

## SDSS-III: MASSIVE SPECTROSCOPIC SURVEYS OF THE DISTANT UNIVERSE, THE MILKY WAY GALAXY, AND EXTRA-SOLAR PLANETARY SYSTEMS

DANIEL J. EISENSTEIN<sup>1,2</sup>, DAVID H. WEINBERG<sup>3,4</sup>, ERIC AGOL<sup>5</sup>, HIROAKI AIHARA<sup>6</sup>, CARLOS ALLENDE PRIETO<sup>7,8</sup>, SCOTT F. ANDERSON<sup>5</sup>, JAMES A. ARNS<sup>9</sup>, ERIC AUBOURG<sup>10,11</sup>, STEPHEN BAILEY<sup>12</sup>, EDUARDO BALBINOT<sup>13,14</sup>, ROBERT BARKHOUSER<sup>15</sup>, TIMOTHY C. BEERS<sup>16</sup>, ANDREAS A. BERLIND<sup>17</sup>, STEVEN J. BICKERTON<sup>18</sup>, DMITRY BIZYAEV<sup>19</sup>, MICHAEL R. BLANTON<sup>20</sup>, JOHN J. BOCHANSKI<sup>21</sup>, ADAM S. BOLTON<sup>22</sup>, CASEY T. BOSMAN<sup>23</sup>, JO BOVY<sup>20</sup>, HOWARD J. BREWINGTON<sup>19</sup>, W. N. BRANDT<sup>21,24</sup>, BEN BRESLAUER<sup>25</sup>, J. BRINKMANN<sup>19</sup>, PETER J. BROWN<sup>22</sup>, JOEL R. BROWNSTEIN<sup>22</sup>, DAN BURGER<sup>17</sup>, NICOLAS G. BUSCA<sup>10</sup>, HEATHER CAMPBELL<sup>26</sup>, PHILLIP A. CARGILE<sup>17</sup>, WILLIAM C. CARITHERS<sup>12</sup>, JOLEEN K. CARLBERG<sup>25</sup>, MICHAEL A. CARR<sup>18</sup>, YANMEI CHEN<sup>27</sup>, CRISTINA CHIAPPINI<sup>28,29,14</sup>, JOHAN COMPARAT<sup>30</sup>, NATALIA CONNOLLY<sup>31</sup>, MARINA CORTES<sup>12</sup>, RUPERT A.C. CROFT<sup>32</sup>, LUIZ N. DA COSTA<sup>33,14</sup>, KATIA CUNHA<sup>1,34</sup>, JAMES R. A. DAVENPORT<sup>5</sup>, KYLE DAWSON<sup>22</sup>, NATHAN DE LEE<sup>23</sup>, GUSTAVO F. PORTO DE MELLO<sup>35,14</sup>, FERNANDO DE SIMONI<sup>33,14</sup>, JANICE DEAN<sup>25</sup>, SAURAV DHITAL<sup>17</sup>, ANNE EALET<sup>36</sup>, GARRETT L. EBELKE<sup>19,37</sup>, EDWARD M. EDMONDSON<sup>26</sup>, JACOB M. EITING<sup>38</sup>, STEPHANIE ESCOFFIER<sup>36</sup>, MASSIMILIANO ESPOSITO<sup>7,8</sup>, MICHAEL L. EVANS<sup>5</sup>, XIAOHUI FAN<sup>1</sup>, BRUNO FEMENÍA CASTELLÁ<sup>7,8</sup>, LETICIA DUTRA FERREIRA<sup>35,14</sup>, GREG FITZGERALD<sup>39</sup>, SCOTT W. FLEMING<sup>23</sup>, ANDREU FONT-RIBERA<sup>40</sup>, ERIC B. FORD<sup>23</sup>, PETER M. FRINCHABOY<sup>41</sup>, ANA ELIA GARCÍA PÉREZ<sup>25</sup>, B. SCOTT GAUDI<sup>3</sup>, JIAN GE<sup>23</sup>, LUAN GHEZZI<sup>33,14</sup>, BRUCE A. GILLESPIE<sup>19</sup>, G. GILMORE<sup>42</sup>, LÉO GIRARDI<sup>43,14</sup>, J. RICHARD GOTT<sup>18</sup>, ANDREW GOULD<sup>3</sup>, EVA K. GREBEL<sup>44</sup>, JAMES E. GUNN<sup>18</sup>, JEAN-CHRISTOPHE HAMILTON<sup>10</sup>, PAUL HARDING<sup>45</sup>, DAVID W. HARRIS<sup>22</sup>, SUZANNE L. HAWLEY<sup>5</sup>, FREDERICK R. HEARTY<sup>25</sup>, JONAY I. GONZALEZ HERNANDEZ<sup>7</sup>, SHIRLEY HO<sup>12</sup>, DAVID W. HOGG<sup>20</sup>, JON A. HOLTZMAN<sup>37</sup>, KLAUS HONSCHEID<sup>38,4</sup>, NAOHISA INADA<sup>46</sup>, INESE I. IVANS<sup>22</sup>, LINHUA JIANG<sup>1</sup>, PENG JIANG<sup>23,47</sup>, JENNIFER A. JOHNSON<sup>3,4</sup>, CATHY JORDAN<sup>19</sup>, WENDELL P. JORDAN<sup>19,37</sup>, GUINEVERE KAUFFMANN<sup>48</sup>, EYAL KAZIN<sup>20</sup>, DAVID KIRKBY<sup>49</sup>, MARK A. KLAENE<sup>19</sup>, JEAN-PAUL KNEIB<sup>30</sup>, G. R. KNAPP<sup>18</sup>, C. S. KOCHANÉK<sup>3,4</sup>, LARS KOESTERKE<sup>50</sup>, JUNA A. KOLLMEIER<sup>51</sup>, RICHARD G. KRON<sup>52,53</sup>, DUSTIN LANG<sup>18</sup>, JAMES E. LAWLER<sup>54</sup>, JEAN-MARC LE GOFF<sup>11</sup>, BRIAN L. LEE<sup>23</sup>, YOUNG SUN LEE<sup>16</sup>, JARRON M. LEISENRING<sup>25</sup>, YEN-TING LIN<sup>6,55</sup>, JIAN LIU<sup>23</sup>, DANIEL C. LONG<sup>19</sup>, CRAIG P. LOOMIS<sup>18</sup>, SARA LUCATELLO<sup>43</sup>, BRITT LUNDGREN<sup>56</sup>, ROBERT H. LUPTON<sup>18</sup>, BO MA<sup>23</sup>, ZHIBO MA<sup>45</sup>, NICHOLAS MACDONALD<sup>5</sup>, CLAUDE MACK<sup>17</sup>, SUVRATH MAHADEVAN<sup>21,57</sup>, MARCIO A.G. MAIA<sup>33,14</sup>, ELENA MALANUSHENKO<sup>19</sup>, VIKTOR MALANUSHENKO<sup>19</sup>, STEVEN R. MAJEWSKI<sup>25</sup>, MARTIN MAKLER<sup>58,14</sup>, RACHEL MANDELBAUM<sup>18</sup>, CLAUDIA MARASTON<sup>26</sup>, DANIEL MARGALA<sup>49</sup>, PAUL MASEMAN<sup>1,25</sup>, KAREN L. MASTERS<sup>26</sup>, CAMERON K. MCBRIDE<sup>17</sup>, PATRICK McDONALD<sup>12,59</sup>, IAN D. MCGREER<sup>1</sup>, RICHARD G. MCMAHON<sup>42</sup>, OLGA MENA REQUEJO<sup>60</sup>, BRICE MÉNARD<sup>61,15</sup>, JORDI MIRALDA-ESCUDE<sup>62,63</sup>, HEATHER L. MORRISON<sup>45</sup>, FERGAL MULLALLY<sup>18,64</sup>, DEMITRI MUNA<sup>20</sup>, HITOSHI MURAYAMA<sup>6</sup>, ADAM D. MYERS<sup>65</sup>, TRACY NAUGLE<sup>19</sup>, ANGELO FAUSTI NETO<sup>13,14</sup>, DUY CUONG NGUYEN<sup>23</sup>, ROBERT C. NICHOL<sup>26</sup>, DAVID L. NIDEVER<sup>25</sup>, ROBERT W. O'CONNELL<sup>25</sup>, RICARDO L. C. OGANDO<sup>33,14</sup>, MATTHEW D. OLMSTEAD<sup>22</sup>, DANIEL J. ORAVETZ<sup>19</sup>, NIKHIL PADMANABHAN<sup>56</sup>, MARTIN PAEGERT<sup>17</sup>, NATHALIE PALANQUE-DELABROUILLE<sup>11</sup>, KAIKE PAN<sup>19</sup>, PARUL PANDEY<sup>22</sup>, JOHN K. PAREJKO<sup>56</sup>, ISABELLE PÂRIS<sup>66</sup>, PAULO PELLEGRINI<sup>14</sup>, JOSHUA PEPPER<sup>17</sup>, WILL J. PERCIVAL<sup>26</sup>, PATRICK PETITJEAN<sup>66</sup>, ROBERT PFAFFENBERGER<sup>37</sup>, JANINE PFORR<sup>26</sup>, STEFANIE PHLEPS<sup>67</sup>, CHRISTOPHE PICHON<sup>66</sup>, MATTHEW M. PIERI<sup>68,3</sup>, FRANCISCO PRADA<sup>69</sup>, ADRIAN M. PRICE-WHELAN<sup>20</sup>, M. JORDAN RADDICK<sup>15</sup>, BEATRIZ H. F. RAMOS<sup>33,14</sup>, CELINE RYLE<sup>70</sup>, I. NEILL REID<sup>71</sup>, JAMES RICH<sup>11</sup>, GORDON T. RICHARDS<sup>72</sup>, GEORGE H. RIEKE<sup>1</sup>, MARCIA J. RIEKE<sup>1</sup>, HANS-WALTER RIX<sup>73</sup>, ANNIE C. ROBIN<sup>70</sup>, HELIO J. ROCHA-PINTO<sup>35,14</sup>, CONSTANCE M. ROCKOSI<sup>74</sup>, NATALIE A. ROE<sup>12</sup>, EMMANUEL ROLLINDE<sup>66</sup>, ASHLEY J. ROSS<sup>26</sup>, NICHOLAS P. ROSS<sup>12</sup>, BRUNO ROSSETTO<sup>35,14</sup>, ARIEL G. SÁNCHEZ<sup>67</sup>, BASILIO SANTIAGO<sup>13,14</sup>, CONOR SAYRES<sup>5</sup>, RICARDO SCHIAVON<sup>75</sup>, DAVID J. SCHLEGEL<sup>12</sup>, KATHARINE J. SCHLESINGER<sup>3</sup>, SARAH J. SCHMIDT<sup>5</sup>, DONALD P. SCHNEIDER<sup>21,57</sup>, KRIS SELLGREN<sup>3</sup>, ALAINA SHELDEN<sup>19</sup>, ERIN SHELDON<sup>59</sup>, MATTHEW SHETRONE<sup>76</sup>, YIPING SHU<sup>22</sup>, JOHN D. SILVERMAN<sup>6</sup>, JENNIFER SIMMERER<sup>22</sup>, AUDREY E. SIMMONS<sup>19</sup>, THIRUPATHI SIVARANI<sup>23,77</sup>, M. F. SKRUTSKIE<sup>25</sup>, ANŽE SLOSAR<sup>59</sup>, STEPHEN SMEE<sup>15</sup>, VERNE V. SMITH<sup>34</sup>, STEPHANIE A. SNEDDEN<sup>19</sup>, KEIVAN G. STASSUN<sup>17,78</sup>, OLIVER STEELE<sup>26</sup>, MATTHIAS STEINMETZ<sup>28</sup>, MARK H. STOCKETT<sup>54</sup>, TODD STOLLBERG<sup>39</sup>, MICHAEL A. STRAUSS<sup>18</sup>, MASAYUKI TANAKA<sup>6</sup>, ANIRUDDHA R. THAKAR<sup>15</sup>, DANIEL THOMAS<sup>26</sup>, JEREMY L. TINKER<sup>20</sup>, BENJAMIN M. TOFFLEMIRE<sup>5</sup>, RITA TOJEIRO<sup>26</sup>, CHRISTY A. TREMONTI<sup>27</sup>, MARIANA VARGAS MAGAÑA<sup>10</sup>, LICIA VERDE<sup>62,63</sup>, NICOLE P. VOGT<sup>37</sup>, DAVID A. WAKE<sup>56</sup>, XIAOKE WAN<sup>23</sup>, JI WANG<sup>23</sup>, BENJAMIN A. WEAVER<sup>20</sup>, MARTIN WHITE<sup>79</sup>, SIMON D. M. WHITE<sup>48</sup>, JOHN C. WILSON<sup>25</sup>, JOHN P. WISNIEWSKI<sup>5</sup>, W. MICHAEL WOOD-VASEY<sup>80</sup>, BRIAN YANNY<sup>52</sup>, NAOKI YASUDA<sup>6</sup>, CHRISTOPHE YÈCHE<sup>11</sup>, DONALD G. YORK<sup>53,81</sup>, ERICK YOUNG<sup>1,82</sup>, GAIL ZASOWSKI<sup>25</sup>, IDIT ZEHAVI<sup>45</sup>, BO ZHAO<sup>23</sup>

Preprint Version, 1/10/11

## ABSTRACT

Building on the legacy of the Sloan Digital Sky Survey (SDSS-I and II), SDSS-III is a program of four spectroscopic surveys on three scientific themes: dark energy and cosmological parameters, the history and structure of the Milky Way, and the population of giant planets around other stars. The Baryon Oscillation Spectroscopic Survey (BOSS) will measure redshifts of 1.5 million massive galaxies and Ly $\alpha$  forest spectra of 150,000 quasars, using the baryon acoustic oscillation (BAO) feature of large scale structure to obtain percent-level determinations of the distance scale and Hubble expansion rate at  $z < 0.7$  and at  $z \approx 2.5$ . SEGUE-2, a now-completed continuation of the Sloan Extension for Galactic Understanding and Exploration, measured medium-resolution ( $R = \lambda/\Delta\lambda \approx 1800$ ) optical spectra of 118,000 stars in a variety of target categories, probing chemical evolution, stellar kinematics and substructure, and the mass profile of the dark matter halo from the solar neighborhood to distances of 100 kpc. APOGEE, the Apache Point Observatory Galactic Evolution Experiment, will obtain high-resolution ( $R \approx 30,000$ ), high signal-to-noise ratio ( $S/N \geq 100$  per resolution element),  $H$ -band ( $1.51\mu\text{m} < \lambda < 1.70\mu\text{m}$ ) spectra of  $10^5$  evolved, late-type stars, measuring separate abundances for

$\sim 15$  elements per star and creating the first high-precision spectroscopic survey of *all* Galactic stellar populations (bulge, bar, disks, halo) with a uniform set of stellar tracers and spectral diagnostics. The Multi-object APO Radial Velocity Large-area Survey (MARVELS) will monitor radial velocities of more than 8000 FGK stars with the sensitivity and cadence ( $10 - 40 \text{ m s}^{-1}$ ,  $\sim 24$  visits per star) needed to detect giant planets with periods up to two years, providing an unprecedented data set for understanding the formation and dynamical evolution of giant planet systems. As of January 2011, SDSS-III has obtained spectra of more than 240,000 galaxies, 29,000  $z \geq 2.2$  quasars, and 140,000 stars, including 74,000 velocity measurements of 2580 stars for MARVELS. In keeping with SDSS tradition, SDSS-III will provide regular public releases of all its data, beginning with SDSS Data Release 8 (DR8) in January 2011.

*Subject headings:* surveys, cosmology: observations, Galaxy: evolution, planets and satellites: detection

<sup>1</sup> Steward Observatory, 933 North Cherry Avenue, Tucson, AZ 85721

<sup>2</sup> Harvard College Observatory, 60 Garden St., Cambridge, MA 02138

<sup>3</sup> Department of Astronomy, Ohio State University, Columbus, OH 43210

<sup>4</sup> Center for Cosmology and Astro-Particle Physics, Ohio State University, Columbus, OH 43210

<sup>5</sup> Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195

<sup>6</sup> Institute for the Physics and Mathematics of the Universe, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, 277-8583, Japan

<sup>7</sup> Instituto de Astrofísica de Canarias, E38205 La Laguna, Tenerife, Spain

<sup>8</sup> Departamento de Astrofísica, Universidad de La Laguna, 38206, La Laguna, Tenerife, Spain

<sup>9</sup> Kaiser Optical Systems, 371 Parkland Plaza Ann Arbor, MI 48103

<sup>10</sup> Astroparticule et Cosmologie (APC), Université Paris-Diderot, 10 rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France

<sup>11</sup> CEA, Centre de Saclay, Irfu/SPP, F-91191 Gif-sur-Yvette, France

<sup>12</sup> Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720

<sup>13</sup> Instituto de Física, UFRGS, Caixa Postal 15051, Porto Alegre, RS - 91501-970, Brazil

<sup>14</sup> Laboratório Interinstitucional de e-Astronomia, - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

<sup>15</sup> Center for Astrophysical Sciences, Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218

<sup>16</sup> Department of Physics & Astronomy and JINA: Joint Institute for Nuclear Astrophysics, Michigan State University, E. Lansing, MI 48824

<sup>17</sup> Department of Physics and Astronomy, Vanderbilt University, Nashville TN 37235

<sup>18</sup> Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544

<sup>19</sup> Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349

<sup>20</sup> Center for Cosmology and Particle Physics, New York University, 4 Washington Place, New York, NY 10003

<sup>21</sup> Department of Astronomy and Astrophysics, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802

<sup>22</sup> Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112

<sup>23</sup> Department of Astronomy, University of Florida, Bryant Space Science Center, Gainesville, FL 32611-2055

<sup>24</sup> Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802

<sup>25</sup> Department of Astronomy University of Virginia P.O.Box 400325 Charlottesville, VA 22904-4325

<sup>26</sup> Institute of Cosmology and Gravitation (ICG), Dennis Sciamia Building, Burnaby Road Univ. of Portsmouth, Portsmouth, PO1 3FX, UK

<sup>27</sup> Department of Astronomy, University of Wisconsin-Madison, 475 N. Charter St., Madison, WI 53706-1582

<sup>28</sup> Astrophysical Institute Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany

<sup>29</sup> 3- Istituto Nazionale di Astrofisica - OATrieste Via G. B. Tiepolo 11 34143 Italia

<sup>30</sup> Laboratoire d'Astrophysique de Marseille, CNRS-Université de Provence, 38 rue F. Joliot-Curie, 13388 Marseille cedex 13, France

<sup>31</sup> Department of Physics, Hamilton College, Clinton, NY 13323

<sup>32</sup> Bruce and Astrid McWilliams Center for Cosmology, Carnegie Mellon University, Pittsburgh, PA 15213

<sup>33</sup> Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

<sup>34</sup> National Optical Astronomy Observatory, 950 N. Cherry Ave., Tucson, AZ 85719

<sup>35</sup> Universidade Federal do Rio de Janeiro, Observatório do Valongo, Ladeira do Pedro Antônio 43, 20080-090 Rio de Janeiro, Brazil

<sup>36</sup> Centre de Physique des Particules de Marseille, Aix-Marseille Université CNRS/IN2P3, Marseille, France

<sup>37</sup> Department of Astronomy, MSC 4500, New Mexico State University, P.O. Box 30001, Las Cruces, NM 88003

<sup>38</sup> Department of Physics, Ohio State University, Columbus, OH 43210

<sup>39</sup> New England Optical Systems, Marlborough, MA 01752

<sup>40</sup> Institut de Ciències de l'Espai (CSIC-IEEC), Campus UAB, 08193 Bellaterra, Barcelona, Spain

<sup>41</sup> Dept. of Physics & Astronomy, Texas Christian University, 2800 South University Dr., Fort Worth, TX 76129

<sup>42</sup> Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK

<sup>43</sup> Osservatorio Astronomico di Padova - INAF, Vicolo dell'Osservatorio 5, I-35122 Padova, Italy

<sup>44</sup> Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstr. 12-14, 69120 Heidelberg, Germany

<sup>45</sup> Department of Astronomy, Case Western Reserve University, Cleveland, OH 44106

<sup>46</sup> Research Center for the Early Universe, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan

<sup>47</sup> Key Laboratory for Research in Galaxies and Cosmology, The University of Science and Technology of China, Chinese Academy of Sciences, Hefei, Anhui 230026, China

<sup>48</sup> Max-Planck-Institut für Astrophysik, Postfach 1, D-85748 Garching, Germany

<sup>49</sup> Department of Physics and Astronomy, University of California, Irvine, CA 92697

<sup>50</sup> Texas Advanced Computer Center, University of Texas, 10100 Burnet Road (R8700), Austin, Texas 78758-4497

<sup>51</sup> Observatories of the Carnegie Institution of Washington, 813 Santa Barbara St., Pasadena, CA 91101

<sup>52</sup> Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510

<sup>53</sup> Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637

<sup>54</sup> University of Wisconsin, Department of Physics, 1150 University Ave., Madison, WI 53706

<sup>55</sup> Institute of Astronomy and Astrophysics, Academia Sinica, Taipei 10617, Taiwan

<sup>56</sup> Yale Center for Astronomy and Astrophysics, Yale Univer-

## 1. INTRODUCTION

The Sloan Digital Sky Survey (SDSS; York et al. 2000) and the Legacy Survey of SDSS-II performed deep imaging of 8400 deg<sup>2</sup> of high Galactic latitude sky in five optical bands, repeat imaging of an equatorial stripe in the southern Galactic cap (roughly 25 epochs on 300 deg<sup>2</sup>), and spectroscopy of more than 900,000 galaxies, 100,000 quasars, and 200,000 stars (Abazajian et al. 2009). In addition to completing the original SDSS goals, SDSS-II (which operated from 2005-2008) executed a supernova survey in the southern equatorial stripe (Frieman et al. 2008b), discovering more than 500 spectroscopically confirmed Type Ia supernovae in the redshift range  $0.1 < z < 0.4$ , and it also performed an imaging and spectroscopic survey of the Galaxy, known as SEGUE (the Sloan Extension for Galactic Understanding and Exploration; Yanny et al. 2009), with 3200 deg<sup>2</sup> of additional imaging and spectra of 240,000 stars selected in a variety of target categories. These surveys were accomplished

sity, New Haven, CT 06520

<sup>57</sup> Center for Exoplanets and Habitable Worlds, 525 Davey Laboratory, Pennsylvania State University, University Park, PA 16802

<sup>58</sup> ICRA - Centro Brasileiro de Pesquisas Físicas, Rua Dr. Xavier Sigaud 150, Urca, Rio de Janeiro, RJ - 22290-180, Brazil  
<sup>59</sup> Bldg 510 Brookhaven National Laboratory, Physics Department, Upton NY, 11973

<sup>60</sup> Instituto de Física Corpuscular IFIC/CSIC, Universidad de Valencia, Valencia, Spain

<sup>61</sup> CITA, University of Toronto, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 3H8, Canada

<sup>62</sup> Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

<sup>63</sup> Institut de Ciències del Cosmos, Universitat de Barcelona/IEEC, Barcelona 08028, Spain

<sup>64</sup> SETI Institute/NASA Ames Research Center, Moffett Field, CA 94035

<sup>65</sup> Department of Astronomy, University of Illinois, 1002 West Green Street, Urbana, IL 61801.

<sup>66</sup> Université Paris 6, Institut d'Astrophysique de Paris, UMR7095-CNRS, 98bis Boulevard Arago, F-75014, Paris - France

<sup>67</sup> Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany

<sup>68</sup> CASA, University of Colorado, 389 UCB, Boulder, Colorado 80309

<sup>69</sup> Instituto de Astrofísica de Andalucía (CSIC), E-18008, Granada, Spain

<sup>70</sup> Institut Utinam, Observatoire de Besançon, Université de Franche-Comté, BP1615, F-25010 Besançon cedex, France

<sup>71</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218

<sup>72</sup> Department of Physics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104

<sup>73</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

<sup>74</sup> UCO/Lick Observatory, University of California, Santa Cruz, 1156 High St. Santa Cruz, CA 95064

<sup>75</sup> Gemini Observatory, 670 N. A'ohoku Place, Hilo, HI 96720

<sup>76</sup> University of Texas at Austin, McDonald Observatory 32 Fowlkes Rd McDonald Observatory, TX 79734

<sup>77</sup> Indian Institute of Astrophysics, II Block, Koramangala, Bangalore 560 034, India

<sup>78</sup> Department of Physics, Fisk University, 1000 17th Ave. N., Nashville, TN

<sup>79</sup> Physics Department, University of California, Berkeley, CA 94720

<sup>80</sup> Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA, 15260

<sup>81</sup> Enrico Fermi Institute, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637

<sup>82</sup> SOFIA Science Center/ USRA, NASA Ames Research Center, MS 211-3, Moffett Field, CA 94035

using a dedicated 2.5-m telescope<sup>83</sup> with a wide field of view (7 deg<sup>2</sup>, 3°-diameter; Gunn et al. 2006), a large mosaic CCD camera (Gunn et al. 1998), a pair of double spectrographs, each fed by 320 optical fibers plugged into custom-drilled aluminum plates, and an extensive network of data reduction and calibration pipelines and data archiving systems. The resulting data sets have supported an enormous range of investigations, making the SDSS one of the most influential astronomical projects of recent decades (Madrid & Macchetto 2006, 2009).

The achievements of SDSS-I and II and the exceptional power of the SDSS facilities for wide-field spectroscopy together inspired SDSS-III, a six-year program begun in July 2008 and consisting of four large spectroscopic surveys on three scientific themes: dark energy and cosmological parameters, the history and structure of the Milky Way, and the population of giant planets around other stars. This paper provides an overview of the four SDSS-III surveys, each of which will be described in greater depth by one or more future publications covering survey strategy, instrumentation, and data reduction software.

The Baryon Oscillation Spectroscopic Survey (BOSS) is the primary dark-time survey of SDSS-III. It aims to determine the expansion history of the universe with high precision by using the baryon acoustic oscillation (BAO) feature in large-scale structure as a standard ruler for measuring cosmological distances (Eisenstein & Hu 1998; Blake & Glazebrook 2003; Seo & Eisenstein 2003). More specifically, the BOSS redshift survey of 1.5 million massive galaxies aims to measure the distance-redshift relation  $d_A(z)$  and the Hubble parameter  $H(z)$  with percent-level precision out to  $z = 0.7$ , using the well established techniques that led to the first detections of the BAO feature (Cole et al. 2005; Eisenstein et al. 2005). Pioneering a new method of BAO measurement, BOSS will devote 20% of its fibers to obtaining Ly $\alpha$  forest absorption spectra of 150,000 distant quasars, achieving the first precision measurements of cosmic expansion at high redshift ( $z \approx 2.5$ ) and serving as a pathfinder for future surveys employing this technique. BOSS is also performing spectroscopic surveys of approximately 75,000 ancillary science targets in a variety of categories. To enable BOSS to cover 10,000 deg<sup>2</sup>, the SDSS imaging camera was used at the start of SDSS-III to survey an additional 2500 deg<sup>2</sup> of high-latitude sky in the southern Galactic cap; this imaging was completed in January 2010. Because BOSS was designed to observe targets 1 – 2 magnitudes fainter than the original SDSS spectroscopic targets, substantial upgrades to the SDSS spectrographs were required. The upgraded spectrographs were commissioned in Fall 2009. As of early January 2011, BOSS had obtained 240,000 galaxy spectra and 29,000 high-redshift ( $z \geq 2.2$ ) quasar spectra.

From July 2008 to July 2009, SDSS-III undertook a spectroscopic survey of 118,000 stars in a variety of target categories, using the original SDSS spectrographs. This survey, called SEGUE-2, is similar in design to the SEGUE-1 spectroscopic survey of SDSS-II, but it used the results of SEGUE-1 to refine its target selection algorithms.<sup>84</sup> While SEGUE-1 included both deep and

<sup>83</sup> The Sloan Foundation 2.5-m Telescope at Apache Point Observatory (APO), in Sunspot, NM.

<sup>84</sup> For clarity, we will henceforth use the retrospective term

shallow spectroscopic pointings, SEGUE-2 obtained only deep pointings to better sample the outer halo, which is the primary reason SEGUE-2 observed fewer stars than SEGUE. Together, the SEGUE-1 and SEGUE-2 surveys comprise 358,000 stars observed along a grid of sightlines totaling 2500 deg<sup>2</sup>, with spectral resolution  $R \equiv \lambda/\Delta\lambda \approx 1800$  spanning  $3800 \text{ \AA} < \lambda < 9200 \text{ \AA}$  (where  $\Delta\lambda$  is the FWHM of the line-spread function). Typical parameter measurement errors are  $5 - 10 \text{ km s}^{-1}$  in radial velocity,  $100 - 200 \text{ K}$  in  $T_{\text{eff}}$ , and  $0.21 \text{ dex}$  in  $[\text{Fe}/\text{H}]$ , depending on signal-to-noise ratio and stellar type (see §3). These data allow unique constraints on the stellar populations and assembly history of the outer Galaxy and on the mass profile of the Galaxy’s dark matter halo. SEGUE-2 observations are now complete.

SDSS-III also includes two bright-time surveys, generally performed when the moon is above the horizon and the lunar phase is more than 70 degrees from new moon. The first of these is the Multi-object APO Radial Velocity Exoplanet Large-area Survey (MARVELS), which uses fiber-fed, dispersed fixed-delay interferometer spectrographs (Erskine & Ge 2000; Ge 2002; Ge, Erskine, & Rushford 2002; van Eyken, Ge, & Mahadevan 2010) to monitor stellar radial velocities and detect the periodic perturbations caused by orbiting giant planets. MARVELS aims to monitor 8,400 F, G, and K stars in the magnitude range  $V = 8 - 12$ , observing each star  $\sim 24$  times over a 2–4 year interval to a typical photon-limited velocity precision per observation of  $8 \text{ m s}^{-1}$  at  $V = 9$ ,  $17 \text{ m s}^{-1}$  at  $V = 10$ , and  $27 \text{ m s}^{-1}$  at  $V = 11$ , with the goal of achieving total errors within a factor of 1.3 of the photon noise. These observations will provide a large and well characterized statistical sample of giant planets in the period regime needed to understand the mechanisms of orbital migration and planet-planet scattering, as well as rare systems that would escape detection in smaller surveys. MARVELS began operations in Fall 2008 with a 60-fiber instrument, which we hope to supplement with a second 60-fiber instrument for the second half of the survey. As of January 2011, it has obtained more than 74,000 radial velocity measurements of 2580 stars.

The Apache Point Observatory Galactic Evolution Experiment (APOGEE) will undertake an  $H$ -band ( $1.51 - 1.70 \mu\text{m}$ ) spectroscopic survey of  $10^5$  evolved late-type stars spanning the Galactic disk, bulge, and halo, with a typical limiting (Vega-based) magnitude of  $H \approx 12.5$  per field. Near-IR spectroscopy allows observations even in regions of high dust extinction, which will allow APOGEE to survey the same kinds of tracer stars (i.e., giants/supergiants) in all regions of the Galaxy. APOGEE spectra will have resolution  $R \approx 30,000$ , roughly 15 times that of SEGUE-2, and will achieve a signal-to-noise ratio  $S/N \gtrsim 100$  per resolution element for most stars. These spectra will enable detailed chemical fingerprinting of each individual program star, typically with 0.1-dex measurement precision for  $\sim 15$  chemical elements that trace different nucleosynthetic pathways and thus different populations of progenitor stars. Once APOGEE begins operations, MARVELS

“SEGUE-1” to refer to the SEGUE survey conducted in SDSS-II, and we will use “SEGUE” to refer to the two surveys generically or collectively.

and APOGEE will usually observe simultaneously, sharing the focal plane with fibers directed to different instruments, although this will not be practical in all fields. APOGEE will use a 300-fiber, cryogenic spectrograph that is currently under construction and scheduled for delivery to APO in early 2011.

SDSS-III will continue the SDSS tradition of releasing all data to the astronomical community and the public, including calibrated images and spectra and catalogs of objects with measured parameters, accompanied by powerful database tools that allow efficient exploration of the data and scientific analysis (Abazajian et al. 2009). These public data releases will be numbered consecutively with those of SDSS-I and II; the first is Data Release 8 (DR8; Aihara et al. 2011) in January 2011, occurring simultaneously with the submission of this paper. To enable homogeneous analyses that span SDSS-I, II, and III, DR8 includes essentially all SDSS-I/II imaging and spectra, processed with the latest data pipelines and calibrations. DR8 also includes all the new imaging data obtained for BOSS, and all SEGUE-2 data. DR9, currently scheduled for July 2012, will include BOSS spectra obtained through July 2011, new SEGUE stellar parameter determinations that incorporate ongoing pipeline and calibration improvements, and MARVELS radial velocity measurements obtained through December 2010. DR10, currently scheduled for July 2013, will include BOSS and APOGEE spectra obtained through July 2012. All data releases are cumulative. The final data release, currently scheduled for December 2014, will include all BOSS and APOGEE spectra and all MARVELS radial velocity measurements.

The four subsequent sections describe the individual surveys in greater detail. We provide a short overview of the technical and scientific organization of SDSS-III in §6 and some brief concluding remarks in §7.

## 2. BOSS

According to general relativity (hereafter GR), the gravity of dark matter, baryonic matter, and radiation should slow the expansion of the universe over time. Astronomers attempting to measure this deceleration using high-redshift Type Ia supernovae found instead that cosmic expansion is accelerating (Riess et al. 1998; Perlmutter et al. 1999), a startling discovery that had been anticipated by indirect arguments (e.g., Peebles 1984; Efstathiou, Sutherland, & Maddox 1990; Kofman, Gnedin & Bahcall 1993; Krauss & Turner 1995; Ostriker & Steinhardt 1995; Liddle et al. 1996) and has since been buttressed by more extensive supernova surveys and by several independent lines of evidence (see, e.g., Frieman, Turner, & Huterer 2008a for a recent review). Cosmic acceleration is widely viewed as one of the most profound phenomenological puzzles in contemporary fundamental physics. The two highest level questions in the field are:

1. Is cosmic acceleration caused by a breakdown of GR on cosmological scales, or is it caused by a new energy component with negative pressure (“dark energy”) within GR?
2. If the acceleration is caused by “dark energy,” is its energy density constant in space and time

and thus consistent with quantum vacuum energy (Zel'dovich 1968) or does its energy density evolve in time and/or vary in space?

For observational cosmology, the clearest path forward is to measure the history of cosmic expansion and the growth of dark matter clustering over a wide range of redshift with the highest achievable precision, searching for deviations from the GR plus cosmological constant model. Supernova surveys measure the distance-redshift relation using “standardized candles” whose luminosities are calibrated by objects in the local Hubble flow. BOSS, on the other hand, employs a “standard ruler,” the BAO feature imprinted on matter clustering by sound waves that propagate through the baryon-photon fluid in the pre-recombination universe (Peebles & Yu 1970; Sunyaev & Zel'dovich 1970; Eisenstein & Hu 1998; Meiksin, White & Peacock 1999). The BAO scale can be computed, in absolute units, using straightforward physics and cosmological parameters that are well constrained by cosmic microwave background (CMB) measurements. BAO are predicted to appear as a bump in the matter correlation function at a comoving scale corresponding to the sound horizon ( $r = 153.2 \pm 1.7$  Mpc, Larson et al. 2010), or as a damped series of oscillations in the matter power spectrum (see Eisenstein, Seo, & White 2007b for a comparison of the Fourier- and configuration-space pictures). When measured in the three-dimensional clustering of matter tracers at redshift  $z$ , the transverse BAO scale constrains the angular diameter distance  $d_A(z)$ , and the line-of-sight scale constrains the Hubble parameter  $H(z)$ .

The first clear detections of BAO came in 2005 from analyses of the 2dF Galaxy Redshift Survey (Cole et al. 2005) and of the luminous red galaxy sample (LRG; Eisenstein et al. 2001) of the SDSS (Eisenstein et al. 2005). The final SDSS-I/II BAO measurements determine the distance to  $z \approx 0.275$  with an uncertainty of 2.7% (Kazin et al. 2010; Percival et al. 2010; improved from the 5% of Eisenstein et al. 2005). Because of the leverage provided by this absolute distance measurement, BAO measurements contribute substantially to the overall cosmological constraints derived from SDSS galaxy clustering (see Reid et al. 2010).

BOSS consists of two spectroscopic surveys, executed simultaneously over an area of  $10,000 \text{ deg}^2$ . The first targets 1.5 million galaxies, selected in color-magnitude space to be high-luminosity systems at large distances. The selection criteria, described further below, produce a roughly constant comoving space density  $n \simeq 3 \times 10^{-4} h^3 \text{ Mpc}^{-3}$  to  $z = 0.6$ , with a slight peak at  $z \simeq 0.55$ , then a declining space density to  $z \simeq 0.8$ . Relative to the SDSS-I/II LRG survey, which contained  $10^5$  galaxies out to  $z = 0.45$ , the higher space density and higher limiting redshift of BOSS yield an effective volume (weighted by signal-to-noise ratio at the BAO scale) seven times larger.<sup>85</sup> The second BOSS survey targets  $1.5 \times 10^5$  quasars, selected from roughly  $4 \times 10^5$  targets (see below), in the redshift range  $2.2 \leq z \leq 4$ , where Ly $\alpha$  forest absorption in the SDSS spectral range can be used as a

tracer of high-redshift structure.<sup>86</sup> The high density and large number of targets will allow BOSS to provide the first “3-dimensional” measurements of large scale structure in the Ly $\alpha$  forest, on a sparsely sampled grid of sightlines that collectively probe an enormous comoving volume. The possibility of measuring BAO in the Ly $\alpha$  forest was discussed by White (2003), and Fisher matrix forecasts were presented by McDonald & Eisenstein (2007), whose formalism was used to motivate and design the BOSS quasar survey. While no previous survey has measured enough quasar spectra to reveal the BAO feature in the Ly $\alpha$  forest, analytic estimates and numerical simulations indicate that it should be clearly detectable in the BOSS quasar survey (McDonald & Eisenstein 2007; Slosar et al. 2009; Norman, Paschos & Harkness 2009; White et al. 2010). The characteristics of BOSS are summarized in Table 1.

Our forecasts, which are described in Appendix A, indicate that BAO measurements with the BOSS galaxy survey should yield determinations of  $d_A(z)$  and  $H(z)$  with  $1\sigma$  precision of 1.0% and 1.8%, respectively, at  $z = 0.35$  (bin width  $0.2 < z < 0.5$ ), and with precision of 1.0% and 1.7%, respectively, at  $z = 0.6$  ( $0.5 < z < 0.7$ ). The errors at the two redshifts are essentially uncorrelated, while the errors on  $d_A(z)$  and  $H(z)$  at a given redshift are anti-correlated (Seo & Eisenstein 2003). BAO are weakly affected by the effects of non-linear structure formation, galaxy bias, and redshift-space distortions. The primary consequence is a damping of oscillations in the power spectrum on small scales, which can be well approximated by a Gaussian smoothing (Bharadwaj 1996; Crocce & Scoccimarro 2006, 2008; Eisenstein, Seo, & White 2007b; Matsubara 2008a,b; Seo et al. 2010). Our forecasts assume that density field reconstruction (Eisenstein et al. 2007a) can remove 50% of the non-linear Lagrangian displacement of mass elements from their initial comoving locations (e.g. Padmanabhan, White & Cohn 2009; Noh, White & Padmanabhan 2009), thereby sharpening the BAO feature and improving recovery of the original signal. Forecasts with no reconstruction would be worse by factors of 1.6-2, while with perfect reconstruction (not achievable in practice) they would improve by factors of 1.3 – 1.5. The uncertainty in BOSS BAO measurements is dominated by cosmic variance out to  $z = 0.6$ ; at these redshifts, a much higher density of targets (eliminating shot noise) would decrease the errors by about a factor of 1.4, while covering the remaining  $3\pi$  steradians of the sky would reduce the errors by a factor of two. Non-linear effects can also generate small shifts in the BAO position, but current theoretical studies indicate that the statistical errors will dominate systematic uncertainties associated with redshift space distortions, non-linear evolution and galaxy bias (see, e.g., Eisenstein, Seo, & White 2007b; Smith et al. 2007; Padmanabhan & White 2009; Takahashi et al. 2009, 2011). To allow some margin over our forecasts — e.g., for reduced sky coverage due to poor weather, or for problems in reconstruction, or for other, unanticipated systematics — we have inflated our projected uncertainties by a factor of 1.2 when defining the measurement goals reported in Table 1.

<sup>85</sup> The SDSS main galaxy sample (Strauss et al. 2002) contains over 700,000 galaxies, but it has a median redshift of 0.1 and therefore a much smaller effective volume for power spectrum measurements on these scales.

<sup>86</sup> SDSS-I/II obtained spectra of 106,000 quasars, but only 17,600 were at  $z \geq 2.2$  (Schneider et al. 2010).

The Ly $\alpha$  forest forecasts, performed with the McDonald & Eisenstein (2007) formalism, indicate errors of 4.5% and 2.6%, respectively, on  $d_A(z)$  and  $H(z)$  at an effective redshift  $z \approx 2.5$  (with significant contributions from  $2 \lesssim z \lesssim 3.5$ ). The errors are again anti-correlated: the forecast error on an overall “dilation factor” that scales  $d_A(z)$  and  $H^{-1}(z)$  in proportion is only 1.8%. These predictions assume 15 quasars per deg<sup>2</sup> over 10,000 square degrees and no density field reconstruction. Reconstruction is less important at high redshift and is unlikely to be possible with a Ly $\alpha$  forest survey as sparse as BOSS. The measurement precision is limited partly by the sparse sampling of the density field and partly by the S/N of the spectra — i.e., at fixed sky area, increasing either the exposure time per quasar or the density of the quasar sample would decrease the errors. Our forecasts could prove somewhat optimistic, as broad absorption-line quasars may be unusable, quasars observed in grey time will have lower signal-to-noise spectra, and we have not included possible systematic uncertainties associated with continuum determination, metal lines, or damped Ly $\alpha$  systems. Conversely, use of additional imaging data sets could improve quasar target selection in some areas of the survey, increasing the surface density and improving the BAO measurement precision. Furthermore, these forecasts are based only on the location of the BAO peak as a function of angle with respect to the line of sight, marginalizing away additional information contained in the *amplitude* of Ly $\alpha$  flux correlations as a function of angle. Including this information — which requires careful theoretical modeling to control systematics — can lead to significant (factor-of-two level) improvements in the  $d_A(z)$  and  $H(z)$  constraints. More generally, the BOSS quasar survey is pioneering a previously untried method of BAO measurement, and performance forecasts are necessarily more uncertain than for the galaxy survey.

The underlying goal of these  $d_A(z)$  and  $H(z)$  measurements is to probe the cause of cosmic acceleration, e.g., to constrain the dark energy equation-of-state parameter  $w$  and its derivative  $w_a$  with respect to expansion factor. BOSS BAO measurements will also yield tight constraints on other cosmological parameters, most notably the Hubble constant  $H_0$  and the curvature parameter  $\Omega_k \equiv 1 - \Omega_m - \Omega_{DE} - \Omega_{rad}$ . Appendix A includes forecasts of BOSS constraints on these parameters in combination with complementary data (Table A1). We also present forecasts incorporating the broad-band galaxy power spectrum measurable with BOSS, which considerably improves dark energy constraints. Controlling systematic effects on the broad-band power to extract the full statistical power of the data set will require new work on the modeling of non-linear galaxy clustering and bias.

Since BOSS observes fainter targets than the original SDSS, it required substantial upgrades to the two dual-channel spectrographs. These were prepared during the first year of SDSS-III and installed during the summer shutdown following completion of SEGUE-2. In the red channel, the two 2048<sup>2</sup>, 24-micron pixel, SITE CCDs were replaced with 4128  $\times$  4114, 15-micron pixel, fully-depleted, 250-micron thick devices from Lawrence Berkeley National Laboratory, with much higher quantum efficiency at the reddest wavelengths, crucial for galaxy redshift measurements. In the blue channel, the two 2048<sup>2</sup>

SITE CCDs were replaced with 4096<sup>2</sup>, 15-micron pixel, e2v devices, with lower read noise and greater sensitivity at the blue wavelengths that are essential for Ly $\alpha$  forest measurements. In both arms, the smaller pixel size and larger format CCDs were selected to match the upgrade of the fiber system from 640 fibers with 3” optical diameter to 1000 fibers (500 per spectrograph) with 2” diameter. The larger number of fibers alone improves survey efficiency by 50%, and because BOSS observes point sources (quasar targets) and distant galaxies in the sky-dominated regime the smaller fibers yield somewhat higher signal-to-noise spectra in typical APO seeing, though they place stiffer demands on guiding accuracy and differential refraction. The original diffraction gratings were replaced with higher throughput, volume-phase holographic (VPH) transmission gratings from Kaiser Optical Systems, and other optical elements were also replaced or recoated to improve throughput. The spectral resolution varies from  $\lambda/\Delta\lambda \sim 1300$  at 3600 Å to 3000 at 10,000 Å. Figure 1 presents a schematic of one of the BOSS spectrographs. While we will not detail them here, we note that the transition to BOSS also involved major upgrades to the instrument and telescope control software, to the infrastructure for fiber-cartridge handling, and to the guide camera, which was replaced with an entirely new system.

BOSS galaxy targets are selected from the SDSS *ugriz* imaging (Fukugita et al. 1996; Stoughton et al. 2002), including the new imaging described below, using a series of color-magnitude cuts. These cuts are intended to select a sample of luminous and massive galaxies with an approximately uniform distribution of stellar masses from  $z \sim 0.2$  to  $z \sim 0.6$ . The sample is magnitude limited at  $z > 0.6$ . As in SDSS-I/II, the selection is the union of two cuts designed to select targets in two different redshift intervals (Eisenstein et al. 2001). Cut I (a.k.a. “LOZ”), aimed at the interval  $0.2 < z < 0.4$ , is defined by

$$r < 13.6 + c_{||}/0.3, \quad |c_{\perp}| < 0.2, \quad 16 < r < 19.5. \quad (1)$$

Cut II (a.k.a. “CMASS” for “constant mass”), aimed at redshift  $z > 0.4$ , is defined by

$$d_{\perp} > 0.55, \quad i < 19.86 + 1.6 \times (d_{\perp} - 0.8), \quad 17.5 < i < 19.9. \quad (2)$$

The colors  $c_{||}$ ,  $c_{\perp}$  and  $d_{\perp}$  are defined to track a passively evolving stellar population with redshift,

$$c_{||} = 0.7 \times (g - r) + 1.2 \times (r - i - 0.18) \quad (3)$$

$$c_{\perp} = (r - i) - (g - r)/4 - 0.18 \quad (4)$$

$$d_{\perp} = (r - i) - (g - r)/8, \quad (5)$$

based on population synthesis models of luminous red galaxies (Maraston et al. 2009). The  $r$ -band and  $i$ -band magnitude limits are imposed using `cmode1` magnitudes rather than the  $r$ -band Petrosian magnitudes used in SDSS-I/II. The 215,000 galaxies observed by SDSS-I/II that pass these cuts are included in the BOSS sample, but they are not reobserved if they had reliable redshifts. Figure 2 shows the space density of BOSS galaxies (including the SDSS-I/II objects) as a function of redshift, based on data obtained through July 2010.

Because the BOSS BAO experiment uses quasars only as backlights for the intervening Ly $\alpha$  forest, there is no

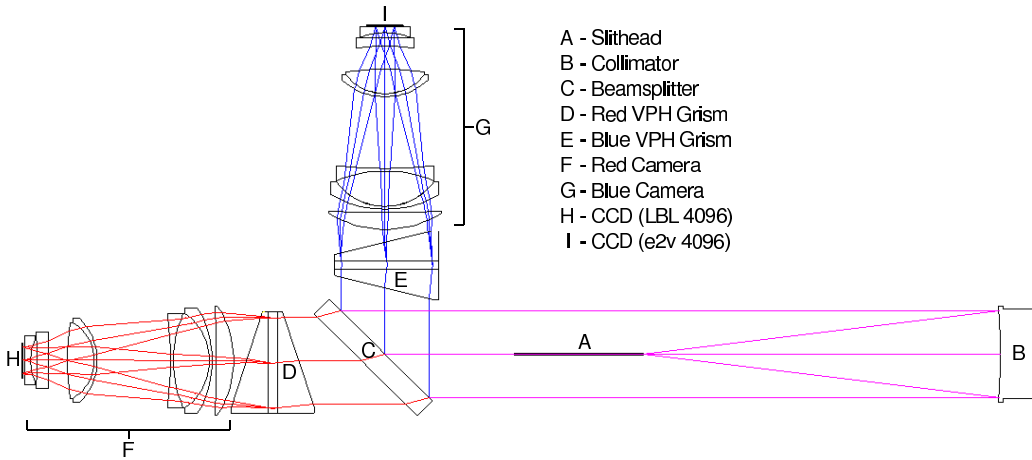


FIG. 1.— Schematic diagram of the BOSS spectrographs, with elements as labeled.

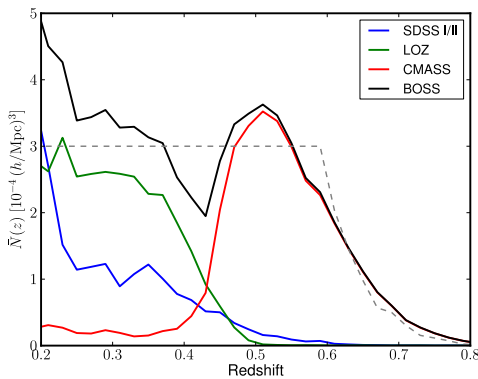


FIG. 2.— The comoving space density of BOSS galaxies from data taken in Spring 2010. The separate contributions of the LOZ cut, CMASS cut, and previously observed SDSS-I/II galaxies are shown, together with the total. The dashed curve shows our “goal” of constant density to  $z = 0.6$  and tapering density beyond. There is a deficit near  $z = 0.45$  at the transition between the two cuts, where obtaining accurate photometric redshifts for target selection is difficult.

need to select the sample homogeneously across the sky. The quasar survey is allocated an average of 40 targets per  $\text{deg}^2$ , and for Ly $\alpha$  forest science the essential criterion is to maximize the surface density of  $z \geq 2.2$  quasars above the practical limit for BOSS spectroscopy ( $g \approx 22$ ). Quasars at  $z < 2.2$  have little or no Ly $\alpha$  forest in the wavelength range covered by the BOSS spectrographs. In detail, the “value” of a quasar for BAO studies is a function of its redshift (which determines the observable Ly $\alpha$  forest path length) and its magnitude (which determines the S/N of the spectrum). Our own recent studies on the SDSS southern equatorial stripe, where deep co-added imaging and variability allow highly complete identification of optically bright (“Type I”) quasars, indicate that the surface density of  $z \geq 2.2$  quasars to the BOSS magnitude limit is approximately 28 per  $\text{deg}^2$  (see Palanque-Delabrouille et al. 2010). However, recovering these quasars from 40 targets per  $\text{deg}^2$  in single-epoch SDSS imaging is challenging because photometric errors are significant at this depth and because the quasar locus (in *ugriz*) crosses the stellar locus at  $z \approx 2.7$  (Fan 1999; Richards et al. 2002). We therefore set the BOSS science

TABLE 1  
SUMMARY OF BOSS

<i>Duration:</i> Fall 2009 - Summer 2014, dark time
<i>Area:</i> 10,000 $\text{deg}^2$
<i>Spectra:</i> 1000 fibers
3600 $\text{\AA} < \lambda < 10,000 \text{\AA}$
$R = \lambda/\Delta\lambda = 1300 - 3000$
$(S/N)^2$ at mag. limit
$\approx 22$ per pix. (averaged over 7000-8500 $\text{\AA}$ )
$\approx 10$ per pix. (averaged over 4000-5500 $\text{\AA}$ )
<i>Targets:</i> $1.5 \times 10^6$ massive galaxies, $z < 0.7$ , $i < 19.9$
$1.5 \times 10^5$ quasars, $z \geq 2.2$ , $g < 22.0$
selected from $4 \times 10^5$ candidates
75,000 ancillary science targets, many categories
<i>Measurement Goals:</i>
Galaxies: $d_A(z)$ to 1.2% at $z = 0.35$ and 1.2% at $z = 0.6$
$H(z)$ to 2.2% at $z = 0.35$ and 2.0% at $z = 0.6$
Ly $\alpha$ Forest: $d_A(z)$ to 4.5% at $z = 2.5$
$H(z)$ to 2.6% at $z = 2.5$
Dilation factor to 1.8% at $z = 2.5$

NOTE. — BOSS imaging data were obtained in Fall 2008 and Fall 2009. BOSS spectroscopy also uses grey time (lunar phase 70–100 degrees) when the NGC is observable. Galaxy target number includes 215,000 galaxies observed by SDSS-I/II. Measurement goals for galaxies are 1.2 times the projected  $1\sigma$  errors, allowing some margin over idealized forecasts. Measurement goals for the Ly $\alpha$  forest are equal to the  $1\sigma$  forecast, but this is necessarily more uncertain because of the novelty of the technique. The “dilation factor” is a common scaling for  $d_A(z)$  and  $H^{-1}(z)$  at  $z = 2.5$ .

requirement at 15 quasars per  $\text{deg}^2$ . Any gains in selection efficiency above this threshold translate into reduced errors on the BAO distance scale measured from the Ly $\alpha$  forest. Because the density field is sparsely sampled, the distance error is (approximately) inversely proportional to the quasar surface density at fixed survey area.

Quasar science — especially global population studies such as luminosity functions, active black hole mass functions, and clustering — benefit greatly from a homogeneous sample. We therefore select 20 of the 40 targets per  $\text{deg}^2$  from single-epoch SDSS imaging using a “core” selection method that remains fixed throughout the survey. This core selection is based on the probability, computed empirically from existing survey data,



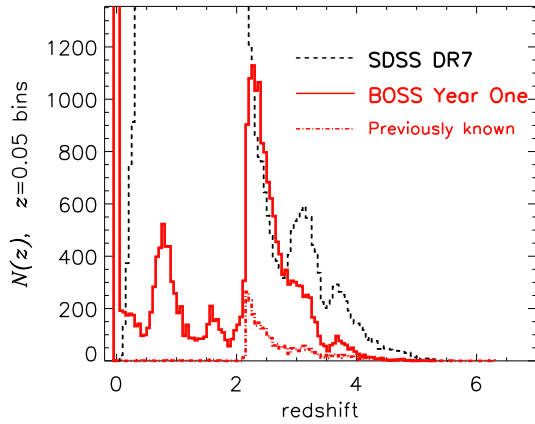


FIG. 3.— Redshift distribution of objects targeted by the BOSS quasar survey and observed between December 2009 and July 2010 (red solid line). There are 12,867 quasars with  $z \geq 2.20$ , obtained from a total of 55,114 targets, of which 32,844 yielded reliable redshifts. The spike at  $z = 0$  represents stellar contaminants, which are 34% of the objects with reliable redshifts. For comparison, the dotted histogram shows all quasars from the quasar catalog of SDSS DR7 (Schneider et al. 2010), and the dot-dashed histogram shows the previously known high- $z$  quasars in the area surveyed, which come mostly but not entirely from DR7 and were reobserved by BOSS.

that a given object is a high-redshift quasar rather than a star, low-redshift quasar, or galaxy (Bovy et al. 2010; J. Kirkpatrick et al., in preparation). The remaining 20 targets per  $\text{deg}^2$ , known as the “bonus” sample, include previously known high- $z$  quasars (including those from SDSS-I/II, reobserved to obtain higher S/N spectra), FIRST radio sources (Becker, White, & Helfand 1995) whose SDSS colors are consistent with  $z \geq 2.2$ , and objects selected by a variety of methods including the KDE method of Richards et al. (2009), the neural network method of Yèche et al. (2010), and lower priority likelihood targets. These targets are selected using additional data where they are available, including additional SDSS epochs (which improve photometric precision where stripes overlap and, on the southern equatorial stripe, provide variability information) and photometry from GALEX (UV; Martin et al. 2005) and UKIDSS (near-IR; Lawrence et al. 2007). The quasar selection criteria have evolved significantly during the first year of BOSS, since BOSS observations themselves provide vastly more training data at these magnitudes than earlier surveys such as 2SLAQ (Croom et al. 2009) and AGES (C. Kochanek et al., in preparation). The BOSS quasar target selection algorithms, including the criteria used during the first year, will be described in detail elsewhere. With single-epoch SDSS imaging we are presently achieving our requirement of 15 quasars per  $\text{deg}^2$ , but with little margin, and we hope to improve future efficiency with incorporation of additional data sets. Figure 3 shows the redshift and apparent magnitude distribution of BOSS quasars observed between December 2009 and July 2010; for this plot, all quasar classifications and redshifts have been checked by visual inspection. As of January 2011, BOSS has obtained spectra of 29,000 quasars with  $z \geq 2.2$  (according to pipeline redshifts), compared to 17,600 from all of SDSS-I and II.

Figure 4 shows several examples of BOSS galaxy spectra (left) and quasar spectra (right), with brighter ob-

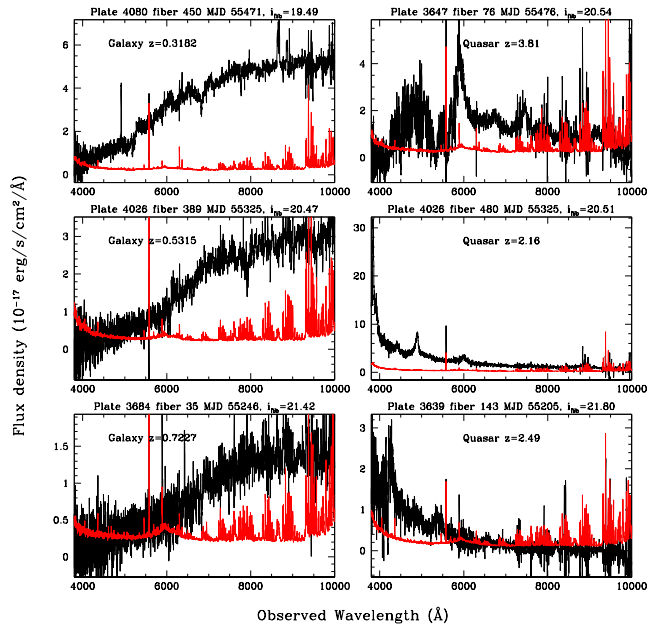


FIG. 4.— Examples of BOSS galaxy spectra (left) and quasar spectra (right), smoothed with a 3-pixel boxcar. In each panel, the black line is the spectrum and the red is the estimated error per pixel. The galaxy redshifts are 0.318, 0.531, and 0.723 (top to bottom). The quasar redshifts are 3.81, 2.16, and 2.47 (top to bottom). The Ly $\alpha$ , CIV, and CIII] emission lines are identifiable features in these quasar spectra. The 2''-fiber  $i$ -band magnitudes of the targets are listed above each panel.

jects at the top and targets near the magnitude limit at the bottom. BOSS observations are done in a series of 15-minute exposures, with additional exposures taken until a regression of  $(S/N)^2$  against magnitude (based on a fast reduction pipeline) yields  $(S/N)^2 \geq 22$  per wavelength pixel ( $1.4 \text{ \AA}$ ) at  $i = 21$  ( $2''$  fiber magnitude) in the red cameras and  $(S/N)^2 \geq 10$  per wavelength pixel ( $1.1 \text{ \AA}$ ) at  $g = 22$  in the blue cameras, where magnitudes are corrected for Galactic extinction (Schlegel, Finkbeiner, & Davis 1998).<sup>87</sup> In transparent conditions, good seeing, and low Galactic extinction, the total exposure time is 45-60 minutes, but the fixed  $(S/N)^2$  criterion ensures homogeneity of redshift completeness across the survey. Our current data reductions, incorporating a spectroscopic reduction pipeline adapted from the one originally developed for SDSS-I/II data by S. Burles and D. Schlegel, meets our science requirement of 95% redshift completeness for galaxy targets. However, we plan to implement the forward modeling techniques described by Bolton & Schlegel (2010) to extract all the information contained in the spectra as accurately as possible. These pipeline improvements will increase our redshift completeness, improve galaxy science, and, most importantly, yield higher S/N and better characterized errors in the Ly $\alpha$  forest, thus maximizing the return of the Ly $\alpha$  forest survey.

<sup>87</sup> Higher  $(S/N)^2$  thresholds, and consequently longer exposure times, were employed during the first year.



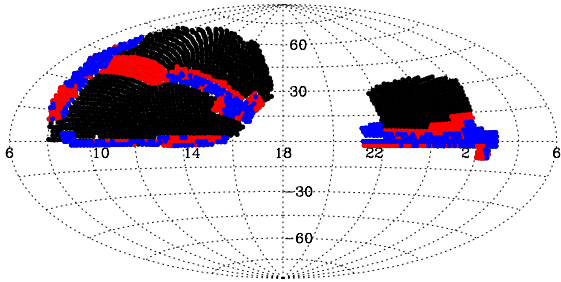


FIG. 5.— Planned footprint of the BOSS spectroscopic survey, showing both the northern (left) and southern (right) Galactic cap regions. Most of the imaging for SGC target selection was done as part of SDSS-III. Each circle marks the location of a spectroscopic plate. Blue circles represent plates that have been observed as of January 2011, while red circles represent plates that have been drilled but not yet observed. The small southern extension of the SGC region is intended to reach fields of the CFHT Legacy Survey.

SDSS I and II imaged  $7646 \text{ deg}^2$  of high-latitude sky in the north Galactic cap (NGC) and three stripes totaling  $777 \text{ deg}^2$  of low extinction sky in the south Galactic cap (SGC).<sup>88</sup> In order to allow BOSS to cover  $10,000 \text{ deg}^2$  with a balance between the fall and spring observing seasons, BOSS used the SDSS camera to image an additional  $2500 \text{ deg}^2$  during the first 18 months of SDSS-III, following the same procedures as SDSS I and II. Figure 5 shows the full footprint for BOSS spectroscopic observations. The total area shown is  $10,700 \text{ deg}^2$ , while our science goal for spectroscopy is  $10,000 \text{ deg}^2$ ; the exact breakdown between NGC and SGC in the spectroscopic survey will depend on the amount of clear weather when these two regions are observable. Assuming historical weather patterns, we anticipate a 5% margin to complete the  $10,000 \text{ deg}^2$  spectroscopic survey by July 2014.

While our BAO measurement goals drive the design and the science requirements of BOSS, the survey will enable a wide range of other science. Redshift-space distortion analyses of BOSS galaxy clustering have the potential to yield strong constraints on clustering growth rates (White, Song, & Percival 2009), while weak lensing by BOSS spectroscopic galaxies measured in SDSS (or deeper) imaging can directly measure the evolution of matter clustering. These methods could substantially increase the impact of BOSS in its “core” science area of testing theories of cosmic acceleration. For large scale power spectrum measurements, the much larger effective volume of BOSS (compared to SDSS-I/II) will enable much stronger constraints on neutrino masses, inflation parameters, and departures from “vanilla”  $\Lambda$ CDM. BOSS galaxy spectra will provide a superb data set for studying the evolution of massive galaxies from  $z \approx 0.7$  to the present, and they are expected to reveal  $\sim 300$  examples of strong gravitational lensing that can be used to constrain the mass profiles of early type galaxies (e.g., Koopmans et al. 2006; Bolton et al. 2008a,b; Koopmans et al. 2009). The high-redshift quasar data set will be ten times larger and approximately 2.5 magnitudes deeper (1.5 magnitudes at  $z > 3$ ) than SDSS-I/II, enabling much stronger constraints on the evo-

lution and clustering of quasars and the black holes that power them. The new BOSS imaging will extend “tomographic” studies of Milky Way structure (e.g., Ivezić et al. 2008) and searches for ultra-faint dwarf galaxy companions (e.g., Belokurov et al. 2006).

Finally, BOSS is devoting about four percent of its fibers to “ancillary” science targets in a variety of categories. These include studies of luminous blue galaxies at high redshifts, brightest cluster galaxies, star forming radio galaxies, remarkable X-ray sources from *Chandra* and *XMM-Newton*, host galaxies of supernovae found in SDSS-II, quasars selected by photometric variability, double-lobed radio quasars, candidate quasars at  $z > 5.6$ , variability in quasar absorption lines, *Fermi*  $\gamma$ -ray sources, distant halo red giants, activity in late-M and L dwarfs, hot white dwarfs, and low mass binary star candidates. Spectra from these ancillary science programs will be included in the public data releases.

### 3. SEGUE-2

The first SDSS imaging maps provided striking confirmation of complex structure in the outer Galaxy (Ivezić et al. 2000; Yanny et al. 2000; Newberg et al. 2002), including the well known tidal tails of the Sagittarius dwarf galaxy (Ibata, Gilmore, & Irwin 1994; Ibata et al. 2001; Majewski et al. 2003) and previously unrecognized streams, rings, and clumps (e.g., Odenkirchen et al. 2001; Yanny et al. 2003; Grillmair 2006; Grillmair & Dionatos 2006; Jurić et al. 2008). The ubiquity of this complex structure (e.g., Belokurov et al. 2006; Bell et al. 2008) supports the view that disrupted dwarf satellites are important contributors to the formation of the Galactic halo (Searle & Zinn 1978), a scenario in qualitative and quantitative agreement with hierarchical, CDM-based models of galaxy formation (Helmi & White 1999; Bullock, Kravtsov & Weinberg 2001; Bullock & Johnston 2005). These initial discoveries motivated the SEGUE-1 survey of SDSS-II (Yanny et al. 2009), which included  $3500 \text{ deg}^2$  of new *ugriz* imaging extending towards the Galactic plane and a spectroscopic survey of 240,000 stars in a variety of target categories. The first year of SDSS-III, during which the upgraded spectrograph components for BOSS were being constructed, offered the opportunity to roughly double the scope of SEGUE, using all of the dark time over one year<sup>89</sup> instead of 1/3 of the dark time over three years. In comparison to the Radial Velocity Experiment (RAVE; Steinmetz et al. 2006; Zwitter et al. 2008; Fulbright et al. 2010), which is roughly comparable in number of stars, SEGUE-1 and SEGUE-2 go much deeper (to  $g \sim 20$  vs.  $I \sim 13$ ) and cover a larger wavelength range ( $3800 - 9200 \text{ \AA}$  vs.  $8410 - 8795 \text{ \AA}$ ), but with lower resolution (1800 vs. 7500) and lower S/N. The SEGUE surveys probe the distant disk and halo, while RAVE provides higher resolution data concentrated in the solar neighborhood.

The defining goal of SEGUE-2 is to map stellar populations and their kinematics over a large volume of the Galaxy, from the inner halo out to large Galactocentric distances where late-time accretion events are ex-

<sup>88</sup> SDSS-II also included  $3500 \text{ deg}^2$  of lower latitude imaging for SEGUE, but these data are not useful for BOSS.

<sup>89</sup> Except for the time devoted to BOSS imaging. SEGUE-2 also observed during grey time, with lunar phase 70 – 100 degrees from new moon.

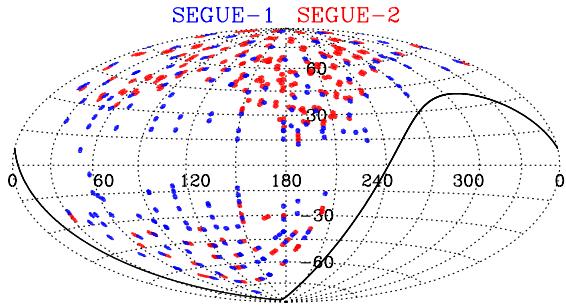


FIG. 6.— Fields of the SEGUE-1 (blue) and SEGUE-2 (red) surveys, in Galactic coordinates.

pected to dominate the halo population. SEGUE-1 and SEGUE-2 are similar enough in strategy and data quality to be treated as a single data set. Both surveys selected targets from the SDSS and SEGUE *ugriz* imaging data along individual 7 deg<sup>2</sup> lines of sight, which are spread out over the imaging survey but do not form a filled area. Both surveys selected spectroscopic targets in several categories designed to map Galactic structure at different distances or to identify classes of objects of particular astrophysical interest. However, the target selection for SEGUE-2 is informed by the lessons from SEGUE-1. The most important strategic difference is that SEGUE-1 paired shorter exposures of brighter targets with deep spectroscopic pointings along the same sightlines, but SEGUE-2 obtained only deep pointings in order to maximize coverage of the distant Galaxy. The survey was designed to obtain 140,000 spectra, but worse than expected weather led to a final sample of 118,151 stars. As with BOSS, SEGUE-2 exposures are accumulated until the S/N crosses a pre-determined threshold. For SEGUE-2, that threshold corresponds to median S/N  $\approx 10$  per pixel ( $\Delta\lambda \approx 1 \text{ \AA}$ ) for metal-poor main sequence turnoff stars at  $r = 19.5$ , with spectral resolution  $R = \lambda/\Delta\lambda \approx 1800$ . Under good conditions, reaching this S/N threshold required approximately three hours of total exposure time. Figure 6 shows the distribution of SEGUE and SEGUE-2 fields in Galactic coordinates.

A detailed description of SEGUE-2 target selection will be provided elsewhere (C. Rockosi et al., in prep.). With one exception, the ultra-cool white dwarf candidates, the selection criteria for all the target categories were adjusted based on what was learned from the SEGUE-1 data in order to obtain a higher success rate for categories like the low metallicity candidates and the blue horizontal branch stars or to push to larger mean distances for samples like the halo main sequence turnoff stars. In brief the SEGUE-2 target categories, selection criteria, and numbers of targets successfully observed are:

- Halo main sequence turnoff (MSTO) stars:  $18 < g < 19.5$ ,  $+0.1 < g - r < +0.48$ ; 37,222 targets.
- Blue horizontal branch (BHB) stars:  $15.5 < g < 20.3$ ,  $-0.5 < g - r < +0.1$ ,  $+0.8 < u - g < +1.5$ ; 9983 targets
- K-giants: selected based on color and low proper motion. with  $15.5 < g < 18.5$  and  $r > 15$ ; 42,973 targets.
- M-giants: selected based on color and low proper

motion, with  $15.5 < g < 19.25$  and  $i > 14.5$ ; 631 targets.

- Halo high velocity stars, selected based on color and high proper motion,  $17 < g < 19.5$ ; 4133 targets.
- Hypervelocity stars: selected based on color and high proper motion,  $17 < g < 20$ , 561 targets.
- Cool extreme- and ultra-subdwarfs: selected based on color and reduced proper motion, with  $15 < r < 20$  and  $g > 15.5$ ; 10,587 targets.
- Ultra-cool white dwarfs : selected based on reduced proper motion, with  $15.5 < r < 20.5$  and  $g > 15.5$ ; no targets.
- Low metallicity candidates, color selected, with  $15.5 < g < 18$  and  $r > 15$ ; 16,383 targets.

The first four categories are aimed primarily at understanding the kinematic and chemical structure of the outer Galaxy, detecting substructures in the stellar halo or outer disk, and constraining the mass profile and shape of the Milky Way’s dark matter halo. These four categories have successively higher characteristic luminosities, so they provide successively deeper but sparser probes, with typical distance limits of 11 kpc (MSTO), 50 kpc (BHB) and 100 kpc (K/M-giants). Hyper-velocity stars (Hills 1988; Brown et al. 2006) are thought to originate in dynamical interactions with the Galaxy’s central black hole, and a systematic census of these stars can probe both the physics of the ejection mechanism and the stellar population at the Galactic Center. Kollmeier et al. (2010) present an analysis of this subset of the SEGUE-2 data. The extreme subdwarf category is designed to find the most metal-poor cool stars in the solar neighborhood. The ultra-cool white dwarf category was modeled after the successful searches for these objects in the larger area SDSS survey (e.g., Gates et al. 2004; Harris et al. 2008) and was carried over unchanged from the SEGUE-1 survey. However, the smaller area covered by SEGUE-2 did not produce any candidates. Finally, the low metallicity category aims to identify candidates for future high-resolution spectroscopy that can probe nucleosynthesis in the first generations of metal-poor or metal-free stars.

Figure 7 shows three examples of SEGUE-2 stellar spectra: a BHB star (top), a very metal-poor main-sequence turnoff star (middle), and a very metal-poor K giant (bottom). The left-hand panels show the flux-calibrated spectra over the entire available spectral range, while the right-hand panels show the blue portion of each spectrum after fitting and removing the continuum, which aids examination of the detailed shape of the individual spectral lines. The estimated atmospheric parameters for each star, obtained as described below, are displayed in the left-hand panels. Prominent spectral lines are identified in the right-hand panels.

Like the SEGUE-1 spectra released in SDSS DR7, SEGUE-2 spectra are first reduced by the `id1spec2d` pipeline (described in the DR8 paper, Aihara et al. 2011), which performs sky subtraction and wavelength and flux calibration, then extracts the 1-d spectrum, carries out a basic classification, and measures the radial

TABLE 2  
SUMMARY OF SEGUE-2

<i>Duration:</i> Fall 2008 - Summer 2009, dark+grey time	
<i>Area:</i> 1317 deg <sup>2</sup> , 118,151 targets	
<i>Spectra:</i> 640 fibers	
3800 Å < λ < 9200 Å	
R = λ/Δλ = 1800	
S/N ≈ 10 per res. element at r = 19.5	
<i>Target Categories:</i>	
Halo main sequence turnoff stars (37,222)	
Blue horizontal branch stars (9983)	
K-giants and M-giants (43,604)	
High-velocity stars (4133)	
Hypervelocity stars (561)	
Cool extreme subdwarfs (10,587)	
Ultra-cool white dwarfs (none selected)	
Low metallicity candidates (16,383)	
<i>Precision:</i> Dependent on stellar type and S/N, but typically	
150 K in T <sub>eff</sub> , 0.23 dex in log g	
0.21 dex in [Fe/H], 0.1 dex in [α/Fe]	

velocity. This pipeline is unchanged from DR7. The radial velocity accuracy is 4 km s<sup>-1</sup> at  $g=18$  (for detailed discussions, see Yanny et al. 2009 for SEGUE-1 and Rockosi et al., in prep., for SEGUE-2). The SEGUE-2 stellar spectra are then processed by the SEGUE Stellar Parameter Pipeline (SSPP; Lee et al. 2008a,b; Allende Prieto et al. 2008; Smolinski et al. 2010), and the three primary stellar atmospheric parameters ( $T_{\text{eff}}$ ,  $\log g$ , and  $[\text{Fe}/\text{H}]$ ) are reported for most stars in the temperature range 4000 – 10,000 K that have spectral S/N ratios exceeding 10:1 per Å (averaged over the entire spectrum). The SSPP estimates stellar atmospheric parameters using several approaches, such as a minimum distance method (Allende Prieto et al. 2006), neural network analysis (Bailer-Jones et al. 2000; Willemsen et al. 2005; Re Fiorentin et al. 2007), auto-correlation analysis (Beers et al. 1999), and a variety of line index calculations based on previous calibrations with respect to known standard stars (Beers et al. 1999; Cenarro et al. 2001a,b; Morrison et al. 2003). We refer the interested reader to Lee et al. (2008a) for more details on the SSPP and to Smolinski et al. (2010) for a description of recent updates. The current best estimates of the precision of the derived parameters  $T_{\text{eff}}$ ,  $\log g$ , and  $[\text{Fe}/\text{H}]$  are 150 K, 0.23 dex, and 0.21 dex, respectively, for SEGUE stars with 4500 K  $\leq T_{\text{eff}} \leq 7500$  K. As described by Lee et al. (2010b), the SSPP has recently been extended to provide estimates of alpha-element-to-iron ratios, with precision of  $\sim 0.1$  dex in  $[\alpha/\text{Fe}]$  for stars with S/N > 20 : 1, 4500  $\leq T_{\text{eff}} \leq 7000$  K, and  $[\text{Fe}/\text{H}] > -2.5$  (or as low as  $-3.0$  for cooler stars). Validation and refinements of the SSPP parameter estimates are still underway, based on a uniform re-analysis of several hundred high-resolution spectra of SDSS/SEGUE stars obtained with the Hobby-Eberly, Keck, Subaru, and VLT telescopes.

#### 4. APOGEE

No survey to date of the chemistry of Milky Way stars has included all Galactic stellar populations in a systematic, consistent manner. This limitation arises primarily because the inner regions of the Galaxy are largely hidden at optical wavelengths by obscuring interstellar dust. Furthermore, obtaining precise, element-by-

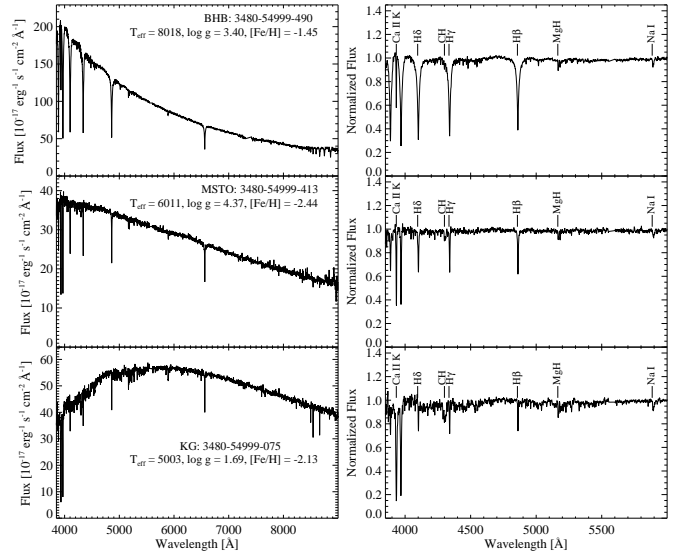


FIG. 7.— Example spectra of several classes of SEGUE-2 stars, all taken from the same spectroscopic plate. The left-hand panels show the flux-calibrated spectra for a halo blue horizontal-branch (BHB) star, a main-sequence turnoff (MSTO) star, and a K giant (KG) star, identified by plate-MJD-fiber. The stellar atmospheric parameter estimates for each object are shown below the label, as determined by the SSPP. The right-hand panels show the blue portion of each spectrum after continuum normalization by the SSPP. Prominent spectral features in each spectrum are labeled. Dereddened apparent magnitudes, from top to bottom, are  $g_0 = 16.27, 17.82, 17.31$  and  $r_0 = 16.35, 17.49, 16.67$ .

element abundances requires high-resolution, high-S/N spectra, well beyond the capabilities of SEGUE or of any large spectroscopic survey to date. APOGEE will address both of these limitations by obtaining  $H$ -band (1.51 – 1.70  $\mu\text{m}$ ) spectra for 10<sup>5</sup> late type, evolved stars with a FWHM resolving power  $R = \lambda/\Delta\lambda \approx 30,000$  and a minimum S/N of 100 per resolution element. Thanks to the greatly reduced extinction at infrared wavelengths ( $A_H/A_V = 1/6$ ), APOGEE will observe efficiently even in heavily obscured regions of the Galaxy. It will be the first large scale, systematic, high-precision spectroscopic survey of *all* Galactic stellar populations — those in the bulge, bar, thin disk, thick disk, and halo — conducted with a homogeneous set of spectral diagnostics, data quality and stellar tracers.

We anticipate that APOGEE will deliver radial velocities with an accuracy of 0.3 km s<sup>-1</sup> or better and star-by-star abundances of  $\sim 15$  key elements, with  $\sim 0.1$ -dex precision (for solar metallicity targets), including the most common metals C, N and O, many of the  $\alpha$ -elements, several iron-peak elements, and two odd- $Z$  elements, Na and Al. These different species form through different nucleosynthetic pathways in stars of different mass and metallicity, and they therefore provide complementary information about chemical evolution of their parent galaxy (the Milky Way or a dwarf progenitor), about the physics of stellar and supernova nucleosynthesis, and about the mixing and enrichment history of the interstellar medium in the Galaxy. APOGEE will increase the total number of high resolution, high-S/N stellar spectra obtained under uniform conditions at any wavelength by more than two orders of magnitude.

APOGEE’s main science objectives are:

- To derive tight constraints on models for the his-

tory of star formation, chemical evolution (including the processes of chemical mixing and feedback in the interstellar medium and dredge-up in individual stars), and mass assembly of the Galaxy.

- To constrain the stellar initial mass function (IMF) in each of the main Galactic components.
- To derive kinematical data at high precision useful for constraining dynamical models of the disk, the bulge, the bar, and the halo and for discriminating substructures in these components, if/where they exist.
- To infer properties of the earliest stars (usually thought to reside or to have resided in the Galactic bulge), by detecting them directly if they survive to the present day in significant numbers or by measuring their nucleosynthetic products in the most metal-poor stars that do survive.
- To unravel the overall formation mechanisms and evolution of the Milky Way by coupling the extensive chemical data to the dynamics of the stars.

The APOGEE sample size will be large enough to provide statistically reliable measures of chemistry and kinematics in dozens of separate zones defined by cuts in Galactocentric radius, Galactic longitude, and height above the plane, at the level of precision currently available only for stars in the solar neighborhood. Detailed chemical fingerprints will allow identification of sub-populations that have a common origin but may now be distributed widely around the Galaxy, providing unique insights into the Galaxy’s dynamical history. In more general terms, APOGEE and SEGUE provide the kinds of data needed to use the Milky Way as a detailed test of contemporary galaxy formation theory (see, e.g., Freeman & Bland-Hawthorn 2002).

The fiber plugplate system for APOGEE is similar to that used for BOSS and SEGUE-2, but the APOGEE spectrograph sits in a building adjacent to the Sloan telescope rather than being mounted on the telescope underneath the mirror. For APOGEE, 300 “short” (2.5-meter) fibers carry light from the focal plane to a connection below the telescope, where they are coupled to “long” (40-meter) fibers that transport the signal to the spectrograph enclosure, penetrate the evacuated dewar, and illuminate a pseudo-slit. The fibers have an outer diameter of  $190\mu\text{m}$  and an inner (light-transmitting) diameter of  $120\mu\text{m}$ , corresponding to  $2''$  on the sky. Figure 8 presents a SolidWorks (TM) model rendering of the APOGEE spectrograph (Wilson et al. 2010). The key elements are:

- A  $2.1\text{ m}^3$  cryostat, maintained at approximately 80 K and a vacuum level below  $10^{-6}$  Torr.
- A  $50.8\text{ cm} \times 30.5\text{ cm}$  volume phase holographic (VPH) grating, the largest yet employed in an astronomical instrument, manufactured with a novel mosaic process by Kaiser Optical Systems.
- A six-element camera manufactured and aligned by New England Optical Systems, with optical elements of silicon and fused silica to a maximum diameter of 40 cm.

TABLE 3  
SUMMARY OF APOGEE

<i>Duration:</i> Spring 2011 - Summer 2014, bright time
<i>Area:</i> $\sim 1575\text{ deg}^2$
$\sim 230$ fields
total exposure times 1-24 hours
<i>Spectra:</i> 300 fibers
$1.51\mu\text{m} < \lambda < 1.70\mu\text{m}$
$R \equiv \lambda/\Delta\lambda = 27,000 - 31,000$
typical S/N $\geq 100$ per resolution element
<i>Targets:</i> $10^5$ 2MASS-selected, evolved stars with $(J - K_s)_0 > 0.5$
Includes RGB, clump giants, AGB, red supergiants
<i>Precision:</i> Abundances to 0.1 dex internal precision
0.2 dex external precision
$\sim 15$ elements per star
Radial velocities to $< 0.3\text{ km s}^{-1}$ per visit.

- Three  $2\text{K} \times 2\text{K}$  HAWAII-2RG infrared detectors (Garnett et al. 2004) mounted on a translational stage that enables precise subpixel dithering in the dispersion dimension.
- A collimator with electromechanical control for focusing and dithering in the spatial dimension (useful for making high quality, “smeared” flatfields).
- Two fold optics used to create a compact spectrograph design and, through the use of dichroic coating on Fold Mirror 2, to divert and sequester out-of-band light.

These components are now being integrated, aligned, and tested at the University of Virginia, and the full spectrograph is scheduled for shipment to APO in early 2011.

The three detectors span the wavelength range  $1.51 - 1.70\mu\text{m}$  with two small gaps. The spectral resolution increases from  $R \equiv \lambda/\Delta\lambda \approx 27,000$  at the shortest wavelengths to  $R \approx 31,000$  at the longest wavelengths. The pixel scale is  $0.35\text{-}0.23\text{ \AA}$  (blue to red spectral end), so there are typically 1.6 pixels per FWHM at the blue end and 2.3 pixels per FWHM at the red end. APOGEE observations will consist of pairs of exposures dithered by half a pixel, which will then be combined to yield fully sampled spectra at the instrument resolution at all wavelengths. In a typical observation,  $\sim 250$  fibers will be devoted to science targets and  $\sim 50$  to calibration stars, telluric standards, and sky. The throughput requirement on the spectrograph is to achieve S/N = 100 per resolution element in three hours of exposure time under good conditions for a star with Vega magnitude  $H = 12.2$ . Our measurements of component throughputs suggest that the total throughput will be somewhat better than this level and will reach the above S/N for  $H = 12.5$  stars in three hours. Visits to each star field will be about one hour, so that most program stars will be observed at least three times. Because internal binary velocities can distort measurements of Galactic kinematics (e.g., by inflating velocity dispersions), the cadences of these repeat visits will be designed so that the majority of the most troublesome binaries can be identified via radial velocity variations.

APOGEE and MARVELS share the focal plane during bright time observations, with separate fibers on the same  $7\text{ deg}^2$  plugplates feeding different instruments.



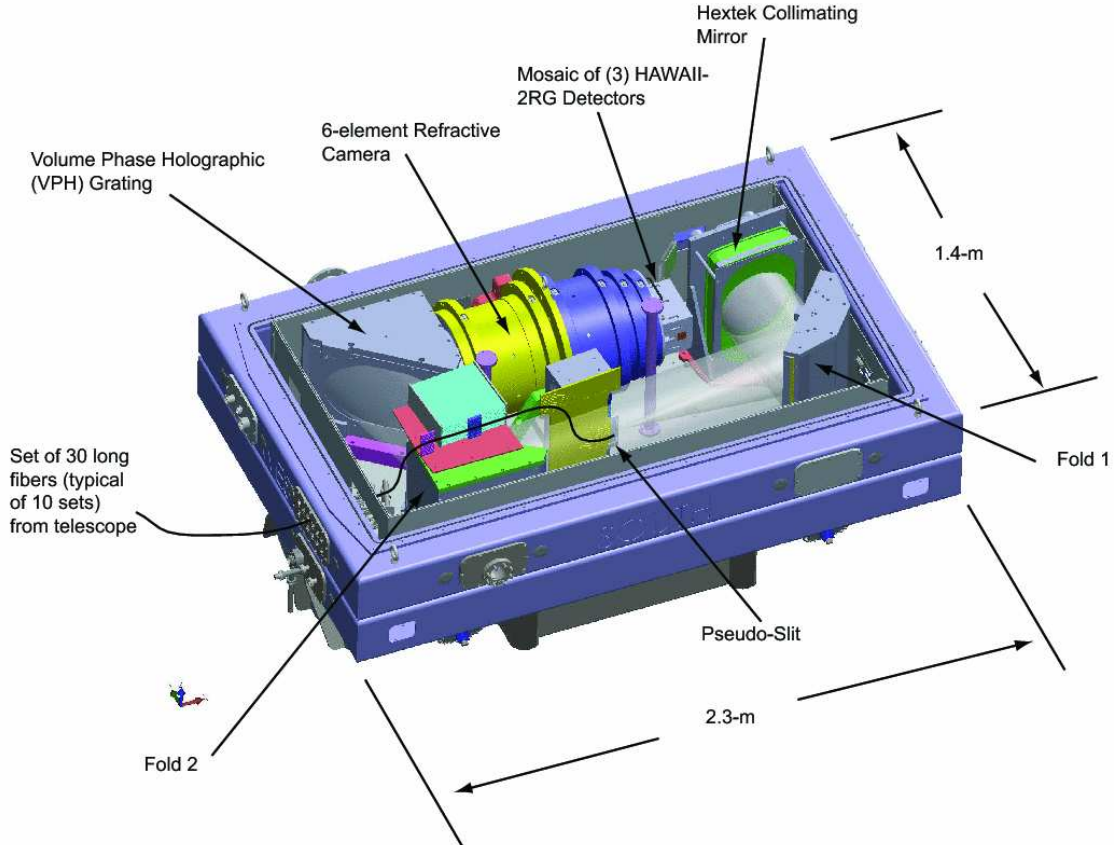


FIG. 8.— A SolidWorks(TM) model rendering of the fiber-fed APOGEE spectrograph. 300 fibers are arranged in a column on a pseudo-slit. Divergent light from the fiber ends is collimated and redirected back towards the pseudo-slit by a collimating mirror. The collimated light passes through the pseudo-slit and is dispersed by a volume phase holographic grating and refocused by a 6-element camera onto a mosaic of three HAWAII-2RG detectors. Fold mirrors are used to package the optical train in an efficient manner.

This scheme nearly doubles the observing time available to both surveys, at the cost of reducing flexibility. In particular, MARVELS observations require visiting the same field as many as 24 times throughout the 3-year APOGEE survey. APOGEE will therefore devote most of its observations to “long” fields, observed for a total of 24 hours or, in cases where earlier MARVELS observations have already accumulated many epochs of data, for smaller total visit times (e.g., 10 hours). The nominal APOGEE exposure time is three hours, but most fields have many more than 250 potential APOGEE targets, so the additional observing time can be used to increase the sampling density along the line of sight by up to a factor of eight. It can also be used to increase depth by observing fainter stars for total exposure times as long as 24 hours, and perhaps to obtain higher S/N spectra for lower metallicity stars that have weaker lines. The detailed mix among these strategies is not yet decided. Roughly 25% of the observing time will be assigned to APOGEE-only, “short” fields that increase the sky coverage of the survey, with three one-hour visits in disk and halo fields and single one-hour visits (concentrating on brighter stars) in bulge fields that are only available for short periods at acceptable airmass from APO.

Figure 9 shows the currently planned distribution of APOGEE fields, although the plan remains subject to (likely minor) changes. Some fields target globular or open star clusters, both for science investigations and to

allow calibration of APOGEE abundances against previous data sets for these clusters. Most fields are designed to provide systematic coverage of the Galaxy, subject to the constraints of observability. Locations of some fields are shifted from a regular grid by modest amounts to increase the number of suitable MARVELS targets. For all fields, target selection is based primarily on 2MASS photometry (Skrutskie et al. 2006), with data from *Spitzer*/IRAC (Churchwell et al. 2009) and *WISE* (Wright et al. 2010) used to correct for dust reddening on a star-by-star basis (see Majewski et al. 2011). At low latitudes, APOGEE targets will be selected with a simple, dereddened, near-infrared color limit — e.g.,  $(J - K_s)_0 > +0.5$ ; such a sample will be dominated by late type, evolved stars like red giant branch, asymptotic giant branch, red clump, and red supergiant stars. These stellar species span all population ages and allow APOGEE to probe to the farthest distances at a given magnitude. At higher latitudes where the ratio of distant giants to nearby dwarf stars is lower, the simple  $(J - K_s)_0$  color selection will be supplemented by newly obtained photometry in the Washington  $M$ ,  $T_2$  and  $DDO51$  filters, which will greatly reduce the contamination of the APOGEE target sample by nearby dwarf stars (see Geisler 1984; Majewski et al. 2000). Because of large differential refraction effects that can occur across the nominal  $3^\circ$  field of view when observing at large airmass, APOGEE fields targeting the inner disk and bulge

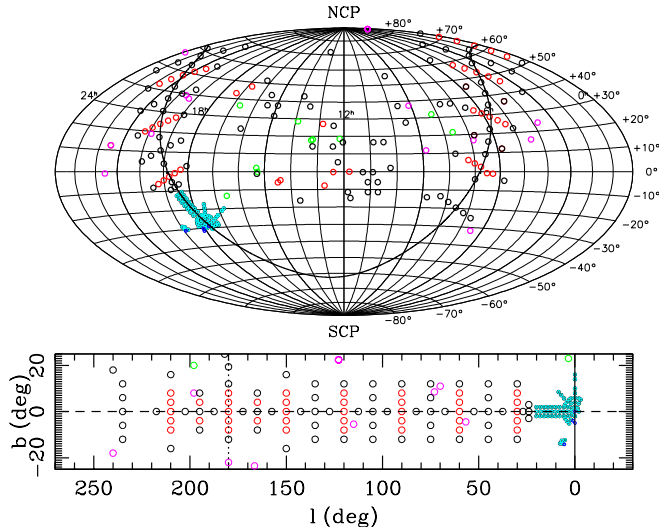


FIG. 9.— Current plan of the APOGEE field distribution in equatorial (top) and Galactic (bottom) coordinates. Fields are represented with circles scaled to represent their actual size ( $7 \text{ deg}^2$  for typical APOGEE fields). Red circles denote “24-hr” APOGEE/MARVELS fields. Green circles denote “10-hr” APOGEE/MARVELS fields. Magenta circles denote fields containing calibration clusters, which will typically be observed for a total of three hours. APOGEE-only fields (also observed for only three hours) are shown with black circles. Fields targeting the inner disk and bulge and the core of the Sagittarius dSph galaxy will have a smaller field of view ( $\sim 1 \text{ deg}^2$ ) to compensate for being observed at high airmass (see text) and are shown as either 3-hr (dark blue) or 1-hr (cyan) fields. The heavy line in the upper plot marks the Galactic plane.

and the core of the Sagittarius dSph galaxy will be observed with fibers placed in a reduced (e.g.,  $1^\circ$ ) field of view. Because of the high stellar density in these fields, there will be no shortage of targets to fill the entire complement of fibers.

In each field, targets will be selected in three different bins of  $H$ -band magnitude to provide a reasonable spread in distance coverage. In the absence of dust extinction, a typical red clump giant has  $H = 12.5$  at a distance of  $\sim 6$  kpc, while a typical star near the tip of the red giant branch has  $H = 12.5$  at a distance of 40 kpc. Observations to  $H = 13.5$  (achievable for some targets in long fields) reach a factor 1.6 farther in distance. While extinction in the  $H$ -band is much lower than in the optical, dust can still significantly reduce the survey depth in the inner disk and bulge in the direction of dense clouds. The dust distribution in the Galactic mid-plane is patchy on the  $3^\circ$  scale of the 2.5-m field-of-view, which increases the incidence of lower reddening windows in the targeted fields. Even where  $A_V = 10$ , APOGEE should be able to probe to the far edge of the Galactic disk with RGB tip stars in ten hours of integration.

Deriving chemical abundances from  $10^5$  high-resolution spectra presents a major analysis challenge. We are developing an optimal spectral extraction and calibration pipeline and a stellar parameters and chemical abundance pipeline, and we have created realistic simulated data to test these pipelines in advance of APOGEE observations. The extraction pipeline performs a number of tasks, including bundling of hundreds of up-the-ramp detector reads for each pixel,<sup>90</sup>

using the up-the-ramp detector operation to correct for cosmic-rays and pixel saturation, and performing sky subtraction, two-dimensional to one-dimensional spectral extraction, wavelength calibration, combination of dithered exposure pairs into single, fully sampled spectra, telluric absorption correction, flux calibration, and the measurement of stellar radial velocities. Spectra from multiple visits are individually corrected to rest-frame wavelengths and optimally combined for each star.

Parameter and abundance determination then proceeds in two stages. First, spectral fitting based on  $\chi^2$  minimization is used to constrain the stellar temperature ( $T_{\text{eff}}$ ), the surface gravity ( $\log g$ ), the microturbulence parameter ( $\xi$ ), and the abundances of elements that have an important effect on stellar atmospheric structure —  $[\text{Fe}/\text{H}]$ ,  $[\text{C}/\text{H}]$ , and  $[\text{O}/\text{H}]$ . Other abundances are then determined by one-at-a-time  $\chi^2$  minimization with the former parameters held fixed. We plan on using Kurucz (Castelli & Kurucz 2004), MARCS (Gustafsson et al. 2008), and, for the warmest targets, TLUSTY (Lanz & Hubeny 2007) model atmospheres. A variety of literature sources are being used to create and cross-check  $H$ -band line lists, using theoretical or laboratory  $gf$ -values when they are available, and otherwise inferring semi-empirical “astrophysical”  $gf$ -values from fitting synthetic spectra to high-resolution observations of standard stars, such as the Sun and Arcturus.

Figure 10 shows a simulated spectrum at APOGEE resolution, sampling, and S/N for two giant stars with  $T_{\text{eff}}=4000$  K,  $\log g=1.0$  and  $[\text{Fe}/\text{H}] = -1.0$  (solid line) and 0.0 (gray line). The spectral regions displayed (totaling only  $\sim 15\%$  of the total APOGEE spectral range) sample some of the absorption lines that will be used to determine elemental abundances from APOGEE spectra, including Fe, the key metals C, N, and O (which will be determined from OH, CO, and CN lines), Al, Mn, Co, and several  $\alpha$  elements (Mg, Si, Ca, Ti).

## 5. MARVELS

Over the last 15 years, the study of extra-solar planets has advanced from first discoveries (Wolszczan & Frail 1992; Mayor & Queloz 1995; Marcy & Butler 1996) to large surveys that have revealed an astonishing diversity of planetary systems. The standard, core accretion scenario of giant planet formation (see, e.g., Lissauer 1987) predicts that planets like Jupiter form in nearly circular orbits in the region beyond the “snow line” in the protoplanetary disk where ices are stable, corresponding to orbital periods of several years or more for solar-type stars. The two greatest surprises of extra-solar planetary discoveries to date have been that many giant planets have periods below one year, sometimes as short as one day, and that many of these planets are on highly eccentric rather than circular orbits. The first finding suggests that many giant planets “migrate” inward after their formation, while the second suggests that some dynamical mechanism must excite

be operated with non-destructive readouts. APOGEE will read the collected charge in each pixel at regular, frequent intervals (“up-the-ramp sampling”) throughout the integration and use this to reduce readout noise and to monitor pixels for the incidence of cosmic rays, the onset of pixel-well charge saturation, and the progress of each integration in the presence of variable observing conditions.

<sup>90</sup> The HAWAII-2RG detectors in the APOGEE instrument can



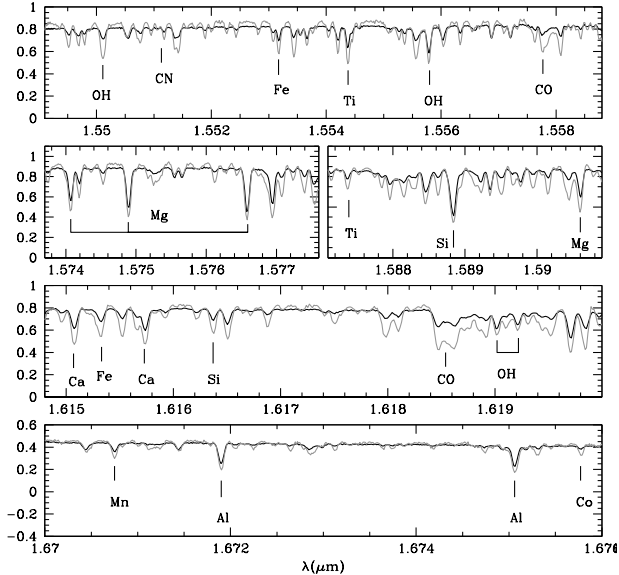


FIG. 10.— Selected sections of simulated APOGEE spectra for two giant stars with  $T_{\text{eff}}=4000$  K,  $\log g=1.0$  and  $[\text{Fe}/\text{H}]=-1.0$  (solid line) and  $0.0$  (gray line) showing some of the lines that will be used to measure element abundances. The vertical scale is in  $F_{\lambda}$  units ( $\text{erg cm}^{-2} \text{s}^{-1} \text{\AA}^{-1}$ ), with an arbitrary normalization. The plotted regions cover only  $\sim 15\%$  of the full APOGEE spectral range.

the planetary eccentricities, probably after the proto-planetary disk has dispersed. Various mechanisms have been proposed to explain planetary migration and the broad eccentricity distribution. These include “smooth” migration via interaction with the proto-planetary gas or planetesimal disk (Lin, Bodenheimer, & Richardson 1996; Murray et al. 1998), violent migration via dynamical processes such as planet-planet scattering (Rasio & Ford 1996; Weidenschilling & Marzari 1996; Ford & Rasio 2008; Jurić & Tremaine 2008), eccentricity pumping via the Kozai mechanism (Holman et al. 1997; Wu & Murray 2003; Fabrycky & Tremaine 2007), and tidal circularization of the highest eccentricity systems to explain the shortest-period giant planets (Wu & Murray 2003; Fabrycky & Tremaine 2007; Nagasawa et al. 2008). It is unclear which, if any, of these mechanisms dominate, and what fraction of systems escape migration and remain on nearly circular orbits.

A large and well characterized sample of giant planets with periods less than a few years is essential for solving the riddles of migration and orbital eccentricities. When coupled with detailed *ab initio* simulations of planet formation, the various proposed migration mechanisms make different predictions for the resulting (post-migration) distributions of planet masses, semi-major axes, and eccentricities. Comparison to the observed distribution of these properties thereby constrains the physical processes involved in planet formation and migration. With the largest homogeneous and statistically complete sample of planets that is currently available (Cumming et al. 2008), it is possible to place constraints on a few specific or extreme migration scenarios (e.g., Schlaufman et al. 2009), but a substantially larger sample is needed to draw strong conclusions (Armitage 2007).

The new generation of planet search experiments are using a variety of technical approaches — high-precision radial velocities (RV), transits, microlensing — to push forward along several distinct dimensions of parameter space, including lower masses and longer periods. Uniquely among these experiments, MARVELS focuses on greatly expanding the target sample for giant planets (roughly speaking, Jupiter mass and larger) in the short-to-intermediate period regime that is most critical for understanding migration and dynamical interaction. It exploits the novel capabilities of fixed-delay dispersed interferometers, which combine interferometers with moderate-resolution spectrographs to enable precision RV measurements with high throughput and a relatively small amount of detector real estate (Erskine & Ge 2000; Ge 2002; Ge, Erskine, & Rushford 2002). This method enables highly multiplexed, multi-fiber searches on moderate aperture telescopes (Ge 2002; Ge et al. 2003), allowing MARVELS to move much further than previous experiments towards large and relatively unbiased target samples.

The basic principles of RV measurement with a dispersed fixed-delay interferometer (DFDI) are reviewed by Ge (2002) and van Eyken, Ge, & Mahadevan (2010). The MARVELS instrument (Ge et al. 2009) uses a monolithic Michelson interferometer and an optical spectrograph with wavelength range  $\lambda \sim 5000-5700 \text{\AA}$ , a  $4k \times 4k$  CCD camera, and resolving power  $R \sim 11,000$ . The two-output interferometer, fed by fibers that subtend  $1''.8$  on the sky, produces two fringing spectra (“beams”) for each of the 60 objects observed in a given exposure. Environmental stabilization keeps temperature drifts below  $\sim 5$  mK during a typical night. The corresponding radial velocity drift is less than  $20 \text{ms}^{-1}$  within a day without any RV drift calibration. Because of this stability, no iodine cell is needed in the stellar beam path during the science exposures. Instead, spectra of a ThAr emission lamp and an iodine absorption cell illuminated by a tungsten continuum lamp are taken before and after each science exposure, and these are used to remove instrumental drifts.

Figure 11 shows an area selected from a MARVELS science exposure, with an expanded region that shows a portion of the fringing spectrum of an individual object, in this case a  $V = 8.5$  star. This region can be compared to the bottom panel of Figure 1 from van Eyken, Ge, & Mahadevan (2010), which shows an idealized case of such a fringing spectrum. The horizontal axis is the wavelength direction, and each spectral line produces a sinusoidally modulated fringe pattern in the vertical direction. The interferometer allows stellar radial velocity signals to be converted from the dispersion direction at very high frequency ( $\sim 2 \text{pixel}^{-1}$ ) to the cross-dispersion direction at low frequency ( $\sim 0.1 \text{pixel}^{-1}$ ) in the MARVELS instrument. Therefore, small shifts in the wavelength direction produce relatively large vertical shifts of these fringes that can be precisely measured. A  $30 \text{ms}^{-1}$  RV change shifts the vertical position of the fringes by  $\sim 0.01 \text{pixel}$ , and it is this mean vertical shift that must be measured by the data pipeline to extract radial velocities (after removing the much larger but computable effects from the Earth’s rotation and orbital motion).

TABLE 4  
SUMMARY OF MARVELS

<i>Duration:</i>	Fall 2008 - Summer 2014, bright time
<i>Spectra:</i>	Dispersed fixed-delay interferometer spectrograph
	60 fibers per 7 deg <sup>2</sup> field (may increase to 120 from Fall 2011)
	5000Å < λ < 5700Å
	$R \equiv \lambda/\Delta\lambda \approx 11,000$
<i>Targets:</i>	8400 FGK stars, $8 \leq V \leq 12.5$
	10% giants
	24 epochs per star, spread over 2-4 years
<i>RV Precision:</i>	10.5 ms <sup>-1</sup> ( $V \leq 9$ )
	22 ms <sup>-1</sup> ( $V = 10$ )
	35 ms <sup>-1</sup> ( $V = 11$ )
	45 ms <sup>-1</sup> ( $V = 11.5$ )

NOTE. — Number of targets assumes 120 fibers from Fall 2011, which will increase the magnitude limit from the current  $V \approx 12$  to  $V \approx 12.5$ . Quoted precision goals are  $1.3 \times$  median photon noise from Years 1+2.

MARVELS aims to survey  $\sim 8400$  stars in the apparent magnitude range  $V = 8 - 12$ , visiting each star approximately 24 times over a two to four year interval. MARVELS began operations in Fall 2008 with a 60-fiber instrument known as the W. M. Keck Exoplanet Tracker (ET). We hope to augment the survey with a second, similar instrument (ET2) before Fall 2011, but we have not yet finalized the funding required to do so. The principal high-level goal of MARVELS is to produce a statistically well defined sample of  $\sim 100$  giant planets with periods up to two years, drawn from a large sample of host stars that have well understood selection biases and encompass a wide range of stellar properties. This data set will be suitable for revealing the diversity in giant exoplanet populations and for testing models of the formation, migration, and dynamical evolution of giant planet systems. In addition, the large stellar sample of MARVELS makes it sensitive to populations of rare systems, which are often signposts of the physical processes at work in planet formation or migration, including very hot Jupiters ( $P < 3$  days), short-period super-massive planets ( $P < 10$  days,  $M \sim 5 - 15 M_J$ ), short-period eccentric planets, planets in extremely eccentric orbits, planets orbiting low metallicity stars, and rapidly interacting multiple planet systems. The systems in which MARVELS identifies giant planets are ideal targets for systematic follow-up campaigns at higher RV precision to quantify the frequency of lower mass or longer period companions in multiple planet systems. Finally, the large size and homogeneity of the target sample makes MARVELS an ideal experiment for quantifying the emptiness of the “brown dwarf desert” at masses  $M \sim 13 - 80 M_{Jup}$  (Grether & Lineweaver 2006) and a unique resource for studying short and intermediate period binary star populations.

During the first two seasons of operation and the first four months of the third season (through December 2010), MARVELS targeted 2820 stars in 47 fields chosen to allow good time coverage across the sky, to have sufficient numbers of stars in the  $8 \leq V \leq 12$  magnitude range, and to have several fields within the *Kepler* (Batalha et al. 2006) survey footprint (see Figure 12). The median number of epochs for these fields was 26,

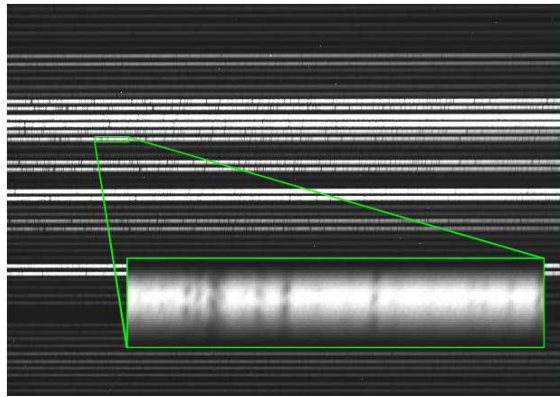


FIG. 11.— Selected region of a MARVELS science exposure. Each horizontal stripe represents one of the two interferometer beam outputs for one of the 60 targets. The expanded region shows a section of the fringing spectrum of HIP 14810 (Wright et al. 2009), a G5 dwarf with  $V = 8.5$ , over the wavelength range  $5293 \text{ \AA} < \lambda < 5316 \text{ \AA}$  (150 wavelength channels out of 4096 in the spectrum). Each spectral line is broken into a series of dark features that represent the minima of the sinusoidal fringe pattern. The shifts in the vertical position of these fringes, fitted over all lines in the spectrum, are measured to extract precision radial velocities.

with a subset of 2580 stars in 43 fields having at least eighteen observation epochs. Figure 12 shows the distribution of observation epochs for each of the target fields. For the remainder of Year 3 and through Year 6, we have selected 58 fields for co-observation with APOGEE; we will not revisit the Year 1+2 fields.

Stars are selected from cross-matched combinations of the NOMAD (Zacharias et al. 2004), UCAC3 (Zacharias et al. 2010), GSC2.3 (Lasker et al. 2008), and 2MASS (Skrutskie et al. 2006) catalogs. Giant stars ( $\log g < 3.5$ ) are separated from dwarfs and subgiants using a reduced proper motion (RPM) diagram,  $RPM_J$  vs.  $J - H$ , using the separation criterion from Collier Cameron et al. (2007). We select the six brightest giants in each field with  $4300 \text{ K} < T_{\text{eff}} < 5100 \text{ K}$ , corresponding to spectral types K2 to G5. We exclude dwarfs with  $T_{\text{eff}} > 6250 \text{ K}$ , which are generally rotating too rapidly and have too few lines to measure precise radial velocities using our instrumentation. Because the brightest dwarfs in a magnitude-selected sample are predominantly earlier spectral types, we require that a maximum of 40% of our dwarfs in a field have  $5800 \text{ K} < T_{\text{eff}} < 6250 \text{ K}$ , corresponding to spectral types G0 to F7. We populate our target list in a given 7 deg<sup>2</sup> field by adding the brightest dwarfs until we have 22 from the G0 to F7 set, then continue in order of decreasing brightness but selecting only those dwarfs with  $T_{\text{eff}} < 5800 \text{ K}$ . Note that for the fields targeted in Years 1+2, we used somewhat different target selection criteria, and also used spectroscopic observations (with the SDSS spectrographs) for giant-dwarf separation. Unfortunately, the SDSS spectra (reduced with an earlier version of the SSPP) proved less effective in separating giants and dwarfs than we had expected, leaving us with a 30% giant fraction compared to our original goal of 10%. The  $RPM_J$  selection we have now implemented should resolve this problem.

MARVELS observes with 50-minute science exposures (which will increase to 60 minutes once co-observing with APOGEE begins) and  $\sim 10$  minutes of overhead per ex-

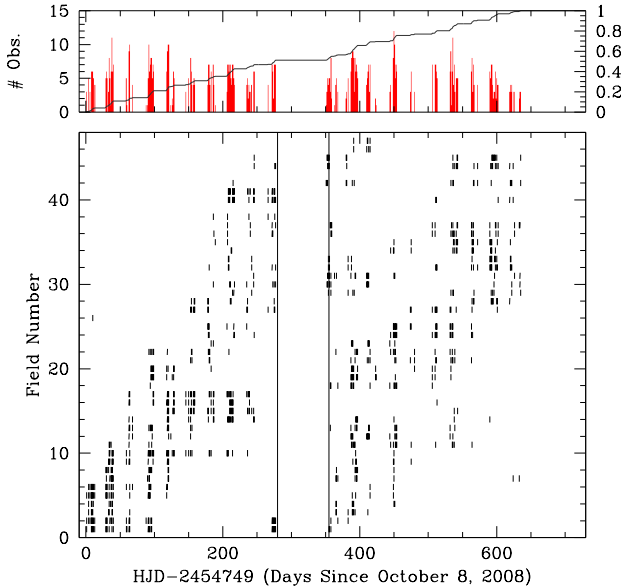


FIG. 12.— (*Bottom*) The distribution of the epochs for each of the 47 MARVELS target fields during Years 1 and 2. Vertical lines mark the summer shutdown. (*Top*) The histogram shows the number of observations as a function of Heliocentric Julian Date (left axis). The line shows the cumulative fraction of the total number of observations as function of HJD (right axis).

posure. The fields for any given night are selected based on observability, the number of previous epochs, and the time since the most recent epoch. The photon-noise limited RV precision for a MARVELS observation depends most strongly on stellar apparent magnitude, but also on other factors that affect fringe visibility including rotation and metallicity. For observations during the first two years, the median RV photon-noise limited precision is approximately  $5 \text{ m s}^{-1}$  at  $V = 8.5$ ,  $8 \text{ m s}^{-1}$  at  $V = 9.0$ ,  $17 \text{ m s}^{-1}$  at  $V = 10.0$ ,  $27 \text{ m s}^{-1}$  at  $V = 11.0$ , and  $35 \text{ m s}^{-1}$  at  $V = 11.5$ . The median rms RV of target stars observed over a one-month timescale is 2 – 3 times the photon noise at bright magnitudes and 1 – 1.5 times the photon noise at  $V \geq 10$ . The current MARVELS data pipeline has shortcomings that lead to worse performance over timescales of several months. We are working on improvements as of this writing, with the eventual goal of achieving total errors within a factor 1.3 of the photon noise errors. The precision goals in Table 4 are defined by this target, using the photon-noise numbers quoted above. We are also implementing changes to the fiber system and data pipeline that we expect to lower the photon noise itself by a moderate factor (e.g., changing from iodine to ThAr calibration, which will allow us to use 100% of the spectral range rather than the current 75%). We therefore regard it as plausible that the final MARVELS RV performance will improve on the goals in Table 4 by  $\sim 30\%$ .

The velocity semi-amplitude of a star of mass  $M_*$  orbited by a planet of mass  $M_p$  with a period  $P$  and inclination  $i$  is

$$K = 28.4 \text{ m s}^{-1} (M_p \sin i / M_{\text{Jup}}) (P / 1 \text{ yr})^{-1/3} (M_* / M_{\odot})^{-2/3}. \quad (6)$$

For  $N$  observations and an rms error of  $\sigma$ , achieving less than one false detection for  $\sim 10^4$  stars requires a total signal-to-noise ratio threshold of  $\sim \sqrt{N/2}(K/\sigma) \sim 13$  (Cumming 2004), or  $\sigma \lesssim K/4$  assuming 24 observations per star. A simple and somewhat conservative forecast, described in Appendix B, indicates that MARVELS should detect approximately 66 planets with  $P < 2$  year and  $M_p < 10 M_{\text{Jup}}$  if the total errors can be reduced to 1.3 times the median photon noise achieved in Year 1+2 data. Of these 66 predicted planets, 53 have periods of  $< 1$  year. The yield falls to 41 planets if the errors are 2.0 times the median photon noise, and it rises to  $\sim 86$  planets if the errors are equal to the median photon noise. With the above mentioned changes to the fiber system and data processing techniques, we may be able to significantly lower the photon noise floor, which could increase the yields by as much as 30%.

These forecasts assume a second MARVELS instrument operating for the final three years of the survey; without it, anticipated planet yields fall by  $\sim 20\%$  (not 50% because it would be observing fainter stars). The predictions are based on a false alarm probability of  $\sim 3 \times 10^{-4}$ , for which we would expect  $\sim 3$  false positives. For a more conservative false alarm probability of  $\sim 3 \times 10^{-5}$ , the yields decrease by  $\sim 15\%$ . We note that we have not attempted to estimate the yield from our sample of giant stars, which constitute  $\sim 10\%$  of our targets, as considerably less is known about the frequency of planets around these systems. Nevertheless, we can reasonably expect to detect additional planets from this sample. Finally, we have only included companions with masses of  $< 10 M_{\text{Jup}}$  in our tally. Extrapolating the planet distribution function found by Cumming et al. (2008) up to larger masses, we estimate that we would detect an additional  $\sim 14$  planets with  $10 M_{\text{Jup}} < M_p < 15 M_{\text{Jup}}$  and periods of  $< 2$  years under the assumption of errors equal to  $1.3 \times$  the photon noise. Because the observing strategy, target selection, and noise characteristics for MARVELS are very well specified, statistical models of the planet population (specifying, e.g., the distribution of masses, periods, and eccentricities as a function of host properties) can be tested statistically against the MARVELS RV measurements even without one-by-one identification of planets.

Figure 13 shows the MARVELS RV curve for the short-period brown dwarf candidate discovered by MARVELS around the star TYC-1240-00945-1 (Lee et al. 2010a). RV measurements from the two interferometer beams are shown separately, as well as measurements from observations with the Hobby-Eberly Telescope (HET) and the SMARTS 1.5m echelle. Supplementary photometric and spectroscopic studies show that the host is a slightly evolved, solar-type star, with an estimated mass of  $1.35 M_{\odot}$  and age of  $\approx 3.0$  Gyr. In this case, the low-mass companion (“MARVELS-1b”) is a likely brown dwarf with minimum mass of  $28.0 \pm 1.5 M_{\text{Jup}}$  at an orbital separation of  $0.071 \pm 0.002$  AU ( $P_{\text{orb}} = 5.9$  days), placing this object squarely within the “brown dwarf desert”. Indeed, MARVELS has already found several more such brown-dwarf “desert dwellers” in the first two-year dataset, enough to robustly quantify the aridity of the desert. The MARVELS team plans similarly detailed follow-up studies and characterization of all MARVELS

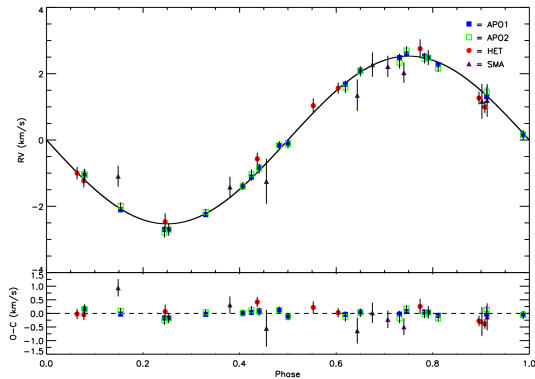


FIG. 13.— The radial velocity curve of the first MARVELS RV companion, a probable brown-dwarf in a 5.9-day orbit around the star TYC 1240-00945-1 (Lee et al. 2010b). Filled and open squares are measurements from the two interferometer outputs of the MARVELS spectrograph. Circles and triangles are measurements from follow-up observations from the Hobby-Eberly Telescope and the SMARTS 1.5m telescope, respectively. Lower panel shows the difference between observed velocities and those calculated from the best-fit model.

hosts and control samples of target stars that did not yield planets. This approach will enable full investigation of the dependence of giant exoplanet and brown dwarf populations on host star properties, including chemical abundances, mass, and evolutionary status. In addition, MARVELS will provide robust statistics on spectroscopic binary star populations, and it will yield a novel sample of eclipsing binary star systems *discovered* spectroscopically.

## 6. SCIENCE ORGANIZATION

An effective collaboration culture is crucial to the successful execution of a large project like the SDSS. Indeed, developing this culture was itself one of the major challenges and significant achievements of SDSS-I. The organization of SDSS-III is, of course, closely modeled on that of SDSS-I and II. We briefly describe this organization here, as it may be of value to those using the SDSS-III data sets and science analyses and to others planning comparably ambitious projects.

Like its predecessors, SDSS-III is being carried out by a large and diverse international collaboration. A wide variety of institutions have joined the project by means of financial or equivalent in-kind contributions, and they all agree to a written set of “Principles of Operation”<sup>91</sup> that serves as the defining policy document of the project. At *Full Member* institutions, all faculty, PhD research staff, and students have access rights to all SDSS-III data and activities. *Associate Member* institutions join with smaller, designated groups of faculty and postdocs. *Participation Groups* are consortia of designated researchers from multiple institutions that act as single members within the SDSS Collaboration. Finally, particular individuals are named as *External Participants* based on their contributions to the SDSS-III project. An up-to-date listing of all the institutions in SDSS-III can be found on the SDSS-III website (<http://www.sdss3.org>). The Apache Point Observatory and the Sloan Foundation Telescope are both owned

by the Astrophysical Research Consortium (ARC), and the ARC Board of Governors has financial authority for the SDSS-III. The Principles of Operation establish an Advisory Council to oversee the survey and to represent the collaboration to the ARC Board of Governors. The Advisory Council consists of one voting member from each full member, participation group, and associate member of sufficient group size.

Figure 14 presents the high-level SDSS-III organization chart, including the individuals who currently hold the indicated positions. Each of the four surveys has its own technical and science team; the number of people who have already made large contributions (i.e., many person-months or more) to the design and execution of the individual surveys ranges from  $\sim 15$  to  $\sim 60$ , with still larger numbers joining in data analysis and quality assurance. The Principal Investigator (PI) of each survey oversees all aspects of the survey’s construction and execution, with the assistance of the Survey Scientist, who has the primary responsibility for defining science requirements and ensuring that the survey data ultimately meet those requirements. BOSS, MARVELS, and APOGEE each have an Instrument Scientist who oversees the design, construction, commissioning, and maintenance of the new instrumentation. The APOGEE spectrograph development has its own Project Manager, as did the MARVELS ET1 spectrograph development prior to delivery of the instrument.

Many tasks, including overall project budgeting and management, span all four surveys of SDSS-III. Organizational responsibility for these tasks lies with the central project office, headed by the Director. The Infrastructure and Observatory Operations teams have responsibility for common facilities (telescope, fiber systems, operations software, etc.) and for performing the observations themselves. The Data Coordinator is responsible for integrating data from the four surveys into the science data archive, the basis both for collaboration science and for public data releases. The Project Spokesperson, elected by the collaboration, promotes scientific coordination within the collaboration and external visibility of SDSS-III in the astronomical community and beyond. The Spokesperson chairs the Collaboration Council, comprised of representatives from all voting institutions, which organizes collaboration meetings and develops and implements collaboration policies, most notably on publications and external collaborators.<sup>92</sup> Over the years, collaborations with non-participants on pre-publication data have been a vital mechanism for bringing additional expertise and resources into SDSS science analyses, and reviewing external collaborator proposals is one of the Collaboration Council’s most important tasks.

The guiding principle of the SDSS-III science collaboration is that all participants have the right to pursue any project they wish with SDSS-III data, but they are required to notify the entire collaboration of their plans and to update them as projects progress. Groups pursuing similar science projects are encouraged to collaborate, but they are not required to do so. There is no binding internal refereeing process, but draft publications using non-public data must be posted to the whole collabora-

<sup>91</sup> Available at <http://www.sdss3.org/collaboration/poo3.pdf>

<sup>92</sup> See <http://www.sdss3.org/collaboration/policies.php>.

## SDSS-III Organization Chart

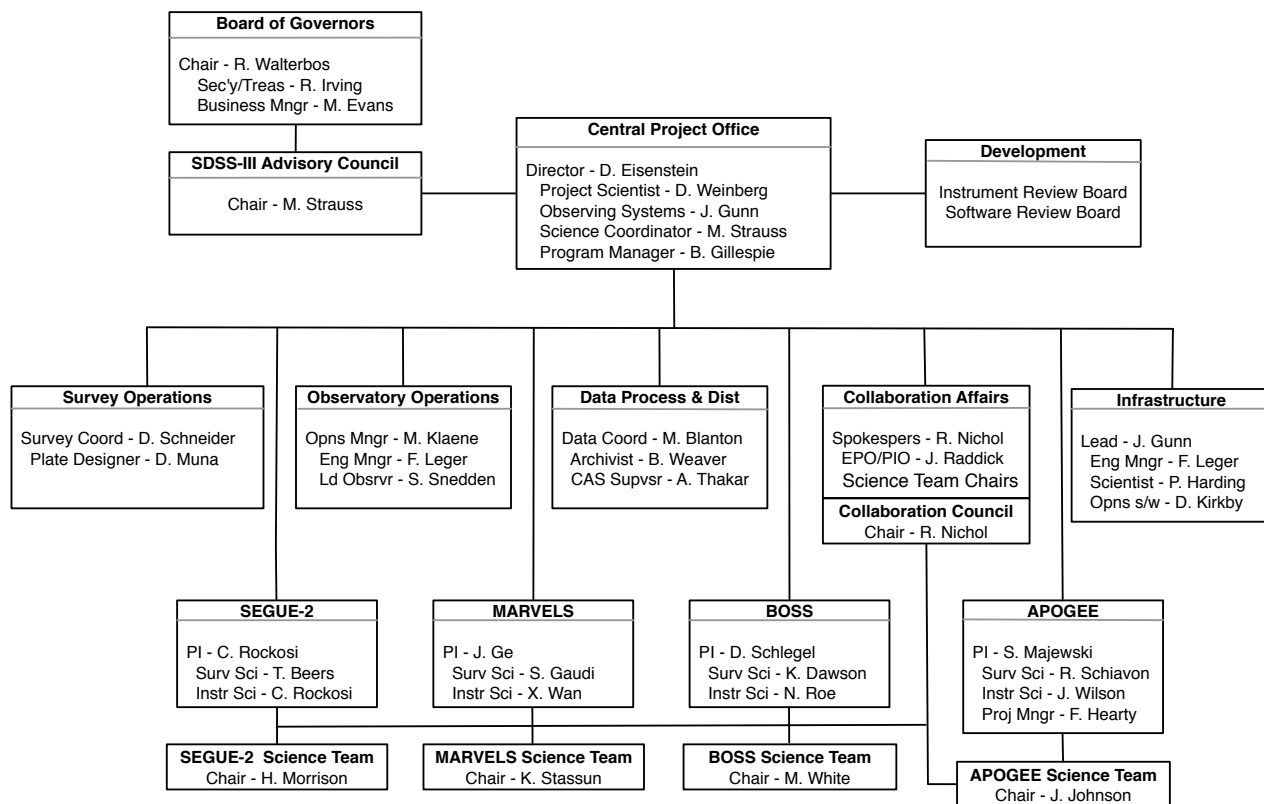


FIG. 14.— The high-level organizational chart for SDSS-III. Named individuals are those filling these positions as of January 1, 2011. As is evident from the author list of this paper, this chart represents the tip of a very large iceberg of SDSS-III contributors.

tion for a review period of at least three weeks prior to submission to any journal or online archive. Manuscripts often undergo significant revision and improvement during this period. Participants outside of the core analysis team may request co-authorship on a paper if they played a significant role in producing the data or analysis tools that enabled it. In particular, scientists who have contributed at least one year of effort to SDSS-III infrastructure development or operations can request “Architect” status, which entitles them to request co-authorship on any science publications for those surveys to which they contributed. All SDSS-III authorship requests are expected to comply with the professional guidelines of the American Physical Society.<sup>93</sup>

Each of the four SDSS-III surveys has its own survey science team (SST), headed by the SST Chair, whose role is to coordinate and promote effective scientific collaboration within the team. Naturally, the science team overlaps and interacts with the survey’s technical team, but the latter is focused on producing the data, while the former is focused on science analysis; data quality assurance is a shared responsibility. Inside an SST, participants may coordinate their efforts on more focused topics via working groups; for example, the BOSS SST presently

has working groups in galaxy clustering, galaxy evolution, Ly $\alpha$  forest cosmology, physics of the intergalactic medium, and quasars. The working groups communicate and collaborate through archived e-mail lists, wiki pages, regular teleconferences, and in-person meetings. As a result, many SDSS papers benefit from the combined efforts and knowledge of many collaborators on the science analysis, as well as on the production of the data that enables it.

Over the years, all of these strategies — representative governing bodies, centralized management overseeing the many technical teams, well defined policies and structures that encourage widespread scientific participation, and extensive communication mechanisms — have proven essential to the successful execution of the SDSS and to producing an enormous range of science from its surveys.

### 7. CONCLUDING REMARKS

The SDSS has demonstrated the extraordinary scientific reach of a moderate aperture telescope equipped with powerful wide-field instruments and operated in an efficient “survey mode,” including sophisticated data pipelines that produce well calibrated and readily usable public data sets. There are now several astronomical imaging cameras in operation or under construction

<sup>93</sup> <http://www.aps.org/policy/statements/02-2.cfm>



that exceed the pixel count of the SDSS imager, which has been officially retired after completing the additional southern imaging for BOSS. However, the Sloan Telescope remains an exceptionally productive facility for wide-field spectroscopic surveys.

The four SDSS-III surveys exploit this capability efficiently to address a wide range of science goals. The BOSS spectroscopic survey requires five years of dark time to cover its 10,000 deg<sup>2</sup> survey area. SEGUE-2 and the BOSS imaging survey used the one year of dark time that was available between the end of SDSS-II and the completion of the BOSS spectrograph upgrades. MARVELS began bright-time operations in the first fall season of SDSS-III. MARVELS and APOGEE will share the focal plane for 75% of the bright-time observing from 2011-2014, allowing each survey to amass a considerably larger sample than it could with a 50% share.

The BOSS galaxy redshift survey will achieve BAO distance scale constraints that are close to the limit set by cosmic variance out to  $z = 0.6$ ; the only substantial (factor-of-two) improvement possible at these redshifts would be to cover the remaining  $3\pi$  steradians of the sky. The BOSS Ly $\alpha$  forest survey will pioneer a new method of measuring 3-dimensional structure in the high-redshift universe and provide the first BAO measurements of distance and expansion rate at  $z > 2$ . Together SEGUE and APOGEE will provide powerful new insights into the formation history and present day structure of the Milky Way. The depth and large sample size of the SEGUE surveys make them especially valuable for characterizing kinematic and chemical structure in the outer Galaxy. The dust-penetrating capacity of APOGEE's infrared observations will make it the first large spectroscopic survey of all Galactic stellar populations, and its high resolution and high precision allow detailed chemical fingerprinting of an enormous sample, orders of magnitude larger than any high resolution sample that exists today. The large sample of stars monitored by the MARVELS RV survey gives it sensitivity to rare planetary systems that are signposts for underlying physical processes, and the combination of sample size and systematic observing strategy will make it a uniquely valuable data set for testing theories of giant planet formation, migration, and dynamical interaction.

Current investigations with SDSS-III data span a

vast range of scales and redshifts, including studies of large-scale structure with massive galaxies and three-dimensional Ly $\alpha$  forest correlations, searches for kinematic and chemical substructure in the Galactic halo and thick disk, and measurement of the incidence of short-period brown dwarf companions to solar-type stars. SDSS-III will continue the long-standing SDSS tradition of public data releases, beginning with the SDSS Eighth Data Release (DR8), which is now available (Aihara et al. 2011). DR8 includes all of the new imaging carried out for BOSS and all of the spectra taken for SEGUE-2. It also incorporates all SDSS-I and II data, processed with the latest versions of our data reduction and calibration pipelines, so that science analyses can incorporate data from all SDSS surveys in a seamless and internally consistent manner. The final data SDSS-III release is scheduled for the end of 2014. Like their predecessors, we anticipate that BOSS, SEGUE-2, MARVELS, and APOGEE will have deep and wide-ranging impacts on many fields of contemporary astronomy and cosmology.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, University of Florida, the French Participation Group, the German Participation Group, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

*Facilities:* Sloan

## APPENDIX

### FORECASTS FOR BOSS

In Table 5 we forecast the constraints on a number of cosmological parameters. To obtain these numbers we first convert our observational parameters into errors on the line-of-sight [ $H(z)$ ] and transverse [ $d_A(z)$ ] distances as a function of redshift using the method of Seo & Eisenstein (2007). This Fisher matrix calculation uses only acoustic oscillation information and no broad-band power, so we believe the error estimates to be robust (and conservative, see the discussion below and the final line of Table 5). To approximate the effects of (partial) reconstruction (Eisenstein et al. 2007a) we suppress the nonlinear smearing ( $\Sigma_{\perp,||}$  in the notation of Eisenstein, Seo, & White 2007b) by a factor of two for the LRG calculation. We use a similar Fisher matrix calculation (McDonald & Eisenstein 2007) to estimate the distance errors that one would obtain for the Ly $\alpha$  forest survey, with no attempt at reconstruction because of the very sparse sampling of the density field. We find errors on  $d_A$  of 1.0% at  $z = 0.35$ , 1.0% at  $z = 0.6$ , and 4.5% at  $z = 2.5$ , with errors on  $H(z)$  of 1.8%, 1.7% and 2.6% at the same redshifts. As noted earlier, current theoretical studies suggest that any shifts in the BAO scale due to non-linearity or galaxy bias are at or below this level. With further work, we should be able to calculate any corrections to a level of accuracy that keeps systematic errors well below these statistical errors.

The constraints on  $d_A$  and  $H$  are then used in a Markov chain to get constraints on the matter density  $\omega_m \equiv \Omega_m h^2$ , the baryon density  $\omega_b \equiv \Omega_b h^2$ , the dark energy density  $\Omega_\chi$ ,  $w_0$ ,  $w_a$ , and the curvature,  $\Omega_K$ , where the dark energy



TABLE 5  
BOSS PARAMETER CONSTRAINT FORECASTS

Expt.	$h$	$\Omega_K$	$w_0$	$w_p$	$w_a$	FoM
Planck+Stage II	0.019	0.0031	0.115	0.036	0.524	53.4
BOSS LRG BAO	0.009	0.0027	0.090	0.031	0.365	87.4
<b>BOSS LRG BAO+LyaF BAO</b>	0.009	0.0019	0.083	0.030	0.320	<b>102</b>
BOSS LRG broad-band+LyaFBAO	0.007	0.0018	0.074	0.019	0.284	188

NOTE. — All constraints assume Planck and the DETF forecasts for “Stage II” experiments. BAO constraints include only the acoustic scale information and are therefore conservative; the final line shows the BOSS forecast that also incorporates broad-band galaxy power information.

equation-of-state parameter is assumed to evolve with expansion factor  $a(t)$  as  $w(a) = w_0 + w_a(1 - a)$ . In addition to the distance constraints from BAO experiments, we apply a (Gaussian) prior on these parameters equal to the Fisher matrices for Planck, HST, and Stage II experiments presented in the technical appendix of the DETF report. The Markov chain allows us to explore the full likelihood surface, without making a “quadratic” assumption like Fisher matrix methods do. This is important when some parameters (e.g.  $w_a$ ) are not well constrained. The variance of each parameter, computed from the chains, is given in Table 5. We also quote the DETF Figure of Merit, which is the inverse of the area of the 95% confidence level region in the  $w_p - w_a$  plane (scaled to correspond to the convention adopted by the DETF). The precise value of the pivot expansion factor depends on which experiments are considered, but it is generally  $a_p \approx 0.8$ , i.e., in this family of models the dark energy experiments best constrain the value of  $w$  at  $z \approx 0.25$ .

#### FORECASTS FOR MARVELS

To estimate the MARVELS survey planet yield, we use a simple, essentially analytic, method. We adopt the following assumptions:

- (1) We consider 9 bins in  $V$  magnitude, from  $V_i = 8$  to  $V_i = 12$  in steps of 0.5 magnitude, where  $i$  is the bin index. For stars monitored in Years 1+2 including the 4 month extension in Year 3, we use the actual distribution of the number of stars per bin  $f_{*,i}$  from a representative subset of the target fields. For stars monitored in Years 3-6, we use an average distribution of  $V$  magnitudes for dwarf stars contained in the 31 preliminary shared fields chosen by APOGEE. We assume that the current MARVELS spectrograph monitors the brightest 60 stars and the second instrument (assumed to begin operation in Fall 2011) monitors the next brightest 60 stars.
- (2) The total number of stars  $N_*$  monitored is set by the number of observations  $N_{\text{obs}}$  per field, the total number of epochs  $N_e$  available per month, the total number of available months  $N_m$  for each observing block, and the fraction of time  $f_{\text{lost}}$  lost to non-MARVELS science (APOGEE commissioning, APOGEE-only fields, etc.),

$$N_* = N_{\text{fiber}} \times \frac{N_e N_m}{N_{\text{obs}}} \times (1 - f_{\text{lost}}). \quad (\text{B1})$$

We assume a total of 46 epochs are available per month, based on the 917 epochs obtained over the 20 months of Years 1+2. This effectively means that for Years 3-6 we assume similar weather and a similar exposure plus overhead time ( $\sim 50 + 10$  minutes) for each observation as we adopted in Years 1+2. We take  $f_{\text{lost}} = 0.25$  starting in 2011 and  $f_{\text{lost}} = 0$  before.

- (3) We assume a log-normal distribution of RV uncertainties  $\sigma$  for each magnitude bin,

$$\frac{df_{\sigma,i}}{d\sigma} \propto e^{-0.5[(\log \sigma - \log \sigma_i)/0.2]^2}, \quad (\text{B2})$$

where the  $\sigma_i$  are the median photon noise uncertainties and RMS scatter for each  $V$  magnitude bin. The value of 0.2 for the log-normal dispersion was chosen to approximately match the distribution of observed scatter. Our final yields do not depend very strongly on this choice, changing by  $\sim 27\%$  over the range 0.05 – 0.4.

- (4) We adopt a power-law distribution for the planet frequency as a function of period and mass,

$$\frac{dN_p}{d \ln m_p d \ln P} = C(P) \left( \frac{m_p}{M_{\text{Jup}}} \right)^{-0.31}, \quad (\text{B3})$$

where  $C(P) \simeq 0.00186$  for  $P < 300$  days and  $C(P) = 0.0093$  for  $P > 300$  days. We assume that there are no planets with  $P < 10$  days and  $m_p > 2M_{\text{Jup}}$ . This distribution is motivated by the results of Cumming et al. (2008), who fit a continuous power-law over the entire period range of  $P < 2000$  days, finding  $dN_p/d \ln m_p d \ln P \propto m_p^{-0.31} P^{0.26}$ . However, since they have a paucity of intermediate period planets ( $P = 10 - 300$  days) relative to this fit, we have conservatively chosen a uniform distribution in  $\ln P$  and step-function truncations. Adopting the continuous power-law distribution would have led to a higher predicted yield.

- (5) We assume a planet is detectable if its periodogram power  $z$ , or equivalently total signal-to-noise ratio  $Q$ , is larger than a given value  $z_0$ . We use an analytic estimate for  $Q$  as a function of  $N_{\text{obs}}$ , the planet semiamplitude  $K$ , and the

RV uncertainty given by

$$Q \equiv \left( \frac{N_{\text{obs}}}{2} \right)^{1/2} \frac{K}{\sigma}. \quad (\text{B4})$$

This is strictly only appropriate for a uniformly-sampled, circular orbit, but it is a good approximation for eccentricities less than  $\sim 0.6$ , and for planet periods less than the time spanned by the observations. The relationship between the periodogram power  $z$  and  $Q$  for a Keplerian orbit fit is given by (Cumming 2004),

$$z \equiv \frac{Q^2}{2} \left( \frac{N_{\text{obs}} - 5}{N_{\text{obs}}} \right), \quad (\text{B5})$$

where  $N_{\text{obs}} - 5$  is the number of degrees of freedom from a Keplerian fit. We use the following simple analytic estimate for the minimum power  $z_0$  for detection (Cumming 2004),

$$z_0 = \frac{3\nu}{4} \left[ \left( \frac{M}{F} \right)^{2/\nu} - 1 \right], \quad (\text{B6})$$

where  $\nu \equiv N_{\text{obs}} - 5$ ,  $M \simeq T\Delta f$  is the number of independent frequencies (periods) searched for planets,  $T = 2$  years is the span of each set of observations,  $\Delta f \equiv P_{\text{min}}^{-1} - P_{\text{max}}^{-1}$  is the range of frequencies searched, and  $F$  is the false alarm probability required for detection. We set  $F = (3 \times 10^3)^{-1}$ . Given that MARVELS will be surveying  $\mathcal{O}(10^4)$  stars, we therefore expect  $\sim 3$  false positives, which will require follow-up to eliminate.

With these assumptions, we can estimate our yield as,

$$N_{\text{det}} = N_* \sum_{i=1}^9 f_{*,i} \int_{\ln P_{\text{min}}}^{\ln P_{\text{max}}} d \ln P \int_{\ln m_{p,\text{min}}}^{\ln m_{p,\text{max}}} d \ln m_p \frac{dN_p}{d \ln m_p d \ln P} \int_0^\infty d\sigma \frac{df_{\sigma,i}}{d\sigma} \Theta[z(m_p, P, \sigma) - z_0], \quad (\text{B7})$$

where  $\Theta[x]$  is the Heaviside step function. In addition, we exclude from the integrals the region where  $m_p > M_{\text{Jup}}$  and  $P < 10$  days.

#### REFERENCES

- Abazajian, K. N., et al. 2009, *ApJS*, 182, 543  
Aihara, H., et al. 2011, submitted to *ApJS*  
Allende Prieto, C., et al. 2006, *ApJ*, 636, 804  
Allende Prieto, C., et al. 2008, *AJ*, 136, 2070  
Armitage, P. J. 2007, *ApJ*, 665, 1381  
Bailer-Jones, C. A. L. 2000, *A&A*, 357, 197  
Batalha, N. M., Borucki, W., Caldwell, D. A., Chandrasekaran, H., Gautier, T. N., Jenkins, J., & Koch, D. G. 2006, *BAAS*, 38, 1188  
Becker, R. H., White, R. L., & Helfand, D. J. 1995, *ApJ*, 450, 559  
Beers, T. C., Rossi, S., Norris, J. E., Ryan, S. G., & Shefler, T. 1999, *ApJ*, 117, 981  
Bell, E. F., et al. 2008, *ApJ*, 680, 295  
Belokurov, V., et al. 2006, *ApJ*, 642, L137  
Bharadwaj S., 1996, *ApJ*, 472, 1  
Blake, C. & Glazebrook, K. 2003, *ApJ*, 594, 665  
Bolton, A. S., Burles, S., Koopmans, L. V. E., Treu, T., Gavazzi, R., Moustakas, L. A., Wayth, R., & Schlegel, D. J. 2008a, *ApJ*, 682, 964  
Bolton, A. S., Treu, T., Koopmans, L. V. E., Gavazzi, R., Moustakas, L. A., Burles, S., Schlegel, D. J., & Wayth, R. 2008b, *ApJ*, 684, 248  
Bolton, A. S., & Schlegel, D. J. 2010, *PASP*, 122, 248  
Bovy, J., et al. 2010, arXiv:1011.6392  
Brown, W. R., Geller, M. J., Kenyon, S. J., & Kurtz, M. J. 2006, *ApJ*, 647, 303  
Bullock, J. S., Kravtsov, A. V. & Weinberg, D. H. 2001, *ApJ*, 548, 33  
Bullock, J. S., & Johnston, K. V. 2005, *ApJ*, 635, 931  
Castelli, F., & Kurucz, R. L. 2004, Proceedings of the IAU Symp. No 210; Modelling of Stellar Atmospheres, eds. N. Piskunov et al., poster A20  
Cenarro, A. J., Cardiel, N., Gorgas, J., Peletier, R. F., Vazdekis, A., & Prada, F. 2001a, *MNRAS*, 326, 959  
Cenarro, A. J., Gorgas, J., Cardiel, N., Pedraz, S., Peletier, R. F., & Vazdekis, A. 2001b, *MNRAS*, 326, 981  
Churchwell, E., et al. 2009, *PASP*, 121, 213  
Cole, S., et al. 2005, *MNRAS*, 362, 505  
Collier Cameron, A., et al. 2007, *MNRAS*, 380, 1230  
Crocce M. & Scoccimarro R., 2006, PRD73, 063519  
Crocce M. & Scoccimarro R., 2008, PRD77, 023533  
Croom, S. M., et al. 2009, *MNRAS*, 399, 1755  
Cumming, A. 2004, *MNRAS*, 354, 1165  
Cumming, A., Butler, R. P., Marcy, G. W., Vogt, S. S., Wright, J. T., & Fischer, D. A. 2008, *PASP*, 120, 531  
Efstathiou, G., Sutherland, W. J., & Maddox, S. J. 1990, *Nature*, 348, 705  
Eisenstein, D. J., Hu, W., 1998, *ApJ*, 496, 605  
Eisenstein, D. J. et al. 2001, *AJ*, 122, 2267  
Eisenstein, D. J., et al. 2005, *ApJ*, 633, 560  
Eisenstein, D. J., Seo, H.-J., Sirko, E., & Spergel, D. N. 2007a, *ApJ*, 664, 675  
Eisenstein, D. J., Seo, H.-J., & White, M. 2007b, *ApJ*, 664, 660  
Erskine, D. J., & Ge, J. 2000, *Imaging the Universe in Three Dimensions*, 195, 501  
Fabrycky, D., & Tremaine, S. 2007, *ApJ*, 669, 1298  
Fan, X. 1999, *AJ*, 117, 2528  
Ford, E. B., & Rasio, F. A. 2008, *ApJ*, 686, 621  
Freeman, K., & Bland-Hawthorn, J. 2002, *ARA&A*, 40, 487  
Frieman, J. A., Turner, M. S., & Huterer, D. 2008a, *ARA&A*, 46, 385  
Frieman, J. A., et al. 2008b, *AJ*, 135, 338  
Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, *AJ*, 111, 1748  
Fulbright, J. P., et al. 2010, *ApJ*, 724, L104  
Garnett, J. D., et al. 2004, *Proc. SPIE*, 5499, 35  
Gates, E., et al. 2004, *ApJ*, 612, L129  
Ge, J. 2002, *ApJ*, 571, 165  
Ge, J., Erskine, D. J., & Rushford, M. 2002, *PASP*, 114, 1016  
Ge, J., Mahadevan, S., van Eyken, J., Dewitt, C., & Shaklan, S. 2003, *Scientific Frontiers in Research on Extrasolar Planets*, 294, 573  
Ge, J., et al. 2006, *ApJ*, 648, 683  
Ge, J., et al. 2009, *Proc. SPIE*, 7440  
Geisler, D. 1984, *PASP*, 96, 723  
Grether, D., & Lineweaver, C. H. 2006, *ApJ*, 640, 1051  
Grillmair, C. J. 2006, *ApJ*, 645, L37  
Grillmair, C. J., & Dionatos, O. 2006, *ApJ*, 643, L17

- Gunn, J. E., et al. 1998, *AJ*, 116, 3040  
 Gunn, J. E., et al. 2006, *AJ*, 131, 2332  
 Gustafsson, B., Edvardsson, B., Eriksson, K., Jørgensen, U. G., Nordlund, Å., & Plez, B. 2008, *A&A*, 486, 951  
 Harris, H., et al. 2008, *ApJ*, 679, 697  
 Helmi, A., & White, S. D. M. 1999, *MNRAS*, 307, 495  
 Hills, J. G. 1988, *Nature*, 331, 687  
 Holman, M. J., Touma, J., & Tremaine, S. 1997, *Nature*, 386, 254  
 Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, *Nature*, 370, 194  
 Ibata, R., Irwin, M., Lewis, G. F., & Stolte, A. 2001, *ApJ*, 547, L133  
 Ivezić, Željko et al. 2000, *AJ*, 120, 963  
 Ivezić, Ž., et al. 2008, *ApJ*, 684, 287  
 Jurić, M., et al. 2008, *ApJ*, 673, 864  
 Jurić, M., & Tremaine, S. 2008, *ApJ*, 686, 603  
 Kazin, E. A., et al. 2010, *ApJ*, 710, 1444  
 Kofman L.A., Gnedin N.Y., Bahcall N.A., 1993, *ApJ*, 413, 1  
 Kollmeier, J. A., et al. 2010, *ApJ*, 723, 812  
 Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, *ApJ*, 649, 599  
 Koopmans, L. V. E., et al. 2009, *ApJ*, 703, L51  
 Krauss, L. M., & Turner, M. S. 1995, *Gen Rel & Grav*, 27, 1137  
 Lanz, T., & Hubeny, I. 2007, *ApJS*, 169, 83  
 Larson D., et al., 2010, preprint [arXiv:1001.4635]  
 Lasker, B. M., et al. 2008, *AJ*, 136, 735  
 Lawrence, A., et al. 2007, *MNRAS*, 379, 1599  
 Lee, B. L., et al. 2010a, *ApJ*, in press, arXiv:1011.5170  
 Lee, Y. S., et al. 2008a, *AJ*, 136, 2022  
 Lee, Y. S., et al. 2008b, *AJ*, 136, 2050  
 Lee, Y. S., et al. 2010b, *AJ*, in press (ArXiv: 1010.2934)  
 Liddle A.R., Lyth D.H., Viana P.T.P., White M., 1996, *MNRAS*, 282, 281  
 Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. 1996, *Nature*, 380, 606  
 Lissauer, J. J. 1987, *Icarus*, 69, 249  
 Madrid, J.P., & Macchetto, F.D. 2006, *BAAS* 38, 1286  
 Madrid, J.P., & Macchetto, F.D. 2009, *BAAS* 41, 913, arXiv:0901.4552  
 Mahadevan, S., van Eyken, J., Ge, J., DeWitt, C., Fleming, S. W., Cohen, R., Crepp, J., & Vanden Heuvel, A. 2008, *ApJ*, 678, 1505  
 Majewski, S. R., Ostheimer, J. C., Kunkel, W. E., & Patterson, R. J. 2000, *AJ*, 120, 2550  
 Majewski, S. R., Skrutskie, M. F., Weinberg, M. D., & Ostheimer, J. C. 2003, *ApJ*, 599, 1082  
 Majewski, S. R., Zasowski, G., & Nidever, D. J. 2011, *ApJ*, submitted  
 Maraston, C., Strömbäck, G., Thomas, D., Wake, D. A., & Nichol, R. C. 2009, *MNRAS*, 394, L107  
 Marcy, G. W., & Butler, R. P. 1996, *ApJ*, 464, L147  
 Martin, D. C., et al. 2005, *ApJ*, 619, L1  
 Matsubara T., 2008a, *PRD*, 77, 063530  
 Matsubara T., 2008b, *PRD*, 78, 083519  
 Mayor, M., & Queloz, D. 1995, *Nature*, 378, 355  
 McDonald, P., & Eisenstein, D. J. 2007, *Phys. Rev. D*, 76, 063009  
 Meiksin A., White M., Peacock J.A., 1999, *MNRAS*, 304, 851  
 Morrison, H. L., et al. 2003, *AJ*, 125, 2502  
 Murray, N., Hansen, B., Holman, M., & Tremaine, S. 1998, *Science*, 279, 69  
 Nagasawa, M., Ida, S., & Bessho, T. 2008, *ApJ*, 678, 498  
 Newberg, H. J., et al. 2002, *ApJ*, 569, 245  
 Noh Y., White M., Padmanabhan N., 2009, *Phys. Rev. D*80, 123501  
 Norman M.L., Paschos P., Harkness R., 2009, in Proceedings of SciDAC 2009, Ed. J. Hules, Journal of Physics: conference series [arXiv:0908.0964]  
 Odenkirchen, M., et al. 2001, *ApJ*, 548, L165  
 Ostriker, J. P. & Steinhardt, P. J. 1995, *Nature*, 377, 600  
 Padmanabhan N., White M., Cohn J.D., 2009, *Phys. Rev. D*79, 063523  
 Padmanabhan N., White M., 2009, *Phys. Rev. D*80, 063508  
 Palanque-Delabrouille, N., et al. 2010, arXiv:1012.2391  
 Peebles P.J.E., 1984, *ApJ*, 284, 439  
 Peebles P.J.E., Yu J.T., 1970, *ApJ*, 162, 815  
 Percival, W. J., et al. 2010, *MNRAS*, 401, 2148  
 Perlmutter, S., et al. 1999, *ApJ*, 517, 565  
 Rasio, F. A., & Ford, E. B. 1996, *Science*, 274, 954  
 Re Fiorentin, P., et al. 2007, *A & A*, 467, 1373  
 Reid, B. A., et al. 2010, *MNRAS*, 404, 60  
 Richards, G. T., et al. 2002, *AJ*, 123, 2945  
 Richards, G. T., et al. 2009, *ApJS*, 180, 67  
 Riess, A. G., et al. 1998, *AJ*, 116, 1009  
 Schlaufman, K. C., Lin, D. N. C., & Ida, S. 2009, *ApJ*, 691, 1322  
 Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525  
 Schneider, D. P., et al. 2010, *AJ*, 139, 2360  
 Searle, L. & Zinn, R., 1978, *APJ*, 225, 357  
 Seo, H. & Eisenstein, D. J. 2003, *ApJ*, 598, 720  
 Seo, H.-J., & Eisenstein, D. J. 2007, *ApJ*, 665, 14  
 Seo, H.-J., et al. 2010, *ApJ*, 720, 1650  
 Skrutskie, M. F., et al. 2006, *AJ*, 131, 1163  
 Slosar, A., Ho, S., White, M., & Louis, T. 2009, *JCAP*, 10, 19  
 Smith R.E., Scocimarro R., Sheth R.K., 2007, *PRD*, 75, 063512  
 Smolinski, J. 2010, *AJ*, in press (ArXiv: 1008.1959)  
 Steinmetz, M., et al. 2006, *AJ*, 132, 1645  
 Stoughton, C., et al. 2002, *AJ*, 123, 485  
 Strauss, M.A., et al. 2002, *AJ*, 124, 1810  
 Sunyaev R., Zel'dovich Y.B., 1970, *Ap&SS*, 7, 3  
 Takahashi, R., et al. 2009, *ApJ*, 700, 479  
 Takahashi, R., et al. 2011, *ApJ*, 726, 7  
 van Eyken, J. C., Ge, J., & Mahadevan, S. 2010, *ApJS*, 189, 156  
 Weidenschilling, S. J., & Marzari, F. 1996, *Nature*, 384, 619  
 White, M. 2003, *The Davis Meeting On Cosmic Inflation*, p. 18, astro-ph/0305474  
 White M., Pope A., Carlson J., Heitmann K., Habib S., Fasel P., Daniel D., Lukic Z., 2010, *ApJ*, 713, 383  
 White, M., Song, Y.-S., & Percival, W. J. 2009, *MNRAS*, 397, 1348  
 Willemsen, P. G., Hilker, M., Kayser, A., & Bailer-Jones, C. A. L. 2005, *A & A*, 436, 379  
 Wilson, J. C., et al. 2010, *Proc. SPIE*, 7735,  
 Wolszczan, A., & Frail, D. A. 1992, *Nature*, 355, 145  
 Wright, E. L., et al. 2010, *AJ*, 140, 1868  
 Wright, J. T., et al. 2009, *ApJ*, 699, L97  
 Wu, Y., & Murray, N. 2003, *ApJ*, 589, 605  
 Yanny, B. et al. 2000, *ApJ*, 540, 825  
 Yanny, B., et al. 2003, *ApJ*, 588, 824  
 Yanny, B., et al. 2009, *AJ*, 137, 4377  
 Yèche, C., et al. 2010, *A&A*, 523, A14  
 York, D. G. et al. 2000, *AJ*, 120, 1579  
 Zacharias, N., Monet, D. G., Levine, S. E., Urban, S. E., Gaume, R., & Wycoff, G. L. 2004, *Bulletin of the American Astronomical Society*, 36, 1418  
 Zacharias, N., et al. 2010, *AJ*, 139, 2184  
 Zel'dovich, Y. B. 1968, *Soviet Physics Uspekhi*, 11, 381  
 Zwitter, T., et al. 2008, *AJ*, 136, 421