. Current facilities lack the velocity resolution required to measure gas kinematics in the hot phase, so only the thermal energy content (and not the kinetic energy) has been measured. Kinematic evidence for outflow lies with observations of much cooler gas, likely (but not definitively) entrained in the hot outflow. The relative speed of these components is model-dependent and not agreed on. Spectral imaging observations of the hot phase should be a key science driver for new X-ray instrumentation and missions.

Requirements for X-Ray Spectral Imaging

To observe a significant sample of perhaps 40 starbursts, the 0.3-10 keV sensitivity must reach fluxes of $\sim 4 \times 10^{-14}$ ergs $\mathrm{s}^{-1}\ \mathrm{cm}^{-2}$. The imaging (5″ and 10-20′ field of view) needs to be sensitive to low surface brightnesses over kpc scales and to separate emission from hot plasmas and stellar sources, mainly bright X-ray binaries and supernova remnants, in galaxies at approximately 10 Mpc. The gas temperature may vary from 10^6 K to 10^8 K depending on the amount of mass-loading (mixing with cool, entrained gas). At the lower end of this temperature range, the plasma emission is dominated by X-ray line emission. Severe blending of these lines hinders the determination of the continuum level in existing spectral imaging data. Existing dispersive grating spectrometers are not designed for spatially extended targets and lack the sensitivity and spectral resolution required to study low surface-brightness emission. Direct measurement of the wind kinetic energy requires line centroiding to approximately 100 km s^{-1}, achievable with full width at half maximum of 1,000 km s^{-1} and $S/N \geq 10$ in the emission line. The required kinematic resolution would automatically enable the use of line-ratio-based temperature and ionization state diagnostics, greatly improving the knowledge of metallicity and thermal energy as well.

Requirements for Ultraviolet Spectral Imaging

Kinematic mapping of warm circumgalactic gas will provide insight into the interaction of a galaxy's disk and halo. It will also reveal the circulation of matter between galaxies and cosmic filaments. Studies of the galactic halo suggest that warm gas constitutes a major component of the infall feeding the galaxy, but they are limited by distance uncertainties, which will not apply to observations of nearby galaxies. Emission-line imaging in the ultraviolet (e.g., O VI, H I Ly-α, C IV, C III, N V, O III, Mg II) will trace filaments and clouds to much larger galactocentric radii than is possible in the optical owing to the lower background. Inferring the electron density will provide the thermal pressure and absolute metal abundance of the CGM. Combining kinematic and line-ratio measurements from rest-frame UV spectroscopy distinguishes shock excitation, photoionization, and cooling radiation. Line ratios determine temperature and metallicity. When spatially mapped over the entire circumgalactic region, these diagnostics can distinguish infalling gas along filaments, clouds condensing and falling toward the galaxy, stripping of gas from merging subunits, outflows from galactic nuclei or star clusters, mixing of layers between phases with different temperatures, and shocks. For example, the collision of the hot wind with cool infalling gas produces radiative shocks, whereas mixing layers between the hot wind and entrained, interstellar gas radiate strong emission lines. The structural information gleaned from this approach would be nearly impossible to obtain from pencil-beam, line-of-sight studies, and it provides useful context for the interpretation of the measured physical information.

Imaging the CGM allows one to map gas topology and structure in all directions. The surface brightness of faint circumgalactic gas is typically 20 to 200 photons cm^{-2} s^{-1} sr^{-1}. Emission from intergalactic gas will be fainter by an amount that varies widely among models. The 1,000- to 2,000-Å band-pass, particularly O VI, Ly-α, and C IV, is critical, and spatial resolution of a few arcseconds should be sufficient to remove point sources. A large field of view (20 arcmin) must be mapped to cover several tens of kpc in a large sample of local galaxies and allows campaigns to map several hundred kpc in selected environments. Spectral resolution must be sufficient to measure line-profile centroids to 30 to 100 km s^{-1} and probe feedback in the low-mass galaxies that appear to dominate intergalactic metal enrichment in the local universe. This spectral resolution may be realized by a FUSE-like spectrometer with a single aperture.

Requirement for Studying Neutral and Molecular Outflows

Cold winds from galactic nuclei may also be important for providing energy input to the IGM, for removing cold gas from massive galaxies, and for returning heavy elements to the IGM. Evidence for cold winds has been revealed from the recent discovery of a fast molecular outflow from the S0 galaxy[1] NGC 1266 and the observation of extraplanar dusty gas (e.g., see Figure 2.2). The importance of these neutral outflows to energize the IGM and clear their host galaxies of star-forming gas depends on how common they are, which is unknown at the present time. Observations with the Enhanced Very Large Array (EVLA), an expanded Allen Telescope Array (ATA), Combined Array for Research in Millimeter-wave Astronomy (CARMA), and Atacama Large Millimeter Array (ALMA), as well as Herschel, will be essential to developing an understanding of what drives these outflows without dissociating and ionizing the cold gas, and to allowing the measurement of how important these cold outflows are on global scales. For starburst galaxies, the above measurements will enable direct tests of the relative contributions of galactic superwinds, gas recycling, stellar-mass loss, and AGN feedback to determining the bulk properties of galaxies. For galactic disks, they test the galactic fountain model, while measurements of galaxies with substantial spheroidal components will provide new insights into the feedback from evolved stars and AGNs.

A full understanding of the multiscale and multiphase CGM requires a synergy of the approaches identified above, plus well-established observing and numerical simulation tools. Some specific questions to be addressed include the following: Are the properties of hot flows consistent with the thermalized energy from supernova explosions and stellar winds? Do cosmic rays and/or radiation pressure impart additional momentum to the low-ionization gas? What is the fate of Supernova Type Ia ejecta? Where is the kinetic and radiative energy from AGN deposited? What is the circulation timescale for ejected gas to be returned to the star-forming disk?

GAN 2. WHAT CONTROLS THE MASS-ENERGY-CHEMICAL CYCLES WITHIN GALAXIES?

How Do Galaxies Build Up Their Stellar Component Over Cosmic Time?

Important strides have been made in the past two decades in determining how galaxies grow from their dark matter seeds to the complex systems of dark matter, stars, and gas that are observed at $z = 0$. Simulations of the formation of structure in a universe dominated by cold dark matter (ΛCDM) have been able to reproduce the spatial distribution of galaxies and clusters of galaxies and have demonstrated the important role played by mergers. Over the same time period, observations of star-forming regions in the Milky Way have revealed some of the characteristics of the star-formation process, including both clustered and individual star formation

---

[1] An “S0 galaxy” shows evidence of a thin disk and a bulge, but has no spiral arms and contains little or no gas.

and the initial mass function. The challenge to astrophysics for the next decade is this: *How do galaxies build up their stellar component over cosmic time?*

To answer this question, the workings of the ISM on large and small scales needs to be understood. The story starts with intergalactic gas falling into the gravitational potential of a growing galaxy. Some key subquestions are these:

- What are the signposts of the inflow?
- How does gas become arranged within a galaxy?
- How does some portion of this gas assemble into molecular clouds that provide sites for star formation?
- How do conditions in and around these clouds determine the rate of star formation and the spectrum of initial stellar masses?
- How does the energy released by stars affect the ISM and conversion of gas into stars?
- How do stellar winds and supernova ejecta enrich the ISM with the heavy elements formed in stars and supernovae?

At the most basic level, there needs to be an understanding of the flow of baryonic matter and energy into galaxies (the accretion of intergalactic matter), within galaxies (secular evolution and flows into and out of stars), and sometimes out of galaxies by way of galactic winds, central jets, tidal interactions with other galaxies, and intergalactic ram pressure effects. Flows in the ISM within galaxies are driven by energy input from stars (radiation, high-velocity winds, supernovae), gravity, and infalling material (Figure 2.3). To understand the development of galaxies as stellar systems, the dynamics of the ISM must be understood on scales ranging from kiloparsecs (spiral structure), to parsecs (formation of giant molecular clouds, starbursts, and star clusters), and down to the sub-parsec level where star formation occurs. To understand the *appearance* of galaxies, there must also be an understanding of obscuration and scattering by dust and of emission by both dust and gas, all of which depend on the complex geometry of the ISM.

The evolution of galaxies at all redshifts depends on a detailed understanding of these physical processes and on knowing which ones dominate in different environments. However, studies of distant galaxies will always be limited by angular resolution. Understanding the dynamics and energy flows at earlier epochs will depend on having the most detailed observations and accompanying theory of nearby galaxies. Progress will require advances on multiple fronts: observations, theory, large-scale simulations, and laboratory astrophysics.

Star formation occurs primarily in clusters and associations, but there is almost nothing known about the conditions in the ISM that separate the production of bound clusters, OB associations, "super star clusters," and globular clusters. Knowl-

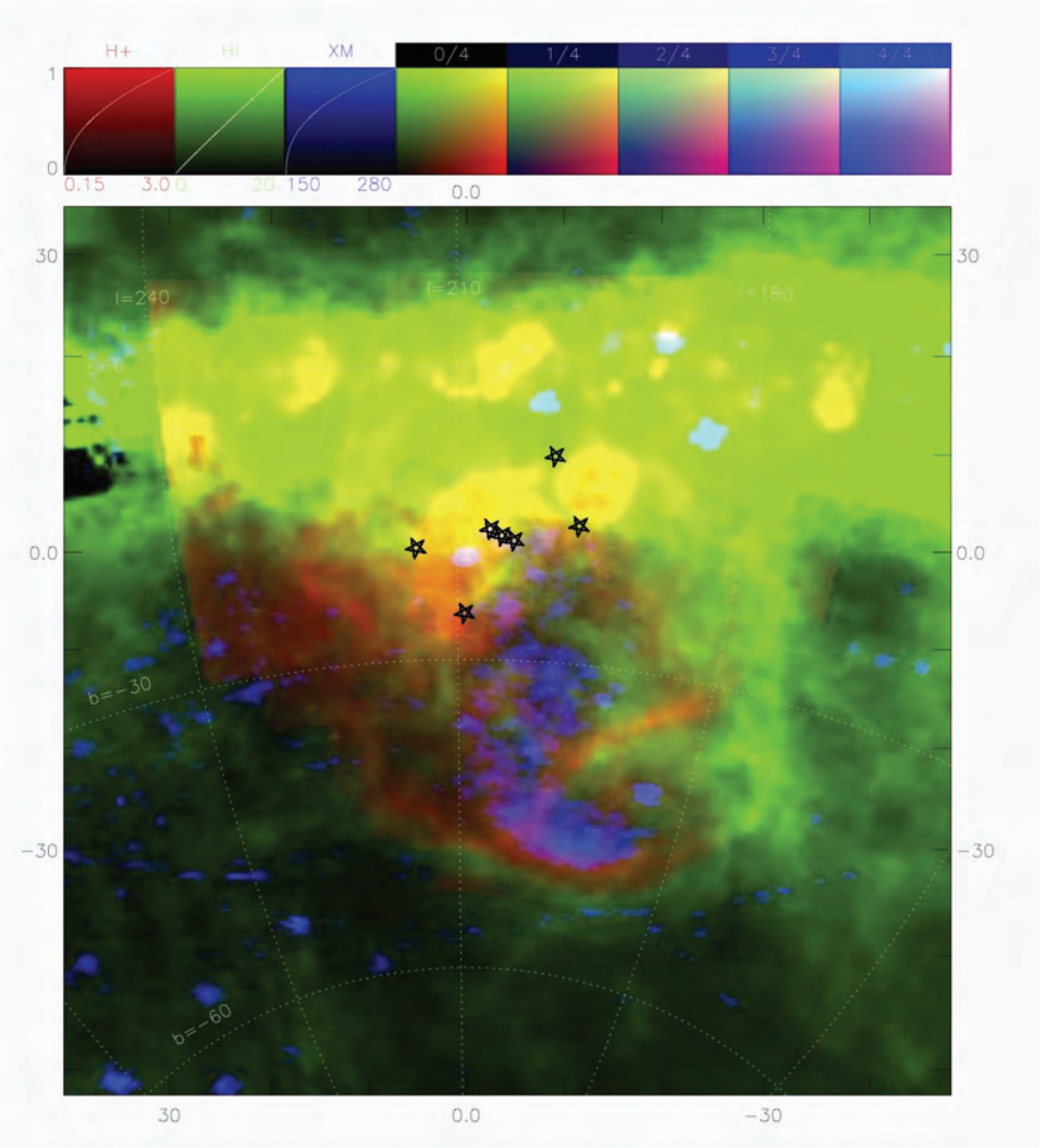


FIGURE 2.3 Large-scale image of the Orion-Eridanus superbubble (~5,000 deg^2) showing X-rays in blue (ROSAT), H II in red (WHAM), and H I in green as well as the brightest stars in Orion for scale. This image shows the importance of large-scale multiwavelength surveys for elucidating structures in the Milky Way interstellar medium. X-rays arise from hot gas heated by supernova explosions that sweep up cool H I. SOURCE: Courtesy of Carl Heiles, University of California, Berkeley.

edge of these conditions is essential, not only for defining the physics of the modes of star formation but for connecting what is learned about star formation in the Milky Way and the nearest galaxies to star formation at higher redshift. After all, beyond $z \approx 0.1$, individual stars are not observed, except possibly for O stars, and knowledge about them comes from observations of clusters and the integrated starlight of galaxies. The conditions that give rise to the stars and clusters that are observed in the nearby universe need to be extrapolated to those that exist at much greater distances.

In the next decade an increasingly clear picture of how stars form in the disk of the Milky Way will be obtained, but there is little understanding of how this relates to star formation in other galaxies. This is especially true for low-mass low-metallicity galaxies that are presumably the analogs of star-forming galaxies in the early universe. The discovery of ultrafaint dwarf galaxies in the Local Group challenges conventional wisdom about the formation of molecular clouds and star formation in such low-mass systems. It is unclear how molecular clouds form in systems with such weak gravity. Are their star-formation processes and histories similar to those observed in nearby low-mass galaxies? The puzzle of how stars form in close proximity to the black hole in the center of the Milky Way is an enigma whose solution can only be probed locally. A suite of millimeter- and submillimeter-wave single-dish telescopes and arrays, all equipped with array receivers, will be necessary to address these questions. Centimeter-wave arrays will also be needed to understand the flows of atomic gas into and out of galaxies; this gas is the raw material for the molecular gas that forms stars.

Star formation, at least in the Milky Way today, produces an initial mass function (IMF) that appears surprisingly insensitive to local conditions. Indeed, it has been difficult to find compelling evidence of variations in the IMF from one region to another. There is now preliminary evidence from nearby galaxies indicating that the IMF does vary, with an apparent deficiency of high-mass stars in outer disks and perhaps an excess of high-mass stars in galactic centers. In the coming decade, infrared and far-infrared imaging with Herschel and continuum observations with ALMA and future large-aperture single-dish submillimeter telescopes will be able to inventory any embedded high-mass stars missed by Hα observations. A 5- to 10-m aperture cooled far-infrared/submillimeter telescope in space would be a powerful tool for this purpose. It is important to carry out further studies to confirm IMF variations and to determine how the IMF depends on conditions in diverse star-forming regions. Understanding IMF variations as a function of environment at low redshift will inform the understanding of star formation at high redshift.

The Multiphase Structure of the Interstellar Medium

The ISM spans a huge range of densities ($\sim 10^{-3}$ cm^{-3} to $>10^{8}$ cm^{-3}) and temperatures (10 K to 10^{8} K) and includes magnetic fields and cosmic rays, both of which are dynamically important. The statistical properties of this fluid system remain poorly determined, including even the geometry and topology of the density field. Because the ISM is a chaotic system, large-scale numerical models can only be tested by comparing the predicted statistical properties (distribution functions for temperature, density, velocity; topological measures; anisotropic power spectrum of turbulence; and others) to those observed. The central question here is simple but sweeping: *What are the structure and physical state of the ISM in the Milky Way and nearby star-forming galaxies?* Only after these conditions in the ISM are known can there be hope of answering key questions such as the following:

What Controls the Radial and Vertical Transport of Mass and Metals?

The Milky Way and nearby galaxies are devouring their molecular star-forming gas at rates that are unsustainable for more than approximately 2 Gyr on average. Either the end of the star-forming era in the universe is near, which is unlikely, or the molecular gas lost to star formation is being replaced from reservoirs of H I in the outer parts of galaxies, or H II from the halo or IGM. The history of star formation and chemical evolution in the Milky Way disk appear to require a continuous delivery of low-metallicity gas from the CGM to the disk, but the process is not understood. The Milky Way's high-velocity clouds are presumably part of this delivery system, but there may be other means by which gas is supplied to the disk. Gas in the outer parts of galaxies may be redistributed through tidal interactions in galaxy groups and/or radial inflows through the disks. Accompanying angular momentum transfer may also redistribute the raw material for molecular cloud formation. The transport of mass and metals out into the halo, and radially in the disk, is also important but poorly understood. Vertical transport will affect the metallicity of outflows from star-forming galaxies, whereas radial transport will alter the radial metallicity gradient in galaxies, which may not be entirely the result of local processing.

What Controls the Disk-Halo Interface?

There is no sharp boundary separating the gaseous disk from the gaseous halo and the flow of matter and energy between them. Superbubbles and "chimneys" resulting from correlated supernova explosions appear to be important for injecting gas, kinetic energy, and cosmic rays into the halo, as well as allowing O stars to photoionize gas far above and below the disk. This interface can be probed with

UV absorption-line spectroscopy and observations of diffuse infrared, optical, UV, and X-ray emission.

What Determines the Energetics and Star Formation of Far Outer Disks?

Several recent discoveries challenge the current understanding of the gas disks of spiral galaxies (Figure 2.4) that extend far beyond their bright optical disks. Atomic hydrogen has been found nearly 100 kpc from the centers of some galaxies, far more distant than had been previously known. Images from the Galaxy Evolution Explorer (GALEX) satellite have shown that star formation occurs, even in the most distant regions of H I disks, challenging present understanding of how and where molecules form and of how star formation takes place on global scales. Older issues such as how outer H I disks maintain nearly constant velocity dispersion with radius and what keeps gas disks thin in the apparent absence of stellar disks remain unanswered. Because most gas accretion likely takes place in the outer disks, understanding these phenomena appears to be critically linked to an understanding of galaxy evolution.

The Need for Multiwavelength Observations

Multiwavelength observations are required to characterize the ISM-IGM complexity.

- *Sensitive, high-resolution, all-sky 21-cm emission surveys* out to $z = 0.1$ will be possible with the Allen Telescope Array if it is expanded to at least 128 and perhaps 256 dishes. The Expanded Very Large Array will be able to survey smaller portions of the sky in H I at yet-higher angular resolution, and the Green Bank Telescope (GBT) equipped with array receivers can survey the sky at more modest resolution, as can the Arecibo telescope with an upgrade to the number of 21-cm receivers that it currently has. These telescopes will be able to survey the Milky Way, galaxies out to moderate redshifts, as well as H I in the IGM. They will be able to determine the H I distribution and kinematics as well as the spin temperature on sight lines with background radio sources. The goal is to map $\geq 3\pi$ steradians over the space of several years with 10 km s^{-1} velocity resolution and ~1,000 deg^2 with 1 km s^{-1} spectral resolution, 1 arcmin spatial resolution, and 20,000 km s^{-1} velocity range. These facilities should be able to achieve $\sigma(T_A) = 0.5$ K per pixel and have bandwidth sufficient to map ~1,000 deg^2 to $z = 0.5$ in a few years of observation.
- *Millimeter- and submillimeter-wave observations of molecular lines and dust continuum in nearby galaxies* using large single-dish telescopes and millimeter-wave arrays equipped with array receivers (in particular, CARMA) will be needed for mapping large areas at high resolution. ALMA will be incomparable for the high-

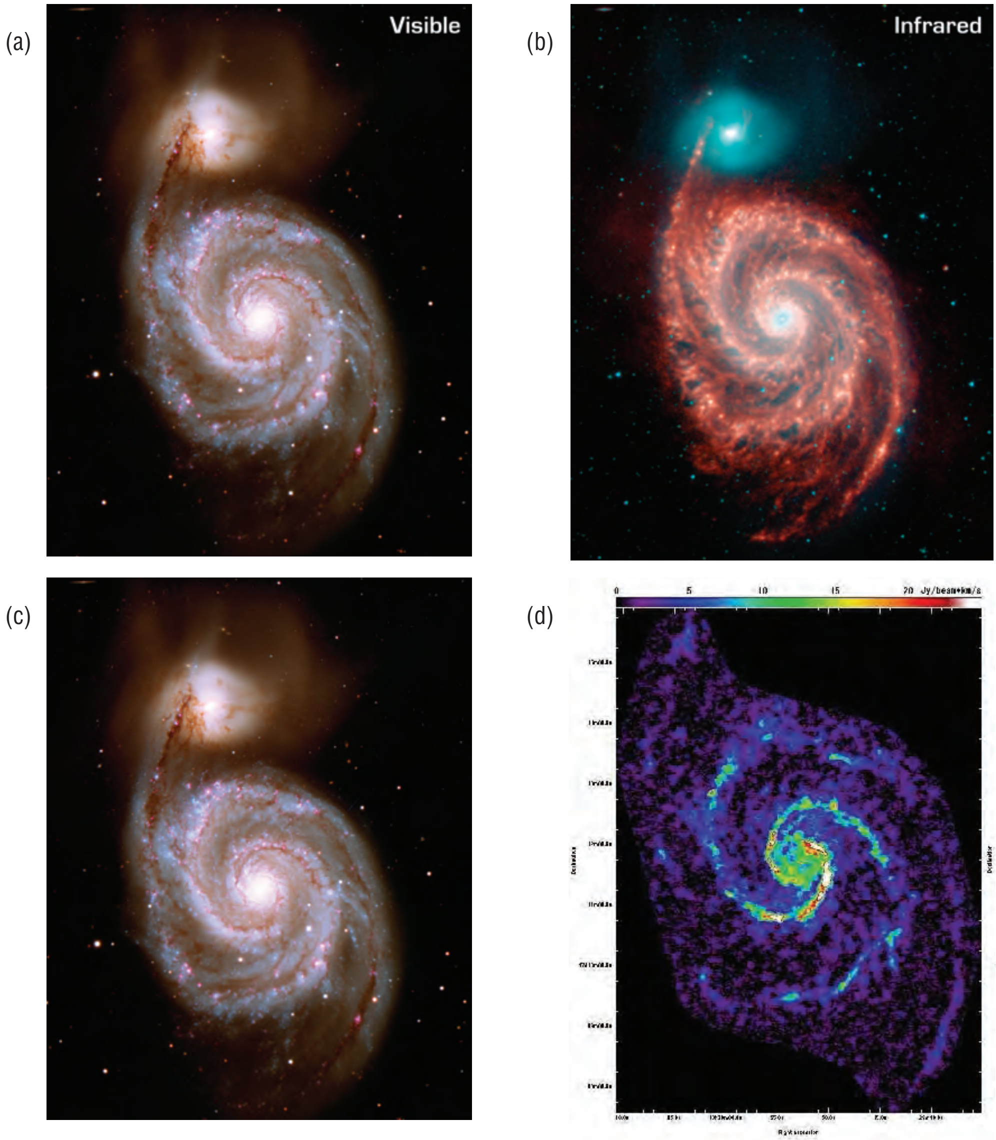


FIGURE 2.4 Three images of the grand-design spiral galaxy M51 at increasing wavelengths from left to right on the same scale. The collection of images shows that the spiral arms, seen in absorption in the visible, are seen in emission in the infrared and millimeter-wave radiation from the CO molecule. The far-infrared image also shows the thermal emission, tracing the distribution of dust heated in the galaxy, peaking in the spiral arms. SOURCE: *Left to right:* NASA/JPL-Caltech/R. Kennicutt (University of Arizona)/DSS; NASA/JPL-Caltech/R. Kennicutt (University of Arizona); ESA and the PACS Consortium; J. Koda, N. Scoville, T. Sawada, M.A. La Vigne, S.N. Vogel, A.E. Potts, J.M. Carpenter, S.A. Corder, M.C.H. Wright, S.M. White, B.A. Zauderer, et al., Dynamically driven evolution of the interstellar medium in M51, *Astrophysical Journal Letters* 700(2):L132, 2009, reproduced by permission of the AAS.

resolution, high-sensitivity mapping of small areas and for high-molecular-line transitions and dust continuum. These telescopes will characterize the distribution and kinematics of molecular gas in structures including individual clouds, giant molecular clouds (GMCs), and dense gas in galactic centers that is likely fuel for energetic activity as well as the distribution of star-forming gas and dust in galaxies to $z = 0.1$. For high-resolution mapping, ALMA will be unsurpassed as long as the present plan to build at least 50 telescopes is maintained. However, the field of view is small, especially at higher frequencies (~5-25 arcsec). Therefore, array receivers and bolometric cameras with a large number of elements should be developed for deployment on CARMA, GBT, and the Large Millimeter Telescope (LMT) so that they may be ultimately installed on ALMA and future large single-dish facilities.

- *All-sky maps from Akari, Planck, and the Wide-field Infrared Survey Explorer (WISE)* will provide new views of the distribution of thermally emitting dust on arcminute scales. As dust and gas are tightly coupled, the dust morphology reveals the distribution of the gas. Spitzer has surveyed star formation throughout the Milky Way and many external galaxies. Herschel (3.5-m aperture) will provide detailed images of selected regions, but a large (~20-m aperture) submillimeter telescope at a high-altitude site would allow high-resolution mapping at 350 μm. A 5- to 10-m cryogenic submillimeter telescope in space would permit dust continuum imaging with better resolution and much greater sensitivity than Herschel can provide.
- *High-resolution ultraviolet absorption-line spectroscopy* remains the best method for determining the composition, ionization, and kinematics of interstellar gas (at $T \leq 3 \times 10^5$ K) along sight lines. Using background stars and AGN allows the three-dimensional structure of the gas to be delineated. The goal is to achieve $\Delta v = 1.5$ km s^{-1} resolution, sensitivity down to $\lambda = 1030$ Å to study H_2 absorption lines ($\lambda < 1110$ Å) and important ionic lines such as O VI 1032, 1038 Å, and the ability to achieve $S/N > 100$ on targets with $m_{AB} = 15$.
- *High-resolution X-ray absorption-line studies* can determine the distribution of abundant elements (such as C, O, Si, Fe) in different ionization states, as well as in solid grains. Such studies are especially important for determining the thermal, chemical, and kinematic properties of the hot gas in a galaxy. The goal is an X-ray telescope with line-detection sensitivity 10 times greater than that of Chandra or XMM-Newton. Energy resolution of ~1 eV would enable studies of grain mineralogy as well as gas kinematics.
- *Observations of diffuse ultraviolet and X-ray emission* will permit high-temperature gas to be mapped and its temperature and ionization state to be diagnosed. Combined with absorption studies, this will determine the gas density and spatial scale of gas in various thermal phases. The required sensitivity and angular resolution are similar to requirements for measuring diffuse emission from the CGM or IGM. The diffuse Hα emission from the ISM is difficult to account

for with known sources of ionization (similar extended Hα emission is seen above and below the disks of other galaxies).

Some Enigmas

It is important to highlight several puzzling phenomena that appear incompatible with current models of the ISM and other galaxies. Because they may be indicative of serious problems with current models, the following phenomena should be pursued both observationally and theoretically: (1) the presence of very-small-scale neutral blobs in the ISM; (2) the so-called extreme scattering events that seem to indicate small, dense blobs of plasma in the ISM; and (3) molecular gas in almost 25 percent of E and S0 galaxies, in many of which no H I is detected.

What Is the Structure of the Magnetic Field in the Interstellar Medium?

Magnetic fields are dynamically important in the interstellar medium: strong enough to control gas motions in H I clouds and star-forming molecular clouds and to affect outflows from the disk into the CGM. Although present-day knowledge of magnetic fields in galaxies remains sparse, technological developments now permit observational progress:

- *Starlight polarimetry.* Using visual wavelengths for diffuse regions and infrared (K band) for dense clouds, starlight polarimetry can provide detailed maps of the magnetic field within individual clouds. This capability will require the construction of imaging polarimeters for large telescopes at V, K, and possibly other wavelength bands.
- *Polarimetry of far-infrared/submillimeter emission from dust.* Planck will provide all-sky maps of polarized submillimeter emission from dust on 5 arcmin scales in bright regions, and on ~1 deg scales from the average infrared cirrus. New instruments such as SCUBA-2—successor to the Submillimetre Common-User Bolometer Array (SCUBA)—and ALMA will be able to map the polarized emission from dust in molecular clouds and protostellar disks. Future instruments may permit far-infrared polarimetry from the Stratospheric Observatory for Infrared Astronomy (SOFIA). These maps, combined with other data, will disclose the three-dimensional structure of the magnetic field within a few hundred pc of the galactic plane and reveal the projected magnetic-field structure in nearby galaxies.
- *Zeeman effect.* The Zeeman effect uses H I, OH, and CN to determine magnetic-field strengths in atomic and molecular gas. High-resolution all-sky surveys are now possible with radio interferometers. Such surveys require polarization purity of 20 dB and velocity resolution of 0.1 km s^{-1}. The sensitivity and resolution are the same as given above for all-sky H I surveys.

- *Faraday rotation.* Faraday rotation uses pulsars and AGN to probe the line-of-sight component of the magnetic field weighted by the electron density. Ultrahigh-bandwidth instrumentation and radio telescopes, now being developed for all-sky Zeeman measurements, will be suitable for Faraday rotation studies.
- *Microwave synchrotron emission.* The C-Band All Sky Survey will produce an all-sky map of polarized synchrotron emission near 5 GHz, where Faraday depolarization is modest, revealing the (projected) galactic magnetic field on ~1-deg scales.

Microphysics of the Interstellar Medium

To understand how galaxies form out of gas, it is necessary to improve the theoretical understanding of the important dynamical processes. Critical areas include the following:

- *Magnetohydrodynamics (MHD) and plasma physics theory.* MHD simulations are increasing in sophistication and spatial resolution but still lack accurate representations of "subgrid" phenomena such as the decay of turbulence, magnetic reconnection, and ambipolar diffusion. Theoretical work must focus on these fundamental processes. Key questions are these: *What processes are responsible for generating galactic magnetic fields? How does the field evolve?*
- *Shock waves.* Shock waves are ubiquitous in the interstellar and IGM, but there are still gaps in the theoretical understanding of phenomena including the following: the structure of collisionless shocks, cosmic-ray acceleration in shocks, magnetic-field amplification in fast shocks, coupling between T_e and T_i in collisionless shocks, thermal conduction, multifluid MHD shocks in neutral clouds, and the role of charged dust grains.
- *Interstellar dust.* Interstellar dust is important because of its role in attenuating and scattering light, its dynamical effects, and its value as a diagnostic tool (tracer of the gas, emission spectra sensitive to the local starlight intensity, and polarized emission and extinction sensitive to the local magnetic field). Using aligned dust as a tracer for magnetic fields requires an understanding of the shapes and optical properties of dust grains and of how variations in the degree of dust alignment depend on local conditions in clouds. The composition of dust varies within the Milky Way and between galaxies; scientists need to understand why. Observational studies, ranging from microwaves (emission from spinning dust) to X-rays (scattering and absorption by dust), provide a growing array of observational constraints which, together with advances in theory, will result in increasingly realistic grain models during the coming decade.
- *Atomic physics and laboratory astrophysics.* Astrophysics is dependent on accurate wavelengths and oscillator strengths, photoionization and photodissocia-

tion cross sections, and rate coefficients for radiative recombination, dielectronic recombination, charge exchange, and collisional excitation. For example, there appears to be a factor-of-two uncertainty in the oscillator strength for the semi-forbidden line, [C II] 2,325 Å, normally used to determine the gas-phase carbon abundance. Next-generation X-ray facilities, as well as ALMA millimeter-wave studies, will require more accurate wavelengths for lines, as well as more accurate X-ray absorption coefficients for likely astrophysical solids. Some quantities can be obtained from calculations, but others may only be obtained from laboratory measurements. It is reasonable to suspect that polycyclic aromatic hydrocarbons (PAHs) might account for the diffuse interstellar bands, but only careful measurement of PAH absorption cross sections in the gas phase in the laboratory can confirm and quantify this. Laboratory measurements are also needed for photoelectric yields from dust grains over a range of sizes, including PAHs.

As the "microphysics" are better understood and computational capabilities continue to increase, it can be anticipated that during the next decade numerical models will be able to account for the statistical properties of the ISM and star formation in the Milky Way and other galaxies. Only at this point will astronomers be able to claim to understand the formation and evolution of galaxies as stellar systems.

GAN 3. WHAT IS THE FOSSIL RECORD OF GALAXY ASSEMBLY FROM THE FIRST STARS TO THE PRESENT?

How galaxies form and evolve over cosmic time into the forms seen today has long been one of the most compelling of the big questions in astrophysics. Modern cosmology provides a theoretical paradigm for galaxy formation and assembly; astronomers can test these ideas with observations of galaxy properties such as morphology, luminosity, and color, and of the distribution of such properties in populations from nearby galaxies to redshifts that possibly overlap the epoch of reionization. As valuable and insightful as those observations will be, the study of ensembles necessarily obscures the physical processes affecting individual galaxies by providing only average properties. *This is the frontier of galaxy formation studies: not only to determine the observed distribution of galaxy properties but also to explain how they came to be.*

The galactic neighborhood can advance the understanding of galaxies with two essential complements to look-back studies. First, it is only in the galactic neighborhood that galaxies are fully resolved and open to intensive study at high spatial resolution. Second, the galactic neighborhood likely contains stars, and remnants of stars, that formed in early epochs and in small galaxies that will remain invisible at high redshift even in the next decade. This rich fossil record of galactic star-formation histories is preserved in the colors of stellar populations and in the kinematics and abundances of individual stars. The basic astronomical techniques that allow the reading of this record of galaxy formation are now in place: the next decade promises to provide the tools needed to obtain and analyze this record in unprecedented detail and over a representative sample of the galactic neighborhood from the Milky Way to beyond 10 Mpc. In sum, these tools and techniques promise to reveal the full history of galaxy formation in galaxies of many different shapes and sizes, going all the way back to their origins in the first stars.

Fossil Record of Galaxies from Resolved Stellar Populations

The fossil record of whole external galaxies can be read in their stellar color-magnitude diagrams (CMDs) (see Figure 2.5). Progress in reconstructing the histories of local galaxies has been hindered primarily by the limited angular resolution available to resolve individual stars in nearby galaxies. From the ground, studies are limited to the nearest galaxies such as the Magellanic Clouds. Even there, limited angular resolution prevents one from determining the main sequence turnoff of the oldest population (which has the faintest turnoff) in the main bodies of these galaxies (where crowding is severe). The principal gains in the number of galaxies that can be studied in such a way came from the use of the Hubble Space Telescope, which opened up much of the Local Group, including the nearby dwarfs, the low-density regions of M31, and the galactic neighborhood out to 4 Mpc.

The variety in CMD morphologies implies a wide range of star-formation histories, which one can associate with such galaxy properties as gas content, environment, and morphology. One can even examine internal patterns, such as the relationship between stellar populations and spiral arms. With the HST, such work on a small number of galaxies promises a physical understanding of the processes that drive star formation in typical galaxies. However, current data rarely reach below the horizontal branch in these systems, do not include giant early-type galaxies (even the nearest ones lie more than 4 Mpc away), and are generally dominated by crowding errors even well above the completeness limits.

Progress in this field is currently limited by angular resolution and survey volume. With the HST, a general rule of thumb is that photometry of main-sequence stars can be done at surface brightness $V \approx 26$ mag arcsec^{-2} until crowding sets in. This limits one to regions beyond 10 kpc along the minor axis or 25 kpc along the major axis of M31. Any gain in resolution will allow one to penetrate correspondingly further in. Analogously, to do comparable work in galaxies that are 10 times more distant would require a telescope with a diameter 10 times larger than that of the HST. Because the increase in aperture matches the decrease in flux from such stars, and because the increase in the background matches the decrease in pixel scale, the achieved S/N remains the same. Larger-aperture telescopes in space

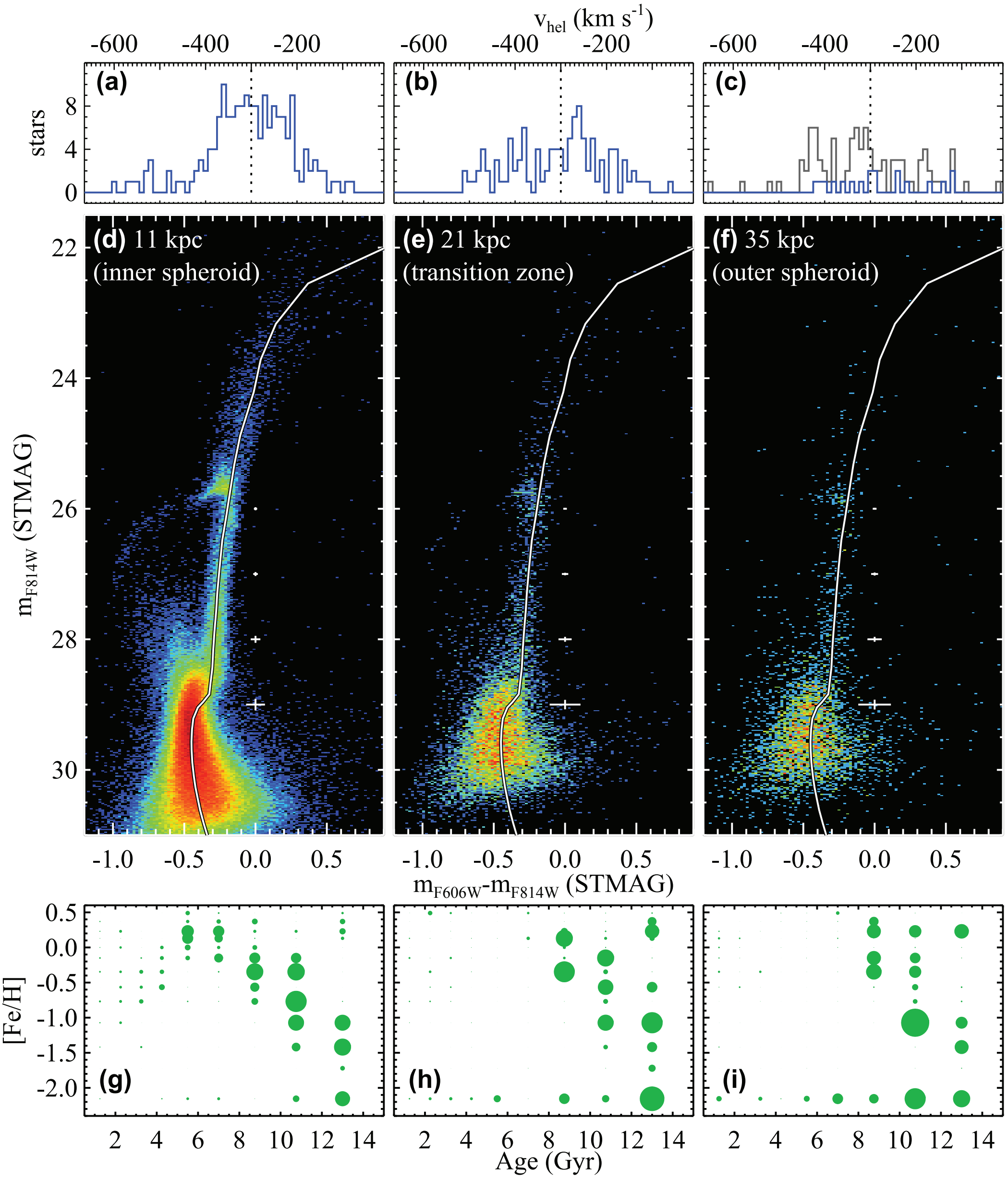


FIGURE 2.5 Reconstruction of star-formation histories of M31. Top panels show radial velocity distributions (M31 systemic velocity is dotted line). Middle panels show stellar color-magnitude diagrams with superposed 47 Tuc fiducial. Bottom panels show reconstructed star formation history. SOURCE: T.M. Brown, R. Beaton, M. Chiba, H.C. Ferguson, K.M. Gilbert, P. Guhathakurta, M. Iye, J.S. Kalirai, A. Koch, Y. Komiyama, S.R. Majewski, et al., The extended star formation history of the Andromeda spheroid at 35 kpc on the minor axis, *Astrophysical Journal Letters* 685:L121, 2008, reproduced by permission of the AAS.

are crucial for obtaining the high angular resolution needed to do the science in crowded fields.

Second, let us examine the case for an increase in survey volume. The two principal drivers for increasing the volume are to probe a variety of environments and to obtain a representative mix of giant elliptical and S0 galaxies that make up much of the stellar mass in the local universe. Access to galaxies out to 10 to 20 Mpc will encompass the Virgo and Fornax clusters as well as numerous galaxy groups (Sculptor, M81, M101). The clusters include ellipticals, and the groups sample different morphological classes across the range of environments; this natural diversity in the galaxy population is inaccessible to the HST. An 8-m space telescope with diffraction-limited capabilities and high point spread function (PSF) stability in the UV and optical would be just able to probe this region, whereas a diffraction-limited 16-m telescope is needed to reach the main sequence effectively. A large space telescope would also dramatically reduce the time required to obtain star-formation histories of nearby galaxies, since the exposure time needed to obtain background-limited photometry on stars at a given distance declines as the fourth power of the aperture.

A number of approaches other than deep, visible-light CMDs can constrain the formation histories of galaxies over these volumes. One of these is to use the red giant branch (RGB) and asymptotic giant branch (AGB) stars as probes of the stellar populations of galaxies. These brighter regions of the Hertzsprung-Russell (H-R) diagram naturally extend the survey volume, but the requirements for angular resolution are still stringent. Adaptive optics (AO) offers dramatic gains in resolution in the near-infrared, where these bright cool stars are best detected, and current work on 8-m telescopes has extended these studies to galaxies at distances of several Mpc. While space telescopes with stable PSFs and low backgrounds are necessary for constraining the star-formation histories with the stellar main sequence at >10 Mpc, a 20-m ground-based telescope with diffraction-limited AO capabilities can reach 2 mag below the main-sequence turnoff in the Local Group, and a 30-m telescope may be able to reach the horizontal branch in the Virgo cluster at 16 Mpc. Stars in late stages of evolution are accessible in the Local Group (≤1 Mpc) with Spitzer and will be studied out to 10 Mpc with the James Webb Space Telescope (JWST). Matching star-formation histories obtained from UV-optical imaging with near-IR and mid-IR data will help unravel the nature of highly evolved stellar populations. The formation history of fair samples of both elliptical and spiral galaxies out to tens of Mpc can also be probed by other stellar population measures, such as globular clusters and spectroscopy of the integrated light of galaxies. These approaches utilize high-quality wide-field imaging and highly multiplexed deep spectroscopy, and they provide both stellar populations and kinematic tests of how individual galaxies in the local universe came to be.

Fossil Record of the Formation of Galactic Halo and Smallest Galaxies

Astronomical techniques for reading the fossil record of the Milky Way have come into their own in recent years in the form of large photometric and spectroscopic surveys. Because stars in the Milky Way can be examined for their individual kinematics and abundances in a way that stars in external galaxies cannot, it can be hoped that a deeper level of understanding can be reached regarding baryonic gas physics and chemical enrichment can be reached, even if the scope is narrowed to one galaxy. The Sloan Digital Sky Survey (SDSS) has made pioneering ventures into the discovery space that will be opened fully by future facilities. In conjunction with follow-up spectroscopy, the SDSS has shown that the Milky Way's stellar halo is a complex structure that belies the traditional view of a smoothly varying density profile. Instead, the stellar halo is now known to possess at least two chemically and kinematically distinct spheroidal components, and within this is a rich array of substructure, including overlapping stellar streams and a dozen new ultrafaint dwarf satellite galaxies (Figure 2.6), some of which have a total luminosity comparable to that of a single giant star. These streams and small galaxies offer the opportunity to gain an understanding of the internal and external influences on galaxy formation at the smallest mass scale and with the most sensitive indicators of such processes as reionization, chemical enrichment, supernova feedback, and tidal disruption. A full unraveling of how gas collapses and forms stars down to these small scales and faint luminosities will illuminate the process of galaxy formation more generally. Lessons learned from local examples can then be applied to an understanding of the early epochs of galaxy formation when large galaxies such as the Milky Way of today were still composed of many such smaller components.

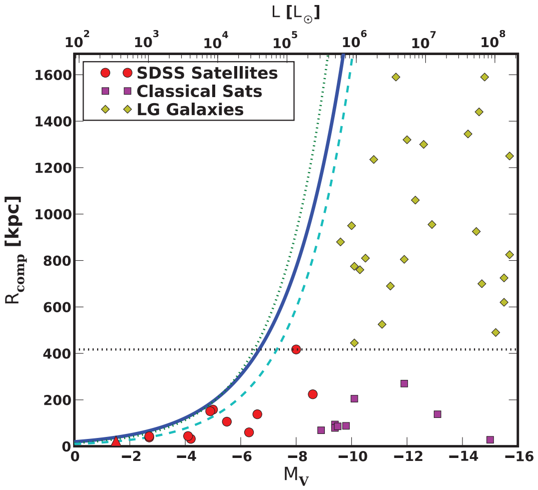


FIGURE 2.6 Detection limits for Local Group dwarf galaxies. The SDSS is only complete to ~50 kpc, while LSST could extend complete sample to 1 Mpc at $M_V = -8$. SOURCE: E.J. Tollerud, J.S. Bullock, L.E. Strigari, and B. Willman, Hundreds of Milky Way satellites? Luminosity bias in the satellite luminosity function, *Astrophysical Journal* 688:277, 2008, reproduced by permission of the AAS.

When corrected for the difficulty of finding such faint objects at large distance, the observed counts of dwarfs imply that there may be hundreds of such dwarf galaxies waiting to be discovered in the Milky Way halo. If so, astronomers will have access to a significant sample of the smallest galaxies, many of which may carry stars formed during the epoch of reionization or even before. Deeper multiband all-sky surveys are needed to extend the search volume for such objects beyond the outer edge of the Milky Way halo and into the Local Group and to turn up additional streams and tidal features left over from the assembly of the Milky Way. At present, 8- to 10-m ground-based telescopes are used to obtain the radial velocities that confirm the structures found in the imaging surveys. The instruments needed to confirm detections by the forthcoming deep surveys will need to be correspondingly larger (20- to 30-m telescopes). Efficient multiobject spectrographs on such telescopes will be needed for measuring the chemical abundances for large samples of stars within these newly discovered galaxies, where 10-m-class telescopes can currently reach only a small handful in the closest faint dwarfs.

Fossil Record of First Stars and Galaxies from Galactic Archaeology

Ground-based observers recognized long ago that surviving remnants of the first stars and galaxies reside today in the Local Group, either directly through long-lived low-mass stars and white dwarfs, or indirectly through chemical abundances in the second-generation stars formed from gas enriched by the first heavy elements. The ability to survey and measure the properties of metal-poor stars throughout the Milky Way and Local Group may lead astronomers to the earliest epoch in the stellar fossil record—the first stars. These objects hold our interest as the ultimate beginning of galaxy formation and chemical enrichment in the universe. Theorists have reached the robust conclusion that the truly first stars formed in small dark matter halos at $z = 10$ to 30 and were likely massive (tens to hundreds of solar masses). Single stars at this epoch, even those with $M > 100$ $M_{\odot}$, are well beyond the reach of our current and forthcoming telescopes such as the JWST. Thus, if the theory is correct, it is certainly very difficult and perhaps impossible to test at high redshift.

A revolution is unfolding, thanks to large photometric and low-resolution spectroscopic surveys such as the Hamburg/European Southern Observatory (ESO) survey (HES) and the SDSS. Spectroscopic follow-up of these surveys may answer many of the open questions about the first stars and galaxies by surveying millions of metal-poor stars in the Milky Way and its dwarf satellites. Hundreds of thousands of metal-poor stars are already known, and hundreds have been subjected to high-resolution spectroscopic follow-up to measure abundances, often to 10 percent precision (Figure 2.7). This pursuit of the fossil record is distinguished from the previous two by a focus on the most metal-poor stars (below 1 percent

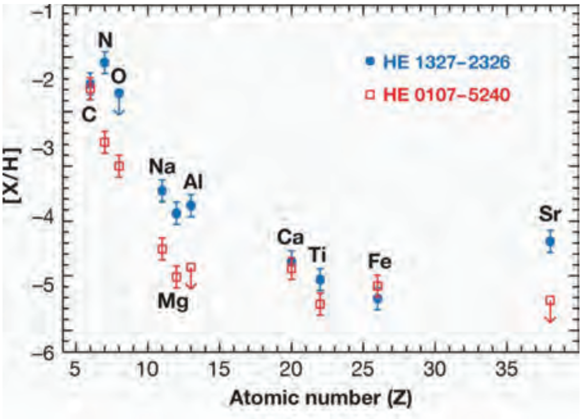


FIGURE 2.7 Abundance patterns for two hyper-metal-poor stars in the galactic halo. Collecting complete abundance patterns with spectroscopy on large telescopes is a key observation. Interpreting these patterns with models of the first stars and chemical enrichment is a theoretical challenge. SOURCE: T.C. Beers and N.B. Christlieb, The discovery and analysis of very metal-poor stars in the galaxy, *Annual Review of Astronomy and Astrophysics* 43:531-580, 2005, copyright 2005 by Annual Reviews, Inc., reproduced with permission of Annual Reviews, Inc. in the format Other book via the Copyright Clearance Center.

solar) and by the goal of obtaining precise abundances for elements from all of the important nucleosynthetic processes that act in stars, from which much information can be obtained about the population of stars that produced the metals. This effort has raised many questions, all ripe for progress in the next decade: How old are the oldest stars in the Milky Way? Where are the lowest-metallicity stars in the Milky Way and when did they form? What were the IMF and chemical yields of the first stars? Did the IMF vary with metallicity and galactic conditions, even after the first stars? When did the IMF become normal and universal? Where are the heavy elements created (particularly neutron-capture elements)? Can chemical tagging of metal-poor stars be used to identify coeval populations, later dispersed around the galaxy?

The first step in the process will be realized by the same large photometric and spectroscopic surveys that map the Milky Way. New data-mining algorithms must be developed to select efficiently the most metal-poor stars from these vast samples, and efficient multiplexed spectrographs on large telescopes will be needed to obtain abundances of the elements that tell with statistical confidence the full story of primordial star formation and chemical enrichment. This field also poses a stiff challenge to theory that must synthesize realistic models of the Milky Way over the full 13- to 14-Gyr history of the universe with mass dynamic range of over 10^7, and calculate chemical evolution in the full abundance space available to modern stellar surveys. This field has a strong link to nuclear astrophysics; supernova yields derived with nuclear physics inputs are crucial for chemical evolution calculations, and the most metal-poor stars promise in turn to reveal the production sites and abundance ratios of the heavy chemical elements and perhaps lead to the discovery of novel supernova mechanisms. Finally, all of the work on resolved stellar populations in the Milky Way could be greatly enhanced by precise astrometric observations for large samples of stars. These measurements could provide precise orbits (see below, the subsection titled "GAN D2. Astrometry as a General Area of Discovery Potential"). This benefit will be enhanced further if future astrometric missions and/or facilities focus their attention on the same regions of the galaxy to be covered by the stellar populations surveys.

In the next decade, the high-redshift frontier of direct look-back studies of galaxies should be advanced beyond $z > 10$ by the JWST, ALMA, and perhaps by some high-z 21-cm experiments. However, theory predicts that there should still be galaxies and stars beyond that frontier. The enormous potential of the fossil record to address this epoch in a complementary way will likely be realized in the next decade, when the synthesis of high- and low-redshift data promise to bring a full understanding of how the stellar components of galaxies are formed and assembled back to the earliest epochs at which stars were formed.

GAN 4. WHAT ARE THE CONNECTIONS BETWEEN DARK AND LUMINOUS MATTER?

Dark matter is the dominant constituent of mass in the universe, but it has only been detected indirectly. Scientists have inferred its existence dynamically in the local universe, and through the perturbations that it has imprinted on the early universe. However, a detailed characterization of the nature of dark matter has defied years of effort, beyond a consensus that it behaves nonrelativistically on large scales and has no significant nongravitational interactions with normal baryonic matter. Here, the panel outlines the importance of the local universe with respect to making progress in one of the greatest unsolved mysteries of modern astrophysics.

Using the Local Universe as a Dark Matter Laboratory

The lambda cold dark matter (ΛCDM) paradigm has had success on a range of astrophysical scales. On the largest scales (>10 Mpc), it can fit the spatial clustering of galaxies, the distribution of temperature anisotropies in the cosmic microwave background (CMB), and the clustering of hydrogen absorption in the Ly-α forest. On smaller scales (0.1-10 Mpc), the density profiles of galaxy halos and cluster dark matter predicted by numerical simulations appear to agree with those inferred from gravitational lensing and galaxy kinematics. On the smallest scales, however, ΛCDM has not yet been verified. These scales offer some of the most sensitive probes of the properties of the dark matter, owing to the close linkage between the innermost structure of collapsed dark matter halos and the high initial phase-space density that characterizes current ΛCDM models. At these scales, there is a well-known tension between the high central densities and cusped profiles predicted by ΛCDM within the centers of galaxies, and the somewhat lower densities and flatter profiles that kinematic observations favor on sub-kiloparsec scales. Furthermore, ΛCDM predicts a large abundance of very small dark matter halos, but to date this substructure remains largely undetected. Whether or not these conflicts require modifications of ΛCDM is a subject of some debate. Although models exist that provide potentially adequate solutions with baryonic physics, these are neither demonstrably correct nor unique.

Complicating these comparisons is the complex interplay between dark matter and baryons within galaxies. The global distribution of baryons within a galaxy is, to first order, controlled by the mass and accretion history of the dark matter. However, the final structure of the dark matter can be strongly affected by the behavior of the baryons, through bars, gravitational contraction, and outflows. Moreover, the baryons can frequently obscure the underlying properties of the dark matter, such as in cases where the baryons dominate the mass in the inner regions of a galaxy.

The effects of baryons must therefore be disentangled before any modification of ΛCDM can be considered truly imperative. As is discussed below, likely approaches include the following: identifying systems where baryons are negligible, constraining the amount of low-mass substructure, improving the understanding of inner halo kinematics, and directly detecting signals from dark matter interactions at high densities. These experiments will inevitably focus on the galactic neighborhood, which is the only environment in which astronomers have sufficient sensitivity and spatial resolution to conduct such tests, including those related to the following questions:

- What is the distribution of dwarf satellite galaxies in the Milky Way? Is it consistent with the predictions of the standard ΛCDM cosmological model, and what does it tell about the relation between the stellar and dark matter substructure?
- What is the distribution of dark matter in the Milky Way? What are the relative densities of baryonic and dark matter, and are they consistent with the standard cosmological model?

What Is the Baryon-Dark Matter Connection at Low Galaxy Masses?

Because lower-mass galaxies appear to be increasingly dominated by dark matter, their kinematics are excellent probes of the dark matter potential. One must find systems of the lowest-possible luminosities and measure their internal kinematics, both radial velocities and proper motions. The first step relies on finding overdensities of stars in deep multicolor photometric surveys and confirming the overdensities with spectroscopy. The SDSS has made substantial progress, primarily within the Milky Way's virial radius. Dramatic increases in identified dwarfs will come with deeper imaging and even-wider-area surveys. The second step requires spectroscopic follow-up of red giant branch stars ($0 > M_I > -4$) with better than

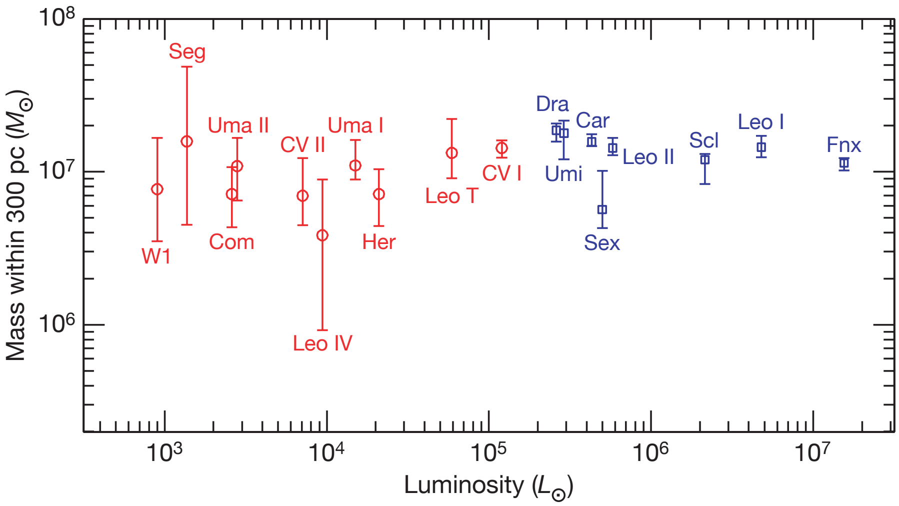


FIGURE 2.8 Mass within 300 pc versus luminosity for the classical (blue dots) and “ultrafaint” (red dots) dwarf galaxies. Remarkably, while the luminosity varies by more than four orders of magnitude in the observations, the mass only varies by factors of order unity in the central region and only 1-2 orders of magnitude for “total” halo mass. SOURCE: Reprinted by permission from Macmillan Publishers Ltd: *Nature*, L.E. Strigari, J.S. Bullock, M. Kaplinghat, J.D. Simon, M. Geha, B. Willman, and M.G. Walker, A common mass scale for satellite galaxies of the Milky Way, *Nature* 454:1096-1097, 2008, copyright 2008.

~1 km s^{-1} resolution, currently possible with 8-m-class telescopes at the distances of the ultrafaint dwarfs found with SDSS (Figure 2.8).

Measuring kinematics becomes challenging for the faintest galaxies, with total luminosities less than that of a single star at the tip of the red giant branch. For such systems, kinematics must be measured for stars of even lower absolute magnitudes than for typical dwarfs with fully populated color-magnitude diagrams.

The baryons in ultrafaint dwarfs are also signposts of processes at the extremes of galaxy environments. Their shallow gravitational potential wells make them susceptible to disruption by feedback from galactic winds, suppression by reionization, and stripping by outflows from nearby massive galaxies. Hints of this fragility can be seen in the variation in mass-to-light ratio of the faintest dwarf galaxies, which span four orders of magnitude in luminosity but have comparable velocity dispersions. The mass limits below which galaxies cannot retain baryons or form stars are currently not known for certain. More-sensitive multiobject spectroscopic or astrometric facilities would open the possibility of measuring hundreds of velocities per galaxy. If extensive spectroscopy is coupled to astrometric measurements in future surveys (~10 μ arcsec), the three-dimensional velocity structure would eliminate uncertainties with respect to the galaxy masses at the limits of galaxy formation. Higher signal-to-noise spectroscopy ($S/N > 20$ at $R \geq 20{,}000$) could constrain the ages and abundance patterns.

How Much Low-Mass Substructure Exists Locally?

ΛCDM predicts an order of magnitude more low-mass dark halos than can be accounted for by the current local inventory of small galaxies (classical and ultrafaint dwarfs). This well-known "missing satellites" problem has two possible solutions. The first appeals to baryonic physics, which can suppress the formation of luminous dwarfs leaving behind a population of dark halos, lacking detectable baryons. The second solution requires the modification of the ΛCDM paradigm itself, which would be necessary if the first solution proves untenable. If more prosaic baryonic physics is responsible, then every ultrafaint dwarf should be the "tip of the iceberg" for a much larger population of invisible dark matter halos. This population is most likely to be detected indirectly, through their gravitational influence on luminous matter: gravitational lensing of background objects, kinematic distortions in apparently isolated galaxies, excess kinematic heating of dwarfs, or seeding of gas accumulations in tidal tails.

Can the Properties of Particle Dark Matter Be Constrained Directly?

If the dark matter is a weakly interacting particle, then it should annihilate with some small probability, producing gamma-ray and/or X-ray radiation in currently favored models. The galactic center is probably the brightest source in the sky for such annihilations, but the astrophysical backgrounds in this region make it difficult to disentangle photons produced by the candidate dark matter particle from those produced by supernova remnants, pulsars, and binaries. The best place to look for the signature of weakly interacting dark matter may be in the heart of the ultrafaint dwarf galaxies. With proportionally fewer stars for their central masses, the astrophysical backgrounds are lower, yet they have high central densities, increasing the rates for any n^2 emission process. Figure 2.9 shows the likelihood for gamma-ray annihilation of a dark matter particle within the Constrained Minimally Super-Symmetric Model; these predicted fluxes are below the current sensitivity of the Fermi Telescope. Convincing detections of emission from dark matter would give the first conclusive maps of the inner dark matter density profiles.

Galactic neighborhood science will also help interpret results from *direct* detection experiments, which are slowly narrowing the possible regimes in which a weakly interacting massive particle (WIMP) dark matter particle can exist. These experiments should prove interesting over the coming decade. The particle physics community has a growing interest in the distribution and kinematics of dark matter within the Milky Way at the solar circle. Such constraints are likely to come

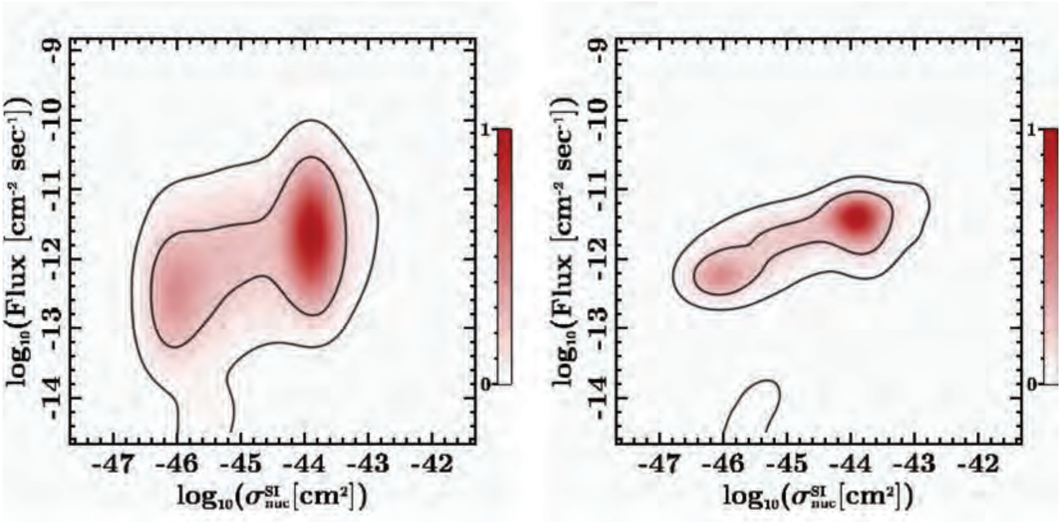


FIGURE 2.9 Contours of likelihood of predicted E > 1 Gev gamma-ray flux versus the dark matter particle cross section, using current data from the dwarf galaxies SEGUE-I (left panel) and Draco (right panel). The inner and outer contours represent the 68 percent and 95 percent confidence regions, with the color and scale bar representing significance normalized to the best-fitted value. These fluxes are below the sensitivity of *Fermi*. SOURCE: G.D. Martinez, J.S. Bullock, M. Kaplinghat, L.E. Strigari, and R. Trotta, Indirect dark matter detection from dwarf satellites: Joint expectations from astrophysics and supersymmetry, *Journal of Cosmology and Astroparticle Physics* 2009:014, 2009.

from a combination of astrometric and kinematic constraints on the Milky Way potential (see below, the subsection titled "GAN D2. Astrometry as a General Area of Discovery Potential") and numerical simulations of the gravitational focusing of dark matter by baryonic disks.

What Are the Mass Contribution and Kinematics of Baryons at Higher Galaxy Masses?

Although dwarf galaxies are superb laboratories for probing dark matter physics, they are only one extreme of the halo mass distribution. It is therefore critical to push the constraints on dark matter structure to higher masses. Doing so requires coping with the population of baryons that dominates the inner galaxy. The current requirements for doing so are (1) two-dimensional kinematic maps of the gas and stars, with arcsecond resolution; (2) accurate constraints on the mass contributed by the stellar and gaseous components, including molecular and hot gas; (3) fully self-consistent kinematic modeling to include the effects of noncircular motions; (4) characterization of the mass distributions at very large radii, using tracers such as globular clusters, distant stars, satellite galaxies, or hot gaseous halos; and (5) constraints on the three-dimensional shape of the halo, using kinematic tracers reaching above the major axis of the galaxy. There are a handful of systems for which more than two of these constraints currently exist.

The first three of these required capabilities can be reached with existing 4- to 8-m facilities, but they require much larger investments of observing time. In contrast, measuring kinematics at large radii and/or off-plane is only marginally feasible with current technology, and it usually requires statistical co-adding of the few available tracers (globular clusters, satellites, sparse stellar clusters, etc). Progress on requirements (4) and (5) above would be accelerated by the addition of spatially resolved X-ray spectroscopy of surrounding X-ray emitting gas with kpc spatial resolution and 100 km s^{-1} spectral resolution, or deployable integral field units on 8-m-class telescopes with arcminute fields of view. Within the Milky Way, three-dimensional large-radius kinematics could be obtained using the velocities and positions of hypervelocity stars, halo RR Lyrae, and blue horizontal branch stars, all of which can be identified in large surveys and followed up astrometrically and spectroscopically. Constraints on the three-dimensional structure of galaxy halos and the full angular momentum distribution of baryons both offer new tests of the predictions of ΛCDM.

A Census of Fundamental Parameters for Black Holes

Massive black holes play an increasingly important position in astronomy and astrophysics. On small scales, their immediate environments probe the extremes of gravitational potentials. On larger scales, massive black holes in galactic nuclei may play major roles in the evolution of galaxies. Over the next decade significant progress is expected in constraining the distribution of the two fundamental properties of black holes—mass and spin—and their relationship to their host galaxies. This progress will come primarily from studies of nearby galaxies and the Milky Way, the only systems that can be observed with the necessary spatial resolution and sensitivity.

What Controls the Masses of Black Holes?

It is now widely accepted that most galaxies host black holes in their nuclei and that the masses of the central black holes tend to be higher in more massive galaxies. However, this simple picture leaves open the basic question, Why? It is currently not known how the central black hole forms, what processes couple the final mass of the black hole to that of the surrounding galaxy, and why some galaxies appear not to have nuclear black holes, or at least appear to have masses below the current sensitivity limits. It is generally accepted that supermassive black holes gain mass from accretion, but the exact process of accretion is far from understood. One of the key questions for understanding the growth of central black holes in galaxies is their initial mass (the seed mass). Models include massive black holes starting as ~10^2 stellar remnants, as 10^4 $M_\odot$ "intermediate-mass" black holes, or as already massive black holes with masses of 10^5 $M_\odot$ and above. Although some evidence suggests that intermediate-mass black hole seeds grow to become supermassive black holes, detections remain sparse and controversial. The local universe offers the most promising avenues for identifying potential seed black holes, by direct detection through dynamical studies of nearby systems, by indirect studies of low-luminosity AGNs, and through the possibility of measuring gravitational waves of black hole inspiral events. Advances in spatial resolution (AO systems) and sensitivity (larger telescopes) will enhance the most-used techniques for measuring black hole masses. Detection of gravitational waves will open up a new avenue for characterizing the demographics of black holes.

The relationship between the mass of the supermassive black hole at the center of a massive galaxy and the surrounding spheroid of stars is one of the most significant discoveries of the past decade. Most current research focuses on the relationship with velocity dispersion or bulge luminosity. These relations have become important benchmarks for theories of galaxy and quasar growth, and it is thus critical to pin down how the masses of central black holes and their host galaxies are related. The local universe provides the anchor for these correlations, with unique capabilities for determining the parameters (slope, normalization, dispersion) of the trends, in much the same way that the Tully-Fisher and Faber-Jackson relationships have been used as diagnostics for galaxy formation as a whole.

Although current studies have opened up this field and provided the basis for a very broad range of work, they are incomplete in a number of key ways. For example, the sample size with highly accurate data is modest, the trends and scatter at low and high stellar masses and dispersion are not well established, and potential systematics in the analysis remain unconstrained. In the next decade, advances in the determination of the black hole relationships will come from a number of areas, including improved spatial resolution using stellar dynamics (from AO-equipped 8-m-class telescopes), more accurate gas kinematics (from AO and improved radio observations of maser disks), improved constraints on kinematics at large radii (requiring longer integrations on 8-m-class telescopes), and more detailed modeling with existing computing facilities (using triaxial and N-body models, and including dark halos).

Finally, significant questions remain at the limits of the nuclear black hole mass function: At the one extreme, what limits the growth of the most massive central black holes? At the other, why do some galaxies appear to have no central black hole at all, even when they host dense nuclear stellar clusters? Is this lack of detection simply a sensitivity issue?

What Is the Distribution of Black Hole Spins?

Understanding the physics of accretion around massive black holes has been a struggle ever since quasars were first discovered. One possible diagnostic of the accretion process is the final spin of the black hole. Observational evidence generally supports the idea that black holes are near maximal spin, although current results are subject to varying interpretations. If true, this result will have significant consequences both for the transfer of angular momentum between the accreting material and the black hole and for the merger histories of black holes. Significant improvements in measures of black hole spin over a wide mass range are possible over the next decade, coming primarily from improved X-ray observations with large collecting areas and good spectral resolution ($\Delta E < 3$ eV at 6 keV) and from variability measures at long wavelengths. Theoretical work related to black hole accretion, such as the proper treatment of magnetohydrodynamics, is also advancing significantly, and the combination of sophisticated numerical modeling with local measures of spin and flux can provide significant advances in the understanding of accretion disk physics.

*The Galactic Center and Sgr A**

The black hole at the center of the Milky Way galaxy is a remarkable object, providing astronomers' best opportunity to observe the environment around a massive black hole and its interaction with the surrounding galaxy. The central parsec of the galaxy demonstrably contains young stars in ordered motion, with a stellar cusp around the central object. These observations have challenged expectations for star formation near a black hole. The existence of young stars in the vicinity of a black hole is hard to understand, given the high stellar density and activity level; this has fueled an active debate as to whether they are formed in situ or fall in from outer radii. The M31 nucleus also shows evidence for very young stars around a black hole, possibly in a disk. Although the Milky Way black hole provides unprecedented detail, other nearby systems need to be explored so that the generality of the trends can be understood. Continued studies in the Milky Way will quantify the range of stellar ages, their orbits, and the properties of the central density cusp (stellar and, potentially, nonbaryonic). These results will reveal how these stars form and how the black hole influences star formation at the galactic center, both of which questions have general implications for the formation of galaxies and stellar associations.

Additional scientific possibilities in the study of the central black hole in the Milky Way would become possible with significant improvements in depth and spatial resolution. These advances would increase the number of stars that could be studied and enable more detailed orbit determinations for stars with smaller pericenters. Broadly speaking, as astronomers measure stars that approach closer to the black hole, they expect to see stronger general relativistic effects and to produce tighter constraints on the central mass distribution (the black hole mass and central cusp). Specific advances that may be enabled include limits on, or the discovery of, a binary black hole, constraints on mechanisms for producing hypervelocity stars, and the possible detection of an event horizon. The needed improvements in spatial resolution and sensitivity could come from high-precision astrometry on systems with higher Strehl ratios (e.g., extreme AO systems) and/or larger-aperture telescopes than are currently available, including improvements in long-wavelength imaging (e.g., submillimeter VLBI images). While important first theoretical steps have been made, truly quantitative predictions require a combination of models of black hole growth that span accretion disk scales as well as galactic scales.

DISCOVERY AREAS

GAN D1. Time-Domain Astronomy as a Galactic Neighborhood Area of Discovery Potential

The panel views the exploration of the transient sky, where enormous swaths of parameter space remain essentially virgin territory, as a potent area of discovery. New areas of parameter space have always led to new discoveries, and there is every reason to think that examining the sky on timescales from nanoseconds to years across the entire electromagnetic spectrum will lead to significant scientific discoveries and new insights. Moreover, the availability of new instruments, along with ever-increasing computational capability and algorithm development, makes the transient sky an area particularly ripe for transforming basic understanding regarding the content of the galactic neighborhood.

The galactic neighborhood is essential for characterizing and interpreting transient phenomena. Measuring the distance, energetics, and demographics of newly observed phenomena is the first essential step in understanding the underlying physics. Transient events observed in nearby galaxies allow one to characterize the luminosities of the events, their rates, and connections with underlying stellar populations and galactic structure with an ease that is not possible in the galaxy where distances are hard to measure, or over cosmological distances, at which high-resolution imaging and spectroscopic follow-up is difficult. Obtaining follow-up spectra is important to enable time-domain studies to reach their full scientific potential.

Time-domain studies of the galactic neighborhood promise a wealth of information on the properties of supernovae, variable stars, late-stage mass loss from evolving stars, binary stars, the disruption of stars near the Schwarzschild radius of central black holes, the flickering of central engines, as well as unanticipated phenomena. Time-domain studies can also address the nature of some unexplained phenomena, such as variable galactic radio sources that have no obvious counterparts at other wavelengths and so-called extreme scattering events seen in the radio. Even well-understood, time-variable stellar phenomena such as RR Lyrae, Cepheids, and long-period variables offer unambiguous tracers of the galactic neighborhood's ancient stellar populations, thereby giving a three-dimensional structure of stellar streams and their orbits, and thus constraining our galaxy's dark matter halo.

In summary, time-domain astronomy within $z < 0.1$ will allow scientists to map the content and evolution of stellar structure in galaxies at a level of detail and precision not easily obtained outside the local universe. It will also secure the base of the distance ladder, providing constraints on H_0 that can take us to a new level of cosmological precision. Insights offered by an expansive view of the local transient sky are likely to lead to a deeper understanding of the structure of dark matter and the values of cosmological parameters.

GAN D2. Astrometry as a General Area of Discovery Potential

Astrometry can open a new window for the discovery of extrasolar planets; discover and characterize vast numbers of Kuiper belt objects, asteroids, and comets; test the weak-field limit of general relativity with unprecedented precision; and measure the aberration of quasars from the centripetal acceleration of the Sun by the galaxy. These surveys can provide a complete inventory of stars near the Sun, with accurate masses for a wide range of stars, particularly for rare objects at the extremes of the Hertzsprung-Russell diagram. They can measure orbits of the globular clusters and satellite galaxies of the Milky Way and galaxies of the Local Group and fix properties of the major stellar components of the Milky Way. The most important tools are large-scale photometric, spectroscopic, and astrometric surveys. Prototypes in past decades were the Two Micron All Sky Survey (2MASS), the SDSS, the Infrared Astronomy Satellite (IRAS), and Hipparcos. Larger, deeper, more accurate surveys have exceptional discovery potential in the next decade, largely from the variety of powerful astrometric techniques now reaching maturity:

- Radio astrometry of masers in massive star-forming regions yields accuracies of a few microarcseconds (μas). These measurements yield the following: (1) accurate (to a few percent) distances and velocities for approximately 20 objects several kiloparsecs away, (2) estimates of the distance to the galactic center and the rotation speed of the local standard of rest that are arguably the most accurate available, and (3) the first proper motions of galaxies other than the Milky Way and its satellites. Radio astrometry of maser sources in AGN accretion disks provides the best masses for black holes at the centers of galaxies, with the possibility of many more, and offers the prospect of determining the extragalactic distance scale (the Hubble constant) with unmatched precision.

- Radio astrometry of the source Sgr A*, believed to coincide with the black hole at the galactic center, now yields the most accurate measurement of the angular speed of the Sun in its galactic orbit, as well as strong constraints on the mass of the black hole and the mass and orbit of any possible companion black hole(s). Infrared astrometry of the stars around Sgr A* proves that it really is a black hole rather than a compact stellar cluster and gives its mass and distance at steadily growing precision.
- Space-based optical astrometry is capable of achieving astrometric accuracies of a few micro arcseconds on hundreds of target stars or ~20 μas for more than 10 million stars, as will be done with the Gaia mission of the European Space Agency.
- Time-resolved ground-based astrometry using large optical surveys can provide proper motions and photometric parallaxes for millions of stars; and can identify unusual high-proper-motion objects that may be faint nearby stars and hypervelocity stars.

SUMMARY OF DESIRED CAPABILITIES

Note that the specifications listed in Tables 2.1 through 2.4 are only approximations of the type of capabilities that would drive galactic neighborhood science in the next decade. This panel was not charged with making detailed analyses of technical capabilities and defers to the Program Prioritization Panels in this regard (see Part II of this volume).

TABLE 2.1 General Capabilities: Theory, Astrometry, and Time-Domain Imaging and Spectroscopy

Capability	Specifications	Scientific Applications
Theory	Support for theory must accompany all major new facilities	Simulations of IGM, CGM, galaxy formation Modeling galaxy feedback Supermassive black hole spin and accretion Microphysics of accretion Connection to large-scale galaxy formation
Astrometry	$\sigma\pi \sim$ 1-10 μas	Mapping the Milky Way Determining fundamental parameters of the Milky Way Determining stellar-mass spectrum Tests of general relativity around Sgr A* Precision measurement (1-3%) of H_0
Time-domain imaging	Multiple visits, $N \sim 100$	Cepheid/RR Lyrae studies in Local Group Optical transients, radio transient sky (almost totally unexplored)
Time-domain spectroscopy	$R \sim 5{,}000$ resolution	Black hole masses

NOTE: Acronyms are defined in Appendix C.

TABLE 2.2 Capabilities: Short-Wavelength Bands (UV, X-Ray, Gamma-Ray)

Capability	Specifications	Scientific Applications
Gamma-ray spectroscopy	$F \sim 10^{-13}$ cm^{-2} s^{-1} (broadband, 10% resolution)	Dwarf galaxies and the dark matter particle
X-ray imaging spectroscopy	0.3-8 keV at $R \approx 300$ FOV ~ 20′ and $\Delta\theta \sim 5''$ $I_{\mathrm{lim}} \sim 10^{-14}$ ergs s^{-1} cm^{-2} arcmin^{-2}	Mapping hot gas outflows, supernova remnants, superbubbles
	0.5-7 keV at 1-3 eV resolution	Mapping hot plasma in SNRs, stellar winds, superbubbles
	Line detection sensitivity ~ 10× Chandra 0.5-7 keV	Black hole mass measurements Dark matter indirect detection
X-ray absorption spectroscopy	0.1-1 keV $R \approx 3{,}000$ $A_{\mathrm{eff}} \sim 1{,}000$ cm^2	Abs-line spectroscopy of hot CGM Abs-line spectroscopy of hot ISM Dust grain mineralogy Inventory of abundant elements in ISM; obtaining kinematics of gas
UV imaging	1,000-3,000 Å $R \approx 1{,}000$-$2{,}000$ FOV ~ 20′ and $\Delta\theta \sim 1''$ $I_{\mathrm{lim}} \sim 100$ LU (ph s^{-1} cm^{-2} sr^{-1}) Now: STIS, GALEX Needed: $\Delta\theta \sim 1''$	Mapping warm ISM and CGM Mapping galaxy outflows Mapping hot plasma in the ISM and at the disk-halo interface Mapping star-forming regions
UV spectroscopy	900-3,000 Å at $R \approx 30{,}000$ $A_{\mathrm{eff}} \sim 10^4$ cm^2 Now: STIS and COS (1,150-3,000 Å) Need: <1,030-3,000 Å $R = 30{,}000$-$100{,}000$ to $m_{\mathrm{lim}} \sim 19$	Abs-line spectra of warm ISM, CGM Diagnose hot plasma in ISM and at disk-halo interface Tomography of ISM, CGM Composition and kinematics of both hot and cold gas, excitation of H_2

NOTE: Acronyms are defined in Appendix C.

TABLE 2.3 Capabilities: Optical and Infrared Bands

Capability	Specifications	Scientific Applications
Optical polarimetry	B, V, R, I imaging polarimeter on large telescope $\Delta p < 0.05\%$ for $m_{AB} = 18$	Mapping magnetic field in diffuse ISM
IR polarimetry	J, H, K imaging polarimeter $\Delta p < 0.05\%$ for m_{AB} (K) = 18	Mapping magnetic field in dense clouds
UV/optical/IR imaging	0.01-0.1″ spatial resolution (HST successor) V ~ 35 or 0.5 mag below main-sequence turnoff Two-color photometry 5-10′ FOV Stable PSF and low backgrounds	CMD-based star-formation histories for galactic neighborhood galaxies (Q3.1, Q4)
	Substantial sky coverage (SDSS successor) V ~ 18-28	
	Multiband photometry $\Delta\theta = 0.04''$ $m_{\mathrm{lim}} \sim 25$ FOV ~ 10′	Locating and identifying faint Milky Way halo substructure (Q3.2, Q4) including streams and dwarf galaxies
	$\Delta\theta = 0.04''$ $m_{\mathrm{lim}} \sim 21$	Probing Milky Way black hole Proper motions of Milky Way stars
Optical/near-IR spectroscopy	R = 10,000-40,000, V ~ 18-19 500-1,000 multiobject spectrograph 10^5-10^6 stars	Kinematics and abundances in Milky Way halo, dwarfs, and streams (Q3.2); power is in sample sizes, and so efficient multiplexing matters
	R = 40,000, V ~ 18-20	Precise (0.1 dex) relative abundances in individual EMP stars (Q3.3) to assess star-formation history and IMF in early galaxies
	R = 2,000 to V ~ 24 10′ FOV, 100 multiobject spectrograph	Kinematics of resolved populations in external galaxy halos (Q3.1) important to separate overlapping spatial components
	R = 400-2,000 V ~ 18 Large FOV (~all-sky coverage) Obtain relative abundances to moderate precision	Efficient preselection of EMP star candidates (Q3.3) based on spectroscopically measured bulk metallicity
NIR spectroscopy	$R \approx 3{,}500$ $\Delta\theta = 0.04''$ 1-3 μm	Black hole mass measurements

NOTE: Acronyms are defined in Appendix C.

TABLE 2.4 Capabilities: Long Wavelengths (IR, Far-IR, Submillimeter, Radio)

Capability	Specifications	Scientific Applications
Far-IR imaging and spectroscopy	Now: Herschel (3.5-m warm telescope)	Mapping dust distribution in nearby galaxies with ~5″ resolution
	Needed: 5- to 10-m cold telescope	Study of major far-IR lines (63 μm [O I], 158 μm [C II], 205 μm [N II])
	2× better angular resolution >10^3× greater sensitivity	
Submillimeter imaging and spectroscopy	Now: ALMA (bands 7-9) for high resolution	High-resolution line and continuum images of star-forming regions' high-excitation molecules, dust
	Needed: ~25-m telescope at high site	Mapping star-forming clouds at galactic center
	$\Delta\theta < 3''$ at 350 μm	Dust maps for nearby galaxies
	350 μm-1 mm	Potential for imaging black holes
	Polarimetry (linear) would be valuable	
Millimeter-wave imaging	Approved: ALMA	High-resolution images of dust in star-forming regions
	Needed: CARMA array receivers (1 mm)	High-resolution images: molecular gas in star-forming regions
		Connecting galactic to extragalactic star formation
Millimeter-wave line and continuum mapping	Now: ALMA in S, CARMA in N	Distribution of CO and other molecules in star-forming galaxies
	Needed: CARMA array receivers (1 mm)	Zeeman measures of $B_{\|\|}$ from CN
	Needed: LMT large-format array receivers	Molecular-cloud properties in different galactic environments
Centimeter-wave line and continuum mapping	Now: EVLA, GBT, Arecibo, ATA	Improved H I maps, deep extragalactic imaging
	Needed: upgrade Arecibo and GBT to array receivers	Zeeman measurement of $B_{\|\|}$ using H I, OH
		Mapping synchrotron and free-free emission
	Needed: upgrade ATA to 128 antennas	Identifying and mapping new heavy molecules and transitions
Gravity waves	Detection	Black hole masses

NOTE: Acronyms are defined in Appendix C.

3

Report of the Panel on Galaxies Across Cosmic Time

> I get wisdom day and night
> Turning darkness into light.
>
> —St. Paul Irish Codex, translated by Robin Flower

SUMMARY

The study of galaxies across cosmic time encompasses the main constituents of the universe across 90 percent of its history, from the formation and evolution of structures such as galaxies, clusters of galaxies, and the "cosmic web" of intergalactic matter, to the stars, gas, dust, supermassive black holes, and dark matter of which they are composed. Matter accretes into galaxies, stars form and evolve, black holes grow, supernovae and active galactic nuclei expel matter and energy into the intergalactic medium (IGM), galaxies collide and merge—and what seemed a static world of island universes only a few decades ago turns out to be a lively dance of ever-changing elements. Across all epochs, these processes are coupled in a complicated evolutionary progression, from the relatively smooth, cold universe at high redshift ($z > 40$ or so) to the highly structured cosmos of galaxies and intergalactic matter today.

The Astro2010 Science Frontiers Panel on Galaxies Across Cosmic Time began its deliberations by reading the extensive set of white papers submitted by the astronomical community to the National Research Council (NRC) at the request of the Committee for a Decadal Survey of Astronomy and Astrophysics. The panel reviewed the substantial advances in the understanding of galaxy and structure evolution that have occurred over the past decade or two. It then identified the four key questions and one discovery area that it believes will form the focus for research in the coming decade:

- How do cosmic structures form and evolve?
- How do baryons cycle in and out of galaxies, and what do they do while they are there?
- How do black holes grow, radiate, and influence their surroundings?
- What were the first objects to light up the universe and when did they do it?
- *Unusual discovery potential:* the epoch of reionization.

To maximize progress in addressing these issues, the panel considered the wide array of observational and theoretical programs made possible by current or future facilities. Observational programs were discussed in sufficient detail to allow an understanding of the requirements (numbers of objects, sensitivity, area, spatial resolution, energy resolution, etc.) so that this panel could provide the most useful input to the study's Program Prioritization Panels (PPPs; see the Preface for further information on this process); however, any assessment of the suitability of existing or proposed facilities to the key science issues outlined here is left to the PPPs and the survey committee.

This report describes the scientific context for the area "galaxies across cosmic time," and identifies the key science questions in this area for the next decade and a set of science programs—observational and theoretical—that will answer the most important questions in the field. Some of these programs would require new observational facilities, whereas others could be done with existing facilities, possibly with a reprogramming of resources. In order to provide more useful input to the Astronomy and Astrophysics 2010 (Astro2010) Survey, the top science programs selected by the panel for purposes of this report are identified in three categories: most important, very important, and important. The panel considered many other programs that were eventually excluded from its list but that remain valuable ways to make progress, and it anticipates that significant progress will also come from unexpected directions.

This Summary addresses each of the four key questions in turn, listing only the programs ranked "most important," plus those "very important" activities that represent unique capabilities. The full set of the panel's top-ranked science programs is summarized in Table 3.1 at the end of this "Summary" section and is presented with rankings and further details in the body of the report.

How Do Cosmic Structures Form and Evolve?

The answer to this question starts with an understanding of the structure of dark matter halos on all scales. The now-standard lambda cold dark matter—ΛCDM—cosmology provides a detailed foundation on which a theory of galaxy formation and evolution can be built and which in turn can be tested by data.

ΛCDM does seem to be validated on the largest scales of the cosmic web and superclusters, but some of its predictions seem to deviate seriously from observations on smaller scales, from clusters of galaxies, down to galaxies themselves. Specifically, theory predicts that, even after small clumps of dark matter have merged to form ever-larger structures, many of the small clumps should survive intact, embedded within the merged halos. Yet observations appear to indicate that the dark matter in halos is much less "lumpy" than predicted by the straightforward calculations. Direct constraints on the dark matter distribution can be derived from observations of gravitational lenses, both weak and strong. The panel therefore concluded:

- It is *most important* to obtain Hubble Space Telescope (HST)-like imaging to determine the morphologies, sizes, density profiles, and substructure of dark matter, on scales from galaxies to clusters, by means of weak and strong gravitational lensing, in lens samples at least an order-of-magnitude larger than currently available. HST can make an important start on this problem, but to develop large statistical samples will require a much larger field of view or more observing time than HST affords.

The best current calculations of cluster formation suggest that gas in the densest regions should cool more than is observed, and that more stars should form in cluster cores, especially in the richest clusters. Perhaps the physical processes that affect baryons in clusters need to be better understood, or perhaps extra energy is injected from supernovae, an active nucleus, or some other source. One critical missing piece of information concerns the dynamics of the hot intracluster gas: how turbulent is the gas, how does it flow through the cluster, what is its ionization and velocity structure, and how do these properties depend on cluster richness and cosmic epoch (redshift)? The panel concluded:

- High-energy-resolution, high-throughput X-ray spectroscopic studies of groups and clusters to $z \sim 2$ are *most important* for understanding the dynamics, ionization and temperature structure, and metallicity of the hot intracluster gas, as well as for studying the growth of structure and the evolution of correlations among cluster properties.

Much is still not known about how galaxies were assembled. The well-defined correlations observed among the shapes, sizes, velocity structures, and compositions of galaxies, observed mainly in the local universe, are poorly understood. A Sloan Digital Sky Survey (SDSS)-size spectroscopic survey at $z \sim$ 1-3 would provide essential information about the evolution of galaxy correlations and should provide essential clues to the process of galaxy formation and evolution. The panel concluded:

- It is *very important* to obtain moderate-resolution multi-slit spectroscopy of SDSS-size galaxy samples at $z \sim 1$-3, in the optical for $z < 1.5$, and in the near-infrared (IR) for $z > 1.5$ (with resolution $[R] \sim 5{,}000$ to allow effective removal of night skylines in the near-IR). For a representative subset of hundreds of galaxies, high-angular-resolution integral field unit (IFU)[1] spectroscopy in the optical or near-IR would help calibrate the slit spectra. To select targets for spectroscopy requires optical/IR pre-imaging over a large area.

How Do Baryons Cycle in and out of Galaxies and What Do They Do While They Are There?

Along with galaxies, clusters, and dark matter, diffuse baryonic gas is a key part of the cosmic web; indeed, it represents most of the baryonic mass in the universe. The metal enrichment of the gas indicates that a great deal of it was processed through stars in the past, yet little is understood about how galaxies acquire gas across cosmic time, convert it to stars, and eject it back into the IGM. To understand this process will require the kind of detailed study of galaxies in the young universe that was done for the local universe with large surveys such as the SDSS and the Two-degree Field Galaxy Redshift Survey.

To create a full evolutionary picture for galaxies, study of the following is needed: the star-formation rate, active galactic nucleus (AGN) activity, star-formation history, stellar mass, and stellar and gas-phase metallicity in galaxies at $z \sim 1$-3, when cosmic star formation and black hole growth rates peaked. By quantifying the correlations of these properties with one another and with the larger-scale environment, astronomers can trace the evolution of galaxies and the baryons within them from the galaxies' origins to the present day. These detailed galaxy properties are accessible through rest-frame optical spectra that have sufficient resolution to measure dynamical and stellar population parameters, sufficient continuum sensitivity to measure absorption lines, and sufficient emission-line sensitivity to measure low levels of star formation (see Figure 3.13 below in this report). The galaxy samples must be large enough to disentangle the covariances among galaxy properties such as luminosity, mass, age, morphology, and metallicity, over volumes large enough to sample representative galaxy environments. A wide-area survey would trace luminous galaxies, while a smaller-volume survey could probe deeper in order to study the fainter progenitors of typical galaxies today.

To develop a complete view of galaxies in the peak epoch of galaxy formation, comparable to the understanding of galaxies in the local universe, the panel concluded:

---

[1] Integral-field units provide spatially resolved spectroscopy, usually across a contiguous field.

- It is *most important* to carry out an SDSS-size near-infrared spectroscopic survey of galaxies at $1 < z < 3$ using multiobject spectrographs. This will require near-infrared advance imaging in the J, H, and K bands (at 1, 1.6, and 2 microns) to select targets for spectroscopy. Properly designed, the same large near-IR spectroscopic survey could serve the first key question as well.

To probe baryons when they are in and around and between galaxies, one can use absorption spectra of background sources along lines of sight passing near galaxies. Such techniques probe both the gas distribution and its velocity field and will yield insights into gas accretion, outflows, chemical enrichment, and the overall cycle of matter between galaxies and the IGM. Theoretical simulations will be critical for connecting such one-dimensional probes to the three-dimensional gas distribution. At $z < 1.5$, the principal absorption lines of gas outflowing from galaxies and quasars are in the ultraviolet (UV). UV absorption-line spectroscopy also provides an alternative to X-rays in searching for the "missing baryons" thought to comprise a warm-hot intergalactic medium (WHIM). It may also be possible to image the WHIM directly using IFUs in the UV. The panel therefore concluded:

- It is *most important* to use extremely large optical/infrared telescopes (ELTs) to map metal- and hydrogen-line absorption from circumgalactic and dense filamentary intergalactic gas, at moderate resolution toward background galaxies and at higher resolution toward background quasars.
- A 4-meter-class, UV-optimized space telescope, equipped with a high-resolution spectrograph and an IFU for spectral mapping, is *very important* for characterizing outflows from galaxies and AGN at $z < 1.5$ and for mapping the WHIM.

A complete inventory of cold gas in and around galaxies is also crucial for understanding baryon cycling. Molecular gas traced by carbon monoxide (CO), neutral carbon atoms (C I), and higher-density probes provides the raw material for star formation. Neutral atomic gas in the circumgalactic medium likely feeds the growth of galaxy mass. Direct observations of cold gas will make it possible to test theoretical models for complex gas physics and predictions for the evolution of gas content. For the construction of a seamless picture of how gas is processed into stars during the epoch $1 < z < 3$, when roughly half of the stellar mass in the universe was formed, a complete inventory of cold gas in and around galaxies is needed. The panel therefore concluded:

- It is *most important* to detect CO emission from a representative sample of typical star-forming galaxies from $z \sim 1$-3, to develop technology for faster spectroscopic follow-up in the (sub)millimeter, and to develop large-collecting-area facilities to study neutral hydrogen (H I) in emission at $z \sim$ 1-3.

Accurately characterizing star formation that is obscured by dust is critical for obtaining a complete view of baryon processing during the epoch of galaxy formation. Complementing $z \sim$ 1-3 rest-frame optical spectroscopy with radio and submillimeter imaging, as well as far-IR spectroscopy of the dustiest systems, would provide a complete synthesis. The panel concluded:

- It is *very important* to do sensitive radio and (sub)millimeter continuum mapping over large areas, preferably coincident with a near-IR (rest-frame optical) spectroscopic survey such as the one described above, and to carry out far-IR spectroscopy of luminous dusty galaxies.

How Do Black Holes Grow, Radiate, and Influence Their Surroundings?

Supermassive black holes (SMBHs), a prediction of Einstein's general theory of relativity, are ubiquitous within our galaxy and throughout the universe. Observations over the past decade suggest that they play an important role in the evolution of galaxies and clusters. It is still uncertain how and when these black holes form, grow, produce relativistic jets, and feed energy back into the environment. The strong correlation between black hole mass and galaxy mass hints at tightly coupled coevolution and possibly a strong regulatory effect of one on the other. In galaxy clusters, there is equally intriguing evidence that energy liberated by accreting black holes—carried by jets or winds—regulates the thermal evolution of the intracluster gas.

Gas swirling into SMBHs in luminous AGN apparently forms a nearly Keplerian, thin accretion disk, much as predicted more than 30 years ago. X-rays reflected from the disk are imprinted with spectral signatures that encode the dynamical state of the gas and the relativistic curvature of space-time around the black hole. Coupled with sophisticated computer simulations, these signatures can be used to probe the physics of black holes and accretion disks directly and to determine the spin distribution function of the local SMBH population. The structure of AGN accretion disks and jets can also be explored through X-ray polarization measurements.

To understand the details of accretion onto supermassive black holes, jet formation, and energy dissipation, the panel concluded:

- It is *most important* to have sensitive X-ray spectroscopy of actively accreting black holes (AGN) to probe accretion disk and jet physics close to the black hole as well as to determine the spin distribution function of the local SMBH population. The effective area should be sufficient to detect the iron Kα emission line on dynamical timescales in a modest sample of the brightest AGN, yielding both spin and mass. To disentangle the effects of absorption in AGN spectra, high resolution ($R > 2{,}000$) is required. The same capabilities will yield time-averaged line profiles of more than a hundred AGN with sufficient signal-to-noise ratios to derive the black hole spin distribution.

Most of the evidence for black hole feedback into the intracluster medium (ICM) is either morphological or based on low-spectral-resolution temperature measurements. But since such feedback is thought to occur primarily by way of the kinetic energy of jets and winds, kinematic measurements would provide a more direct test. High-throughput, high-resolution X-ray spectroscopy will reveal bulk motions and turbulence in the ICM, allowing the AGN/ICM coupling to be explored. In order to seek evidence of black hole feedback, the panel concluded:

- It is *most important* to measure turbulence and/or bulk flows using X-ray imaging spectroscopy of the ICM of nearby galaxy clusters and groups, with sufficient image quality, field of view, energy resolution, and signal to noise to provide ionization and velocity maps on the scale of the interaction between the AGN outflow (e.g., radio source) and the gas.

A census of black holes across cosmic time is fundamentally important for understanding when and how black holes formed and grew and for assessing whether the energy liberated is adequate to the feedback task. Various multiwavelength survey techniques have been effective at sampling large fractions of the SMBH population, although no one technique yields a full census. Hard X-rays are most directly connected to the energy-generation mechanism in AGN and penetrate all but the highest line-of-sight column densities; IR observations effectively capture radiation reprocessed by dust; optical narrow-line surveys find faint AGN, even when heavily obscured; and radio surveys are completely insensitive to obscuration and can readily detect jets. The panel concluded:

- It is *very important* to do complementary multiwavelength surveys to track the growth of black holes across cosmic time. A hard X-ray, all-sky survey for AGN is an essential complement to the deep pencil-beam surveys of active galaxies expected from the upcoming NuSTAR Explorer. Long-wavelength IR surveys capture the total energy output, and rest-frame optical spectroscopic surveys allow black hole mass determinations.

The next decade offers the prospect of detecting gravitational radiation from merging SMBHs in the 10^5 to 10^7 M_{Sun} range out to $z \sim 10$. While the restricted mass range and the possibility of small-number statistics will prevent a detailed reconstruction of the SMBH merger tree, such observations can discriminate between small- and large-seed scenarios for early SMBH growth and determine the masses and spins of some objects. The panel concluded:

- The search for gravitational radiation from merging supermassive black holes, at lower frequencies than are probed with the Laser Interferometer Gravity Observatory, is *very important* for an understanding of the buildup of supermassive black holes.

What Were the First Objects to Light Up the Universe and When Did They Do It? —and Discovery Area: The Epoch of Reionization

Concerning the first objects to light up the universe, when and where did these objects form? When did the first galaxies emerge and what were they like? How was the universe reionized? This very early phase of galaxy evolution occurred during the epoch of reionization, which the panel designates as its discovery area because of its great discovery potential. This epoch lies at the frontier of astronomy and astrophysics for the next decade.

The first objects to light up the universe could be stars, black holes, galaxies, and/or something less obvious, such as dark matter annihilation. What these objects are and when and where they formed are almost completely unknown. They and subsequent generations provided enough light to reionize the universe by a redshift of $z \sim 6$, but the topology of the ionization is unconstrained at present. The expectations of astronomers are guided almost entirely by theory.

The first stars should have been essentially metal-free and extremely massive ($M > 100\ M_{Sun}$), with a radiation field that is very efficient at ionizing hydrogen and helium. For redshifts $z < 11$, key emission features will appear in the J band. While individual stars will be much too faint (AB ~ 38-40) to be detected directly with the James Webb Space Telescope (JWST) or an ELT, aggregates of stars may be visible in JWST deep fields, especially with the aid of gravitational lensing. Hypernovae and/or gamma-ray bursts (GRBs), which may be the first individual stellar objects to be observed, can be found through time-domain surveys. GRBs can be used as a probe of the high-redshift intergalactic medium provided that several dozen with $z > 8$ are detected; this would take several years for a facility with an order-of-magnitude-higher detection rate (which depends on the product of field of view and sensitivity) than is possible with the Swift satellite.

To find and characterize the first-generation aggregates of stars, the panel concluded:

- It is *most important* to use JWST to make deep surveys, followed up with near-IR spectroscopy on an ELT.

- It is *very important* to develop a next-generation GRB observatory to search for the first explosions, with an order-of-magnitude-greater GRB detection rate than is possible with Swift, augmented by a rapid follow-up capability for infrared spectroscopy of faint objects.
- It is *very important* to do time-domain surveys to identify the first stars from their supernova or hypernova explosions.

One of the most tantalizing probes of the epoch of reionization is the redshifted 21-cm H I line. Ionization pockets in the cold intergalactic gas are expected to cause fluctuations in the 21-cm brightness temperature. Existing experiments (e.g., the Low Frequency Array for radio astronomy [LOFAR], the Murchison Widefield Array [MWA]) may be able to detect these fluctuations to $z \sim 10$. Ultimately, with future large-area low-frequency radio arrays, it should be possible to map the entire history of reionization by means of an all-sky map of redshifted 21-cm emission.

Absorption-line spectroscopy along sightlines toward the first stars, GRBs, or supernovae will allow the detection of the presence of metals and the ionization level throughout the epoch of reionization. Such observations require the collecting area and spectroscopic capability of an ELT.

The panel concluded that to explore the discovery area of the epoch of reionization:

- It is *most important* to develop new capabilities to observe redshifted 21-cm H I emission, building on the legacy of current projects and increasing sensitivity and spatial resolution to characterize the topology of the gas at reionization.
- It is *very important* to do near-infrared absorption-line spectroscopy with JWST, ELTs, and 10-m-class telescopes to probe the conditions of the IGM during the epoch of reionization.

Although this discussion has so far focused on the first objects, it is very important to find and identify objects residing in the later stages of the epoch of reionization, including radio-loud AGN, quasars, galaxies, supernovae, and GRBs. The panel concluded:

- It is *very important* to do multiwavelength surveys to detect galaxies, quasars, and GRBs residing in the late stages of reionization at $6 < z < 8$, including near-infrared surveys for galaxies and quasars, hard X-ray or gamma-ray monitoring for GRBs, and time-variability surveys for supernovae or hypernovae.

Theory and Laboratory Astrophysics in the Next Decade

Underlying all of astronomy and astrophysics is critical work in theory and other intellectual infrastructure, such as laboratory astrophysics. Theory is at the heart of astronomical inference, connecting observations to underlying physics within the context of a cohesive physical model. The past decade has seen great advances in theoretical aspects of galaxy formation and black hole astrophysics, particularly in the computational arena, which is driven by technological advances (much as with observations). To understand the universe better, to reap the full value of new observational capabilities as they become available, and to guide the next observations, the panel concludes that investments are needed in the following theoretical areas:

- *Cosmological context.* Hydrodynamical simulations within a hierarchical structure-formation context, expanding the dynamic range to study detailed galaxy and cluster assembly within a representative volume.
- *Galactic flows and feedback.* Central to galaxy assembly; requires understanding of the associated two-phase interfaces and instabilities as gas moves through the inhomogeneous intergalactic medium and of how energy, momentum, and relativistic particles feed back into ambient gas.
- *Magnetohydrodynamics (MHD) and plasma physics.* Studies of how magnetic fields channel and transport energy over a large dynamic range, including developing a better understanding of magnetic reconnection, particle acceleration, and cosmic-ray transport.
- *Radiation processes.* Coupling radiative transfer models to dynamical galaxy-formation simulations, and incorporating radiation hydrodynamics and nonthermal processes into models of jets and accretion disks.

Summary of the Panel's Conclusions

The panel's conclusions and top-rated science programs are summarized in Table 3.1.

INTRODUCTION

Galaxies are complex systems that evolve dramatically across cosmic time (Figure 3.1). Their critical constituents—not only stars, gas, and dust, but also supermassive black holes and dark matter—are strongly coupled to one another. During the past decade scientists have learned that no galaxy is an island: they constantly influence and are influenced by their environments.

TABLE 3.1 Summary of Conclusions of the Panel on Galaxies Across Cosmic Time

	Key Questions			
	How do cosmic structures form and evolve?	How do baryons cycle in and out of galaxies, and what do they do while they are there?	How do black holes grow, radiate, and influence their surroundings?	What were the first objects to light up the universe and when did they do it?
Existing or approved facilities and *Important Observational Programs*	SPT, ACT, VLBA, EVLA, e-Merlin, ALMA, HST, Chandra, Spitzer, JWST, eROSITA-SRG, Astro-H, *HST-like imaging of gravitational lenses, OIR MOS and IFU spectra of galaxies, black hole masses with ELT/ radio, long-baseline radio imaging of gravitational lenses, deep multiband cluster surveys*	ALMA, EVLA, JWST, Herschel, 8- to 10-m OIR telescopes, *SDSS-like near-IR (rest-frame optical) spectroscopic survey, CO emission in high-z star-forming galaxies, radio continuum survey, far-IR spectroscopy of luminous dusty galaxies, JWST IFU spectra*	Chandra, XMM-Newton, Suzaku, NuStar, GEMS, Astro-H, VLBA, (sub) millimeter VLBI, Fermi, VERITAS, HESS, MAGIC, HST, *IR and O III surveys, time-resolved UV spectra with HST COS, multiwavelength surveys for AGN, multiwavelength spectra blazars/AGN*	HST, JWST, Planck, LOFAR, MWA, ALMA, EVLA, 8- to 10-m telescopes, *time-domain surveys*
New facilities needed	XRSO,[a] high-resolution optical imaging of lenses, OIR MOS and IFUs on 8- to 10-m telescopes, ELT[b]	Highly multiplexed near-IR MOS on 8- to 10-m, high-z H I 21-cm, ELT, 4-m UV space telescope, far-IR spectra, large (sub) millimeter telescope, XRSO, ELT IFU	XRSO, hard X-ray survey telescope, GWO,[c] ELT, X-ray polarimeters	ELT, sensitive high-z 21-cm capability, GRB observatory, hard X-ray and gamma-ray and optical time-domain survey telescopes, CMB E-mode polarization, new ALMA receivers
Theory needed	Hydrodynamical simulations of Gpc^3 volumes, studies of magnetohydrodynamics and plasma physics, studies of gas flows and feedback, calculations incorporating radiative transfer, dynamical simulations within the hierarchical structure-formation context.			

NOTE: Red, blue, and green are associated, respectively, with the scientific investigations that the Panel on Galaxies Across Cosmic Time deemed most important, very important, and important for addressing the key questions that it identified. Black indicates existing projects or facilities. *Italics* indicate observational programs. Acronyms are defined in Appendix C.

[a] High-throughput high-resolution X-ray spectroscopy observatory.

[b] Extremely large (20- to 40-m) ground-based optical/infrared telescope, with adaptive optics.

[c] Gravitational wave observatory sensitive in the 10^{-1} to 10^{-3} Hz range.

The understanding of galaxies and galaxy evolution has changed radically over the past two decades, thanks to major advances in instrumentation across the electromagnetic spectrum, leaps in conceptual and computational theory, and innovative observational programs. These exciting advances can be illustrated with just a few recent examples:

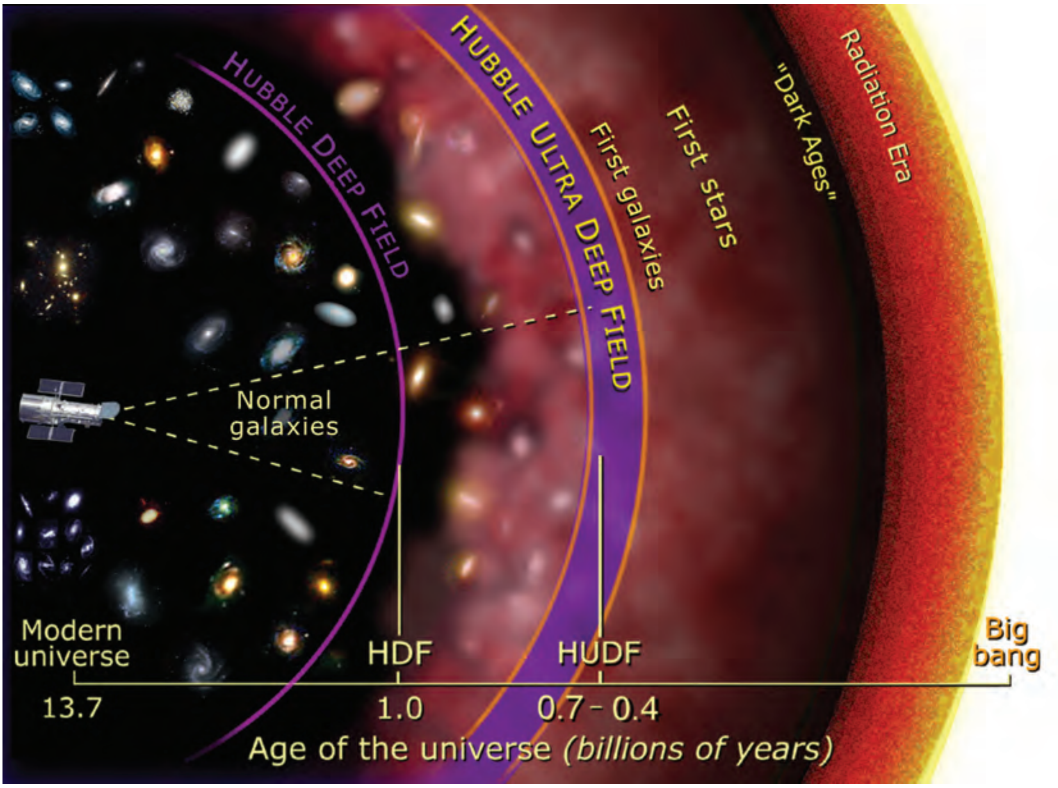


FIGURE 3.1 The universe of galaxies today has developed over the 13.7 billion years since the big bang. Images from the Hubble Space Telescope's deep surveys extend from the nearby cosmos to the current frontier of the observable universe at $z \sim 6$, and even beyond in a few fortuitous cases (e.g., when enhanced by gravitational lensing). Detection of the first objects and first structures condensing out of the "Dark Ages" is a key goal for the coming decade. SOURCE: Courtesy of NASA and A. Feild (STScI).

- Astrophysical objects are now detected out to extraordinarily high redshifts (current record: $z \sim 8.2$), probably before the universe is fully reionized (see Figure 3.1).
- Scientists have discovered that most baryons in the universe—and most metals—do not lie in galaxies, implying that galaxy formation is relatively inefficient and/or winds are very important (Figure 3.2).
- The strong correlation observed between the stellar velocity dispersions in galaxies and the masses of their central black holes—the so-called M-σ relation—indicates that the evolution of galaxies must be closely tied to the growth of supermassive black holes.
- The simplest theory of galaxy formation predicts that big galaxies should be forming lots of stars, but this is not observed, which in turn has prompted theorists

FIGURE 3.2 M82, a starburst galaxy driving a galactic superwind. In similar galaxies at high-redshift, this kind of wind or an active galactic nucleus-driven wind may provide the feedback that inhibits star formation. The galaxy's stellar disk is shown in yellow-green; orange and red trace ionized gas and dust being driven out of the galaxy by the starburst, and blue indicates gas heated by the outflow to such a high temperature that it emits X-rays. The metals in this wind will enrich the intergalactic medium. SOURCE: *X-ray:* Courtesy of NASA/CXC/Johns Hopkins University/D. Strickland. *Optical:* Courtesy of NASA/ESA/STScI/AURA/The Hubble Heritage Team. *IR:* Courtesy of NASA/JPL-Caltech/University of Arizona/C. Engelbracht.

to consider the role of feedback (see Figure 3.2; see also Figure 3.6 in the next section) from AGN and supernovae in galaxy and cluster evolution.

- For the first time, a quantitative standard model of cosmology—ΛCDM—provides a detailed foundation on which a theory of galaxy formation and evolution can be built and which in turn can be tested by data.

Advances like these are driven by observations across a very broad wavelength range, by surveys and studies of individual objects, by temporal studies, and by theory. In the past decade, large surveys have provided unprecedented statistical coverage of galaxy properties, leading to such breakthroughs as the evidence for reionization at $z > 6$ and the quantification of the bimodal distribution of galaxies in mass-color space. Studies of deep fields have changed the understanding of AGN evolution and revealed a population of dust-enshrouded galaxies with high rates of star formation and/or accretion onto black holes. Computational power has reached a level at which large enough volumes of the universe can be simulated to do meaningful statistical studies, address a range of scales in a single calculation, and model increasingly sophisticated physics.

The study of galaxies across cosmic time draws on and informs the subject matter of nearly every major theme in contemporary astrophysics. To connect the current universe with its origins billions of years ago, there must be understanding of how stars formed under conditions ranging from the extremes of density, pressure, and metal abundance that prevailed as galaxies were being assembled, to the conditions typical of present-day galaxies such as the Milky Way. As stars age and die, they enrich the cosmos with heavy elements, and their explosions as supernovae and gamma-ray bursts help probe the cosmos at high redshift. Nearby galaxies demonstrate how to translate distant observables into physical parameters, how galaxies acquire mass from their surroundings, and how they expel mass, energy, and metals into their environments. The evolution of clusters of galaxies and the formation of structure with redshift provide powerful new tools for cosmology, constraining the properties of dark matter and dark energy. Black holes and relativistic jets provide insights into strong gravity, particle acceleration, and perhaps the nature of the highest-energy cosmic rays. Astronomy pushes the envelope of high-performance computing and ties in with the latest developments in areas such as atomic physics, numerical relativity, and plasma physics.

The purview of this Panel on Galaxies Across Cosmic Time encompasses the main constituents of the universe across 90 percent of cosmic time. Overlapping responsibilities with other Astro2010 Science Frontiers Panels—for example, responsibility for the epoch of reionization (with the Panel on Cosmology and Fundamental Physics) or for the study of nearby supermassive black holes (with the Panel on the Galactic Neighborhood)—were resolved in discussions among the panel chairs, taking into account the continuity with the other material considered by each panel. With the Panel on the Galactic Neighborhood, this panel agreed to a fuzzy boundary at $z \sim 0.1$, but with this panel including within its study supermassive black holes in nearby galaxies (not the Milky Way). With the Panel on Cosmology and Fundamental Physics it was agreed that this panel would focus on astrophysics (e.g., What is the distribution of dark matter?), whereas the Panel on Cosmology and Fundamental Physics would focus on fundamental physics (e.g., What is the dark matter made of?), and that this panel would cover the epoch of reionization.

Many of the cosmic objects of interest to science in the area of galaxies across cosmic time are faint and distant, yet dramatic progress has been made in the past decade in understanding how the components fit together. To set the stage for this panel's assessment of the most pressing questions in this field for the next decade, and the observational and theoretical capabilities needed to address them, it is first necessary to summarize the context inherent in present understanding.

Cosmic structures develop against the background of the expanding universe. Simulations show that the ΛCDM model successfully reproduces many of the properties of the universe on large scales (Figure 3.3). The largest structures seen—superclusters of galaxies and the cosmic web of filaments—evolve straightforwardly from tiny quantum fluctuations in the early universe, in a way that agrees well with the predictions of ΛCDM. The amount and distribution of dark matter have also been measured in systems from dwarf galaxies to clusters, and it has been verified that dark matter dominates over normal matter on the largest scales (Figure 3.4).

However, on the small scales of galaxies and for some properties of clusters, the data seem to disagree with the simple ΛCDM predictions. The distribution of baryons—the matter that makes the stars and gas and also collapses to form super-

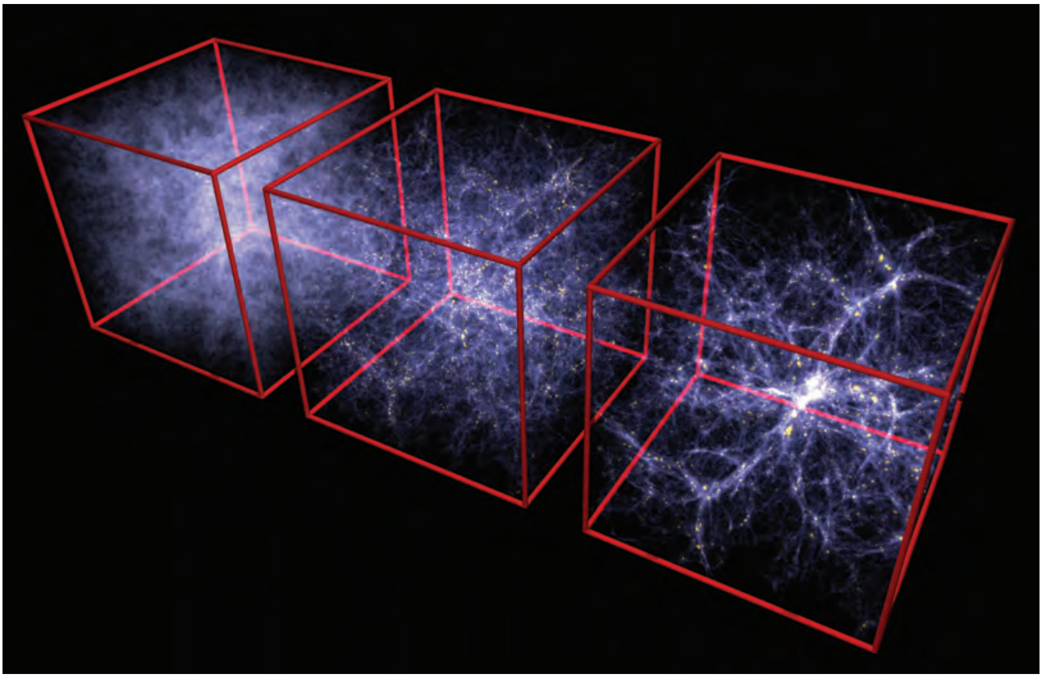


FIGURE 3.3 Simulation showing the evolution of large-scale structure in the universe's gas component, from redshift $z = 6$ to 2 to 0 (*left to right*). Each box is 100 Mpc/h (co-moving) on a side. Locations of galaxies where stars have formed are shown in yellow. The cosmic web of dark matter and gas condenses through gravitational instability, and galaxies form in the overdense regions that trace the underlying large-scale structure. SOURCE: Courtesy of Volker Springel, Heidelberg Institute for Theoretical Studies.

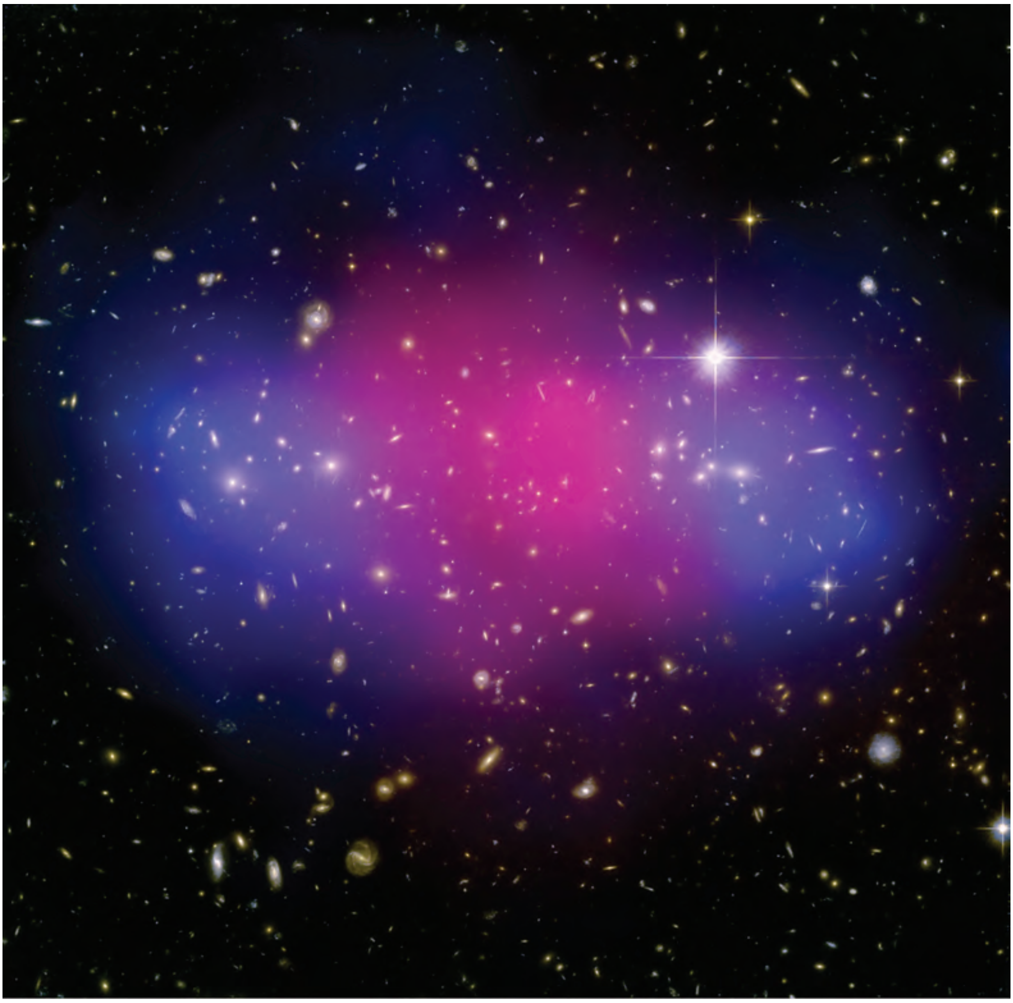


FIGURE 3.4 A composite image of the matter in the galaxy cluster MACSJ0025.4–1222. Individual galaxies are seen in the underlying optical image from the Hubble Space Telescope, the hot X-ray-emitting gas is shown in red, and the dark matter mapped by gravitational lensing is shown in blue. The gas cloud in the center is distorted by the collision between the two clusters, whereas the dark matter has passed through the cluster gas without interacting. The clear separation between gas and dark matter is direct evidence that dark matter exists (as opposed to modified gravity) and that it does not interact with baryonic matter. SOURCE: *X-ray:* Courtesy of NASA/CXC/Stanford University/S. Allen. *Optical/Lensing:* Courtesy of NASA/STScI/University of California, Santa Barbara/M. Bradac; see http://www.nasa.gov/mission_pages/chandra/news/08-111.html.

massive black holes at the centers of galaxies—turns out to be more complicated than predicted: if one assumes a simple relationship between the masses of dark matter halos and the luminosities of their galaxies, models of cosmic structure predict the wrong number of galaxies, with the wrong masses and colors, forming at the wrong times. Similarly, clusters of galaxies should show simple scaling relations between, for example, X-ray luminosity and temperature, yet the observations deviate from the predictions. An additional energy source besides gravity—perhaps coming from active galaxies and/or supernovae—appears to be necessary in order to understand cosmic structures, but so far there is little direct observational evidence for this so-called feedback (see, in the next section, Figure 3.7).

The history of star formation plays a crucial role in the evolution of galaxies, clusters of galaxies, and the intergalactic medium. The IGM contains most of the baryons at all epochs and as early as $z \sim 6$ was enriched with heavy elements that must have been made in stars. Astronomers are just learning how star-formation rates, heavy-element abundances, and galaxy masses change with cosmic time. The preliminary results indicate that the stars in massive galaxies formed when the universe was very young and that these galaxies stopped forming stars many billions of years ago, whereas smaller, less massive galaxies, like the Milky Way, have continued to form stars until the present. Galaxies, at least at $z < 0.5$, come in a bimodal distribution of mass and color and exhibit relatively simple correlations linking mass, metallicity, and galactic structure.

It has also been learned that supermassive black holes, at least at low redshifts, lie at the centers of virtually all massive galaxies and that there are relatively simple relations between properties of the galaxy—such as its bulge mass and velocity dispersion—and the mass of the black hole. The reasons for these correlations are not understood, nor is it known how these relationships evolve and whether they extend to lower masses. Active galaxies—manifestations of accreting supermassive black holes—were vastly more abundant and luminous in the past: their numbers and integrated luminosity peaked roughly when star formation was at its peak.

Both black hole growth and star formation are implicated in feedback and in the evolution of structures on all scales. Yet the current inventory of these processes is far from complete. Cold gas and dust enshroud many star-forming galaxies and AGN. Roughly half of all the light emitted in the universe has been absorbed by dust and gas and re-emitted in the mid- to far-IR, and hard X-ray surveys with the latest generation of satellites have shown that approximately two-thirds of all AGN are obscured and that these obscured AGN emit most of the energy radiated by supermassive black holes. Infrared and submillimeter studies have revealed that most early star formation occurred behind a thick screen of dust, as exemplified spectacularly by the ultraluminous infrared galaxies that are the most prodigious sites of star formation in the universe (Figure 3.5).

Before the first stars and galaxies formed, the universe was dark. The reioniza-

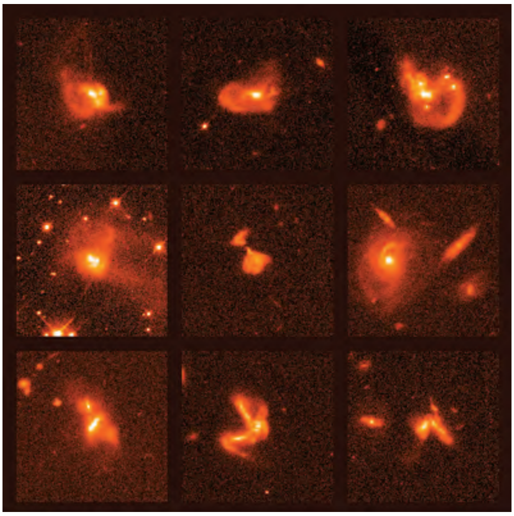


FIGURE 3.5 Hubble Space Telescope images of ultraluminous infrared galaxies make clear that their prodigious star-formation rates are associated with violent mergers of two or even more galaxies. Often these are dusty, obscured environments best imaged at far-infrared or X-ray wavelengths. SOURCE: Courtesy of NASA and K. Borne (Raytheon ITSS and NASA Goddard Space Flight Center), H. Bushouse (STScI), L. Colina (Instituto de Fisica de Cantabria, Spain), and R. Lucas (STScI); see http://hubblesite.org/newscenter/archive/releases/1999/45/image/a/format/web_print/.

tion of the intergalactic medium occurred as the first generations of luminous objects ionized their surroundings, somewhere between redshifts of (very roughly) 6 and 20. Analysis of the cosmic microwave background (CMB) and absorption-line studies of gas along the line of sight to distant quasars and gamma-ray bursts have begun to probe this epoch. But it is still unknown which kinds of objects—stars or quasars, or something else (e.g., decaying dark matter)—ended the cosmic "Dark Ages." Computer simulations are providing theoretical tools needed to push toward an answer, while upcoming space missions and ground-based instruments, such as the James Webb Space Telescope and the Atacama Large Millimeter Array (ALMA), may provide crucial clues.

Against this background of rapid-fire and unexpected discoveries during the past decade, the panel has identified four questions and one discovery area that are particularly ripe for addressing in the next decade:

- How do cosmic structures form and evolve?
- How do baryons cycle in and out of galaxies, and what do they do while they are there?
- How do black holes grow, radiate, and influence their surroundings?
- What were the first objects to light up the universe and when did they do it?
- *Unusual discovery potential:* The epoch of reionization.

In the following sections, the panel discusses each of these issues and outlines a science program of observations and theory to address them during the coming decade. The program is ambitious: to help make the inevitable hard choices, the panel classifies each approach in one of three categories: as "most important," "very important," or "important." Many other ideas considered but eventually excluded from the list remain valuable as ways to make progress, and the panel anticipates that significant progress will also come from unexpected directions. The history of astronomy and astrophysics attests that key breakthroughs often come from serendipitous discoveries with broadly capable facilities.

GCT 1. HOW DO COSMIC STRUCTURES FORM AND EVOLVE?

- What is the structure of dark matter halos on galaxy, group, and cluster scales?
- What is the origin of the observed correlations among the fundamental properties of galaxies and of clusters, and how do they evolve with time?

Progress to Date

The origin of the cosmic structures seen on different scales throughout the universe depends on the underlying physics. While ΛCDM is validated on large scales, it is not fully understood how luminous structures on smaller scales arise and whether the observed properties of galaxies and clusters are consistent with the cold dark matter cosmology.

The challenges on small scales are significant. In and around galaxies, the dark matter appears to be much less lumpy than predicted, as evidenced by the mismatch between the number of low-mass satellites detected around the Milky Way in the Sloan Digital Sky Survey and the abundance of subhalos predicted by dark matter simulations. The internal mass density profiles of galaxies and clusters appear to deviate from the simple universal shape predicted by theory in the central regions of dark matter halos. Also the limited amount of ongoing star formation in elliptical galaxies, along with the relative abundance of large, fragile disks in the local universe, is not understood in the context of current theory (Figure 3.6).

The origin of the well-defined correlations among the shapes, sizes, and compositions of galaxies and clusters is also not understood. For example, the "fundamental plane" linking properties of elliptical galaxies or the Tully-Fisher relation for spirals, should be predicted by a successful theory of galaxy formation, as should the tight relationship between the mass of the central supermassive black hole and that of the host galaxy, as well as associations among stellar mass, total mass, and chemical composition of galaxies. Similar tight relationships are observed among the properties of clusters: for example, correlations among cluster mass and the luminosity and temperature of the hot intracluster gas. It is not yet understood what processes cause these relations and govern their evolution over cosmic time.

Although there is some information, albeit limited, on the evolution of galaxies out to $z \sim 6$, scientists are only beginning to glimpse how clusters evolve out to $z \sim 1$. Comparison with hydrodynamical simulations reveals significant departures

FIGURE 3.6 *Top:* Simulated cluster-scale (*left*) and galaxy-scale (*right*) dark matter halos look quite similar, despite differing in mass by two orders of magnitude. The regions shown are at comparable contrast and resolution and correspond to approximately 2 virial diameters around each halo. *Bottom:* Hubble Space Telescope images of galaxy cluster Abell 1689 (*left*) and of galaxy NGC4458 (*right*). In the actual data, the observed number of satellites is dramatically different for the two observed systems, even though the dark matter substructure is predicted to be virtually indistinguishable. SOURCE: *Top:* Courtesy of Andrey Kravtsov (University of Chicago), Anatoly Klypin (New Mexico State University), and Stefan Gottloeber (AIP, Potsdam, Germany). *Bottom left:* Courtesy of NASA, ESA, Richard Ellis (Caltech), Jean-Paul Kneib (Observatoire Midi-Pyrenees, France), A. Fruchter, and the ERO Team (STScI and ST-ECF). *Bottom right:* Courtesy of NASA, ESA, E. Peng (Peking University, Beijing), and the ACS Virgo Cluster Survey Team.

from theory, underscoring the importance of baryonic physics—including cooling and heating, gas ejection, and turbulence—that have not yet been modeled adequately. For example, recent X-ray spectroscopic studies of supposedly cool cores in the centers of many massive clusters show that the amount of gas cooling is considerably less than simple models predict (Figure 3.7), raising important questions: What prevents these cool cores from cooling further? What does the

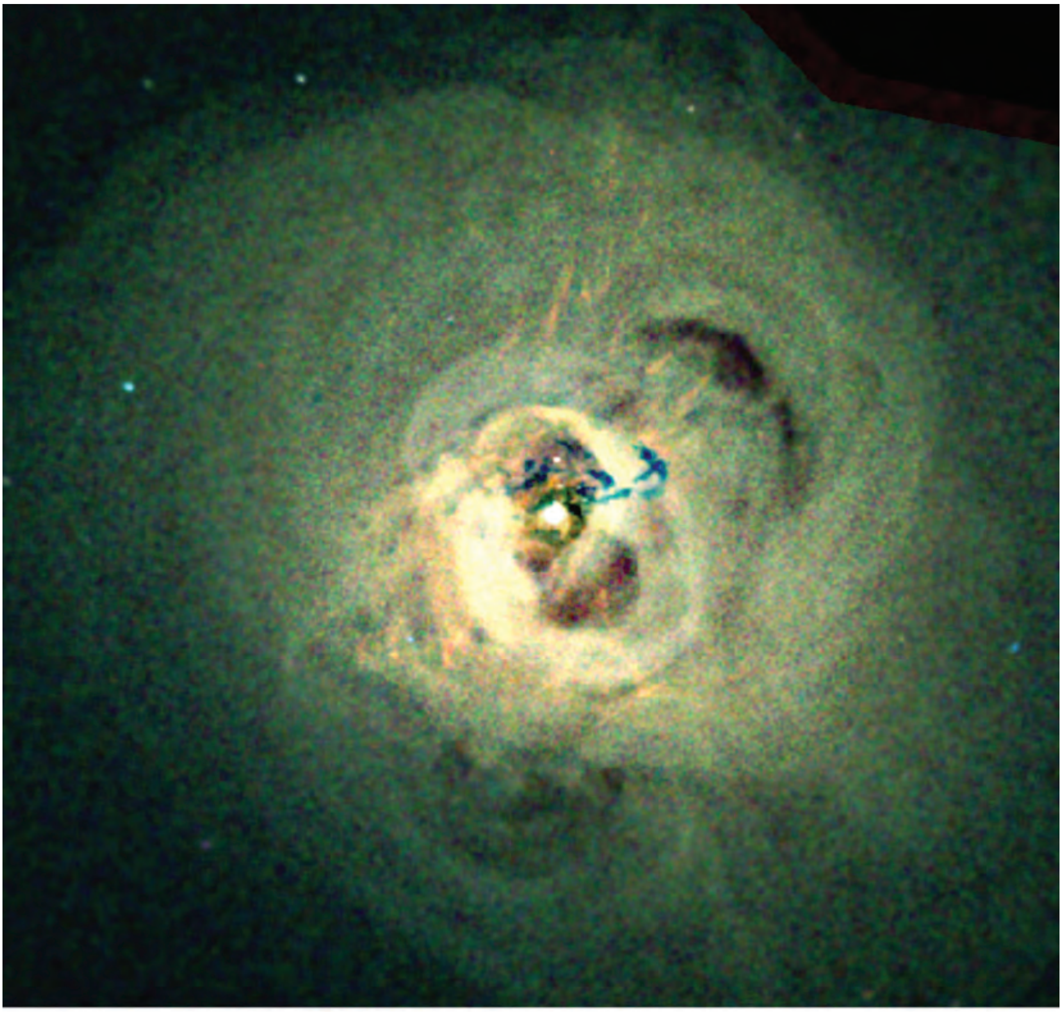


FIGURE 3.7 Chandra X-ray image of the Perseus cluster. Low-density regions are seen as dark bubbles surrounded by bright regions of hot, X-ray-emitting gas, believed to be generated by outbursts of activity from the central black hole. This may be direct evidence for feedback between black hole growth and cluster evolution. SOURCE: Courtesy of NASA/CXC/IoA/A. Fabian et al.

early metal enrichment of the intracluster medium, to ~1/3 solar, reveal about structure formation (as well as baryon cycling)? Understanding the properties of clusters and the ICM will also help improve the understanding of why galaxies in clusters differ from those in the field.

Exploring the observed correlations among galaxy and cluster properties and measuring their evolution with time will provide powerful probes of structure formation and of the complex processes of star formation, metal enrichment, feedback, and radiative cooling that are underway in these systems.

Steps for the Next Decade

Understanding how cosmic structures form and evolve on galaxy, group, and cluster scales starts with studying the dark matter distribution on a range of scales, measuring the characteristics of intracluster gas, and probing the interplay of dark and luminous matter. Here the panel outlines the highest-priority activities for the coming decade, on dark matter substructure, hot gas dynamics, galaxy properties, and black hole masses.

Dark matter substructure and the overall shape of dark matter halos can be probed directly, on both galaxy and cluster scales, through its gravitational lensing effects (Figure 3.8). Large samples of lenses, at least an order-of-magnitude larger than present samples, are required to constrain the mass function of halo substructure, and high spatial resolution (HST-like or better) is needed to detect the lensing signal in the positions and flux ratios of multiple images (in the case of strong lensing). JWST and ELTs with adaptive optics can probe the full range of the mass function using even higher resolution. Detailed imaging and spectroscopic studies of individual objects would reveal the evolution of dark matter profiles over time, through a combination of lensing (weak and strong) and dynamics. In order to determine the mass function of substructure and the profile of the dark matter in galaxies and clusters, the panel concluded:

- It is *most important* to obtain HST-like imaging to determine the morphologies, sizes, density profiles, and substructure of dark matter, on scales from galaxies to clusters, by means of weak and strong gravitational lensing, in lens samples at least an order-of-magnitude larger than currently available. HST can make an important start on this problem, but to develop large statistical samples will require a much larger field of view or more observing time than HST affords.

Studies of the evolution of cluster properties are currently limited by small high-redshift samples with a variety of selection biases. Further progress requires uniformly selecting clusters over a wide redshift and mass range using coordinated wide-field X-ray, Sunyaev-Zel'dovich effect (SZE), and optical/infrared (O/IR) sur-

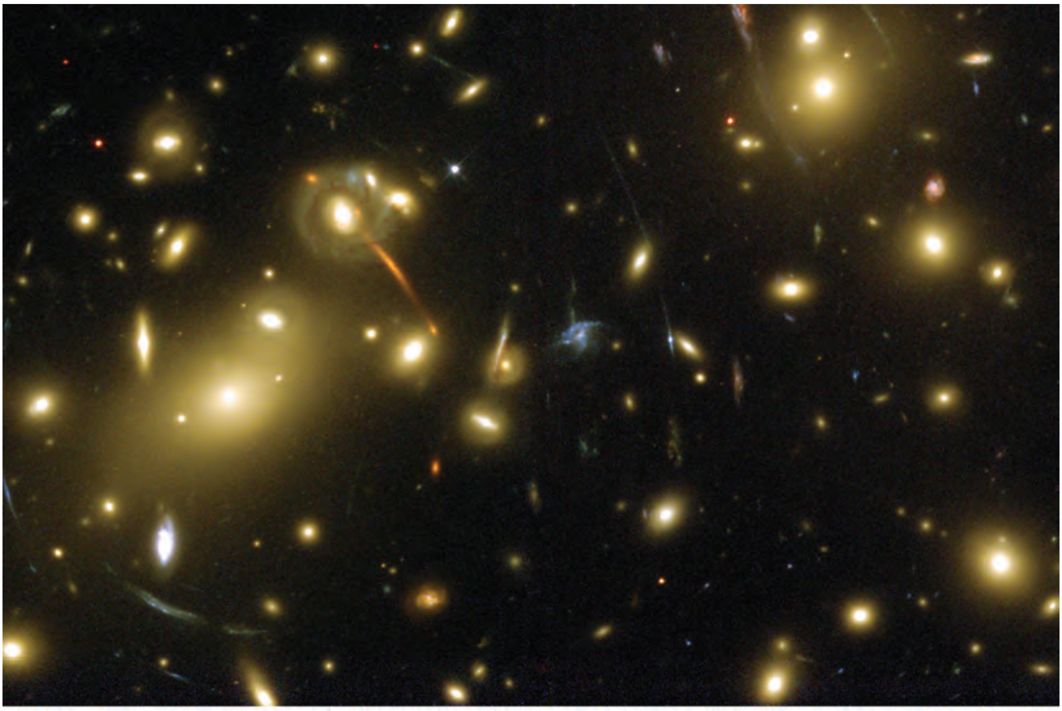


FIGURE 3.8 The gravitational potential of the cluster Abell 2218 (yellowish galaxies) acts as a cosmic lens, bending the light from background sources, magnifying them, and creating multiple and/or distorted images (bluish arcs). The curvature and displacement of the images contain enough information to reconstruct the smooth mass distribution and that associated with cluster galaxies. A gravitational lensing experiment on the smaller scale of galaxies will constrain the dark matter substructure predicted by numerical simulations. SOURCE: Courtesy of NASA, N. Benitez (Johns Hopkins University), T. Broadhurst (Racah Institute of Physics/Hebrew University), H. Ford (Johns Hopkins University), M. Clampin (STScI), G. Hartig (STScI), G. Illingworth (UCO/Lick Observatory), the ACS Science Team, and ESA.

veys (Figure 3.9). Some surveys that can deliver these samples are either operating or appear to be already funded, including SZE surveys with the South Pole Telescope (SPT) and Atacama Cosmology Telescope (ACT) and a future X-ray survey with eROSITA/Spektrum-RG. Both SZE and X-ray surveys require deep multiband (*griz*) optical surveys, sufficient to detect galaxies a couple of times fainter than the typical galaxy at $z = 1$ (e.g., the Dark Energy Survey [DES], funded to begin in 2011) to deliver optical confirmation and redshifts to $z \sim 1$, and near-infrared surveys (e.g., the Visible and Infrared Survey Telescope for Astronomy [VISTA]) pushing to 0.5 L* for clusters at $z > 1$.

The most critical study of new cluster samples concerns the dynamics of the gas component of clusters, which dominates the baryonic mass and contributes to the

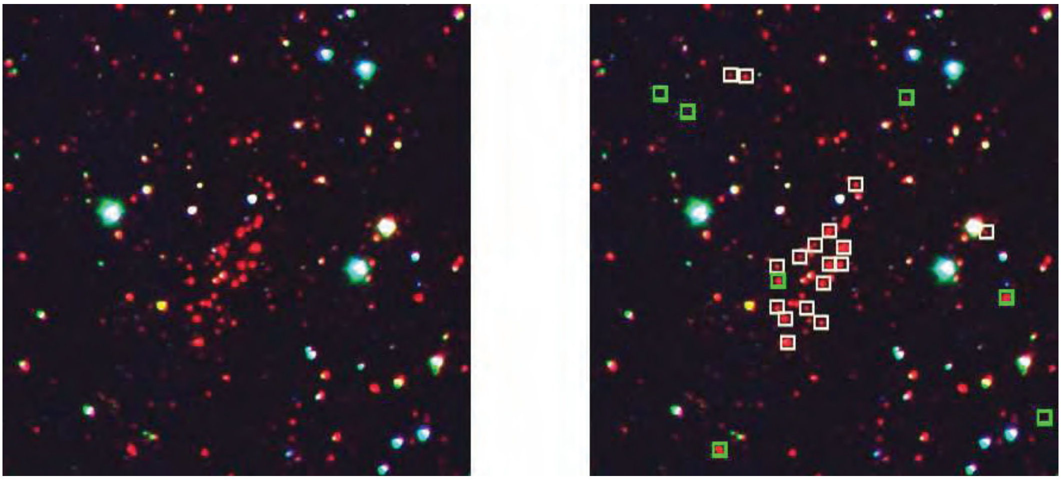


FIGURE 3.9 Massive red-sequence-selected galaxy cluster at $z \sim 1.2$, detected in an optical/infrared survey. *Left:* Color composite of the cluster combining Canada-France-Hawaii Telescope Megacam *R*- and *z*-band images with Spitzer Infrared Array Camera 3.6-μm images. The cluster of massive, old red galaxies is easily visible. *Right:* Spectroscopically confirmed cluster members are marked with white squares; background and foreground galaxies are indicated with green circles. The photometric redshifts agree well with the spectroscopic redshifts, validating their use and enabling large surveys for high-redshift clusters; however, the mass of the cluster estimated from the light in the galaxies does not seem to agree with scaling relations from lower-redshift clusters. This color selection technique makes it possible to study cluster evolution out to quite high redshifts. SOURCE: A. Muzzin, G. Wilson, H.K.C. Yee, H. Hoekstra, D. Gilbank, J. Surace, M. Lacy, K. Blindert, S. Majumdar, R. Demarco, J.P. Gardner, M. Gladders, and C. Lonsdale, Spectroscopic confirmation of two massive red-sequence-selected galaxy clusters at $z \sim 1.2$ in the SpARCS-North Cluster Survey, *Astrophysical Journal* 698(2):1934-1942, 2009, reproduced by permission of the AAS.

pressure. Very little is known about the velocity or ionization structure of this gas, although it is suspected that turbulence may be a very important source of energy dissipation within it. High-resolution X-ray spectroscopy of cluster gas is the best means of directly measuring the turbulence, bulk flows, and AGN-driven shocks within this gas. The planned Astro-H mission will make the first few measurements of this kind, in the brightest local clusters. Similar observations over a broader range of cluster mass and redshift will allow fresh insights into puzzles such as the excess entropy and heating mechanisms within clusters. Additional information on gas dynamics could come from detailed radio and/or submillimeter mapping of clusters detected at lower resolution by SZE surveys. The panel concluded:

- High-energy-resolution, high-throughput X-ray spectroscopic studies of groups and clusters to $z \sim 2$ are *most important* for understanding the dynamics, ionization and temperature structure, and metallicity of the hot intracluster gas, as well as for studying the growth of structure and the evolution of the correlations among cluster properties.

Clues to the relationship between baryons and dark matter are encoded in correlations among galaxy properties like mass, luminosity, surface brightness, and velocity structure. A great deal has been learned about galaxies in the local universe from large spectroscopic surveys in the rest-frame optical, such as the Sloan Digital Sky Survey, including the bimodal nature of the galaxy color-magnitude distribution. Comparable information at higher redshifts, where most of the star formation occurs, lags well behind; available samples come from deep imaging surveys such as the Great Observatories Origins Deep Survey (GOODS), Deep Extragalactic Evolutionary Probe 2 (DEEP2), or the Cosmological Evolution Survey (COSMOS), which cover at most a few square degrees and for which the spectroscopic follow-up (with single-slit near-IR spectrometers) is very slow and therefore limited.

Considerable progress would result from a near-infrared spectroscopic survey of galaxies at the epoch of peak star formation, $1 < z < 3$. With SDSS-size samples, galaxy properties can be sorted by color, mass, redshift, morphology, and age, as was done very effectively at $z \sim 0$. Spatially resolved spectroscopy of a representative subset, to obtain velocity dispersions and/or rotation curves, would trace the detailed internal velocity fields and help in understanding the slit spectroscopy for the larger sample (Figure 3.10). Currently, substantial samples could be studied with existing large telescopes, using upcoming multiobject near-IR spectrometers with moderate resolution ($R \sim$ few thousand), although a truly SDSS-like survey would require additional resources. The panel concluded:

- It is *very important* to obtain moderate-resolution multi-slit spectroscopy of SDSS-size galaxy samples at $z \sim 1$-3, in the optical for $z < 1.5$, and in the near-IR for $z > 1.5$ (with $[R] \sim 5{,}000$ to allow effective removal of night skylines in the near-IR). For a representative subset of hundreds of galaxies, high-angular-resolution integral field unit (i.e., two-dimensional) spectroscopy in the optical or near-IR would help calibrate the slit spectra. To select targets for spectroscopy requires optical/IR pre-imaging over a large area. It should be noted that, properly designed, the same large near-IR spectroscopic survey could serve the second key question as well.

Measurements of supermassive black hole masses over the past decade have revealed a remarkable correlation between black hole mass and the stellar kinematics, mass, and luminosity of the galaxy bulge. This was unexpected because it means that stars very far outside the sphere of influence of the black hole somehow “know about” its mass, suggesting that the process of galaxy formation and evolution is closely tied to the growth of the central supermassive black hole. At present, astronomers have few indications of how this might happen, and they are

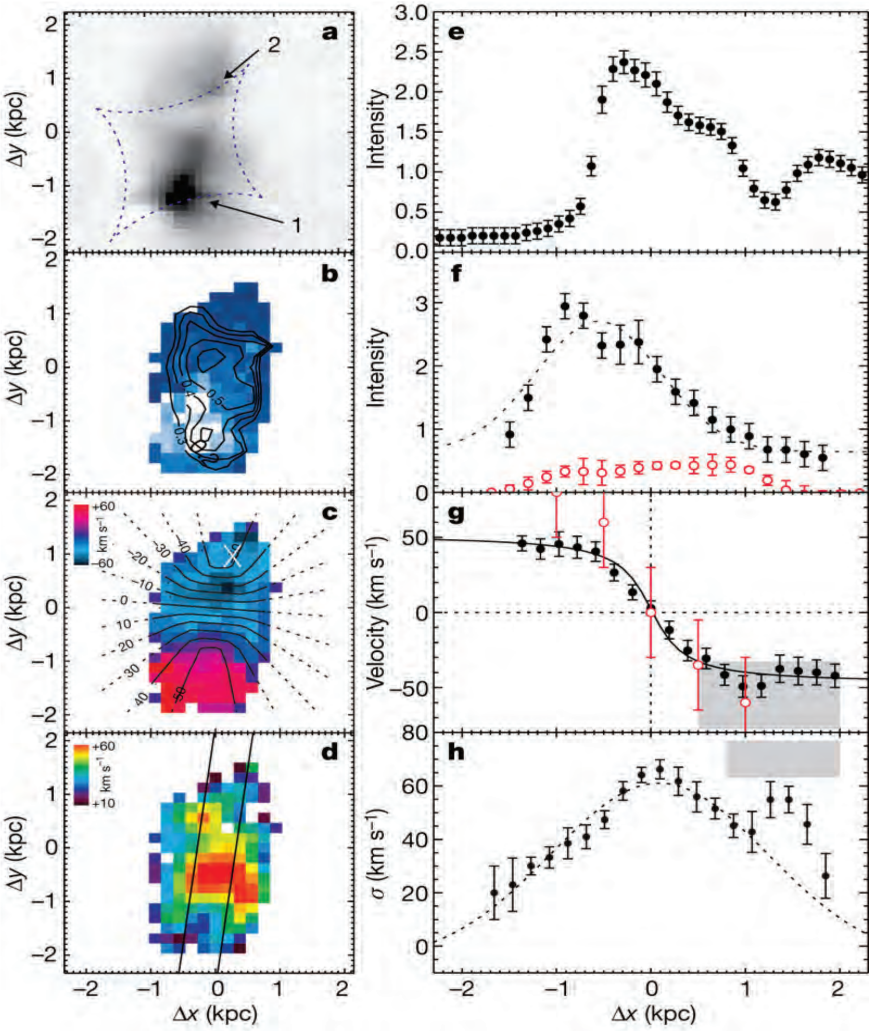


FIGURE 3.10 Integral field unit (IFU) spectroscopy of a gravitationally lensed galaxy at $z = 3.07$, demonstrating the range of information provided by IFU measurements. All panels are source-plane reconstructions of the lensed galaxy called the Cosmic Eye. With the next generation of ground- and space-based large telescopes, spatial resolution of hundreds of parsecs will be obtained for galaxies at $z > 1$, enabling the study of spatially resolved continuum (panels *a* and *e*) and emission-line (panels *b* and *f*) intensities, velocities (panels *c* and *g*), and velocity dispersions (panels *d* and *h*); the *left panels* show the two-dimensional galaxy reconstruction, and the *right panels* show the one-dimensional slit extraction. Such measurements are crucial for determining the overall dynamical state (disk/merger) and detailed kinematic structure of galaxies as they form. Furthermore, mapping the spatially resolved emission from multiple rest-frame optical emission lines will trace out not only the distribution of star-forming regions but also the related gradients in chemical abundances. SOURCE: Reprinted by permission from Macmillan Publishers Ltd.: *Nature*, D.P. Stark, A.M. Swinbank, R.S. Ellis, S. Dye, I.R. Smail, and J. Richard, The formation and assembly of a typical star-forming galaxy at z ~ 3, *Nature* 455:775, 2008, copyright 2008.

struggling to obtain reliable black hole mass estimates for more than a few dozen galaxies. Defining this relation over a larger mass range and a broader redshift range would help improve the understanding of the interplay of the stellar assembly and black hole growth. The panel therefore concluded that, in order to extend the M-σ relation to low-mass local systems and massive high-redshift systems:

- It is *important* to measure supermassive black hole masses in hundreds or even thousands of systems using spatially resolved spectroscopy with adaptive optics on ELTs, centimeter-wave maser observations with very long baseline arrays supplemented by large aperture dishes, and/or reverberation mapping of AGN. (See similar conclusion for the third key question, GCT 3.)

A precise study of the mass function of substructures on galaxy scales can also be done with radio interferometers, which can image strong galaxy-scale gravitational lenses to constrain the mass distribution within galaxies. Compared to optically imaged lenses, fewer radio lenses are bright enough for this technique, but the spatial resolution is far superior to even optical imaging from space. Because the incidence of lensing rises sharply with increasing spatial resolution, it will eventually be possible to study very large samples of lenses with high-resolution radio imaging, with sensitive facilities that survey a large fraction of the sky. This kind of study can start now with existing and/or upgraded radio facilities and will reach truly powerful levels with future facilities. The panel concluded:

- It is *important* to obtain Very Long Baseline Array (VLBA), Expanded Very Large Array (EVLA), e-Merlin, and/or ALMA imaging of at least a few hundred new galaxy-scale lenses. The development of a large-collecting-area radio facility that can survey a large fraction of the sky with subarcsecond resolution will increase the number of accessible lenses by several orders of magnitude.

GCT 2. HOW DO BARYONS CYCLE IN AND OUT OF GALAXIES, AND WHAT DO THEY DO WHILE THEY ARE THERE?

- How do galaxies acquire gas across cosmic time?
- What processes regulate the conversion of gas to stars as galaxies evolve?
- How are the chemical elements created and distributed?
- Where are the baryons as a function of redshift?

Progress to Date

Galaxies grow by obtaining fuel from the intergalactic medium, converting it into stars, and returning gas, heavy elements, and energy back to their surroundings. These baryonic processes are poorly understood: current models fail to reproduce even basic observed properties of galaxies without invoking many ad hoc parameters. To understand how galaxies form and evolve across cosmic time, it is critical to understand how baryons cycle into, within, and out of galaxies.

There has been a radical change in the understanding of galaxy growth in the past decade, driven by the recognition of its complex dynamics. Much of this progress has been obtained by optical measures of distant star formation, near-infrared measures of stellar-mass growth, submillimeter measures of dust-enshrouded star formation, and radio measures of molecular gas content. These data have yielded cosmic rates of star formation, accretion, and enrichment during the heyday of cosmic star formation at $1 < z < 3$, at least for the brightest galaxies, which appear to be quite different from today's systems. These bright galaxies are forming stars much more vigorously, are more compact and gas-rich, and are ubiquitously expelling gas. Studying this epoch is crucial, since many properties of present-day galaxies were established at that time.

While typical galaxies are expected to grow mainly by accreting material from the IGM, mergers are probably important for morphological transformations and central black hole growth and may drive the most powerful episodes of early star formation. Quantitatively characterizing the role of gas inflows and mergers in the evolution of galaxies is important for understanding galaxy growth (Figure 3.11).

The IGM holds key signatures of baryon cycling. Not only is it a repository of gas for accretion, but it also contains heavy elements expelled by galaxies. The mass and energy associated with outflows help to modulate star formation and chemical enrichment in galaxies. The cosmic distribution of metals, from diffuse intergalactic gas to intracluster gas to galaxies, provides crucial constraints on baryon cycling. A complete census of cosmic baryons and metals, including those in difficult-to-detect phases such as the warm-hot intergalactic medium ($10^5 < T < 10^7$ K), is important for connecting all of these data. Indeed, galaxies, their halos, and the IGM must be viewed as forming an interconnected cosmic ecosystem within which galaxies form, grow, and die.

Steps for the Next Decade

Two key constituents of galaxies are stars and gas. In the next decade, astronomers are poised to make dramatic advances that will, for the first time, comprehensively trace the evolution of stars and gas within and around galaxies over the majority of cosmic time. This will involve technological developments such as improved near-IR detectors, increased far-IR and radio sensitivities, higher efficiency and spectral resolution in the UV and X-ray, along with anticipated new data from EVLA, Herschel, JWST, and ALMA. Sophisticated numerical simulations will be critical for providing physical insight and enabling observations to be synthesized into a cosmological framework.

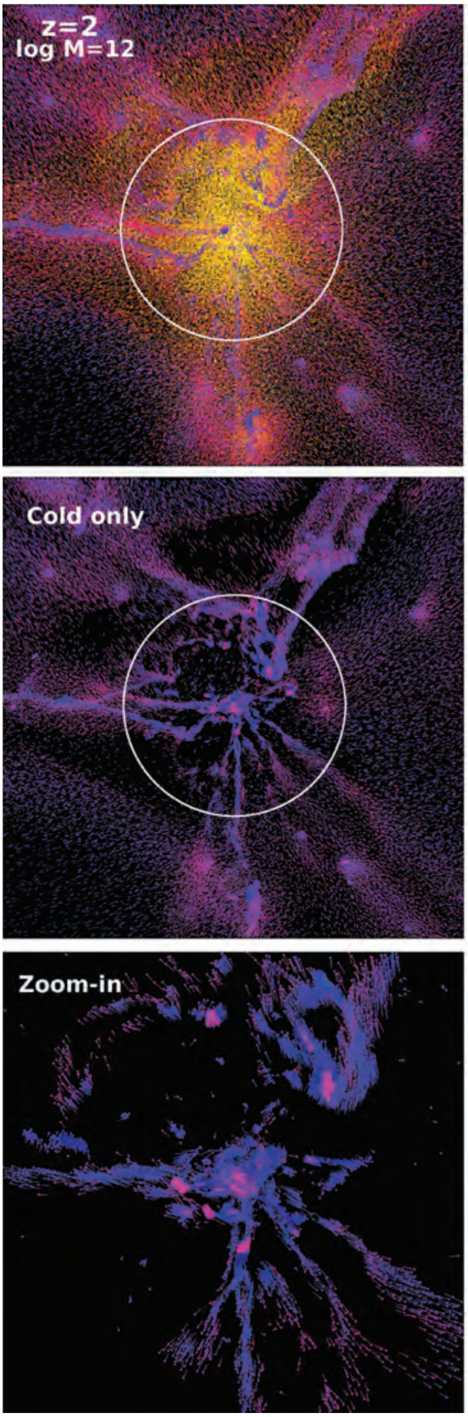


FIGURE 3.11 Computer simulation illustrating cold-mode accretion—namely, accretion from the intergalactic medium in which the material remains well below the virial temperature through the virial radius—into a large galaxy at redshift z = 2. Owing to the high cosmic gas density in the early universe and the density enhancement provided by filamentary structure, cold-mode accretion is expected to dominate the global accretion budget during the peak epoch of cosmic star formation. Understanding this fundamental mode of galaxy growth is a key motivation for spectroscopically characterizing the young universe at $z \sim$ 1-3. SOURCE: Reprinted with permission from D. Kereš, N. Katz, M. Fardal, R. Davé, and D.H. Weinberg, Galaxies in a simulated Lambda-CDM Universe. I: Cold mode and hot cores, *Monthly Notices of the Royal Astronomical Society* 395:160-179, 2009, copyright 2009 Royal Astronomical Society.

Building on the comprehensive view of galaxies at $z \sim 0$ afforded by the SDSS survey (Figure 3.12) a full evolutionary picture for galaxies calls for comparable rest-frame optical spectroscopy of galaxies and intergalactic gas during the peak epoch of galaxy formation, at $z \sim$ 1-3. The rest-frame optical is better than rest-frame UV for probing light from the bulk of the stellar population, and is less sensitive than is redder light to poorly understood late phases of stellar evolution. The sample size should be large enough to disentangle the covariances among galaxy properties such as luminosity, age, morphology, metallicity, mass, redshift, and so on, over a large enough volume to sample a representative range of galaxy environments. Today, only about a hundred $z > 1$ galaxies have rest-frame optical spectra, and most of these are strongly star-forming objects, so this small sample gives an incomplete and biased picture of the baryonic contents of the galaxy population at early epochs. Furthermore, these spectra typically consist of emission lines, with little or no continuum detected. With improvements in near-IR detector technology, and a factor of approximately 40 to 50 in multiplexing capability (e.g., Multi-Object Spectrometer for Infra-Red Exploration [MOSFIRE], Flamingos II, EMIR[2], or other multiobject infrared spectrometers), it should be possible to obtain spectra of roughly a thousand high-redshift galaxies, spanning a broader range of star-formation histories, with present 6- to 10-meter telescopes. However, at least another order of magnitude in sample size is needed to begin to disentangle covariances among the salient suite of galaxy properties and environments; this may be possible with new instrumentation and/or a future reallocation of current telescope resources.

Over smaller areas one could probe substantially below L*, to study the progenitors of typical galaxies today, with sufficient resolution to measure dynamical and stellar population parameters, sufficient continuum sensitivity to measure absorption lines, and sufficient emission-line sensitivity to measure low levels of star formation. From the ground, $R \sim 5{,}000$ spectra are required to resolve out infrared skylines. These data would establish the evolution of the distribution functions of star-formation rate, AGN activity, star-formation history, stellar mass, and stellar- and gas-phase metallicity, and would quantify the correlations of these properties with one another and with the large-scale environment across the most active epoch of galaxy formation (Figure 3.13).

Targeted follow-up using ELTs and JWST will be crucial for studying the faintest and highest-redshift objects and for measuring features that are shifted into the thermal infrared. For maximum synergy with ALMA and possible future high-sensitivity H I observations, survey fields would be equatorial or southern. Deep optical and near-IR imaging surveys for target selection over this area are a prerequisite.

---

[2] EMIR (Espectrografo Multiobjeto Infrarrojo) is a multiobject infrared spectrograph for the Gran Telescopio Canarias.

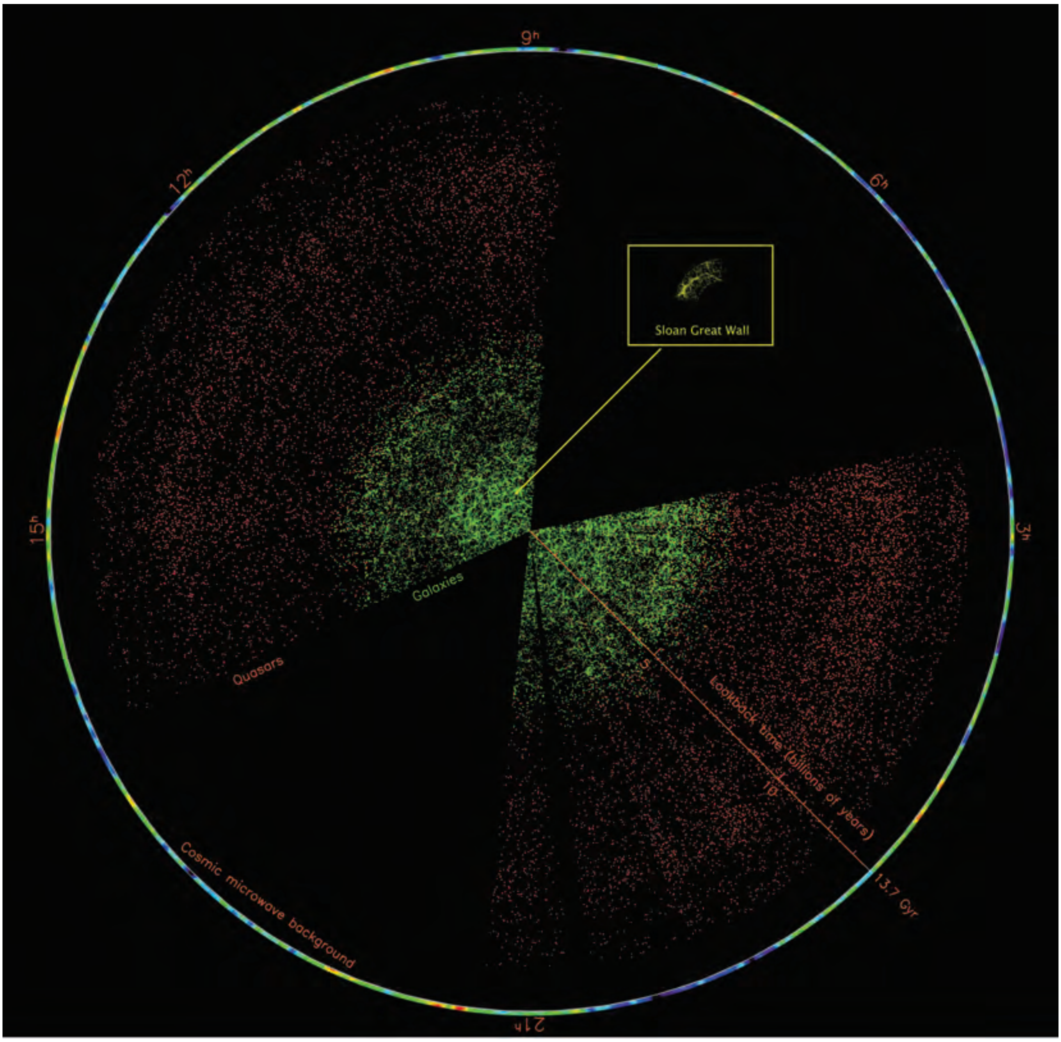


FIGURE 3.12 The large-scale distribution of galaxies ($z < \sim 0.4$; green) and quasars ($z < \sim 6.5$; red) observed in the Sloan Digital Sky Survey (SDSS) provides a well-sampled view of the local universe, but a much sparser sampling at large distances, during the key epoch of star formation ($1 < z < 3$, or look-back times of approximately 8 to 11 Gyr). The circular boundary represents the cosmic microwave background at 13.7-Gyr look-back time. An SDSS-size spectroscopic survey at $z \sim$ 1-3 will reveal stellar, gas, and dynamical properties of millions of galaxies in the early universe, at the peak of cosmic star formation and galaxy assembly. SOURCE: W. Colley, J.R. Gott, M. Juric, R.J. Vanderbei, and the SDSS (Sloan Digital Sky Survey) team, to appear in *Sizing Up the Universe*, by J. Richard Gott and Robert J. Vanderbei, National Geographic.

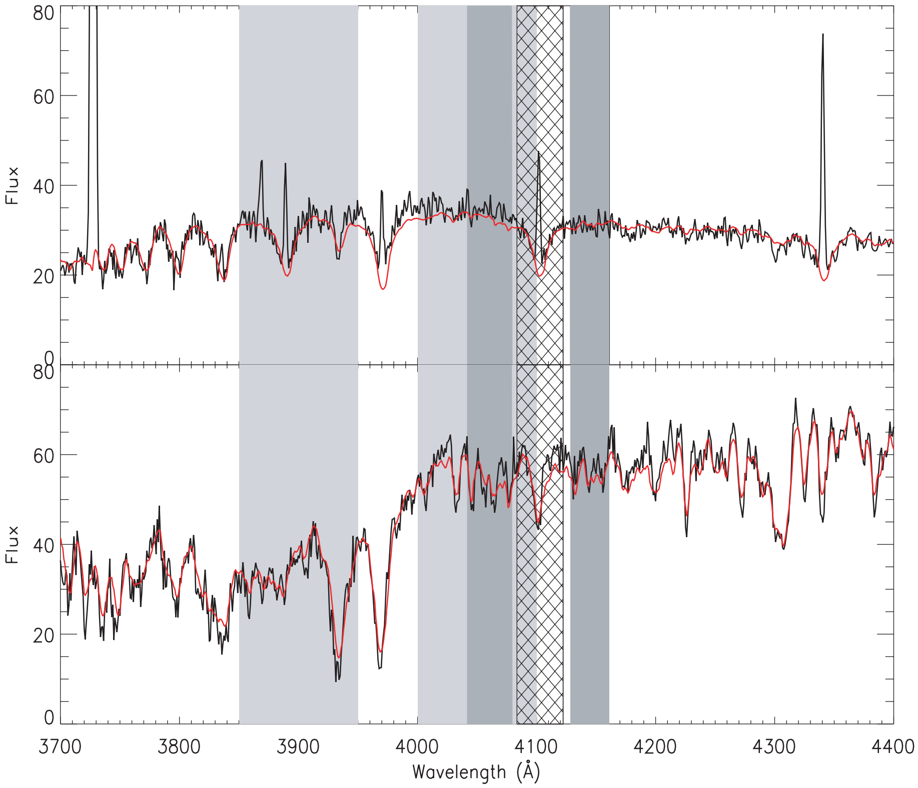


FIGURE 3.13 Constraints on stellar populations and the past history of star formation can be derived from rest-frame optical continuum and absorption-line spectroscopy. Metallicities and dynamical properties are reflected in both rest-frame emission and absorption features. With few exceptions, rest-frame optical spectroscopy for galaxies at $z > 1$ has been focused on emission-line measurements for the most actively star-forming systems. In the next decade these studies should be extended to more typical systems, tracing both emission and absorption lines for large Sloan Digital Sky Survey (SDSS)-like samples. Shown here are SDSS spectra of a local late-type galaxy (*top*) and a local early-type galaxy (*bottom*) plotted over the interval 3,700 to 4,400 Å in the rest frame. The red line shows a best-fitting stellar population model. The shaded regions indicate bandpasses over which indices sensitive to the past history of star formation are measured. The hatched region shows where the Hδ equivalent width is estimated. SOURCE: Reprinted with permission from G. Kauffmann, T.M. Heckman, S.D.M. White, S. Charlot, C. Tremonti, J. Brinchmann, G. Bruzual, E.W. Peng, M. Seibert, M. Bernardi, M. Blanton, et al., Stellar masses and star formation histories for 10^5 galaxies from the Sloan Digital Sky Survey, *Monthly Notices of the Royal Astronomical Society: Letters* 341(1):33-53, 2003, copyright 2003 Royal Astronomical Society.

For a representative subsample of galaxies, rest-frame optical IFU spectroscopic maps will be crucial for characterizing the detailed baryonic processes of galaxy assembly, feedback, and chemical enrichment in action. With planned near-IR IFUs on an ELT or the Near-Infrared Spectrometer (NIRSpec) IFU on JWST, it will be possible to map extended warm gas associated with outflows and inflows, as well as the dynamics probed by higher surface-brightness regions of star formation. Sensitivities to the star-formation rate surface density probed with Hα emission will be roughly an order-of-magnitude higher than is currently possible with 10-meter-class ground-based telescopes. Detailed chemical abundance measurements will require the longest integration times, as they depend on weaker nebular lines.

On the basis of these considerations, the panel concluded that, to understand the full cycling of baryons in and out of galaxies:

- It is *most important* to carry out an SDSS-size near-infrared spectroscopic survey of galaxies at $1 < z < 3$. This will require near-infrared pre-imaging in the J, H, and K bands (at 1, 1.6, and 2 microns) to select targets for spectroscopy. It should be noted that, properly designed, the same large near-IR spectroscopic survey could serve the first key question as well.
- It is *important* to obtain spatially resolved spectroscopy of a representative sample of star-forming galaxies using IFUs on ground-based ELTs and NIRSpec on JWST.

To probe baryons when they are in and around and between galaxies, absorption spectra of background sources along lines of sight passing near galaxies can be used. Such tomographic probes of galaxies and intergalactic gas on transverse scales <1 Mpc from galaxies will yield revolutionary new insights into gas accretion, outflows, chemical enrichment, and the cycle of matter between galaxies and the IGM during the peak epoch of galaxy growth.

Currently, the best-understood transitions for probing inflowing and outflowing diffuse gas lie in the rest-UV, and include transitions of H I, C IV, O VI, and S IV. Until now, the background sources used for high-resolution IGM studies have necessarily been rare bright objects (typically quasars), providing only a sparsely sampled view of the gas around foreground galaxies. In the coming decade, the larger collecting areas of ELTs will open up a new frontier for the study of baryon cycling. At $z > 1.5$, the key transitions can be probed in the optical, providing a census of baryons and metals around galaxies. High spectral resolution is required to infer accurate ionization conditions for metals and to resolve the photoionized Ly-α forest to the thermal width limit. Yet even with the many faint background sources accessible with ELTs, there will be at most a few galactic absorption systems along any one sight line. Larger samples can be obtained at $R \sim 5{,}000$ using typical star-forming galaxies as background probes with much higher surface density.

At redshifts $z < 1.5$, these same tracers migrate from the optical into the UV. To increase dramatically the sampling density of background probes of warm baryons in the cosmic web requires the development of a 4-meter-class UV telescope with high-resolution spectroscopic capabilities, which can access significantly fainter quasars and galaxies than is currently feasible with HST/Cosmic Origins Spectrograph. Such a facility would help provide a comprehensive view of baryons around galaxies, groups, clusters, and in filaments out to $z \sim 1.5$.

At the present epoch, a large fraction of baryons is thought to be shock-heated and in the WHIM, the bulk of which is best traced by soft X-ray absorption lines. Greater collecting areas in the X-ray are needed to obtain a statistical sample (approximately 100) of background sources for absorption studies of oxygen (O VII, O VIII) and carbon (C V, C VI).

A new and potentially ground-breaking area of study consists of mapping the metal and hydrogen-line emission from circumgalactic and dense filamentary intergalactic gas, using IFUs in the UV. Such observations will provide direct information about the kinematics, density, and chemical content of baryons cycling between galaxies and the IGM.

In short, absorption-line tomography toward background galaxies and quasars and direct emission-line imaging can yield unprecedented information about the kinematics, density, and chemical content of baryons cycling between galaxies and the IGM. Simulations will be critical for connecting such one-dimensional probes to the three-dimensional gas distribution. The panel concluded:

- It is *most important* to use ELTs to map metal- and hydrogen-line absorption in circumgalactic and dense filamentary intergalactic gas, at moderate resolution toward background galaxies and at higher resolution toward background quasars.
- A 4-meter-class UV-optimized space telescope, equipped with a high-resolution spectrograph and an IFU for spectral mapping, is *very important* for characterizing outflows from galaxies and AGN at $z < 1.5$ and for mapping the WHIM.
- It is *important* to obtain moderate-resolution soft-X-ray spectra for approximately 100 quasar sightlines.

A complete inventory of cold gas in and around galaxies is also crucial for understanding baryon cycling. Molecular gas traced by CO, C I, and higher-density probes provides the raw material for star formation. Neutral atomic gas in the circumgalactic medium likely feeds the growth of galaxy mass. Direct observations of cold gas will make it possible to test theoretical models for complex gas physics and predictions for the evolution of gas content, and will permit the extension of empirical scaling relations (such as the Schmidt-Kennicutt relation) to high redshifts. These data will allow the construction of an accurate picture of how gas is processed into stars during the epoch $1 < z < 3$, when roughly half of the stellar mass in the universe was formed.

Facilities coming online in the next decade will enable major advances toward a cosmic census of cold gas. Continuous frequency coverage and wide bandwidths below 50 GHz (EVLA) and above 84 GHz (baseline ALMA) will make it possible to map the same molecular emission lines in different galaxies across broad, albeit incomplete, ranges in redshift. Baseline ALMA could determine molecular gas masses for a representative sample of several hundred typical star-forming galaxies at $z \sim$ 1-3. These data would provide a dramatic improvement in sample size and breadth over the approximately 60 highly luminous or lensed objects at $z > 1$ that have been detected in molecular gas with current facilities (Figure 3.14).

There are two additional priorities for investment over the next decade. The first is the need for faster (sub)millimeter spectroscopic follow-up of large samples with sparse sky distributions relative to the typical interferometer fields of view; this could be achieved by equipping interferometers or large single-dish telescopes

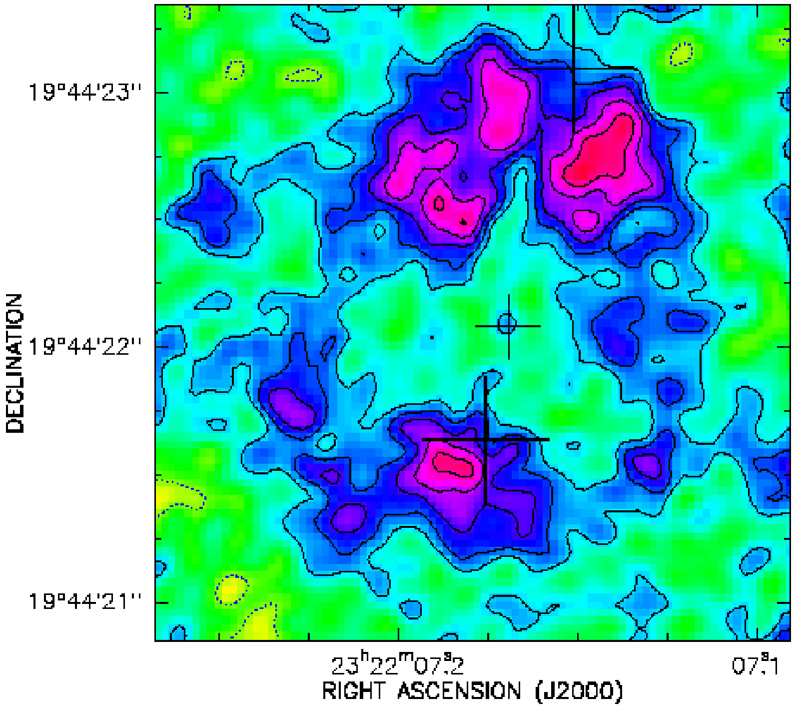


FIGURE 3.14 High-resolution (0.3″) Very Large Array map of CO ($J = 2 \rightarrow 1$) emission from the $z = 4.12$ quasar PSS J2322+1944, tracing cold molecular gas in its host galaxy. Gravitational lensing into a full Einstein ring makes this system easier to detect and map. The new capabilities of the Expanded Very Large Array and the Atacama Large Millimeter Array will make it possible to extend such work to wider ranges in redshift and to target more typical high-redshift galaxies that are neither extremely luminous nor highly lensed. SOURCE: D.A. Riechers, F. Walter, B.J. Brewer, C.L. Carilli, G.F. Lewis, F. Bertoldi, and P. Cox, A molecular Einstein ring at $z = 4.12$: Imaging the dynamics of a quasar host galaxy through a cosmic lens, *Astrophysical Journal* 686(2):851, 2008, reproduced by permission of the AAS.

with multiobject spectroscopic capability. The resulting observations would allow more efficient detection of emission lines from the obscured galaxies detected in large continuum surveys and would help reduce the orders-of-magnitude discrepancy between the large spectroscopic sample sizes for high-redshift galaxies in the optical and those in the radio, which is a major impediment to synthesizing a full picture of baryon cycling.

The other priority is to work toward large-collecting-area facilities that can detect H I emission at $z > 0.4$ by means of the 21-cm line. Enhancing the low-frequency performance of existing interferometers (to reach $z \sim 0.8$ in emission) may be a fruitful initial step in this direction. For studying neutral gas in galaxies and the cosmic web, the key frequency range for comparison with studies at other wavelengths and theoretical models of cold accretion will be 200 to 500 MHz (corresponding to $z = 2 \sim 6$ for the 21-cm line). At $z \sim 3$, substantially larger collecting areas than those of current facilities will allow for the detection of H I masses comparable to that of the Milky Way, an ambitious but essential long-term goal. For starting an inventory of cold gas in and around galaxies, the panel concluded:

- It is *most important* to detect CO emission from a representative sample of typical star-forming galaxies from $z \sim$ 1-3, to develop technology for faster spectroscopic follow-up in the (sub)millimeter, and to develop large-collecting-area facilities to study H I in emission at $z \sim$ 1-3.

The majority of star formation at $1 < z < 3$ is obscured by dust, and hence accurately characterizing reprocessed emission is critical to obtaining a full view of star formation across cosmic time. To probe obscured star formation, deep and wide surveys at radio and submillimeter wavelengths are essential; with its new correlator, the EVLA will offer substantial gains in continuum sensitivity in the early part of the next decade. These could be complemented by surveys in the far-IR and (sub)millimeter with apertures large enough to probe below the confusion limits of today's facilities, and push down to the lower flux densities at which star-forming galaxies selected at other wavelengths are expected to appear. (Sub)millimeter mapping of deep spectroscopic survey fields, in two or more submillimeter bands to minimize biases in dust temperature and redshift, would significantly enhance the understanding of early gas processing. For the most luminous and obscured sources, far-infrared spectroscopy could be used to break the degeneracy between the contributions of star formation and black hole accretion to the bolometric luminosity. The panel concluded:

- It is *very important* to do sensitive radio and (sub)millimeter continuum mapping over large areas, preferably coincident with a near-IR (rest-frame optical) spectroscopic survey such as the one described above, and to carry out far-IR spectroscopy of luminous dusty galaxies.

GCT 3. HOW DO BLACK HOLES GROW, RADIATE, AND INFLUENCE THEIR SURROUNDINGS?

- How do black holes grow over cosmic time?
- How do the quantity and form of energy production in accreting systems depend on black hole mass, accretion rate, and spin?
- How does black hole feedback shape the evolution of cosmic structures?

Progress to Date

Supermassive black holes appear to play a unique—and surprising—role in shaping the structure and evolution of galaxies and clusters. If the basic idea of AGN feedback is borne out, it represents a truly remarkable example of multiscale physics—an object a few light-minutes across influencing structure hundreds of thousands of light-years in extent. The central engines of AGN are also interesting in their own right, providing a rare window into the physics of strong gravity and high-energy particle acceleration.

Accretion onto black holes liberates gravitational energy and may also tap energy stored as black hole spin. The energy can be released in various forms, including radiation from an accretion disk and the ejection of powerful jets of relativistic plasma (Figure 3.15).

The past decade has seen dramatic advances in both observational and theoretical studies of the accretion and jet-formation processes. X-ray studies have revealed spectral features originating close to the black hole event horizon. These broad iron lines confirm the presence of thin accretion disks; measuring their profiles has led to the first indications of black hole spin. This places astronomers at the threshold of being able to correlate black hole spin with other fundamental properties, such as the power of relativistic jets, as well as to use spin as a probe of the evolutionary history of SMBHs. X-ray absorption spectroscopy has led to the identification of high-velocity, high-ionization accretion disk winds with mass flow comparable to or greater than the mass accretion rate onto the black hole. Multi-epoch interferometric radio observations have imaged AGN jets with spatial resolutions down to a few tens of Schwarzschild radii, allowing observers to watch ejection events evolving along the jet in real time (Figure 3.16). New ground- and space-based X- and gamma-ray observatories promise to deliver fresh insights into the processes of particle acceleration by AGN and the structure and energy content of AGN jets.

On the theoretical side, the past decade has seen a significant deepening in understanding of the mechanism of accretion disks, in part from fully general relativistic MHD simulations of disks based on first principles. These simulations have progressed to the point of including the jet-launching region around rapidly rotating black holes (see Figure 3.15).

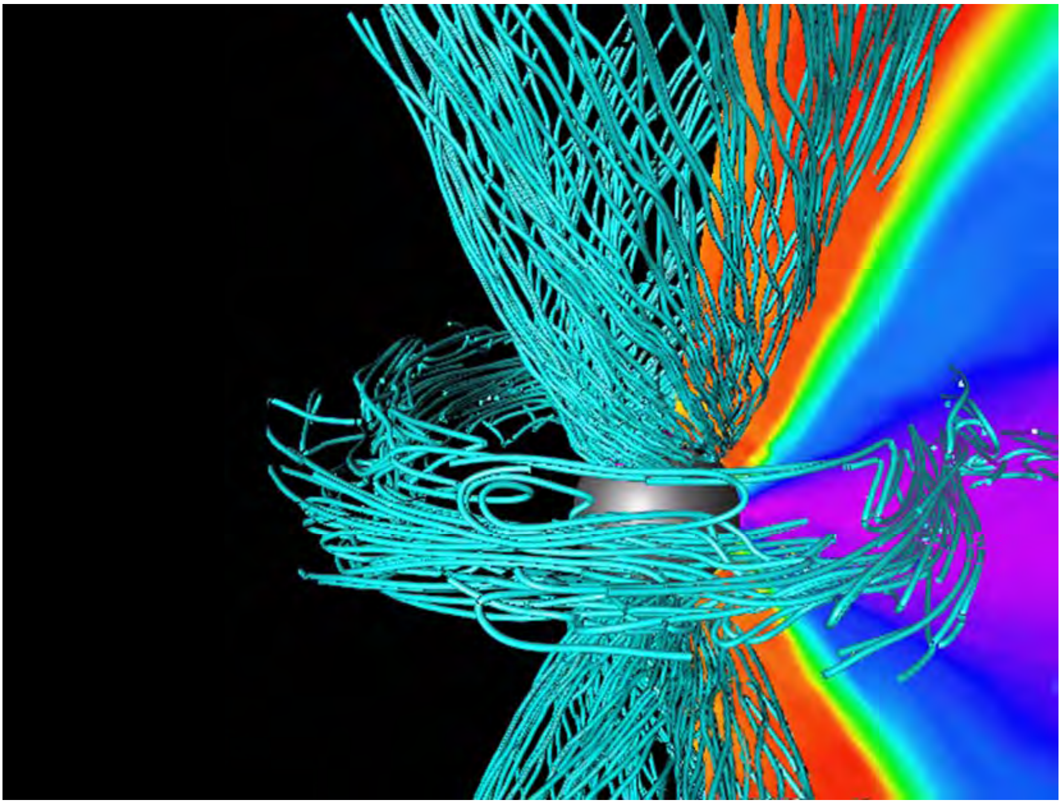


FIGURE 3.15 Illustration of the cross section of a black hole accretion disk from a global hydrodynamical simulation (colors) with magnetic field lines (shown in light blue) threading the black hole. SOURCE: Courtesy of S. Hirose, Japan Agency for Marine-Earth Science and Technology; J.H. Krolik, Johns Hopkins University; J.-P. De Villiers, University of Alberta; J.F. Hawley, University of Virginia; and Richard Sword, University of Cambridge, Institute of Astronomy.

There has been remarkable progress in the development of knowledge of black hole demographics over the past decade. For AGN, this is principally due to deep panchromatic surveys from the radio to the X-ray. These new data have made clear that the majority of accretion occurs in obscured AGN, in which lines of sight are blocked by high column densities of dust and gas that extinguish much of the optical, UV, and soft X-ray radiation and reprocess it into the infrared. It is now known that the first quasars appeared before a redshift of $z \sim 6$ and that SMBHs are present at the centers of massive galaxies.

There is evidence that the injection of momentum, mass, and energy from SMBHs plays a role in the development of structure on galactic and cluster scales.

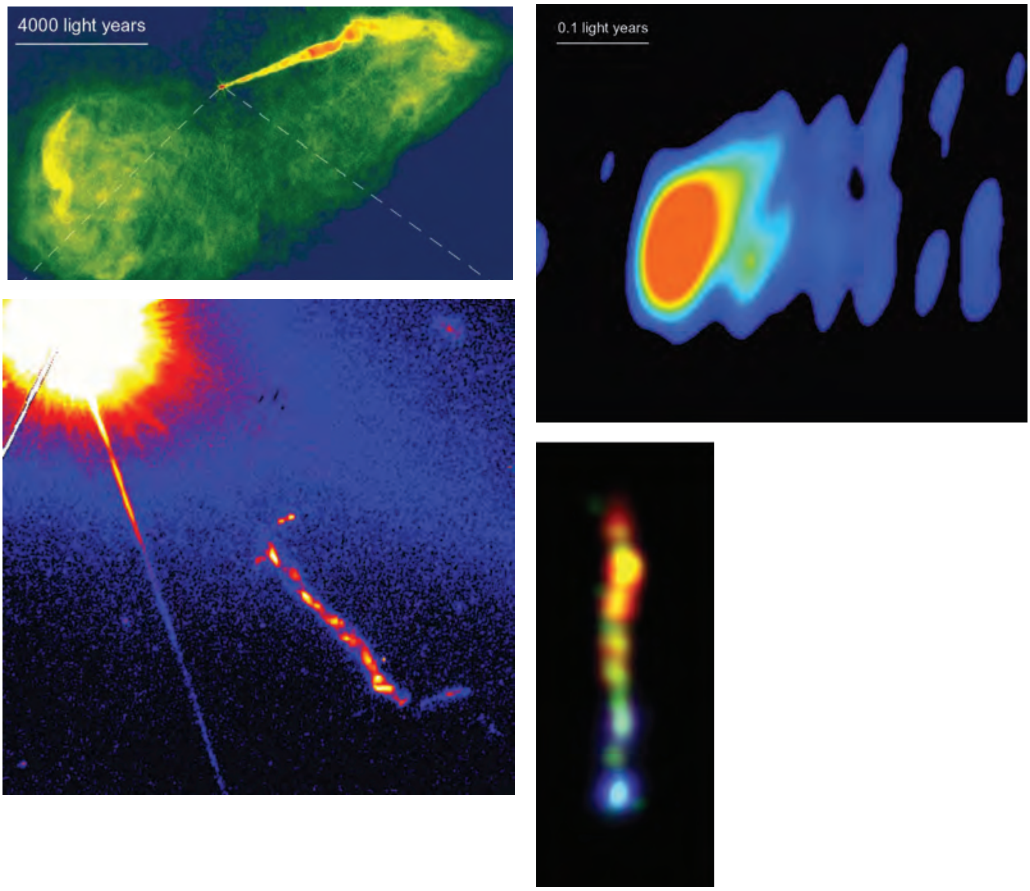


FIGURE 3.16 *Top:* Radio images of the galaxy M87 at different scales show (*top left*, Very Large Array [VLA]) giant, bubble-like structures of radio emission powered by the jets from the galaxy's central black hole and (*top right*) Very Long Baseline Array image of the jet being formed into a narrow beam within a hundred Schwarzschild radii of the black hole. The scale of each image is shown by white bars; 1,000 light-years is about 60 million times the distance from Earth to the Sun, and 0.1 light-year is about 1,000 Schwarzschild radii for M87's black hole mass of $3 \times 10^8 M_{\odot}$. *Bottom:* Hubble Space Telescope (HST) optical image of the quasar 3C 273 and its kiloparsec-scale jet (*bottom left*) and a multiwavelength Chandra/HST/Spitzer/VLA image of the jet alone (*bottom right*). The jet emission is synchrotron radiation from energetic electrons accelerated by the jet's magnetic field, and extends nearly 40,000 light-years across the sky. The highest-energy particles, which radiate X-rays (blue), lose their energy quickly, whereas the lower-energy electrons that radiate optical (green), infrared (red), or radio (yellow) light persist to the end of the jet. Unresolved gamma-ray emission is also detected (e.g., with the Fermi gamma-ray space observatory) from these and other active galactic nuclei and is important for understanding the kinetic energy of jets. Calculating the electron energy from modeling the emission constrains the jet power, which is essential to understanding jet formation and propagation. SOURCE: *Top:* Courtesy of NASA, National Radio Astronomy Observatory/National Science Foundation, John Biretta (STScI/Johns Hopkins University), and Associated Universities, Inc. *Bottom left:* Courtesy of NASA/STScI. *Bottom right:* NASA/JPL-Caltech/Yale University.

In the local universe, SMBH mass is correlated with properties of the host galaxy, suggesting regulated growth. Theoretical studies of structure formation show that, without the injection of energy from SMBHs and/or supernovae at the right times and places, models predict galaxies that are too massive, too blue, and have the wrong angular momenta, as well as baryons in clusters that cool far too much. SMBH feedback appears to be essential in resolving the problem whereby theory predicts overmassive blue galaxies and overcooling in clusters, whereas some mixture of SMBH and supernova feedback may be needed to explain other galaxy properties. Furthermore, theory indicates the need for at least two forms of AGN energy injection: strong outflows or radiation that inhibits star formation in massive galactic bulges when the black hole is growing, and a "maintenance mode" that gently heats baryons in massive halos and prevents further galaxy growth (Figure 3.17).

Direct evidence for either feedback process has been elusive. Radio and X-ray observations of galaxy clusters and groups clearly show the interaction of AGN jets with the ICM (see Figure 3.7), but definitive indications of direct ICM heating or the existence of a regulated feedback loop have yet to be found. High-velocity outflows from AGN at both high and low redshift indicate appreciable ejection of energy, mass, and metals, but degeneracies in modeling the current data prevent this from being quantified. There appears to be coevolution of SMBHs and galaxies, with the spatially averaged star-formation rate and black hole luminosity roughly tracking each other at $z < 2$, but its physical origin is unknown.

Steps for the Next Decade

Relativistically broadened X-ray reflection features, principally the Kα line of iron, are the only clean probes of the inner regions of SMBH accretion disks. Due to their limited throughput, current X-ray observatories (principally Chandra, X-ray Multi-mirror Mission (XMM)-Newton and Suzaku) can only determine time-averaged iron-line profiles (Figure 3.18). Major advances in the understanding of black holes and relativistic accretion disks will be enabled by a next-generation large X-ray telescope with high-throughput high-resolution spectrometers, sensitive in the 0.5- to 10-keV band. Such a facility would allow the characterization of the variability of broad iron lines that is expected to be associated with orbiting coronal features (on dynamical timescales) and X-ray light echoes associated with rapid flares (i.e., reverberation mapping on light-crossing timescales). General relativistic effects such as the Shapiro delay are imprinted on these line variations and would allow the measurement of the masses and spins of the black holes. Characterizing short-timescale iron-line variability would also provide the first robust constraint on the geometry and size of the X-ray-emitting corona, a structure that is responsible for processing a significant fraction (20 percent or more) of the total

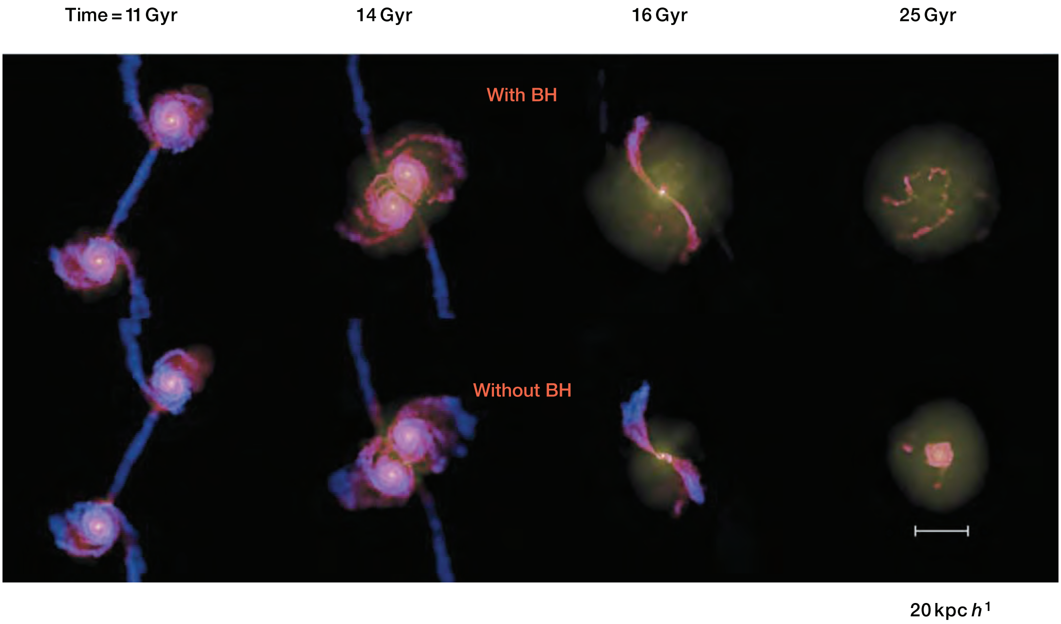


FIGURE 3.17 Snapshots of the time evolution of a merger of two galaxies with (*top*) and without (*bottom*) supermassive central black holes illustrate the possible role of active galactic nuclei (AGN) feedback in galaxy formation. The images visualize the projected gas density color-coded by increasing temperature (blue to red). The supermassive black holes accrete from the surrounding gas, undergoing a quasar phase close to the merger event, and the associated feedback energy heats and expels a large fraction of the gas after the major merger. In this simulation, AGN feedback turns the galaxy into a red-and-dead elliptical. SOURCE: Reprinted by permission from Macmillan Publishers Ltd.: *Nature*, T. Di Matteo, V. Springel, and L. Hernquist, Energy input from quasars regulates the growth and activity of black holes and their host galaxies, *Nature* 433:604-607, 2005, copyright 2005.

accretion power. In particular, it would be possible to distinguish disk coronae from the X-ray-bright bases of jets.

A complementary survey of time-averaged broad iron lines in fainter AGN will yield the spin distribution function of local SMBHs, a powerful constraint on SMBH evolution models. Furthermore, comparison of spins in radio-loud and radio-quiet subsamples could provide the first direct observational test of whether black hole spin powers jets. These data will also allow a determination of the properties of the disk as a function of black hole mass, spin, accretion rate, and jet power. The technical requirements for this spin survey are similar to those needed for the study of time-variable iron lines. Astro-H will be a first step in this direction, yielding high-quality average spectra of broad iron lines in approximately 20 AGN. It is also essential that improved theoretical models of accretion disks continue to

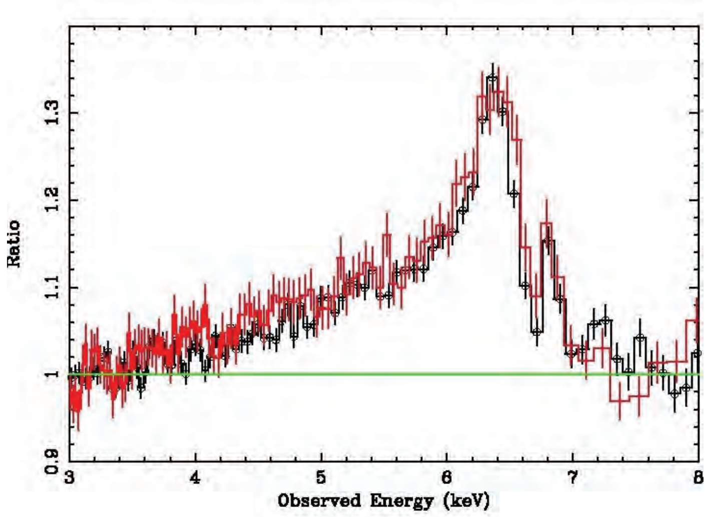


FIGURE 3.18 Broadened iron Kα emission line in the time-averaged X-ray spectrum of the active galaxy MCG-6-30-15, represented as a ratio of the data against the underlying power-law continuum. Shown here are data from X-ray Multi-mirror Mission-Newton (red) and Sukazu (black). This emission line is believed to be produced in the surface layers of the inner accretion disk around the central supermassive black hole (SMBH), with the highly broadened and skewed line profile caused by relativistic Doppler broadening and a strong gravitational redshift. In this galaxy, the extreme line broadening implies that the SMBH must be rapidly rotating. SOURCE: Reproduced with permission from G. Miniutti, A.C. Fabian, N. Anabuki, J. Crummy, Y. Fukazawa, L. Gallo, Y. Haba, et al., Suzaku observations of the hard x-ray variability of MCG-6-30-15: The effects of strong gravity around a Kerr black hole, *Publ. Astron. Soc. Japan* 59:S315-S325, 2007, copyright 2007, Astronomical Society of Japan.

be developed concurrently. The ability to model thin accretion disks dominated by radiation pressure is a particularly important goal.

The structure of AGN accretion disks and jets can be explored by means of X-ray polarization measurements. Energy-dependent X-ray polarimetry of radio-quiet (disk-dominated) AGN will constrain the geometry of the X-ray-emitting accretion disk corona. The X-ray polarization of blazars will probe the structure of the jet magnetic field. The results obtained with the recently approved Gravity and Extreme Magnetism Small Explorer are expected to motivate further investigations of X-ray polarization, and some other future X-ray missions have additional polarization capabilities.

To understand the details of accretion onto supermassive black holes, jet formation, and energy dissipation, the panel concluded:

- It is *most important* to have sensitive X-ray spectroscopy of actively accreting black holes (AGN) to probe accretion disk and jet physics close to the black hole as well as to determine the spin distribution function of the local SMBH population. The effective area should be sufficient to detect the iron Kα emission line on dynamical timescales in a modest sample of the brightest AGN, yielding both spin and mass. In order to disentangle the effects of absorption in AGN spectra, high resolution ($R > 2{,}000$) is required. The same capabilities will yield time-averaged line profiles of more than a hundred AGN with sufficient signal-to-noise ratios to derive the black hole spin distribution.

- It is *important* to measure X-ray polarization down to a few percent in AGN and blazars to explore disk structure and jet magnetic field geometry.

A major limitation on the understanding of AGN feedback in clusters/groups is ignorance of the dynamical state of the ICM. High-throughput, high-resolution X-ray spectroscopy will reveal bulk motions and turbulence in the ICM, allowing the AGN/ICM coupling to be explored. The dependence of AGN-induced motions on a group or cluster mass scales will be determined, providing a powerful new observational constraint on the mass dependence of SMBH feedback. Astro-H should be able to detect bulk motion in a few nearby clusters with central AGN. A vigorous program of theoretical research on AGN/ICM interactions will be needed in order to interpret these data, including sophisticated simulations of jet-induced feedback and studies of the basic physics of the ICM plasma.

High-resolution time-resolved spectroscopy in the rest-frame UV and X-ray bands will enable the mass, metal, and energy fluxes from quasar winds to be assessed. This is vital for understanding AGN feedback, since such winds are potentially prime sources of feedback energy. Time-resolved spectroscopy is needed to determine the recombination time of the (photoionized) absorbing outflow and hence break the degeneracy that exists in a single spectrum between the distance of the absorbing material and its density. Sufficient spectral resolution is needed to measure wind velocities and ionization states.

Therefore, in order to seek evidence of black hole feedback, the panel concluded:

- It is *most important* to measure turbulence and/or bulk flows using X-ray imaging spectroscopy of the ICM of nearby galaxy clusters and groups, with sufficient image quality, field of view, energy resolution, and signal to noise to provide ionization and velocity maps on the scale of the interaction between the AGN outflow (e.g., radio source) and the gas.
- It is *very important* to obtain time-resolved spectroscopy of AGN in order to study quasar winds and their possible role in feedback. The Cosmic Origins Spectrograph on HST has the necessary sensitivity in the UV; in X-rays, capabilities beyond XMM-Newton and Chandra are required.

Complementary multiwavelength surveys to perform a census of active and inactive black holes across cosmic time and to track their growth are crucial for determining the numbers and masses of black holes, the prime mode of their energy generation, and the relationship of black hole and galaxy properties. Because a large fraction of black hole growth is obscured by dust and gas, an AGN census must be sensitive to both obscured and unobscured objects. Since no single selection method detects all such AGN, a combination of at least four techniques is needed: (1) wide and deep hard X-ray ($E > 10$ keV) surveys, because observations in the rest-frame hard X-ray band ($E \sim$ 10-50 keV) are unbiased with respect to column densities up to 3×10^{24} cm^2; (2) infrared color and emission line surveys, because the absorbed energy is reprocessed and re-emitted in the IR; (3) sensitive spectroscopic surveys in the rest-frame optical/UV, focusing on detection of high-ionization narrow lines such as O III, O IV, and Ne V, because these are mostly generated in regions outside the obscuring material and their luminosity seems to track the AGN bolometric luminosity; and (4) radio continuum surveys, which are completely insensitive to obscuration and can readily detect jets (which, however, occur in a minority of AGN).

Of these four techniques, the hard X-ray survey is the most important because hard X-rays are most directly connected to the energy-generation mechanism in AGN. The observation of hard X-rays ($E >$10 keV) unequivocally signals the presence of an AGN, whereas infrared continuum, some optical emission lines, and radio continuum can be produced by, for example, star formation. The panel anticipates major discoveries in this area with the 2011 launch of NuSTAR, a sensitive narrow-field instrument that will find moderate luminosity obscured AGN at $z \sim 1$. This is analogous to the deep GOODS survey in the optical/IR, which found mainly lower-luminosity galaxies and AGN, and is no substitute for an SDSS-type wide-area survey, which is sensitive to the most-luminous quasar-like objects. To study the full range of galaxies and AGN, NuSTAR must be complemented by a sensitive hard X-ray survey over an appreciable fraction of the sky.

Because of their great sensitivity, infrared and O III surveys can detect large numbers of objects in short observing times over a wide redshift range. Optical spectroscopy can also be used to estimate black hole masses. At present radio surveys go extremely deep, but to understand the survey content, follow-up studies will be needed in the optical and IR bands. Optical and IR follow-up studies will also be useful to estimate black hole masses from broad-line widths.

The panel concluded:

- It is *very important* to do complementary multiwavelength surveys to track the growth of black holes across cosmic time. A hard X-ray all-sky survey for AGN is an essential complement to the deep pencil-beam surveys of active galaxies expected from the upcoming NuSTAR Explorer. Long-wavelength IR surveys capture the total energy output, and rest-frame optical spectroscopic surveys allow black hole mass determinations.

Gravitational radiation from merging SMBHs can be detected with a space-based gravitational wave observatory. Although the restricted mass range and the possibility of small-number statistics will prevent a detailed reconstruction of the SMBH merger tree, gravitational radiation results will allow the discrimination of small-seed (10^2 to 10^3 $M_\odot$) and large-seed (10^5 $M_\odot$) scenarios for early SMBH growth and will determine the masses and spins of some objects. It will, however, be impossible to relate these SMBH mergers to the host galaxy properties without finding a significant number of electromagnetic counterparts. On balance, the panel concluded:

- The search for gravitational radiation from merging supermassive black holes, at lower frequencies than are probed with the Laser Interferometer Gravity Observatory, is *very important* for an understanding of the buildup of supermassive black holes.

Correlations between the masses of supermassive black holes and the properties of the galaxies that host them are an important fossil of the feedback processes that couple SMBH and galaxy evolution. This is evidenced by the remarkable M-σ relation locally, which was established via the dynamical measurement of the masses of SMBHs in several dozen galaxies. Adaptive optics on ELTs will allow dynamical measurement of SMBH masses in thousands of low-redshift galaxies (including low-mass systems for which current constraints are poor), as well as searches for SMBHs with masses $>10^8$ $M_\odot$ at $z > 0.1$ and for extremely massive SMBHs ($\geq 10^9$ $M_\odot$) at any redshift. These new data will yield the intrinsic mass-dependent scatter in the M-σ relation and perhaps its evolution, as well as any dependence on host galaxy morphology, bulge velocity dispersion, or galaxy luminosity. The only other probe of inactive galactic nuclei is the detection of transient events caused by the accretion of stars. Recent UV and soft X-ray studies have found several normal galaxies whose nuclei brightened greatly, presumably due to the tidal disruption of stars. Large time-domain surveys will find many more such objects, allowing the detection of normally quiescent black holes. The panel concluded that, in order to extend the M-σ relation to low-mass local systems and massive high-redshift systems:

- It is *very important* to measure dynamical masses of a thousand or more SMBHs, using spatially resolved spectroscopy with adaptive optics on ELTs. SMBH masses in smaller samples are possible with centimeter-wave maser observations with very long baseline arrays supplemented by large aperture dishes, and/or reverberation mapping of AGN (see the discussion for the first key question).

Finally, the bases of jets can be studied by way of multiwavelength observations of nearby radio galaxies and blazars. Exciting discoveries are just beginning to come from multiwavelength campaigns involving the VLBA, (sub)millimeter very long baseline interferometry (VLBI) observations, and the current generation of ground-based and spaceborne gamma-ray observatories. The observations will show which improvements in spatial resolution and sensitivity are needed for further progress. The panel concluded:

- Multiwavelength observations of blazars, including ground-based and spaceborne gamma-ray observatories, the VLBA, and millimeter VLBI (and eventually submillimeter VLBI), are *important* for determining jet composition and constraining formation models.

GCT 4. WHAT WERE THE FIRST OBJECTS TO LIGHT UP THE UNIVERSE, AND WHEN DID THEY DO IT?

- Where and when did the first objects form?
- When did the first galaxies emerge, and what were they like?
- How did these first objects reionize the universe?

Progress to Date

The epoch between the last scattering of the cosmic microwave background radiation, at $z \sim 1{,}100$, and the current high-redshift frontier at $z \sim 8$, where the most distant quasars, GRBs, and galaxies have been observed, remains completely unexplored. This epoch contains the first stars, first galaxies, and first massive black holes. These objects must produce the ionizing photons that led to the reionization of the neutral hydrogen (formed at recombination, $z \sim 1{,}100$) in the universe. The formation of the first astrophysical objects and the subsequent epoch of reionization are at the frontiers of astrophysical research in the next decade.

The study of the first objects and the epoch of reionization (Figure 3.19) is an *area of discovery*, leading into uncharted territory. Theoretical efforts are the only guide to understanding what the first luminous objects were and how they manifested themselves. Rapid progress is being made, and the interplay of theory and observations promise to answer key questions if, in the next decade, a comprehensive array of empirical means to study this epoch are developed.

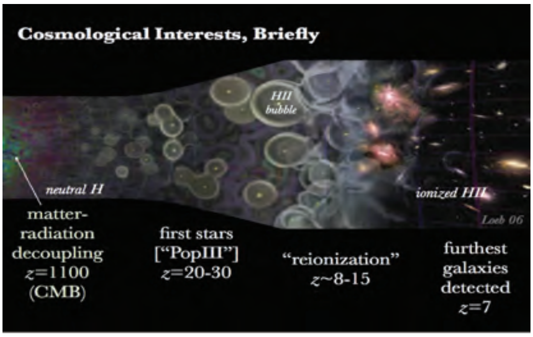


FIGURE 3.19 Cosmic timeline showing the formation of structure in the universe from the Dark Ages (*left*) to the formation of the first stars, galaxies, and black holes that started the reionization of the intergalactic medium, which is transparent in the present day (*right*). SOURCE: Courtesy of Abraham Loeb and Jean-Francois Podevin, adapted from Abraham Loeb, The dark ages of the universe, *Scientific American* 295:46-53, 2006.

GCT Discovery Area—The Epoch of Reionization

The discovery area identified by the Panel on Galaxies Across Cosmic Time is closely entwined with the panel's fourth key question, What were the first objects to light up the universe, and when did they do it? Originally the panel intended to separate the two, with the fourth key question addressing that part of the early universe for which there is existing knowledge (e.g., objects discovered out to $z \sim 8$), and the discovery area referring to the less-well-understood part of the epoch of reionization (e.g., collapse of the first H I structures). However, it proved simpler and more straightforward to combine the two; hence this section includes both the focused questions and the broader inquiry appropriate to a discovery area.

Recent years have seen significant progress both in theoretical understanding and in observational probes of this transition epoch in the early universe. Observations of high-redshift quasars have uncovered a number of luminous objects at $z \sim 6$, the spectra of which imply that the process of reionization was complete by then. There are hints of objects at $z \sim 7$ and a few candidates at even higher redshifts that have been detected with the aid of gravitational lensing. GRBs have been discovered out to $z \sim 8.2$. The recent measurements of the optical depth to electron scattering from the Wilkinson Microwave Anisotropy Probe (WMAP) polarization studies suggest that the first sources of light had already significantly reionized the IGM at $z > 11$. There is no doubt from these observations that the first galaxies and quasars were well in place by $z \sim 6$ and may have started to appear as early as $z \sim 15$-20. The actual history of reionization is still highly uncertain (as illustrated in Figure 3.1); Figure 3.20 summarizes current observational constraints on the inferred neutral hydrogen fraction of the IGM.

Theoretical work suggests a general framework for the formation of the first objects: ΛCDM simulations indicate that the first objects formed as early as $z \sim 40$, out of gas that cooled via molecular hydrogen in the first dark matter halos. Most

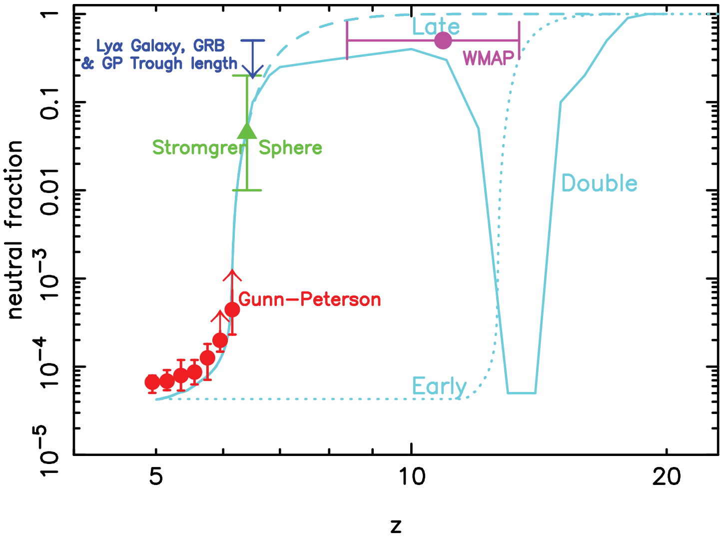


FIGURE 3.20 Current constraints on the volume average neutral fraction of the intergalactic medium versus redshift. The red points indicate measurements based on the highest-redshift quasars. The green triangle shows the constraints from cosmic Strömgren spheres and surfaces around the highest-redshift quasars, and the magenta point indicates the Wilkinson Microwave Anisotropy Probe constraint. The curves show the expectations for different assumptions about early star formation. SOURCE: Adapted from X. Fan, C.L. Carilli, and B. Keating, Observational constraints on cosmic reionization, *Annual Review of Astronomy and Astrophysics* 44:415-462, 2006.

simulations suggest that the first stars, called Population III, were predominantly very massive; their demise should have created the first black holes. The soft UV and X-rays emitted by these first stars and black holes provided significant feedback to the chemical and thermal state of the IGM by driving expanding ionized H II regions into it (Figure 3.21). These events may have marked the beginning of reionization. Having dissociated the molecular hydrogen, radiative feedback from the first stars led to a second generation of stars and galaxies in which the gas cooled through radiative transitions in atomic hydrogen and helium. The ionized regions surrounding these primordial galaxies grew and merged until they fully overlapped, marking the end of reionization.

It is still very uncertain whether stars were the first significant ionizing sources. An alternative possibility is that a similar fraction of this gas formed massive black holes directly (see Figure 3.21).

Moreover, massive metal-free stars should (in most cases) leave behind stellar-mass seed black holes that accrete gas as mini-quasars. Depending on the efficiency with which ionizing photons escape into the IGM, the first halos might blow away their gas after a single episode of star formation and might not form any stars until

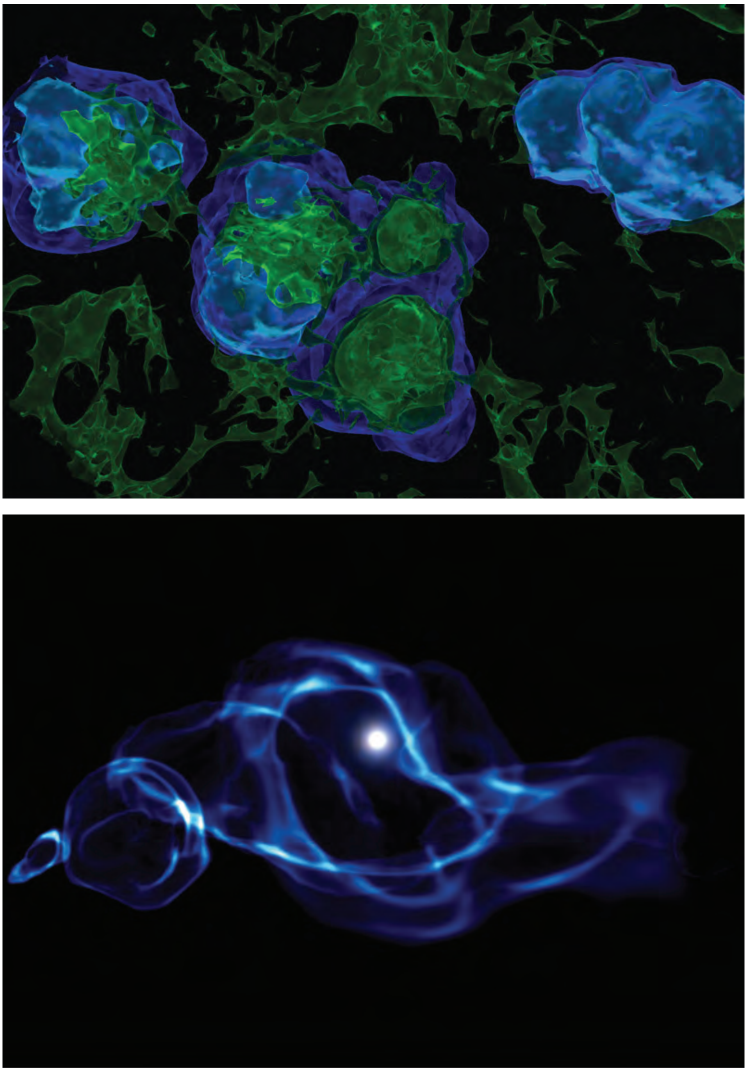


FIGURE 3.21 Computer simulations of first stars and black holes illustrating the initial steps in the process of reionization. *Top:* Ionizing radiation emanates from the first massive stars that form inside dark matter mini-halos, creating ionized bubbles (blue) interspersed with regions of high molecular abundance (green), both embedded in the still-neutral cosmic gas. The large residual free-electron fraction inside relic H II regions, left behind after the central star has died, rapidly catalyzes the formation of molecules. Simulations were performed by Bromm et al. (2009), and visualization is courtesy of the Texas Advanced Computing Center. *Bottom:* Simulation showing the X-rays produced by a black hole (white), created when a first-generation star collapses, and their ionizing effect on nearby gas (blue). SOURCE: *Top:* Reprinted by permission from Macmillan Publishers Ltd.: *Nature*, V. Bromm, N. Yoshida, L. Hernquist and C.F. McKee, Formation of the first stars and galaxies, *Nature* 459:49-54, 2009, copyright 2009. *Bottom:* Courtesy of KIPAC/SLAC/M. Alvarez, T. Abel, and J. Wise.

deeper potential wells are formed. In general, several feedback processes are likely to be important at these stages, from internal feedback (e.g., due to the presence of supernovae) to external global feedback provided by the UV and X-ray background produced by the first sources. Numerical simulations are the most promising way to address many of these issues; however, accommodating the dynamic range required to resolve the small-scale physics in mini-halos and the large cosmological volumes needed to capture the emergence of rare density peaks is extremely challenging, especially since radiative transfer effects need to be included.

Steps for the Next Decade

Existing observations provide only a preliminary glimpse into the late stages of the epoch of reionization (Figure 3.22). Because this is an *area of discovery*, it is difficult to outline specific future science programs, and the panel was guided primarily by uncertain theoretical predictions. The first stars should have been essentially metal-free and extremely massive ($M > 100\ M_{Sun}$), with a radiation field that is very efficient at ionizing hydrogen and helium. Thus, direct observation of the first stars depends on the detection of the rest-frame UV continuum and line emission, particularly the He II (1,640 Å) and Ly-α emission lines, which for redshifts $z < 11$ will appear in the J band. The ability to detect the first stars depends critically on the physics of star formation in a metal-free environment (e.g., initial mass function, star-formation efficiency, etc.) and the extent to which these first objects are clustered. While individual stars will be much too faint to be detected directly with JWST or an ELT, if the first stars formed in larger aggregates then the resulting H II regions are expected to have line fluxes approximately 10^{-21} erg cm^{-2} s^{-1}.[3] The likeliest way forward is to use JWST deep fields, especially with the aid of gravitational lensing, to find the earliest stars and protogalaxies. Simply finding the first objects is not enough: to probe the astrophysics of the first objects requires follow-up near-IR spectroscopy ($R \sim 5{,}000$) and to determine the properties of the IGM requires absorption-line spectroscopy.

The first explosions are key probes of the beginning of the epoch of reionization. Detection of transient objects during this epoch will allow the study of the intervening IGM through absorption spectroscopy. GRBs can be used as probes of the high-redshift universe, provided that several dozen with $z > 8$ are detected; this would take 2 to 5 years for a facility that detects GRBs at 10 times the rate that Swift does. It is critical to understand stellar astrophysics in metal-free conditions in order to predict the properties of the first supernovae. A Type IIn supernova should

---

[3] See Astronomy and Astrophysics Advisory Committee, *GSMT and JWST: Looking Back to the Future of the Universe*, available at http://www.nsf.gov/mps/ast/aaac/reports/gsmt-jwst_synergy_combined.pdf. Accessed February 2010.

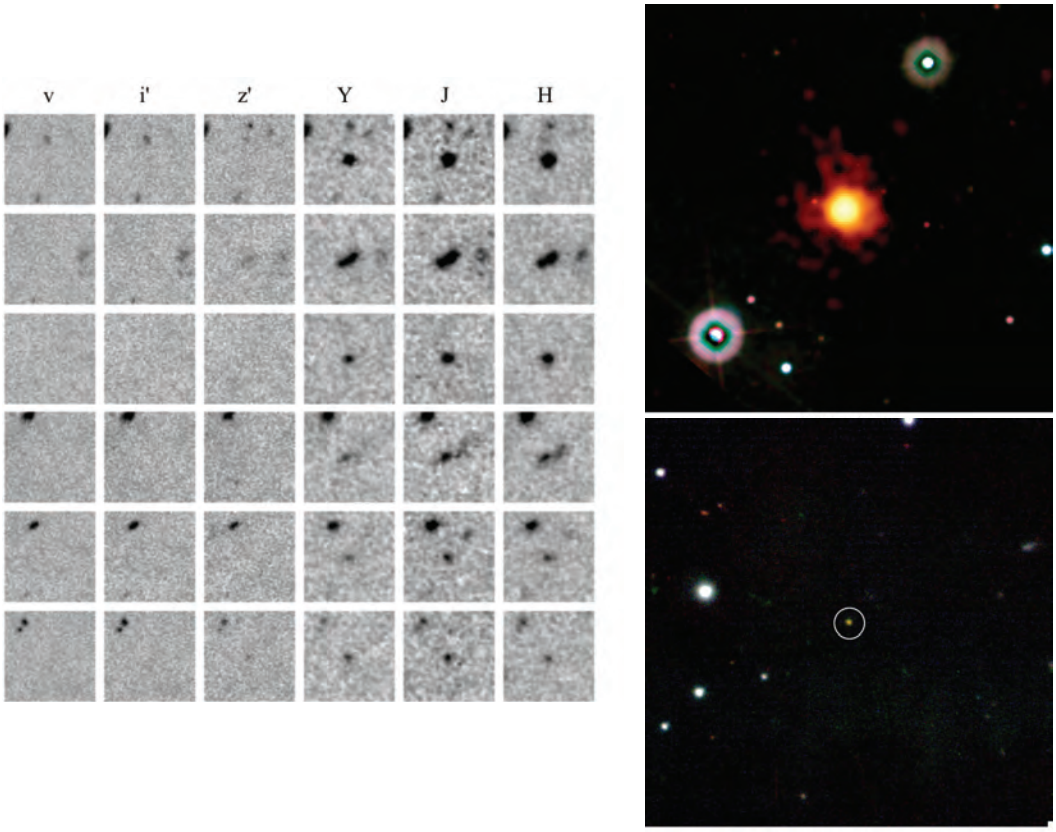


FIGURE 3.22 *Left:* Candidate redshift $z \sim 7$ galaxies identified by the "dropout" technique whereby each object is detected in the near infrared (J and H bands) but not at optical wavelengths (viz bands). The goal for the next decade is to find galaxies at even higher redshifts, confirm them spectroscopically, and study their chemical composition to identify primordial stellar populations. *Top right:* Swift ultraviolet/X-ray image (orange, red) of a gamma-ray burst at redshift $z \sim 8.2$, assuming the extremely red color in the optical/infrared image (*bottom right*) is due to the Lyman break at rest-frame 912 Å. This is the highest redshift cosmic explosion detected to date. No visible or ultraviolet light (green, blue) was detected at the position of the burst, but near-infrared emission was detected with the United Kingdom Infrared Telescope Facility and (*bottom right*) the Gemini North telescope. More such events, and more distant ones, will be visible with sensitive wide-field gamma-ray satellites and prompt follow-up observations in the near-infrared and optical. SOURCE: *Left:* R.J. Bouwens, G.D Illingworth, M. Franx, and H. Ford, z ~ 7-10 galaxies in the HUDF and GOODS fields: UV luminosity functions, *Astrophysical Journal* 686(1):230-250, 2008, reproduced by permission of the AAS. *Top right:* Courtesy of NASA/Swift/Stefan Immler. *Bottom right:* Courtesy of Gemini Observatory/NSF/AURA/D. Fox, A. Cucchiara (Pennsylvania State University), and E. Berger (Harvard University).

have a rest-frame UV magnitude of M ~ −21, giving an apparent magnitude of 27 at $z \sim 6$. More distant supernovae would be correspondingly fainter. Confirmation of the brighter events at the end of reionization should be possible with the current generation of 10-meter telescopes, but more distant sources require an ELT.

To find and characterize the first-generation aggregates of stars, the panel concluded:

- It is *most important* to use JWST to make deep surveys, followed up with near-IR spectroscopy on an ELT.
- It is *very important* to develop a next-generation GRB observatory to search for the first explosions, with an order-of-magnitude-greater GRB detection rate than is possible with Swift, augmented by a rapid follow-up capability for infrared spectroscopy of faint objects.
- It is *very important* to do time-domain surveys to identify the first stars from their supernova or hypernova explosions.

Ionization of the neutral IGM during the epoch of reionization creates bubbles that should correlate with the location of the first sources. These bubbles can be detected by means of fluctuations in the brightness temperature of the 21-cm line of H I; because the universe is nearly transparent at these frequencies, the 21-cm line should be an outstanding probe of the entire history of reionization. At high redshift, the 21-cm emission should display angular as well as frequency structure due to inhomogeneities in the gas density field, H II fraction, and H I spin temperature. This prospect has already motivated the design and construction of arrays of low-frequency radio telescopes (e.g., LOFAR, MWA) that aim to search for this signal from $z \sim 6$-15, redshifted to wavelengths of approximately 1.5 to 4 meters. Ultimately, it will be possible to map the entire history of reionization by way of an all-sky map of redshifted 21-cm emission.

Absorption-line spectroscopy along sight lines toward the first objects—be they stars, GRBs, or supernovae—will allow the detection of the presence of metals and the ionization level throughout the epoch of reionization. Such observations necessarily require the collecting area and spectroscopic capability of an ELT, which can determine redshifts, stellar masses, and chemical compositions in early galaxies, quasars, and transient events. Detection of ultralow abundances of heavy elements—the definitive indication of the first generation of stars and galaxies—should be feasible within the next decade.

It is also desirable to measure the mean fraction of neutral hydrogen with independent probes. The WMAP data were used to derive the first constraints on the redshift range of the epoch of reionization based on E-mode polarization measurements. The Planck satellite will substantially improve on these results. The Planck results might motivate additional measurements with a follow-up mission.

The panel concluded that to explore the discovery area of the epoch of reionization:

- It is *most important* to develop new capabilities to observe redshifted 21-cm H I emission, building on the legacy of current projects and increasing sensitivity and spatial resolution to characterize the topology of the gas at reionization.
- It is *very important* to do near-infrared absorption-line spectroscopy with JWST, ELTs, and 10-meter-class telescopes to probe the conditions of the IGM during the epoch of reionization.
- It is *important* to measure the CMB E-mode polarization with Planck and possibly follow-on missions.

Although this discussion has so far focused on the first objects, it is very important to find and identify objects residing in the later stages of the epoch of reionization, including radio-loud AGN, quasars, galaxies, supernovae, and GRBs. These populations connect in a directly observable way to the universe today and thus hold the key to a full understanding of its evolution. There is tremendous promise in high-frequency (>30 GHz) searches for CO in the high-redshift universe with both the EVLA and future radio facilities. Cooling by means of the C II and O I fine structure lines might be detectable if metal enrichment occurs early enough. The panel concluded:

- It is *very important* to do multiwavelength surveys to detect galaxies, quasars, and GRBs residing in the late stages of reionization at $6 < z < 8$, including near-infrared surveys for galaxies and quasars, hard X-ray or gamma-ray monitors for GRBs, and time-variability surveys for supernovae or hypernovae.
- It is *important* for ALMA to have the capability to search for C II and O I fine structure line emission at redshifts $z > 6$.

THEORY AND LABORATORY ASTROPHYSICS: THE NEXT DECADE

Underlying all of astronomy and astrophysics is critical work in theory and other intellectual infrastructure, such as laboratory astrophysics. Theory is at the heart of astronomical inference, connecting observations to underlying physics within the context of a cohesive physical model. The past decade has seen great advances in theoretical aspects of galaxy formation and black hole astrophysics, particularly in the computational arena driven by technological advances. Theory has become more interconnected with data as the fundamental physical processes are better understood and as models are placed within the context of the overall cosmological paradigm. Many areas are now driven by theoretical goals, such as testing the structure of halos predicted by simulations, observing the cold accretion predicted to be a main channel for supplying galaxies with fuel for star formation, and testing whether jets are powered by black hole spin.

To understand our universe better, to reap the full value of new observational capabilities, and to guide the next observations, investments should be made in the following theoretical areas:

- *Cosmological context.* To compare ΛCDM predictions with surveys and understand the connections among galaxies, intergalactic gas, and large-scale structure require hydrodynamical simulations within a hierarchical structure-formation context. The key challenge is to expand the dynamic range of current simulations to study detailed galaxy and cluster assembly within a representative volume, for which developing new algorithms, improving subgrid physics models, and taking advantage of new technologies will be crucial.
- *Galactic flows and feedback.* Studies of galaxy formation require understanding accretion and feedback processes, which are central to galaxy assembly. A key issue here is to properly understand two-phase interfaces and instabilities that occur during the motion of gas through ambient media of strongly differing temperatures and densities, and how energy is exchanged and released across such interfaces (e.g., mixing processes). Understanding the injection of energy, momentum, and relativistic particles into ambient gas (i.e., feedback) will require simulations of MHD turbulence, with accurate treatment of transport processes such as viscosity and heat conduction.
- *Magnetohydrodynamics and plasma physics.* Studies of accretion disks, jets, and their interactions with ambient plasma require a better understanding of how magnetic fields channel and transport energy over a large dynamic range. Solving some of the key questions will require a better understanding of particle processes such as magnetic reconnection, particle acceleration, and cosmic-ray transport. These processes will have to be incorporated into the next generation of codes in a physically realistic way.
- *Radiation processes.* Calculations incorporating radiative transfer are critical for studying the epoch of reionization, the escape of ionizing and Ly-α radiation from galaxies, and the evolution of the IGM. The challenge arises in coupling such models to dynamical galaxy-formation simulations, to understand fully how radiation interacts with the highly clumpy and asymmetric gas distribution. Models of jets and accretion disks must include nonthermal radiation processes such as synchrotron radiation and inverse Compton scattering in order to compute spectra and must include radiation hydrodynamics to capture disk structure accurately.

Addressing these questions should be done using a tiered approach, with an emphasis on computational work that takes advantage of rapidly developing technologies. Small teams led by individuals are in a good position to push the frontiers of algorithms and new computing architectures. Medium-size groups of several experts are important to make coherent and concerted progress on key numerical issues outlined above. Large, heroic simulations that push the limits of available technology are critical to drive forefront work; these are best facilitated by large computing centers. All of these areas are informed by analytic theory, which is crucial for providing the basic physical underpinnings for more complex models. Together these aspects are fundamentally important for making progress toward understanding the universe, and the galaxies, clusters, black holes, gas, and dark matter within it.

Laboratory astrophysics is clearly important to understanding galaxies, black holes, and clusters across cosmic time. For example, the details of absorption by dust are not understood, even though most cosmic objects are substantially impacted by dust. Similarly, still lacking are important cross sections for hot gas cooler than 4 million degrees, which is significant in clusters and the IGM. Scientists are not sure of recombination rates that determine ionization equilibria. Spectral features at millimeter to infrared wavelengths are especially poorly known and difficult to calculate, as they arise in molecules and clusters of atoms; ALMA will see a forest of lines that, without new laboratory measurements, will be difficult to interpret. Although the lack of laboratory measurements may not today be a limiting factor in the studies outlined earlier in this report, it may well become the limiting factor as the data improve.

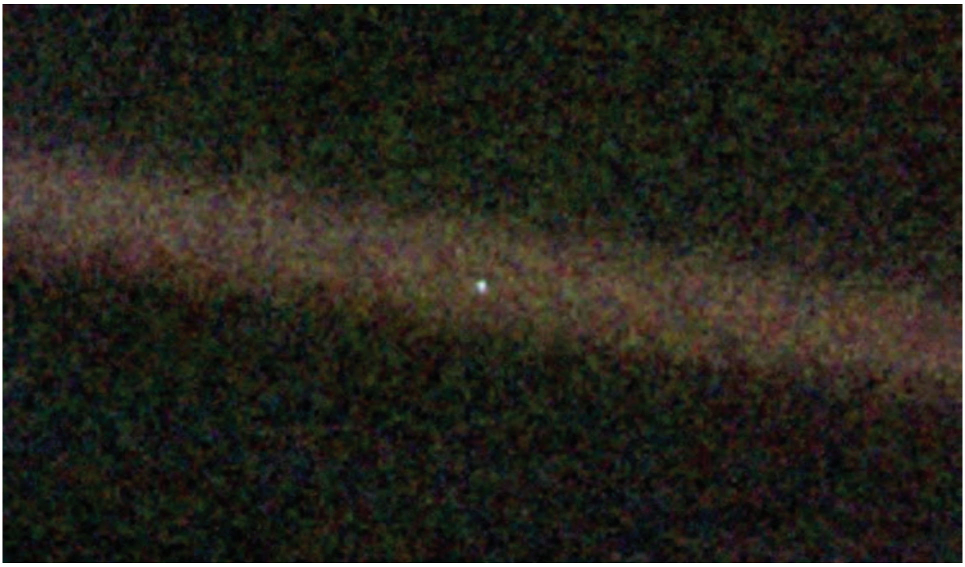
CONCLUSION

We now stand at the threshold of being able to observe the full history of cosmic structure, across all mass scales, from earliest times to the present day. With a judiciously chosen set of new facilities, instruments, and observing strategies, we can learn how galaxies, clusters, and black holes form, evolve, interact, and influence each other, to a degree of accuracy undreamed of just 20 years ago. This unprecedented wealth of data should be accompanied by a concomitant increase in understanding through the development of new analytic and computational theoretical tools. The past decade has brought dramatic advances in cosmology and an emerging picture of the growth of galaxies and black holes across the immensity of cosmic time. The next decade will open up the entire universe to detailed study and will revolutionize our understanding of the processes by which the observable cosmos came to be.

4

Report of the Panel on Planetary Systems and Star Formation

> Only one generation in the history of the human species
> is privileged to live during the time
> those great discoveries are first made; that generation is ours.
>
> —Carl Sagan



The Pale Blue Dot: Earth, as seen in 1990 from a distance of 40.6 AU, by Voyager 1.
SOURCE: NASA/JPL.

SUMMARY

There is an opportunity in the coming decade to make fundamental advances in understanding the origins of stars and planets, and to ascertain the frequency of potentially habitable worlds. These compelling scientific opportunities have far-reaching implications in areas ranging from cosmic evolution and galaxy formation to the origins of life. The paths by which star-forming clouds produce stars and planet-forming disks have become much clearer over the past decade, and a startling diversity of planets orbiting nearby stars has been discovered. We now stand on the verge of determining whether habitable worlds are common in the galaxy. Moreover, there exists the immediate possibility of identifying any such worlds circling nearby very cool stars and of characterizing their physical properties and atmospheres as the search for signs of habitation is carried out. Now is the time to take advantage of this progress to answer some of the key questions of our cosmic origins that have inspired scientists and fascinated the public.

The Astro2010 Science Frontiers Panel on Planetary Systems and Star Formation was charged to consider science opportunities in the domain of planetary systems and star formation—including the perspectives of astrochemistry and exobiology—spanning studies of molecular clouds, protoplanetary and debris disks, and extrasolar planets, and the implications for such investigations that can be gained from ground-based studies of solar system bodies other than the Sun.[1] The panel identifies four central questions that are ripe for answering and one area of unusual discovery potential, and it offers recommendations for implementing the technological advances that can speed us on our way. The questions and the area of unusual discovery potential are these:

- How do stars form?
- How do circumstellar disks evolve and form planetary systems?
- How diverse are planetary systems?
- Do habitable worlds exist around other stars, and can we identify the telltale signs of life on an exoplanet?
- *Discovery area:* Identification and characterization of nearby habitable exoplanets.

How Do Stars Form?

The process of star formation spans enormous ranges of spatial scales and mass densities. The first stage involves the formation of dense structures that con-

---

[1] The Astronomy and Astrophysics 2010 Survey of which this panel report is a part does not address solar system exploration, which is the subject of a parallel decadal survey.

stitute only a small fraction of the volume and mass of a typical molecular cloud. Knowing how these dense regions form and evolve is vital to understanding the initiation of star formation, and it has implications for galactic and cosmic evolution. Yet the mechanisms controlling these processes are not well understood. To make further progress in characterizing the internal dynamical states of molecular clouds over a wide range of spatial scales and environments, the panel recommends the following:

- Extensive dust and molecular-line emission surveys of massive giant molecular clouds spanning spatial scales from 100 to 0.1 parsec (pc) at distances greater than 5 kiloparsec (kpc), and
- Complementary studies of the young stellar populations spawned in these regions, conducted by means of infrared surveys with spatial resolution at least 0.1 arcsec to reduce source confusion in clusters, with probing sufficiently faint to detect young brown dwarfs.

In the next stage of star formation, the dense structures in molecular clouds fragment into self-gravitating "cores" that are the direct progenitors of stars. There is mounting evidence from nearby star-forming regions that the distribution of core masses may be directly related to the resulting distribution of stellar masses, although some subsequent fragmentation likely produces binaries and very low mass objects. This may occur especially during the final stage of star formation through disk accretion. To explore this evolution and to improve core-mass spectra and characterize the core properties that may lead to subsequent fragmentation into stars, the panel recommends the following:

- Deep surveys of cores down to sizes of 0.1 pc at millimeter and submillimeter wavelengths in diverse star-forming environments out to distances of several kiloparsecs, using both interferometers and large single-dish telescopes, far-infrared imaging and spectroscopy from spaceborne telescopes, and polarimetry to determine the role of magnetic fields.

An essential test of the understanding of star formation requires a definitive answer to the question of whether the initial mass function (IMF)—that is, the relative frequency with which stars of a given mass form—is independent of environment. This is a topic of great importance to an understanding of the development of galaxies and the production of heavy elements over cosmic time (see the discussion in Chapter 2, "Report of the Panel on Galaxies Across Cosmic Time") in this volume. Initial investigations of massive young clusters using the Hubble Space Telescope (HST) and other large instruments have suggested that the IMF may be "top-heavy" (with larger fractions of massive stars) in very dense regions, such as might prevail in starburst galaxies. To explore IMFs in more extreme environments, such as dense galactic regions and the nearest low-metallicity systems (the Magellanic Clouds), the panel recommends the following:

- Near-infrared surveys with less than 0.1-arcsec resolution to limit source confusion in the galaxy and 0.01-arcsec resolution for the Magellanic Clouds.

Major theoretical efforts will be necessary to develop a fundamental understanding of these new observations, including improved treatment of thermal physics for an understanding of fragmentation and the origin of the IMF, along with better models for the chemical evolution of collapsing protostellar cores. More realistic calculations of the effects of massive stars on their environments (most dramatically in supernova explosions) are also needed to contribute to an understanding of how this feedback limits star-formation efficiencies. To facilitate these advances in the theoretical understanding of star formation and to enable the interpretation of complex data sets, the panel recommends the following:

- The development of improved algorithms, greater computational resources, and investments in laboratory astrophysics for the study of the evolution of dynamics, chemistry, and radiation simultaneously in time-dependent models of star-forming regions.

How Do Circumstellar Disks Evolve and Form Planetary Systems?

Circumstellar disks are the outcome of the collapse of rotating protostellar cores. Both central stars and planets are assembled from disks. Major advances were made over the past decade in characterizing evolutionary timescales of protoplanetary disks, but their masses and structure are much less certain. In the coming decade, improved angular resolution will routinely yield resolved images of disks, providing keys to their mass, physical and chemical structure, and mass and angular momentum transport mechanisms, crucial to the understanding of both star and planet formation.

The superb new high-resolution, high-contrast imaging capabilities of the Atacama Large Millimeter Array (ALMA), the James Webb Space Telescope (JWST), and large optical/infrared ground-based telescopes with adaptive optics (AO) will revolutionize the present understanding of disks. Resolved submillimeter-wavelength measurements of dust emission will help constrain dust opacities and improve the understanding of disk masses and mass distributions. The direct detection of spiral density waves resulting from gravitational instabilities would enable independent estimates of disk masses and establish their role in mass and angular momentum transport. Spiral waves and gaps can also be produced by forming planets; the latter may be directly detected within these gaps owing to their high luminosities during formation. Detections of forming planets would enable monumental advances in the understanding of planet formation. To achieve these goals, the panel recommends the following:

- Studies of protoplanetary disks in nearby star-forming regions at resolutions below 100 milliarcsec, with every effort to achieve 10-milliarcsec resolution, at millimeter, submillimeter, infrared, and optical wavelengths, in order to map disk structure on spatial scales of approximately 1-10 AU;
- Searches for infant planets in disk gaps using JWST, extreme-AO near-infrared (near-IR) imaging on 8- to 10-m-class telescopes, and eventually extreme-AO imaging with 30-m-class telescopes.

Improved imaging will also revolutionize the understanding of later-stage debris disks, illuminating planetary system architectures through the detection of structure in the debris formed by the collisions of numerous solid bodies undergoing dynamical evolution. To exploit these possibilities, the panel recommends the following:

- Imaging debris disks in optical and near-IR scattered light on 8-m-class telescopes and in thermal dust emission at submillimeter wavelengths with ALMA and other interferometric arrays in order to search for resonant structures, gaps, and other features caused by the gravitational perturbations produced by planets, allowing the inference of unseen bodies and constraining their masses.

Similar dynamical instabilities also occurred early in the evolution of our own solar system, as indicated by resonant structures in the Kuiper belt. To improve vastly the understanding of the evolution of our solar system as well as to provide an essential link to the understanding of extrasolar debris disk systems, the panel recommends the following:

- Systematic, whole-sky, synoptic studies to *R* magnitude ≥24 of Kuiper belt objects (KBOs).

The physics and chemistry of disks, particularly those in the protoplanetary phase, are extremely complex. To make progress in understanding these topics, the panel recommends the following:

- Expanded theoretical efforts and simulations, with a detailed treatment of observational tracers to test theories, to develop an understanding of mass transport within disks and of the processes of coagulation and accretion that lead to planet formation; and

- Major new efforts in chemical modeling and laboratory astrophysics to contribute to the understanding of the chemistry underlying molecule formation in the wide-ranging conditions in disks. In particular, laboratory studies of molecular spectra in the poorly studied far-infrared and submillimeter-wavelength regions of the spectrum are urgently needed to allow understanding and interpretation of the vast new array of spectral lines that are being detected by the Herschel mission and will be found by ALMA.

How Diverse Are Planetary Systems?

The past decade has seen a dramatic increase in the knowledge of the population and properties of planets orbiting nearby stars. Many more than 300 such exoplanets are now known, along with direct estimates of the densities and atmospheric temperatures for several dozen of these worlds. What has been learned from these exoplanets—mostly gas and ice giants—makes it clear that planetary systems are far from uniform. Yet these results apply just to the 14 percent of stars with close-in giant planets detectable with current techniques. The actual frequency of planetary systems in the galaxy and the full extent of their diversity, especially for small, rocky worlds similar to Earth, await discovery in the coming decade.

The recently commissioned Kepler mission is expected to yield the first estimate for the population of terrestrial exoplanets. However, the scientific return will be fully realized only if mass estimates can be obtained for a significant number of such planets. Therefore, the panel recommends the following:

- Both a substantial expansion of the telescope time available to pursue radial-velocity work, and the development of advanced radial-velocity techniques with a target precision sufficient to detect an Earth-mass planet orbiting a Sun-like star at a distance of 1 AU.

This investment in radial-velocity precision will also augment the understanding of more massive worlds located at distances of 1-10 AU from their stars, which is the region of giant planets in our own solar system. Another promising approach is the detection of microlensing, which does not require that data be gathered over a full orbital cycle and can thus relatively rapidly provide detailed statistics on the masses and orbital separations of planets in the outer as well as inner reaches of planetary systems.

Thus the findings from Kepler combined with the results of a space-based microlensing survey will provide the essential statistics to test astronomers' grand picture of how planetary systems form and whether the solar system is a commonplace occurrence or a cosmic rarity. Although the fundamental basis for understanding exoplanet diversity rests on measuring orbits and masses, and radii when possible, the chemistries, structures, and dynamics of exoplanet atmospheres can be explored with spectra. Therefore, the panel recommends the following:

- Extension of the eclipsing techniques currently employed with HST and the Spitzer Space Telescope to JWST, and
- Extreme-contrast-ratio imaging with both the extant ground-based observatories and the next generation of giant segmented-mirror telescopes (GSMTs) in order to image planets with dynamical mass estimates and to calibrate models predicting emission from planets as a function of mass and age.

Do Habitable Worlds Exist Around Other Stars, and Can We Identify the Telltale Signs of Life on an Exoplanet?

One of the deepest and most abiding questions of humanity is whether there exist inhabited worlds other than Earth. Discovering whether or not such a planet exists within the reach of Earth's astronomical observatories will have ramifications that surpass simple astronomical inquiry to impact the foundations of many scholarly disciplines and irrevocably to alter our essential picture of Earth and humanity's place in the universe.

The goal of detecting life on other worlds poses daunting technological challenges. A Sun-like star would be 100 times larger, 300,000 times more massive, and 10 million to 10 billion times brighter than a terrestrial planet with an atmosphere worthy of studying. Although several techniques have been proposed to achieve detection of biomarkers, it is currently premature to decide the technique and scope of such a mission. Rather, the panel endorses the finding of the National Science Foundation-National Aeronautics and Space Administration-U.S. Department of Energy (NSF-NASA-DOE) Astronomy and Astrophysics Advisory Committee (AAAC) Exoplanet Task Force[2] that two key questions that will ultimately drive the technical design must first be addressed:

1. What is the rate of occurrence of Earth-like planets in the habitable zones of Sun-like stars, and hence at what distance will the target sample lie?
2. What is the typical brightness of the analogs of the zodiacal light disks surrounding solar analogs; in particular, do a significant fraction of stars have dust disks that are so bright as to preclude the study of faint Earth-like planets?

Kepler will address the first question, but the means to answer the second, perhaps through ground-based interferometry or space-based coronagraphy, has

---

[2] The full report, ExoPlanet Task Force, *Worlds Beyond: A Strategy for the Detection and Characterization of Exoplanets,* Washington, D.C., May 22, 2008, is available at http://www.nsf.gov/mps/ast/aaac/exoplanet_task_force/reports/exoptf_final_report.pdf.

yet to be fully developed. Provided that Earth analogs are sufficiently common, the panel recommends the following as the preferred means to identify targets with appropriate masses:

- A space-based astrometric survey of the closest 100 Sun-like stars with a precision sufficient to detect terrestrial planets in the habitable zones.

The characterization effort lies beyond the coming decade, but it could be achieved in the decade following, provided that the frequency of Earth analogs is not too low. The panel recommends the following:

- A strong program to develop the requisite technologies needed for characterization should be maintained over the coming decade.

Discovery Area: Identification and Characterization of Nearby Habitable Exoplanets

An exciting possibility in the coming decade is the detection of possibly habitable, large, rocky planets (super-Earths) orbiting the abundant and nearby stars that are much less massive than the Sun (less than 0.3 solar masses). *The panel deems this to be the single greatest area for unusual discovery potential in the coming decade, as it can be carried out with current methods provided that the necessary resources are made available.*

The low luminosities of these cool, low-mass stars in the solar neighborhood ensure that the conditions for liquid water to exist on the surface of an orbiting planet occur at a small separation of planet and star. The small stellar size, low stellar mass, and small orbital separation for habitable conditions all conspire to facilitate the discovery of super-Earths by a combination of the two detection methods that have proven the most successful to date: stellar radial velocities and timing of obscuration due to planetary transits of host stars. These techniques, currently refined for the study of Sun-like stars, need to be adapted for cooler, low-mass stars. Therefore, the panel recommends the following:

- Increasing the amount of observing time available for radial-velocity studies,
- Investing in precision radial-velocity techniques at longer wavelengths, and
- Developing novel methods to calibrate the new, longer-wavelength spectrographs.

A far-reaching outcome of this investment is that the atmospheres of transiting super-Earths would be amenable to spectroscopic study with JWST and a future

GSMT, permitting a search for biomarkers in the coming decade. Thus, the panel recommends the following:

- The closest 10,000 M-dwarfs should be surveyed for transiting super-Earths in their stellar habitable zones in time to ensure that the discoveries are in hand for JWST.

The discovery of even a handful of such worlds would present an enormous scientific return, fundamentally alter our perspective on life in the universe, and offer a hint of what might be expected for the properties of terrestrial worlds around Sun-like stars.

Summary of Requirements

The conclusions of this panel report are summarized in Table 4.1.

TABLE 4.1 Summary of Conclusions of the Panel on Planetary Systems and Star Formation

	Question 1: How Do Stars Form?	Question 2: How Do Disks Evolve and Form Planetary Systems?	Question 3: How Diverse Are Planetary Systems?	Question 4: Can We Identify the Telltale Signs of Life on an Exoplanet?	Discovery Area: Identification and Characterization of Nearby Habitable Exoplanets
Facilities expected	EVLA, ALMA, Herschel, SOFIA, JWST, 8- to 10-m telescope with AO	EVLA, ALMA, Herschel, JWST, 8- to 10-m telescope with AO, UV/visible synoptic surveys	1 m sec^{-1} RV surveys and transit follow-up; Kepler, JWST, Spitzer transits; Gaia astrometry	1 m sec^{-1} RV surveys and transit follow-up; Kepler, JWST, Spitzer transits	JWST transiting-exoplanet spectroscopy
New facilities needed	30-m submillimeter telescope; 8- to 10-m telescope with MCAO; GSMT with extreme AO; deep centimeter-wave interferometry on very long baselines	GSMT with extreme AO; near-IR synoptic surveys	0.2 m sec^{-1} RV; microlensing surveys; GSMT with extreme AO	Earth-like planet frequency ($\eta_\oplus$); 10-zody limits on exozodies; 0.1-μas astrometry	Census and transit survey, 10^4 nearest M-dwarfs; visible/near-IR RV follow-up
Always needed	Support for theoretical efforts, including high-performance computational resources, and laboratory-molecular astrophysics with an emphasis on far-infrared, submillimeter, and millimeter line identifications, along with chemical studies ranging from surface reactions relevant to cold clouds to processes in planetary atmospheres.				

NOTE: Acronyms are defined in Appendix C.

INTRODUCTION

Human beings exist, in part, because we live on a rocky planet that has an atmosphere and water and is warmed appropriately by a long-lived star. As knowledge of the universe has expanded over the centuries, so has speculation about the possibility of life elsewhere. Remarkably, we now stand on the threshold of developing the technology needed to detect other habitable planets, and to determine how common they are in the galaxy. Primed by the discovery over the past 15 years of significant numbers of other planetary systems, by the realization that planet-forming disks result as a natural and frequent by-product of star formation, and by major advances in characterizing the properties of the gas clouds that form stars, scientists are now ready to make fundamental progress on the central questions related to the birth of stars and planets. Following its charge, the Panel on Planetary Systems and Star Formation identified four questions that it considers as ripe for answering in the coming decade, as well as one discovery area:

- How do stars form?
- How do circumstellar disks evolve and form planetary systems?
- How diverse are planetary systems?
- Do habitable worlds exist around other stars, and can we identify the telltale signs of life on an exoplanet?
- *Discovery area:* Identification and characterization of nearby habitable exoplanets.

In the sections that follow, the panel explores each of these questions and identifies key observational and theoretical advances (set apart with bullets and summarized in tables at the end of each major section) that are necessary to make the associated, fundamental advances.

PSF 1. HOW DO STARS FORM?

Star formation plays a crucial role in many important astrophysical processes, ranging from galactic evolution to planet formation. Major investments over the past decade in both observational and computational facilities have brought astronomers to the verge of developing a quantitative understanding of how, where, and when stars form; why planet-forming disks result from protostellar cloud collapse; the ways in which the formation of massive stars differ from that of solar-type stars; and how the energy input from massive stars drives the evolution and destruction of star-forming clouds. New facilities enabling both highly detailed studies of individual, nearby objects and large-scale surveys of diverse star-forming regions will allow scientists to address three key aspects of the star formation process: What sets the overall rate and efficiency of star formation? What determines the properties of star-forming cloud cores? And finally, is the initial distribution of stellar masses universal or a function of environment?

What Determines Star-Formation Rates and Efficiencies in Molecular Clouds?

The rates at which stars form over the age of the universe strongly affect galactic structure and cosmic evolution (see also, the second key science question, GCT 2, in Chapter 3, "Report of the Panel on Galaxies Across Cosmic Time"). Increasingly detailed studies of external galaxies have led to improved Kennicutt-Schmidt "laws" relating large-scale gas content and other global galactic properties to star-formation rates. To develop a comprehensive theory of how star formation depends on environment, large-scale extragalactic studies need to be complemented by investigations on the much smaller spatial scales on which clouds actually fragment into clusters and stars (Figure 4.1). Here the panel focuses on the opportunities

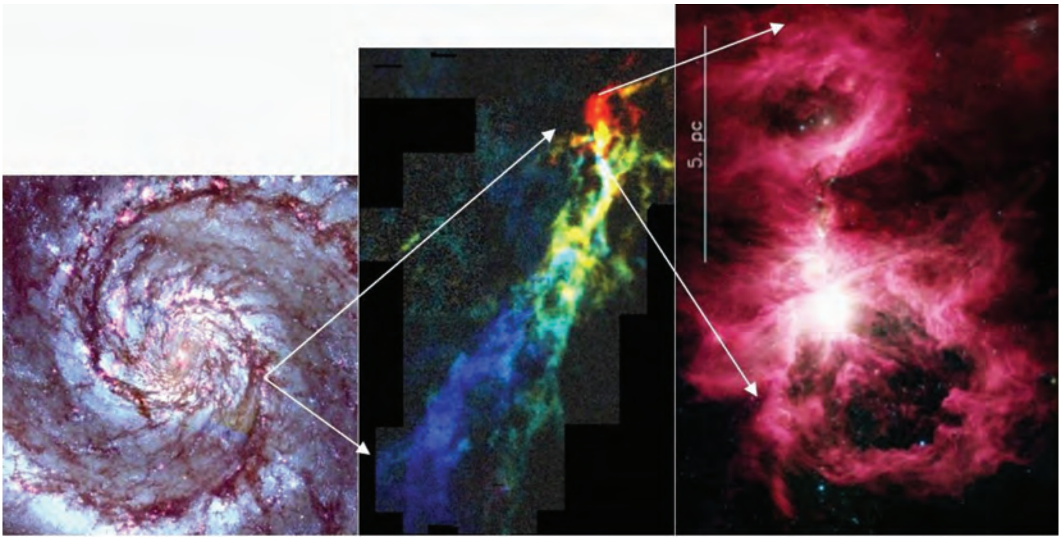


FIGURE 4.1 Schematic of the hierarchy of star formation. *Left:* Hubble Space Telescope image of the spiral galaxy M51, with Hα emission (in red) tracing the massive star-forming regions with a pixel scale equivalent to 5 pc. *Center:* Nearby Orion A molecular cloud traced by ^{13}CO J = 1 $\rightarrow$ 0 emission, where the colors represent radial Doppler velocity. *Right:* Orion Nebula Cluster as seen by the IRAC infrared camera on the *Spitzer* Space Telescope. SOURCE: *Left:* N.Z. Scoville, M. Polletta, S. Ewald, S.R. Stolovy, R. Thompson, and M. Rieke, High-mass, OB star formation in M51: Hubble Space Telescope Hα and Paα imaging, *Astronomical Journal* 122(6):3017-3045, 2001, reproduced by permission of the AAS. *Center:* J. Bally, Overview of the Orion Complex, in *Handbook of Star Forming Regions, Vol. I.* (B. Reipurth, ed.), Astronomical Society of the Pacific, San Francisco, Calif., 2008, reproduced by kind permission of the Astronomical Society of the Pacific. *Right:* NASA/JPL-Caltech/University of Toledo.

over the next decade to advance the understanding of star-formation efficiencies and rates from the study of local (galactic) star-forming regions.

Local molecular clouds convert only a few percent of their mass into stars before they are dispersed, probably due to stellar energy input; efficiencies may be still lower in the massive giant molecular clouds (GMCs) that contain most of the galaxy's molecular mass. These low star-formation efficiencies are clearly related to the complex structures of molecular clouds, in which only a small fraction of the mass resides in high-density regions prone to gravitational collapse. Understanding how these dense regions arise within much larger volumes of low-density, magnetized, and supersonically "turbulent" gas is an essential first step in star-formation theory, and poses major challenges. Numerical simulations of molecular clouds over the past decade have made great strides, yielding a fraction of gravitationally collapsing dense gas in rough agreement with observations. However, the nature and origin of the supersonic turbulence that produces dense structures are uncertain, and the final efficiency of star formation depends in part on cloud dispersal, which has only recently begun to be addressed numerically.

A quantitative understanding of how dense star-forming cloud structures are produced will rest on the observational characterization of the physical and dynamical states of molecular clouds, from the typical ~100 pc sizes of GMCs to the ≤0.1 pc scales of the densest structures. The panel recommends the following:

- Surveys to enable a combination of wide- and narrow-field studies of molecular clouds in millimeter and submillimeter dust continuum emission and a range of molecular tracers from low to high density using current and new telescopes under development (e.g., ALMA, Herschel, EVLA), along with the implementation of large-format heterodyne and bolometer arrays on single-dish telescopes and interferometers.
- Studies beyond the solar neighborhood out to ~8 kpc, which would allow access to the Milky Way's "molecular ring" and the galactic center, to measure star-formation efficiencies and rates in massive ($\geq 10^5$ $M_\odot$) GMCs, which are the sites of most star formation. A spatial scale of 0.1 pc, which would resolve the densest gas, corresponds to 3 arcsec at the galactic center; surveys need to be complete for cores of mass ≥1 $M_\odot$ to probe near the peak of the core and stellar mass functions (see the next two subsections below).
- Polarimetric imaging from millimeter through infrared wavelengths and measurements of the Zeeman effect using key molecules at radio wavelengths to characterize the dynamical role of magnetic fields and relate these measures of gas properties to star-formation rates and efficiencies.
- Infrared surveys with high spatial resolution, ~0.1 arcsec to minimize source confusion in clusters (assuming conditions similar to the Trapezium region in the

Orion Nebula Cluster), and probing down to K = 19 mag to detect young brown dwarfs to characterize the associated stellar populations.

Deriving stellar masses requires comparison to pre-main sequence models and corrections for differential extinction and binarity. Understanding in this area will be greatly advanced by the European Space Agency's Gaia mission, which will provide accurate distances to vast numbers of stars. In addition, the proper motion measurements from Gaia will provide new insight into the dynamical states of star-forming regions. The Very Long Baseline Array (VLBA) will also contribute by measuring parallaxes for those young stars having strong nonthermal emission, and will be able to map out the locations of star-forming regions throughout the galaxy using water and methanol masers.

Energy input from young stars—ionizing radiation, winds, radiation pressure, and supernova explosions—generates dynamically important turbulence, and clearly limits star-formation efficiencies and rates by disrupting and dispersing molecular clouds. To understand how stellar energy input or feedback inhibits star formation in some regions and triggers it in others, the panel recommends the following:

- Detailed investigations of GMC complexes—including low-frequency radio surveys for supernova remnants and high-resolution imaging of embedded protostars to measure their ionized winds, H II regions, infrared luminosities, and molecular outflows—to separate the various processes. Studies characterizing the effects of feedback should be able to resolve ultracompact H II regions (typically 0.1 pc) and CO outflows (down to a few times 0.01 pc) in a variety of different environments including the galactic center (8 kpc away), requiring sub-arcsecond imaging at all wavelengths. These measurements are therefore best achieved using interferometers at radio through to submillimeter wavelengths (e.g., ALMA, CARMA, EVLA, future centimeter-wavelength instruments), and near- and mid-infrared imaging and spectroscopy using ground-based telescopes up to 30-m class and JWST.

Improved theoretical tools will be needed to interpret these observations and to develop quantitative theories of the origin of cloud structure leading to star formation. As enormous dynamic ranges in size and density are involved, the panel recommends the following:

- Investment in improved algorithms and major computational resources, as well as insight into which subproblems can be tackled separately. Major improvements are required in treating radiative transfer and thermal physics in simulations (critical to fragmentation), along with better chemical models to facilitate comparison with observations.

Incorporating more complex physics, including magnetic fields and stellar energy input (ionizing radiation, winds, radiation pressure, supernova explosions) is essential to the development of a deep understanding of how pre-stellar cloud cores are assembled.

What Determines the Properties of Pre-Stellar Cloud Cores?

Stars form from dense cores within molecular clouds. Cores represent the mass reservoirs from which stars form, and their angular momentum content is responsible for the formation of protoplanetary disks and probably fragmentation into multiple star systems. Developing a quantitative theory of star formation requires an understanding of how cores form, which in turn necessitates the characterization of individual core properties and the dependence of these properties on environment.

Over the past decade, surveys with millimeter and submillimeter bolometer-array cameras and infrared extinction mapping have identified large enough samples of cores to begin investigating the distribution of core masses. Encouragingly, several different studies have found core mass functions (CMFs) that appear to be similar to the stellar IMF for stars at moderate masses (~1-10 $M_\odot$; Figure 4.2). These findings suggest that, at least in some regimes, there may be a relatively direct mapping from the CMF to the IMF. However, there are significant systematic uncertainties in determining core masses, not least of which is that most cores do not have well-defined boundaries. Fragmentation into multiple systems during the evolution from cores to stars may be especially important in defining the transition from the CMF to the IMF at low masses, while the accretion of intracluster gas may be particularly important in making the highest-mass stars.

Better characterization of the properties of cores and their immediate environments will require measurements of both dust emission and molecular line tracers to disentangle local chemical and radiative transfer effects. The accuracy of core mass estimates from continuum measurements is dependent on the dust emissivity and temperature; multiwavelength ($\lambda \sim 0.4$-1.1 mm) continuum observations are needed to provide better estimates of dust emissivities and temperatures in cores that will yield more accurate mass estimates. Kinematic studies, with molecular line observations, of both cores and their immediate environments are especially important, not only to probe the dynamic states of cores and their angular momentum content, but to eliminate "false" cores resulting from line-of-sight projections, distinguish multiple superposed cores, and measure continued accretion of

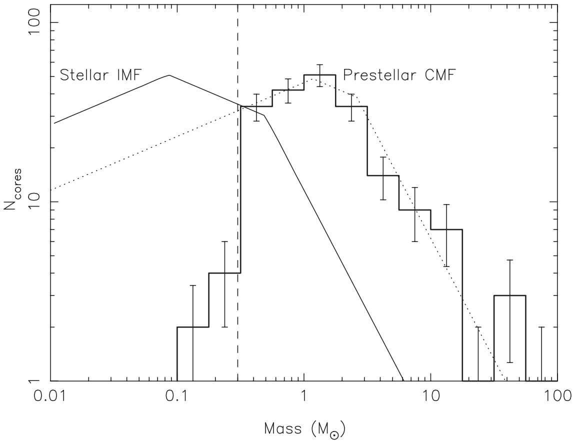


FIGURE 4.2 The core mass function (CMF) of the Orion B molecular cloud, derived from dust continuum emission. As has been found in other studies, the CMF for $M > 1\ M_\odot$ (dotted line) is similar in *slope* to the high-mass slope of the stellar IMF (solid line), but the possible turnover at lower masses is uncertain, as completeness limits become important (vertical dashed line). SOURCE: Reprinted with permission from D. Nutter and D. Ward-Thompson, A SCUBA survey of Orion—The low-mass end of the core mass function, *Monthly Notices of the Royal Astronomical Society* 374:1413-1420, 2007, copyright 2007 Royal Astronomical Society.

matter onto cores from their surroundings. To accomplish these studies the panel recommends the following:

- Millimeter- and submillimeter-wavelength interferometers such as ALMA, observations with large (>15 m) single-dish telescopes at millimeter and submillimeter wavelengths, far-infrared (far-IR) imaging and spectroscopy from spaceborne telescopes, and polarimetry to determine the role of magnetic fields; and
- High-resolution imaging of cores in far-IR/submillimeter continuum and in high-density molecular line tracers is needed to explore how nonspherical, asymmetric core geometries may produce further fragmentation into multiple star systems. These studies should be complete for cores of ~0.01 $M_\odot$ and extend to regions forming the full range of stellar masses, to encompass those cores with the potential for forming massive planets.

As is the case for star-formation efficiencies, it is crucially important to extend the current set of CMF determinations, which are derived mostly from the nearest, low-density star-forming regions, to the more distant sites where massive stars and star clusters are being formed. These regions are responsible for most star formation; their study is needed to develop an understanding of whether the CMF (and ultimately the IMF; see the subsection below titled "What Is the Origin of the Stellar Mass Function?") varies with environment. Initial steps toward this goal have been taken, based on the detection in mid-infrared absorption of massive filaments in distant (few kiloparsec) clouds—the infrared dark clouds (IRDCs)—using the Spitzer Space Telescope (Figure 4.3). Only a relatively small number of IRDCs have been studied in any detail so far, but further studies may lead to an understanding

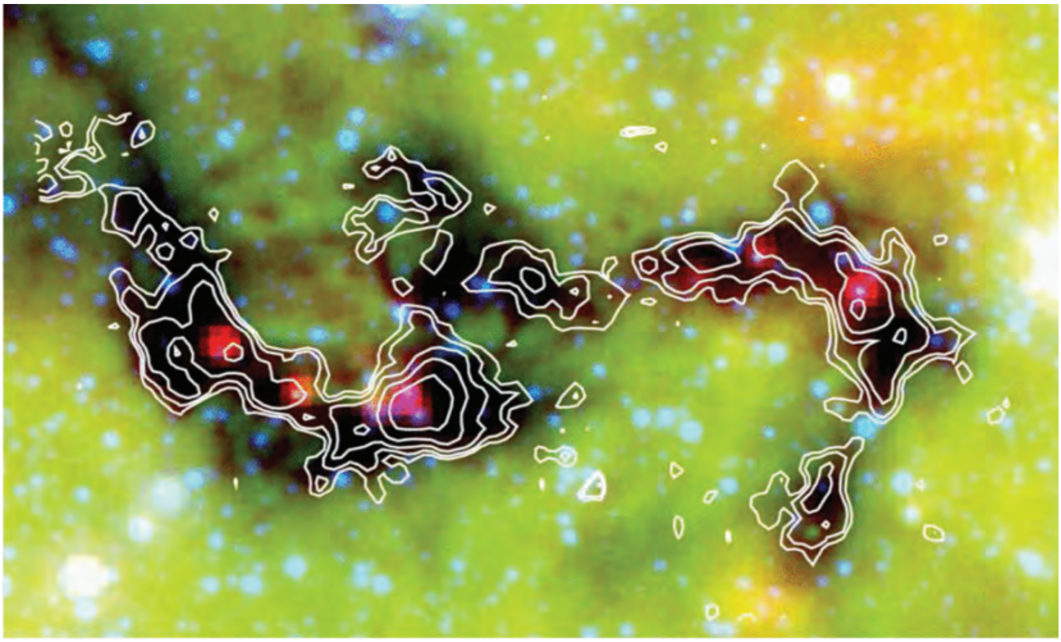


FIGURE 4.3 Three-color (RGB = 24, 8, 3.6 microns) Spitzer image of an Infrared Dark Cloud (IRDC) with white CARMA N_2H^+ (J = 1-0) contours. Such IRDCs are likely the sites of future massive star and cluster formation. SOURCE: C. Brogan, T. Hunter, R. Indebetouw, and A. Wootten, in preparation. Courtesy of C. Brogan, personal communication.

of how molecular clouds fragment to produce stellar clusters and how the dense gas fraction (see the subsection above titled "What Determines Star Formation Rates and Efficiencies in Molecular Clouds?") relates to the properties of the surrounding molecular material on scales up to about 100 pc.

For studies of IRDCs and similar regions the panel recommends the following:

- A combination of interferometers and array cameras on single-dish millimeter and submillimeter telescopes in order to provide the full dynamic range of spatial scales needed to minimize biases introduced by observing technique.
- Extending core detections to sufficiently low masses that the turnover in the CMF—and not just a high-mass power-law "tail"—can be detected definitively. Surveys will need to be complete to below 1 $M_\odot$ for the most distant regions and to ~0.03 $M_\odot$ in nearby regions to achieve this goal.

In the nearest IRDCs at distances of 2 kpc, a core of spatial scale of 0.01 pc corresponds to a resolution of 1 arcsec. For galactic-center star-forming regions at distance of 8 kpc, the corresponding resolution is 0.3 arcsec.

Numerical studies over the past decade have demonstrated that turbulence within molecular clouds is likely to be responsible for creating many of the observed (transient) structures at moderate to high densities, with gravity further compressing these structures into protostellar condensations. However, realistic multi-physics models of cloud formation and internal evolution (including feedback) are only now becoming possible. CMFs seen in numerical simulations—including turbulence, self-gravity, and diffusion of partially ionized gas through magnetic fields—increasingly resemble observations, but much work is still needed in order to develop a predictive theory connecting large-scale cloud properties (which vary widely) and the CMF. As discussed in the subsection above on star-formation rates, the panel recommends the following:

- Improvements in the radiation transfer and thermal physics of simulations to contribute to the understanding of gravitational fragmentation and consequent core formation; and
- Better models of chemical evolution to enable detailed tests of theories by comparison with observations.

Almost all far-infrared, submillimeter, and radio observations rely on the use of molecular lines as probes of molecular gas evolution. The interpretation of these data therefore requires knowledge of the relevant molecular spectra in the laboratory and is particularly important for understanding warm regions of the interstellar medium where the spectral density in the millimeter-wave and submillimeter-wave regions is high, and precise wavelengths are required. Comparisons of line observations and chemical simulations containing both gas-phase and grain-surface chemistry have been used to constrain estimates of the lifetimes and physical conditions of evolutionary stages leading to the formation of low-mass stars (e.g., pre-stellar cores) and, to a lesser extent, high-mass stars (e.g., hot cores). Therefore the panel recommends expanded laboratory work in two areas:

- Spectroscopy in the far-infrared, a poorly studied region, which will be opened up by Herschel and SOFIA; and
- Studies to improve knowledge of the relevant chemical processes on small grains, especially the surface chemistry, which also requires a detailed understanding of stochastic effects. Chemical simulations will have to be coupled more strongly to heterogeneous and dynamically evolving cloud models.

What Is the Origin of the Stellar Mass Function?

The distribution of masses with which stars form—the IMF—is an essential ingredient in studies of galaxy formation and the evolution of heavy-element abundances; the proportion of high-mass stars to the total production of stellar mass controls the cycles of mass and energy, as well as chemical cycles within galaxies. In the preceding subsection the panel discussed the starting point for understanding the origin of the IMF: specifically, the properties of star-forming cores and how they are produced. Here the transformation from CMF to IMF and the dependence of the IMF on the star-forming environment are addressed.

Although the similarity in shape of the CMF to the stellar IMF hints at a direct mapping between the two, dispersal and subsequent core fragmentation into binary and multiple stars must occur; in particular, formation of the lowest-mass objects may require fragmentation in disks. To probe the transformation from the CMF to the IMF will require the following:

- High-angular-resolution studies of molecular line emission in collapsing cores with ALMA to resolve the transition zones between the infalling envelope of cores and their protostellar, possibly circumbinary, disks on scales of ~100 AU, where further gravitational fragmentation may occur;
- Further laboratory studies of molecular spectra in the far-infrared and submillimeter regions to contribute to the understanding and interpretation of the vast new array of spectral line observations that will form the basis for sophisticated physical and chemical models; and
- Improved numerical simulations, including magnetic fields and feedback.

In the nearest star-forming clouds, near-IR observations have reached sufficient angular resolution to distinguish individual stars in crowded regions and to start to resolve wide binaries. As new facilities such as JWST and adaptive optics systems for ground-based 8- to 10-m telescopes become available, such studies can be extended to more distant regions, thereby probing a wider variety of environments. Typical stellar separations in a region with crowding similar to the Trapezium region in the Orion Nebula Cluster are ~0.1 arcsec if viewed at a distance of 8 kpc; a wide binary with separation of 100 AU subtends ~10 mas at this distance. The greater resolving power of GSMT plus AO would make significant improvements in the knowledge of the embedded stellar content of cores out to the distances of massive galactic star-forming regions, although the resolution of protostellar multiplicity will likely remain confined to relatively nearby clouds. Gaia will complement this effort by providing accurate astrometry for the detection of binary companions.

Estimates of star-formation rates in high- and intermediate-redshift galaxies typically assume a specified, usually invariant, form of the IMF. It is crucial to determine whether the IMF is in fact universal or whether it is top-heavy in the densest regions, as suggested by studies of the galactic center and some galactic massive clusters (Figure 4.4). To eliminate major uncertainties introduced by this

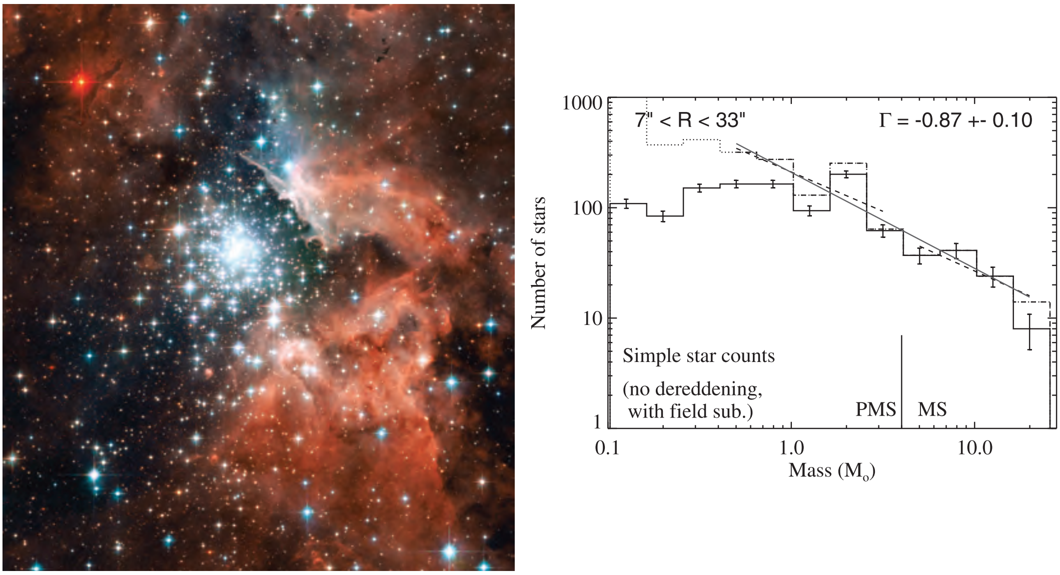


FIGURE 4.4 *Left:* Hubble Space Telescope image of the dense star cluster NGC 3603, at a distance of ~6 kpc. *Right:* the mass function derived for NGC 3603 is slightly flatter than that derived for the local stellar population. SOURCE: *Left:* NASA, ESA, and the Hubble Heritage (STScI/AURA)-ESA/Hubble Collaboration. Acknowledgment: Jesús Maíz Apellániz (Instituto de Astrofísica de Andalucía, Spain) and Davide de Martin (skyfactory.org). *Right:* A. Stolte, W. Brandner, B. Brandl, and H. Zinnecker, The secrets of the nearest starburst cluster. II. The present-day mass function in NGC 3603, *Astronomical Journal* 132:253-270, 2006, reproduced by permission of the AAS.

assumption, much larger and more sensitive surveys of the IMF in a wider variety of environments are needed, probing down to substellar masses. In particular, the panel recommends the following:

- Measurements of the IMF in the lower-metallicity environment of the Large Magellanic Cloud. At ~50 kpc, approximately 10 times better angular resolution is needed compared with what is required for the galactic center, and sensitivities down to K ~ 23 mag are needed to probe a similar range of the IMF.

It is challenging to obtain an accurate local IMF across the full mass range. High-mass stars are rare except in massive star-forming regions, but at the distance of most such regions, lower-mass stars are difficult to resolve. In older clusters, dynamical effects on the population, such as mass segregation and ejection, are important. Thus investigations must be done in young—often, heavily extinguished—regions. Furthermore, while the IMF of massive stars on radiative tracks may be determined adequately from photometric infrared luminosity functions, accurate mass distributions for the low-mass stars and brown dwarfs require theoretical and observational calibrations of mass-luminosity relations as a function of age, which also requires good distance determinations using trigonometric parallaxes and spectroscopic identification.

To obtain improved determination of IMFs down to substellar masses over a range of environments, the panel recommends the following:

- High-resolution infrared imaging in the galaxy on 8- to 10-m-class telescopes with wide-field AO (Figure 4.5) and near-IR integral-field spectrographs (IFUs);

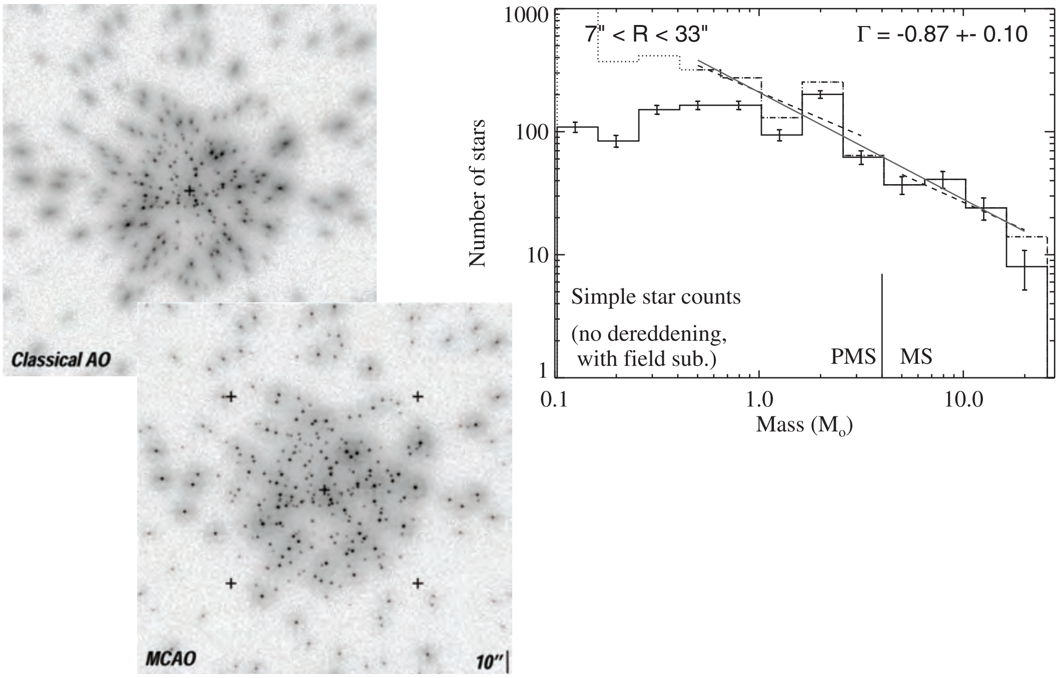


FIGURE 4.5 Wide fields, high spatial resolution, and high image contrast are needed to cover large areas while separating clusters into individual stars. Multiconjugate adaptive optics can potentially achieve these goals. *Left:* Simulated observation of an open stellar cluster at *H* (1.65 μm) with a Gemini 8-m telescope, using classical (single-phase-reference) adaptive optics and the proposed Gemini MCAO instrument. The image is well corrected over the central 30 arcsec of the classical AO image, and over the entire 2.75-arcmin field in the MCAO image. *Right:* Slice through the center of a typical point-spread function (PSF) at *H*, simulated for Gemini MCAO. The PSF diameter is 0.05 arcsec (FWHM). Also plotted are the relative brightnesses of stars with that of an O5 star scaled to the peak of the PSF, indicating the proximity to O stars at which fainter stars can be detected. SOURCE: Courtesy of Gemini Observatory/AURA.

- Similar instrumentation on a 30-m-class telescope to extend observations to the Magellanic Clouds;
- More investment in theoretical pre-main sequence evolutionary tracks and their calibration;
- Deep X-ray surveys to identify low-mass populations, particularly for members without infrared excesses, with sensitivities to X-ray luminosities down to a few times 10^{28} erg s^{-1} for emission at $T > 10^7$ K at the galactic center to detect low-mass members without disks and thus establish the full range of the IMF; and
- Wide-field, high-precision relative astrometry of the cluster with respect to the background field for establishing cluster membership.

Conclusions: Star Formation

Table 4.2 summarizes the panel's conclusions on activities to address its first science question.

PSF 2. HOW DO CIRCUMSTELLAR DISKS EVOLVE AND FORM PLANETARY SYSTEMS?

Disks are integral both to the formation of stars over a wide range of masses and to their planetary systems. Formed as a consequence of the angular momentum in a molecular cloud core undergoing gravitational collapse, disks initially build stars during a phase of active accretion. Later, as the mass reservoir from the collapsing core is dissipated, the thinning disk provides a fertile environment for the growth of planets. Images of disks around young stars, both from optical/near-infrared light scattered by small grains along the surface and from submillimeter- and millimeter-wavelength emission from large grains in the midplane, show that many disks have sizes at least as large as that of our own solar system if not much larger. The ubiquity of potentially planet-forming disks around the youngest stars, demonstrated most recently by a complete photometric census of young clusters closer than a kiloparsec with Spitzer, surpasses the current exoplanet detection frequency, suggesting that the majority of nearby planetary systems await discovery.

Astronomers' understanding of the evolution of protoplanetary disks advanced dramatically in the past decade. It is now known that by 3 Myr, dust in the inner few tens of astronomical units is strongly depleted in about half the stars in a cluster (Figure 4.6), in pace with the depletion of disk gas and the diminution of stellar accretion. By 6 Myr, nearly all inner disks have disappeared; this limits the timescale for planet formation. About 10 percent of disks around stars younger than 3 Myr have breaks in their spectral energy distributions indicative of dust-free inner holes or gaps 1 to 40 AU across, with a few of the largest confirmed by submillimeter and

TABLE 4.2 Panel's Conclusions Related to the Study of Star Formation

Observation Type	Area	Angular Resolution	Sensitivity	Instrumentation	Science Area
Dust continuum emission survey	Galactic plane (10^3 deg^2); detailed studies of selected star formation (SF) regions (10^2 deg^2)	1-20″	Galactic-plane surveys: 1 $M_\odot$ (≤5 mJy/beam at λ = 0.4-1.1 mm for dust at 2 kpc); detailed studies: 0.01 $M_\odot$	ALMA; Herschel; large-format (10^3-10^4 pixels) bolometer array cameras on large, single-dish millimeter/submillimeter/far-IR telescopes	SF rates/SF efficiencies; core properties, CMF; physics of CMF/IMF connection
Molecular and atomic line emission surveys	Galactic plane (10^3 deg^2); detailed studies of selected SF regions (10^2 deg^2)	1-20″	0.05 K at Δv = 0.1 km/s	ALMA; Herschel; SOFIA; EVLA: heterodyne focal plane arrays on millimeter interferometers and large single-dish telescopes; wide-field imaging capabilities and large collecting areas for centimeter-wave interferometers	SFRs/SFEs, feedback; core properties, CMF; physics of CMF/IMF connection
Radio continuum survey	Selected SF regions (10^2 deg^2)	≤1″	1-100 μJy/beam at 1-50 GHz	EVLA; large collecting areas for other centimeter-wave interferometers	Role of ionized gas in feedback; physics of CMF/IMF connection
Infrared, submillimeter, and millimeter polarization imaging	Selected SF regions (10^2 deg^2)	0.5″-1″	1-σ IR polarization fraction ~0.1% to infer magnetic field direction from measured position angle; 1-σ millimeter/submillimeter polarization fraction <1%	O/IR polarimeter on 8-m-class telescope; MIR polarimeter on SOFIA; polarimeters on submillimeter/millimeter telescopes	Role of B-fields in SFR/SFE; core properties; physics of CMF/IMF connection
Zeeman measurements	Selected SF regions (10^2 deg^2)	~10 mas for masers; 0.5″-30″ for disks to cores	1 mK rms with Δv = 0.1 km/s to measure small Zeeman shifts for typical magnetic fields in star-forming clouds	VLBI, ALMA, EVLA, and larger-area radio interferometers, heterodyne focal plane arrays on millimeter interferometers and large single-dish telescopes	Role of B-fields in SFR/SFE; core properties; physics of CMF/IMF connection
Near-IR stellar census imaging and spectroscopic survey	Galactic plane (10^3 deg^2), Magellanic Clouds	≤20 mas	K ~ 19, 23 mag for hydrogen burning limit in the galactic center, LMC, respectively	30-m-class telescope plus AO with multiple object IFU; JWST	SFR/SFEs; IMF studies
X-ray imaging of clusters	All clusters within Local Group	0.1″	10^{28} erg s^{-1} for T > 10^7 K at galactic center	New X-ray satellite	Stellar content of clusters for SFR/SFEs; IMF studies

TABLE 4.2 Continued

Observation Type	Area	Angular Resolution	Sensitivity	Instrumentation	Science Area
Trigonometric parallaxes	Selected SF regions	few mas	≤2-µas positional accuracy	Increased sensitivity on very long centimeter-wave interferometer baselines; optical interferometer in space (Gaia)	Stellar populations, distance measurements for IMF calibration
Theory support	Development of new efficient algorithms for following molecular cloud evolution, core formation, and collapse to protostars spanning many orders of magnitude in size and density with improvements to radiation transfer and thermal physics, and inclusion of magnetic fields; support for improvements to stellar evolutionary tracks				
Laboratory astrophysics	Studies of far-infrared and submillimeter molecular spectra, ranging from diatomics to prebiotic molecules, especially to aid in-line identification; determination of chemical rates involving granular surface processes and unusual gas-phase reactions such as radiative association for star-forming regions				

NOTE: Acronyms are defined in Appendix C.

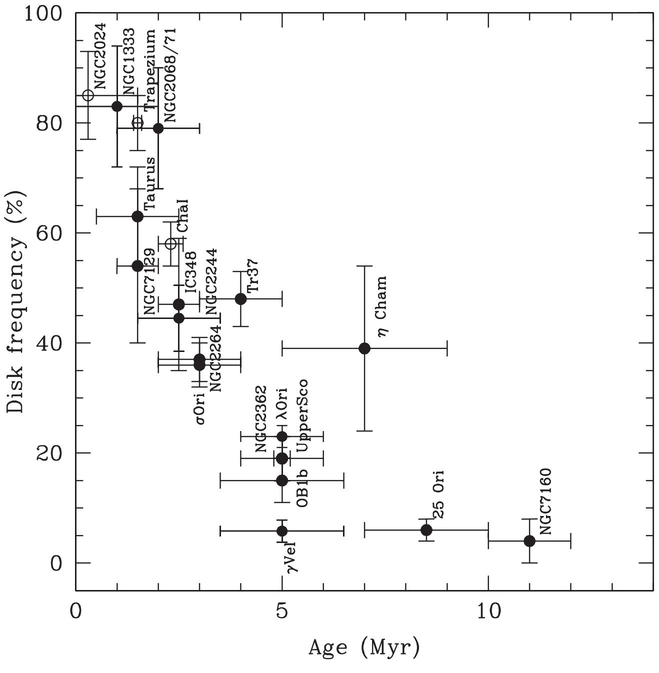


FIGURE 4.6 Declining frequency of disks in clusters of ages 0.5 to 11 Myr. SOURCE: J. Hernandez, L. Hartmann, N. Calvet, R.D. Jeffries, R. Gutermuth, J. Muzerolle, and J. Stauffer, A Spitzer view of protoplanetary disks in the γ Velorum Cluster, *Astrophysical Journal* 686:1195-1208, 2008, reproduced by permission of the AAS.

millimeter imaging at approximately 50 AU resolution (Figure 4.7). The gaps in at least some of these *transitional* disks may be carved by newborn, giant planets.

The past decade also witnessed major advances in the understanding of later-stage debris disks, composed of optically thin dust generated from collisions between macroscopic bodies. Among solar-type stars of ages 10 to 100 Myr, about 15 percent possess debris disks. Such disks—youthful analogues of the solar system's Kuiper belt—also hold keys to the process of planet formation. A dozen debris systems are close enough to have been imaged at a resolution of several astronomical units in submillimeter emission and visible/near-IR scattered light, revealing intricate structures: gaps, clumps, and eccentric rings. These features are probably sculpted by giant planets, which await direct detection in the coming decade; some may be imaged already.

Another revolution in the understanding of protoplanetary and debris disks impends, as significant increases in angular resolution are anticipated from near-IR to centimeter wavelengths initially with ALMA, the EVLA, and JWST, and later with a GSMT, along with orders-of-magnitude improvement in the dynamic range, or *contrast*, of images; the detection and mapping of complex molecules in disks;

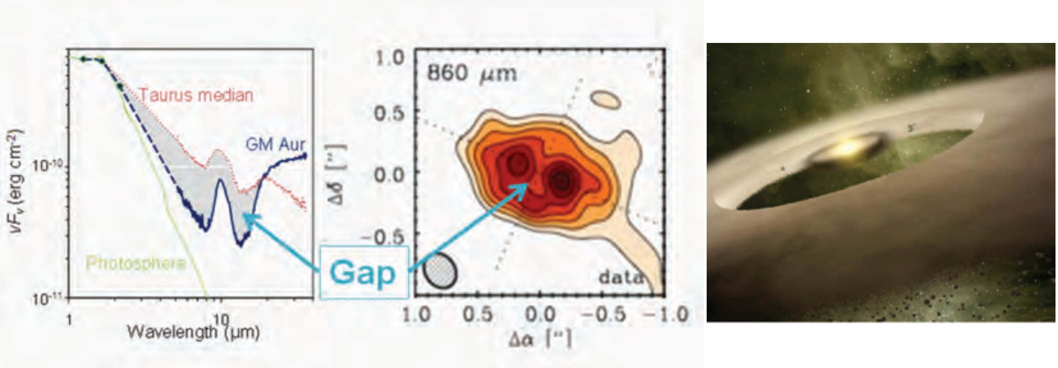


FIGURE 4.7 Gap in a transitional disk, probably caused by the gravitation of one or more newborn giant planets, around GM Aurigae, a solar-mass star 140 pc away. *Left:* Spitzer-IRS spectrum reveals a flux deficit shortward of 20 μm relative to the median in its cluster, for which models indicate an absence of small dust grains in the 5- to 24-AU radial range, with dusty material at smaller and larger radii. *Center:* Truncation of the disk at 24 AU is verified by central brightness minimum in this SMA image of dust emission at 860 μm. *Right:* Artist's conception of the disk, and the gas-giant protoplanets hypothesized to have created the gap. SOURCE: *Left:* Data from N. Calvet, P. D'Alessio, D.M. Watson, R. Franco-Hernández, E. Furlan, J. Green, P.M. Sutter, W.J. Forrest, L. Hartmann, K.I. Uchida, L.D. Keller, et al., Disks in transition in the Taurus population: Spitzer IRS spectra of GM Aurigae and DM Tauri, *Astrophysical Journal Letters* 630:L185-L188, 2005. *Center:* A.M. Hughes, S.M. Andrews, C. Espaillat, D.J. Wilner, N. Calvet, P. D'Alessio, C. Qi, J.P. Williams, and M.R. Hogerheijde, A spatially resolved inner hole in the disk around GM Aurigae, *Astrophysical Journal* 698:131-142, 2009, reproduced by permission of the AAS. *Right:* NASA/JPL-Caltech/T. Pyle (SSC).

and synoptic surveys in both the optical and the IR. Here the panel poses some key questions about disks that can be answered definitively in the coming decade, provided the requisite observational capabilities and theoretical developments are mustered.

What Is the Nature of the Planet-Forming Environment?

There are still many unknowns regarding the mass, structure, and evolution of disks in the pre-planetary state. How does disk mass evolve with time? How are mass and angular momentum transported to form stars and cause planets to migrate? To what extent is steady accretion punctuated by violent episodes? What are the thermal and chemical structures of disks? Where are the molecules regarded as prebiotic—that is, where are the molecules formed of the simple sugars and amino acids from which living organisms derive? Are the most massive stars also assembled from disks and thus likely to harbor planets?

Essential to answering these questions are improvements in angular and spectral resolution and imaging contrast, especially at wavelengths from the near-IR through the millimeter. To date some 60 protoplanetary disks have been imaged at >20-AU resolution by means of emission from gas and dust with the current generation of millimeter/submillimeter aperture-synthesis arrays (SMA, CARMA, PdBI) and by means of scattered light with HST and AO on ground-based telescopes. A major goal for the next decade is to improve significantly the spatial resolution in imaging such disks. Initially ALMA and JWST will achieve resolutions around 10 AU for the thousand nearest disks, but the panel recommends the following:

- Every effort should be made to push to 1 AU (10 mas) resolution, with upgrades for ALMA and with a GSMT, in order to achieve the biggest scientific return.

One outcome of AU-resolution imaging of protoplanetary disks would be a clear definition of gaps and small-scale structures that will illuminate the dynamics of planet-disk interaction and early planetary migration. Of special significance would be the detection in disks of gravitational instabilities, such as spiral density waves, that would indicate self-gravitating disks. Compression amplitudes in gravitational instabilities should exceed a factor of 10, making their contrast with respect to the disk very large—brighter than the unperturbed disk—in molecular lines that trace high-density gas (Figure 4.8). Currently the role of self-gravity in disk evolution is a controversial topic, but if density waves are discovered, disk masses could be derived independently of uncertain gas-to-dust ratios and grain-size distributions, and the role of such instabilities in giant-planet formation could be assessed. Detecting waves in disks with the normal range of accretion rates will

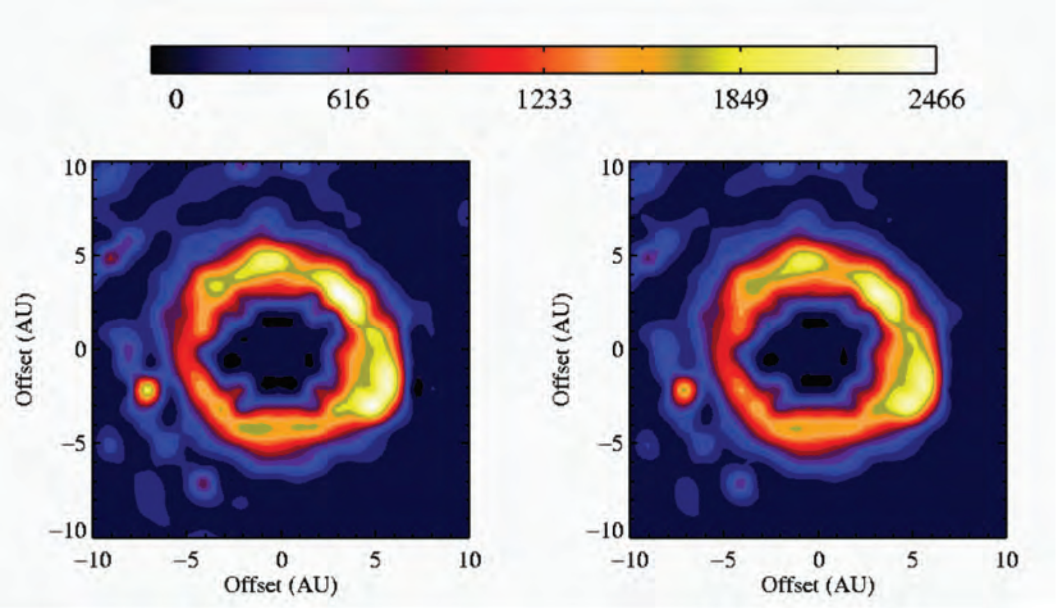


FIGURE 4.8 A gravitationally unstable disk imaged *in line emission from very dense gas* in HCO^+ J = 7-6 (λ = 480 µm) for a 0.09 M disk around a 1 M star at a distance of 140 pc, viewed face-on. *Left:* Three-dimensional model at full resolution shows a Jovian-mass collapsing fragment in the outer ring. *Right:* The same scene convolved to the resolution of ALMA in its longest-baseline configuration, with 0.007-arcsec resolution (1 AU at 140 pc). At ALMA's projected sensitivity, the noise level in this image would be 100 K km/sec in 16 hours of observation; the protoplanetary fragment would be marginally detected (3 sigma) in about 1 hour. SOURCE: D. Narayanan, C.A. Kulesa, A. Boss, and C.K. Walker, Molecular line emission from gravitationally unstable protoplanetary disks, *Astrophysical Journal* 647:1426-1436, 2006, reproduced by permission of the AAS.

also enable an evaluation of the role of self-gravity in the disk's accretion flow. Breakthroughs in the theory of accretion disks will require disk thermodynamics to be treated more realistically. For example, gravito-turbulent accretion is driven by net cooling, so the next generation of disk models should include in their energy budgets the feedback from stellar irradiation.

The panel recommends several key measurements and measurement capabilities that will advance the understanding of the structure and evolution of preplanetary disks, as follows:

- Sensitive millimeter-wave imaging with ALMA at few-AU resolution plus high-spectral-resolution observations by ALMA and Herschel in column-density tracers and density-sensitive molecular lines to provide the first maps of the opacity, density, thermal, and chemical structure of disks. In turn this will improve estimates of disk masses, enable measurements of the abundances of molecular ions and complex prebiotic molecules that have so far proved impossible, and test chemical models.

- Longer-wavelength imaging with the EVLA, and high-resolution, molecular-vibration spectroscopy with SOFIA, to provide important constraints on the chemical and radiation environments. These observations should also yield measurements of the disk ionization fraction, which is crucial for assessing the role of magnetic fields in disk transport and distinguishing actively accreting gas from dead zones. The latter may promote the survival of giant planets by acting as barriers to planet migration.
- High-resolution imaging at near-IR wavelengths, initially with JSWT and later with a GSMT to trace disk structure by means of scattering by small grains residing on the disk surface. Gap contrasts can be several times higher than for thermal emission from the midplane (Figure 4.9). The near-IR will also be a fertile ground for identifying gravitationally induced density waves, which can appear at a contrast of about 0.1 relative to the rest of the disk.
- High-contrast imaging with extreme adaptive-optical (ExAO) correction to see the disk against the glare of the star in the near-IR. Combining extreme AO with the sensitivity and resolution of a GSMT will be required to probe the innermost regions, where disk/star contrast factors at 10-AU range from about 10^{-3} at 1 μm (see Figure 4.9) to about 10^{-1} at 10 μm.
- Combining high spatial and spectral resolution ($\lambda/\Delta\lambda > 50{,}000$) in the near- and mid-IR, coupled with chemical models to characterize the chemistry and ionization in the warm inner disk, which is where our own terrestrial planets reside. An integral field spectrometer with $\lambda/\Delta\lambda \geq 5{,}000$ working at 10-mas resolution in the optical and near-IR would be able to profile the mineralogy and location of small solids in partially depleted disks, enabling an assessment of dust grain growth and crystallization, and testing models of coagulation by grain-grain collisions. This capability would also be invaluable in mapping the structure and ionization of accretion-powered jets on astronomical unit scales, elucidating the onset of collimation and possibly also the origin of the accretion-outflow connection that has so far eluded scientists' grasp.

Achieving 10-mas resolution at submillimeter and near-IR wavelengths would also illuminate the role of disks in the formation of high-mass stars, by resolving distant, crowded young clusters (100 AU at 10 kpc). Disks around young, high-mass stars could be identified by means of direct imaging, and the properties of these necessarily massive, short-lived, and externally ionized disks could be studied for the first time and could be used to discriminate among models for massive star formation.

A further goal in the study of protoplanetary disks is to characterize their accretion history by means of wide-field long-term synoptic surveys in both the

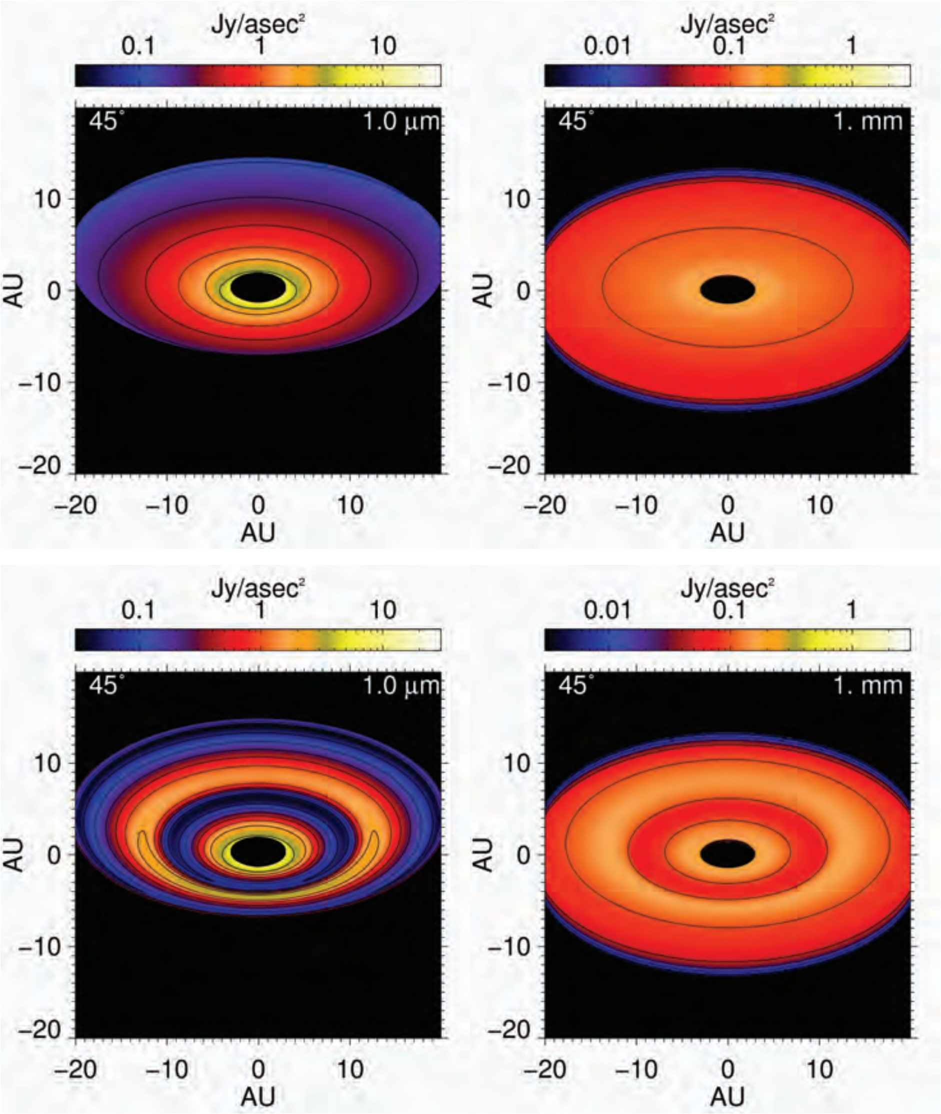


FIGURE 4.9 Surface brightness distributions and gap contrasts in both scattered light and thermal emission of the inner 20 AU of a protoplanetary disk around a 1-Myr-old solar mass star, tipped by 45° to the line of sight. The upper panels show the unperturbed disk, at 1 μm (left) and at 1 mm (right). In the lower panels the disk has a 4-AU-wide gap at 10 AU, created by a 100-Earth-mass planet, showing a contrast ratio of about 0.1 in scattered light and 0.5 in thermal emission with respect to the rest of the disk. At the distance of the Taurus star-formation region each image is 0.28 arcsec wide. Assuming a GSMT with diameter 30 m operating at 1 μm, the instrumental point-spread function would be 1.1 AU in diameter (FWHM), and the surface brightness of 1 Jy arcsec^{-2} corresponds to a contrast, with respect to the stellar image, of about 3×10^{-3}. Expected noise levels at 1 Jy arcsec^{-2} are on the order of 1 hour with ALMA; GSMT times are much shorter, depending on backgrounds. SOURCE: Courtesy of H. Jang-Condell, personal communication.

optical and near-IR, probing both optically visible and heavily extincted systems. This monitoring can be carried out with conventional resolution and sensitivity (limiting *K* magnitude of 18 per epoch), but it would ideally include measures of stellar accretion rate, such as near-IR hydrogen recombination lines. Such studies would clarify how important eruptive FU Ori-like events are in the buildup of stellar mass, and whether planet formation may be periodically disrupted.

High-resolution spectroscopy in the ultraviolet (UV) and X-ray regimes would significantly sharpen the understanding of how accretion flows arrive on the surface of a star, with implications for the star's angular momentum, and for its accretion luminosity, which controls the ionization of the planet-forming regime in the disk. Ground-based near-IR interferometry also has a vital role in the study of protoplanetary disks, probing the inner 0.1 AU, where the stellar magnetosphere and radiation from stellar accretion spots can greatly affect the disk structure. In the coming decade it will be important to extend the sensitivity of this technique so that more than just the few brightest systems can be observed.

How Do Giant Planets Accrete from and Interact with Disks, and What Are These Young Planets Like?

An especially exciting discovery potential that is presented with high-angular-resolution imaging in the near-IR is the possibility of directly detecting infant giant planets (<5 Myr) within disk gaps that they have excavated. In nearby young clusters (<500 pc), there are more than 200 disks whose spectra suggest gaps or inner holes, for which near-IR direct detection and characterization of companions are realizable goals. This is illustrated in Figure 4.10 for the case of hypothetical 1-M_J protoplanets residing in known disk gaps. At this young age, Jovian planets are relatively luminous (~10^{-5} $L_\odot$).

- Thus the biggest detection challenge for directly detecting infant giant planets is imaging contrast, which needs to be 10^{-5} or better. A handful of targets may be detectable with JWST and extreme-AO near-IR images on 8- to 10-m telescopes, but most of the known candidates will require the higher resolution of 30-m telescopes, fitted with similar high-contrast instrumentation.

Finding newborn giant planets will enable astronomers to test, for the first time, theories of how these planets nucleate from the disk and feed from it. The infrared luminosity from a young giant planet traces the planet's age and primordial heat of formation, and as such can be used to decide how these planets formed: whether in a "cold-start" nucleated instability around a rock-ice core (known as "core-accretion" formation), or in a "hot-start," high-entropy, gravitationally unstable clump of disk gas. Characterization of the atmospheres and neighbor-

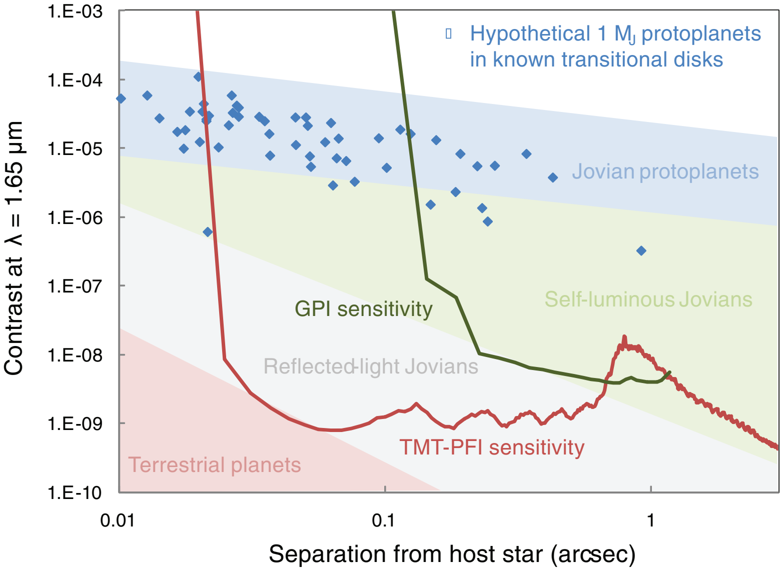


FIGURE 4.10 Photometric contrast with respect to parent star required for detection in the astronomical *H* band of hypothetical, ringless, Jovian-mass protoplanets in the gaps of known transitional disks (blue diamonds), compared to the 5σ contrast sensitivity predicted for two proposed extreme-adaptive-optical instruments (solid curves); detection space lies above these curves. The protoplanets are assumed to be the same age as the host star and to lie just inside the outer edges of the inferred gaps. Also shown are the domains in which one could observe older Jovian planets and smaller terrestrial planets. SOURCE: Sensitivities for the Gemini Planet Imager (GPI) and the proposed Thirty Meter Telescope's Planet Formation Instrument (PFI) were adapted from B. Macintosh, M. Troy, R. Doyon, J. Graham, K. Baker, B. Bauman, C. Marois, D. Palmer, D. Phillion, L. Poyneer, I. Crossfield, et al., Extreme adaptive optics for the Thirty Meter Telescope, *Proceedings of SPIE* 6272:62720N, 2006, reproduced by permission of SPIE. Additional data from J.M. Brown, G.A. Blake, C.P. Dullemond, B. Merín, J.C. Augereau, A.C.A. Boogert, N.J. Evans II, V.C. Geers, F. Lahuis, J.E. Kessler-Silacci, K.M. Pontoppidan, and E.F. van Dishoeck, Cold disks: Spitzer spectroscopy of disks around young stars with large gaps, *Astrophysical Journal Letters* 664:L107-L110, 2007; K.H. Kim, D.M. Watson, P. Manoj, E. Furlan, J. Najita, W.J. Forrest, B. Sargent, C. Espaillat, N. Calvet, K.L. Luhman, M.K. McClure, J.D. Green, and S.T. Harrold, Mid-infrared spectra of transitional disks in the Chamaeleon I Cloud, *Astrophysical Journal* 700:1017-1025, 2009; K.H. Kim, personal communication; B.A. Macintosh, J.R. Graham, D.W. Palmer, R. Doyon, J. Dunn, D.T. Gavel, J. Larkin, B. Oppenheimer, L. Saddlemyer, A. Sivaramakrishnan, J.K. Wallace, B. Bauman, D.A. Erickson, C. Marois, L.A. Poyneer, and R. Soummer, The Gemini Planet Imager: From science to design to construction, *Proceedings of. SPIE* 7015:701518-701518-13, 2008.

hoods—protosatellite disks or rings—of infant giant planets could be carried out with low-resolution spectrometers and polarimeters on the same facilities that imaged them.

Some infant giant planets may be detectable along with their cooler accretion streams, by ALMA, as illustrated in Figure 4.11. Imaging of the planet's immediate environment, together with high-resolution theoretical studies of planet-disk

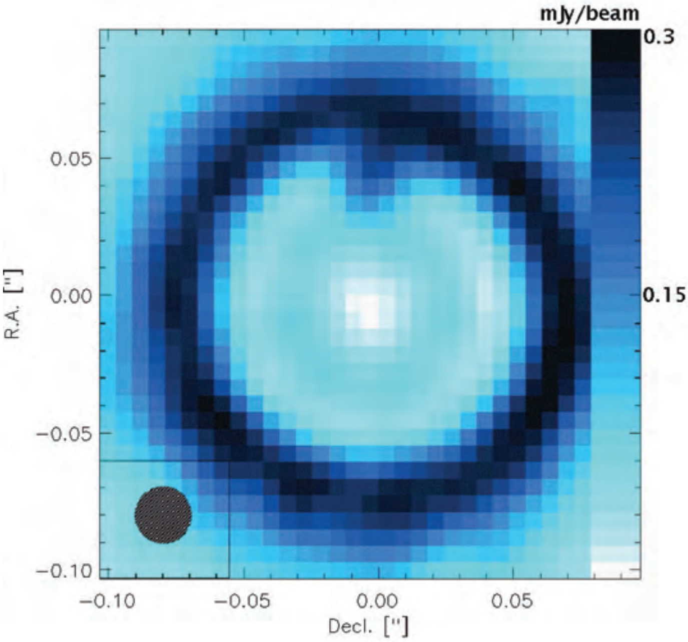


FIGURE 4.11 Simulation of an 8-hour, 345 GHz ALMA observation of a 0.01-$M_\odot$ circumstellar disk with a recently formed planet of 1 M_J, at the 12 o'clock position, orbiting 5 AU away from a 0.5-$M_\odot$ star (not shown). The distance to this hypothetical system is 100 pc. Most of the flux detected at the planet's position is from material about to be accreted by the planet. SOURCE: S. Wolf and G. D'Angelo, On the observability of giant protoplanets in circumstellar disks, *Astrophysical Journal* 619: 1114-1122, 2005, reproduced by permission of the AAS.

interaction, should help to decide how disk non-axisymmetries torque planetary orbits, and should help to determine the sign and speed of orbital migration and eccentricity evolution.

An important milestone will be the measurement of the masses of infant planets through stellar reflex motion. This will be extremely challenging even in the nearest young clusters, either through radial-velocity ($\Delta v \approx 10$ m sec^{-1}) or astrometric ($\Delta r \approx 0.001$ AU) means. Magnetic activity and high rotation rates impose fundamental limits to radial-velocity precision on very young stars, which would only be sensitive to close-in or very massive planets. Astrometry faces less severe limitations, but the small reflex motions involved (≤7 μas) will require space-based facilities. Future space-based astrometric surveys could also further extend the census of infant planets in the 1-5 AU region, where direct imaging is most challenging.

Somewhat older giant planets (10-100 Myr), including those producing gaps in nearby debris disks, also await direct detection by high-contrast imagers. Non-axisymmetric structures seen in these disks ($d < 50$ pc)—clumps, warps, and eccentricities—are attributed to sculpting by planets. Planets at stellocentric distances of 20-200 AU, where a handful of debris disks have been resolved and where high-contrast imagers operate best, could be the result of long-range migration, but also might provide long-sought examples of planet formation by gravitational instability. It is here, in the outermost reaches of disks, that fragmentation by gravitational instability during the protoplanetary phase is most viable, since cooling times are shortest and orbital times longest. Indeed a system of such planets has recently been reported, orbiting an A4V star with a debris disk, HR 8799. In the future, as higher angular resolution and high contrast imaging become available, the inner regions of these systems can also be searched for young planets. In this age range, a Jovian planet luminosity is predicted to be $\sim 3 \times 10^{-6} - 2 \times 10^{-7}\ L_{\odot}$. Planets in some 200 nearby debris-disk systems are thus within the grasp of the next generation of high-contrast instruments on 8- to 10-m telescopes (see Figure 4.10).

Major advances in the theoretical understanding of planet formation, migration, and interactions with their disks—both gaseous and debris—will be required to take advantage of the wealth of observational data that will be obtained in the coming decade. The required efforts span a wide range of problems, including investigations of turbulence driven by magnetic fields in partially ionized gas; coagulation of solids with settling; disk-driven migration; and fragmentation in irradiated, self-gravitating disks. Many if not all of these problems will require major computational resources.

What Can Debris Disks and the Kuiper Belt Demonstrate About the Dynamical Evolution of Planetary Systems?

The orbits of known extrasolar planets reflect a host of dynamical processes—some violent—including migration, planet-planet scattering, and forcing by stars. Our own solar system did not emerge unscathed, as is evident in the hot thick disk of the Kuiper belt, filled with resonant orbits and large-perihelion bodies that suggest a dynamically violent epoch when 99 percent of the solar nebula's solids were ejected. Similar dynamical evolution is currently underway in extrasolar debris disks, and the variety of structures observed to date in a handful of systems already hints at the diversity of their planetary systems. With increased angular resolution and sensitivity in the O/near-IR and in the submillimeter-wavelength regions, unprecedentedly fine structures in the dust will be revealed (Figure 4.12), and for systems with asymmetric structures, disk pattern speeds will be measured. Careful modeling of the observed disk morphology will enable the masses and orbits of sculpting bodies to be inferred. New theoretical approaches need development, since extrasolar debris disk densities are so large that interparticle collisions outweigh radiation drag, with consequences for disk structure that are only beginning to be explored. The panel recommends the following:

- Future models should track not only forcing by planets, but also particle fragmentation in collisional cascades and momentum exchanges between colliding particles.

Multi-epoch observations will break modeling degeneracies by measuring proper motions of debris disk clumps—that is, speeds of wave patterns driven by plan-

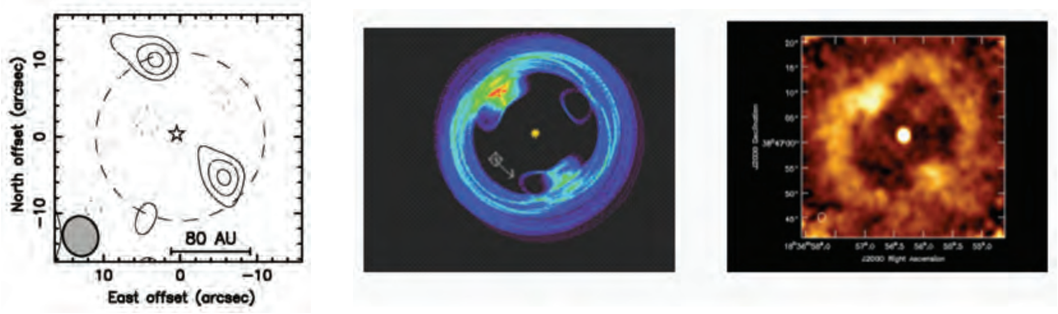


FIGURE 4.12 Planet-sculpted features in Vega's debris disk. *Left:* IRAM image at 1.3 mm with 2-deg resolution. *Center:* Model based on resonant trapping of planetesimals by migration of Neptune-mass planet. *Right:* Simulated ALMA observation at 350 GHz with a 1.5-arcsec beam from the 12-element ACA neutron star configuration, plus three additional compact configurations, for 11 hours. Total flux density in the simulated image is about 42 mJy. SOURCE: *Left:* D.J. Wilner, M.J. Holman, M.J. Kuchner, and P.T.P. Ho, Structure in the dusty debris around Vega, *Astrophysical Journal Letters* 569:L115-L119, 2002, reproduced by permission of the AAS. *Center:* M.C. Wyatt, Resonant trapping of planetesimals by planet migration: Debris disk clumps and Vega's similarity to the solar system, *Astrophysical Journal* 598:1321-1340, 2003, reproduced by permission of the AAS. *Right:* Courtesy of R. Reid, L. Mundy, and A. Wooten, personal communication.

ets (for planets orbiting at 100 AU, the pattern speed is approximately 1 AU/yr). Characterizing debris disks of different ages will enable the understanding of how planets and destructive collisions scour disks of their smallest planetesimals, leaving only large bodies.

A more secure understanding of the dynamical history of our own solar system will come from an unbiased characterization, both kinematic and physical, of the Kuiper belt. Evidence that the orbits of the ice giants Uranus and Neptune have undergone dramatic changes, involving migration and/or violent planet-planet scatterings, is imprinted in the distribution of orbits of known Kuiper belt objects (KBOs), especially in the well-populated mean-motion resonances. The current census of KBOs is pieced together from a patchwork of surveys, each with different areal coverage and limiting magnitude, giving an incomplete picture at best.

- The panel recommends a repeated systematic whole-sky synoptic survey to a limiting red magnitude of $R = 24$, sensitive to 30- to 50-km bodies at 40 AU that will provide a large enough unbiased sample (>20,000 KBOs) to test dynamical evolution models, yielding the relative populations of the resonances, the abundance of scattered and detached populations, and the incidence of retrograde objects and binaries.

Surface compositions will also provide important clues to the chemistry of planetesimal formation. KBOs show a broad range of colors, suggesting a wide variety of compositional types from ice-rich to organic-rich. About half possess ultra-red matter that may be extremely pristine, preserved only in the deep freeze of the Kuiper belt. Key questions include whether there is a radial composition gradient across the belt and whether resonantly trapped KBOs are of similar composition or have migrated from different source regions. To answer these questions requires the following:

- Spectra at resolving power $\lambda/\Delta\lambda \approx 500$ from 0.3 to 2.5 μm to measure the depth and width of bands such as methane for the brightest objects, and multi-wavelength precision photometry for the faintest ones. Optical colors need to be accurate to ±0.03 magnitudes, and O/near-IR color indices to ±0.06 magnitudes, for objects as faint as $R = 24$ in order to discriminate among surface properties and examine correlations between colors and dynamical properties. Time-resolved photometry with similar precision can reveal the shapes and rotation states of the KBOs, two further constraints on their origins and collisional histories.

Conclusions: Disk Evolution and Planet Formation

Table 4.3 summarizes the panel's conclusions about activities to address its second science question.

PSF 3. HOW DIVERSE ARE PLANETARY SYSTEMS?

The great potential of the coming decade is to achieve a census of the planetary bodies orbiting the Sun's neighboring stars. With a commitment of sufficient resources, scientists' efforts will be rewarded with no less a prize than the knowledge of the population of small, rocky, Earth-like planets. The detection and characterization of these worlds will invigorate the efforts both to understand their origin in the grand context of star formation (science question PSF 1) and protoplanetary disk evolution (science question PSF 2), and to compare their compositions, structures, and atmospheres with those of the planets of the solar system. It is this great quest for detection and characterization that concerns the panel in the present section. In the subsequent section on science question PSF 4, the panel focuses on the particular question of life on these worlds, arguably one of the greatest intellectual endeavors ever undertaken by humanity. The panel then identifies a fast-track opportunity that may permit the study of habitable worlds within 5 years (the PSF discovery area), albeit for stars very different from the Sun.

The dominant methods of discovery of planets orbiting other stars (known as exoplanets) are measurements of radial velocities (wherein one infers the existence of a planet through its acceleration of the central star) and photometric transits (in which one measures the periodic dimming of the central star due to the intervening

TABLE 4.3 Panel's Conclusions Regarding the Study of Disk Evolution and Planet Formation

Recommendations	Goals for Observational Capabilities		
Map physical and chemical composition of protoplanetary disks on AU scales.	**AU-scale imaging,** **λ = 1 μm to 1 cm** Resolution: 10 mas Field of view: few arcsec 1-mm sensitivity 0.01 Jy arcsec^{-2} 1-μm sensitivity 0.1 Jy arcsec^{-2}	**Spectroscopy,** **λ = 1 μm to 1 cm** $0.1 < \lambda < 13$ μm: Δv = 5 km s^{-1} $\lambda > 0.2$ mm: Δv = 0.1 km s^{-1} **Laboratory astrophysical spectroscopy,** **λ = 1 μm to 1 cm** Prebiotic and organic molecules, molecular ions	**X-ray, UV spectroscopy** Resolving power $\lambda/\Delta\lambda$ = 5,000 **Integral-field spectroscopic imaging** $\lambda/\Delta\lambda$ = 5,000 at 1 μm
Characterize disk gaps and outer disk structure in transitional disks.			
Find and characterize infant giant planets in protoplanetary disk gaps and debris disks.			**High-contrast infrared imaging** 1-μm planet/star contrast 10^{-5}-10^{-7}
Identify disk features associated with gravitational instability, such as spiral waves.			**Broadband polarimetry** 1% accuracy on all four Stokes parameters, 0.1-13 μm
Resolve images of debris disks; look for changes on orbital timescales.			
Characterize accretion history of young stars, whether continuous or episodic.	**Wide-field synoptic surveys** Visible and near-IR wavelengths: interval ≤ weeks, duration ≥ years		Limiting magnitude K = 18
Obtain a de-biased census of Kuiper belt objects.			Limiting magnitude R = 24
Theory	Studies of planet formation, planet-disk interactions, using new algorithms, methods; investment in large-scale computational resources		
Laboratory	Identification of rotational molecular spectra in terahertz range		

passage of an orbiting planet). Together these methods account for nearly all of the more than 400 known planetary systems, in which more than 500 exoplanets have been detected. The nascent methods of microlensing (in which the gravitational field of a planetary system causes the magnification of an unassociated background star) and direct imaging (wherein one spatially separates the light from the planet from the glare of the central star) have each yielded a handful of detections.

Toward an Understanding of Planet Formation

A mere 15 years ago, the knowledge of "normal" planetary systems was restricted to the eight planets orbiting the Sun. The prevailing model for the formation of gas giant planets was that of core accretion (see above the subsection "How Do Giant Planets Accrete"), predicting that such planets would take 10-100 Myr to form and would be found on circular orbits at distances of 5-10 AU from their central star.

In contrast, the past decade of discovery has revealed much shorter timescales for giant-planet formation (see the subsection on giant-planet accretion) and startling diversity in the architectures of mature planetary systems (Figure 4.13). Astronomers have learned of the existence of "hot Jupiters," planets with masses similar to that of Jupiter, yet orbiting a mere 0.05 AU from their central stars. Although many stars indeed host Jupiter-mass planets at greater separations, the median eccentricity of the population is roughly 0.2, exceeding that of all the solar system planets. Furthermore, many of these eccentric planets are members of multiplanet systems, and in some cases the mutual dynamical interactions of the member planets can be directly observed. There are several strong pieces of

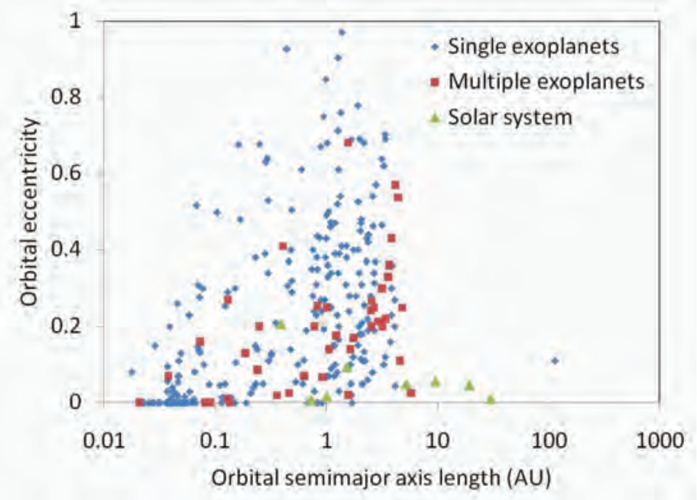


FIGURE 4.13 Orbital eccentricity as a function of orbital semimajor axis, for single exoplanets (diamonds), exoplanets in multiplanet systems (squares), and the solar system (triangles). Data from the Extrasolar Planets Encyclopaedia, hosted by Paris Observatory.

evidence that the fundamental tenets of the core-accretion model are correct, such as the metal enrichment of exoplanet host stars (Figure 4.14) and the rather high densities of some of the giant exoplanets (Figure 4.15), which demand heavy-element enrichment. Yet it is clear that the hot Jupiters could not have formed in place by core accretion. It also appears that gas and ice giants of our own solar system have moved significantly from their birth place. Such results have underlined the importance of studies of planetary migration and re-opened questions of the basic mechanisms of the formation of giant planets (see the subsection on giant-planet accretion).

Nature has proved far more inventive in forming planets than scientific theories have in predicting their properties. *Exoplanetary studies have progressed admirably during the past decade, but it still is not known whether solar systems like ours are commonplace or are cosmically rare. For a true understanding of the process of planet formation, the end states need to be understood: the planetary census must be completed.* The fundamental starting point for characterizing planetary systems is the determination of orbital properties and masses; this is vital not only for understanding the dynamics and composition of these systems but also for understanding how planetary systems originate. In addition, a further step should be taken to develop the sensitivity to detect and characterize the bulk properties of rocky planets akin to the solar system terrestrial planets. Novel methods need to be encouraged for the detection of planets far from their stars, at distances where the radial velocity and transit methods are precluded due to the very long orbital periods and low transit probabilities. Astrometry with Gaia will also contribute to the census of massive exoplanets in the coming decade.

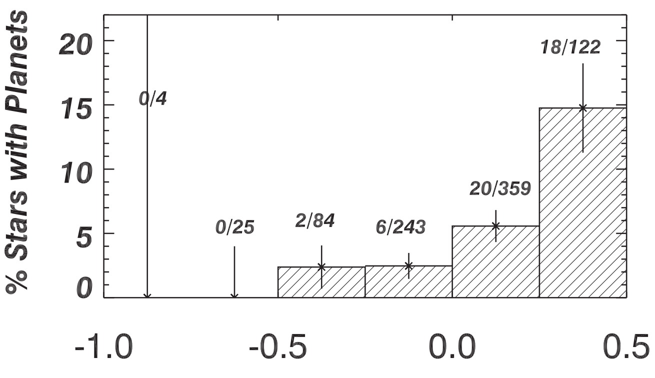


FIGURE 4.14 Correlation of stellar Fe/H relative abundance with possession of exoplanets. SOURCE: D.A. Fischer and J. Valenti, The planet-metallicity correlation, *Astrophysical Journal* 622:1102-1117, 2005, reproduced by permission of the AAS.

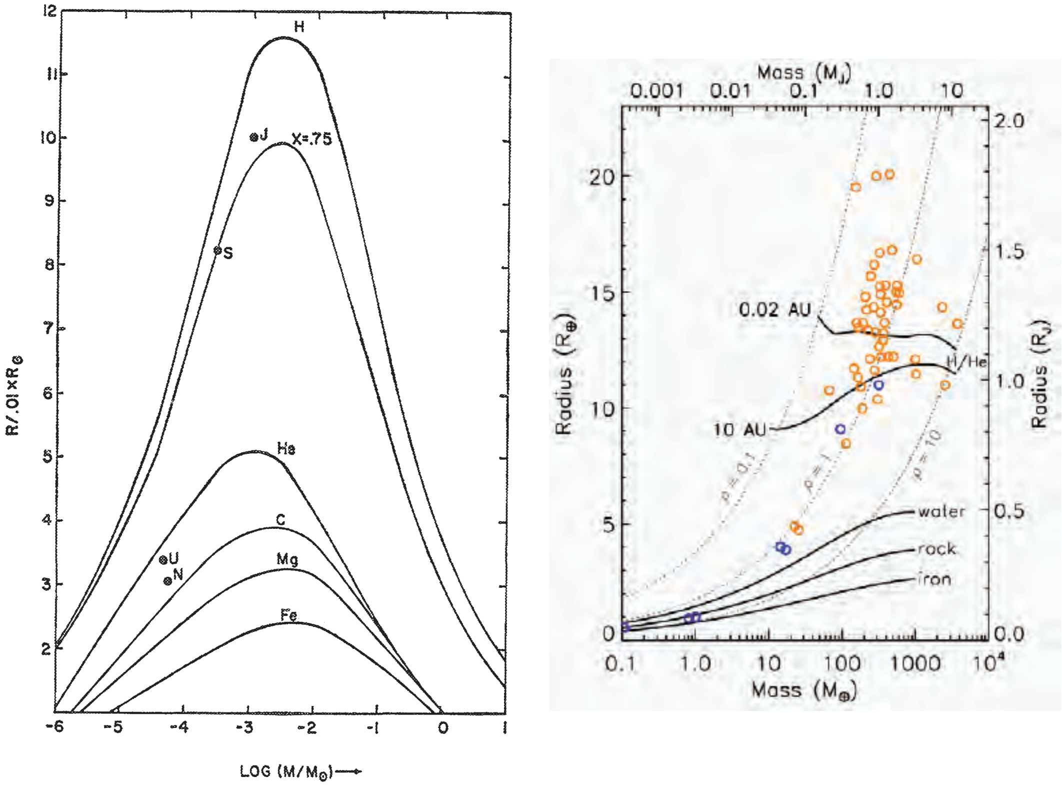


FIGURE 4.15 Observed masses and radii of planets both inside and outside the solar system, compared to models based on differing bulk composition. *Left:* The original demonstration by Zapolsky and Salpeter that Jupiter and Saturn are predominantly hydrogen and helium (i.e., a near-solar composition), whereas Uranus and Neptune are dominated by heavier elements. *Right:* A recent rendition in which the solar system's planets (blue points) are joined by a large number of exoplanets (orange points). The predicted radii of pure H/He bodies are shown for two representative orbital separations to show the expected effect of strong irradiation, but the observed variation exceeds the span of these two models. The two smallest exoplanets shown lie near Neptune and similarly have densities that require a bulk composition dominated by ice and rock. SOURCE: *Left:* H.S. Zapolsky and E.E. Salpeter, The mass-radius relation for cold spheres of low mass, *Astrophysical Journal* 158:809-813, 1969, reproduced by permission of the AAS. *Right:* J. Fortney, University of California, Santa Cruz.

Completing the Planetary Census, 2010-2020

The recently commissioned NASA Kepler mission, with a modest contribution from the CoRoT spacecraft, will offer humanity its first opportunity to study the population of terrestrial exoplanets within 1.5 AU of their stars. This mission employs the transit method and hence will deliver estimates of the planetary radii and orbital periods. However, the scientific return will be fully realized only if mass estimates can be obtained for a significant number of these planets. The determination of the radial-velocity orbit serves both as a robust confirmation that a Kepler signal is indeed planetary in origin and also yields the planetary density and thus constrains the composition, a topic of significant importance (see below). The challenges for radial velocities are that the targets are faint and the signal is small: An Earth-mass planet 1 AU from a Sun-like star produces a peak-to-peak radial-velocity difference of only 0.18 m sec^{-1}, five times smaller than the current state-of-the-art precision. *Even if the limiting precision were to remain fixed*, the panel recommends the following:

- A substantial expansion of the telescope time available to pursue radial-velocity work, since the mass determination for rocky bodies several times the mass of Earth or at smaller orbital separations is already within reach.

However, given the 10-fold improvement in radial-velocity precision of the past 15 years, the panel finds that it is not unreasonable to aim to accomplish the following:

- Develop advanced radial-velocity techniques capable of achieving 0.2 m/s precision as a basic requirement (which will only be possible for the most stable stars), with a long-term goal to do better. This will require the development of dedicated and highly specialized spectrographs and novel means for wavelength calibration.

Beyond the vital follow-up of Kepler-detected worlds, this investment in radial-velocity observations will also drive the understanding of more massive worlds located at distances of 1.5-10 AU from their stars, straddling the ice line that is thought to have played so crucial a role in their formation. There is not space here to mention the large number of ongoing radial-velocity projects that will continue to make important contributions to the planetary census; the panel necessarily groups all these efforts together in its endorsement of the expansion of ground-based radial-velocity measurements. If implemented on next-generation large-aperture telescopes (needed to gather a sufficient number of photons), this could permit a survey of the closest 100 Sun-like stars to find the nearest Earth-like planets orbiting within their stellar habitable zones. These systems would be invaluable for subsequent efforts to search for atmospheric biomarkers (science question PSF 4).

The method of microlensing complements that of radial velocities and transits. It does not require that data be gathered over a full orbital cycle and thus can provide in relatively short order the detailed statistics on the masses and orbital separations of planets in the outer reaches of planetary systems. Microlensing is demanding of instrumental stability: a satellite borne survey instrument, for which the point spread function profile would be stable over times very long compared to the duration of a planetary microlensing event, would—compared with a ground-based survey covering the same field of view—detect about 100 times as many planets of given mass and reach planetary orbits about 10 times larger. The results of a space-based microlensing survey would be particularly dramatic for several reasons. First, the sensitivity of Kepler to exoplanets declines sharply beyond 1.5 AU, whereas a microlensing survey in space could reasonably detect a statistically significant population of exoplanets from 0.5 to 15 AU. Second, microlensing will naturally survey a wider range of host stars and will include a large number of low-mass stars, as they are the most numerous type of star in the galaxy. Kepler will survey few low-mass stars owing to their intrinsic dimness. Importantly, a space-based microlensing survey can determine the planetary masses and projected separations in physical units and can frequently detect the light emitted by the host star. This is in contrast to ground-based microlensing surveys, which must generally rely on HST imaging to accomplish this, and there is insufficient time available to follow up a significant number of detections.

To obtain the fundamental data set to test the current picture of how planetary systems form, how their properties depend on the properties of the central star and, by inference, the conditions of the circumstellar disk, the panel recommends a combination of the following:

- A space-based microlensing survey, which will provide a statistical picture of planetary architectures, and
- The Kepler findings on terrestrial planet frequency within 1.5 AU of their central stars, augmented by
- Expanded radial-velocity measurements on 4- to 10-m-class telescopes to provide masses.

More generally, these efforts will demonstrate how typical our own solar system is.

Inferring the Bulk Compositions of Exoplanets

When both photometric and Doppler signals are measured for the same planet-hosting star, the planetary mass and radius are determined directly. These permit, in turn, an estimate of the density and, by inference, the bulk composition and the formation history. Fifty such transiting worlds are now known, and these objects have proven immensely valuable for constraining models of the physical structures of planets as a function of mass, composition, and external environment. Indeed, many show significant discrepancies from the theoretical prediction for a degenerate body with solar abundances (see Figure 4.15). In many cases the planets are larger than expected for their age, indicating that a mechanism to prevent cooling and contraction is in operation. A handful of systems are smaller than expected, indicating the presence of a large fraction of heavy elements, usually interpreted as a core. For nontransiting planets, the radial-velocity method provides only a lower bound on the mass, but

- The true masses could be determined with space-based astrometry. Since these would be the systems nearest to Earth, knowledge of the true masses for each of the planets in these systems would be extremely valuable.

Progressing downward through the mass range corresponding to Neptune-like bodies, two transiting examples of which are known at the time of writing, the interpretation of radii becomes increasingly ambiguous, because there are three ingredients—gas, ice, and rock—and thus multiple ways to obtain the observed radius and mass. Even lacking unique solutions, the observed diversity yields important constraints on the range of planetary-formation conditions. As one approaches masses that are only modestly larger than Earth's, one could distinguish between a body that is half ice–half rock (for example) and a body that is almost entirely rock, assuming that the atmosphere contributes only a small fraction of the radius. Deciding whether the atmosphere is small in this sense may prove to be very challenging, but for sufficiently low-mass bodies at modest orbital radii (e.g., 1 AU), strong theoretical arguments can severely limit the presence of a thick hydrogen-rich atmosphere. At the rather low level of accuracy attainable for determining bulk compositions, the present understanding of thermodynamics and pressure-density relationships for candidate materials is adequate for the task. The expected diversity of observations will thus be traceable to the diversity of conditions and environments of planet formation.

- The tremendously exciting opportunity to make informed estimates of the compositions of perhaps hundreds of Earth-like planets detected by Kepler serves as a compelling motivation for increased radial-velocity precision and the expansion of available observatory time to undertake these measurements.

Characterizing the Atmospheres of Exoplanets

It is essential that spectra be gathered to determine the chemistries, structures, and dynamics of exoplanetary atmospheres. Nearly all of the available data on the atmospheres of exoplanets comes from one of two techniques that are possible only for transiting systems: (1) In transmission spectroscopy, one takes the ratio of a spectrum gathered when the planet is in front of the star with a spectrum of the unocculted star, interpreting any residual absorption features as arising in the atmosphere of the planet. (2) In occultation spectroscopy, one observes the difference between a spectrum gathered when both the planet and star are in view and a spectrum gathered when the star occults the planet. The residual emission is that from the planet. Together these studies have proven immensely valuable, permitting the detection of numerous atoms and molecules including Na, H, H_2O, and CH_4, as well as clouds, and the direct determination of atmospheric temperatures and studies of temperature inversions. These studies are rendered all the more penetrating by the fact that the masses and radii of the planets are accurately known. More recently, planetary emission has been studied as a function of planetary longitude (Figure 4.16), permitting a direct constraint on the atmospheric dynamics, namely, the degree to which winds circulate energy from the irradiated dayside to the night side.

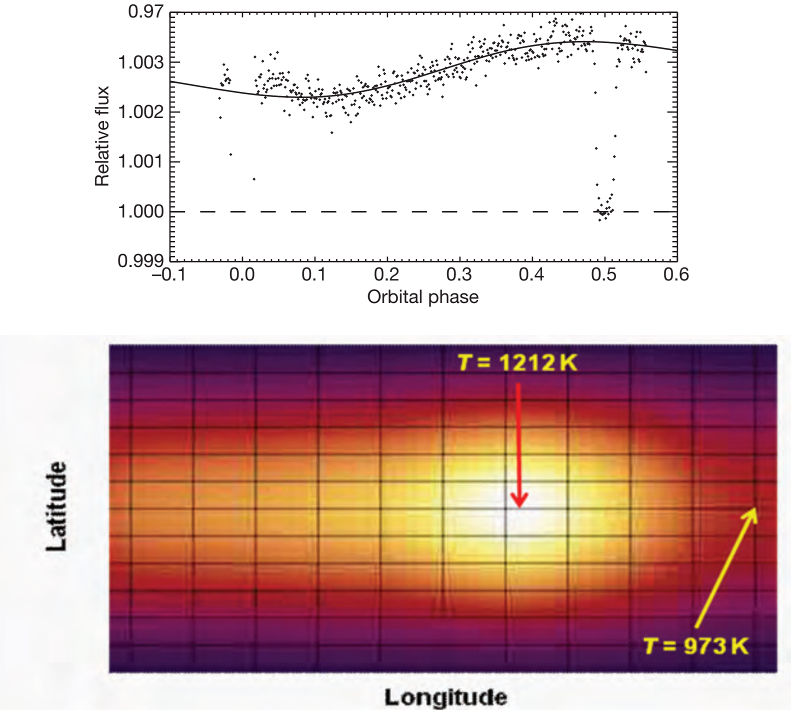


FIGURE 4.16 A map of the temperature over the surface of the exoplanet HD 189733b. *Upper:* The light curve at mid-infrared wavelengths, showing the passage of the planet in front of (phase 0) and then behind (phase 0.5) the star. Since the planetary emission is not observed at phase 0.5, the dashed line indicates the stellar flux, and any flux in excess of this value at other phases originates from the planet. *Lower:* The temperature map of HD 189733b derived from the light curve. The substellar point (corresponding to high noon on the planet) is in the center of the diagram. The hottest point on the planet (indicated with the red arrow) is eastward of the substellar point, indicating the action of strong super-rotating winds on the planet. SOURCE: Reprinted by permission from Macmillan Publishers Ltd.: *Nature*, H.A. Knutson, D. Charbonneau, L.E. Allen, J.J. Fortney, E. Agol, N.B. Cowan, A.P. Showman, C.S. Cooper, and S.T. Megeath, A map of the day-night contrast of the extrasolar planet HD 189733b, *Nature* 447:183-186, 2007.

These techniques hold great promise for the future, and indeed the panel foresees a fast-track opportunity to use these methods to search for atmospheric biomarkers such as molecular oxygen. These methods have been developed for HST and Spitzer, and surely will prove even more powerful with JWST. The panel encourages the following:

- A detailed study of the optimal observing methods with JWST, and if these studies confirm their feasibility, the acquisition of very high signal-to-noise-ratio spectra of hot Jupiters and even-lower-mass planets that are predominantly ice and rock in composition. An extension of the warm phase of Spitzer would bridge the gap, permitting the continued study of exoplanet atmospheres until the launch of JWST.

The methods described above will be limited to short period planets for which transits are likely and frequent. The natural complement is direct imaging, in which the faint light from the orbiting planet is spatially separated from that of the star. These efforts are buoyed by the recent images of two nearby, potentially, planetary systems. Moreover, both systems have debris disks, permitting an unprecedented glimpse at the early phases of planetary systems orbiting stars more massive than the Sun. The panel encourages both of the following:

- The development of extreme-contrast-imaging capabilities on the currently available 8- to 10-m-class telescopes, as well as their counterparts on the next generation of extremely large observatories. Space-based coronographic spectroscopy may also play an important role in this effort.

A particularly compelling avenue would be to image planets for which dynamical mass estimates can be obtained through radial velocities or astrometry, providing an essential calibration of models that predict the emission from planets as a function of mass and age. The current range of these theoretical tracks varies greatly, depending upon differing assumptions of the starting conditions and the rate of cooling, making it currently unreliable to estimate masses when only brightness measurements are available. The direct spectroscopic study of radial-velocity-detected exoplanets that orbit at distances of several astronomical units from nearby stars will require contrast ratios of 10^{-9} in visible light and 10^{-6} at infrared wavelengths.

Investments in Understanding Planetary Structures and Atmospheres

The anticipated rich yield of exoplanet discoveries and characterization efforts call out for the support of a vibrant research program in theory to interpret the data with respect to models of planet formation, evolution, dynamics, and planetary structures and atmospheres. This includes the following:

- The development of models to predict the outcome of theories of planet formation in terms of orbital elements, rates of occurrence, multiplicity, and bulk composition;
- The development of models of exoplanet atmospheres of varying composition and under differing levels of irradiation, including the study of clouds and photochemistry; and
- The tabulation of molecular-line data supported by laboratory astrophysics as needed, in support of these modeling efforts.

Theoretical Investments in Planetary Dynamics

The abundant harvest of exoplanet discoveries has made it clear that planetary systems are diverse. Commensurate with this observed richness is an increasingly complex array of planet-planet and disk-planet interactions that must be disentangled in attempts to infer the origins of these systems. Algorithmic advances and investments in parallel computing resources are needed to resolve delicate resonant interactions between planets and disks. The most lasting theoretical contributions will be those that can be validated by observations, many of which can be made most readily in the solar system—for example, in planetary rings, which afford miniature, accessible versions of protoplanetary disks, or in the Kuiper belt, which furnishes our own local debris disk.

Conclusions: Planetary-System Diversity

Table 4.4 summarizes the panel's conclusions on activities to address its third science question.

PSF 4. DO HABITABLE WORLDS EXIST AROUND OTHER STARS, AND CAN WE IDENTIFY THE TELLTALE SIGNS OF LIFE ON AN EXOPLANET?

The gleaming jewel in the crown of exoplanet discoveries would be a world that could sustain life, which would initiate the study of astrobiology on habitable exoplanets as an experimental science. Detection of life elsewhere in the universe remains a compelling goal supported by widespread fascination. However, this specific goal cannot be achieved with only incremental steps in the current areas of progress. The Kepler mission promises the sensitivity to detect Earth-size planets orbiting Sun-like stars at 1 AU. With concerted follow-up observations, the mass, radius, and density of these planets will be known, but precious little, if any, knowl-

TABLE 4.4 Panel's Conclusions Regarding the Study of Planetary-System Diversity

Technique	Requirements
Radial velocity (RV)	A target of 0.2 m sec^{-1}, requiring novel wavelength calibration methods (gas cells, laser combs) Substantial expansion of currently available observing time on 4- to 10-m-class telescopes for Kepler follow-up and other RV survey
Precision photometry/ spectroscopy	Microlensing surveys: long-duration near-IR coverage of at least 10^8 stars, with resolution <0.3 arcsec, and PSF stability <10% of FWHM over >1 month; JWST primary/secondary transit spectroscopy, 10^{-4}-10^{-5} of host-stellar signal
Direct imaging astrometry	Image contrast ~10^{-9} in optical, 30 mas—1 arcsec from host star using Gaia; other space mission with 0.1 mas—mission relative positional accuracy
Theory, numerics	Advances in planet/brown dwarf atmospheric models permitting inclusion of dynamics (clouds, zonal flows, etc.) and chemistry Studies of planetary dynamics, resonances; major computational resources required

NOTE: Acronyms are defined in Appendix C.

edge, of their detailed composition, surface, or atmosphere will exist. One can easily imagine these planets to be just like Earth, or completely unlike Earth: one need not look beyond Earth and Venus to know that similar worlds can differ markedly in their habitability. The faintness of the Kepler sample will prohibit determination beyond such imaginings. However, Kepler will observe a sufficiently large sample to measure accurately the fraction $\eta_\oplus$ of stars that host Earth-like planets. The section below titled "PSF Discovery Area—Identification and Characterization of a Nearby Habitable Exoplanet" discusses a fast-track approach to enabling some studies of habitable worlds within the coming decade, but this approach is limited to stars that are very different from the Sun. To truly understand life's origins and to place the life on Earth in context, a new capability is required to search for biomarkers on a twin of Earth orbiting a nearby Sun-like star. This challenge will require sustained investment in technology development and supporting scientific areas.

How Could We Discover the Planets and Characterize Their Atmospheres?

In *Astronomy and Astrophysics in the New Millennium*[3] (AANM), the previous decadal survey committee identified three important precursors to enable a successful Terrestrial Planet Finder (TPF) mission: measurement of the abundance of Earth-like planets, $\eta_\oplus$; measurement of the level of exozodiacal light in the range 10-100 times the solar system's zodiacal light ("zody"); and a space interferometry mission to discover Earth-mass planets.

The Kepler mission is meeting its required photometric accuracy and is well

---

[3] National Research Council, *Astronomy and Astrophysics in the New Millennium*, National Academy Press, Washington, D.C., 2001.

on the way to delivering the first of these pivotal results. The panel concludes the following:

- It is essential that sufficient resources be devoted to the analysis and follow-up of Kepler photometry through ground-based observations, to provide an accurate and definitive estimate of $\eta_{\oplus}$.

Figure 4.17 compares the expected strength of the exozodiacal light to the sensitivity of various observatories. AANM envisioned that the exozodiacal measurements would be accomplished by SIRTF (i.e., Spitzer), the Keck Interferometer (KI), and the Large Binocular Telescope Interferometer (LBTI). Spitzer has achieved its sensitivity goals but did not reach the requisite level for the exozody TPF precursor requirement owing to the extreme accuracy with which the background must be removed and stellar photospheres must be modeled. Measurements with unresolved IR/submillimeter excess from Herschel will provide addi-

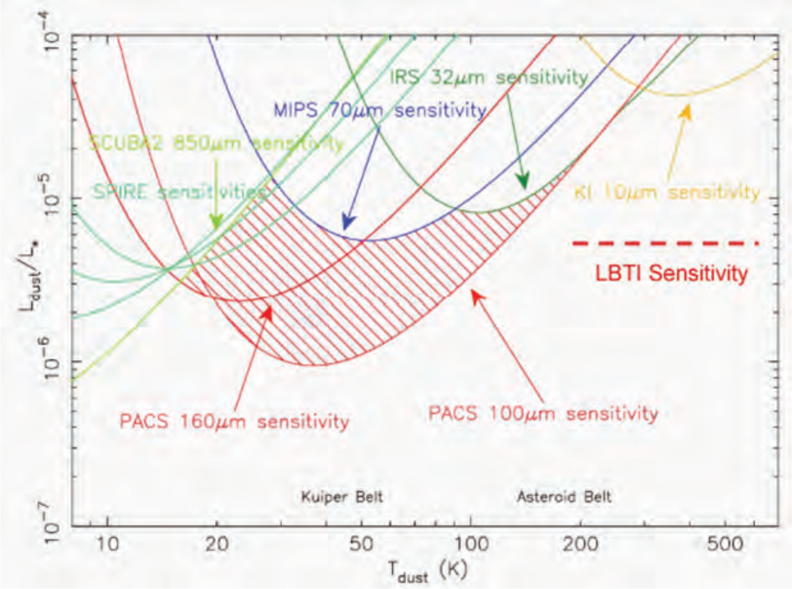


FIGURE 4.17 The sensitivity limits to dust around other stars achieved with current state-of-the-art techniques. MIPS and IRS are instruments on Spitzer, and PACS and SPIRE are instruments on Herschel; SCUBA is a ground-based submillimeter camera on JCMT. These are all sensitive to dust levels only several hundred to many thousand times that of the cool Kuiper belt. LBTI and KI are ground-based interferometers and are sensitive to only tens to hundreds of times the zodiacal-dust equivalent. The labels for Kuiper and asteroid belts indicate the approximate level of the zodiacal light. SOURCE: Exoplanet Task Force, *Worlds Beyond: A Strategy for the Detection and Characterization of Exoplanets*, Washington, D.C., 2008.

tional constraints, but, ultimately, resolved measurements are required to eliminate the stellar modeling uncertainty. The KI has achieved the nulling of light from the central star but is unlikely to deliver results at the requisite <10-zody level (i.e., 10 times the solar system zodiacal light). The LBTI has not yet been commissioned, but with lower background and higher throughput it will deliver better sensitivity than KI; a low-background GSMT with shearing/nulling interferometry would perform much like LBTI. The panel concludes as follows:

- Substantial improvement with ground-based instrumentation might require exploiting the cold, dry, and stable atmosphere of the high Antarctic plateau with a nulling interferometer. Owing to the dramatic benefits in IR background and improved seeing, 10-zody sensitivity could be achieved with a relatively modest instrument.
- Space-based coronagraphy might be the only means to obtain meaningful constraints on the exozodiacal light, but even here relatively large apertures are required to constrain emission at radii of 1 AU for any but the very nearest stars. The problem is that dust is likely sculpted into belts, as in the solar system, so extrapolation inward from large radii is unjustified.
- The observation of exozodiacal light in candidate targets will eliminate a risk to the scientific viability of direct-detection missions.

The exploration of additional approaches to this problem is warranted to retire the risk of a single-point failure on this scientific path.

The Astronomy and Astrophysics Advisory Committee's Exoplanet Task Force (ExoPTF) identified contingent strategies to achieve characterization of Earth analogs based on results of $\eta_\oplus$ and exozody measurements.

- Central to this strategy was the launch of a space-based astrometric mission to identify specific targets for a planet characterization mission. A planet-finding astrometric mission will provide specific stars with known planets of measured mass.

The knowledge obtained from such a mission would almost certainly have a major impact on the design of a direct-detection mission and would provide certainty that targets exist within the survey sample, rather than simply that it is statistically likely. Moreover, studies of the spectra of these worlds will be far more penetrating if the masses are known.

The recent proliferation of candidate mission concepts offers multiple windows to change fundamentally the present vision of planetary systems. Technology development through the next decade is needed to bring the varied mission concepts to a level at which the scientific priority of competing approaches can be assessed.

Which Measurable Characteristics Define Habitability?

The ubiquity of the organized chemistry that is called life remains a subject of intense interest and speculation. The current working definition of habitability is that liquid water be stable, energy be available, and certain chemical species have non-equilibrium concentrations. Of these factors, the stability of liquid water is the most important diagnostic for planetary habitability. The complexities of this subject are beyond the scope of this report but are explored in great detail in the ExoPTF report[4] and the NRC report *The Limits of Organic Life in Planetary Systems.*[5] Ultimately we won't know how parochial we are until we find the answer.

Given the ultimate focus on finding planets as sites of biology, careful development and understanding of the limitations of biosignatures are required. *This will be an activity that requires interdisciplinary collaboration with theory, astrochemistry, biology, and planetary science.*[6] The interpretation of molecular oxygen or ozone (Figure 4.18) as biological in origin requires the presence of another signature that rules out an abiotic origin. Detection of the red edge of spectra of vegetation would be a particularly remarkable discovery.

During the next decade, observations of exoplanets should be used to direct concepts for future missions. A mission to study an Earth analog will rely on the comprehensive and concerted synergy of the understanding of planetary systems in the universe. To launch a mission that minimizes scientific risk, the efforts of many astronomers will need to be synthesized on questions of $\eta_\oplus$, exozodiacal emission, and global understanding of planetary formation and evolution.

Kepler will yield quantitative results on $\eta_\oplus$, which are essential for estimating what number of stars will need to be surveyed. High-precision (0.1-μas) astrometry can identify specific host stars, decoupling the task of detection from that of characterization and reducing the risk of scientific failure for a mission that may be correctly scoped based on measurements of $\eta_\oplus$ but unlucky in the distribution of planets around the target stars. Moreover, astrometry will provide dynamical estimates of planetary masses that will be essential to interpreting the spectra of these distant worlds.

---

[4] Exoplanet Task Force, *Worlds Beyond: A Strategy for the Detection and Characterization of Exoplanets*, Washington, D.C., 2008.

[5] National Research Council, *The Limits of Organic Life in Planetary Systems*, The National Academies Press, Washington, D.C., 2007.

[6] For a more extensive discussion of such cross-disciplinary opportunities, see National Research Council, *The Astrophysical Context of Life*, The National Academies Press, Washington, D.C., 2005.

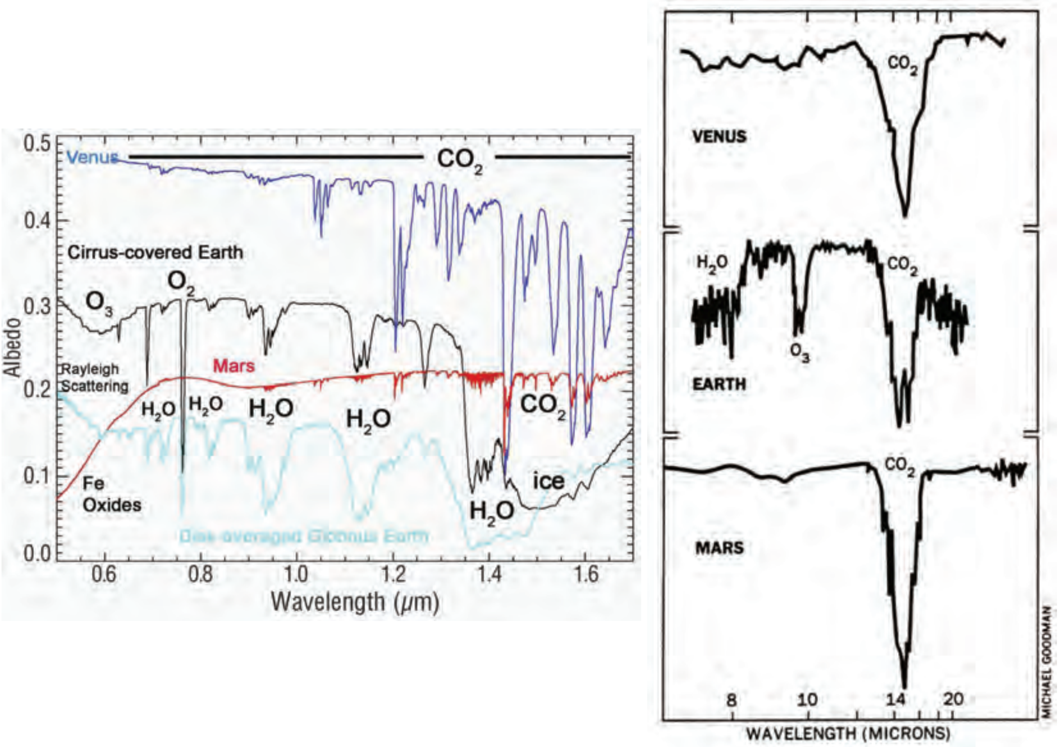


FIGURE 4.18 Spectral signatures of habitability? *Left:* Full-resolution synthetic disk-averaged albedo spectra of Venus, Earth, and Mars. Synthetic Earth spectra are shown for both uniform high cirrus cloud cover, and as a fit to Earthshine observations of the gibbous Earth. The Venus spectrum was approximated to a disk average and has been multiplied by 0.6 to fit the plot. The Mars and Earth spectra are disk averages of three-dimensional spatially and spectrally resolved Virtual Planetary Laboratory models of Earth and Mars. For the observed Earth, which is ocean dominated with relatively little cloud cover, the Rayleigh scattering (0.45-0.6 μm) is pronounced, but the ozone is less apparent. The ozone absorption is much more pronounced for Earth with cloud cover, increasing the difficulty of identifying the Rayleigh scattering component. *Right:* Thermal-infrared spectra of Venus, Earth, and Mars. SOURCE: Exoplanet Task Force, *Worlds Beyond: A Strategy for the Detection and Characterization of Exoplanets,* Washington, D.C., 2008.

Conclusions: Discovering Habitable Exoplanets

Table 4.5 summarizes the panel's conclusions about activities to address its fourth science question.

TABLE 4.5 Panel's Conclusions Regarding Discovery of Habitable Planets

Technique	Requirements
Radial velocity (RV)	0.2 m $\sec^{-1}$, or better, requiring novel wavelength calibration methods (gas cells, laser combs) Substantial expansion of currently available observing time on 4- to 10-m-class telescopes for Kepler follow-up and other RV survey, consistent with ExoPTF report[a]
Precision photometry/ spectroscopy	$\eta_\oplus$ (Kepler) 10-zody sensitivity to exozodiacal emission, on as small a scale as possible JWST primary/secondary transit spectroscopy, 10^{-4}-10^{-5} of host-stellar signal
Astrometry	0.1-µas relative positional accuracy
Theory, numerics	Advances in planet/brown dwarf atmospheric models permitting positive identification of biomarkers

NOTE: Acronyms are defined in Appendix C.

[a] Exoplanet Task Force, *Worlds Beyond: A Strategy for the Detection and Characterization of Exoplanets,* Washington, D.C., 2008.

PSF DISCOVERY AREA—IDENTIFICATION AND CHARACTERIZATION OF A NEARBY HABITABLE EXOPLANET

The question of whether or not there exist habitable worlds besides Earth is one of the deepest and most abiding ever posed by humanity. The answer would have ramifications that extend far beyond astronomy. One possibility is that no other inhabited world exists within the reach of modern astronomical observatories and thus that humanity is effectively alone in the cosmos. Another is that ours is but one of many planets where life has flourished. The knowledge of either reality would profoundly affect the scientific fields of astronomy, planetary science, prebiotic chemistry, and evolutionary biology, but it would also inform each individual's sense of place in the universe. In its discussion above of science question PSF 4, the panel outlines a strategy to search for life on Earth-like planets orbiting in the habitable zones of Sun-like stars. That plan will require a dedicated effort spanning at least two decades. The panel believes that the effort is justified by the transformative power of the question that it seeks to answer. It is natural to wonder what shorter-term opportunities exist that could sustain our efforts in the years ahead.

There is a remarkable, although speculative, fast-track opportunity to finding and characterizing habitable exoplanets: namely, the search for large Earth-analogs (super-Earths) orbiting nearby low-mass, M-dwarf stars, the most common kind of star in the galaxy. This opportunity capitalizes on the most successful techniques for detecting and characterizing exoplanets from the past decade; as such it was unforeseen at the time of the previous astronomy and astrophysics decadal survey. Changes can now be detected in stellar radial velocities as small as 1 m $\sec^{-1}$ and decreases in starlight as small as 0.3 percent due to planetary transits. Applied to Sun-like stars, these sensitivities cannot detect solid planets orbiting in their stellar habitable zones. However, applied to low-mass stars, they are sufficient for the detection of habitable super-Earths. Moreover, transiting-based follow-up work has yielded a rich set of studies of the structures, compositions, and atmospheres of exoplanets and has fueled the nascent field of comparative exoplanetology. Here again, these techniques are not feasible for the study of habitable, Earth-like planets orbiting Sun-like stars, as the signal of such small planets would be overwhelmed by the photon flux from the star. They *would* be feasible if applied to nearby M-dwarf stars. It is noted that this opportunity was highlighted in the recent report of the ExoPlanet Task Force to the Astronomy and Astrophysics Advisory Committee.[7]

This discovery opportunity differs from that related to science question PSF 4 above in several respects. The primary distinction is the scientific question of whether M-dwarfs can indeed form habitable-zone super-Earths and whether or not the properties of M-dwarfs would preclude life on those worlds. On this last point, there are many ways in which M-dwarfs appear unwelcoming to life as we know it. These include their large UV and X-ray fluxes at young ages and the expectation that such planets would be tidally locked; these questions are discussed in more detail later. If indeed it is learned that M-dwarf planets sustain life despite these very different conditions, a fundamental piece of information will have been learned about the robustness of life in the cosmos. If, however, it is found that M-dwarf planets are lifeless, an important bound will have been placed on the conditions in which life can flourish. Clearly the present theoretical understanding of the origin of life is not remotely sufficient to predict whether M-dwarf planets will be inhabited, but the coming 5 years afford the opportunity to move this question into the realm of observational study. The panel also notes two practical ways in which this path differs from that related to PSF 4: First, the methods are based on separating the light of the planet from that of the star in time (through transit and occultation spectroscopy), as opposed to detecting spatial separation of planetary and stellar light (through imaging and interferometry). Second, the timescale for discovery is 5 years (as opposed to 20 years). It is this last point—the idea that the study of life outside the solar system could begin by 2014—that perhaps provides the greatest motivation for this avenue as a separate path from that discussed above in relation to science question PSF 4.

The Small Star Opportunity

Roughly 70 percent of stars in the immediate solar neighborhood are M-dwarfs. For a typical M-dwarf—spectral type M4V, corresponding to a radius and mass only 25 percent of the solar values—the habitable zone is located only 0.07 AU

---

[7] ExoPlanet Task Force, *Worlds Beyond: A Strategy for the Detection and Characterization of Exoplanets,* Washington, D.C., 2008.

from the star (Figure 4.19). As a result of the small planet-star separation, small stellar radius, and low stellar mass, the discovery of a large terrestrial planet orbiting within the stellar habitable zone could be achieved using only current radial-velocity and photometric precision, provided that this precision, demonstrated for Sun-like stars, can be obtained on these M-dwarfs. In particular, the transits of an Earth-size body would be 0.13 percent deep, and the radial-velocity signal would be 1.4 m sec^{-1}. Moreover, the transits would occur more frequently—the orbital period would be a mere 15 days—and would be thrice as likely as a habitable-zone planet orbiting a Sun-like star. Perhaps most intriguingly, the planet-to-star contrast ratio in the Rayleigh-Jeans limit would be 0.012 percent, facilitating the study of the infrared spectrum of the planet by occultation spectroscopy. For an M8V primary (with a mass and radius only 10 percent of the solar values), the situation is even more favorable: the habitable zone lies at 0.017 AU, the transits would be 0.84 percent deep and would recur every 2.5 days, and the reflex radial-velocity amplitude would be 4.4 m sec^{-1}. Moreover, the planet-to-star contrast ratio would be 0.11 percent, a helpful increase over the contrast ratio for a super-Earth of ~0.001 percent.

Several authors have considered the signal-to-noise ratio for a mock observing campaign of such a system with various instruments on JWST. These authors conclude that it would require a major investment of JWST time to achieve the requisite signal-to-noise ratio to be able to infer the presence of biomarkers but that it is possible for a super-Earth: an example is shown in Figure 4.20. The prospect that an observatory currently under construction would have the sensitivity to

FIGURE 4.19 Habitable zones around solar-type and M-dwarf stars. The shaded regions denote the range of distances from a G2V star (left) and an M5V star (right) for which the equilibrium temperature of the planet is greater than 0°C and less than 100°C, and hence for which water might be liquid at the surface. This naïve definition ignores the greenhouse effect, which maintains the surface temperature of Earth at roughly +30°C above the equilibrium temperature. Earth's orbit is indicated by the dashed circle, and the orbit at which a planet would receive the same amount of energy per unit area and unit time is shown as a dashed circle in the right plot. SOURCE: J. Irwin, Harvard-Smithsonian Center for Astrophysics.

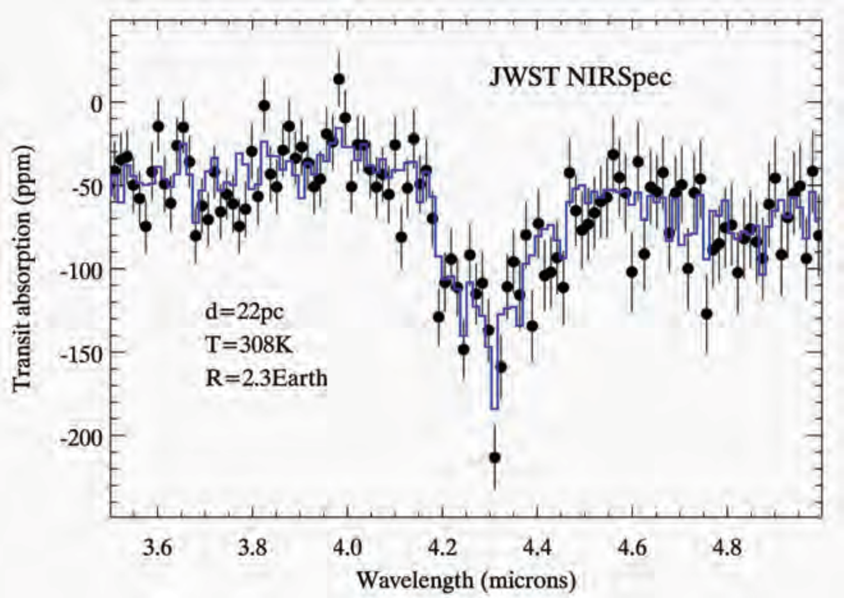


FIGURE 4.20 A simulated JWST NIRSpec observation (shown as black points) of a habitable-zone super-Earth with a radius of 2.3 $R_\oplus$ and a temperature of 308 K located at 22 pc. The absorption feature due to carbon dioxide is detected with a signal-to-noise ratio of 28 for 85 hours of data in transit and an equal number outside of transit. Both the data and the model (blue curve) are shown at a sampling of λ/300, which would support a spectral resolution of 100 assuming three samples per optical resolution element. SOURCE: Courtesy of D. Demming, personal communication. Adapted from D. Deming, S. Seager, J. Winn, E. Miller-Ricci, M. Clampin, D. Lindler, T. Greene, D. Charbonneau, G. Laughlin, G. Ricker, D. Latham, and K. Ennico, Discovery and characterization of transiting super Earths using an all-sky transit survey and follow-up by the James Webb Space Telescope, *Publications of the Astronomical Society of the Pacific* 121:952-967, 2009.

search for atmospheric biomarkers in a terrestrial planet is tremendously exciting, and a spectrum with the fidelity to identify atmospheric composition is clearly at the extremes of the reach of JWST. The more modest goal of simply obtaining a brightness temperature measurement would already be exceptionally interesting. An extension to this would be the determination of temperature as a function of planetary longitude, since a small day-night contrast would be strong evidence for the existence of an atmosphere enshrouding this distant rocky world.

There are two primary reasons why such M-dwarf planets, if they exist, might not be habitable. First, the proximity of the habitable zone places it within the tidal-locking radius, implying that the planet *could* have fixed day and night sides. This may in turn lead to a freeze-out of volatiles on the night side, at least for thin atmospheres. However, some planets, such as Mercury, may avoid this fate by becoming trapped in spin-orbit resonances, and in any event a thick atmosphere or deep oceans could serve to ensure a moderate day-night temperature contrast. Second, all M-dwarfs are chromospherically active, and their intense stellar winds, combined with a weak or absent magnetic field on the planet owing to its slow rotation, may cause the atmosphere to be stripped away.

The panel takes these points as caveats; the only sure means for progress is measurement. The task of predicting habitability is surely no less complicated than that of modeling planet formation, and thus it bears noting that on the latter question, imagination—prior to the first exoplanet discoveries in 1995—proved wholly inadequate. Even a null result would be quite important: if numerous M-dwarf terrestrial planets are discovered that are similar to Earth in bulk composition, age, and insolation, yet are uninhabited, a fundamental fact about the requirements for life will have been learned.

The Hunt for Inhabited Worlds, 2010-2020

In order to pursue the exciting opportunity described above, the panel recommends the following:

- Resources to export the radial-velocity precision of 1 m sec^{-1}, currently achieved for Sun-like stars, to large numbers of M-dwarfs;
- The development of near-IR spectrographs or red optical (700 nm < λ < 1,000 nm) charge-coupled-device-based spectrographs, both suited to capitalize on the spectral regions where M-dwarfs are brightest;
- The development of new calibration techniques, such as novel gas absorption cells or laser frequency combs, and possibly very high resolution infrared spectrographs on large telescopes;[8] and, as noted earlier,
- A substantial increase in the amount of observatory time to undertake radial-velocity surveys, which was also a key finding for the study of Sun-like stars (science question PSF 3).

The community should pursue a transit survey of the closest 10,000 M-dwarfs for Earth-size planets (or, if conditions dictate, planets with a radius twice Earth's value) orbiting within the stellar habitable zone. This may require the following:

- Large-scale ground-based synoptic surveys, or a space-based survey, but the survey by necessity will need to cover a large fraction of the sky;

---

[8] Internal precision of 10 m/sec has recently been reported for near-infrared radial velocity measurements of a very low mass star. See J.L. Bean, A. Seifahrt, H. Hartman, H. Nilsson, A. Reiners, S. Dreizler, T.J. Henry, and G. Weidemann, The proposed giant planet orbiting VB 10 does not exist, *Astrophysical Journal Letters* 711:L19, 2010.

- A census of M-dwarfs out to 50 pc, and stellar astrophysical studies of these stars to understand their mass, radius, metallicity, distance, and if possible age; and
- Detailed study of the atmospheric activity of M-dwarfs, to understand the limitations that such variability will impose on searches for planetary companions.

The community should also undertake intensive studies to determine the best observing practices for the study of such objects with JWST, and to determine if similar studies can be pursued with large ground-based telescopes. Accurate estimates of the likely signal-to-noise ratios from JWST (e.g., see Figure 4.20) will be crucial as preparations for the interpretation of such data are made. In conjunction with this work, theoretical expertise in the following areas needs to be fostered:

- Modeling, in collaboration with the geophysics community, of the physical structures of Earth-size and super-Earth exoplanets with a range of composition, and an emphasis on the interpretation of anticipated data. Such studies may require laboratory investigations into the equations of state of materials that could be significant components in super-Earths.
- Modeling the atmospheres of habitable Earths and super-Earths. Such studies will include photochemistry and cloud modeling and will require the development of detailed molecular databases.

These studies will likely be interdisciplinary, requiring expertise at the interface of astronomy, planetary science, and biology.

JWST is currently planned for launch in 2014, and it will have a finite lifetime. Because the spectroscopic studies envisioned here will likely extend over several years, it is crucial that the target planets be identified prior to or soon after the launch of the observatory. There is considerable expertise in the exoplanet community with respect to what needs be accomplished to extend the current radial-velocity and transit-survey precision to a sufficient number of M-dwarfs stars to enable this very exciting path toward detecting the first habitable worlds and undertaking a spectroscopic study of its atmosphere to search for biomarkers.

Conclusions: Habitable Planets Around M-Dwarfs

Table 4.6 summarizes the panel's conclusions on activities to address its selected general area with unusual discovery potential.

TABLE 4.6 Panel's Conclusions Regarding Detection of Habitable Planets Around M Dwarfs

Technique	Requirements
Radial velocity (RV)	1 m sec^{-1}, on M dwarfs, requiring red-near-IR spectroscopy and new wavelength calibration methods (gas cells, laser combs) Substantial expansion of currently available observing time on 4- to 10-m-class telescopes for M dwarf survey
Precision photometry/spectroscopy	Transit survey (precision 10^{-4}-10^{-5} of host-stellar signal) for 10^4 closest M dwarfs Characterization of activity of M dwarfs within 50 pc JWST primary/secondary transit spectroscopy, 10^{-4}-10^{-5} of host-stellar signal

NOTE: Acronyms are defined in Appendix C.

5

Report of the Panel on Stars and Stellar Evolution

SUMMARY

The science frontier for stars and stellar evolution is as close as the Sun and as distant as exploding stars at redshift 8.3. It includes understanding processes of exquisite complexity that connect the rotation of stars with their magnetic fields and areas of nearly total ignorance about phenomena that have been imagined but not yet observed, such as accretion-induced collapse. Because astronomers understand stars well, they have the confidence to use them as cosmic probes to trace the history of cosmic expansion; but because this understanding is not complete, there is much to learn about the subtle interplay of convection, rotation, and magnetism or the not-so-subtle violent events that destroy stars or transform them into neutron stars or black holes. Although the topics of stars and their changes over time comprise great chunks of introductory astronomy textbooks, and although the tools for these investigations are tested and sharp, many of the simplest assertions about the formation of white dwarfs, mass loss from giant stars, and the evolution of binary stars are based on conjecture and a slender foundation of facts.

The future is promising. X-ray and radio observations allow astronomers to probe stars where strong gravity is at work. These settings stretch the understanding of fundamental physics beyond the range of laboratory investigations into unknown areas of particle interactions at higher densities than those produced in any nucleus or terrestrial accelerator. By testing three-dimensional predictions against the evidence, more-powerful computers and programming advances put astronomers on the brink of understanding the violent events that make stars explode and collapse. Advances in laboratory astrophysics lead to a better understanding of the underlying nuclear, atomic, and magnetohydrodynamic (MHD) processes. New technology for optical and infrared (IR) spectropolarimetry and for interferometry open up the possibility of seeing magnetic fields and resolving the disks of stars. When well-sampled imaging is coupled to powerful systems for processing vast quantities of data sampled over time, subtle features of stellar interiors can be inferred. Similarly, rare and rapid transients that have eluded surveys to date will surely be found, and may be connected with gravitational waves.

These advances are certain to open up a new and unexplored world of investigation on timescales from seconds to decades. In this report, the Astro2010 Science Frontiers Panel on Stars and Stellar Evolution sketches the most fertile opportunities for the coming decade in the field of stars and stellar evolution. The panel is confident that it will prove a fruitful decade for this field of astronomy, with the resolution of today's questions producing many new problems and possibilities.

As requested by the Astro2010 committee, the panel formulated its report around four science questions and one outstanding discovery opportunity. The panel is under no illusion that this short list is complete: the field is so rich that there will surely be advances in areas not emphasized here. The panel does, however, have every reason to believe that these questions capture some of the most promising areas for advances in the coming decade. The four questions and discovery opportunity are as follows:

1. How do rotation and magnetic fields affect stars?
2. What are the progenitors of Type Ia supernovae and how do they explode?
3. How do the lives of massive stars end?
4. What controls the mass, radius, and spin of compact stellar remnants?
5. *Unusual discovery potential:* time-domain astronomy—in which the technology on the horizon is well matched to the many timescales of stellar phenomena.

The subsections below summarize the main points.

How Do Rotation and Magnetic Fields Affect Stars?

There's an old chestnut about a dozing theorist at the weekly colloquium who opens his eyes at the end of every talk and rouses himself to ask, to great approbation for his subliminal understanding, "Yes, all very interesting, but what about rotation and magnetic fields?"

Astronomers are now in a position to address this question in a serious way. In the Sun, the effects are visible; in many other stars they are likely to be much more important. It is not sufficient to think of rotation and magnetism as perturbations on a one-dimensional star. These are fundamental physical phenomena that demand a three-dimensional representation in stars.

Astronomers are poised to learn how stars rotate at the surface and within and how that rotation affects mass loss and stellar evolution. They seek a better understanding of how magnetic fields are generated in stars across the mass spectrum and of how these fields power the chromospheres and coronae that produce observed magnetic activity. Finally, the origin of highly magnetized main sequence stars, in which surface fields approach 10^4 gauss, remains mysterious, and the investigation of these stars promises to shed light on the star-formation process that produced them as well as on the origin of even more highly magnetized compact objects.

The prospects for progress on this question in the next decade stem from the emergence of greatly improved tools for measuring magnetic fields from polarization, for resolving the atmospheres of some stars with interferometry, for probing the interiors of stars through their vibration spectra, and for extending observations into X-rays and gamma rays. When combined with more thorough understanding of the static and dynamic properties of magnetic atmospheres, astronomers will learn how stellar atmospheres really work and how rotation and magnetism affect the evolution of stars.

What Are the Progenitors of Type Ia Supernovae and How Do They Explode?

Many lines of evidence converge on the idea that Type Ia supernovae are thermonuclear explosions of white dwarfs in binary systems. Because of their high luminosity, and with effective empirical methods for determining their distances from light-curve shapes, Type Ia supernovae have acquired a central role not just in stellar astrophysics but also in tracing the history of cosmic expansion and in revealing the astonishing fact of cosmic acceleration. Because this result points to a profound lack in the understanding of gravitation, a problem right at the heart of modern physics, completing the astronomical story of Type Ia supernovae is a pressing priority for the coming decade.

First of all, the provenance of the exploding white dwarfs seen in other galaxies is not known with certainty. The prevailing picture is that Type Ia explosions arise in binary systems in which a white dwarf accretes matter until it approaches the Chandrasekhar limit, simmers, and then erupts in a thermonuclear flame. But it is not known how this picture is affected by chemical composition or age, two essential ingredients in making a precise comparison of distant events with those nearby. Events that are precipitated by the merger of two white dwarfs are not excluded. The Type Ia supernovae in star-forming galaxies and in ellipticals are at present treated in the same way, but this is the result of small samples, not of evidence that they should be analyzed together. More broadly, these gaps in knowledge illuminate the need for a better understanding of the evolution of interacting binary stars, which are responsible for a variety of crucial, yet poorly understood, phenomena.

It can be expected that both theory and improved samples will place this work on a firmer foundation. The complex, turbulent, unstable nuclear flame that rips through the star and incinerates its core is at the present time impossible to compute fully in three dimensions. But the prospects for achieving that goal in the coming decade are intriguing. Samples today amount to a few hundred objects at low redshift and similar numbers beyond redshift 0.5. Much larger and significantly more uniformly discovered samples are coming soon through targeted aspects of time-domain surveys. They will create a much tighter connection between chemistry, binary stellar populations, and supernova properties. Inferences on dark-energy properties are at present limited by inadequate understanding of the intrinsic properties of Type Ia supernovae as distorted by interstellar dust. Observing in the rest-frame infrared will expand the basis for comparing observations with computations and provide more accurate measurements of dark energy.

How Do the Lives of Massive Stars End?

Ninety-five percent of stars will end their lives as white dwarfs. For the rest, stellar death is spectacular and dramatic: these massive stars can explode as supernovae, emit gamma-ray bursts (GRBs), and collapse to form neutron stars or black holes. The elements that they synthesize and eject become the stuff of other stars, planets, and life. The energy and matter that they produce are crucial for the evolution of galaxies and clusters of galaxies.

Despite a basic understanding that gravity is the energy source for these events, a clear connection between the mass and metallicity of the star that collapses, the nature of the collapse and explosion, and the properties of the compact remnant remain mysterious. The rotation of the progenitor and its mass loss, areas of uncertainty highlighted in the panel's first question, seem to be essential aspects of the link between core collapse supernovae and GRBs. Exploring these frontiers will require continued thoughtful analysis and full-blown first-principles calculations.

The full range of outcomes for stellar deaths may not be well represented in current observational samples. Deeper, faster, wider surveys will surely detect rare or faint outcomes of stellar evolution that have not yet been seen. These could include pair-instability supernovae and other types of explosions that have only been imagined.

The role of massive stars in the evolution of the universe is coming into view. The fossil evidence of massive stars is embedded in the atmospheric abundance patterns of our galaxy's most metal-poor stars. The most distant object measured so far is a gamma-ray burst, presumably from a massive star, at redshift 8.3. In the coming decade, the direct observation of the first generation of stars, which are predicted to be exceptionally massive, will be within reach with the James Webb Space Telescope (JWST).

Massive stars could be the source of gravitational wave signals, a neutrino flash, or nuclear gamma-ray lines. All of these novel messages from stars are within reach for very nearby cases, and could be exceptionally important in shaping the future understanding of the deaths of massive stars.

What Controls the Mass, Radius, and Spin of Compact Stellar Remnants?

Unanswered questions about the magnetic fields and rotation of stars carry through to similar questions about the exotic remnants that they leave behind as neutron stars and black holes. These are exceptional places in the universe where understanding of physics is extended beyond the reach of any laboratory.

For example, the equation of state for nuclear matter sets the relation between mass and radius for neutron stars. Theoretical understanding of the forces at work is uncertain where the density exceeds that of the densest nuclei. Prospects for measuring masses for radio pulsars and neutron star radii from X-ray techniques promise a glimpse into the strange world of quantum chromodynamics and the possibility of hyperons, deconfined quark matter, or Bose condensates.

The spins of neutrons stars and black holes are rich areas for future work. It is known that millisecond pulsars are spinning much faster than when the neutron stars were formed, and it is understood how accretion in a binary system can accelerate their rotation, but the mechanism that limits how fast these neutron stars can whirl is not known. The answer is expected to come from new pulsar surveys that are less biased against detecting the fastest pulsars. Similarly, there are now plausible measurements that imply that black holes are spinning rapidly. It seems very likely that these black holes are telling us about the conditions in which they formed, during the collapse of a massive star—one of the key points in the panel's third question. In the coming decade, X-ray spectroscopy should be a powerful technique for expanding the slender sample of spinning black holes, all identified in binaries.

Most stars surely become white dwarfs, but present understanding of the white dwarf mass for a main sequence star of a given initial mass is seriously incomplete. How stars lose mass is not understood well enough to predict which stars will become white dwarfs. Important details of the white-dwarf population remain unsolved and could lead to types of supernovae that have not yet been recognized. Large surveys will be powerful tools for finding these objects, making it possible to fill in these embarrassing gaps in understanding.

Discovery Area: Time-Domain Astronomy

For poets, stars are symbols of permanence. But astronomers know that this is not the whole story. Stars reveal important clues about their true nature by their rotation, pulsation, eclipses and distortions, mass loss, eruptions, and death. Across a wide range of timescales from seconds to years, stars are changing, and scientific knowledge has been obtained through narrow windows of time set by practical matters of telescope time, detector size, and the ability to sift the data. The panel foresees a rich flood of data from specialized survey instruments capable of exploring this new frontier in astronomy, across the electromagnetic spectrum. These instruments will provide new time-domain data, with the potential for major impact on stellar astronomy, ranging from the precise understanding of stars through seismological data and the periodicities that rotation produces, to the detection of rare transient events that have not yet been revealed in extant surveys. An example provides a glimpse of the excitement: wide, deep, and frequent surveys will be the way to find the electromagnetic counterparts of gravitational wave events. The broad problems of binary star evolution, about which so much is assumed and so little is known, can be sampled by such an undertaking, perhaps advancing the knowledge of the progenitors of thermonuclear supernovae from being a plausible story to becoming an established fact. The range of stellar phenomena that will be addressed with large, accessible, time-domain databases goes far beyond the four questions of the panel.

Summary of Panel's Conclusions

The conclusions of this panel report are summarized in Table 5.1.

INTRODUCTION

The advent of quantum mechanics and the study of nuclear fusion in the 1930s led to the first successful models for energy generation in stars and to a basic understanding of the distribution of stars in the Hertzsprung-Russell diagram. This work, along with the study of stellar atmospheres, laid the foundation for modern astrophysics. In just the past decade, the resolution of the solar neutrino problem demonstrated conclusively the presence of new neutrino physics and the accuracy of solar models originally developed in the 1960s. The success of this work relied in part on the approximate spherical symmetry of stars, which enabled accurate one-dimensional models of their structure.

The opening of new parts of the electromagnetic spectrum dramatically broadened astronomers' views of stellar phenomena, leading to a number of breakthrough discoveries. Newly discovered radio pulsars in binaries, including the unique double-pulsar system, provided some of the most stringent tests of general relativity. Advances in X-ray astronomy have led to new discoveries related to accreting neutron stars and black holes, compact remnants of massive stars, and energetic phenomena, such as coronae and flares, on normal stars. The characterization of brown dwarfs, cool low-mass objects at the border between stars and planets, was due to the development of new infrared surveys. And the

TABLE 5.1 Summary of Conclusions of the Panel on Stars and Stellar Evolution

	Question 1: How Do Rotation and Magnetic Fields Affect Stars?	Question 2: What Are the Progenitors of Type Ia Supernovae, and How Do They Explode?	Question 3: How Do the Lives of Massive Stars End?	Question 4: What Controls the Mass, Radius, and Spin of Compact Stellar Remnants?
Current and expected facilities	ATST, SDO, HST, 4-m and 8-m telescopes, Kepler, CoRoT, Gaia	PTF, PanStarrs-1, KAIT, PAIRITEL, JWST, Swift, HST, Chandra, 8- to 10-m telescopes	PTF, PanStarrs-1, Swift, NuSTAR, EVLA, 8- to 10-m telescopes, HST, GALEX, Chandra	EVLA, ALMA, LOFAR, Gaia FAST, LIGO, FRIB, Chandra, XMM, Suzaku, RXTE, Fermi
New facilities needed	High spatial and synoptic solar magnetometry; helio- and asteroseismology; OIR interferometry; OIR time-domain, large-FOV observations; high-resolution multiobject OIR spectroscopy; plasma physics experiments	OIR time-domain, large-FOV, high-cadence observations; precise IR follow-up; X-ray spectroscopy; 20- to 30-m telescope	Multiwavelength (radio to X-ray) time-domain, large-FOV, high-cadence observations; post-Swift GRB studies; X-ray spectroscopy; neutrino and gravitational wave observatories; 20- to 30-m telescope	Large-area decimeter-wavelength telescope; large-effective-area X-ray timing and spectroscopy; gravitational wave observatory
Crucial capabilities	Detailed solar and stellar studies of internal rotation and magnetism; surveys of stellar surface rotation, activity, magnetism, and mixing diagnostics; three-dimensional MHD simulations; pulsation theory; UV and X-ray spectroscopy	Panchromatic spectroscopy of a representative sample; large sample for finding diverse objects and correlations with environment; advanced three-dimensional simulation capability; progenitor surveys; nuclear cross sections	Large-scale three-dimensional simulations; discovery of broad range of transients and multiwavelength follow-up; nuclear data and oscillator strengths; abundance studies of extremely low metallicity stars	High-sensitivity X-ray timing and spectral observations of known neutron stars and black holes; laboratory measurements of nuclear equation of state; deep radio pulsar searches; star-cluster white-dwarf searches; gravitational wave detection of compact binary inspirals
Other priorities	Dedicated follow-up for inferring fundamental stellar properties; progress on abundance determinations; laboratory measurements of opacities; support for basic theory and computational astrophysics			

NOTE: Acronyms are defined in Appendix C.

conclusive determination that long-duration gamma-ray bursts are associated with the deaths of massive stars showed that the central collapse of some massive stars can produce relativistic jets.

In parallel with the continued exploration of the electromagnetic spectrum, high-energy neutrino and gravitational-wave views of the universe will likely be unveiled in the next decade. Observational techniques such as astrometry, interferometry, and time-domain surveys, all of which are well suited to studies of stellar phenomena, will reach maturity. Given this influx of new data, there are tremendous opportunities in stellar astrophysics, from the lowest-mass stars to compact objects. In "traditional" topics such as stellar structure and evolution and stellar seismology, it is crucial that the United States support the intellectual infrastructure—observational, experimental, and theoretical—to take full advantage of these new opportunities.

Stellar astrophysics has informed many other areas of physics, including nuclear physics, particle physics, and general relativity. Moreover, an understanding of stellar astrophysics is needed for many other problems in astronomy. The study of galaxies at high redshift relies critically on an understanding of the stellar populations that make up those galaxies. The formation of stars, galaxies, and the intracluster medium in galaxy clusters is strongly influenced by the heavy elements, ionizing photons, and explosions produced by massive stars. The Sun continues to be a working template for understanding magnetohydrodynamics and plasma physics "in practice"—physics that is crucial in many other arenas, including that of compact objects. The most distant known object in the universe is now a GRB. These GRBs are (temporarily) much brighter than quasars in the optical-ultraviolet (UV), allowing unique studies of the intergalactic medium at high redshift. A distinct and complementary probe of star formation at these early times is provided by the discovery of nearby extremely metal-poor stars, which constrain the nucleosynthetic products of the first generations of stars.

Just as stellar astrophysics has a significant impact on other branches of physics and astrophysics, it also requires input from other disciplines, notably laboratory experiments. For example, the stellar interior models now in use rely on purely theoretical opacity calculations, but these will be tested by laboratory data in the next decade. Next-generation solar-neutrino experiments can measure the central solar temperature and potentially constrain solar abundances. Key nuclear cross sections are needed for a quantitative understanding of nucleosynthesis and energy generation in stars and stellar explosions. More recently, experiments focused on studying complex hydrodynamic processes have provided key insights into aspects of stellar physics. For example, laboratory experiments and space-physics measurements contributed to the recognition that fast magnetic reconnection occurs only in plasmas with low collision rates, with potential applications to stellar and accretion disk coronae. Continued laboratory studies of basic physical processes important in stellar astrophysics, such as reconnection, angular-momentum transport, and combustion, would complement more traditional observational and theoretical work. More generally, a transition to models grounded in experimental data has the capability to open up new realms of precision stellar astrophysics that could have a major impact on astronomy and physics.

One-dimensional stellar models have yielded considerable insight and are quantitatively sufficient for many applications. However, there are also crucial aspects of stellar structure and evolution that require a three-dimensional approach. In these areas, a combined effort involving numerical simulations informing observations, and vice versa, is needed for progress. Studying these inherently three-dimensional problems computationally has become feasible in the past 10 to 15 years, with the tremendous advances in computational resources. Similarly, observational progress on many of these problems is just now feasible. This includes the realization of interferometric techniques in the past decade to illuminate the rotationally distorted shapes of massive stars, the enlarged radii of active M-dwarfs, and the details of how stars lose mass. In addition, observations of the Sun show that convection (inherently multidimensional) plays a critical role in shaping both its rotation profile and magnetic structure. The observational and theoretical study of the solar case provides a point of departure for the study of rotation and magnetism in other stars (science question SSE 1).

Other problems in which a full three-dimensional understanding is crucial are the thermonuclear explosions of white dwarfs in Type Ia supernovae, and the core collapse and explosion of massive stars (science questions SSE 2 and SSE 3). These topics will benefit from the development of new time-domain surveys (SSE discovery area). Already, surveys are finding supernovae with extreme properties in terms of luminosity and energy (both high and low). The evolution leading up to stellar explosions—in particular, mass loss and the dynamics of binaries—is crucial for an understanding of the variety of observed explosions. The study of compact stellar remnants has become mature, but fundamental questions remain about the basic properties of compact objects, such as what determines their masses, spins, and radii (science question SSE 4).

The entire subject of stellar astrophysics will benefit tremendously from large time-domain surveys, given that the variability of the timescales to which large surveys are sensitive matches well with those of many stellar phenomena (see the section “SSE Discovery Area: Time-Domain Surveys”). In addition to explosive events often related to compact objects, a variety of variable and binary stars can be studied. Many time-variable events may be found that have been predicted but not yet discovered (e.g., orphan afterglows of gamma-ray bursts).

THE SCIENCE FRONTIERS

SSE 1. How Do Rotation and Magnetic Fields Affect Stars?

The standard models of stellar structure are traditionally one-dimensional. Observational and theoretical results in the past decade have demonstrated that the study of rotation and magnetic fields, coupled with convection, demands a dynamic, three-dimensional approach that is now achievable. This complex reality is most vividly illustrated by the Sun's bewildering array of multiscale inhomogeneity, which is now *seen* with remarkable magnetic sensitivity, and with temporal and spatial resolution, and *understood* with the help of detailed simulations. It is sobering that this rich structure and its few-gauss average dynamo fields result from the interaction of only a few-kilometers-per-second surface rotation with comparable subsurface convective velocities. The much broader range of *stellar* rotation, convection, and magnetism poses both a challenge and an opportunity to complement the detailed views of the Sun.

Beyond the Standard Picture

Rotation is now recognized as fundamental to the understanding of stellar evolution. Photometric monitoring programs have mapped the mass-dependent range of pre-main sequence rotation rates and have shown that star-disk interactions (planet and star formation processes) are key to the subsequent evolution of stellar rotation. It is known that Sun-like stars spin down from magnetized winds. Yet recent large samples of active stars now show that the stellar-mass dependence of these winds is uncorrelated with the boundary at which stars become fully convective, contrary to theoretical expectations. Helioseismology has revealed the internal solar rotation, invigorating debate over dynamo theories and confirming the strong coupling between the radiative core and convective envelope. However, mechanisms for angular-momentum transport in stellar interiors, the related mixing, and relevant timescales are still poorly understood. Nascent asteroseismic constraints on internal stellar rotation exist, but crucial knowledge, such as the core rotation rate in pre-supernova stars, is lacking. Understanding rotation-induced mixing in stellar interiors and its effect on stellar and chemical evolution demands that these uncertainties be resolved.

It is known that magnetic fields in cool stars are generated by a dynamo mechanism, but its precise nature, and the heating mechanism for chromospheres and coronae, remain subjects of vigorous debate. Computational advances now allow three-dimensional, radiative MHD models of the outer solar atmosphere that are consistent with observations showing no meaningful "average" or homogeneous chromosphere. Such models are finally making progress on the important question of how magnetic energy from the convective regions makes its way into the magnetically dominated layers. Strong magnetic fields are also detected in hot, higher-mass stars without strong surface convection, and in their stellar remnants. Their origin and evolution, possibly from relic fields, is a mystery.

The prospect of exciting breakthroughs in the understanding of these issues arises from great leaps in stellar observational capabilities, such as interferometry, spectropolarimetry, magnetometry, asteroseismology, and extended waveband surveys from radio and infrared to X-rays and gamma rays. Stars can now be imaged and modeled as three-dimensional objects. Realistic three-dimensional atmosphere simulations are a fruitful new frontier in the understanding of the envelopes of stars. Powerful diagnostics of internal properties are available, and time-domain surveys will allow the study of stellar rotation in samples of unprecedented size, precision, and duration. The complementary approaches of large surveys and detailed studies of smaller samples promise new observational constraints on the origin, nature, and consequences of stellar rotation and magnetic fields. Gaia will provide precise astrometric data and spectroscopic information for an unprecedented sample, which will be invaluable for characterizing stellar properties. These will be combined with unprecedented new capabilities for high-resolution solar observations, particularly with the Advanced Technology Solar Telescope (ATST) and the Solar Dynamics Observatory (SDO). On the theoretical side, fully three-dimensional, time-dependent MHD solar simulations will provide the template for more realistic two-dimensional and three-dimensional stellar models. Against this backdrop, the following subsections describe five important problems involving stellar rotation and magnetic fields that seem particularly ripe for progress.

How Are Magnetic Fields Generated in Stars?

Helioseismology, spectropolarimetry, and radio gyrosynchrotron observations offer new tools for measuring the complex global magnetic structures in the Sun and stars. Local helioseismology with the SDO (a satellite to be launched in early 2010) will revolutionize astronomers' vision of subphotospheric magnetic fields, while sensitive polarimetry obtained with ATST (now under construction) will unveil the dynamic magnetized atmosphere from photosphere to corona (Figure 5.1). Lower-spatial-resolution coronal magnetometry over the full disk using microwave gyrosynchrotron or IR Zeeman measurements could effectively sample intermittent solar atmospheric explosive events. This would also be a powerful trigger for more sensitive ATST observations. This new generation of high temporal- and spatial-resolution magnetometry will help disentangle the influence of flares and mass ejections on chromospheric and coronal structure. For stars, accurate spectropolarimetry with high Stokes Q/U/V sensitivity is the demonstrated key to measuring surface magnetism. A significant increase in the stellar sample size and duration of these observations could be made with just a single, dedicated 4-m-class telescope and suitable instrumentation equipped with a spectropolarimeter.

A new generation of sophisticated simulations that include realistic treatments of radiation and non-equilibrium ionization can be harnessed to interpret these data. The panel urges the provision of support for such theoretical investigations and the extension of detailed solar models to the broader stellar regime. This re-

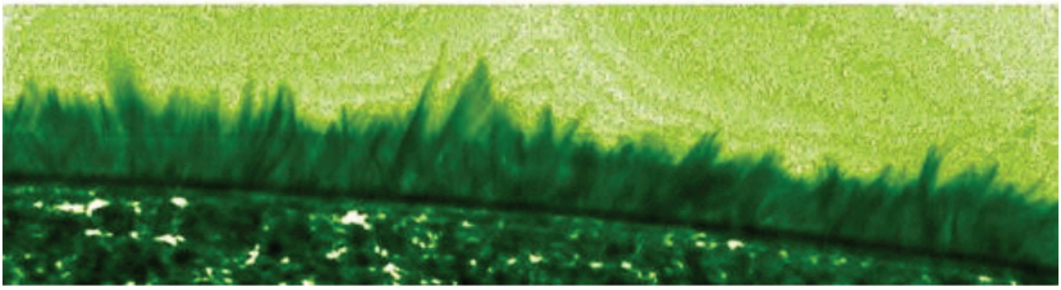


FIGURE 5.1 The solar chromosphere has a fundamentally dynamic structure. This spectacular high-resolution CaII image from the Hinode satellite shows how fine-scale magnetic structures (spicules) dynamically set local physical boundary conditions with the photosphere below. A new generation of high-resolution ground-based measurements (ATST) will directly measure this so-far-invisible driving magnetic field from the photosphere far into the solar atmosphere. SOURCE: Y. Suematsu, K. Ichimoto, Y. Katsukawa, T. Shimizu, T. Okamoto, S. Tsuneta, T. Tarbell, and R.A. Shine, High resolution observations of spicules with Hinode/SOT, p. 27 in *First Results from Hinode*, Astronomical Society of the Pacific Conference Series (S.A. Matthews, J.M. Davis, and L.K. Harra, eds.), Vol. 397, Astronomical Society of the Pacific, San Francisco, Calif., 2008. Courtesy of Dr. Y. Suematsu, National Astronomical Observatory of Japan.

search includes exploring dynamos in the cores of massive stars and examining a range of rotation rates, masses, and evolutionary states, including fully convective stars and giants and supergiants. Precise radius measurements in stars with differing rotation and starspot properties, along with detailed seismic studies, will allow the measurement of how interior stellar structure is affected by magnetic activity. Understanding magnetized winds in close-binary systems and how they differ from those of single stars may prove important in understanding phenomena such as blue stragglers and cataclysmic variable stars. It is essential to future progress to train and support theorists who work across the boundaries of solar and stellar studies.

What Are the Stellar Surface and Internal Rotation Distributions?

Rotation periods from spot modulation can be now be inferred for large samples of cool stars and brown dwarfs, including slow rotators (Figure 5.2). From stars with a broad range of masses and ages, astronomers can understand angular-momentum evolution and learn about the star-formation process. Obtaining rotation rates and activity measurements in samples of known age will characterize rotation-age relationships, whereas measuring rotation in numerous old-field stars is a new capability with real discovery potential. Large samples including rotation are especially important for understanding binary and interacting-binary evolu-

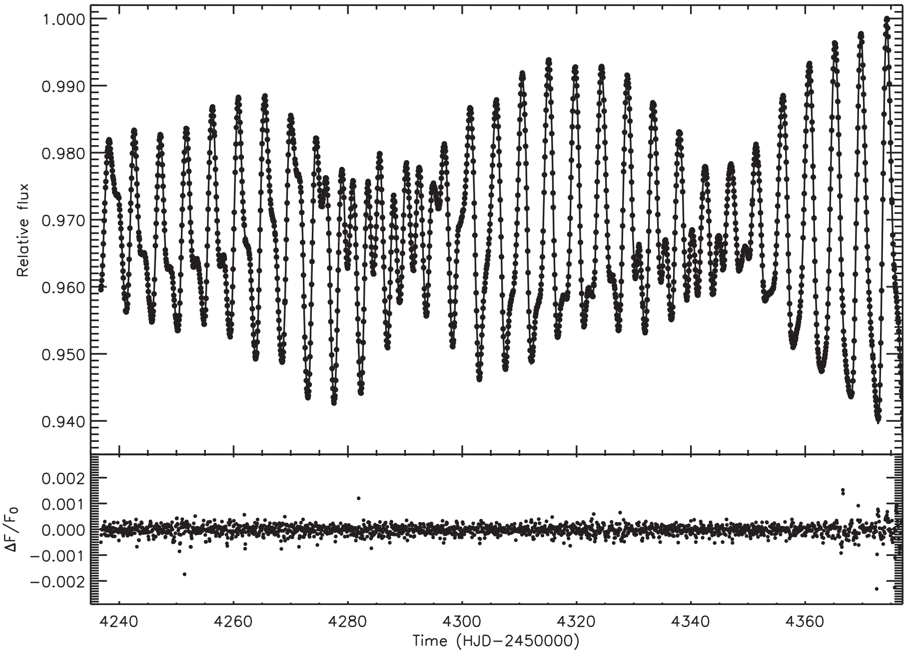


FIGURE 5.2 The light curve of a solar-like star fitted with a multiple-spot model—with two sets of spots at different angular velocities—shows the potential for studying the dependence of magnetic activity patterns on stellar age, metallicity, and rotation using time-series data of relatively high precision and high cadence from planet search programs. In the top panel, CoRoT data are fitted by a model; in the lower panel, the residuals are shown. SOURCE: A.F. Lanza, I. Pagano, G. Leto, S. Messina, S. Aigrain, R. Alonso, M. Auvergne, A. Baglin, P. Barge, A.S. Bonomo, P. Boumier, A. Collier Cameron, M. Comparato, G. Cutispoto, J.R. De Medeiros, B. Foing, A. Kaiser, C. Moutou, P.S. Parihar, A. Silva-Valio, and W.W. Weiss, Magnetic activity in the photosphere of CoRoT-Exo-2a—Active longitudes and short-term spot cycle in a young Sun-like star, *Astronomy & Astrophysics* 493:193-200, 2009, reproduced with permission © ESO. Courtesy of A.F. Lanza, INAF-Osservatorio Astrofisico di Catania, based on data obtained with CoRoT, a space project operated by the French Space Agency, CNES, with participation of the Science Programme of ESA, ESTEC/RSSD, Austria, Belgium, Brazil, Germany, and Spain.

tion. Rotation samples of early-type stars will require moderate-resolution, multi-object spectroscopy.

Direct tests of internal stellar rotation through interferometry (resolved imaging of distorted massive stars) and asteroseismology (in a variety of stellar masses and evolutionary regimes) are now possible. Kepler and the Convection, Rotation and Planetary Transits (CoRoT) spacecraft will provide extensive asteroseismic data in the first few years of the 2010-2020 decade. These should be complemented with next-generation, midsize space missions or networks of modest-aperture, ground-based telescopes using long time-baseline and high-cadence campaigns for modest-size and carefully selected samples. Support for theoretical pulsation studies, especially of massive and evolved stars, will also be required so that the full benefits of the new seismic data can be reaped. Extending sophisticated solar simulations of convection and angular momentum transport to the stellar regime (including rotationally induced mixing) should be a priority, as should the training and supporting of young investigators to take advantage of the computational advances and new data.

What Is the Impact of Rotation on Stellar Evolution and Mass Loss?

Rotation can impact stellar structure directly and induce mixing in radiative regions. The effects will be most dramatic in massive, rapidly rotating stars. Evolved massive supergiants often show enhanced N/O and N/C in their atmospheres and winds, implying interior mixing and the dredging up of CNO-processed material from the core. Strong rotation-induced interior mixing can also alter evolution, weakening redward loops. Rotation in massive stars is even observed to approach the critical rate, generating distortion and the associated equatorial gravity darkening that can now be directly measured interferometrically (Figure 5.3). The relatively bright poles drive a strong wind, leading to the bipolar, prolate mass-loss nebulas seen in luminous blue variable (LBV) stars such as Eta Carina. In classical Be and supergiant B[e] stars, near-critical rotation can induce the centrifugal ejection of material into equatorial "excretion disks." A key issue for future studies is the relative importance of these competing mass-loss effects on the associated angular-momentum loss and its role in spin-down, or in limiting spin-up, during massive star evolution. The insensitivity of centrifugal excretion to metallicity means that it could be particularly important in the first stars.

In intermediate-mass stars (~2-8 $M_\odot$ [solar masses]), asteroseismic studies of convective-core sizes will establish the degree of overshooting and whether it depends on rotation. Progress in theoretical pulsation studies may permit the measurement of internal stellar rotation rates in intermediate-mass stars. Such data will be the most direct diagnostic of the impact of rotation on stellar lifetimes. Multi-object, high-resolution spectroscopy can map out the mixing of nuclear-processed material to the surface and can detect subsurface mixing from changes in the degree of dredge-up. In addition to serving as tests of stellar physics, such abundance data will be a powerful asset in assigning stellar ages and in investigations of chemical evolution. Both photometric and spectroscopic abundance diagnostics will also need to be refined to elucidate mixing patterns as a function of metallicity.

How Are Chromospheres and Coronae Formed?

Much of the knowledge of solar and stellar magnetism comes from observing the activity that it induces in their atmospheres. A key goal is to identify the heating

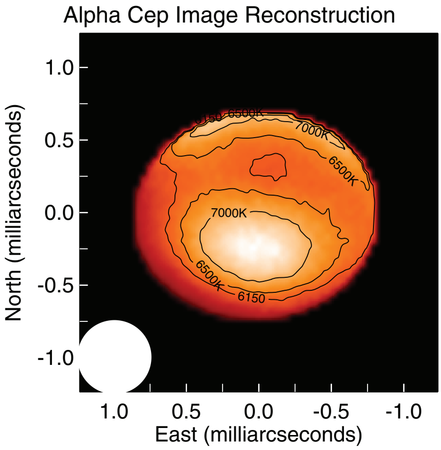


FIGURE 5.3 Temperature variation across the surface of α Cep (Alderamin) imaged with the CHARA interferometry array, interpreted through modeling. The apparent surface brightness is the result of both rotation (hot poles and cool equator) and limb darkening (seeing into higher, cooler layers away from the disk center) and thus does not peak inside the latitude oval surrounding the rotational pole. Yellow lines are constant latitude lines from the standard model used by the authors to interpret the data. The white circle indicates the convolving beam used to produce the image, 0.68 mas. SOURCE: M. Zhao, J.D. Monnier, E. Pedretti, N. Thureau, A. Mérand, T. Ten Brummelaar, H. McAlister, S.T. Ridgway, N. Turner, J. Sturmann, L. Sturmann, P.J. Goldfinger, and C. Farrington. Imaging and modeling rapidly rotating stars: α cephei and α ophiuchi, *Astrophysical Journal* 701:209, 2009, reproduced by permission of the AAS.

mechanism that transfers energy from the magnetic field to heat the chromosphere and corona, providing observable diagnostics of magnetic activity. Comprehensive, multiwavelength observations of stellar activity from radio to X-ray wavelengths are being carried out at present with such facilities as the Expanded Very Large Array (EVLA), 4-m-class and 8-m-class ground-based telescopes, the Galaxy Evolution Explorer (GALEX), Chandra, and the Hubble Space Telescope (HST), with the newly refurbished Space Telescope Imaging Spectrograph (STIS) instrument. New facilities, especially those focused on ultraviolet and X-ray spectroscopy, will allow such studies to expand to cover much larger samples of stars. The Sun can be observed at very high spatial and temporal resolution (particularly with the expected advent of ATST in the near future), and—along with its own intrinsic interest—also acts as a template for the stellar observations. Solar observations from radio to X-ray wavelengths are needed. Laboratory experiments and plasma simulations should also contribute to this effort.

Both steady and episodic components of magnetic activity and photometric variability are of interest. Steady components illustrate how activity changes on long timescales; flares and coronal mass ejections trace very short-lived energetic events; and spots and activity cycles address changes on timescales from weeks to years. The study of changes in magnetic activity as a function of stellar mass and age will benefit from large time-domain surveys, especially in clusters. Phenomena to be understood include the following: the spindown time and its connection to magnetic activity; the origin of a seeming upper limit to activity (often termed saturation); the still-inexplicable presence of strong, long-lived magnetic fields and magnetic activity in fully convective stars; and the appearance of chromospheres in objects as disparate as red giants and brown dwarfs. All of these diagnostics will help constrain the dynamo mechanism(s) that produce the fields. Additionally, late

M-dwarfs are promising targets for finding Earth-mass planets, but the impact of their persistent activity on habitability needs to be understood.

What Is the Origin of Highly Magnetized Stars?

Spectropolarimetry reveals a subset (~5 to 10 percent) of massive, hot stars that possess large-scale magnetic fields with global strengths up to 10^4 G, dwarfing what is seen on the Sun or other later-type stars. In marked contrast to the dynamo activity cycles of lower-mass stars, the global properties of massive-star fields seem stable for years and even decades. This, and an apparent decline of occurrence fraction with age, hint at a fossil origin during star formation. These strong fields can profoundly affect the mass and angular-momentum loss from winds. The rotational modulation associated with magnetic clouds and surface abundance patterns allows the determination of very accurate rotation periods, with potential for the direct measurement of stellar spindown. Improved spectropolarimeters on larger (10-m-class) telescopes should make possible field detection at the 1-10 gauss level in statistical samples of massive stars. Comprehensive surveys are needed to characterize accurately the underlying stellar populations and thus to provide clearer constraints on the origin of both strong and weak fields.

Not only are main-sequence stars highly magnetized; about 10 percent of field white dwarfs have strong fields ($B \sim 10^6$-10^9 G). The fraction is even larger, approximately 25 percent, for white dwarfs in close binaries. There is also a significant population of neutron stars with magnetic fields ~10^{14}-10^{15} G. The birthrate of these "magnetars" is uncertain, but they probably represent at least 10 percent, and perhaps up to 50 percent, of all neutron stars. The connection between highly magnetized compact objects and their progenitors remains uncertain. Binarity is clearly implicated for some magnetized white dwarfs, but its potential role in neutron-star magnetism is unknown. Numerical simulations of dynamos during different stages of stellar evolution would clarify the likely origin of compact-object magnetism. Improved statistics on the numbers of magnetized hot stars and observational constraints on internal stellar rotation may also distinguish between fossil-field and dynamo models.

SSE 2. What Are the Progenitors of Type Ia Supernovae and How Do They Explode?

The evidence is strong that Type Ia supernovae (SNe Ia) are thermonuclear explosions of white dwarfs. Yet there is much that is not known. A physical understanding for the mechanism of the explosions and an astronomical understanding of which stars become SNe Ia are sought. The next decade presents opportunities for major advances on these questions.

The Standard Picture

A Type Ia supernova is probably an accreting carbon-oxygen white dwarf in a binary system. Theoretical models of those explosions generally match the data, although direct evidence for this mechanism has proved elusive. The nature of the white dwarf's companion is not established. Although single-degenerate binary progenitors are favored, merging white dwarfs (double-degenerate binaries) are not ruled out. Merging white dwarfs are disfavored because there is not direct evidence for sufficiently many appropriate binary progenitor systems, and the mergers may not lead to explosions with the uniform properties of SN Ia. SN Ia explosions triggered by merging white dwarfs are not ruled out. Type Ia supernovae are an important area of stellar astrophysics and cosmic evolution: they are violent end points of stellar evolution, and they create much of the iron in the universe. Using SN Ia light curves provides strong evidence for the accelerating expansion of the universe. Subtle effects of age and composition on stellar evolution are likely to modify the properties of SNe Ia over cosmic time. Understanding these will strengthen inferences about dark energy. Because these challenging questions in stellar physics have such broad significance, elevating the understanding of SNe Ia to a new level of precision will be an important task for the next decade.

The basic picture of an exploding white dwarf seems sound: only SNe Ia occur in elliptical galaxies, indicating long-lived progenitors for these SNe Ia. The spectrum at maximum light shows no hydrogen, as expected for a highly evolved star, whereas the spectrum seen months after the explosion is a mass of blended lines from iron, as expected for a star that burns its interior to the iron peak. Theory indicates that fusion ignites near the center and that a thermonuclear fusion flame then rips through the star. But theory also shows that the flame is unstable and creates turbulence. This complex, unstable, turbulent flame burns much of the star to radioactive ^{56}Ni. The radioactive decay of that nickel is responsible for the 4×10^9 solar luminosities that allow SNe Ia to be seen halfway across the universe. Although there is a range of outcomes from nuclear burning, and SNe Ia show a factor-of-three range in peak luminosity, there is an effective way to determine the intrinsic brightness of an SN Ia from its light curve in many colors. This is thought to be the result of diffusion of energy from various amounts of ^{56}Ni through the expanding envelope, but a theoretical model that can account quantitatively for the observed relationship is still lacking. The precision of corrections based on light-curve shape makes SNe Ia the most powerful extragalactic distance indicators. But the path of evolution to an accreting white dwarf, the sites of ignition, the turbulent nuclear burning, and a possible transition from subsonic burning to detonation all remain uncertain in this model. Empirical work may be lumping together explosions that have different chemical composition and progenitor paths. We can do better.

Opportunities for the Coming Decade

Major advances in understanding SNe Ia lie just ahead: supernova samples will improve dramatically, enabling progenitor evolution to be traced, and it can be expected that computations will cope more effectively with the underlying physics of the explosions. Improved observational capabilities across a wide range of wavelengths will be applied effectively to the unsolved puzzles of SNe Ia.

- *Samples.* The world's sample of well-observed SNe Ia is only a few hundred objects. Ground-based surveys underway—such as the Katzman Automatic Imaging Telescope (KAIT), Palomar Transient Factory (PTF), and PanStarrs-1—and planned will discover orders-of-magnitude more events. Light curves in several colors will be determined for most. A large number need firmly measured redshifts and types; a smaller number will have detailed spectroscopic histories. With large samples, the diversity of this class of explosion will be more fully explored to determine the connection between supernova properties and the chemistry and star-forming history of the host galaxies. At present, uncertainties in dust properties pose the limit to the precision of distance measurements with Type Ia supernovae. During the coming decade, measurements in the near IR with the Peters Automated Infrared Imaging Telescope (PAIRITEL) and JWST should provide much better understanding of this problem and provide the most precise and accurate distance measures. Studies may be extended to higher redshift with wide-field IR surveys from space.
- *Tracing evolution.* Progress is also expected in tracing the long path of stellar evolution to an SN Ia explosion. This path includes mass transfer and common-envelope phases: matching the numbers and lifetimes of progenitors with the rates and types of stellar explosions is an incomplete task that future synoptic surveys will probe. The close binaries that precede mergers will be sources of gravitational radiation in the galaxy and have broad interest beyond the question of which stars become SNe Ia (or possibly other types of stellar explosions not yet categorized). Some of these binaries will be detected in synoptic surveys that search the timescales associated with their orbits, which range from hours to years. Matching the observed populations with the evolutionary paths and the observed rates of stellar death requires large, intensive, and carefully characterized surveys to find and count the stars at each stage. In the next decade, surveys in the galaxy should aim to discover binaries that contain white dwarfs approaching the Chandrasekhar mass, on their way to becoming SNe Ia. Progress in binary evolution models, especially poorly understood phases such as common envelope evolution, is needed along with the observational developments.
- *Computation.* Major advances in modeling will come from three-dimensional simulations of ignition and burning physics with spatial resolution that can fol-

low the turbulence (Figure 5.4). Light curves and time-dependent spectra will be calculated from three-dimensional models. Improvements in three-dimensional radiative-transfer calculations will be important for matching simulations to observations. The required improvement of roughly a factor of 10 in resolution is a challenging, yet reasonable goal. Although the factor-of-10^4 increase in central processing unit (CPU) capability to do this by brute-force computing alone is not feasible, significant improvements in codes may achieve this goal. As the large observational and model databases accumulate, finding the best matches will demand new ways to analyze the results—perhaps machine intelligence will help. Time series of spectra will map the distribution of ejected composition with velocity for many events. A fruitful confrontation of models with data will ensue. In a decade, today's conceptual issues for SN Ia explosions, including the possible transition from a deflagration to a detonation, should be settled. This is not the same as saying that the accuracy of theoretical predictions by themselves will exceed the requirements for the most demanding applications of SN Ia. Since the goal in dark-energy studies is to determine the distance to an ensemble of SNe Ia with an accuracy better than 1 percent, a combination of theoretical insight and better empirical evidence will be required. Major advances are also expected in the understanding of important nuclear physics model input such as the $^{12}C+^{12}C$ fusion rate and electron capture rates. Progress depends on the development of experimental facilities as well as of nuclear theory, because not all relevant transitions can be measured in the laboratory.

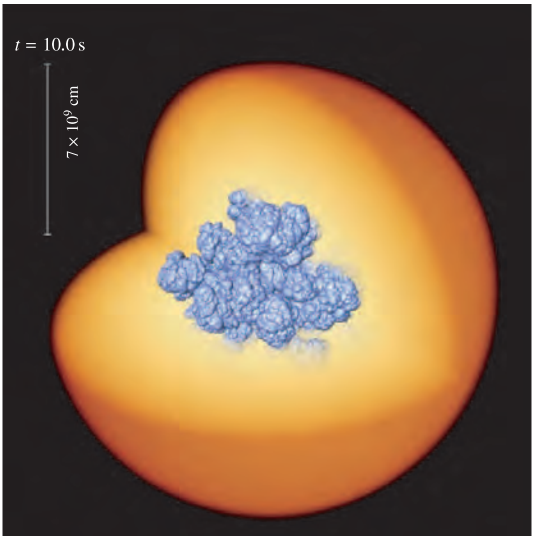


FIGURE 5.4 This hydrodynamic simulation of the turbulent development of a Type Ia supernova explosion by means of deflagration, starting with a C+O white dwarf, shows the potential of three-dimensional codes now becoming available. SOURCE: F.K. Röpke, W. Hillebrandt, W. Schmidt, J.C. Niemeyer, S.I. Blinnikov, and P.A. Mazzali, A three-dimensional deflagration model for Type Ia supernovae compared with observations, *Astrophysical Journal* 668:1132, 2007. Reproduced by permission of the AAS. Courtesy of F. Röpke, Max Planck Institute for Astrophysics, Garching, Germany.

- *Many wavelengths.* Major advances in observing SNe Ia from improved searches and more effective follow-up over a wide range of wavelengths from the radio to gamma rays are expected. More-sensitive radio observations with the EVLA and future instruments can pose strict limits on circumstellar matter near the explosion. Recent near-IR observations demonstrate that SNe Ia are better standard candles at 1.6 microns than in the visible, and the infrared corrections for dust extinction are also significantly smaller. This leads to the prospect that SNe Ia, already the best standard candles, can be made more precise and accurate through infrared follow-up. Follow-up in the rest-frame IR can be carried out at low redshifts from the ground, but a space-based survey would be extremely valuable to reach the rest-frame infrared beyond $z \sim 0.5$. Rolling searches that have a cadence of a few days will provide excellent sampling of thousands of SNe Ia light curves. Spectroscopic follow-up remains a high scientific priority: queue observing on Gemini has been effective, and improved access would be useful. Samples that are drawn from well-defined searches with carefully prescribed selection criteria can be used to establish rates and can be compared across cosmic time to learn more about the gestation time for supernova progenitors.

The ultraviolet is directly observable from the ground for high-redshift samples, where observations hint that the UV emission from SNe Ia is more variable than the optical. There are some observations from HST and Swift, but overall the situation in the UV is not very satisfactory and offers a rich opportunity for future work.

Current 8- to 10-m telescopes have allowed the detailed study of SNe Ia. The next generation of large telescopes (20 to 30 m) will provide opportunities to study distant supernovae in much the same way that the nearby sample is studied. This will provide direct comparisons for observing any drifts in the properties of supernovae with cosmic epoch that may result from shifts in the mean age of the stellar population and chemical evolution in galaxies. Very large telescopes will be more capable for the high-resolution spectroscopy needed to search for residual gas in the neighborhood of the explosions and for the photon-starved work of spectropolarimetry. Polarization measurements teach astronomers about asymmetries in explosions that will be fruitful to compare with predictions of multidimensional explosion models. Near-IR spectroscopy is still poorly explored for SNe Ia, and the advantages of large telescopes with adaptive optics (AO) systems for this work are substantial.

In our galaxy, Chandra has produced beautiful X-ray observations of Tycho's and Kepler's SNe (Figure 5.5), giving both spatial and spectral information on the distribution of heavy elements. Explosion models do not yet match the observations well. Detection of faint X-ray lines with future missions is promising for providing diagnostics of the explosion mechanism and the metallicity of the progenitor. ^{56}Ni

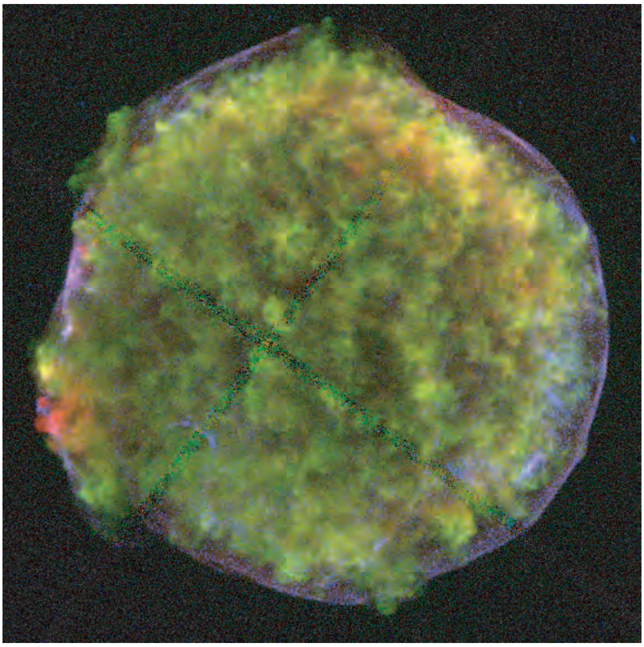


FIGURE 5.5 In this X-ray image of Tycho's supernova remnant (known to be a normal Type Ia supernova from echo emission), the blue outer emission is from the interstellar shock wave and the flocculent emission is from ejecta, which are Si and Fe rich. Red, green, and blue colors are assigned to bands 0.95-1.26, 1.63-2.26, and 4.1-6.1 keV of the Chandra X-ray telescope. SOURCE: J.S. Warren, J.P. Hughes, C. Badenes, P. Ghavamian, C.F. McKee, D. Moffett, P.P. Plucinsky, C. Rakowski, E. Reynoso, and P. Slane, Cosmic-ray acceleration at the forward shock in Tycho's supernova remnant: Evidence from Chandra X-ray observations, *Astrophysical Journal* 634(1):376, 2005, reproduced by permission of the AAS.

and ^{56}Co are powerful gamma-line emitters, and SNe Ia are copious producers of these isotopes (more than half the mass of the explosion is converted to iron-peak elements). In principle, a detailed map of the speed and abundance of radioactive ^{56}Ni and ^{56}Co would sharply constrain the explosion physics. To get a reasonable rate of events, the sensitivity of a useful gamma-ray instrument needs to be much better than has been achieved so far. The detection of the integrated gamma-ray background from all the SNe Ia that have ever occurred over cosmic time would be a major accomplishment and warrants further study. Instrument development is needed in the coming decade for this form of astronomy to realize its potential.

SSE 3. How Do the Lives of Massive Stars End?

What they lack in numbers, massive (initial mass $M \geq 8\ M_\odot$) stars make up for in energy and extreme conditions. In their deaths, they produce spectacular fireworks—supernovae and gamma-ray bursts—that leave behind exotic objects: neutron stars and black holes. They create the elements necessary for life. The interior of a "core-collapse supernova" is a physical laboratory with conditions not seen elsewhere in the universe. The neutrino burst that announces the central core collapse is one of the most powerful events in the universe, and the jet that makes a GRB is one of the fastest flows of matter. Yet no one is really sure how it all works.

The Standard Picture

Most massive stars end their lives when their inner cores of heavy elements collapse to form neutron stars or black holes. This releases an enormous amount of energy, chiefly radiated away as neutrinos. As already noted in the late 1930s, if a small fraction—approximately 1 percent—of this binding energy could be tapped, this could account for the comparatively modest kinetic energy of common supernovae. For nearly 70 years, however, finding a robust mechanism to drive the explosion has proven elusive. Today, the picture is both complicated and made more tractable by the realization that probably not every star blows up (or collapses) in the same way. In particular, stars that die with rapidly rotating cores have an additional source of energy to tap, and the conduit may involve magnetic fields. This has probably been illustrated by observations showing that "long-soft" gamma-ray bursts are connected with massive-star death and are often accompanied by supernovae. In the next decade, driven especially by advances in computation and large observational surveys, it is likely that major advances will occur in the understanding of how massive stars of all masses, metallicities, and rotation rates end their lives.

Opportunities for the Coming Decade: The Progenitor Stars

The outcome of core collapse depends on the properties of the pre-supernova star. The greatest uncertainties lie at the extremes of mass, metallicity, and rotation rate. On the low-mass end ($M \approx 8$ to $10\ M_\odot$), what is the main-sequence mass that separates white-dwarf remnants from supernovae? Might some progenitor stars even have a white dwarf (carbon and oxygen or oxygen and magnesium) core? At the high-mass end, does the well-established metallicity dependence of line-driven winds for lower-mass stars carry over to the episodic ejections that now appear to dominate for very massive stars? A deeper understanding of mass loss is essential to extend what is learned from relatively close Type II SNe (those with observed hydrogen) to those that exploded when the universe was young. It is also important for determining what main-sequence mass separates stars that form neutron stars from those that form black holes.

Recent models suggest that stars of only 10 $M_\odot$ explode with neutrino heating alone but that more massive stars are more tightly bound and are harder to blow up this way. About a dozen Type II supernovae currently have reliably determined progenitors based on pre-supernova archival imaging, especially with HST. These progenitors generally have properties consistent with main-sequence masses around 10 to 15 $M_\odot$, as expected for average core-collapse SNe, although a Type IIn (narrow line) SN has been found to have a massive, luminous blue variable progenitor (Figure 5.6). The strong episodic mass loss in LBVs affects the appearance of the

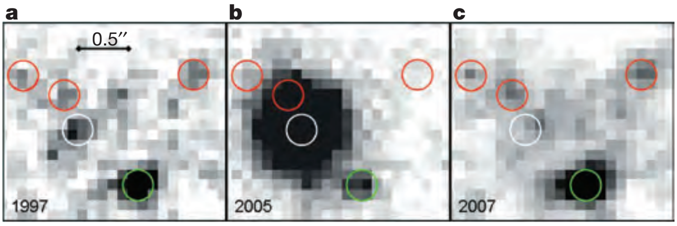


FIGURE 5.6 Three views of the location of a supernova in galaxy NGC 266. (a) In 1997, five stars can be located in the field, one of them a luminous blue variable (LBV), inside the white circle. (b) The supernova, SN 2005gl, is observed in October 2005, at the location of the LBV. (c) When the supernova has faded, the location of the LBV is now empty, clear evidence that the star seen in 1997 was the progenitor. Using multicolor imaging of nearby galaxies with a space telescope or a large ground-based telescope, many more progenitor stars should be identified. SOURCE: Reprinted by permission from Macmillan Publishers Ltd.: *Nature*, A. Gal-Yam and D.C. Leonard, A massive hypergiant star as the progenitor of the supernova SN 2005gl, *Nature* 458:865, 2009, copyright 2009.

supernova, and an understanding of the relation between stellar properties, episodic mass loss, and supernova type is needed. Observations of young massive star clusters show that highly magnetized neutron stars, magnetars, form from stars with main-sequence masses above 40 to 50 $M_\odot$. One possibility is that rotation and magnetic fields become a bigger factor in the explosion as the mass increases. Some ultrabright supernovae such as SN 2007bi suggest that the most massive stars may not even die by core collapse at all, but by an electron-positron "pair instability." Binary evolution can also affect the appearance of a supernova through mass transfer to or mass loss from the progenitor star. Supernovae with stripped H envelopes (Types IIb, Ib, Ic) may have binary, as well as single-star, progenitors. Better understanding of binary-star populations is needed to discern these possibilities.

New surveys, such as PTF and the Catalina Real-Time Transient Survey, are revealing the surprising diversity of core-collapse supernovae (Figure 5.7). In the next decade, large-volume optical and near-IR time-domain surveys will increase the observational database of supernovae by orders of magnitude. In particular, surveys with approximately 5-day cadence and volumes 10 to 100 times larger than current surveys could double the known SNe population in about 2 years. These surveys will find the very rare events that are often keys to progress. They will also increase the number of known pre-supernova stars, detect large numbers of LBV-like outbursts, and determine whether there is a class of "orphan afterglows" (long-wavelength emission from GRBs for which the higher-energy emission is beamed out of our line of sight). These discoveries will help to constrain directly the white dwarf–neutron star transition mass of the progenitor star, the maximum stellar mass for main-sequence stars, the uncertain stellar evolution prior to core collapse (e.g., the relative importance of steady winds versus LBV eruptions), and the role of rotation and magnetic fields in core-collapse explosions.

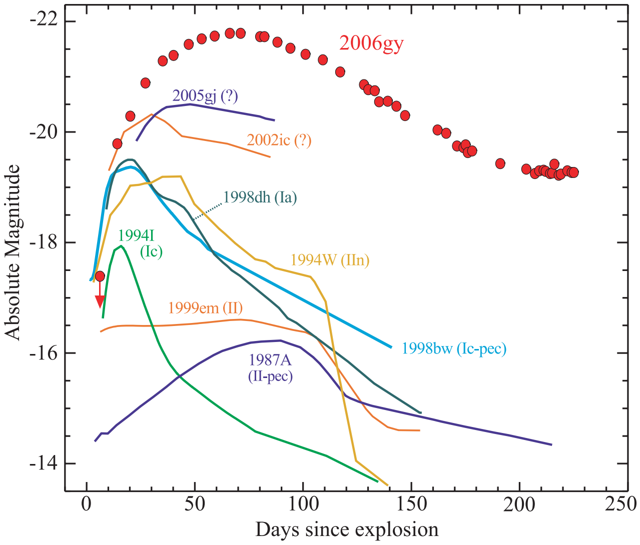


FIGURE 5.7 The wide variety of light curves of likely core-collapse supernovae is shown, along with that of one Type Ia supernova. The light curves provide physical constraints on the explosion parameters. SOURCE: N. Smith, W. Li, R.J. Foley, J.C. Wheeler, D. Pooley, R. Chornock, A.V. Filippenko, J.M. Silverman, R. Quimby, J.S. Bloom, and C. Hansen, SN 2006gy: Discovery of the most luminous supernova ever recorded, powered by the death of an extremely massive star like η Carinae, *Astrophysical Journal* 666(2):1116, 2007, reproduced by permission of the AAS. Courtesy of Nathan Smith, Astronomy Department, University of California, Berkeley.

Radio and X-ray observations will also probe shock interactions in the circumstellar medium and mass loss in the hundreds to thousands of years leading up to the explosion. The EVLA will allow detailed radio observations of all types of core-collapse supernovae. For young galactic supernova remnants, the Chandra X-ray observatory is a powerful tool and has yielded useful imaging and composition information on objects such as Cassiopeia A. Higher sensitivity (by a factor of approximately 10) is required to obtain spectra of extragalactic supernovae. Wide-field X-ray and UV surveys should yield observations of shock breakout from core-collapse supernovae, which can be used to determine the radii of the progenitor stars and to tie down the times of the explosions. To date, this has been observed only for the Type Ib supernova SN 2008D, which had a relatively compact progenitor and produced fainter, harder shock-breakout emission. The more common Type II supernovae emit longer, brighter transients, chiefly in the ultraviolet. Swift, which discovered SN 2008D, and GALEX have made a start on this problem.

How Do the Stars Explode?

On the theoretical front, it is not currently understood which stars leave behind neutron stars and which leave behind black holes, nor is it understood what determines the observed masses, spins, and magnetic fields of these compact objects (see discussion of science question SSE 4). In the next decade, the exponential increase in computer power, plus the wide availability of multidimensional codes that scale well on these big machines, will lead to major advances in the understanding of how all kinds of stars explode. It is just now becoming feasible to do three-dimensional simulations of stellar core collapse with multigroup neutrino transport coupled to the hydrodynamics, an essential for a first-principles model. By the middle of the decade, such calculations will be commonplace and should, at a minimum, answer the long-standing problem of whether the simplest (non-rotating) neutrino-transport model works for any range of stellar masses. What about the rest? Driven in part by the interest in gamma-ray bursts, codes that realistically couple magnetic fields, rotation, and general relativity in the context of the core-collapse problem are also in an advanced stage of development. Neutrinos have not yet been included in these codes in a realistic way, but before the end of the decade they will. By the end of the decade there should be a quantitative theory of how massive stars of all masses die—given the pre-supernova characteristics. While computational modeling of core-collapse supernovae may produce the most spectacular results, analytic "pencil and paper" theory will also remain crucial for the interpretation of numerical results, for refining the microphysics that goes into the simulations, and for exploring uncertain aspects of stellar evolution and core collapse.

On the observational front, spectroscopic and spectropolarimetric observations of supernovae will ensure major advances in studies of the composition, kinematics, and asphericity of the explosion at a time when numerical models will become increasingly predictive. When combined with time-domain surveys that will discover new and rare events, these observations will significantly advance the understanding of core-collapse explosion physics. As multidimensional simulations become, it is hoped, increasingly successful at producing supernovae, comparison to the observed polarimetry is likely to become particularly productive. Large telescopes (10 to 30 m) are needed for spectropolarimetric observations, while smaller telescopes, which can spectroscopically study broad lines (1,000 km/s resolution), are important for providing input for modeling supernova parameters. Estimates of supernova energy are useful constraints on the supernova mechanism. The observation of neutron-star kicks received during a supernova provides another constraint on the symmetry of the explosion mechanism.

A gravitational wave and/or neutrino detection of a nearby SN in the next decade would provide a wealth of unique information about the dynamics in the central approximately 100 km during core collapse. Such a detection would provide stringent and unique tests of theoretical models as well as a laboratory for studying the properties of the neutrino itself. Unfortunately, a neutrino detection can only be made for a very nearby (e.g., in the galaxy or Large Magellanic Cloud [LMC]) supernovae, and the prospects for gravitational wave detections depend on uncertain core-collapse physics.

Only about 1 in 100 massive-star deaths end in a GRB. The distinguishing properties of the stars that end this way are not known, although a high central angular momentum is likely to play a role. A combination of theoretical studies of stellar evolution and larger observational samples of GRB-SN associations and more "orphan afterglows" are needed for progress on this question.

Gamma-ray line signals from radioactive decays offer a way of sampling SN nucleosynthesis and dynamics in a most direct way, if the signals can be seen. The strongest gamma-ray line expected from galactic supernova remnants that are centuries old comes from the decay of ^{44}Ti. The only remnant detected to date is Cas A, which is about 330 years old. The Nuclear Spectroscopic Telescope Array (NuSTAR) should provide more detections of ^{44}Ti through the 68- and 78-keV hard X-ray lines. Determining the amount of ^{44}Ti in an explosion is important for determining the mass cut between the central compact object and the ejected material. More generally, nuclear physics is a crucial ingredient in core-collapse SN calculations. This includes the nuclear equation of state, rates for electron capture on nuclei, properties of neutron-rich nuclei produced in the r-process, and, potentially, the physics of neutrino oscillations in the core-collapse environment. In the next decade, advances in observations, theory, and laboratory nuclear physics will likely determine whether some core-collapse SNe can produce r-process nuclei (approximately half of the nuclei above the iron peak); this would solve one of the major puzzles regarding the origin of heavy elements.

Blasts from the Past

The most distant point source in the universe is now a GRB at redshift 8.3. GRBs can be used as tools for exploring stellar evolution throughout cosmic time because of their brightness and the penetrating power of gamma rays. GRBs may be the deepest direct probe of first-generation stars. In addition, observations of the composition of *nearby metal-poor stars* provide information on the chemical traces of early supernovae and their evolution, and thus provide a critical piece of information for understanding supernovae. Currently, the number of very metal-poor stars known is extremely small, and larger samples are critical for progress. This requires large-scale surveys and high-resolution spectroscopic follow-up with optical telescopes up to 30 m in diameter.

SSE 4. What Controls the Mass, Radius, and Spin of Compact Stellar Remnants?

The deaths of stars give rise to compact stellar remnants—neutron stars, black holes, and white dwarfs—that produce the most exotic and energetic phenomena in the universe, from the brightest known sources of radiation to the steadiest astrophysical clocks. The properties of compact stellar remnants provide not only unique information about the late stages of stellar evolution, but also a testing ground for studying physics that is otherwise inaccessible to astronomers and physicists alike.

Many fundamental questions about compact stellar remnants center on understanding, both empirically and theoretically, their basic physical properties, including mass, radius, spin, and magnetic field. Here several key questions in which major progress can be made in the coming decade are highlighted.

What Is the Equation of State of Ultradense Matter?

The basic properties of neutron stars are closely coupled to the physics of their interiors. Neutron star cores have densities up to several times that found in atomic nuclei ($\rho \sim 3 \times 10^{14}$ g cm^{-3}) and are therefore the densest objects, without event horizons, known. Because the equation of state for such ultradense matter is still poorly constrained, the basic compositions of neutron star cores are unknown, with exotic new states of matter (e.g., deconfined quark matter or Bose condensates) possible. This high-density regime is mostly inaccessible to terrestrial laboratories, but its properties determine the mass-radius relation of neutron stars, providing an astrophysical probe (Figure 5.8). In the past decade, the ability to measure masses of binary radio pulsars has dramatically improved, and several X-ray techniques for constraining neutron-star radii have been developed.

Radio pulsars are the most commonly observed neutron stars known, with almost 2,000 cataloged to date. Precise neutron-star-mass determinations can be obtained in "recycled" binary millisecond pulsars by the measurement of relativistic orbital effects, but only a handful of suitable systems are known. Pulsars with unusually high or low mass directly constrain the dense matter equation of state, and in fact the measurement of a single neutron star with mass $>2\ M_\odot$ would rule out most forms of exotic material in neutron star cores. Current and future galactic surveys for millisecond radio pulsars are crucial for detecting new systems for neutron-star-mass measurements (currently only approximately 1 percent of pulsars are suitable), as well as finding other exotic pulsars for gravitational-wave-detection experiments and strong-field gravity tests. The past 5 years have seen the discovery of several eccentric binary millisecond radio pulsars that contain neutron stars likely more massive than $1.7\ M_\odot$.

Radio surveys with current or soon-to-be-available facilities (such as the Green Bank Telescope, Arecibo Observatory, Parkes Observatory, Effelsberg Radio Telescope, and the EVLA) should double the number of known radio pulsars in the next decade and (because of computational and instrumentation improvements) should increase the number of millisecond radio pulsars by an even larger factor. Further in the future, the Chinese Five-hundred-meter Aperture Spherical Telescope (FAST) could find approximately 5,000 new radio pulsars by 2020, while a

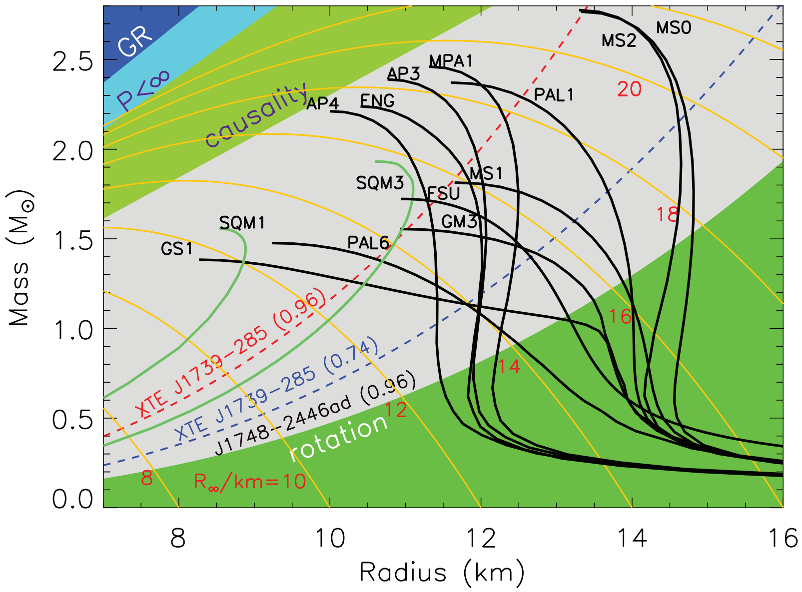


FIGURE 5.8 The radius of a neutron star is sensitive to details of the equation of state for matter at near-nuclear densities. Measurement of masses and radii for neutron stars in the next decade should suffice to rule out some of the equations of state displayed here (as black and green curves) in a figure from Lattimer and Prakash, 2007. The green region at the lower right is excluded from observations of pulsar J1748-2446ad. The recent discovery of pulsars with $M > 1.6\ M_\odot$ rules out equation-of-state models that do not extend to such high masses. SOURCE: Reprinted from J.M. Lattimer and M. Prakash, Neutron star observations: Prognosis for equation of state constraints, *Physics Reports* 442(1-6):109-165, copyright 2007, with permission from Elsevier.

large-area decimeter-wavelength radio telescope could find approximately 20,000 new radio pulsars, including thousands of new millisecond pulsars. Gamma-ray surveys with facilities such as Fermi Gamma-ray Space Telescope will likely also contribute, albeit with smaller numbers. Current-generation pulsar surveys and timing observations are completely sensitivity-limited, and larger telescope collecting areas are needed.

The most extreme outliers from such surveys may be of greatest interest. A recent example is the so-called Double Pulsar J0737-3039, an exceptionally relativistic pulsar binary discovered in 2003. That system has provided the most precise tests of general relativity in the strong-field regime to date, and long-term timing (which may require a span of 10 years or more) will eventually allow the measurement of spin-orbit effects on periastron advance, which in turn will determine the neutron-star moment of inertia, strongly constraining the neutron-star equation of state (since the neutron-star mass is already precisely known in that system).

X-ray observations hold great promise as well. Measurements of neutron-star radii by means of X-ray timing are possible by observing certain bright low-mass X-ray binaries (LMXBs) as well as nearby faint isolated millisecond pulsars. For the LMXBs, observations with the Rossi X-ray Timing Explorer (RXTE) have identified both millisecond oscillations during thermonuclear X-ray bursts and longer-lasting accretion-powered millisecond pulsations that encode information about neutron star parameters. Pulse shape and spectral modeling of these phenomena with a larger-area X-ray timing instrument can strongly constrain both neutron-star radius and mass. Broadband X-ray spectroscopy of Eddington-limited radius-expansion bursts in many of these same systems can also independently constrain neutron-star parameters. Neutron-star radii and masses can also be constrained through pulse shape modeling of the faint thermal pulsations seen from some isolated millisecond pulsars. Recent observations with X-ray Multi-Mirror Mission (XMM)-Newton have demonstrated this technique, but a more-sensitive, focusing X-ray telescope is required to obtain strong constraints. In addition, soft X-ray spectroscopic observations of transiently accreting neutron stars in quiescence show thermal X-ray spectra from the cooling neutron star surfaces that yield the neutron star radius when the source distance is known and the atmospheric model is correct. The measurement of these faint targets requires a sensitive focusing X-ray telescope with moderate spectral resolution.

Finally, laboratory nuclear physics experiments are expected to provide complementary constraints on the nuclear-matter equation of state. These measurements will constrain some aspects of the neutron-star equation of state that can serve as input to neutron-star models, allowing one to interpret observations, probe models, and constrain the regimes of the neutron-star equation of state not accessible in the laboratory. Major advances will be possible in the next decade. Examples include constraints on the nuclear symmetry energy around nuclear density from precision measurements of the neutron skin thicknesses of heavy nuclei, using parity-violating electron scattering, and—for higher densities—from heavy-ion collisions at a range of energies and asymmetries at various advanced rare isotope facilities, including the Facility for Rare Isotope Beams (FRIB). Nuclear-theory work to identify the most useful signatures of equation-of-state properties in heavy-ion collisions and how to interpret them quantitatively is also needed.

What Is the Spin Distribution and Maximum Spin of Neutron Stars and Black Holes?

Millisecond pulsars (MSPs) are neutron stars that have undergone mass transfer from companion stars in LMXBs and have been "spun up" to rapid rotation rates in the process. The physical processes that stop the transfer of angular momentum and thereby establish the maximum spin rates of neutron stars are currently unknown. One of the primary candidates is the emission of gravitational radiation which, if correct, would be of major importance to current and future gravitational-wave-detection facilities such as the Advanced Laser Interferometer Gravitational Wave Observatory (LIGO). Identifying the correct spin frequency distribution of MSPs will help to determine both the maximum spin rates of neutron stars and the limiting physical processes. Current and future radio and X-ray timing surveys have many fewer selection effects toward rapidly rotating pulsars than those in the past. If such systems are detected, they would directly limit the neutron-star equation of state by determining the maximum radius of the neutron star as a function of its mass for which it does not shed material at its equator. New constraints on neutron-star physics would come from the detection of a neutron star spinning more rapidly than 1,000 Hz.

Astrophysical black holes are completely described by just two quantities, their mass and spin. Although the masses of stellar black holes in X-ray binaries have been measured dynamically for decades, it is only in the past few years that it has become possible to constrain the spins of black holes. The spin is constrained by determining the inner radius of the accretion disk, either by fitting the thermal disk component of the X-ray continuum spectrum, or through the relativistically broadened shape of the Fe K disk fluorescence line. The radius inferred by these methods is believed to be comparable to that of the last stable orbit in general relativity, but there are systematic uncertainties in this association that limit the precision of current constraints on black hole spin. Inferences about spin have now been made in 10 systems (using a variety of X-ray missions, most recently including Chandra, XMM-Newton, and Suzaku) and most are rotating significantly, with a wide variety of black hole spins measured, and several are believed to be spinning near the maximal amount allowed by general relativity. A slowly spinning, disk-accreting black hole must double its mass in order to spin rapidly, which is impossible for a black hole in an X-ray binary; thus, the measured spin distribution is essentially sampling the birth properties of these black holes. An alternative way of measuring black hole spin is through the spin-orbit coupling of a pulsar with a black hole. This method will require pulsar searches to discover pulsars with black hole companions. Gravitational-wave detection of stellar black-hole/black-hole or black-hole/neutron-star inspirals offers still another promising method for measuring black hole spins (and masses).

A larger sample of black hole spin measurements will provide very strong constraints on models of massive star evolution and core-collapse supernovae. More broadly, an improved understanding of black hole spin can be used to address a number of important issues, including the role of black hole spin in producing jets and in powering GRBs. Accurate knowledge of the black hole spin distribution is also crucial for designing theoretical search templates required for the direct detection of gravitational waves from black-hole/black-hole and black-hole/neutron-star mergers. In order to make continued progress, soft X-ray continuum spectroscopy and medium- to high-resolution X-ray line spectroscopy of a larger sample of black holes in X-ray binaries is needed, requiring a more sensitive telescope than Chandra or XMM-Newton. In addition, further numerical and theoretical work on general-relativistic MHD models of black hole accretion disks is essential for interpreting X-ray continuum observations, and better theoretical models of the X-ray irradiation and fluorescent Fe line emission from the inner accretion disk are needed to interpret Fe line spectra.

What Determines the Initial-Final Mass Relation Connecting Progenitors to White Dwarfs?

Observational constraints on the initial-final mass relation come primarily from white dwarfs in open clusters and require accurate main-sequence turnoff ages plus white-dwarf cooling ages (to infer the initial mass) and precise final masses (Figure 5.9). The largest source of error is the variance in theoretical mass-lifetime relationships for stars with convective cores (and, by extension, open cluster ages). Gaia will provide precise distances and membership information for open clusters, and asteroseismology from missions such as CoRoT and Kepler may constrain the sizes of convective cores. Uncertainties in the white-dwarf cooling timescale also need to be addressed, in particular the properties of the atmospheric "blanket" that governs heat transport and is a poorly understood by-product of asymptotic giant branch (AGB) evolution. The mass of the blanket is a result of the processes that end the AGB evolution of the star. The recent discovery of carbon-atmosphere white dwarfs is as yet unexplained and points to interesting discovery areas in white-dwarf formation. Also crucial to understanding white-dwarf properties is the onset of crystallization, which alters the internal energy structure and causes an abrupt change in effective temperature and luminosity. Both have been constrained by asteroseismology on a small number of stars, giving a partial picture and great promise for future advances.

The mass of a white dwarf originating from a single star such as the Sun is related to the luminosity of the star as it leaves the AGB. This luminosity, and thus the white-dwarf mass, is determined by the mass-loss process. Theoretical and observational studies of the dependence of mass-loss rates on stellar parameters have not reached a consensus, and prescriptions advocated and used differ dramatically from one another. Empirical and theoretical formulas span a wide range of slopes. To match observed initial-final mass relations with evolutionary and population models, empirical laws have been "corrected" with a variety of tuning parameters. However, there is no widely accepted or demonstrably correct mass-loss formula for the mass loss that produces white-dwarf stars, and thus no predictive power for extrapolating to understudied populations (such as young, low-metallicity cases).

To measure mass-loss rates for large numbers of stars, infrared surveys and

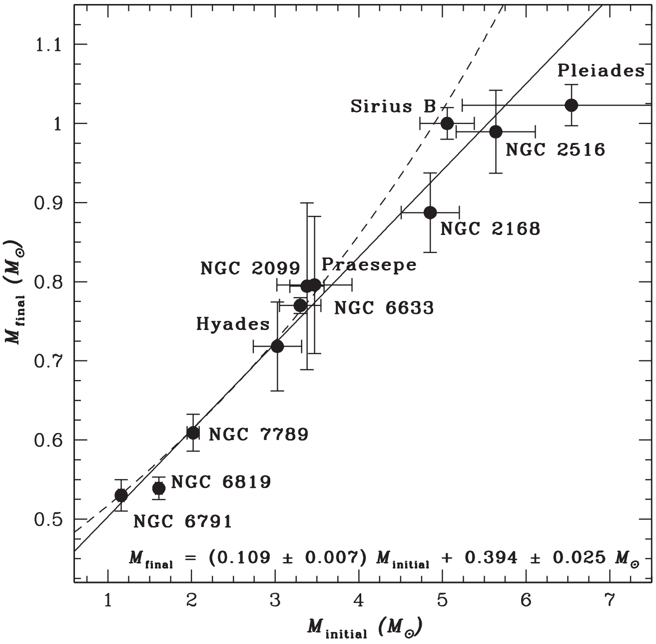


FIGURE 5.9 The fates of stars—the relationship between their initial and their final mass—for stars that lose mass and thereby (if they start with >1.4 $M_\odot$) avoid exploding as supernovae, are observationally constrained by the study of white dwarf stars in clusters of known age. Considerable uncertainty exists for intermediate-mass stars (2 to 8 $M_\odot$), including how much initial composition affects the result. With the detection of many more white dwarfs in clusters, enough information to constrain mass-loss models and perhaps enough to extrapolate to unobservable populations (low metallicity, high mass—as in the early universe) may become available. SOURCE: J.S. Kalirai, B.M.S. Hansen, D.D. Kelson, D.B. Reitzel, R.M. Rich, and H.B. Richer, The initial-final mass relation: Direct constraints at the low-mass end, *Astrophysical Journal* 676(1):594, 2008, reproduced by permission of the AAS. Courtesy of Jason Kalirai, Space Telescope Science Institute.

molecular-line surveys have already proven useful, although the IR measures require reliable gas/dust ratios and the radio lines measure mass loss at a different time (farther out in the flow); many mass-losing stars have variable outflows. The modeling of mass-loss processes requires non-local thermodynamic equilibrium hydrodynamics with shocks, non-equilibrium chemistry, and grain nucleation and growth. Strong tests of models are coming from interferometric studies of the structures of the atmospheres of mass-losing stars, as, for example, the discovery of "molecular shells" at about twice the stellar radius and coincident with the region where dust is expected to form. Molecular lines are ideal for mass-loss studies because (1) they trace the gas, and (2) they carry velocity information. CO has been widely used in our galaxy. The Atacama Large Millimeter Array (ALMA) should be able to detect CO and thus measure mass-loss rates for AGB and red giant branch (RGB) stars in the Magellanic Clouds, a vital laboratory for the study of these populations. Additionally, CO traces mass loss for both C-rich and O-rich giants, which is not the case for tracers currently used to study the highest-mass-loss-rate objects in the Small Magellanic Cloud and LMC.

Most white-dwarf stars are believed to be composed of carbon and oxygen. Observations of novae show that some O-Ne-Mg white dwarfs are formed, presumably from relatively high mass progenitors, with details of the formation channel(s) as yet unclear. There are also He white dwarfs, including a surprisingly large population of He white dwarfs in very metal-rich clusters and a population of single-field He white dwarfs. Explaining the origin and evolution of these different classes of objects should be illuminating. The formation of He white dwarfs is understood only in the context of binaries, so further understanding of other ways of forming them potentially by single stars or through disrupted binaries or ejections from dense star clusters is needed. Stars that ignite C in their degenerate cores before losing their envelopes to mass loss may also produce unusual thermonuclear supernovae. Very little is known about this potential channel, but large surveys should yield valuable information.

SSE DISCOVERY AREA: TIME-DOMAIN SURVEYS

Astronomical timescales evoke the long stretches of time, reckoned in gigayears, that characterize cosmic expansion and most phases of stellar evolution. For these phenomena, a single comprehensive survey can reveal the essential facts, as in a Hertzsprung-Russell diagram for a cluster. But there are phenomena of rotation and pulsation, of orbiting binaries, of explosions and mass loss, and most spectacularly, of stellar death, for which the physical timescales are measured in seconds, days, or months. For a wide range of stellar events, knowledge has been obtained by observing through narrow windows in time, often set by single-investigator observing strategies or by the technical capabilities of the detectors being used. Narrow windows produce limited views. The panel anticipates that in the coming decade the burgeoning technological change due to detector development, fast computers, automated pipelines, and the ability for the entire community to interact with large-volume public databases (from a distance, over the Internet) will lead to significant scientific progress in revealing and exploring a wide range of stellar phenomena. For these reasons, time-domain surveys represent a significant discovery potential for the study of stars and stellar evolution.

Discovery in the time domain in the next decade will be driven by detectors with large fields of view, which scan the sky with approximately daily-weekly cadence and provide all-sky data sets. In addition to unanticipated discoveries (Figure 5.10), there are expected events; Table 5.2 gives a sample of the wide range

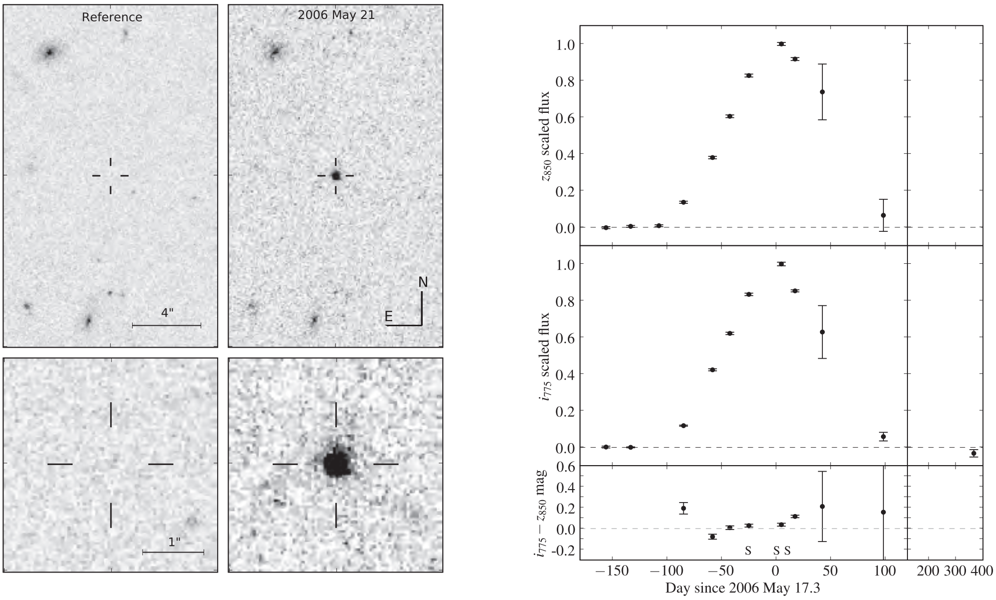


FIGURE 5.10 An unusual optical transient with a nearly symmetric light curve was discovered during the Hubble Space Telescope Cluster Supernova Survey. Most explosive events have a rapid rise and a slower decline; this symmetry is anomalous. From its redshift and comparison with other recent supernova discoveries, the object was found to be an unusual, luminous supernova. Surprising, rare events continue to be discovered as time-domain surveys expand in reach and duration and improve in cadence and precision. SOURCE: K. Barbary, K.S. Dawson, K. Tokita, G. Aldering, R. Amanullah, N.V. Connolly, M. Doi, L. Faccioli, V. Fadeyev, A.S. Fruchter, G. Goldhaber, A. Goobar, A. Gude, X. Huang, Y. Ihara, K. Konishi, M. Kowalski, C. Lidman, J. Meyers, T. Morokuma, P. Nugent, S. Perlmutter, D. Rubin, D. Schlegel, A.L. Spadafora, N. Suzuki, H.K. Swift, N. Takanashi, R.C. Thomas, and N. Yasuda for the Supernova Cosmology Project, Discovery of an unusual optical transient with the Hubble Space Telescope, *Astrophysical Journal* 690(2):1358, 2009, reproduced by permission of the AAS. Courtesy of Kyle Barbary, University of California, Berkeley, Lawrence Berkeley National Laboratory.

of stellar science that can be addressed with this type of time-domain survey, at many wavelengths. Follow-up observations of various types are often essential to carry out the science goals. For example, evolved giants and brown dwarfs observed interferometrically show evidence for time-variable spatial structures, possibly associated with dust-cloud formation and weather-like phenomena; follow-up observations with new interferometric facilities will provide important constraints on the physical mechanisms behind the observed time variations. Other kinds of time-domain studies not mentioned in Table 5.2 will also be valuable—particularly, continuous monitoring observations with high cadence for extended duration on individual objects, as in the case of asteroseismology.

This panel's four science questions all mention time-domain observations as

TABLE 5.2 Time-Domain Surveys, Large Field of View, All-Sky Coverage, Daily-Weekly Cadence

Science	Survey Capabilities	Follow-up
Variable Stars and the Sun		
Starspots and rotation	Optical/multicolor	Spectroscopy, spectropolarimetry
Massive stars—LBVs	Optical/IR/UV	Spectroscopy, photometry, spectropolarimetry
Eclipsing binaries	Optical/multicolor, long duration	Photometry, spectroscopy, radial velocities
Clouds and weather on brown dwarfs	Optical/IR	Multiwavelength photometry, spectroscopy; improved models
Pulsating variables—classical, rare	Optical/IR, long duration, Milky Way and nearby galaxies, helio- and asteroseismology	Photometry, spectroscopy; interferometry; improved models
Rare stages of stellar evolution	Optical/IR, long duration, clusters and nearby galaxies	Photometry, spectroscopy; improved models; interferometry
Stellar mergers on dynamical or thermal timescale	Optical/IR, long duration, resolved stellar populations	Photometry, spectroscopy; improved models
Stellar flares	Optical (blue, U/u filter), UV, X-ray, radio	Time-resolved, multiwavelength photometry, spectroscopy; improved models
Solar corona	Optical/IR/radio magnetometry	ATST follow-up of energetic events
Pulsars—rare types	Radio	Radio
Magnetar flares	X-ray, γ-ray	Multiwavelength photometry, spectroscopy
Variable Accreting Systems		
CVs, novae	Optical, UV	Spectroscopy—optical, UV, IR
Tidal disruption of stars	Optical, UV, X-ray, radio	Multiwavelength photometry, spectroscopy; host galaxy properties
LMXBs (black-hole/neutron-star novae, X-ray bursters, superbursters)	X-ray, wide-field	Time-resolved X-ray photometry, spectroscopy
Supernovae and GRBs		
Milky Way supernova	Optical/IR, ν, gravitational waves, γ-ray, IR, radio	Multiwavelength photometry, spectroscopy; rapid response
Supernova searches including rare forms, optical transients	Optical/IR	Multiwavelength photometry, spectroscopy
Shock breakout in supernovae Types II and Ibc	UV, X-Ray	Multiwavelength photometry, spectroscopy
Electromagnetic counterparts to gravitational wave sources	Optical/IR, UV, X-ray	Photometry, spectroscopy—radio; rapid response
Gamma-ray bursts	γ-ray, X-ray	Multiwavelength photometry, spectroscopy; rapid response
Orphan afterglows of GRBs	Optical/IR, radio	Multiwavelength photometry

NOTE: Acronyms are defined in Appendix C.

essential to making progress in the next decade (e.g., supernova searches, rotation, episodic mass loss from massive stars, pulsar searches). A few illustrative examples of additional stellar topics from Table 5.2 that have particular discovery potential are described in more detail below.

Accretion-Induced Collapse (Rare Type of Supernovae)

How often does a white dwarf approaching the Chandrasekhar mass in a binary undergo an accretion-induced collapse (AIC) to form a neutron star, rather than blowing up as an SN Ia? An understanding of this question is essential for understanding the evolution of white dwarfs in binary systems and would dramatically constrain the allowed progenitors of SNe Ia. AIC has also been proposed as one of the most promising sites for third-peak *r*-process nucleosynthesis. AICs are predicted to be accompanied by the ejection of up to approximately 0.01 to 0.03 $M_\odot$ of Ni, produced in an accretion disk around the newly formed neutron star. This outflowing Ni produces a short, approximately 1 day, optical/near-IR SN-like transient with a peak luminosity of 10^{41} to 10^{42} ergs/s, significantly fainter and of shorter duration than ordinary SNe Ia.

Electromagnetic Counterparts to Gravitational Wave Sources

In the coming decade, it is likely that transient gravitational wave sources will be discovered by experiments such as Advanced LIGO and VIRGO (a gravitational wave detector at the European Gravitational Observatory), with lower-frequency gravitational wave sources perhaps becoming detectable toward the end of the decade. To optimize the astrophysics that results from such detections, it is critical to have nearly simultaneous electromagnetic observations. Wide-field-of-view cameras are essential given the rather poor localizations provided by gravity-wave detectors (fractions of a square degree). A unique electromagnetic counterpart temporally and spatially coincident with a gravitational wave source would provide more confidence in the gravitational wave detection. Combined gravitational wave and electromagnetic observations could potentially provide detailed information about stellar sources, including neutron-star/neutron-star mergers, black-hole/black-hole mergers, short gamma-ray bursts, and (perhaps) core-collapse supernovae. This information is unique given the gravitational wave constraints on the masses and spins (magnitudes and direction) of the objects.

Rare Stages of Stellar Evolution (He Core Flash)

At the tip of the first-ascent RGB, for stars of $M < {\sim}2.5\ M_\odot$, He ignition in the degenerate core leads to the He core flash. Very little is known observationally about what happens during the He core flash: for example, does the star lose mass? The ultimate fate of Earth—into the Sun or backing away—depends on whether the Sun will lose 20 percent of its mass before or during this event. Yet one cannot point to a single object in the sky that is currently undergoing an He core flash. The probable duration of this phase is about a thousand years, assuming that the star's response resembles its response to He shell flashes, evolving on a thermal timescale. One probable observational signature will be erratic pulsation with rapid period changes. Large samples, including stars in clusters of various ages, are needed to identify individual objects in this critical stage of stellar evolution. These data will provide essential input to theoretical models of late-stage stellar evolution.

Eclipsing Binaries and Binary Star Evolution

Eclipsing binary stars are powerful diagnostics of stellar structure and evolution, and they are relatively easy to find in time-domain surveys (Figure 5.11). They also provide a secure way to measure masses and radii for stars of all spectral types, metallicities, and ages. Longer-period eclipsing binaries, although geometrically less favorable, should be found in large surveys with long durations. Such stars will be valuable because their evolution is less likely to be impacted by the presence of the companion (synchronous rotation, enhanced activity), which may influence the stellar radius. Large samples will permit tests of the mass ratios and distributions of orbital separations of close binaries, which will in turn inform theories of binary-star formation and evolution. Studies of interacting binaries will also benefit, particularly for unusual and rare systems such as contact binaries and common-envelope systems, probable precursors to stellar mergers (leading, for example, to blue stragglers). Investigations of white-dwarf/white-dwarf and white-dwarf/massive-star systems are important for understanding the origin of SNe Ia and cataclysmic variables. For very wide eclipsing binaries and stars with large planets, detailed information on the resolved stellar surface can be obtained during the eclipse/transit.

Radio Transients

Stars of all kinds produce a surprisingly wide variety of nonthermal radio emission from timescales of nanoseconds (the giant pulses from the Crab pulsar), to months (the radio afterglows of supernovae). Recently, several relatively small-scale radio surveys have uncovered new forms of transients from known sources, such as extremely rare millisecond-duration pulses from rotating neutron stars (the so-called RRATs, or rotating radio transients), and bright coherent emission from brown dwarfs. Other surveys have found unidentified radio transients in extragalactic blank fields and toward the galactic center. Yet these surveys have covered

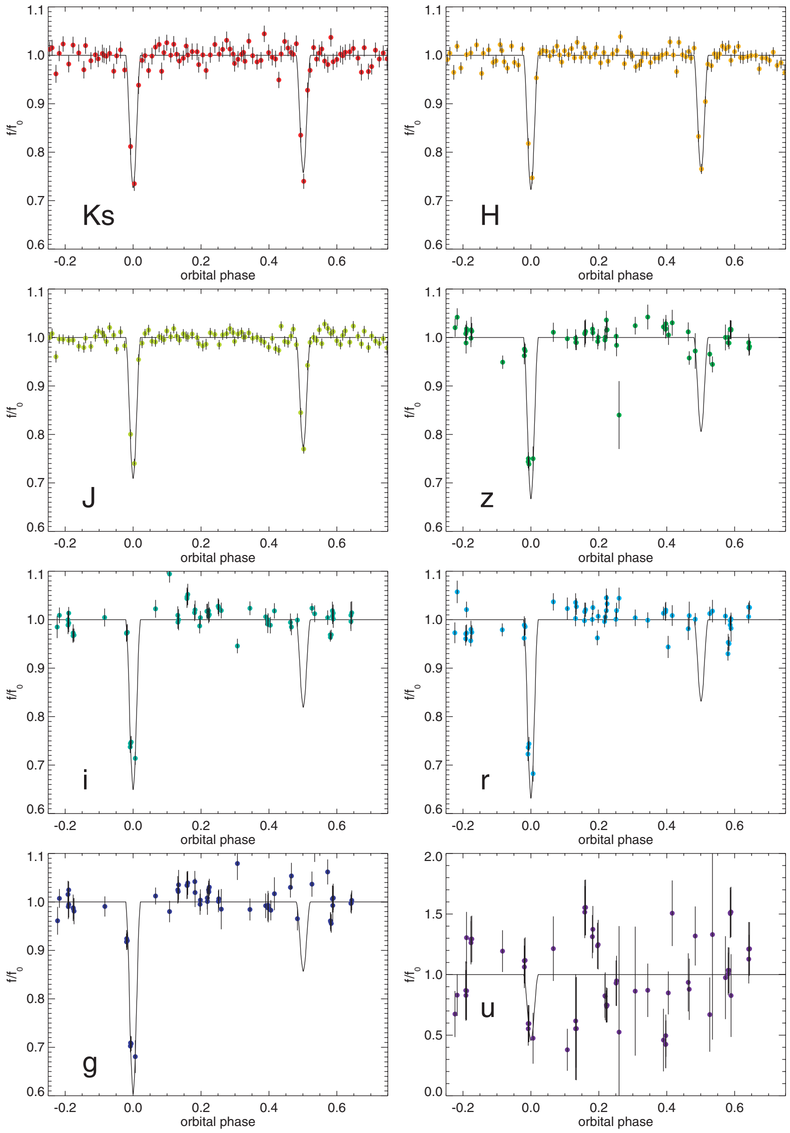


FIGURE 5.11 Multicolor light curves for an eclipsing binary, from Sloan Digital Sky Survey and 2 Micron All Sky Survey data. The two stars are low-mass, M0 + M1, dwarfs with an orbital period of 2.639 days. Time-domain surveys will provide large numbers of new eclipsing binary systems for stars across the Hertzsprung-Russell diagram, allowing for much-improved basic data for unusual as well as common types of stars. Follow-up for systems such as this one—for example to get radial velocities—will require large telescopes and/or substantial telescope time. SOURCE: Reprinted with permission from A.C. Becker, E. Agol, N.M. Silvestri, J.J. Bochanski, C. Laws, A.A. West, G. Basri, V. Belokurov, D.M. Bramich, J.M. Carpenter, P. Challis, et al., Two-Micron All-Sky Survey J01542930+0053266: A new eclipsing M dwarf binary system, *Monthly Notices of the Royal Astronomical Society* 386:416, 2008, copyright 2008 Royal Astronomical Society.

either only tiny fractions of the sky or a very small range of timescales. The situation is improving with the development of the Low Frequency Array (LOFAR), the Long Wavelength Array, the Murchison Widefield Array, and ATA-42 (the Allen Telescope Array, configured with 42 radio-telescope dishes). As radio fields of view continue to increase and computing capability grows to allow wide-field, rapid-cadence, radio imaging, new surveys will uncover many more transient events of both known and unknown origin. These events have the potential to tell about particle acceleration, stellar magnetic fields and rotation, strong-field gravity, the interstellar and intergalactic media, the violent deaths of stars, and possibly physics beyond the standard model.

Summary of SSE Discovery Area

In summary, the time domain represents great discovery potential well matched to the timescales that are relevant for stellar phenomena during their lifetimes and their death throes. Astronomers look forward to the next decade as a period of renaissance for stellar astronomy as time information is added to the new advances in three-dimensional spatial resolution and the idealization of a star as a static, spherical object is put to bed.

Summary Findings

TABLE I Summary of Science Frontiers Panels' Findings: Identification of Key Science Questions and Areas of Unusual Discovery Potential

Panel	Science Questions		Area(s) of Unusual Discovery Potential
Cosmology and Fundamental Physics	CFP 1	How did the universe begin?	Gravitational wave astronomy
	CFP 2	Why is the universe accelerating?	
	CFP 3	What is dark matter?	
	CFP 4	What are the properties of neutrinos?	
Galactic Neighborhood	GAN 1	What are the flows of matter and energy in the circumgalactic medium?	Time-domain astronomy, astrometry
	GAN 2	What controls the mass-energy-chemical cycles within galaxies?	
	GAN 3	What is the fossil record of galaxy assembly from the first stars to the present?	
	GAN 4	What are the connections between dark and luminous matter?	
Galaxies Across Cosmic Time	GCT 1	How do cosmic structures form and evolve?	The epoch of reionization
	GCT 2	How do baryons cycle in and out of galaxies, and what do they do while they are there?	
	GCT 3	How do black holes grow, radiate, and influence their surroundings?	
	GCT 4	What were the first objects to light up the universe, and when did they do it?	
Planetary Systems and Star Formation	PSF 1	How do stars form?	Identification and characterization of nearby habitable exoplanets
	PSF 2	How do circumstellar disks evolve and form planetary systems?	
	PSF 3	How diverse are planetary systems?	
	PSF 4	Do habitable worlds exist around other stars, and can we identify the telltale signs of life on an exoplanet?	
Stars and Stellar Evolution	SSE 1	How do rotation and magnetic fields affect stars?	Time-domain surveys
	SSE 2	What are the progenitors of Type Ia supernovae and how do they explode?	
	SSE 3	How do the lives of massive stars end?	
	SSE 4	What controls the mass, radius, and spin of compact stellar remnants?	

Part II

Reports of the Astro2010 Program Prioritization Panels

6

Report of the Panel on Electromagnetic Observations from Space

SUMMARY

NASA's support of astrophysics research is an essential element in the world-class accomplishments of U.S. astronomers in their exploration of the cosmos. In addition, as was aptly expressed in the first finding of the congressionally requested National Research Council (NRC) study *An Enabling Foundation for NASA's Earth and Space Science Missions*, "The mission-enabling activities in SMD [NASA's Science Mission Directorate]—including support for scientific research and research infrastructure, advanced technology development, and scientific and technical workforce development—are fundamentally important to NASA and to the nation."[1]

The Astro2010 Program Prioritization Panel on Electromagnetic Observations from Space (the EOS Panel) reviewed current astrophysics activities supported primarily by NASA's Science Mission Directorate—specifically, those activities requiring electromagnetic observations from space as distinct from observations of particles or gravitational waves. The charge of the panel was to study possible future activities and to recommend to the Astro2010 Survey Committee a scientifically compelling, balanced, affordable, and relatively low-risk program for the 2010-2020 decade.

---

[1] National Research Council, *An Enabling Foundation for NASA's Earth and Space Science Missions*, The National Academies Press, Washington, D.C., 2010, p. 2.

Guided by the science opportunities identified in the reports of the decadal survey's five Science Frontiers Panels (Chapters 1 through 5 in this volume) and within the framework of current and in-process facilities and programs available to the astrophysics community, the panel formulated the program described below for electromagnetic space missions for the 2010-2020 decade. In the process of formulating this program, the panel reviewed nearly 100 written submissions from the astronomy and astrophysics community describing a broad range of potential facilities, required tools, and needed technology development, as well as thought-provoking manifestos on process and principles.

The program recommended by the panel reflects its judgment that, in the 2010-2020 decade—with many scientifically compelling space missions to choose from but with a tightly constrained budget—the highest priority is for programs that will have a major impact on many of the most important scientific questions, engaging a broad segment of the research community.

The panel's recommended program is divided into large activities and moderate/small activities. The panel expresses emphatic support for a balanced program that includes both. The three large initiatives—the Wide-Field Infrared Survey Telescope (WFIRST) Observatory mission, the International X-ray Observatory (IXO) mission, and an exoplanet mission—are presented in prioritized order.

The four moderate/small activates are not prioritized. The panel's recommended program calls for strong support of all four activities, although one—the Space Infrared Telescope for Cosmology and Astrophysics and the Background-Limited Infrared-Submillimeter Spectrograph (SPICA/BLISS)—has de facto priority because of its time-critical nature. The moderate/small initiatives are the SPICA/BLISS initiative, augmentation of NASA's Explorer program for astrophysics, technology development for a Hubble Space Telescope (HST) successor, and augmentation of NASA research and analysis (R&A) programs in technology development and suborbital science. The relative levels of support for these activities would depend on factors that cannot be forecast in detail, such as (1) the future funding level of NASA's Astrophysics Division base budget and (2) science opportunities and cost-benefit trade-offs. In the final section ("Funding a Balanced Program") of this report the panel recommends funding levels across the program that address these issues for three different budget projections for the Astrophysics Division.

Large Initiatives

Wide-Field Infrared Survey Telescope

The WFIRST Observatory is a 1.5-m telescope for near-infrared (IR) imaging and low-resolution spectroscopy. The panel adopted the spacecraft hardware of the Joint Dark Energy Mission (JDEM)/Omega mission as proposed to NASA and the Department of Energy (DOE) and substantially broadened the program for this facility. In addition to two dedicated core programs—cosmic acceleration and microlensing planet finding—WFIRST would make large-area surveys of distant galaxies and the Milky Way galaxy, study stellar populations in nearby galaxies, and offer a guest observer program advancing a broad range of astrophysical research topics.

International X-ray Observatory

The IXO mission, a proposed collaboration of NASA, the European Space Agency (ESA), and the Japan Aerospace Exploration Agency (JAXA), will revolutionize X-ray astronomy with its large-aperture and energy-resolving imager. IXO will explore the role of feedback in galaxy evolution by connecting energetic processes within galaxies with the physical state and chemical composition of hot gas around and between galaxies and within galaxy clusters and groups. Time-resolved, high-resolution spectroscopy with IXO will probe the physics of neutron stars and black holes. IXO will measure the evolution of large-scale structure with a dynamic range and detail never before possible.

Exoplanet Mission

One of the fastest-growing fields in astrophysics is the study of planets beyond our solar system. NASA's current Kepler and this panel's recommended WFIRST mission will advance knowedge of the demographics of other planetary systems, but further steps will have to be taken to investigate the properties of individual planets around nearby stars. A micro-arcsecond astrometry mission such as the Space Interferometry Mission (SIM) Lite could detect nearby systems of planets and measure their masses. SIM Lite could even detect Earth-like planets, which are particularly difficult to find, that would be near enough to allow detailed study in more ambitious, future spectroscopic missions. Alternatively, rapid advances in starlight-suppression techniques could enable a moderate-size facility that could image and characterize giant planets (and perhaps some smaller ones) and investigate the debris and dust disks that are stages in the planet-forming process. Discovering even smaller planets and studying their atmospheres with transit photometry and spectroscopy employ another powerful, rapidly improving technique. The panel urges increased technology development for these techniques and recommends that one of these missions, or a yet-to-be-developed approach, be competitively selected around mid-decade and, if the budget permits, started before the end of the decade.

Moderate/Small Initiatives

Background-Limited Infrared-Submillimeter Spectrograph—U.S. Collaboration on the JAXA-ESA SPICA Mission

The tremendous success of the Spitzer Space Telescope has spurred the development of a yet-more-powerful far-IR mission, the Japanese-led Space Infrared Telescope for Cosmology and Astrophysics. The U.S. community should join this project by making the crucial contribution of a high-sensitivity spectrograph covering far-IR to submillimeter wavelengths, capitalizing on U.S. expertise and experience in detectors and instruments of this kind. Joining SPICA is time-critical and needs to be a priority. Such participation would provide cost-effective access to this advanced facility for the U.S. research community. Because JAXA and ESA are currently moving ahead with SPICA, the panel recommends that NASA commit to participation and begin to fund this activity now.

Augmenting the Explorer Program for Astrophysics

NASA's Explorer program is arguably the best value in the space astrophysics program. After years of reduced funding, increased support for astrophysics Explorers is essential to a balanced program of research and development (R&D). The panel recommends a substantial augmentation of funding dedicated to astrophysics Explorers with the goal of returning to a flight rate of one Explorer per year by the end of the decade.

Technology Development for a Hubble Successor

The imperative of understanding the history of the "missing baryons," as well as the evolution of stars and galaxies, requires ultraviolet (UV) spectroscopic observations that are more sensitive, and at shorter wavelengths, than are possible with the new Cosmic Origins Spectrograph (COS) on the HST. Key advances could be made with a telescope no larger than Hubble but equipped with high-efficiency UV and optical cameras having greater areal coverage than Hubble's. These would support a very broad range of studies. Achieving these same capabilities with a 4-m or larger aperture, in combination with an exoplanet mission capable of finding and characterizing Earth-like worlds, is a compelling vision that requires further technology development. The panel recommends a dedicated program of major investments in several essential technologies to prepare for what could be the top priority in astrophysics for the 2021-2030 decade.

Augmenting Research and Analysis Programs in Technology Development and Suborbital Science

The NASA Research and Analysis (R&A) programs support diverse activities that are crucial to the astrophysics program. This includes research grants in both observation and theory, as well as for laboratory astrophysics, technology development, and the Suborbital program. The panel, recognizing that these are core activities that underlie the NASA astrophysics program, recommends as urgent the augmentation of R&A funding that targets technology development and the Suborbital program. The panel calls for (1) a new initiative of focused technology development for projects that are likely to move ahead in the 2021-2030 decade, (2) a more aggressive program of technology development for missions in their conceptual phase, and (3) greater support for the most promising, possibly transformational, ideas that are not necessarily tied to a particular mission. The panel also recommends an augmentation of the Suborbital program, which also plays a critical role in developing and testing new technologies while providing a nearly space-like environment for low-cost science and—crucially—the training of new instrumentalists.

CONTEXT FOR ELECTROMAGNETIC OBSERVATIONS FROM SPACE IN THE 2010-2020 DECADE

An Impressive Suite of Current Missions

The remarkable advances in astronomy, astrophysics, cosmology, and fundamental physics during the past few decades have been achieved to a considerable extent through a broad array of space facilities that cover much of the electromagnetic spectrum. Scientists currently have access to 15 operating space missions (Figure 6.1). Most are NASA-led missions, but some—for example, Planck, Herschel, and Astro-H—are international collaborations in which NASA is a partner. The crucial roles that these facilities play in advancing the field are discussed in Part I of this volume, the reports of the Science Frontiers Panels of the Astro2010 survey. Although a review of that material is beyond the scope of this panel's report, their contributions, through the scientific priorities identified in the SFP reports, have guided the thinking and the recommendations of the EOS Panel.

This broad and powerful suite of capabilities will not continue: most of these missions will reach the ends of their planned lifetimes early in this decade. An important issue early in the decade will be the choice to extend some of these missions that will be made in the NASA Astrophysics Division's senior review process. The astrophysics community can and should play a role in weighing the scientific

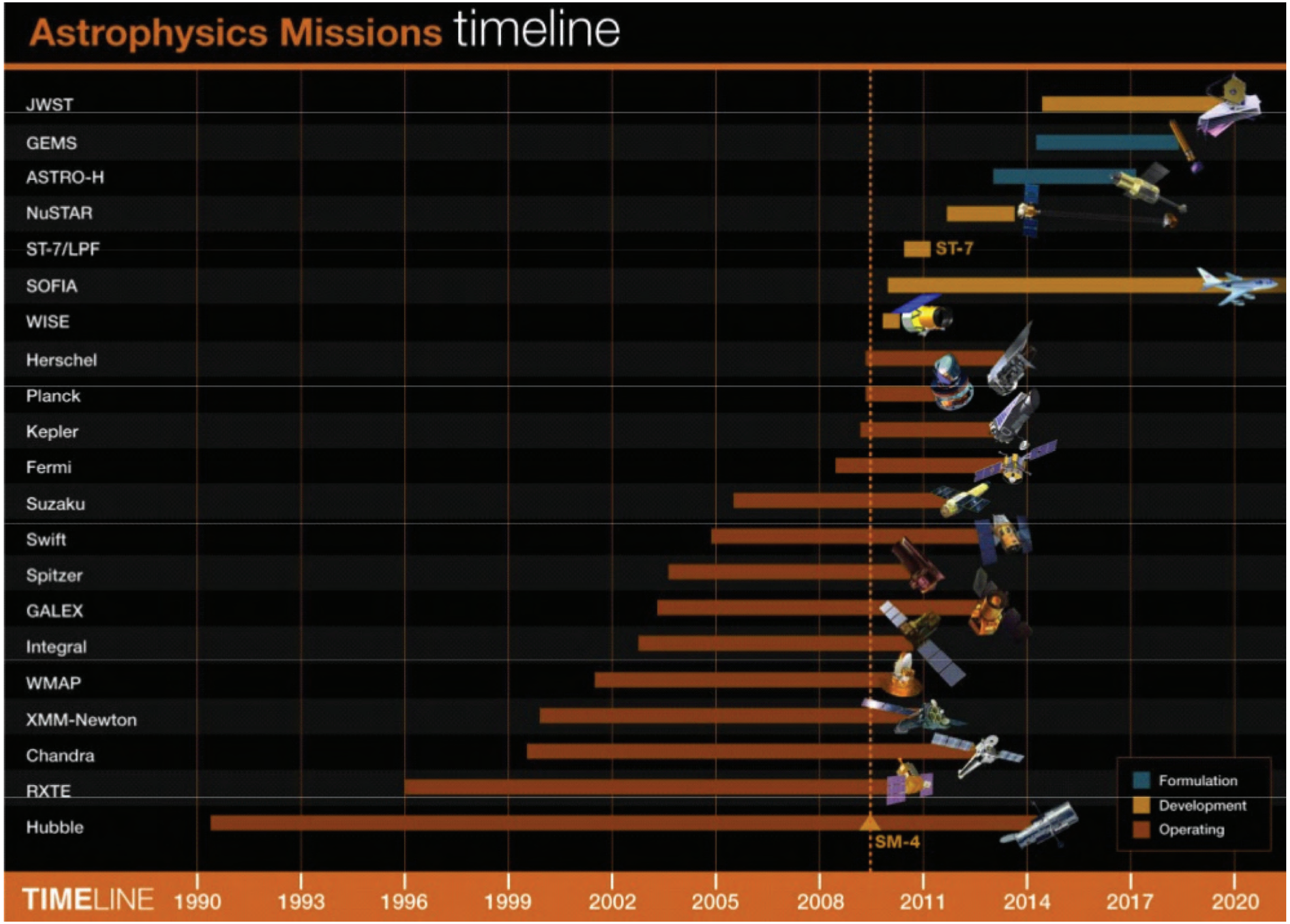


FIGURE 6.1 Astrophysics missions continuing into the 2010-2020 decade. The WISE mission, shown as under development, was launched in December 2009 and is now in operation. The termination dates shown are planned at this time and are not a recommendation of this panel. SOURCE: NASA.

productivity of each mission against the potential of the initiatives recommended in this panel's report.

New and Developing Missions for the Next Decade

The Next Generation of Broadly Capable Observatories

NASA's four Great Observatories—Compton, Hubble, Chandra, and Spitzer—are becoming legends in the history of science. Covering the majority of the electromagnetic spectrum—from gamma rays to the far-IR—these space telescopes have greatly increased our ability to learn what happens in the universe as well as how and why it happens. Not only by providing windows to light that does not reach the ground, but also by offering exquisite spatial resolution and a dark sky, these facilities have been our guides for astronomy's greatest adventures.

A *second generation* of highly capable, broad-purpose observatories is emerging. They are Fermi, the James Webb Space Telescope (JWST), a proposed new X-ray Observatory, and future large-aperture, UV-optical and far-IR telescopes. To realize all will take decades, but the first—the Fermi Gamma-ray Space Telescope (formerly GLAST)—is already up, surveying the full sky every 3 hours since August 2008. Fermi's Large Area Telescope (LAT) covers 20 MeV to ~300 GeV, and its gamma-ray-burst monitor is sensitive over the range 8 keV to 40 MeV. Fermi has already detected active galaxies, pulsars, supernova remnants, compact binaries, globular clusters, and gamma-ray bursts (GRBs). New classes of gamma-ray sources have also been discovered: starburst galaxies, high-mass X-ray binaries, and new varieties of gamma-ray-emitting pulsars. Fermi has also made important measurements of the galactic diffuse radiation and precise measurements of the high-energy spectrum of cosmic-ray electrons and positrons. Fermi is a highly successful example of interagency cooperation—NASA and DOE—that includes international partners as well, a matter of relevance to the panel's first-priority recommendation.

The next such observatory will be JWST, the top priority of the previous decadal survey, *Astronomy and Astrophysics in the New Millennium* (AANM).[2] Without a doubt, JWST will be *the* big space mission of the next decade, independent of the program advanced by the Astro2010 survey. JWST is a broadly capable observatory in the mold of Hubble, a collaborative project led by NASA, with major participation by the European Space Agency (ESA) and the Canadian Space Agency. JWST will cover a broad range of visible-to-mid-IR wavelengths, 0.6 μm to 28 μm, and will provide orders-of-magnitude greater sensitivity than previous IR space

---

[2] National Research Council, *Astronomy and Astrophysics in the New Millennium,* National Academy Press, Washington, D.C., 2001.

telescopes and with 10 times better angular resolution—comparable to Hubble's resolution in visible light. Its large 6.5-m segmented primary mirror will view deep space from the Earth-Sun L2 Lagrange point—a million miles from Earth—where the telescope will cool to a frigid 50 K, key for high sensitivity at long wavelengths.

JWST's prime mission is to open the last frontier in galaxy evolution—the earliest generations of stars and the birth of galaxies in the first billion years of cosmic history—and to follow the growth and maturation of galaxies through to the modern era. It will also be a powerful platform for studying how stars and their systems of planets are born in the Milky Way galaxy, taking crucial steps to probe the origin of life itself.

The U.S./German Stratospheric Observatory for Infrared Astronomy (SOFIA), a 2.4-m telescope peering into space from the side door of a Boeing 747 aircraft, began science operations in 2010. SOFIA's near- to far-IR *high-resolution* spectroscopic observations are likely to have the greatest scientific impact, but its imaging instruments have unique capabilities as well. SOFIA offers 20 years of operations above the absorbing effects of Earth's atmosphere, flexible operations from different locations over the globe to access both sky hemispheres, and the capability of updating and replacing instruments as technologies improve.

Not all of this panel's recommended missions fit into the "broadly capable observatory" category—for example, WFIRST (which will devote most of its lifetime on two dedicated programs and surveys with broad application) and an evolving exoplanet mission do not—but one that certainly does is IXO, a revolutionary X-ray telescope proposed as a collaboration of NASA, ESA, and JAXA. The panel also recommends U.S. participation in the Japanese-led SPICA mission, a large, cold telescope for sensitive far-IR observations of primeval galaxies and of disks where planets form. As a successor to the Spitzer Space Telescope, SPICA is another example of a space observatory that will have substantial impact on a wide range of astrophysics questions. The panel also recommends a dedicated technology-development program that could lead to a large UV-optical telescope—a successor to Hubble—in the 2021-2030 decade. The panel views the 2010-2020 decade as a time of great opportunity to capitalize on the Great Observatories program and leverage its success to this next generation.

Smaller Missions in Development for the Next Decade

Much astrophysical research requires ambitious, technologically sophisticated, and complex missions like Hubble and JWST. However, progress depends equally on the ingenuity of scientists and the success of parts of the program that emphasize specific objectives, specialized capabilities, and more flexible and rapidly evolving science programs.

The jewel of this approach is the Explorer program. Two decades of astrophysics Explorers have compiled a stunning record of achievement. In this tradition, the Wide-field Infrared Survey Explorer (WISE), launched successfully in December 2009, has already begun an all-sky survey from 3 to 25 μm that will be hundreds of times more sensitive than that of the Infrared Astronomical Satellite (IRAS) and nearly two orders of magnitude deeper than the JAXA mission Akari. The WISE survey will help search for the origins of planets, stars, and galaxies and create an infrared atlas whose legacy will endure for decades. The Nuclear Spectroscopic Telescope Array (NuSTAR), now in development, will be the first mission to focus high-energy X-rays, pioneering sensitive studies of the "hard" X-ray sky. Scheduled for launch in 2011, NuSTAR will search for black holes, map supernova explosions, and study the most extreme active galactic nuclei (AGN). NASA recently selected for a formulation-phase study the Gravity and Extreme Magnetism Small Explorer (GEMS), a mission that could measure polarization of cosmic X-ray sources and provide unique evidence of a black hole and data on its spin.

The United States is also a major participant in the Japanese Astro-H mission, a moderate-aperture X-ray telescope, the novel feature of which will be its relatively high sensitivity to moderately hard X-rays, E >10 keV. Among its important instruments will be an imaging array of microcalorimeter spectrometers, built by a U.S. team. Astro-H is expected to give a preview of the revolutionary capabilities of the more ambitious IXO mission.

Table 6.1 links this discussion of new and developing missions to recommendations of the two preceeding National Research Council decadal surveys: *Astronomy and Astrophysics in the New Millennium* (AANM; Taylor-McKee)[3] and *The Decade of Discovery in Astronomy and Astrophysics* (Bahcall).[4]

SCIENCE DRIVERS FOR KEY NEW FACILITIES

The Science Frontiers Panels (SFPs) of the Astro2010 survey have posed key questions that will require new observational capabilities. In Table 6.2 the panel lists these questions and correlates them with the program of activities recommended by the EOS Panel.

The Cosmology and Fundamental Physics (CFP) and the Planetary Systems and Star Formation (PSF) SFPs shine their spotlights on what are arguably the most exciting two areas in astrophysics: the study of the extraordinary *acceleration* of our expanding universe, and *exoplanets*—expanding the search for and study of the rapidly growing population of planets known to orbit stars other than the Sun.

---

[3] National Research Council, *Astronomy and Astrophysics in the New Millennium*, National Academy Press, Washington, D.C., 2001.

[4] National Research Council, *The Decade of Discovery in Astronomy and Astrophysics*, National Academy Press, Washington, D.C., 1991.

TABLE 6.1 Status of Previously Recommended Space Astrophysics Programs

Program	Current Status and EOS Panel's Recommendation
2001 AANM[a] Recommended	
Major	
Next Generation Space Telescope (NGST)	Renamed James Webb Space Telescope (JWST), in development for a scheduled launch in 2014.
Constellation-X Observatory	Reconfigured as International X-ray Observatory (IXO). EOS Panel recommends for this decade (pending ESA selection).
Terrestrial Planet Finder (TPF)	Several concepts studied and considerable technology development, but not ready for this decade. EOS Panel recommends further technology development.
Single Aperture Far-Infrared Telescope (SAFIR)	Not ready. EOS Panel recommends further technology development and contribution of far-infrared/submillimeter spectrometer to JAXA-led SPICA mission for this decade.
Moderate	
Gamma-ray Large Area Space Telescope (GLAST)	Developed jointly by NASA, DOE, and foreign partners. Renamed Fermi. Began operations in August 2008.
Laser Interferometer Space Antenna (LISA)	Not reviewed by EOS Panel. Reviewed by Astro2010 Panel on Particle Astrophysics and Gravitation
Energetic X-ray Imaging Survey Telescope (EXIST)	Expanded mission proposed to Astro2010 survey in "Major Category." EOS Panel judged the science of insufficient priority to justify high cost and schedule risk as determined by the Astro2010 independent cost appraisal and technical evaluation process. Not recommended by EOS Panel.
Small	
R&A program	R&A budgets cut over past decade. EOS Panel recommends budget augmentation for R&A programs, including technology development, theory, laboratory astrophysics, and the Suborbital program.
Ultralong-Duration Balloon program and Sounding Rocket to Orbit program	EOS Panel acknowledges great scientific potential of these longer-duration programs beyond the present Suborbital program and suggests possible payload support from Missions of Opportunity or Explorer lines.
1991 Bahcall Survey[b] Recommended	
Stratospheric Observatory for Infrared Astronomy (SOFIA)	Under development by NASA and German Aerospace Center, DLR Telescope installed and first door-fully-open flight successful. First science in 2010.
Space Interferometry Mission (SIM)	Re-endorsed by AANM. Reconfigured and proposed as SIM Lite Astrometric Observatory to Astro2010. EOS Panel recommends as a candidate for exoplanet mission with possible start late in the decade.

[a] National Research Council, *Astronomy and Astrophysics in the New Millennium*, National Academy Press, Washington, D.C., 2001.

[b] National Research Council, *The Decade of Discovery in Astronomy and Astrophysics*, National Academy Press, Washington, D.C., 1991.

TABLE 6.2 The Questions Posed by the Astro2010 Science Frontiers Panels, Correlated with the Activities Recommended by the EOS Panel

Science Frontiers Panel Question/Discovery Area		WFIRST	IXO	EXO-PLANET	BLISS SPICA	UV-Optical Telescope
PSF-1	How do stars form?					■
PSF-2	How do circumstellar disks evolve and form planetary systems?	▨		■	■	▨
PSF-3	How diverse are planetary systems?	■		■		🟥
PSF-4	Do habitable worlds exist around other stars, and can we identify the telltale signs of life on an exoplanet?	▨		■		🟥
PSF-D	Identification and characterization of nearby habitable exoplanets			■		🟥
SSE-1	How do rotation and magnetic fields affect stars?		■			
SSE-2	What are the progenitors of Type Ia supernovae and how do they explode?	▨	▨			
SSE-3	How do the lives of massive stars end?	▨	■			■
SSE-4	What controls the mass, radius, and spin of compact stellar remnants?	▨	■			
SSE-D	Time-domain surveys	■				
GAN-1	What are the flows of matter and energy in the circumgalactic medium?		■			■
GAN-2	What controls the mass-energy-chemical cycles within galaxies?		■		■	■
GAN-3	What is the fossil record of galaxy assembly from the first stars to the present?	▨	▨		▨	■
GAN-4	What are the connections between dark and luminous matter?	■	■			■
GAN-D1	Time-domain astronomy	■				▨
GAN-D2	Astrometry	▨		▨		■
GCT-1	How do cosmic structures form and evolve?	■	■		■	■
GCT-2	How do baryons cycle in and out of galaxies, and what do they do while they are there?	▨	■			■
GCT-3	How do black holes grow, radiate, and influence their surroundings?	▨	■	▨	■	▨
GCT-4	What were the first objects to light up the universe, and when did they do it?	▨	■		■	▨
GCT-D	The epoch of reionization					

continued

TABLE 6.2 Continued

Science Frontiers Panel Question/Discovery Area		WFIRST	IXO	EXO-PLANET	BLISS SPICA	UV-Optical Telescope
CFP-1	How did the universe begin?					
CFP-2	Why is the universe accelerating?	■	□			■
CFP-3	What is dark matter?	□		□		□
CFP-4	What are the properties of neutrinos?	■	□			
CFP-D	Gravitational wave astronomy					

NOTE: Darker color indicates a strong impact of the facility on answering the question. "Exoplanet" entries correlate the PSF questions with several proposed missions, as described in the panel report text. The maroon squares under "UV-optical telescope" refer to a possible planet-finding and characterization capability.

The EOS Panel assigned highest priority to a major new facility that will address both of these exciting fields along with many other SFP questions. The *Wide-Field Infrared Survey Telescope, WFIRST*, is a modest-aperture, near-IR telescope—the JDEM/Omega design—that will make precise measurements of the universe's evolving geometry and structure over cosmic time. Using three powerful methods, WFIRST will lead the effort to unravel a great astrophysics mystery of our time—why is the universe not just expanding, but expanding at an accelerating rate? Is this cosmic acceleration due to a strange, previously unknown "dark energy" that is defeating the pull of gravity on the vast scale of the universe, or is it possible that we have discovered that Einstein's formulation of the law of gravity is not quite right? WFIRST's cosmology experiments will help us find out.

With the same instrumentation, WFIRST will search for planets around distant stars—from giant planets down to those smaller than Earth—using the technique of microlensing, a consequence of Einstein's general theory of relativity. Combining the results of NASA's ongoing Kepler mission with the WFIRST planet search will produce the first comprehensive survey of other worlds and begin to answer the fundamental questions posed by the PSF Panel: How diverse are planetary systems and how often are they like our own solar system? How common are small, rocky worlds like Earth, and how often are they found in orbits that lead to temperate conditions favorable for life?

WFIRST will also carry out near-IR surveys of extensive areas of the sky, including the plane of the Milky Way galaxy, and support guest observer programs of *pointed* observations, for example studying the stellar populations of individual nearby galaxies. Such programs will be crucial for progress in the study of galaxy evolution and the development of large-scale structure in the universe, and for understanding how our galaxy and its neighbors were assembled and grew. WFIRST will do both focused and broad scientific programs at a cost substantially less than several focused programs would require.

Another critical subject that is almost untouched, for lack of resources and need of further technological advances, is the exploration of the vast store of "normal" atomic matter—to astrophysicists, baryonic matter. Galaxies, stars, planets, and people make up some fraction of this, but most baryonic matter is between the galaxies, its location and condition largely unknown. The connection between "missing baryons" and the known baryons threads through the reports of the Galaxies Across Cosmic Time (GCT) SFP and the Galactic Neighborhood (GAN) SFP and is a major component in the key science questions they pose.

Astronomy was born and has matured in the glow of visible light, but now it is known that most baryons exist as gas that emits essentially no visible light at all. In fact, from cosmological measurements and the study of the abundances of chemical elements, it is known that these "invisible" baryons outnumber the visible ones by 10 to 1. But little has been learned about them, progress so far only showing where, and at what temperature, the missing baryons are *not.* They are not cold gas, which would be "seen" in radio or millimeter waves. They must be hot—most, very hot. Why is it important to know about this plasma (ionized gas) sea made of electrons, protons, and helium nuclei, loitering around the outskirts of galaxies, and between them? Because this vast reservoir of baryonic matter is a key to learning how and why the universe formed stars and galaxies. We want to understand how what was once only hot gas grew into the complexity of matter seen today—up to and including life itself.

We are well along in, and well equipped for, studying the baryons in stars and galaxies: JWST, Spitzer, Hubble, Herschel, SOFIA, and WISE focus on the cooler, lumpier parts of the universe—such as planets and forming and dying stars—while Chandra, XMM, and Fermi observe hotter baryons in a menagerie of exotic forms—neutron stars, black holes, quasars, pulsars, supernovae, starbursts, novae, and so on. But new facilities are essential for understanding the role played by the invisible baryons and how they continue to interact and influence what happens in the universe. The panel agrees with the conclusion of the GCT and GAN SFPs that high-resolution UV and X-ray spectroscopy are the only efficient tools for observing these baryons and studying their dynamic flows into and out of galaxies, a process that profoundly affects galaxy evolution.

Both are essential: X-rays probe the hottest gas between galaxies and within galaxy clusters, while UV allows much higher spectral and spatial resolution for measuring the physical state of the gas, chemical element abundances, and microphysics. The hottest plasmas can be observed only in X-rays, whereas warm gas—a transition between the cold gas in galaxies and the hot plasma between them—requires the much higher resolution afforded by UV spectrographs. A complete picture of galaxy "ecosystems," for which there are many theoretical studies but few observational constraints, requires a combination of both techniques.

The panel thus recommends as its second priority *IXO, the International X-ray Observatory*, a mission that will make giant strides in studies of the hot baryonic component of the universe as well as providing unprecedented capabilities for studying high-energy phenomena such as black holes and supernovae. IXO will collect much more light than its predecessors and carry a revolutionary microcalorimeter—a unique and powerful instrument that produces X-ray images and high-resolution spectra simultaneously. This information is exactly what is needed to explore the interplay of the baryonic matter in stars with the hot baryonic matter beyond—for example, the supernova explosions and energy eruptions from accreting massive black holes (active galactic nuclei, AGN) that drive gas flows to the outskirts of galaxies. IXO is a broad-purpose astrophysical observatory that will make giant strides in this subject and in many others.

The Hubble, with its new COS, will make the most sensitive UV search to date for hot gas around and between the galaxies, but a new-technology UV-optical telescope will be needed to go beyond what HST can do. The panel recommends an aggressive program of technology development to develop such a facility. It is quite possible that this UV-optical capability can be combined with a major exoplanet mission: a telescope that can detect directly small planets like Earth orbiting nearby stars, study what they are like, and possibly even find evidence of life on them. The EOS Panel's recommendation includes technology development for this possibility—in particular, to show whether the two capabilities can be successfully combined.

The search for and study of exoplanets is one of the most exciting and fastest growing fields in astrophysics. Hundreds of planets have been detected by their gravitational "tugs" on parent stars, and progress has been made in the difficult job of describing what some of these planets are like, another priority for the PSF Panel. Kepler and WFIRST will compile basic data for thousands of exoplanets for a statistical study of the diversity of planetary systems, but most will be too distant for detailed study—the next step in this exciting field. The PSF Panel gives special priority as a "discovery area" to the next step of characterizing planets around nearby stars, with the goal of finding planets like Earth.

The panel strongly endorses this goal and recommends as its third priority an exoplanet mission to learn more about the planets of nearby stars. There are already several alternatives for this mission. SIM Lite, a micro-arcsecond astrometry mission that could detect the "wobble" of nearby stars induced by their planets, could discover nearby worlds as small as Earth and offer a unique program of general astrophysics based on precision distance/motion measurements. Earth-like worlds found with SIM Lite would be targets for more ambitious future missions to study their atmospheres and search for evidence of life. At a somewhat smaller scale, a moderate-sized telescope with excellent starlight suppression could directly detect and take spectra of giant planets, and perhaps even some supersized Earth-like planets. Probing the dusty disks around stars that are in the process of forming planets, and measuring the amount of dust left over in mature systems—data needed to build a mission capable of detecting a true "Earth-twin"—are principal tasks for such a facility. Another rapidly advancing approach is to search for planets that transit nearby stars and, once they are found, to obtain precision spectroscopy during transits in order to study a planet's atmosphere by the absorption of starlight. An exoplanet mission could not begin until late in the decade, so as its third priority, the panel recommends a competition between these and other possible approaches that are likely to develop over the next 5 to 10 years.

The Spitzer Space Telescope made remarkable progress in studies of galaxy evolution and star formation with its astonishing sensitivity in the far infrared. Spitzer has shown that dust-laden galaxies glowing brightly in the infrared, powered by starbursts and accreting black holes, become increasingly important elements in galaxy evolution as we probe further back in time toward the big bang. Spitzer also made significant breakthroughs in studying the cool, dusty environments where stars are born, as well as the disks around young stars, which are planet nurseries. To build on this critical work, the panel recommends—as its highest priority for a modest-scale investment—U.S. participation in SPICA, the JAXA-led successor to Spitzer, which will make far-IR imaging and spectroscopic observations with even greater sensitivity and spatial resolution. Exploiting U.S. leadership in detector technologies to build a background-limited, far-IR/submillimeter spectrometer for SPICA will secure a cost-effective U.S. share of this world-class facility for U.S. astronomers.

The study of the cosmic microwave background (CMB)—the light left from the big bang—has revolutionized the field of cosmology. The past decade's combination of ground-based, balloon-borne, and satellite facilities has pushed CMB temperature measurements to what had once been unthinkable precision. These data have provided the basis for highly accurate methods of measuring the size, age, mass-energy content, and expansion history of the universe and for detecting the earliest signs of structures that led to stars and galaxies. The state of the art has been achieved with the Wilkinson Microwave Anisotropy Probe, WMAP, a stunning example of how comparatively inexpensive Explorer missions can make giant steps in forefront fields of astrophysics. The Explorer program also builds critical relationships among university groups, national laboratories, and aerospace companies, where the next generation of instrumentalists, engineers, and managers are trained. To ensure that NASA's Explorer program continues its remarkable record of highly cost-effective, cutting-edge science, the panel proposes a substantial funding *augmentation* for astrophysics missions of the Explorer program.

NASA's collaboration with ESA on the newly launched Planck mission will allow U.S. astrophysicists to take the next step in CMB research. Planck will map with even greater precision than WMAP the CMB anisotropies and possibly CMB polarization, a powerful diagnostic of early-universe physics. While the U.S. community prepares to analyze Planck data, it continues to develop better technology for an even more powerful mission aimed at polarization measurements. Balloon-borne payloads—part of the NASA Suborbital program—are a key component in the invention of new detectors and methods, allowing low-cost development of components for future satellite missions. The EOS Panel recommends an *augmentation* of the Suborbital program, together with technology development to support CMB and other research programs that depend on invention, innovation, and experimentation. These activities deliver cutting-edge science while providing irreplaceable hands-on experience with technologies that are often destined for large space missions, training the next generation of instrument builders in the process. They are essential to the health of the NASA astrophysics program.

WIDE-FIELD INFRARED SURVEY TELESCOPE—WFIRST

Remarkable opportunities for new space initiatives are at hand. However, resources in the decade 2010-2020 could be substantially less than in the previous two decades. Although excellent science is always the first priority, the panel believes that new space initiatives encourage activities that will have a major impact on a wide range of high-priority research questions, engaging a broad segment of the community. It was accordingly with considerable excitement that the panel recognized in the input received from the U.S. astronomical community several strong scientific programs describing essentially the same facility—a 1.5-m space telescope with a focal plane covered with near-IR array detectors. Because of the absorption and emission of Earth's atmosphere in the near-infrared, a modest-size space telescope can have great impact on many subjects and complement and leverage observations from ground-based facilities. Furthermore, the ~3 times better spatial resolution, compared to the best ground-based (non-adaptive optics-corrected) seeing, offers tremendous gains in sensitivity that are clearly "enabling" for many important programs.

Accordingly, the EOS Panel recommends as its highest priority a facility it calls WFIRST—the Wide-Field Infrared Survey Telescope. The panel chose one particularly well-studied concept—Joint Dark Energy Mission/Omega (JDEM/Omega)—as a "hardware template" for the WFIRST mission. WFIRST will first and foremost tackle two of the biggest questions in astrophysics: Why is the universe accelerating? (SFP science question CFP 2) and, Do habitable worlds exist around other stars, and can we identify the telltale signs of life on an exoplanet? (SFP science question PSF 4). The first 5 years of a 10-year mission would be dedicated largely to answering these questions. The cosmic acceleration research program will make unique measurements of the dark-energy equation-of-state parameter "w" and its time evolution, w', while also addressing whether cosmic acceleration might be due instead to an imperfect understanding of gravity. In addition, an unprecedentedly powerful extrasolar planet search will be carried out by monitoring a large sample of stars in the central bulge of the Milky Way galaxy for microlensing events, producing a demographic survey that will sensitively sample a wide range of planet masses—down to planets smaller than Earth—and of distances from the parent stars.

A significant fraction of the first 5 years will also be used for surveys and smaller peer-reviewed guest-observer projects that will investigate, for example, galaxy evolution, stellar populations of nearby galaxies, and the plane of the Milky Way galaxy. The combination of depth, area, and quality of WFIRST data in the infrared will easily surpass that of any other ground-based or space-based facility. WFIRST research bears substantially on 10 of the 20 key questions posed in the Astro2010 Science Frontiers Panel reports (see Table 6.2). The WFIRST mission will contribute to answering some of these questions through extensive surveys. A deep infrared galaxy survey will cover the key epoch of galaxy assembly, $1 < z < 2$, with multiband imaging and low-resolution (grism) spectroscopy (redshifts and line strengths) via Hα (hydrogen emission line) observations. A complete survey of the galactic plane in the near-infrared will also be undertaken. As the mission unfolds, especially as the cosmic acceleration and planet-finding programs complete their first campaigns, these surveys—made immediately available to the entire community—and pointed observations through guest observer (GO) programs, will assume more importance. Peer review will determine the balance between these elements and the continuation of the cosmology and planet-finding programs. By combining this double-core focus with a broad vision that addresses the diverse and changing research priorities of astrophysics, WFIRST will serve as a dedicated facility and a broader-use observatory.

The ability of a single facility to have such a broad impact—together with its combination of affordability, technical readiness, and low risk—is why the EOS Panel recommends WFIRST as the next large U.S. space mission. The scientific questions WFIRST addresses will certainly evolve during its development. But WFIRST's versatility—providing wide and deep imaging and wide-field slitless spectroscopy, with diffraction-limited imaging at 2 μm and a high level of system stability—guarantees that this will be a very productive and important facility for astrophysics research.

Probing Cosmic Acceleration

The discovery of cosmic acceleration is among the most exciting science results of our time. After the discovery of the expansion of the universe, astronomers worked for nearly a century to measure its deceleration, to find out how big and how old the universe is and how much matter it contains. When tools and methods finally became good enough to do that job properly, astronomers discovered—to the astonishment of most—that the expansion of the universe is *accelerating*. Is this due to the repulsive force of a previously unknown, pervasive mass-energy component—named dark energy but not currently understood—or does it signal a breakdown on the largest scales of Einstein's description of gravity in his general theory of relativity? Or is this perhaps a clue to something even more exotic and about which even less is understood? The evidence for acceleration is itself compelling, from measurements of galaxy distances used to track the expansion and from studies of ripples in the cosmic microwave background that record the mass-energy density early in the universe's history. Whether it is dark energy, or a revision in the law of gravity, or something else altogether, these observations are telling us something new about fundamental physics.

The WFIRST program to study cosmic acceleration described here is based on the JDEM (IDECS and Omega) program that was presented to the panel, but it also benefits from the work of a group study within the Astro2010 survey that considered a combined space and ground approach. Specifically, the WFIRST program provides unique contributions to the three methods most strongly endorsed by the Astronomy and Astrophysics Advisory Committee's (AAAC's) Dark Energy Task Force: (1) weak lensing; (2) baryon accoustic oscillations; and (3) infrared photometry of $0.2 < z < 0.8$ supernovae.

The apparent cosmic acceleration can be determined by measuring either the apparent brightness of "standard candles" or the apparent size of "standard rulers" over the history of the universe. In the general theory of relativity, the detailed expansion history is governed by the equation of state, which describes the relation between pressure and energy density in the universe. The best measurements to date indicate the presence of something like a vacuum "dark energy" that is supposedly driving the acceleration of the expansion of the universe. The most accurate standard candles are supernova explosions (SN Ia), which can readily be measured back to a time when the universe was half its present size. The most promising standard ruler is a "bump" in the galaxy power spectrum—a preferred scale. This is thought to arise from so-called baryon acoustic oscillations (BAOs)—peaks and valleys that were imprinted on the dark-matter distribution by sound waves carried in the primordial plasma of the early universe. A third way of measuring the equation of state hinges on measuring the growth of structures much smaller than the universe itself, for example, clusters of galaxies: in the presence of dark energy their growth is terminated sooner than it would be otherwise. By virtue of their self-gravity, the masses of these structures produces small distortions—weak gravitational lensing—of the shapes of galaxies that lie behind them.

While some of these observations are being attempted with ground-based facilities, observing in the near-infrared from space offers powerful advantages, especially in the $1 < z < 2$ redshift range where these cosmological measurements are most effective. This includes better angular resolution for defining galaxy shapes (weak lensing) and the accessibility of the Hα emission line of hydrogen gas for redshift measurements (BAO) over the maximum volume that can be targeted.

Why should WFIRST employ all three methods? Supernovae (in particular, type SNe Ia) give the best measurements of cosmic acceleration parameters at low redshift due to their greater precision per sample or per object. BAO excels over large volumes at higher redshift. Together SNe Ia and BAO provide the most precise measurements of the expansion history for $0 < z < 2$ and place significant constraints on the equation of state. Weak lensing provides a complementary measurement through the growth of structure. Comparing weak-lensing results with those from supernovae and BAO could indicate that "cosmic acceleration" is actually a manifestation of a scale-dependent failure of general relativity. Combining all three tests provides the greatest leverage on cosmic-acceleration questions. WFIRST can do all three. The panel thinks it would be far less sensible to build a smaller, somewhat-less-costly facility to carry out one or at most two of these tests.

Can WFIRST "solve" the problem of cosmic acceleration? Only to the extent that it can rule out alternative explanations. The leading contender among these is Einstein's cosmological constant, with equation-of-state parameter $w = -1$, giving a negative pressure equal to the energy density. A measurement of a value of w significantly different from -1 would imply the existence of a "dark energy" that is not "merely" a cosmological constant. At present, w is consistent with -1 to an accuracy of about 10 percent, and its rate of change, w', is known to ~100 percent. WFIRST could lower the uncertainty in w to ~1 percent and in w' to ~10 percent, subjecting the cosmological constant hypothesis to a test 100 times more stringent (using the figure of merit proposed by the Dark Energy Task Force) than current observations.

The Architectures of Extrasolar Planetary Systems

Searching for planets orbiting other stars—exoplanets—is a relatively new field, of burgeoning interest to the astrophysics community and the public alike. Most of the 400+ known planets have been found, beginning in the mid-1990s, by observing the "to-and-fro" velocity of neigboring stars as they are tugged by the gravity of orbiting planets. This method is most sensitive to giant planets and especially those that are close to their parent stars. Detection of a solar system like our own remains beyond our reach: the detection of far-out giant planets requires measurements over decades, and detection of a true Earth-twin requires technology beyond the current state of the art. Over this decade, progress in a variety of approaches will begin to change this situation, but for now very little is known about how common are "solar systems" like the one we live in.

NASA's recently launched Kepler space telescope is up and running and looking for distant planets crossing in front of their parent stars, minutely diminishing their brightness. Kepler's survey should answer the crucial question of how common Earth-like worlds are in the "habitable" zones of the stars it is monitoring—mostly galactic inner-disk and bulge stars. By the nature of the experiment, the planets that will be detected by Kepler are thousands of light-years from Earth—too distant for direct study, although the masses of some will be determined by ground-based radial velocity measurements.

Microlensing is another consequence of Einstein's general theory of relativity: a mass, for example a star, aligning almost exactly along the line of sight to another, more distant star, makes what amounts to a gravitational telescope—the light from the background star is intensified. If there are planets orbiting the star that is "lensing" the background star, they too will amplify the light—briefly—making their presence known. These microlensing events typically last a few days, and the detailed rise and fall of the light yields the planet-to-host-star mass ratio and separation. For most of the observed events, subsequent observations of the host star can determine its mass, which then gives the planet masses directly—accurate to about 20 percent. A spectacular example of the power of the technique is one such event observed with a ground-based telescope, shown in Figure 6.2.

Because planets even smaller than Earth can be detected in this way, at a wide range of distances from the parent star, space-based microlensing is an extremely effective way to conduct a survey of what other solar systems are like: How many planets are there? What are their masses? At what distances do they orbit their stars? In Figure 6.3 the panel shows the exoplanet discovery space of a microlensing planet search compared to those of Kepler and several ground-based techniques. The complementarity of WFIRST (illustrated by the MPF curve in Figure 6.3) to Kepler—the greater sensitivity to planets farther from the star, and its deeper reach compared to ground-based microlensing—is readily apparent.

An exoplanet microlensing program requires continuous monitoring of a few fields containing tens of millions of stars in the galactic bulge for long contiguous periods. In the optimistic scenario that every star has an Earth-like planet (fraction of stars hosting an Earth $\eta_\oplus = 1$), a 500-day microlensing campaign (spread over the first 5 years of the mission) would find ~200 Earth-mass planets (and many thousands of larger planets). In fact, this is the only technique yet proposed that could find planets with masses smaller than Earth's and is the only likely way to obtain a good statistical sample of planets with masses less than 5 Earth masses. On the other hand, finding no Earth-like planets at all in such a campaign would indicate a well-measured upper limit to the fraction of stars with an Earth-like world, $\eta_\oplus < 0.05$, and would be a stunning result. Figure 6.4 shows the expected number of planets of different masses found over 250 days of monitoring for microlensing events with WFIRST.

Because microlensing observations using WFIRST can detect planets covering a wide range in planet masses and orbits, this will be the most complete demographic survey of other planetary systems—an early but giant step in a fascinating program to study and compare other planetary systems to our own. It is the next stage of a decades-long journey to understand whether our inhabited planet is one of many in the cosmos.

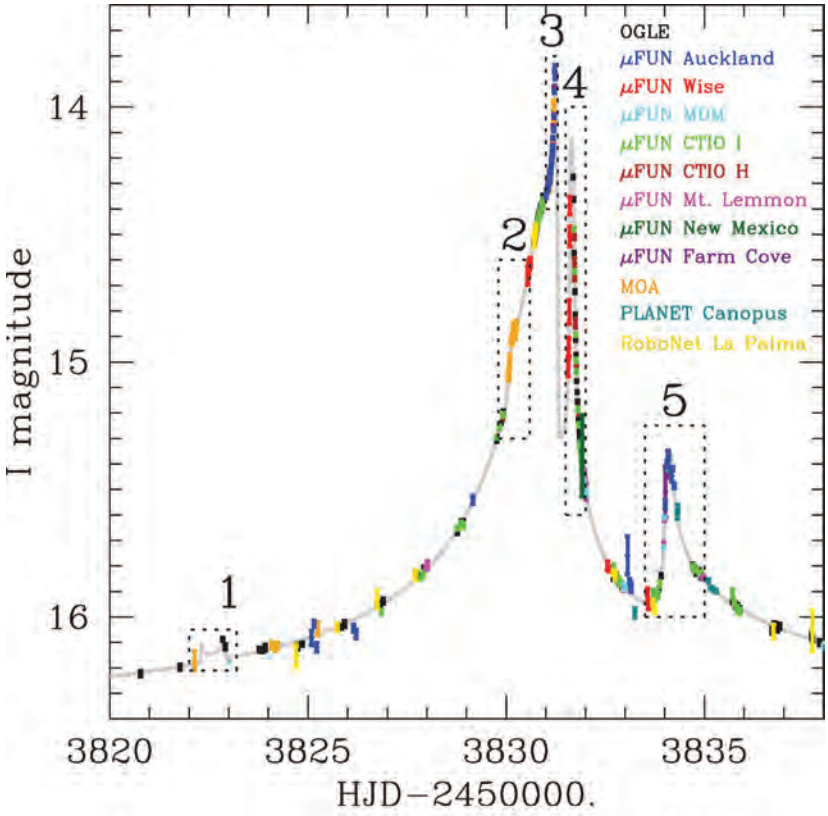


FIGURE 6.2 Microlensing event showing how a wealth of information on orbiting planets is contained in a detailed light curve. The different color data represent photometric observations over the course of the event from numerous ground-based observatories. The sharp features in the light curve, marked 1 through 5, result from two orbiting planets, analogous to Saturn and Jupiter in our solar system, each with about half their mass. SOURCE: B.S. Gaudi, D.P. Bennett, A. Udalski, A. Gould, G.W. Christie, D. Maoz, S. Dong, J. McCormick, M.K. Szymanski, P.J. Tristram, S. Nikolaev, B. Paczynski, et al., from the PLANET and RoboNet Collaborations, B. Chaboyer, A. Crocker, S. Frank, and B. Macintosh, Discovery of a Jupiter/Saturn analog with gravitational microlensing, *Science* 319(5865):927-930, 2008. Reprinted with permission of AAAS.

The High Value of Near-IR Imaging from Space

The sensitivity of WFIRST for near-IR imaging will be unrivaled. The requirements for a deep, multiband, infrared-imaging survey are essentially the same as those for the weak-lensing program, and so a by-product will be a spectacular, unprecedented, distant-galaxy survey with 0.2-μJy sensitivity and 0.2-arcsecond resolution over a large fraction of the sky—a boon for studies of galaxy evolution, large-scale structure, searches for high-redshift quasars, and galactic white and brown dwarfs, just to name a few examples. WFIRST will sample as deeply as the

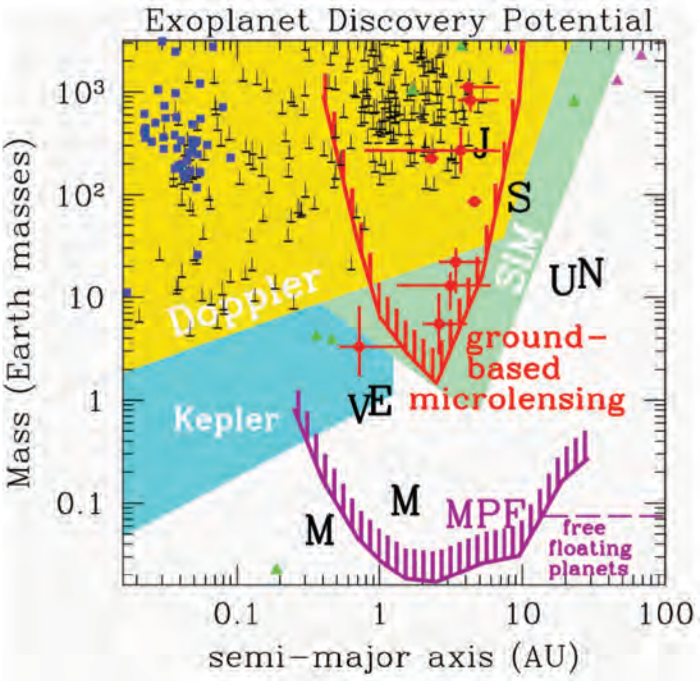


FIGURE 6.3 Comparison of the ability of Kepler (cyan band) and of a space-based microlensing mission, Microlensing Planet Finder (MPF; purple hatched) to detect planets of different masses and distances from their stars. The WFIRST microlensing survey (fairly represented here by "MPF") is well suited for finding Earth-like worlds in the habitable "sweet spot" around K and G stars, and is more sensitive than ground-based microlensing surveys (red-hatched) by factors of 10 to 100. The black capital letters indicate where the planets in our solar system fall in this diagram. "Doppler" indicates the ranges now detectable by the velocity shift of the parent star, and "SIM" (light green) indicates the ranges accessible with the SIM Lite mission for nearby stars. Ground-based microlensing discoveries are in red; Doppler detections are inverted "T's"; transit detections are blue squares; and timing and imaging detections are green and magenta triangles, respectively. SOURCE: D. Bennett, J. Anderson, J.-P. Beaulieu, I. Bond, E. Cheng, K. Cook, S. Friedman, B.S. Gaudi, A. Gould, J. Jenkins, R. Kimble, D. Lin, et al., "Completing the Census of Exoplanets with the Microlensing Planet Finder (MPF)," Astro2010 white paper, available by request from the National Academies Public Access Records Office at http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=48964.

deepest IR surveys yet done—for example, the UK Infrared Deep Sky Survey—over an area 10,000 times larger. Compared to other surveys that have sampled wide areas, WFIRST will reach several thousand times deeper than 2MASS at 2.2 μm and a thousand times deeper than the WISE mid-IR surveyor. The community's imperative for such a facility can be gauged by the painfully inefficient way existing space telescopes have been used to map deeply areas of less than a square degree—similar studies would require only four exposures for WFIRST.

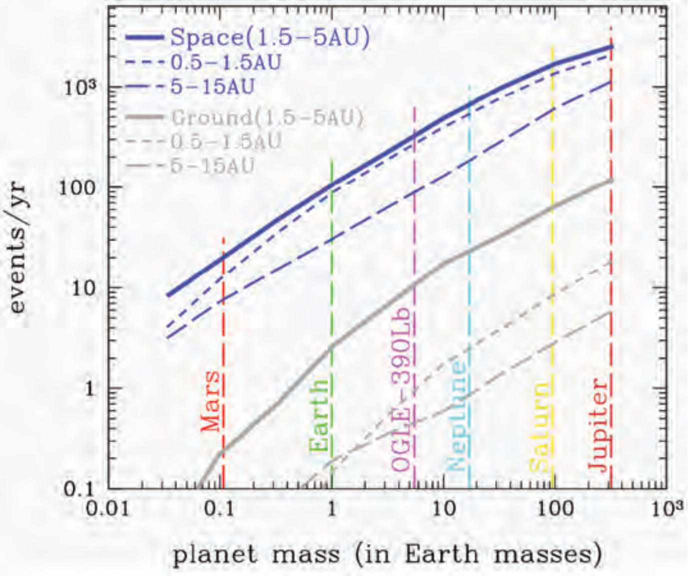


FIGURE 6.4 The expected yield of an exoplanet survey for a 250-day campaign, in numbers of detected planets. The wide range in planet masses and orbit sizes is shown, as well as the factor of 10 to 100 gain in sensitivity afforded by a space telescope. If Earth-like planets are common, a good statistical sample is achieved in this relatively short time, but if the discovery rate is low, the time spent on this program can be increased accordingly. In a given time WFIRST's yield would be comparable to or greater than that of MPF. SOURCE: D. Bennett, J. Anderson, J.-P. Beaulieu, I. Bond, E. Cheng, K. Cook, S. Friedman, B.S. Gaudi, A. Gould, J. Jenkins, R. Kimble, D. Lin, et al., "Completing the Census of Exoplanets with the Microlensing Planet Finder (MPF)," Astro2010 white paper, available by request from the National Academies Public Access Records Office at http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=48964.

By providing Hα detections for most galaxies, and photometric redshifts for all, WFIRST will produce a three-dimensional map of the evolving universe over the redshift interval of $1 < z < 3$. Such data will also constrain the history of star formation over this epoch. This will link the evolution of galaxies to the growth of large-scale structure during the critical epoch when galaxies grew from adolescence to adulthood. Its sub-micro-Jansky sensitivity would be a perfect match for optical surveys by the Large Synoptic Survey Telescope (LSST), greatly enhancing their utility, and WFIRST could perform time-resolved surveys over more limited areas of sky, complementing LSST here, too. The ability to resolve stellar populations of nearby galaxies—measuring the detailed shape of the stellar giant-branch—will provide unique information on the histories of star formation for a wide variety of galaxy types. A deep survey of the galactic halo will be the most complete study of the history of satellite-galaxy accretion by the Milky Way, and a similar study can be done for the neighboring Andromeda galaxy. A survey of the galactic *plane*, with its treasure trove of sites of vigorous star formation pumping energy and heavy chemical elements into the intergalactic gas, will be our most detailed look at galaxy evolution in action.

In addition to synergy with LSST, the surveys done by WFIRST—large and small—will provide high-quality source material for studies of smaller areas of the sky, in concert with ground-based telescopes such as the Thirty Meter Telescope, the Giant Magellan Telescope, and the Atacama Large Millimeter Array, and premier space facilities such as JWST and SPICA. Working with these facilities, WFIRST will play a key role in answering the wide range of questions on galaxy and cosmic evolution identified by the Astro2010 Science Frontiers Panels.

WFIRST Programmatics

The broad capabilities of WFIRST—carrying out two, very different, dedicated programs and a wide variety of surveys and smaller programs—mean that this facility will be of critical interest to astronomers in a variety of research areas. To contain the cost and risk of this facility, however, the panel recommends that the architecture of JDEM/Omega be adopted and modified only as is necessary to opimize the two core programs: cosmic acceleration and the microlensing search for planets. Mindful of the priority of these two programs, planning for the operation of WFIRST should incorporate broader interests, including those of galactic and extragalactic surveys, stellar populations, and diverse guest observer programs; the panel imagines a newly appointed science working group to address these issues. In summary, the design of the telescope, spacecraft, and focal-plane instrumentation should be left to the project team and focused around the programs of cosmic acceleration and microlensing planet finding, with the science working group helping to construct an operating plan for the facility that can accomplish the combination of the two dedicated programs together with surveys and pointed observations.

The EOS Panel chose to incorporate the study of cosmic acceleration within a broadly capable space mission because it believes that this approach is scientifically compelling and the best use of limited resources by and for our discipline. It hopes that those who have developed concepts for “dark energy” missions with NASA and DOE, including those within these agencies, will find the far greater science productivity of WFIRST a compelling reason to support a broader mission. The panel also recognizes that international collaboration offers the possibility of further benefits. However, it strongly believes that the design, broad purpose, and multiple allocation of resources envisioned in WFIRST should be retained through any process of international partnering.

As a straw-man example for the first 5 years of a 10-year mission, the panel imagines 2+ years dedicated to the cosmic-acceleration program. These observations will provide more than 8,000 deg^2 for the BAO survey (grism) and 4,000 deg^2 for the weak lensing (single-band imaging) survey (about half of the JDEM/Omega program) and will produce a large multiband galaxy survey for public archives.[5] Dedicated microlensing campaigns of 100 days duration in each of the 5 years could accumulate a significant sample, even within the first few years of the mission. A galactic-plane survey of one-half year, together with about 1 year allocated by open competition, would fill the initial 5-year timeline. Barring any operational problems, WFIRST should continue for another 5 years: peer review would compete augmentations of the cosmic acceleration or planet-survey programs with new or larger surveys and smaller guest observer programs. The initial allocations for cosmic acceleration and microlensing in the panel's straw-man plan should produce dramatic advances and allow for informed judgment as to the benefit of additional observations. It has not been the panel's intention here to be overly prescriptive but simply to show how WFIRST might accomplish all that is envisioned.

WFIRST Technical Issues

The concept for WFIRST adopts the hardware configuration of JDEM/Omega, a 1.5-m-aperture, three-mirror focal anastigmat with 2×10^8 pixels from ~50 HgCdTe detectors (JWST and HST-WFC3 heritage) for both imaging (0.33 deg^2) and low-resolution slitless grism spectroscopy (0.55 deg^2), with high sensitivity from 0.8 to 2 μm. The panel recommends that the focal plane be optimized for the cosmic acceleration and microlensing programs, while avoiding a substantial increase in cost. For example, a change in the number of field segments or the number dedicated to slitless spectroscopy may enhance a broad science program without significantly increasing the cost or complexity of the facility. Mechanisms for exchanging filters and grisms should be the only moving parts in the instrument. To take full advantage of the benefits of observing from space, imaging pixels should be no larger than 0.18 arcsecond, as planned for JDEM/Omega. This will critically sample the diffraction-limited point-spread function at $\lambda = 2.1$ μm wavelength; next-generation HgCdTe detectors, if available, will have smaller 15-μm pixels that would lower the wavelength of diffraction-limited sampling to $\lambda = 1.7$ μm. With WFIRST deployed to L2 and no consumables beyond propulsion, its planned lifetime could be at least 10 years.

The JDEM/Omega configuration is somewhat different from that of the Micro-

---

[5] The weak-lensing/galaxy survey could be interleaved with about a half-year's worth of repeated observations of polar fields to monitor high-redshift supernovae.

lensing Planet Finder (MPF; Bennett[6]) mission whose capabilities are represented in Figures 6.3 and 6.4. Although the wavelength range and detectors are the same, WFIRST's larger aperture would provide higher resolution with better sampling of the point-spread function by smaller pixels. The factor-of-two smaller field coverage of WFIRST compared to MPF is compensated by the larger aperture, such that equal areas can be sampled in equal time. A significant difference is WFIRST's location at L2 compared to the proposed geosynchronous orbit for MPF, which requires that the microlensing planet survey be carried out in campaigns of approximately 100 days per year instead of the continuous 9-month monitoring allowed by MPF's orbit.

The more-general, near-IR imaging surveys and targeted observations envisioned for WFIRST would resemble those described in the Near-Infrared Sky Surveyor (NIRSS) document (Stern[7]). Like JDEM/Omega, NIRSS is also proposed as a 1.5-m-aperture telescope, with a wide-area focal plane populated with 36 HgCdTe array detectors. NIRSS is configured for a full-sky survey sampling 33 percent more coarsely than WFIRST's (JDEM/Omega) imaging-camera layout: four filters ranging from 1.2 μm to 3.4 μm are to be exposed simultaneously with nine detectors dedicated to each band. In comparison, WFIRST would sample about twice as much area per exposure in a single color.[8] The bottom line is that a three-band-survey (~J, ~H, ~K) would require 3 to 4 years of dedicated NIRSS use to cover the whole sky, while a substantial fraction of the sky is readily incorporated into WFIRST's diverse program.

It is very clear that WFIRST would be a fully subscribed and extremely productive facility. For this reason, and to avoid cost growth, the panel strongly discourages additional instrumentation, such as optical imaging with charge-coupled devices (CCDs) (for which the advantage of space is far smaller than for the near-IR), coronographs, integral field units, and so on. Observing modes other than the "staring" imaging mode described here would increase complexity, cost, and management problems, disproportionate with any benefits. The panel notes that JDEM/IDECS—basically WFIRST (or JDEM/Omega) but including a CCD array camera—*was not ranked highly* by the panel for such reasons.

---

[6] D. Bennett, J. Anderson, J.-P. Beaulieu, I. Bond, E. Cheng, K. Cook, S. Friedman, B.S. Gaudi, A. Gould, J. Jenkins, R. Kimble, D. Lin, et al., "Completing the Census of Exoplanets with the Microlensing Planet Finder (MPF)," Astro2010 white paper, available by request from the National Academies Public Access Records Office at http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=48964.

[7] D. Stern et al., "The Near-Infrared Sky Surveyor (NIRSS)," Astro2010 white paper, available by request from the National Academies Public Access Records Office at http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=48964.

[8] There is a possibility of using the WFIRST spectroscopy channels to add a factor-of-two greater area, albeit at a lower sampling scale of 0.35″ per pixel.

WFIRST Cost, Risk, and Trade-offs

The JDEM/Omega spacecraft was analyzed as part of the independent cost assessment carried out by the Astro2010 decadal survey. As a design template for WFIRST, JDEM/Omega rated well in terms of maturity and schedule risk. Technical risk was rated as "medium low" and cost risk as "medium." The *life-cycle* cost (including launch, with all costs in FY2009 dollars) according to the project was set at \$1.1 billion; the independent assessment came in at \$1.6 billion. Based on this input, the panel is convinced that a similar facility with a broad scientific program could be launched in the next decade with a total lifetime cost of ~\$1.5 billion, and it urges that this targeted cost be retained.

A concern was raised in the independent assessment about the acquisition of such a large number of HgCdTe detectors. The panel notes, however, that with a Phase B start following the launch of JWST, there should be many years to accumulate these detectors. A more serious concern, also expressed in the assessment, is achieving and maintaining the image quality required for weak lensing over the full focal plane. The JDEM/Omega team recently demonstrated significant progress verifying the performance of the JWST HgCdTe detectors for the specific application of weak lensing, but knowledge of the point-spread function and its time variation remains a mission-driving requirement. Because weak lensing provides the only component of the WFIRST dark-energy program that is sensitive to general relativity and the growth of structure, the panel believes it is vital to commit sufficient resources to address this risk in the WFIRST design as soon as possible. It will then be important to have another independent assessment of the result of this effort and for the WFIRST project, once constituted, to evaluate its merit-to-cost quotient in the context of ground-based measurements of weak lensing (e.g., LSST) and the benefits to the full scientific program. The anticipated schedule of a start in mid-decade should at least prevent this from becoming a risk for significant cost growth, assuming that adequate funds are provided to properly address this risk.

A final concern expressed in the independent cost-risk assessment is the required telemetry rate, which is challenging for a mission at L2 for both of the dedicated programs. Many future missions are likely to demand increased bandwidth, but in the case of WFIRST, this need may become urgent, especially considering the requirements of JWST and other future missions. Onboard data storage and processing, and the development of task-specific data-compression algorithms, may be required. This issue also needs early attention.

IXO AND THE COMPELLING CASE FOR X-RAY ASTRONOMY

Less than 50 years ago the Sun was the only known source of cosmic X-rays. In 1962 a group led by Riccardo Giacconi sent an X-ray detector on a rocket high above Earth's atmosphere. The first detected source turned out to be a neutron star—a structure as dense as the nucleus of an atom but as massive as the Sun, the remnant of a supernova explosion.

From its birth, X-ray astronomy has offered a startlingly different view of the universe. X-rays come from extreme objects because extreme conditions are needed to produce an X-ray photon, typically 1,000 times more energetic than a photon of visible light. X-ray photons get their high energies by emerging from gas with temperatures exceeding a million degrees, or from particles moving at nearly the speed of light. An abundant source of X-rays from space, therefore, is always something interesting and unusual. Such sources include supermassive black holes at the centers of distant galaxies, hot bubbles of newly minted elements from supernova explosions, swirling gas around the neutron stars and black holes left by old supernovae, and immense clusters of galaxies. Many such X-ray sources are "invisible"—they cannot even be seen with an ordinary optical telescope—so X-ray astronomy provides unique information that informs a wide range of cosmic questions. However, because of their high energies, X-rays are completely blocked by Earth's atmosphere—fortunately for life on Earth—and observing them requires a space telescope.

It has been only a few decades since it was learned that the stars, the focus of astronomical research for centuries, account for less than 1 percent of all matter. Ten times as much normal (atomic) matter is in the form of gas, most drifting between the galaxies, never to become stars; X-ray observations are essential for detecting and studying this majority component. Similarly, gas on the brink of tumbling into a black hole emits more X-ray light than visible light, making X-ray observations key for probing the fundamental physics of gravitation.

The Next Step in X-Ray Telescopes

Over the last few decades, NASA has launched many powerful X-ray telescopes, culminating in the Chandra X-ray Observatory, one of the four highly successful Great Observatories. Figure 6.5 shows a Chandra picture of the million-degree gas in the center of a cluster of galaxies and compares it to a Hubble picture of the starlight from the galaxies themselves: X-ray telescopes produce entirely different views of the same part of the universe. But pictures of the sky, whether in X-rays or visible light, record only the most basic information—how the universe "looks." Astronomers have long recognized that understanding the how and why—the *astrophysics*—comes mainly from *spectroscopy*, separating the light into a *spectrum* of different energies, which our eyes sense as color in visible light. Spectra tell us how the world works.

X-ray cameras have always had a ready ability to do this to some extent: CCD detectors (which produced most of the X-ray pictures on display in this panel re-

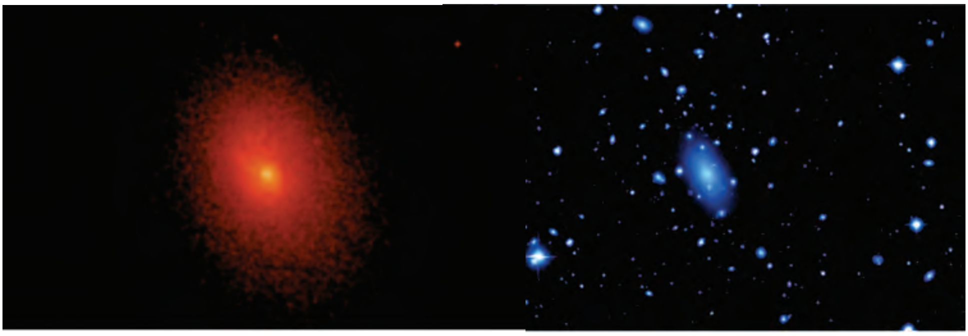


FIGURE 6.5 The cluster of galaxies Abell 2029. Most of the normal matter in the clusters is in the hot X-ray-emitting gas between the galaxies (*left*), not in the galaxies of billions of stars (*right*). SOURCE: *Left:* NASA/CXC/IoA/S. Allen et al. *Right:* NOAO/Kitt Peak/J. Uson, D. Dale, S. Boughn, J. Kuhn.

port) record not just the location but also the energy of each and every individual X-ray photon. A typical X-ray "picture" is made up of thousands to millions of photons, sorted by energy, and displayed and analyzed in many different ways. However, this is crude spectroscopy. For example, Chandra is arguably the most powerful X-ray telescope ever built because it focuses X-rays with great precision, revealing details smaller than 1 arcsecond on the sky. But even Chandra has only a modest capability to sort photons by their energies: its CCD cameras can separate photons with an energy difference of 2 percent, a spectral resolution of $E/\Delta E \sim 50$. This is generally too low to distinguish important spectral features, such as the emission lines of distinct chemical elements and their ions. In comparison, even a comparatively low-resolution spectrum in visible light has $E/\Delta E \sim 1{,}000$.[9] The next big step in X-ray telescopes will come from cameras that produce images with higher energy resolution and from traditional spectroscopic observations (with a diffraction grating) at much greater sensitivity for a vastly larger number of point-sources—for example, faint quasars and X-ray binary stars.

Such enhanced capability is essential for addressing many key questions in astrophysics. A more sensitive X-ray spectrometer will probe the early universe: the first clusters of galaxies and the first massive black holes. X-ray maps—with better spectral resolution of extended sources like supernova remnants, galaxies, and galaxy clusters—will show how the abundances and ionization stages of the elements change and provide precise information on temperatures, densities, and motions of the gas. For example, the Doppler shift of an emission line and its shape

---

[9] The X-ray grating on Chandra does better—$E/\Delta E \sim 1{,}000$—but its use is limited to bright X-ray point sources, like stars neighboring the Sun, or Milky Way binaries containing a black hole or a neutron star, and a few very bright and relatively nearby quasars.

reveal how fast the emitting gas is moving and its level of turbulence. Such diagnostics are essential for decoding the important physical processes occurring in any astronomical object. Visible-light telescopes have done this for decades, but they cannot study the many interesting objects observable only with X-ray telescopes.

Advancing Our Understanding of the Universe with the International X-Ray Observatory—IXO

The panel recommends *IXO, the International X-ray Observatory*, as a maturing project that will effectively address key science questions, open a vast discovery space, and provide for the community a versatile tool that will support all of astrophysics.

The heart of IXO is a large-aperture (3 m^2), lightweight, focusing X-ray mirror—more than 10 times that of any previous mission—that should achieve 5-arcsecond angular resolution. The key component of the IXO focal plane is the X-ray Microcalorimeter Spectrometer—a 40×40 array of transition-edge sensors (TESs) covering several arcminutes of sky with energy-dependent spectral resolution of $E/\Delta E = 250\text{-}3{,}000$. TESs are state-of-the-art superconducting devices that accurately measure the energy deposited by, in this case, X-ray photons. IXO also carries traditional X-ray gratings, optimized for point sources, that will provide a factor-of-three greater spectral resolution compared to current X-ray grating spectrometers, at 10 times the effective aperture (between 0.5 and 2.0 keV). Figure 6.6 compares IXO to contemporary X-ray facilities in terms of effective aperture (light-gathering power) and energy resoluton. A schematic view and layout of basic components is shown in Figure 6.7.

The microcalorimeter instrument is, in effect, a powerful X-ray integral-field spectrometer—a revolutionary new capability: moderate spatial and spectral resolution over substantial areas of sky. IXO's large collecting area also delivers sufficient photons to resolve the time-variability of spectra from black hole accretion disks, neutron stars, and AGN over intervals of tens of seconds. IXO's combination of spatial and spectral resolution and large throughput will enable scientific breakthroughs much beyond the capabilities of current X-ray telescopes.

IXO Science

The science questions IXO will answer are broad and penetrating and, for the most part, impossible to answer without X-ray spectroscopy. For example:

- Does energy feedback from a supermassive black hole—an AGN or a quasar—suppress star formation in its host galaxy? How does the energy pouring from supermassive black holes affect intergalactic and intracluster gas?

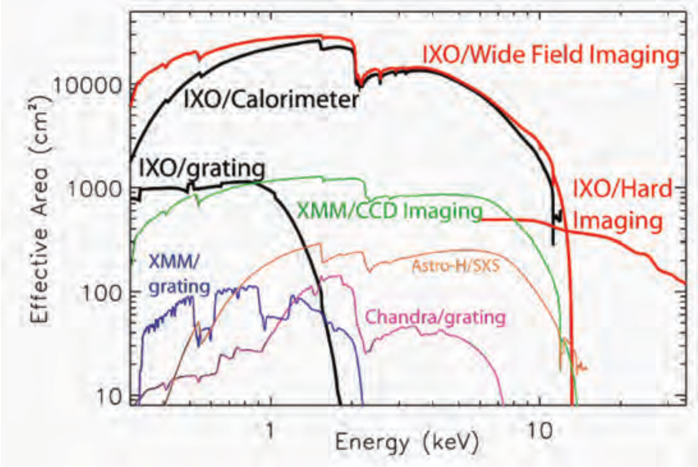


FIGURE 6.6 IXO compared to contemporary X-ray missions. IXO provides an order-of-magnitude more aperture (light-gathering power) over a vast energy range. Chandra CCD imaging falls significantly below the XMM CCD imaging capability in terms of effective area. SOURCE: J. Bookbinder, on behalf of the IXO Study Coordination Group, "The International X-ray Observatory," Astro2010 white paper, available at http://constellation.gsfc.nasa.gov/decadal_references/IXOAstro2010RFI2Web.pdf.

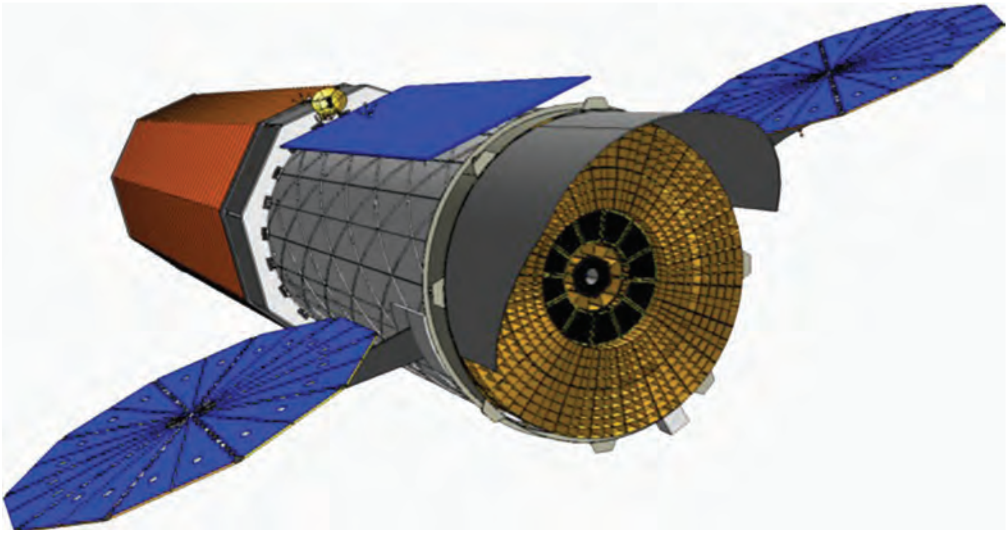


FIGURE 6.7 Conceptual image and schematic layout of IXO. The telescope is viewed from the mirror end, which, like a refracting optical telescope, is the front. A schematic layout of basic components is shown on the right. The field of view covered by the calorimeter is represented by the large circle. Four other instruments, with smaller fields of view—including the high-energy imager and high-time-resolution spectrometer—are attached to a movable instrument platform. The XGS instrument uses diffraction gratings that illuminate a CCD camera. SOURCE: J. Bookbinder, on behalf of the IXO Study Coordination Group, "The International X-ray Observatory," Astro2010 white paper, available at http://constellation.gsfc.nasa.gov/decadal_references/IXOAstro2010RFI2Web.pdf.

- What are the energetics and heavy-chemical-element content of galactic winds, and how do they affect a galaxy's surroundings? How much energy and heavy-element material escape from galaxies into the intergalactic medium?
- How much intergalactic gas is there, and how hot is it?
- What are the masses and spins of black holes? What spins up black holes, and what determines their masses?
- Are the progenitors of Type Ia supernovae always the same kinds of objects?
- At what mass threshold does an exploding star leave a remnant black hole instead of a neutron star?
- What does the neutron star equation of state tell us about the theory of quantum chromodynamics?
- How does structure formation proceed in the densest regions of the universe—clusters of galaxies—gauged by the evolution of the hot gas component that dominates the baryonic mass?

In the following, the panel expands on some of these science questions and the role IXO will play in answering them. The discussion is organized by the themes represented by four of the five Science Frontiers Panels of Astro2010.

Galaxies Across Cosmic Time (GCT)

Computer simulations show that the paradigm for structure growth, the ΛCDM model, successfully predicts the mass and distribution of dark matter on large scales. However, the simple assumption that the normal matter—baryons—starts out cold and falls into the dark-matter halos fails to explain key observations—for example, the ratio of numbers of bright to faint galaxies, the colors of the most massive galaxies, and the correlation between masses and X-ray luminosities of clusters of galaxies.

At least one additional ingredient must be added to our models, and this ingredient is generically called "feedback." Baryons falling into gravitational potential wells of galaxies shock, compress, condense into clumps, cool, and form stars. Star formation itself creates winds of hot gas and leads to supernova explosions that dramatically rearrange the gas, enrich it, and suppress and/or stimulate further star formation. This is a key ingredient of the feedback process.

Another effect results from the relatively small amount of gas that reaches the galaxy's center, where it can join an accretion disk surrounding a *supermassive* black hole—millions or even billions of solar masses. As this gas feeds into the black hole, some 10 percent of its rest mass is transformed into energy, producing huge outbursts of kinetic energy and radiation—feedback—that can transform the environment of the host galaxy and beyond, to distances of thousands, perhaps millions, of light-years.

Chandra has revealed powerful shocks arising in the cores of galaxy clusters in the nearby universe—a central black hole is the only plausible explanation. Figure 6.8 shows such a galaxy with radio jets and a hot X-ray halo. But how can a radio galaxy, ejecting energy in two directions, heat the intracluster gas in all directions? Theorists have proposed various mechanisms to transport the energy through the cluster, including magnetic fields, conduction, and cosmic-ray heating, but there are few data to test even the most rudimentary models.

The IXO high-resolution imager will provide diagnostics, from measurements of the total masses and galaxy contents of clusters to the turbulence and bulk motions of the intracluster gas. The violent deaths of massive stars in galaxies also drive winds, and the kinetic and thermal energies and the distribution of elements in the winds can point back to different types of supernovae. Such previously unobtainable data are crucial for progress in the study of feedback.

Also key to the story of galaxy evolution across cosmic time is the answer to the fundamental question, Where are the baryons? Although undetected, most of

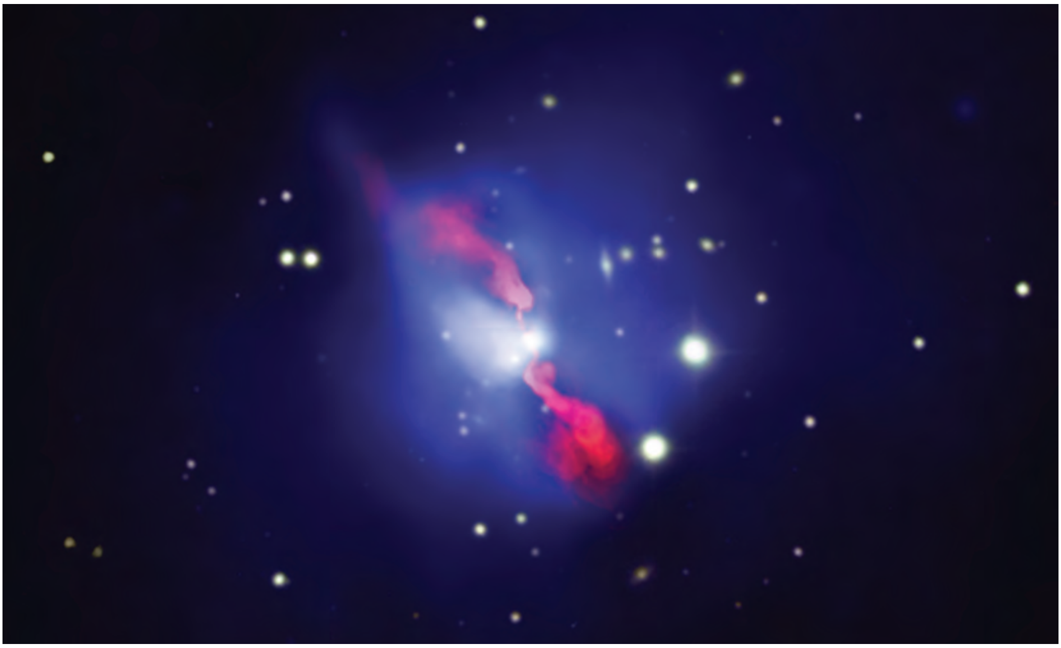


FIGURE 6.8 Composite X-ray/radio image of Hydra-A, a radio galaxy. The red jets are radio emission from hot gas streaming from the nucleus at close to the speed of light. The blue halo shows X-ray emission from very hot, pressure-supported gas. Note the two apparent “holes” where the red jets meet the X-ray halo. Studying energy transfer in a situation like this is crucial to our understanding of galaxy feedback. SOURCE: *X-ray:* NASA/CXC/University of Waterloo/C. Kirkpatrick et al. *Radio:* NSF/NRAO/VLA. *Optical:* Canada-France-Hawaii-Telescope/DSS.

the baryons in the universe must nevertheless lie between the galaxies, in a hot, tenuous gas only slightly polluted by heavy elements. The detection of this hot, diffuse, intergalactic gas has been a prize pursued with XMM and Chandra spectroscopy, but the goal has remained elusive. IXO, with its high-resolution grating, will be able to provide definitive answers as to whether this hot gas is common or rare, what is its physical state, and how it relates to the ubiquitous cooler gas seen by O VI absorption in UV light.

Galactic Neighborhood (GAN)

Circumgalactic matter is heated, shocked, and enriched with heavy chemical elements by powerful galactic winds driven by the explosions of stars. IXO will detect and study these winds by using quasars and X-ray binary systems to "backlight" the gas and measure its absorption. H- and He-like lines of O and Ne, and L-shell Fe ions (0.5-2.0 keV) broader than 100 km/s will be resolved, and crucial measurements of absorption-line velocities established to 10- to 20-km/s resolution.

Galactic winds are also "seen" in X-ray emission lines between 0.5 and 2.0 keV. Imaging with the microcalorimeter, sensitive to a few 10^{-15} erg s^{-1} cm^{-2} $arcmin^{-2}$, will allow studies of the energy content and chemical compositions of the winds emanating from ~40 starburst galaxies and, for the first time, the kinematics of the X-ray gas—bulk flows and turbulence. Figure 6.9 shows a composite X-ray/optical image of a local starburst galaxy, M82, and illustrates the capabilities of Chandra and IXO to study the hot gas erupting from such a galaxy.

Stars and Stellar Evolution (SSE)

X-ray observations provide possibly the best, if not the only, opportunity to determine the masses and radii of neutron stars. The mass distribution can tell us whether there are different paths of stellar evolution that make neutron stars. The mass-radius relation for neutron stars constrains the equation of state of matter at supra-nuclear density. This is a vitally important test of quantum chromodynamics, QCD, which predicts the state of matter in the early universe, quark confinement, the role of gluons, and the structure of the proton and neutron. Figure 6.10 shows the kind of data that IXO will obtain and its application to the problem.

Massive stars end their lives as supernovae—stupendous explosions that are the most violent events in the universe. Figure 6.11 shows the emission from the Tycho supernova remnant, comparing a composite Spitzer/Hubble/Chandra image to a simulated IXO calorimeter image. IXO will provide critical information, such as the Mn/Cr ratio, which is sensitive to the electron-to-nucleon ratio in the progenitor white dwarf, and abundances of trace elements beyond O, Ne, C, Si, S, and Fe in about a half-dozen Milky Way supernova remnants, and it will be able

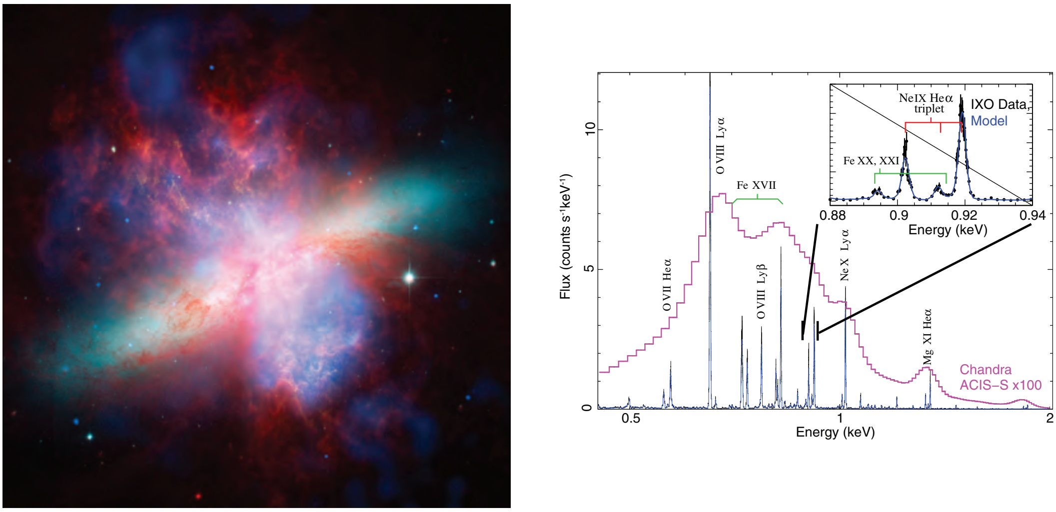


FIGURE 6.9 *Left:* Composite X-ray/optical image of M82, a nearby starburst galaxy. The red clouds are Hα-emitting filaments from radiating outflows. The blue haze is from a Chandra image that shows the hotter X-ray-emitting gas of the outflow. *Right:* IXO will have the ability to take high-resolution spectra to study the properties of this gas, as shown in the comparison of the Chandra spectral resolution (magenta line) to a simulated IXO spectrum of sharp emission lines. SOURCE: *Top: X-ray*—NASA/CXC/Johns Hopkins University/D. Strickland; *Optical*—NASA/ESA/STScI/AURA/The Hubble Heritage Team; *IR*—NASA/JPL-Caltech/University of Arizona/ C. Engelbracht. *Bottom:* Courtesy of Dave Strickland.

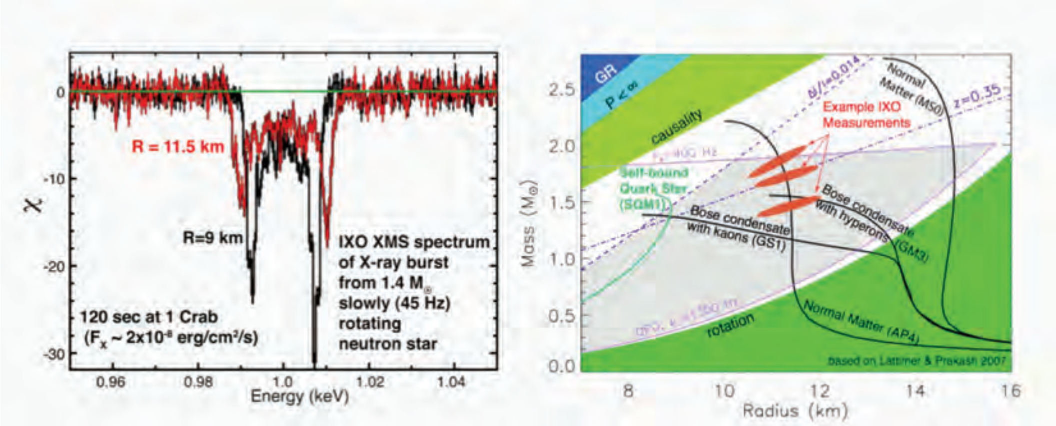


FIGURE 6.10 *Left:* Resolved line-profile during a burst of an accreting neutron star, plotted as normalized signal intensity χ as a function of energy. By measuring the response of a neutron star to material falling on it, the star's radius and mass can be independently estimated. Such measurements could constrain QCD models of the neutron star equation of state. *Right:* The mass-radius plane for neutron stars with representative theoretical curves for different assumptions for the equation of state. The observational constraints that IXO could provide are indicated by the small red ellipses. SOURCE: *Left:* M. Méndez. *Right:* Reprinted from J.M. Lattimer and M. Prakash, Neutron star observations: Prognosis for equation of state constraints, *Physics Reports* 442(1-6):109-165, copyright 2007, with permission from Elsevier.

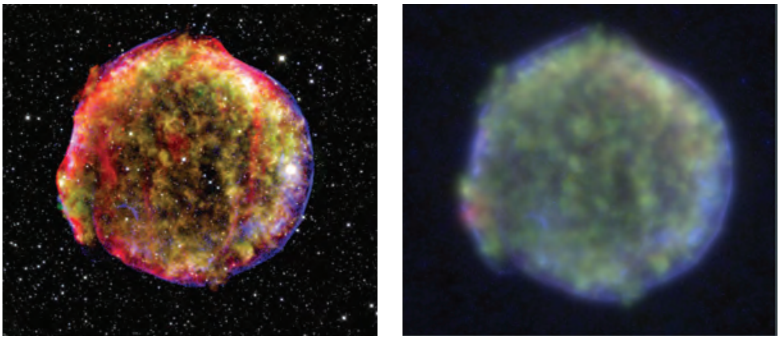


FIGURE 6.11 Composite mid-IR (Spitzer)/optical (Hubble)/X-ray (Chandra) image of the Tycho Supernova (*left*), compared to a Chandra X-ray image degraded to show IXO spatial resolution (*right*). Although the IXO image suffers in comparison to the high spatial resolution of Chandra, the IXO calorimeter will provide full spectroscopic data for each point in the image, as shown by the simulated emission in Fe L (red), Si K (green), and the 4- to 6-keV continuum (blue), at a much increased spectral resolution (see Figure 6.9). SOURCE: *Left: X-ray*—NASA/CXC/SAO; *Infrared*—NASA/JPL-Caltech; *Optical*—MPIA, Calar Alto, O. Krause et al. *Right:* Courtesy of John P. Hughes.

to acquire spectra of remnants in other Local Group galaxies. Such data can be used to make a definitive determination of the pre-explosion mass that separates a core-collapse supernova forming a black hole from one that forms a neutron star.

Cosmology and Fundamental Physics (CFP)

As mentioned above in the WFIRST section, cosmic acceleration as a manifestation of "dark energy" could in fact be an incorrect diagnosis: a departure from Einstein's general theory of relativity on cosmic scales is an alternate explanation. WFIRST will test this using weak lensing. IXO can provide an independent assessment by counting the number/mass distribution of clusters of galaxies as a function of lookback time. Essential for this test are better mass determinations than have been previously possible, attainable through the greater sensitivity and spatially extended temperature measurements (Figure 6.12) provided by IXO. Furthermore, comparing the number of galaxy-cluster-size dark-matter halos to the number of lower-mass dark-matter halos (providing a better measurement of the shape and normalization of the halos' mass power spectrum of halo mass) would be a sensitive test of the neutrino mass, an important goal in fundamental physics.

As discussed above, IXO will measure the mass, radius, and spin of neutron stars, data that has great relevance to fundamental physics. Similarly, IXO's ability to measure the profile of the Fe K line will give the spin of a black hole of any mass, from stellar-size objects to supermassive black holes, up to billions of times the mass of the Sun.

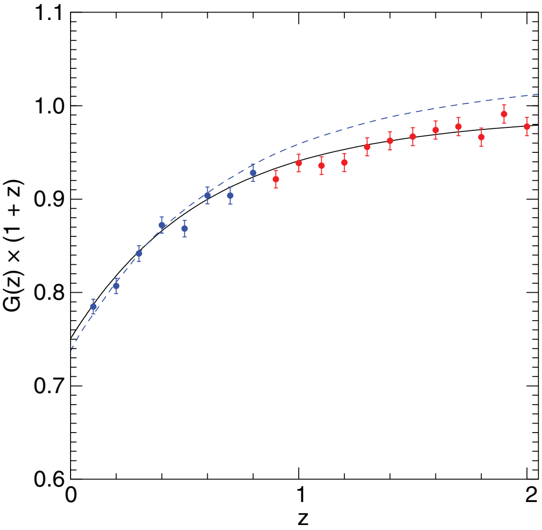


FIGURE 6.12 Testing departures from standard general relativity (GR) using cluster counts to measure the growth of structure. The normalized number density as a function of redshift is modeled to show the sensitivity to departures from standard GR. Using supernovae distance measurements in combination with IXO's measurements of cluster growth would provide highly accurate measures of $w = P/\rho$ and Ω_{DE}. SOURCE: A. Vikhlinin, S.W. Allen, M. Arnaud, M. Bautz, H. Boehringer, M. Bonamente, J. Burns, A. Evrard, J.P. Henry, C. Jones, B.R. McNamara, D. Nagai, D. Rapetti, and T. Reiprich, "Cosmological Studies with a Large-Area X-ray Telescope," Astro2010 white paper, available at http://adsabs.harvard.edu/abs/2009astro2010S.304V.

IXO Programmatic Issues

IXO enjoys the enthusiastic support of the X-ray community and the keen interest of a broad group of multiwavelength astronomers. U.S. astronomers have, over decades, developed close collaborations with ESA and JAXA astronomers that have produced stunning successes. The panel recognizes, however, that a large, international collaboration also carries with it potential inefficiencies and challenges. It is essential that the U.S. IXO team lead a concerted effort to streamline the organization of the IXO project, to clarify the relationships among the partners, and to simplify interfaces, both personal and technical. The panel also thinks it crucial that the international collaboration for IXO share risks—for example, cost growth—as well as opportunities.

IXO Cost, Risk, and Trade-offs

The independent cost appraisal process of Astro2010 descibed IXO as a project with two medium-high risks: (1) an insufficient mass margin for a project at this stage, which could lead to a larger launch vehicle, and (2) the challenge of successfully manufacturing the large-aperture mirror and achieving a spatial resolution of 5 arcseconds. The panel believes that, were these risks to lead to signficant cost growth, they would endanger the entire space astrophysics program. The panel therefore evaluated the impact on some key science programs of a 30 percent reduction in mirror area—a substantial mass reduction—and a spatial resolution of only 10 arcseconds, the state of the art. It concluded that the ability of the mission to meet its primary science goals would not be heavily compromised; degradation would be mostly quantitative rather than qualitative. Generally, the reason for this is that the principal science goals are photon limited rather than background limited; as such they can usually be compensated for simply by longer exposures. In particular, the premier instrument—the calorimeter—would still provide breakthrough science even with a factor-of-two lower spatial resolution.

The panel's interpretation of the independent cost appraisal and technical evaluation was that by rescoping IXO in this or similar ways the technical risk can be reduced to "medium." Further possibilities for reducing cost and schedule risk include the deletion of one or more instruments and/or elimination of the instrument turntable in favor of fixed stations. The panel does not recommend a specific set of rescoping options—its purpose here is to express its conclusion that options are readily available and that they would not substantially degrade the value of the mission. The key to keeping IXO's scientific priority is to feed a calorimeter with data from a much larger collecting area than has been done before.

The XGS (the grating spectrometer) was also identified in the independent cost appraisal and technical evaluation as a significant advance beyond current capabilities. Losing this capability would significantly degrade IXO's study of the intergalactic and interstellar media through absorption of the light of background AGN—the calorimeter cannot reach the required energy resolution of several thousand at the energy of interest. Although the loss of this one science program would be significant, the mission would remain as a giant step in the study of the X-ray universe. The panel recommends that the XGS be retained in the context of a "best effort" but that it not be allowed to drive the cost of the mission in a significant way. Again, the panel's point is that the mission remains a very desirable one even if the proposed performance of the XGS cannot be met.

The IXO project presented a life-cycle cost (including launch; all costs in FY2009 dollars) of $3.2 billion, with a 50 percent U.S. share being $1.6 billion. The independent cost appraisal process reported a life-cycle cost for the full IXO as $5 billion, which at 50 percent would imply a $2.5 billion U.S. share. The panel believes that this level of investment is too high, given projected resources and the necessity of a balanced program. The panel therefore concluded that any needed rescoping must be done as soon as possible, with a target U.S. share not larger than $2.0 billion to be validated by a subsequent independent cost and risk assessment. The rescoped IXO, the one that meets this target, is the one that the panel recommends.

Previous experience has also revealed significant hidden costs in major international collaborations, often at the level of tens of percent. Thus, not only the design issues, but also the definitions of interfaces and responsibilities, both technical and managerial, challenge the ability of the project to keep the U.S. contribution at the

$2.0 billion level. Given the delayed availability of funding, the panel believes that there is sufficient time for the project to demonstrate IXO's essential technologies, rescope as necessary, reduce major risks, and develop a detailed, convincing plan for the international partnership, all before committing to a new start. Adequate funds are essential to carry out these activities successfully.

MISSIONS TO SEARCH FOR AND STUDY EXOPLANETS

State of the Art

There has been remarkable progress in the discovery and understanding of exoplanets in the current decade. At the time of this writing, more than 400 have been discovered. Nearly all have masses (modulo system inclination) determined by ground-based radial velocity measurements; many are in multiplanet systems. Radial velocity measurements are now made to an accuracy of better than 1 m s^{-1}, and many systems have been monitored for a decade or more. Steady improvements in precision, which may eventually reach 20 cm s^{-1} or better have already led to the discovery of planets with orbital periods from less than 1 day to more than 10 years and with masses as small as ~5 $M_\oplus$. However, most known planets have large masses, about that of Jupiter or larger, and many orbit their stars in only days or weeks.

This impressive progress is eclipsed, so to speak, only by what has been learned from the 70 or more planets discovered to transit their host stars. Transit measurements yield a planet's radius and orbital inclination; adding a mass from radial velocity measurements allows the planet's density to be determined, which is key to modeling its nature. Visible and near-IR spectroscopic measurements during transit using the Hubble and Spitzer telescopes have detected molecules in the planetary atmospheres, and planetary thermal emission has been measured using Spitzer observations when planets pass behind their stars (secondary eclipse). The derived measurements of temperature profiles, compositions, and winds are far advanced from the situation at the start of the decade, when there was only a single known transiting exoplanet, and there were no spectroscopic or thermal emission measurements.

The recently begun CNES COROT (Convection, Rotation, and Planetary Transits) mission has already discovered the first transiting super-Earth, a planet with less than twice Earth's diameter, with more likely to follow. NASA's Kepler Discovey mission will continue to monitor a field of more than 10^5 stars for at least the 3.5-year duration of its prime mission. Kepler should detect more than 100 Earth-size, habitable-zone (HZ) planets transiting their host stars if $\eta_\oplus \sim 1.0$ (the fraction of stars with an Earth-like planet). A null result—no Earth-size planet detections—

would be very significant, constraining $\eta_\oplus$ to less than 0.1 with high confidence. The actual number of Earth-size planets in habitable zones will not be known until 2013 or later, because all candidate detections by Kepler must be followed up with high-spatial-resolution images and radial velocity measurements to rule out blended stars and spectroscopic binaries. The stars monitored by Kepler are relatively faint, and so radial velocity measurements will not be easy, and further study of the properties of these planets will be very challenging.

JWST's large aperture, stable environment, and powerful instruments will provide significant capabilities for exoplanet imaging, allowing JWST to detect directly even gas giants that are old and cool, at separations of 1 to 200 AU from their host stars. Transit spectroscopy with JWST will reveal molecular constituents of gas-giant and ice-giant atmospheres and will measure temperatures and abundances, providing insight into planet formation and accretion processes. JWST may even be able to detect molecular constituents in the atmospheres of super-Earths transiting nearby M stars, as well as thermal emissions from their surfaces. This would constitute significant progress toward the PSF Panel's "discovery area" of characterizing a nearby, habitable exoplanet.

Exoplanet Missions for the Next Decade

Together with ground-based observations, these missions will provide significant data on a wide variety of exoplanet properties, but they are just first steps toward answering important scientific questions about exoplanets. For example, the mass distribution of terrestrial planets will still not be known, because Kepler will measure only sizes. The sample of masses for small planets from radial velocity observations will probably remain small (and biased?) because of the difficulty of measuring such a small velocity shift (~10 cm/s for Earth), the accuracy limited ultimately by stellar variability. Spectra of a true Earth-twin or Jupiter-twin (similar masses at similar distances from a Sun-like G star) are well beyond the capabilities just described.

Without new facilities, little will be learned about the statistical distribution of planetary masses, and virtually nothing will be learned about the properties of individual planets—numbers, masses, densities, atmospheres, and so on—around the closest stars.

WFIRST Microlensing Survey: A Census of Planet Orbits and Masses

The microlensing exoplanet survey proposed for the WFIRST mission will measure the distribution of planetary masses to less than 1 $M_\oplus$ from semimajor axes less than 0.5 AU to beyond 10 AU. This approach will provide high sensitivity to planets on both sides of the “snow line” at ~3 AU, the orbital distance beyond which gas giants like Jupiter are believed to form. Figure 6.13 compares the sensitivities of WFIRST and Kepler: when combined they will be sensitive to planets with sub-Earth size and sub-Earth mass, from ~0.1 AU to beyond 10 AU, effectively completing the statistical survey strongly urged in the PSF Panel’s report (Chapter 4 of this volume). The prospective WFIRST upper limit of $\eta_\oplus < 0.05$ discussed above is somewhat tighter than the constraint from Kepler. Furthermore, while Kepler will be sensitive to Earth-size planets in the habitable zones (~1 AU) of F/G/K stars, WFIRST will be more sensitive to colder Earth-mass planets 0.5 to 5 AU from K and M stars. Since WFIRST results will improve with each succeeding campaign, there is a good chance that WFIRST will eventually succeed in measuring $\eta_\oplus$ over the entire range of orbital separations.

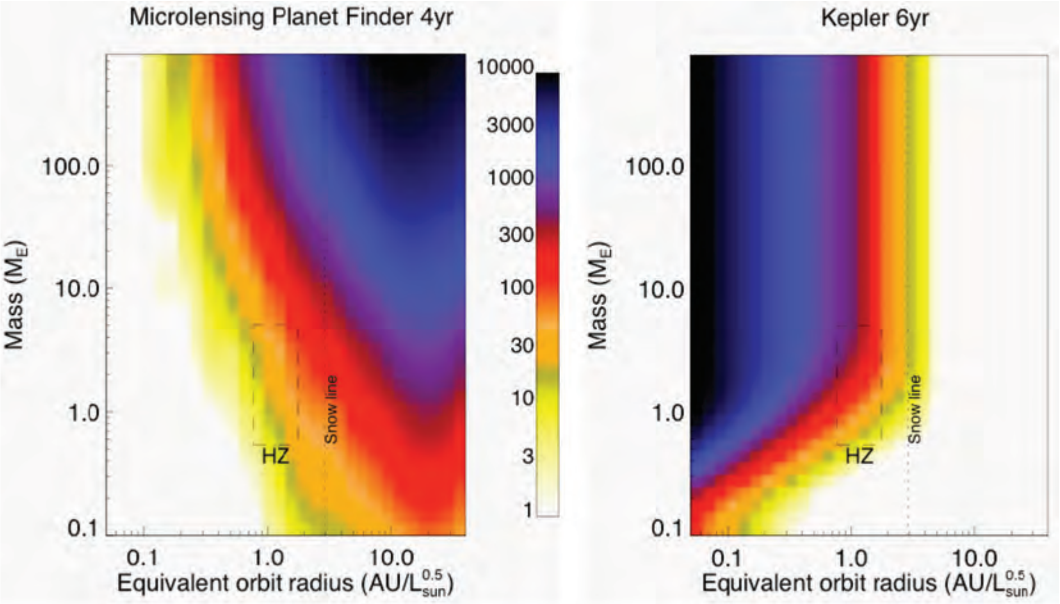


FIGURE 6.13 The yield of planets anticipated from the Microlensing Planet Finder (MPF) for a 4-year program, compared to that of Kepler, as a function of orbit radius (scaled by the star’s luminosity) (from ExoPlanet Task Force, *Worlds Beyond: A Strategy for the Detection and Characterization of Exoplanets*, Washington, D.C., May 22, 2008). The color bar between the images indicates the number of detections. The straw-man program for WFIRST microlensing observations would yield approximately half as many detections as the MPF program in the first 5 years. Boxes (HZ) indicate the habitable zone. SOURCE: D. Bennett J. Anderson, J.-P. Beaulieu, I. Bond, E. Cheng, K. Cook, S. Friedman, B.S. Gaudi, A. Gould, J. Jenkins, R. Kimble, D. Lin, et al., “Completing the Census of Exoplanets with the Microlensing Planet Finder (MPF),” Astro2010 white paper, available by request from the National Academies Public Access Records Office at http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=48964.

Surveying and Characterizing the Closest Planetary Systems

The Kepler and WFIRST surveys will determine the statistical distributions of exoplanets, while future transit and secondary-eclipse observations will reveal much about the atmospheres of giant planets and perhaps something about terrestrial planets as well. However, new missions are needed to take the next crucial steps: to search for planets around the closest stars, to measure their masses, and to characterize their atmospheres.

The panel is enthusiastic about one such program, SIM Lite—a 5-year mission to measure planet masses down to ~1 $M_\oplus$ for the nearest ~60 stars. Because these stars are mostly within 10 pc, planets orbiting them would be good candidates for future direct-detection (imaging) and characterization missions. The SIM Lite sensitivity to nearby planets is shown in Figure 6.14. Its single-pointing precision of ~10 microarcseconds or better would also allow astrometric measurements of other targets, such as massive stars, neutron stars, AGN, and other targets too faint

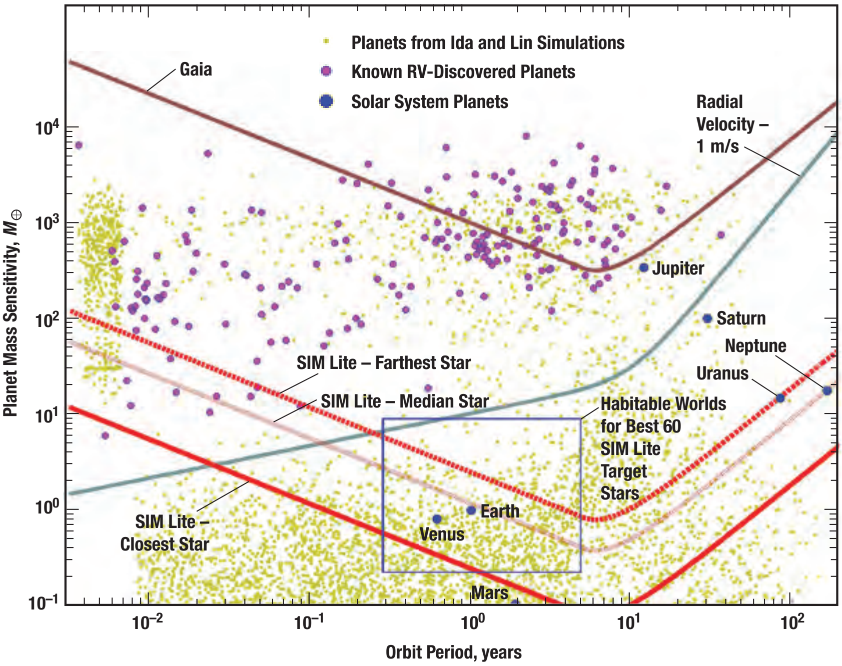


FIGURE 6.14 SIM Lite planet-search-space sensitivity comparison. The red lines bracket the parameter space where SIM Lite can detect planets, a function of both orbital period and planetary mass. The dots show planets from theoretical simulations (green), known planets from detections of radial velocity (purple), and planets in our own solar system if placed at a distance of 10 pc. SOURCE: J. Marr, et al., "SIMLite: Astrometric Observatory," Astro2010 white paper, available by request from the National Academies Public Access Records Office at http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=48964.

for the all-sky ESA Gaia mission. SIM Lite could detect stellar streams in the Milky Way's halo and map the shape of the galaxy's dark-matter halo through precise measurements of space motions of stars in the halo.

In the independent cost-appraisal process of Astro2010, SIM Lite was judged to be low-risk and technically ready for Phase C, with a lifetime cost (including launch; all costs in FY2009 dollars) of $1.9 billion, to be compared with the project's estimate of $1.4 billion. The panel agrees with the Astronomy and Astrophysics Advisory Committee's Exoplanet Task Force (ExoPTF) and the Astro2010 PSF Panel that a sub-microarcsecond astrometry mission like SIM Lite would be the current best choice for a new exoplanet mission. However, as the third-ranking priority in the panel's program, any exoplanet mission could start only late in the decade, at the earliest. Because exoplanet science, observation techniques, and technologies are advancing rapidly, the panel recommends that the present preference for an astrometry mission be re-evaluated by an expert panel if and when a new start becomes possible. A mission that is capable of precision astrometry and high-contrast imaging and spectroscopy—a concept emerging recently and still very immature—could be especially powerful. Near-term technology development will be crucial to assess its feasibility.

Alternative missions to study the planetary systems of nearby stars are developing rapidly. Direct detection and spectroscopy are compelling for both their scientific and their "exploration" value. Even low-resolution spectra can distinguish between planetary atmospheres dominated by water or by methane, and can identify non-equilibrium abundances that could indicate the presence of life on other worlds, as highlighted in the PSF Panel report. However, the efficient starlight suppression systems that are needed are not yet mature. It is the judgment of the panel that missions capable of characterizing a significant number of Earth-like planets could not be started before 2020. However, the panel did evaluate, and found appealing, several "probe-class" concepts employing ~1.5-m primary mirrors and internal starlight suppression systems, often coronagraphs with advanced wavefront control. Each was judged to be technically feasible after completion of a several-year technology development program, and could cost significantly less than a precision-astrometry mission like SIM Lite.[10] Such a mission could image about a dozen known (radial velocity) giant planets and search hundreds of other nearby stars for giant planets. Importantly, it could also measure the distribution and amount of exozodiacal disk emission to levels below that in our own solar system (1 zodi) and detect super-Earth planets in the habitable zones of up to two

---

[10] However, the independent cost appraisal and technical evaluation for one representative example—a moderate-size telescope with a coronagraph—judged that the risk of immature technologies would raise the cost to $1.6 billion (FY2009 dollars) compared to the $0.8 billion estimated by the project.

dozen nearby stars. These would be extremely important steps, both technically and scientifically, toward a mission that could find and characterize an Earth-twin.

The level of zodiacal-dust emission from disks around nearby stars is an unknown but crucial factor for designing planet-finding missions. Figure 6.15 is a simulated image from a 1.5-m-aperture coronagraphic space telescope of the radial velocity giant planet 47 UMa b; an exozodiacal dust disk with 3 zodis of material has been added. A terrestrial planet would be much more difficult to detect with this level of exozodiacal dust. Measuring exozodiacal dust to levels of 1 to 10 zodi in the habitable zones of nearby stars is critical for understanding the content and

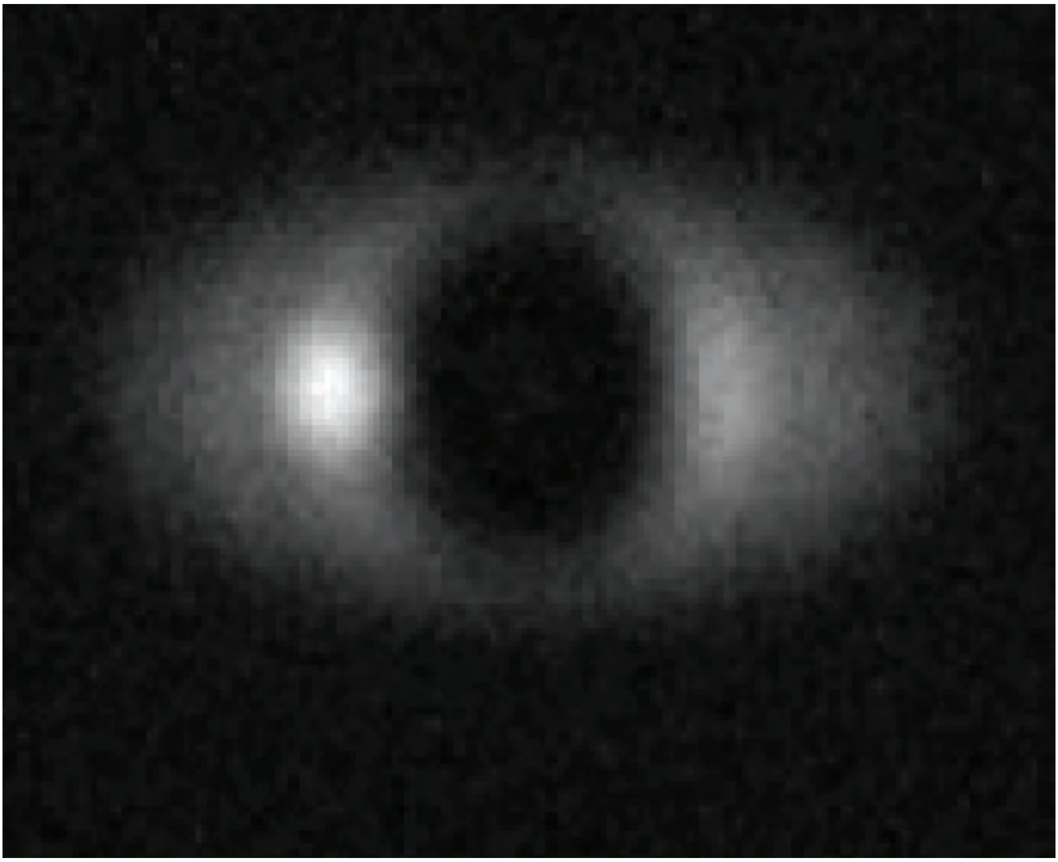


FIGURE 6.15 Simulated image of the giant planet 47 UMa b including a 3-zodi dust disk, as observed with a small coronagraphic space telescope. In this case the planet is easily detected against the zodi background, but moving to smaller planets and/or higher zodi levels rapidly diminishes the contrast. The zodi levels for typical stars in the solar system have not yet been measured, but that could be done with a small coronagraphic mission. SOURCE: Tom Greene, NASA Ames Research Center.

dynamics of small rocky bodies in planetary systems and for planning ambitious future imaging missions. It is highly unlikely that such measurements will be done with extant missions or from the ground, but the probe-class mission described above—and perhaps even an Explorer-class mission—will do the job. The panel believes that making this important measurement within the decade would be an important step on the path to a future misson to find and characterize planets around nearby stars. Another rapidly developing technique is to discover transiting planets and make sensitive spectroscopic observations of the host star's light during transit to probe the planet's atmosphere. The PSF Panel rated this among its highest priorities, and the EOS Panel endorses that position. JWST will be highly capable of studying transiting giant and super-Earth planets around very nearby stars, but it would be inefficient to use JWST to find them. Ground-based surveys are fully capable of discovering transits of giant planets like Jupiter, which would produce a ~1 percent dimming of the star's light. However, an Earth-size planet would dim the light by only 0.1 percent to 0.01 percent, depending on the size of the star, too difficult a detection for ground-based observations. For small planets, a survey with a high-precision space telescope covering a substantial area of sky will thus be needed to assemble a suitable target list for JWST. A space survey is also needed to find planets sufficiently far from their stars to be cool enough to support life—these would orbit once every few weeks or months. Such a list of nearby stars with small transiting planets will enable JWST to address the PSF discovery area of identifying and characterizing nearby habitable exoplanets.

A recent Small Explorer (SMEX) mission-concept study described a space-based, all-sky transit survey in visible light that could find super-Earth planets for JWST follow-up. The panel also reviewed a concept for a similar survey at infrared wavelengths, which would be more sensitive to planets around the small and numerous M stars. Such a facility might be implemented within the Explorer program, with a development time short enough to yield targets sufficiently early in the lifetime of JWST. The panel believes that the Explorer-class scale is well suited for this activity and looks forward to strong proposals for this approach.

In summary, the panel believes it is too early to choose an exoplanet mission that might be started late in the decade. An astrometric, direct-detection, or transit misson, or something not yet developed, are all prime candidates. Therefore, the panel recommends a generic exoplanet mission as its third priority, after WFIRST and IXO, but leaves this choice to later judgment. Any of the missions briefly described here will make good progress toward addressing the SFP science question PSF 3, How diverse are planetary systems? The more intriguing question, PSF 4, Do habitable worlds exist around other stars, and can we identify the telltale signs of life on an exoplanet? is considerably more challenging. The panel discusses this challenge further in the following section.

AN ADVANCED ULTRAVIOLET-OPTICAL TELESCOPE

The science and mission discussions above frequently resonate with the idea of a more capable UV-optical telescope to follow the Hubble. As discussed above, scientists now know that most baryonic matter is in the form of low-density gases and plasmas in galaxy halos and the intergalactic medium. High-resolution UV spectrosocopy is essential for detecting gas in these plasmas and probing low-density inflows and outflows of matter and energy that profoundly affect galaxy evolution. Figure 6.16 shows real and schematic examples of O VI absorption from such gas in the light of a background quasar (QSO).

Of course, as the Hubble has shown, deep, high-resolution imaging in the near-UV and visible is a touchstone for a good deal of astronomical research, including those programs that are built mainly around other wavelengths, from radio to gamma rays. It is difficult to imagine an advanced UV-optical telescope that would not, in addition to spectroscopic capability, include high-resolution cameras, probably with considerably greater areal coverage than has been possible with the Hubble.

Furthermore, as discussed above in the section "Missions to Search for and Study Exoplanets," answering what is arguably the ultimate question in this field, Are there other worlds like Earth, and do they harbor life?, is a formidable challenge, and so it is important to acknowledge that a promising way to search for habitable planets and life employs the platform of a large-aperture, UV-optical telescope.

What these three elements share is great promise for a facility that cannot happen without significant technology advances to make these science goals attainable and—especially if a large aperture is required—affordable. The panel views this as a unique opportunity that requires a *dedicated technology development program*. Because it believes that a UV-optical telescope is a particularly strong candidate for a new start in the 2021-2030 decade, the panel recommends pursuit of several different technologies over the coming decade that will, together, help define what form this keystone facility will take:

1. Previous UV spectrographs on HST and the decommissioned Far Ultraviolet Spectroscopic Explorer (FUSE) have provided only glimpses of the missing baryons. Significant progress will be enabled by the COS on the Hubble, but such recent UV instruments are still hampered by low-quantum-efficiency detectors and significant light-loss from non-optimal optics coatings. It is estimated that an order-of-magnitude improvement in efficiency is possible. With significant improvements in detectors and coatings, substantial progress in "baryon science" could be achieved with an aperture no bigger than that of the Hubble.
2. Measuring the masses and spectra of Earth-like planets around a significant number of nearby stars is a challenging technological problem because the contrast

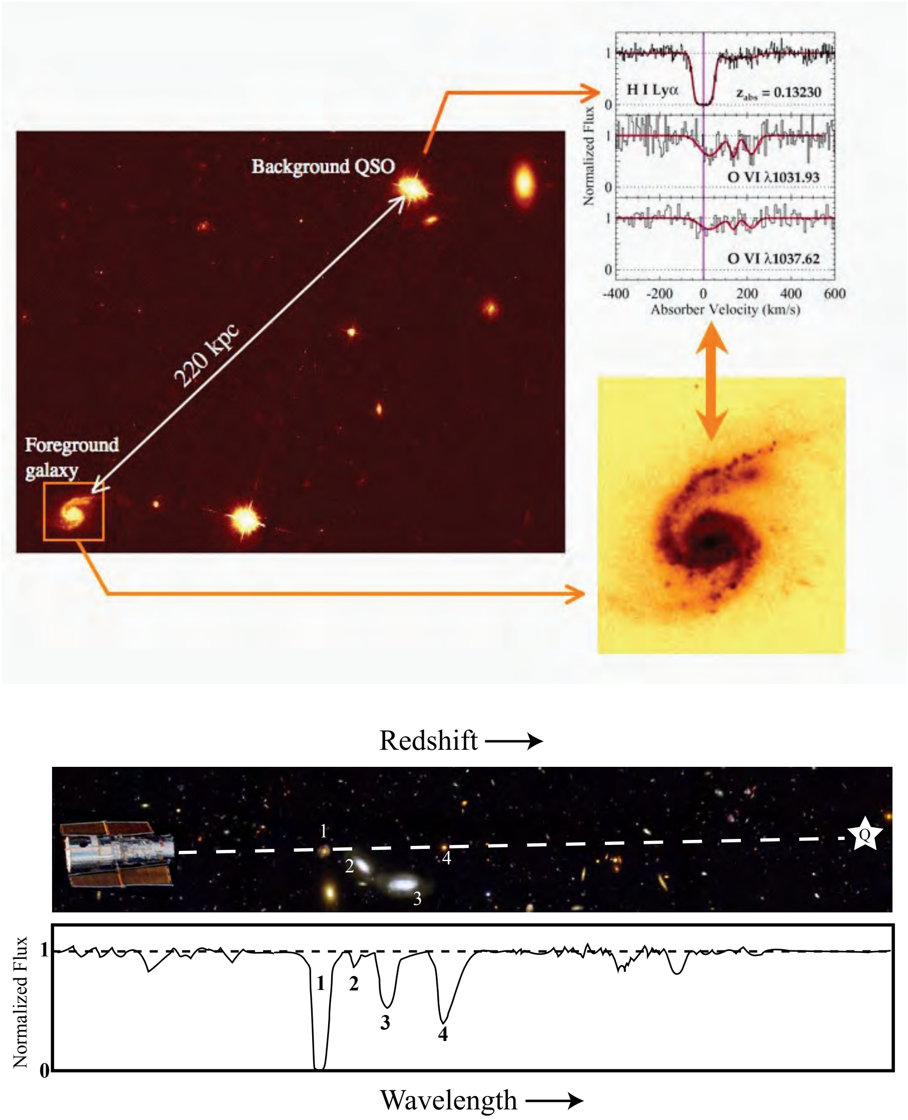


FIGURE 6.16 *Top:* Ultraviolet O VI absorption features in circumgalactic gas seen from back illumination from a quasi-stellar object. The gaseous halo is vastly larger than the optical extent of the galaxy and likely contains most of the baryonic matter. *Bottom:* Cartoon example of the absorption signatures from intervening galaxies as quasar light passes through. Absorption often arises far from the visible boundaries of a galaxy. SOURCE: *Top, clockwise:* Todd Tripp, University of Massachusetts, Amherst, adapted from Figure 29 in T. Tripp, K.R. Sembach, D.V. Bowen, B.D. Savage, E.B. Jenkins, N. Lehner, and P. Richter, A high-resolution survey of low-redshift quasi-stellar object absorption lines: Statistics and physical conditions of O VI absorbers, *Astrophysical Journal Supplement Series* 177:39, 2008; Todd Tripp, University of Massachusetts, Amherst. *Bottom:* J.C. Howk, University of Notre Dame.

between the star's light and that of an Earth-twin would be enormous—a factor of ~10 billion. Although significant advances in coronagraphic and like techniques have been made in the laboratory toward this level of "starlight suppression," substantial work remains to demonstrate that this is practical and that it can be made sufficiently robust for a space mission. Such concerns have led other groups to promote an alternative: external star shades that serve as artificial moons to eclipse the light of distant stars so that their much fainter planets can be studied. This is an approach with a technical challenge as large as its promise. Concerted development of all types of starlight suppression techniques is essential if the large-optical-telescope approach to exoplanet research is to prove feasible, buildable, and successful.

Could a single telescope do both the baryon and exoplanet science, and include the broad astrophysics program that the Hubble has supported with its superb cameras? It is not obvious. Among the chief concerns, for the case of internal starlight suppression, are the uncertain characteristics of putative high-performance UV coatings: wavefront-cancellation techniques place stringent demands on mirror coatings. External star shades avoid this problem, but the necessity of a second spacecraft that must be launched, precisely deployed, and able to fly in formation at a great distance from the UV-optical telescope could take such a mission beyond the $5 billion expenditure that, at this time, seems to be the upper bound for the cost of an astrophysics space mission.

3. Since both UV spectroscopy and exoplanet science will be "photon starved," the quest for a larger-aperture telescope—4 m or greater—will be hard to resist. However, cost is likely to be the limiting factor, and so it was good to learn that several engineering groups are experimenting with radical new ways to achieve large-aperture mirrors for space telescopes at a fraction of the weight and cost of the present generation. Deployable segmented mirrors are but one option.

It is abundantly clear that each of these technology challenges—UV coatings and detectors, starlight-suppression techniques, and bigger, less expensive mirrors—needs to be vigorously supported if a path is to be cleared for building the most capable UV-optical telescope, the Hubble successor. By combining these technology initatives with the science results that will come in this decade—for example, the "missing baryon" search with the COS, the Kepler survey that will tell us how common Earth-like planets are, and possible measurements of exozodi levels for nearby stars—the astrophysics community would be in a far better position to know in what direction to move and how far to reach.

The EOS Panel believes that, if technology developments of the next decade show that a UV-optical telescope with a wide scope of observational capabilities can also be a mission to find and study Earth-like planets, there will be powerful reason to build such a facility. If a single facility cannot effectively carry out both programs, the panel recommends separate paths, including renewed attention to mid-IR interferometric planet finders (e.g., TPF-I), and the prompt development of a plan for a UV-optical telescope for the 2021-2030 decade.

MODERATE AND SMALL PROGRAMS

"BLISS": U.S. Participation in the JAXA SPICA mission

The EOS Panel recommends proceeding expeditiously with a high-sensitivity, moderate-spectral-resolution spectrometer for the JAXA-led Space Infrared Telescope for Cosmology and Astrophysics (SPICA). The panel reviewed a description of a specific concept for such an instrument, the Background-Limited Infrared-Submillimeter Spectrograph (BLISS), but the panel's recommendation applies to any instrument with similar capabilities and cost. For convenience, the panel uses the BLISS acronym in the following discussions, without taking a position on the competitively selected approach.

BLISS represents a consensus approach, as summarized in the U.S. Far-Infrared Astrophysics Community Plan that was submitted to the Astro2010 decadal survey. It provides an exciting opportunity for new science, building on results from Spitzer, Akari, and Herschel. Contributing BLISS will also allow testing of the rapidly advancing far-IR-detector and cooling technologies, in terms of both development for flight and of providing experience and heritage with respect to in-flight performance. The endorsement of BLISS by the infrared community is based on its ability to achieve these goals within a strongly resource-constrained environment. BLISS will help build scientific and technical infrastructure that will enable the more powerful far-IR missions of the future.

SPICA promises breakthroughs in far-IR astronomy. Compared to Spitzer, it will have a much broader instrument suite, with advanced detector arrays having orders-of-magnitude improvements in performance. The 3.5-m aperture of SPICA enables directly both higher sensitivity and resolution, and it combats the confusing noise that was the ultimate limitation for Spitzer in deep, far-IR observations. Compared to Herschel, SPICA will use improved detector arrays (both higher sensitivity and larger formats) and will combine them with a sufficiently cold and low-background telescope to take full advantage of their capabilities. In concert with the other instruments planned for SPICA, BLISS can provide the spectroscopic measurements of faint sources that are essential for many studies. For example, BLISS will measure the strengths of far-IR fine-structure lines in distant galaxies. These lines are important coolants for the interstellar medium, and they characterize both the level of energy input (e.g., from star formation) and the potential chemical evolution of the interstellar gas with redshift. Mid-IR fine-structure lines that signal AGN will fall within the BLISS operating range even at moderate redshifts ($z > 0.5$), and so BLISS will have unique power to identify AGN from very early times ($z \sim 6$) to the epoch when they were most common ($z \sim 2$), and nearly to the present. The high sensitivity and angular resolution of BLISS/SPICA will let astronomers use far-IR fine-structure lines to probe conditions in star-forming regions within the Milky Way galaxy, for a range of different density and metallicity environments distributed over the galactic disk. It will also be possible to explore protoplanetary disks at the time of formation of gas- and ice-giant planets; giant planets appear to be common, but our current understanding of how they form is very incomplete.

The Necessity of a Healthy Explorer Program

The Explorer program has been a key component of the NASA portfolio since the earliest days of the space program. The relatively low-cost astrophysics Explorers have been highly productive and have provided much of the transformational science of the past 50 years, for example: (1) UHURU, the first X-ray sky survey; (2) IRAS, the first all-sky infrared survey, which discovered protoplanetary dust disks around nearby stars; (3) COBE, finding compelling evidence for the big bang by demonstrating the exact blackbody spectrum of the CMB, detecting the primordial density fluctuations that have led to stars and galaxies, and discovering a cosmic infrared background of starlight absorbed by dust over cosmic time and re-radiated in the IR; and (4) WMAP, mapping the CMB with sufficient precision to measure accurately long-sought cosmological parameters. There are many others.

The Explorer program is also an essential programmatic element of the space-science program: (1) it allows focused investigation of key questions not readily addressable with general-purpose missions; (2) it enables relatively rapid response to changing scientific knowledge; and (3) it permits the use of highly innovative technologies at lower risk than would be the case for large missions.

The Explorer program was very active between 1995 and 2003, with six MIDEX and five SMEX missions selected for flight (although two were subsequently canceled) at an average program cost per year of about $200 million (real-year dollars). However, since 2003 there has been a steep drop in the frequency of Explorers: only one astrophysics Explorer has been selected since 2003—the recently chosen Gravity and Extreme Magnetism Small Explorer (GEMS), which is currently in the formulation phase. WISE was launched in December 2009 and NuSTAR is in development, with a launch planned for 2011. Projected budgets for the combined Astrophysics and Heliophysics Explorer programs are projected to increase $170 million per year (real-year dollars) for the next 5 years, significantly below the 1995-2003 level in real terms. Further compounding the problem is the lack of a reliable, affordable launcher for MIDEX-scale missions since the shutdown of the

Delta II program. The panel strongly recommends that NASA encourage initiatives in private industry to find a robust replacement. Re-establishing a higher launch rate for Explorers would be an important incentive toward that end.

The 2000 decadal survey endorsed “a vigorous Explorer program” (AANM, p. 28). The panel endorses that objective as well, reiterating the goal of the astrophysics community, and NASA, of an annual astrophysics Explorer launch. The 2006 NRC report *Principal-Investigator-Led Missions in the Space Sciences* (The National Academies Press, Washington, D.C., 2006) offers extensive advice for strengthening the Explorer program, together with insights from the community that should help to improve its effectiveness as it is re-invigorated. The EOS Panel’s endorsement includes a recommended augmentation of the present support level for astrophysics Explorers that would restore a launch rate of one Explorer per year by the end of the decade (see below the section “Funding a Balanced Program”).

The Suborbital Program

The Suborbital program is a small but essential part of NASA’s overall program of science and technology development. Over the past decade it has had multiple successes across broad areas of science, including measurements of the CMB, IR studies of the early universe, and detection of cosmic rays and even neutrinos. Along the way, suborbital experiments have tested technologies and techniques that enable future missions. Because of the relatively low cost, the program also offers an excellent environment in which to train graduate students and young postdoctoral scientists. Many leading astrophysicists, including many leaders within NASA, gained invaluable early experience in the program. The generic utility of the Suborbital program is widely recognized by the community. Of the programs submitted to the EOS Panel by the community, ~25 percent of them incorporated suborbital work (the bulk of which were for the balloon program), ranging from simple checks of technology, to crucial pathfinders, to full suborbital-based science programs. Easy, cost-effective access to near space is essential.

The study of the CMB is a case in point. Our remarkable progress in cosmology is due primarily to innovation in detectors and instruments that continue to push measurements of CMB anisotropies to astonishingly small levels. The CMB community informed the panel that it is working toward a next-generation satellite (a Planck successor named CMBPol) to tackle the exceedingly difficult measurements of CMB B-mode polarization that probe cosmic inflation and gravitational radiation from the big bang. The community stated that informing the design of this next mission requires continuation of a combination of technology development, ground-based, and balloon-borne experiments for the next few years. Support is currently at an adequate level, but it is essential that this support, and flight opportunities, continue. The contingency of the detailed characteristics of CMBPol made it difficult for the panel to prioritize this mission; nevertheless, the panel recognizes how vitally important this research is to astrophysics: the panel gives unqualified, vigorous endorsement to this program.

More generally, the Suborbital program has come under increasing stress in recent years by being tapped to subsidize the technology-development program, particularly in the area of new detectors. At the same time, in the case of the balloon program, the payloads have become ever more capable and consequently more costly. The net result has been a drastic reduction in the number of payloads that can be supported, and a corresponding reduction in the number of flights.

In the coming decade, when budgetary constraints will limit the number of much-more-expensive satellite programs, increasing support for the Suborbital program is a priority. By doubling the access to near space, multiple areas of study will remain vibrant at relatively low cost. This will naturally require a corresponding investment in the program offices. In addition to addressing many of the key science objectives of the current decadal survey, it will help to ensure that NASA will enter the following decade in a strong position in terms of technology and expertise.

As the capabilities of the Suborbital program increase, the sophistication of the proposed instruments also increases. For example, groups are now proposing coronagraphs from balloon platforms. This bold initiative could provide unique measurements while obtaining invaluable experience with the technology. New initiatives in the ballooning and sounding-rocket programs also have the potential to increase greatly the amount of time available to payloads. The ultralong-duration balloon (ULDB) program will provide 100-day, mid-latitude flights. With sufficient support, this relatively mature program could be returning science early in this decade.

Over the last decade, the balloon program has dominated the science return from the Suborbital program for astrophysics missions. It seems unlikely that this situation will change without a significant advance in capabilities, for example, with multiorbit sounding rockets with vastly increased integration times. While still only a concept, this could transform the scientific and technological value of the sounding-rocket program. Implementing a program of multiorbit sounding rockets is estimated at $25 million, with subsequent missions costing approximately $15 million. Similarly, while the cost of ULDB missions will be incrementally larger than that of previous ballooning programs, the sophistication of their payloads will require funding levels beyond the program's current budget. Given the scope of these initiatives, the panel recommends that they should be funded outside the current program, perhaps competed within the Mission of Opportunity or Explorer lines. This will help to ensure that the low-cost management style of traditional suborbital missions is retained. In the meantime, the EOS Panel recommends that the return on science and technology development be taken as a significant consideration when determining the balance of funding between the rocket and balloon programs in the next decade.

THE R&A PROGRAM: INCREASING SUPPORT FOR TECHNOLOGY DEVELOPMENT AND THE SUBORBITAL PROGRAM

NASA's R&A program funds a wide variety of essential activities. A $70 million investment in 2008 supported the processing and archiving of data from NASA missions and provided research grants to guest observers to produce science. Of a further $68 million, half was allocated to development programs, including research-grant support for theory and fundamental physics, data analysis tools, and laboratory astrophysics, and half to technology development and the Suborbital program. In this section the panel focuses on the vital contribution of these latter activities and provides a rationale for their increased funding, a step that will, in fact, raise the level of support for all R&A activities.

Three Types of Technology Development Activities

NASA has built its successes of the last five decades on a foundation of new technology. Continued progress depends on fully developing the technology for high-priority missions, supporting development for future missions, and providing for new enabling technologies for yet-to-be-conceived missions.

Missions selected for development support their individual technology needs from the mission line. However, prior to starting formal development, a thorough understanding of the technology is necessary for accurate budgeting and scheduling. For this reason, the viability of the high-priority missions recommended in this panel report depends on investments made early in the decade for *specific and targeted* technologies to bring them to appropriate technology readiness levels (TRLs).[11] Examples of the technologies that could be developed under this program are described below.

The missions recommended by the panel will not address some key science goals described in the reports of the Astro2010 Science Frontiers Panels; therefore, some specific areas of technology effort are necessary to develop future missions that will. The longer time-horizon argues against highly targeted development—a more general approach is better at achieving the required advances. Specific examples, such as detector and optics development, are called out in the sections below.

Finally, imaginative and innovative approaches to solving problems have always prepared the way for significant leaps in capability. Support of novel ideas is

---

[11] TRL levels range from TRL 1, a basic concept, to TRL 9, proven through a successful mission. TRL 6 is desirable for a technology to be selected for a mission.

crucial to the future vitality of NASA, even when some approaches are not directly applicable to planned missions. Indeed, it would have been impossible to predict a decade ago today's advances in areas such as detector technologies and optics.

Technology Development for Recommended Missions

Most of the high-priority missions for the next decade already have well-developed programs and technology (Figure 6.17). For example, the WFIRST mission has direct technology heritage from the WFC-3 on Hubble and NIRCam on JWST. While a substantial number of detectors is required, the HgCdTe near-IR detectors for JWST meet all the current requirements for WFIRST.

As described above, IXO has been in development in one form or another for more than a decade. Techniques for building efficient large-area mirrors, X-ray microcalorimeters, and gratings are advancing at a pace to allow IXO to enter Phase B by mid-decade.

A micro-arcsecond astrometry mission is considered a strong candidate for the next exoplanet mission. At present, SIM Lite is the most thoroughly studied mission, with more than 13 years of design studies. It is ready to enter Phase C and would therefore require minimal technology development. In the EOS Panel's recommended program, however, an exoplanet mission cannot start until late in the decade. In particular, NASA should ensure that the sophisticated technologies already developed in the SIM program are not lost. However, by later in the decade—when an exoplanet mission might be selected—SIM Lite may not be the best way forward. NASA should continue to respond to new technologies in this area to provide a broad range of options when the time comes.

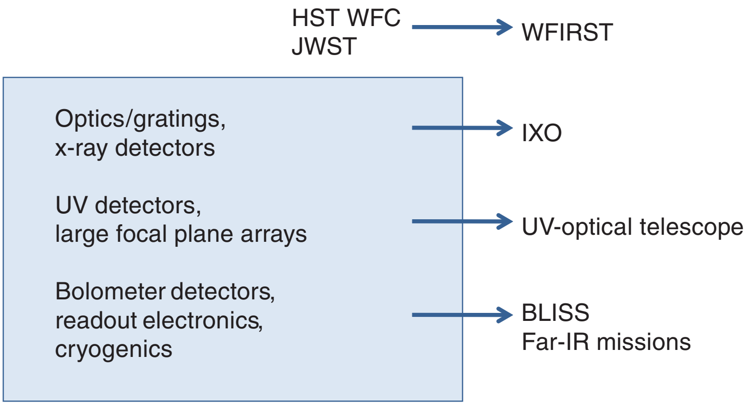


FIGURE 6.17 Near-term technology development for missions recommended by the EOS Panel.

A new UV-optical observatory requires significant progress in detector manufacturing and testing to build large focal plane arrays. These types of detectors might also benefit from suborbital testing. Along the same lines, the BLISS-like instrument for SPICA requires early investment in bolometer detectors, readout electronics, and cryogenics in 2010-2012 to maintain the JAXA/ESA schedule. It is also important to note that such development could feed directly into a U.S.-led far-IR mission in the 2021-2030 timeframe.

Technology Development in Support of Future High-Priority Science

Increased investment in technologies for missions with possible starts late in the coming decade serves several critical needs for NASA. While WFIRST, IXO, and BLISS for SPICA are well along their technology development paths, there remain key technologies for exoplanet and other possible future missions that will require substantially more work. Attention must also be paid to maintaining expertise: insufficiency of resources risks the loss of significant advances already made—many of which would be irretrievable. Finally, testing new technologies on ground-based and suborbital platforms can raise TRLs while returning high-quality science and supporting the field as a whole.

Technology for millimeter to far-IR bands has made considerable advancement, fueled early in the decade by work on missions such as Planck and Herschel—and more recently by ground-based and suborbital CMB experiments—and with facility instruments such as SCUBA2 on the JCMT. However, to prepare for future missions like the proposed CMBPol, CALISTO/SAFIRE, and SPIRIT, several key areas need to be addressed. These include large multiplexed arrays of detectors and their associated electronics, cryogenic coolers, and optical designs that address specific systematic effects (Figure 6.18). This field, in particular, benefits substantially from suborbital testing of technology.

A number of future planet-finding missions will require significant investments in technology in the coming decade to determine feasibility, risk, and cost. Missions for direct detection of planets in visible light require suppression of starlight to better than 1 part in 10^{10}. Although this level of suppression is deemed a tractable technological problem, significant investments in traditional and nontraditional coronographs, coatings, and deformable mirrors will have to be made.

Technology for the Future

The directed program outlined above is meant to close technology gaps in order to move ahead with high-priority missions in key science areas. However, new ideas have the potential to revolutionize the field, and so it is important to maintain the current level of support in the R&A program for advanced technolo-

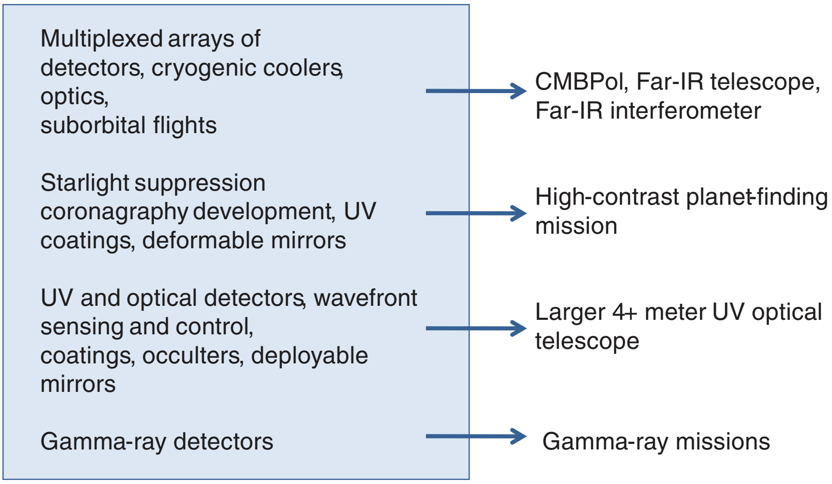


FIGURE 6.18 Technology development in support of future high-priority science.

gies, particularly ones with the potential for paradigm shifts. Which programs to support will, as always, be determined through the peer-review process. Potential areas of advancement include new mirror and optical technologies, star shades, and new types of nulling interferomety (Figure 6.19).

To pick one example, external star shades may allow enhanced planet searches with telescopes that are also well suited for ultraviolet observations, thus serving two high-priority science programs and a wide array of science investigations. However, to make star shades feasible it may be necessary to invent new approaches to manufacturing, testing, deployment, and formation flying. These initial explorations are appropriate under the current R&A approach, where new ideas are funded based solely on peer review and not on specific development goals. If these issues can be resolved and star shades can approach TRL 3 or 4, a program focused on their application in future missions would become a high priority. This example

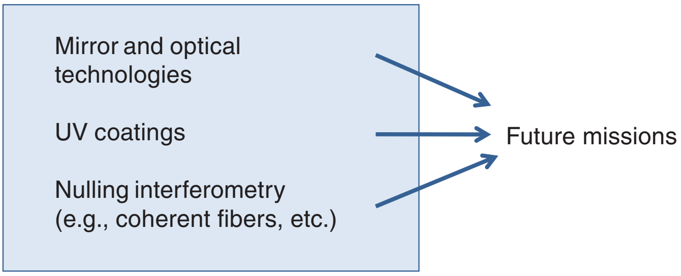


FIGURE 6.19 Technology development for the future.

illustrates both the importance of an open, peer-reviewed technology program and the necessity of a mechanism to transfer a successful paradigm-shifting technology development to the more focused efforts needed for applications in future missions.

An advanced UV-optical telescope would address many of the key questions identified by the Astro2010 Science Frontiers Panels. However, achieving a viable program by the end of the decade will require advances in detector technology, wave-front sensing and control, and UV-optical mirror coatings. Most importantly, new approaches to very lightweight yet highly accurate monolithic or segmented mirrors may be needed.

Increased Funding for Technology Development and the Suborbital Program

Funding levels largely determine whether required technologies are developed successfully—closing the technology-readiness level gap has always been challenging. Seed money has in the past come from the R&A program. However, the program budget has dropped from $79 million in 2004 to a projected $65 million in 2010 (real-year dollars, based on an annual inflation index of 1.03). Because the R&A program supports a wide variety of research, including technology development, the Suborbital program, laboratory astrophysics, and theory, all of this research has been adversely affected by declining funding levels. While the discussion here focuses on the technology and suborbital programs, all areas will benefit from the recommended budget increases.

The panel believes that the R&A program has been strained beyond its ability to meet the needs of NASA: the panel thus recommends a significant increase to correct for declining funding in the 2001-2010 decade. Specifically, $20 million per year inside the current budget should be targeted to technologies that address high-priority science questions not covered by the EOS Panel's recommendations for new starts in this decade. Based on the documents submitted to the panel by the community, an additional $20 million per year should be distributed more broadly to foster new ideas, with the aim of bringing them to TRL 3 to 5. The panel's evaluation of the needs for suborbital efforts, based again on documents submitted to the panel, shows that at least an additional $15 million per year is required. For the Suborbital program, the expectation is that $15 million per year in new funding combined with part of the $20 million in recommended new-technology funding will greatly enhance the number of missions. In recognition of the dual role of many suborbital missions, those that have a significant technology component should have access to both lines of resources through a single proposal. There are also some ideas for new, more costly suborbital initiatives (ULDB and sounding-rockets-to-orbit) that would need funding outside the traditional Suborbital program, possibly through the Mission of Opportunity or Explorer lines. In summary, the panel recommends an annual augmentation of $35 million to the existing R&A program: $20 million for new technologies, and $15 million for demonstration of technologies and science programs in the Suborbital program.

FUNDING A BALANCED PROGRAM

Projected NASA budgets provided to the panel show little or no funding for new initiatives before 2014. Thereafter, the completion of JWST leads to substantial, increasing resources in excess of the "base budget," which includes mission operations, R&A, Missions of Opportunity, and program management. The integral of the funds available for new initiatives from 2014 through 2020 could be as large as $4.0 billion (real-year dollars) with present projections. However, although this figure includes operating funds for JWST and SOFIA, it has no allowance for possible cost growth to complete them, and it does not include continued operations of existing missions such as HST and Chandra after 2015. The panel was also shown a projected budget that included assumptions about these possible added expentitures that arrived at a much lower figure of $2.3 billion for new initiatives in this decade. Conversion of these figures to FY2009 dollars would mean even less money for new initiatives by ~15 percent, unless adjustments for inflation occur.

In this final section, the panel attempts to describe the execution of the panel's recommended program bracketed by these estimates of available funds for new initiatives in the 2010-2020 decade. In Table 6.3 the panel shows funding levels and priorities for the recommended program. The panel takes as a nominal budget the middle ground, the center (green) allocations, $3.1 billion FY2009 dollars, which it considers the minimum balanced program. Sufficient funds are identified to build WFIRST, to augment the Explorer program for astrophysics by $500 million over the period, to start IXO,[12] and to fund a small exoplanet mission or start a larger one. The Explorer program, in fact, is part of the base budget, but the panel includes the augmentation it is recommending in the "new initiatives" category because the amount it recommends is too large to be accommodated by changing priorities within the base budget. When combined with the present level of funding for astrophysics Explorers, this is a sufficient allocation to raise the launch rate to one per year after 2015. The R&A augmentation (shown in green), which encompasses the additional investments discussed in this section, is intended to be phased in by rearranging priorities of the base budget and as such is not counted here as a new initiative. The U.S. contribution of a BLISS-like spectrometer to SPICA is also as-

---

[12] LISA, the Laser Interferometer Space Antenna, was reviewed and prioritized by the Astro2010's Particle Astrophysics and Gravity Program Prioritization Panel, and was not subject to EOS Panel review. However, since both LISA and IXO are collaborations with ESA and part of its competition for an L-class mission start, it is unlikely that both will move forward. For purposes of exploring budget profiles and program elements, the IXO slot could therefore be considered to represent either IXO or LISA.

TABLE 6.3 Funding the EOS Panel's Recommended Program at Different Levels

	$3.8 Billion Level	$3.1 Billion Level	$2.4 Billion Level
WFIRST	1.6	1.6	1.6
IXO (start)	1.0	0.7	0.3
Exoplanet	0.7	0.3	—
Explorer	0.5	0.5	0.5
R&A	**0.3**	**0.3**	**0.3**

sumed to be within the base budget (a NASA solicitation for science-investigation concept studies for SPICA has already been issued); approximately half of the estimated $200 million cost of SPICA participation is projected to occur before 2014.

An enhanced budget (blue) would increase the amount spent on IXO through 2020 to $1.0 billion, and allocate $700 million for an exoplanet mission. The exoplanet mission could be a SIM Lite start, a fully funded probe-class mission, or an investment in a more ambitious mission for the following decade. As described above, this choice would be made later in the decade by selective competition between SIM Lite and one of the alternatives discussed in the section "Missions to Search for and Study Exoplanets."

In Table 6.3, the budget below "nominal"—in yellow—would provide a bare start for IXO, but no funding for an exoplanet mission. In this scenario, it is crucial to maintain a healthy Explorer program so that some diverse science goals can still be pursued, albeit at a much more modest level. However, the panel does not consider this level of funding sufficient for a balanced program.

In any scenario, the panel gives highest priority to building WFIRST as expeditiously as possible, with the goal of launching it within the decade, and then to augment the funding for astrophyscis Explorers as rapidly as possible. Any additional funding should next go to IXO, and then to the Exoplanet mission—consistent with the priorities of this panel's report. Even in the best case, building IXO will require substantial further investment in the 2021-2030 decade. In the minimal budget case, funding is only sufficient to ensure that IXO would have enough support for technology development to qualify for a new start, a recommendation that would likely defer IXO to the 2020 decadal survey.

7

Report of the Panel on Optical and Infrared Astronomy from the Ground

SUMMARY

The celebration of the 400th anniversary of Galileo's first use of an astronomical telescope provides a fitting context for planning new goals and directions for ground-based optical and infrared (OIR) astronomy in the 21st century. The revolutionary improvement over the unaided eye that Galileo's telescope provided in angular resolution and sensitivity began a transformation and expansion of our knowledge of the universe that continues to this day. The OIR ground-based projects and activities recommended for the 2010-2020 decade are the next step that will open up unprecedented capabilities and opportunities, ranging from discovery in our solar system and the realms of exoplanets and black holes to understanding of the earliest objects in the universe and the foundations of the cosmos itself.

The vital science carried out by optical and infrared telescopes on Earth is at the core of the challenging astrophysics program laid out by the Astro2010 Science Frontiers Panels (SFPs). With the federal support recommended in this report by the Program Prioritization Panel on Optical and Infrared Astronomy from the Ground (the OIR Panel) for the construction of a Giant Segmented Mirror Telescope (GSMT), the Large Synoptic Survey Telescope (LSST), the development of ever-more-capable and technically advanced instrumentation, and renewed strategic stewardship of the nation's suite of telescopes, the United States will maintain a leading role in the pursuit of science that probes to the farthest corners of the known universe. With the generation of extremely large and rich data sets, the system of telescopes and facilities envisioned will continue the transformation of astrophysical research. They will build on the success of programs identified by previous decadal surveys and lay the foundations for astronomical research far beyond 2020 by supporting the next generation of telescopes for which the astronomical community has been planning and preparing over the past two decades.

This panel recommends new programs to optimize science opportunities across astronomy and astrophysics in ways that will support work at all scales: from the inspired individual to teams of hundreds of astronomers and billion-dollar projects. These recommendations combine to reinvigorate the U.S. system of OIR telescopes and facilities, heralding a new, expanded era of federal and nonfederal partnership for astronomical exploration.[1]

The Astro2010 survey occurs at a time of great challenge and great opportunity for OIR astronomy in the United States, which has led the world for the past century. In addition to the technical and intellectual challenges of OIR research, Europe, through its European Southern Observatory, is achieving parity with the United States in telescopes with apertures greater than or equal to 6 m and is poised to take a leading position with its plans for a 42-m Extremely Large Telescope project. The opportunity exists for U.S. OIR astronomy to marshal and coordinate its great resources and creativity and build on its successes and accomplishments to answer the fundamental questions posed by Astro2010.

Large Projects

The frontiers of astronomy and astrophysics have been advanced over the course of the 20th century, starting with the Mount Wilson 60-inch (1.5-m) telescope in 1908, by each decade's suite of ever-more-capable OIR telescopes and instruments. Continuing into the 21st century, the science opportunities in the coming decade promise to be equally great, as the OIR community stands ready to build the next generation of facilities.

A GSMT, with a collecting area exceeding 100 times that of the Hubble Space Telescope and with a 10-times-better angular resolution, will open up discovery space in remarkable new directions, probe dense environments within the Milky Way and in nearby galaxies, and—coupled with advanced adaptive optics (AO)—will map planetary systems around nearby stars. A GSMT's capabilities for astrometry will offer an unparalleled ability to probe the kinematics of galaxies, stars, and planets at the very highest angular resolution, offering sensitivities that are, in some

---

[1] The previous decadal survey—*Astronomy and Astrophysics in the New Millennium* (AANM; National Academy Press, Washington, D.C., 2001)—advocated a system perspective toward the sum of all U.S. OIR facilities in order to encourage collaborations between federally funded and independent observatories so that federal funds would be leveraged by private investment. The system today is an emerging network of public and private ground-based observatories with telescopes in the 2- to 10-m-aperture range.

cases, almost 10 magnitudes better than that achievable in space in the next decade. With a suite of spectroscopic and imaging instrumentation covering the optical and near-infrared (IR) bands, GSMT will be crucial for detailed follow-up investigations of discoveries from existing and planned facilities, including the James Webb Space Telescope (JWST) and the Atacama Large Millimeter Array (ALMA).

The promise of the next decade lies also in the capability of building a telescope to conduct systematic, repeated surveys of the entire available sky to depths unobtainable before now. Combining repeated survey images will provide composite wide-field images extending more than 10-fold fainter. Readily available synoptic data will revolutionize investigations of transient phenomena, directly addressing the key discovery area of time-domain astronomy, as well as being invaluable in surveys of regular and irregular variable sources, both galactic and extragalactic. At the same time, the combined images will provide a multi-waveband, homogeneous, wide-field imaging data set of unparalleled sensitivity that can be used to address a wide range of high-impact scientific issues. As the 48-inch Schmidt Telescope did with respect to the 200-inch Palomar Observatory, so also will LSST play a fundamental role in detecting the most fascinating astronomical targets for follow-up observations with GSMT.

Having considered proposals from the research community for new large facilities, the panel's conclusions with respect to large projects are as follows:

- The science cases for a 25- to 30-m Giant Segmented Mirror Telescope and for the proposed Large Synoptic Survey Telescope are even stronger today than they were a decade ago.
- Based on the relative overall scientific merits of GSMT and LSST, the panel ranks GSMT higher scientifically than LSST, given the sensitivity and resolution of GSMT.
- Both GSMT and LSST are technologically ready to enter their construction phases in the first half of the 2010-2020 decade.
- The LSST project is in an advanced state and ready for immediate entry into the National Science Foundation's (NSF's) Major Research Equipment and Facilities Construction (MREFC) line for the support of construction. In addition, the role of the Department of Energy (DOE) in the fabrication of the LSST camera system is well defined and ready for adoption.
- LSST has complementary strengths in areal coverage and temporal sensitivity, with its own distinct discovery potential. Indeed, GSMT is unlikely to achieve its full scientific potential without the synoptic surveys of LSST. Consequently, LSST plays a crucial role in the panel's overall strategy.
- GSMT is a versatile observatory that will push back today's limits in imaging and spectroscopy to open up new possibilities for the most important scientific problems identified in the Astro2010 survey. This exceptionally broad and powerful

ability over the whole range of astrophysical frontiers is the compelling argument for building GSMT.

- Given the development schedules for GSMT, and in order to ensure the best science return for the U.S. public investment, it is both vital and urgent that NSF identify one U.S. project for continued support to prepare for its entry into the MREFC process.

Based on these conclusions, the panel recommends the following ordered priorities for the implementation of the major initiatives that form part of the research program in optical and infrared astronomy from the ground for the 2010-2020 decade:

1. Given the panel's top ranking of the Giant Segmented Mirror Telescope based on its scientific merit, the panel recommends that the National Science Foundation establish a process to select which one of the two U.S.-led GSMT concepts it will continue to support in preparation for entering the GSMT as soon as practicable into the MREFC line. This selection process should be completed within 1 year from the release of this panel report.
2. The panel recommends that NSF and DOE commit as soon as possible but no later than 1 year from the release of this report, to supporting the construction of the Large Synoptic Survey Telescope. Because it will be several years before either GSMT project could reach the stage in the MREFC process that LSST occupies today, the panel recommends that LSST should precede GSMT into the MREFC approval process. The LSST construction should start no later than 2014 in order to maintain the project's momentum, capture existing expertise, and provide critical synergy with GSMT.
3. The panel recommends that NSF, following completion of the necessary reviews, should commit to supporting the construction of its selected GSMT through the MREFC line at an equivalent of a 25 percent share of the total construction cost, thereby securing a significant public partnership role in one of the GSMT projects.
4. The panel recommends that in the longer term NSF should pursue the ultimate goal of a 50 percent public interest in GSMT capability, as articulated in the 2001 decadal survey (*Astronomy and Astrophysics in the New Millennium*). Reaching this goal will require (most likely in the decade 2021-2030) supporting one or both of the U.S.-led GSMT projects at a cost equivalent to an additional 25 percent GSMT interest for the federal government. The panel does not prescribe whether NSF's long-term investment should be made through shared operations costs or through instrument development. Neither does the panel prescribe whether the additional investment should be made in the selected MREFC-supported GSMT in which a 25 percent partnership role is proposed already for the federal government.

But the panel does recommend that, in the long run, additional support should be provided with the goal of attaining telescope access for the U.S. community corresponding to total public access to 50 percent of the equivalent of a GSMT.

Medium Projects and Activities

In assembling its prioritized program, the panel became convinced of the strategic importance of the entire national OIR enterprise, including all facilities—public and private. The panel crafted its program to maximize the scientific return for the entire U.S. astronomical community and to maintain a leading role for OIR astronomy on the global stage.

- The panel recommends as its highest-priority medium activity a new medium-scale instrumentation program in NSF's Division of Astronomical Sciences (AST) that supports projects with costs between those of standard grant funding and those for the MREFC line. To foster a balanced set of resources for the astronomical community, this program should be open to proposals to build (1) instruments for existing telescopes and (2) new telescopes across all ground-based astronomical activities, including solar astronomy and radio astronomy. The program should be designed and executed within the context of, and to maximize the achievement of science priorities of, the ground-based OIR system. Proposals to the medium-scale instrumentation program should be peer reviewed. OIR examples of activities that could be proposed for the program include massively multiplexed optical/near-IR spectrographs, adaptive optics systems for existing telescopes, and solar initiatives following on from the Advanced Technology Solar Telescope. The panel recommends funding this program at a level of approximately $20 million annually.
- As its second-highest-priority medium activity, the panel recommends enhancing the support of the OIR system of telescopes by (1) increasing the funds for the Telescope System Instrumentation Program (TSIP) and (2) adding support for the small-aperture telescopes into a combined effort that will advance the capabilities and science priorities of the U.S. ground-based OIR system. The OIR system includes telescopes with apertures of all sizes, whereas the TSIP was established to address the needs of large telescopes. The panel recommends an increase in the TSIP budget to approximately $8 million (FY2009) annually. Additional funding for small-aperture telescopes in support of the recommendations of the National Optical Astronomy Observatory (NOAO) Renewing Small Telescopes for Astronomical Research (ReSTAR) committee (approximately $3 million per year) should augment the combined effort to a total of approximately $11 million (FY2009) to encompass all apertures. The combined effort will serve as a mechanism for coordinating the development of the OIR system. To be effective, the funding level

and funding opportunities for this effort must be consistent from year to year. Although it is possible that the total combined resources could be administered as a single program, the implementation of such a program raises difficult issues, such as formulas for the value of resources or the need to rebuild infrastructure. The panel considers the administration of two separate programs under the umbrella of "system development" to be a simpler alternative. The expanded TSIP and the mid-scale instrumentation program both provide opportunities to direct these instrumentation funds strategically toward optimizing and balancing the U.S. telescope system.

- The U.S. system of OIR telescopes currently functions as a *collection* of federal and nonfederal telescope resources that would benefit from collaborative planning and management—for example, to avoid unnecessary instrument duplication between telescopes. The panel recommends that NSF ensure that such a mechanism exists, operating in close concert with the nonfederal observatories, for the management of the U.S. telescope system. The panel recommends that a high priority be given to renewing the system of ground-based OIR facilities, requiring a new strategic plan and a broadly accepted process for its implementation.

Small Programs

The panel concluded that initiating a tactical set of small targeted programs (each between $1 million and $3 million per year) would greatly benefit ground-based OIR science in the coming decade and would provide critical support for some of the medium and large programs. The panel recommends the small programs in the following, unprioritized list:

- An adaptive optics technology-development program at the $2 million to $3 million per year level;
- An interferometry operations and development program at a level of approximately $3 million per year;
- An integrated ground-based-astronomy data-archiving program starting at a level of approximately $2 million per year and ramping down to approximately $1 million per year; and
- A "strategic theory" program at the level of approximately $3 million per year.

Recommendations for Adjustments to Continuing Activities

The panel makes the following recommendations for continuing activities:

- NSF should continue to support the National Solar Observatory (NSO) over the 2010-2020 decade to ensure that the Advanced Technology Solar Telescope

(ATST) becomes fully operational. ATST operations will require a ramp-up in NSO support to supplement savings that accrue from the planned closing of current solar facilities.

- Funding for NOAO facilities should continue at approximately the FY2010 level.
- The governance of the international Gemini Observatory should be restructured, in collaboration with all partners, to improve the responsiveness and accountability of the observatory to the goals and concerns of all its national user communities. As part of the restructuring negotiations, the United States should attempt to secure an additional fraction of the Gemini Observatory, including a proportional increase in the U.S. leadership role. The funding allocated for any augmentation in the U.S. share should be at most 10 percent of FY2010 U.S. Gemini spending. The United States should also seek improvements to the efficiency of Gemini operations. Efficiencies from streamlining Gemini operations, possibly achieved through a reforming of the national observatory to include NOAO and Gemini under a single operations team, should be applied to compensate for the loss of the United Kingdom from the Gemini partnership, thereby increasing the U.S. share. The United States should support the development of medium-scale, general-purpose Gemini instrumentation and upgrades at a steady level of about 10 percent of the U.S. share of operations costs. U.S. support for new large Gemini instruments (greater than approximately $20 million) should be competed against proposals for other instruments in the recommended mid-scale instrumentation program—a program aimed at meeting the needs of the overall U.S. OIR system discussed elsewhere in this panel report.
- The NSF-AST grants program (Astronomy and Astrophysics Research Grants [AAG]) should be increased above the rate of inflation by approximately $40 million over the decade to enable the community to utilize the scientific capabilities of the new projects and enhanced OIR system.
- NSF-AST should work closely with the NSF Office of Polar Programs to explore the potential for exploiting the unique characteristics of promising Antarctic sites.

The above program and the funding recommendations, presented in additional detail in the following sections of the panel's report, represent a balanced program for U.S. OIR astronomy that is consistent with historical federal funding of astronomy and, more importantly, is poised to enable astronomers to answer the compelling science questions of the decade, as well as to open new windows of discovery. The proposed program involves an increased emphasis on partnerships, including NSF, DOE, NASA, U.S. federal institutions, state and private organizations, and international or foreign institutions. These partnerships not only are required by the scale of the new projects, which are beyond the capacity of any one institution or even one nation to undertake, but also are motivated by the key capabilities that each of the partners brings to ensuring a dynamic scientific program throughout the decade.

The revolution in human understanding that began with Galileo's telescope 400 years ago has not slowed down or lost its momentum—in fact it is accelerating—and the panel believes that it has identified the most promising areas for future investment by the United States in optical and infrared astronomy from the ground.

INTRODUCTION AND CONTEXT

Ground-based optical and infrared astronomy provides a fundamental basis for our knowledge of the universe at almost all astronomical scales. Moreover, OIR observations and facilities render astronomy accessible and inspirational to the general public. In the last decade, the development of new technologies has expanded our capabilities in many ways. In the time domain we are embarking on multi-epoch sky surveys, while in terms of spatial resolution ground-based telescopes are obtaining diffraction-limited images that surpass the angular resolution of current space telescopes.

The previous (2001) decadal survey, *Astronomy and Astrophysics in the New Millennium* (AANM), recommended two large activities and two medium activities for OIR: a giant (30-m-class) segmented-mirror, adaptive-optics-equipped, ground-based optical-infrared telescope, now known simply as the Giant Segmented Mirror Telescope (GSMT), and a large-aperture (6.5-m-class), very-wide-field (~3-deg) synoptic survey telescope to achieve an unprecedented combination of sky coverage, faint limiting magnitude, and time-domain coverage. The medium activities were an Advanced Technology Solar Telescope (ATST, called AST in AANM), support for developing the concept of treating the federally supported and independent observatories in the United States as a system, and using the system concept to increase the instrumentation capabilities and community access to U.S. OIR telescopes through a program called the Telescope System Instrumentation Program.

Progress has been made on all four initiatives. ATST has just entered its MREFC-funded construction phase. TSIP has operated at a low but significant level for almost the entire decade. Projects have been formed and developed as candidates for GSMT and the large-aperture, very-wide-field synoptic survey telescope.

The new observational capabilities introduced in the last decade fostered an exciting period of scientific discovery for OIR, fundamentally changing the way the contents and history of the universe are understood and capturing the imagination and interest of the general public. Notable examples include

- *Exoplanets.* The discovery of a diverse set of extrasolar planets has defied many preconceived notions of the properties of other solar systems. Planet hunt-ing is now pursued with an increasingly rich variety of survey techniques (each of which can probe different kinds of systems), including increasingly sensitive radial velocities, transit timing, microlensing, and adaptive optics imaging (Figure 7.1).

- *Dark energy and structure formation.* The discovery of the acceleration of the expansion rate of the universe has profoundly altered our view of fundamental physics. Rapidly improving measurements of the acceleration, along with a detailed view of the large-scale structure of the universe, have established the lambda cold dark matter (ΛCDM) model as the standard model of cosmology.
- *Galactic-center black hole.* Definitive proof for the existence of a supermassive black hole and the first detailed kinematic look at the way black holes interact with their stellar environments have been obtained through measurements of individual stellar orbits at the galactic center (Figure 7.2)
- *Gamma-ray bursts.* The study of optical-IR afterglows of gamma-ray bursts (GRBs) has provided detailed light curves and redshifts of these events. This in turn has generated an understanding of their connection with supernovae and the birth of black holes, extending as far back as $z = 8.2$.
- *Milky Way satellite and streams.* The discovery of new components, remnants, and companions to the Milky Way has significantly altered our picture of halo formation, the early stages of galaxy formation, and the importance of mergers (Figure 7.3).
- *Quasars and GRBs at first light.* Quasars, powered by early supermassive black holes, have been discovered back to redshifts of 6. Observations of a $z = 8.2$ GRB have opened a new window for studying the deaths of massive stars when the universe was only 4 percent of its current age. The first detection of the effect of hydrogen absorption on quasar light from $z = 6.4$ now provides the signature of the last phases of the reionization of the universe (Figure 7.4).
- *Galaxies and massive black holes across cosmic time.* The discovery and characterization of a tight (and unexpected) correlation between the mass of a supermassive black hole and the velocity dispersion of the host galaxy's bulge has driven new ideas about the evolution of these objects. Determinations of the histories of cosmic star formation, chemical enrichment, and massive black hole accretion have provided additional input, supplemented by the discovery of a bimodal color-magnitude distribution in the galaxy population at the present epoch.
- *Brown dwarfs.* Sky surveys have revealed hundreds of field brown dwarfs, objects that have a direct connection to the more difficult study of giant planets. This large sample has enabled the application of theoretical atmospheric models yielding a thorough understanding of their structure and composition, direct measurements of dynamical masses, and definitive elimination of brown dwarfs as dark matter candidates.
- *Kuiper belt objects.* The discovery of objects within the solar system that are comparable to, or more massive than, Pluto has revolutionized our understanding of the constituents of the solar system.

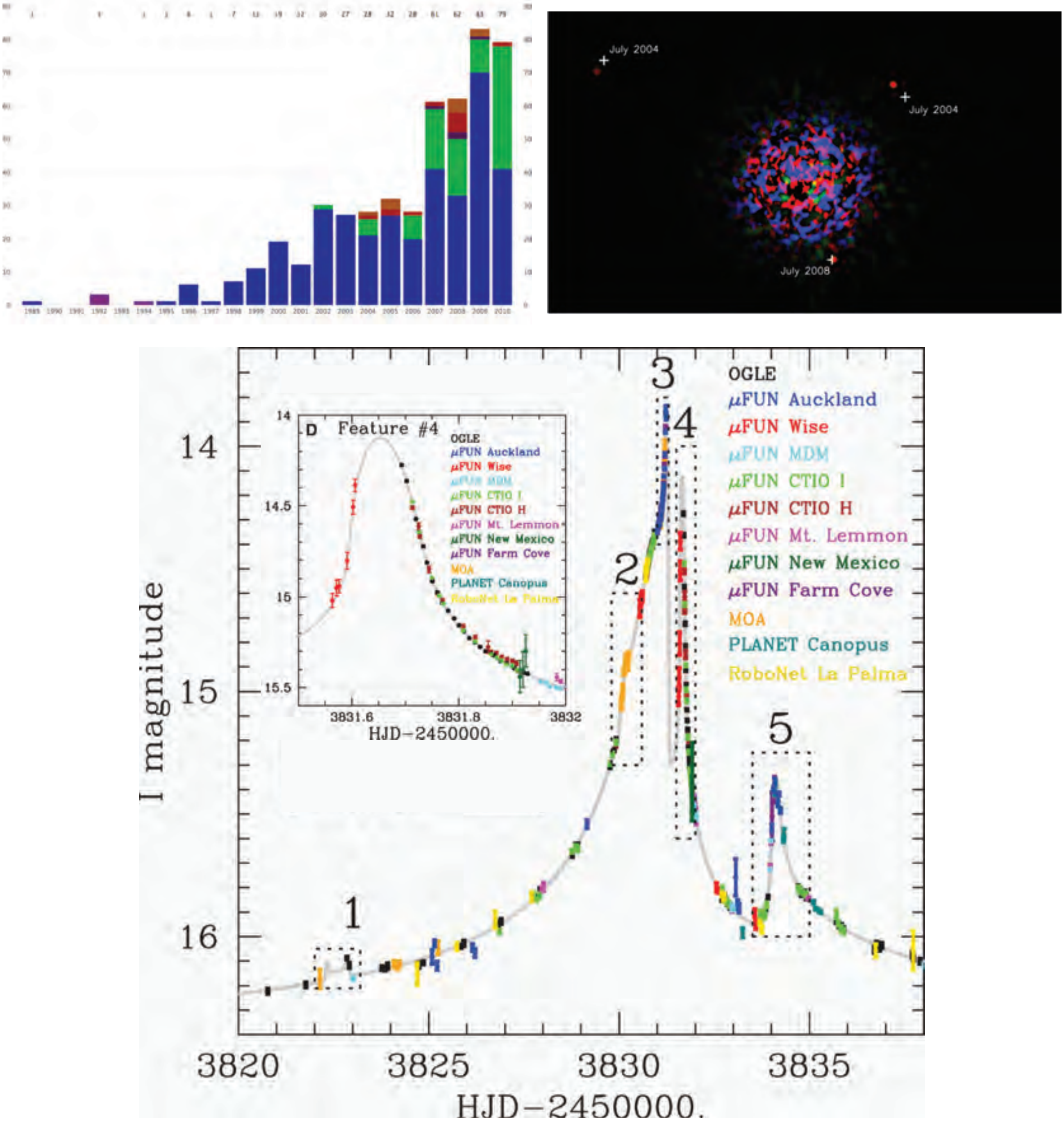


FIGURE 7.1 *Top left:* Exoplanet discoveries by year, color coded by discovery technique: radial velocity (blue), transit (green), timing (dark purple), astrometry (dark yellow), direct imaging (red), microlensing (orange), and pulsar timing (light purple). *Top right:* The first directly imaged multiple planet system (HR 8799; adaptive optics imaging in 2008). *Bottom:* Planetary system discovered by microlensing. SOURCE: *Top left:* Available at http://commons.wikimedia.org/wiki/File:Exoplanet_Discovery_ Methods_Bar.png#filehistory (19:35; October 3, 2010). *Top right:* National Research Council of Canada—Herzberg Institute of Astrophysics, C. Marois and Keck Observatory. *Bottom:* B.S. Gaudi, D.P. Bennett, A. Udalski, A. Gould, G.W. Christie, D. Maoz, S. Dong, J. McCormick, M.K. Szymaski, P.J. Tristram, S. Nikolaev, et al., Discovery of a Jupiter/Saturn analog with gravitational microlensing, *Science* 319(5865):927-930, 2008, reprinted with permission of AAAS.

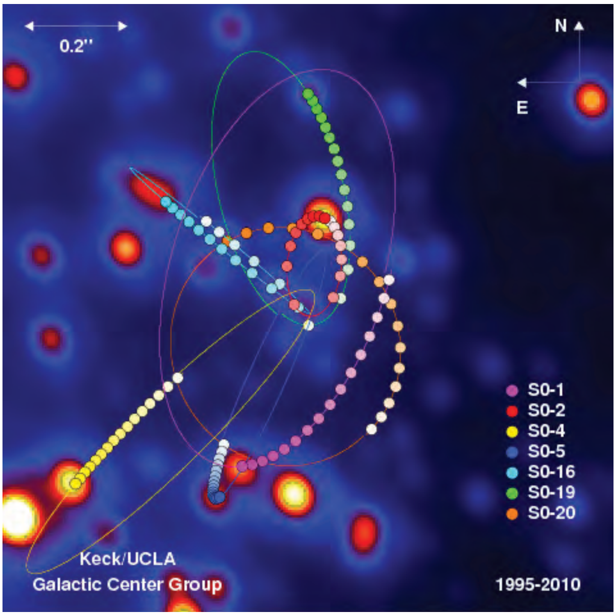


FIGURE 7.2 Stellar orbits at the galactic center that have demonstrated the existence of a supermassive black hole and revealed the kinematic structure of surrounding stellar population, which is key to understanding the growth of black holes. SOURCE: Image courtesy of and created by Andrea Ghez and her research team at UCLA from data sets obtained with the W.M. Keck Telescopes.

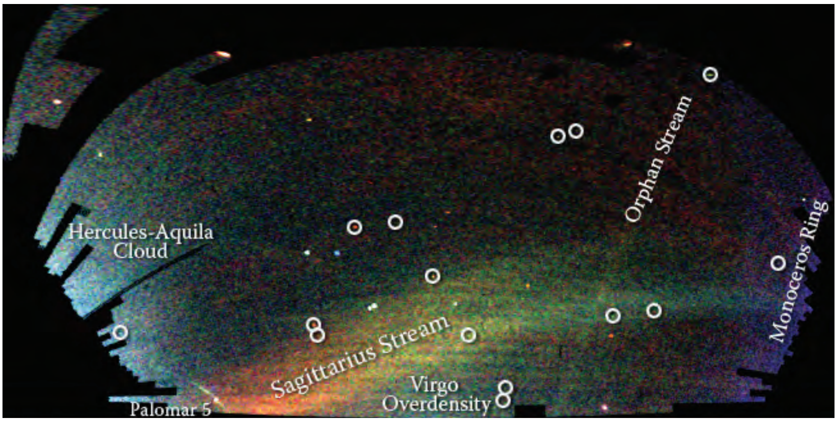


FIGURE 7.3 Discovery of Milky Way streams. SOURCE: V. Belokurov and the Sloan Digital Sky Survey.

- *Inside the Sun.* Helioseismology revealed unanticipated temperature and velocity structures just beneath sunspots, measured interior flows that constrain solar-dynamo action throughout the convection zone, and made routine the detection of active regions on the far side of the Sun.

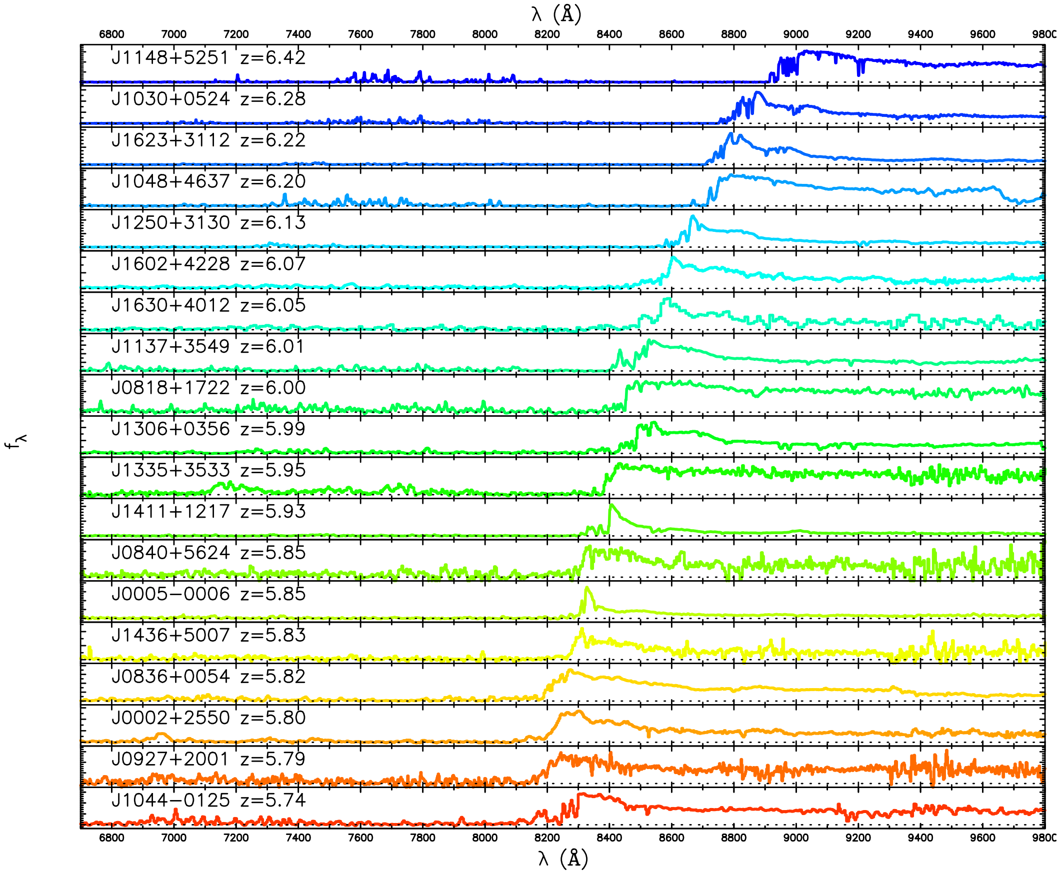


FIGURE 7.4 Stack of quasar spectra showing Gunn-Peterson absorption trough. SOURCE: X. Fan, C.L. Carilli, and B. Keating, Observational constraints on cosmic reionization, *Annual Review of Astronomy and Astrophysics* 44:415-62, 2006.

- *A complex stellar atmosphere.* An incessant flurry of mixed-polarity magnetic elements appears everywhere on the solar surface and above, down to the smallest resolvable scales. Energetic features and flows now observable in this transitional regime are being matched by increasingly realistic models of turbulent magneto-convection near the photosphere and in the exceedingly heterogeneous atmosphere (Figure 7.5).

These scientific achievements and many others have been enabled by the development of astronomical technologies. Two that stand out as particularly important are (1) the evolution of adaptive optics systems into scientifically ver-

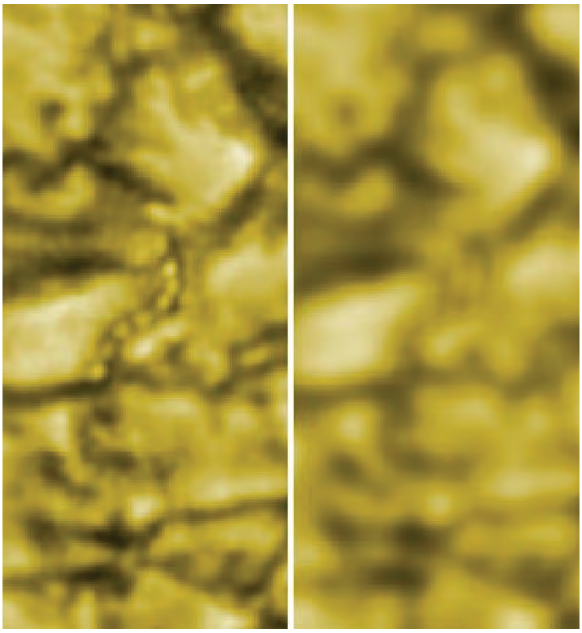


FIGURE 7.5 This high-resolution solar image from the 1.6-m New Solar Telescope (NST) at Big Bear Solar Observatory demonstrates the advantage of a large aperture. Even without adaptive optics, this 4″ x 2″ speckle-reconstructed image from March 30, 2009, reveals seven side-by-side bright points in a dark lane at the center. Bright points are probably associated with magnetic field concentrations and are each about 0.1″ (75 km) in diameter. The beads cannot be seen in the image on the right that was degraded to the resolution of the old 0.6-m telescope. ATST, scheduled for first light in 2017, will have a 4-m aperture. SOURCE: Big Bear Solar Observatory/ New Jersey Institute of Technology.

satile and reliable user instrumentation and (2) the development of the hardware and software infrastructure to produce, distribute, and analyze large and reliable astronomical surveys.

The development of adaptive optics (AO) into a workhorse technology on most large ground-based telescopes (e.g, the Keck, VLT, Gemini, MMT, the Hale telescope, and the Dunn solar telescopes) has enabled diffraction-limited, near-IR imaging and spectroscopy previously possible only from space. Key technology developments include the production of stable lasers capable of generating a bright artificial guide star (Figure 7.6), large-format deformable mirrors, lower-noise wavefront sensors that allow fainter tip-tilt stars and natural guide stars to be detected, and the engineering of these components into reliable, optimized, and rugged systems that meet scientific needs. For example, the laser-guide-star AO system on the 10-m W.M. Keck II telescope routinely achieves an angular resolution of 0.04 arcsecond (with Strehl ratio >0.4 at 2.2 microns) in its laser-guide-star mode and can achieve this performance over a large fraction of the sky. In 2009, more than 150 refereed scientific papers were published based on AO data, 29 using laser-guide-star systems. Other critical developments important for both current and future AO systems include adaptive secondaries and three-dimensional tomographic analysis of atmospheric turbulence. Areas of scientific research benefiting from AO systems now range from studies within our own solar system (including

FIGURE 7.6 Three laser adaptive optics systems operating on Mauna Kea. SOURCE: © Subaru Telescope, National Astronomical Observatory of Japan. All rights reserved. Reprinted with permission.

the Sun itself) to the most distant galaxies and include a number of truly transformative research results over the last decade (e.g., see Figures 7.1 and 7.2). The pace of technological development and scientific utilization of AO systems has occurred quite rapidly since the 2001 decadal survey, and astronomers now have powerful systems as part of the basic facility infrastructure on many telescopes. New, more specialized and powerful systems will soon be operational, including dedicated planet-finding, high-contrast imagers, wide-field multi-conjugate AO systems, and the powerful facility designed for the ATST.

Survey astronomy also has undergone a transformation in the past decade, with a profound impact not only on the science, but also on the culture of astronomy. Large, well-reduced, statistically rigorous data sets are now standard tools in most topics of OIR astronomy. These data sets are enabled by powerful instruments and a renewed commitment to data pipelines and archives. A key aspect of survey astronomy is that many science interests can be served simultaneously, including serendipitous discoveries and unanticipated opportunities. Surveys have the greatest impact when the full reduced data are made public, greatly increasing the com-

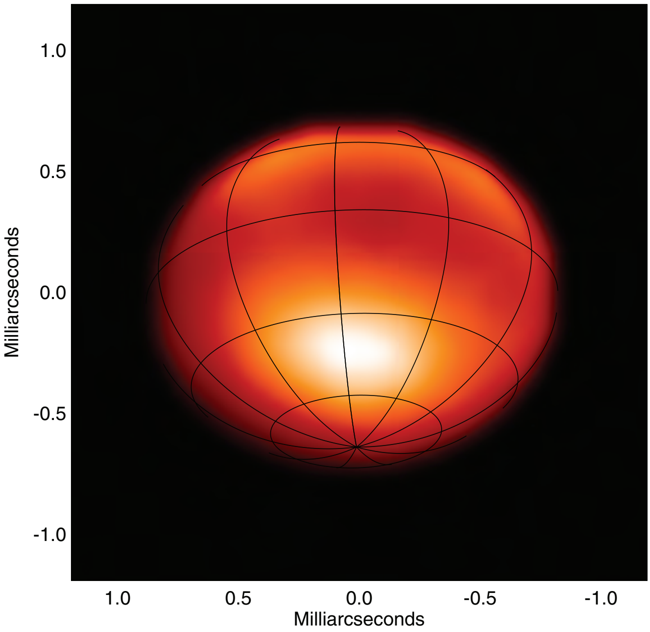


FIGURE 7.7 Imaging at milliarcsecond resolution is now becoming routine with today's infrared interferometers. This recent image of Alderamin (α Cep) from the CHARA Array reveals the centrifugally distorted photosphere of this rapidly rotating star, strong "gravity darkening" along the equator, and hot emission at the poles. SOURCE: M. Zhao, J.D. Monnier, E. Pedretti, N. Thureau, A. Mérand, T. Ten Brummelaar, H. McAlister, S.T. Ridgway, N. Turner, J. Sturmann, L. Sturmann, P.J. Goldfinger, and C. Farrington, Imaging and modeling rapidly rotating stars: α Cephei and α Ophiuchi, *Astrophysical Journal* 701:209, 2009, reproduced by permission of the AAS.

munity of people who can use the data. An increasing number of astronomers rely on data archives for their work, and new tools and protocols have been developed to serve, search, and cross-link these data sets and to find complex trends in the results.

Besides these two areas of major development, advances in many other technologies have provided significant improvement in the capabilities of ground-based OIR telescopes. These technologies include new generations of large-format optical and IR detector arrays; interferometry, with pairs of large telescopes and arrays of smaller ones (Figure 7.7); new, durable, low-emissivity mirror coatings; volume-phase holographic gratings; and AO-fed coronagraphs.

The structure of the astronomical enterprise has served the pursuit of groundbreaking science well. It will need to evolve further in the decades to come in order to make the most efficient use of, and provide the broadest access to, observational capabilities and thereby continue to advance the frontiers of our knowledge and answer the urgent science questions identified in this Astro2010 decadal survey of astronomy and astrophysics.

OPPORTUNITIES IN OIR SCIENCE

The Astro2010 Science Frontiers Panels have identified key research questions and discovery areas for the next decade. Ground-based OIR astronomy laid the foundations in many of these areas and is poised to enable transformative studies to address these questions over the next decade (Table 7.1). Larger telescopes—with greater light grasp, AO-enabled resolution, and multiplex instruments—will tackle high-contrast investigations of objects ranging from the Sun to faint sources in crowded fields in nearby galaxies and at cosmological distances. Large-scale, high-sensitivity imaging surveys at high cadence will open the way for a thorough exploration of time-domain astronomy. New facilities, coupled with upgraded existing public and private resources, will enable astronomers to seek answers to some of the key scientific questions of our era—from the nature of dark matter and dark energy to the formation and evolution of galaxies, stars, and planets.

Determining the properties of exoplanetary systems and their disk progenitors remains a prime science objective of ground-based OIR astronomy (PSF 3 and PSF 5,[2] Table 7.1; see Figure 7.8). The first insights are being gained into the structure of exoplanet systems with increased observational statistics and improved radial-velocity precision. Systematic campaigns with new high-precision radial-velocity spectrometers, particularly at near-IR wavelengths, can extend coverage to longer-period and lower-mass planets around later-type stars, approaching the level needed to detect Earth-like planets under the best conditions. Besides radial-velocity surveys, detection techniques now include transit surveys, gravitational microlensing surveys and direct imaging. Deepening the census of planets is a crucial step toward understanding formation processes. Transit techniques that combine discovery from the ground with space-based characterization and direct imaging and spectroscopy with advanced AO coronagraphs will let us measure the temperature and composition of the planets found. These new abilities have just begun to be exploited, and our knowledge of the physical properties of exoplanets will explode in the coming decade. Realizing progress over the next decade will require dedicated planet-detecting instruments (from IR Doppler spectrographs to high-contrast AO coronagraphs) on telescopes of all sizes from 2-m to GSMT-scale facilities. Ultimately, ground-based characterization and synoptic studies, complemented by space-based observations, will help us understand how solar systems form and whether systems like our own—including Earth-like planets—are common or rare.

High-contrast imaging and spectroscopy on 8-m and GSMT-class telescopes will map the structure and evolution of protoplanetary disks (PSF 2, Table 7.1) at angular separations exceeding 0.15 to 0.04 arcsecond, respectively, and contrast ratios from 10^{-7} to 10^{-9}. GSMT-class telescopes offer the potential of detecting Jupiter analogs around nearby solar-type stars, and will directly complement ALMA submillimeter surveys investigating the kinematics and dust in protoplanetary

---

[2] The questions and discovery areas developed by the five SFPs are identified throughout this report by the three-letter panel acronym plus question number and the letter "D" (for discovery area) as given in Table 7.1.

TABLE 7.1 Contributions by OIR Facilities and Activities to Addressing Key Science Questions Identified by the Astro2010 Decadal Survey Science Frontiers Panels

Science Frontiers Panel	Science Questions and Discovery Areas (D)	GSMT	LSST	Existing Facilities	Mid-Scale Instrumentation	Interferometry
Planetary Systems and Star Formation (PSF)	1. How do stars form?	▒	░	▓	▒ AO	
	2. How do circumstellar disks evolve and form planetary systems?	■	▒	▒	■ AO	■
	3. How diverse are planetary systems?	▒		▒	■ ExAO, HPRV	░
	4. Do habitable worlds exist around other stars, and can we identify the telltale signs of life on an exoplanet?					
	D. Identification and characterization of nearby habitable exoplanets	▒		░	▒ HPRV	
Stars and Stellar Evolution (SSE)	1. How do rotation and magnetic fields affect stars?	▒	▒	■	■ Solar	■
	2. What are the progenitors of Type Ia supernovae and how do they explode?	■	■	░	░ AO	▒
	3. How do the lives of massive stars end?	▒	■	░	▒ MOS, AO	▒
	4. What controls the mass, radius, and spin of compact stellar remnants?					
	D. Time-domain surveys	▒	■	▒	■ Solar	░

continued

TABLE 7.1 Continued

Science Frontiers Panel	Science Questions and Discovery Areas (D)	GSMT	LSST	Existing Facilities	Mid-Scale Instrumentation	Interferometry
Galactic Neighborhood (GAN)	1. What are the flows of matter and energy in the circumgalactic medium?	░		▨		
	2. What controls the mass-energy-chemical cycles within galaxies?	░		▨		■
	3. What is the fossil record of galaxy assembly from the first stars to the present?	■	■	■	■ MOS	
	4. What are the connections between dark and luminous matter?	■	■	▨		
	D. Astrometry[a]	▨	■	░		▨
Galaxies Across Cosmic Time (GCT)	1. How do cosmic structures form and evolve?	■	■	▨	■ MOS, DWFIR	
	2. How do baryons cycle in and out of galaxies, and what do they do while they are there?	■		■		
	3. How do black holes grow, radiate, and influence their surroundings?	■	░	▨	■ AO	
	4. What were the first objects to light up the universe, and when did they do it?	■		▨	▨ MOS	
	D. The epoch of reionization	▨	░	░	▨ DWFIR	

Cosmology and Fundamental Physics (CFP)	1. How did the universe begin?	░	■	░	▓ MOS	
	2. Why is the universe accelerating?	▓	■	■	■ MOS	
	3. What is dark matter?	▓	▓	░		
	4. What are the properties of neutrinos?		▓		▓ MOS	
	D. Gravitational wave astronomy	░	░	░	░ DWFIR	

---

[a] The GAN SFP also identified time-domain astronomy as a discovery area for the next decade.

NOTE: The gray-intensity coding depicts the level of contribution(s) made by GSMTs, LSST, currently existing (public/private) OIR telescope facilities, mid-scale instrumentation, and current interferometry capabilities to addressing the science questions and discovery (D) areas identified by the five Astro2010 Science Frontiers Panels: Black indicates strong contributions to an SFP question or discovery area, dark gray indicates an important contribution, light gray indicates a minor contribution, and no coloring indicates no or no significant contribution. In the Mid-Scale Instrumentation column, it is noted where the contribution will come from: ExAO, future coronagraphic high-contrast, adaptive optics systems, including GSMT-class instrumentation; AO, future high-Strehl/visible-light laser-guide-star adaptive optics systems on 5- to 10-m telescopes; DWFIR, Deep, Wide-Field IR Survey; HPRV, high-precision radial-velocity spectroscopy; MOS, wide-field multiobject spectroscopy; and Solar, new solar instrumentation.

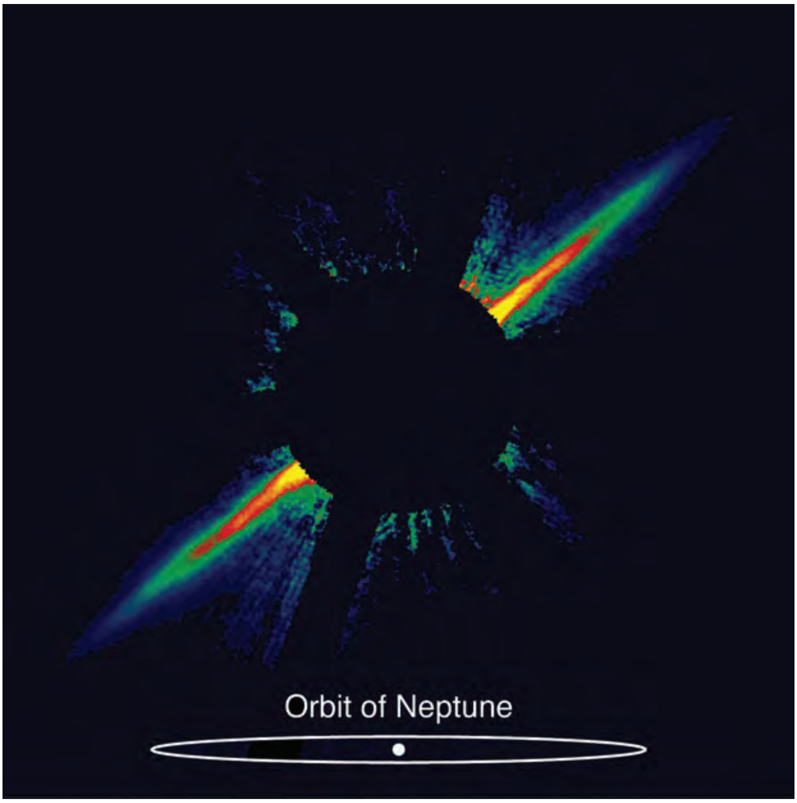


FIGURE 7.8 Keck K-band image of the protoplanetary disk around the nearby young M dwarf, AU Mic. New-generation adaptive optics systems on 8- to 10-meter-class telescopes and GSMT will extend such observations to older, more distant systems. SOURCE: M. Liu, IfA-Hawaii/W.M. Keck Observatory.

disks. Thermal mid-IR observations trace warm dust in young systems, while high-resolution spectroscopy at blue and IR wavelengths probes gas kinematics at sub-AU (astronomical unit) resolution. Optical and near-IR measurements of scattered light map colder material and remnant debris disks. AO-assisted coronagraphy and polarimetry will be powerful tools for probing disks around young stars in the solar neighborhood, deriving dust composition, radial density, and size distributions. These observations will be complemented by synoptic surveys that will provide the first census of the size-frequency and compositional distributions of trans-neptunian objects in the outermost regions of our own solar system, reaching magnitude r ~ 24.5 and trans-neptunian objects with diameters as small as 25 km.

New OIR ground-based capabilities will also open new windows on one of the most critical astrophysical processes: the formation of stars (PSF 1, Table 7.1).

Detailed observations of star clusters spanning a range of ages, coupled with wide-angle surveys of nearby field stars, have resulted in broad consensus as to the overall form of the stellar initial mass function (IMF). Three key areas of uncertainty remain: at high masses (>10 solar masses) there are indications of variations in either the shape of the IMF or the upper-mass cutoff; at the opposite extreme it remains unclear whether there is a well-defined low-mass cutoff; finally, uncertainties remain concerning binary frequency and the role played by binary systems at all masses. Tackling these issues demands high-sensitivity, high-resolution imaging. AO-equipped 8-m telescopes can detect jovian-mass brown dwarfs in the nearest star-forming regions and can resolve binaries at separations exceeding 30 mas. GSMT-class telescopes can resolve high-mass stars in a wide range of star clusters in the nearby Andromeda galaxy and in star-forming regions in the nearer irregular galaxies, setting constraints on variations in the upper IMF within a range of environments.

Stellar-evolution theory represents a triumph of 20th century astrophysics. Recent technological developments and observational innovations put us in a position to refine and test that theory in unprecedented detail over the next decade. In particular, scientists will quantify the influence of magnetic fields, mass loss, and rotation through high-resolution, high-sensitivity spectroscopy and spectropolarimetry of individual stars (SSE 1, Table 7.1), as well as through direct imaging via long-baseline OIR interferometry.

The Sun is a key target, providing scope both for continuous, comprehensive measurements of global temporal changes (SSE D, Table 7.1) and for detailed observations of physical processes operating at the finest spatial and temporal scales, where plasma and magnetic fields interact. The 4-m ATST will reveal the Sun's magnetic field at sub-arcsecond resolution and probe the complex transition region between the photosphere and the corona. Solar observations are key for three topics of broad astrophysical interest: (1) the origin and evolution of large- and small-scale magnetic flux elements; (2) the formation of chromospheres and coronae and the physics of energy transport in inhomogeneous, non-equilibrium atmospheres; and (3) energy storage and release in the dynamic corona, which lead to a range of events including flares, coronal mass ejections, and particle acceleration. Activity in the third has a significant impact on our understanding of Earth-Sun interactions, including space weather.

Within the galaxy, detailed surface-magnetic-field geometries can be reconstructed for active stars using (Zeeman) spectropolarimetry and Doppler tomography on high-throughput echelle spectrographs on large-aperture telescopes. The surfaces of nearby stars can be directly imaged using long-baseline infrared interferometry, probing photospheric distortions caused by rapid rotation or binary interactions. In the wider field, repeated full-sky monitoring by synoptic surveys promises a complete statistical view of stellar cycles and rare stellar events, including flares that can tell us about magnetic-field properties for stars and brown dwarfs spanning the full range of masses. Maturing asteroseismology will provide insight into stars' internal structure and dynamics.

Supernovae and gamma-ray bursts are two examples of stellar behavior whose characterization has implications that extend well beyond evolutionary theory (PSF 1; SSE 2, 3, and D; GAN 1, 2; see Figure 7.9). In neither case is there a good understanding of the progenitor population or of the full range of evolutionary pathways. Insight into both will be gained from studying the dramatic mass-loss and strong binary interactions seen in massive stars in the Milky Way, its satellites, and nearby galaxies, coupling AO-based spectroscopic and photometric analyses with interferometric data. Large-scale synoptic surveys will identify numerous supernovae and GRB afterglows, and detailed follow-up observations of well-chosen subsets will pinpoint their precise locations, providing either direct identification of the progenitor or insights gained through investigating the properties of its immediate stellar neighbors. Similar resources can also be deployed to identify and study the nature of sources of gravitational waves (CFP D). Investigations at different cosmic epochs will be supplemented by "galactic archeology," using today's abundance patterns of metal-poor stars—identified from multiobject spectroscopic follow-up of wide-field imaging surveys—to probe supernova frequencies and the enrichment history of the early Milky Way.

On larger scales, basic questions center on feedback mechanisms and energy transport in the interstellar medium and the recycling of circumgalactic gas in galaxies (GAN 1, 2; GCT 2). At high redshifts, OIR observations in the rest-frame ultraviolet detect winds from Lyman-break galaxies and directly measure the resulting spread of metals; in the lower-redshift universe, interstellar optical and near-ultraviolet lines also allow the tracking of galactic winds as a function of galaxy properties.

The star formation histories of galaxies, the growth of central black holes, the properties of dark matter halos, and the way interactions between those quantities are reflected in present-day galactic morphology are all crucial areas for research in the next decade (GAN 3, 4, and D; GCT 1, 3). In recent years, precision measurements across cosmic time have led to a widely accepted cosmological paradigm for galaxy assembly and evolution, the ΛCDM model (see Figure 7.10). Within this theory, galaxies form "bottom-up," with low-mass objects ("halos") collapsing earlier and merging to form larger and larger systems over time. Ordinary matter follows the dynamics dictated by the dominant dark matter until radiative, hydrodynamic, and star-formation processes take over (GCT 2). Although ΛCDM has had great success in explaining the observed large-scale distribution of mass in the universe, the nature of the dark matter particle is best tested on small scales, where its interaction properties manifest themselves by modifying the structure of galaxy halos and their clumpiness (CFP 4). On these scales, detailed comparisons

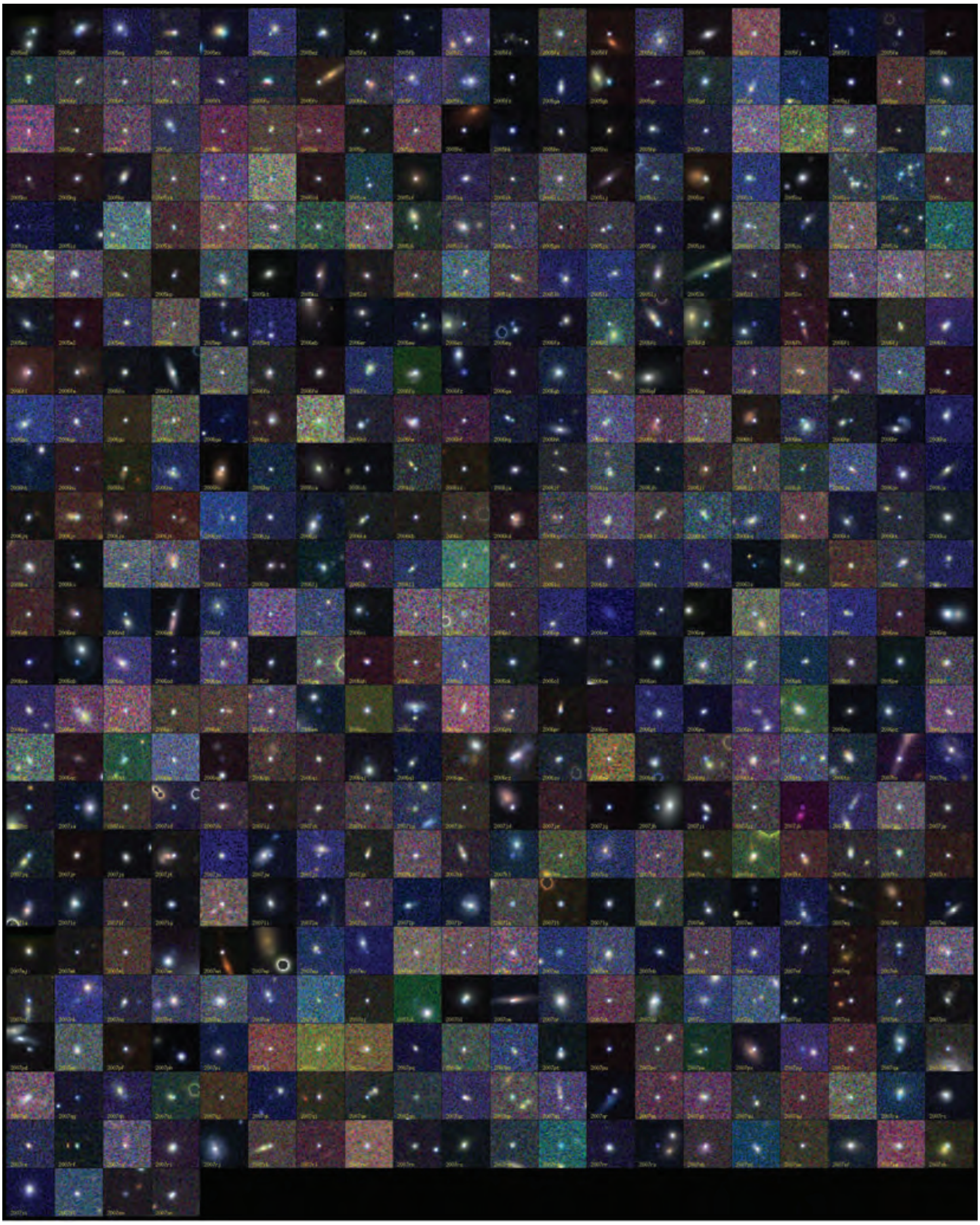


FIGURE 7.9 A menagerie of supernovae from the SDSS survey. The LSST program will identify tens of thousands of such objects. SOURCE: Ben Dilday and the Sloan Digital Sky Survey (SDSS) Collaboration.

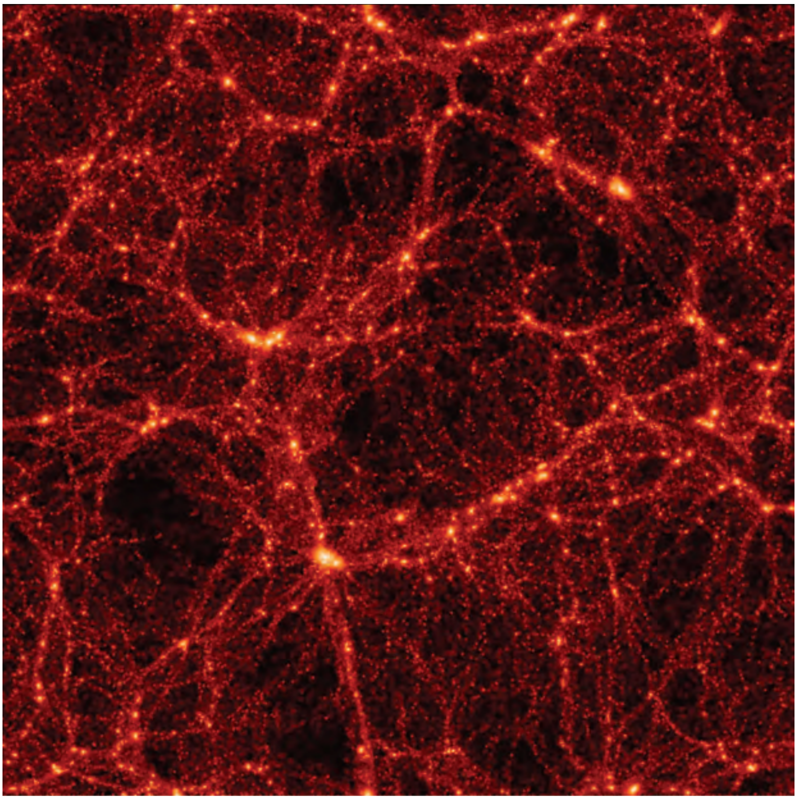


FIGURE 7.10 Simulations of the ΛCDM cosmic web. SOURCE: A. Jenkins, C.S. Frenk, F.R. Pearce, P.A. Thomas, J.M. Colberg, S.D.M. White, H.M.P. Couchman, J.A. Peacock, G. Efstathiou, and A.H. Nelson, Evolution of structure in cold dark matter universes, *Astrophysical Journal* 499:20-40, 1998. Courtesy of Joerg Colberg and the Virgo Consortium.

between observation and theory reveal several discrepancies, such as the apparent mismatch between the galaxy's relatively smooth stellar halo and the extremely clumpy dark matter distribution predicted by numerical simulations (the "missing satellite problem"). Ground-based OIR observations will discover fainter and more distant satellites and streams, confirming their existence, measuring velocity dispersions and dark matter properties, and dramatically constraining the formation of structure in ΛCDM.

Optical and near-infrared color-magnitude diagrams remain a highly effective means of probing the star formation history in nearby galaxies, and the next generation of telescopes will extend observations to more crowded environments and greater distances (PSF 1, GAN 2). Those investigations complement the early star-formation history inferred from galactic archaeology (GAN 3), and a GSMT-class telescope offers the potential of resolving horizontal-branch stars in the nearest ellipticals. High-angular-resolution imaging and spectroscopy can focus on galactic cores (including the galactic center); measure stellar and gaseous motions to 1-kms^{-1} accuracy, more than an order of magnitude better than current measurements; and probe the detailed structure of the stellar cusps and the mass of nuclear black holes. At the same time, spectroscopic and astrometric measurements (with 30-μas accuracy) of stars in the Milky Way's outer halo and its satellites can measure space motions at accuracies better than 5 kms^{-1} and map the distribution of dark matter on larger scales (GAN 4, 5).

Large-scale spectroscopic surveys using multiobject instruments on 4- to 10-meter-class telescopes will be vital in probing the growth of structure at low and intermediate redshifts and tracing the demographics of black holes and active galaxies (GCT 1, 3). Follow-up observations on a GSMT-class telescope can use emission lines to investigate the chemical enrichment of distant galaxies at the peak of their star formation activity (redshift $z \sim$ 2-4). The light-gathering power of a GSMT at low-background, near-IR wavelengths is such that its AO-assisted spectroscopic performance at moderate resolution (R = 100-1,000) is 10 to 100 times better than JWST. AO-fed integral-field-unit spectrographs will reveal the internal structures, kinematics, and metallicities of those systems. Absorption-line studies using background probes will enable a detailed understanding of the topology, ionization state, and chemical enrichment of the intergalactic medium (GCT 2).

The first dwarf-size galaxies will contain massive stars formed from primordial gas (Population III): these subgalactic stellar systems, aided perhaps by a population of accreting black holes in their nuclei, generated the UV radiation that reheated and reionized the universe at the end of the cosmic dark ages (GCT 4, 5). Although JWST has a clear advantage in near-IR imaging of high-redshift sources, once identified, near-IR spectroscopy on an AO-assisted GSMT-class telescope is significantly more efficient and will provide a more detailed story of the properties and influence of the first stars on the intergalactic medium.

OIR will remain in the forefront in our quest for the mapping of cosmological initial conditions over the widest possible dynamic range (CFP 1, 2, 4), through observations of large-scale structure (using galaxies, intergalactic gas, and gravitational lensing) and of standard candles (primarily Type Ia supernovae). While supernovae and baryonic acoustic oscillations provide complementary methods for measuring the distance-redshift relation, weak lensing is sensitive both to distances and to the growth of dark matter clustering. Improved distance measurements can test whether the distance-redshift relationship follows the form expected for vacuum energy or whether the dark energy evolves with time. Measurements of the growth rate of large-scale structure provide an independent probe of the effects of dark energy. The combination of distance and growth constraints tests the validity of general relativity on large scales. These goals can be reached by future, large, ground-based optical-imaging surveys that will provide high-signal-to-noise (S/N > 20), multiband, optical data for $>10^9$ galaxies to I = 25, permitting measurements of shape and photo-z for galaxies at redshifts $z < 2$, angular-correlation measurements to $z \sim 4$, and detections to $z \sim 5$. Additionally, ground-based optical facilities with highly multiplexed, wide-field spectrographs can survey several million galaxies at $z > 1$, and use baryonic acoustic oscillations to constrain $H(z)$. Standard candles will remain crucial for cosmological investigations at lower redshifts, where large-scale-structure methods are limited by cosmic variance (GCT 1). Future synoptic imaging surveys will identify tens of thousands of well-measured Type Ia supernovae, accumulating millions of observations by the end of the decade. Spectroscopic follow-up on 8-m and, for a very limited subset, GSMT-class telescopes will be crucial for characterizing the subclasses of those objects and testing for systematic deviations from the norm.

In summary, ground-based OIR observations remain essential in key science areas ranging from exoplanet research to dark energy and from solar physics to large-scale cosmology. The new facilities and programs outlined in the following section will be integral to successful progress in the next decade.

FUTURE PROGRAMS IN OIR ASTRONOMY

This section describes the projects and activities that the OIR Panel (henceforth the panel) is recommending for support, including those of large, medium, and small scope, as well as continuing activities and U.S. participation in the Gemini Observatory. The large and medium activities are presented in this section in priority order within their respective categories. The section "Recommended Priorities and Plan for the Next Decade" brings the large, medium, small, and continuing activities together, presenting a prioritized implementation plan that describes a balanced program in U.S. ground-based OIR astronomy that fits within the budgetary guidelines given to the panel.

Large Programs

The 200-inch telescope on Mt. Palomar was constructed over a period of more than 20 years. It dominated ground-based astronomy for two decades from 1950 and continued as a leading instrument until the next generation of giant telescopes was constructed in the 1990s. Like the 200-inch, the large programs described here are at the forefront of engineering and technical imagination and will be central to astrophysics for decades to come. Whether the topic is circumstellar disks, exploding stars, the fossil record of galaxy assembly, black holes, or cosmic acceleration, present knowledge shows that there are tantalizing and decisive observations that lie just beyond the grasp of today's telescopes. To answer the questions highlighted by the Science Frontiers Panels (see Table 7.1), America's astronomers need to build the telescopes of tomorrow. The panel is convinced that the community knows how to build the instruments needed to answer these pressing science questions. The panel urges a prompt start on these great enterprises.

The Giant Segmented Mirror Telescope (GSMT)

The GSMT projects currently under development in the United States are 25- to 30-m-class telescopes for optical and near-IR observations with instruments that will open new frontiers of research across the entire optical and near-IR spectral regions observable from the ground.

GSMT is a versatile observatory that will push back today's limits in imaging and spectroscopy to open up new possibilities for the most important scientific problems identified by the Science Frontiers Panels. This exceptionally broad and powerful ability over the whole range of astrophysical frontiers is the compelling argument for building GSMT. To make this explicit, in Table 7.1 the panel indicates the ways in which the GSMTs being designed today will address 18 of the 20 science questions and 4 of the 5 discovery areas that the SFPs have identified as today's most important scientific frontiers. These span the range from exoplanets around nearby stars to galaxies beyond redshift 7, where new abilities afforded by GSMT to image at its diffraction limit and to obtain extremely deep multiobject spectra will lead to decisive advances in our field. While it is no coincidence that these telescope designs address the most important scientific problems understood today, a GSMT also has the potential to adapt to new technology and to new scientific questions in the future. As with the other great ground-based observatories, an ongoing program of instrument development will ensure that the telescopes built in the coming decade will have capabilities that remain at the technological forefront and can be focused on the most pressing scientific questions of the coming decades.

The great strength of a GSMT lies in the breadth of its capabilities and the breathtaking new regions of angular resolution, sensitivity, and speed that it will open up for the first time. Two U.S.-led projects for a GSMT, the Giant Magellan Telescope and the Thirty Meter Telescope are technologically ready. They collect 5 to 9 times as much light as the most powerful telescopes operating today. With the capable AO systems planned for early operation, these telescopes will make images that are 3 times sharper than those obtained with existing ground-based telescopes (and 10 times sharper than HST's!) In many situations where this type of imaging is useful, the exposure time it takes to make an image of a given quality scales as $1/D^4$, where D is the telescope diameter. This means that the gain in speed over the largest existing telescopes is an astonishing factor of 70. Five nights on GSMT would be worth a year on today's telescopes, opening up new possibilities for ambitious work and making the study of fast-changing events possible for the first time.

GSMT will open up discovery space in remarkable new directions. They will probe dense environments within the Milky Way and in nearby galaxies and, coupled with advanced AO, map planetary systems around nearby stars. Achieving an astrometric precision better than 30 to 50 micro-arcseconds, GSMT capabilities will approach those of planned space missions while offering sensitivities that are, in some cases, almost 10 magnitudes better. With rapid-response capabilities coupled to sensitive, high-resolution spectroscopy, GSMTs will be vital to characterizing the physical properties, and the environments, of new variable sources discovered in synoptic surveys, including flare stars, novae, supernovae, GRBs, and, eventually, gravitational-wave sources. GSMT sensitivity will be crucial in detailed investigations of discoveries from existing and planned facilities—including JWST, ALMA, and LSST—in science areas ranging from probing star and planet formation in the solar neighborhood to understanding the properties of the first stars and the reionization era. Table 7.2 provides five specific examples of key science programs that are made possible only by GSMTs.

Given the range of crucial and unique capabilities that are integral to addressing so many high-priority questions, the panel concluded that GSMTs must be pursued vigorously in the coming decade.

TABLE 7.2 Five of the Key Science Programs That Can Be Addressed Only with a GSMT

SFP Question	Science	Capability	Synergies
PSF 2, 3	Direct detection and spectroscopy of giant exoplanets; orbital measurement and characterization of disk environments	High-sensitivity adaptive optics (AO)-assisted coronagraphy and IFU spectroscopy	JWST, ALMA
GAN D, GCT 3	Orbital characteristics of faint sources near the galactic center; measuring the black hole mass and R_0 to 1 percent and testing GR in the medium-field regime	High-sensitivity, high-precision AO-assisted astrometry in crowded fields	
SSE 2, D; CFP 2	Spectroscopy and imaging of supernovae and GRBs and their environments; characterizing the progenitors of Type Ia supernovae and testing	High-sensitivity, high-precision AO-assisted photometry and spectroscopy in crowded fields	LSST, JWST
GAN 4, GCT 1	Radial-velocity and proper-motion measurements for hundreds of stars in dwarf galaxies; probing velocity anisotrop, and the nature and form of the underlying dark matter	High-sensitivity, high-precision AO-assisted astrometry and spectroscopy in crowded fields	
GCT 4, 5	Near-infrared spectroscopy of galaxies and large, forming star clusters that are gravitationally lensed or that are rich in massive stars at redshifts $z > 7$; probing the physical properties of the first stars	High-sensitivity, high-precision spectroscopy at the faintest magnitudes	JWST, ALMA

The U.S. GSMT Projects As mentioned above, there are two U.S. GSMT projects: the Giant Magellan Telescope and the Thirty Meter Telescope. They are described next, in alphabetical order.

The Giant Magellan Telescope (GMT) The GMT project design makes use of seven 8.4-m-diameter mirrors to achieve an overall primary mirror with the resolving power of a 24.5-m diameter mirror (Figure 7.11). The baseline project includes the telescope, an adaptive secondary mirror system, and an initial suite of three to four instruments to be selected in 2011 from eight concepts currently under development (Figure 7.12). An international partnership has been formed to construct and operate GMT. It consists of the Carnegie Institution for Science, Harvard University, the Smithsonian Astrophysical Observatory, Texas A&M University, the University of Texas at Austin, and the University of Arizona from the United States; the Korea Astronomy and Space Science Institute, representing Korea; and the Australian National University, together with Astronomy Australia, Ltd., on behalf of Australian astronomers.

The GMT project office and staff have been established in Pasadena, and the project completed its conceptual design phase in 2006, following a successful conceptual design review. More than 30 people from across the partnership are actively involved in the project. The project-management team has developed a full work breakdown structure and schedule and produced a total cost appraisal for the project. The project is now in the design development phase, which is scheduled to end in 2011 with a system preliminary design review and submission of the implementation plan for construction and commissioning. The project recognized early on that production of the primary mirror segments is on the critical path and acted in 2005 to cast the first 8.4-m primary mirror segment in the Steward Observatory Mirror Laboratory (SOML). Since then SOML has been working on figuring and polishing the mirror segment, and the overall fabrication effort is 85 percent complete. The site for the telescope has been selected. It will be located at the Las Campanas Observatory in Chile, which already has six operating telescopes, including the two Magellan 6.5-m telescopes.

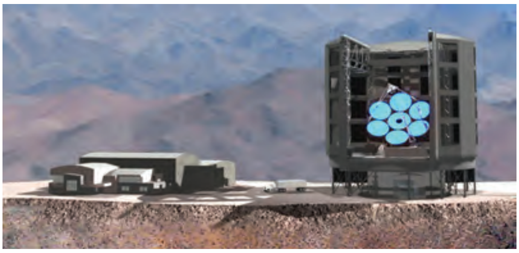


FIGURE 7.11 The current GMT observatory design shown as it would appear on Cerro Las Campanas. SOURCE: Courtesy of GMTO; image by Todd Mason/Mason Productions.

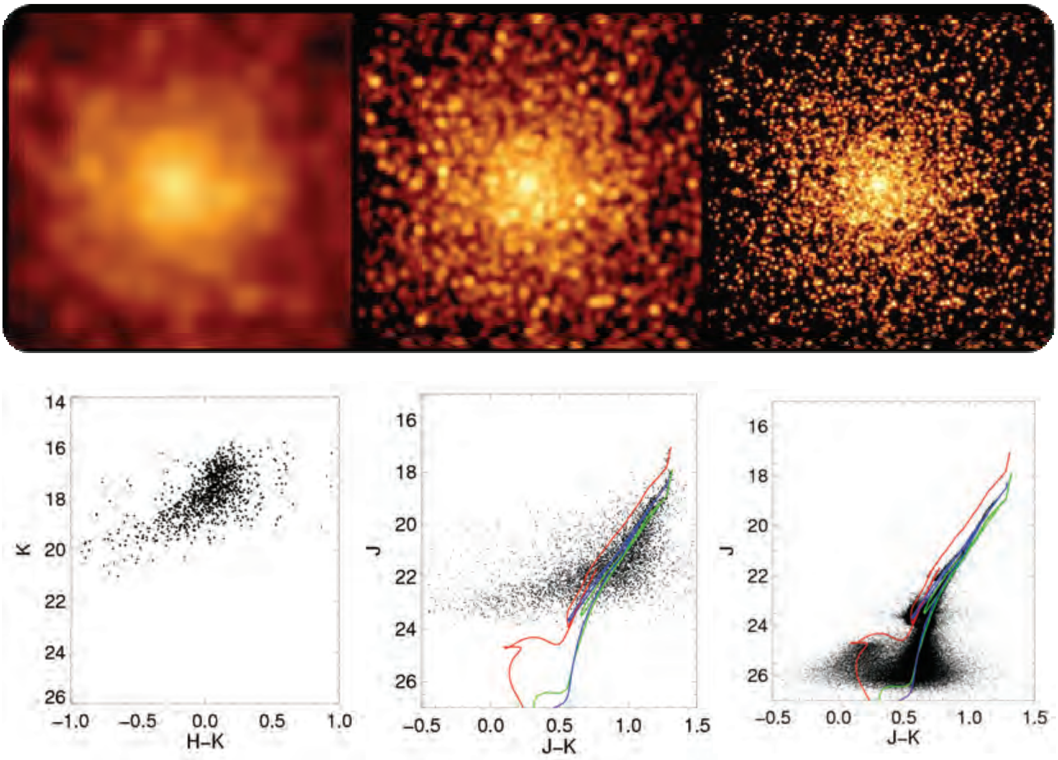


FIGURE 7.12 Examples of simulations of gains to be realized from the GSMT's aperture and adaptive optics. *Upper:* Simulated H-band images of a globular cluster at the distance of NGC5128 (Cen A) with a 3-pc core radius. The left panel shows a simulated image with the resolution of HST; the center panel corresponds to an 8-m aperture, and the right image uses the GMT PSF and 4-mas pixels. Each panel is 2″ on a side. *Lower:* Color-magnitude diagrams for M32. Gemini observations (left) are compared to simulated data for JWST (center) and GSMT with adaptive optics (right). The power of the large aperture in resolving crowded regions is clearly demonstrated in this simulation. SOURCE: *Upper:* P. McCarthy et al., "Giant Magellan Telescope Project: Response to the DS2010 Activity RFI," Astro2010 white paper, available by request from the National Academies Public Access Records Office at http://www8.nationalacademies.org/cp/ManageRequest.aspx?key=48964. *Lower:* K. Olsen et al., "The Star Formation Histories of Disk and E/S0 Galaxies from Resolved Stars," Astro2010 white paper, available at http://sites.nationalacademies.org/BPA/BPA_050603, last accessed February 2011.

The GMT project estimates the cost to complete the project—including corporate and project office and systems engineering, telescope system and optics, adaptive optics, instrumentation, enclosure and facilities—to be $686 million (FY2009 dollars), a value that includes an overall 20 percent reserve. The operations budget, including continuing facility and instrumentation development, is estimated to be $36 million per year in FY2009 dollars. The current schedule calls for early science operations beginning in 2018, with the project phase ending near the end of the decade, 1 year later.

The Thirty Meter Telescope (TMT) The TMT project plans to build a 30-m primary mirror consisting of 492 segments, each 1.45 m in size (Figure 7.13). The telescope, located at the Hawaiian Mauna Kea Observatory, will have an AO system and an initial suite of three instruments, including an infrared-imaging spectrograph, an infrared multi-slit spectrograph, and a wide-field optical spectrograph (Figure 7.14). An international partnership comprising a group of Canadian universities, Caltech, and the University of California has been formed to carry out the project. The National Observatory of Japan is a collaborating institution in the project, and China has recently joined the effort as an observer. The TMT Project Office and staff have been established, and in March 2009 the project successfully completed its 5-year, $77 million design development phase. It subsequently entered its early construction phase with the goal of beginning construction in 2011.

Approximately 40 full-time technical personnel are engaged in TMT design and early construction studies at the Project Office in Pasadena, with a nearly equal number distributed at the various partner institutions. The technical level of TMT is enhanced by a work breakdown structure that is five to seven levels deep for the majority of the project, a complete and well-staffed systems-engineering approach, and organized and "well-positioned" software development. In collaboration with the European ELT project, which has substantially similar mirror segments, four polishers are currently engaged in studies of large-scale segment production. Extensive science flowdown has been established to the engineering requirements for the project, and a detailed science case has been in place since 2007. A draft state

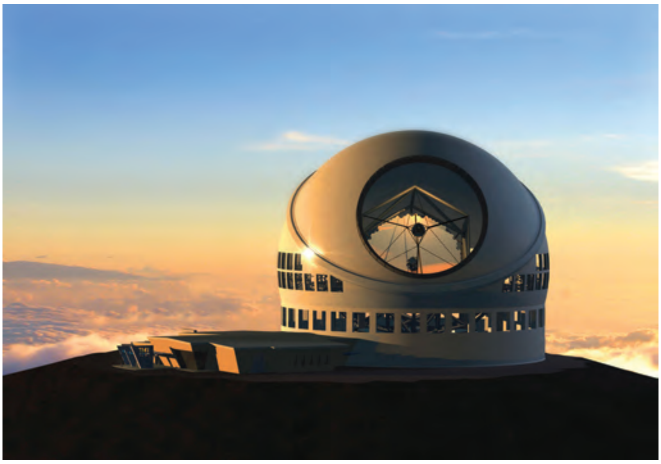


FIGURE 7.13 The current TMT observatory design. SOURCE: TMT Observatory Corporation.

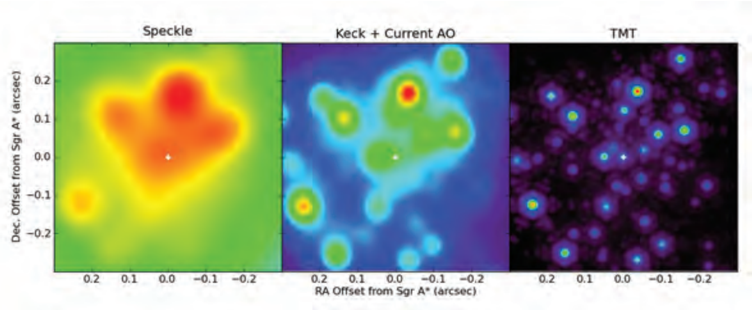


FIGURE 7.14 Imaging of the galactic center with a speckle interferometer, the Keck Observatory with its current adaptive optics system, and the simulated performance of TMT with adaptive optics, showing individual stars orbiting the supermassive black hole. The TMT observations detect enough stars with sufficient precision that their orbital paths can be used to test the predictions of general relativity. SOURCE: Image courtesy of and created by Andrea Ghez and her research team at UCLA from data sets obtained with the W.M. Keck Telescopes.

environmental impact study for the chosen Hawaiian site was published in May 2009, and a final study is expected in early 2010.

The TMT project estimates that the cost to complete the project—including facilities telescope, AO and instruments, operations design and observatory software, and project management and system engineering—is $987 million (FY2009 dollars), a value that includes a 30 percent contingency. Operating costs are estimated to be $54 million per year, including $20 million per year for new instrumentation. The project goal is to achieve first light in 2018.

Assessment of Technical Readiness, Cost, and Schedule for GSMT The panel had available to it two reviews of the U.S. GSMT projects. First, a non-advocate community assessment of the status of both GMT and TMT (GCAR) was carried out in 2009 by the National Optical Astronomy Observatory (NOAO) as GSMT program manager for NSF. A main goal of the review was to determine the progress of both projects toward the NSF-mandated preliminary design review needed for a major facility. Second, as part of the Astro2010 survey process, an independent analysis of GMT and TMT was carried out by Aerospace Corporation, a consulting firm hired by the National Research Council as part of the survey's assessment of technical readiness and schedule/cost risks. This independent assessment was called the cost appraisal and technical evaluation (CATE). The CATE process did not evaluate operating costs.

Having assessed these two sources of independent information on the GSMT projects, the OIR Panel found that the technology needed to build a GSMT exists. In summary, the panel found as follows.

The GMT builds on the successful heritage of the 6.5-m Magellan Telescopes and the twin 8.4-m Large Binocular Telescope (LBT). The TMT builds on the successful design of the 10-m Keck Telescopes with their segmented mirrors, a technology that the European Extremely Large Telescope has also selected.

Adaptive optics is a key area for both projects, since the justification for building a large (20- to 30-m) telescope hinges on achieving diffraction-limited images to enable revolutionary science. AO for GSMT-scale telescopes is practical given current technology. Deformable mirrors, wavefront sensors, tomographic reconstruction algorithms, and lasers all exist in prototype forms that will be straightforward to scale to the requirements for a GSMT. The TMT project has completed a preliminary design for its first-light AO system, NFIRAOS, which could be constructed today using existing technologies. GMT plans to use deformable mirrors that are very similar in size and complexity to the adaptive secondary mirrors now being completed for the LBT. Although AO for a GSMT is an order of magnitude more challenging than the first AO systems on 5- to 10-m telescopes, a new generation of instruments bridges the gap. Specialized high-performance facilities—extreme AO designed for planet-imaging, multi-laser wide-field AO, or prototype visible-light AO systems—are being deployed on 5- to 10-m telescopes in the next 1 to 5 years. To achieve extreme performance on an 8- to 10-m telescope requires technology close to that required for general-purpose AO on 20- to 30-m telescopes. Table 7.3 compares the requirements of NFIRAOS to AO systems commissioned at the beginning of the last decade and to the state of the art at the beginning of this decade. Other AO modes (e.g., high-contrast ExAO) require

TABLE 7.3 Comparison of Adaptive Optics (AO) Component Specifications for Three Systems

Component	Keck AO LGS (2003)[a]	Current State of the Art (2011)[b]	TMT NFIRAOS (2018)[c]
N = Diameter of deformable mirror (DM) in actuators and actuator spacing	N = 16 actuators at 1-cm spacing	N = 41 actuators at 0.5-cm spacing N = 64 actuators at 0.25-cm and 0.04-cm spacing	N = 64 and N = 76 actuators at 0.5-cm spacing
Visible-light wavefront sensor (WFS) detector	64 x 64 pixels at 700 Hz, 8 e^- noise	160 x 160 pixels at 1500 Hz, 3 e^- noise	360 x 360 pixels at 800 Hz, 3 e^- noise
Lasers	1 x 15 W	1 x 50 W	6 x 25 W
Wavefront control algorithm	Least-squares reconstruction between single DM and single WFS	Least-squares reconstruction between 5 WFSs and 3 DMs	Maximum-likelihood tomography between 6 WFSs and 2 DMs
Wavefront reconstruction computational scale, rate, and hardware	480 x 349 at 700 Hz (high-performance CPUs)	6,000 x 3,000 at 2,000 Hz (multiple GPUs)	35,000 x 7,000 at 800 Hz (field-programmable gate array architecture)

[a] First-generation Keck laser-guide-star system.

[b] Next-generation systems being commissioned on 5- to 8-m telescopes in 2011 (Palomar PALM3000, Gemini Multi-conjugate AO and Gemini Planet Imager, and VLT SPHERE instruments).

[c] Requirements for AO for a GSMT.

further technology development that would be supported by an AO development program (see below). The panel believes that there is a clear path for scaling to the requisite level and extensive opportunities for testing before final implementation on a completed GSMT.

The independent CATE assessment found that the technical risk of GMT, independent of cost, is medium. For the GMT, key technical and project challenges noted in the two external assessments include:

1. Production and metrology of primary mirror segments; the six off-axis 8.4-m segments are a particular challenge;
2. Alignment and phasing of primary mirror segments;
3. Design and production of segments for an adaptive secondary mirror;
4. Other AO components, particularly lasers;
5. Insufficient instrument maturity and cost allocation; and
6. The need to strengthen project management and system engineering.

The independent CATE assessment is that cost and schedule risk for the GMT is medium-high in comparison to the project's estimates. Based on the available information, the CATE assessment concluded that a full cost appraisal is not possible at this time. Therefore a cost sensitivity analysis was carried out for the primary mirrors and instruments. That analysis yielded a cost appraisal (at a 70 percent confidence level) that was roughly 60 percent higher than the cost appraisal provided by GMT project personnel, which would take the cost without operations from $689 million to $1.1 billion. The cost risks for the primary mirrors are increased polishing and metrology costs above current estimates and the cost growth associated with a 24-month schedule delay. The result for the growth in instrument costs is based on experience with 20 space projects at a similar level of development and the assumption that the cost growth would be similar.

The CATE assessment rated the schedule risk to be 4 years, with primary mirror fabrication and completion of the AO system and science instruments being the main risk items. Note that the independent assessment used the completion of all initial instrumentation and the AO system as defining the end of the project.

For the TMT, the community assessment review places the TMT project at a preliminary design review level essentially consistent with the initial readiness reviews necessary for NSF major facilities consideration. The CATE independent assessment noted the following key challenges:

1. Meeting the very ambitious production plan for primary optics—492 segments for the primary mirror, 82 spare segments, and 82 different optical prescriptions;
2. Segment alignment and phasing;

3. AO components, particularly lasers; and
4. Insufficient instrument maturity and cost allocation.

A key finding of the independent CATE assessment for TMT was that the typical reduction in manufacturing time for each doubling of a mass-production line, as proposed for the primary mirror segments, is likely only 20 percent. Applying this to the production rate demonstrated for Keck segments yields 152 months as opposed to 82 months to produce the 574 segments if only a single production facility is used. Use of a second vendor or production line would reduce this to 104 months. The overall technical risk for TMT, independent of cost, is judged to be medium high.

The CATE assessment of the cost and schedule risk for TMT is judged to be high in comparison to the project's estimates. A cost-sensitivity analysis for the primary mirrors and instruments developed by the independent contractor yielded a cost appraisal (at a 70 percent confidence level) about 42 percent higher than the cost appraisal provided by the TMT project personnel, which would increase the cost from $987 million to about $1.4 billion. The main factors in the cost risk for the primary mirrors are the uncertainties in the ability of external vendors to meet the cost and schedule estimates of the project. As with GMT, the result for the growth in instrument costs is based on experience with space projects at a similar level of development and the assumption that the cost growth would be similar. The independent estimate of the schedule risk is 3 to 6 years, depending on the approach used to fabricate the primary mirror segments.

Based on the outcomes of the independent assessments, the panel concluded that both U.S.-led GSMT projects should in the near term undergo an independent full cost and schedule review as they proceed through any NSF down-select, preliminary design review, or major facility review process.

The Community and Federal Participation in GSMT Because of the vast array of key science questions GSMTs address and the large, broad, and diverse U.S. community that will use the telescopes, the panel concludes that the United States should ideally amass a total share of 50 percent of the equivalent of a GSMT, made up from one or both of the GSMT projects, consistent with the recommendation of AANM, the 2001 decadal survey.

It is expected that U.S. community access to a GSMT project would be awarded in a competitive process and would be commensurate with the share in construction and operations costs provided by the federal government. Maximizing community access would also require participation in the decision-making process for future development of the telescope (especially instrumentation), as discussed elsewhere in this panel report.

The panel finds that national participation in a GSMT must be well integrated into the U.S. system. As a result, the panel concludes that the only tenable outcome is national participation in at least one GSMT as a full, decision-making (>25 percent) partner that is able to guide instrument choices and operational considerations. *Given the GSMT development schedules, the panel believes that it is both vital and urgent that NSF identify one U.S. project to join as a partner at this level.*

The European Southern Observatory (ESO) has announced plans to build a 42-m GSMT, the European Extremely Large Telescope (E-ELT), as its highest priority. E-ELT design work has been progressing rapidly, and the U.S. community must consider its role in light of these concrete plans. The panel does not recommend joining the E-ELT project, in part because the panel recognizes the importance of leveraging the private and nonfederal contributions to TMT and GMT for the benefit of U.S. astronomy, which will likely total well over $1 billion. At the initial proposed contribution level of a 25 percent share in a GSMT, the role the United States would play in the E-ELT would be too small to secure a scientific and technical leadership position for the nation in the project. Further, if the United States joined the E-ELT project it would simultaneously compete with the GMT and TMT projects, which would undercut U.S. aspirations to be a leader in ground-based astronomy. In addition, as a minor partner in the E-ELT, which is tied to the European system of telescopes, integration with the U.S. system would not be possible.

For the past century, U.S. astronomers have enjoyed access to the most powerful telescopes in the world, and this technological edge has led to transformative discoveries and a vibrant community. Maintaining this momentum is vital to the health of science in this country, as serious international challenges are faced. The panel concluded that without community involvement in a GSMT project, U.S. astronomy is in danger of becoming uncompetitive in the next, and succeeding, decades.

The Large Synoptic Survey Telescope (LSST)

"Wide-fast-deep" are the key words for LSST, a project that developed from a high-priority recommendation of the 2001 AANM: "a large-aperture (6.5-m-class), very-wide-field (~3 deg) synoptic survey telescope" (p. 107). The project has four main science goals: to explore the nature of dark energy, study the solar system, investigate optical transients, and study galactic structure. LSST will perform systematic, repeated surveys of the entire available sky to depths of 24th magnitude at optical wavelengths. Combining repeated survey images will provide composite wide-field images extending more than 10-fold fainter. LSST synoptic data will revolutionize investigations of transient phenomena, directly addressing the key discovery area of time-domain astronomy highlighted by two of the Astro2010 SFPs. LSST data will be invaluable in surveys of regular and irregular variable sources, both galactic and extragalactic, and in astrometry of objects in the solar system and the solar neighborhood (Table 7.4). At the same time, the combined

TABLE 7.4 Five of the Key Science Programs That Can Be Accomplished Only with LSST

SFP Question	Science Program	Capability	Synergies
PSF 2, GAN 5, SSE D	Composition and orbital structure of the outer solar system: 10- to 100-fold increased census of trans-neptunian objects and other small bodies to r ~ 24.5	High-sensitivity, high-cadence, multiband, wide-field synoptic imaging	JWST, ALMA, GSMT
SSE 2, D; CFP 2	Identifying and characterizing Type Ia supernovae to $z \sim 0.5$-0.7 and using those observations to constrain the dark-energy equation of state	High-sensitivity, high-cadence, multiband, synoptic, wide-field imaging	OIR System, GSMT
GAN 3	Architecture and formation history of the MW: identifying very metal-poor stars and mapping star streams in the galactic halo to ~100 kpc using main-sequence subdwarfs and RR Lyraes	High-sensitivity, high-cadence, multiband, wide-field imaging	OIR System, GSMT
GCT 1, CFP 1	Measurement of baryonic acoustic oscillations and cluster growth: large-scale structure from photometric redshifts for $\sim 10^{10}$ galaxies	High-sensitivity, multiband imaging over ~20,000 square degrees to r ~ 27	
CFP 1, 2	Constraints on dark energy from weak lensing measurements for ~3 billion galaxies	High-sensitivity, multiband imaging over ~20,000 square degrees to r ~ 27	

images will provide a multiwaveband, homogeneous, wide-field-imaging data set of unparalleled sensitivity that can be used to address a wide range of high-impact scientific issues. All told, LSST will address 14 of the 20 questions raised by the SFPs and 4 of the 5 discovery areas (Table 7.1).

Substantial progress toward realizing the LSST vision has occurred during the past decade. The LSST project, a consortium of 30 universities and research institutions, plans to build a survey telescope with an 8.4-m primary mirror and a 6.7-m effective aperture on Cerro Pachon, Chile. Its 3.2-gigapixel camera, covering a 9.6-deg^2 field of view, will give LSST an effective entendue of 319 m^2 deg^2, which will be more than 10 times greater than that of any other telescope (Figure 7.15). The telescope will survey the sky over a period of 10 years, during which time each field will be observed 1,000 times in six filter bands and produce 30 terabytes of data per night (Figure 7.16). The project has a credible plan for making the data and tools to analyze them available to the scientific community and the public.

The project is nearing the end of its design-development phase and expects to be ready for construction beginning in 2011. More than 100 technical personnel from across the partner institutions are currently involved in the project. The project uses state-of-the-art project-management techniques for design, system engineering, and cost and schedule planning. Highlights of achievements to date reported by the project include (1) preparation of preliminary engineering designs for major subsystems to estimate both performance and cost; (2) development of a detailed work breakdown structure for the project to the sixth level; (3) estimation

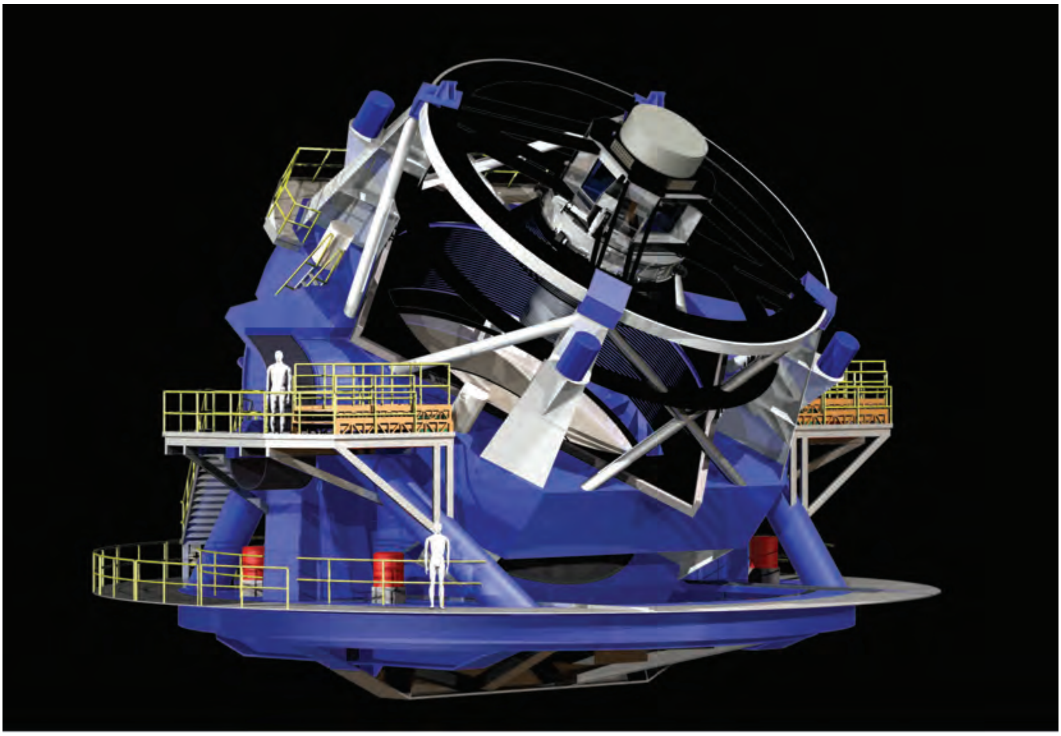


FIGURE 7.15 The 8.4-meter LSST will use a special three-mirror design creating an exceptionally wide field of view, and it will have the ability to survey the entire sky at any given time of year in a single filter in only three nights. SOURCE: LSST Corporation.

of costs based on a detailed work breakdown structure, estimates from potential manufacturers, and prior experience with similar projects; (4) completion of all prerequisites for Chilean governmental permission to construct the telescope, including environmental permitting; (5) placing contracts for casting and polishing the primary/tertiary mirrors; (6) casting the 8.4-m primary/tertiary mirror; (7) contracting for and casting the secondary mirror blank; (8) completion of the science requirements, and their flowdown to engineering/system requirements; (9) and producing the 600-page LSST Science Book,[3] which demonstrates the level of community involvement and reflects the level of scientific promise.

The project estimates that the cost of LSST—including the project-management office, telescope and site, camera, data management, education and outreach, and commissioning—to be $455 million (2009 dollars). This value in-

---

[3] LSST Science Collaborations and LSST Project 2009, LSST Science Book, Version 2.0, arXiv: 0912.0201, http://www.lsst.org/lsst/scibook.

FIGURE 7.16 A simulated image of one 15-second exposure with the LSST charge-coupled device (4K × 4K) with 0.2″ pixels, 0.4″ seeing, and a field of view of 13.7′ × 13.7′, representing roughly 0.5 percent of the LSST focal plane. The brightest stars in the image are 12th magnitude. An object of brightness 33rd magnitude would record 1 photon in a 15-second exposure. The image is a true-color composite of three images, with the g, r, and i filters mapped into B, G, and R colors, respectively. Each color channel is on a logarithmic intensity scale. LSST will produce 2 billion single-band images of the same size. SOURCE: LSST Corporation © 2009. Reprinted with permission.

cludes a contingency of $101 million. Of the total, $299 million is to be requested from NSF, $84 million from DOE, and the balance from other and nonfederal resources. Operations costs are estimated to be $40.9 million per year, with NSF and DOE shares each being one-third, or $13.7 million per year. The remaining third will be sought from other sources.

Assessment of Technical Readiness, Cost, and Schedule for LSST The independent contractor rated the overall technical risk of the LSST project as medium low. The CATE process identified the main remaining risks and concerns as:

1. Support structure for the secondary mirror and camera;
2. Maturity of focal plane arrays for the camera and their procurement;
3. Production of the large camera elements;
4. Camera mechanisms;
5. Data management challenges; and
6. Achieving the mirror surface figure specifications.

The LSST team acknowledges these challenges and should be able to achieve the performance goals and commission the telescope system prior to 2020.

The independent assessment of the LSST project cost produced a value that was slightly below the project values, primarily in the areas of the telescope, optics, and software. However, the contractor was unable to assess the costs of instruments for the ground-based facilities because of a lack of analogous data, and so the costs associated with the camera were not assessed. Because the camera is both a technically challenging item and on the critical path for the LSST schedule, the panel finds it prudent to use the project's cost appraisal pending the detailed cost review that will occur following the NSF preliminary design review. The independent assessment of the LSST construction schedule yielded an increase of 15 months and a completion date of the end of 2018 relative to the project estimate, with the production schedule for the detectors and camera being the main risk item. No assessment was made of the $41 million annual operating cost appraisal.

The panel concluded that LSST is technologically ready and is the most advanced toward construction readiness of the three large projects under consideration by the panel.

Additional Comments and Conclusions for LSST The panel notes that follow-up observations are critical to many LSST science goals. Time-critical spectroscopy of variable and transient objects is the most important need. Photometric follow-up of variables, particularly those that brighten beyond the saturation limit of LSST, is also important. The U.S. system in its current form, even if augmented by a giant 30-m telescope, is likely insufficient to provide enough follow-up capability. Strong consideration should be given to strategic use of 4- and 8-m-class telescopes in direct support of the LSST project. The LSST project is understandably devoting its efforts to creating the LSST facility itself, but the panel concluded that additional attention to arranging the necessary follow-up observational capabilities to fully exploit LSST is essential.

LSST will result in significant changes in the modes of research carried out by much of the United States and, indeed, the global astronomical community. If properly supported, the archive will enable frontier research by scientists at institutions without their own telescope resources. LSST will foster large collaborative teams of researchers probing the broader science goals. This project has very significant computational challenges, in terms of data acquisition, processing, and storage, and data-mining algorithms. Solutions to these problems will require input from the information technology community and may well generate positive payback to society well beyond the astronomical research community.

The high value and transformative nature of LSST derives from the production of a high-quality public archive and the investment of effort by a wide spectrum of astronomical users. The panel believes that the LSST project has vigorously and effectively engaged this aspect of its mission and concludes that the project is well scoped and costed, with a detailed science-operations plan. Based on its own analysis of the information provided to it by the proponents of LSST and by the survey's independent contractor, the panel concluded that the LSST project is in an advanced state and ready for immediate implementation.

GSMT and LSST as a Coordinated Program

Rather than view its prioritization as a competition between GSMT and LSST, the panel stresses the synergy of these two projects. Scientific coordination between GSMT and LSST will be robust regardless of the ultimate sites selected for the projects; even if the projects are located in opposite hemispheres there will be substantial overlap in the sky accessible to both. Further, each would be greatly enhanced by the existence of the other, and the omission of either would be a significant loss. The combination of wide-area photometric surveys and large-aperture spectroscopy has a long, productive history in OIR astronomy, grounded in the 1950s with the combination of the Palomar 48-inch Schmidt and the Hale 200-inch and supplemented in later years by the NOAO 4-m telescopes. Interesting sources identified in the wide-field survey are studied in detail with the larger telescope. An aperture ratio of ~5:1 is particularly useful: the area ratio of ~25 combined with an exposure-time ratio of up to 2 orders of magnitude appropriately couples the capabilities of broad-band photometry to low/medium-resolution spectroscopy. More recently, the combination of the Sloan Digital Sky Survey and new 8- to 10-m class telescopes provided a 4-fold increase in aperture, coupled with order-of-magnitude improvements in detector quantum efficiency.

Thus the *combination* of GSMT and LSST would create a particularly powerful opportunity for another series of breakthroughs. The power of this combination can be seen in Figure 7.17, which demonstrates the wide range of SFP questions to which *both* GSMT and LSST apply. The panel notes that these are not overlapping capabilities—rather they are in all cases complementary. LSST will provide photometry and time-series data for large numbers of objects, while GSMT will provide highly detailed studies of critical targets.

The panel concluded that a crucial goal for ground-based OIR astronomy in the coming decade should be to realize the potential of the combination of these facilities as linchpins for an upgraded comprehensive U.S. OIR system of telescopes.

Medium Programs

Mid-Scale Projects and Instrumentation

The OIR Panel's highest priority for a medium-scale initiative is support for the NSF Mid-Scale Projects and Instrumentation program. Construction pathways are needed for powerful new specialized capabilities that will provide the sensitive measurements needed to answer the highest-priority science questions. Those capabilities include new instruments for existing and new telescopes, new mid-size facilities, large surveys, and technical-development programs. The cost of construction and operation of such capabilities falls in the ~$8 million to $120 million gap between the existing NSF Major Research Instrumentation (MRI) and MREFC programs; they are not easily funded in the current ground-based astronomical landscape. This Mid-Scale Projects and Instrumentation program is essential for the United States to effectively realize the full capabilities of the astronomical system.

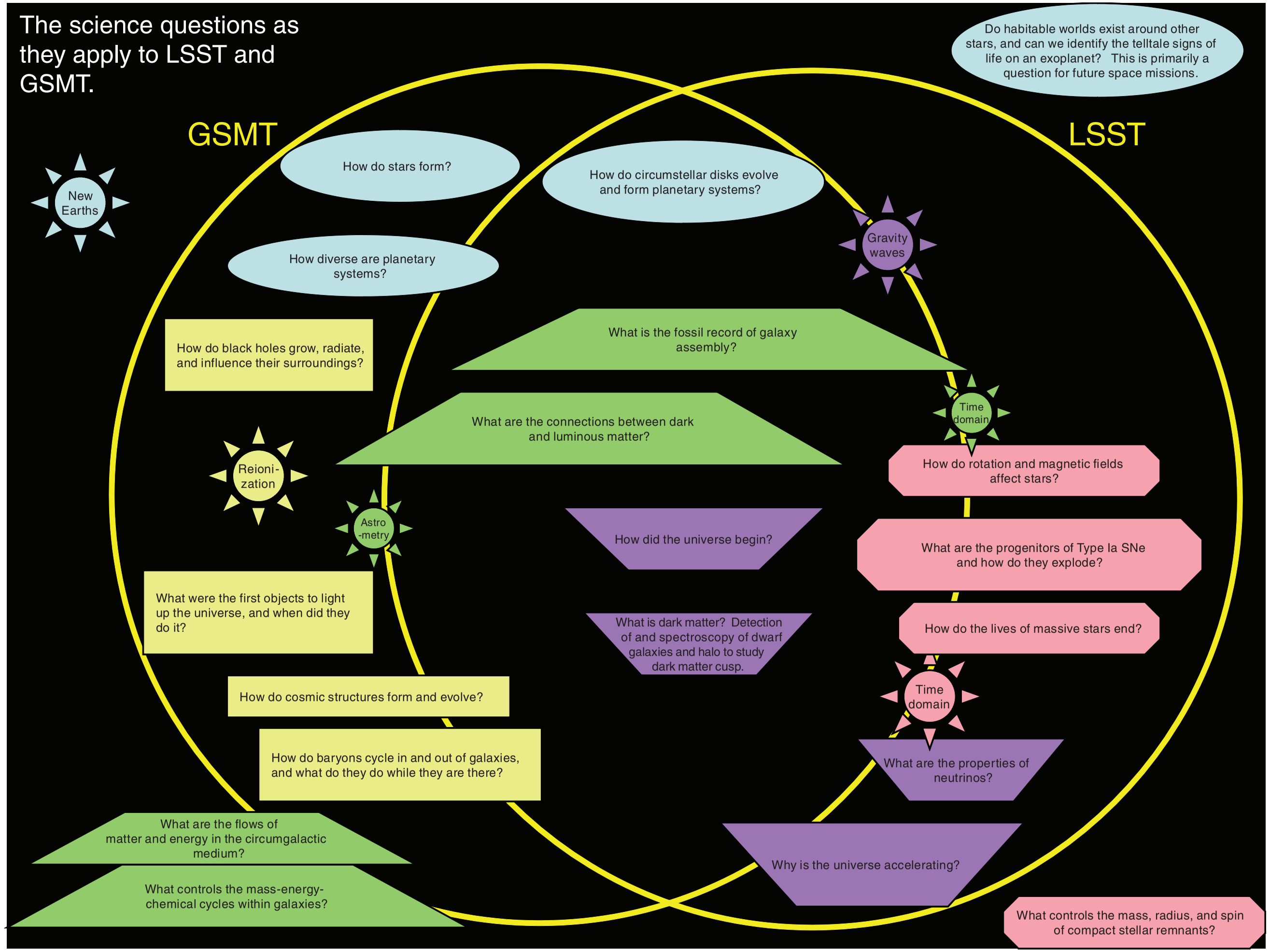


FIGURE 7.17 Venn diagram showing the interrelationship of GSMT and LSST for addressing the main science questions identified by the Astro2010 Science Frontiers Panels (see Table 7.1).

Cutting-edge astronomical instrumentation has significantly increased in cost over the past 20 years. As examples, the DEIMOS spectrograph for the Keck Observatory (completed in 2002) cost approximately $14 million in FY2009 dollars ($21 million with indirect costs), while the proposed Gemini WFMOS concept exceeded $65 million. There are several reasons for this:

- *Scale.* Instruments for 8- to 10-m telescopes are necessarily larger and more sensitive than those for smaller telescopes.
- *Increased scientific capabilities.* Most telescopes already are equipped with simple imagers and moderate-resolution spectrographs. Providing significant enhancements over current capabilities requires complex systems that employ new technologies, such as large-scale multiplexed spectrographs or advanced adaptive optics.
- *Data management and software complexity.* As instrumental output increases from megabytes to gigabytes and terabytes, costs also increase, and this is all the more applicable to survey instruments where sophisticated pipelines are needed to provide fully reduced data products.

In addition, projects of this scale require appropriate management and contracting approaches to control cost and schedule and to reduce ultimate cost overruns. This can increase initial, up-front costs and in some cases requires involvement of capable and more expensive non-university institutions such as national laboratories or aerospace companies.

Another factor is that some of the most important and exciting instrument proposals concentrate on a particular scientific question or must be optimized for a specific capability, rather than being general-purpose facilities. Again, this is in part due to progress during the previous decade—most observatories have complete suites of general-purpose instruments, and the questions within their reach are being answered; new scientific questions will often require very challenging observations necessitating instruments designed specifically to address those questions.

Several new OIR instruments or facilities proposed for the next decade that would provide revolutionary scientific capabilities include:

- Massively multiplexed optical/near-IR spectrographs and spectroscopic surveys on 4- to 8-m telescopes to map large-scale structure for the study of dark energy and cosmology (CFP 1, 2, 3), to measure the evolution of galaxies across redshift and environment using spectral diagnostics (GCN 1), and to study the chemical and dynamical history of the Milky Way with large spectroscopic samples of stars (GCN 4, GAN 3, SSE 3). Several SFPs identified the need for surveys at least an order of magnitude larger than those currently underway; such surveys require new instrumentation for either fully or highly dedicated facilities, as well as large survey teams. BigBOSS and HETDEX are compelling examples of next-generation projects in this category.

- Next-generation AO systems, ranging from wider-field-of-view systems based on ground-layer and multi-conjugate AO to improved narrow-field systems, to provide diffraction-limited capability at visible-light wavelengths. Such systems will enable diffraction-limited spectroscopy of the cores of nearby galaxies to study the environments of massive black holes (GCT 3), integral-field spectroscopy to study the stellar constituents and kinematics of galaxies during the crucial epoch at $z = 1.5\text{-}2.5$ when most galaxies and stars are forming (GCT 1), and high-resolution studies of the multiplicity of solar system minor bodies to provide important information about the history of their formation, in addition to more familiar applications to the detection and characterization of planets and protoplanetary disks (PSF 2, 3) and stellar birth and death (PSF 1; SSE 2, 3)
- New solar instruments and observatories to complement the unprecedented high-resolution view of ATST. A robust suite of more-capable, full-Sun, synoptic observations is needed to characterize the evolving global context, explore the solar interior, map the coronal magnetic field (SSE 1), and enable long-term time-domain studies (SSE D). Proposals for new facilities such as a coronal magnetism observatory that will provide measurements of the coronal magnetic field to understand structure, dynamics, and particle acceleration in a stellar atmosphere are examples of solar projects for the mid-scale program. Second-generation instrumentation for the ATST telescope is another important category.

Additional examples include:

- Precision IR Doppler measurements to detect Earth-analog planets orbiting red dwarf stars and enhanced visible-light Doppler capabilities, perhaps with a dedicated 4-m telescope (PSF 3 and D).
- Specialized scientific capabilities for the future GSMTs, such as ultrahigh-contrast planet imaging that could directly detect mature giant or "super-Earth" planets (PSF 3).
- IR surveys leveraging the enormous advances in IR array technology to provide several-orders-of-magnitude increase in sensitivity over 2MASS and synoptic sky coverage (GCT 1 and D).

These projects are not well matched to existing NSF astronomical funding programs. They are too large for the Advanced Technologies and Instrumentation (ATI), MRI, and TSIP programs. Moreover, the mid-scale program is intended to (1) cover all ground-based astronomy, including solar and radio, millimeter, and submillimeter projects as well as OIR, (2) support proposals for new telescopes, and (3) allow for instrument upgrades to federally supported telescopes. TSIP, in contrast, is for existing, large, nonfederal telescopes in the nighttime OIR system. In the past, projects in the range of the mid-scale program have been funded by NSF in response to unsolicited proposals, but this approach also creates problems, with no clearly defined process for principal investigators to follow, no easy way to select projects that match to particular scientific goals, and an opaque selection process.

The panel recommends supporting the Mid-Scale Projects and Instrumentation program at the level of $20 million per year to address the needs outlined above.

The approach would be analogous to the NASA Explorer missions, where projects are driven by the scientific vision of a team (rather than an imposed set of outside requirements), focused around a set of well-articulated high-priority science goals, cost-capped, but allowing for innovation. Typically this would fund instruments, telescopes, or projects with an NSF cost of $10 million to $30 million; some might exceed $50 million. The panel notes that in many of these cases, costs would be shared with other federal agencies or partners. In a steady state, the program might be funded at the level of $20 million per year, allowing it to support two to five projects at any given time, with projects initially funded only through a conceptual design for later down-select.

Selection criteria would include scientific merit, connections to research priorities set out in agency reports such as those of the AAAC Task Forces, project management, quality of the team, realism of the project plan, workforce development benefits, and the outcome of a conceptual design review. Public access to data products after an appropriate period (most likely through the archive centers recommended below in this report) would be expected unless explicitly restricted for scientific reasons. Public access to telescopes involved would depend on the nature of the project—this program is not intended to replace TSIP funding of general-purpose instruments for the broad community but instead to fund advanced facilities judged on their scientific merit.[4]

It is anticipated that this program would largely replace the current ad hoc funding of major unsolicited proposals. The panel notes that while it has focused on possible OIR concepts for such a program, it would encourage the program to include all wavelengths.

Strengthening the U.S. Telescope System in the Next Decade

The OIR Panel's second-highest-priority, medium-scale initiative is enhancing funding to strengthen the U.S. telescope system in the next decade. The number of research telescopes of all apertures available to U.S. astronomers surpasses that of ESO and other countries, yet these facilities do not deliver their full combined

---

[4] For example, the Sloan Digital Sky Survey provided public access to its data but not to the telescope and instruments, which were not designed for, or used in, a guest-observer mode.

potential because technical effort is duplicated and not every facility has the resources to run at peak efficiency.

The 2001 decadal survey (AANM) advocated a "system" approach toward the entire suite of U.S. OIR facilities in order to encourage collaborations between federally funded and independent observatories so that federal funds are leveraged by private investment. AANM went on to recommend a Telescope System Instrumentation Program (TSIP) as its "highest-priority moderate initiative . . . [TSIP] would substantially increase NSF funding for instrumentation at large telescopes owned by independent observatories and provide new observing opportunities for the entire U.S. astronomical community" (pp. 12-13). AANM recommended a funding level of $50 million over the decade. A productive TSIP was established by NSF in 2002 and has since awarded $25 million, securing public observing time on nonfederal facilities with apertures >6 meters in return for new instrumentation (Figure 7.18). However, these investments have not strategically guided instrument choices or effectively coordinated activities across the portfolio of large telescopes.

Recently NOAO has engaged the community via two studies, ReSTAR and ALTAIR,[5] that have assessed the science goals and U.S. telescope system performance for telescopes with apertures below and above 6 m, respectively. Combined, these reports suggest that a true system is manifold, including telescopes from 1-m aperture up to the 30-m-class GSMTs with telescopes at every level providing fundamental input toward answering the leading science questions of the next decade. Smaller telescopes are an integral component because they provide survey and time-domain support complementing large-telescope science, execute time-intensive programs unique to small apertures, provide the hands-on training ground for the next generation of observers and instrumentalists, and can serve as dedicated, single-instrument facilities, focused on specific scientific questions.

As the cost of instrumentation grows and as the GSMTs begin construction that will not be completed until the end of the decade, the opportunity remains ripe for federal investment in nonfederal facilities and for these investments to play a role in shaping a more efficient federal/nonfederal telescope system. Coordinating the unique power of the U.S. system, however, is a complex task. Strategies cannot be imposed centrally but must be implemented by generating incentives that foster community and efficiency across highly diverse groups and institutions. Investments have to be coordinated without expecting the individual components of the system to relinquish the leadership opportunities that they value, but should encourage independent observatories to consider the broader national perspective in framing their programs. Not surprisingly, the panel's view of the next steps in

---

[5] ReSTAR report, available at http://www.noao.edu/system/restar/files/ReSTAR_final_14jan08.pdf. Accessed May 2010. ALTAIR report, available at http://www.noao.edu/system/altair/. Accessed August 2010.

FIGURE 7.18 The network of federal and nonfederal observatories, allied for excellence in scientific research, education, and public outreach, that enable experimentation and exploration throughout the observable universe. Telescopes pictured from top: *First row:* Gemini North with its laser-guide-star (LGS) system (courtesy of Gemini Observatory); Magellan Clay and Baade Telescopes (used with permission of the Observatories of the Carnegie Institution for Science); Lick Observatory Shane Telescope (courtesy of Lawrence Livermore National Laboratory); Large Binocular Telescope Observatory (courtesy of Large Binocular Telescope Corporation). *Second row:* Hobby-Eberly Telescope (courtesy of Thomas A. Sebring, *West Texas Time Machine: Creating the Hobby-Eberly Telescope*, Little Hands of Concrete Productions, 1998); Multiple Mirror Telescope (MMT) Observatory (courtesy of Howard Lester, MMTO); Keck Telescopes (courtesy of W.M. Keck Observatory). *Third row:* Small and Moderate Aperture Research Telescope System (SMARTS)/Cerro-Tololo International Observatory (courtesy of NOAO/AURA/NSF); Kitt Peak 4-meter Mayall telescope (courtesy of NOAO/AURA/NSF); Wisconsin-Indiana-Yale-NOAO (WIYN) Telescope (courtesy of Mark Hanna/NOAO/AURA/NSF); Palomar Observatory's 200-inch Hale Telescope (courtesy of Scott Kardel/Caltech/Palomar Observatory). *Fourth row:* The Southern Astrophysical Research (SOAR) Telescope (© Southern Astrophysical Research Consortium, Inc.; all rights reserved; reprinted with permission); Ritter Observatory (courtesy of Erica N. Hesselbach); WIYN 0.9-meter Observatory (courtesy of NOAO/AURA/NSF); Astrophysical Research Consortium 3.5-meter Telescope (courtesy of Dan Long, Apache Point Observatory). *Fifth row:* Dunn Solar Telescope (courtesy of NSO/AURA Inc. and NSF); Wilcox Solar Observatory (courtesy of Wilcox Solar Observatory at Stanford University); Mt. Wilson Solar Observatory (© 2003 The Regents of the University of California; all rights reserved; reprinted with permission); Big Bear Solar Observatory (courtesy of Big Bear Solar Observatory/New Jersey Institute of Technology); McMath-Pierce/SOLIS solar telescope (courtesy of NOAO/AURA/NSF).

the evolution of the system was varied. At the two extremes were a more ambitious version of the current approach to managing the system and a new, more-strategic approach. In detail, these extremes would entail:

1. A more ambitious version of the status quo would include an augmented TSIP providing substantially increased support (both directly via current programs and also via the proposed expansion of "mid-scale" instrumentation). Support would also be extended to telescopes with smaller apertures. Awarding the increased support would be governed by the free market of peer review, and modest brokering by NSF/NOAO could provide guidance to balance these resources in the name of efficiency.
2. An aggressively managed system would entail creating an entity (with suggestions ranging from an NOAO advisory panel to an openly competed "Telescope System Institute") which—working with the same augmented TSIP/mid-scale resources—would be charged with providing active guidance to maximize the utility of the existing suite of telescopes, with the goal of avoiding duplication of expensive instrumentation. Another goal would be to broker federal and nonfederal time trades to relieve the pressure for instrument duplication.

By ceding a modest amount of autonomy, all players can ultimately exploit a system that interprets broad community needs and manages the available incentives to the greatest benefit while preserving the individual, entrepreneurial, and multi-component aspects of U.S. astronomy that have made it so successful in the past century. Doubling TSIP funding and introducing a mid-scale instrumentation program provides the resources needed to develop the incentives to realize these goals. An increase in the annual TSIP to $8 million would allow several instruments in the $5 million to $10 million cost range to be funded continuously. This range is typical of instruments funded by TSIP, such as the Keck instrument MOSFIRE and the WIYN One Degree Imager. The operators of the large nonfederal facilities have indicated that they would be willing to provide time to the community up to approximately this level. Given the strong community oversubscription for TSIP time on private facilities, often exceeding time on other national facilities, the panel believes that the community should obtain whatever time can be made available in this way. Furthermore, the large nonfederal facilities have indicated that this ongoing level would provide stability and predictability that would motivate them to seek longer-term commitments to partner with the federal side. This stability was not characteristic of the program in the past decade but is essential for building the system in the next. These incentives should also extend to embrace smaller telescopes, as outlined in the ReSTAR report.

As the era is approached in which the OIR system evolves to include the new initiatives—GSMT and LSST—at its apex, the panel believes that it is critical that the supporting and complementary capabilities are maintained and improved. The panel expects that any future sales of shares of U.S.-led facilities will be made with consideration of the system so that opportunities to make those facilities available to the U.S. user community are always given consideration. At the same time, care must be taken to ensure that the TSIP/mid-scale programs provide adequate incentives for private observatories to actively partner with the national system rather than simply trading time for cash. Examples of incentives for partnering might include trading or sharing time on different instruments in the system and establishing partnerships to enable new projects

The panel recommends the following for the OIR system:

- NSF-AST should reaffirm the national observatory's role of leading and coordinating the development of the U.S. ground-based OIR system, with the understanding that all relevant groups will be involved in both planning and implementing its development.
- The system should provide ongoing guidance toward fulfilling identified needs and seizing new scientific and technical opportunities.
- The panel recommends enhancing the support of the OIR system of telescopes by (1) increasing the funds for the Telescope System Instrumentation Program (TSIP) and (2) adding support for the small-aperture telescopes into a combined effort that will advance the capabilities and science priorities of the U.S. ground-based OIR system. The OIR system includes telescopes with apertures of all sizes, whereas the TSIP was established to address the needs of large telescopes. The panel recommends an increase in the TSIP budget to ~$8 million (FY2009) annually. Additional funding for small-aperture telescopes in support of the recommendations of the NOAO ReSTAR committee (~$3 million per year) should augment the combined effort to a total of ~$11 million (FY2009) to encompass all apertures. The combined effort will serve as a mechanism for coordinating the development of the U.S. OIR system. To be effective, the funding level and funding opportunities for this effort must be consistent from year to year. Although it is possible that the total combined resources could be administered as a single program, the implementation of such a program raises difficult issues, such as formulas for the value of resources or the need to rebuild infrastructure. The panel considers the administration of two separate programs under the umbrella of system development to be a simpler alternative.
- The expanded TSIP and the mid-scale instrumentation program both provide opportunities to direct these instrumentation funds strategically toward optimizing and balancing the U.S. telescope system. The system currently functions as a collection of federal and nonfederal telescope resources that would benefit from collaborative planning and management—for example to avoid unnecessary duplication of instruments between telescopes. The panel recommends that NSF

ensure that a mechanism exists, operating in close concert with the nonfederal observatories, for the management of the U.S. telescope system.

Small Programs

The OIR Panel identified the following smaller programs and activities as essential for maintaining a balanced program of U.S. astronomy and astrophysics and for providing the technological base and the support needed for the large and medium activities mentioned above. The panel did not prioritize these activities.

Optical Interferometry

OIR interferometry enables milli-arcsecond resolution and thus can reveal signs of planet formation around young stars, study rapid rotation and magnetic fields in nearby stars, image mass loss in massive and solar-mass stars, contribute to studies of supernova Type Ia progenitors, and deliver precision astrometry for exoplanet and galactic center studies—all in aid of addressing key SFP questions.

OIR interferometry is still a young and growing discipline, currently pursued by a handful of specialized groups. Support for interferometry in the coming decade should focus on advancing the technology while making interferometry more accessible to mainstream astronomers. Doing so sets the stage for the development of a facility-level interferometer a decade hence. The NSF University Radio Observatory program provides dedicated support to university radio observatories (UROs) in exchange for community access.

The panel recommends that NSF establish an equivalent program for OIR interferometry at a level of ~$3 million per year that would primarily fund competitive proposals for partial operations support of OIR interferometers in return for public access.

This program would also fund community resources supporting the development of new public users of these facilities. The program should also support technology development for interferometry in a manner analogous to the adaptive optics development program (AODP; discussed below). If funding for this program is not available, the TSIP should be broadened to include optical interferometers.

Adaptive Optics Development Program

As discussed above in the section "Introduction and Context" the near-IR adaptive optics systems are routinely used on the world's largest telescopes, and as discussed in the section "Future Programs in OIR Astronomy," the technological framework exists to extend this capability to a GSMT. AO at visible-light wavelengths, however, is currently feasible only on very bright targets and will require significant development. AO currently has sufficient technological maturity to support first-light AO systems on GSMT. There are several areas of investment needed to move toward higher-performance AO on GSMT and existing telescopes, including development of lasers, innovative wavefront sensors and detectors, high-density deformable mirrors, and very large, adaptive secondary or curved tertiary mirrors. The AO community has prepared a technology development roadmap to reach these ambitious goals. A dedicated AO technology-development program is required for success. NSF briefly funded such a program in the past decade, but without a sustained investment, progress has almost halted; the existing NSF-AST grant programs are not well suited to strategic development of a particular technology.

The panel recommends an AO technology-development program at a level of $2 million to $3 million per year.

Next-Generation Technology

Technology development continues to drive observational progress in OIR astronomy. Beyond AO and interferometry programs, other evident technological needs—such as next-generation detectors, astro-photonics, lightweight mirrors, and fiber positioners—must be supported. These challenges can be met and prioritized through the peer-review process within the NSF ATI and MRI programs and NASA and DOE technology programs, assuming that these programs continue to receive the same fraction of the overall growing budgets. The panel endorses the current balance and mix of technology and instrumentation within NSF ATI/MRI, emphasizing that robust investment in technology development is essential to continued astronomical advances and future breakthrough instruments. The panel also notes that the Mid-Scale Projects and Instrumentation program could fund technology development in support of larger objectives that are beyond the scope of ATI. As important as the technology itself, appropriate support for the development and strengthening of the technology workforce should be part of technology efforts during the coming decade.

Archives and Archive Curation

Over the past decade, astronomy has become data-intensive. Open-access archives are commonplace for space missions and can double scientific productivity for a small fraction of the original mission cost. Archival research is a growing component of ground-based OIR astronomy, particularly for student and postdoctoral research. In Europe, ~15 percent of publications that use observations from the ESO include archival data. In the United States, archives have been established for large-scale surveys, notably the Sloan Digital Sky Survey (SDSS) and the 2-Micron

All-Sky Survey (2MASS); individual observatories (e.g., Keck, Gemini) are establishing archives for selected instruments; and the Virtual Astronomical Observatory (VAO, which is the operational phase of the National Virtual Observatory) will provide the infrastructure for cross-linking individual data archives. The widely used Virtual Solar Observatory provides the community with convenient data identification, search, and retrieval from most solar instruments.

The panel highlights the following issues:

- Archives are most effective when they provide access to higher-level data products; this requires appropriate investment in data pipelines.
- Systematically collected, well-calibrated data sets offer the greatest potential for broad scientific exploitation.
- Developing an appropriately configured data archive requires deliberate investment of resources; established archives can incorporate new data sets without requiring substantial additional resources if the appropriate reduction and analysis pipelines are available.
- There is no formal mechanism for obtaining long-term support for data archives extending beyond the operational timelines of surveys such as SDSS.

The panel recommends support for one or more data curation centers for ground-based data. The agencies should coordinate support for archive development and maintenance, so that observatories can focus their individual resources on data pipelines. Broad public access to data and data products must be a strong consideration when significant public funding is provided for private facilities.

The panel recommends an integrated ground-based astronomy data-archiving program starting at a level of ~\$2 million per year for construction and ramping down to ~\$1 million per year for ongoing operations costs.

Data Analysis Software Infrastructure

Astronomers use a variety of software tools on a diverse set of computer platforms to carry out their research. The past decade has seen significant developments using several common data analysis environments, including IRAF, IDL, and Python. Current and forthcoming OIR telescopes utilize instruments of increasing complexity and data volume, including multiobject spectrographs, integral field units, and optical interferometers. Development of dedicated software packages and specialized analysis techniques requires significant resources, but costs can be shared between observatories and users within commonly accepted environments.

The panel concluded that the provision of adequate software tools will be crucial to maximizing the scientific gains from community investment in new instrumentation for telescopes at all wavelengths.

The panel recommends that resources should be invested to plan and develop common-use infrastructures that permit cooperative development of key software to support astronomical instrumentation.

Theoretical Astrophysics and OIR Astronomy

Theoretical astrophysics plays multiple pivotal roles in all OIR research activities. Theorists generate new ideas about the universe, provide the conceptual framework for observational discovery, and seek to synthesize a coherent world view of astrophysical phenomena. They mine large astronomical data sets and interpret pre-existing observations using sophisticated algorithms and modeling. They run supercomputer simulations that incorporate key physical processes and strive to provide a realistic imitation of reality. Theory guides the conception and design of new observational programs and surveys, posing specific questions to investigate and identifying crucial predictions for observers to confirm or refute. An increasingly large number of major breakthroughs in astronomy come from observations directly motivated by theory.

As examples, the LSST science theme to probe dark energy and dark matter will require detailed cosmological simulations to interpret the data. The same is true of GSMT investigation of the early generation of galaxies and the reionization of the universe. In the realm of helioseismology, theory and observations have repeatedly challenged each other to measure and understand better the flow patterns in the Sun and the way they generate magnetic-field variability.

The SFPs identified a number of key theoretical studies associated with almost all of the scientific themes defining the frontier of research in astronomy and astrophysics during the next decade.

The panel recommends a significant increase in the funding for theoretical astrophysics.

A subset of the new funding opportunities should be devoted to a new strategic theory program (STP). The STP will fund (1) investigations aimed at generating new observatory and mission concepts that address the science questions highlighted by the Astro2010 Science Frontier Panels and (2) more focused studies on behalf of observatories and missions that are approaching or have already entered the design phase.

The panel recommends a strategic theory program at ~$3 million per year.

High-Priority Continuing Activities

To maintain and enable a balanced and strong U.S. program in OIR astronomy during the 2010-2020 decade the panel assessed the level of NSF-AST support for continuing activities and recommends continued support at the current levels in two areas:

- NSF should continue to support the National Solar Observatory (NSO) over the 2010-2020 decade to ensure that the Advanced Technology Solar Telescope (ATST) becomes fully operational. ATST operations will require a ramp-up in NSO support to supplement savings that accrue from the planned closing of current solar facilities.
- Funding for NOAO facilities should continue at approximately the 2010 level.

In addition, the panel identified two continuing activities that need enhanced support in the coming decade: OIR astronomy grants programs and the Gemini Observatory.

OIR Astronomy Grants Programs

The highest-priority activities in the continuing category are the grants programs at NSF and at NASA—they are the lifeblood of U.S. astronomy. The individual grants program at NSF is particularly critical for the ground-based OIR enterprise, as in most cases there is not a direct connection between awards of telescope time and awards of funding to support people to analyze the data. In many cases, scientific progress is actually paced by the availability of labor for the analysis.

The grants program is the primary mechanism for pursuing a diverse portfolio of smaller projects. Such projects, whether from individuals or small teams, are a highly effective investment, as one can maximize the use of competitive selection and avoid the encumbrances and management overheads of larger projects. In a time when astronomical facilities are growing in scope, the grants program is the principal counterbalance to maintain a culture for the field in which new, riskier ideas can flourish and individuals can tailor a science program based on multiple data sets. Scientific research proceeds as much from punctuated breakthroughs as from persistence, and astronomy would be at great risk if individual investigations were diminished in emphasis.

The grants program is also critical for the development of the next generation of astronomers and instrumentalists. In addition to the student and postdoctoral support and mentoring provided on project grants, funding for conferences, focused schools, and other programs can be handled through a nimble small-grants program.

The panel concluded that the individual grants programs at all agencies are as critical as ever.

The panel recommends that the NSF-AST grants program (AAG) should be increased above the rate of inflation by ~$40 million over the decade to enable the community to utilize the scientific capabilities of the new projects and enhanced OIR system.

The Gemini Observatory

The second high-priority continuing activity for OIR astronomy is the Gemini Observatory, which provides on a competitive basis the largest number of nights on 8-m telescopes that are open to the general U.S. community. The Gemini Observatory, properly instrumented, can address many of the science questions identified by the Science Frontiers Panels and provide ground-based spectroscopic follow-up for LSST discoveries and support for space missions.

The Gemini Observatory, however, is a partnership of seven nations, with the 51.12 percent U.S. share representing the bulk of community open access to large apertures. The complex management structure that has evolved from the international nature of the Gemini partnership prevents the U.S. National Gemini Office from serving as an advocate for U.S. interests at a level commensurate with its partnership share. The ALTAIR report found community dissatisfaction with its current instrumentation (dominated by narrow-field infrared capabilities), its operation model emphasizing queue observing, and the lack of transparency of its governance.

The Gemini Observatory should be viewed as an integral component of the U.S. telescope system. Specifically, Gemini would be more effective with a stronger connection between observatory management and the U.S. community. The renegotiation of the international Gemini agreement in 2012 will provide an opportunity to simplify management to achieve these goals while still respecting the sensitivities of the international partners. In addition, the NSF-AST senior review and internal Gemini reviews, have noted that Gemini operations costs are high compared to those of other national facilities and the nonfederal observatories. A Gemini Observatory better integrated with the U.S. system and its resources could streamline operations. Although the panel has no means to estimate the savings that could be realized, decreased operating costs could provide a means to augment the U.S. Gemini share if the opportunity arose and if the Gemini Observatory were seen as more representative of U.S. interests.

Continued support for Gemini operations into the next decade is a priority in maintaining a balanced OIR research program. The panel recommends that the following actions be taken to enhance Gemini's contribution to the U.S. OIR system:

1. The governance of the international Gemini Observatory should be restructured, in collaboration with all partners, to improve the responsiveness and accountability of the observatory to the goals and concerns of all its national user communities. Project share should be an important factor in the development of strategic plans. U.S. participants in governance (for example, Gemini Board representatives) should be selected with the goal of accurately representing the broad interests of the U.S. OIR system.

2. NSF and its international partners should convene an independent panel to solicit detailed advice from management experts, existing observatories, instrument builders, experienced Gemini users, and other parties on the design of the new Gemini governance and management structure.

3. As part of the restructuring negotiations, the United States should attempt to secure an additional fraction of the Gemini Observatory, including a proportional increase in the U.S. leadership role. The funding allocated for any augmentation in the U.S. share should be at most 10 percent of FY2010 U.S. Gemini spending.

4. The United States should also seek improvements to the efficiency of Gemini operations. Efficiencies from streamlining Gemini operations, possibly achieved through a re-forming of the national observatory to include NOAO and Gemini under a single operations team, should be applied to compensate for the loss of the United Kingdom from the Gemini partnership, thereby increasing the U.S. share.

5. The panel recommends that the United States support the development of medium-scale, general-purpose Gemini instrumentation and upgrades at a steady level of about 10 percent of the U.S. share of operations costs. U.S. support for new large Gemini instruments (>~$20 million) should be competed against proposals for other instruments in the recommended mid-scale instrumentation program—a program aimed at meeting the needs of the overall U.S. system discussed elsewhere in this panel report.

Additional Activity

During its work the OIR Panel identified an additional activity with significant potential importance: optical and infrared astronomy in Antarctica.

New sites for astronomical observations in Antarctica that are located away from the U.S. facilities at the South Pole hold significant promise. The unique climate and weather of Antarctica offer extraordinary opportunities for OIR observations. While submillimeter observations have been executed successfully from the South Pole, the best sites for OIR work are being explored at high-altitude plateau sites such as Dome C. Promising results for unusually dark IR skies and for excellent seeing above a low-lying ground layer are being investigated by international programs based in Europe, Australia, and China. For some applications, the potential exists to rival space and airborne platforms for science in long-wave IR astronomy and OIR interferometry. Despite the rich scientific potential of these sites, no significant U.S. site development plans exist. NSF is the lead agency for scientific work in Antarctica, but NSF-AST and the NSF Office of Polar Programs need to work collaboratively to facilitate the best overall scientific program for the continent, including improving the site assessments and potentially establishing a presence on the plateau sites, either through partnership or on our own.

The panel recommends internal coordination at NSF to encourage U.S. investigators to participate in the development of the new, and potentially revolutionary, arena for OIR astronomy represented by excellent observing sites in Antarctica.

OIR Astronomy and Demographics

The panel's charge emphasizes the need to recommend a balanced program for OIR for the next decade. The programs highlighted above strike that balance in terms of the division between large, medium, and small-scale projects, but most importantly in balancing the access to these projects and facilities across the complete astronomical demographic. Central to the panel's recommendations is an endorsement and strengthening of the U.S. telescope system providing widespread access to the full range of telescope apertures and capabilities, as well as ensuring public access to GSMTs and LSST, the large-project recommendations of this panel, which are integral components of this system. This level playing field, particularly in the context of a vigorous AAG program, invites and enables the growth of those with budding astronomical interests. Sky surveys, such as LSST, create databases as accessible to high-school students as they are to professional researchers at well-endowed institutions. The Astro2010 survey's Infrastructure Study Group for demographics noted the severe underrepresentation of minorities in astronomy. A strong, diverse U.S. system—with investment at all aperture sizes and with emphasis on surveys and data archiving—provides scientific opportunities to engage entry-level students in research, and to retain them in the career-development pipeline, at virtually any institution during their undergraduate years. The demographics group also noted that training of the next generation of instrumentalists may be at risk, particularly with the dominance of large-scale projects limiting opportunities for hands-on end-to-end training on projects of modest scale. The expanded TSIP, and particularly ReSTAR, instrumentation opportunities inherent in the recommendations above directly address this need and ensure the continued existence of innovative and flexible instrumentation programs at federal and nonfederal institutions alike.

Interagency Collaboration

As the principal agency supporting ground-based OIR astronomy in the United States, NSF should continue to lead the federal efforts for the OIR system as described in this section and work with DOE as it becomes increasingly involved in system activities. In addition, NASA should be engaged as a full collaborator in supporting the construction, operation, and analysis activities of ground-based OIR facilities where such joint interchange will advance the science and programmatic objectives of both agencies.

RECOMMENDED PRIORITIES AND PLAN FOR THE NEXT DECADE

This section presents (1) a summary of the conclusions for the activities described above in “Future Programs in OIR Astronomy”; (2) the recommended priorities; and (3) a prioritized implementation approach for the next decade for the budget guidelines provided by Astro2010.

Conclusions and Recommendations for Large Programs

Having considered proposals from the research community for new large facilities, the OIR Panel reached the following conclusions for large programs:

- The science cases for a 25- to 30-m Giant Segmented Mirror Telescope and for the proposed Large Synoptic Survey Telescope are even stronger today than they were a decade ago.
- Based on the relative overall scientific merits of GSMT and LSST, the panel ranks GSMT higher scientifically than LSST, given the sensitivity and resolution of GSMT.
- Both GSMT and LSST are technologically ready to enter their construction phases in the first half of the 2010-2020 decade.
- The LSST project is in an advanced state and ready for immediate entry into NSF’s MREFC line for the support of construction. In addition, the role of DOE in the fabrication of the LSST camera system is well defined and ready for adoption.
- LSST has complementary strengths in areal coverage and temporal sensitivity to GSMT, with its own distinct discovery potential. Indeed, GSMT is unlikely to achieve its full scientific potential without the synoptic surveys of LSST. Consequently LSST plays a crucial role in the panel’s overall strategy.
- GSMT is a versatile observatory that will push back today’s limits in imaging and spectroscopy to open up new possibilities for the most important scientific problems identified in the Astro2010 survey. This exceptionally broad and powerful ability over the whole range of astrophysical frontiers is the compelling argument for building GSMT.
- Given the development schedules for GSMT and in order to ensure the best science return for the U.S. public investment, it is both vital and urgent that NSF identify one U.S. project for continued support to prepare for its entry into the MREFC process.

Based on these conclusions, the panel recommends the following ordered priorities for the implementation of the major initiatives that form part of its recommended OIR research program for the 2010-2020 decade:

1. Given the panel's top ranking of the Giant Segmented Mirror Telescope based on its scientific merit, the panel recommends that the National Science Foundation establish a process to select which one of the two U.S.-led GSMT concepts it will continue to support in its preparation for entry as soon as practicable into the MREFC line. This selection process should be completed within 1 year from the release of this panel report.

2. The panel recommends that NSF and DOE commit as soon as possible, but no later than 1 year from the release of this report, to supporting the construction of the Large Synoptic Survey Telescope. Because it will be several years before either GSMT project could reach the stage in the MREFC process that LSST occupies today, the panel recommends that LSST should precede GSMT into the MREFC approval process. The LSST construction should start no later than 2014 in order to maintain the project's momentum, capture existing expertise, and provide critical synergy with GSMT.

3. The panel recommends that NSF, following the completion of the necessary reviews, should commit to supporting the construction of its selected GSMT through the MREFC line at an equivalent of a 25 percent share of the total construction cost, thereby securing a significant public partnership role in one of the GSMT projects.

4. The panel recommends that in the longer term NSF should pursue the ultimate goal of a 50 percent public interest in GSMT capability, as articulated in the 2001 decadal survey (Astronomy and Astrophysics in the New Millennium). Reaching this goal will require (most likely in the decade 2021-2030) supporting one or both of the U.S.-led GSMT projects at a cost equivalent to an additional 25 percent GSMT interest for the federal government. The panel does not prescribe whether NSF's long-term investment should be made through shared operations costs or through instrument development. Neither does the panel prescribe whether the additional investment should be made in the selected MREFC-supported GSMT in which a 25 percent partnership role is proposed already for the federal government. But the panel does recommend that, in the long run, additional support should be provided with the goal of obtaining telescope access for the U.S. community corresponding to total public access of 50 percent of a GSMT.

The panel has chosen not to prescribe in this report a process by which NSF could choose between the GSMT projects for a federal investment. However, the panel would expect such a process to include the setting and application of criteria by a committee of U.S. researchers (with the possible participation of non-OIR astronomers and researchers from other fields cognizant of the challenges of implementing billion-dollar-class facilities), as well as a thorough and independent assessment of cost and technical and schedule risk.

Conclusions and Recommendations for Medium and Small Programs

The panel concluded that mechanisms are limited or non-existent for funding major instruments and projects with costs above the funding guidelines of standard programs such as the NSF AST or MRI programs but below the level of the NSF-wide MREFC.

- The panel recommends as its highest-priority medium activity a new medium-scale instrumentation program in NSF's AST that supports projects with costs between those of standard grant funding and those for the MREFC line. To foster a balanced set of resources for the astronomical community, this program should be open to proposals to build (1) instruments for existing telescopes and (2) new telescopes across all ground-based astronomical activities, including solar astronomy and radio astronomy. The program should be designed and executed within the context of, and to maximize the achievement of, science priorities of the ground-based OIR system. Proposals to the medium-scale instrumentation program should be peer reviewed. OIR examples of activities that could be proposed for the program include massively multiplexed optical/near-IR spectrographs, adaptive optics systems for existing telescopes, and solar initiatives following on from the Advanced Technology Solar Telescope. The panel recommends funding this program at a level of approximately $20 million annually.

The panel concluded that the Telescope System Instrumentation Program represents a successful model for enhancing federal and nonfederal telescope facilities while providing expanded public access to nonfederal facilities. The panel concluded further that TSIP, given sufficient resources, has the potential to help configure and balance the overall U.S. OIR system to maximize the efficiency of observing resources while providing instruments that will enable observations at the limits of current technology.

- As its second-highest-priority medium activity, the panel recommends enhancing the support of the OIR system of telescopes by (1) increasing the funds for the Telescope System Instrumentation Program and (2) adding support for the small-aperture telescopes into a combined effort that will advance the capabilities and science priorities of the U.S. ground-based OIR system. The OIR system includes telescopes with apertures of all sizes, whereas the TSIP was established to address the needs of large telescopes. The panel recommends an increase in the TSIP budget to approximately $8 million (FY2009) annually. Additional funding for small-aperture telescopes in support of the recommendations of the National Optical Astronomy Observatory (NOAO) Renewing Small Telescopes for Astronomical Research (ReSTAR) committee (approximately $3 million per year) should augment the combined effort to a total of approximately $11 million (FY2009) to encompass all apertures. The combined effort will serve as a mechanism for coordinating the development of the OIR system. To be effective, the funding level and funding opportunities for this effort must be consistent from year to year. Although it is possible that the total combined resources could be administered as a single program, the implementation of such a program raises difficult issues, such as formulas for the value of resources or the need to rebuild infrastructure. The panel considers the administration of two separate programs under the umbrella of "system development" to be a simpler alternative. The expanded TSIP and the mid-scale instrumentation program both provide opportunities to direct these instrumentation funds strategically toward optimizing and balancing the U.S. telescope system.

- The U.S. system of OIR telescopes currently functions as a collection of federal and nonfederal telescope resources that would benefit from collaborative planning and management—for example, to avoid unnecessary instrument duplication between telescopes. The panel recommends that NSF ensures that a mechanism exists, operating in close concert with the nonfederal observatories, for the management of the U.S. telescope system. The panel recommends that a high priority be given to renewing the system of ground-based OIR facilities, requiring a new strategic plan and a broadly accepted process for its implementation.

The panel concluded that initiating a tactical set of small targeted programs ($1 million to $3 million per year each) would greatly benefit ground-based OIR science in the coming decade and provide critical support for some of the medium and large programs.

- The panel recommends the small programs in the following, unprioritized list:

—An adaptive optics technology development program at the $2 million to $3 million per year level;

—An interferometry operations and development program at a level of approximately $3 million per year;

—An integrated ground-based-astronomy data-archiving program starting at a level of approximately $2 million per year and ramping down to approximately $1 million per year; and

—A "strategic theory" program at the level of approximately $3 million per year.

Recommendations for Continuing Activities

The panel makes the following recommendations for continuing activities:

- NSF should continue to support the National Solar Observatory (NSO) over the 2010-2020 decade to ensure that the Advanced Technology Solar Telescope

(ATST) becomes fully operational. ATST operations will require a ramp-up in NSO support to supplement savings that accrue from the planned closing of current solar facilities.

- Funding for NOAO facilities should continue at approximately the FY2010 level.
- The governance of the international Gemini Observatory should be restructured, in collaboration with all partners, to improve the responsiveness and accountability of the observatory to the goals and concerns of all its national user communities. As part of the restructuring negotiations, the United States should attempt to secure an additional fraction of the Gemini Observatory, including a proportional increase in the U.S. leadership role. The funding allocated for any augmentation in the U.S. share should be at most 10 percent of FY2010 U.S. Gemini spending. The United States should also seek improvements to the efficiency of Gemini operations. Efficiencies from streamlining Gemini operations, possibly achieved through a reforming of the U.S. national observatory to include NOAO and Gemini under a single operations team, should be applied to compensate for the loss of the United Kingdom from the Gemini partnership, thereby increasing the U.S. share. The United States should support the development of medium-scale, general-purpose Gemini instrumentation and upgrades at a steady level of about 10 percent of the U.S. share of operations costs. U.S. support for new large Gemini instruments (greater than approximately $20 million) should be competed against proposals for other instruments in the recommended mid-scale instrumentation program—a program aimed at meeting the needs of the overall U.S. OIR system discussed elsewhere in this panel report.
- The NSF-AST grants program (Astronomy and Astrophysics Research Grants [AAG]) should be increased above the rate of inflation by approximately $40 million over the decade to enable the community to utilize the scientific capabilities of the new projects and enhanced OIR system.
- NSF-AST should work closely with the NSF Office of Polar Programs to explore the potential for exploiting the unique characteristics of the promising Antarctic sites.

Implementation Plan

The supplement at the end of this report presents an implementation plan for the panel's recommendations for the projects and activities summarized in this section. The plan is consistent with the historical funding pattern for NSF-AST and fits within the budgetary planning guideline provided by the Astro2010 Survey Committee.

SUMMARY COMMENTS

In this report the panel presents a balanced program for U.S. OIR astronomy that is consistent with historical federal funding of the field. More importantly, the program recommended will enable astronomers to answer the compelling science questions that can now be formulated, and it will open new windows of discovery for the future. The two large projects recommended, GSMT and LSST, will each advance fundamentally important observing capabilities by one or two orders of magnitude beyond the present. Together, they address the core of astronomy and astrophysics, including 18 of the 20 science questions and all 5 of the discovery areas identified by the Astro2010 Science Frontiers Panels. The OIR Panel's two medium-scale recommendations address clear gaps in the current approach and will support a range of specialized instruments and telescopes that will provide a variety of tools to address central science questions from many angles and to strengthen the overall U.S. system of OIR telescopes. The panel values the small-scale programs because they advance OIR science and provide essential support for the medium- and large-scale activities. OIR astronomy has an exceptionally complex administrative structure. The panel believes that this can be turned to advantage through an increased emphasis on partnerships that include NSF, DOE, and NASA; U.S. federal, state, and private institutions; and international partners. The scale of the new large projects demands cooperation: if this task is approached with imagination and good will, we can gain from the key capabilities that each partner brings.

The revolution in human understanding that began with Galileo's telescope 400 years ago has not slowed down or lost its momentum—in fact it is accelerating. This panel has identified the most promising areas for the United States to invest in right at the center of a thriving scientific adventure. It looks forward to the rich flow of science that will come from implementing these recommendations.

SUPPLEMENT: IMPLEMENTING THE PANEL'S RECOMMENDATIONS WITHIN THE PROJECTED FUNDING CONTEXT

The NSF Division of Astronomical Sciences (NSF-AST) is the federal steward and principal source of federal funding for ground-based OIR astronomy in the United States. NASA provides some support for ground-based efforts in connection with its space missions (e.g., partial support of Keck, IRTF, and interferometry). DOE is becoming an increasingly important source of funding, as the scientific interests of high-energy physics and cosmology become deeply interconnected. Here the panel concentrates on NSF funding because of its central importance to ground-based funding and because it is the primary source for most of the projects and efforts under consideration by the panel.

As stated above in the section "Interagency Collaboration" at the end of the section "Future Programs in OIR Astronomy," NSF should continue to lead the federal efforts for the OIR system and work with DOE as it becomes increasingly involved in system activities. In addition, NASA should be engaged as a full collaborator supporting construction, operation, and analysis activities of ground-based OIR facilities where such joint interchange will advance the science and programmatic objectives of both agencies.

Total Costs and Schedules for Large Projects

The technical, cost, and schedule risks for the large projects are described in "Large Programs" under "Future Programs in OIR Astronomy." For ease of reference, Table 7.5 provides a summary of the construction and operations costs and construction schedules provided by the individual projects and the results of the

TABLE 7.5 Summary of Estimates of Construction and Operations Costs (FY2009 dollars) and Schedule Estimates for OIR Large Projects

	Project-Identified Values					CATE Values[a]		
Project	Cost Appraisal	Reserve	Total	Operations Costs per Year	Construction Schedule (months)	Construction Schedule (months)	Cost Analysis	Cost Sensitivity 70% value
LSST	$354M	$102M	$456M	$41M	97	112	$398M[b]	n/a
GMT	$563M	$113M	$676M	$37M[c]	128	178	n/a	$1.1B[d]
TMT	$760M	$227M	$987M	$54M[e]	168	191-239[f]	n/a	$1.4B[d]

[a] The calibration data available for ground-based projects were less than those available for space-based projects; this was a limiting factor in the independent cost appraisal and technical evaluation (CATE) assessment, particularly for the two GSMT projects. The methodology is discussed in more detail in the section "Opportunities in OIR Science" of this panel report. It is important to note that this is a cost-sensitivity analysis rather than a cost appraisal, showing the effects of cost variations in key subtasks.

[b] For LSST, the cost appraisal was based on applying a parametric model to a level-3 work breakdown structure submitted by project personnel, including optics and facility fabrication. The cost appraisal did not examine instrument or data-handling costs, and this estimate does not include operation costs. The independent CATE assessment produced a single cost appraisal for LSST, and so no distribution is noted here.

[c] Includes instrument and facility upgrades of $16 million per year.

[d] For the two GSMTs, cost-sensitivity analysis was restricted primarily to the costs of manufacturing the optics (discussed in the section "Opportunities in OIR Science") and the science instruments (assumed to increase by ~100 percent based on the contractor's experience with space missions). The independent CATE assessment produced a cost-sensitivity distribution for the GSMTs, and costs at 70 percent confidence levels are given here.

[e] Includes instrument upgrades of $20 million per year.

[f] Range of values reflects the independent contractor's assessment of schedule uncertainty for fabrication of the primary mirror segments.

assessment of the independent contractor. All projects explained the methodology and basis for their estimates and the basis for a reserve (contingency) amount that they included in their total. The independent CATE process utilized the information provided by the projects and its own methods and data to develop independent assessments for costs, risks, and schedules.

The available calibration data for ground-based projects were less than those available for space-based projects; this was a limiting factor in the independent CATE assessment, particularly for the two GSMT projects. The independent CATE assessment also estimated schedule risk in all three projects, and these values are listed in Table 7.5.

The panel concluded that both GSMT projects involve a substantial risk to cost and/or schedule that need to be carefully examined and considered as NSF executes a process to select a potential partner for public investment. It is likely that such a commitment would have to be sustained for more than a decade prior to the completion of commissioning and initial scientific utilization of a GSMT concept by the community. The panel notes and endorses the NSF procedure for obtaining detailed cost and schedule estimates as part of its Major Research Equipment and Facilities Construction (MREFC) process.

Proposed NSF Share of Large Project Costs

It is important to note that none of the three large projects recommended by the OIR Panel is requesting full support from NSF. The LSST project has developed a proposed cost-sharing plan involving NSF, DOE, and other partners. GMT and TMT also have other partners, and the panel assumed NSF participation at the 25 percent level. Table 7.6 shows the proposed cost to NSF for the large projects, based on the total values shown in Table 7.5.

Large Projects and MREFC Opportunities

The construction costs of the GSMT and LSST projects recommended by the panel are of a scale that an NSF investment in construction will come from the agency's MREFC program. This program supports the construction of new major

TABLE 7.6 Proposed NSF Share of Costs of Construction of Large Projects (FY2009 dollars)

Project	NSF Share	Notes
LSST	$261 to $298M	Balance from DOE and other sources
GMT	$169 to $250M	25% share for NSF, balance from other partners
TMT	$247 to $325M	25% share for NSF, balance from other partners

equipment and facilities with costs that exceed the capabilities of a single NSF division and require investments on an NSF-wide basis.

Both GSMT and LSST have advanced to the stage at which they can be considered for construction and operations funding during the period covered by Astro2010. LSST is already placed in the final preparatory stages of the MREFC approval process.

While the MREFC line offers the primary opportunity for implementing federal support for these two large projects in the coming decade, there are significant factors beyond the requirements of MREFC that enter into the planning process. First, both GSMT projects will involve funding and partnering outside of NSF. GMT and TMT are established as private, nonfederal projects with international components. Second, LSST is to be a partnership with DOE and also utilizes private, nonfederal resources. Third, MREFC does not cover operating costs, which will be significant over the lifetime of the projects and will have to be covered by NSF-AST and the partners in the projects. The plans presented below account for both proposed MREFC funding and the projected operating costs to be borne by NSF-AST. The assumed total MREFC funding for GSMT and LSST combined is consistent with the inflation-adjusted MREFC funding for ALMA.

LSST Funding

The panel notes that the LSST project has advanced to the final design phase, which will take it to the point of readiness for construction and also prepare it for the NSF critical design review. As noted above, this panel report recommends that NSF should commit to supporting the construction of LSST, with construction starting no later than 2014 in order to maintain the project's momentum, capture existing expertise, and achieve critical synergy with GSMT.

The level of NSF support requested by the project would be $298.5 million. The Department of Energy would provide support in the form of construction of the LSST camera system, with some support coming from other sources. LSST operations costs (projected at $40 million per year) would be split evenly among NSF, DOE, and an additional partner.[6]

GSMT Funding

The panel reaffirms the goal of the previous decadal survey, *Astronomy and Astrophysics in the New Millennium*, of a 50 percent public share of a GSMT. However, the panel recognizes that the projected availability of public funding will re-

---

[6] The additional partner was not identified to the panel in the documentation provided to it by the LSST project.

quire a distributed investment over more than the present decade to reach that goal. The panel believes that a 50 percent public share of a GSMT is in the best interest of the nation, maximizing the scientific yield of the OIR system and GSMT itself. Below, the panel outlines one possible scenario to achieve this objective, although of course other approaches exist.

Elsewhere in this report, the panel recommends that the U.S. government become a full partner in a GSMT project at the 25 percent level as soon as possible and that NSF should establish a process to choose which of the two U.S.-led GSMT concepts will receive a founding-member public investment.

After selection, negotiations to establish U.S. participation can begin and the particular mechanisms and commitments worked out. The investment made by NSF should ensure a significant and ongoing public role in the governance and operation of the chosen telescope. The panel's preferred mechanism is to support this participation through the MREFC process; 25 percent of a GSMT (approximately $250 million in FY2009 dollars) is consistent with a typical MREFC funding envelope, and significantly below U.S. ALMA participation. NSF-AST would also support operations at a level consistent with its share (25 percent of total costs.) Alternatively, the balance between operations and construction funding could be negotiated as part of the initial buy-in.

In the case of 7 percent per year AST budget growth over the decade,[7] additional participation, most likely in the other GSMT project, could be done through the NSF-AST budget process, providing support for instrumentation and possibly construction. This could be done either formally through a negotiated participation or through a substantially expanded TSIP-like program. The goal would be a 25 percent share in observing time in the other GSMT, but without a formal governance role.

Funding Profile for the 2010-2020 Decade

Table 7.7 and Figure 7.19 show (as an existence proof) a representative overall budget profile for this panel's OIR recommendations that follows guidelines provided by the Program Subcommittee of the Astro2010 Survey Committee. The guidelines assume 3.7 percent per year growth above inflation for the NSF-AST budget for the years beyond the currently known or projected values and also take into account the continuing costs of existing AST programs such as grants, national centers, instrumentation, and special projects. The OIR plan described in Table 7.7 and Figure 7.19 is for the continuing and new funds for OIR activities that are projected to become available during the decade.

---

[7] This level of increase is consistent with NSF-AST keeping pace with a doubling of the NSF budget.

TABLE 7.7 Budget Profile for Panel's OIR Recommendations

Activities	Funds for the Decade (FY2009 dollars)
Large Activities	Requested from MREFC; see Table 7.6
Medium Activities (in priority order)	
Mid-scale instruments and projects	$190M
TSIP and ReSTAR	$106M ($76M + $30M)
Small Activities (alphabetical order)	
Adaptive optics development program	$18M
Ground-based archive	$14M
Interferometry operations and technology development	$27M
Strategic theory	$26M
Continuing Activities	
NSF-AST research grants program (AAG)	$490M (includes $40M increase)
Gemini operations and instrumentation	$250M (increases U.S. funding by 10%)
NOAO operations	$260M
NSO ATST increase	$117M (includes $26M increase to support LSST)
Large Projects Operations and Additional GSMT Share (in priority order)	
NSF share of LSST and GSMT operations resulting from MREFC investments	$83M ($14M/year and $8M/year)
Additional GSMT share	$97M (construction or instrumentation support to acquire additional share)

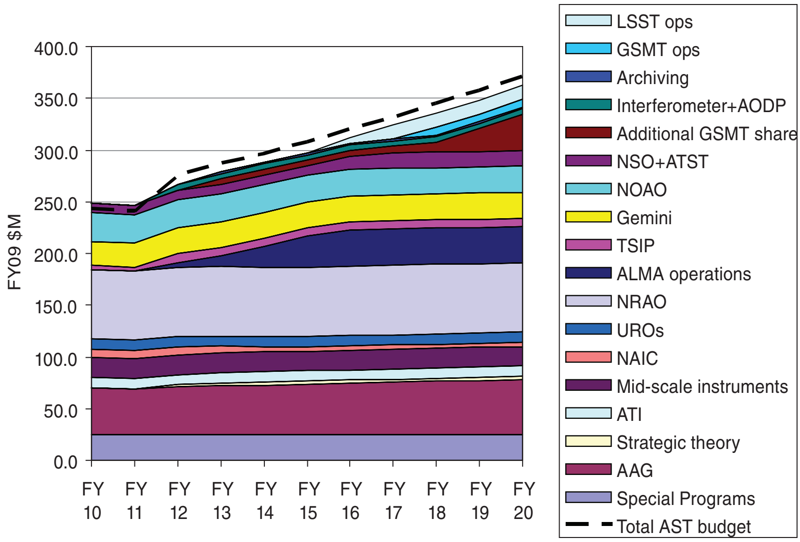


FIGURE 7.19 Chart showing straw-man scenario for overall NSF-AST spending, assuming 3.7 percent real growth per year.

8

Report of the Panel on Particle Astrophysics and Gravitation

SUMMARY

The fertile scientific ground at the intersection of astrophysics, gravitation, and particle physics addresses some of the most fundamental questions in the physical sciences. For example, the unexplained acceleration of the expanding universe leads scientists to question their understanding of cosmology. There may be an as-yet-uncharacterized component of the mass-energy that drives the dynamics of the universe—a cosmological constant or a new type of field—called "dark energy." Or, gravity may be described not by Einstein's general theory of relativity but rather by a different theory altogether. Solving these puzzles will require new astrophysical observations.

Another unsolved mystery is the origin of the initial conditions at the beginning of the universe, the first density fluctuations that grew into the structures seen today. There is evidence that these initial conditions were set down during the period of inflation in the very early universe. That leaves open the question, What caused inflation? Again, gravity may provide the clue. Measurements of the stochastic background of gravitational waves that formed at the same time as the initial density perturbations provide an important tool that might probe the inflationary period. Connecting physics and astronomy, the initial density perturbations set the stage for structure formation: How and when did the first structures form in the universe? Observations of gravitational waves from black hole mergers at high redshift will provide unique information about this era, complementing other probes.

Another puzzle is that of the laws of nature in the environments that harbor the most extreme gravitational fields. Supermassive black holes inhabit the centers of galaxies, and they somehow—following the laws of gravity—generate tremendous outflows of energetic particles and radiation, twisting magnetic fields into concentrated pockets of magnetism. Scientists cannot help but strive to understand these extreme environments and to take advantage of them as laboratories to put theories of gravitation to their most demanding tests.

Gravitation is a unifying theme in nearly all of today's most pressing astrophysics issues. Much of the precursor work of the past decade was motivated by the scientific imperative of understanding gravitation, and an intense period of technology development to build the necessary tools is reaching fruition and must now be exploited. Scientists now have ground-based laser interferometric detectors that are on a path to reaching the level of sensitivity at which the detection of gravitational waves is virtually assured. They have a plan and a design for a network of spacecraft that will measure long-wavelength gravitational waves where astrophysical sources are predicted to be the most abundant. They have developed high-precision techniques for pulsar observations that are promising probes of the gravitational waves associated with inflation and with supermassive black holes. Recognizing these developments, the Panel on Particle Astrophysics and Gravitation presents a program of gravitational-wave astrophysics that will bring the investments in technology to fruition. The panel recommends that the Laser Interferometer Space Antenna (LISA) be given a new start immediately; that ground-based-laser gravitational-wave detectors continue their ongoing program of operation, upgrade, and further operation; and that the detection of gravitational waves through the timing of millisecond pulsars move forward. To complement the use of gravitational waves as a beacon for astrophysics and fundamental physics, the panel recommends that the theoretical foundations of gravity themselves be put to stringent test, when such tests can be carried out in a cost-effective manner. These tests of gravitation will be provided by LISA's observations of strong-field astrophysical systems, by electromagnetic surveys to characterize dark energy (considered by other Astro2010 Program Prioritization Panels), by precise monitoring of the dynamics of the Earth-Moon system, and by controlled tests of gravity theories done in the nearly noise-free environment of space. The time has arrived to explore the still-unknown regions of the universe with the new tool of gravitation.

Understanding the nature of three-quarters of the universe is an important goal, but the other one-quarter, which is known to be some form of matter, must not be overlooked. Scientists have identified one-sixth of this matter: it is in the form of stars, galaxies, and gas that have been studied extensively for centuries. However, the nature of the other five-sixths is still a mystery. Evidence exists that the unknown part is *not* made up of familiar materials but rather must be a diffuse substance that interacts only weakly with ordinary matter. The leading candidates for this so-called dark matter are new families of particles predicted by some theories of fundamental particle physics. There are three complementary approaches to attacking the dark-matter problem: direct detection in the laboratory, indirect detection by way of astronomical observations, and searches for candidate particles in human-made high-energy particle accelerators. The panel's recommendations concern only indirect detection by astrophysics, although all three approaches will be important in ultimately resolving this mystery.

The indirect detection of dark matter involves searching not for the dark-matter particles themselves, but rather for products of the annihilation or decay of dark-matter particles. These may be gamma rays, cosmic rays, or neutrinos. The sources will be places in the cosmos where scientists believe that dark matter concentrates, such as in the gravitational potential wells of galaxies. Therefore, the panel recommends a program of gamma-ray and particle searches for dark matter.

The field of high-energy and very-high-energy particle astrophysics has blossomed in the past decade, with a wealth of results from spaceborne and ground-based gamma-ray telescopes and cosmic-ray detectors, and it is hoped that similar exciting results will come soon from neutrino telescopes. These instruments provide unique views of astronomical sources, exploring the extreme environments that give rise to particle acceleration near, for example, supermassive black holes and compact binary systems. The panel recommends continued involvement in high-energy particle astrophysics, with particular investment in new gamma-ray telescopes that will provide a much deeper and clearer view of the high-energy universe, as well as a better understanding of the astrophysical environment necessary to disentangle the dark-matter signatures from natural backgrounds. The panel's highest-priority recommendation for ground-based instrumentation is significant U.S. involvement in a large international telescope array that will exploit the expertise gained in the past decade in atmospheric Čerenkov detection of gamma rays. Such a telescope array is expected to be an order-of-magnitude more sensitive than existing telescopes, and it would for the first time have the sensitivity to detect, in other galaxies, dark-matter features predicted by plausible models.

The panel also recommends a broad program for particle detectors to be flown above the atmosphere, making use of the cost-effective platforms provided by balloons and small satellites. Major developments in large ground-based detectors for neutrinos are in progress already. These programs are an important component of dark-matter and astrophysical particle characterization and should be continued, along with the research and development that will improve the sensitivity of neutrino detectors in decades to come.

The above recommendations are possible only because there is now available a suite of new instruments that have recently achieved technical readiness. In the program areas that the panel considered, a significant component of the technology development has been done outside the United States. To maintain the nation's ability to participate in research in astrophysics in the future, the panel recommends that the technology-development programs of all three funding agencies relevant to particle astrophysics and gravitation be augmented. To enable missions to test theories of gravitation and to carry out timely and cost-effective experiments in particle astrophysics, gravitation, and other areas of astrophysics, the panel recommends an augmentation of NASA's Explorer program. It is expected that such missions will compete in a forum of peer review. To enable particle-detection experiments, the panel recommends an augmentation of NASA's balloon program to support ultralong-duration ballooning.

Finally, on an even more fundamental level, the panel recognizes that the ultimate goal of all these activities is the advancement of knowledge, for the achievement of which the culminating activities are the interpretation and dissemination of results, and which in turn lead to new frameworks for subsequent exploration. Therefore, the panel supports a strong base program in all areas of astronomy and astrophysics. This base program must include theory as one of its components.

The program in particle astrophysics and gravitation that this panel recommends includes missions, projects, and activities that will result in new tools for attacking many of the outstanding problems of astronomy and astrophysics, both in this decade and in the future. The recommended program will launch the new discipline of gravitational-wave astrophysics. It will develop new detectors for cosmic rays, gamma rays, and neutrinos that—working in tandem with gravitational-wave and longer-wavelength electromagnetic detectors—will enable multi-messenger astrophysics. It will confront theories of gravitation with new data, in the context of understanding the strong fields around black holes and the nature of dark energy on cosmological scales. It will seek to identify the elusive dark matter. It will elucidate the remarkable dynamics of black holes and their fields and outflows. All in the astronomy and astrophysics community look forward to the discoveries of the next decade.

THE SCIENCE CASE

The scope of this panel's deliberations is defined by those areas of astronomy and astrophysics that use experimental and observational techniques at the intersection of astronomy and physics. In particular, the panel addresses the windows on the universe offered by high-energy particles (including gamma rays) and gravitational waves. Each provides a radically new and, especially in the cases of gravitational waves and high-energy neutrinos, so-far-unexplored window on astrophysical objects and processes. The panel also considers tests of general relativity and other theories of gravity. The scientific questions explored address regimes of fundamental physics not accessible to laboratory experiment, and they span the largest cosmological scales. They also probe the highest energies accessible to science, and they probe the smallest scales that may characterize a unified theory of the basic interactions. The answers to these questions reveal clues to the most fundamental origins—of space and time, of matter and energy, of the universe itself. This quest will stretch our best tools and most cherished ideas about physics; it will require robust investment in both theory and observational infrastructure in the coming decade; and it will allow us literally to see the universe in new ways.

Gravitational-Wave Astrophysics

One of the remarkable predictions of Einstein's theory of general relativity—and, more generally, any theory that describes non-instantaneous gravitational forces—is the existence of gravitational waves. The direct observation of gravitational waves will provide much more than just another confirmation of Einstein's theory: it will open the study of the astrophysical sources in an entirely new way (Figure 8.1). When gravitational-wave astrophysics progresses to the point of

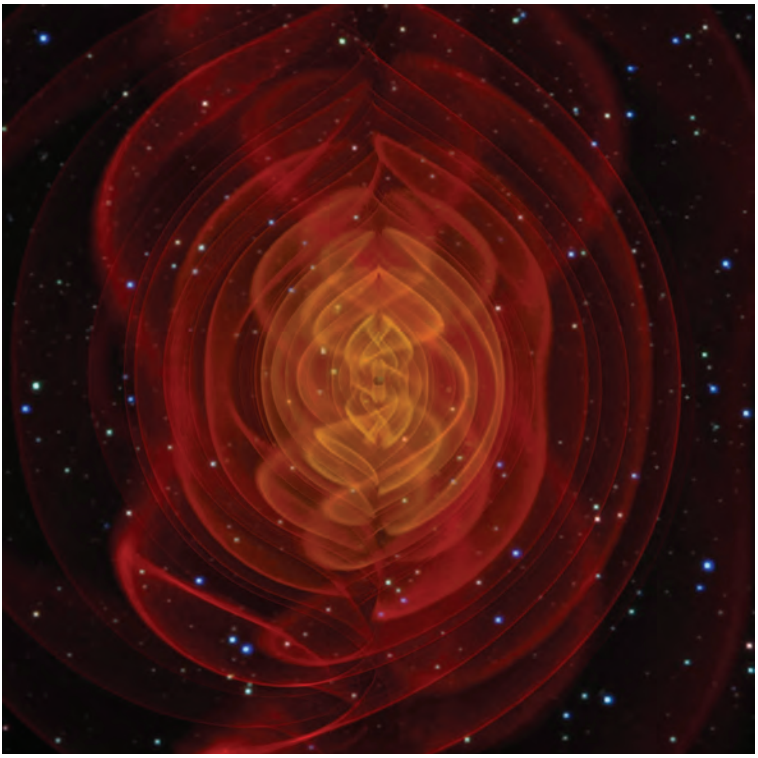


FIGURE 8.1 Gravitational waves computed from a numerical simulation of the merger of two black holes. The orange and red contours indicate the amplitude of the waves; the peak amplitude is reached around the time of the merger, when a single black hole is formed. This figure shows the configuration just after the black holes have merged. SOURCE: Courtesy of C. Henze, NASA.

detecting sources at a significant rate—a milestone expected to be reached in the next decade—it will create a vision of the universe as different and revolutionary as those that radio, X-ray, and gamma-ray telescopes provided in past decades.

Gravitational wave astrophysics specifically addresses many of the key science questions laid out by the Astro2010 Science Frontiers Panels (SFPs). Direct detection of gravitational waves was called out as an area of great discovery potential in astrophysics that would initiate a new era of gravitational wave astrophysics. Gravitational wave astrophysics will also inform two other SFP discovery areas: time-domain astronomy, by elucidating the process of gravitational collapse; and the epoch of reionization, by uniquely recording black hole mergers at high redshift. See Box 8.1.

Gravitational waves are a distinctive cosmic messenger. They carry information about the universe that cannot be obtained from electromagnetic radiation or particle detectors. They are generated by the motions of massive objects, rather than by the glow of plasma, and thus give information about the dynamics of black holes and other massive compact objects. The possibility of obtaining new information about cosmic sources, and of combining information from the different windows (electromagnetic, particle, gravitational), can advance astronomy into a new era (see Box 8.2).

Currently, indirect measurements of gravitational radiation as seen in the orbital dynamics of pulsars are available: the Hulse-Taylor pulsar B1913+16 and the double pulsar J0730-3039B, both binary systems in our galaxy, have provided exquisite tests of the theory of general relativity, including the energy emitted in gravitational waves shrinking their orbits. Although the weak coupling of gravitational waves makes direct detection very difficult, high-precision instruments have been under development for more than a decade in the quest to capture signals from these elusive waves. Today, the prognosis for direct detection is excellent.

BOX 8.1
Science from Gravitation and Gravitational Waves

Science Frontiers Panel (SFP) Discovery Areas	Relevant SFP Panel
Gravitational-wave astronomy	Panel on Cosmology and Fundamental Physics (CFP)
Time-domain astronomy	Panel on Galactic Neighborhood (GAN) Panel on Stars and Stellar Evolution (SSE)
The epoch of reionization	Panel on Galaxies Across Cosmic Time (GCT)

BOX 8.2
Science from Gravitational Waves

SFP Questions Addressed		**Measurements Addressing the Questions**
GCT 1	How do cosmic structures form and evolve?	Tracing galaxy-merger events by detecting and recording the gravitational-wave signatures
GCT 3	How do black holes grow, radiate, and influence their surroundings?	Using gravitational-wave inspiral waveforms to map the gravitational fields of black holes
GCT 4	What were the first objects to light up the universe, and when did they do it?	Identifying the first generation of star formation through gravitational waves from core-collapse events
SSE 2	What are the progenitors of Type Ia supernovae and how do they explode?	Detecting and recording the gravitational wave signatures of massive-star supernovae, of the spindown of binary systems of compact objects, and of the spins of neutron stars
SSE 3	How do the lives of massive stars end?	
SSE 4	What controls the mass, radius, and spin of compact stellar remnants?	
CFP 1	How did the universe begin?	Detecting and studying very-low-frequency gravitational waves that originate during the inflationary era
CFP 2	Why is the universe accelerating?	Testing of general relativity—a deviation from general relativity could masquerade as an apparent acceleration—by studying strong-field gravity using gravitational waves in black hole systems, and by conducting space-based experiments that directly test general relativity

The universe is bright in gravitational waves, because many highly energetic sources exist. For example, the Hulse-Taylor binary pulsar emits the same power in gravitational waves as the Sun does in electromagnetic waves. Even more remarkably, the gravitational-wave luminosity during the final coalescence of a black hole binary is about 10^{23} solar luminosities; for a stellar black hole binary (10 solar masses) this luminosity lasts for about 5 milliseconds, whereas for a massive black hole binary (about 10^{6} solar masses) it lasts about 10 minutes. This is far more energy than is emitted electromagnetically by a gamma-ray burst and in fact exceeds the total energy emitted as electromagnetic waves by all the stars in the observable universe during these same time periods! However, the physical effect of gravitational waves in stretching space-time or accelerating other masses is minuscule: the universe is almost transparent to gravitational waves. Hence, by detecting gravitational waves we can hope to see to black holes through the obscuring matter surrounding them and possibly even to the earliest instants in the history of our universe through the plasma that obscures this epoch for electromagnetic radiation.

The gravitational-wave spectrum extends from the lowest frequencies (about 10^{-16} Hz for the waves that alter the polarization of the cosmic microwave background) to very high frequencies (above 1 kHz for oscillations of stellar black holes). In terms of wave periods, these correspond to a range from a few percent of the Hubble time to fractions of milliseconds. This vast span of physical scales encompasses an impressive variety of astrophysical sources, including massive black hole binaries, quantum fluctuations in the early universe, neutron-star-binary mergers, the final spiral of a stellar black hole into a massive black hole in the center of a galaxy, and stellar explosions.

According to most cosmological models, gravitational waves were produced in the early universe through amplification of the vacuum fluctuations that occurred during inflation or through phase transitions of cosmological fields. These waves are accessible at the ultralow- and very-low-frequency end of the spectrum—10^{-16} to 10^{-8} Hz (Figure 8.2). Detection of these waves would identify the energy of symmetry breaking in the inflationary phase transition. It would provide crucial information to help us understand the physics that drove inflation and that set the stage for the subsequent evolution of the universe and its contents.

Pulsar surveys are in the process of identifying a sample of short-period pulsars that can be timed with 100-nanosecond precision. In the next decade, these pulsars—observed from the ground with radio telescopes—will emerge as a tool for detecting gravitational waves in the range of 10^{-10} to 10^{-8} Hz, a band that contains signals from supermassive black hole binaries as well as signals from the early universe. Observing an array of these pulsars, each with 100-ns root mean square residual timing noise in each observation, is expected to yield a confusion-limited gravitational-wave background signal from binaries with masses $>10^8$ $M_\odot$. On top of this confusion-limited signal should be a handful of individually resolvable higher-mass binaries at distances out to $z \sim 2$. In addition, the pulsar-timing array can detect possible contributions to the stochastic background from cosmic strings, superstrings, and inflation. All of the contributions to the stochastic background, including that from massive black holes, have different spectral indices and so might be separately identified. In addition, pulsar-timing observations may detect violations of general relativity resulting from quantum-gravity corrections that accumulate over long distances, or from additional polarization states not predicted by Einstein's theory.

Extending somewhat higher in frequency, the low-frequency band from 10^{-5} to 10^{-1} Hz can be only probed by instruments in space, due to the obscuration by seismic noise in Earth's crust and by gravity-gradient noise in ground-based arrays. This part of the gravitational-wave spectrum is exceptionally rich in astrophysical sources, including mergers of massive and seed black hole binaries, of inspirals

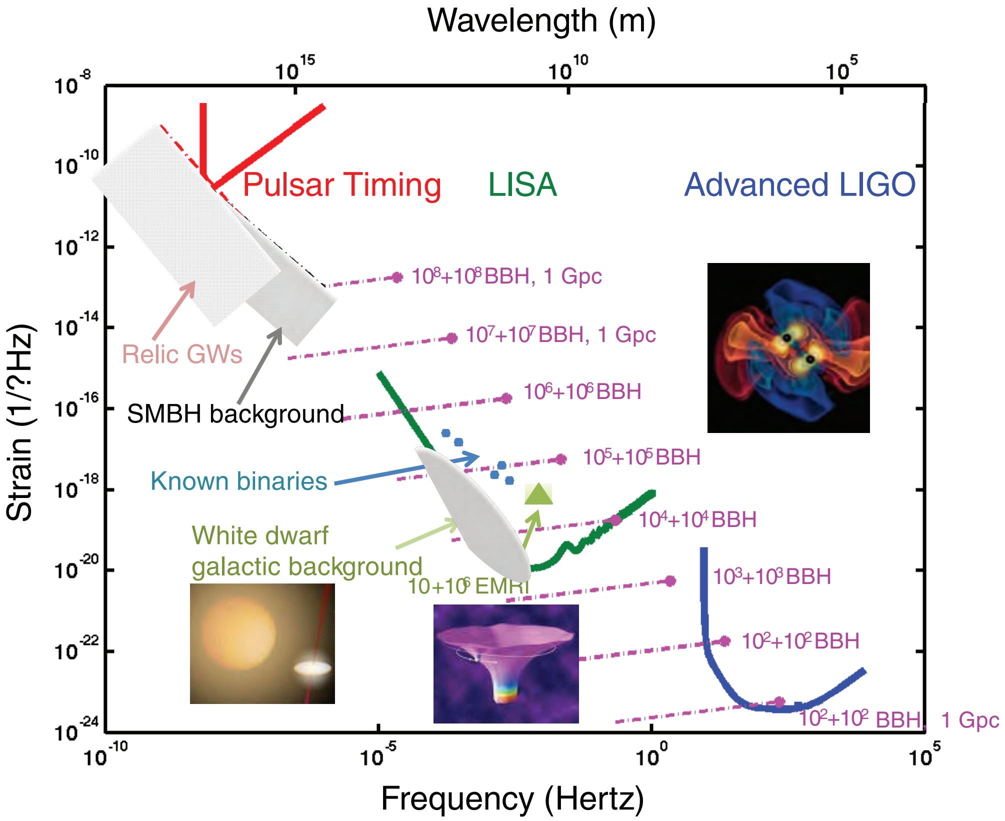


FIGURE 8.2 Strain amplitude sensitivity expected for pulsar timing (red), LISA (green), and Advanced LIGO (blue). The continuous curves show strain-noise-amplitude spectral density. The pulsar-timing sensitivity assumes the use of 20 pulsars with 100-ns timing residuals. The dashed magenta curves show the instantaneous strain of gravitational waves emitted by binary black hole (BBH) systems 1 Gpc away, evolving in frequency to the final coalescence frequency. In the LISA band, the figure shows an estimate of the unresolved background from galactic white-dwarf binary systems in shaded light green; the amplitude of some of the known binary systems; and a representative amplitude of the coalescence of an extreme-mass-ratio inspiral (EMRI) system. In the pulsar-timing band, the expected background is shown from relic gravitational waves (GWs) and from the unresolved signals of supermassive binary black hole systems.

of compact objects into massive central black holes, and of compact binary stars within the galaxy.

The Laser Interferometry Space Antenna (LISA) has the goal of placing a cluster of three spacecraft in solar orbit to detect gravitational waves in this low-frequency band (Figure 8.3). Precise measurements of the relative positions of the spacecraft will detect gravitational waves from this extremely rich collection of astrophysical sources, thereby providing unprecedented scientific opportunities as the result of

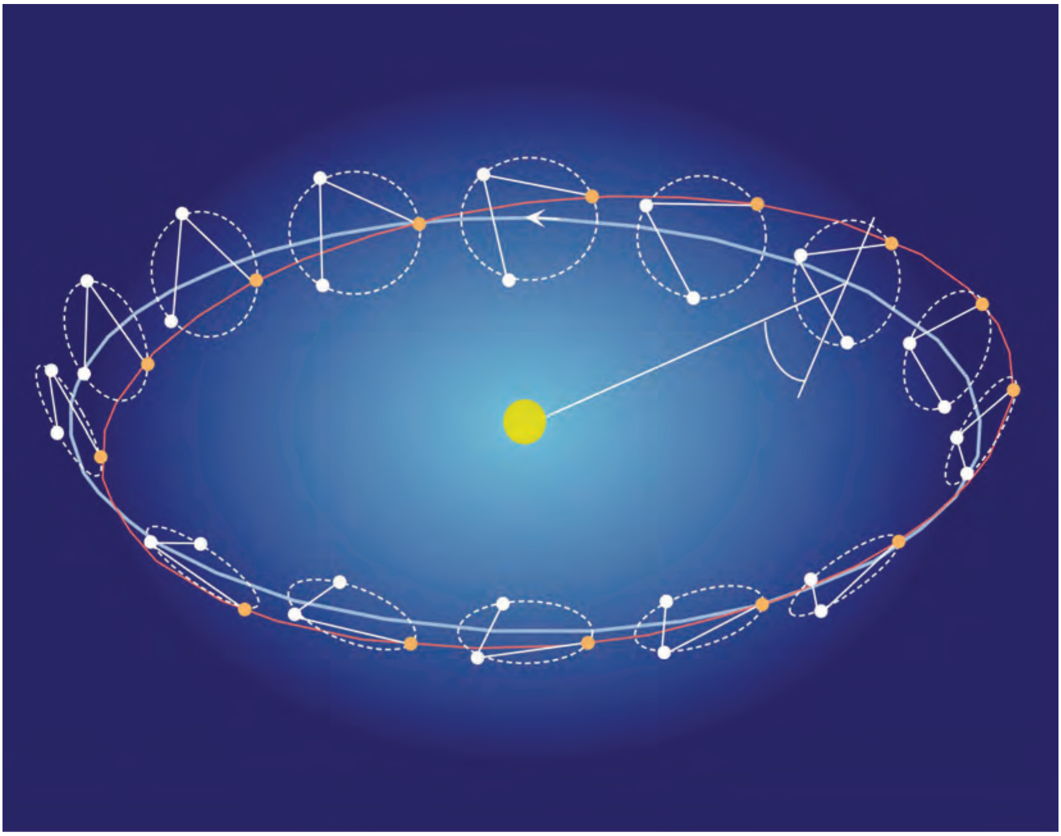


FIGURE 8.3 Schematic of the orbits of the three LISA spacecraft around the Sun. SOURCE: Courtesy of NASA/JPL.

an entirely new technique. Table 8.1 summarizes key astrophysics sources for the LISA mission described below.

- *Mergers of massive black holes*—black holes at the centers of galaxies with masses of 10,000 to 10 million times the Sun's mass—provide the most promising sources. LISA's sensitive frequency band captures the late-inspiral phases of these systems, lasting weeks to years, with very high signal-to-noise ratios. For many of these sources, it is likely that there will be predictions of the time and sky location of the future merger, providing opportunities for simultaneous electromagnetic observations. The detection of the merger waveforms will provide unprecedented tests of general relativity in the regime of very strong and highly dynamical fields: such a measurement is qualitatively different from all the current tests of the theory. The expected detection rates are about 1 per year for close systems with small redshift (less than 2), and up to 30 per year for systems with higher redshifts (up to 15). The study of these sources will provide detailed information on the spins and masses of the massive black hole population at different redshifts, as well as the history of the growth of black holes and the mergers of galaxies.

- *Capture of stellar-mass compact objects by massive black holes.* "Small" compact objects near galactic centers (black holes, neutron stars, and white dwarfs) will inspiral and be captured by the black holes at the centers of the galaxies. The orbital periods will be short (hundreds of seconds), but the signals may last for years. The rate of capture of the small (less than 10 solar masses) objects will be observed up to small redshifts but at a high rate (about 50 per year); intermediate-mass black holes can be detected up to a redshift of 10. Again, the signals with large signal-to-noise ratios will provide very precise tests of general relativity. The measured properties of the black holes captured will provide much-sought-after information about abundances, masses, and spins—contributing to understanding of the history of black hole growth.
- *Close binaries of stellar-mass compact objects in our galaxy.* Tens of thousands of white dwarf binary systems in our galaxy have orbital periods of hundreds to thousands of seconds and produce gravitational waves in LISA's band. Some of the known binaries are close enough to be used as verification systems. Most of the systems will produce signals forming a diffuse foreground, sometimes called confusion noise. The resolved systems will provide a 100-fold increase in the number of known binaries, as well as information on their evolutionary pathways. It is likely that conclusions will be possible regarding white-dwarf binaries as possible progenitors of Type Ia supernovae, and about the physics of tidal interactions and mass transfer before merger.

Finally, LISA may detect signals from cosmological backgrounds (for example, from an early-universe phase transition), bursts from cusps on cosmic (super-) strings, and unforeseen sources.

The LISA mission will release its science data products to enable use and analysis by the astronomy and astrophysics community. In addition, the mission will make available the algorithms, the software, and the models (including a physical model of the gravitational reference sensor) used for processing the data, and will ensure that the data-processing history for any data published by the LISA Science Data Center is traceable and retrievable. With its rich and abundant sources, LISA will usher in a new era in gravitational-wave and multi-messenger astronomy. Therefore, the case is excellent for giving LISA the highest priority for a new start in the next decade.

At the high-frequency end of the spectrum, compact stars of one to hundreds of solar masses (neutron stars and small black holes) near coalescence will produce gravitational waves with frequencies of a few hertz to a few kilohertz, with "chirp"

TABLE 8.1 Key Astrophysics Sources for LISA

Source Type and Details	
Massive black hole (MBH) mergers	
Characteristics	Mergers of binaries involving 2 MBHs, with masses in the range of 10^4 to 10^7 $M_\odot$, orbital periods of 10^2 to 10^5 s, signal durations of ~ weeks to years, amplitude signal-to-noise ratios up to several thousand
Detection rate	~1/yr at $z < 2$, ~30/yr out to $z \sim 15$
Observables	Masses, $\Delta M/M < 1\%$; spins, $\Delta S/S < 2\%$ (typical detections); luminosity distances, $\Delta D_L/D_L \sim (5\text{-}20)\%$ (typical detections), $\Delta D_L/D_L < 3\%$ (at $z = 1$, limited by weak lensing)
Science payoffs	Nature of black hole seeds at high z; history of MBH growth and galaxy mergers as function of z; tests of general relativity in strong-field, highly dynamical regime
Capture of stellar-mass compact objects by MBHs	
Characteristics	Compact object (black hole, neutron star, white dwarf) spirals into MBH; MBH mass of 10^4 to 10^7 $M_\odot$; orbital period of 10^2 to 10^3 s; signal duration ~ years.
Detection rate	~50/yr, mostly captures of 10-$M_\odot$ black holes at $z \sim 1$; captured intermediate-mass black holes detected to $z > 10$
Observables	Masses, $\Delta M/M < 0.1\%$; spins, $\Delta S/S < 0.1\%$ (typical detections); luminosity distances, $\Delta D_L/D_L < 4\%$ (typical detections)
Science payoffs	Measure MBH spins, which reflect their growth history; populations and dynamics of compact-object populations in galactic nuclei; precision tests of general relativity and Kerr nature of black holes
Close binaries of stellar-mass compact objects in the galaxy	
Characteristics	Close binary systems of black holes, neutrino stars, and white dwarfs in Milky Way; primarily white-dwarf/white-dwarf binaries, mass-transferring or detatched; orbital periods of 10^2 to 10^4 s
Detections	~20,000 individual sources, including ~10 known "verification binaries"; diffuse galactic foreground at frequencies below ~2 mHz
Observables	Orbital frequency; sky location to approximately a few degrees; chirp mass and distance from df/dt for some high-f binaries
Science payoffs	~100-fold increase in census of short-period galactic binaries; white dwarf-white dwarf binaries as possible supernova Ia progenitors; evolutionary pathways (e.g., outcomes of common-envelope evolution); physics of tidal interactions and mass transfer

signals that span the spectrum up to a coalescence frequency inversely proportional to the total mass of the system. Pulsars—and spinning stars in general, if not perfectly axisymmetric—radiate gravitational waves at twice their spin frequency. This band also includes mergers of stellar and intermediate-mass black holes, as well as binary neutron stars and supernovae. All these sources can be probed by ground-based interferometric detectors.

Tests of General Relativity and Other Theories of Gravity

General relativity is important both on the large distance scales of astronomy and cosmology and on the small scales that may characterize a unified theory of the basic interactions. Relativistic gravity is therefore a two-way bridge between astronomy and fundamental physics. The SFP science questions *How did the universe begin?* and *Why is the universe accelerating?* require an accurate understanding of relativistic gravity, as does the discovery area of the SFP on cosmology and fundamental physics—gravitational wave astronomy.

Table 8.2, which summarizes the current status of tests of general relativity, organized by length scale and by whether the tests probe weak or strong gravitational fields, shows clearly that general relativity has been well tested on solar system scales in the weak-field regime. In contrast, for stronger fields and larger scales, there are mostly qualitative tests. In cosmology, for instance, general relativity has been assumed, and data have been used to determine the cosmological parameters that characterize our universe. Now that these are known to a good accuracy, it is reasonable to ask whether it is possible to begin to test general relativity on cosmological scales.

There is no compelling classical alternative to general relativity that has survived the solar system tests. Theorists generally agree that deviations from classical

TABLE 8.2 Current Limits on Deviations from General Relativity

Distance Scale	Weak Gravity	Strong Gravity
Laboratory	Weak equivalence principle, 10^{-13} Limits on fifth force and compact extra dimension size, 56 μ Gravitational redshift, 10^{-8}	
Solar system	Weak equivalence principle, 10^{-13} Strong equivalence principle, 10^{-4} Gravitational redshift, 10^{-4} Bending of light, 10^{-4} Shapiro time delay, 10^{-5} Precession of perhihelion, 10^{-3} Lense-Thirring precession, 5 to 15%	Gravitational radiation from binary pulsars, 10^{-3} *Black holes*
Galactic	Lensing bending of light, 10%	*Black holes*
Cosmological		*Observations fit to values of the Hubble constant and the densities of matter, radiation, and dark energy*

NOTE: The numbers in this table are order-of-magnitude values for the current accuracies with which the measured effects currently agree with general relativity. For instance, for most solar-system tests, the accuracy to which the parametrized post-Newtonian parameters have the general-relativistic values is listed. Italics are used to indicate tests that are only qualitative at this time.

general relativity can be expected at the Planck scales that characterize quantum gravity (10^{-33} cm, 10^{19} GeV), but these scales are very far from those that can be probed directly with today's experiments. Rather, the question is whether physics at the Planck scale leaves imprints at low energies that can be tested by our observations and experiments (Figure 8.4).

Contemporary ideas of fundamental theory allow the construction of many different four-dimensional theories of gravitation which might govern the results determined by observation and experiment. Some of these theories predict deviations from general relativity, but while many models produce deviations from general relativity that might be tested by experiment, there are no secure predictions to provide targets for tests.

The science case for tests of general relativity rests generally on the importance of the theory for astronomy, fundamental physics, and the connection between them. The following aspects of the current experimental and theoretical situation strongly motivate tests of the theory in the next decade (Figure 8.5):

FIGURE 8.4 The 4-cm-high torsion pendulum at the University of Washington is at the heart of probably the most accurate test of principle in all of physics—the equivalence principle, which states that all masses fall with the same acceleration in a gravitational field regardless of their composition. This principle is central to general relativity, and any deviation would entail a significant revision of the ideas about gravitation. The pendulum is suspended from a rotating platform by a fiber barely visible at the top of the original figure. Any difference in the accelerations of the test masses would appear as a twisting of the pendulum. The experiment confirms the equivalence principle for beryllium and titanium masses in the gravitational field of Earth to an accuracy of less than 1 part in 10^{13} (S. Schlamminger, K.-Y. Choi, T.A. Wagner, J.H. Gundlach, and E.G. Adelberger, Test of the equivalence principle using a rotating torsion balance, *Physical Review Letters* 100:041101, 2008). Similar experiments provide accurate tests of the gravitational inverse square law and Lorentz invariance. SOURCE: Courtesy of E. Adelberger, University of Washington.

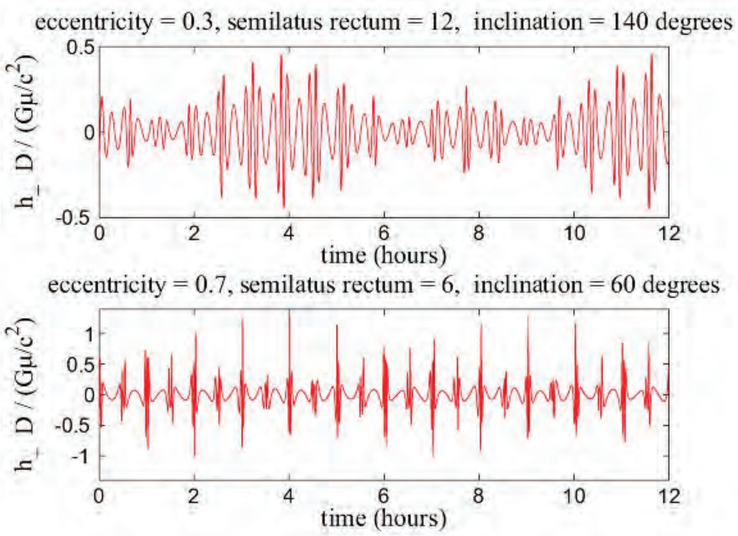


FIGURE 8.5 Waveform for one polarization of gravitational waves produced by a test mass orbiting a million-solar-mass black hole that is spinning at 90 percent of the maximum rate allowed by general relativity. The two panels correspond to different configurations of the test-mass orbit. The top panel assumes a slightly eccentric and inclined retrograde orbit moderately far from the horizon. The bottom panel assumes a highly eccentric and prograde orbit much closer to the horizon. The amplitude modulation visible in the top panel is due mostly to Lense-Thirring precession of the orbital plane. The bottom panel's more eccentric orbit produces sharp spikes at each pericenter passage. SOURCE: Steve Drasco, Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Potsdam, and Scott A. Hughes, Department of Physics and Kavli Institute for Astrophysics and Space Research, MIT.

- *Much remains untested at large scales and in strong fields.* Detailed astronomical observations in the next decade will probe the regimes of strong field and large scale in which general relativity is quantitatively untested. This also provides strong motivation to complete the missing parts of Table 8.2 at the same time. The LISA mission will, for the first time, probe these gaps in current knowledge by characterizing gravitational waves across cosmological distances from sources governed by strong-field gravity.
- *Much remains unknown about general relativity.* Is there a non-zero cosmological-constant term in the Einstein equation, or is the observed acceleration of the universe due to a new type of field? Electromagnetic surveys that measure cosmological distances as a function of redshift, and which map the growth of structure as a function of redshift, will test the hypothesis of a cosmological constant. Are the massive central objects in galaxies indeed black holes, described by the Kerr metric of general relativity? Is the "cosmic-censorship" conjecture true, so that black holes rather than naked singularities form in sufficiently advanced gravitational collapse? LISA will observe the inspirals of compact stellar remnants

spiraling for years into the massive dark objects at the centers of galaxies, tracking the last 100,000 cycles of gravitational radiation emitted in the strong-field region and producing a high-precision map of the space-time that will reveal any small deviations from the Kerr metric (see Figure 8.5). If the central massive object is not a black hole, but rather an object with no horizon, then the radiation will continue long after it would have turned off in the black hole space-time.

- *Much remains unknown about fundamental theory.* Modifications of general relativity on accessible scales are not ruled out by today's fundamental theories and observations. It makes sense to look for them by testing general relativity as accurately as possible. Cost-effective experiments that increase the precision of measurements of parametrized post-Newtonian (PPN) parameters, and which test the strong and weak equivalence principles, should be carried out. For example, improvements in Lunar Laser Ranging promise to advance this area.

Dark-Matter Detection and Characterization

The inferred existence of dark matter, the mysterious substance that makes up 83 percent of the matter in the universe, raises profound questions for both astronomy and physics. What is dark matter? Where does it come from? What are the connections between dark matter and "ordinary" matter? (See Box 8.3.) The detection of astrophysical signatures of dark-matter particles will play an essential role in the discovery of dark matter, along with direct detection and collider experiments. Astrophysical signatures are expected to include gamma rays, antiparticles (such as positrons and anti-protons), anti-nuclei, and neutrinos.

The interaction of dark matter with the ordinary universe is so tenuous that it is known only through gravitational interactions. In spite of the indirect nature of the evidence, galactic rotation curves, strong gravitational lensing, and large-scale-structure measurements, combined with big bang nucleosynthesis constraints, all continue to support the dark matter hypothesis with increasingly convincing data.

BOX 8.3
Science from Dark Matter Detection and Characterization

SFP Questions Addressed		**Measurements Addressing the Questions**
CFP 3	What is dark matter?	Indirect astrophysical searches for dark-matter annihilation and decay signatures
GAN 4	What are the connections between dark and luminous matter?	

Furthermore, the power spectrum of the cosmic microwave background agrees with the scenario that dark matter constitutes 83 percent of the matter and 20 percent of the energy density in the universe. The matter power spectrum is measured at smaller scales by galaxy surveys, yielding consistent results.

In spite of these accomplishments, the nature of the dark-matter particles is not yet known. The 2001 decadal survey emphasized three leading candidates: massive neutrinos, axions, and weakly interacting massive particles (WIMPs).[1] Massive neutrinos now appear less likely, as the most massive neutrino flavor is thought to be in the 0.04- to 2-eV range. This mass range cannot explain large-scale structure formation, both because such neutrinos remain relativistic for too long in the early universe, and because the mass density of relic neutrinos is far too small. The axion was postulated to solve the "strong CP problem"—that is, to prevent the violation of charge-parity symmetry in strong interactions. If axions exist, they may constitute a significant fraction of the dark matter. Because axions are light they are unlikely to produce high-energy particles via annihilation or decay in astrophysical scenarios, and signatures of axions are the subject of a separate class of direct-search experiments targeted for the axion's coupling to two photons. Thus in this report the panel concentrates on indirect searches for WIMPs.

A WIMP is a thermal relic of the big bang—a massive particle in thermal equilibrium in the early universe that decouples when its number density falls so low that the annihilation timescale equals the Hubble time. This establishes a relationship between annihilation cross section and mass density that gives the correct dark-matter density today if the annihilation cross section is appropriate for the scale of the weak interactions, and the mass of the WIMP is within about two orders of magnitude of the weak scale (say, 1 GeV to 10 TeV).

Because the existence of dark matter has been inferred only from its gravitational effects, the expected astrophysical signals can be estimated only in the context of some models. A leading WIMP candidate is the lightest supersymmetric particle in the minimal supersymmetric standard model (MSSM). For the parts of MSSM parameter space that yield the correct relic density, the lightest particle is the superpartner of the neutral gauge bosons, known as the neutralino. In such a model interaction cross sections and masses can be estimated (at least to within several orders of magnitude), yielding a starting point for searches. However, this basic scenario may be extended in many ways.

Although any thermal relic WIMP must annihilate at some level, it is not yet known whether such annihilation produces a signal that will be detectable in the coming decade. The signals generally expected from conventional WIMP models are small; however, recent reports of a possible excess (over those expected from

---

[1] National Research Council, *Astronomy and Astrophysics in the New Millennium*, National Academy Press, Washington, D.C., 2001.

combined astrophysical and particle physics models) of particles and photons in the ~10- to 1,000-GeV range triggered interest in models of dark matter with an annihilation cross section significantly higher than that of the "standard" thermal relic WIMP. These "Sommerfeld-enhanced" models, which were proposed as a way to understand otherwise contradictory results, involve a new force of nature that acts on dark matter, but only very weakly on ordinary matter. This force causes a significant (1 to 3 orders of magnitude) enhancement in the annihilation cross section and favors annihilation to leptons rather than to hadrons. These models are of interest because they can fit the recently observed particle and photon spectra, which is difficult to do with a more conventional WIMP. Such models were originally invented to explain the annual modulation of the scattering signal in the DAMA/LIBRA sodium-iodide direct-detection experiment. In this scenario, the new force mediates an inelastic scattering between two WIMP states, allowing DAMA/LIBRA to be marginally compatible with other experiments. Such ideas will be thoroughly tested by the current generation of direct-detection experiments with target mass at the 100-kilogram scale.

Improved capabilities in spectral measurements of cosmic rays and gamma rays and in background rejection can provide evidence for dark matter. In fact, the most distinctive signature of dark matter may be in the gamma-ray spectrum from the inner Milky Way, Milky Way satellite galaxies, or other nearby galaxies. Interpretation of both gamma-ray and particle signals will also require improved modeling of sources and propagation. If the apparent excesses described above result from dark matter, this is of fundamental importance to particle physics. If not, it is essential to understand them as astrophysical signals, both in their own right and so that future projects will be able to isolate the faint dark-matter signals from the much larger astrophysical background. Experiments that push the limits of detector mass and exposure time are called for. The next generation of these experiments motivates a strong program in support of balloon payloads and flights, particularly ultralong-duration ballooning, and opportunities for cost-effective experiments on small to midsize satellites.

High-Energy Particle Astrophysics

High-energy particles, including gamma rays, cosmic rays, and neutrinos, bring new and complementary views of astronomical sources and probe physical processes under extreme conditions throughout the universe. Gamma rays and cosmic rays (charged particles or particles whose nature is not known) are produced in cosmic accelerators over a vast range of scales, from the solar system to powerful extragalactic sources. Cosmic rays are a major contributor to the energization of the interstellar medium in our galaxy and in others, as well as in galaxies undergoing formation, where jets from active galactic nuclei play major roles in regulating star formation. High-energy cosmic neutrinos have yet to be detected; their detection will add invaluable information to the study of cosmic accelerators, as the neutrinos arise deep inside the acceleration regions and travel undeflected from cosmologically distant sources. Advancing the physical understanding of cosmic accelerators through observation, simulation, and theory feeds directly into several major themes of galactic and extragalactic astrophysics and cosmology. (See Box 8.4.)

Particle-astrophysics observations have the important ability to probe physics beyond the standard model of particle physics. A striking example of this potential is the search for dark-matter annihilation and decay products, discussed above, which could in principle involve any of the species of particles or gamma rays studied using the tools of particle astrophysics. Particle astrophysics extends the high-energy frontier well above energies accessible to laboratory accelerators, through study of ultrahigh-energy cosmic rays, photons, and neutrinos. To use high-energy particles effectively for probing dark-matter signatures and physics beyond the Large Hadron Collider scale requires understanding their origin and their propagation to Earth.

Gamma rays are produced both in cosmic electron and in hadron accelerators; they provide essential constraints for the study of electron accelerators and will be key in unveiling the sites of hadronic acceleration. Recent advances in high-energy

BOX 8.4
Science from High-Energy Particle Astrophysics

SFP Questions Addressed		**Measurements Addressing the Questions**
GAN 2	What controls the mass-energy-chemical cycles within galaxies?	Gamma rays, cosmic rays, and neutrinos from active galactic nuclei and gamma-ray bursts
GCT 3	How do black holes grow, radiate, and influence their surroundings?	
SSE 1	How do rotation and magnetic fields affect stars?	Gamma rays and neutrinos from supernovae, gamma rays from gamma-ray bursts, and gamma rays from stars and binary systems
SSE 3	How do the lives of massive stars end?	
CFP 4	What are the properties of neutrinos?	Detection of neutrinos from cosmic accelerators
		Detection of cosmogenic (GZK) neutrinos
		Tests of physics above the TeV scale
How are ultrahigh-energy particles accelerated? (not an SFP question)		Studies of ultrahigh-energy cosmic rays, neutrinos, and gamma rays (composition, spectra, sources)

gamma-ray observations have revealed a plethora of gamma-ray sources, including gamma-ray pulsars, compact binaries, the galactic center, and extragalactic sources such as starburst galaxies and radio galaxies. During the next decade, improved sensitivity and spectral coverage of the new generation of gamma-ray observatories will provide detailed spatial and spectral information on known sources to determine how binaries and pulsars produce gamma rays, how supermassive black holes power jets, what powers gamma-ray bursts at low and high redshifts, what is the extragalactic background light, and what is the origin of cosmic rays. Gamma-ray observations may also address the nature of dark matter and may discover new, completely unexpected sources. Much progress has been made in the detection of high-energy gamma rays by way of atmospheric Čerenkov emission. There is a strong science case now for a factor-of-10 improvement in sensitivity and for instruments with a wider field of view. There is also a strong science case for improvement by a factor of 2 to 3 in angular resolution. These requirements motivate U.S. involvement in an international Čerenkov telescope array with an effective area of approximately 1 square kilometer. Finally, the importance of observations that cover a wide field of view with a large duty cycle is also recognized, as demonstrated by the Fermi Gamma-Ray Space Telescope (formerly GLAST; ranked first among medium space-based missions in the 2001 decadal survey). These observations are particularly important for transient and extended sources.

The origin of cosmic rays is still a mystery, although much has been learned in the past decade. The observations span more than 30 orders of magnitude in flux and track acceleration and propagation processes at scales from the solar system to powerful extragalactic sources. The standard paradigm for the origin of galactic cosmic rays involves Fermi acceleration in non-relativistic shock waves in supernova remnants. Low- and intermediate-energy cosmic rays are consistent with a galactic origin. At ultrahigh energies, cosmic rays show a sky distribution and spectral features consistent with an extragalactic origin. The transition between galactic and extragalactic origins and the locations of the principal sites of highly efficient acceleration (in excess of 10 percent) are poorly understood. The observed attenuation of the ultrahigh-energy spectrum is consistent with the predicted Greisen-Zatsepin-Kuzmin (GZK) cutoff due to the interaction of cosmic-ray protons with the cosmic microwave background, but it could alternatively be due to the limits of acceleration in sources, especially if heavier nuclei dominate. The indication of an anisotropic distribution of the highest-energy cosmic rays, as reported by the Auger South Observatory, suggests a dominantly proton flux. More-precise determination of the degree of anisotropy, which may be possible with improved statistics by adding Auger North in the next decade, will help determine the nature of the ultrahigh-energy cosmic rays and thereby help establish expected GZK fluxes of ultrahigh-energy neutrinos and photons.

Neutrinos were identified by the Astro2010 SFPs as a particularly fruitful area of research for the next decade. They are weakly interacting messengers that, unlike gamma rays and cosmic rays, can reach Earth undeflected from the edge of the observable universe, independent of energy. They should be produced at sites of hadronic acceleration or interaction and thus provide a unique tool to study high-energy accelerators and particle interactions. The main challenge is their detection.

The expectations for neutrino astronomy are deeply connected with gamma-ray and cosmic-ray observations, and all three types of particle must be considered together. Gamma- and cosmic-ray sources are likely also to produce neutrinos, and neutrinos are produced by the interaction of ultrahigh-energy cosmic rays with the cosmic background radiation. Cosmogenic neutrinos (the GZK neutrinos) are often considered to be a guaranteed source of ultrahigh-energy neutrinos, but the predicted fluxes depend on the proton fraction and the cosmological evolution of cosmic-ray sources. New physics above the TeV scale may be tested by comparing rates of horizontal neutrino air showers initiated in Earth's atmosphere with Earth-skimming neutrinos.

Connections to Other Areas of Physics and Astrophysics

The science questions discussed in this panel report can also be approached by methods that complement those that this panel considered. For example, to address the question *What are the particles that make up the dark matter?* three approaches are needed: indirect evidence from observations of high-energy particles produced in dark-matter annihilation or scattering processes; direct searches with detectors sensitive to the rare scattering of dark-matter particles by normal matter; and studies of production and decay cascades in collider experiments. While only astrophysical detection is within the purview of this report, it plays a critical role in establishing that any new signals found by the other approaches are dark-matter related, as opposed to unrelated new physics. Similarly, two of the most fascinating fundamental science questions relevant to particle astrophysics and general relativity are the SFP questions, *Why is the universe accelerating?* and *How did the universe begin?,* yet many of the experiments that address these questions, such as large electromagnetic surveys, fall in the purview of other Astro2010 panels and hence are not discussed in any depth in this panel report. Also excluded are experiments addressing the physics of low- to intermediate-energy gamma rays, which were considered by another panel. These exclusions are not a statement about the interest in this science, but rather that it is not probed by the types of experiments that this panel is charged to consider. As the panel presents its recommendations, connections to the experiments and programs considered by other panels are indicated, as appropriate.

THE PROGRAMMATIC CONTEXT

Gravitational-Wave Astrophysics

The 2001 decadal survey recognized gravitational-wave astrophysics as a promising area; LISA was given the highest priority by the Panel on Particle, Nuclear, and Gravitational-Wave Astrophysics (PNGA)[2] and was ranked second overall among the space-based medium-class missions. The PNGA recommended a technology-development program to reduce the risk in the LISA mission. It also noted the major challenge to theory of calculating waveforms from black hole mergers, a key LISA source, using numerical relativity. Much progress has been made in both these areas in the past decade.

Ground-based interferometric gravitational-wave detectors have improved sensitivity and are now collecting data with some potential for discovering gravitational waves from binary systems as far away as the Virgo cluster of galaxies. In the United States, the Laser Interferometer Gravitational Wave Observatory (LIGO) consists of three detectors, 2 and 4 km long (Figure 8.6). Two instruments in Europe are VIRGO, 3 km long, and GEO, 0.6 km long. Even more importantly, advanced designs for the LIGO detectors have been funded and are being built by the United States, and a similar advanced design is being pursued by VIRGO. With these advanced detectors operating at their design sensitivities, the predicted rate for observations of neutron stars and stellar black holes is in the range of dozens per year. Even now the existing detectors are producing results of astrophysical interest. For example:

- A search with LIGO and VIRGO in temporal and direction coincidence with 22 gamma-ray bursts found no statistically significant candidates. Thus, neutron-star/black-hole progenitors are excluded from our galaxy or neighboring galaxies.
- A LIGO search for gravitational radiation from the Crab pulsar limits the power radiated in gravitational waves to less than 2 percent of the spindown power.
- Null results in a search for a stochastic gravitational-wave background with LIGO places an upper limit on the density of such a background at less than 1 part in 100,000 of the critical cosmological density, ruling out some models of early-universe evolution.

The current program in gravitational-wave astrophysics includes technology development for other advanced and future detectors around the world. An underground cryogenic prototype in Japan (CLIO) is operating, and a large Japanese

---

[2] National Research Council, *Astronomy and Astrophysics in the New Millennium: Panel Reports*, National Academy Press, Washington, D.C., 2001.

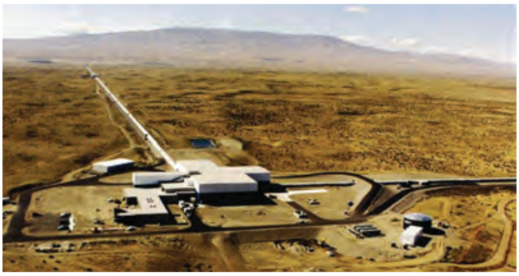


FIGURE 8.6 The Laser Interferometer Gravitational Wave Observatory (LIGO) facility in Hanford, Washington. SOURCE: LIGO Laboratory.

cryogenic telescope (LCGT) is under consideration. There are also plans for a European third-generation underground detector known as the Einstein Telescope.

Over the past decade, technology for LISA has matured, and the European Space Agency (ESA) is planning a LISA Pathfinder mission that will provide critical tests of several of LISA's subsystems in a space environment. Significant progress has also been made in planning for LISA data reduction. Numerical relativists have conquered the theory challenge with breakthroughs in simulating the merger of two black holes and in computing the resulting gravitational waveforms. This has enabled detailed modeling of source waveforms, development of analysis software, and several community-wide "mock data challenges." Addressing detection at even lower frequencies, there has also been much progress in developing pulsar-timing techniques and in completing surveys to identify suitable pulsar-timing systems.

Tests of General Relativity and Other Theories of Gravity

Relativistic gravitation underlies modern astrophysics and cosmology and is a cornerstone of fundamental physics, yet a coherent program to test general relativity that would systematically explore the ranges of scale and strength shown in Table 8.2 has not been established. Tests of gravity can be carried out using ground-based instrumentation, where a high degree of control is available, or in space where quiet conditions prevail. The targets of tests may involve either weak gravity, as in laboratory or solar-system sources, or strong gravity, as in some astrophysical sources and cosmology. As a consequence, the testing of general relativity is spread across different science-support agencies with little coordination among them. Many tests have exploited opportunities presented by missions developed primarily for other purposes.

Despite this fragmentation, significant progress in testing general relativity has been made in the past decade. To give a flavor of what has been achieved, the panel presents the following highlights:

- Measurements of the parameterized post-Newtonian parameter γ on solar system scales using the differential Shapiro time delay, as a by-product of the Cassini mission, and on kiloparsec scales from a comparison of gravitational-lensing and velocity-dispersion measurements;

- The direct detection of gravitomagnetic effects (the Lense-Thirring precession) from Lageos/Grace, Gravity Probe B, and lunar laser ranging;
- The measurement of geodetic spin precession in the double pulsar system J0737-3039A/B (see Figure 8.7 for an illustration of five tests of general relativity in that system);
- The ongoing monitoring of the binary-pulsar orbital decay, consistent with the emission of gravitational radiation predicted by general relativity to within a fraction of a percent;

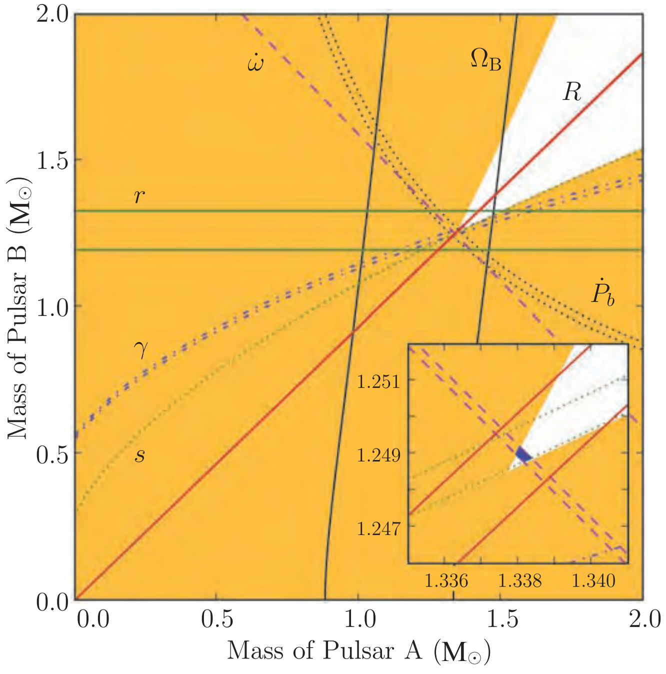


FIGURE 8.7 Five tests of general relativity using the double pulsar J0737-3039A/B. This figure illustrates both the variety and the precision of tests of general relativity that are possible today in one system. The post-Keplerian parameters are fitted along with the masses of the two pulsars. The parameter $\dot{\omega}$ is the periastron advance, $\dot{P}_b$ is the period decrease due to gravitational radiation, and s and r are related to the precession of pulsar B. The white area corresponds to allowed inclination angles of the orbit. The intersection of the allowed ranges in the tiny box at the center of the inset provides five tests of general relativity at the 10^{-4} to 10^{-5} level. SOURCE: From R.P. Breton, V.M. Kaspi, M. Kramer, M.A. McLaughlin, M. Lyutikov, S.M. Ransom, I.H. Stairs, R.D. Ferdman, F. Camilo, and A. Possenti, Relativistic spin precession in the double pulsar, *Science* 321(5885):104-107, 2008. Reprinted with permission from the AAAS.

- The verification of the gravitational inverse square law down to ranges of 56 microns in laboratory experiments;
- The lunar-laser-ranging verification of the strong equivalence principle to 10^{-4}, meaning that the triple-graviton vertex is now known to better accuracy than is the triple-gluon vertex;
- Limits on the fractional rate of change of the gravitational constant G ($<10^{-12}$/yr) from lunar laser ranging;
- Atomic experiments limiting time variation of the fine structure constant to 10^{-16}/yr over periods of several years;
- The supernova and cosmic microwave background measurements of the acceleration of the universe; and
- Experiments in progress that include the MICROSCOPE equivalence principle experiment, the APOLLO lunar-laser-ranging observations, and tests of general relativity using torsion balances and atom interferometry.

The above list shows that much has been done to test general relativity, and Einstein's theory is consistent with all experimental tests to date! However, a glance at Table 8.2 shows that much remains to be done to test the theory in the domains of strong gravity and on scales larger than the solar system. The accomplishments of the past decade set the stage for the next decade; in Table 8.3 the panel presents goals for the next decade. In making its final recommendations, the panel returns to those that fall within its purview.

TABLE 8.3 Possible Tests of Relativistic Gravity in the Next Decade

Distance Scale	Weak Gravity	Strong Gravity
Laboratory	Improved equivalence-principle limits and measurements of parametrized post-Newtonian (PPN) parameters from atom interferometry	Better constraints on extra dimensions, for example, from accelerator experiments
Solar system	Improved strong- and weak-equivalence-principle limits; better determination of PPN parameters and the rate of change of the gravitational constant from next-generation lunar laser ranging	Gravitational waves detected directly and their predicted speed and polarization confirmed; properties of rotating black holes confirmed quantitatively by gravitational waves and X-ray reverberation
Galactic	Measurement of PPN parameters by lensing and velocity dispersion	Predicted connections between sources and gravitational waves confirmed quantitatively
Cosmological	Better bounds on variations of fundamental constants, e.g., α, m_e/m_p	Gravitational waves detected by cosmic microwave background polarization; relation between expansion history and growth of structure tested with supernova, baryon acoustic oscillations, and weak lensing

Dark Matter Detection and Characterization

The possibility that a new class of fundamental particles could make up the dark matter gives the search for WIMPs in the galactic halo a very high scientific priority. Direct detection of dark matter would be the most definitive way to determine that WIMPs make up the missing mass. The study of WIMP candidates in accelerator experiments is also critical for determining the relic density of these particles. The indirect detection of astrophysical signals due to WIMP-WIMP self-annihilation may also provide important clues, but in many cases such signals may be difficult to separate unambiguously from more mundane astrophysical sources. That leaves direct detection as playing a central role in establishing the presence of WIMPs in the universe today. Also, given both the technical challenge and the fundamental importance of direct detection of WIMPs, it is vital to have the means to confirm a detection in more than one type of detector. While evaluation of future direct-detection experiments is outside the charge of this panel, current and possible future experiments are summarized as part of the context for indirect astrophysical detection experiments.

Experiments based on cryogenic noble liquids are scaling up rapidly to provide large detector mass with very low background (for example, the U.S.-led XENON100 experiment and LUX (Figure 8.8); the Japanese-led XMASS experiment; and the Italian-led WArP experiment. The current generation is using 100-kg-scale detectors, and there are realistic prospects for a ton-scale experiment by the middle of the next decade (XENON1T). Experiments using silicon and germanium crystals cooled to millikelvin temperatures continue to scale up (the U.S.-led CDMS and the French-led EDELWEISS). The bubble chamber technique pioneered by the U.S.-led COUPP has carved out a niche in the spin-dependent model space and may be scalable to ton-scales in the future. As experiments using these various techniques increase substantially in size, it is quite possible that one or more will make a convincing detection of WIMP-nucleon scattering. If this happens, it will be possible to search for the annual modulation of the signal as Earth goes around the Sun. Direction-sensitive detectors would seek the changing direction of the particle “wind” as Earth moves in its orbit. Furthermore, by comparing nuclear-recoil spectra from multiple experiments, information about the WIMP mass could be derived.

These direct-detection experiments provide an essential step in establishing the existence of a WIMP candidate. However, whatever particle they may discover need not constitute the majority of dark matter. There may be many species of WIMPs, and because the annihilation cross section and scattering cross section are parametrically related (in tree-level interactions), the scattering probability may scale as the inverse of the thermal relic freeze-out density. This raises the possibility that there may be many kinds of WIMPs with densities and cross sections span-

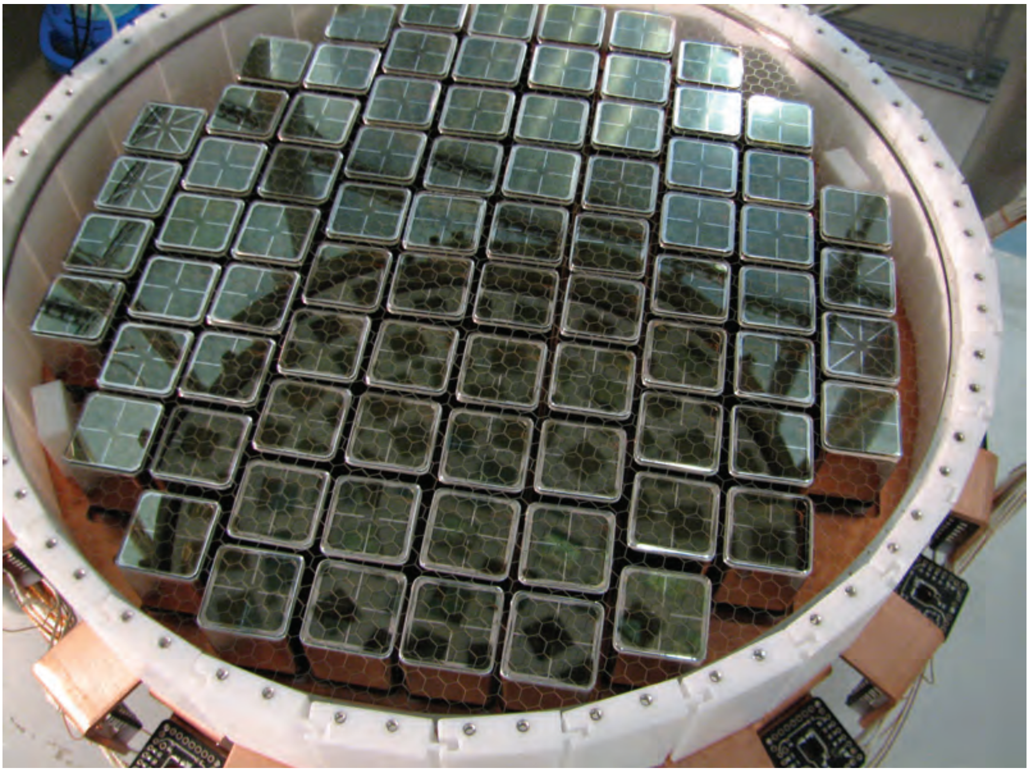


FIGURE 8.8 The XENON100 experiment. One-inch-square photomultipliers, highly sensitive to the vacuum-ultraviolet scintillation of liquid xenon, detect the faint signals expected from a WIMP interaction in the liquid xenon time-projection chamber (TPC). The photo shows the bottom array of 80 such photomultipliers and the reflected image of the top array of 98 photomultipliers. Also visible is the hexagonal wire mesh, which serves as the TPC's cathode for drifting electrons through the 30-cm-deep liquid-xenon target. SOURCE: XENON100 Collaboration.

ning several orders of magnitude. Therefore, proof that the first WIMP detected is indeed *the* dark-matter particle will require astrophysical detection of annihilation or decay signals compatible with the direct-detection signals. The most obvious such signals are cosmic rays, neutrinos, and gamma rays.

Recent high-energy particle results, while tantalizing, have failed to paint a coherent picture consistent with a plausible dark matter candidate. The Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics (PAMELA) spacecraft has measured the e^+ fraction as 10 percent at 100 GeV, twice what it is at 10 GeV and much larger than that expected from models of secondary particles produced by proton cosmic rays interacting with the interstellar medium.

However a similar effect is not observed for anti-protons. At still higher energies, the Advanced Thin Ionization Calorimeter (ATIC) balloon-borne experiment claimed a peak in the electron spectrum (ATIC cannot distinguish positrons from electrons). Fermi also measured the electron-plus-positron spectrum up to 1 TeV, finding a broad-spectrum excess above the expected spectrum, but not the ATIC peak. These signals are much larger than expected for a thermal WIMP annihilating through a hadronic cascade, but they are possible in more exotic models. Models requiring leptophilic dark matter are required to justify the lack of an excess in the PAMELA anti-proton measurements. The situation indicates that it is important to improve knowledge of secondary production of cosmic rays to distinguish between dark-matter models or alternative scenarios, such as cosmic rays produced by local sources.

Anti-nuclei are also a target for indirect dark-matter detection experiments. The Alpha Magnetic Spectrometer (AMS) experiment is designed to characterize anti-nuclei using an orbiting magnetic spectrometer. The AMS01 prototype detector provided a limit on the existence of anti-helium, in addition to measuring the proton and electron-positron spectra in near-Earth orbit. Scheduled for the International Space Station, the full AMS experiment, with a magnetic spectrometer, will provide spectra of electrons, positrons, protons, anti-protons, and various nuclei up to the TeV region. The panel notes that the best limits on anti-helium to date are those due to the BESS-polar balloon experiments, illustrating the potential of this relatively low-cost platform. AMS and balloon experiments under development will also have the potential to detect anti-deuterons, which may be produced in dark-matter annihilations but are not made by any other known process.

High-energy photons can also provide valuable information, and, since they are not deflected by magnetic fields, they have the advantage of pointing straight back to their sources. Dark-matter annihilation can produce gamma rays in three ways: prompt photons, produced directly or via pion decays in hadronic cascades; final-state radiation, in which electron-positron pairs produce bremsstrahlung as they are created; and inverse Compton scattering, in which high-energy electrons or positrons scatter ambient starlight photons to high energies. The first two trace the dark-matter annihilation, while the third involves particles that have diffused significant distances over millions of years. Such signals are sought in the galactic center, in Milky Way subhalos, and in other galaxies. Fermi is mapping the sky from 100 MeV to 300 GeV and will either detect such signals or place interesting limits. Ground-based data from HESS, VERITAS, and other imaging atmospheric Čerenkov telescopes provide useful constraints in the high-energy range, above 100 GeV.

Finally, neutrinos from WIMP annihilations in the Sun are being sought with IceCube, a gigaton neutrino detector currently under construction at the South Pole. For a particular mass range, WIMPs that happen to pass through the Sun do so at sufficiently high speeds that they can lose enough energy in a nuclear scattering to become bound. The resulting WIMP over-density in the Sun leads to a substantial annihilation rate, which could be detectable.

High-Energy Particle Astrophysics

In high-energy particle astrophysics the past decade has brought into operation major observatories of gamma rays and cosmic rays that have provided a new view of the high-energy universe, probing the astrophysics of particle acceleration and non-thermal processes in a wide variety of physical regimes. Neutrino observatories are reaching the sensitivity level necessary to make the first detections of high-energy neutrinos from astrophysical sources, which would directly signal the acceleration of hadrons. The theory of astrophysical particle-acceleration processes, exploiting advances in computational plasma physics, has advanced to the point where useful contact with observations has become possible. As a result of the successes of the current instruments, there is great worldwide interest in initiating new, much more capable, high-energy observatories in the upcoming decade.

Over the past decade, ground-based arrays of imaging atmospheric Čerenkov telescopes have discovered very-high-energy ($E > 100$ GeV) gamma-ray emission from a wide variety of astrophysical sources, both galactic and extragalactic, increasing the source catalog by more than an order of magnitude (to almost 100 established sources). As important as the increase in number of detected sources, the quality of the data from Čerenkov telescope arrays has provided spatially resolved images of galactic sources such as supernova remnants, pulsar-wind nebulae, and the galactic center region, as well as spectral measurements of high-energy emission from active galactic nuclei with excellent (minute-scale) time resolution. The supernova-remnant observations suggest hadronic acceleration and subsequent emission of gamma rays, possibly confirming these as the main sites of proton acceleration in the galaxy, complementing their role as electron accelerators. The major Čerenkov telescopes currently operating are the U.S.-led VERITAS (a recommendation of the 2001 decadal survey) on Mt. Hopkins in Arizona (Figure 8.9), and the European-led instruments HESS in Namibia and MAGIC at La Palma.

The air-shower technique, as exemplified by the Milagro telescope, which operated until 2007 in New Mexico, is complementary to Čerenkov telescopes for detecting very-high-energy gamma rays. The latter achieve better angular and energy resolution, and hadron-background discrimination, than did Milagro. However, the air-shower technique permits a much wider field of view and continuous sky monitoring due to the essentially 100 percent duty cycle. Milagro carried out a survey of the Northern Hemisphere sky, detecting a number of sources in the galactic plane at energies above 10 TeV as well as diffuse emission from the plane at these energies. The Milagro sources were unexpected because of their relatively

FIGURE 8.9 The VERITAS atmospheric Čerenkov telescope array for gamma-ray astronomy on Mt. Hopkins in southern Arizona. SOURCE: Steve Criswell, Smithsonian Astrophysical Observatory.

high flux and hard spectra; they may signal the acceleration of hadrons. Another interesting result from Milagro, not yet fully understood, is the detection of anisotropy in the cosmic-ray arrival directions on the scale of 10 to 30 degrees (also reported by IceCube, the Tibet Array, and other experiments).

High-energy gamma-ray astronomy at GeV energies can be carried out very effectively by satellite telescopes. Fermi has worked flawlessly, providing continuous all-sky coverage of the gamma-ray sky, and is expected to continue operation through at least 2013. Early exciting results from Fermi include the discovery of many new pulsars not known from other wavebands, the detection of many new high-energy blazars, multi-GeV emission from gamma-ray bursts, and new measurements of the high-energy isotropic diffuse radiation and of the spectrum of cosmic-ray electrons up to 1 TeV. It is important to note that the Large Area Telescope (LAT), the main instrument of Fermi, was built by a successful international partnership that included astrophysics groups supported by NASA and particle-physics groups supported by DOE.

The spectrum of high-energy cosmic rays extends from GeV energies up to 10^{20} eV and perhaps beyond. Spaceborne instruments such as PAMELA and AMS-01—and balloon-borne experiments such as ATIC, CREAM, and TIGER—have played essential roles in measuring the composition of the cosmic rays up to energies of $\sim 10^{14}$ eV, just below the knee in the energy spectrum, and efforts are underway to extend the reach of such instruments to even higher energies. The existence of ultrahigh-energy ($E > 10^{17}$ eV) cosmic rays has been known for several decades, but their origin remains a deep mystery. At these energies, cosmic rays are detected by ground arrays of detectors or fluorescence telescopes, such as Auger South (described below) and the Telescope Array, that observe the air showers generated by cosmic-ray interactions in the atmosphere. Modern observatories using both techniques allow for hybrid detection of events. Recent worldwide efforts have begun to develop alternative techniques for detecting cosmic rays using radio signals.

In the previous decade, the AGASA ground array in Japan found surprising evidence suggesting a continuation of the energy spectrum past the GZK cutoff near 6×10^{19} eV that is expected from the interaction of protons with the cosmic microwave background. In contrast, the Fly's Eye HiRes fluorescence experiment in Utah reported a rollover in the spectrum at the highest energies. This confusing situation has now been clarified. Auger South, a hybrid array in Argentina with unprecedented collection area for ultrahigh-energy cosmic rays, started full operations in 2008 (Figure 8.10). It has reported several important new results: (1) a clear confirmation of the rollover of the energy spectrum; (2) an indication that the

FIGURE 8.10 Pierre Auger Observatory of ultrahigh-energy cosmic rays in Mendoza Province, Argentina. Shown is one of the particle detector tanks of the 1,660 units covering 3,000 square kilometers and one of the four fluorescence telescopes that overlook the array. SOURCE: Courtesy of the Pierre Auger Observatory.

highest-energy cosmic rays are distributed anisotropically; (3) a strong limit on the photon fraction that rules out, in large part, top-down models for the production of ultrahigh-energy cosmic rays; and (4) the strongest limit to date on the flux of cosmogenic neutrinos between 10^{17} and 10^{19} eV.

Neutrino telescopes offer a compelling avenue for understanding high-energy astrophysics and to probe for dark-matter annihilations. The large IceCube detector under construction at the South Pole will be fully operational in 2012 and will search for TeV and PeV astrophysical sources of neutrinos with an expected capability that should permit the first high-energy-neutrino source detections. The detection of ultrahigh-energy neutrinos is challenging for existing detectors. New techniques are needed to establish their existence firmly and to extract useful physics and astrophysics. For cosmogenic neutrinos, the best energy range is around 10^{18} to 10^{19} eV where the flux is maximal. Various R&D activities are ongoing, led by radio-detection techniques such as radio antennas in ice (e.g., RICE and prototypes in the IceCube holes); antennas on balloons (e.g., ANITA) and on the ground to measure the Čerenkov signal produced by the Askaryan effect; and antennas at extensive-air-shower arrays (e.g., Auger and LOPES) for the detection of radio emission from neutrino-initiated atmospheric showers. R&D programs on acoustic detections of pulses due to heat produced by particle cascades in sea water, ice, and salt are ongoing.

FUTURE PROGRAM

The activities considered by the panel receive federal support from four sources: the NASA Astrophysics Division, the NSF Division of Astronomical Sciences, the NSF Division of Physics, and the High Energy Physics program within the DOE Office of Science. Therefore, the funding environment is complex. The panel presents a science program that cuts across all four agency funding units and constitutes a complete program with complementary contributions from the units. It also presents recommendations on infrastructure issues. The panel's recommendations are based primarily on considerations of whether a particular project, mission, or activity addresses a high priority and a compelling science question or questions. In addition, the panel considered whether an activity merits what it would cost, would produce results related to theoretical predictions or opens a new capability with high discovery potential, is technically feasible or incorporates verifiable technology development, and is of interest in a worldwide context and does not unnecessarily duplicate efforts outside the United States. Table 8.8 at the end of this report indicates the relationship between the panel's recommended activities and the science priorities identified by the Astro2010 Science Frontiers Panels.

Science Program

In gravitational-wave astrophysics the panel recommends complementary ground- and space-based programs supported by NSF and NASA, respectively. Ongoing and expected improvements in ground-based detectors may well give us the first detection of gravitational waves in the next decade. The progression from detection to astrophysical insight will require subsequent exploitation of the higher-frequency portion of the spectrum accessible from the ground. In addition, opening the lower-frequency portion of the spectrum is predicted to provide access to many detectable astrophysical sources, different from those relevant to ground-based detectors. The key frequency range for many of the most exciting investigations is 10^{-4} to 10^{-1} Hz, which requires laser-interferometric measurements over extremely large baselines, which are possible only in space. The only option for opening this window is LISA. The LISA technology to enable these measurements has advanced considerably over the past decade, and the scientific case is compelling for a gravitational-wave mission to be the flagship space mission for astrophysics in the coming decade. While all of the risk cannot be eliminated from such a mission (or any space mission), even with thorough ground testing and a space precursor, the scientific case is strong enough to warrant moving forward even at a medium level of technical risk. To achieve that level of residual risk, however, requires successful completion of the LISA Pathfinder (LPF) precursor. Given the great astrophysical importance of opening up the low-frequency gravitational wave band, the panel recommends that the LISA mission be given the highest priority for a new start in the next decade. Furthermore, given the extensive technology development that has already been completed, the expected short time until LPF launch, and the need to maintain momentum in the U.S. community and guarantee a smooth transition to a joint NASA-ESA mission, the panel recommends that NASA funding of LISA begin immediately, with continuation beyond LPF contingent on the success of that mission.

The LPF is an ESA technology-demonstration mission, scheduled for launch in 2012, that will test several of the LISA subsystems in the space environment. This includes the Gravitational Reference Sensor (GRS), the Interferometry Measurement System (IMS), and two micro-Newton thruster designs. The tests to be performed are end-to-end performance of a shortened version of one LISA arm, a complete test-mass-to-local-spacecraft measurement, and modeling of physical disturbances. As described in more detail below, the spacecraft environment for LPF will not be the same as that for LISA; LPF will be placed at the Earth-Sun L1 Lagrangian point rather than in an Earth-trailing orbit. LPF development is complete, and flight hardware is under construction. It is intended that analysis of the LPF results be finished before the beginning of LISA Phase B.

The panel identified four major areas of technical risk within the LISA project: (1) the Disturbance Reduction System (DRS), consisting of the GRS, micro-Newton thrusters, and accompanying control system; (2) the IMS; (3) the spacecraft environment; and (4) long-range interferometry. Risks in the first two areas are either entirely or partially mitigated by a successful LPF. Although the DRS must demonstrate a 5-orders-of-magnitude reduction in system noise compared to previous similar implementations, extrapolation from current ground-based tests indicate that it will meet requirements. The GRS and thrusters (noise and minimum-impulse bit) will be tested to LISA levels on LPF. However, the DRS must perform as a system, and the LPF architecture and space environment are different from those of LISA. Also, on LPF the second mass will be slaved to the first, rather than acting as a free test mass. With these differences, the overall DRS performance on LPF is predicted by models of the system to be a factor of 10 worse than on LISA, even if the subsystems perform as designed (i.e., at LISA levels). The performance of the LISA system must also be evaluated through modeled extrapolation from LPF performance of the subsystems. The panel also notes that two technologies for the thrusters are under study, and both will be tested on the LPF, although the necessary long lifetime will not be verified.

The second critical area is the interferometric phase measurement system. This has been extensively studied on the ground, including efforts at the Jet Propulsion Laboratory, and will be verified at LISA performance levels on LPF (Figure 8.11). The main residual risks after LPF relate to the differing thermal, magnetic, and gravitational environment and the extent to which LPF results can be used to extrapolate to LISA performance. Likewise, while the LPF will test the operation of the lasers, it will not demonstrate the enormous path lengths planned for LISA and the relevant spacecraft subsystems and software for control. The current program calls for ground testing with space performance evaluated by analysis and extrapolation.

There are a number of programmatic issues related to LISA that remain to be resolved. Given the cost of LISA, the concentration of technology development in Europe, and the fact that LPF is being executed entirely as a European program, a collaboration with ESA is appropriate and recommended. The nature of the collaboration, in particular the division of responsibilities, needs to be established. The panel recommends commitments based on statements of work, and not cost caps (as is currently proposed by ESA), for each partner. Its review of costs and schedule led the panel to believe that those provided by an independent evaluation are more plausible than those provided by the project, and the panel adopts the former numbers in its budget analysis below.

A major consideration in the panel's deliberations has been the level and nature of risk associated with the LISA mission. The LISA components are integrated to a degree unprecedented in an astrophysics space mission, making the risk of total mission failure relatively high compared to typical science missions. A significant

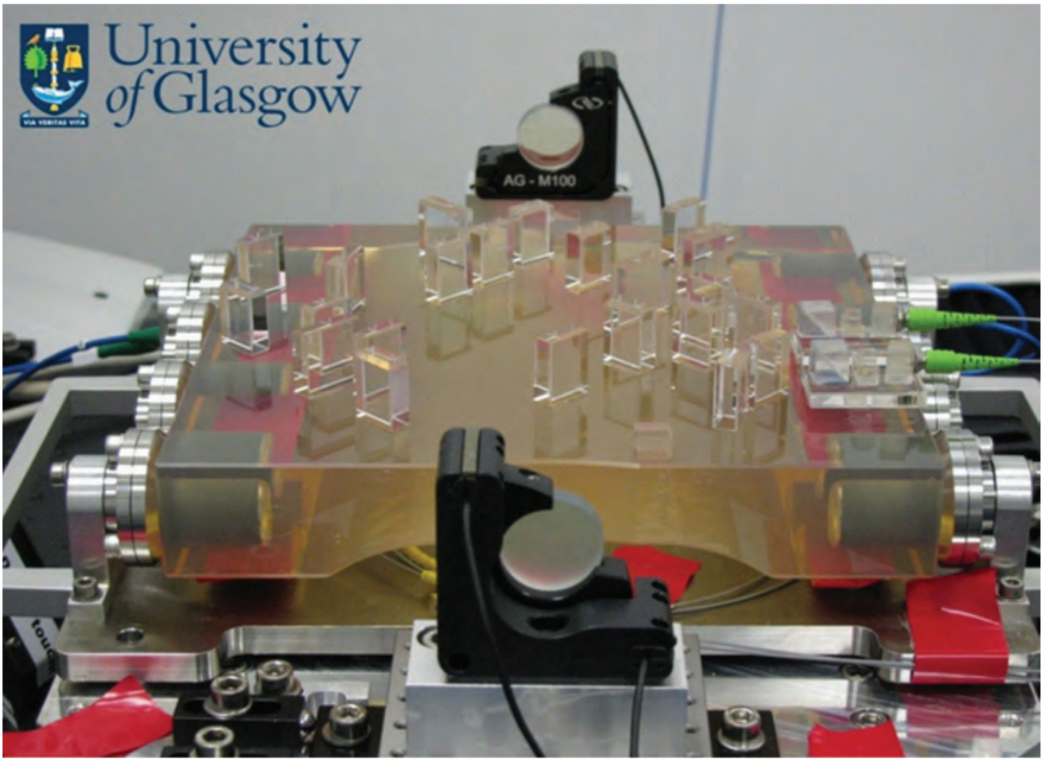


FIGURE 8.11 Fully bonded optical bench. SOURCE: H. Ward, Department of Physics and Astronomy, University of Glasgow.

investment in systems engineering early in the program, with clear responsibilities assigned to NASA and ESA, is essential. In addition, even with a successful LPF test of the critical subsystems at the LISA level (the GRS, the IMS, and the micro-Newton thrusters), certain risks will remain associated with the different environment and the lack of multiple spacecraft in LPF to test the long-range interferometry. Nevertheless, if LISA fails to meet the required strain sensitivity, the degradation in science is graceful. Consider, for instance, a situation in which LPF succeeds and all LISA subsystems perform at the required level for LPF, but the low-frequency acceleration is only as extrapolated from LPF performance. Here, the number of detectable galactic binaries below 0.7 mHz would decrease (which is a small loss), and a few verification binaries would be eliminated. The main effect is that the angular positional accuracy for black hole mergers would decrease substantially, and the likelihood of finding electromagnetic counterparts to massive black hole mergers would worsen substantially. However, most of the other science is retained, including accurate measurements of the masses, spins, and luminosity distances of massive black hole mergers. Now consider a situation in which LISA sensitivity is significantly worse than the extrapolation from LPF (by a factor of 3), perhaps due to errors in modeling, failure to control the spacecraft environment, or unexpected problems in the GRS or IMS. In this case, the main effect would be on the expected rate of detection for extreme-mass-ratio inspirals, which would decrease from ~50 per year to ~2 per year or less. This could result in no detections of extreme-mass-ratio inspirals over the life of the mission, given the astrophysical modeling uncertainty. However, while detections of massive black holes would be reduced by a factor of ~2 and detections of galactic white dwarf binaries by a factor of 10, the science from those detected (such as accurate massive black hole masses and spins) would be unchanged.

LISA science is rather robust against failure of one of the gravitational-reference sensors, which would result in the loss of a single arm of the interferometer; in fact, there is no science requirement that LISA maintain six working links throughout the mission. Such a failure would decrease the detection rates by a factor of only 1.5 to 3, because the system is designed for two redundant laser links. The most significant impact would be degradation in position measurements of merging black holes, resulting in the loss of positional information critical for electromagnetic follow-up observations. Another situation might be that the thruster lifetime is not as long as predicted; the mission lifetime would be correspondingly curtailed, with a reduction in the number of events detected. Nevertheless, in any of these scenarios, gravitational waves from coalescing black holes and the centers of galaxies will still be observed, greatly advancing astrophysics studies and tests of strong-field general relativity.

For completeness, the panel considered a more extreme failure scenario: the complete loss of one of the three spacecraft would result in virtually no remaining science. All three spacecraft must operate, with two DRSs operational on at least two of them, for a successful mission.

The panel is of the opinion that the enormous potential science return of the LISA mission, and its robustness against degradations in performance, justify accepting these risks. The panel also thinks that, after a successful LPF mission, further testing and analysis would have limited return, and therefore that LISA should be funded now without further delay but with a gate to continuation through the final project phases to launch based on the performance of LPF. The spending rate during the initial years of the LISA project is low, and support in the pre-LPF phase will maintain the project teams. The panel also notes that LPF may be neither a complete success nor an outright failure. Its expectation is that an evaluation of expected LISA performance will have to be evaluated in light of LPF test results, enabling a re-evaluation of costs and scientific benefits. If for some reason LPF were a complete failure, then the panel would not support a continuation of the LISA mission unless a decision were made to repeat the pathfinder.

Despite LISA's impressive capabilities, LISA will not reach the very lowest gravitational-wave frequencies that probe the stochastic background produced in the early universe via relic gravitational waves. A promising approach for detecting this cosmological background is provided by high-precision timing of millisecond pulsars, using a network of pulsars distributed in the galaxy as a gravitational-wave detector. As shown in Figure 8.2, a network consisting of 20 pulsars, each with a variance of 100 nanoseconds and observed for 5 years, can achieve sensitivities comparable to the stochastic background produced by the radiation from binary systems of merging supermassive black holes. As for LISA, this astrophysical source of "noise" would be a discovery in itself. The amplitude and shape of such a spectrum depend on a hierarchical galaxy-formation model, and so the astrophysical information derived from positive results, or the upper limits in the absence of discovery, will advance knowledge of the evolution of the universe.

To explore these exciting new areas, the panel recommends that NSF support a coherent program in gravitational-wave detection through the timing of a sample of millisecond pulsars. Such a program must begin with searches (which are already underway) for the additional millisecond pulsars with small timing noise that are necessary to complete the pulsar "array" required for detection. Technological challenges include compensation for the effects of the interstellar medium (especially critical for extending the pulsar sample to larger distances) and refining and extending algorithms to improve sensitivity and reduce pulsar-timing residuals. It will also be necessary in the long term to develop techniques to use effectively future array telescopes, which are attractive because of their planned large collecting areas, for pulsar timing. Progress toward these technical goals should be monitored over the course of the decade. The panel was provided with an estimate of $66 million as the cost of such a program over this period. The program is a collection of activities that include observing with radio telescopes, computing and signal processing hardware, salaries, and travel. Such an estimate is difficult to verify but probably is indicative of the level of support needed. The panel notes that the pulsar-timing program requires that Arecibo, or a future telescope with similar capabilities for pulsar timing, be available to the program.

Finally, the panel also recognizes that while the ground-based gravitational-wave detectors are likely to produce their first detections in the coming decade with expansions already underway, exploitation of their full scientific potential will require improvements in sensitivity and extension of the sensitive band to lower frequencies. The ground-based detectors are sensitive to a unique mass range for sources in our galaxy. The panel therefore recommends broad agency support for continued improvement of technologies for ground-based gravitational-wave detectors, including instrument technologies and data analysis and interpretation.

As complements to a vigorous program in gravitational-wave astrophysics, the panel believes that other activities that test general relativity and theories of gravity are scientifically valuable. However, as stated in the summary of the science case above, theoretical guidance is lacking as to how one might expect general relativity to break down. Therefore, the panel recommends that these activities receive support based on the following principles:

- Favor tests of general relativity on scales and domains where it has not been well tested so far. That means emphasizing (1) quantitative tests of its strong-field predictions, for example, black holes, gravitational waves, and cosmology, and (2) tests in either strong- or weak-field regimes that are on scales larger than the solar system.
- Favor clean tests in which the physics of the testing system is simple and calculable. That is, emphasize observations that test general relativity with a minimum of astrophysical parameters to be modeled and determined. Examples are binary pulsars and extreme-mass-ratio black hole inspirals.
- Favor tests that substantially improve the precision of basic parameters such as the parametrized post-Newtonian parameters, the rate of change of G, and increasing the accuracy of tests of the strong and weak equivalence principles, *provided* that results can be obtained at moderate cost so that there is high rate of return of science for the cost.

LISA is the panel's top priority for testing relativistic gravity. The direct detection of gravitational waves, the verification of their propagation speed, and the nature of their polarization would in itself provide a significant test of general relativity. Advanced LIGO is likely to provide the first tests of strong-field gravity by measuring waveforms of black-hole and neutron-star mergers. However, LISA's unprecedented sensitivity to a wide range of astrophysical sources will extend these tests to many objects, including black holes at the centers of galaxies, and to much higher signal-to-noise measurements. LISA is the most direct route to testing theories of gravitation in the strong-field regime and will provide the most data on the effects of gravitation on mostly unexplored galactic scales.

High-precision measurements of distance and growth of structure as a function of redshift offer exciting opportunities to test general relativity on cosmological scales. Data deviating from the precise relationship between distance and growth predicted by general relativity would signal its breakdown. Tests of general relativity on cosmic scales are deeply entwined with fundamental questions related to cosmic acceleration and dark energy. A strong electromagnetic-survey program providing distance/growth tests of general relativity as well as high sensitivity to the dark-energy equation of state would have a profound impact on our understanding of general relativity. These survey programs are being considered by other Astro2010 panels.

A new lunar laser ranging (LLR) program, if conducted as a low-cost robotic mission or an add-on to a manned mission to the Moon, offers a promising and cost-effective way to test general relativity and other theories of gravity (Figure 8.12). So far, LLR has provided the most accurate tests of the weak equivalence principle, the strong equivalence principle, and the constancy in time of Newton's gravitational constant. These are tests of the core foundational principles of general relativity. Any detected violation would require a major revision of current theoretical understanding. As yet, there are no reliable predictions of violations. However, because of the importance of these principles, the panel favors pushing their limits when it can be done at a reasonable cost. The installation of new LLR retroreflectors to replace the 40-year-old ones might provide such an opportunity.

The panel emphasizes again its opinion that experiments to improve measurements of basic parameters of gravitation theory are justified only if they are of moderate cost. Therefore, it recommends that NASA's existing program of small- and medium-scale astrophysics missions address this science area by considering, through peer review, experiments to test general relativity and other theories of gravity. The panel notes that a robotic placement of improved reflectors for LLR is likely to be consistent with the constraints of such a program. It returns to this recommendation below in the context of a recommendation to augment the Explorer program.

A balanced program in particle astrophysics and gravitation must include experiments designed to identify the particle or particles that make up the dark

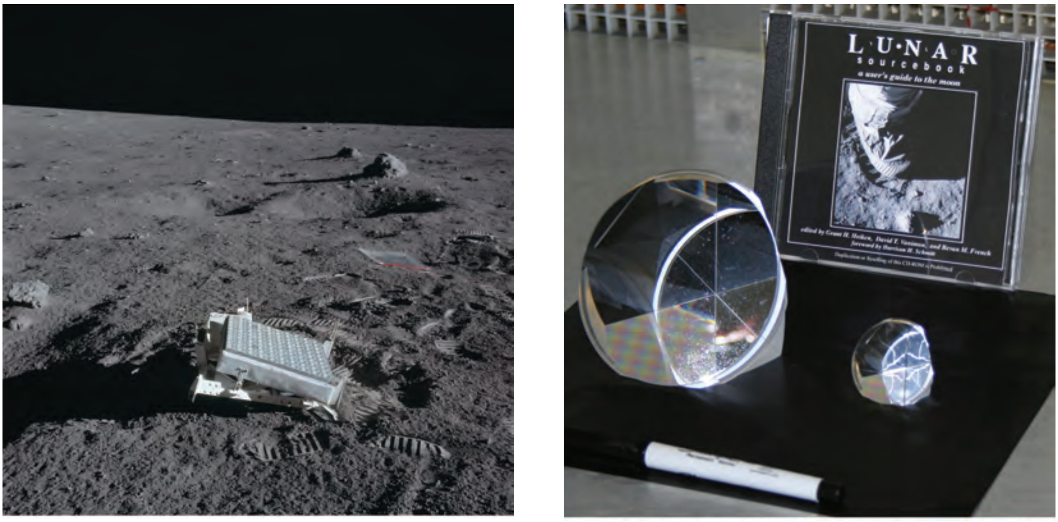


FIGURE 8.12 *Left:* Retroreflector placed on the lunar surface by Apollo astronauts. *Right:* A 10-centimeter solid corner cube developed for the next generation of lunar-laser-ranging experiments, shown next to a 3.8-centimeter engineering model. SOURCE: *Left:* NASA. *Right:* Douglas Currie, University of Maryland, College Park.

matter. The panel's recommendations address only indirect astrophysical searches for dark matter, although these recommendations reflect cognizance of the context provided by direct laboratory experiments and accelerator experiments. Astrophysical indirect searches for dark matter involve both particle and gamma-ray experiments. To distinguish between a dark-matter signature and a previously unknown, local astrophysical source of high-energy particles, it is important to measure the positron spectrum to higher energies—if possible up to 1 TeV, where the expected (combined electron and positron) spectrum rolls off—to see whether the positron excess persists through that entire energy range. Because of the high cost of putting a larger magnet into orbit, positron measurements up to 1 TeV are most likely to be done best with ultralong-duration balloon flights. Improved sensitivity to anti-protons and heavier anti-nuclei, up to helium, is also desirable. Again, the masses of the needed detectors argue for the value of the ultralong-duration balloon option. Improved modeling of particle and anti-particle production from astrophysical sources is also needed.

Much can also be done by extending the electron spectrum to higher energies and better sensitivities. In the 1- to 10-TeV electron-energy range, atmospheric Čerenkov telescopes have so far yielded the best results. With an experiment having an order of magnitude more sensitivity than VERITAS or HESS, the secondary tail of local e^+e^- production should be visible at energies above a few TeV. This would provide helpful information about very local production and propagation and would give a baseline for measurements at 100 to 1,000 GeV to help interpret any excess gamma rays there from dark matter or pulsars.

The gamma-ray measurements of atmospheric Čerenkov telescopes also provide essential information. Current constraints on dark matter annihilation models by HESS (and at lower energies by Fermi) are within an order of magnitude of either detecting or ruling out most dark-matter annihilation scenarios that can produce the PAMELA positrons. The two plots in Figure 8.13 demonstrate the potential of a next-generation Čerenkov array. A significant region of parameter space could potentially be excluded (or the effort might result in a detection!) through observations of nearby dwarf galaxies. Therefore, increasing the sensitivity of atmospheric Čerenkov telescopes by another order of magnitude is the panel's top priority for exploring the nature of dark matter.

The future gamma-ray astronomy instruments under development are expected to improve the sensitivity by about an order of magnitude over the current generation of experiments and to cover a wider energy range. These improvements are well motivated scientifically, as the success of VERITAS, HESS, MAGIC, and Milagro has led to important progress in TeV gamma-ray astronomy over the past decade. These instruments have greatly increased the number of known sources and have studied the high-energy astrophysical processes in these objects in great detail and with good temporal, spectral, and angular resolution. The performance

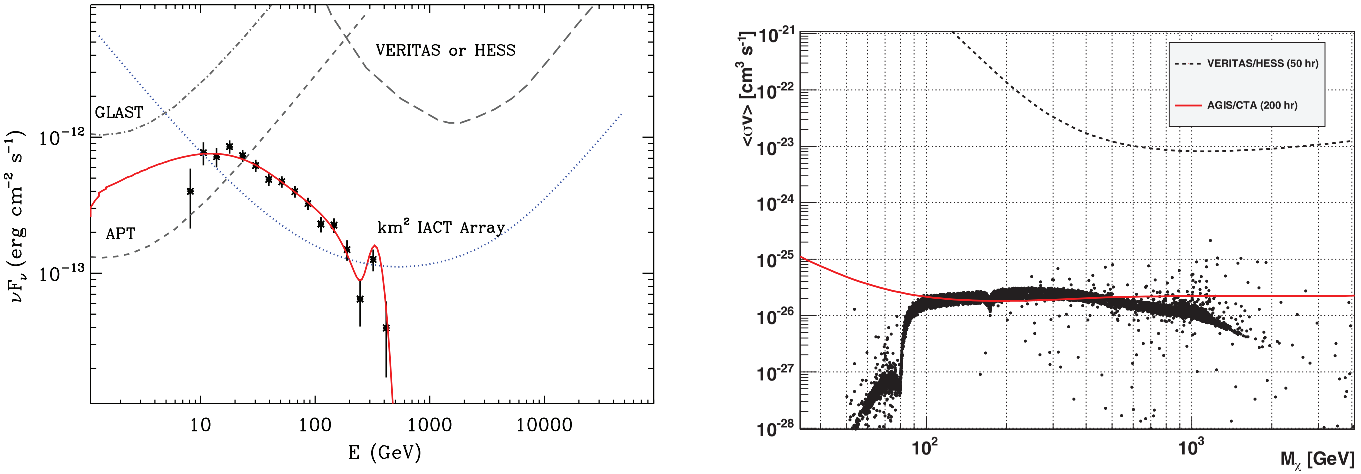


FIGURE 8.13 *Left:* Predicted gamma-ray signal from the dwarf spheroidal galaxy Ursa Major for a dark-matter neutralino with mass of 330 GeV. *Right:* Plot showing the ability of the next-generation Čerenkov array to exclude predicted dark matter candidates. Each point represents a prediction of a model that is a supersymmetric extension of the standard model of particle physics. SOURCE: *Left:* F. Aharonian, J. Buckley, T. Kifune, and G. Sinnis, High energy astrophysics with ground-based gamma ray detectors, *Reports on Progress in Physics* 71:096901, 2008. *Right:* Courtesy of Matthew Wood, University of California, Los Angeles.

of current detectors can be improved by one or two (in opposite hemispheres) atmospheric Čerenkov telescope arrays with a collection area of the order of 1 square kilometer, a larger field of view, and improvement by a factor of two to three in angular resolution. Europe is moving forward with the Čerenkov Telescope Array (CTA), and compelling concepts for future improvements in capability are being studied in the United States. Proposed for development in the United States is the Advanced Gamma Ray Imaging System (AGIS; Figure 8.14), a large atmospheric Čerenkov array comprising 36 new-concept telescopes in a regular grid, separated by 120 to 150 meters. The collecting area of AGIS as proposed is about a factor of 10 larger than that of VERITAS. The sensitivity of this array is improved by a factor of 10 to 20 with respect to the previous generation. The telescope concept is a double-mirror Schwarzschild-Couder design that has the potential to widen the field of view of a single telescope while reducing aberrations that limit the angular resolution.

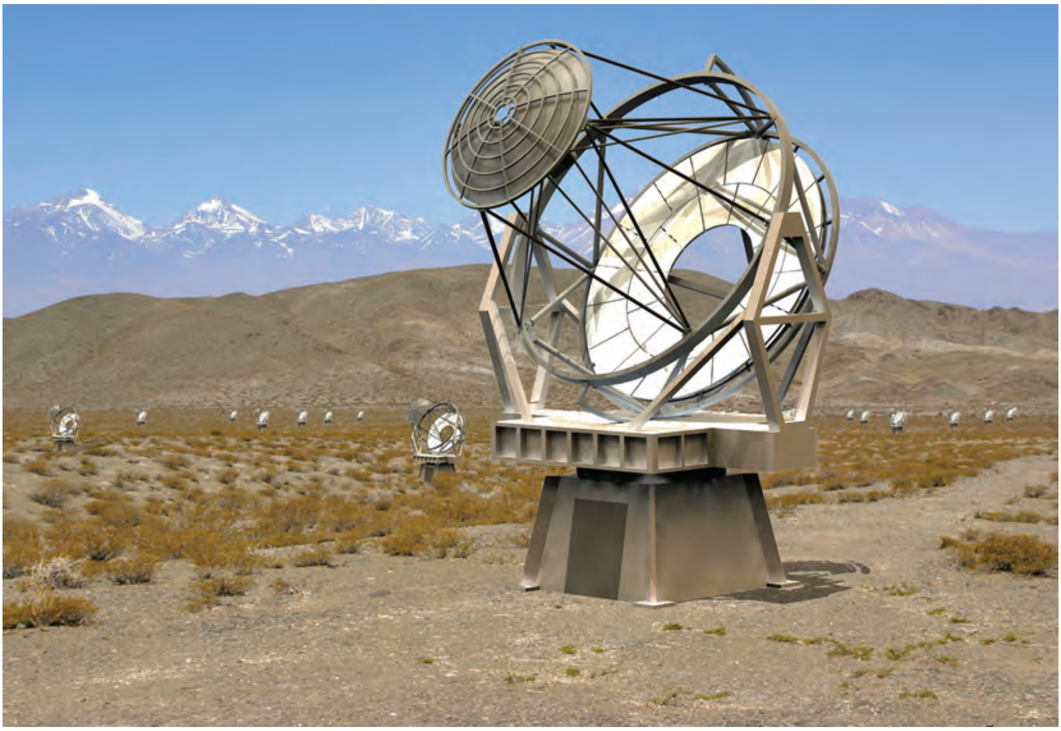


FIGURE 8.14 The AGIS telescope concept, based on a Schwarzschild-Couder optical design, which has the advantages over traditional designs of a wider field and shorter focal length. This design enables compact, multi-pixel, wide-area detector arrays, reducing cost and increasing reliability. SOURCE: Jim Buckley, Washington University.

Although the panel strongly endorses the goals of increasing the field of view and improving the angular resolution, the AGIS proposal raised concerns about the project schedule and cost. The R&D phase required by such a new telescope, with its tight constraints on the roughness of the mirror surfaces, will inevitably introduce a time lag with respect to the European schedule. CTA has adopted the well-established Davies-Cotton telescope concept to ensure timely readiness of the array and its science outcome. Moreover, the CTA strategy, which combines different mirror sizes and varying baselines, enables lowering the energy threshold and covering a wider energy range than the AGIS array as proposed.

The panel is concerned about the total cost of a stand-alone U.S. effort. It is also concerned that AGIS will not be competitive with the European CTA experiment, which is on a considerably shorter time schedule. Given these considerations, the panel recommends that there be U.S. involvement, jointly supported at a significant level by DOE and NSF, in an international Čerenkov array with a square-kilometer effective area. An international collaboration seems necessary for the full potential of the next-generation array to be realized. In addition, R&D by the U.S. group on new telescope technologies should be encouraged. This could lead to the possibility of adding or replacing telescopes with a new design in the future. The panel also encourages R&D on the cameras and photodetectors. Finally, the planning of the future AGIS-CTA science program should include provisions for broad access to the data. This should most certainly include a public data archive and might also include observing time being made available to the community on a competitive basis.

Unfortunately, the planning and development of the next-generation atmospheric Čerenkov arrays are at such an early stage that an accurate cost appraisal is not possible. The European CTA consortium has provided the panel with an estimate of the cost of the CTA of €180 million. However, a joint European-U.S. array might involve technologies substantially different from those now being considered for the CTA. Furthermore, European cost accounting typically does not include significant in-kind contributions, largely in the form of salaries, that quite possibly could double the true cost. The independent cost evaluation carried out as part of the decadal survey review activities concluded that a cost appraisal made at this time would be very uncertain. Therefore, the panel doubled the CTA estimate to account for in-kind contributions, and then took half that as the U.S. contribution, arriving at approximately $200 million for the U.S. share. Although the panel strongly supports the goals of an international Čerenkov array, it is quite clear that a thorough review of cost and technical feasibility will have to be carried out by the funding agencies before a decision can be made to proceed.

In addition to a strong program in atmospheric Čerenkov detection, the panel recognizes the promise of HAWC, a higher-energy, wide-field-of-view, and high-duty-cycle observatory. Based on a compact array of water Čerenkov detectors at high altitude, HAWC would operate at a lower energy threshold and would be an order of magnitude more sensitive than the previous generation of wide-field TeV gamma-ray telescopes, allowing it to carry out deeper surveys and to catch weaker high-energy transient events. This kind of capability will be best utilized in conjunction with the operation of Fermi to guarantee the ability of observing transient sources over a wide energy range. A new approach to atmospheric Čerenkov detection, aimed at the very highest energies, is space-based observation of air showers that enables the use of a very large volume of the atmosphere as a detector. Participation in international efforts to develop this technique is a promising path for the future. The panel concluded that a balanced program must include support for these smaller, peer-reviewed programs.

Recent cosmic-ray observations have challenged the current understanding of local cosmic ray sources and their propagation to Earth. Dark-matter models have been proposed to explain these findings. A continuing ability to observe galactic cosmic rays is important in order to disentangle dark-matter signatures from nearby astrophysical sources. As described above, experiments carried by balloons have made an important contribution to dark-matter and cosmic-ray studies. The panel expects that the capability of balloons to transport the large masses required for achieving interesting event rates—enabling a number of different measurements that allow cross-checking of results with different systematic errors—will be critical to future progress. One example of a promising experiment is a search for anti-deuterons, with a very low expected background; it is anticipated that an ultralong-duration balloon (ULDB) experiment would improve limits to be set by AMS by an order of magnitude. Another example is the radio detection of ultrahigh-energy neutrinos, with expected improvements in sensitivity of an order of magnitude provided by the long exposures of ULDB flights, particularly if trajectory control enables the directing of flights away from populated areas with higher backgrounds. Because of its promise, the panel recommends that NASA maintain support for the ULDB program to provide the capability for a strong program in cosmic-ray and neutrino detection.

At ultrahigh energies, the ability to point back to the most powerful cosmic accelerators has been reported recently by the Pierre Auger Observatory. The next decade may see the identification of the sources of ultrahigh-energy cosmic rays and the ability to test hadronic interactions at energies above those at which laboratory accelerators operate. The limited statistics of the events observed at the highest energies have led the international Pierre Auger collaboration to propose a much larger northern observatory in Colorado to augment the results obtainable from continued operation of Auger South. Funding for the U.S. fraction of this international effort could be accommodated in a high-budget scenario for the decade, as discussed in the panel's budget analysis below, in a program that preserves the balance between small and large programs.

The construction of the IceCube detector at the South Pole is currently underway. IceCube may open the window of neutrino astronomy early in the next decade. In addition to the continuation of this effort in TeV/PeV neutrino astronomy, the panel supports the development of complementary techniques for neutrino detection at the highest energies in order to study neutrino interactions at energies above those achieved by laboratory accelerators.

Infrastructure Issues

The Base Program

All four government funding-agency units reserve a fraction of their budgets for a "base" program that supports the operation of existing projects and facilities, technology development, laboratory programs, data analysis, computation, and theory. At NASA and NSF, a substantial fraction of the support to university groups is provided through peer-reviewed programs open to individual investigators. In the DOE program, the individual-investigator component is relatively small, with support for particle astrophysics targeted chiefly at established efforts at DOE laboratories and university groups engaged in related project development and analysis. In all cases, the panel considers the base program to be critical to the realization of the scientific goals of the projects and facilities and to carrying out the precursor studies and prototyping that enable future activities. The panel supports continued investment in these areas with interagency coordination that fosters peer-reviewed competition. Under no plausible budget scenario would the panel recommend a reduction in the fraction of the program devoted to the base program; in fact, it recommends augmentations in certain areas as discussed below.

Technology Development

In all areas of astronomy and astrophysics research, the invention and the development of innovative technology have been key to progress. For the field to remain vital, it is essential that adequate funding be made available to encourage the birth of new technical concepts and to allow promising concepts to be brought to sufficient maturity so they can be incorporated into new experiments with minimal risk. However, enabling technology development has been a challenge for all federal agencies, given competition for resources and the uncertain payoffs of investing in technology the application of which is not guaranteed. Technologies are evaluated according to their technology readiness level (TRL): TRLs 1 to 4 are associated with the validation of basic concepts, analytical evaluations of expected performance, and breadboard testing of key components; TRLs 5 to 7 are associated with the fabrication and testing of prototypes in a representative environment; and TRLs 8 to 9 apply to fully representative systems that have been successfully deployed. The cost of technology development grows steeply with increasing TRL. NASA, NSF, and DOE have all struggled with diverse mechanisms to meet those costs in their own way, with mixed success. In the view of the panel, the situation has not been optimal at any of the three agencies.

Over the years, NASA has had various grant-funding mechanisms to seed development of new technologies at TRLs 1 to 3. However, a significant gap has existed for mid-TRLs (4 to 7). The costs of development of representative prototypes typically requires funding at a level of multimillion dollars per year, even for a single subsystem. There has been no mechanism to provide that level of funding for projects that have not yet been awarded Phase A approval. Yet "technical immaturity" has been cited as a rationale for the denial of Phase A approval for projects that were otherwise ranked very highly on scientific grounds. In the President's Budget for FY2011, significant funding was made available for far-term investment in space technology. This is a very promising development; however, it remains to be seen how responsive that new funding line will be to the needs of astrophysics missions, both in the Explorer and major mission lines. In the panel's view, an augmentation of $300 million over the decade to the base program at NASA is required for technology development to support missions addressing the science areas of particle astrophysics and gravitation. The panel believes that a further augmentation for other science areas is also justified, but a detailed analysis is beyond the scope of this panel. This would represent a reasonably small fraction of the total level of funding identified for the new space technology investment and so may not be incompatible with current NASA plans.

At NSF, the gap in technology-development funding parallels the well-known gap in project funding between the Major Research Instrumentation (MRI) awards and the Major Research Equipment and Facility Construction (MREFC) awards. MRIs can be approved at relatively low TRL values, and so the cost of technology development is appropriately absorbed into the cost of the project. However, given increasing attention to cost and schedule validation in the MREFC approval process, all future MREFCs will be required to demonstrate a high level of technical maturity in all enabling subsystems before they can proceed into development.

Because costs for MREFC construction do not come out of the divisional budgets, they are not, in principle, in competition with grants to university-based investigators. But in the present system pre-project-approval technology-development costs must come from the sponsoring division. It is not unusual for costs associated with bringing key technologies to final design readiness to amount to 30 percent or more of the construction costs. The panel therefore recommends that the division's budgets include an augmentation for technology development.

Funding for particle astrophysics projects at DOE has come mostly from the Office of High Energy Physics (OHEP) within the Office of Science. Until recently, funding for technology development for future projects was handled naturally through the university grants programs and through base funding at the national laboratories, and this system worked reasonably well. In the last few years, however, this system has evolved in a more conservative direction, and funding for technology development is very limited until a given project reaches a relatively advanced state of approval. The rationale has been to avoid spending significant amounts of money on projects that may never be constructed. That has led to a state of affairs similar to that described above for NASA and NSF: there is a noticeable gap in the availability of funding for mid-TRL technologies. This panel recommends that OHEP should also increase the funding available for candidate projects that are still awaiting approval. Therefore, the panel recommends an augmentation in the base program at OHEP to support technology development for experiments addressing the science areas of particle astrophysics and gravitation.

NASA's Explorer Program

For several decades, NASA's Explorer program has supported relatively small missions with focused scientific goals. In the current implementation of the program, proposals for small-scale missions (SMEX; capped at around $150 million, excluding launch), mid-scale missions (MIDEX; capped at around $250 million, excluding launch), and Missions of Opportunity (instruments to be flown on non-NASA missions) are selected according to peer review. The missions are led by a principal investigator who has the ultimate responsibility for scientific leadership, management, and the overall success of the mission. In the past, the goal of the Explorer program—in astrophysics and heliospheric science combined—has been a launch rate of about one mission per year, responding to new scientific opportunities in a timely way. The program has produced some remarkable recent successes, such as the Wilkinson Microwave Anisotropy Probe, the Swift high-energy transient mission, and the Galaxy Evolution Explorer. In the past decade, however, cost overruns in large missions and overall NASA budget constraints led to severe cutbacks in the Explorer program, and only three new missions in astrophysics were approved. The current oversubscription rate is very large, with a 6 percent success rate in the most recent Explorer competition. The panel recommends that NASA's Explorer program be restored to its previous funding level and launch rate. Within the particle astrophysics and gravitation science area, the panel suggests that these funds be used to carry out—if justified by peer review—innovative missions that address tests of general relativity and other theories of gravity. It might also be possible to carry out indirect searches for dark matter on the Explorer platform. One of the strengths of the Explorer program is that missions are chosen based on the strength of the science case and on technical feasibility, independent of the specific science topic. The panel thus recommends an augmentation to the Explorer program in the next decade. Adding three Explorer opportunities over this period, at an average cost of $300 million per opportunity (including an allowance for

Missions of Opportunity and launch costs), would require an augmentation of $900 million. The panel intends that this recommendation pertain to all areas of astrophysics, exclusive of the heliospheric science portion of the Explorer program, which is outside the scope of this study.

NASA's Balloon Program

NASA's balloon program supports the astrophysics community in several unique ways. It provides a testing environment that qualifies new technology for spaceflight missions, a critical step in the development of a payload. It provides a platform for certain experiments that can be conducted more cheaply and quickly on suborbital balloon flights than in space. It provides a means to conduct experiments whose large masses would make an orbital implementation prohibitively expensive. Finally, it provides an environment in which students and experimentalists can be trained for future leadership in space missions and suborbital experiments. In the discussion above, the panel recommends that ULDB development be completed and that ULDB flights be supported for the purposes of conducting experiments in the detection and characterization of dark-matter and cosmic-rays. The balloon program cuts across many scientific areas and is also being considered by other Program Prioritization Panels. For this panel's purposes, it recommends that the development of technologies needed for ULDB flights be completed and that a ULDB program of one or two flights per year be supported, including their payloads, possibly replacing some long-duration balloon flights. The panel estimates that the cost for this capability requires an augmentation to the balloon program of about $250 million over the next decade.

Theory

The panel recognizes the important role that theoretical investigations play in the advancement of knowledge in astronomy and astrophysics. "Theory" encompasses a wide range of activities and includes "blue sky" theory, pencil-and-paper analysis, simulation, and investigation of data analysis techniques. The recent success of NASA-supported theoretical investigations of astrophysical gravitational-wave sources also highlights the importance of strategic theory in enabling new avenues of investigation. The panel supports a strong base program at the funding agencies that includes theory as one of its components and that is flexible enough to include new interdisciplinary topics. In addition, the panel draws attention to two particular issues. The first is the difficulty of properly supporting areas that straddle the traditional boundary between physics and astronomy. The second is the lack of support at universities and research centers for cluster computing, which is necessary as a testbed for parallel-computing programs at the national supercomputing centers.

Most of modern astronomy leans on physics and chemistry for basic conceptual and technical tools needed to interpret the universe and its contents. Astronomy also initiates new fields of inquiry in physics and chemistry. For example, astronomical discoveries such as dark matter and dark energy have opened new directions in fundamental physics beyond the standard model of particle physics. Similarly, studies of cosmic rays and of compact objects have driven new directions in plasma physics. The development of new concepts and tools in these and other areas of physics and chemistry now proceed with astronomical observation and modeling as an integral part of the studies of the basic physics. There is a concern that the mechanisms that funding agencies use to allocate resources to theoretical activities are not always effective in funding new areas of intersection in physics and astronomy, and the panel urges that the agencies structure their reviews so that these new areas receive full consideration.

In theoretical astrophysics, computation has become essential for progress, with analytic estimates and models still important as first steps in formulating concepts and for developing computational models of observed phenomena. Much effort has gone into developing computing resources at the national scale that are open to all users and that are located in a variety of supercomputer centers sponsored by NSF, NASA, and DOE. However, resources for intermediate-scale, massively parallel computing are also needed to provide the rapid turnaround essential for algorithm and code development and for student education in computational techniques. Access to such resources has been difficult at the national centers. They tend to be provided through various local funding strategies, such as startup packages for new faculty with and interest in computational techniques, but such strategies are clearly inadequate to keep up with replacements of and improvements in equipment, which are needed on a 3- to 5-year cycle. Funding of such medium-scale facilities needs to be stabilized. Recognizing that development programs on medium-scale cluster computers are essential for the effective use of supercomputing centers, the panel is of the opinion that a reasonable fraction of the resources devoted to supercomputing should be targeted to medium-scale facilities suitable for algorithm development and rapid exploration of concepts. Such funds should be competed following normal peer-review procedures. The panel also thinks it would be beneficial for an organization such as the American Astronomical Society to provide a forum that encourages efforts in student training in computational astrophysics.

Laboratory Astrophysics

An important component of research in astronomy and astrophysics is experimentation that is not an end in itself, but that provides data necessary for the interpretation of experiments and observations aimed at scientific issues. (Thus, experiments whose main motivation is to probe fundamental physics would not fall under the category of laboratory astrophysics.) Traditional areas include, for example, laboratory measurements of atomic cross sections and rates. The panel notes that the experimental techniques of particle astrophysics and gravitation are used in some measurements critical to the interpretation of experiments and observations under consideration by this panel and by other panels. Measurements of nuclear cross sections are needed to inform studies of nucleosynthesis and stellar evolution. Laboratory experiments in plasma and fluid physics are needed to inform quantitative modeling of astronomical systems in, for example, the areas of fluid and magnetohydrodynamic turbulence, high-energy particle acceleration in turbulent plasmas, magnetic reconnection, and collisionless shocks. When assessing their programs the funding agencies must include consideration of the indirect yet important benefit of such experiments.

RECOMMENDATIONS

The first detection of gravitational waves is likely to take place in the next decade with ground-based detectors. Substantial progress in testing general relativity and studying astrophysical sources requires exploiting the lower-frequency part of the spectrum, which can be done only from space. The potential scientific benefit is enormous, because quantitative strong-field tests of general relativity will be possible for the first time, and a qualitatively new window for studying astrophysical systems at a broad range of redshifts will be opened. A great deal of technical progress has been made in the past decade, and a successful LISA pathfinder will eliminate much of the remaining risk. *The panel recommends that the LISA mission be given the highest priority for a new start in the next decade, given the extensive technology development that has already been completed, the expected short time until the LISA Pathfinder (LPF) mission launch, and the need to maintain momentum in the U.S. community and guarantee a smooth transition to a joint NASA-ESA mission. The panel recommends that NASA funding of LISA begin immediately, with continuation beyond LPF contingent on the success of that mission.*

LISA will not be sensitive to the lowest frequencies of gravitational waves that are predicted. This portion of the spectrum probes the fundamental question, How did the universe begin? It might also provide signals from merging supermassive black holes. Pulsar timing is a promising technique for detecting very-low-frequency gravitational waves, and *the panel recommends that NSF provide support for a coherent program in gravitational-wave detection through timing of millisecond pulsars.*

Although much progress has been made in the past decade in testing general relativity in the weak-field limit and on scales of the solar system, little has been done to test strong-field general relativity and gravitation on large (cosmological) scales. The discovery that the universe is apparently accelerating may be a manifestation of a breakdown of general relativity. As yet, theory provides little guidance as to where best to search for deviations from general relativity. Therefore, it makes sense to carry out experiments that favor scales and domains where the theory has not yet been tested, tests that are relatively unambiguous in their predictions, and experiments that improve the precision of measurement of basic parameters in a cost-effective way. Therefore, *the panel recommends that NASA's existing program for small- and medium-scale astrophysics missions (the Explorer program) include consideration, through peer review, of experiments to test general relativity and other theories of gravity.*

Over the past decade, ground-based arrays of imaging Čerenkov telescopes have discovered gamma-ray emission from a wide variety of astrophysical sources. There is a strong case for moving forward with studies of the extreme physics of these sources, probing the highest-energy particles and photons, the strongest magnetic fields, and the strongest gravitational fields. Furthermore, gamma-ray signatures of dark-matter annihilation and decay provide a promising signal for dark-matter searches. *The panel recommends U.S. involvement, supported jointly at a significant level by DOE and NSF, in an international Čerenkov array with an effective area of a square kilometer.* The future science program should include broad access to the data by the community.

For several decades, NASA's Explorer program has supported small missions with focused and timely science goals, and the program has produced some remarkable successes. Yet the current program has been cut back to the point that the oversubscription rate is very large. To support its recommendation for opportunities for space experiments focused on tests of gravitation, and possibly dark-matter searches and particle astrophysics, *the panel recommends that the Explorer program be restored to its previous launch rate. To support reestablishing the launch rate for astrophysics missions in particular, the panel recommends that the Explorer program be augmented by $900 million for astrophysics missions over the next decade.*

Recent measurements of certain particle excesses in the spectra of cosmic rays may be the first indirect evidence for dark-matter particles. Or these results may signify new, nearby sources of cosmic rays. Future progress in indirect detection of dark matter requires a better understanding of cosmic-ray acceleration mechanisms and propagation processes, and further improvements in sensitivity to dark-matter annihilation and decay products. Ultralong-duration ballooning could enable new experiments by providing a combination of long integration times and the transport of massive detectors. *The panel recommends that NASA support ultralong-duration balloon technology development and augment the balloon program to support ULDB missions for indirect detection of dark matter and for cosmic-ray physics and astrophysics.*

All three funding agencies that support activities related to particle astrophysics and gravitation maintain base programs that are critical to the realization of science goals and that enable future projects, missions, and facilities. It is important that investment continue in these areas. *The panel recommends that the funding agencies maintain their levels of base funding, and that, in addition, certain specific components of the base programs be augmented as identified below.*

In all areas of astronomy and astrophysics, the invention and development of innovative technology are critical to success. It is the view of the panel that in the past decade the level of funding for technology development has been lower than optimal, with a particular shortfall in the development that addresses preparation of mission-specific technologies. At NSF, there is no good mechanism for supporting the technology development necessary to fully design and cost major facilities. *The panel recommends that the base programs in astronomy and astrophysics be augmented for technology development at all three agencies—NASA, NSF, and DOE—across divisional boundaries.*

The panel recognizes the important role that theoretical investigations play in the advancement of knowledge in astronomy and astrophysics, and that theory encompasses a wide range of activities. *The panel supports a strong base program that includes theory as one of its components.* It draws attention to two issues—supporting new physics and astronomy topics that fall outside traditional boundaries, and supporting mid-range computing facilities that provide for training and development.

In Table 8.4 the panel presents its recommendations, including costs, for space-based activities in particle astrophysics and gravitation (except for the Explorer

TABLE 8.4 Panel's Recommendations for Space-Based Activities

Activity	Project's Cost Estimate ($M)	Panel's Cost Estimate ($M)
Large		
LISA (U.S. cost only)	900[a]	1,500[b]
Other (unranked)		
Explorer augmentation	—	900
NASA technology development augmentation	—	300[c]
Ultralong-duration balloon R&D and augmentation	305	250[c]
TOTAL		2,950

NOTE: All costs are in fixed-year 2009 dollars.

[a] Based on an assumption of a €650 million contribution from Europe in a joint U.S.-European project.

[b] Cost appraisal by an independent assessment, assuming 50 percent U.S. participation in a joint U.S.-European project and including launch costs and 5 years of operations.

[c] These augmentations are the panel's recommendations for particle astrophysics and gravitation-related activities only, and not for the entire astrophysics program.

TABLE 8.5 Panel's Estimates of Total Funds Allocatable for New Initiatives for the Next Decade Under Two Assumed Budget Scenarios

	Optimistic Budget Scenario ($M)	Pessimistic Budget Scenario ($M)
NSF/Astronomy	324	324
DOE Particle Astrophysics	390	221

NOTE: All costs are in fixed-year 2009 dollars.

augmentation, which includes all astrophysics Explorer missions). For the "other" category, the panel recognizes that all are important components of a balanced program, and the list is not a ranked list.

For the ground-based program, the panel makes its recommendations within the context of two assumed budget scenarios for the Department of Energy. For both budget scenarios the panel assumes a "budget doubling" profile for NSF's Division of Astronomical Sciences; without such an assumption, no new activities are possible for that division. Budget projections for NSF's Physics Division were not available to the panel. For the DOE Office of Science (particle astrophysics only) budgets the panel assumed an "optimistic" scenario and a "pessimistic" scenario. Table 8.5 shows the total funds that can be allocated over the next decade for new initiatives under the two different budget assumptions.

It is not realistic to assume that all the funds projected in Table 8.5 will be allocated only to the activities under consideration by the panel, which therefore makes the assumption that 50 percent of the funds will be available for activities in particle astrophysics and gravitation.

The costs for the recommended ground-based activities under the two budget scenarios are tabulated in Tables 8.6 and 8.7. The list in the "medium" category is a ranked list. Table 8.6 does not include small projects.

SUMMARY TABLE

Table 8.8 shows the relationship between the capabilities of activities endorsed by the Panel on Particle Astrophysics and Gravitation and the scientific priorities identified by the Astro2010 Science Frontiers Panels.

TABLE 8.6 Panel's Recommendations for Ground-Based Activities, Optimistic Budget Scenario

Activity	Agency	Project's Cost Estimate ($M)	Panel's Cost Estimate ($M)
Large			
AGIS-CTA (U.S. portion)	DOE	—	100[a]
	NSF/Astronomy	—	50[a]
	NSF/Physics	—	50[a]
Medium (ranked)			
Pulsar timing array for gravitational wave detection	NSF/Astronomy	70	70[b]
Technology development augmentation	DOE	—	65
	NSF/Astronomy	—	42
	NSF/Physics	—	42
Auger North (U.S. portion)	DOE	60	30[c]
	NSF/Physics		30[c]
TOTAL			479
DOE			195
NSF/Astronomy			162
NSF/Physics			122

NOTE: All costs are in fixed-year 2009 dollars.

[a] The AGIS and CTA projects are at a preliminary stage, and the cost appraisals are not reliable. Therefore, these numbers should be taken to be illustrative only, although they are consistent with the independent cost appraisal carried out as part of the survey.

[b] The cost is the project's estimate for costs associated with technique development, observing, and data analysis. No independent cost appraisal is available.

[c] No independent cost appraisal for Auger North was carried out, and the panel adopted the project's cost appraisal. The cost listed in this table assumes $42 million for the U.S. share of construction costs ($40 million increased by 6.1 percent for inflation), and 10 years of operations at $1.8 million per year (one-third of the total project operations cost of $5.4 million per year).

TABLE 8.7 Recommendations for Ground-Based Activities, Pessimistic Budget Scenario

Activity	Agency	Project's Cost Estimate ($M)	Panel's Cost Estimate ($M)
Large			
AGIS-CTA (U.S. portion)	DOE	—	90[a]
	NSF/Astronomy	—	50[a]
	NSF/Physics	—	50[a]
Medium (ranked)			
Pulsar timing array for gravitational wave detection	NSF/Astronomy	70	70[b]
Technology development augmentation	DOE	—	20
	NSF/Astronomy	—	42
	NSF/Physics	—	42
Auger North (U.S. portion)	DOE	60	0
	NSF/Physics		0
TOTAL			394
DOE			110
NSF/Astronomy			162
NSF/Physics			92

NOTE: All costs are in fixed-year 2009 dollars.

[a] The AGIS and CTA projects are at a preliminary stage, and the cost appraisals are not reliable. Therefore, these numbers should be taken to be illustrative only, but they are consistent with the independent cost appraisal carried out as part of the survey.

[b] The cost is the project's estimate of costs associated with technique development, observing, and data analysis. No independent cost appraisal is available.

TABLE 8.8 Activities in Particle and Gravitational Astrophysics That Address Astro2010 Science Frontiers Panel Questions

	Missions						
Science Question	LISA	Pulsar Timing Array	Lunar Laser Ranging	AGIS/CTA	HAWC	ULDB	Auger N
Planetary Systems and Star Formation							
PSF 1: How do stars form?	—	—	—	—	—	—	—
PSF 2: How do circumstellar disks evolve and form planetary systems?	—	—	—	—	—	—	—
PSF 3: How diverse are planetary systems?	—	—	—	—	—	—	—
PSF 4: Do habitable worlds exist around other stars, and can we identify the telltale signs of life on an exoplanet?	—	—	—	—	—	—	—
Discovery area: Identification and characterization of nearby habitable exoplanets	—	—	—	—	—	—	—

continued

TABLE 8.8 Continued

	Missions						
Science Question	LISA	Pulsar Timing Array	Lunar Laser Ranging	AGIS/CTA	HAWC	ULDB	Auger N
Stars and Stellar Evolution							
SSE 1: How do rotation and magnetic fields affect stars?	—	—	—	Gamma rays from stars, binary systems, supernova remnants...	Gamma rays from stars, binary systems, supernova remnants...	—	Ultrahigh-energy cosmic-ray probe of galactic magnetic field relevant for star formation
SSE 2: What are the progenitors of Type Ia supernovae?	White-dwarf/white-dwarf binaries in galaxy, which can be progenitors of Type Ia supernovae	Provides pulsar survey	—	—	—	—	—
SSE 3: How do the lives of massive stars end?	Black holes from the first generation of stars	Provides millisecond pulsar survey	—	Gamma rays from gamma-ray bursts (GRBs), supernova remnants	Gamma rays from GRBs, supernova remnants	—	Cosmic rays and neutrinos from GRBs
SSE 4: What controls the mass, radius, and spin of compact stellar remnants?	$>10^4$ compact binaries in the galaxy	Requires discrete sources	—	—	—	—	—
Discovery area: Time-domain surveys	Binary black hole mergers and extreme mass ratio inspirals	—	—	—	Gamma-ray transients	—	—

Galactic Neighborhood							
GAN 1: What are the flows of matter and energy in the circumgalactic medium?	—	—	—	—	—	—	—
GAN 2: What controls the mass-energy-chemical cycles within galaxies?	—	—	—	Local interstellar medium (ISM)	Local ISM	Local ISM	Ultrahigh-energy cosmic-ray probe galactic magnetic field
GAN 3: What is the fossil record of galaxy assembly from the first stars to the present?	Mergers of the first black holes, relics of the first generation of stars; important in determining mass of “seed” black holes	Background of gravitational waves from very massive black hole mergers	—	—	—	—	—
GAN 4: What are the connections between dark and luminous matter?	Distribution of black hole mass and spin via mergers and extreme mass ratio inspirals	—	—	Dark matter indirect searches	Dark matter indirect searches	Dark matter indirect searches	—
Discovery area: Time-domain astronomy	Binary black hole mergers and extreme mass ratio inspirals	Requires discrete sources	—	High-energy flares (e.g., GRBs, blazars, magnetars)	High-energy flares (e.g., GRBs, blazars, magnetars)	High-energy flares	Ultrahigh-energy (neutrino) flares

continued

TABLE 8.8 Continued

	Missions						
Science Question	LISA	Pulsar Timing Array	Lunar Laser Ranging	AGIS/CTA	HAWC	ULDB	Auger N
Galaxies Across Cosmic Time							
GCT 1: How do cosmic structures form and evolve?	Black hole mergers out to $z > 10$; distribution of masses and spins with z	Background of gravitational waves from very massive black holes constrains models	—	—	—	—	—
GCT 2: How do baryons cycle in and out of galaxies, and what do they do while they are there?	—	—	—	—	—	—	—
GCT 3: How do black holes grow, radiate, and influence their surroundings?	Black hole mergers at redshifts up to $z > 10$ and extreme mass ratio inspirals out to $z \sim 1$	Provides pulsar survey	—	Gamma rays from active galactic nuclei (AGN), GRBs	Gamma rays from AGN, GRBs	Neutrinos from AGN, GRBs	Ultrahigh-energy cosmic rays from AGN, GRBs
GCT 4: What were the first objects to light up the universe, and when did they do it?	Seed black holes from the first generation of stars	—	—	—	—	Neutrinos from GRBs	Neutrinos from GRBs
Discovery area: The epoch of reionization	Black hole mergers at high $z > 10$	—	—	—	—	Neutrinos from GRBs	Neutrinos from GRBs

Cosmology and Fundamental Physics							
CFP 1: How did the universe begin?	Gravitational waves are direct probe of early universe	Gravitational waves are direct probe of early universe	Tests of general relativity	—	—	—	—
CFP 2: Why is the universe accelerating?	Strong, direct tests of general relativity from gravitational wave sources	Gravitational wave background model dependent	Tests of general relativity	—	—	—	—
CFP 3: What is dark matter?	—	—	—	Indirect dark matter searches (gamma rays from dark matter halo)	Indirect dark matter searches (gamma rays from dark matter halo)	Indirect dark matter searches (positrons, anti-nuclei, …)	—
CFP 4: What are the properties of neutrinos?	—	—	—	—	—	Greisen-Zatsepin-Kuzmin (GZK) ultrahigh-energy neutrinos	Ultrahigh-energy GZK cosmic rays and neutrinos
Discovery area: Gravitational wave astronomy	Open low-frequency window—very rich in astrophysical sources	Open very-low-frequency window—possibly best chance to detect cosmological background	—	—	—	—	—

NOTE: Shaded entry, direct connection to science question. Unshaded entry, indirect or possible connection but not guaranteed.

9

Report of the Panel on Radio, Millimeter, and Submillimeter Astronomy from the Ground

SUMMARY

Astronomy at radio, millimeter, and submillimeter (RMS) wavelengths is poised for a decade of discoveries. The Atacama Large Millimeter Array (ALMA) will be commissioned in 2013, enabling detailed studies of galaxies, star formation, and planet-forming disks, with spectral coverage from 0.3 to 3 mm, at a resolution approaching 4 milli-arcseconds at the shortest wavelengths. Soon, the Expanded Very Large Array (EVLA) will have an order-of-magnitude more continuum sensitivity than the original Very Large Array (VLA), with continuous spectral coverage from 0.6 to 30 cm. The Herschel Space Observatory, with coverage from 60 to 670 μm, is delivering catalogs of tens of thousands of new "submillimeter-bright" galaxies. The Green Bank Telescope (GBT) operates over a broad range of centimeter and millimeter frequencies and has the potential for vastly improved mapping speeds with heterodyne and large-format bolometric-array cameras. With upgrades, the Very Long Baseline Array (VLBA) will improve astrometric distances critical to studies of star formation, galactic structure, and cosmology. It is possible that gravitational waves will be detected by timing arrays of pulsars, with the Arecibo Observatory playing a crucial role. The University Radio Observatories (UROs) will produce steady streams of excellent science, provide training grounds for graduate students, and remain at the cutting edge of science and technological development. The sizes of detector arrays at millimeter and submillimeter wavelengths and the computational capabilities of digital correlators are both experiencing exponential growth.

The foundation for further advances in this field must be laid in this decade.

The crucial scientific questions and themes identified today can be addressed if the necessary steps are taken to lead to the instruments of tomorrow. RMS projects of modest cost will provide insights into the origins of the first sources of light that re-ionized the universe and led to the first galaxies. With truly large-format detector arrays on single-dish telescopes, large-scale surveys for galaxies forming stars intensely will inform the origin of the cosmic order observed today. An RMS project will provide insights into fundamental processes on the Sun and use the Sun as a laboratory for understanding the role of magnetic fields in astrophysical plasmas. Upgrades of modest cost to existing RMS facilities may allow the first discovery of gravitational waves and imaging of the event horizon around a black hole. The steps taken during this decade can lead to the next great advance in future decades, a telescope capable of studying the atomic gas flows that fed galaxies back in cosmic time and capable of studying the inner parts of circumstellar disks, where Earth-like planets may be forming. With continued, robust support for studies of the cosmic microwave background (CMB), RMS science extends from the Sun to recombination and the physics of inflation.

The Panel on Radio, Millimeter, and Submillimeter Astronomy from the Ground has identified key capabilities that are needed to answer the science questions posed by the five Astro2010 Science Frontiers Panels. By comparing those key capabilities to existing capabilities, the panel identified three new projects for mid-scale funding that will provide critical capabilities. The panel further identified enhancements to existing or imminently available facilities that fulfill other requirements, and this report presents a balanced program with support for small facilities, technology development, laboratory astrophysics, theory, and algorithm development. Priorities and phasing are discussed in the panel report's final section, "Recommendations." Those recommendations are summarized here.

Recommended New Facilities for Mid-Scale Funding

The Hydrogen Epoch of Reionization Array (HERA) will provide unique insight into one of the last remaining unknown eras in the history of the universe. The panel recommends continued funding of the two pathfinders (collectively HERA-I) and a review mid-decade to decide whether to build HERA-II. The panel identified specific milestones to be met by HERA-I activities. If those are met, HERA-II is the panel's top priority in this category of recommended new facilities for mid-scale funding. HERA-I requires about \$5 million per year, as is currently being spent, and HERA-II construction is estimated to cost \$85 million.

The Frequency-Agile Solar Radiotelescope (FASR) will scan conditions in the chromosphere and corona across the full solar disk once a second, all day, every day. It is a vital complement to the Advanced Technology Solar Telescope (ATST) and provides essential ground truth for studies of magnetic fields on other stars. The estimated construction cost for FASR is $100 million, and operations will cost $4 million per year, both of which the panel assumes will be evenly split between the National Science Foundation's (NSF's) Division of Astronomical Sciences (AST) in the Directorate for Mathematical and Physical Sciences and the Division of Atmospheric and Geophysical Sciences (AGS) in NSF's Directorate for Geosciences.

CCAT (formerly the Cornell-Caltech Atacama Telescope) will provide the capability for rapid surveys of the submillimeter sky, essential for the optimal exploitation of ALMA. CCAT is a 25-m-diameter telescope located on a very high, dry site and equipped with megapixel detector arrays; it will address many of the questions posed by the Science Frontiers Panels. CCAT is estimated to cost $110 million, with $33 million coming from NSF. NSF's share of operating expenses would be about $7.5 million per year, a net increase of $5 million per year, assuming that current funding for the Caltech Submillimeter Observatory (CSO) is recycled.

FASR and CCAT have equal, and very high, priority in this category, but different phasing.

Development of Current and Imminent Activities

Studies of the CMB have delivered much of the most valuable information about the universe at large. The panel strongly recommends a continued robust program at the current funding levels of ground-based CMB studies, with multiple approaches that are driven by individual investigators.

An expansion of the Allen Telescope Array to 256 antennas (ATA-256) would significantly improve astronomers' ability to find and study transient sources and to detect gravitational waves by timing an array of pulsars. The ATA can test ideas needed for the development of next-generation telescopes such as the Square Kilometer Array (SKA). The estimated cost of construction for the expansion is about $44 million. The panel recommends that NSF explore collaboration with other agencies and private foundations for the enhancement of ATA-42.

The National Radio Astronomy Observatory (NRAO) telescopes (and soon, ALMA) provide a broad range of scientific capabilities needed to answer many of the SFP questions, but all will need instrument development, especially the completion of frequency coverage, multibeam capability, and electronics improvements to enable much higher data rates. The panel recommends a sustained and substantial program to enhance the NRAO telescopes and ALMA capabilities, amounting to $90 million for NRAO and $30 million for the U.S. share for ALMA over the decade.

The Arecibo telescope is essential for science with pulsars, which test general relativity, constrain the neutron star equation of state, and may lead to the detection of gravitational waves. The telescope can also make the deepest maps of galactic and extragalactic neutral hydrogen currently possible. A future multi-pixel upgrade would dramatically speed up surveys at centimeter wavelengths. The panel recommends that support for Arecibo be enhanced by $2 million per year over projected levels.

The UROs provide cost-effective capabilities, testbeds for technology, and training grounds for young scientists. The panel recommends a modest enhancement ($2 million per year) in the budget for the *current* program, and it recommends that FASR ($2 million per year, starting in 2015) and CCAT (net $5 million per year, starting in about 2017) be operated under the URO program.

Small Projects

To achieve a balanced program, the panel recommends that a range of small and moderate projects be supported through a combination of funding from the Advanced Technologies and Instrumentation (ATI) program at NSF/AST and from NSF's Major Research Instrumentation (MRI) program. Examples of such projects include an enhancement of the VLBI's millimeter-wave capabilities to allow imaging of the event horizon around a black hole and multifeed receivers for the Combined Array for Research in Millimeter-wave Astronomy (CARMA). A program of technology development in a number of areas, and a focused program of laboratory astrophysics, are both vital needs. Support for theoretical work is crucial to realizing the investment in RMS facilities, as is a program of algorithm development. Both will allow observations to confront theory, essential to moving science forward. The panel recommends enhancements to NSF/AST's ATI program of $1 million per year and an added program of laboratory astrophysics at $2 million per year.

Looking to the Future

The SKA has remarkable discovery potential, including studies of the epoch of reionization (SKA-low), determination of the gas content of galaxies at z of 1 to 2 (SKA-mid), and studies of the terrestrial-planet zones of planet-forming disks (SKA-high). However, substantial technology development is needed to define an affordable instrument. Many areas that the panel recommends for technology development will be crucial for this effort. The HERA project provides a development pathway for SKA-low, and the North American Array (NAA) project (part of NRAO development) develops technology for SKA-high. The panel recommends the continued development and exploration of options for realizing SKA-mid.

THE SCIENCE CASE

Here the panel identifies the RMS capabilities (in italic) that are needed to answer the science questions raised by the Science Frontiers Panels (SFPs) and summarizes them (as numbered below) in Table 9.1 at the end of this section and also maps them in Figure 9.6 there. The panel lists only the SFP questions that RMS facilities can address. It also poses some additional, unnumbered, relevant questions. In later sections, the panel matches these capabilities to specific activities and concludes with recommendations.

Cosmology and Fundamental Physics (CFP)

Q1: How did the universe begin? The CFP report requires a suite of instruments that characterize the cosmic microwave background (CMB) to determine properties of the early universe. CMB observations probe the potential-energy function of the primordial field (or fields), the type of primordial fluctuation (adiabatic or iso-curvature), whether the fluctuations were Gaussian, and the degree to which gravitational waves (tensors) played a role in the infant universe. CMB observations may yield the only observational constraint on theories of quantum gravity. These measurements will complement the Planck satellite, which will have exquisite sensitivity to temperature fluctuations for $2 < l < 2{,}500$, where l is the multi-pole index of a spherical harmonic; the corresponding angular scale is $\theta \sim 180/l$ in degrees. With new CMB polarization measurements in the range $2 < l < 200$, one can find or limit primordial gravitational waves and probe reionization. The polarization for $200 < l < 3{,}500$ sheds light on early-universe physics, neutrino mass, and the helium abundance. The temperature spectrum for $1{,}000 < l < 10{,}000$ should be pursued from the ground to understand the high-l tail of the primary CMB spectrum and, for example, to identify clusters through the Sunyaev-Zel'dovich effect (SZE). The search for large-angular-scale B-modes (from tensor fluctuations) will discover them or decrease the current limit on the tensor-to-scalar ratio from about 0.3 to 0.01. These techniques lay the foundation for a future satellite focusing on polarization.

Q2: Why is the universe accelerating? It is known that the universe is accelerating but not why or how. To make progress, the z-dependence of the dark-energy equation of state, $w(z)$, and the Hubble parameter, $h(z)$, must be better determined. Intimately related to these are possible changes in the gravitational coupling constant (G) and the growth rate of structure. The CFP report calls for precision tests of general relativity, which is central to our understanding of cosmology. Studies of pulsars in relativistic binary systems provide the strongest constraints on possible changes in G, and timing may allow detection of gravitational radiation from the early universe. Our capabilities to find relativistic binary pulsars and time them must be enhanced. Additional tests will come with measurements of the supermassive black hole (SMBH) in the center of our galaxy with ultrahigh spatial resolution and through gravitational lensing of radio by SMBHs in distant galaxies, requiring sensitive centimeter-wave imaging. The growth of structure will be measured a number of ways with the CMB. Through the SZE, measurements of clusters can provide a mass-selected, almost z-independent sample of clusters of galaxies over large regions of sky, tracing the development of clusters from the early to nearby universe.

Q4: What are the properties of neutrinos? The presence of relativistic neutrinos affects the growth of structure at early times. High-l CMB-lensing B-modes are particularly sensitive to this effect, as are other aspects of the measurements. CMB experiments are expected to determine the sum of neutrino masses to 0.05 eV. A vigorous ground-based CMB program is a requirement in this area.

Discovery area: Gravitational wave astronomy—listening to the universe. The CFP report calls for *pulsar timing arrays* to probe the nanohertz frequency range to detect the stochastic background of gravitational waves from SMBH binaries (Figure 9.1). A project focusing on this goal is discussed in the PAG report, and so only the consequences for RMS facilities are summarized here. A rough requirement for gravitational-wave detection is careful timing of the arrival of pulses from a set of ~20 stable, millisecond pulsars distributed over the sky to an accuracy of 100 ns over a period of 5 years. This task may be divided into two parts: discovery and timing. Discovery requires large collecting area, wide bandwidths, high-throughput back-end electronics, and computational power. In the United States, pulsar-discovery capabilities exist primarily at Arecibo and the GBT. They must be sustained and upgraded. Timing requires about a day per week of observations by a telescope with 10^4 m^2 collecting area and with back-end hardware and software capable of coherent dispersion removal. Observations across a wide bandwidth or at multiple simultaneous frequencies are essential to correct for effects due to interstellar dispersion. Timing requires either the reallocation of existing facilities or ideally a facility with a large fraction of the time dedicated to timing.

Galaxies Across Cosmic Time (GCT)

Enabled in part by RMS observations with new facilities and instrument technologies, an understanding of the formation and evolution of galaxies and large-scale structure is beginning to form. However, fundamental questions remain. The observed structures of galactic dark-matter halos challenge structure-formation theories. The interaction of gas and stars in the galaxy-building process is poorly understood, driving a need for inventories of the cold atomic and molecular gas contents of galaxies. Supermassive black-hole growth and feedback must be characterized to understand the correlation between black-hole masses and stellar-bulge velocity dispersions. At the highest redshifts, the nature of the first objects that reionized the universe remains unconstrained.

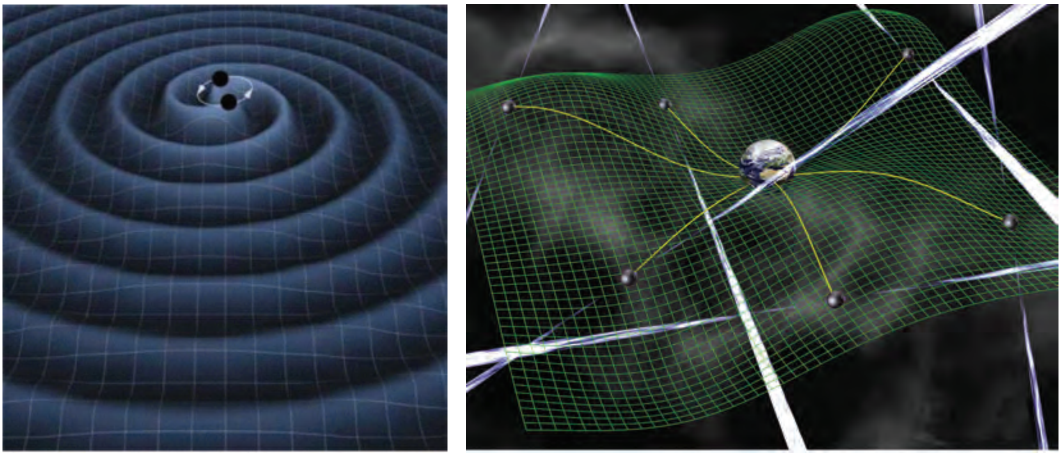


FIGURE 9.1 Black holes orbiting one another produce ripples in space-time or gravitational waves. *Left:* The ripples from a single supermassive black-hole binary cause the signals traveling from two pulsars to the telescope to arrive faster or slower than expected if they were traveling through flat space. *Right:* The superposition of many supermassive black-hole binaries distributed throughout the universe gives rise to a low-frequency stochastic background. This background will induce correlated shifts in the arrival times of an array of pulsars. SOURCE: *Left:* D. Backer/JPL/NASA. *Right:* David Champion, Max-Planck-Institut für Radioastronomie.

Q1: How do cosmic structures form and evolve? RMS facilities can characterize the shapes and substructures of galactic halos using high-angular-resolution observations of gravitational lensing. The EVLA will find hundreds of lenses, but ultimately approximately 10^6 lenses will be needed, requiring a capability for very sensitive centimeter-wave imaging. Unbiased by redshift, SZE surveys associated with *the CMB program* will discover large samples of clusters out to their epoch of formation. Understanding the physics of these clusters will require multiwavelength observations, including detailed centimeter and millimeter observations (Figure 9.2).

Q2: How do baryons cycle into and out of galaxies, and what do they do while they are there? A comprehensive model of galaxy formation requires an understanding of accretion, mergers, and evolution of the gas in galaxies from the formation of the first galaxies to the present day. For studies of dusty, distant, star-forming galaxies, large-area (tens of square degrees), continuum submillimeter surveys will sample the galaxy luminosity function over a broad range of redshifts. Large, single-dish submillimeter and millimeter telescopes equipped with large-format detector arrays are needed. To characterize the gas associated with star formation, spectroscopy of CI, CO, and molecular tracers of dense gas is essential. The current ALMA and EVLA facilities will offer broad but incomplete coverage in redshift for the detection of CO in distant galaxies. New facilities and upgrades can enable fast

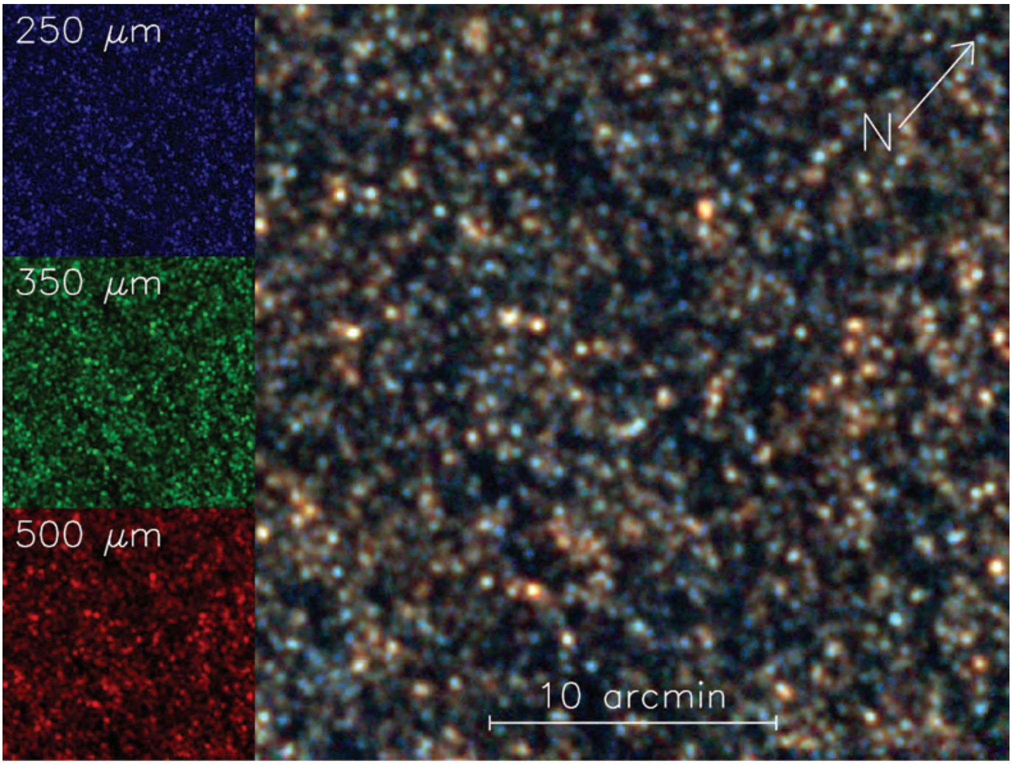


FIGURE 9.2 Deep submillimeter extragalactic image in the GOODS-N region obtained with the Hershel Space Observatory SPIRE camera. The three images at the left are in three bands, which are coded in those colors in the large composite. All the sources are dusty galaxies. Bluer galaxies are warmer and/or more nearby, while redder galaxies are cooler and/or at higher redshift. The submillimeter sky is extremely rich in galaxies—thousands of (unresolved) high-redshift galaxies appear, and the image is highly confusion-limited. SOURCE: Courtesy of the European Space Agency and the SPIRE and HerMES consortia.

spectroscopic follow-up observations of galaxies detected in continuum surveys, particularly in the $z = 2$ to 4 range. This goal requires complete submillimeter-millimeter frequency coverage, broadband correlators, and multiobject spectroscopic capability for single-dish telescopes. Large-scale intergalactic gas flows and early accretion onto galaxies will consist largely of atomic hydrogen. H I 21-cm line observations can assess the total masses and atomic gas contents of galaxies, circumgalactic streams, and the cosmic web, mapping the gaseous origins of galaxies. The current suite of centimeter-wave telescopes are limited to observing H I structures in the local ($z < 0.2$) universe. Detecting H I in galaxies at redshifts as high as $z \sim 1$ will require very sensitive centimeter-wave imaging, requiring a large increase in collecting area over current facilities and substantially increased correlator capabilities.

Q3: How do black holes grow, radiate, and influence their surroundings? Supermassive black holes are ubiquitous in the centers of galaxies. However, much still needs to be learned about how they become active galactic nuclei (AGN), and how they interact with their host galaxies. Ultrahigh-spatial-resolution interferometric observations can resolve regions close to AGN: jet-formation mechanisms can be constrained by VLBA and submillimeter VLBI observations of blazars, and submillimeter VLBI will enable studies of circumnuclear disks in galaxies, including our own.

Q4: What were the first objects to light up the universe, and when did they do it? The epoch of reionization (EoR) is the frontier in understanding galaxy formation. Understanding cosmic reionization at $z > 6$ requires mapping the H I distribution via the hyperfine transition, which is redshifted to meter-wavelengths for $z > 5$. The first-generation instruments (PAPER, MWA, EDGES, and international efforts) are designed to detect fluctuations and constrain the redshift range of reionization. Precise measurement of the power spectrum of H I emission requires a sensitive meter-wave array with a factor-of-10 larger collecting area ($A_{eff} \sim 10^5$ m^2). Imaging hydrogen structures will require another order of magnitude in collecting area ($A_{eff} \sim 10^6$ m^2). Whether stars in galaxies or other sources are primarily responsible for cosmic reionization remains controversial. The roughly constant apparent brightness of dusty galaxies of a given luminosity as a function of redshift in submillimeter and millimeter bands provides an advantage for fast surveys at millimeter/submillimeter wavelengths. Observations of redshifted CO, C II, N II, and O I have the potential to measure the radiative cooling of galaxies in the later stages of the EoR, requiring *complete wavelength coverage* in spectral windows accessible from the ground.

The Galactic Neighborhood (GAN)

The galactic neighborhood (GAN) science frontier covers galaxy build-up and evolution back to $z \sim 0.1$. RMS facilities are well suited for GAN studies through their ability to probe cold flows, star formation, the environment of SgrA*, and magnetic fields. RMS interferometers provide high resolution and high astrometric accuracy (Figure 9.3).

Q1: What are the flows of matter and energy in the circumgalactic medium? Nearly all massive galaxies are accreting material, observed as H I and stellar streamers from tidally disrupted satellite galaxies. Radio facilities are required to assess the kinematics and distributions of H I on both large (low-resolution) and small (high-

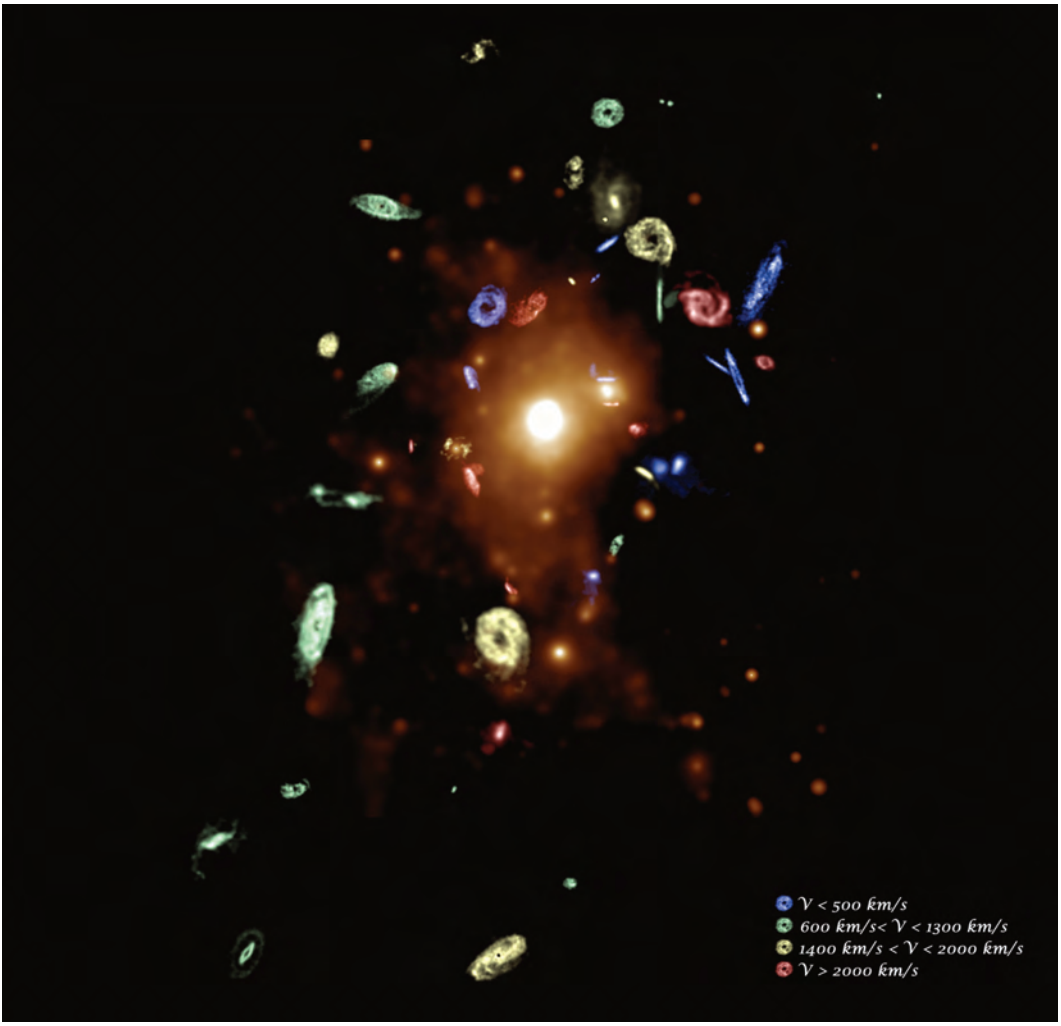


FIGURE 9.3 Neutral hydrogen (H I) gas in the Virgo cluster of galaxies taken with the VLA, colored by velocity (galaxy sizes have been increased by a factor of 10 to make them more visible). The background image is from ROSAT. SOURCE: NRAO/AUI and Chung et al., Columbia University.

resolution) scales. Feedback in galaxies from star formation or AGN activity may enhance or halt additional star formation. By imaging radio synchrotron emission, H I, and molecular lines, one can assess the impact of feedback on the surrounding medium. Sensitive, low-resolution observations with radio facilities play a crucial role in studying feedback on large scales, but efficient, high-resolution, centimeter-wave capabilities are needed to provide maps of energy deposition into the ISM.

Q2: What controls the mass-energy-chemical cycles within galaxies? RMS facilities play a critical role in studying the build-up of stellar mass by gas inflow and accretion over cosmic time. H I is a prime tracer of the inflow of gas into galaxies, some of which becomes star-forming H_2 clouds. Observations of dust emission and molecular lines (millimeter-submillimeter) trace the gas, and synchrotron emission from supernovae traces star-formation rates. Very sensitive 2-cm imaging and fast surveys at submillimeter wavelengths and full wavelength coverage from meter to submillimeter will be required to achieve the goals.

Q4: What are the connections between dark and luminous matter? The H I disks of galaxies extend well beyond the stellar light of galaxies and were instrumental in providing evidence of dark matter. Sensitive, high-resolution H I observations are required to trace the kinematics of the faint, dark-matter-rich dwarf galaxies and the outer parts of disk galaxies. Accurate measurements of the properties of SMBHs and their interplay with their immediate gaseous environments are required to understand how SMBHs fit into the build-up and evolution of galaxies, including nuclear star-forming rings. These science goals require high-resolution ($<0.05''$) radio and millimeter observations to probe the non-thermal emission and state of the star-forming gas, respectively. Ultrahigh resolution at millimeter/submillimeter wavelengths tantalizes us with the possibility of imaging the event horizon of Sgr A*.

Discovery area: time-domain astronomy. Temporal RMS observations are a largely unexplored area of astronomy likely to show significant progress in the next decade. Such observations require a sensitive, dedicated instrument designed to scan the available sky rapidly enough to sample variable (or moving) objects. A dedicated transient-search telescope covering a wide range of wavelengths is needed to explore this discovery space.

Discovery area: astrometry. One of the great strengths of RMS interferometers is the direct detection of electromagnetic phase and the ability to do astrometry routinely to a fraction of a beamwidth. Parallaxes and proper motions can be measured within the galaxy, and for galaxies within the Local Group and beyond, allowing for a better assessment of dynamics and everything that comes with better distances. An ultrahigh-resolution capability with improved sensitivity is needed for precision astrometry.

Stars and Stellar Evolution (SSE)

Q1: How do rotation and magnetic fields affect stars? For many years, radio observations have provided clues to magnetically driven solar and stellar activity. In the next decade, they have the potential to do much more. New data on magnetic fields in the Sun's corona and chromosphere can help characterize how magnetism powers stellar atmospheres. Spatially resolved, high-time-cadence broadband imaging spectroscopy is required to obtain vector magnetic field measurements and to detect sites of magnetic reconnection and particle acceleration. Full-Sun, continuous observations are also needed both to map the global magnetic field and to survey magnetically driven dynamics. Such observations, in conjunction with the full range of multiwavelength observations of chromospheric and coronal plasma, together with three-dimensional models and laboratory analyses of magnetic reconnection, can synthesize an understanding of solar magnetism (Figure 9.4). That understanding can be extended outward to stars and throughout plasma astrophysics. A dedicated solar radio telescope is needed for this synthesis.

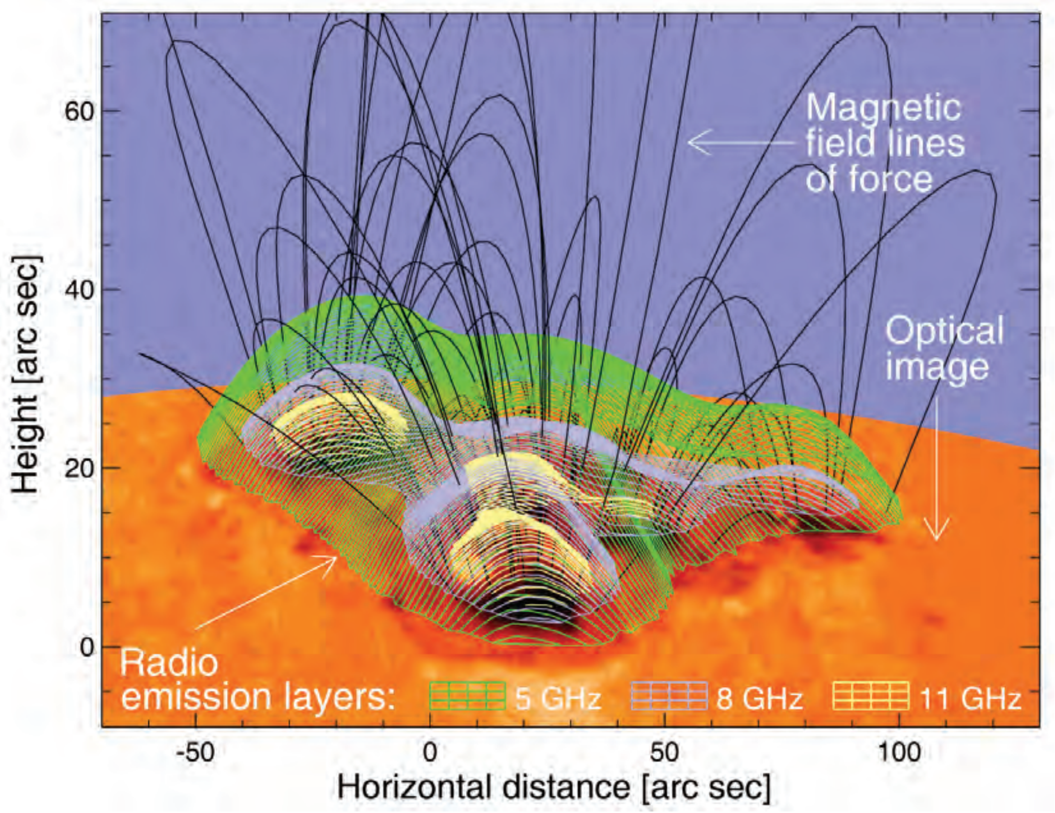


FIGURE 9.4 Models for magnetic fields on the Sun. SOURCE: Jeongwoo Lee, New Jersey Institute of Technology.

Q2: What are the progenitors of Type Ia supernovae and how do they explode? Radio observations are important probes into the origins of exploding stars, such as Type Ia supernovae (SNe). In particular, radio emission arises from circumstellar material disturbed by the explosion. Therefore, sensitive and rapid follow-up radio detections of Type Ia SNe with flexible radio interferometers can provide crucial insight as to the progenitor phase of these SNe.

Q3: How do the lives of massive stars end? Episodes of extreme mass loss preceding the deaths of massive stars can hide the explosions from observations at shorter wavelengths, but they can be traced in the radio with sensitive interferometric observations, spaced at appropriate intervals. Radio and X-ray observations can also be combined to trace the shock physics occurring when these massive ejecta propagate into the interstellar medium. Such a synthesis of radio and X-ray observations requires high spatial resolution in the radio (at least ~1″), comparable to Chandra. "Orphan afterglows" of gamma-ray bursts are best studied in the radio using dedicated transient survey instruments.

Q4: What controls the mass, radius, and spin of compact stellar remnants? Nearly 2,000 neutron stars are known today, the majority of which have been discovered and characterized through radio observations. The measurement of relativistic effects through binary pulsar timing yields accurate neutron-star and (likely compact-object) companion masses. Relativistic spin-orbit coupling should be detectable for the double-pulsar system within the next decade and will yield the neutron-star moment of inertia, providing an important constraint on the equation of state. Pulsar searches with excellent short-period sensitivity will determine whether gravitational-wave damping is important for limiting minimum spin periods. The spin of a black hole could be determined through the detection of relativistic spin-orbit coupling in a pulsar/black-hole binary. Large-scale surveys, with short-period sensitivity, and search algorithms that correct for binary acceleration are necessary to discover these systems. Frequent and long-term timing of these discoveries will then be needed to yield fundamental constraints on the equation of state. Surveys with large telescopes and the ability to time multiple pulsars simultaneously and commensally with other surveys are necessary.

Discovery area: time-domain astronomy. The next decade promises discoveries of both new source classes and new phenomena from known sources. Transient radio emission is expected from the Sun, flaring stars, neutron stars, tidal disruptions, local and cosmological supernovae, gamma-ray bursts, and, possibly, extraterrestrial civilizations. Radio observations probe energetic and/or explosive processes associated with high magnetic fields and provide unique information not available at other wavelengths. To fully exploit the expected parameter space, large-field-of-view surveys with time resolutions of a millisecond or less are required. Optimally, the same region of sky would be surveyed multiple times. New facilities must also enable prompt radio follow-up of sources found at other energies. Effective monitoring of solar transients also requires a field of view sufficient to image the entire Sun.

Planetary Systems and Star Formation (PSF)

Q1: How do stars form? With transformative new RMS instruments coming on line early in the next decade, and key missing capabilities under development, we are poised to address fundamental questions about how, where, and under what physical conditions stars form across the whole spectrum of the stellar mass function. The capabilities required are a combination of fast-survey capabilities at millimeter/submillimeter wavelengths and fast, sensitive surveys at centimeter wavelengths to conduct wide-field surveys at 5- to 20-arcsecond resolution, coupled with more sensitive, sub-arcsecond resolution follow-up of selected regions. Accurate distances obtained via ultrahigh-resolution capability are fundamental.

What determines star-formation rates and efficiencies in molecular clouds? In the solar neighborhood, typically only a few percent of the mass of a giant molecular cloud (GMC) is converted into stars. Understanding what controls star-formation rates and efficiencies in clouds in our galaxy can inform models of galaxy evolution (Figure 9.5). Large-area surveys of GMCs and their deeply embedded star-forming regions require fast millimeter/submillimeter surveys. Diagnostics of the physical properties of star-forming regions require spectroscopy of molecular and ionized gas, as well as observations of dust, free-free emission, non-thermal continuum emission, and polarization on spatial scales ranging from 0.1 to 100 pc, corresponding to 0.05 to 200 arcseconds, to extend the study to nearby galaxies at ~10 Mpc.

What determines the properties of pre-stellar cloud cores, and what is the origin of the stellar mass function? Stars form in dense cores of size ~0.1 pc. Similarities between the core mass function (CMF) and the mass distribution of stars (stellar IMF) raise the question, Is the CMF, and potentially the stellar IMF, dependent on environment? Studies of more distant regions, with comparable spatial resolution, are needed. Exploration of the CMF is limited by current sensitivity and angular resolution, as well as by uncertainties in the dust emissivity and temperature. Dramatic progress can be made with sensitive, high-angular-resolution, multiwavelength spectral-line and dust-continuum data, from the centimeter to far-infrared, on 0.01-pc scales, requiring a best resolution of 0.05 arcseconds (to access the LMC and SMC). Fast-survey capability at submillimeter wavelengths is needed to assemble complete samples.

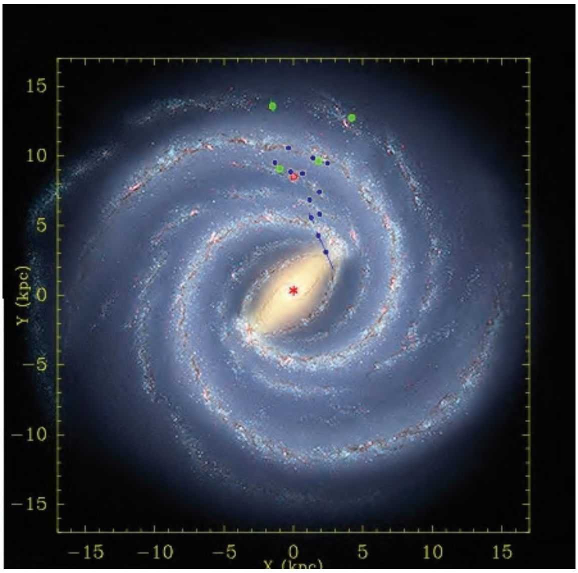


FIGURE 9.5 Model of the Milky Way with star-forming regions indicated; before recent VLBA measurements, these regions did not align with spiral arms. SOURCE: Robert Hurt, IPAC; Mark Reid, CfA, NRAO/AUI/NSF.

Q2: How do circumstellar disks evolve and form planetary systems? Circumstellar disks are ubiquitous by-products of low-mass star formation, an inevitable consequence of angular-momentum conservation and gravitational collapse. A suite of panchromatic observations has been used to characterize the bulk properties of the disks around pre-main-sequence stars, establishing them as potential sites for planet formation. Both ALMA and a large, single-dish, millimeter/submillimeter telescope play key roles because these disks are impenetrable at shorter wavelengths. A complete understanding of the origins and diversity of planetary systems will require multiwavelength observations of large samples of disks to probe their structure, to discover features associated with planets, and to identify evolutionary trends.

What is the nature of the planet-forming environment? How do giant planets accrete from disks, and what are these young planets like? While current RMS facilities offer a glimpse of disk structure on scales >20 AU, imminently available new facilities are poised to resolve dust and gas tracers on scales down to 1 AU with high fidelity. Such revolutionary observations have the potential to directly reveal spiral density waves, accretion streams, snow lines, and regions prone to rapid gravitational instabilities. A new frontier will be the inner-disk regions of terrestrial-planet formation at milli-arcsecond scales, where low dust opacities are reached only at long millimeter and short centimeter wavelengths. Imaging these small regions will require the development of very sensitive centimeter-wave telescopes with excellent resolution.

What can debris disks and the Kuiper belt reveal about the dynamical evolution of planetary systems? Imaging- and modeling-resolved features in debris disks—such as central cavities, offsets, and clumps—(and their proper motions) can locate unseen planets, particularly at separations >10 AU, where classical indirect planet-detection techniques become difficult. The debris features may encode a historical record of the system dynamics, similar to the scattered and resonant populations of the Kuiper belt. Large-format bolometer arrays on single-dish telescopes at millimeter/submillimeter wavelengths will be required to assess both the full range of debris-disk properties—as well as the frequency of large-separation planets—and the habitability of terrestrial planets in the heavy bombardment phases that create debris disks.

Q4: Do habitable worlds exist around other stars, and can we identify the telltale signs of life on an exoplanet? The ultimate question is whether humans are alone. Discovery of potentially habitable worlds is imminent, and the PSF report explores the options for finding "telltale" signs of life on exoplanets. Of course, the most certain sign of extraterrestrial life would be a signal indicative of intelligence. An RMS facility that devoted some time to the search for extraterrestrial intelligence would provide a valuable complement to the efforts suggested by the PSF report on this question. Detecting such a signal is certainly a long shot, but it may prove to be the only definitive evidence for extraterrestrial life.

Summary of Needed Capabilities

The panel collected in Table 9.1 and illustrated in Figure 9.6 the RMS capabilities noted in the previous discussion and identified the SFP questions for which they are most crucial. These capabilities will impact many other questions besides those noted here.

THE PROGRAMMATIC CONTEXT

The RMS region covers more than five decades in wavelength, from the boundary of the far-infrared, at $\lambda = 200\ \mu\text{m}$, to the ionospheric cutoff, $\lambda = 30$ m. Table 9.2 lists the RMS facilities open to U.S. investigators—either currently operating or under construction—excluding dedicated experiments, such as CMB or EoR experiments. Below the panel summarizes their capabilities, noting links to SFP questions and needed capabilities (see Table 9.1).

TABLE 9.1 Capabilities Needed to Address RMS-Related Science Questions Identified by the Astro2010 Science Frontiers Panels

Capability	CFP	GCT	GAN	SSE	PSF
Cosmic microwave background program	Q1, 2, 4	Q1			
Sensitive meter-wave array		Q4			
Solar radio telescope				Q1, Discovery	
Fast millimeter/submillimeter surveys		Q2, Q4	Q2		Q1, 2
Fast centimeter surveys			Q2	Discovery	Q1
Efficient high-resolution imaging at centimeter/millimeter	Q2	Q2	Q1, Q4	Q3	Q1, 2
Very sensitive centimeter imaging		Q1, 2			
Dedicated pulsar timing, transients	Q2, Discovery		Discovery	Q2, 3, 4, Discovery	Q4
Ultrahigh resolution	Q2	Q3	Q4, Discovery		Q1
Complete wavelength coverage		Q2, 4			

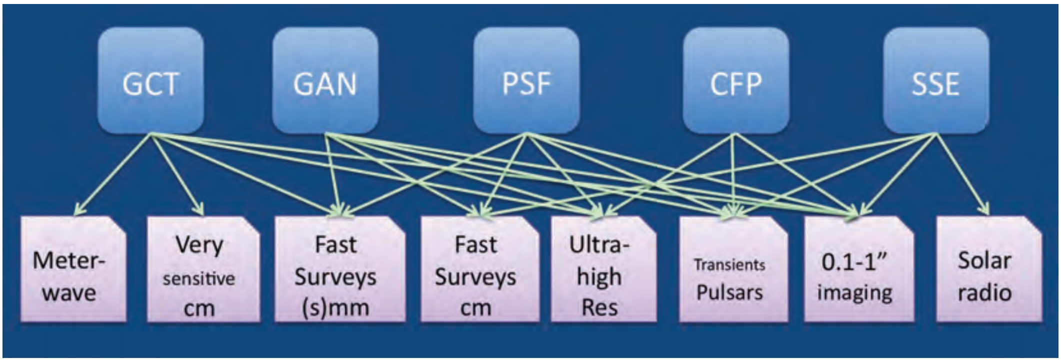


FIGURE 9.6 Mapping of Astro2010 Science Frontiers Panel areas to needed radio, millimeter, and submillimeter capabilities.

Status of Current U.S. National RMS Facilities

National RMS facilities are funded primarily by NSF-AST (Table 9.2). They may be used on an equal basis by the entire international astronomical community under an "open skies" policy. The National Radio Astronomy Observatory operates a number of telescopes under the management of Associated Universities, Inc.,

TABLE 9.2 U.S. Radio Astronomy Facilities

Observatory	Telescope	Wavelength (cm)	Aperture	Collecting Area (sq m)	Angular Resolution	Field of View	Commissioned	Source of Funding
Existing								
NRAO	Very Large Array (VLA)	0.7-400	27 x 25 m	13,250	1″–44″ (λ/21 cm)	30′ (λ/21 cm)	1982	NSF, Mexico, Canada
NRAO	Expanded Very Large Array (EVLA)	0.7-21	27 x 25 m	13,250	1″–44″ (λ/21 cm)	30′ (λ/21 m)	2012	NSF
NRAO	Very Long Baseline Array (VLBA)	0.3-90	10 x 25 m	4,900	0.005″ (λ/21 cm)	30′ (λ/21 m)	1993	NSF
NRAO	Green Bank Telescope (GBT)	0.3-100	100 m	7,850	9′ (λ/21 cm)	9′ (λ/21 cm)	2000	NSF
NAIC	Arecibo Observatory	3-100	225 m	40,000	3′ (λ/21 cm)	3′ (λ/21 cm)	1963	NSF
Hat Creek Radio Observatory	Allen Telescope Array (ATA)	3-60	42 x 6 m	1,200	4′ x 2′ (λ/21 cm)	2.5° (λ/21 cm)	2007	Berkeley, USAF, DARPA, private, NSF
CARMA	Combined Array for Research in Millimeter-wave Astronomy (CARMA)	0.1-0.3	6 x 10 m, 9 x 6 m, 8 x 3.5 m	850	0.2″–4″ (λ/1.3 mm)	44″ (λ/1.3 mm)	2007	NSF, Caltech, Berkeley, Illinois, Maryland
Caltech Submillimeter Observatory	Caltech Submillimeter Observatory (CSO)	0.13-0.035	10.4 m	85	25″ (λ/870 μm)	25″ (λ/870 μm)	1988	NSF, Caltech, Texas
Arizona Radio Observatory	12-Meter Telescope	0.3-0.13	12 m	110	55″ (λ/3 mm)	55″ (λ/3 mm)	1985	Arizona
Arizona Radio Observatory	Submillimeter Telescope	0.13-0.035	10 m	78	25″ (λ/870 μm)	25″ (λ/870 μm)	1993	Arizona

CARA	South Pole Telescope (SPT)	0.1-0.3	10 m	78	42″ (λ/1.5 mm)	42″ (λ/1.5 mm)	2007	NSF
SAO/ASIAA	Submillimeter Array (SMA)	0.035-0.2	8 x 6 m	226	0.3″-3″ (λ/870 μm)	30″ (λ/870 μm)	2003	Smithsonian/ ASIAA
Under Construction								
INAOE/ University of Massachusetts	Large Millimeter Telescope (LMT)	0.13-0.3	50 m	1,960	12″ (λ/3 mm)	12″ (λ/3 mm)	2010+	INAOE and University of Massachusetts
Long Wavelength Array	Long Wavelength Array (LWA)	300-3,000	256 dipoles		2″ (λ/1.5 m)	8° (λ/1.5 m)	2012+	University of New Mexico, NRL, LANL, Virginia Tech, NRAO, USAF, University of Iowa
NRAO	Atacama Large Millimeter/ submillimeter Array (ALMA)	0.3-0.035	50 x 12 m	5,700	0.02″ (λ/1 mm)	15″ (λ/870 μm)	2013	NSF, Canada, ESO, Japan, Chile

with NSF support. NASA occasionally contributes to ground-based RMS facilities for mission-oriented purposes—as, for example, in the NASA-funded upgrade of the VLA that enabled it to join the Deep Space Network to receive transmissions from Voyager 2 at Neptune. The panel lists the national facilities below in order of decreasing age since commissioning.

Arecibo Observatory

The 305-m Arecibo Telescope is a facility of the National Astronomy and Ionosphere Center (NAIC), currently operated by Cornell University under cooperative agreement with NSF. Arecibo is the largest telescope in the world. It is equipped with a 7-element 21-cm array (ALFA) and participates in VLBI. Arecibo is the most sensitive H I telescope, and feasibility studies have begun on a 40-beam survey array for H I surveys of the nearby universe (GAN 1, 2; GCT 2). Arecibo is the most sensitive telescope for pulsar astronomy. Discovery of millisecond pulsars is critical to the pulsar-timing-array effort (CFP Discovery). Arecibo radar, with the NASA Goldstone antenna, is critical to the characterization of potentially catastrophic incoming near-Earth objects (NEOs), the subject of the recent NRC study, *Defending Planet Earth: Near-Earth Object Surveys and Hazard Mitigation Strategies.*[1]

Very Large Array (VLA)

Dedicated in 1980, the NRAO VLA is a centimeter-wave array of 27 25-m dishes, with four configurations corresponding to maximum baselines of 1 to 36 km. A nearby VLBA antenna can be added to give a best angular resolution of 0.04″. The VLA has been a productive instrument, yielding 170 refereed papers per year since the mid-1980s. It is in the process of a major equipment upgrade, the "Expanded VLA," described below.

Very Long Baseline Array (VLBA)

The NRAO VLBA consists of 10 25-m dishes, spread over baselines up to 8,000 km, allowing centimeter-wave imaging at resolutions to 80 micro-arcseconds and astrometry to 10 micro-arcseconds. The VLBA has major scientific impact in science areas that require high astrometric accuracy, such as parallaxes and proper motions. The VLBA has been used to study the proper motions of Sgr A* and

---

[1] National Research Council, *Defending Planet Earth: Near-Earth-Object Surveys and Hazard Mitigation Strategies*, The National Academies Press, Washington, D.C., 2010.

nearby galaxies (GAN Discovery), Keplerian maser disks and black-hole masses in nearby galaxies (GCT 3; GAN 4), precision parallaxes for star-forming regions in the Milky Way, and superluminal motions in AGN jets (GCT 3). Upgrading of receivers and bandwidth is essential to improve the sensitivity of this ultrahigh-resolution capability.

Robert C. Byrd Green Bank Telescope (GBT)

The Robert C. Byrd Green Bank Telescope of the NRAO is a 100-m dish operating from 3 mm to 1 m. Dedicated in 2000, it is the largest fully steerable telescope in the world. In addition to heterodyne receivers, the GBT has the 3-mm bolometer array MUSTANG. The GBT is an efficient millisecond-pulsar machine and a sensitive instrument for mapping H I and recombination lines. It can detect CO and HCN emission from high-redshift galaxies (GCT 2). MUSTANG is being used to study the SZE toward massive clusters (GCT 1). With upgrades (S4.3), the GBT can provide much of the needed capability for fast centimeter-wave surveys (PSF 1, 2).

Current U.S. University Radio Facilities

University radio facilities are an integral part of the U.S. RMS capabilities. Those receiving NSF/AST operations funding are part of the URO program, and the fraction of their observing time funded by NSF is subject to an "open skies" policy. Capital funding for construction of university telescopes is a mix of NSF money and alternative sources such as endowments, gifts, and state funding. Since the closure of the NRAO 12-Meter Telescope in 2000, there has been no national U.S. facility in the millimeter and submillimeter portion of the spectrum. U.S. access for observing at these wavelengths has been provided exclusively by the university facilities, which have also made major contributions to the technological developments leading to ALMA. The panel discusses below the facilities that currently receive operations support from the URO program, and then those currently without URO funding.

Caltech Submillimeter Observatory (CSO)

The CSO operates a 10.4-m telescope on Mauna Kea with an active surface. It can be used as an element of the Submillimeter Array, along with the JCMT. Since 1984, the CSO has pioneered the development of submillimeter heterodyne receivers. Continuum cameras include Bolocam, operating at $\lambda = 1.1$ and 2.1 mm, and SHARC II, at $\lambda = 350$ and 450 μm. Science from the CSO is wide-ranging:

atmospheric chemistry of Earth and other planets, comet chemistry, turbulence in molecular clouds, CO and CI lines in other galaxies, magnetic-field mapping of star-forming clouds, interstellar-cloud chemistry (GAN 2), and the SZE (GCT 1). The CSO has recently completed a 1-mm continuum survey of the galactic plane with Bolocam (PSF 1). These areas of research can now be followed up using ALMA. The CSO will be decommissioned in 2016.

Combined Array for Research in Millimeter-wave Astronomy (CARMA)

The Berkeley-Illinois-Maryland Association Array and the Owens Valley Millimeter Array were merged to form CARMA in 2003, a joint effort of the California Institute of Technology, the University of California, Berkeley, the University of Illinois, and the University of Maryland. The SZE array of the University of Chicago has recently been added. CARMA was a recommendation of *Astronomy and Astrophysics in the New Millennium* (AANM), the 2001 decadal survey report. CARMA consists of six 10.4-m, nine 6.1-m, and eight 3.5-m antennas at a high site, providing arcsecond imaging in the 3-mm and 1-mm atmospheric windows. CARMA images molecular line emission from comets, star-forming regions, nearby and distant galaxies, and the SZE. CARMA can be used for VLBI, including recent observation of Sgr A* at 1 mm (GAN 4; GCT 3). CARMA is just beginning to operate at full capacity. Operations are funded in part by NSF, and as a result, 30 percent of the observing time is "open skies." Upgrades to CARMA (S4.4) can test multibeam systems on interferometers and enhance survey capability (see Table 9.1).

Allen Telescope Array (ATA-42)

The ATA consists of 42 6-m antennas acting together as a fast, wide-field, centimeter-wave mapper, a joint project of the SETI Institute and the University of California, Berkeley. High resolution and a large field of view make the ATA an excellent instrument for mapping degree-scale H I structures such as tidal tails in the local universe (GAN 1, 2). The ATA currently devotes part of its observing time to the search for radio transients over wide fields of view (GAN Discovery; SSE Discovery), and it carries out SETI monitoring (PSF 4) simultaneously with other observing. Technology under development at the ATA, such as large-N correlators, beam-forming techniques, commensal observing strategies, and data management, will influence directions taken in future centimeter-wave arrays, such as the SKA. Construction of the first 42 antennas was funded by Paul Allen. The operations budget is funded modestly by NSF ($200,000 annually) and the state of California, and through partnerships with the USAF and the USNO. These arrangements limit the time available for astronomy.

Arizona Radio Observatory (ARO)

The ARO comprises two telescopes in Arizona run by Steward Observatory, the 12-m telescope (formerly operated by NRAO) on Kitt Peak, and the 10-m Submillimeter Telescope on Mt. Graham, Arizona. Science done at the ARO includes the discovery of new molecular lines, the molecular properties of evolved stars and protoplanetary nebulae (SSE 3), characterization of ion chemistry, refractory chemistry, and organic chemistry in interstellar space, and the kinematics and chemistry of star-forming regions (PSF 1). As one of the few telescopes in the submillimeter VLBI network, the SMT was critical to the detection of Sgr A* at 1.3 mm, which constrained the orientations for an accretion disk (GCT 3; GAN 4). The ARO is currently not supported by the URO program. Observing time is available through collaborations with members of Steward Observatory.

Submillimeter Array (SMA)

The SMA consists of eight moveable 6-meter antennas on Mauna Kea, Hawaii. It is the first telescope capable of providing sub-arcsecond imaging at submillimeter wavelengths. The SMA studies the surfaces and atmospheres of solar system objects, magnetic-field structure in star-forming regions (PSF 1), properties of protoplanetary disks (PSF 2), the chemistry of evolved star envelopes, the supermassive black hole at the galactic center (including VLBI; GAN 4; GCT 3), the cool interstellar medium in nearby galaxies (GAN 2), and starbursts at cosmological distances (GCT 2). The SMA is a joint project of the Smithsonian Astrophysical Observatory and the Academia Sinica Institute of Astronomy and Astrophysics (ASIAA). It has not been supported by NSF. Up to 30 percent of the Smithsonian share of observing time is open to the international community.

South Pole Telescope (SPT)

The 2001 decadal review committee (AANM) recommended construction of the South Pole Submillimeter-wave Telescope (SPST) for observations between 0.2 and 1 mm. The SPT is a project of the University of Chicago, the University of California at Berkeley, Case Western Reserve University, the University of Illinois, and the Smithsonian Astrophysical Observatory. A 10-m telescope has been constructed and is currently observing. The first instrument, a multifrequency bolometric camera, is aimed at detecting thousands of clusters of galaxies through the SZE and measuring the small-scale anisotropy in the CMB. Results of initial surveys are very encouraging. The Office of Polar Programs at NSF funds the SPT as an experiment, not as a facility that will be open to guest investigators.

RMS Facilities Currently Under Construction with U.S. Participation

Enhanced Very Large Array (EVLA)

A long overdue upgrade of the NRAO VLA, with help from Canada and Mexico, will be completed in 2012, and will result in a factor-of-10 improvement in continuum sensitivity and in broadband spectral-line access. The centimeter-wave sources studied with the EVLA will include jet emission associated with AGN and galactic micro-quasars (GCT 3), masers, and extragalactic H I emission, radio supernovae (GAN Discovery) and supernova remnants, free-free emission and radio recombination lines from star-forming regions in the galaxy (PSF 1) and beyond, and CO emission from high-z galaxies (GCT 2). The improvement in continuum sensitivity will make the EVLA sensitive to thermal sources of emission at the higher frequencies, allowing observations of free-free emission from compact objects and stellar photospheres. Upgrades to improve wavelength coverage and imaging would improve the capability for efficient high-resolution centimeter-wavelength imaging.

Atacama Large Millimeter Array (ALMA)

ALMA will be the largest ground-based astronomical facility ever built (Figure 9.7). Under construction in the Atacama Desert in northern Chile at an altitude of 5,000 meters, ALMA consists of a main array of 50 12-meter antennas, which are reconfigurable to allow a wide range of angular resolutions down to 4 milliarcseconds at submillimeter wavelengths. ALMA is supplemented by the Atacama Compact Array (ACA), which can image large-scale structures. ALMA is a partnership of Europe, North America, and East Asia in cooperation with the Republic of Chile. ALMA is funded in Europe by the European Southern Observatory (ESO), in North America by NSF in cooperation with the National Research Council of Canada and the National Science Council of Taiwan (NSC), and in East Asia by the

FIGURE 9.7 Five antennas at ALMA's high-elevation Array Operations Site (December 2009). SOURCE: Nick Whyborn, ALMA (ESO/NAOJ/NRAO).

National Institutes of Natural Sciences (NINS) of Japan in cooperation with the Academia Sinica (AS) in Taiwan. ALMA construction and operations are led on behalf of Europe by ESO, on behalf of North America by NRAO, which is managed by Associated Universities, Inc. (AUI), and on behalf of East Asia by the National Astronomical Observatory of Japan (NAOJ). The North American partners are investing \$500 million for construction and \$30 million per year for operations and are entitled to ~35 percent of the observing time. The North American ALMA Science Center (NAASC) will provide full-service user support for the North American community. Three key science goals drive the ALMA specifications: (1) detect spectral line emission from CO or CII in normal galaxies at a redshift of $z = 3$, in 24 hours (GCT 2); (2) image the gas kinematics in protoplanetary disks around young Sun-like stars at the distance of the nearest star-forming clouds (PSF 1); and (3) provide high-fidelity images at an angular resolution of 0.1″ to match current OIR capabilities. With ALMA, one can use diagnostics such as thermal dust continuum emission, molecular rotational lines, and atomic fine-structure lines to study sources as diverse as planetary atmospheres, cometary nuclei, molecular cloud cores, protostellar jets, circumstellar and protoplanetary disks, the immediate environment of the galactic center supermassive black hole, and star-forming galaxies from the present day to $z > 10$ (PSF 1, 2; SSE 1, 3; GAN 1, 2; GCT 2, 4). Early science is expected to start in 2011, transitioning to full operations in 2013. Upgrades to achieve full wavelength coverage, when combined with EVLA coverage would, enhance this capability.

Large Millimeter Telescope (LMT)

The LMT is a 50-m telescope under construction at 4,600-m altitude on Sierra Negra in Puebla, Mexico. It is a collaboration of the Instituto Nacional de Astrofísica Óptica y Electrónica (INAOE) and the University of Massachusetts at Amherst. The LMT will observe dust and molecular gas from the solar neighborhood to high-redshift galaxies, providing key information on how stars form (PSF 1), the stellar initial mass function (PSF 1), how giant molecular clouds form and evolve in galaxies (GAN 2), how matter cycles into and out of galaxies (GAN 2), and how cosmic structures form and evolve via observations of the SZE (GCT 1). Large surveys and fast mapping speeds are essential to these science goals. First-light instruments have been completed and used as guest instrumentation on existing telescopes. The antenna structure and 32 m of the planned 50-m-diameter primary mirror surface have been completed, and integrated testing has begun. Commissioning was anticipated to begin in 2010. The LMT has a small amount of funding through the URO program.

Long Wavelength Array (LWA)

The long-wavelength portion of the electromagnetic spectrum is not well explored, particularly the transient sky. The Long Wavelength Array (LWA) is a unique U.S. facility aimed at this science. Situated in New Mexico, the LWA has a relatively quiet radio-frequency-interference environment, with partial access to the southern sky, and thus complements the European LOFAR effort. The primary goals of LWA are to survey the long-wavelength sky, to search for decametric emission from exoplanets, and to map the circumgalactic medium. The LWA is a joint effort of the University of New Mexico, the Naval Research Laboratory, Virginia Tech, the Jet Propulsion Laboratory, LANL, and the University of Iowa. The initial station, out of a planned 52, is under construction near the VLA site with funding from NRL; Phase 1 operations are expected in 2011. The proposed 52-station LWA would require additional funding; the 16-station core could be operating as early as 2017, and the full array by 2019.

The International Context

In the previous section and in Table 9.2, only facilities with U.S. funding are described. In this section the panel describes the international context in terms of providing the U.S. community with the capabilities it needs to meet the decadal science goals.

ALMA will be a powerful new facility, giving unparalleled sensitivity and resolution to the submillimeter and millimeter sky. But it has a restricted instantaneous field of view, several arcseconds across at the shortest wavelengths. The ACA provides some wide-field mapping capability for extended fields on arcminute scales at the longer wavelengths. However, taking full advantage of ALMA requires a finder scope, a very-wide-field submillimeter telescope for wide-area surveys, playing the role of the LSST for the GSMT in the millimeter/submillimeter region. During the past two decades the Europeans and the Japanese have moved actively into millimeter and submillimeter research, and recently into southern ALMA pathfinder telescopes. The French-German-Spanish IRAM operates two Northern Hemisphere facilities, a 30-m single-dish telescope located on Pico Veleta, Spain, and the Plateau de Bure Interferometer (PdBI), with six 15-m telescopes in the French Alps, for observing in the millimeter atmospheric windows. PdBI has 40 percent more collecting area than CARMA but fewer baselines for imaging, with restricted access to southern sources such as the galactic center. IRAM is seeking funding to expand the PdBI by doubling the number of antennas and the longest baselines, which would bring its sensitivity to within a factor of 3 of ALMA at 3 mm. The APEX (Atacama Pathfinder EXperiment) Telescope is a 12-m ALMA prototype antenna operating at Chajnantor near the ALMA site, equipped with heterodyne receivers operating from 1.3 mm to ~210 μm, and bolometer arrays at 870 and 400 μm. APEX is a collaboration between the Max-Planck-Institut für Radioastronomie, Onsala Space Observatory, and ESO. The ASTE (Atacama Submillimeter Telescope) is a 10-m telescope operated by Japan (NAOJ) at Chajnantor at 870 μm. The Chajnantor telescopes are important submillimeter pathfinders for the southern sky. Access by U.S. astronomers to these facilities is limited.

The 45-Meter Telescope of the Nobeyama Radio Observatory is a facility of the NAOJ in Japan that operates at 3 mm and can be used by U.S. astronomers with a Japanese collaborator (the Nobeyama Millimeter Array is no longer open for general observing). The James Clerk Maxwell Telescope, run by the Joint Astronomy Centre (United Kingdom, Canada, the Netherlands) is a 15-m telescope on Mauna Kea operating in the submillimeter. Its bolometer camera—the highly successful SCUBA, operating at 870 and 450 μm—was responsible for the discovery of submillimeter galaxies in the distant universe. Its successor, SCUBA-2, with a projected mapping speed a thousand times SCUBA's, has recently received its science detectors. ESA's Herschel Space Observatory is opening the part of the far-infrared-submillimeter spectrum blocked from the ground with deep surveys and spectroscopy; although this is a European-led mission, NASA has funded some instrumentation, and some key projects are led by U.S. principal investigators.

Pulsar observing is a key area for decadal research. For pulsar-timing experiments, sensitivity is important, requiring telescopes with large collecting areas, large bandwidths, and sites with low radio-frequency interference. Many centimeter-wave telescopes around the world can contribute to this effort, but Arecibo affords the highest precision time-of-arrival measurements. The GBT and the Parkes telescope (of the Australia National Telescope Facility) have the advantage of being at low-radio-frequency-interference sites. After 2014, the FAST Telescope (Five-hundred-meter Aperture Spherical Telescope), under construction in Guizhou Province, China, will allow very sensitive pulsar work, with a collecting area twice that of Arecibo, albeit initially with less bandwidth.

The ability to determine celestial positions and proper motions with a high degree of precision is a key decadal science goal. The VLBA is unique in providing a platform for precision astrometry at centimeter wavelengths. None of the worldwide VLBI arrays remotely approaches the capabilities of the VLBA—especially when it is augmented with the VLA, GBT, and Arecibo as station elements—in terms of sensitivity, high-frequency coverage, rapid switching for phase calibration, intrinsic resolution, and throughput of a dedicated and highly tuned instrument, as compared to a network such as the European VLBI Network, which is based on part-time use of inhomogeneous, general-purpose radio telescopes. Improvements to the VLBA as part of the NRAO development plan will put the VLBA even farther ahead of other arrays.

The future of centimeter-wave observing has been dominated by visions for the Square Kilometer Array. As discussed below, the SKA has evolved into three instruments. Demonstration arrays for SKA-mid (3 to 100 cm) are under construction. The South African MeerKAT (Meer Karoo Array Telescope) is based on 12-m dishes with broadband, single-pixel feeds, with plans to have 80 dishes on 8-km baselines by mid-decade. The Allen Telescope Array (ATA) has taken a similar approach. The Australian SKA Pathfinder (ASKAP) uses 12-m antennas with a 30-beam array operating at L-band ($\lambda \sim 16$ to 40 cm) and plans for 36 antennas. The European project A3IV (Aperture Array Astronomical Imaging Verification) uses aperture arrays at L-band to synthesize very large fields of view. All three SKA-mid concepts (broadband single-pixel, L-band focal plane array, L-band aperture array) have great potential to inform the final design of SKA-mid (~3,000 15-m antennas).

The meter-wave portion of the spectrum is divided between epoch of reionization experiments and general-purpose facilities. There are currently four first-generation EoR observatories: LOFAR in the Netherlands, GMRT of the Tata Institute of Fundamental Research in India, and PAPER and the MWA by U.S. teams in western Australia (an initial deployment of PAPER will occur in South Africa). All four arrays are striving to detect, and provide rough characterization of, the power spectrum of EoR H I emission—qualitatively similar to COBE's CMB power-spectrum constraints. The main concern for EoR power-spectrum instruments is achieving the extremely high spatial, polarimetric, and spectral purity needed to subtract foreground emission. Collectively PAPER and the MWA form the first-generation HERA-I experiment and are leading instruments in terms of both sensitivity and the development of the precision calibration needed for foreground subtraction. HERA-II is envisioned as a second-generation EoR experiment that will use lessons learned from the first-generation instruments of HERA-I to measure the EoR power spectrum in detail. There are no plans to expand either the GMRT in India or LOFAR in Europe into a HERA-II competitor.

The GMRT and LOFAR are also capable multipurpose low-frequency arrays, in which only the shortest baselines (<<1 km) and the 110- to 200-MHz band ($z = 12$ to 6) are used in EoR observing. The GMRT spans $\lambda = 20$ to 600 cm with 30 45-m dishes and baselines up to 25 km. LOFAR is under construction and features two aperture arrays covering $\lambda = 4$ to 30 m and $\lambda = 1.2$ to 2.7 m with baselines up to ~1,000 km. The lower-frequency capabilities of LOFAR are similar to those of the LWA project in New Mexico, although LOFAR is sited far north in a challenging radio-frequency-interference environment.

Current Capabilities Within the RMS Observatory Suite

The current suite of RMS facilities provides a broad set of functionalities that can address some of the science questions of the next decade. The panel briefly summarizes these capabilities below.

- *Ultrahigh-resolution imaging and astrometry.* Imaging at milli-arcsecond resolution is a requirement for many areas of astrophysics. The VLBA in concert with the EVLA provides such capability, but some science questions require higher frequencies (GCT 3, 4). The ability to measure astronomical coordinates with micro-arcsecond precision (GAN Discovery) allows the measurement of distances across the galaxy, including the galactic center, and the three-dimensional motions of nearby galaxies. This precision-astrometry capability is provided exclusively by the VLBA, and upgrades to receivers and bandwidth are needed.
- *Pulsar timing.* Pulsar-timing experiments provide a promising way to measure gravitational waves (CFP Discovery). Centimeter-wave observatories with large collecting areas can do this: Arecibo, Green Bank Telescope, and the EVLA. However, the restrictions imposed by the observing cadences required for pulsar-timing experiments complicate observing on general-purpose telescopes, and while Arecibo is very sensitive, the declination limits are restrictive. Progress in pulsar-timing experiments would accelerate with a dedicated facility.
- *Efficient, high-resolution centimeter-wave imaging.* Interstellar and circum-galactic gas structures require sensitive imaging in the centimeter continuum, for studies of synchrotron emission and magnetic fields (GAN 1, 2), and for the relatively weak 21-cm line of hydrogen (GAN 1, 2; GCT 2). Arecibo is the most sensitive H I mapper, but it has declination limits and more importantly, its 3′ beam at 21 cm is a good match only for galactic and local-universe H I studies. The EVLA would require two orders of magnitude more sensitivity to perform cosmological H I studies on arcsecond scales, and it has a relatively limited field of view. The ATA-42 has a wide field of view, suitable for large-scale mapping, but it also lacks sensitivity to detect H I on small scales and would require an increase of three orders of magnitude in collecting area to reach the level called for in Table 9.1.
- *Efficient, high-resolution millimeter/submillimeter-wave imaging.* Studies of the emission from CO in galaxies at redshifts up to $z \sim 3$ and above comprise an essential element of studies of galaxy evolution (GCT 2). Redshifted CO lines can be observed with the EVLA and with ALMA, with sufficient spectral resolution to do galaxy kinematics. Many of the galaxies observable by ALMA will be nearly as bright at high redshifts as they are at lower redshifts, owing to the shift of the far-infrared spectral peak into the millimeter/submillimeter band, the "negative K correction" (GCT 2). Full wavelength coverage is needed to cover all redshifts.
- *Transient sources.* The ATA-42 is currently the only RMS observatory with dedicated time for transient science, as opposed to transient follow-up. The time available for this purpose is limited by the minimal level of NSF support.

Potential to Lose Capabilities

In 2005, NSF initiated a senior review of its ground-based telescopes. Its recommendations with regard to RMS facilities, which are nearly all under the jurisdiction of NSF, are the following:[2]

> *Radio-Millimeter-Submillimeter Astronomy Transition Program.* The National Astronomy and Ionosphere Center and the National Radio Astronomy Observatory, which are heavily subscribed by other communities, should seek partners who will contribute personnel or financial support to the operation of Arecibo and the Very Long Baseline Array respectively by 2011 or else these facilities should be closed.

In response to the senior review, NSF recommended a ramp-down in NSF funding of Arecibo from $10.5 million in 2007 to $4 million in 2011. The panel understands that the Arecibo Observatory leaders are actively pursuing other funding sources, but the prospects are very uncertain. The future funding of the VLBA is also uncertain, but efforts to find other funding are ongoing.

The panel has noted earlier that important capabilities identified by the SFPs are provided by these facilities. The pursuit of gravitational waves (CFP Discovery) is dramatically enhanced by Arecibo (Figure 9.8). The large sky visibility of the GBT is strongly complemented by the better sensitivity, though limited sky coverage, of Arecibo; both are needed to detect and study nanohertz gravitational waves. High-precision astrometry (GAN Discovery) is the exclusive province of the VLBA. If these two facilities do not obtain other sources of funds, capabilities that have been identified by two of the SFPs as key discovery areas will be lost. Since the time of the senior review, developments in the areas of transient astronomy and astrometry have dramatically changed the context in which these instruments are viewed. Because of the importance of Arecibo for pulsar timing and galaxy evolution, the panel recommends restoring $2 million per year to its baseline budget. The importance of astrometry likewise justifies continued support of the VLBA.

Capabilities Missing or Inadequate Within the Current Portfolio

Consideration of the science goals set forth by the SFPs reveals the following capability gaps in the current RMS portfolio (Table 9.3).

- *Sensitive meter-wave imager.* Studies of the epoch of reionization require new capabilities at meter wavelengths.
- *A dedicated solar radio telescope.* A radio complement to ATST would deliver full-disk images with rapid cadence over a wide range of wavelengths.

---

[2] Information on the NSF 2005-2006 AST Senior Review is available online at http://www.nsf.gov/mps/ast/ast_senior_review.jsp. Accessed February 2011.

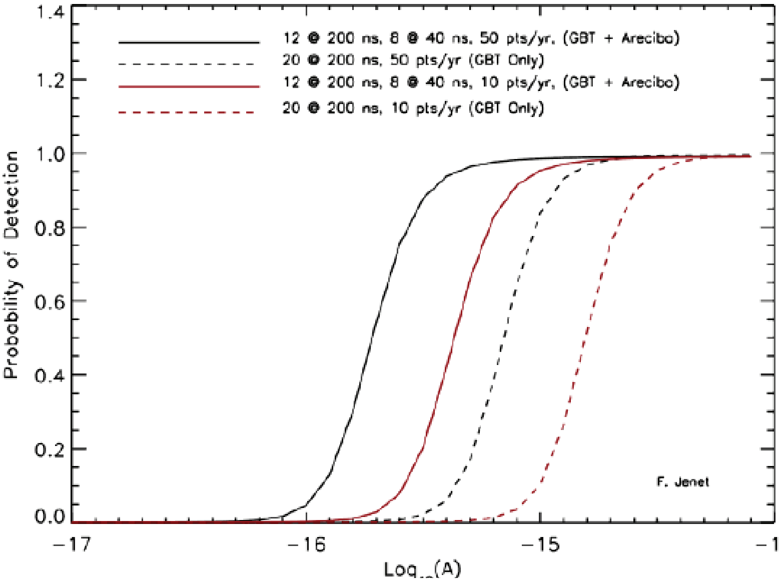


FIGURE 9.8 The probability of detecting a stochastic background generated by an ensemble of supermassive black-hole binaries as a function of the energy density in gravitational waves (normalized by the closure density of the universe). The solid line assumes 20 pulsars, 12 observed with the Green Bank Telescope and 8 observed with Arecibo. The dashed line assumes 20 pulsars observed with the GBT alone. Both lines assume 5 years of observations, with two observations of each pulsar per month, for the same net telescope time. SOURCE: Fredrick Jenet, University of Texas, Brownsville.

TABLE 9.3 Capabilities Provided by Current Facilities

Capability	Current Facilities	Needed Development
Cosmic microwave background program	Existing program	Continue successful program
Sensitive meter-wave array	Demonstrators only	Factor-of-10 more area
Solar radio telescope	Demonstrators only	Full range, fast imager
Fast millimeter/submillimeter surveys	CSO until 2016	Bigger, faster, southern skies
Fast centimeter surveys	ATA-42, GBT, Arecibo	Expand ATA, multibeam GBT, Arecibo
Efficient high-resolution imaging at centimeter/millimeter	VLA, soon EVLA, ALMA, CARMA	Enhance EVLA, ALMA, CARMA
Very sensitive centimeter imaging	None	About 1-square-kilometer area array
Dedicated pulsar timing, transients	Nothing dedicated, partial support from Arecibo, GBT	Expand ATA, enhance and dedicate time on Arecibo, GBT
Ultrahigh resolution	VLBA	Improve sensitivity, astrometric accuracy, develop millimeter/submillimeter VLBI
Complete wavelength coverage	EVLA, ALMA	Add missing bands

- *Fast millimeter and submillimeter survey instrument.* ALMA will have a very limited instantaneous field of view, only arcseconds across at the highest frequencies. It can map larger fields at the longer wavelengths in combination with the ACA, but large-scale surveys—such as surveys of the galactic plane—would be extremely time-consuming and an unwise use of ALMA time. A large-field mapper operating at millimeter and submillimeter wavelengths is required to pave the way for higher-resolution follow-up observations with ALMA.
- *A very sensitive centimeter-wave imager.* To image H I in galaxies beyond z of about 0.1 requires much greater sensitivity than even the EVLA achieves. Because receivers have reached their sensitivity limits, only collecting areas on the order of a square kilometer will provide this capability.
- *Dedicated transient instruments across the RMS spectrum.* Transients by their nature require large amounts of dedicated survey time and flexibility in observing cadences. Predictable sources such as pulsars can be scheduled, albeit with some difficulty, along with traditional programs at existing telescopes. Serendipitous transients require a dedicated instrument.
- *Ultrahigh-resolution imaging in the millimeter and submillimeter.* The ability to image Sgr A* in the radio regime free of scattering requires more sensitive submillimeter VLBI observations than are currently possible.

FUTURE PROGRAM

Introduction

The panel presents a future program that is balanced: across scientific fields; across wavelength regions; between capabilities for fast, wide-field surveys—enabled by large, single-dish telescopes equipped with new generations of large-format detector arrays—and for high-resolution, high-sensitivity studies of objects found in surveys. It is balanced among large national/international facilities, small university facilities, and PI-driven projects. The panel includes recommendations for technological developments and relevant laboratory and theoretical programs. The program includes innovative new facilities and cost-effective upgrades to existing facilities.

The RMS Panel was presented with a wide range of possible activities. In prioritizing the possible activities, the panel has been guided by the work of the SFPs. The panel used the reports of the SFPs to identify needed RMS capabilities (see Table 9.1). Then the panel summarized the capabilities that are supplied by existing facilities and identified the capabilities that are missing (see Table 9.3). In this section, the panel shows how the recommended program provides these capabilities, employing a combination of building new facilities and of sustaining and developing current capabilities. The panel discusses some small and moderate programs that add needed capabilities and address some system issues. This panel's final prioritization is done in the "Recommendations" section below.

New Facilities for Mid-Scale Funding

Hydrogen Epoch of Reionization Array (HERA)

The study of cosmic reionization is currently at the forefront of astrophysical research and is highlighted by the GCT Panel. In the epoch of reionization, fluctuations in neutral hydrogen trace directly the fluctuations in the matter density of the universe; observing these fluctuations provides new constraints on the physics of the early universe and the earliest luminous sources. These fluctuations can be observed only in the redshifted 21-cm emission from neutral hydrogen; hence, radio observations in the meter-wave range are a unique probe of this previously unexplored epoch of cosmic evolution.

To explore the EoR, a meter-wave array (1 to 3 m, or $z = 5$ to 15 for the H I hyperfine line) is needed (see Table 9.1). Such a capability would further enable studies of the solar corona and solar wind as well as searches for variable sources (another potential area of discovery). HERA provides a program to achieve such capabilities over the next decade and beyond in a staged approach.

The HERA program consists of three major steps (or stages). The goal of HERA-I is to detect the reionization signal and to measure a few of its most general properties, such as the power spectrum, over a limited range of spatial scales and cosmic redshifts. The HERA-I program is currently being actively pursued in the United States, spearheaded by the Murchison Widefield Array (MWA) and the Precision Array to Probe the Epoch of Reionization (PAPER), which are testing alternative approaches (Figure 9.9). The collecting area of each of these experiments is on the order of 0.01 square kilometers or less. Continued support of MWA and PAPER at about $5 million per year for about 5 years will be needed to complete the HERA-I stage.

The second stage of the program (HERA-II) aims at detailed characterization of the power spectrum of the fluctuations and other statistical measures of the signal. The HERA-II experiment will require approximately a factor-of-10 increase in the collecting area (to about 0.1 square kilometers).

The HERA-III stage aims at direct imaging of neutral hydrogen during the reionization epoch. Such an instrument would require on the order of 1 square kilometer of collecting area and is a natural candidate for the long-wavelength ($\lambda >$ 1 m) component of the Square Kilometer Array project. Even in the most optimistic scenario, construction of such a telescope cannot start earlier than 2020. HERA-II will set the stage for HERA-III.

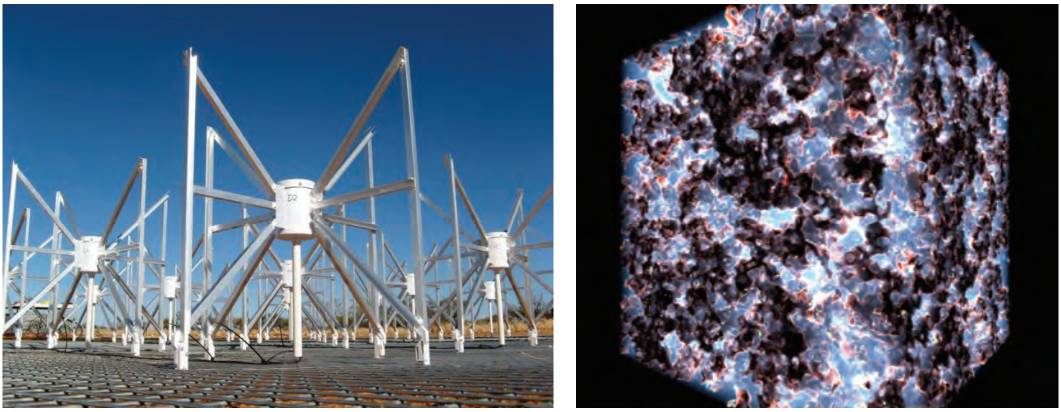


FIGURE 9.9 *Left:* Some elements of the Murchison Widefield Array. *Right:* Tilted data cube from a simulation of the reionization of the universe. The box is 1 Gpc h^{-1} on a side and is a periodic volume at redshift $z \sim 9$, when the universe was about half ionized in this model. Ionized regions are blue and translucent, ionization fronts are red and white, and neutral regions are dark and opaque. A random sampling of 5 percent (about 40,000) of all halos at $z = 0$ is shown as yellow points. Reionization is still quite inhomogeneous on these large scales, with large regions ionizing long before others. SOURCE: *Left:* MIT Haystack Observatory. *Right:* Marcelo Alvarez, Canadian Institute for Theoretical Astrophysics, University of Toronto.

A crucial technological challenge for the HERA program is presented by the exquisite dynamic-range requirements. The expected cosmic signal is dominated by the system temperature and the foreground emission from the Milky Way and external galaxies. To remove these contaminants, the system calibration and the removal of the cosmic foregrounds must be achieved at about 1 part in a million precision. Such precision has so far only occasionally been obtained in radio observations; the success of both the HERA-I and HERA-II experiments will hinge on the ability of the project teams to achieve this precision routinely over extended periods of time.

The cost of HERA-II remains uncertain. The panel estimates a construction cost of $85 million, balanced between the cost in the plan presented by the HERA consortium and an independent cost analysis. The U.S. portion of the cost will depend on the in-kind and financial contributions from international partners, in particular Australia, where the HERA-II experiment may be located. The panel recommends an evaluation of HERA-I results in about 5 years. If specific milestones are met, there should be an open competition for HERA-II.

Frequency-Agile Solar Radiotelescope (FASR)

The study of stellar magnetism is the first question raised by the Panel on Stars and Stellar Evolution. In the words of that panel, the Sun "continues to be a working template for understanding magnetohydrodynamics and plasma physics 'in practice'—physics that is crucial in many other arenas" ("Introduction" in Chapter 5, this volume). The panel identified a major contribution to this goal: a dedicated radio-survey instrument studying the Sun (see Table 9.1). Existing solar radio facilities such as NRAO's Green Bank Solar Radio Burst Spectrometer (BSRBS) and NJIT's Owens Valley Solar Array (OVSA) have demonstrated the potential of solar radio diagnostics for characterizing the strength and topology of magnetic fields in the Sun's corona and for capturing the dynamics of solar eruptions and associated particle acceleration. However, to probe solar magnetism in detail with radio observations, the high-resolution imaging spectroscopy over three-and-a-half decades of wavelength (6 m to 1.5 cm) of the Frequency Agile Solar Radiotelescope is required (Figure 9.10).

FASR consists of three arrays: the A array covers 1.4 to 15 cm with an array of ~100 steerable 2-m antennas, spread over about 4 km; the B array covers 12 to

FIGURE 9.10 Artist's conception of the FASR Array. SOURCE: Isaac A. Gary, New Jersey Institute of Technology.

100 cm with an array of ~70 steerable 6-m antennas, spread over about 4 km; the C array covers 1 to 6 m with 50 non-steerable antenna stations, each with a small array of electronically steered dipole antennas. All arrays are fixed in position. A prototype antenna for array A has been evaluated, and a costed design exists for the 6-m antenna. The correlator technology is not challenging.

FASR will produce a full Stokes spectrogram of the complete solar atmosphere with arcsecond resolution (λ/1.5-cm arcseconds), providing a CAT-scan-like probe of the temperature, density, and magnetic field, from the chromosphere through the corona, once every second. These observations will be unprecedented and will address the question of how magnetic fields power a star's chromosphere and corona. Their full-Sun, continuous nature will enable FASR to act as a solar survey instrument. Its operational lifetime will span a 22-year solar magnetic-activity cycle, throughout which it will gather a rich database of solar eruptions. Thus FASR will be an essential complement to the Advanced Technology Solar Telescope, which provides high-spatial-resolution, but small field-of-view, coronal-magnetic-field observations. It will be unique in its ability to monitor magnetically driven transient activity across the solar disk and limb. Data-pipeline products, such as two-dimensional magnetograms of the coronal base to be produced each second, are planned to ensure broad access to the science observations. This will help to ensure synthesis with coronal observations at longer wavelengths and coordination with the modeling efforts needed for effective interpretation of the data.

FASR is a mature effort that has been recommended in two previous NRC decadal surveys. Extensive development and design reviews have been funded and achieved, including the testing of prototype instrumentation and the detailed planning of FASR operations, maintenance, and management. Independent analysis of FASR characterized it as "doable today." It could be completed by 2015-2018. The ATST current schedule is for scientific operations to begin in 2017, and so the complementarity between ATST and FASR should be exploitable. The project cost was estimated at $68 million for construction; independent estimates put the cost at $109 million based primarily on higher estimates for project management, antennas, and reserves. The operations costs were estimated at $3 million per year (project) and $4 million per year (independent).

Because FASR science is intrinsically interdisciplinary—the fundamental astrophysical processes of magnetism and stellar dynamics that it probes have a direct impact on Earth's space environment—it has excellent prospects for cross-directorate funding. Indeed, support to date has been split between two NSF directorates, Mathematical and Physical Sciences (which includes the AST division) and Geosciences (which includes the ATM—now AGS—division). FASR has a current "Pathfinder" proposal submitted to AGS with a budget of about $8 million that would demonstrate a limited subset of FASR's science goals and could be upgraded in stages to achieve the full FASR implementation. The evaluation of FASR has been based on its full implementation, which the panel judges to be ultimately necessary to achieve critical science goals. The panel adopted cost appraisals closer to those of the independent estimate, allowing for possible higher costs for antennas and a larger reserve than estimated by the project. The panel assumed a total cost of \$100 million for construction and \$4 million for operations, both split equally between NSF-AST and NSF-AGS. The panel recommends FASR with priority equal to that of CCAT.

CCAT

The GCT, GAN, and PSF reports identified a need for a fast millimeter/submillimeter survey instrument, which would provide essential input to ALMA observations (see Table 9.1). This need is met by a Southern Hemisphere, large-aperture submillimeter telescope equipped with large-format continuum detector arrays and spectroscopic instrumentation. The revolution in incoherent detector array technology will enable arrays with 10^5 to 10^6 detectors, each detector very sensitive because it is broadband. Such arrays cannot be used in interferometers, but on single-dish telescopes they can provide surveys of large areas, revealing rare objects—such as the recently discovered starburst QSO at $z > 6$—and large samples for statistical analysis. Such a survey telescope plays a role relative to ALMA analogous to that of the 48-inch Schmidt relative to the 200-inch Palomar telescope.

The proposed CCAT is a 25-m-diameter-aperture facility to be located near the ALMA site and designed to operate at submillimeter wavelengths ($\lambda = 0.2$ to 3 mm) (Figure 9.11). CCAT will be equipped with megapixel-scale detector arrays, enabling it to execute large-scale surveys and mapping with an exceptional continuum sensitivity of 2 mJy $s^{0.5}$ per pixel at 1 mm. About half the observing time will be devoted to large surveys that will provide essential source catalogs for ALMA, with much less source confusion than smaller telescopes operating at longer wavelengths. Thus CCAT will complement surveys at millimeter wavelengths. NSF participation will ensure community access to the survey results.

In terms of SFP science drivers, CCAT will be able to (1) study structures of molecular regions; (2) identify young circumstellar disks via their submillimeter excesses; (3) survey young embedded submillimeter sources in dense molecular clouds and assess their relationship with the stellar initial mass function; (4) map thermal dust emission from nearby galaxies; (5) identify primeval submillimeter galaxies out to the epoch of stellar bulge and supermassive black hole build-up; and (6) detect large-scale structures in the early universe via the SZE. *These CCAT wide-field surveys will drive much of the science to be done with ALMA.* The eventual inclusion of multiobject spectrometers will enable early-universe galaxy-

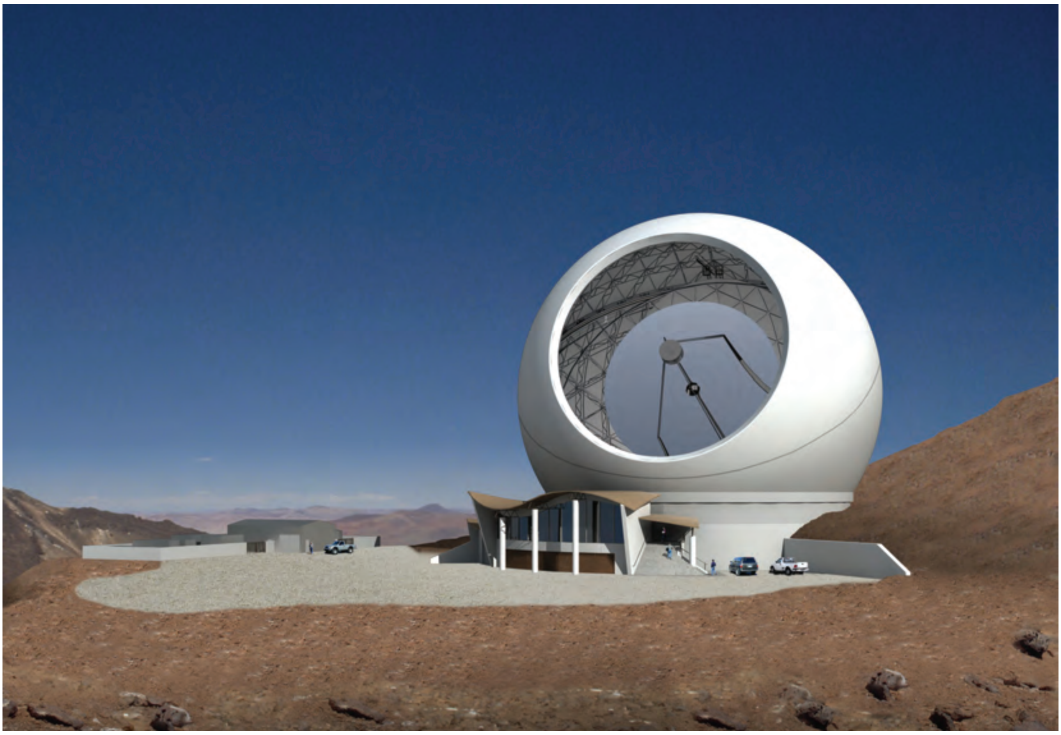


FIGURE 9.11 Design of CCAT at its planned location 600 meters above the ALMA site. SOURCE: CCAT; see www.submm.org.

redshift searches and detailed studies of star-forming regions and starburst galaxies. The potential technology risks are the maturity of focal-plane design and the pointing-control-system performance. These risks are correlated with image quality, which impacts sensitivity and the confusion limit of data. Detector development is progressing rapidly, however, as discussed below in the subsection "Technology Development."

The project expects first-light in 2017, but an independent estimate suggested 2024. The panel believes that 2017 is more nearly correct if funding is available on the schedule requested. While ALMA will start operations before this time, it will operate for decades, and the panel expects ALMA to focus first on the brightest submillimeter galaxies, which will be found by other surveys. CCAT's lower confusion limit will provide samples of galaxies to lower levels in the luminosity function that will be needed as ALMA matures.

The construction costs estimated by the project ($110 million) are similar to an independent estimate ($138 million with a 30 percent reserve). These estimates include the costs of the telescope and optics, facility, software, initial instrumentation, and both project management and system engineering. Operations costs are estimated to be \$10 million per year (project) to \$11 million per year (independent). Note that only \$33 million of the construction costs will be requested from NSF, with the rest coming from university and foreign partners. Based on independent estimates, the panel assumes an NSF share for operations and production of public databases at \$7.5 million per year. Because the CSO, funded at \$2.5 million per year, will close when CCAT opens, the cost increase will be \$5 million per year.

The panel recommends participation at this level with priority equal to that of FASR.

Sustaining and Developing Current Activities

The new facilities recommended above leave some key requirements, as identified in Table 9.1, unmet. There is an excellent suite of existing telescopes, some new telescopes will be completed, and others will be upgraded in the coming decade. There are also ongoing programs that meet key science goals. What remaining needs could be satisfied by continuing successful programs or with upgrades and development on these telescopes?

The remaining capabilities that are needed to address the science questions are the following: a vigorous program of ground-based CMB studies; an instrument dedicated to transients and pulsar timing; fast-survey capability at centimeter wavelengths; improved sensitivity and wavelength coverage on high-resolution (0.1″ to 1″ resolution) imaging telescopes; and improved sensitivity at ultrahigh resolution (micro-arcseconds). In this section, the panel lays out a plan to achieve these capabilities by sustaining and developing current activities.

Studying the Cosmic Microwave Background

As a high priority, the RMS Panel recommends continuing a suite of measurements of the CMB temperature and polarization anisotropy. Although much has been learned from the study of the CMB, there is much more still to be learned. The CMB community is actively pursuing observations that will complement those from the Planck satellite and deliver exciting science throughout the decade (Figure 9.12).

A primary goal of large-angular-scale polarization measurements ($\theta > 1$ deg, $l < 200$) is to measure the presence of primordial gravitational waves as revealed through the B-modes. The limit on B-modes may be improved by roughly a factor of five before multifrequency techniques are required to remove astrophysical foregrounds. Telescopes for the $l < 200$ measurements are ~1 to 2 m in size. The measurements are done from balloon (NASA) and ground (NASA/NSF/DOE)

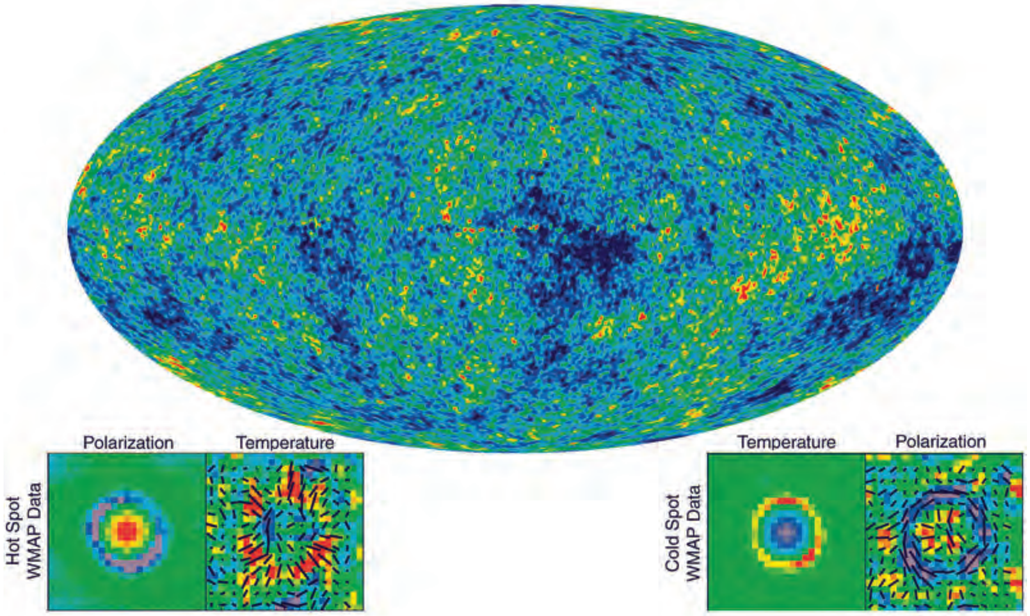


FIGURE 9.12 The WMAP full-sky map of temperature fluctuations of the cosmic microwave background. The galactic signal has been subtracted. The lower-left inset shows the averaged temperature and polarization maps around temperature hot spots, and the lower-right inset shows those around cold spots. The characteristic polarization patterns—radial around hot spots and tangential around cold spots at about 1 degree from the temperature peaks—were predicted in 1994 and first observed in 2010. The patterns show the motion of baryonic gas into cold spots and out of hot spots in response to gravitational potential wells laid down at the birth of the universe. SOURCE: NASA/WMAP Science Team.

platforms. The successful detection of the B-modes, which may occur within the decade, will require substantial confirmation. As the field matures, larger and more sophisticated detection techniques will have to be employed.

The primary goals of the small-angular-scale measurements of temperature and polarization ($\theta < 0.1$ deg, $l > 2{,}000$) are to determine the parameters that characterize the primordial field (or fields), to quantify the cosmic neutrino content, and to measure the process of cosmic evolution with the SZE in thousands of clusters of galaxies. These angular scales are largely unexplored, but great strides can be made before the cosmology is limited by astrophysical foregrounds. The existing ACT (6 m) and SPT (10 m) efforts are part of this program and should continue to receive support.

The excitement of the science has driven the development of thousand-element cameras operating at 0.3 K, of superconducting electronics, and of compact coherent polarimeters. These developments will pave the way for a future satellite, CMBPol, aimed at measuring the B-modes over the full sky with the accuracy attainable only with a dedicated satellite. The comparatively modest investment in the current ground-based and balloon program will significantly mitigate the risk for a future space program. The panel strongly endorses continuation of funding by NASA, DOE, NIST, and NSF's Office of Polar Programs at least at current levels.

Allen Telescope Array Expansion (ATA-256)

The ATA plans to expand its current array from 42 elements to 256 elements (ATA-256) to provide roughly the collecting area of the GBT. Five key science surveys are planned to take advantage of the increase in sensitivity and resolution that the expansion will provide. In particular, the ATA-256 will be well equipped to time an array of pulsars in three to four different wave bands simultaneously, which is crucial for removing effects due to the interstellar medium, allowing the highest-precision measurements. This timing can also lead to valuable secondary science, such as measurements of neutron-star masses, tests of general relativity from relativistic binaries, and constraining the neutron star equation of state.

Currently, the ATA-42 is carrying out high-speed and wide-field-of-view surveys of the transient sky, but the ATA-256 will offer significantly improved sensitivity for transient science (Figure 9.13). The ATA-256 will be the only radio facility committed and dedicated to doing time-domain astronomy (GAN Discovery and SSE Discovery areas). In addition, transient searches will be able to be done commensally with pulsar timing, sensitive mapping of H I in the local universe, and SETI monitoring. Finally, the ATA-256 can play an important role as a precursor, large-*N* array for the SKA through lessons learned in operations, maintenance, signal processing, data management, and data archiving.

The ATA-256 consortium hopes to obtain funding for the expansion by early 2011, with full science operations to begin in late 2013. The ATA-256 consortium has estimated costs (based on the expenses incurred by building the ATA-42) to be $44 million over the decade for construction of the expanded array. After construction, it estimates $6 million per year for operations, maintenance, and survey science, of which up to 30 percent would come from private partner contributions. Independent cost appraisals were not available. The ATA-42 will carry out transition science during the construction and commissioning period as elements are added to the array at a lower operating cost of $3 million per year.

The ATA-256 has lower priority for mid-scale funding than HERA-II, CCAT, and FASR. If more money is available, or if HERA-I does not meet the milestones needed to proceed with HERA-II, ATA-256 would be the next priority after CCAT and FASR. Alternatively, expansion to 256 antennas could be treated as an upgrade, similar to those discussed below. The panel recommends that NSF explore partnerships with other agencies and private foundations to advance ATA-256.

FIGURE 9.13 ATA-42 at the Hat Creek, California, site. SOURCE: Seth Shostak/SETI Institute.

NRAO Development

An ongoing program of moderate, cost-effective, science-driven enhancements to NRAO telescopes can provide missing capabilities (see Table 9.1) that are essential to confront many of the key Science Frontiers Panel questions. The following non-prioritized developments will leverage the substantial capital invested in these telescopes and keep them on a path to new discoveries through the next decade and beyond.

EVLA

The panel supports a program of proposed enhancements to broaden greatly the EVLA capabilities: (1) E configuration, (2) Pie-Town link, (3) water vapor radiometers, and (4) long-wavelength-receiver system. The first three of these enhance the capability for efficient, high-fidelity imaging at centimeter wavelengths, while the fourth provides some meter-wave capability. The total cost is less than

$15 million plus $360,000 per year for operations. The technologies are mature, the timelines are short, and they carry little risk.

The E configuration ($8 million construction, $60,000 per year operations) will add 20 new antenna stations for higher surface-brightness sensitivity and better image fidelity at large angular scales, allowing new studies of H I in the galaxy and nearby galaxies (PSF 1; GAN 1, 2, 4), dense molecular cores and H II regions (PSF 1), and the SZE (GCT 1).

The Pie-Town link ($3 million construction, $80,000 per year operations) will improve the angular resolution by a factor of two, allowing studies of obscured and dense H II regions (GAN 2), protoplanetary disks (PSF 2), the mass-loss environment of massive stars, supernovae, and supernova remnants (SSE 2, 3), the AGN-starburst connection (GCT 3), and the structures, cores, and cusps of gravitational lenses (GCT 1; GAN 4).

The water-vapor radiometers ($1.2 million construction, $90,000 per year operations) will greatly enhance the phase stability at short wavelengths and long baselines, including the Pie-Town link, allowing studies of thermal emission from disks (PSF-2) and jets (GAN 1), the photospheres of supergiant stars (SSE 1), obscured H II regions and supernovae (GAN 2), and molecular lines at high redshift (GCT 3).

The long-wavelength receiver system ($1.8 million construction, $130,000 per year operations) will open up the observing window from 6- to 0.3-m wavelengths to allow studies of radio transients (GAN 5), magnetospheres of extrasolar planets (PSF 3), supernova-cloud interactions, (GAN 1, 2; SSE 2), steep-spectrum sources (GCT 3), radio relics and radio lobes (GCT 3), and atomic hydrogen and magnetic fields across cosmic time (GCT 1, 2). It is, however, not well suited to EoR observations.

GBT

The GBT combines large collecting area and good surface quality for sensitive, filled-aperture centimeter-wave observations (Figure 9.14). The panel supports the proposed, staged, array receiver and camera-development program. These upgrades will provide some of the needed capability for fast centimeter-wave surveys. The research and development costs are $31 million, and the production budget is $28 million, with $600,000 per year for operations. Much of the research, development, and production will be done at universities, supported by NSF grants. With sufficient funds for fundamental technologies, these new instruments can be realized in the next decade without high risk.

Three types of cameras are envisioned, including (1) a 100-pixel heterodyne camera for wavelengths of 2.6 to 4.3 mm and a 64-pixel heterodyne array for 10 to 17 cm, (2) a phased-antenna array for 20 cm, and (3) a 1,000-pixel bolometer

FIGURE 9.14 The Green Bank Telescope in West Virginia. SOURCE: NRAO/AUI.

array for 3 mm, as well as associated efforts in integration and packaging of receiver elements, high-speed analog-to-digital conversion, and data transmission. These new instruments will enhance searches for gravitational waves using millisecond pulsars (CFP Discovery), studies of atomic and molecular gas content and evolution and astrochemistry throughout the Milky Way (PSF 1) as well as in galaxies nearby (GAN 1, 2) and at cosmological distances (GCT 2, 3), the characterization of galaxy clusters through the SZE (GCT 1), and the statistics of stellar remnant spins from pulsar timing (SSE 4, 5).

VLBA

The VLBA is a unique facility that provides ultrahigh angular resolution at radio wavelengths and has proved powerful for micro-arcsecond astrometry. Modest upgrades will significantly expand this capability. The panel supports the proposed development package that includes (1) EVLA-style 4- to 8-GHz receivers, (2) wider data-acquisition bandwidths—to 4 GHz per polarization, and (3) water-vapor radiometers. The cost is about $16 million and carries little risk. These upgrades will significantly improve VLBA capabilities for astrometry (GAN Discovery) and will obtain accurate distances to large samples of star-forming regions throughout the galaxy, using 6.7-GHz methanol masers (PSF 1); measurements of Local Group motions to probe dark-matter content (GAN 4); and precision cosmology through a 1 percent determination of H_0 using megamasers (CFP 2).

NAA

Looking to the future, the North American Array (NAA) initiative will pave the way for the U.S. community to lead the development and prototyping of the SKA-high. This $40 million investment is divided roughly equally between (1) developing enabling technologies, such as low-cost antenna concepts, wideband receivers, and data-processing capabilities, and (2) the implementation of a prototype NAA antenna station, perhaps at a location in New Mexico, that leverages the existing EVLA/VLBA infrastructure. This development package, in concert with international and national efforts aimed at longer wavelengths, will lay the groundwork for the United States to propose to grow SKA-high from the NAA in the 2020 decade.

ALMA Development

ALMA will be the world's premiere facility for high-resolution imaging at millimeter/submillimeter wavelengths. It is essential that a program of upgrades be supported to maintain its vitality. While the upgrades will be determined by agreement of the international consortium, there are some obvious examples. Some receiver bands are not included in the first-light complement; adding these is important for obtaining complete wavelength coverage within atmospheric windows, which is especially needed for line studies of galaxies over the full range of redshifts. Adding capability to join the millimeter-wave VLBI network would greatly enhance the sensitivity for ultrahigh-resolution millimeter/submillimeter studies (see Table 9.1).

The consortium plans a program costing $90 million over a decade, of which $30 million would come from the North American partners. The panel fully supports this plan.

Small or Moderate Missions and Other Activities

Overview

A healthy RMS landscape will include small- and moderate-cost activities in addition to the major new initiatives and development of existing facilities. The panel supports in general a balanced program; a few particular activities that submitted responses to the survey's request for information are discussed here. The panel looked among these for those that satisfied needs that were unmet by the larger projects.

The remaining capabilities that are needed to address the SFP science questions are the following: improved sensitivity at millimeter wavelengths for ultrahigh resolution (micro-arcsecond); and still better resolution at centimeter wavelengths. In this section, the panel lays out a plan to sustain and develop current activities.

Event Horizon Telescope

The Event Horizon Telescope will outfit and combine millimeter/submillimeter telescopes worldwide to directly image the black hole event horizon of SgrA* and the nearest active galactic nucleus in M87. It will also measure the black hole spin and constrain accretion and jet-launching models. The project is separated into three phases of development. The primary construction components include outfitting a number of existing telescope facilities with the requisite 0.85- and 1.3-mm receivers and VLBI capability, along with higher-bandwidth backends and data recorders. Observations would happen through few-week coordinated campaigns several times a year. Many of the required EHT upgrades will augment the science capabilities of the host facilities beyond the EHT program.

The current model-dependent measurements give the size of SgrA* at 1.3 mm as 3.7 (+1.6/–1.0) R_{sch}, where R_{sch} = 10 micro-arcseconds. The EHT aims for a resolution approaching 10 micro-arcseconds with sufficient sensitivity to make detailed images of the event-horizon environment. This has unique capability to image the region around an event horizon and explore fundamental black hole properties and physics. These goals are crucial for science questions GAN 4, *What are the connections between dark and luminous matter?*, and GCT 3, *How do black holes grow, radiate, and influence their surroundings?* The costs come in three phases (Phase I: $15.5 million; Phase II: $20.1 million; and Phase III: $11.5 million, anticipated to begin in 2019). The panel supports Phase I, with a reassessment at mid-decade. Phase I, with seven antennas operating at 1.3 mm, is expected to allow imaging and characterization of the central shadow predicted by general relativity, caused by the orbiting optically thin plasma, and to measure the orbital periods of material orbiting the central black hole to constrain the spin.

FIGURE 9.15 CARMA with all three telescope sizes shown. SOURCE: John Carlstrom.

CARMA Development

Hoped-for expansion of CARMA's capabilities in the next decade includes a new, broadband, flexible, digital-correlator system, array receivers at 1 and 3 mm, and ultrawide-bandwidth receivers at 3 mm (Figure 9.15). The total cost of these development projects is \$16 million. These upgrades will yield fast mosaicing speeds with a unique combination of high angular resolution and good surface-brightness sensitivity. The fast speeds will enable surveys of statistically large samples of molecular clouds for studies of star formation from the solar neighborhood (PSF 1) to nearby galaxies (GAN 2). The ultrawide-bandwidth 3-mm receivers will probe a broad range of redshifts in the early universe; for example, for $z > 3$, at least one CO line is always present in the proposed ultrawide 3-mm band. The flexible correlator would also allow sensitive, high-angular-resolution observations of the SZE (GCT 1) at 1 cm with unprecedented angular dynamic range (from 0.05 to 5 arcminutes using the OVRO, BIMA, and SZA antennas together). Large surveys are essential to achieve all of these science goals.

By enabling large surveys, by providing essential access for the U.S. community to interferometric observations to bolster ALMA observing-time proposals in a highly competitive environment, and by providing a crucial hands-on training ground for U.S. astronomers and students, CARMA will remain complementary to ALMA in the next decade. With its smaller size, CARMA has the ability to prototype and test future technology upgrades efficiently.

SAMURAI

The SAMURAI (Science of AGNs and Masers with Unprecedented Resolution in Astronomical Imaging) project will substantially augment the scientific capabilities of the second-generation space VLBI station, VSOP-2, which is expected to be launched by the Japanese perhaps as early as 2013. VSOP-2 will observe at $\lambda= 7$ mm to 4 cm and achieve resolutions a factor of ~3 higher than ground-based VLBI. The U.S. project would focus on two science goals. The first is to determine the structure of radio emission from the accretion envelopes of black holes; at $\lambda = 7$ mm, SAMURAI will have a resolution of 40 mas, corresponding to 2 Schwarzschild diameters for Sgr A* and 7 for M87. The second goal is to provide precision astrometry of H_2O masers in AGN that can be used to estimate distances precisely enough to improve the accuracy of H_0 to 1 to 2 percent (CFP 3). SAMURAI and the VLBA provide complementary approaches to this critical measurement. The project requires construction of two ground-based tracking stations, use of existing U.S. stations, and other mission support. It also represents a contribution that will allow U.S. scientists access to the VSOP-2 program. In this sense the U.S. contribution is highly leveraged by Japan and the international community. The cost, $44 million, would be funded by NASA; thus the project falls outside this panel's purview. The panel can only comment that this would help to provide the ultrahigh-resolution capability listed in Table 9.1 as necessary to addressing RMS-related science questions identified by the SFPs.

The RMS System and Community

The RMS System

The RMS "system" is quite different from the OIR "system." NRAO operates national and international facilities that include only unique telescopes; older, smaller telescopes have been ruthlessly retired. A small number of University Radio Observatories provide complementary capabilities and hands-on training for students and postdoctoral scholars in both observing and instrumentation. Aside from the needed new capabilities discussed above, the major gaps in the RMS system are adequate funding for individual investigators and adequate support for the UROs.

Archives and User Support for RMS Facilities

With the data flow from ALMA, RMS science will enter a new era. The ALMA Science Archive will be the first full-service archive for RMS astronomy, and it will set the standard for such archives. Archival research, combining RMS data with data from other wavelength regions, will grow as the decade progresses. Full participation in Virtual Observatory protocols will be vital. NRAO has made some recent efforts to help its users in this area. Computing was chronically underfunded in the early years of the VLA, and the VLA/VLBA has yet to achieve a user-friendly archive. There are efforts underway to provide an improved aperture-synthesis reduction-and-mapping package, CASA, for both the EVLA and ALMA, as well as for URO synthesis telescopes, and to provide a more user-friendly image archive.

Unlike the situation at the NASA great observatories, observing time on RMS facilities is not generally accompanied by funding for analysis and publication. NRAO supports page charges and has instituted a program to provide stipend and travel support for students using NRAO telescopes. While these efforts help, they do not provide the same level of support as the NASA facilities do for their users. The NASA support of guest investigators on HST alone dwarfs all user support for RMS facilities. This situation diminishes exploitation of expensive resources and makes it difficult for the United States to compete on the international stage. The panel reiterates the recommendation by the 2001 decadal survey AANM that NSF provide funding support for U.S. observers on its telescopes, whether international (e.g., ALMA), national, or a URO.

Training of Students and Postdoctoral Scholars

The training of the next generation of RMS scientists and instrumentalists takes place both at national facilities, such as the NRAO and NAIC telescopes, and at university radio facilities. National facilities mostly provide data, but there are some opportunities for hands-on operations by students, particularly at Arecibo and the GBT. The Jansky Postdoctoral Fellowships are the only prize fellowships directed toward RMS research. ALMA funds postdoctoral fellowships worldwide (mainly at ESO, NRAO, and NAOJ). The VLA and the GBT have modest programs that can support thesis research by graduate students. The GBT supports some visitor instruments and the concomitant instrumental experience back at their home universities. Research experience for undergraduates programs exist at a number of RMS sites. In the past decade, about 160 graduate students and 100 postdoctoral fellows have learned their craft while observing at the UROs.

Support for the UROs

The UROs are chronically underfunded. The RMS Panel met with a group representing currently funded UROs, along with groups that would like to be funded now or in the future. It was clear from the discussion and from later information gathered that the UROs provide not only vital capabilities but also the fundamental training grounds for future RMS scientists. Some scientists trained at UROs have been instrumental in transferring technology and techniques to other wavelengths. The panel strongly supports enhancements to the URO budget of $2 million per year for existing facilities and suggests that the URO system provide a mechanism for operating new mid-scale facilities for which the panel recommended construction in this decade.

Technology Development

A strong and secure technology-development program is an essential component of a balanced RMS program. The panel recommends that the Advanced Technology and Instrumentation (ATI) program be enhanced by at least $1 million per year.

There are four major technology developments that can increase the speed of radio measurements by increasing the fields of view and the available bandwidths of both existing and future instruments.

- *Detector arrays.* Millimeter and submillimeter bolometric arrays are opening several new scientific frontiers. Larger and more sensitive cameras promise to continue this revolution. A diversity of approaches in the power-detection, RF-coupling, and readout technologies has been a strength and should be supported, along with the continued development of total-power coherent detector arrays. Ongoing support is needed to continue the growth in detector elements shown in Figure 9.16.
- *Wideband digital systems.* In the gigahertz range, advances in digital processing enable replacing analog RF mixers and filters with higher-performance digital complements, potentially decreasing costs and improving stability. At higher frequencies, the advent of 80-GHz samplers will lead to an increase in the available continuum and spectral bandwidth.
- *Large-N correlators.* Several science goals require correlating thousands of inputs. This is especially important for observations of H I in high-redshift galaxies and epoch of reionization measurements, where the development of very-large-N correlators is the primary technical hurdle.
- *Phased-array feeds.* Fully sampling the electric field across focal planes could significantly enhance the survey speed of both single-dish and interferometric instruments.

The panel encourages, whenever possible, including astronomical observations as part of technology assessment. Scientific observations with new instruments can significantly enhance the careers of students and postdoctoral associates, helping to replenish the pool of those with knowledge of state-of-the-art instrumentation.

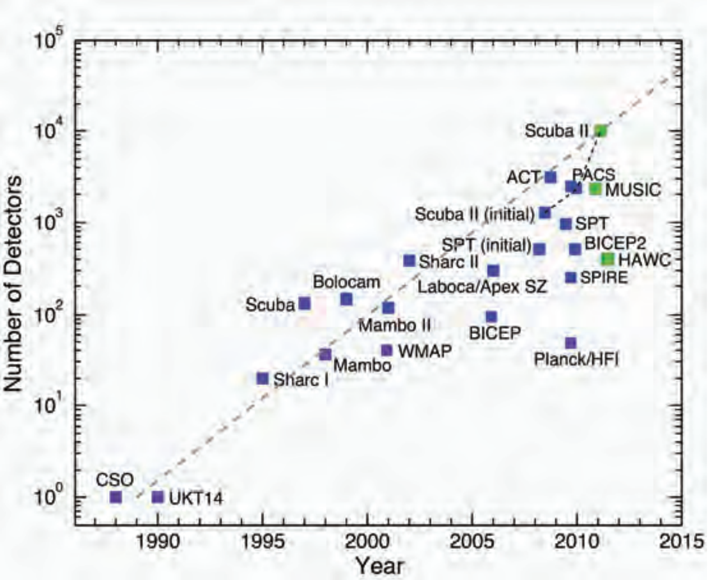


FIGURE 9.16 “Moore’s law” for submillimeter detectors showing a doubling time for a number of elements of about 20 months (dashed line). SOURCE: Jonas Zmuidzinas, California Institute of Technology.

Laboratory Astrophysics

A wide range of new RMS facilities that are currently operating (e.g., spectrometers on the Herschel Space Observatory) or coming on line in the next decade will provide unprecedented access to diagnostic astrophysical spectral lines at sensitivities that may be 10 to 100 times better than those of current instruments (Figure 9.17). Even with current sensitivities, 30 to 50 percent of lines remain unidentified. To interpret anticipated data, improved information on frequencies, collision rates, and chemical reaction rates is needed. Key measurements include rest-frequency assignments for all known molecules (including vibrationally excited states and isotopologues); predicted spectral-line intensities and their temperature dependence; collisional excitation rates; isotopic-fractionation, photolysis, and dissociative-recombination branching ratios; binding and diffusion energies for molecules on ice surfaces; and activation energies for reactions. Some of these parameters (notably collisional excitation rates and potential surfaces for molecules) rely primarily on theoretical calculations.

Historically, many of the advances in laboratory astrophysics have been funded outside typical astronomy funding lines, for example in NSF’s physics and chemistry divisions. Unfortunately, these divisions are understandably reluctant to divert significant funds for activities that are directed primarily at astronomy. NASA has

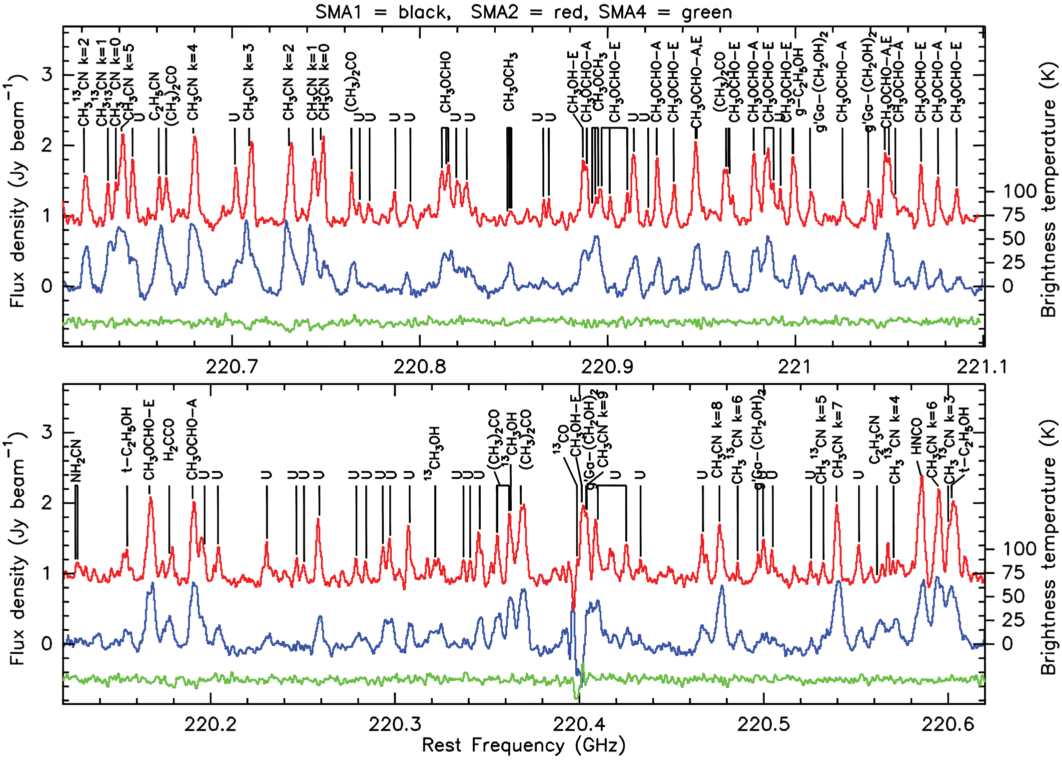


FIGURE 9.17 Submillimeter Array 1.3-mm spectra from three dust cores (each plotted in a different color offset in the y-axis for clarity) in the massive star-forming region NGC6334I (sources are separated by only 2 arcseconds or about 4,000 AU) show tremendous chemical complexity over small size scales as well as a large fraction of unidentified lines. SOURCE: Todd R. Hunter, National Radio Astronomy Observatory.

provided some directed funding for laboratory astrophysics in support of a few missions, for example the Herschel Space Observatory, but that funding has been limited and of relatively short duration. In the coming decade, it is critically important that the astronomical community find a way to provide sustained support for laboratory astrophysics, potentially through cross-divisional NSF funding initiatives or directed funding lines within the astronomy division. The full scientific potential of the next generation of RMS facilities will be severely compromised without a commensurate dedication of resources to laboratory astrophysics. The panel recommends a program funded at $2 million per year.

Theory for RMS-Related Science

The panel lists below in two categories some theoretical needs that are especially important for RMS science.

The theoretical developments *required* for extracting science from RMS observations are the following:

- Theoretical calculations of fluctuations in the intensity and polarization of the cosmic microwave background are an inherent part of deriving cosmological constraints from the CMB data.
- Numerical simulations of cosmic reionization are needed to predict observables for redshifted 21-cm experiments (especially for the first stage, HERA-I). Given the potential complexity of foreground removal and accurate calibration, without such modeling the observational measurements of statistical properties of the cosmological signal (such as the power spectrum) will remain inconclusive until actual imaging capabilities are developed.
- Full three-dimensional modeling of circumgalactic environments is crucial for the correct interpretation of the spectroscopic data. Without such modeling, differentiation between hot-gas accretion, outflows, virial shocks, and gas flows along filaments is extremely difficult or impossible.
- Chemo-dynamical models of star-forming regions that trace the simultaneous evolution of density, gas and dust temperature, and molecular abundances through the evolution of a particular dynamical model are necessary for a correct interpretation of molecular-line observations.
- Spectropolarimetric and magnetohydrodynamic forward modeling of the solar atmosphere in three dimensions will be essential to exploit the full potential of FASR data.
- Interference mitigation will be critical, especially for low-frequency observations.

In several other areas, major theoretical development will be required to fully realize the investment in an RMS facility:

- Modeling cosmological structure formation in representative cosmological volumes with resolution adequate to include the most of the important physical processes in the ISM.
- Modeling black hole accretion on a wide range of scales, from the inner edges of accretion disks to global galactic environments.
- Theoretical studies and numerical modeling of MHD processes on a wide range of scales, from the small-scale physics of reconnection and particle acceleration to the effects of magnetic fields on galactic and extragalactic scales.

- Numerical modeling of radiative transfer in galactic simulations, in simulations of star-forming regions, and in models of black-hole environments.
- Precision modeling of stellar pulsations.
- Theoretical studies and numerical modeling of planet formation, evolution, and dynamics in a range of environments.
- Modeling of radiative and dynamical processes in planetary atmospheres.
- Modeling of gravitational-wave signatures expected from stochastic backgrounds and from individual sources which could be detected by pulsar-timing arrays.

Algorithm Development

Many of the scientific goals for the next decade rely on the development of new algorithms that will facilitate the processing of large datasets, allow the discovery of weak signals, and foster cross-disciplinary data sharing. The most critical needs are as follows:

- Foreground removal for epoch of reionization studies.
- Optimal detection and characterization of gravitational-wave signals in pulsar data.
- Computationally efficient pulsar-search algorithms to handle large amounts of data with sensitivity to the most relativistic binary systems and weak transients. Real-time search algorithms are imperative given the expected data rates of new correlators and multi-pixel receivers.
- Spectral-line-analysis tools that aid in the identification of lines, extraction of line parameters, and analyses of physical conditions using laboratory astrophysics results and radiative-transfer algorithms.
- Automated source-detection algorithms in the spectral domain to aid in the creation of source catalogs in formats compatible with Virtual Observatory standards.
- Imaging algorithms that keep pace with the cutting edge of possible data rates through parallelization and high-performance computing possibilities.

A crucial need for RMS is greater access to high-speed data transmission for data acquisition and retrieval as well as access to long-term storage. Many of the proposed facilities can produce data at rates approaching a petabyte an hour. Innovative solutions for storage and public access, in keeping with observatory policies, are necessary. Strong partnerships between the NSF-AST division and the NSF Office of Cyberinfrastructure, as well as international agreements, will facilitate these goals.

Spectrum Management

The radio spectrum is a precious resource for radio astronomy and for communication in the modern world. Without continued vigilance to protect some of this resource for passive scientific use, radio astronomy from Earth's surface will become increasingly difficult. The traditional approach of seeking protection for small, defined segments of the spectrum for radio astronomy is no longer adequate because of the wide bandwidths needed for sensitivity and the broad frequency coverage needed for spectral-line studies at high redshifts. Resources must be made available to develop modern technologies for radio-interference mitigation and for sharing the spectrum through time- and frequency-multiplexing methods.

Looking to the Future

The Square Kilometer Array is seen as the next-generation centimeter-wave and meter-wave telescope, which would address many fundamental science questions. This project has already garnered significant international support, with 55 institutions in 19 countries participating. The 2001 decadal survey recommended that the U.S. participate in a program of technology development funded at $22 million; $12 million has been funded, starting in 2007.

Over the past decade, the SKA has evolved into three instruments covering three wavelength regimes: SKA-low (1 to 3 m); SKA-mid (3 to 100 cm); and SKA-high (0.6 to 3.0 cm). The panel believes that it is very important for the United States to play a role in this international project. However, based on the information received from the projects and from independent analysis, none of the parts of this project have reached maturity sufficient to recommend construction at this time. Defining the way forward in this context requires a mix of technology development, demonstrator projects (e.g., LWA, MWA, PAPER, LOFAR, MeerKAT, ASKAP), and careful consideration of priorities. The results of the demonstrators will not be available for a number of years.

The long-wavelength (1.0 to 3 m) part of the spectrum covered by SKA-low provides the only way to study the process of reionization (H I at $z = 5$ to 14); through that, it is one of the most promising ways to study the first luminous objects (GCT 4). The HERA activity provides a step-by-step path for the U.S. community to lead in this exciting aspect of SKA science. It lays out a sequence of intermediate science and associated technology goals that address this key science area.

The mid-wavelength (3 to 100 cm) part of SKA provides the capability for very sensitive centimeter-wave imaging. SKA-mid is essential to study the role of atomic gas in galaxy evolution (GCT 2; GAN 1); it could provide spectroscopic imaging of the H I emission for a billion galaxies out to $z \sim 1$. This cannot be done with present facilities and is strong justification in itself for this ambitious instrument. SKA-mid could address other key scientific questions as well: it would enable a powerful pulsar machine that would enable a census of the galactic pulsar population (SSE 4), test general relativity in the strong-field regime (CFP 2), and almost certainly detect low-frequency gravitational waves and constrain gravitational-wave sources (CFP 1, Discovery). It would provide excellent sensitivity to the transient radio sky (SSE Discovery).

In spite of the compelling science case for it, the panel finds that there are substantial issues of technical readiness and cost for SKA-mid. The total construction cost of the project, already $2.2 billion in the project's estimate, was raised to $5.9 billion by independent analysis. SKA-mid was considered not technologically ready in the independent analysis, and the panel concurs. Further development and study of alternative options for this wavelength range are needed and could be funded in open competition within the ATI program, potentially in conjunction with other international efforts. Pathfinders, such as ATA-256, could help to test technical concepts that would lead to the final design of the SKA. Alternative approaches to constraining dark energy, such as 21-cm intensity mapping, could be explored to see if they can be useful on shorter timescales and at lower cost. The panel recommends revisiting the SKA design costs in 5 years to assess end-of-decade feasibility.

The short-wavelength (0.6 to 3.0 cm) part of SKA helps constrain dark energy (CFP 2), dark matter (CFP 3), galaxy evolution via CO and other molecules at z above 1.3 (GCT 2), planet formation (PSF 2), and the ends of massive stars (SSE 3). Because of the U.S. heritage with EVLA, GBT, VLBA, and ALMA, it is natural for the United States to build on these in developing this part of SKA. A modest program of technology development and prototyping should begin in this decade. The NAA activity discussed above provides an attractive way to proceed.

RECOMMENDATIONS

The panel recommends a program with three new major initiatives for mid-scale funding, upgrades to existing and imminent facilities, and increased funding for smaller facilities. The panel identifies a need for technology development in four main areas and an interdisciplinary laboratory astrophysics program, along with theory and algorithm development relevant to RMS science. The panel's recommendations are made in the context of the following assumptions: a 7 percent per year increase in the NSF-AST budget and the augmentation of a funding line for mid-scale construction projects of at least $20 million per year. The panel recommends no projects for MREFC funding.

Major New Initiatives

The major new initiatives are HERA, CCAT, and FASR. In terms of scientific importance, the panel ranks HERA first, with CCAT and FASR tied and close behind. However, there are issues of readiness that require a distinction between ranking and phasing of project starts.

Of these, FASR is the most ready to proceed and, as the panel has noted, has excellent prospects for cross-directorate funding. The panel recommends that a funding strategy for FASR be developed with core contributions from both NSF-AST and NSF-AGS (the panel assumes here an even split). The FASR Pathfinder proposed by the project to AGS ($8 million) would be a good way to begin and resolve any remaining cost questions, but it should proceed in a manner that is ultimately compatible with the full implementation of FASR. The $50 million from AST would come from the mid-scale line starting in 2010 and ending in 2015. The 50 percent AST share of operations ($2 million per year) could come from an increase in the URO budget.

CCAT is also far along in design. The project, a consortium of U.S. and foreign institutions, estimated a total cost of $110 million, and the independent estimate was $138 million for construction. Of this, the consortium requested only $33 million from NSF. The panel recommends proceeding with this project as soon as the design is finalized and the consortium has the balance of the funding or suitable guarantees. There is substantial urgency in this project, because it will provide sources to optimize use of ALMA. CCAT would be an ideal candidate for funding by a mid-scale funding source when it becomes available, but it should start by 2012. It should phase in after FASR and complete funding by 2017. Operations costs from NSF will be $7.5 million per year, but shutting the CSO will save $2.5 million per year, and so $5 million per year extra will be needed in the URO budget by 2018 (see below).

HERA has the panel's top science ranking, but it comes in three phases. The first phase, HERA-I, is underway with two parallel efforts engaged in testing techniques. The panel strongly recommends continued funding for both these efforts at a combined rate of approximately $5 million per year to about 2015, at which time a review would be needed. If the HERA-I projects achieve certain milestones, the panel would strongly favor funding of HERA-II, or a similar project selected in open competition, from the mid-scale funding line. The milestones are demonstration of successful techniques for calibration and foreground removal; detection of the power spectrum of H I in the epoch of reionization; a decision on the optimum design for HERA-II; and development of a full proposal for HERA-II with credible costs. The current estimates of cost for a mid-decade start for HERA-II range from $80 million to $115 million, but they depend strongly on future correlator developments. There should be further technology development toward HERA-III, which would be a future-decade project. For HERA-I, the costs are a continuation of current funding, and so the panel counts only HERA-II as new funding and assumes a total cost of $85 million from the mid-scale line, starting in 2015 and ending in 2020. The panel assumes no operations funding this decade.

Sustaining and Upgrading Current and Imminent Facilities

The panel's top priority in the facilities area is continued funding of a vigorous, diverse program of ground-based research on the cosmic microwave background. Detection of the so-called B-modes, which trace primordial gravitational waves, is a primary goal. The program should also constrain cosmological parameters, determine or limit the sum of neutrino masses, and measure large-scale structure. Since this is an ongoing program, it does not represent new costs, but the panel emphasizes its absolute importance.

The current ATA-42 is in serious need of funding. This project has pioneered the concept of large arrays of inexpensive antennas with broadband imaging response. Its current NSF support is inadequate to keep the current array running, much less to continue the technological evolution of the array to 256 antennas. The observing capabilities provided by the ATA-256 would provide major advances in the ability to find transients and detect gravitational waves. Moreover, this project can provide valuable technological developments for a mid-range SKA, in the area of wideband feeds, large-field imaging, and large-correlator development. The panel could not fit the funding for ATA-256 into a $20 million per year mid-scale line, but it would be the preferred back-up for such funding should HERA-I not meet its milestones. ATA-256 may be able to attract further funding by private foundations or other agencies. The panel recommends an effort to explore ways to move forward with a modest investment of NSF funds. This is the second priority in this category.

The panel recommends a regular program of upgrades for existing facilities, including ALMA, NRAO facilities, Arecibo, and elements of the UROs program. These upgrades will provide some of the capabilities identified as needed to answer the science questions. In particular, such upgrades are the most cost-effective way to obtain the capability for efficient high-resolution imaging at centimeter wavelengths and improved sensitivity with ultrahigh resolution (see Table 9.1). Multifeed arrays on the GBT and CARMA provide test beds for new techniques. Convincing cases were made for a total of $90 million over the decade for each of ALMA (the U.S. share is $30 million) and NRAO. The current UROs need more operating funds to improve their utility to the larger community (the panel recommends an extra $2 million per year) and further increases (ending the decade with an increment of $9 million per year) once FASR and CCAT become operational.

Because of the importance of Arecibo for pulsar timing, the panel recommends restoring $2 million per year in funding to its baseline budget.

Small Projects

Keeping a balance between large projects and national/international facilities and smaller projects is vitally important. Examples of excellent projects of this kind are the enhancement of millimeter-wave VLBI to create the Event Horizon Telescope, adding the huge collecting area of ALMA, and the addition of multifeed receivers to the CARMA telescopes. The panel recommends a total of $25 million for this effort over the decade, most likely funded by ATI or MRI.

The RMS System

Funding for user support and archive exploitation is important to the operation of the RMS system. The panel also recommends enhancements to the ATI program (by $1 million per year) that will allow more technology development for the future. The panel also recommends a program of laboratory astrophysics ($2 million per year) in which similar programs can be evaluated in context. RMS-related science depends heavily on laboratory and theoretical advances. Strategic theory and algorithm development should be supported to maximize the return on investments in facilities.

Summary

With a combination of new facilities and the sustenance and invigoration of existing programs and facilities, almost all the RMS capabilities needed to answer the science questions posed by the Astro2010 Science Frontiers Panels can be realized (Table 9.4, Figure 9.18). The most notable exception is the capability for very sensitive centimeter-wave imaging, needed for the study of H I at redshifts of 1 to 2. That requires something like SKA-mid, and the panel recommends some steps toward that goal. The panel summarizes in Table 9.5 the *additional* costs to NSF-AST for construction and operations. Table 9.5 indicates which items would be suitable for mid-scale funding and prioritizes projects costing at least $30 million.

TABLE 9.4 Needed RMS Capabilities and the Panel's Recommended Actions

Capability	Recommended Action
Cosmic microwave background program	Continue successful program
Sensitive meter-wave array	Continue HERA-I, fund HERA-II mid-decade
Solar radio telescope	Construct FASR
Fast millimeter/submillimeter surveys	Participate in construction, operations of CCAT
Fast centimeter surveys	Enhance ATA, GBT, Arecibo
Efficient high-resolution imaging at centimeter/millimeter	Enhance EVLA, ALMA, CARMA
Very sensitive centimeter imaging	Cannot meet this decade, technology development
Dedicated pulsar timing, transients	Enhance ATA, Arecibo, GBT
Ultrahigh resolution	Enhance VLBA, millimeter-wave VLBI
Complete wavelength coverage	Enhance ALMA

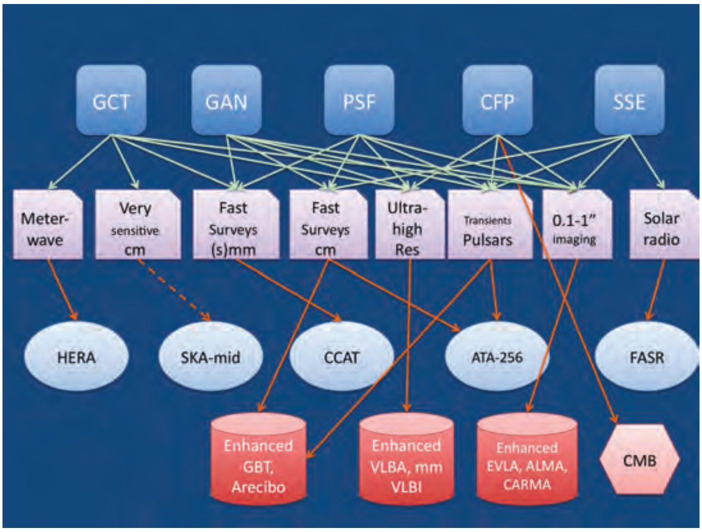


FIGURE 9.18 Mapping of required capabilities to new initiatives, upgrades of existing facilities, and continuation of successful programs. Dashed arrow indicates that need cannot be met this decade.

TABLE 9.5 *Added* Costs to NSF-AST for Construction and Operations (FY2009 million dollars)

Action	Construction (total)	Priority	Operations (per year)
Continue HERA-I, construct HERA-II	85[a]	1	0
Construct FASR (50 percent AST)	50[a]	2, tie	2, starting 2015
Participate in construction, operations of CCAT	33[a]	2, tie	5, starting 2017
Enhance ATA if possible	44[b]	4	3, increasing to 6 in 2015
Enhance GBT, EVLA, VLBA	90[c]	5, tie	1
Enhance ALMA	30[c]	5, tie	
Enhance CARMA, EHT	25[c]		
Enhance UROs support			2
Enhance Arecibo support, if possible			2, starting in 2012
Enhance ATI			1
Laboratory astrophysics			2
Total over decade	357		131

[a] Mid-scale funding.

[b] Mid-scale or other funds for upgrades.

[c] Some elements could be mid-scale instruments; others could be MRI or ATI.

Appendixes

A

Statements of Task for the Astro2010 Panels

COSMOLOGY AND FUNDAMENTAL PHYSICS STATEMENT OF TASK

The Cosmology and Fundamental Physics (CFP) Panel will identify and articulate the scientific themes that will define the frontier in CFP research in the 2010-2020 decade. Its scope will encompass cosmology and fundamental physics, including the early universe, the microwave background, the reionization and galaxy formation up to virialization of protogalaxies, large scale structure, the intergalactic medium, the determination of cosmological parameters, dark matter, dark energy, tests of gravity, astronomically determined physical constants, and high energy physics using astronomical messengers. Its assessment will play a key role in the Astronomy and Astrophysics 2010 study, which will survey the field of space- and ground-based astronomy and astrophysics, recommending priorities for the most important scientific and technical activities of the decade 2010-2020. The CFP Panel will prepare a report that will identify the scientific drivers of the field and the most promising opportunities for progress in research in the next decade, taking into consideration those areas where the technical means and the theoretical foundations are in place for major steps forward.

More broadly, this panel will be charged (as will each of the five science panels) with the following tasks:

1. Identify new scientific opportunities and compelling scientific themes that have arisen from recent advances and accomplishments in astronomy and astrophysics;

2. Describe the scientific context of the importance of these opportunities, including connections to other parts of astronomy and astrophysics and, where appropriate, to the advancement of our broader scientific understanding;
3. Describe the key advances in observation and theory necessary to realize the scientific opportunities within the decade 2010-2020; and
4. Considering the relative compelling nature of the opportunities identified and the expected accessibility of the measurement regimes required, call out up to four central questions that are ripe for answering and one general area where there is unusual discovery potential and that define the scientific frontier of the next decade in the SFP's sub-field of astronomy and astrophysics.

In completing this task, each Science Frontiers Panel will provide the Astronomy and Astrophysics 2010 Committee's Subcommittee on Science with its inputs in the Spring of 2009 and complete its panel report thereafter. The panel reports will be published following the release of the main survey committee's report in 2010. The Subcommittee on Science will issue a request for community input to ensure broad community participation in the process of identifying the scientific frontiers.

GALACTIC NEIGHBORHOOD STATEMENT OF TASK

The Galactic Neighborhood (GAN) Panel will identify and articulate the scientific themes that will define the frontier in GAN research in the 2010-2020 decade. Its scope will encompass the galactic neighborhood, including the structure and properties of the Milky Way and nearby galaxies, and their stellar populations and evolution, as well as interstellar media and star clusters. Its assessment will play a key role in the Astronomy and Astrophysics 2010 study, which will survey the field of space- and ground-based astronomy and astrophysics, recommending priorities for the most important scientific and technical activities of the decade 2010-2020. The GAN Panel will prepare a report that will identify the scientific drivers of the field and the most promising opportunities for progress in research in the next decade, taking into consideration those areas where the technical means and the theoretical foundations are in place for major steps forward.

More broadly, this panel will be charged (as will each of the five science panels) with the following tasks:

1. Identify new scientific opportunities and compelling scientific themes that have arisen from recent advances and accomplishments in astronomy and astrophysics;
2. Describe the scientific context of the importance of these opportunities, including connections to other parts of astronomy and astrophysics and, where appropriate, to the advancement of our broader scientific understanding;

3. Describe the key advances in observation and theory necessary to realize the scientific opportunities within the decade 2010-2020; and
4. Considering the relative compelling nature of the opportunities identified and the expected accessibility of the measurement regimes required, call out up to four central questions that are ripe for answering and one general area where there is unusual discovery potential and that define the scientific frontier of the next decade in the SFP's sub-field of astronomy and astrophysics.

In completing this task, each Science Frontiers Panel will provide the Astronomy and Astrophysics 2010 Committee's Subcommittee on Science with its inputs in the Spring of 2009 and complete its panel report thereafter. The panel reports will be published following the release of the main survey committee's report in 2010. The Subcommittee on Science will issue a request for community input to ensure broad community participation in the process of identifying the scientific frontiers.

GALAXIES ACROSS COSMIC TIME STATEMENT OF TASK

The Galaxies across Cosmic Time (GCT) Panel will identify and articulate the scientific themes that will define the frontier in GCT research in the 2010-2020 decade. Its scope will encompass galaxies across cosmic time, including the formation, evolution, and global properties of galaxies and galaxy clusters, as well as active galactic nuclei and QSOs, mergers, star formation rate, gas accretion, and supermassive black holes. Its assessment will play a key role in the Astronomy and Astrophysics 2010 study, which will survey the field of space- and ground-based astronomy and astrophysics, recommending priorities for the most important scientific and technical activities of the decade 2010-2020. The GCT Panel will prepare a report that will identify the scientific drivers of the field and the most promising opportunities for progress in research in the next decade, taking into consideration those areas where the technical means and the theoretical foundations are in place for major steps forward.

More broadly, this panel will be charged (as will each of the five science panels) with the following tasks:

1. Identify new scientific opportunities and compelling scientific themes that have arisen from recent advances and accomplishments in astronomy and astrophysics;
2. Describe the scientific context of the importance of these opportunities, including connections to other parts of astronomy and astrophysics and, where appropriate, to the advancement of our broader scientific understanding;
3. Describe the key advances in observation and theory necessary to realize the scientific opportunities within the decade 2010-2020; and

4. Considering the relative compelling nature of the opportunities identified and the expected accessibility of the measurement regimes required, call out up to four central questions that are ripe for answering and one general area where there is unusual discovery potential and that define the scientific frontier of the next decade in the SFP's sub-field of astronomy and astrophysics.

In completing this task, each Science Frontiers Panel will provide the Astronomy and Astrophysics 2010 Committee's Subcommittee on Science with its inputs in the Spring of 2009 and complete its panel report thereafter. The panel reports will be published following the release of the main survey committee's report in 2010. The Subcommittee on Science will issue a request for community input to ensure broad community participation in the process of identifying the scientific frontiers.

PLANETARY SYSTEMS AND STAR FORMATION STATEMENT OF TASK

The Planetary Systems and Star Formation (PSF) Panel will identify and articulate the scientific themes that will define the frontier in PSF research in the 2010-2020 decade. This panel will consider science opportunities and themes surrounding planetary systems and star formation, including solar system bodies (other than the Sun) and extrasolar planets, debris disks, exobiology, the formation of individual stars, protostellar and protoplanetary disks, molecular clouds and the cold ISM, dust, and astrochemistry. Its assessment will play a key role in the Astronomy and Astrophysics 2010 (Astro2010) study, which will survey the field of space- and ground-based astronomy and astrophysics, recommending priorities for the most important scientific and technical activities of the decade 2010-2020. The PSF Panel will prepare a report that will identify the scientific drivers of the field and the most promising opportunities for progress in research in the next decade, taking into consideration those areas where the technical means and the theoretical foundations are in place for major steps forward.

More broadly, this panel will be charged (as will each of the five science panels) with the following tasks:

1. Identify new scientific opportunities and compelling scientific themes that have arisen from recent advances and accomplishments in astronomy and astrophysics;
2. Describe the scientific context of the importance of these opportunities, including connections to other parts of astronomy and astrophysics and, where appropriate, to the advancement of our broader scientific understanding;
3. Describe the key advances in observation and theory necessary to realize the scientific opportunities within the decade 2010-2020; and
4. Considering the relative compelling nature of the opportunities identified

and the expected accessibility of the measurement regimes required, call out up to four central questions that are ripe for answering and one general area where there is unusual discovery potential and that define the scientific frontier of the next decade in the SFP's sub-field of astronomy and astrophysics.

In completing this task, each Science Frontiers Panel will provide the Astronomy and Astrophysics 2010 Committee's Subcommittee on Science with its inputs in the Spring of 2009 and complete its panel report thereafter. The panel reports will be published following the release of the main survey committee's report in 2010. The Subcommittee on Science will issue a request for community input to ensure broad community participation in the process of identifying the scientific frontiers.

STARS AND STELLAR EVOLUTION STATEMENT OF TASK

The Stars and Stellar Evolution (SSE) Panel will identify and articulate the scientific opportunities and themes that will define the frontier in SSE research in the 2010-2020 decade. Its scope will encompass stars and stellar evolution, including the Sun as a star, stellar astrophysics, the structure and evolution of single and multiple stars, compact objects, supernovae, gamma-ray bursts, solar neutrinos, and extreme physics on stellar scales. Its assessment will play a key role in the Astronomy and Astrophysics 2010 study, which will survey the field of space- and ground-based astronomy and astrophysics, recommending priorities for the most important scientific and technical activities of the decade 2010-2020. The SSE Panel will prepare a report that will identify the scientific drivers of the field and the most promising opportunities for progress in research in the next decade, taking into consideration those areas where the technical means and the theoretical foundations are in place for major steps forward.

More broadly, this panel will be charged (as will each of the five science panels) with the following tasks:

1. Identify new scientific opportunities and compelling scientific themes that have arisen from recent advances and accomplishments in astronomy and astrophysics;
2. Describe the scientific context of the importance of these opportunities, including connections to other parts of astronomy and astrophysics and, where appropriate, to the advancement of our broader scientific understanding;
3. Describe the key advances in observation and theory necessary to realize the scientific opportunities within the decade 2010-2020; and
4. Considering the relative compelling nature of the opportunities identified and the expected accessibility of the measurement regimes required, call out up to four central questions that are ripe for answering and one general area where

there is unusual discovery potential and that define the scientific frontier of the next decade in the SFP's sub-field of astronomy and astrophysics.

In completing this task, each Science Frontiers Panel will provide the Astronomy and Astrophysics 2010 Committee's Subcommittee on Science with its inputs in the Spring of 2009 and complete its panel report thereafter. The panel reports will be published following the release of the main survey committee's report in 2010. The Subcommittee on Science will issue a request for community input to ensure broad community participation in the process of identifying the scientific frontiers.

ELECTROMAGNETIC OBSERVATIONS FROM SPACE STATEMENT OF TASK

The Astronomy and Astrophysics 2010 Program Prioritization Panel on Electromagnetic Observations from Space (EOS) will identify and recommend a prioritized program of federal investment in research activities that involve space-based observations of astrophysical phenomena primarily by means of electromagnetic radiation. In formulating its conclusions, the EOS panel will draw on several sources of information: (1) the science forefronts identified by the Astro2010 science frontiers panels, (2) input from the proponents of research activities, and (3) independent cost and technical readiness assessments. The EOS panel's recommendations will be integrated into a program for all of astronomy and astrophysics by the Astro2010 Committee.

In particular the Astro2010 Programmatic Prioritization Panel on Electromagnetic Observations from Space will:

1. Report on the status of existing EOS research activities to set the context for future research activities, incorporating findings of the Study Groups.
2. Preview and compare proposed EOS research activities including those carried forward from previous surveys that have not been given a formal construction start.
3. State the relative importance of (a) smaller projects and generic research programs that involve competitive peer review and (b) programs that leverage public and private infrastructure investments, where appropriate.
4. Assess and describe best available estimates of the construction costs and lifetimes for each recommended research activity together with their full running costs (operations, science, and upgrades).
5. Identify particular risks for each research activity that would adversely affect the projected cost, technical readiness, or schedule of the activity. Identify those factors that could change an activity's priority and/or scope.
6. Informed by (a) the recommendations of the science frontier panels and (b) the panel's own research activity assessments, recommend a prioritized, bal-

anced, and integrated research program which includes a rank ordering of research activities and a balanced technology development program. A preliminary recommended program will be used to identify activities that will be subject to an independent technical evaluation and cost estimate. The panel's final recommendation to the Survey Committee will include consideration of the results of the independent technical evaluation and cost estimate.

In completing this task, each PPP will provide the Astronomy and Astrophysics 2010 Committee's Subcommittee on Programs with an interim internal and confidential summary report of its recommended program in the Fall of 2009 and complete its panel report thereafter. The panel reports will be published following the release of the survey committee's report in 2010.

OPTICAL AND INFRARED ASTRONOMY FROM THE GROUND STATEMENT OF TASK

The Astronomy and Astrophysics 2010 Program Prioritization Panel (PPP) on Optical and Infrared Astronomy from the Ground (OIR) will identify and recommend a prioritized program of federal investment in research activities that involve observations of astrophysical phenomena primarily by means of optical and infrared measurements from the ground. In formulating its conclusions, the OIR panel will draw on several sources of information: (1) the science forefronts identified by the Astro2010 science frontiers panels, (2) input from the proponents of research activities, and (3) independent cost and technical readiness assessments. The OIR panel's recommendations will be integrated into a program for all of astronomy and astrophysics by the Astro2010 Committee.

In particular, the Astronomy and Astrophysics 2010 Program Prioritization Panel on Optical and Infrared Astronomy from the Ground will:

1. Report on the status of existing OIR research activities from the ground to set the context for future research activities, incorporating findings of the Study Groups.
2. Preview and compare proposed OIR research activities from the ground including those carried forward from previous surveys that have not been given a formal construction start.
3. State the relative importance of (a) smaller projects and generic research programs that involve competitive peer review and (b) programs that leverage public and private infrastructure investments, where appropriate.
4. Assess and describe best available estimates of the construction costs and lifetimes for each recommended research activity together with their full running costs (operations, science, and upgrades).

5. Identify particular risks for each research activity that would adversely affect the projected cost, technical readiness, or schedule of the activity. Identify those factors that could change an activity's priority and/or scope.

6. Informed by (a) the recommendations of the science frontier panels and (b) the panel's own research activity assessments, recommend a prioritized, balanced, and integrated OIR research program which includes a rank ordering of research activities and a balanced technology development program. A preliminary recommended program will be used to identify activities that will be subject to an independent technical evaluation and cost estimate. The panel's final recommendation to the Survey Committee will include consideration of the results of the independent technical evaluation and cost estimate.

In completing this task, each PPP will provide the Astronomy and Astrophysics 2010 Committee's Subcommittee on Programs with an interim internal and confidential summary report of its recommended program in the Fall of 2009 and complete its panel report thereafter. The panel reports will be published following the release of the survey committee's report in 2010.

PARTICLE ASTROPHYSICS AND GRAVITATION STATEMENT OF TASK

The Astronomy and Astrophysics 2010 Program Prioritization Panel on Particle Astrophysics and Gravitation (PAG) will identify and recommend a prioritized program of federal investment in research activities exploring areas at the interface of physics and astronomy such as gravitational radiation, TeV gamma-ray astronomy, and free-flying space missions testing fundamental gravitational physics. In formulating its conclusions, the PAG panel will draw on several sources of information: (1) the science forefronts identified by the Astro2010 science frontiers panels, (2) input from the proponents of research activities, and (3) independent cost and technical readiness assessments. The PAG panel's recommendations will be integrated into a program for all of astronomy and astrophysics by the Astro2010 Committee.

In particular, Astronomy and Astrophysics 2010 Program Prioritization Panel on Particle Astrophysics and Gravitation will:

1. Report on the status of existing PAG research activities to set the context for future research activities, incorporating findings of the Study Groups.

2. Preview and compare proposed PAG research activities including those carried forward from previous surveys that have not been given a formal construction start.

3. State the relative importance of (a) smaller projects and generic research programs that involve competitive peer review and (b) programs that leverage public and private infrastructure investments, where appropriate.

4. Assess and describe best available estimates of the construction costs and lifetimes for each recommended research activity together with their full running costs (operations, science, and upgrades).

5. Identify particular risks for each research activity that would adversely affect the projected cost, technical readiness, or schedule of the activity. Identify those factors that could change an activity's priority and/or scope.

6. Informed by (a) the recommendations of the science frontier panels and (b) the panel's own research activity assessments, recommend a prioritized, balanced, and integrated research program which includes a rank ordering of research activities and a balanced technology development program. A preliminary recommended program will be used to identify activities that will be subject to an independent technical evaluation and cost estimate. The panel's final recommendation to the Survey Committee will include consideration of the results of the independent technical evaluation and cost estimate.

In completing this task, each PPP will provide the Astronomy and Astrophysics 2010 Committee's Subcommittee on Programs with an interim internal and confidential summary report of its recommended program in the Fall of 2009 and complete its panel report thereafter. The panel reports will be published following the release of the survey committee's report in 2010.

RADIO, MILLIMETER, AND SUBMILLIMETER ASTRONOMY FROM THE GROUND STATEMENT OF TASK

The Astronomy and Astrophysics 2010 Program Prioritization Panel on Radio, Millimeter, and Submillimeter (RMS) Astronomy from the Ground will identify and recommend a prioritized program of federal investment in ground-based research activities that primarily operate in the radio, millimeter, and submillimeter portions of the electromagnetic spectrum. In formulating its conclusions, the RMS panel will draw on several sources of information: (1) the science forefronts identified by the Astro2010 science frontiers panels, (2) input from the proponents of research activities, and (3) independent cost and technical readiness assessments. The RMS panel's recommendations will be integrated into a program for all of astronomy and astrophysics by the Astro2010.

In particular, the Astronomy and Astrophysics 2010 Program Prioritization Panel on Radio, Millimeter, and Submillimeter (RMS) Astronomy from the Ground will:

1. Report on the status of existing RMS research activities to set the context for future research activities, incorporating findings of the Study Groups.

2. Preview and compare proposed RMS research activities including those car-

ried forward from previous surveys that have not been given a formal construction start.

3. State the relative importance of (a) smaller projects and generic research programs that involve competitive peer review and (b) programs that leverage public and private infrastructure investments, where appropriate.

4. Assess and describe best available estimates of the construction costs and lifetimes for each recommended research activity together with their full running costs (operations, science, and upgrades).

5. Identify particular risks for each research activity that would adversely affect the projected cost, technical readiness, or schedule of the activity. Identify those factors that could change an activity's priority and/or scope.

6. Informed by (a) the recommendations of the science frontier panels and (b) the panel's own research activity assessments, recommend a prioritized, balanced, and integrated RMS research program which includes a rank ordering of research activities and a balanced technology development program. A preliminary recommended program will be used to identify activities that will be subject to an independent technical evaluation and cost estimate. The panel's final recommendation to the Survey Committee will include consideration of the results of the independent technical evaluation and cost estimate.

In completing this task, each PPP will provide the Astronomy and Astrophysics 2010 Committee's Subcommittee on Programs with an interim internal and confidential summary report of its recommended program in the Fall of 2009 and complete its panel report thereafter. The panel reports will be published following the release of the survey committee's report in 2010.

B

Glossary

21-cm line: The hyperfine transition of neutral hydrogen (H I) with a corresponding rest-frame wavelength of 21 cm, or a frequency of 1.420 GHz. If the line is optically thin it is a direct measure of the neutral hydrogen content along a given line of sight.

^{56}Ni: Radioactive isotope of nickel, produced in large quantities in Type Ia supernovae.

α Cep: A rapidly rotating A-type star in the constellation of Cepheus. It has the traditional name Alderamin.

f_{NL}: A dimensionless parameter describing the amplitude of non-gaussianity of the fluctuations in the CMB.

Absorption line spectroscopy: A spectroscopic technique to measure the absorption of radiation as a function of frequency or wavelength in order to probe the nature of foreground material. Material along the line of sight absorbs light from the background source at wavelengths corresponding to atomic transitions that depend on its chemical composition, temperature and velocity. The pattern of absorption lines in the spectrum therefore yields information on the composition, temperature, and motion of the intervening matter.

Accretion disk: A flattened cloud of material that accretes onto a central gravitating mass (star or black hole).

Accretion-induced collapse (AIC): The collapse of a white dwarf, initially formed in a binary system, into a neutron star.

Acoustical peaks: Peaks in the temperature fluctuations of the cosmic microwave background as functions of the multipole moment *l*.

Active galactic nucleus/nuclei (AGN): Active accretion of mass onto the supermassive black hole at the center of a galaxy. Such black holes appear to be a feature of most galaxies, but in most cases, the rate of accretion is minimal. When the accretion rate is high, the AGN releases energy well in excess of the host galaxy starlight, all of it originating near the center of the galaxy, hence the name.

Adaptive optics (AO): A technology used to improve the resolution of ground-based telescopes by compensating for atmospheric distortions in real time.

Adiabatic fluctuations: Adiabatic fluctuations are fluctuations in the density of dark matter, ordinary matter, and radiation with the ratio of the density of matter components to the density of the radiation components held spatially constant.

AGN (quasar) winds: Outflows or jets generated by quasars or AGN. See also *Jets.*

Akari: Japanese satellite for infrared astronomy.

Allen Telescope Array (ATA): ATA-42 is an array of 42 6-meter-diameter radio telescope dishes at Hat Creek Observatory, California.

Arecibo: A 305-meter radio telescope in Puerto Rico, operated by the National Astronomy and Ionosphere Center (NAIC).

Asphericity: Deviation from a spherical shape.

Asteroseismology: The study of the internal structure of stars through observation of surface vibrations.

Astro2010: The National Research Council's decadal survey of astronomy and astrophysics published in 2010 as *New Worlds, New Horizons in Astronomy and Astrophysics.*

Astro-H: Japanese X-ray satellite, scheduled for launch in 2013; see http://astro-h.isas.jaxa.jp/ for details.

Asymptotic giant branch (AGB): A phase of stellar evolution undergone by a low to intermediate mass star (0.6 to 10 solar masses) late in its life and appearing in a second red giant phase. They have a dormant, helium-filled core surrounded by a helium-fusing shell and then a hydrogen-fusing shell.

Atacama Cosmology Telescope (ACT): A 6-meter-diameter, millimeter-wavelength telescope on Cerro Toco in the Atacama Desert of northern Chile, used to study the cosmic microwave background radiation (CMB).

Atacama Large Millimeter/submillimeter Array (ALMA): An international interferometer located at Llano de Chajnantor in the Atacama Desert of northern Chile.

Auger: The Pierre Auger Cosmic-Ray Observatory. Array of light detectors and water tanks in Argentina used to detect Čerenkov radiation produced by relativistic charged particles as they pass through Earth's atmosphere to study extremely high energy cosmic rays.

Axion: A postulated elementary particle. One of the mysteries of the standard model of physics is the lack of CP (charge conjugation and parity) symmetry violation in the strong interaction (that holds nuclei together). One of the solutions of the "CP problem" is to postulate a new symmetry, the Peccei-Quinn symmetry. If this symmetry exists in nature and is at the appropriate scale, then copious numbers of axions would be produced in the early universe. The axion is suggested as a possible explanation of dark matter.

B-mode polarization: Component of the CMB polarization map that can be expressed as the curl of a vector field and has zero divergence. This contrasts with the E-mode polarization, which can be expressed as the gradient of a scalar field (and hence is curl free). B-mode polarization can arise because of gravity waves from the inflation at the beginning of the universe.

Baryon: Strictly speaking, "baryon" is the collective term given to particles composed of three quarks (with the neutron and proton being the two most common examples). In the present context, the term is used to encompass all normal matter that interacts via electromagnetic forces. The vast majority of dark matter is believed to be non-baryonic.

Baryon acoustic oscillation (BAO): Sound waves in the early universe generate fluctuations both in the cosmic microwave background spectrum and the matter (baryon) fluctuation spectrum.

Be stars: Luminous blue stars, often rapidly rotating, with prominent hydrogen emission lines.

Big bang nucleosynthesis (BBN): The production of nuclei during the first 10 minutes of the universe.

Binary millisecond pulsars: Orbiting pairs of rapidly rotating neutron stars that emit radio signals.

Black hole: A body so massive and dense that its gravity prevents even light from escaping.

Black hole feedback: The release of energy generated by an AGN into the surrounding interstellar or intergalactic medium in clusters of galaxies, and the effect of that energy on processes like star formation or galaxy formation.

Blazars: AGN with relativistic jets aligned close to the line of sight, such that the jet emission is Doppler beamed, an effect in Einstein's theory of special relativity whereby radiation from a moving source appears to be concentrated in the forward direction. Blazar emission is therefore strongly jet-dominated, emitting at radio through gamma-ray wavelengths, and blazars appear to be highly variable and polarized.

Blue stragglers: Stars in open or globular clusters that are hotter and bluer than other cluster stars whose initial mass should have evolved it away from the main sequence and thereby exhibiting atypical stellar evolution.

Bose condensate: A state of matter of a dilute gas of weakly interacting bosons confined in an external potential and cooled to temperatures very near to absolute zero.

Brown dwarfs: Cool, low-mass objects roughly at the boundary of mass between stars and planets.

Cadence: Regular time interval of observation.

Cas A: Young supernova remnant (exploded near 1680 AD) in the constellation Cassiopeia.

Cataclysmic variable (CV): Binary star systems that have a white dwarf and a companion star. The companion star loses material onto the white dwarf and forms an accretion disk. Strong UV and X-ray emission is often seen from the accretion disk. The accretion disk may be prone to instability leading to outbursts of rapidly burning hydrogen.

Centrifugal decretion disk: Disk produced around a rapidly rotating (usually type Be) star (see also *Decretion disk*).

Čerenkov radiation: Electromagnetic radiation emitted by a charged particle moving faster than the phase velocity of light in a dielectric medium such as water or air.

Chandra: The Chandra X-ray Observatory, launched in 1999, one of NASA's four Great Observatories.

Chandrasekhar limit/mass: The mass (about 1.44 solar masses) supported by quantum-mechanical electron degeneracy pressure above which a star will ultimately collapse into a neutron star or black hole.

Charge parity symmetry: The combination of charge conjugation symmetry and parity symmetry, states that the laws of physics should be the same if a particle were interchanged with its antiparticle (charge conjugation symmetry) and left and right were swapped (parity symmetry).

Chemical enrichment/Metal enrichment: Processes related to the creation and distribution of all elements heavier than helium, which are synthesized during the evolution of stars.

Chirp: A signal with rapidly increasing frequency.

Chromosphere: The thin, outer layer of a star at T ~ 10,000 K that lies above the photosphere.

Circumgalactic: Around a galaxy.

Circumgalactic medium (CGM): The matter pervading the space and the gaseous medium located 100-200 kpc around a galaxy that are strongly influenced by their gravity and by chemical and mechanical feedback.

Clusters of galaxies: A large group of galaxies bound together gravitationally.

Clusters of stars: A group of stars formed at about the same time.

CMBPol: A next-generation successor to the Planck spacecraft designed to measure CMB B-mode polarization.

Coherent emission from brown dwarfs: Emission process in the magnetized region around a low-mass dwarf in which bunches of electrons produce very intense radio radiation.

Column density: The total amount of material along a path (e.g., the path line of sight from a quasar or galaxy to us), expressed as the number of atoms of hydrogen or mass per square centimeter.

Compact binary: A close two-star system that contains a compact object—a white dwarf star, neutron star, or black hole.

Compact stellar remnant: The endpoint of stellar evolution—a white dwarf star, neutron star, or black hole.

Compton: The Compton Gamma-Ray Observatory, launched in 1991 and deorbited in 2000.

Confusion limit: Brightness level in a given image below which it becomes impossible to determine the properties of individual sources with any confidence, because of fluctuations in a large underlying population of still fainter unresolved sources. At a particular wavelength, confusion can be ameliorated by re-observing the same target field at higher angular resolution.

Convective star: A star in which most of its energy is transported by moving matter rather than by the diffusion of radiation or conduction.

Cool cores: The centers of clusters of galaxies in which the temperature drops below the temperature in the surrounding gas. Clusters with cool cores have high central surface brightnesses and high central gas densities and short apparent cooling times.

Core-collapse supernova: A supernova explosion produced when the core of a massive star runs out of nuclear fuel and collapses to neutron-star densities.

Corona: The outermost part of the stellar atmosphere, a tenuous but extremely hot (million K) plasma.

Correlator: A device employed to correlate two signals, used for imaging and spectroscopy. It is necessary to combine signals from separate antennas in a radio interferometer.

Cosmic microwave background (CMB): Blackbody radiation pervading the universe, which is left over from the hot big bang (the initial state of the universe) and cooled by the expansion to a current temperature of 2.73 K above absolute zero. Fluctuations in the CMB intensity correspond to density fluctuations that grow over time through gravity.

Cosmic Origins Spectrograph (COS): UV spectrograph on HST, installed during the 2009 servicing mission.

Cosmic strings: Filaments formed by symmetry breaking in various particle physics models that may stretch across the visible universe.

Cosmic variance: The statistical uncertainty of a limited sample size in observations of the universe at very large distances, based on the idea that it is possible to observe only part of the universe at a particular time, thus making it difficult to draw statistical statements about cosmology on the scale of the entire universe.

Cosmic web: The structure of the universe created from intergalactic gas and galaxies, embedded in filaments, stretching between voids.

Dark ages: Term for the long period when the universe is opaque between the emission of the CMB at redshift $z \sim 1,100$ and the formation of the first radiating objects at redshift $z \sim 30$ to 50.

Dark energy: Hypothetical form of energy that causes the present-day expansion of the universe to accelerate.

Dark Energy Survey (DES): A ground-based survey to study the nature of dark energy by observing the distributions of galaxies and clusters of galaxies, weak gravitational lensing, and supernovae.

Dark matter: The hypothetical form of matter that dominates the mass budget of galaxies, clusters, and the universe and is thought to interact only through gravitational forces. The presence of dark matter is inferred from its gravity; experiments to detect the dark matter particles directly are underway.

Dark matter halo: Collection of dark matter particles that has evolved from an initially overdense region of the universe, separated from the global expansion of the universe, and collapsed into a virialized configuration. Dark matter halos range in mass from subgalactic to cluster scales, and their associated gravitational potential wells bind together the baryonic components of galaxies and clusters.

Debris disk: A disk of matter around a young star that contains dust grains, generated by collisions of protoplanets or asteroids.

Deconfined quark matter: At sufficiently high temperatures or densities, hadronic matter should evolve into a new phase of matter containing almost free quarks and gluons, also known as a quark-gluon plasma or quark matter. Deconfined quark matter may have existed in the first few microseconds after the big bang and might exist inside the cores of dense neutron stars.

Decretion: The centrifugal ejection of material.

Decretion disk: An equatorial disk formed by the centrifugal ejection of material (see also *Centrifugal decretion disk*).

Degenerate core: The dense central region of an evolved star supported by electron degeneracy pressure.

Dome C: A high-altitude site on the ice sheet atop the Antarctic Plateau. Site of the Concordia Research Station operated by France and Italy, more than 1,000 km from the coast.

Doppler tomography: Rotation velocities in stars or disks produce Doppler shifts, which can be used to reconstruct the geometry of the rotating object.

Double pulsar: The system J0737-3039, an exceptionally relativistic binary pulsar discovered in 2003.

Dynamical dark energy: If the energy density of the dark energy can evolve with time and the dark energy can cluster, then the dark energy is called "dynamical." Quintessence models are an example of a model with dynamical dark energy. If the dark energy is vacuum energy, then it is not dynamical.

Early-type stars: Stars hot enough to keep hydrogen ionized at their surface, typically with spectral classifications O, B, and A.

Einstein ring: The image formed when a background source located directly behind a massive galaxy undergoes gravitational lensing.

Electromagnetic Observations from Space (EOS): One of the four Astro2010 Program Prioritization Panels.

Ellipticals: Galaxies with a smooth brightness profile and an ellipsoidal shape that ranges from nearly spherical to highly flattened. Most ellipticals show signs that they are made primarily of old stars and have little current star formation.

e-MERLIN: The enhanced version of the United Kingdom's Multi-Element Radio Linked Interferometer Network (MERLIN), an interferometer array of radio telescopes. In e-MERLIN, the microwave links between dishes are replaced by optical fibers.

EMIR: Near-infrared multiobject spectrograph for the Gran Telescopio de Canarias.

E-mode polarization: Component of the CMB polarization map that can be expressed as the gradient of a scalar field (and hence is curl free). This contrasts with the B-mode polarization, which can be expressed as the curl of a vector field and has zero divergence. Temperature and density fluctuations at the last scattering surface generate a pure E-mode polarization.

Epoch of reionization: The period during which the baryonic content of the universe is gradually ionized. The onset of reionization coincides with the emergence of the first ionizing (luminous) sources, at the end of the dark ages, and ends when the intergalactic medium is fully ionized.

eRosita/Spectrum RG: Extended Röntgen Survey with an Imaging Telescope Array, a German X-ray instrument designed to fly on the Russian Spectrum RG satellite in 2012.

Event horizon: The boundary in space-time, most often surrounding a black hole, from within which no signals can reach an outside observer.

Exoplanet: An extrasolar planet, i.e., a planet orbiting a star other than the Sun.

Exozodiacal light: Emission from dust in debris disks.

Expanded Very Large Array (EVLA): Upgrade to the Very Large Array radio interferometer in New Mexico, which will yield greatly increased sensitivity, spectral capabilities, and other improvements.

Extremely Large Telescope (ELT): A generic reference to the kind of 20-meter to 40-meter-class optical and infrared telescopes currently being designed in the United States and Europe.

Faber-Jackson relation: An empirical correlation between the central velocity dispersions and luminosities of elliptical galaxies

Facility for Rare Isotope Beams (FRIB): An ongoing Department of Energy project to establish an accelerator facility to study rare, unstable nuclei.

Far-infrared: Wavelengths from 20 to 200 microns (also see *Infrared*).

Feedback: Energy input from various physical processes—such as supernovae explosions, ionizing radiation, stellar and galactic winds, black hole activity, jets, and more—that is emitted back into the surrounding medium around stars, galaxies, and the intracluster and intergalactic medium, affecting its energetics and distribution.

Fermi GLAST: The Fermi Gamma-Ray Space Telescope, formerly GLAST (Gamma-ray Large Area Space Telescope), a large-area gamma-ray satellite launched by NASA in 2008.

Field of view (FOV): The maximum angular extent observable through an optical or other device.

Filaments: See *Cosmic web.*

Flamingos II: A near-infrared, multiobject spectrometer and wide-field imager built for Gemini South by the University of Florida.

Fossil field: Magnetic field left from a previous evolutionary phase.

Fundamental plane: A relation between effective radius, surface brightness, and central velocity dispersion for elliptical galaxies.

Gaia: A European Space Agency astrometry space mission to provide precise astrometric data and spectroscopic information.

Galactic Neighborhood (GAN): One of the five Astro2010 Science Frontiers Panels.

Galaxies Across Cosmic Time (GCT): One of the five Astro2010 Science Frontiers Panels.

Galaxy Evolution Explorer (GALEX): An orbiting space telescope that observes galaxies and gas in ultraviolet light.

Galaxy kinematics: The motions of stars and gas in a galaxy.

Gamma ray (γ ray): Light with wavelength shorter than X-ray light, roughly 0.1 MeV (100 keV) to more than 100 TeV, or frequency range log $\nu \sim 20.5$ to 28.5 Hz.

Gamma-ray burst (GRB): Extremely bright flashes of gamma rays, probably produced by compact objects in distant galaxies. Long-duration (>2 seconds) GRBs are probably associated with the explosions of very massive stars. Short-duration (<2 seconds) GRBs may be produced by merging neutron stars.

Gas accretion: Infall of gas from the intergalactic medium onto star-forming regions of galaxies or extended gaseous halos. Such infall provides a source of raw material to form new stars.

Gauss: A unit of magnetic field strength.

Gaussian fluctuations: Fluctuations in the matter distribution or CMB temperature are Gaussian if the probability distribution of deviations follows a standard bell curve, with $p(\Delta)$ proportional to $\exp(-\Delta^2/2\ s^2)$.

Gemini: The Gemini Observatory, which consists of two 8-m telescopes, one in Hawaii and one in Chile.

General relativity: Einstein's theory of gravitation in terms of the curvature of space-time produced by mass and energy density.

Gigayear (Gyr): One billion years.

Globular cluster: A nearly spherical cluster of 10^5 to 10^6 stars.

Grand-design galaxy: A spiral galaxy in which the spiral arms are especially prominent.

Gravitational lensing: Deflection of light by the gravitational potential associated with astronomical objects, such as stars, galaxies, and groups and clusters of galaxies. Two distinct regimes are important: (1) strong lensing, where the gravitational potential is deep enough to produce multiple images of a background source, and (2) weak lensing, where the gravitational potential distorts the appearance of background sources but does not create multiple images.

Gravitational microlensing: An intensification of light from a background star produced by the gravity of a mass, such as another star or an exoplanet, that is aligned almost exactly along the line of sight.

Gravitational radiation: Propagating waves of space-time curvature predicted by Einstein's theory of general relativity. Gravitational waves propagate at the speed of light.

Gravitational wave astronomy: Branch of observational astronomy that uses gravitational waves (fluctuations in space-time) to collect data about the source.

Greisen-Zatsepin-Kuzmin (GZK) cutoff: Ultrahigh-energy cosmic rays (mostly protons) with energies above 5×10^{19} eV interact with cosmic microwave background radiation to produce pions that decay into muons, electrons, neutrinos, and photons. Because of this interaction, cosmic rays above this energy cannot propagate over very large distances (50-100 Mpc or 150-300 million light-years).

Gunn Peterson absorption trough: A broad absorption feature shortward of Lyman alpha emission in QSO spectra produced by redshifted H I in the intergalactic medium.

Gyrosynchrotron: The electromagnetic radiation emitted by a charged particle moving at relativistic velocities in a magnetic field.

H II regions: Emission nebulae created when young, massive stars photoionize nearby gas clouds, composed mostly of hydrogen.

Helioseismology: The study of the interior of the Sun through observation of surface vibrations.

Herschel Space Observatory: A European Space Agency satellite launched in 2009 to study the formation of stars and galaxies at far-infrared and submillimeter wavelengths.

Hertzsprung-Russell diagram (HR diagram): A plot of the luminosity versus surface temperature (or magnitude versus color) of a set of stars. Location in the HR diagram is a diagnostic of the mass and evolutionary state of a star.

HESS: The High Energy Stereoscopic System, an array of four telescopes located in Namibia, designed to map very high energy (TeV) gamma-ray radiation from cosmic sources. It works by tracing the Čerenkov light produced by airshowers of secondary charged particles when gamma rays traverse the atmosphere.

High-cadence observations: A sequence of short observations taken in rapid succession.

Hinode: A Japanese solar spacecraft equipped with advanced, high-resolution solar telescopes for optical and X-ray wavelengths and an extreme ultraviolet imaging spectrometer.

HiRes: The High Resolution Fly's Eye cosmic-ray detector experiment in Utah.

Hubble constant: Constant of proportionality between the distance to a galaxy and its recession velocity (redshift); describes the rate of the expansion of the universe.

Hubble Space Telescope (HST): Orbiting optical/UV/near-IR space telescope launched in 1990 and refurbished at periodic intervals. Current Hubble Space Telescope cameras can image with ~0.1-arcsecond resolution and take UV spectra between 1,100 and 3,000 Å.

Hydrodynamic processes (for stars): Physical behavior of bulk stellar gas, taking into account gravity, shocks, radiative cooling, photoionization, and other effects.

Hydrodynamical simulations: Numerical simulations that attempt to model directly the observable baryonic component of the universe and dark matter by incorporating processes associated with gas physics in addition to N-body gravity, including shocks, radiative cooling, and photoionization, along with selected other processes relevant to the system being modeled.

Hypernovae: Explosions of energy 10^{52} ergs of extremely massive stars at the end of their lives. The endpoints of somewhat smaller stars are supernovae of energy 10^{51} ergs.

Hyperon: A type of baryon containing one or more strange quarks, but no charm quarks or bottom quarks.

IceCube: A neutrino-detection experiment that uses the ice in Antarctica as the detection medium.

Inflation: The theorized exponential expansion of the universe that began ~10^{-35} seconds after the big bang and helps to explain the observed flatness of space and seeds for the large-scale spatial distribution of matter and galaxies.

Infrared (IR): Light with wavelength longer than visible light but shorter than radio waves. Near-infrared refers to the wavelength range from ~1 to 5 microns, mid-infrared is ~5 to 20 microns, and far-infrared extends from ~20 microns to ~200 microns, where the submillimeter begins.

Initial mass function (IMF): An empirical function that describes the mass distribution of a population of stars in terms of their initial mass.

Inspiral: The decaying orbit of a binary system.

Integral field unit (IFU): Instrument that performs imaging spectroscopy, such that a "data cube" gives the spectrum at multiple locations across a source. Typically these have fields of view in the arcsecond-arcminute range, with 10 to 100 pixels per IFU.

Interferometry: Technique using the wave interference of light from two or more telescopes to infer information about the spatial distribution of the light source.

Intergalactic medium (IGM): The matter pervading the space between galaxies and likely to contain a significant fraction of ordinary baryonic matter.

Interstellar medium: The matter that exists between the stellar systems in a galaxy.

Intracluster medium (ICM): The matter pervading the space between galaxies in clusters and groups of galaxies, usually referring to the superheated gas at the center of a cluster.

Iron K-alpha (Fe Kα): X-ray spectral line emitted when an inner-shell electron of iron drops from the 2p orbital of the "L" shell (principal quantum number $n = 2$) to the innermost "K" shell ($n = 1$). The line provides diagnostics of the thermal and dynamical state of matter close to neutron stars and black holes and may be useful for measuring black hole spin.

IR Zeeman: Determination of the magnetic field of stars achieved through observing shifted infrared spectral lines.

Isocurvature fluctuations: Fluctuations in the density of dark matter, ordinary matter, and radiation with the total density held spatially constant. For example, dark matter isocurvature models have regions of high (low) dark matter density compensated by regions of low (high) photon density.

James Webb Space Telescope (JWST): An infrared space observatory under construction, scheduled for launch in 2016 or later, which is the successor to HST.

Jansky (Jy): A unit of electromagnetic flux density equal to 10^{-26} W m^{-2} Hz^{-1}.

Jets: Collimated outflows from stars or AGN, probably generated in the vicinity of compact objects (black holes or neutron stars). Jets are "relativistic" when they are moving at nearly the speed of light.

Joint Dark Energy Missions (JDEM/IDECS and JDEM/Omega): Two versions of a joint NASA-DOE dark energy mission. JDEM/Omega omits the CCD imager that is included in JDEM/IDECS.

Kepler: Kepler, a NASA satellite mission, is a photometric monitor of 150,000 stars looking for transits of extra-solar planets ranging in size from Jupiter down to sub-Neptune and Earth-like in radius.

Kuiper belt: An annulus of the thick outer disk of the solar system, at distance beyond 40 AU from the Sun, that contains thousands of ice bodies more than 100 km in diameter and trillions of comets.

Kuiper belt objects (KBOs): Bodies within the Kuiper belt.

Laboratory astrophysics: The investigation in terrestrial laboratories of materials or physical processes that are of fundamental importance to interpreting and/or modeling astronomical observations.

Lambda cold dark matter (ΛCDM): This model currently provides the best fit to the observational data. It posits a flat universe composed of ordinary (baryonic) matter, cold (heavy) dark matter, and a vacuum energy (like the cosmological constant, Λ), and with adiabatic nearly scale-invariant Gaussian random phase fluctuations.

Large-angular-scale B-modes: See *B-Mode Polarization.*

Large-scale structure (LSS): Refers to the characterization of observable distributions of matter and light on the largest scales.

Lepton number: A measure of light, weakly interacting particles and antiparticles (electrons, muons, tauons, and their neutrinos). Electrons and neutrinos have lepton number +1; positrons and anti-neutrinos have lepton number −1. The universe has a lepton number asymmetry if there are more neutrinos than anti-neutrinos (or vice versa). A lepton number asymmetry alters the isotope abundances predicted by big bang nucleosynthesis.

Life-cycle cost: Total project cost, including launch and operations, for a space mission.

Light curve: The observed apparent intensity/brightness of an object as a function of time.

Local Group: The Milky Way, Andromeda and M-33 spiral galaxies, the Magellanic Clouds, and perhaps 30 dwarf galaxies and some two dozen other galaxies within about a megaparsec of the Milky Way.

Low Frequency Array for radio astronomy (LOFAR): A large-area, interferometric array of radio telescopes, based in the Netherlands, that will conduct a survey of the universe at frequencies below 250 MHz.

Luminosity: The amount of energy an object radiates per unit time (measured in erg s^{-1}).

Luminosity function: A mathematical function describing the number of stars or galaxies (or other objects) per luminosity interval.

M dwarf: A small, cool, low-mass star on the main sequence (0.1 to 0.3 solar masses).

$M_\odot$: One solar mass, equal to the mass of the Sun.

M-σ relation: The strong correlation observed between the stellar velocity dispersion in galaxies, σ, and the masses of their central black holes, M.

Magellanic Clouds: Two small galaxies (Large and Small Magellanic Clouds) that are gravitationally bound satellites of the Milky Way Galaxy located 50-60 kpc in the galactic halo.

Magnetar: A neutron star with an extremely strong magnetic field, the decay of which powers high-energy emission.

Magnetic reconnection: The process whereby magnetic field lines from different magnetic domains are spliced to one another, changing their patterns of connectivity with respect to the sources.

Magnetohydrodynamics (MHD): Theory of electrically conducting fluids or gases, believed to be broadly applicable to ionized gases in astrophysical systems. In the idealized case usually assumed in numerical simulations, the electrical conductivity of the gas is so high that it shorts out the electric field. However, the gas is still subject to forces due to magnetic fields, which can exert both pressure and tension. The proper treatment of MHD effects is essential to the study of black hole accretion and jets, and is important for many other problems of galaxy formation and evolution, such as star formation and the transport of energy in the atmospheres of galaxy clusters.

Magnetometry: Measurement of the strength and direction of a magnetic field.

Main-sequence turnoff: The high-mass end of the locus of hydrogen-burning stars (the main sequence) in the Hertzsprung-Russell diagram. Stars with higher masses have already completed core-hydrogen burning, whereas those with lower masses have not yet exhausted central hydrogen.

Major Atmospheric Gamma-ray Imaging Čerenkov (MAGIC) Telescope: An imaging atmospheric Čerenkov telescope located on La Palma in the Canary Islands. Its two dishes detect gamma rays with energies from 50 GeV to 30 TeV.

Mass function: A mathematical function describing the number of stars or galaxies (or other objects) per mass interval.

Mass-loss nebula: The nebulous emission around a star resulting from its ejection of mass through an eruption or stellar wind.

Mass-loss processes: Mechanisms for propelling the loss of matter from a star.

Metal-poor stars: Stars with an iron-to-hydrogen abundance ratio of less than 1/10 that of the Sun.

Metallicity: The proportion of a star's matter that is made up of chemical elements other than hydrogen and helium.

Microlensing: See *Gravitational microlensing.*

Mid-infrared: Light with wavelengths between 5 and 20 microns (see also *Infrared*).

Milagro: An observatory in New Mexico that detects gamma rays and Čerenkov radiation from cosmic-rays: A water Čerenkov system that detects airshower particles produced by ultrahigh-energy cosmic rays passing through Earth's atmosphere.

Milky Way: The spiral galaxy in which we live.

Millimeter: Wavelength region from 1 to ~5 millimeters.

Mode: See *E-mode polarization, B-mode polarization.*

Multiscale physics: The study of how microscopic properties of systems can lead to macroscopic behavior. The thermal regulation of the gas in a cluster of galaxies (millions of light-years across) by accretion onto a central black hole (similar in size to the solar system) is one example of multiscale physics.

Murchison Widefield Array (MWA): Formerly the Mileura Widefield Array, this is a low-frequency radio (at frequencies from 80 to 300 MHz) interferometer array to be built in Western Australia.

Near-infrared: Wavelengths from ~1 to 5 microns (also see *Infrared*).

Neutrinos: Weakly interacting elementary particles that are emitted, for example, in beta decay. They come in three "flavors" with differing masses and oscillate between those flavors as they travel. Neutrinos emitted by nuclear reactions in the core of the Sun or in the formation of a neutron star in a core-collapse supernova can be detected on Earth.

Neutron skin: Thin layer of pure neutron matter covering the surface of many nuclei.

Neutron star: A collapsed star with a mass comparable to that of the Sun crammed into the dimensions of a city, giving it a density comparable to that of an atomic nucleus.

Newton: See *XMM-Newton.*

Non-thermal radio emission: Radiation emitted by particles for reasons other than the high temperature of the source. The spectrum of non-thermal radiation is different from that predicted by Planck's law for a blackbody.

Nuclear cross section: The effective size (area in cm^2) of a nucleus for a reaction of a given type for a given collision energy.

OB association: A group of high-mass stars, looser than a cluster.

Omega (Ω): The ratio of the density of the universe to the critical density of the universe. If omega is greater than one, the universe will turn around and begin collapsing; if omega is one, the universe will eventually stop expanding but will not collapse; and if omega is less than one, the universe will continue expanding.

Opacity: A measure of the ability of a gas to absorb or scatter radiation.

Optical: Wavelength range of light to which the human eye is sensitive, namely, 3,500 to 8,000 Å.

Optical and Infrared Astronomy from the Ground (OIR): One of the four Astro2010 Program Prioritization Panels.

Outflows: Gas flowing outward and escaping from gravitating bodies. Examples include stellar winds, galactic winds powered by supernovae in star-forming galaxies, winds from the surfaces of accretion disks, and jets emerging from the vicinity of accreting black holes.

PanStarrs-1: Panoramic Survey Telescope and Rapid Response System, a wide-field (1.8 meter) imaging facility being developed at the University of Hawaii's Institute for Astronomy.

Parity-violating electron scattering: An experimental technique to determine the neutron distribution in a nucleus by weak-force scattering of polarized electrons.

Particle Astrophysics and Gravitation (PAG): One of the four Astro2010 Program Prioritization Panels.

Periastron advance: Angular change in the periastron point in the orbital plane;

the periastron is the point in the orbit of one component of a binary system where it is nearest the other component.

Phased-array feed: A way to provide multibeaming with a radio telescope.

Photometric redshifts: The recession velocity based on measuring galaxy photometry in multiple optical and near-IR wavelength bands. These measurements of the shape of the galaxy energy emission can be obtained more quickly than spectroscopic redshifts. However, they are less accurate.

Photometry: Measurement of the brightness of an object.

Planck: An ESA satellite launched in 2009 to map tiny fluctuations in the cosmic microwave background radiation.

Planetary Systems and Star Formation (PSF): One of the five Astro2010 Science Frontiers Panels.

Point spread function (PSF): The point spread function describes the angular response of a telescope to a point-like source of light. Because of atmospheric turbulence and the finite resolving power of a telescope, light from point-like object is spread out in angle.

Polarimetry: The measurement and interpretation of the polarization of transverse electromagnetic waves.

Population III: The first stars of zero metallicity, which formed before the heavy elements were produced by stars.

Post-Swift GRB: Gamma-ray burst studied after the completion of the Swift mission.

Pre-main sequence: Collapsing protostellar cores that have not yet ignited hydrogen fusion in their interiors.

Program Prioritization Panels (PPPs): Four panels of scientists constituted by the National Research Council to write reports on priorities across specific disciplines for the Astro2010 Decadal Survey.

Prolate: A spheroid with a shape like an ellipse rotated about its major axis, i.e., football-shaped.

Protoplanetary disk: A primordial disk around a young star that has just begun to form planets.

Pulsar: A rotating neutron star that emits periodic radio pulses of periods ranging from milliseconds to seconds.

Pulsar/black-hole binary: A binary system in which one star is a pulsar and the other is a black hole.

Quantum chromodynamics (QCD): A quantum field theory that describes the strong nuclear forces and interactions of interaction of quarks and gluons.

Quantum gravity: A physical theory that describes the gravitational interactions of matter and energy, which are themselves described by quantum theory.

Quasar (AGN) winds: Outflows or jets generated by quasars or AGN. See also *Jets.*

Quasar sightline: Line of sight through interstellar and intergalactic material of interest toward a background quasar.

Radiative cooling: The loss of energy by a gas as the result of the emission of radiation.

Radio: Light that has wavelengths longer than the far-infrared, roughly beginning at 200 microns.

Radio galaxy: Galaxy that emits strong radio emission from powerful jets and lobes.

Radio jet: See *Jets.*

Radio lobes: Diffuse extended radio-emitting plasma fed by jets and formed when the material in the jet decelerates and interacts with the intergalactic medium.

Radio, Millimeter, and Submillimeter Astronomy from the Ground (RMS): One of the four Astro2010 Program Prioritization Panels.

Radio transients: Time variable radio source.

Red giant branch (RGB): A phase of stellar evolution in which the stars have expanded their outer envelope in response to the extra energy generated by gravitational contraction, hydrogen shell fusion, and ultimately helium fusion in the core.

Redshift: The increase in wavelength of electromagnetic radiation (doppler shift) caused by the motion of an object as described by the theory of special relativity, or by luminous material in a gravitational field as described by the general theory of relativity. In cosmology, it refers to the fractional increase in the wavelength of a photon received from a distant object, because of the expansion of the universe between emission and reception. Coupled with a cosmological model, redshifts can be used to determine the distance and look-back time of phenomena in the universe.

Redshift space distortions: Effects of inhomogeneities in light propagation. On small scales, random motions will cause particles at the same distance to have slightly different redshifts. This elongates structures along the line of sight. On very large scales, the opposite happens and overdensities are enhanced.

Redward loops: Paths of post-main-sequence stellar evolution that proceed toward the red spectral range in the Hertzsprung-Russell diagram.

Reionization: See *Epoch of reionization.*

Relativistic jet: See *Jets.*

Relativistically broadened line: A spectral line that appears wider to an observer because of the movement of its source at speeds approaching the speed of light.

Rotation curve: Representation of the orbital velocity of the stars or gas in a galaxy as a function of radius from the center of the galaxy.

S0 galaxy: A disk galaxy similar to a spiral galaxy, but shows no spiral arms.

Schmidt-Kennicutt relation: An empirical relation between the surface density of star formation rate and the surface density of gas in galaxies.

Schwarzschild radius: A characteristic radius associated with every quantity of mass, generally referring to a collapsing celestial object, which is (3 km) times the black hole mass (measured in solar-mass units). It is the radius of the object at which the force of gravity would be so great that no known force could stop the mass from continuing to collapse into a point of infinite density.

Science Frontiers Panel (SFPs): Five panels of scientists constituted by the National Research Council to write reports on science priorities for the Astro2010 Decadal Survey.

SFSR inflation: Single-field, slow-roll inflation, a simple model for inflation whereby inflation is driven by the displacement of a scalar field from the minimum of its potential. In slow-roll inflationary scenarios, the homogeneous part of the inflation field rolls slowly down its potential toward a minimum.

Shapiro delay: The delay induced in the travel time of a light ray passing close to a massive object by the presence of space-time curvature; named after Irwin Shapiro, who first suggested searching for the delay associated with the space-time curvature induced by the Sun's mass using radar ranging to Venus.

Shock: Supersonic interactions with ambient gas, as happens when the supersonic wave of gas from a supernova hits the ambient gas.

Shock-heated: Heated by supersonic interactions with ambient gas, as happens, for instance, when infalling gas accelerated by gravity hits ambient gas within a galaxy's halo.

SMBH (supermassive black hole) merger tree: History of black hole mergers. Galaxies form as small pieces that gradually merge to create larger and larger galaxies. Black holes in the centers of these galaxies are thought to merge in the same interactions, although the details of evolution from a tight black hole binary to a single merger remnant are not understood.

Solar-neutrino experiments: Experiments designed to detect weakly interacting neutrinos that are emitted during nuclear reactions in the core of the Sun.

South Pole Telescope (SPT): A microwave/millimeter-wave telescope with a 10-m diameter, located at the Amundsen-Scott South Pole Station, Antarctica. Its main scientific mission is to survey several thousand clusters of galaxies, in order to constrain the nature of the dark energy and other aspects of cosmic evolution.

Spectroscopic abundance: The chemical composition of a star as inferred from analysis of absorption and emission features formed in the star's outer layers.

Spectroscopic redshift: The recession velocity based on measuring the shift of spectral lines from their usual position.

Spindown time: The characteristic time for a star to slow its rotation due to loss of angular momentum through its stellar wind and/or magnetic field. It is typically much shorter in magnetically active stars.

Spitzer: NASA's Space Infrared Telescope Facility, launched in 2003.

Standard model of particle physics: The theory of fundamental interactions and elementary particles that make up all the visible matter in the universe; it does not include gravitation, dark matter, or dark energy.

Starburst galaxy: A galaxy, or a region of a galaxy, with an exceptionally high rate of star formation.

Stars and Stellar Evolution (SSE): One of the five Astro2010 Science Frontiers Panels.

Starspot: The equivalent of sunspots (cool surface features) on other stars.

Stokes Q/U/V sensitivity: Measurement of the surface magnetism of stars by measuring the polarization of light in various spectral bands. Stokes parameters describe the polarization state of electromagnetic radiation.

Strömgren sphere: A sphere of ionized hydrogen gas produced by intense UV radiation from a massive, young star balanced by ion-electron captures (recombinations).

Subgrid: Below the resolution of the grid in a numerical calculation.

Submillimeter: Wavelengths between ~0.2 millimeters (200 microns) and 1 millimeter (1,000 microns).

Suborbital: NASA's program of sounding rockets and balloons.

Subsonic burning: A flame that moves slower than the sound speed, i.e., a deflagration, in contrast to a detonation.

Sunyaev-Zel'dovich effect (SZE): The scattering of cosmic microwave background photons by the electrons in hot, ionized gas, which appears as a deficit or enhancement (depending on observing frequency) in radio continuum maps. Clusters of galaxies can be detected through the change they produce in the CMB spectrum. The SZ signal is proportional to the integral of the pressure (density of electrons times temperature) through the cluster and is independent of distance. The magnitude of the effect depends on the path integral of the electron density, and so cluster detection roughly scales as the cluster mass.

Superbubble: An extremely large bubble of very hot gas produced in the interstellar medium by multiple supernova explosions or stellar winds.

Supergiant: An extremely luminous, massive star.

Supermassive black hole (SMBH): A black hole with a mass in excess of a million solar masses.

Supernova neutrino background: When a massive star collapses down to a neutron star, it produces a core collapse supernova explosion. Most of the energy in a supernova explosion is radiated in neutrinos. The neutrinos radiated from all of the core collapse supernovae make up the neutrino background.

Supersymmetry: A particle-physics symmetry relating elementary particle differing by one-half unit of spin.

Superwind: A powerful outflow of hot gas from a starburst galaxy, driven by the combined effects of supernova explosions and winds from massive stars.

Suzaku: Japanese X-ray satellite, launched in 2005, featuring high spectroscopic resolution and a very wide energy band.

Swift: A NASA medium-class gamma-ray (and UV/X-ray) Explorer mission to provide rapid identification of GRBs and multiwavelength follow-up.

Synoptic solar magnetometry: Long-term and daily measurements of the solar photospheric and coronal magnetic fields.

Tau neutrino: One of the three flavors of neutrinos. The tau neutrino, along with the associated tauon, forms the third generation of leptons.

TeV experiment: An experiment designed to measure radiation from astronomical sources in the TeV (10^{12} eV) energy range of gamma rays. Most modern TeV experiments operate from the ground using the imaging atmospheric Čerenkov techniques (see *HESS*, *MAGIC*, and *Veritas*).

Thermal disk component of the X-ray continuum spectrum: Emission from the hot surface of an accretion disk around a compact object.

Time-domain survey: A survey of objects that focuses on obtaining their fluxes as a function of time, perhaps in several wavelength bands. The critical elements are high sampling rate, signal-to-noise ratio, and time span of coverage with as few gaps as possible.

Transitional disk: A protoplanetary disk with a gap, or in other usage, a disk in which the dust has grown in size and settled.

Tully-Fisher relation: An empirical correlation between the rotation velocities and luminosities of spiral galaxies.

Type Ia supernova: Observationally, a supernova without hydrogen and with strong lines of ionized silicon; theoretically the thermonuclear explosion of an accreting white dwarf in a binary system with a light curve powered by the decay of radioactive ^{56}Ni and ^{56}Co.

Type II supernova: Observationally, a supernova with spectroscopic evidence of hydrogen; theoretically, typically the result of core collapse to a neutron star or black hole in a star with an extended hydrogen envelope, though thermonuclear varieties are possible.

Uhuru: The first satellite dedicated to X-ray astronomy.

Ultradense matter: Exotic new states of matter, e.g., deconfined quark matter or Bose condensates.

Ultrahigh-energy cosmic rays (UHECRs): Fast-moving particles with detected energy occasionally over 10^{20} eV. They come from unidentified cosmic accelerators outside our galaxy. We detect them when they collide with atoms in the upper atmosphere and produce a shower of particles and faint blue light from fluorescence. The flux of UHECRs is a sharply declining function of energy.

Ultraviolet (UV): Light with wavelength slightly shorter than visible light, from 3,000 Å down to 912 Å (far ultraviolet or FUV) and ~100 Å (extreme ultraviolet or EUV).

Velocity dispersion: The spread of the random velocities of stars, gas, or galaxies.

Velocity map: An image of the distribution of speeds and directions of stars or gas in a galaxy.

Very Energetic Radiation Imaging Telescope Array System (VERITAS): Located in Arizona, this is an imaging atmospheric Čerenkov array of four telescopes used for gamma-ray astronomy in the 100-GeV to 10-TeV energy range.

Very Large Telescope (VLT): An array of four 8.2-m-diameter optical-IR telescopes operated by European Southern Observatory (ESO) at the Paranal Observatory in the Atacama Desert of northern Chile. The telescopes can be used independently or combined into an interferometer.

Very Long Baseline Array (VLBA): An array of 10 radio telescopes, operated by the U.S. National Radio Astronomy Observatory, which functions as the world's largest dedicated, full-time astronomical instrument. By using the technique of very long baseline interferometry, the VLBA attains extremely high angular resolution.

Virial radius: The radius of a sphere around a galaxy or cluster within which virial equilibrium holds (a balance between kinetic and potential energy).

Visible and Infrared Survey Telescope for Astronomy (VISTA): A 4-m-class, wide-field telescope located at the Paranal Observatory in Chile and currently operated by the European Southern Observatory.

Warm-hot intergalactic medium (WHIM): Gas in the intergalactic medium residing at sufficiently high temperature that hydrogen is fully ionized (very little neutral gas remains), and the gas is too diffuse for detectable X-ray emission; a nominal temperature range is ~10^5 to 10^7 K. This term sometimes refers to "missing baryons," since it is estimated theoretically that 30 to 50 percent of the cosmic baryons at the present epoch may reside in this phase, hidden from current observational capabilities.

Weak lensing: Fluctuations in the large-scale distribution of matter produce variations in the "shape" of space that alter the path that light takes from a distant galaxy to our telescopes. This gravitational lensing changes the shape of a galaxy and makes it appear more elliptical. By measuring the shapes of large numbers of galaxies, astronomers can detect this "weak lensing" effect and infer the large-scale distribution of matter.

Weakly interacting massive particle (WIMP): A broad class of hypothetical particles that have small interaction cross-sections with ordinary matter and radiation. Postulated WIMPs include the lightest supersymmetric particle. WIMPs are suggested as a possible explanation of dark matter.

White dwarf: The compact remnant of an evolved star from approximately 0.1

to 8 solar masses. The star is supported against gravitational collapse by electron degeneracy pressure.

XMM-Newton: ESA's X-ray Multi-Mirror mission, named after Sir Isaac Newton, is an orbiting X-ray observatory launched in 1999. It features a large collecting area for X-ray spectroscopy.

X-ray: Light with wavelength considerably shorter than visible light, roughly ~100 Å to ~0.01 Å, often referred to by the equivalent energy, 0.1 keV to 1 MeV, or frequency range log ν ~16.5 to 20.5 Hz.

X-ray polarimetry: The measurement and interpretation of the polarization of transverse electromagnetic waves, at X-ray wavelengths. The main physical mechanisms producing polarization in the X-ray band are thought to be synchrotron radiation (primarily from blazars), Compton scattering from accretion disks, radiative transfer in strong magnetic fields, and geometrical effects due to light bending in strong gravitational fields.

X-ray spectroscopy: The study of the distribution of X-ray intensity as a function of energy, used to analyze the characteristics of the emitted radiation and chemical composition of the gas.

Zodiacal light: Scattered light from dust particles orbiting the Sun, the dominant source of mid-infrared sky brightness seen from Earth.

Zody or zodi: See *Zodiacal light.*

C

Acronyms

2MASS—Two-Micron All-Sky Survey
A3IV—Aperture Array Astronomical Imaging Verification, European SKA-mid demonstrator
AAAC—Astronomy and Astrophysics Advisory Committee
AAG—Astronomy and Astrophysics Research Grants, an NSF program
AANM—*Astronomy and Astrophysics in the New Millennium*, the 2001 NRC decadal survey
ACA—Atacama Compact Array, an array of smaller dishes that complement the main array of ALMA
ACT—Atacama Cosmology Telescope
AGASA—Akeno Giant Air Shower Array
AGB—Asymptotic giant branch
AGIS—Advanced Gamma-ray Imaging System
AGN—Active galactic nuclei
AGS—NSF Division of Atmospheric and Geospace Sciences (formerly ATM)
AIC—Accretion-induced collapse
ALFA—Multi-receiver system for H I studies at Arecibo
ALMA—Atacama Large Millimeter/submillimeter Array
ALTAIR—*Access to Large Telescopes for Astronomical Instruction and Research*, a 2009 NOAO report
AMS—Alpha Magnetic Spectrometer
ANITA—Antarctic Impulsive Transient Antenna
AO—Adaptive optics

AODP—Adaptive optics development program
APEX—Atacama Pathfinder Experiment, 12-m telescope at ALMA site run by Max-Planck-Institut für Radioastronomie, ESO, and Onsala Space Observatory
ARO—Arizona Radio Observatory
AS—Academia Sinica, Taiwan
ASIAA—Academica Sinica Institute of Astronomy and Astrophysics
ASKAP—Australian SKA Pathfinder
AST—The Division of Astronomical Sciences in the MPS Directorate at NSF
ASTE—Atacama Submillimeter Telescope Experiment, 10-m telescope at ALMA site run by the National Astronomical Observatory of Japan (NAOJ) in collaboration with universities in Japan and Chile
ATA—Allen Telescope Array
ATA-42—Current version of ATA with 42 antennas
ATA-256—Allen Telescope Array with 256 antennas
ATI—Advanced Technologies and Instrumentation program at NSF-AST
ATIC—Advanced Thin Ionization Calorimeter
ATST—Advanced Technology Solar Telescope
AUI—Associated Universities, Inc.
AURA—Association of Universities for Research in Astronomy

BAO—Baryon acoustic oscillation
BBH—Binary black hole
BBN—Big bang nucleosynthesis
BESS—Balloon-borne Experiment with Superconducting Spectrometer
BIMA—Berkeley-Illinois-Maryland Association
BLISS—Background-Limited Infrared-Submillimeter Spectrograph
BSRBS—NRAO's Green Bank Solar Radio Burst Spectrometer

CALISTO—Cryogenic Aperture Large Infrared Space Telescope Observatory
CARA—Center for Atmospheric Research in Antarctica
CARMA—Combined Array for Research in Millimeter-wave Astronomy
CASA—Common Astronomy Software Applications, a data reduction package led by NRAO for EVLA and ALMA use
CATE—Cost appraisal and technical evaluation
CCAT—Formerly the Cornell-Caltech Atacama Telescope
CCD—Charge-coupled device
CDMS—Cryogenic Dark Matter Search
CFP—Cosmology and Fundamental Physics, an Astro2010 Science Frontiers Panel
CGM—Circumgalactic medium

Chandra—X-ray telescope in space, a NASA mission
CHARA—Center for High Angular Resolution Astronomy, a six-telescope optical-infrared interferometric array
CI—neutral carbon
CLIO—A prototype, underground, cryogenic gravitational-wave detector in Japan
CMB—Cosmic microwave background
CMBPol—Concept for space mission to study the polarization of the CMB
CMD—Color-magnitude diagram
CMF—Core mass function
CNES—Centre National d'Études Spatiales
COBE—Cosmic Background Explorer
COROT—Convection Rotation and Planetary Transits satellite
COS—Cosmic Origins Spectrograph on HST
COSMOS—Cosmological Evolution Survey
COUPP—Chicagoland Observatory for Underground Particle Physics
CP—Charge-parity symmetry
CREAM—Cosmic Ray Energetics and Mass
CSA—Canadian Space Agency
CSO—Caltech Submillimeter Observatory
CTA—Čerenkov Telescope Array
CV—Cataclysmic variable
CXC—Chandra X-ray Observatory Center

DAMA/LIBRA—Dark Matter Experiment/Large Sodium Iodide Bulk for Rare Processes, a Gran Sasso detector
DARPA—Defense Advanced Research Projects Agency
DEEP2—Deep Extragalactic Evolutionary Probe 2
DEIMOS—Deep Imaging Multi-Object Spectrograph, a Keck Observatory spectrograph
DES—Dark Energy Survey
DM—Deformable mirror
DOE—Department of Energy
DRS—Disturbance Reduction System for LISA
DWFIR—Deep, Wide-Field IR Survey

EDGES—Experiment to Detect the Global EoR Step, an MIT project
EDELWEISS—Expérience pour Detecter Les Wimps En Site Souterrain
E-ELT—European Extremely Large Telescope project
EHT—Event Horizon Telescope
ELT—Extremely Large Telescope

EMIR—Near-infrared multiobject spectrograph for the Gran Telescopio de Canarias
EMP—Extremely metal poor
EMRI—Extreme mass ratio inspiral
EoR—Epoch of reionization
EOS—Elecromagnetic Observations from Space, an Astro2010 Program Prioritization Panel
ESA—European Space Agency
ESO—European Southern Observatory
EVLA—Enhanced Very Large Array
ExAO—Extreme adaptive optics
EXIST—Energetic X-ray Imaging Survey Telescope
ExoPTF—Exoplanet Task Force

FASR—Frequency-Agile Solar Radiotelescope
FAST—Five-hundred-meter-Aperture Spherical Telescope, under construction in Guizhou Province, China
FOV—Field of view
FRIB—Facility for Rare Isotope Beams
FUSE—Far Ultraviolet Spectroscopic Explorer
FWHM—Full-width, half-maximum

G—Gravitational coupling constant
GALEX—Galaxy Evolution Explorer
GAN—Galactic Neighborhood, an Astro2010 Science Frontiers Panel
GBM—Gamma-ray-burst monitor
GBT—Robert C. Byrd Green Bank Telescope
GCT—Galaxies Across Cosmic Time, an Astro2010 Science Frontiers Panel
GEMS—Gravity and Extreme Magnetism Small Explorer
GEO—NSF Geosciences Directorate; also, gravitational wave detector
GLAST—Gamma-ray Large Area Space Telescope
GMC—Giant molecular cloud
GMRT—Giant Metre-wave Radio Telescope, operated by the Tata Institute of Fundamental Research, India
GMT—Giant Magellan Telescope
GMTO—Giant Magellan Telescope Office
GO—Guest observer
GOODS-N—The northern deep field of the Great Observatories Origins Deep Survey
GPI—Gemini Planet Imager

GR—General relativity
GRB—Gamma-ray burst
GRS—Gravitational Reference System for LPF and LISA
GSMT—Giant Segmented Mirror Telescope
GUT—Grand Unified Theory
GZK—Greisen-Zatsepin-Kuzmin mechanism

H I—Neutral atomic hydrogen
H II—Ionized hydrogen
$h(z)$—Hubble parameter
HAWC—High-Altitude Water Čerenkov experiment
HERA—Hydrogen Epoch of Reionization Array
HerMES—Herschel Multi-tiered Extragalactic Survey
HESS—High Energy Stereoscopic System
HiRes—High Resolution Fly's Eye
HPRV—High-Precision Radial-Velocity Spectroscopy
H-R diagram—Hertzsprung-Russell diagram
HST—Hubble Space Telescope
HZ—Habitable zone (not to be confused with Hz [hertz], a frequency of one cycle per second)

ICM—Intracluster medium
IDECS—International Dark Energy Cosmology Survey
IDL—Interactive Data Language
IFU—Integral field unit
IGM—Intergalactic medium
IGW—Inflationary gravitational wave
IMF—Initial mass function, the distribution of stars over mass before any stars have died
IMS—Interferometry Measurement System for LPF and LISA
INAOE—Instituto Nacional de Astrofisica Optica y Electronica
IPAC—Infrared Processing and Analysis Center
IR—Infrared
IRAF—Image Reduction and Analysis Facility
IRAM—Institut de Radioastronomie Millimetrique
IRAS—Infrared Astronomy Satellite
IRDC—Infrared dark cloud
IRS—Infrared Spectrometer, on Spitzer
ISM—Interstellar medium
IXO—International X-ray Observatory

JAXA—Japan Aerospace Exploration Agency
JCMT—James Clerk Maxwell Telescope
JDEM—Joint Dark Energy Mission
JPL—Jet Propulsion Laboratory
JWST—James Webb Space Telescope

KAIT—Katzman Automatic Imaging Telescope
KBO—Kuiper belt object
KI—Keck Interferometer

ΛCDM—Lambda cold dark matter
LAGEOS—Laser Geodynamics (satellite)
LANL—Los Alamos National Laboratory
LAT—Large Area Telescope, an instrument on the Fermi satellite
L-band—An IEEE-designated band covering roughly 16 to 40 cm
LBT—Large Binocular Telescope
LBTI—Large Binocular Telescope Interferometer
LBV—Luminous blue variable
LCGT—Large Cryogenic Gravitational-wave Telescope, under consideration in Japan
LGS—Laser guide star
LHC—Large Hadron Collider
LIGO—Laser Interferometer Gravitational Wave Observatory
LISA—Laser Interferometer Space Antenna
LLR—Lunar Laser Ranging
LMC—Large Magellanic Cloud
LMT—Large Millimeter Telescope
LMXB—Low-mass X-ray binary
LOFAR—Low Frequency Array for radio astronomy, a low-frequency antenna array centered in the Netherlands
LOPES—LOFAR prototype station
LPF—LISA Pathfinder, a NASA mission
LSS—Large-scale structure
LSST—Large Synoptic Survey Telescope
LUX—Large Underground Xenon detector
LWA—Long Wavelength Array

MAGIC—Major Atmospheric Gamma-ray Imaging Čerenkov Telescope
MBH—Massive black hole
MCAO—Multiconjugate adaptive optics
MeerKAT—Meer Karoo Array Telescope, a SKA-mid pathfinder in South Africa

MERLIN—Multi-Element Radio Linked Interferometer Network
MHD—Magnetohydrodynamic
MIDEX—Mid-size Explorer mission, capped at $250 million, excluding launch
MIPS—Multiband Imaging Photometer for SIRTF, on Spitzer
MMT—Multi-Mirror Telescope
MMTO—MMT Office
MOS—Multi-Object Spectrograph
MOSFIRE—Multi-Object Spectrometer for Infrared Exploration
MPF—Microlensing Planet Finder
MPS—NSF directorate of Mathematical and Physical Sciences
MREFC—Major Research Equipment and Facilities Construction, an NSF program
MRI—Major Research Instrumentation, an NSF program
MSP—Millisecond pulsar
MSSM—Minimal supersymmetric standard model
MUSTANG—Multiplexed Squid TES Array at Ninety Gigahertz, a bolometer array on the GBT
MWA—Murchison Widefield Array, a low-frequency radio (at frequencies from 80 to 300 MHz) interferometer array to be built in Western Australia

NAA—North American Array
NAASC—North American ALMA Science Center
NAIC—National Astronomy and Ionospheric Center
NAOJ—National Astronomical Observatory of Japan
NASA—National Aeronautics and Space Administration
NEO—Near-Earth object
NFIRAOS—Narrow-Field Infrared Adaptive Optics System for the Thirty-Meter Telescope
NGST—Next-Generation Space Telescope
NINS—National Institutes of Natural Sciences, Japan
NIRCam—Near-Infrared Camera on JWST
NIRSpec—Near-Infrared Spectrograph IFU on JWST
NIRSS—Near-Infrared Sky Surveyor
NIST—National Institute of Standards and Technology, Department of Commerce
NOAO—National Optical Astronomy Observatory
NOvA—A neutrino-oscillation detector under construction in Minnesota
NRAO—National Radio Astronomy Observatory
NRC—National Research Council, USA; also, National Research Council of Canada
NRL—Naval Research Laboratory, Department of the Navy

NSC—National Science Council, Taiwan
NSF—National Science Foundation
NSO—National Solar Observatory
NST—New Solar Telescope
NuSTAR—Nuclear Spectroscopic Telescope Array, a NASA Small Explorer mission featuring an imaging hard-X-ray telescope, to be launched in 2011

OHEP—Office of High Energy Physics, at DOE
OIR—Optical-infrared; also, Optical and Infrared Astronomy from the Ground, an Astro2010 Program Prioritization Panel
OPP—Office of Polar Programs, NSF
OVRO—Owens Valley Radio Observatory
OVSA—New Jersey Institute of Technology's Owens Valley Solar Array

PACS—Photodetector Array Camera and Spectrometer, on Herschel
PAG—Particle Astrophysics and Gravitation, an Astro2010 Program Prioritization Panel
PAH—Polycyclic aromatic hydrocarbon
PAIRITEL—Peters Automated Infrared Imaging Telescope, which conducts robotic transient surveys
PAMELA—Payload for Antimatter Exploration and Light-nuclei Astrophysics
PAPER—Precision Array to Probe the Epoch of Reionization
PdBI—Plateau de Bure Interferometer, a millimeter wave array in France, operated by IRAM
PFI—Planet Formation Instrument for the proposed Thirty-Meter Telescope
PNGA—Particle, Nuclear, and Gravitational-Wave Astrophysics, a panel from AANM
PPN—Parametrized Post-Newtonian
PPP—Program Prioritization Panel
PSF—Planetary Systems and Star Formation, an Astro2010 Science Frontiers Panel; also, point-spread function
PTF—Palomar Transient Factory
Python—An interactive data-analysis language

QCD—Quantum chromodynamics
QSO—Quasi-stellar object

R&A—Research and analysis
R&D—Research and development
ReSTAR—*Renewing Small Telescopes for Astronomical Research*, a 2007 NOAO report

RFI—Radio-frequency interference
RGB—Red-giant branch
RICE—Radio Ice Čerenkov Experiment
RMS—Radio, millimeter, and submillimeter; also, Radio, Millimeter, and Submillimeter Astronomy from the Ground, an Astro2010 Program Prioritization Panel
ROSAT—Röntgen Satellite
RRAT—Rotating radio transient
RV—Radial velocity
RXTE—Rossi X-ray Timing Explorer

SAFIRE—Submillimeter and Far-Infrared Explorer
SAMURAI—Science of AGNs and Masers with Unprecedented Resolution in Astronomical Imaging
SCUBA—Submillimetre Common-User Bolometer Array, operated on the JCMT
SCUBA-2—Successor to SCUBA with many more detectors
SDO—Solar Dynamics Observatory
SDSS—Sloan Digital Sky Survey
SETI—Search for Extraterrestrial Intelligence
SF—Star formation
SFE—Star-formation efficiency
SFP—Science Frontiers Panel
SFR—Star-formation rate
SFSR—Single-field, slow-roll inflation
Sgr A*—Sagittarius A*, a radio source at the center of the Milky Way Galaxy
SHARC II—Submillimeter High Angular Resolution Camera II, at CSO
SIM—Space Interferometry Mission
SIRTF—Space Infrared Telescope Facility, now Spitzer
SKA—Square Kilometer Array
SLAC—Stanford Linear Accelerator Center
SMA—Submillimeter Array
SMARTS—Small and Moderate Aperture Research Telescope System
SMBH—Supermassive black hole
SMC—Small Magellanic Cloud
SMEX—Small-scale Explorer mission, costing less than $150 million, excluding launch
S/N—Signal-to-noise ratio
SNe—Supernovae
SOAR—Southern Astrophysical Research Telescope
SOFIA—Stratospheric Observatory for Infrared Astronomy
SOLIS—Synoptic Optical Long-term Investigations of the Sun

SOML—Steward Observatory Mirror Laboratory
SPHERE—Spectro-Polarimetric High-contrast Exoplanet Research, a VLT instrument
SPICA—Space Infared Telescope for Cosmology and Astrophysics
SPIRE—Spectral and Polarimetric Imaging Receiver, on Herschel
SPIRIT—Space Infrared Interferometric Telescope
SPST—South Pole Submillimeter-wave Telescope
SPT—South Pole Telescope
SSE—Stars and Stellar Evolution, an Astro2010 Science Frontiers Panel
STIS—Space Telescope Imaging Spectograph
STScI—Space Telescope Science Institute
SZ—Sunyaev-Zel'dovich
SZA—Sunyaev-Zel'dovich Array
SZE—Sunyaev-Zel'dovich effect, a distortion in the spectrum of the CMB caused by gas in clusters

TES—Transition-edge sensors
TIGER—Trans-Iron Galactic Element Recorder
TMT—Thirty-Meter Telescope
TPC—Time-projection chamber for the XENON100 experiment
TPF—Terrestrial Planet Finder
TRL—Technology readiness level
TSIP—Telescope System Instrumentation Program

UHECR—Ultrahigh-energy cosmic rays
ULDB—Ultralong-duration balloon
URO—University Radio Observatories (currently includes CSO, CARMA, UMASS, and the ATA)
USAF—U.S. Air Force
UV—Ultraviolet

VAO—Virtual Astronomical Observatory, the operational phase of the National Virtual Observatory
VERITAS—Very Energetic Radiation Imaging Telescope Array System, a cosmic ray detector
VIRGO—A European gravitational-wave detector
VISTA—Visible and Infrared Survey Telescope for Astronomy
VLA—Very Large Array
VLBA—Very Long Baseline Array
VLBI—Very Long Baseline Interferometry
VLT—ESO's Very Large Telescope

VO—Virtual Observatory
VSOP-2—A Japanese radio telescope satellite that could be used with ground-based VLBI antennas

$w(z)$—Dark energy equation-of-state parameter
WArP—WIMP Argon Program
WD—White dwarf
WFC-3—Wide Field Camera 3, a Hubble instrument
WFIRST—Wide-Field Infrared Survey Telescope
WFMOS—Wide-Field Fiber-Fed Optical MOS, a proposed Gemini instrument
WFS—Wave-front sensor
WHIM—Warm-hot intergalactic medium
WIMP—Weakly interacting massive particle
WISE—Wide-field Infrared Survey Explorer
WIYN—Wisconsin-Indiana-Yale-NOAO Telescope
WMAP—Wilkinson Microwave Anisotropy Probe

XENON100—A WIMP detector
XENON1T—A WIMP detector
XGS—X-ray grating spectrometer
XMASS—A WIMP detector
XMM-Newton—ESA's X-ray Multi-Mirror Mission
XRSO—X-ray spectroscopy observatory