Precision Cosmology With Large Scale Structure

David Weinberg, Ohio State University

ICTP Cosmology Summer School 2015

Suggested background reading

Observational Probes of Cosmic Acceleration (*OPCA*), by Weinberg, Mortonson, Eisenstein, Hirata, Riess, and Rozo 2013, Phys Rep 530, 87-255 (arXiv:1201.2434), especially chapters 1, 2, 9.

Lecture 1: Questions and Methods

Cosmological Questions

Broad categories

"Fundamental" questions: matter and energy contents of the universe, initial conditions, global structure, origin

"Astrophysical" questions: physics of galaxy formation, intergalactic medium, clusters, etc.

These lectures will focus mainly on the former category.

Less closely intertwined than they used to be, since uncertainties in the cosmological model are no longer a major source of uncertainty in galaxy formation, and we have many more probes of cosmology besides galaxy clustering.

Fundamental cosmology questions, circa 1990

• Is the gravitational instability picture basically correct?

• What were the properties of the initial fluctuations, and where did they come from?

• What is the dark matter?

• What is Ω?

——

• What is the relation between the distribution of galaxies and the underlying distribution of mass?

Largely answered!! Or at least the range of possible answers has been sharply curtailed in all cases.

Fundamental cosmology questions, today

• Why is the universe accelerating?

• What is dark matter?

• What are the masses of neutrinos?

• What are the departures of initial conditions from scale-invariant, Gaussian, adiabatic, scalar perturbations?

• What is the physics of inflation? (Is inflation correct?)

The last two bullets are observational and theoretical formulations of the same theme.

More generally, we can search for cracks in the "standard cosmological model."

The Standard Cosmological Model

Over the past two decades, cosmologists have converged on a "standard cosmological model," known as ΛCDM, that explains an impressively wide range of observational data with relatively few adjustable parameters.

The physics of ΛCDM

ΛCDM can be summarized as "inflationary cold dark matter with a cosmological constant."

In its simplest "vanilla" form, it assumes:

• Primordial fluctuations as predicted by simple inflation models: adiabatic (present equally in matter, radiation, etc.), Gaussian, and nearly scale-invariant — $P(k) \propto k^{n_s}$ with $n_s \approx 1$

• Dark matter that is weakly interacting (or has purely gravitational interactions), non-baryonic, and "cold" in the sense that its initial thermal velocities were too small to affect galaxy formation

• A cosmological constant Λ, with energy density that is constant in space and time

• A flat universe ($\Omega = 1$), as predicted by inflation

The model has some adjustable parameters. In minimal form, these can be reduced to six, which can be taken as:

$\Omega_c h^2$ – physical density of CDM

$\Omega_b h^2$ – physical density of baryons

Ω_Λ – energy density of cosmological constant

A_s – amplitude of primordial power spectrum

n_s – index of primordial power spectrum ("tilt" is $n_s - 1$)

τ – the Thomson optical depth for CMB photons, depends on redshift of reionization

The radiation energy density $\Omega_r h^2$ is known from the CMB temperature plus standard neutrino physics to give the neutrino contribution (I'm glossing over neutrino masses for the moment).

With a flat universe $\Omega_c + \Omega_b + \Omega_r + \Omega_\Lambda = 1$, so the Hubble constant h is not a separately adjustable parameter.

τ is a (physically interesting) "astrophysical nuisance parameter," necessary because the observed amplitude of the CMB power spectrum is (on most scales) proportional to $A_s e^{-2\tau}$, not simply A_s.

At low redshift, we often refer to the matter fluctuation amplitude in terms of σ_8, the rms fluctuation in $8h^{-1}$ Mpc spheres as predicted by linear perturbation theory.

The empirical basis of the standard model

Vanilla ΛCDM gives an impressively good match to:

• CMB temperature and polarization anisotropies over the full observed range

• Measurements of the cosmic expansion history from supernovae and baryon acoustic oscillations (BAO)

• The shape of the power spectrum of matter clustering as inferred from galaxy clustering and the Lyα forest

• The amplitude of matter clustering as inferred from the abundance of galaxy clusters, weak gravitational lensing, redshift-space distortions, the Lyα forest

With plausible assumptions about baryonic physics it yields reasonably good explanations for the observed properties and evolution of galaxies, galaxy clusters, and the intergalactic medium, over redshifts $z \approx 0 - 6$.

Caveats:

Multiple tensions at the roughly 2σ level:

• the lowest CMB multipoles

• Lyα forest BAO measurement

• direct (distance ladder) measurements of H_0

• low redshift measurements of matter clustering, most of which yield lower amplitudes than predicted by forward extrapolation from the CMB

All of these may go away with improved data and/or modeling. I think the last one has the best chance of being a real discrepancy with interesting physical implications.

Also:

• There is long-standing controversy about whether CDM predicts the correct properties of galaxies, particularly the inner density profiles and the abundance of low mass galaxies.

The galaxy formation tensions could indicate something interesting about the properties of dark matter (warm, self-interacting), but the uncertainties in the baryonic physics are large enough that it is hard to make this case convincing.

Alternatives: What kinds of physically interesting variants might there be?

Many cosmological models that would be plausible *a priori* (e.g., a universe with only baryonic matter) are firmly ruled out by the rich array of high-precision measurements that underpins the standard model.

Any alternative model has to pass the tests that ΛCDM has already passed.

Observations place ever tighter constraints on the parameters of the standard model, and we can hope that they will yield inconsistencies that point to new physics.

A "guaranteed" extension of the standard model is

• Neutrinos have non-zero mass, change from "radiation" to "matter." $\Omega_\nu h^2$ should be measurable.

A "plausible" extension of the standard model is

• tensor contributions to the CMB from primordial gravity waves, introducing new parameters $r = A_t/A_s$ and n_t

Other extensions that are not "expected" but would be physically interesting revisions within the standard framework are:

• dynamical dark energy, not constant in space and time, described by $w(z)$ not $w = -1$

• extra relativistic species, usually referred to by $N_\nu \neq 3.046$

• significant curvature or features in the inflationary power spectrum, "running" or steps

• non-Gaussianity of primordial fluctuations, with implications for inflation

• non-flat universe, $\Omega_k \neq 0$, important implications for inflation and global structure of universe

More radical revisions include:

• departure from GR, a gravitational explanation for cosmic acceleration

• decaying dark matter

• other variations of DM properties that change galaxy formation physics

• isocurvature fluctuations

Cosmological Probes and Their Information Content

Brief overviews here. Some treated in more detail in other lectures.

The cosmic microwave background (CMB)

Suggested reading: *Cosmic Microwave Background Anisotropies* by Hu & Dodelson 2002, Ann Rev Astron Astrophys 40, 171-216. You can get a copy from Wayne Hu's web page.

Hu, Sugiyama, & Silk 1997, Nature 386, 37 is a classic short introduction to CMB physics

For current observations, Planck 2015, especially paper XIII.

The virtues of the CMB as a cosmological probe are:

• Controllable observational systematics, usually limited by detector noise and/or cosmic/sample variance. Polarization systematics (instrumental and foregrounds) remain challenging.

• Linear physics, usually not limited by theoretical uncertainty

• CMB power spectrum responds to many different physical effects, lots of power to constrain cosmological parameters

Rough summary of information content:

• Heights of peaks constrain $\Omega_c h^2$, $\Omega_b h^2$. Given these, and radiation content (CMB + standard neutrinos), can compute sound horizon

$$r_s = \int_0^{t_*} \frac{c_s(t)}{a(t)} dt, \quad c_s(t) = \frac{c}{\sqrt{3}} \left[1 + \frac{3\rho_b(z)}{4\rho_\gamma(z)} \right]^{-1/2}$$

• Angular scale of peaks determines $r_s/D_A(z_*)$ with near perfect accuracy. Very sensitive to curvature, also depends on other parameters.

• Overall shape of spectrum constrains n_s, other departures from scale invariance

• Amplitude constrains $A_s e^{-2\tau}$

• Low multipole polarization constrains τ

• Damping tail constrains history of recombination, any departures from standard model prediction

• Low multipoles constrain tensor/scalar ratio

• Lensing (4-point correlations, details of power spectrum shape, B-mode polarization) constrains amplitude of matter clustering, most sensitive to $z = 1 - 4$

• High-l sensitive to SZ contributions, from clusters and IGM, depends on matter clustering and gas physics.

• Large angle B-mode polarization can detect inflationary gravity waves

Interlude: Some Important Equations

The Friedmann equation

$$\frac{H^2(z)}{H_0^2} = \Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_k(1+z)^2 + \Omega_\phi \frac{u_\phi(z)}{u_\phi(z=0)}$$

Evolution of Ω_m

$$\Omega_m(z) \equiv \frac{\rho_m(z)}{\rho_{\rm crit}(z)} = \Omega_m(1+z)^3 \frac{H_0^2}{H^2(z)}$$

Comoving distance

$$D_C(z) = \frac{c}{H_0} \int_0^z dz' \frac{H_0}{H(z')}$$

Comoving angular diameter distance

$$D_A(z) \approx D_C \left[1 + \frac{1}{6}\Omega_k \left(\frac{D_C}{c/H_0}\right)^2\right]$$

Linear growth of fluctuations

$$\delta(\mathbf{x}, t) \;\equiv\; \frac{\rho_m(\mathbf{x}, t) - \bar{\rho}_m(t)}{\bar{\rho}_m(t)} \;=\; \delta(\mathbf{x}, t_i) \times \frac{G(t)}{G(t_i)}$$

$$\ddot{G}_{\rm GR} + 2H(z)\dot{G}_{\rm GR} - \frac{3}{2}\Omega_m H_0^2 (1+z)^3 G_{\rm GR} = 0$$

Approximations for linear growth rate and growth factor

$$f_{\rm GR}(z) \equiv \frac{d \ln G_{\rm GR}}{d \ln a} \approx [\Omega_m(z)]^\gamma$$

$$\gamma = 0.55 + 0.05[1 + w(z=1)]$$

$$\frac{G_{\rm GR}(z)}{G_{\rm GR}(z=0)} \approx \exp\left[-\int_0^z \frac{dz'}{1+z'} [\Omega_m(z')]^\gamma\right]$$

Evolution of dark energy density

$$\frac{u_\phi(z)}{u_\phi(z=0)} = \exp\left[3 \int_0^z [1 + w(z')] \frac{dz'}{1+z'}\right]$$

Type Ia supernovae and other standard candles

Suggested reading: *OPCA* ch. 3, Betoule et al. 2014 (1401.4064)

Type Ia supernovae measure the relative distance scale, i.e., $D_L(z_2)/D_L(z_1)$.

With good observations, the rms distance error per SNIa is 5-8%, so many supernovae can yield very high precision.

Key challenges are observational systematics — photometric calibration, dust extinction — and possible redshift evolution of the SNIa population.

Current data yield $\sim 1-2\%$ relative distance measurements over range $z = 0 - 0.8$, limited mainly by observational systematics.

Absolute distance measurements require absolute calibration of the SNIa luminosity scale, which is difficult.

Baryon acoustic oscillations (BAO)

Suggested reading: *OPCA* ch. 4, Aubourg et al. 2015 (1411.1074)

BAO measure absolute distances and expansion rates, $D_A(z)$ and $H(z)$, calibrated to r_s, which (for standard matter and radiation content) is known to 0.4% precision from current CMB data.

Achievable precision grows with redshift because more volume available to measure high precision clustering.

Same standard ruler measured by CMB, providing long redshift baseline.

Current measurements are $1-2\%$ precision, limited by statistics.

Likely to remain statistics limited even for cosmic variance limited surveys to $z \sim 3$.

Much more in Lecture 3.

Clusters of galaxies

Suggested reading: *OPCA* ch. 6, Mantz et al. 2014 (1407.4516)

Abundance of clusters as a function of mass constrains a parameter combination that is approximately $\sigma_8(z)\Omega_m^\alpha$ with $\alpha \approx 0.3 - 0.5$. (Note that Ω_m is the $z = 0$ value.)

Can find clusters by X-ray emission, optical richness, SZ decrement.

Biggest challenge is accurate calibration of cluster mass scale.
Weak lensing is most promising route to achieving percent-level precision.

Current measurements are 5-10% precision in $\sigma_8\Omega_m^\alpha$, limited by mass calibration uncertainty.

Weak lensing

Suggested reading: *OPCA* ch. 5, Heymans et al. 2013 (1303.1808)

Cosmic shear weak lensing measures combination of matter clustering amplitude and distances.

Intuitively useful formula for linear theory shear variance for sources at z_s and lenses at $z_s/2$ in a flat universe:

$$\Delta^2(l) \equiv \frac{l^2}{2\pi} C_{EE}(l) = 1.8 \times 10^{-4} \sigma_8^2 \left[\left(1 + \frac{z_{\rm s}}{2}\right) G\left(\frac{z_{\rm s}}{2}\right) \right]^2 \Omega_m^2 [H_0 D_C(z_{\rm s})]^{2.3} l^{0.7}$$

At present, especially useful as a constraint on matter clustering amplitude given expansion history constrained by other probes.

With future surveys like LSST, Euclid, WFIRST, also becomes very high precision test of expansion history because of sensitivity to distances.

Key challenges are shear measurement systematics, photo-z uncertainties, contamination of signal by intrinsic galaxy alignments, baryonic effects on predicted signal.

Very tough but great potential.

Galaxy-galaxy lensing and cluster-galaxy lensing are also potentially powerful, with somewhat different systematics.

More on these in Lecture 2, and more on weak lensing in Jain lectures.

Redshift-space distortions (RSD) and the Alcock-Paczynski (AP) effect

Suggested reading: *OPCA* §7.2, Reid et al. 2012 (1203.6641)

In linear theory, redshift-space distortions of galaxy clustering constrain a parameter combination $f(z)\sigma_8(z)$ where $f(z)$ is the clustering growth rate $d\ln G/d\ln a$.

Non-linear effects allow some degeneracy breaking, but also complicate modeling.

Key challenge, especially going forward, is modeling RSD at level of precision achievable with observations. Constraining power is a strong function of $k_{\rm max}$ or $s_{\rm min}$.

Current constraints at $\sim 10\%$ level in $f(z)\sigma_8(z)$, limited by combination of statistics and modeling systematics.

In absence of peculiar velocities, clustering would be isotropic if one assumes the correct ratio of $D_A(z)/H^{-1}(z) = D_A(z)H(z)$ to relate angular and redshift separations (AP test).

BAO automatically gets $D_A(z)H(z)$ by measuring both separately, but isotropy can constrain product using higher precision clustering measurements on smaller scales.

Key challenge is modeling RSD well enough to remove its effect.

Measurements that extend to BAO scale are good for breaking AP/RSD degeneracy.

More in Lecture 2.

Galaxy clustering

In addition to RSD and AP, galaxy clustering can constrain parameters through shape of power spectrum.

Analogous to CMB, but baryonic effects on galaxy $P(k)$ shape are much weaker.

Turnover scale probes Ω_m/Ω_r.

Tilt probes n_s.

Large vs. small scale can probe Ω_ν.

Sensitivity to h because observed scales are h^{-1} Mpc.

Key challenge is modeling non-linear evolution of matter power spectrum and, especially, scale-dependent bias between galaxies and matter.

Current measurements in good agreement with ΛCDM, strengthen constraints for flexible models.

Further discussion in Lecture 2.

The Lyα forest

The Lyα forest traces fluctuating neutral hydrogen absorption, which in turn traces underlying matter fluctuations.

Powerful probe of structure at $z = 2 - 4$, where Lyα is accessible to ground-based observations and forest absorption is not saturated.

Basic physics is straightforward, though details are complex.

Before SDSS-III/BOSS, Lyα was a "1-d" subject, with each quasar spectrum providing an independent map of structure along its line of sight.

BOSS was designed to allow 3-d measurements from sightline cross-correlations.

BOSS has enabled first measurements of BAO in the Lyα forest.

Smaller scale measurements, 1-d and 3-d, can probe shape and amplitude of matter power spectrum at sub-Mpc to multi-Mpc scales. Growth factor at high z, neutrino masses, tilt, etc.

Current BAO measurement precision is 2%, limited (probably) by statistics.

1-d $P(k)$ measurements are very high precision, yielding tightest upper limit on neutrino masses ($\sum m_\nu < 0.15$ eV, Palanque-Delabrouille et al. 2014). Limited by accuracy of theoretical modeling.

Direct measurement of H_0

Reading: *OPCA* §7.1.

Value of H_0 determines the present day critical density. Maximum lever arm compared to CMB.

Many challenges in direct distance ladder measurement, including photometric calibration, extinction, blending, basic calibration of Cepheid distance scale.

Intriguing tension between values $H_0 \approx 70 - 75\,\mathrm{km\ s^{-1}\ Mpc^{-1}}$ from best direct studies and $H_0 \approx 67$ predicted for ΛCDM with Planck CMB constraints.

BAO+SN "inverse distance ladder" measurement in Aubourg et al. (2015) yields $H_0 = 67.3 \pm 1.1\,\mathrm{km\ s^{-1}\ Mpc^{-1}}$, excellent agreement with ΛCDM predictions.

Could be reconciled with direct distance ladder by modifying *pre-recombination* physics to decrease sound horizon (e.g., adding radiation or early dark energy).

But probably indicates that direct measurements are too high.

21cm at high redshift

A new frontier in observational cosmology is mapping very high redshift structure ($z \sim 6 - 20$) using (redshifted) 21cm emission and absorption.

In principle this opens access to an enormous comoving volume and thus huge numbers of Fourier modes.

This could ultimately yield cosmological measurements that are much higher precision than those possible with galaxy or Lyα forest surveys, as well as access to a new redshift range.

There are many instrumental, astrophysical foreground, and modeling uncertainties to overcome.

At $z = 0.5 - 3$, "21cm intensity mapping" offers a potentially inexpensive way to achieve cosmic variance limited BAO measurements.

Forecasting Experimental Performance

Forecasting the ability of an experiment or survey to constrain cosmological parameters or test models has become a major area of research in itself.

Forecasting can be useful for making the case for a survey or space mission, for optimizing strategy or instrument design, for understanding the interplay among different empirical constraints, and for understanding the influence of systematic and statistical uncertainties.

Basic idea: Assume a fiducial model, and estimate the errors and error covariances for your proposed experiment. Translate these errors into errors on the parameters of interest.

The most frequently used approach is Fisher matrix forecasting, which relies on some simplifying assumptions about the errors.

There is a pretty good high-level discussion of this topic in section 2 of Tegmark, Taylor, & Heavens (1997, ApJ, 480, 22) and a valuable but dense presentation in very different notation by Gould (2003, arXiv:astro-ph/0310577).

Caution: The output of a forecast is only as good as the assumptions that go into it.

Warmup:

Suppose we have an observable y_1 that we can predict given some model parameter θ_1, and that we measure y_1 with some observational error $\sigma(y_1)$.

Simple "chain rule" error propagation then tells us that the error on θ_1 is

$$\sigma(\theta_1) = \left(\frac{d\,y_1}{d\theta_1}\right)^{-1} \sigma(y_1).$$

Often we are interested in the fractional error

$$\frac{\sigma(\theta_1)}{\theta_1} = \sigma(\ln\theta_1) = \left(\frac{d\ln y_1}{d\ln\theta_1}\right)^{-1} \sigma(\ln y_1).$$

For example, if $y_1 \propto \theta_1^3$, then the fractional error on θ_1 is only 1/3 the fractional error on y_1.

If we can anticipate the observational error we will get from some measurement or data set, then we can forecast the error we will get on the parameter θ.

More general case:

If we have a parameter vector $\vec{\theta}$, the Fisher information matrix is defined by

$$F_{ij} = -\left\langle \frac{\partial^2 \ln L}{\partial\theta_i \partial\theta_j} \right\rangle.$$

To the extent that the likelihood is well described by a quadratic Taylor expansion about the maximum likelihood value, the expected error on parameter θ_i is

$$\sigma_i \equiv \sigma(\theta_i) = (F^{-1}_{ii})^{1/2}$$

if all of the parameters are being estimated from the data set and

$$\sigma_i \equiv \sigma(\theta_i) = (F_{ii})^{-1/2}$$

if all parameters other than θ_i are known.

Under more general conditions, the error of any unbiased estimator must be greater than or equal to these values, a result known as the *Cramér-Rao Bound*.

In *Fisher matrix forecasting*, we assume a fiducial model and properties of a data set to predict the Fisher matrix and thereby forecast the errors that will be obtained on model parameters.

Fisher matrix for Gaussian likelihoods

Suppose we have a Gaussian likelihood function for N data points

$$-\ln L = \frac{N}{2}\ln(2\pi) + \frac{1}{2}\ln[\det(\mathbf{C})] + \frac{1}{2}\Delta_k C^{-1}_{kl}\Delta_l,$$

where $\Delta_k = y_k - y_{\rm mod}(x_k)$ and the covariance matrix is

$$C_{ij} = \langle (y_i - \langle y_i\rangle)(y_j - \langle y_j\rangle)\rangle \ .$$

Assume that we can ignore any dependence of the covariance matrix on the model parameters. This is a non-trivial assumption that will not always hold. For example, in cosmological applications we sometimes have "cosmic variance" errors that depend on the amplitude of matter or galaxy clustering, and the expected size of these errors depends on the cosmological parameters.

If we make this assumption, the derivative of det$\mathbf{C}$ with respect to parameters vanishes, and the Fisher matrix becomes

$$F_{ij} = \frac{1}{2}\left\langle \frac{\partial^2 \Delta_k C^{-1}_{kl}\Delta_l}{\partial\theta_i\partial\theta_j}\right\rangle .$$

which after some manipulation can be reduced to

$$F_{ij} = \frac{\partial\Delta_k}{\partial\theta_i} C^{-1}_{kl}\frac{\partial\Delta_l}{\partial\theta_j}.$$

(Einstein summation convention in use.)

Though notationally different, I think this is equivalent to equation (15) of Tegmark et al. (1997), except that the term $\mathbf{A}_i\mathbf{A}_j$ in that equation has vanished because we have assumed that the dependence of C_{ij} on the parameters can be neglected.

Sensitivity and observational errors

We can decompose a Fisher matrix into a matrix product:

$$F_{ij} = -\left\langle \frac{\partial^2 \ln L}{\partial\theta_i\partial\theta_j}\right\rangle = -\left\langle \frac{\partial y_k}{\partial\theta_i}\cdot\frac{\partial^2 L}{\partial y_k\partial y_l}\cdot\frac{\partial y_l}{\partial\theta_j}\right\rangle .$$

If the errors on the observables are Gaussian *and* independent of the model parameters, then

$$\frac{\partial^2 \ln L}{\partial y_k\partial y_l} = C^{-1}_{kl},$$

the inverse covariance matrix.

Thus, the Fisher matrix has an "outer" piece $\partial\vec{y}/\partial\vec{\theta}$ that represents the sensitivity of the observables to the parameters and an "inner" piece that represents the errors on the observables themselves.

If we consider the 1-parameter, 1-observable case, we get

$$F_{11} = \frac{\partial y}{\partial \theta} \cdot \frac{1}{\sigma_y^2} \cdot \frac{\partial y}{\partial \theta},$$

implying

$$\sigma^2(\theta) = 1/F_{11} = \sigma_y^2 \left(\frac{\partial y}{\partial \theta}\right)^{-2},$$

in agreement with our earlier chain rule result.

For a Fisher matrix forecast of parameter errors, we compute the parameter sensitivity from our model, and we take the expected values of the observable errors (and their covariances).

As far as I know, $\partial\mathbf{y}/\partial\theta$ doesn't have a special name, but we can think of it as an "influence matrix" or "sensitivity matrix."

While computing the Fisher matrix requires assumptions about the data set, the sensitivity matrix requires only knowledge of the model, and it can be an interesting quantity to compute even if one doesn't have a specific data set in mind.

Adding Fisher matrices

Suppose we have two data sets that are statistically independent.

In this case, the joint likelihood (or posterior probability) is just the product of the individual likelihoods (or posterior probabilities), since $p(x, y) = p(x)p(y)$ for independent variables.

Therefore, one obtains $\langle \ln L \rangle$ for the two data sets by adding the two individual values of $\langle \ln L \rangle$, and the Fisher matrix for the two data sets is just the sum of the Fisher matrices for the individual data sets.

This still holds even if the data sets are quite different in character provided they constrain the same underlying parameters.

For example, one can forecast cosmological parameter errors that will be obtained by joint fits to CMB data, supernova data, and a direct measurement of H_0 by adding the Fisher matrices for the three data sets.

This is a powerful technique.

Lecture 2: Theoretical Approaches

Dark Matter Clustering

Linear perturbation theory

N-body simulations

Higher order perturbation theory

Treated in much more detail in Zaldarriaga lectures

Halo model of dark matter clustering

A Sketch of Galaxy Formation Theory

Modeling Galaxy Clustering

Hydrodynamic simulations

(Methods discussed in Borgani lecture)

Populating DM halos via semi-analytic models

Abundance matching and age matching

Halo occupation distribution (HOD) modeling

Galaxy-galaxy lensing and cluster-galaxy lensing

Theory of redshift-space distortions (RSD)

Linear theory and its failings

Ways to go further

The Lyα forest

Basic phenomenology and physics

Hydro simulations

1-d and 3-d measures of Lyα forest structure

Methods for modeling large volumes

How accurate does theory need to be?

Lecture 3: Observational Prospects

Where are we now?

Current major data sets in CMB, supernovae, BAO, galaxy clustering, Lya forest, weak lensing, clusters

BAO surveys and analysis

These are interesting in their own right, and they illustrate points relevant to other observational probes.

Basic BAO methodology

Sampling and volume: shot noise vs. sample variance

Reconstruction

Fitting and nuisance parameters

Estimating errors and covariance matrices

Observational systematics

Theoretical systematics

BAO with the Lyman-alpha forest

BAO with 21cm intensity mapping

Principles of survey design

Figures of Merit and their perils

Sources of statistical error

Observational systematics

Theoretical systematics

Statistical and systematic errors

What will dominate your total errors and how do you control it?

Mitigating systematics by marginalization

Survey design: observational aspects

Imaging vs. spectroscopic surveys

Telescope aperture and field of view

Pixel size and sampling

Spectroscopic multi-plexing

Object noise, sky noise, read noise

The basic trade: area vs. depth

What's on the horizon?

I'll spend most time on the ones marked with $^+$, as I know them the best and they illustrate most of the principles

Dark Energy Survey$^+$

Hyper-Suprime Cam

eBOSS

HETDEX

DESI$^+$

Subaru PFS

LSST

Euclid

WFIRST$^+$

Where might we be in 2030?