SLOAN DIGITAL SKY SURVEY: RECENT RESULTS

Stephen Kent
Fermi National Accelerator Laboratory *
M.S. 127, P.O. Box 500, Batavia, IL 60510

for the SDSS Collaboration

ABSTRACT

The status of and recent results from the Sloan Digital Sky Survey are reviewed. The survey has completed 78% of the imaging and 48% of the spectroscopic observations of its 5 year baseline plan. Data from the first year are now publicly available. A crucial advantage of the SDSS dataset is its size and homogeneity, which allows for a systematic investigation of the dependence of galaxy clustering as a function of galaxy properties such as luminosity, color, and morphology. Analyses of the SDSS data by SDSS collaboration members are corroborating current cosmological models in which the universe is critically bound with 30% of the density in the form of baryons and dark matter and 70% in the form of dark energy.

*Supported by DOE Contract No. DE-AC02-76CH03000

1 Goals and Status of SDSS

The Sloan Digital Sky Survey (SDSS) is a project to map 1/4 of the sky in 5 bands and obtain the spectra of 1 million galaxies and quasars.[1] A particular goal of the SDSS is to measure the large scale structure of galaxies on scales up to 100 Mpc and to measure the distribution of quasars on scales comparable to the visible size of the universe. A 2.5 m telescope at the Apache Point Observatory in New Mexico is equipped with a mosaic imaging camera and a pair of dual-channel spectrographs capable of collecting 640 spectra simultaneously. The imaging camera[2] operates in drift-scan mode and collects data as long strips on the sky. The strips are interleaved and stitched together in order to produce a seamless map of the sky. The data are processed to produce corrected pixel maps and object catalogs[6] along with photometric[7–9] and astrometric[10] calibrations. The brightest galaxies[3,4] and quasar candidates[5] are selected for followup spectroscopy. A variety of data products that made available to the collaboration and the astronomical community.[11,12]

While the original goal of the survey was to cover 1/4 of the sky in 5 years, the actual data collection rate has been somewhat less than anticipated, and a reduced five year baseline plan was set up that covers somewhat less area of sky. Table 1 shows the status of the survey relative to this reduced baseline plan as of August 2003. The imaging survey is well on its way to completion, while the spectroscopic survey is still lagging somewhat. Nevertheless, the existing data are providing a wealth of results already.

Table 1. SDSS North Galactic Cap Status

Dataset	5 year baseline	Completed
Imaging	8252 sq. deg.	78%
Spectroscopy	1688 tiles	48%
PT (Calibrations)	1563 patches	93%

Data from the first year of operations (through June, 2001) plus some earlier commissioning data have been released[12] to the public as Data Release 1 (DR1) and are available online through the main SDSS website at www.sdss.org. Additional releases are planned to occur at roughly year increments. The full DR1 is

roughly 3 terabytes in size.

2 Strength of SDSS

While the SDSS is not the only imaging or spectroscopic survey in existence, it has unique advantages compared to others. Table 2 compares the SDSS with a couple of other large imaging and spectroscopic surveys. For imaging, the only surveys of comparable area or depth are still the Schmidt surveys done using photographic plates.[13,14] The big advantage of the SDSS is its high photometric precision, both for single objects and for systematics across the sky. This precision allows one to measure galaxy and quasar colors to high precision, which is useful in several contexts, and to measure large scale structure on the largest angular scale of the survey with unprecedented accuracy. For spectroscopy, the 2dF survey[15,16] has nearly as many redshifts, but because it selects galaxy and quasar candidates from photographic plates, analyses of the data are limited by the coarser photometric accuracy.

Table 2. Strengths of SDSS

IMAGING	DPOSS[14]	SDSS
Depth	$r = 20.5$	r=22.5
Accuracy	20%	2%
Bands	JFI	$ugriz$
Area	30000 sq. deg.	5000 sq. deg.
SPECTROSCOPY	2dF[15,16]	SDSS
Galaxies	220,000	400,000
Quasars	30,000	40,000
Matching Photometry(?)	Plates	CCD

A key problem for using galaxies as a probe of large scale structure is that galaxies are imperfect tracers of the mass distribution since they are the end product of a complex process involving the intricacies of gas dynamics and star formation, which are only imperfectly understood, whereas ideally one would like to observe the distribution of total mass (or at least the dark matter) directly. One says that the distribution of galaxies is biased relative to the dark matter,

and this bias is a multivariate function of a galaxy's luminosity, morphology/color, local environment (density), surface brightness, and redshift. Characterizing these biases fully requires a large sample of galaxies, and one of the motivations for the SDSS was to provide such a sample.

Galaxy samples are typically constructed by selecting galaxies brighter than some apparent brightness and surface brightness in a single filter band. This procedure already imposes a selection function on galaxies that depends on intrinsic luminosity, color and surface brightness, and it is a function of redshift. The importance of correcting for these selection biases will be demonstrated in the analyses that follow.

3 Galaxy Properties

One of the simplest but most important characterizations of galaxies is the luminosity function. An analysis of the properties of a statistically complete sample of galaxies[17] from the SDSS is shown in Figure 1. This figure gives the number counts of galaxies as a function of the r band luminosity in a volume limited sample along with a best fitting Schechter[18] function. This function is a semi-empirical way of parameterizing the luminosity function. While the function appears to fit the data extremely well, this result is rather fortuitous; Figure 2 shows the luminosity function for the galaxies divided into five bins based on $g - r$ color; it is obvious that the luminosity function of galaxies is a strong function of color, and no subsample is as well described by a Schechter function as the ensemble as a whole. Since one normally selects galaxies based on apparent magnitude in a single fixed bandpass, the mix of galaxies that one selects will likewise depend on the choice of bandpass.

Luminous galaxies are thought to be embedded in dark matter halos, and an important piece of information for using galaxies in cosmology is the relationship between galaxy luminosity and the halo mass. This relationship is difficult to determine with any reliability because of the scarcity of optical tracers such as satellite galaxies or extended gaseous disks that reach to large radii round single galaxies. The technique of weak lensing has proven to be a powerful method for filling this void.[19] The gravitational field of a nearby galaxy causes a small but measurable distortion in the shapes of background galaxies, and by combining the shapes of distant galaxies around an ensemble of nearby galaxies all of similar

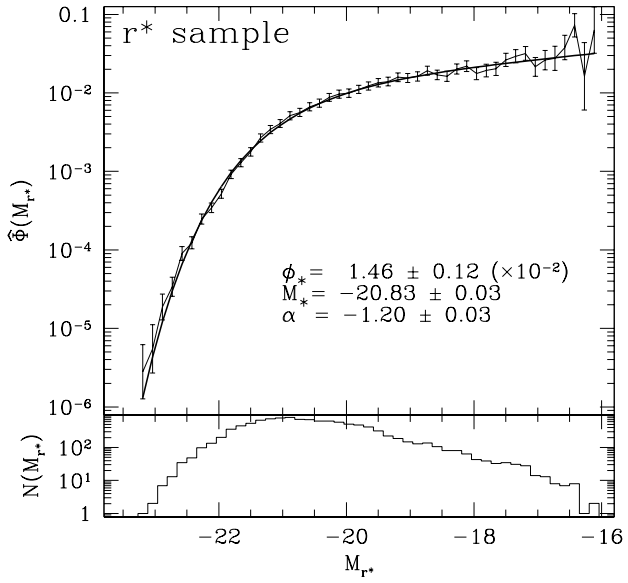

Fig. 1. Top Panel: Galaxy luminosity function in SDSS r band along with best-fitting Schechter function. Bottom panel: number of galaxies in each bin.[17]

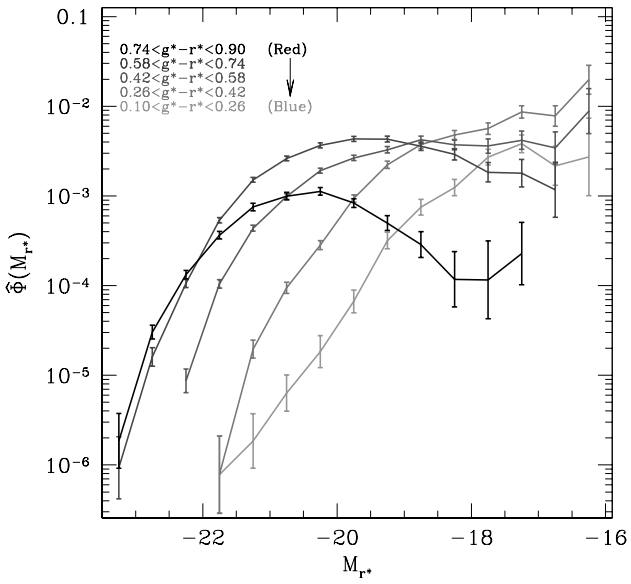

Fig. 2. Galaxy luminosity function for subsamples selected by color.[17]

luminosity and type, one can derive approximate mass profiles for different classes of nearby galaxies. Figure 3 shows the shear signal (converted to projected mass density) measured for background galaxies behind 31,000 foreground galaxies. Table 3 (Ref. 19), gives the ratio of mass to light for galaxies as a function of the bandpass used to measure the luminosity and morphological type (which is roughly correlated with color). It is seen that the mass/light ratio is a function of type, but the difference is minimized by observing in a red or near-infrared bandpass.

Table 3. Calibration of M/L versus Morphological Type and Bandpass

Galaxy type	M/L_u (h $M_\odot/L_\odot$)	M/L_r (h $M_\odot/L_\odot$)	M/L_z (h $M_\odot/L_\odot$)
Spirals	167	143	97
Ellipticals	647	221	123

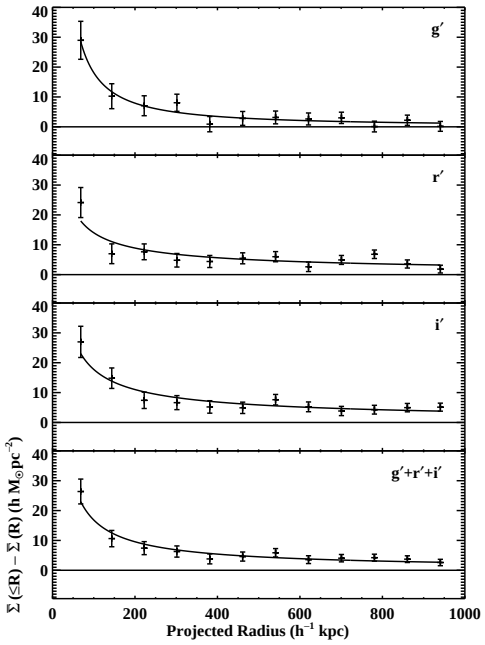

Fig. 3. Projected mass density for 31,000 nearby galaxies measured by weak lensing.[19]

4 Galaxy Clustering

One of the simplest measures of galaxy clustering is the 2-point correlation function $\epsilon(r)$, which is defined as the probability that a given galaxy will have a neighboring galaxy at a distance r relative to a random distribution. Mathematically, one has

$$\epsilon(r) = \langle \delta(r')\delta(r'+r) \rangle,$$

where $\delta = \Delta\rho(r')/\rho_0$ is the fraction excess density of the universe relative to the mean at position r', and the angle brackets indicate the ensemble average over r'. This function can be approximated by a power law:

$$\epsilon(r) = \left(\frac{r}{r_0}\right)^{-\gamma},$$

with parameters $r_0 \approx 8h$ Mpc^{-1} and $\gamma = 1.8$. The Fourier transform of $\epsilon(r)$ is the power spectrum $P(k)$.

Figure 4 shows the correlation function for a sample drawn from the SDSS data.[20] In detail, the correlation function has some curvature relative to a power law. The sample is once again divided into two subsamples based on color (which is roughly the same as morphological type here). It is clear that the red (elliptical/S0) galaxies are more strongly clustered than the blue (spiral/irregular) galaxies, which is a manifestation of the well-known morphology density relation for galaxies.[21] It also appears that the differentiation disappears on scales larger than r_0 where density fluctuations are still in the linear regime.

Galaxy clusters are the largest gravitationally bound objects, and one can study their properties in detail just as for individual galaxies. Identifying clusters is more of a challenge, however, because they are not as cleanly delineated as galaxies. Traditionally they have been identified by searching for excess in the projected density of galaxies along with some model for the luminosity function and density profile. Such algorithms work decently at low redshifts but suffer at higher redshifts due to contamination by line-of-sight projection of unrelated small groups of galaxies. The projection effects can be largely eliminated if one has redshifts, but even with surveys like the SDSS, redshift coverage is still limited to relatively bright galaxies. Because the galaxies in cluster centers are predominantly E and S0 types with very uniform old stellar populations, one can obtain an approximate redshift of such galaxies quite easily from photometry alone if the

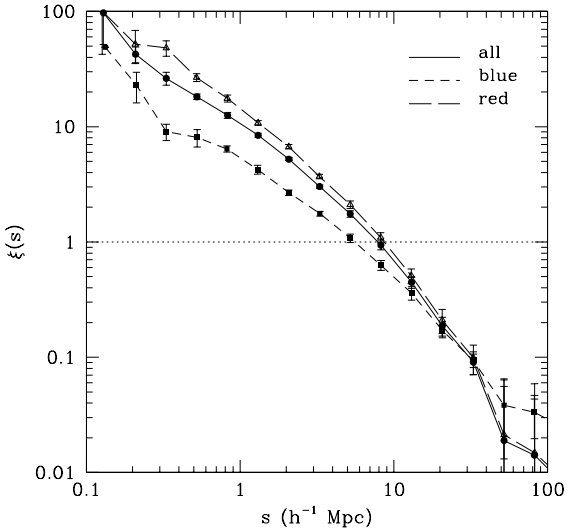

Fig. 4. Two point correlation function for blue and red subsamples of galaxies as well as the combined sample.[20]

accuracy is high enough. The SDSS data can determine such photometric redshifts out to $z \approx .7$ (well beyond the limits of the spectroscopic survey) and have led to a minor renaissance in cluster finding. Several algorithms are being tried by various SDSS investigators, but the most effective involve using the colors and magnitudes of galaxies to identify a core of red galaxies all at the same redshift.

Figure 5 shows how one technique, called the maxBCG method[22–24] works. The top panel shows the $g - r$ and i magnitudes of galaxies in a field with a rich cluster, and the lower panel shows the $g - r$ versus $r - i$ colors for the same galaxies. An idealized model of a galaxy cluster, depicted in the figure by an ellipse and two straight lines in the upper panel and by an ellipse in the lower panel, is filtered through the data to search for an excess of galaxies. The two colors are particularly sensitive to the redshift of old type galaxies. The parameters measured for a cluster are the number of galaxies n_{gals} and redshift z. An analysis of the SDSS data finds of order 20 clusters per square degree out to $z = 0.7$, with the caveat that most of the clusters are smallish systems and the search is incomplete beyond $z = 0.3$ to 0.4.

While one expects n_{gals} to be correlated with total cluster mass, the slope and zero-point of the relation must be determined empirically. Once again one can measure the weak lensing of galaxies behind the clusters to calibrate the cluster mass. Figure 6 shows such a correlation for about 40 clusters.[25] Using this correlation, one can now compute the cluster mass function and compare it with predictions from n-body simulations and derived various cosmological parameters. Figure 7 shows one such comparison.[22] The theoretical cluster mass distribution function[27] has been converted to a distribution as a function of n_{gals} and compared to the measured cluster number counts. It is interesting to note that clusters found by Xray and S-Z surveys are typically more massive than 2×10^{14} $M_\odot$ whereas maxBCG can identify clusters 10 times smaller. In a more extensive analysis using two different cluster detection algorithms,[23] one finds the parameter combination $\sigma_8 \Omega_M = 0.33 \pm 0.03$, where σ_8 is the (linearized) amplitude of the galaxy fluctuations on $8h^{-1}$ Mpc scales, and Ω_M is the fraction of mass in baryonic plus dark matter. This value is smaller than the canonical value 0.5 found from other studies.

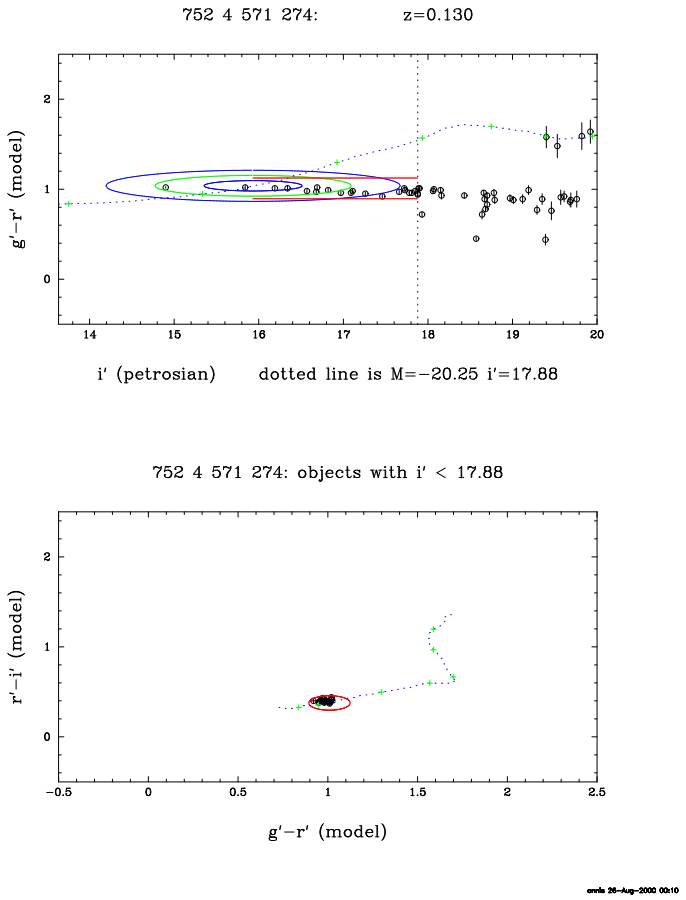

Fig. 5. MaxBCG technique used to detect a galaxy cluster.[22] Top panel: Color magnitude diagram, showing how red cluster galaxies fall in a well-defined locus. Bottom panel: Color color diagram, showing how red clusters galaxies fall in a tight ellipse.

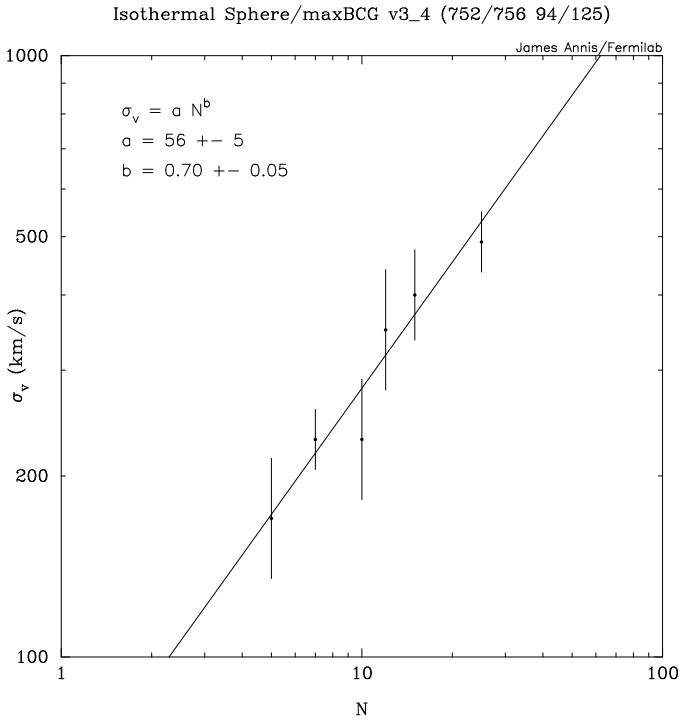

Fig. 6. Calibration of cluster masses as a function of n_{gals} using weak lensing.[25,26]

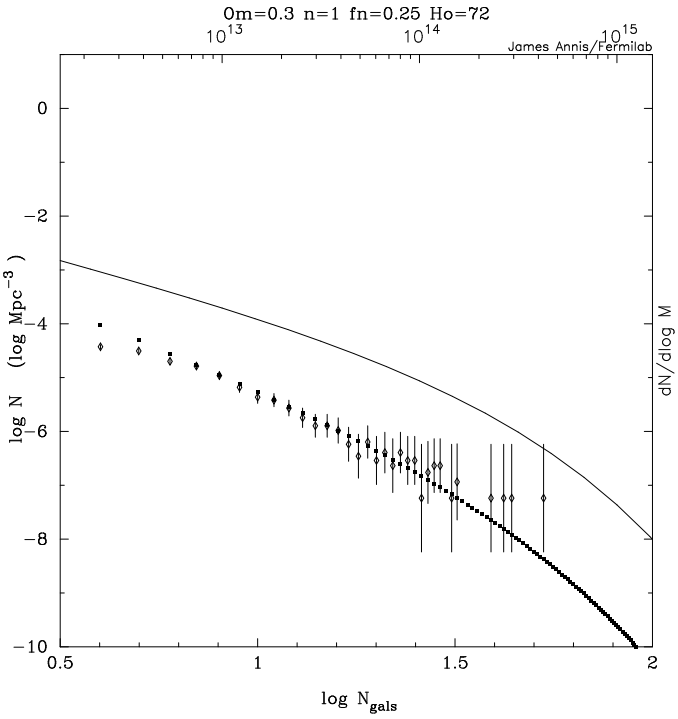

Fig. 7. Number of galaxy clusters as a function of mass and n_{gals}.[22]

5 Large Scale Structure

The most straightforward prediction from theories for the formation of structure in the universe is the power spectrum of mass fluctuations as a function of scale. On scales larger than $8h^{-1}$ Mpc today, these density fluctuations are still in the linear growth regime, so tying galaxy and density fluctuations to the primordial power spectrum ought to be relatively straightforward. On smaller scales, the density fluctuations are nonlinear (i.e. structures have collapsed into bound systems) and one must understand the intricacies of galaxy formation and nonlinear dynamics to connect observations to theory.

Figure 8 shows the power spectrum derived from a subset of the SDSS spectroscopic galaxy survey.[28] Also plotted in the figure is a prediction for a standard ΛCDM model. Three parameters, corresponding roughly to the amplitude, characteristic scale length, and the slope of the power spectrum, were solved for. A key part of the analysis was to correct for biases in the power spectrum as a function of galaxy luminosity. In general, the bias parameter b gives the ratio of the amplitude of the power spectrum between two classes of objects and might be a function of scale. Figure 9 shows the measured bias for galaxies as a function of luminosity using $L*$ galaxies as a reference. It is clear that high luminosity galaxies are more clustered than low luminosity objects, and the degree of bias is roughly independent of scale. The parameter σ_8 obtained from the fit in Figure 8 refers to the amplitude of $L*$ galaxies. The bias between $L*$ galaxies and the actual mass itself is thought to be close to unity but must be determined by other means.

Figure 10 shows the same power spectrum plotted along with measurements determined using other techniques. The agreement among the measurements in the regions where they overlap is quite impressive. The SDSS measurements are also in reasonably good agreement with the spectrum determined by 2dF[29] although the SDSS data show no evidence for a change in slope of the power spectrum with scale.

The WMAP experiment[30] measures fluctuations in the cosmic microwave background that are largely due to density fluctuations at $z \approx 1400$ whereas the SDSS measures fluctuations in galaxies at $z \approx 0$. However, a slight correlation between the two measurements can be induced by the ISW (integrated Sachs-Wolfe) effect, whereby the energy of a photon changes as it traverses a gravitational potential

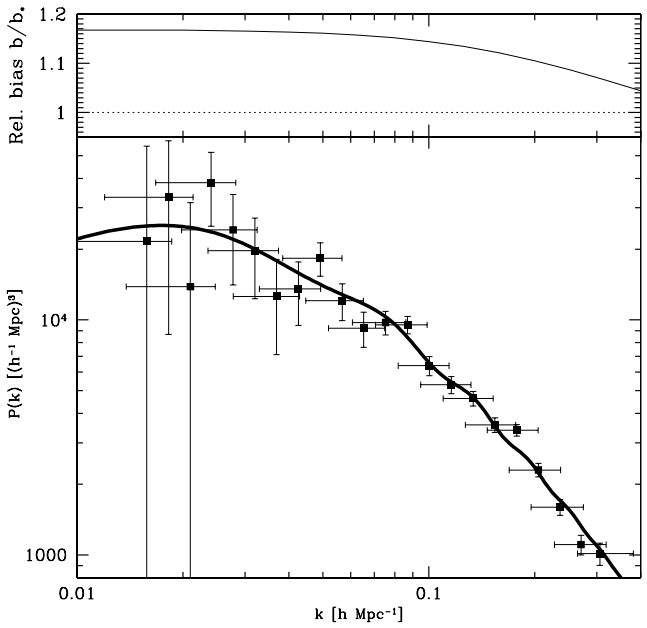

Fig. 8. Large scale structure power spectrum from SDSS galaxies. Cosmological parameters include $\Omega_m h = 0.213$, $\sigma_8 = 0.96$, and $n = 0.995$. σ_8 refers to the fluctuation amplitude of L_* galaxies.[28]

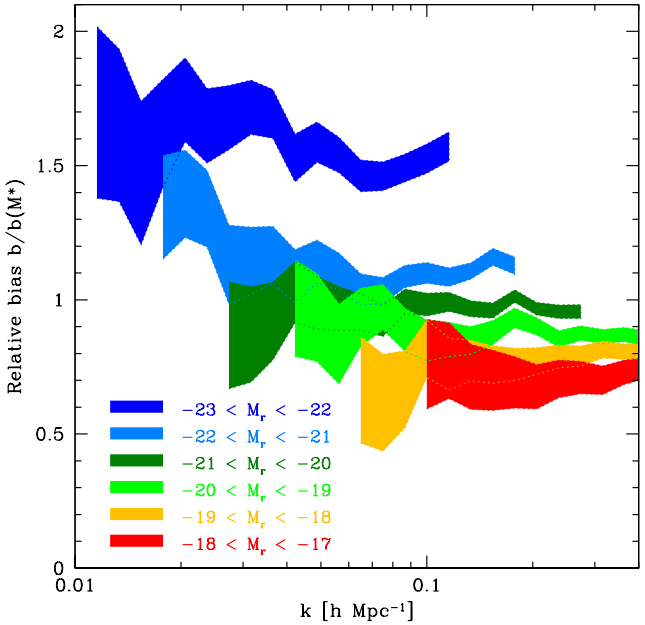

Fig. 9. Bias in galaxy power spectrum as a function of galaxy luminosity.[28]

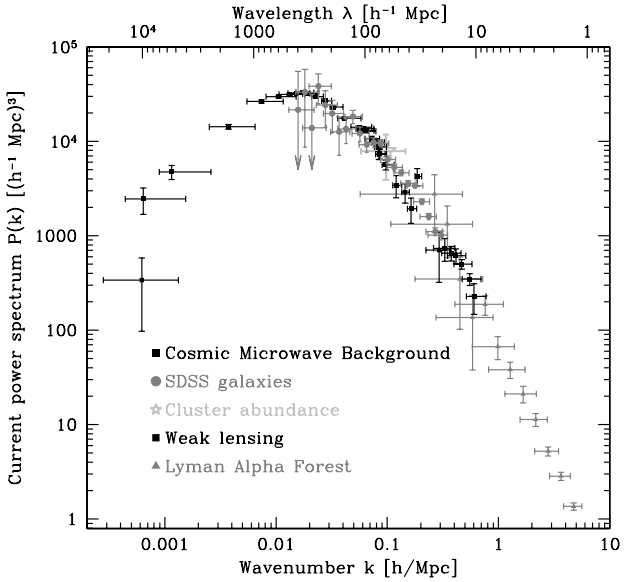

Fig. 10. Large scale structure power spectrum measured by several techniques.[28]

well that is itself changing with time in an expanding universe. For a flat universe made of CDM and other cold matter only, the net correlation is zero, but in a universe where the CDM component is less than the closure density (including ΛCDM), a positive correlation is introduced. Recently such a correlation has indeed been found using the SDSS data[31] at about the 3 σ level (figure 11). Although this detection is not accurate enough to provide new information about cosmological parameters, it does add yet more evidence for the existence of dark energy.

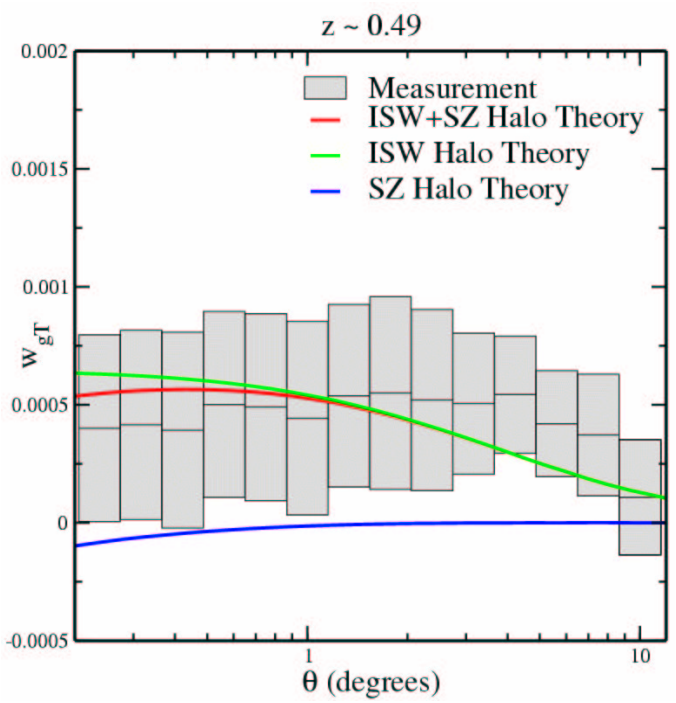

Fig. 11. Cross-correlation between WMAP temperature fluctuations and SDSS galaxies counts.[31]

6 High Redshift Quasars

Quasars are point-like objects and thus have shapes that are indistinguishable from stars. However, their spectra have power law continua, strong emission lines, and (at high redshift), strong absorption troughs shortwards of the Lyman Alpha emission line due to intervening clouds of hydrogen. The multicolor SDSS

filter system is designed in part to allow identification of quasars over a wide range in redshifts,[32,33] with an upper limit of about $z = 6.5$ (Ref. 34). At present most known quasars with $z > 4$ have been discovered by the SDSS, and quasars with redshifts greater than 6 have now been identified[35] (Figure 12). These objects give one the ability to probe the intergalactic medium at high redshifts and search for the so-called epoch of reionization where newly formed stars have generated enough ultraviolet light that the universe is reionized from the neutral state it had after recombination. The signature of this transition is a sudden increase in optical depth of the Lyman Alpha trough at redshifts greater than that of reionization. Figure 13 shows the optical depth of the trough for a set of high redshift quasars.[36] There is a hint that the optical depth jumps suddenly for $z > 6$, which is consistent with reionization occurring at this redshift. This result is inconsistent with the WMAP[30] estimate for reionization at $z = 20$, but confirmation or refutation will require a larger sample of quasars.

7 Conclusions

The SDSS is performing successfully and will continue to accumulate additional data for two more years. It is testing and corroborating current cosmological models for the expansion of the universe and growth of structure in a plethora of ways and is significantly increasing the precision with which optical data can be applied to these problems. Finally, although not addressed in this contribution, the SDSS is producing leading edge science in a wide range of other areas of astrophysics.

8 Acknowledgements

The author is grateful to the members of the SDSS collaboration, who actually most of the work reported here.

Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is http://www.sdss.org/.

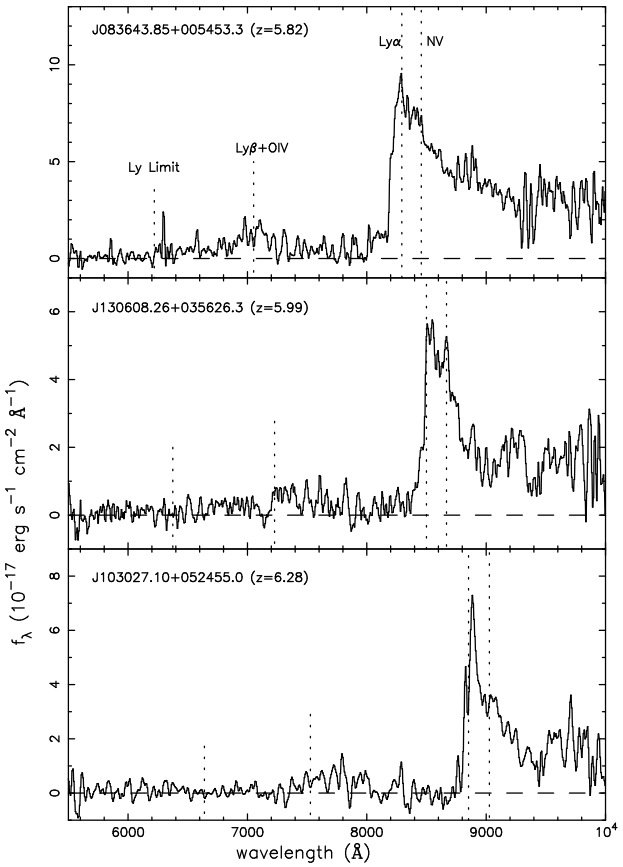

Fig. 12. Spectra of three high redshift quasars, showing the absorption trough to the left of the prominent Lyman Alpha emission line.[35]

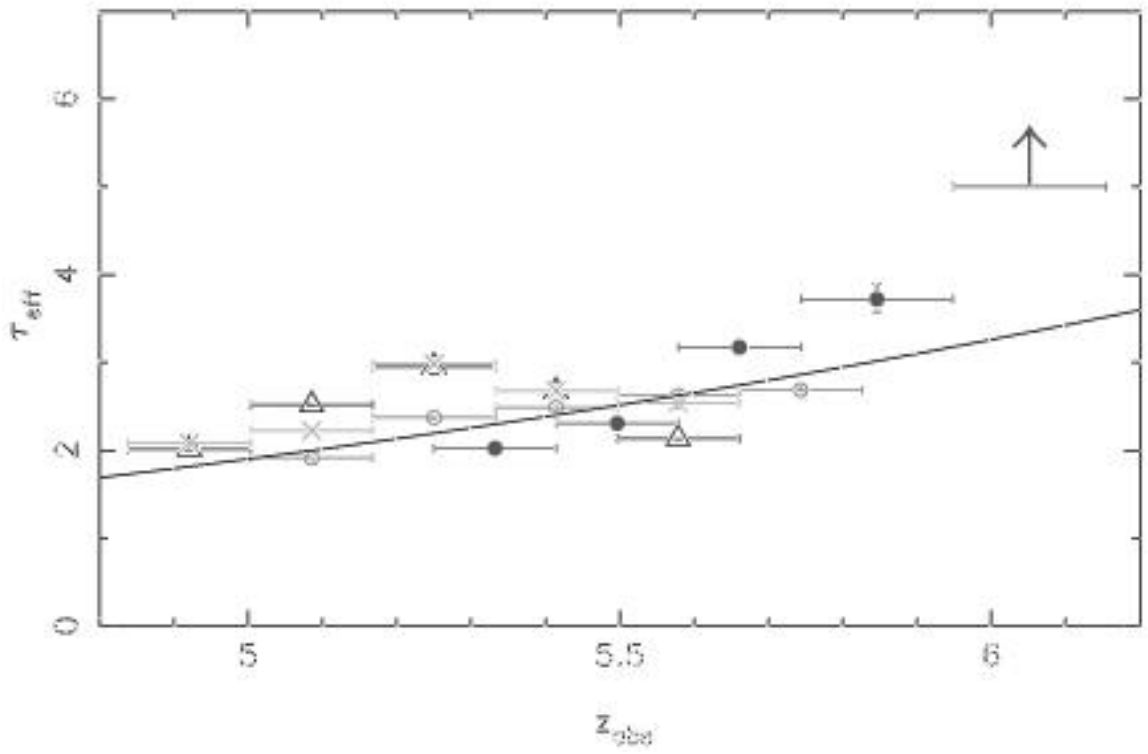

Fig. 13. Optical depth in the Lyman Alpha absorption trough for several quasars of different redshifts. A lower limit is indicated for the one quasar with $z = 6.28$.[36]

The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, Princeton University, the United States Naval Observatory, and the University of Washington.

References

[1] D. G. York *et al.*, Astron. J. **120**, 1579 (2000).

[2] J. E. Gunn *et al.*, Astron. J. **116** 3040 (1998).

[3] M. Strauss *et al.*, Astron. J. **124**, 1810 (2002).

[4] D. Eisenstein *et al.*, Astron. J.**122**, 2267 (2001).

[5] G. Richards *et al.*, Astron. J.**123**, 2945 (2002).

[6] Lupton, R. L. *et al.*, in *ASP Conf. Ser. 238, Astronomical Data Analysis Software and Systems X*, edited by F. R. Harnden Jr., F. A. Primini, and H. E. Payne (San Franciso, Astr. Soc. Pac., 2001), p. 269.

[7] M. Fukugita *et al.*, Astron. J. **111**, 1748 (1996).

[8] D. W. Hogg *et al.*, Astron. J. **122**, 2129 (2001).

[9] J. A. Smith *et al.*, Astron. J. **123**, 2121 (2002).

[10] J. R. Pier *et al.*, Astron. J. **125**, 1559 (2003).

[11] C. Stoughton *et al.*, Astron. J. **123**, 485 (2002).

[12] K. Abazajian *et al.*, Astron. J. **126**, 2081 (2003).

[13] S. J. Maddox, G. Efstathiou, W. J. Sutherland, and J. Loveday, Mon. Not. Roy. Astron. Soc. **243**, 692 (1990).

[14] S. G. Djorgovski *et al.*, in *Wide Field Surveys in Cosmology*, edited by S. Colombi *et al.* (Paris, Ed. Frontière), p. 89 (1999).

[15] M. Colless *et al.*, Mon. Not. Roy. Astron. Soc. **328**, 1039 (2001).

[16] B. J. Boyle *et al.*, Mon. Not. Roy. Astron. Soc. **317**, 1014 (2000).

[17] M. Blanton *et al.*, Astron. J. **121**, 2348 (2001).

[18] P. Schechter, Astrophys. J. **203**, 297 (1976).

[19] T. McKay *et al.*, Astron. J., submitted (astro-ph/0108013) (2001).

[20] I. Zehavi *et al.*, Astrophys. J. **571**, 172 (2002).

[21] A. Dressler, Astrophys. J. **236**, 351 (1980).

[22] J. Annis *et al.*, in preparation (2003).

[23] N. Bahcall *et al.*, Astrophys. J. **585**, 182 (2003).

[24] N. Bahcall *et al.*, Astrophys. J. Suppl. **148**, 243 (2003).

[25] E. Sheldon *et al.*, Astrophys. J. **554**, 881 (2001).

[26] T. A. McKay, private communication (2003).

[27] A. Jenkins *et al.*, Mon. Not. Roy. Astron. Soc. **321**, 372 (2001)

[28] Tegmark, M., *et al.*, Astrophys. J. submitted (2003).

[29] W. J. Percival *et al.* Mon. Not. Roy. Astron. Soc., **327**, 1297.

[30] C. L. Bennett *et al.*, Astrophys. J. **583**, 1 (2003).

[31] R. Scranton *et al.*, *Phys. Rev. Lett.* submitted (2003).

[32] H. Newberg *et al.*, Astrophys. J. Suppl. **113**, 89 (1997).

[33] X. Fan *et al.*, Astron. J. **117**, 2528 (1999).

[34] X. Fan *et al.*, Astron. J. **120**, 1167 (2000).

[35] X. Fan *et al.*, Astron. J. **122**, 2833 (2001).

[36] R. Becker *et al.*, Astron. J. **122**, 2850 (2001).

[37] C. L. Bennett *et al.*, Astrophys. J. Suppl. **148**, 1 (2003).