Cosmology in the Era of Large Surveys

Ryan Scranton

Google

13 March 2007
A Long, Long Time Ago...
The State of Cosmology in 1996

- COBE measurements had shown that the Cosmic Microwave Background was constant to one part in $10^5$ at recombination. COBE DMR had measured CMB anisotropy correlations on very large scales. Ground and balloon based experiments were pushing to $\sim 1^\circ$ scales.

- Measurements of galaxy clustering in the local universe were largely limited to photometric data. Spectroscopic surveys were mostly focused on measuring bulk velocity flows.

- Big Bang Nucleosynthesis measurements had put strong constraints on the baryon density ($\Omega_b h^2$) but not much else.

- Fundamental questions about the size and shape of the Universe were still very much open.
1996 Great Age Debate

Two camps on the value of Hubble’s Constant:

- Cepheid variable stars in nearby galaxies gave $H_0 \sim 80$ km/s/Mpc.
- Measurements using nearby supernovae as standard candles yield $H_0 \sim 55$ km/s/Mpc.
- In both cases error bars are small ($\sim 5 - 10$).

Sidney van den Bergh: $H_0 = 81 \pm 8$ !!!
Lurking in the background, the Age Crisis:

- For an Einstein-deSitter universe ($\Omega_M = 1$), the age of the universe is $t_0 = 2/3 \, H_0^{-1}$
- Measurements of stellar ages in globular clusters put $t_0 > 11.5 \text{ Gyr}$ (as high as $t_0 > 17 \text{ Gyr}$ in some estimates)
- Either $\Omega_M < 1$ or $H_0 < 50 \text{ km/s/Mpc}$

Gustav Tammann:

$H_0 = 55 \pm 10$ !!!
1996 Great Age Debate

Lurking in the background, the Age Crisis:

- For an Einstein-deSitter universe ($\Omega_M = 1$), the age of the universe is $t_0 = 2/3 \ H_0^{-1}$

- Measurements of stellar ages in globular clusters put $t_0 > 11.5$ Gyr (as high as $t_0 > 17$ Gyr in some estimates)

- Either $\Omega_M < 1$ or $H_0 < 50$ km/s/Mpc or cosmology is broken.

Gustav Tammann: $H_0 = 55 \pm 10$ !!!
The Shape of the Universe

- Peebles 1995 summary on the state of $\Omega_M$

<table>
<thead>
<tr>
<th>Observation</th>
<th>$\Omega_M = 1$</th>
<th>$\Omega_M \sim 0.1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamics &amp; biasing on scales $\leq 3$ Mpc</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Dynamics on scales $\geq 10$ Mpc</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Expansion time $H_0 t_0$</td>
<td>???</td>
<td>YES</td>
</tr>
<tr>
<td>Radial and angular size distances</td>
<td>NO (?)</td>
<td>YES (?)</td>
</tr>
<tr>
<td>Plasma mass fraction in clusters</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Models for structure formation</td>
<td>YES (?)</td>
<td>YES (?)</td>
</tr>
</tbody>
</table>

- Theoretical bias toward $\Omega_M = 1$, given COBE & inflation

- Strong gravitational lensing: $\Omega_\Lambda < 0.65$ at 95% confidence for flat universes (Kochanek 1996)
Ten Years Later...

Two Major Improvements

• Better Data
  ★ More Supernovae
  ★ Finer CMB Anisotropies
  ★ Bigger Galaxy Surveys

• Better Analysis
  ★ Fisher Matrix
  ★ Markov Chain Monte Carlo
  ★ Machine Learning & Data Mining

Permutter et al. (1998)
Ten Years Later...

Two Major Improvements

- **Better Data**
  - More Supernovae
  - Finer CMB Anisotropies
  - Bigger Galaxy Surveys

- **Better Analysis**
  - Fisher Matrix
  - Markov Chain Monte Carlo
  - Machine Learning & Data Mining

Hu (2000)
Ten Years Later...

Two Major Improvements

- **Better Data**
  - More Supernovae
  - Finer CMB Anisotropies
  - Bigger Galaxy Surveys

- **Better Analysis**
  - Fisher Matrix
  - Markov Chain Monte Carlo
  - Machine Learning & Data Mining
Ten Years Later...

Two Major Improvements

- **Better Data**
  - More Supernovae
  - Finer CMB Anisotropies
  - Bigger Galaxy Surveys

- **Better Analysis**
  - Fisher Matrix
  - Markov Chain Monte Carlo
  - Machine Learning & Data Mining

Vogeley (1997)
Ten Years Later...

Two Major Improvements

- Better Data
  - More Supernovae
  - Finer CMB Anisotropies
  - Bigger Galaxy Surveys

- Better Analysis
  - Fisher Matrix
  - Markov Chain Monte Carlo
  - Machine Learning & Data Mining

Tegmark et al. (2003)
Ten Years Later...

Two Major Improvements

- **Better Data**
  - More Supernovae
  - Finer CMB Anisotropies
  - Bigger Galaxy Surveys

- **Better Analysis**
  - Fisher Matrix
  - Markov Chain Monte Carlo
  - Machine Learning & Data Mining
Ten Years Later...

Two Major Improvements

- **Better Data**
  - More Supernovae
  - Finer CMB Anisotropies
  - Bigger Galaxy Surveys

- **Better Analysis**
  - Fisher Matrix
  - Markov Chain Monte Carlo
  - Machine Learning & Data Mining

Tegmark et al. (2006)
Current Picture of Reality

- Very good evidence that the universe is flat ($\Omega = 1$), but that $\Omega_M \sim 0.3$, $\Omega_\Lambda \sim 0.7$.

- Doing astronomy with millions of objects (instead of dozens) gives us power to investigate dark matter and dark energy in ways that were previously impossible.

- Examples of Precision Cosmology: Cosmic Magnification & Integrated Sachs-Wolfe Effect
Tracking Dark Matter:

Cosmic Magnification with Galaxies and Quasars
Two Effects of Gravitational Lensing

Wittman (2000)

Light from distant sources is magnified and distorted by dark matter
Two Effects of Gravitational Lensing

Magnification ($\mu$) increases flux (amplification); decreases density (dilution)
Quantifying Cosmic Magnification

- If we are in the weak lensing regime ($\mu \approx 1$),

$$w_{GQ}(\theta) = 12\pi^2\Omega_M (\alpha(m) - 1) \int d\chi dk k \mathcal{K}(k, \theta, \chi) P_{gm}(k, \chi)$$

$$= (\alpha(m) - 1) \times w_0(\theta), \quad (1)$$

where $\alpha(m)$ is the power-law slope of the QSO number counts, $\mathcal{K}$ depends on the foreground and background redshift distributions and $P_{gm}(k)$ is the galaxy-dark matter power spectrum.

- For $\alpha(m) > 1$, increasing amplification outweighs the dilution effect, inducing a positive cross-correlation between foreground and background objects. For $\alpha(m) < 1$, dilution wins and the cross-correlation is negative.

- The lensing magnification is less than 1% per object, so we need to average over many, many QSOs.
Controversy – Is $\Omega_M \approx 1$?

- First lensing motivated measurements in late 1980s and early 1990s
  - Lick, IRAS & APM galaxies, Abell & Zwicky clusters
  - UVX and radio-selected QSOs

- More recently, Guimares, Myers & Shanks (2003) used 2dF QSOs + APM & SDSS galaxy groups

- Consistently detect signal
  $\sim 3 - 10 \times$ the expected lensing

Guimares, Myers, & Shanks (2003)
The Four Horsemen

- **Photometric Calibration**
  - Small amplification effect requires excellent photometry
  - Photographic plates not up to the challenge

- **Uniform Selection**
  - Photographic plates have variable depth of field and numerous defects
  - Spectroscopic surveys require detailed selection function

- **Redshift Overlap**
  - Physical Clustering dominates lensing signal
  - Require either spectroscopy or photometric redshifts for each object

- **Object Density**
  - Poisson errors dominate
  - When object density is low, only systematic signal is detected.
Photometric QSO Selection

- Traditional QSO selection involves cuts in 2-D color projections

- Kernel Density Estimation (KDE) using full 4-D color space
  - 2 training sets: QSOs & stars
  - compute distance in color space to assign new objects

- SDSS spectroscopic selection 85% efficient for $i' < 19$

- KDE selection $> 95\%$ for $g' < 21 \Rightarrow 10 \times$ density

Richards et al. (2004)
Measurement in $g'$

- Select 5 magnitude bins in $g'$:
  \[ 17 < g' < 19,\ 19 < g' < 19.5,\ 19.5 < g' < 20,\ 20 < g' < 20.5,\ 20.5 < g' < 21 \]

- Calculate $\langle \alpha - 1 \rangle$ in each bin:
  \[
  \langle \alpha - 1 \rangle = \frac{\int N(m) (\alpha(m) - 1)}{\int N(m)}
  \]

- Expect to see positive correlation for $g' < 19.5$ and negative correlation for $g' > 20$

Scranton et al. (2005)
Measurement in $g'$

- Select 5 magnitude bins in $g'$:
  17 < $g'$ < 19, 19 < $g'$ < 19.5, 19.5 < $g'$ < 20, 20 < $g'$ < 20.5, 20.5 < $g'$ < 21

- Calculate $\langle \alpha - 1 \rangle$ in each bin:

$$\langle \alpha - 1 \rangle = \frac{\int N(m) (\alpha(m) - 1)}{\int N(m)} \quad (3)$$

- Expect to see positive correlation for $g' < 19.5$ and negative correlation for $g' > 20$

Scranton et al. (2005)
Optimal Signal

- Magnitude bin measurements track expected signal as $\langle \alpha - 1 \rangle$ varies from bright to faint QSOs

- $\langle \alpha - 1 \rangle \approx 0$ for full QSO sample

- To extract the full statistical significance for lensing measurement, use second moment:
  - Re-calculate estimator weighting each QSO by $\alpha(m) - 1$
  - Expected signal:
    $$w_{GQ,0}(\theta) = \langle (\alpha - 1)^2 \rangle \times w_0(\theta)$$
    \hspace{1cm} (4)

- Instead of canceling, positive and negative correlations add coherently
Optimal $g'$

- 105,000 QSOs & 13 million galaxies
- $8\sigma$ detection of lensing against null signal
- Excellent match to expected signal
- For $z \sim 0.3$, detecting lensing on scales from 60 kpc/$h$ to 10 Mpc/$h$

Scranton et al. (2005)
Flip the Script

- Correlate QSO flux with galaxy density
- Differences between bands indicates wavelength-dependent extinction
- Reconstruct dust halo profile of average galaxy & environment
- Key systematic consideration for SNe missions

Menard & Scranton et al. (2007)
Flip the Script

- Correlate QSO flux with galaxy density
- Differences between bands indicates wavelength-dependent extinction
- Reconstruct dust halo profile of average galaxy & environment
- Key systematic consideration for SNe missions

Menard & Scranton et al. (2007)
Flip the Script

- Correlate QSO flux with galaxy density
- Differences between bands indicates wavelength-dependent extinction
- Reconstruct dust halo profile of average galaxy & environment
- Key systematic consideration for SNe missions

Menard & Scranton et al. (2007)
Flip the Script

- Correlate QSO flux with galaxy density
- Differences between bands indicates wavelength-dependent extinction
- Reconstruct dust halo profile of average galaxy & environment
- Key systematic consideration for SNe missions

Menard & Scranton et al. (2007)
The Future of Cosmic Magnification

- Using photometric QSOs and galaxies from SDSS DR3, we observe a signal with the expected lensing amplitude. The signal also exhibits the expected variation in amplitude and sign with changing $\alpha(m)$.

- Optimally combining all of our $g'$-selected QSOs, we detect cosmic magnification of QSOs at $> 8\sigma$. Earlier conflicting $\Omega_M$ values are resolved.

- Correlating QSO flux and galaxy density gives us the first ever measurement of galaxy dust halo shapes & will be critical for flux-based observations like future SNe.

- The techniques used for efficient QSO selection are readily applicable to next generation of large, multi-band surveys. Cosmic magnification with galaxies or QSOs is an excellent (free!) complement to planned cosmic shear surveys (same physics & cosmology, different systematics).
Detecting Dark Energy:
Integrated Sachs-Wolfe Effect
Integrated Sachs-Wolfe Effect in 2 Minutes

- After matter-radiation equality, dark matter falls into potential wells set up during inflation.

- For open or $\Lambda$CDM universes, universe expands faster than potentials, leading to potential decay.

- CMB photons passing through potentials see net blue-shift in energy $\Rightarrow$ positive correlation with foreground structure.
From WMAP, we know that the overall geometry of the universe is very close to flat ($\Omega = 1$).

Hence, detecting ISW signal is a clear signature of dark energy.

Orthogonal to SNe detection. ISW signal related to growth of structure, while SNe signal is due to expansion history.
The Galaxy Data Set

- Initially began with LRGs from SDSS (Scranton et al., 2003); detection at $\sim 3\sigma$
- Increased sample to include galaxies from 2MASS, FIRST and NVSS
- New galaxy sample contains 15 galaxy maps spanning $0 < z < 2.5$ and wavelengths from radio to IR to optical to near-UV.
The Galaxy Data Set

- Initially began with LRGs from SDSS (Scranton et al., 2003); detection at $\sim 3\sigma$

- Increased sample to include galaxies from 2MASS, FIRST and NVSS

- New galaxy sample contains 15 galaxy maps spanning $0 < z < 2.5$ and wavelengths from radio to IR to optical to near-UV.
Map Comparison
Results

Scranton et al. (2003)
Results

Scranton et al. (2007)
Results

Scranton et al. (2007)
Results

Scranton et al. (2007)
Results

Scranton et al. (2007)
Results

- Angular bins within measurements and between measurements highly correlated.

- Individual surveys have $2-3\sigma$ detections.

- Combining measurements from all 15 galaxy maps, ISW is detected at $>5\sigma$.

- Part of the (S/N) comes from magnification of high redshift samples by foreground structure.

Scranton et al. (2007)
Results

- Angular bins within measurements and between measurements highly correlated
- Individual surveys have 2-3σ detections
- Combining measurements from all 15 galaxy maps, ISW is detected at $>5\sigma$
- Part of the (S/N) comes from magnification of high redshift samples by foreground structure.

Scranton et al. (2007)
The Future of ISW

- Using all current large scale galaxy surveys (2MASS, SDSS, NVSS, & FIRST) covering $0 < z < 2.5$, we detect ISW at $> 5\sigma$. This dark energy signature is completely independent of SNe evidence based on acceleration.

- Efforts to turn this detection in cosmological constraints are underway (Scranton et al., 2007).

- ISW signal is sensitive to changes in dark energy over time, but noise from primary CMB anisotropies keeps (S/N) low ($\sim 6 - 10\sigma$).

- Future galaxy surveys with larger areas and deeper samples can constrain dark energy equation of state to 5% (Hu & Scranton, 2004).

- Measurements combining ISW & magnification effects may also offer another lever arm for describing dark energy.
Our Dark Energy Future...
A Riddle Wrapped in a Mystery Inside an Enigma

• The current crop of surveys have largely resolved the issues surrounding $\Omega_M$ and $H_0$ from 10 years ago.

• Along the way, however, they discovered the existence of dark energy (so named in 1999), which is even more puzzling.

• Currently, dark energy theory is in a state of maximal ignorance. We don’t know

  ★ what the dark energy equation of state ($w \equiv P_{DE}/\rho_{DE}$) is
  ★ whether $w = w(z)$
  ★ whether dark energy clusters (Hu & Scranton, 2004; Bean & Dore, 2003)
  ★ whether “dark energy” is actually a change in gravity (DGP, 2000; Knox, Song & Tyson, 2005; Linder & Huterer, 2006)
A Riddle Wrapped in a Mystery Inside an Enigma

- With no theoretical guidance, the two new questions to answer are
  - What is the Universe’s expansion history over the last 10 Gyr?
  - What is the rate of large scale structure growth over the last 10 Gyr?

- No longer 2 parameters; now we have to constrain two functions.

- We will need measurements that handle the first question (supernovae, baryon acoustic oscillations), the second question (ISW) and both (weak lensing, cluster abundance).

- Just as importantly, we will need to be able to combine these measurements in a statistically meaningful way.
New Instruments, New Realities

- Next generation of surveys to focus on weak lensing, supernovae, baryon acoustic oscillations (BAO), and galaxy clusters.

- Ground based galaxy surveys (LSST), CMB cluster finders (SPT), space-based surveys (SNAP)

- Much larger data sets with more complicated geometries, selection functions, time domain data, etc.
New Instruments, New Realities

- Greater data size & complexity demands new statistical & algorithmic tools and a new way of looking at data.

- Maximizing primary and secondary science (galaxy evolution, cluster physics, stellar physics) will require moving easily between surveys & between measurements.

- Need a unified survey language (STOMP) and a single point of access & exploration (The Google Thing).
STOMP: Space and Time Ordered Mapping Package

http://nvogre.phyast.pitt.edu/gestalt/

- All cosmological statistics are measurements of spatial properties (area, angular distance, density)

- Describe complex geometries on the sphere and possibly spatial variations

- Find unions, intersections, and overlaps between large numbers of observations

- It has to be fast
STOMP: Space and Time Ordered Mapping Package

- Pixelize arbitrary survey footprints with 1” resolution
- Hierarchical scheme: extremely rapid localization & angular statistics
- Spatial information (completeness, flux, temperature, etc) & geometry
Lingua Franca – A World of Applications & Results

- **Angular Correlations**
  - Integrated Sachs-Wolfe Effect
  - $w(\theta)$ (Scranton et al., 2007)
  - Higher Order (Ross et al., 2006)

- **Survey Simulations**
  - Dark Energy Survey (DES)
  - LSST

- **Galaxy Evolution**
  - Galaxy Environment (Welikala et al., 2007)
  - LRG luminosity function (Loh et al., 2007)

- **Weak Lensing**
  - Magnification & Extinction
  - Shear Lensing (Sheldon et al., 2004; Mandelbaum et al., 2005)

- **Spectroscopic Surveys**
  - BAO (Eisenstein et al., 2005)
  - Halo Multiplicity Function (Berlind et al., 2005)

- **Galaxy Clusters**
  - MaxBCG (Koester et al., 2007)
  - Optical/X-ray Counterparts (Miller et al., *in preparation*)
About that Google thing, I can neither confirm nor deny that something very, very cool is about to come out that will change the way that you do astronomy...
Summary

- The current generation of surveys has succeeded in solving the central questions in cosmology from a decade ago.

- Along the way, the richness of the data gathered drove new scientific results that were largely unanticipated prior to the beginning of these surveys.

- The next generation of surveys has the potential to tell us a great deal about the nature of dark energy, but the unavoidable size & complexity of these surveys will be a problem for outside users.

- By unifying the data and analysis sides into a common framework, STOMP surmounts these obstacles, allowing astronomers to easily move within and between surveys and measurements.

- Keep your eyes open for the next couple months. Something very interesting is on the way...