Large Scale Structure: comparison of theory with observation (I)

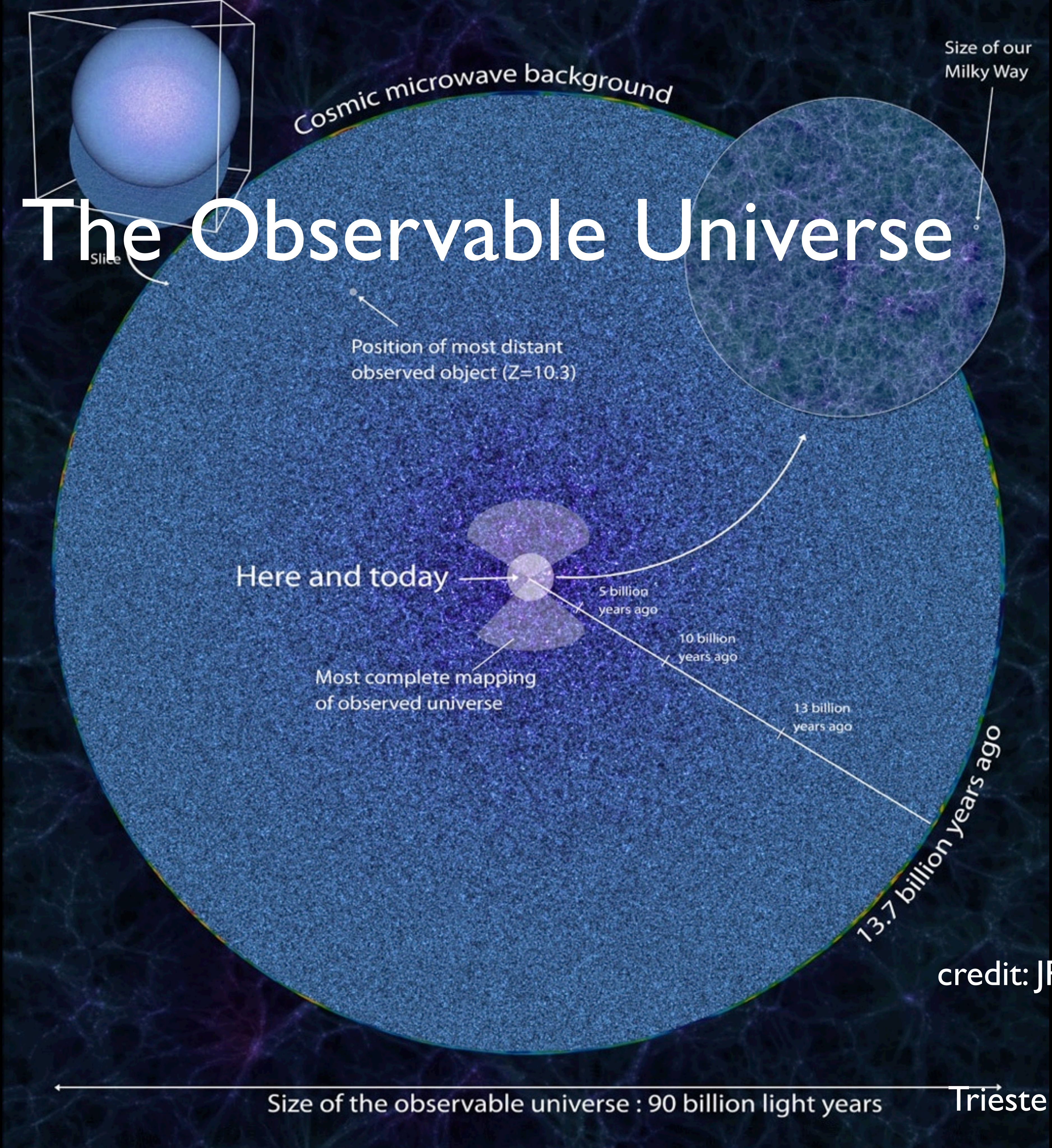
Beth Reid
Cosmology Data Science Fellow
UC Berkeley Center for Cosmological Physics/LBNL

The Sloan Digital Sky Survey 2.5m telescope
Apache Point, New Mexico

Outline

- The Big Picture
- What can we measure?
- What is our theory?
- Statistical Interlude
- Fitting a Toy Universe

The Observable Universe



credit: JP Kneib/eBOSS

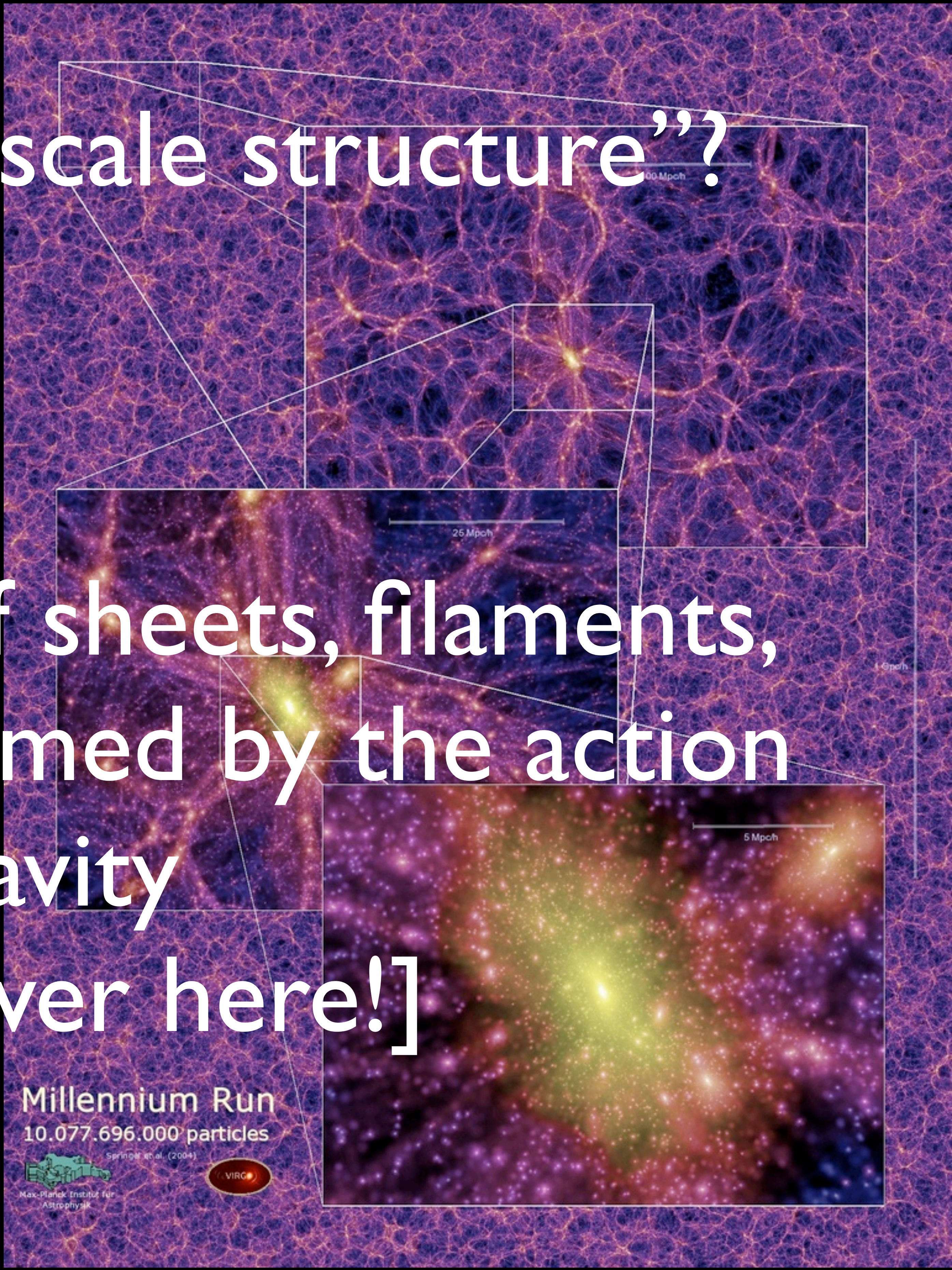
SDSS Fly-through

credit:
Miguel A Aragon (JHU), Mark Subbarao (Adler Planetarium), Alex Szalay (JHU)

Q: What is “large scale structure”?

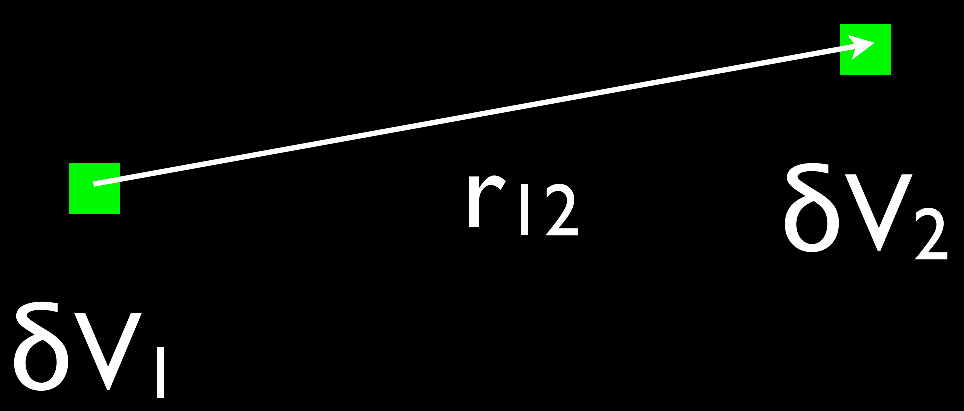
A: The network of sheets, filaments, halos, voids, etc formed by the action of gravity

[Your Answer here!]



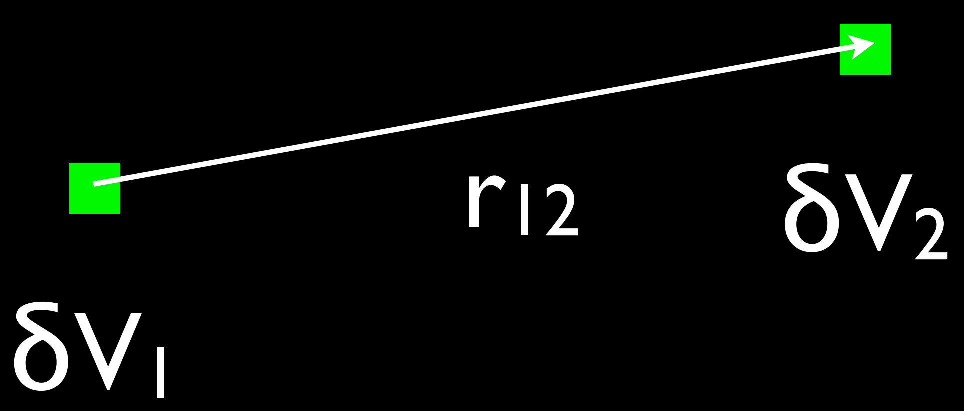
Q: How do we characterize "large scale structure"?

- One profitable way is the two-point correlation function.
- Start with a sample of galaxies with number density n. The probability to find a galaxy within a tiny volume δV_1 is $n \, \delta V_1$.



Q: How do we characterize "large scale structure"?

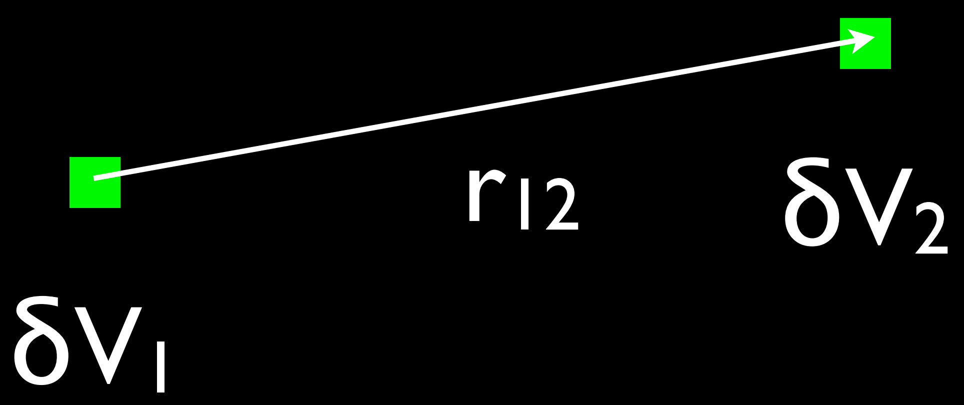
- What is the probability to find a galaxy in δV_1 and δV_2?
- If they are independent events, then $P(1,2) = n^2\, \delta V_1\, \delta V_2$
- The two-point correlation function ξ measures the probability in excess of random to find a galaxy at both δV_1 and δV_2: $P(1,2) = n^2\, (1+ \xi(r_{12}))\, \delta V_1\, \delta V_2$



Q: How do we characterize "large scale structure"?

Exercise: What is the probability of P(~1,~2), i.e., there is no galaxy in δV_1 and there is no galaxy in δV_2?

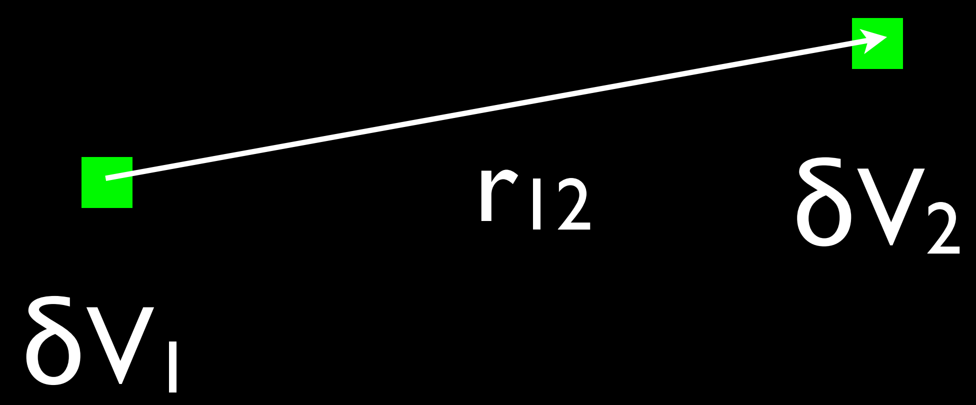
Answer: $1 - n\,\delta V_1 - n\,\delta V_2 + n^2\,(1+\xi)\,\delta V_1\,\delta V_2$



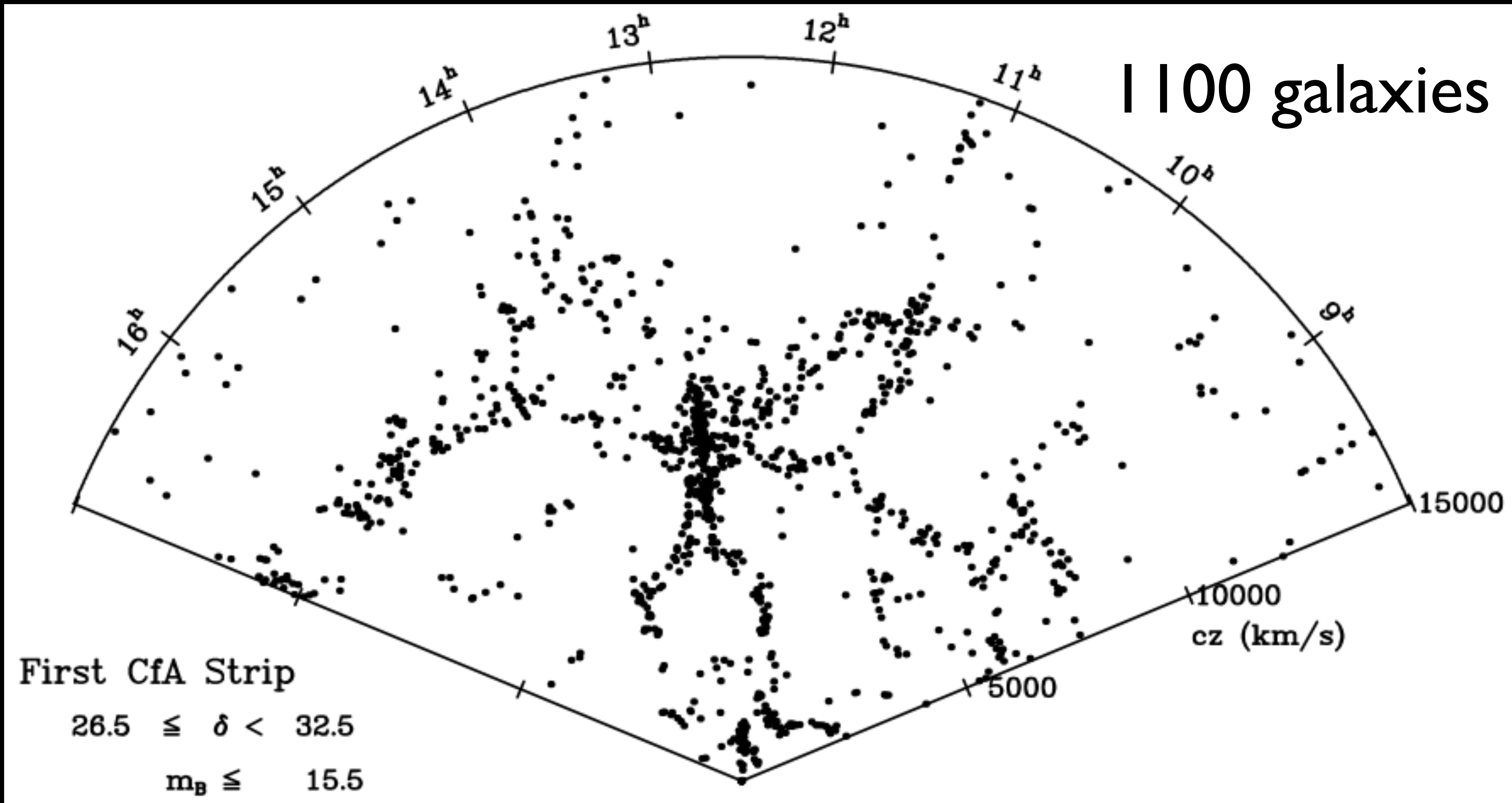
Q: How do we characterize "large scale structure"?

For many applications, it is convenient to work with the Fourier transform of ξ, the power spectrum:

$$P(k) = \int d^3x\, \xi(r)\, e^{-i\, k \cdot r}$$



CfA2 redshift survey



De Lapparent, Geller, Huchra, 1986, ApJL 302, L1

CfA2 redshift survey

"The best available model for generating the bubble-like structures observed in the survey is the explosive galaxy formation theory of Ostriker and Cowie, in which galaxies form on the surfaces of expanding shock waves. Most of the current models ideally assume that the structures form directly from the action of gravity on the matter perturbations, but the sharpness of the transition between the high-density regions and the voids in the survey indicates that hydrodynamic processes must be important in the formation of galaxies."

De Lapparent, Geller, Huchra, 1986, ApJL 302, L1

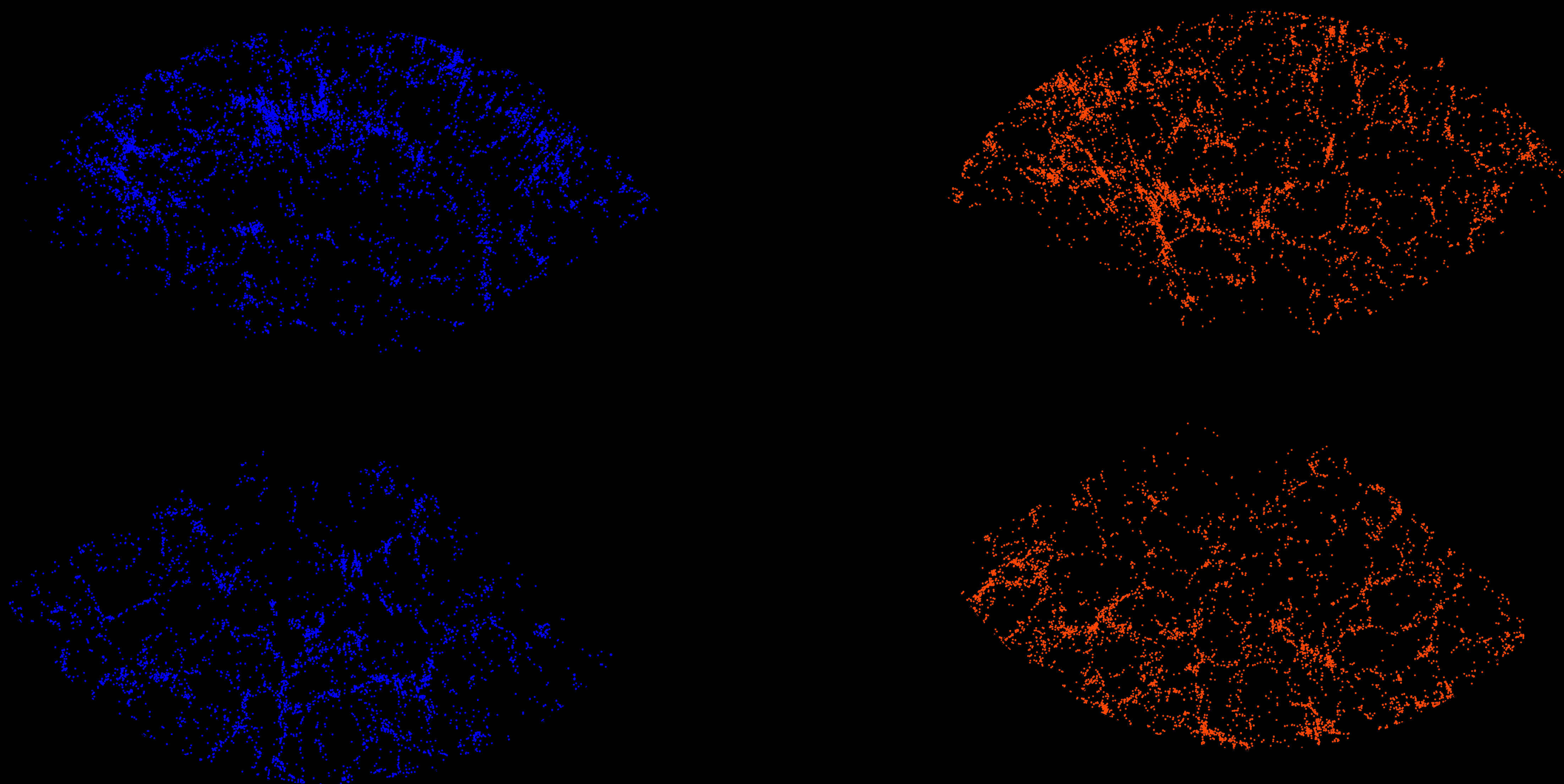
CfA2 redshift survey

"...Most of the current models ideally assume that the structures form directly from the action of gravity on the matter perturbations, but ..."

No but (yet!)

De Lapparent, Geller, Huchra, 1986, ApJL 302, L1

SDSS Main Galaxy Sample -- which is the real universe?



Credit: Andreas Berlind and Cameron McBride

Tracing out the large scale structure (LSS)

- Our theoretical models robustly predict the behavior of mass under the evolution of gravity
- With the exception of gravitational lensing sourced by underlying mass fluctuations, everything we use to study LSS is a *tracer*
- Examples: galaxies/quasars, Ly-α forest, 21 cm, intensity mapping, etc.
- In these lectures, we will develop the tools to connect tracer maps with the underlying cosmological theory
- Focus will be on galaxy redshift surveys

Outline

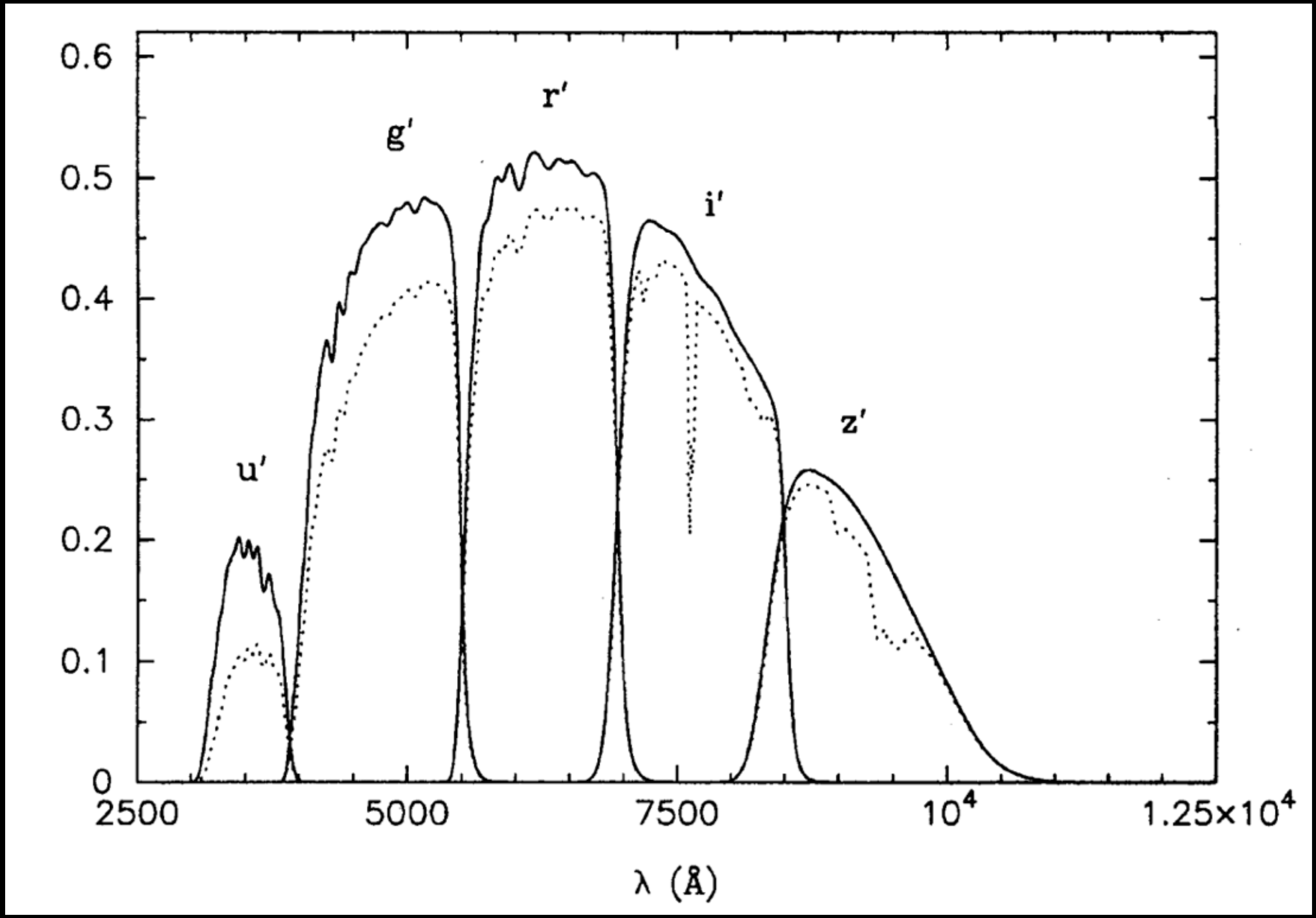
- The Big Picture
- What can we measure?
- What is our theory?
- Statistical Interlude
- Fitting a Toy Universe

What can we measure?

Images -- collect all the light in broad filters

SDSS images are just 54 seconds/filter!

Response function for SDSS filters



Fukugita et al. 1996

What can we measure?

Images -- two dimensional maps

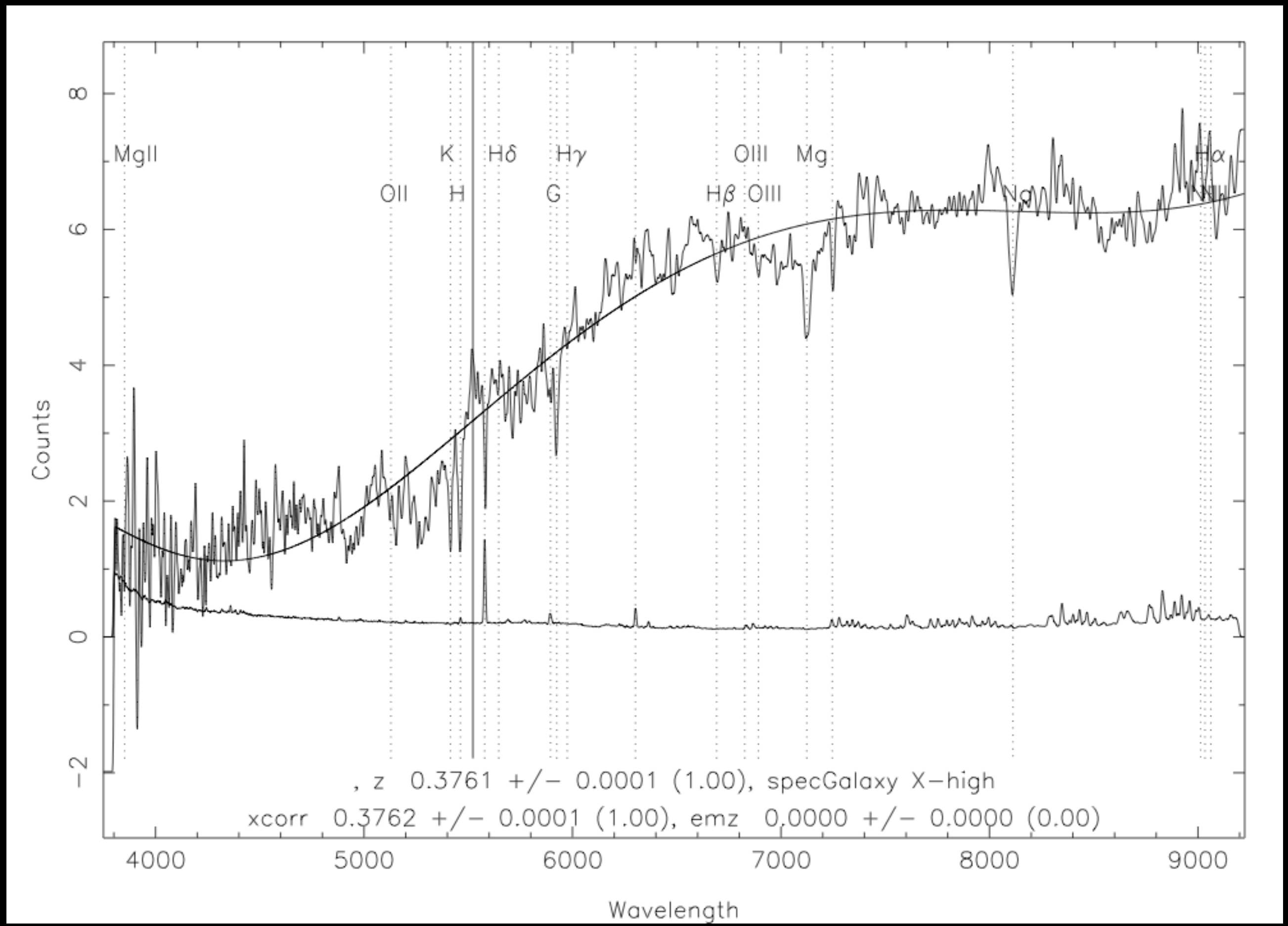
SDSS run 756

Imaging -- what can we measure?

- Galaxy positions in two dimensions (on celestial sphere)
- Galaxy magnitudes in each band --> "photometric redshifts" (allows coarse separation into redshift slices)
- Clustering in 2d (angular correlation function/power spectra)
- Galaxy shapes (allows lensing measurements)
- Look for overdensities of red-sequence galaxies to find clusters (e.g. Rykoff et al. ApJ 785 2014)
- Multiple passes allow search for transients (SN, quasars, other less interesting things..)

What can we measure? Spectra.

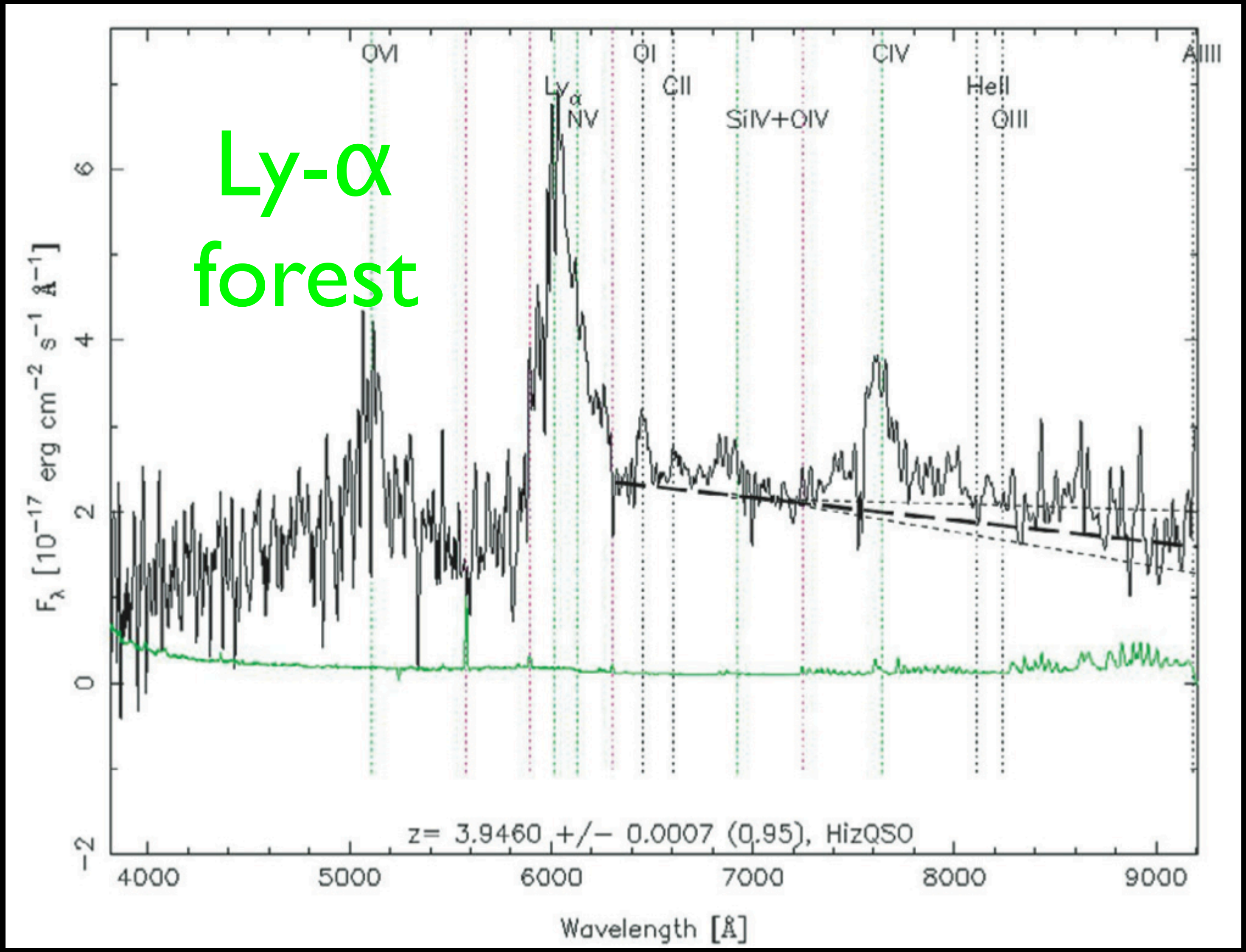
SDSS-III BOSS spectra require ~75 minutes to collect enough photons to accurately measure redshifts. We must choose wisely which objects to target for spectroscopy!



SDSS z = 0.38 galaxy spectrum, after sky subtraction (~99% of photons!)

What can we measure? Spectra.

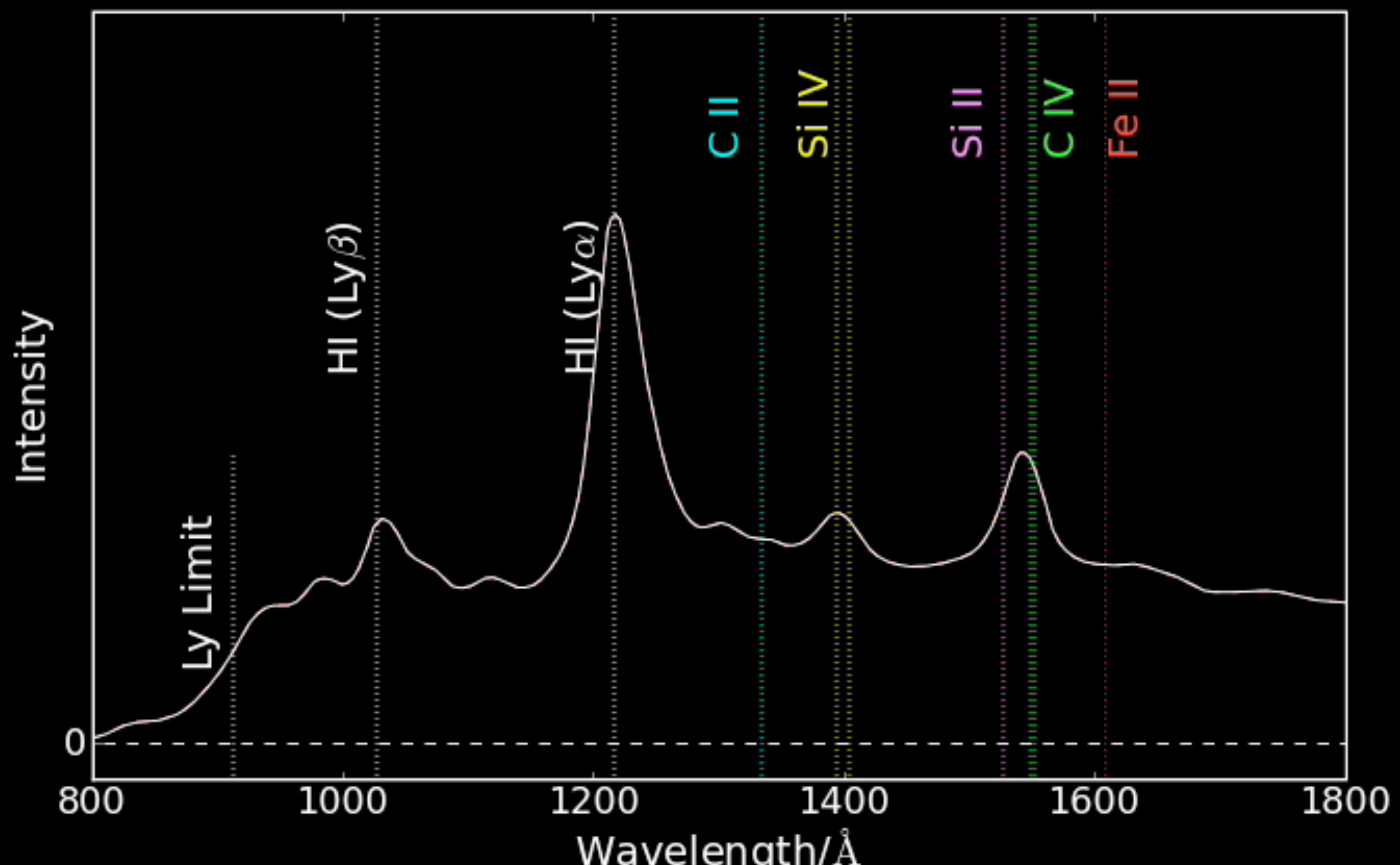
SDSS-III BOSS also targets quasars to measure the Ly-α forest absorption by neutral hydrogen along the line of sight



Giommi et al., A&A 468, 97

SDSS z = 3.95 quasar spectrum

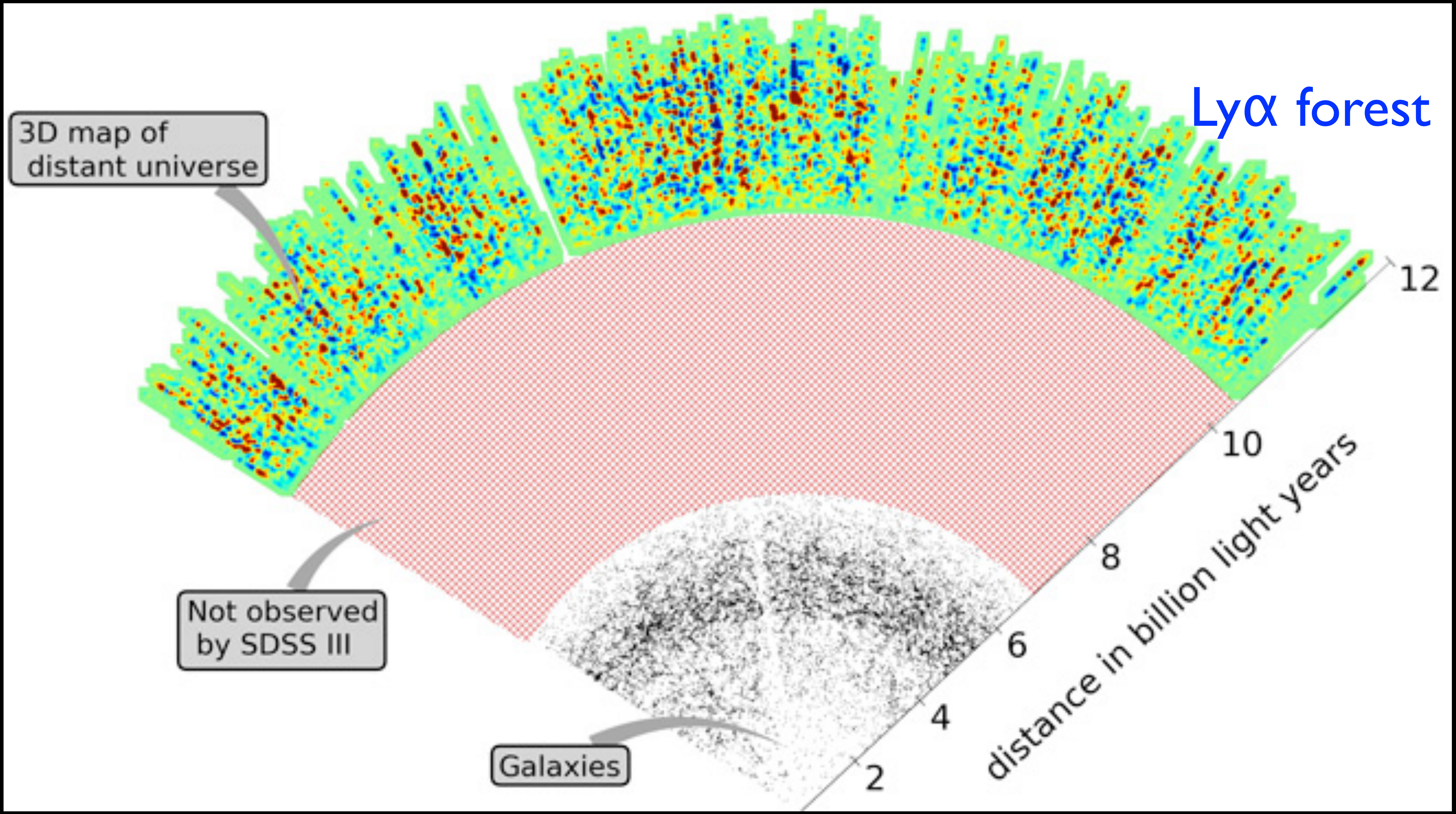
Spectroscopy -- Ly-α forest



Spectroscopy -- what can we measure?

- Galaxy positions in three dimensions (on celestial sphere)
- Lyα forest: absorption of neutral gas along many quasar lines of sight --> 3d absorption field

Spectroscopy -- what can we measure?



The target galaxies of current and future LSS galaxy redshift surveys

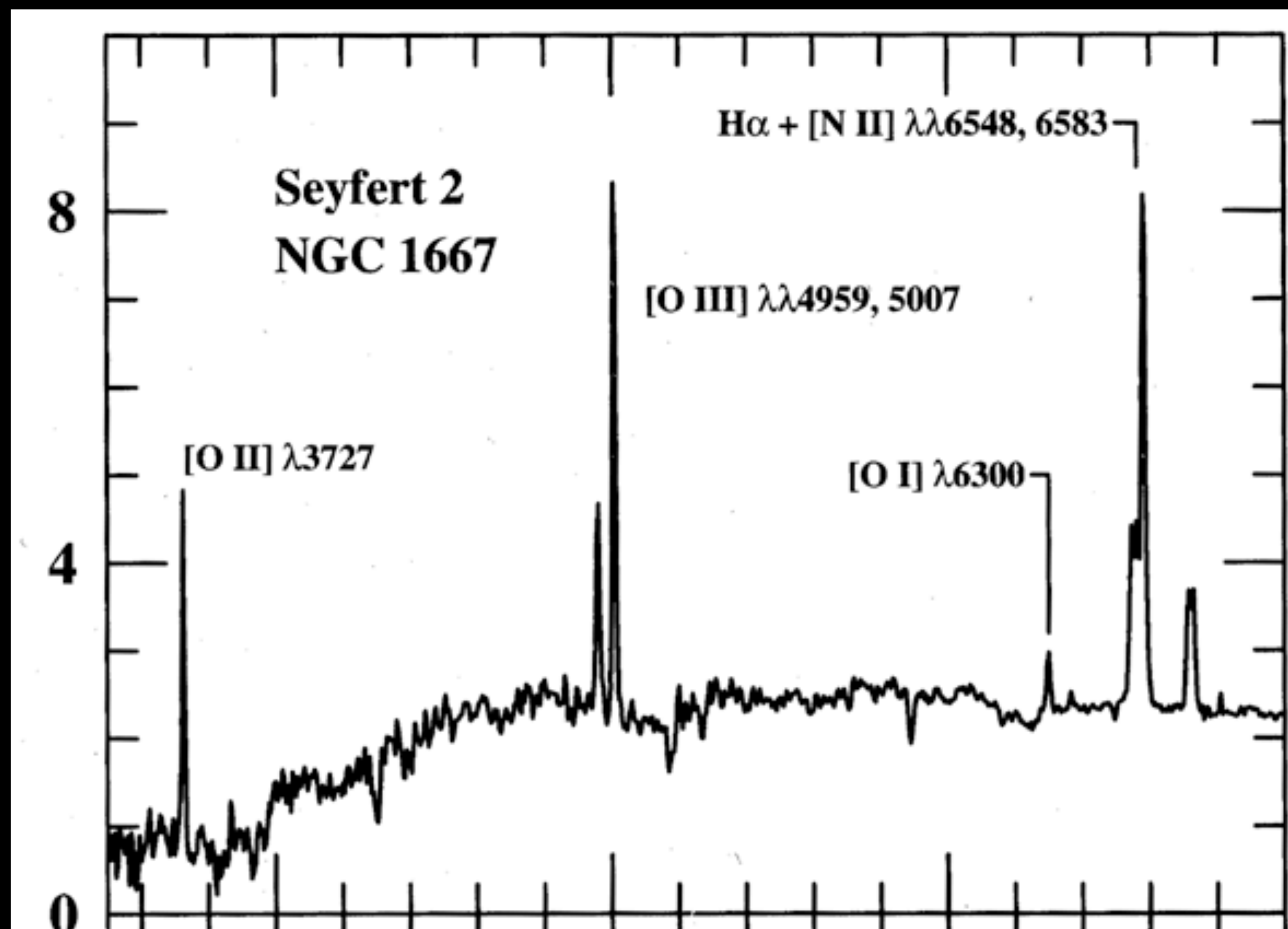
- For a fixed amount of telescope time, we want to minimize our measurement errors on the cosmic density fluctuations
  - Primary: choose targets that take the smallest amount of time to obtain a robust redshift from the spectrum
  - Secondary: choose targets with the most signal per object (highly biased objects; we'll define that later)
  - When designing an experiment, Fisher matrices are a quantitative way to answer such questions

The target galaxies of current and future LSS galaxy redshift surveys

- Luminous Red Galaxies (LRGs): SDSS I/II
- CMASS ($0.4 < z < 0.7$) and LOWZ ($0.2 < z < 0.4$) targets in SDSS-III BOSS are similar types of objects
- These galaxies are “red and dead”: their spectra contains the prominent 4000 Angstrom break which makes obtaining redshifts quick
- Quasars: eBOSS, DESI [new -- as LSS tracers, not just backlights for Ly-α forest]

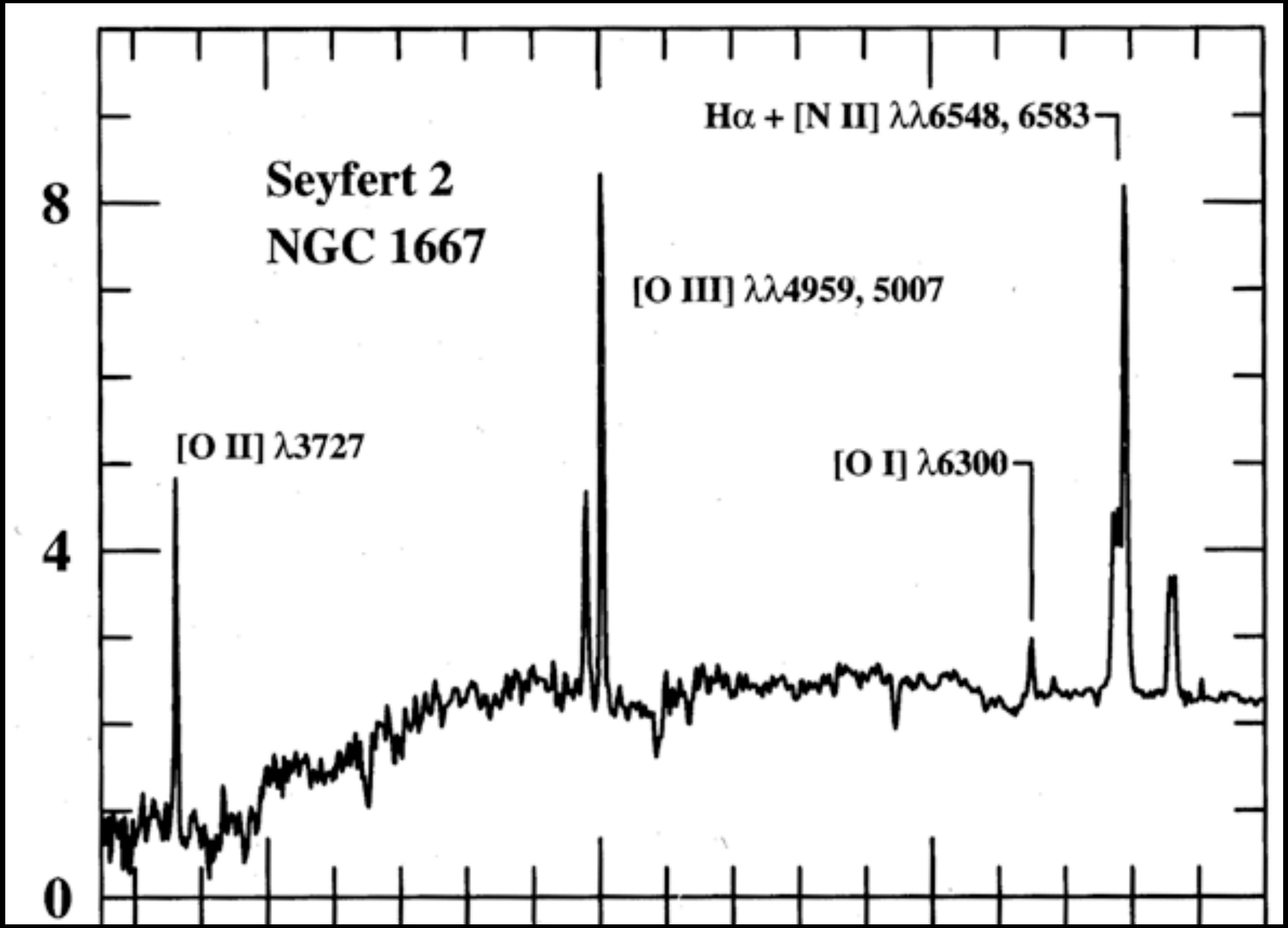
The target galaxies of current and future LSS galaxy redshift surveys

- Emission Line Galaxies: young bright stars in star-forming galaxies photoionize surrounding gas



The target galaxies of current and future LSS galaxy redshift surveys

- WiggleZ, DESI (groud-based optical): [OII] doublet, Euclid/ WFIRST (space-based infrared): Hα, [OIII]



Outline

- The Big Picture
- What can we measure?
- What is our theory?
- Statistical Interlude
- Fitting a Toy Universe
- Estimating LSS covariance matrices

LSS theory

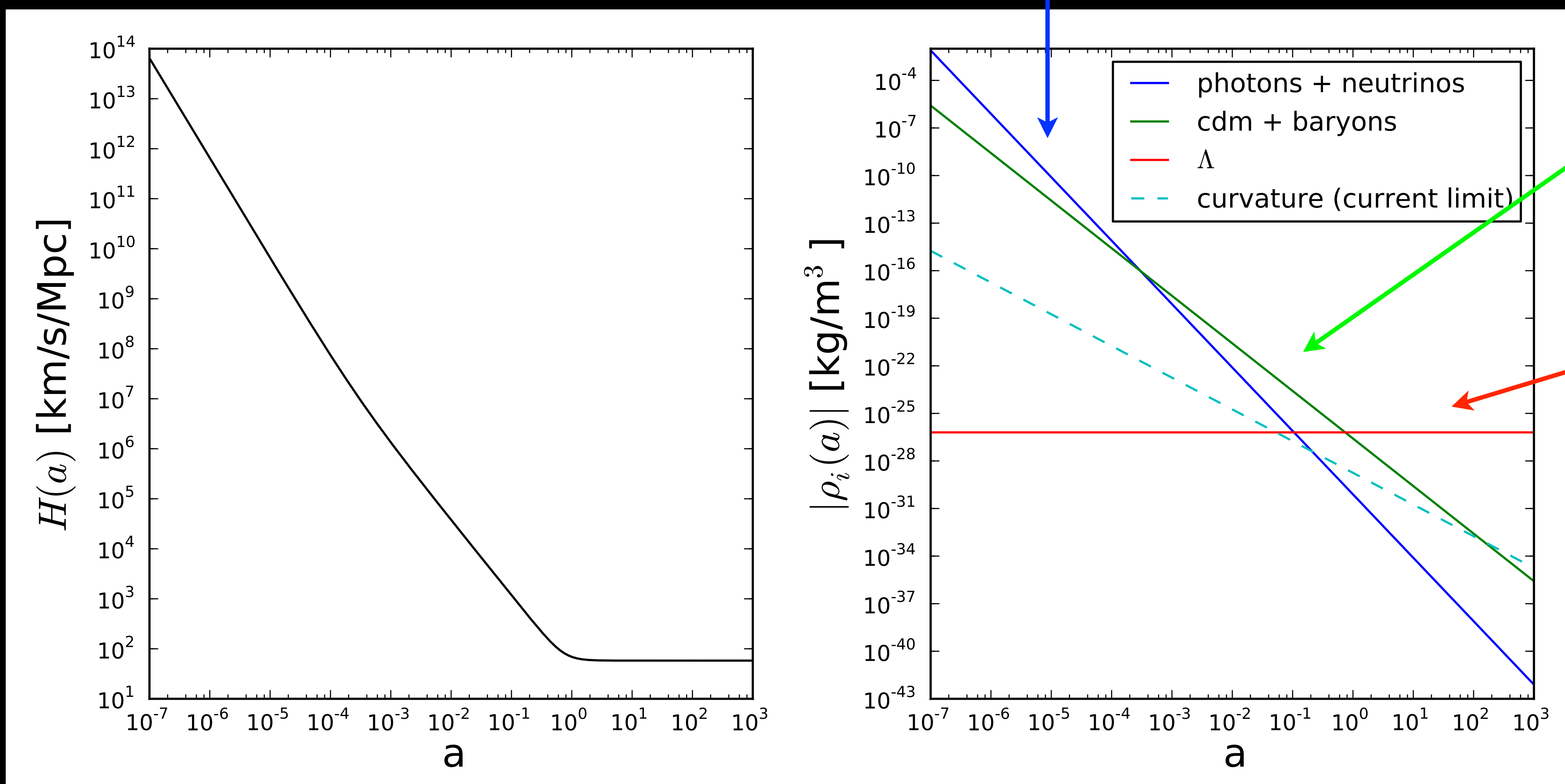
- Initial conditions: adiabatic, Gaussian, nearly scale invariant spectrum of perturbations [$P_i(k) = A_*(k/k_*)^{n_s}$]
- Relevant physics: gravity (GR) and relativistic pressure in an expanding background
- See Wayne Hu's tutorials (http://background.uchicago.edu/~whu/beginners/introduction.html) and lectures on CMB and large scale structure next week

LSS theory

radiation domination

matter domination

dark energy domination



[see IPython examples for plot source code]

LSS theory

- In general relativity, the expansion history determines the growth of linear perturbations: $\sum \Omega_i = 1$. The physical energy densities are $\omega_i = \Omega_i h^2$ (so $\sum \omega_i = h^2$).

$$\left[\frac{1}{a}\,\frac{da}{dt}\right]^2 = H_0^2\left[\Omega_\gamma a^{-4} + \Omega_\nu(a) + \Omega_m a^{-3} + \Omega_K a^{-2} + \Omega_{DE}(a)\right]$$

$$\left[\frac{1}{a}\,\frac{da}{dt}\right]^2 = \left(\frac{1}{9.785\ \mathrm{Gyr}}\right)^2\left[\omega_\gamma a^{-4} + \omega_\nu(a) + \omega_m a^{-3} + \omega_K a^{-2} + \omega_{DE}(a)\right]$$

- Linear perturbations grow at different rates in the radiation, matter, and dark energy dominated epochs

LSS theory

$$P(k,z) = D^2(z)\, T^2(k)\, P_i(k)$$

Inflation physics

CMB-epoch physics: radiation domination and baryon acoustic oscillations

Matter/Dark energy dominated era growth function

LSS theory

$P(k,z) = D^2(z)\, T^2(k)\, P_i(k)$

Matter/Dark energy dominated era growth function obeys

$$\frac{d^2G}{d\ln a^2} + \left(2 + \frac{d\ln H}{d\ln a}\right)\frac{dG}{d\ln a} = \frac{3}{2}\Omega_{\mathrm{m}}(a)G$$

[I use D and G interchangeably]

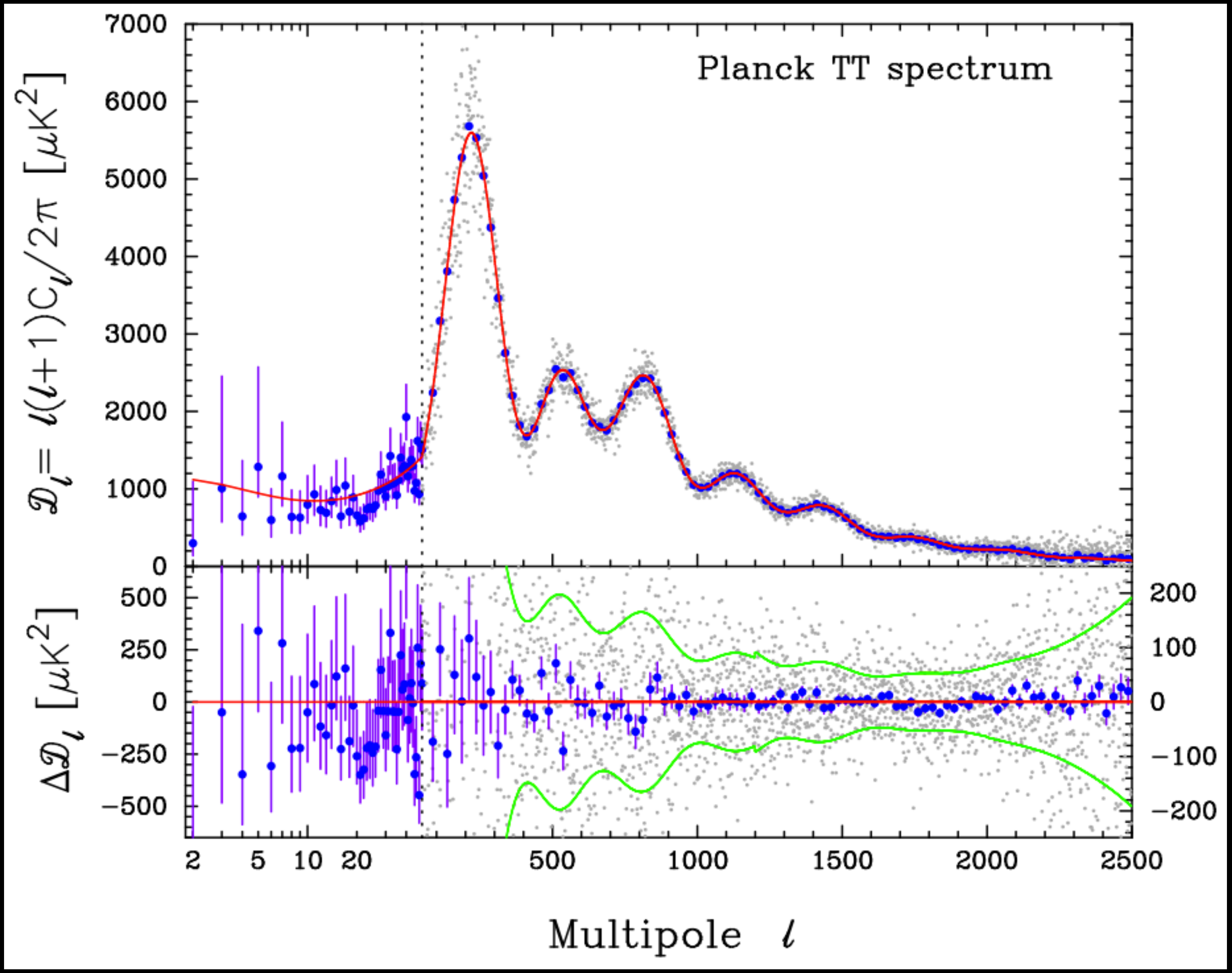
$\delta(z) = D(z)\, \delta(z_i)$

Analytic solution for w = -1 [see Sirko astro-ph/0503106 Eqn 3, Dodelson 2003 Eqn 7.77]

LSS theory

- The transfer function encapsulates impact of z >~1000 physics on the linear matter power spectrum: $P(k,z) = D^2(z)\, T^2(k)\, P_i(k)$
- Modes inside the horizon during radiation domination are damped
- Pressure breaks adiabaticity, generates baryon acoustic oscaillation (BAO) feature
- After decoupling, linear perturbation structure is “frozen”; all k-modes are amplified by gravity at the same rate [growth function, D(z)]
- How do we know all of this? Measurements of the cosmic microwave background

CMB provides LSS initial conditions

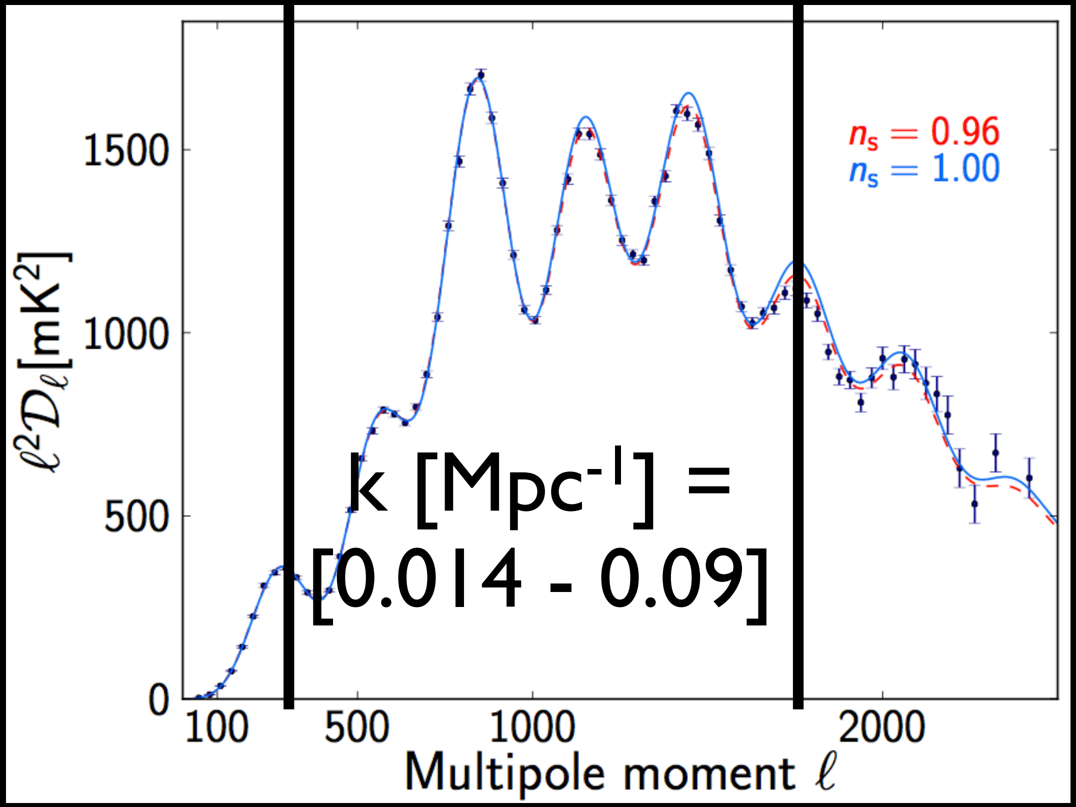


Initial Conditions from the CMB

- COBE/FIRAS measured the temperature of the CMB extremely precisely: $T\gamma = 2.7255 \pm 0.00057$ K.
- This measurement determines the energy density in both photons and (thermal relic) neutrinos. These particles were relativistic at z >~ 1000 [$\Omega_\gamma \propto a^{-4}$]
- Ratios of first to second (third) peak heights in the CMB determine $\Omega_b h^2$ ($\Omega_c h^2$)
- This simple picture works (so far); but allowing for new physics generates uncertainty in the prediction of P(k) from CMB (from e.g., Neff, r, isocurvature modes, m_ν, etc), and P(k) measurements can constrain those parameters

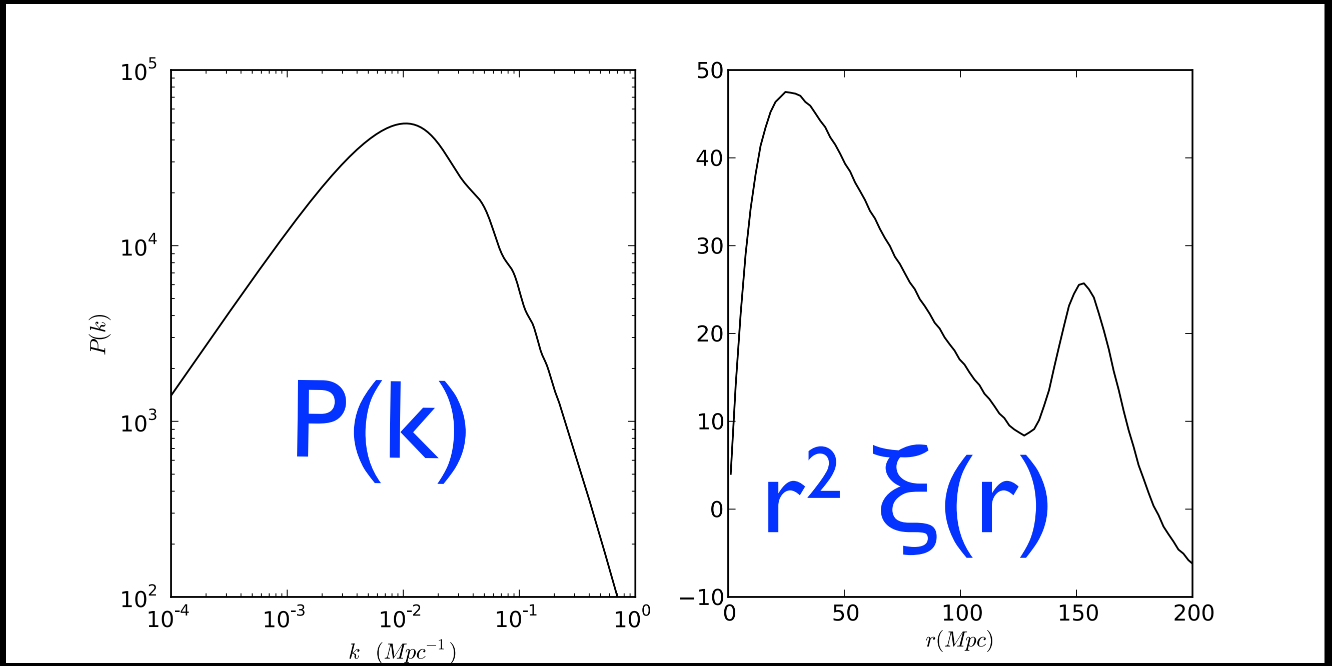
CMB constraints on initial conditions: n_s is measured on the same physical scales as LSS

photon-baryon fluid



Planck 2013 #16

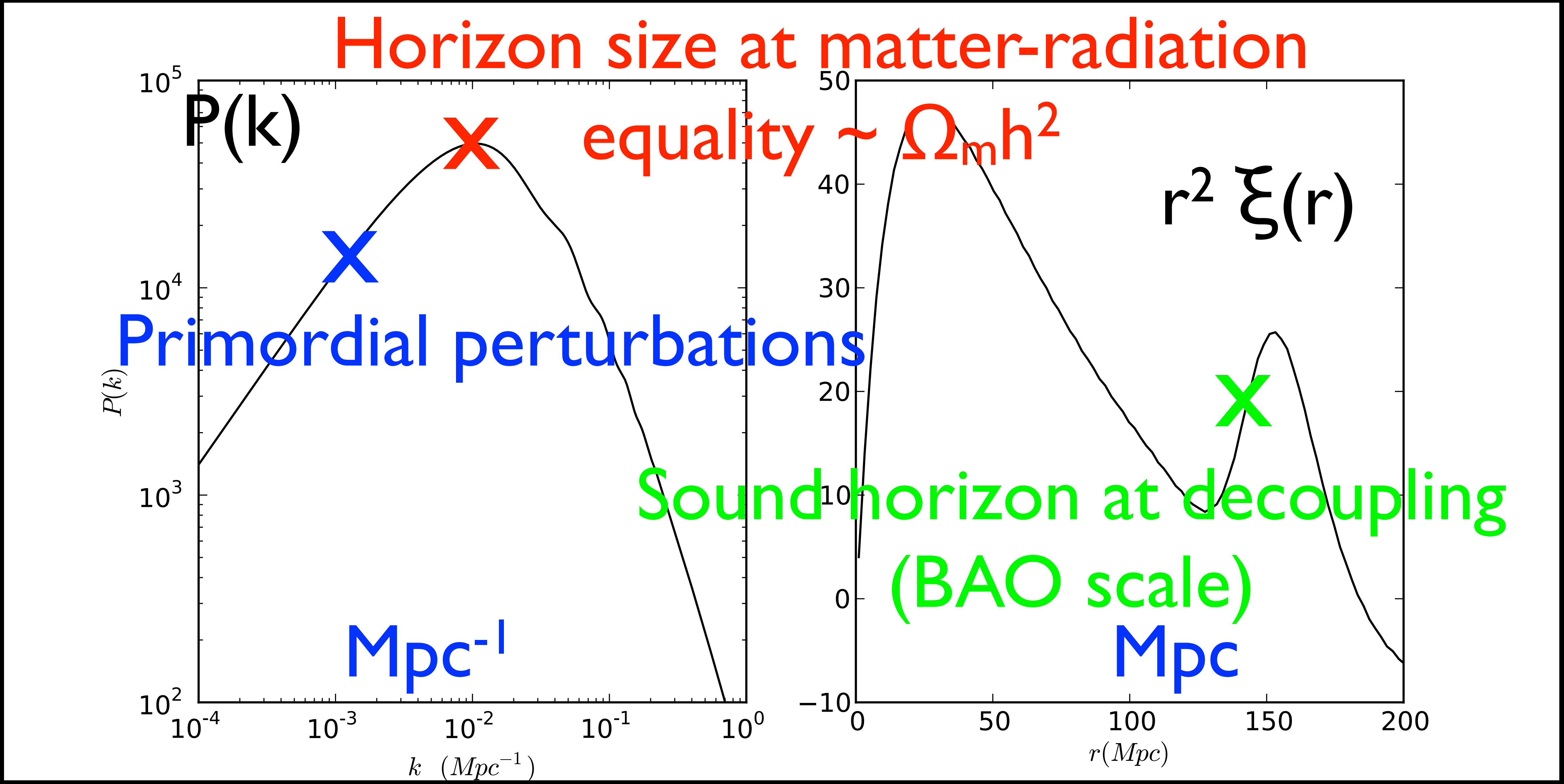
dark matter dominated



Mpc^{-1} Mpc

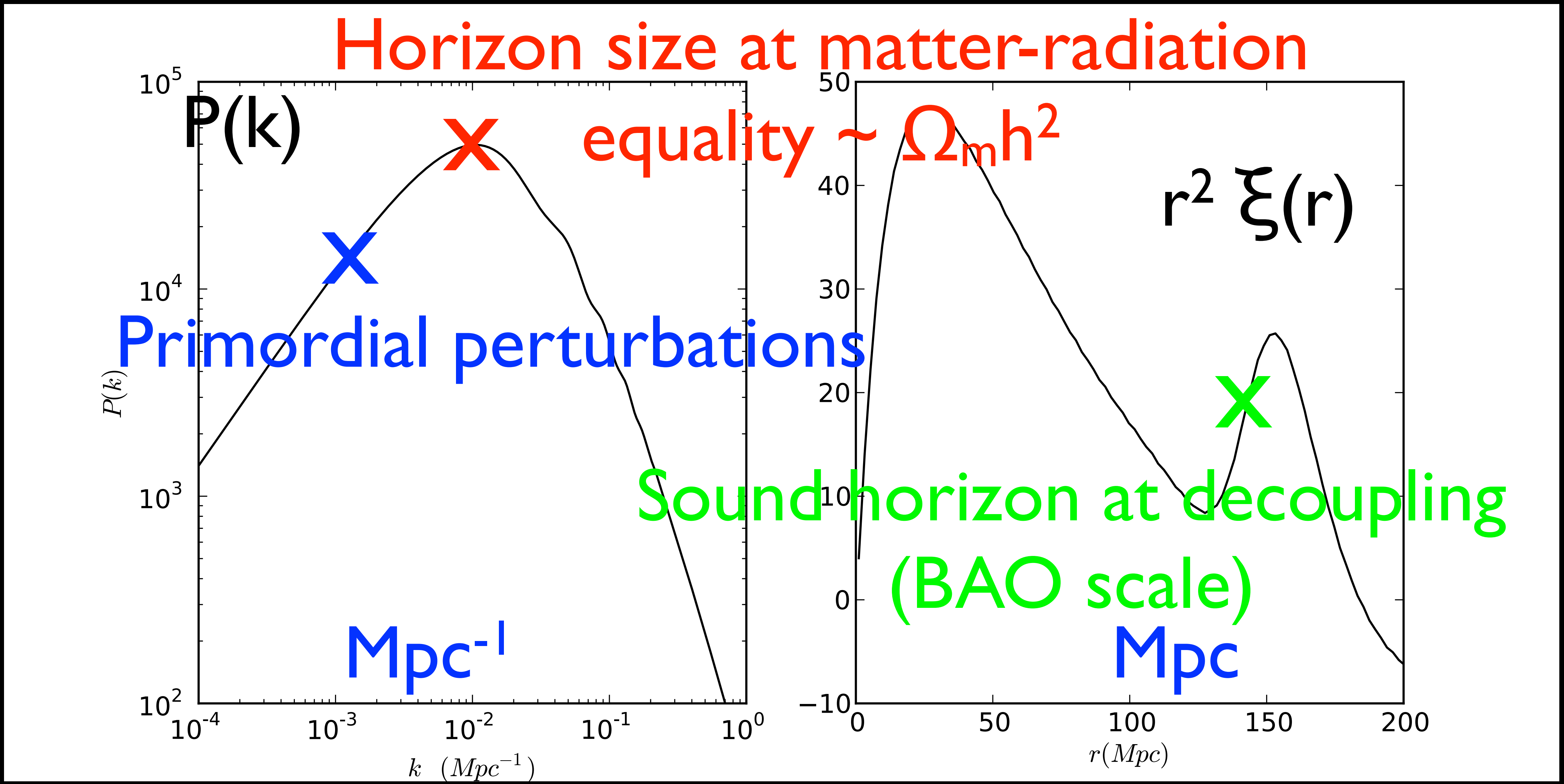
CMB constraints on initial conditions

- depends on $\Omega_{c,b,\gamma}\, h^2$ and n_s, but NOT $D_A(z_{CMB})$



CMB constraints on initial conditions

- Most important take-away: natural units are Mpc, not h^{-1} Mpc!!



CMB constraints on initial conditions

- Most important take-away: natural units are Mpc, not h^{-1} Mpc!!

Quiz:
Why does everyone always plot h^{-1} Mpc then?

CMB constraints on initial conditions

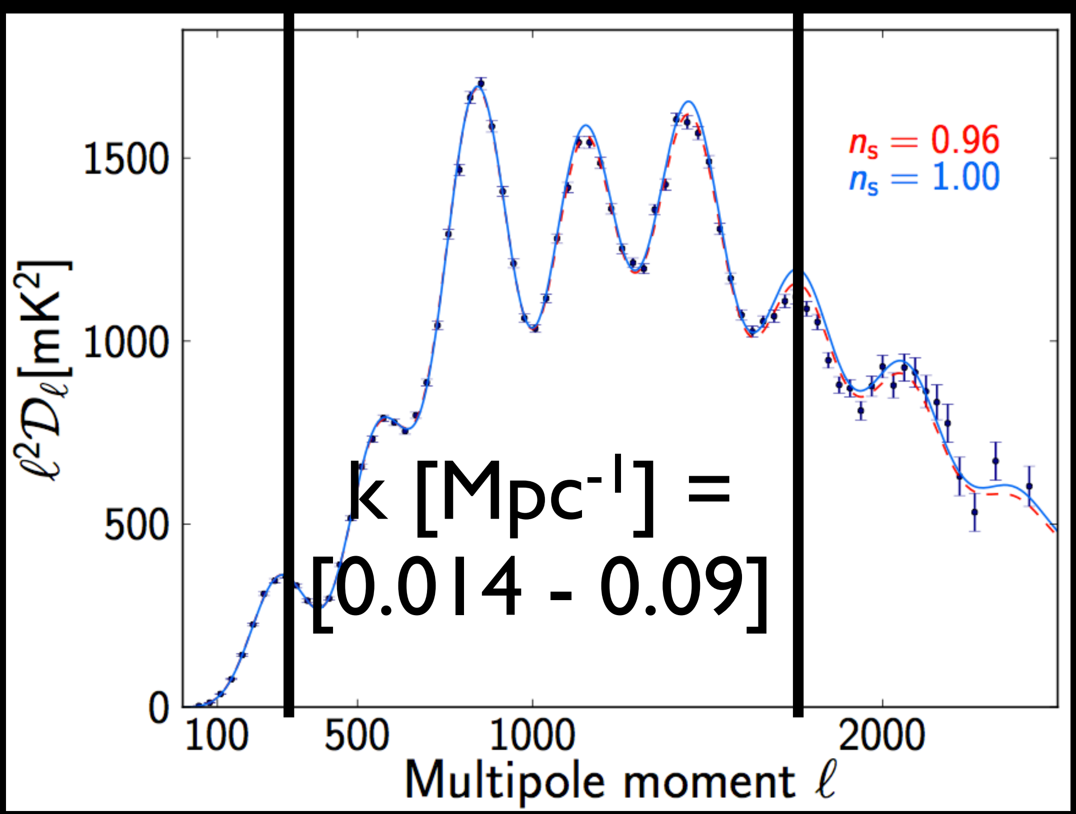
- Most important take-away: natural units are Mpc, not h^{-1} Mpc!!

Quiz:
Why does everyone always plot h^{-1} Mpc then?

Answer:
comoving distance $\chi = \int_0^z c\, dz/H(z)$
$\approx 3000z\ h^{-1}$ Mpc at $z \approx 0$
At higher redshifts, cannot neglect $H(z)$ dependence on Ω_m, Ω_{DE}, etc.

Comparison of Planck prediction with SDSS-II LRG halo power spectrum

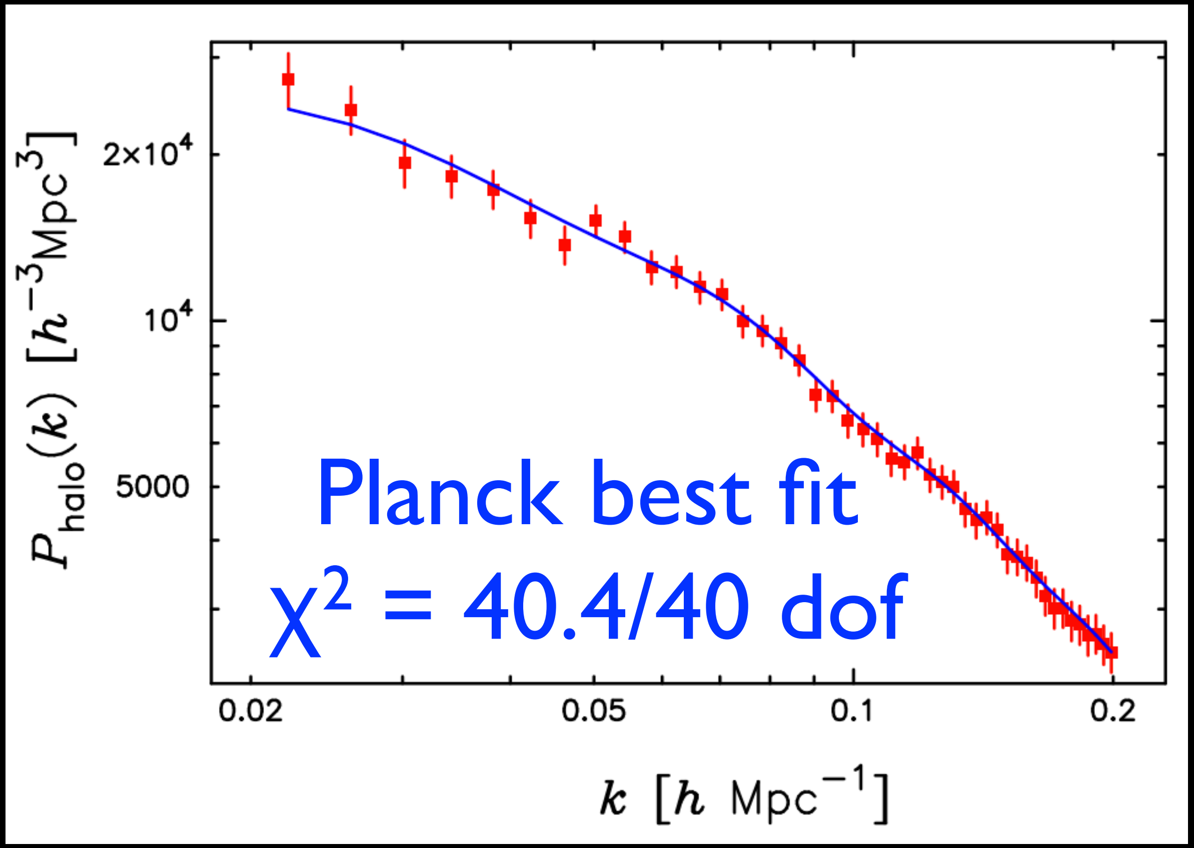
photon-baryon fluid



Planck 2013 #16

$\longrightarrow$

DM halo P(k) from SDSS-II LRGs
Reid et al 2010



Statistical Interlude

http://nbviewer.ipython.org/github/bareid/Trieste/blob/master/TriesteReid1.ipynb

Statistics Summary

- Galaxy redshift surveys are three-dimensional maps of the distribution of galaxies
- The simplest model for the observed galaxy density field is $\delta_g = b_g \delta_m + \varepsilon$
- δ_m is a Gaussian random field which the galaxy distribution samples as a Poisson process, so $\langle \varepsilon(k)\, \varepsilon(-k) \rangle = \bar{n}^{-1}$
- All cosmology information contained in $\langle \delta_m(k)\, \delta_m(-k) \rangle = P(k)$; next lecture we will begin extracting it!