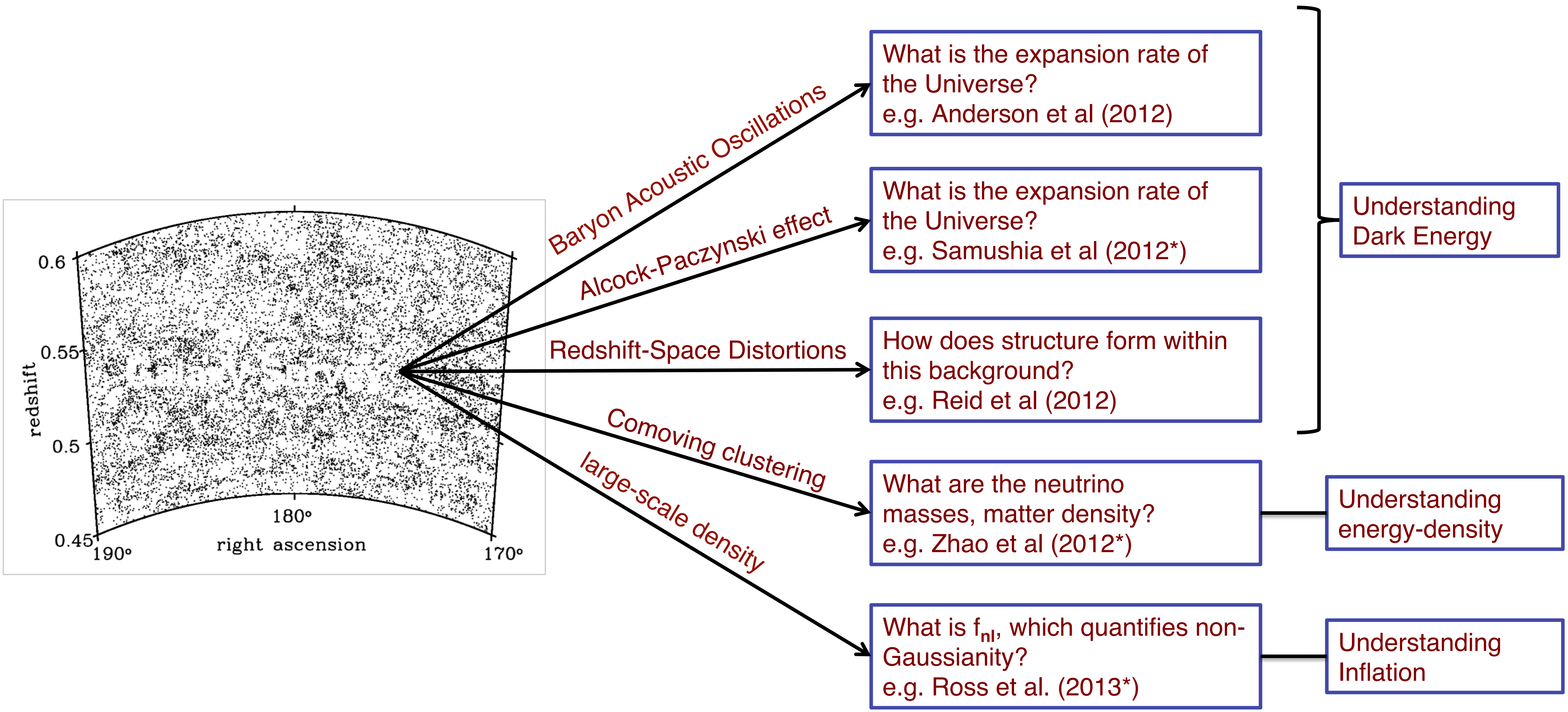
Cosmological Measurements from Galaxy Clustering

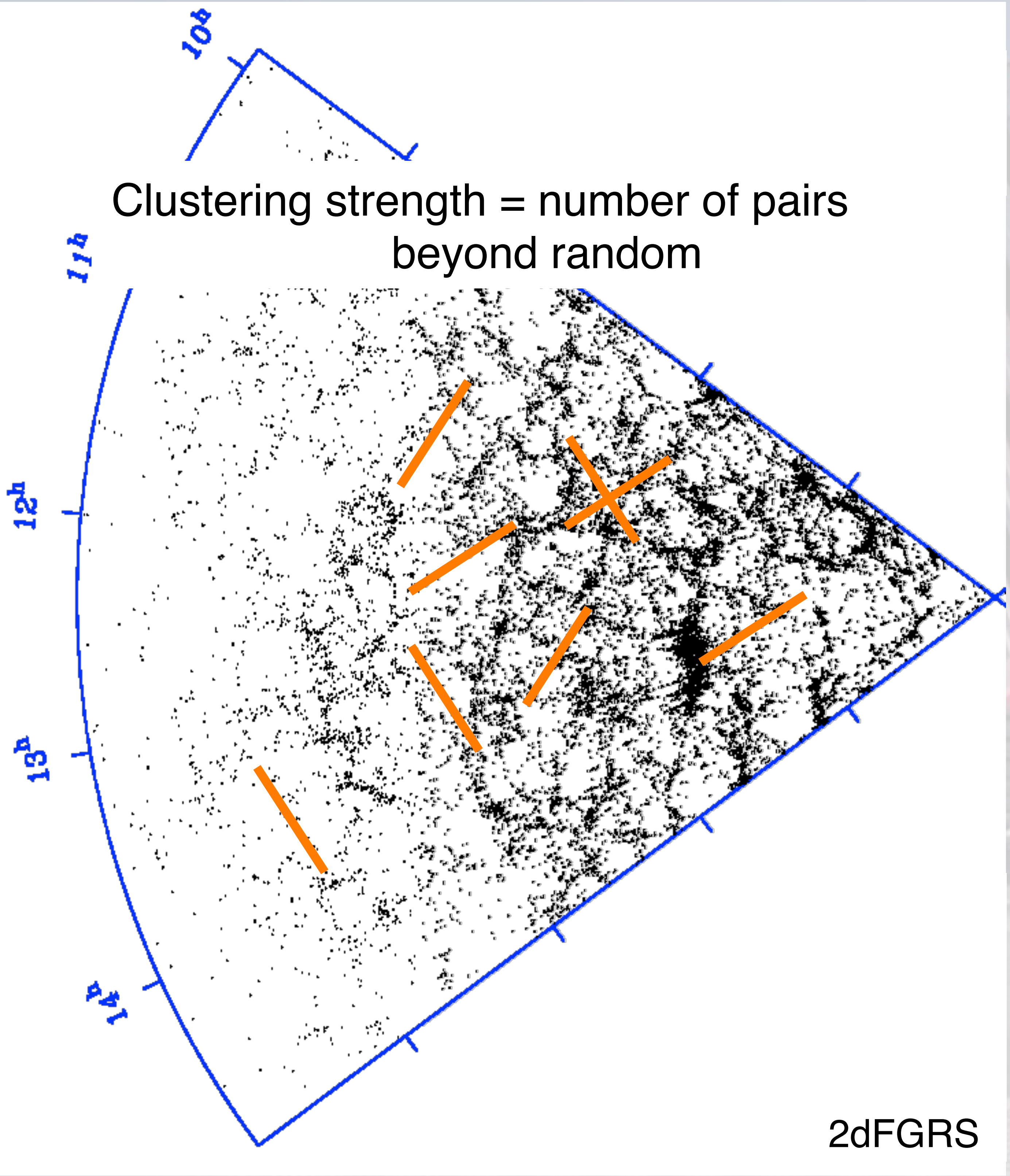
Will Percival

+ Hector Gil-Marin, Ashley Ross, Davide Bianchi, Beth Reid
+ BOSS, eBOSS, & Euclid GC SWGs

Cosmology from spectroscopic galaxy surveys



2-point clustering



"probability of seeing structure", can be cast in terms of the overdensity

$$\delta = \frac{\rho - \rho_0}{\rho_0}$$

The correlation function is simply the real-space 2-point statistic of the field

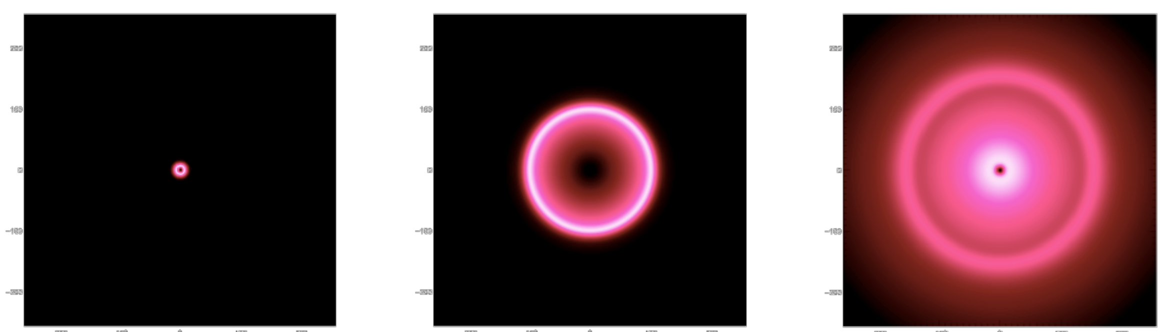
$$\xi(r) = \langle \delta(\mathbf{x}) \delta(\mathbf{x} + \mathbf{r}) \rangle$$

Its Fourier analogue, the power spectrum is defined by

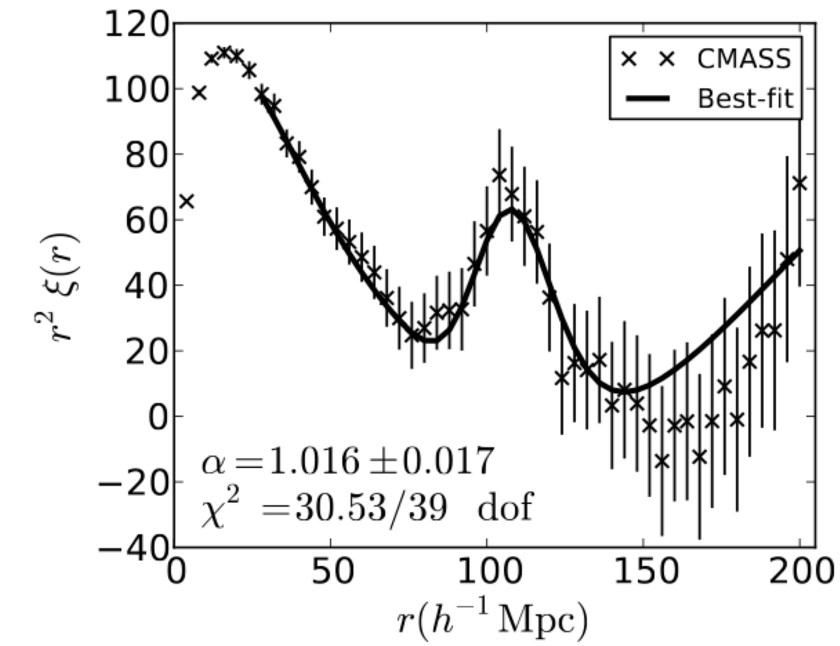
$$P(k) = \langle \delta(\mathbf{k}) \delta(\mathbf{k}) \rangle$$

By analogy, one should think of "throwing down" Fourier modes rather than "sticks"

Baryon Acoustic Oscillations (BAO)

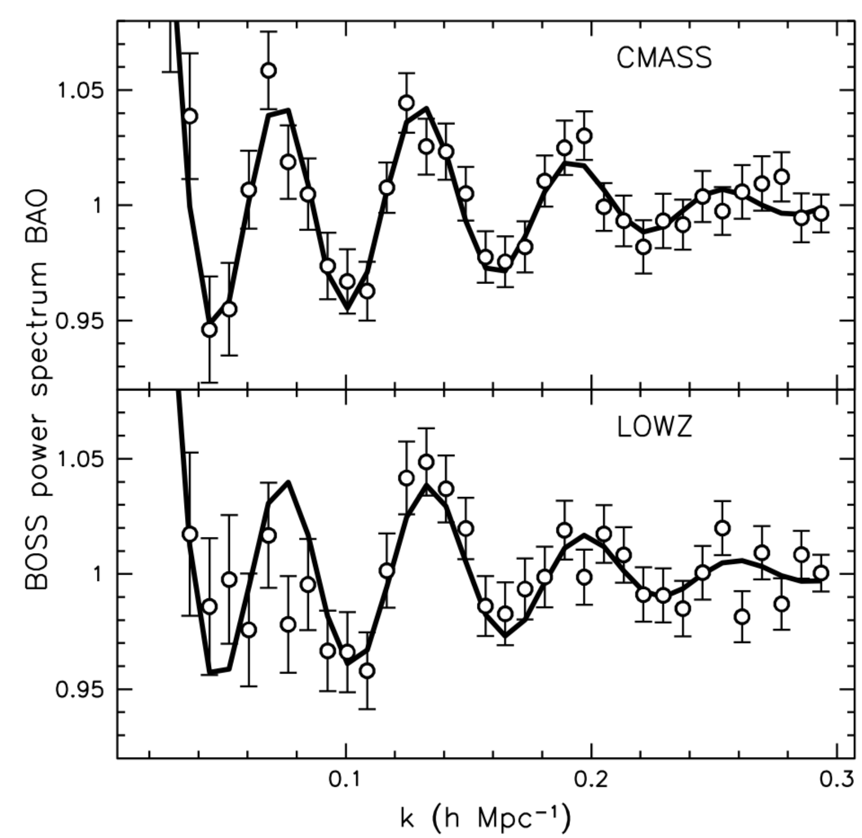


(images from Martin White)



To first approximation, comoving BAO wavelength is determined by the comoving sound horizon at recombination

$$r_s = \frac{1}{H_0 \Omega_m^{1/2}} \int_0^{a_*} da \frac{c_s}{(a + a_{eq})^{1/2}}$$



comoving sound horizon ~110h^{-1}Mpc, BAO wavelength 0.06hMpc^{-1}

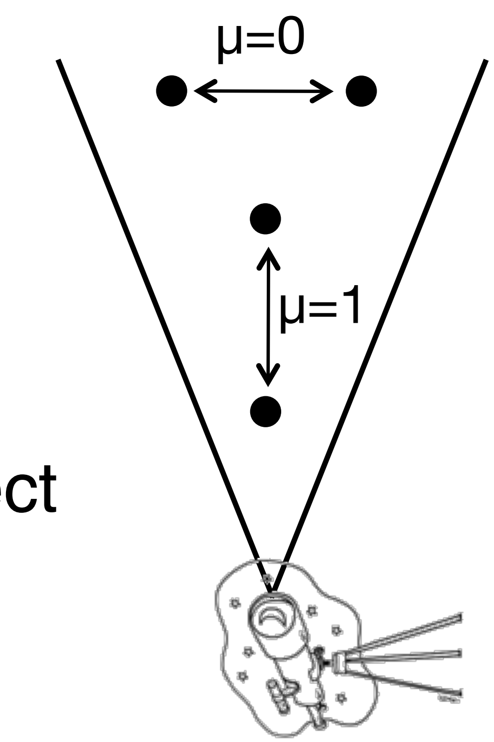
Anderson et al. 2012; arXiv:1203.6594

BAO as a standard ruler

Surveys measure angles and redshifts, and we have to use a fiducial model (denoted “fid”) to translate to comoving coordinates

$$\mu = cos(\alpha)$$

Changes in apparent BAO position ($\Delta d_{\mathbf{comov}}$) depend on:



Radial direction

$$\alpha_{\parallel} = \frac{H(z)_{\rm fid}}{H(z)_{\rm true}}$$

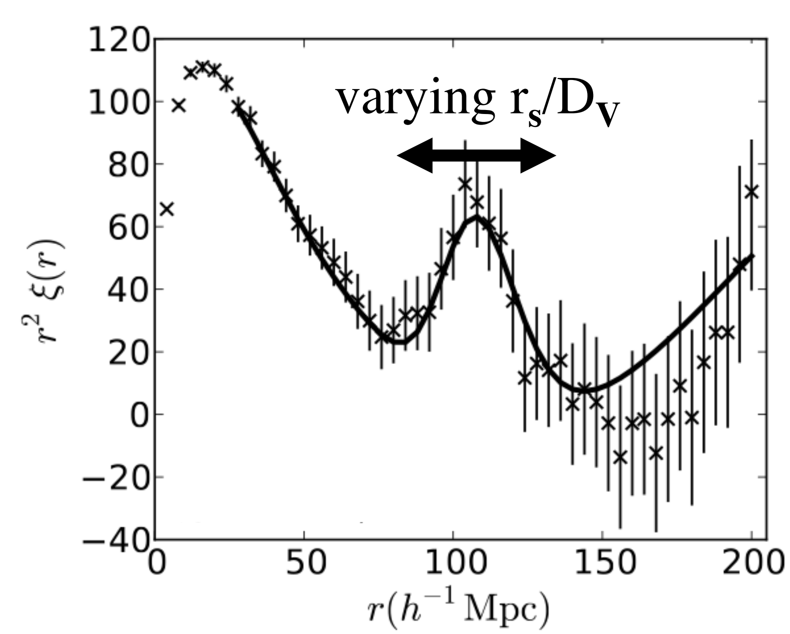
Angular direction

$$\alpha_{\perp} = \frac{D_A(z)_{\rm true}}{D_A(z)_{\rm fid}}$$

(i.e. these terms anisotropically stretch clustering - the relative effect known as Alcock-Paczynski Effect)

We see from geometrical arguments that a set of random pairs constrains

$$D_V = \left[(1+z)^2 D_A^2(z) \frac{cz}{H(z)}\right]^{1/3}$$



Anisotropic projection

BAO scale depends on angle

$$\alpha(\mu) = \sqrt{\mu^2\alpha_{||}^2 + (1-\mu^2)\alpha_\perp^2}$$

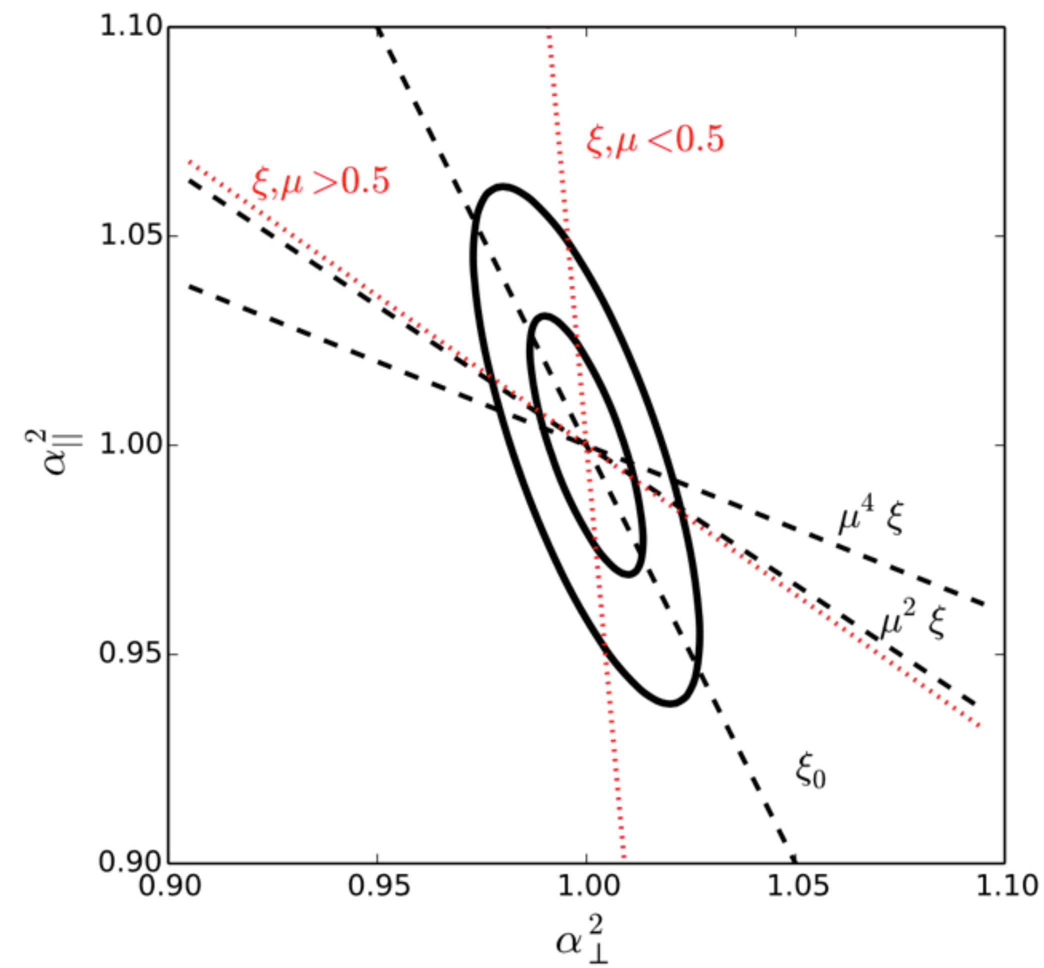
Define a (general) moment of the power spectrum

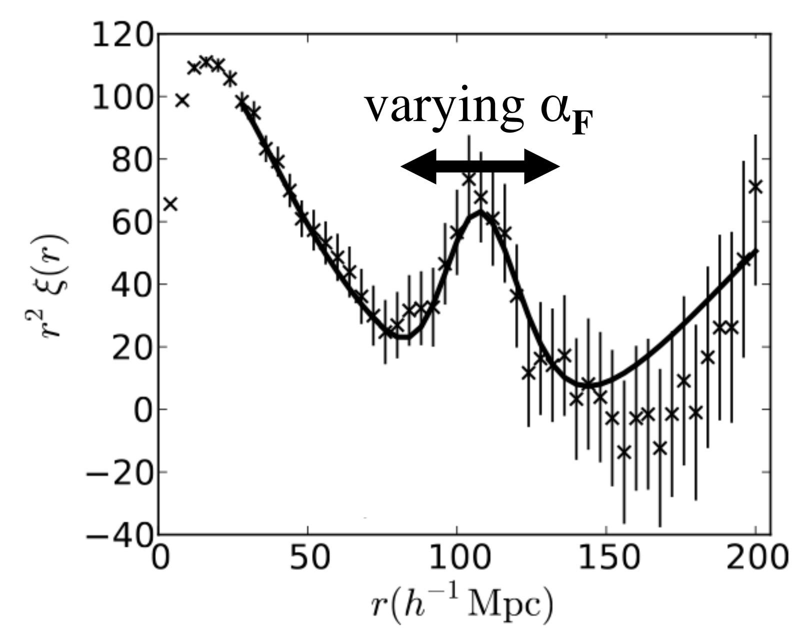
$$P_F(k) = \int_0^1 d\mu\, F(\mu) P(k,\mu)$$

Then the BAO scale measured in this moment depends on a combination given by

$$\alpha_F = \int_0^1 d\mu\, F(\mu) \left[\mu^2\alpha_{||}^2 + (1-\mu^2)\alpha_\perp^2\right]^{\frac{1}{2}}$$

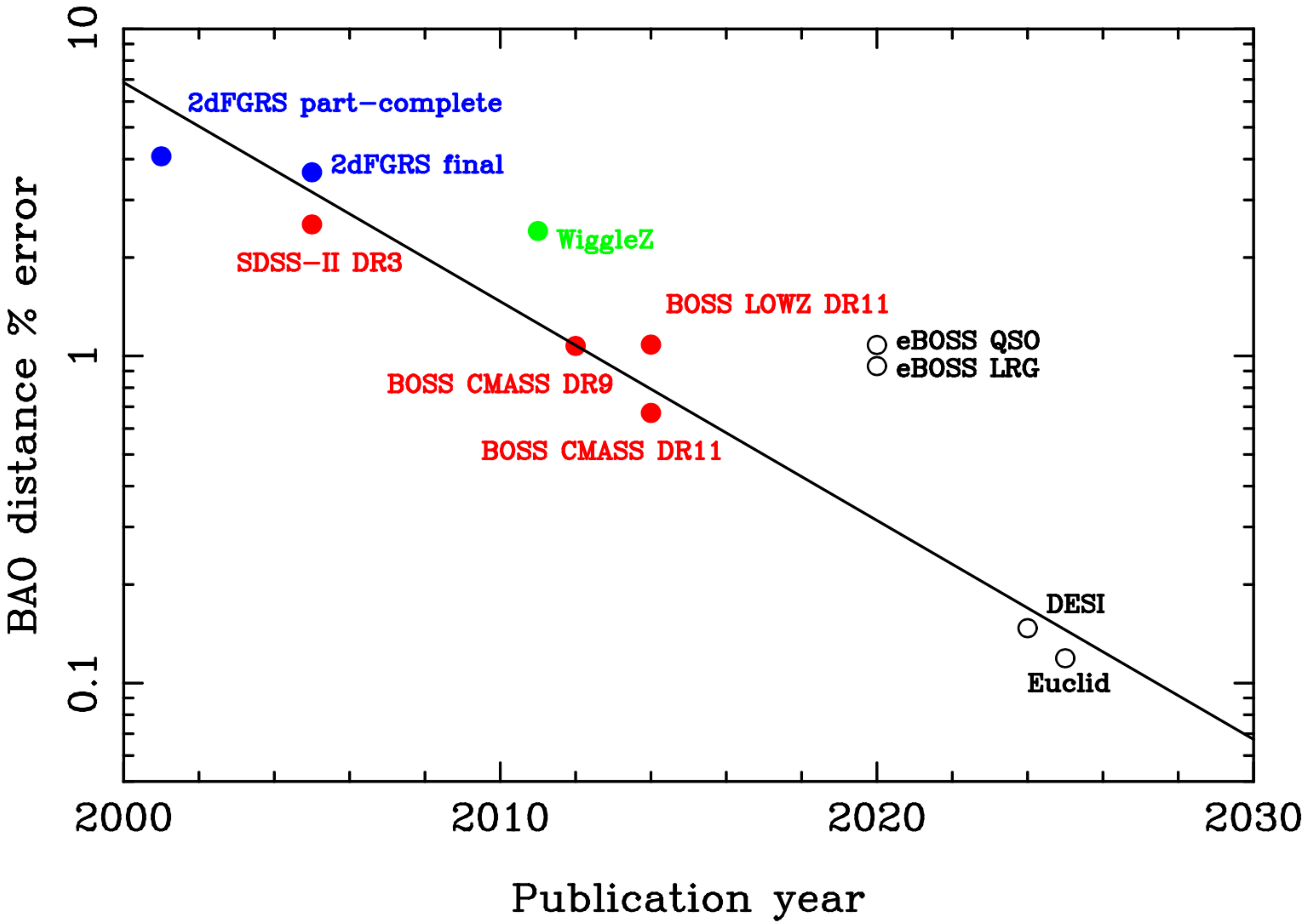
Given linear dependence, get same information from monopole and quadrupole $F(\mu)=1, 3\mu^2-1$, as $F(\mu)=1,\mu^2$





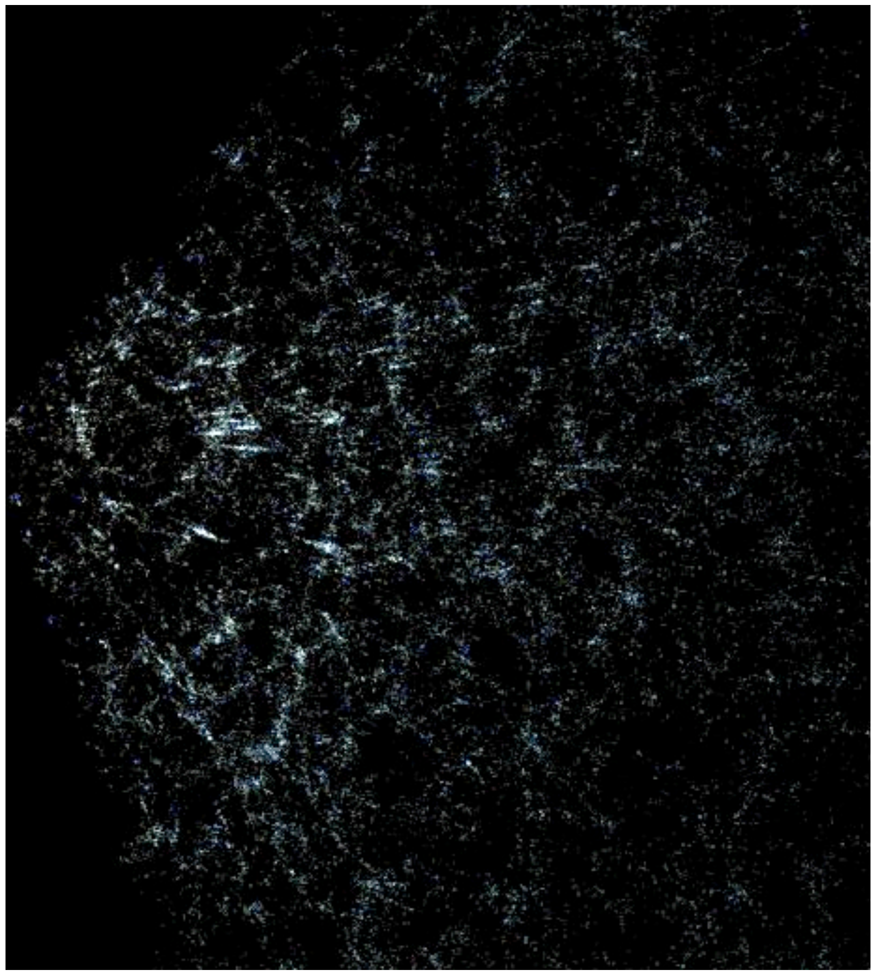
Ross et al. (2015); arXiv:1501.05571

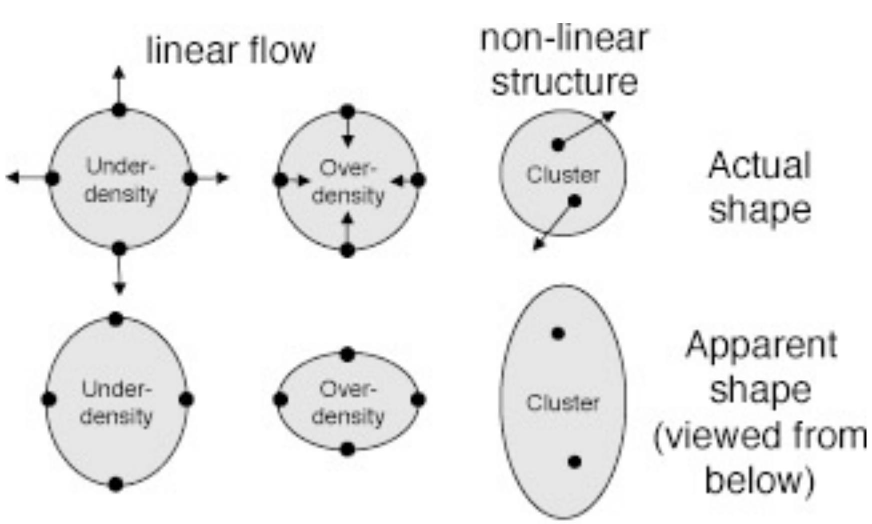
BAO errors from current / future galaxy surveys



Reid et al. 2015, arXiv:1509.06529

Redshift-Space Distortions





Observed redshift depends on both Hubble expansion and additional "peculiar velocity"

Galaxies move because cosmological structure is growing

Resulting change in redshift is coherent with structure

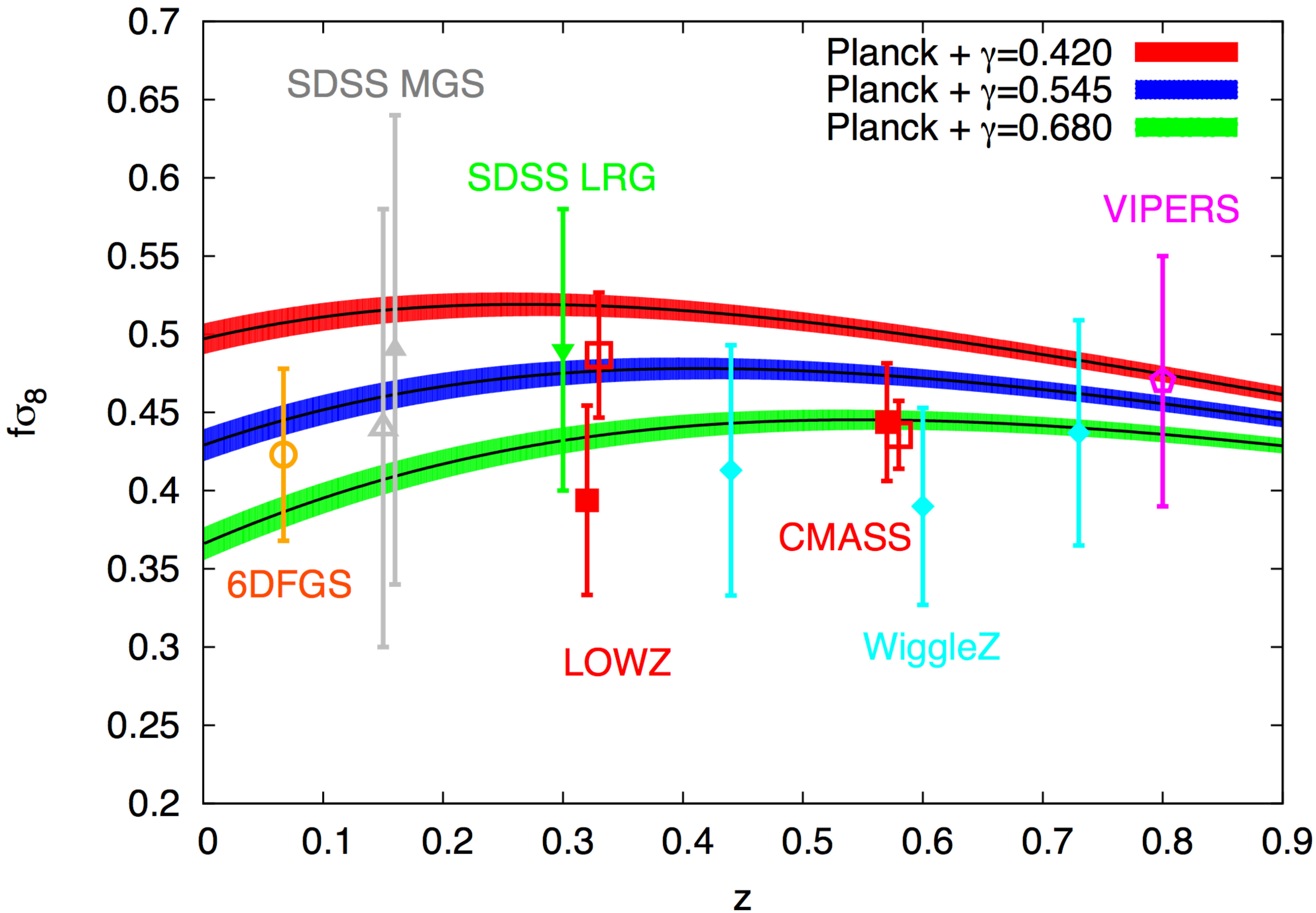
Extra component of 2-point clustering amplitude depends on peculiar velocities: additive term to δ

$$\mu^2 f(z)\sigma_8(z) \propto \mu^2 \frac{dG}{d\log a}$$

where G is the linear growth rate

Increases the correlation function quadrupole relative to monopole

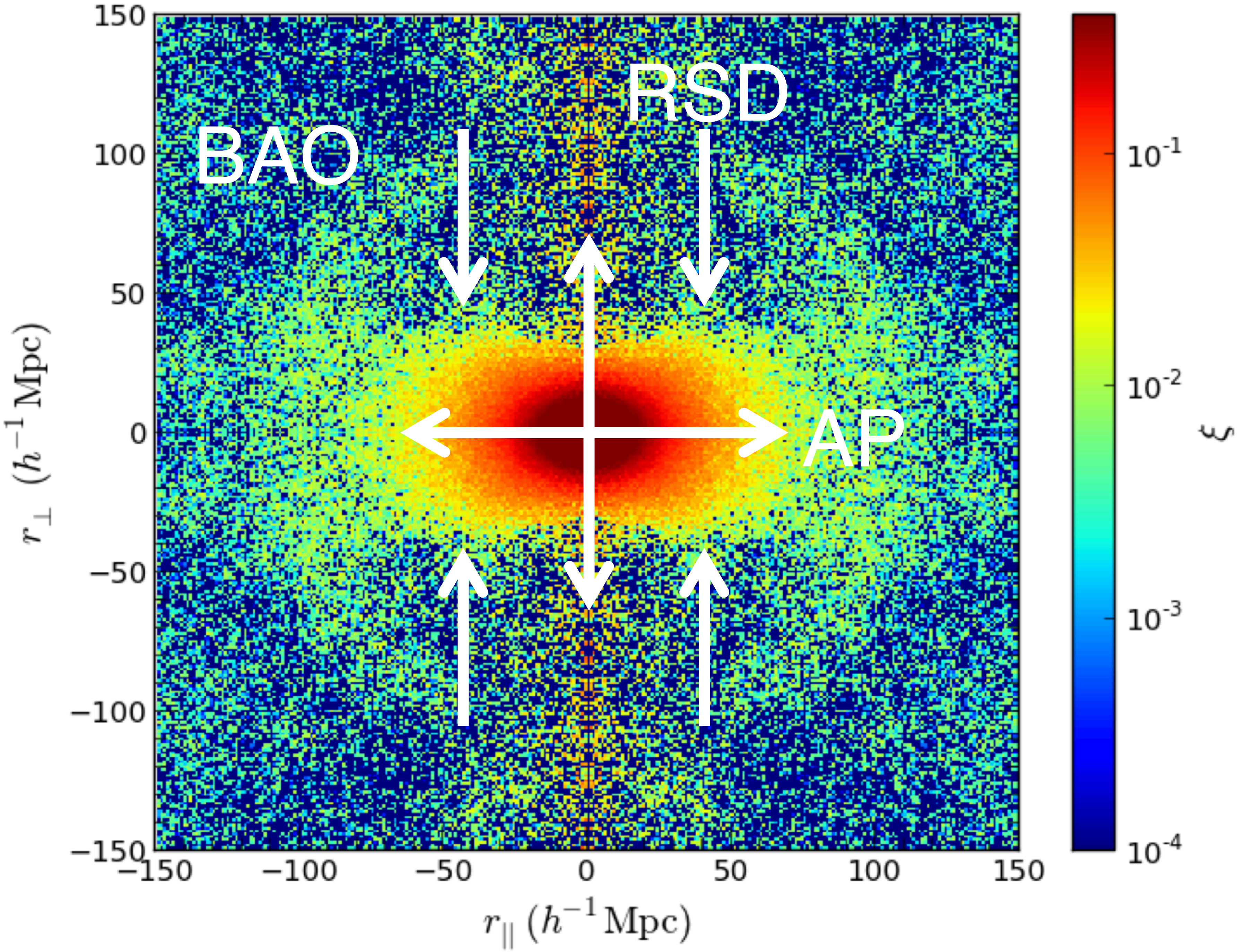
Testing GR through RSD



Testing a phenomenological model with $f = \Omega_m^{\gamma}$ (GR has $\gamma \sim 0.55$)

Gil-Marin et al. 2015, arXiv:1509.06386

The full clustering signal



BOSS, Samushia et al. 2013; MNRAS, 439, 3504

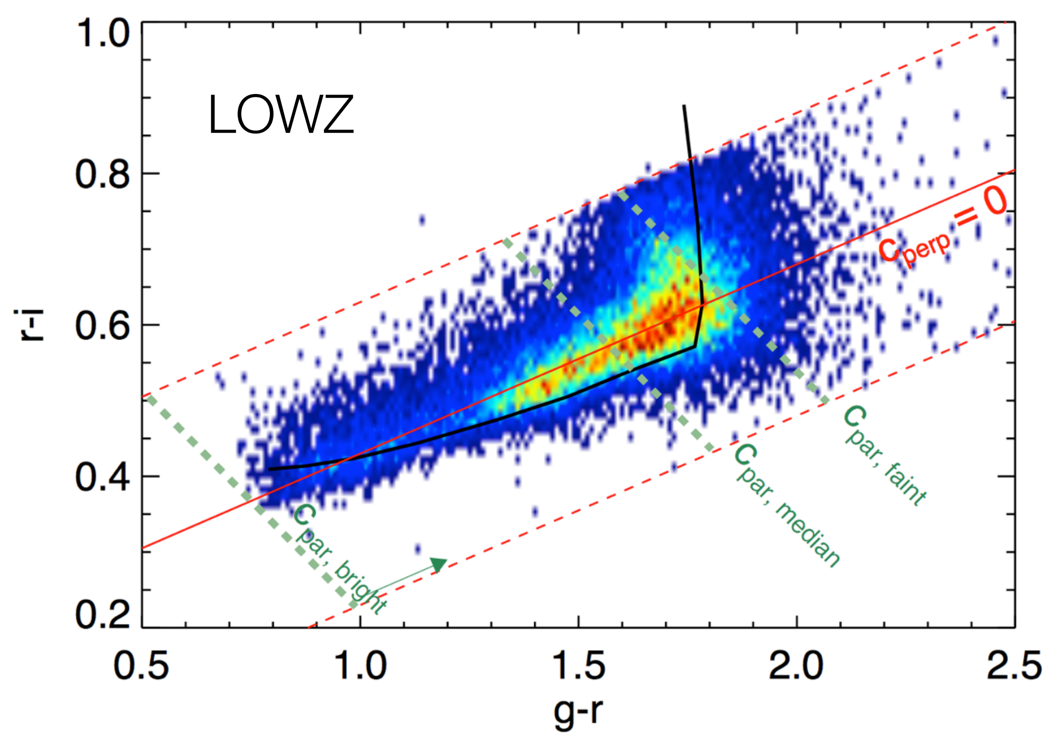
Baryon Oscillation Spectroscopic Survey

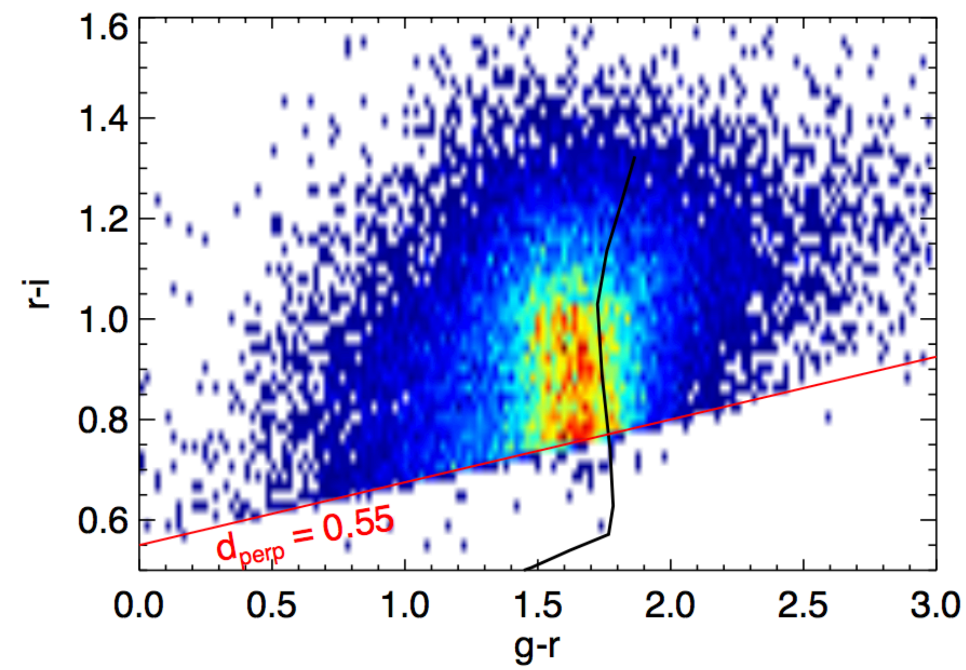
- Duration: Fall 2009 - Summer 2014, dark time
- Telescope: 2.5m Sloan
- Upgrade to SDSS-II spectrograph
  - 1000 smaller fibers
  - higher throughput
- Spectra:
  - 3600Å < λ < 10, 000Å New spectrograph
  - R = λ/Δλ = 1300 – 3000
  - (S/N) at mag. limit
    - 22 per pix. (averaged over 7000-8500Å)
    - 10 per pix. (averaged over 4000-5500Å)
- Area: 10,000 deg^2
- Targets:
  - **1.5 × 10^6 massive galaxies, z < 0.7, i < 19.9**
  - 1.5×10^5 quasars, z>2.2, g<22.0
  - 75,000 ancillary science targets, many categories

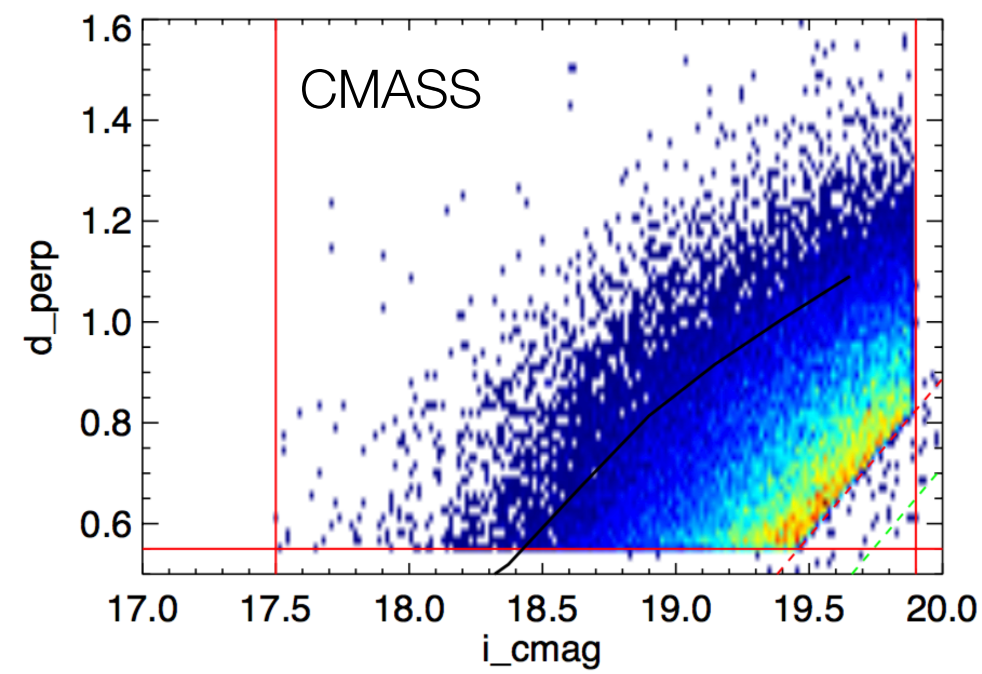
Dawson et al. 2012, AJ, 145, 10

BOSS Data Release 12 galaxies

- Two galaxy samples targeted: LOWZ and CMASS
- Colour cuts to select old, massive galaxies for easy redshift measurement and high bias
- Based on locus of passive galaxies
- CMASS broader (in colour) than LOWZ with a cut
  $d_{\perp} = (r_{\mathbf{mod}} - i_{\mathbf{mod}}) - (g_{\mathbf{mod}} - r_{\mathbf{mod}})/8 > 0.55$
  to select to an approximate stellar mass limit





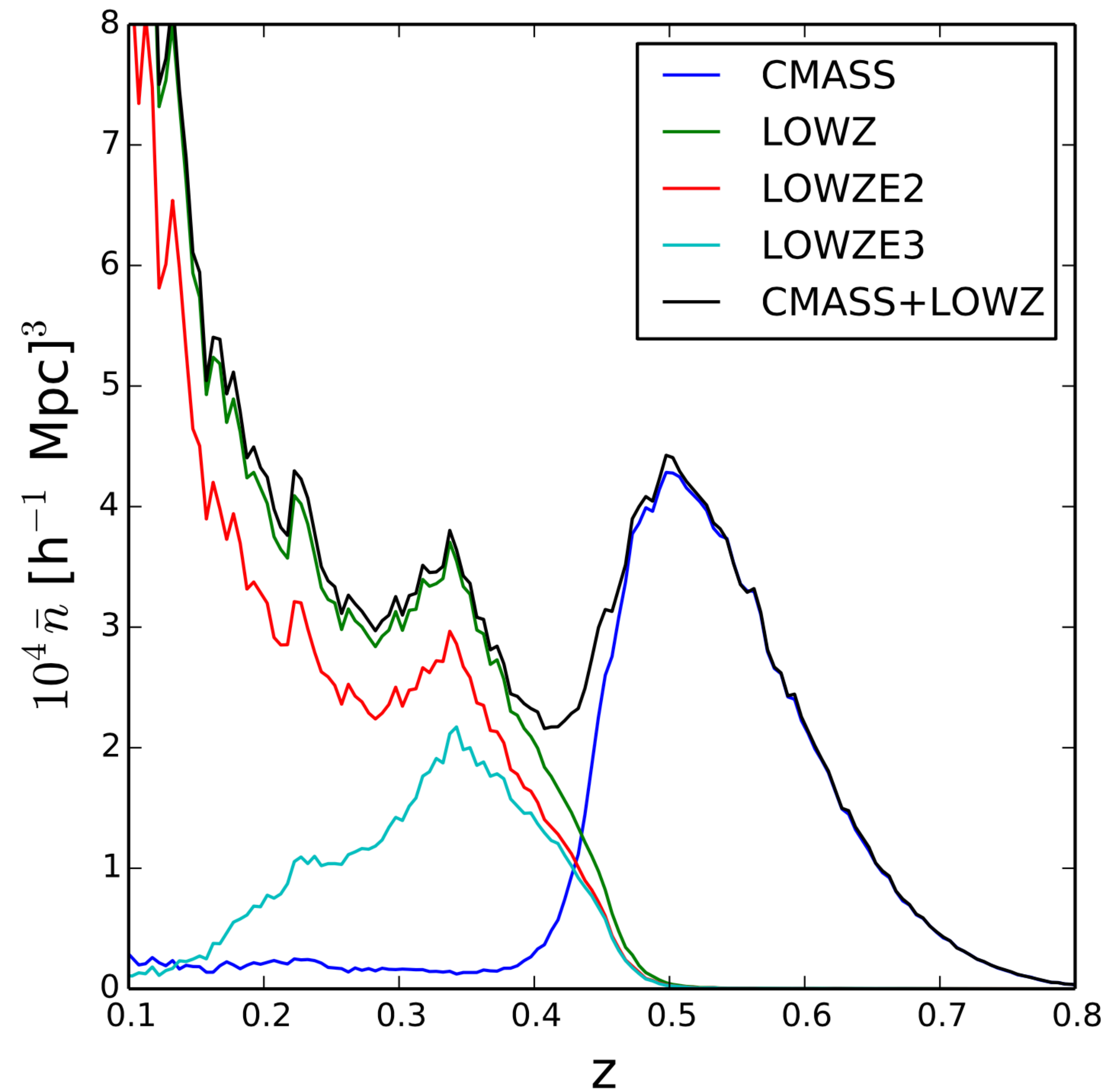


Reid et al. 2015, arXiv:1509.06529

BOSS Data Release 12 galaxies

- Some problems with early LOWZ observations, where an incorrect star-gal separation was used (LOWZE2 and LOWZE3)

- Catalogues are created for both samples, quantifying the mask with a Monte-Carlo sampling correcting for:
  - angular and radial distribution of targets
  - unobserved, previously known redshifts
  - observed stars
  - spectra that didn't result in a redshift (including angular and radial distribution)
  - target galaxies that were not observed
  - galaxies not observed due to fiber collisions



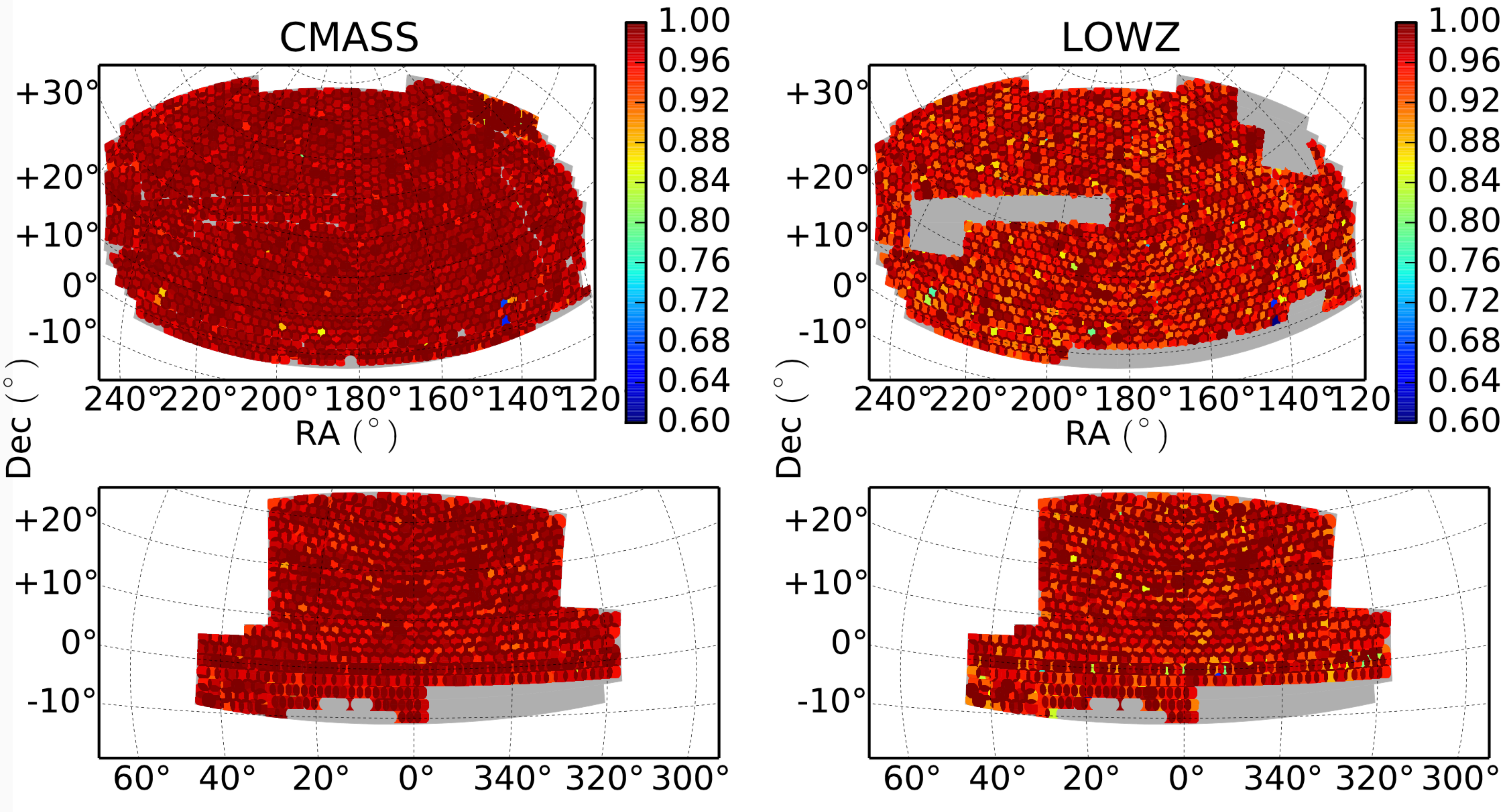
Reid et al. 2015, arXiv:1509.06529

BOSS Data Release 12 galaxies

Property	NGC	SGC	total	NGC	SGC	total	NGC	NGC
Sample		CMASS			LOWZ		LOWZE2	LOWZE3
$\bar{N}_{\rm gal}$	607,357	228,990	836,347	177,336	132,191	309,527	2,985	11,195
$\bar{N}_{\rm known}$	11,449	1,841	13,290	140,444	13,073	153,517	2,730	6,371
$\bar{N}_{\rm star}$	14,556	8,262	22,818	1,043	976	2,019	24	61
$\bar{N}_{\rm fail}$	10,188	5,157	15,345	868	602	1,470	21	55
$\bar{N}_{\rm cp}$	34,151	11,163	45,314	4,459	4,422	8,881	16	167
$\bar{N}_{\rm missed}$	7,997	3,488	11,485	10,295	3,499	13,794	114	609
$\bar{N}_{\rm used}$	568,776	208,426	777,202	248,237	113,525	361,762	4,336	15,380
$\bar{N}_{\rm obs}$	632,101	242,409	874,510	179,247	133,769	313,016	3,030	11,311
$\bar{N}_{\rm targ}$	685,698	258,901	944,599	334,445	154,763	489,208	5,890	18,458
Total area (deg^2)	7,429	2,823	10,252	6,451	2,823	9,274	144	834
Veto area (deg^2)	495	263	759	431	264	695	10	55
Used area (deg^2)	6,934	2,560	9,493	6,020	2,559	8,579	134	779
Effective area (deg^2)	6,851	2,525	9,376	5,836	2,501	8,337	131	755
Targets / deg^2	98.9	101.1	99.5	55.6	60.5	57.0	43.4	23.5

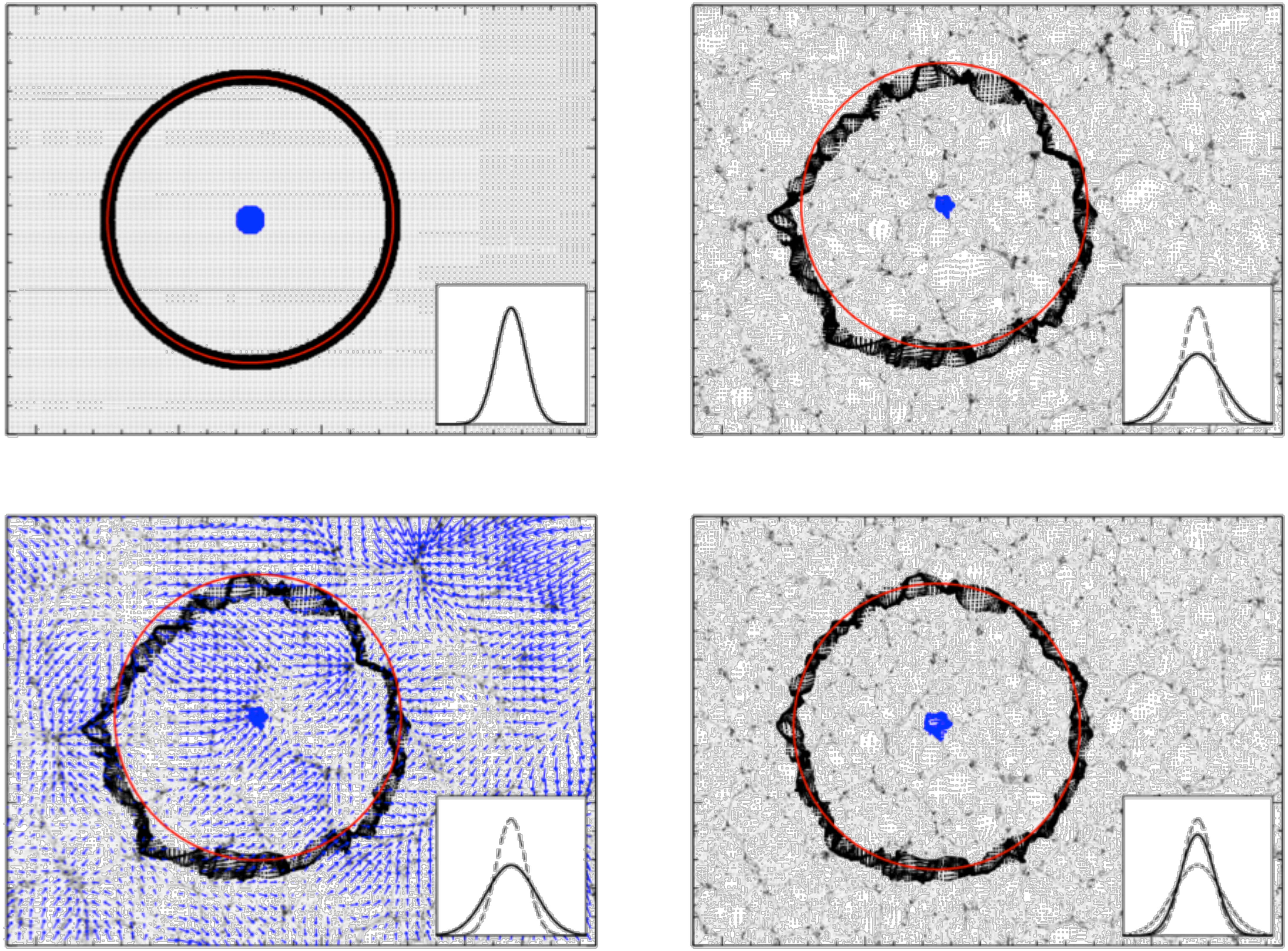
Reid et al. 2015, arXiv:1509.06529

BOSS Data Release 12 galaxies



Reid et al. 2015, arXiv:1509.06529

Reconstruction of “linear” signal



Padmanabhan et al. 2012; arXiv:1202.0090

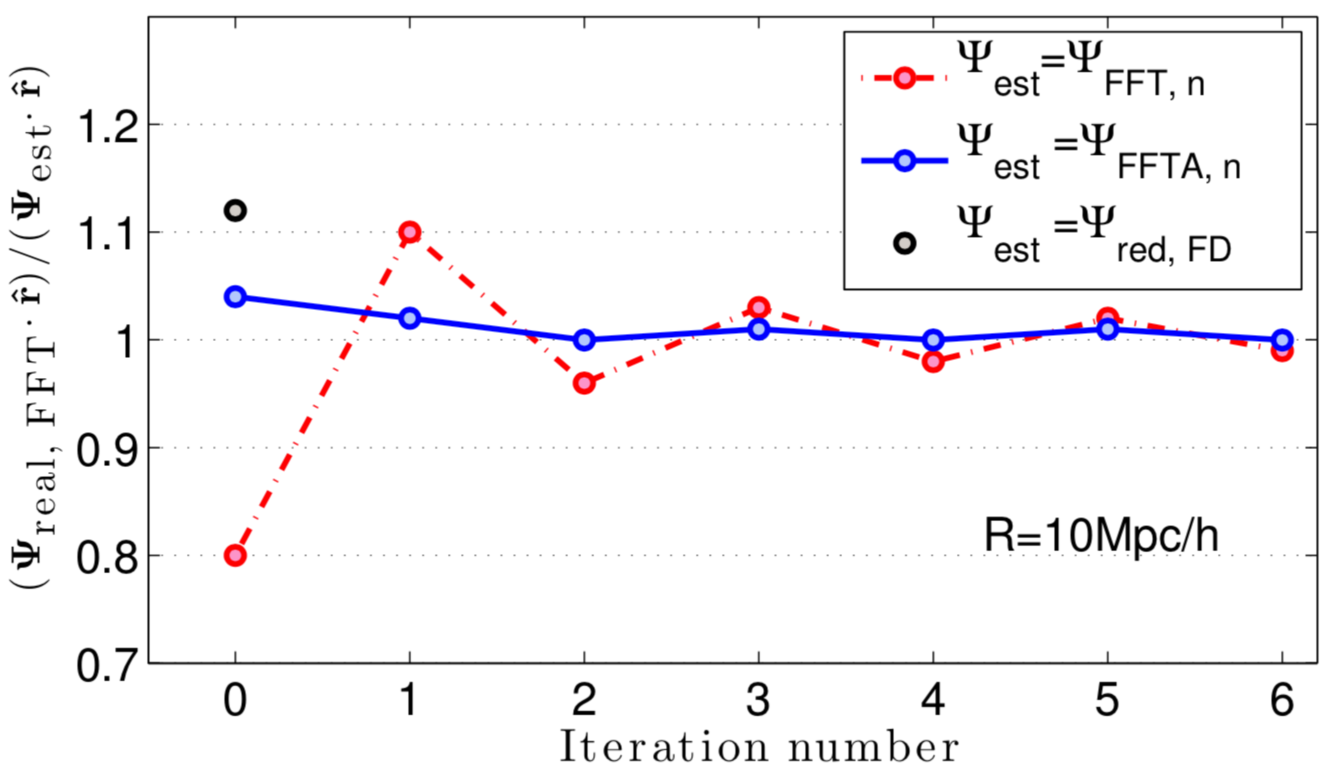
Reconstruction of “linear” signal

Problem for reconstruction is RSD and dealing with varying line-of-sight across a survey: displacements **Ψ** are (linear theory) relates to overdensities by

$$\nabla \cdot \boldsymbol{\Psi} + \frac{f}{b}\nabla \cdot (\boldsymbol{\Psi} \cdot \hat{\mathbf{r}})\hat{\mathbf{r}} = \frac{-\delta}{b}$$

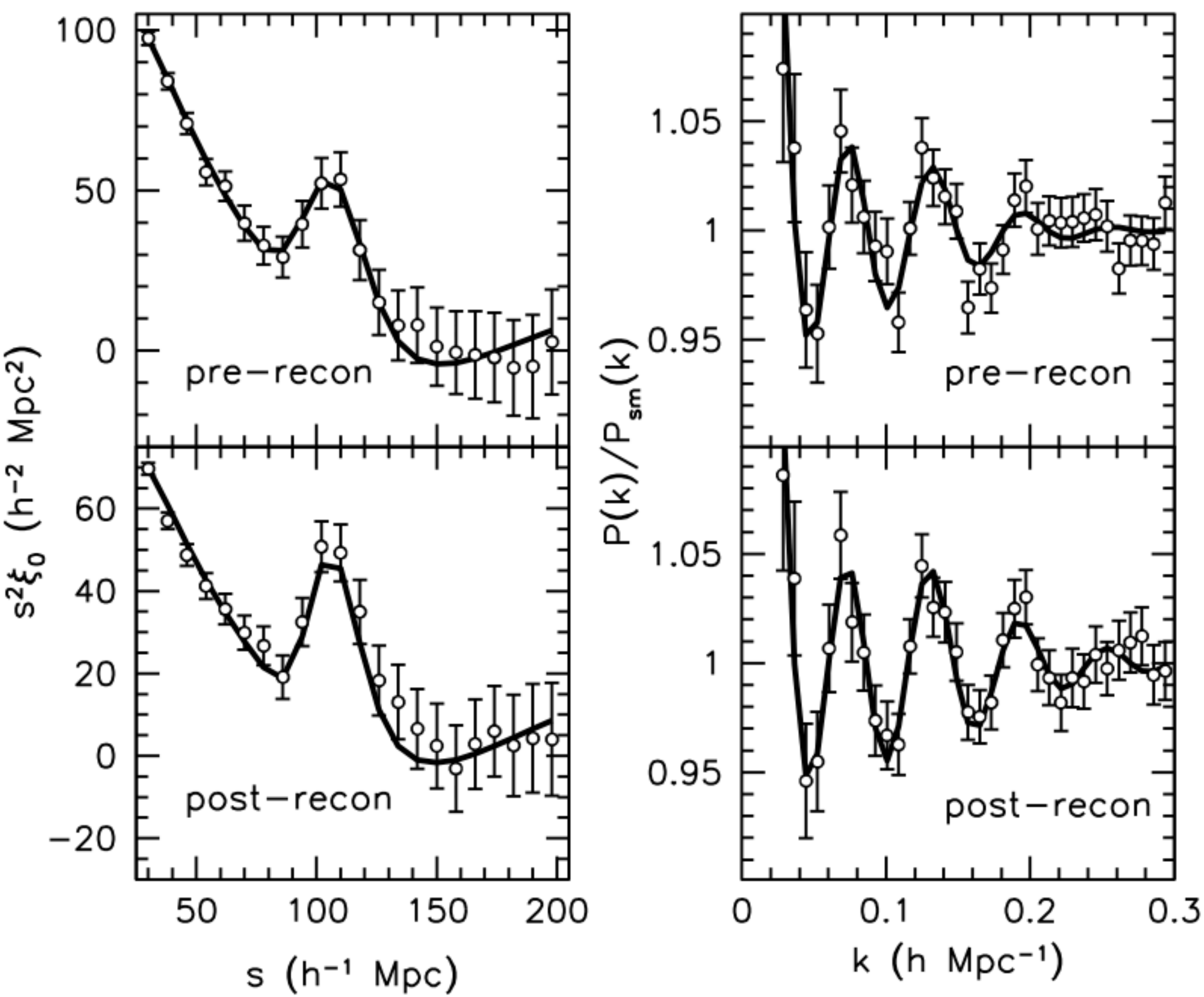
The RSD term limits fast calculation of the expected displacements as it is not irrotational, and depends on a varying line-of-sight

Introduce a new iterative method, allowing use of FFTs



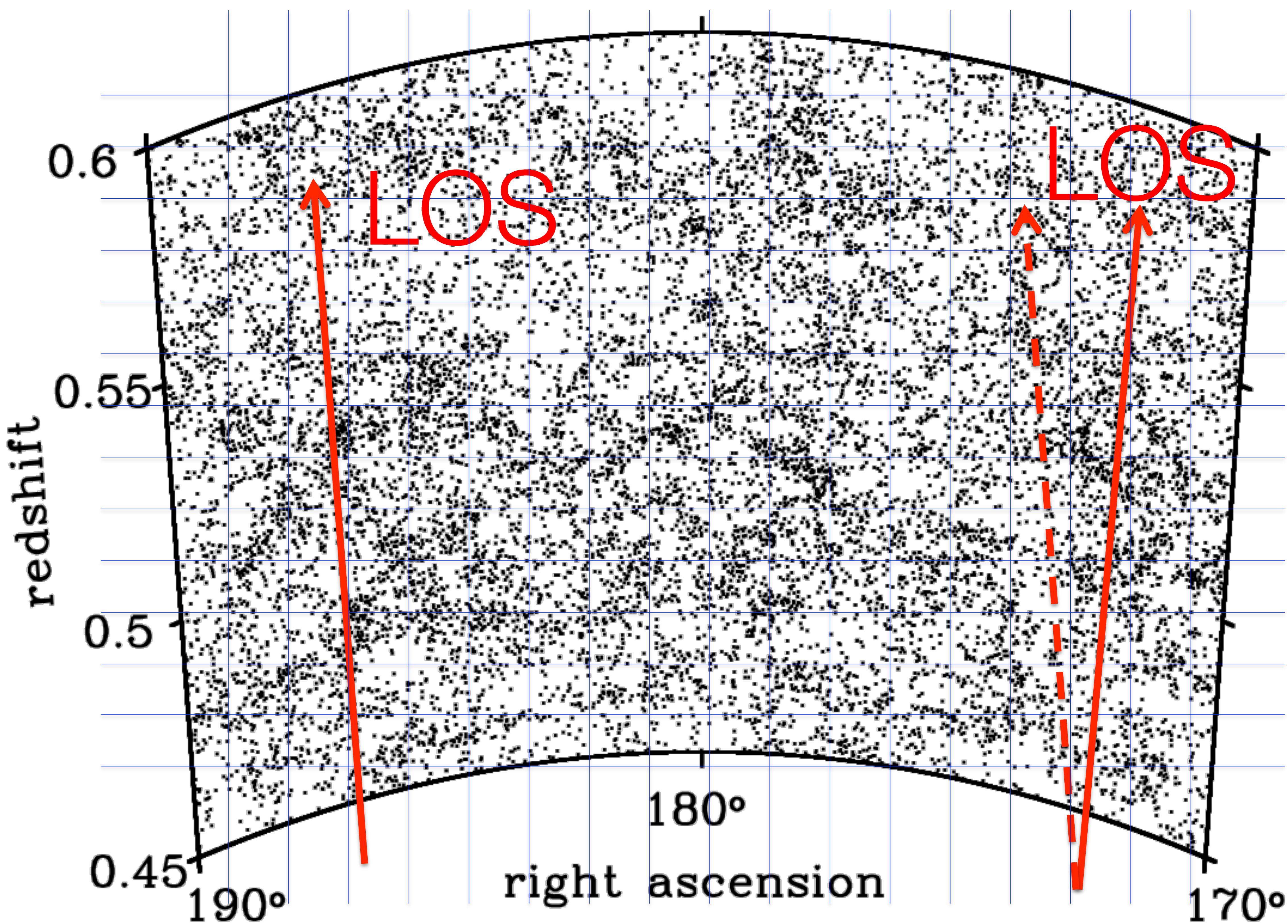
Burden et al. 2014; arXiv:1408.1348, Burden et al. 2015; arXiv: 1504.02591

The improvement from reconstruction



Anderson et al. 2013; arXiv:1312.4877

Measuring anisotropic clustering in Fourier space



Measuring power as a sum over pairs

- Define the overdensity field

$$N(\mathbf{r}) = \frac{w(\mathbf{r})}{\left[\int d\mathbf{r}\, w^2 \bar{n}^2(\mathbf{r})\right]^{1/2}} [n(\mathbf{r}) - \alpha n_s(\mathbf{r})]$$

- Power spectrum moments can be written as a integral over pairs

$$\hat{P}_F(k) = \frac{(2\ell+1)}{I_2} \int \frac{d\Omega_k}{4\pi} \left[\int d\mathbf{r}_1 \int d\mathbf{r}_2\, N(\mathbf{r}_1) N(\mathbf{r}_2) \times e^{i\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} F(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}_{\rm pair}) - P_\ell^{\rm noise} \right] ,$$

- The clever part is defining the LOS to the pair as LOS to one galaxy

$$\hat{P}_F(k) = \frac{(2\ell+1)}{I_2} \int \frac{d\Omega_k}{4\pi} \left[\int d\mathbf{r}_1\, N(\mathbf{r}_1) e^{i\mathbf{k}\cdot\mathbf{r}_1} \times \int d\mathbf{r}_2\, N(\mathbf{r}_2)\, e^{-i\mathbf{k}\cdot\mathbf{r}_2} F(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}_2) - P_\ell^{\rm noise} \right] ,$$

Yamamoto et al. 2005; astro-ph/0505115

Writing this in terms of FFTs

- For power-law $F(\mu)=\mu^n$, the "unit" to be solved is

$$A_n(\mathbf{k}) = \int d\mathbf{r}\, (\hat{\mathbf{k}} \cdot \hat{\mathbf{r}})^n N(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}$$

- We can expand the dot product on a Cartesian basis

$$\mu \equiv \hat{\mathbf{k}} \cdot \hat{\mathbf{r}} = \frac{k_x r_x + k_y r_y + k_z r_z}{kr}$$

- So that (for example) A_2 is decomposed (similarly for n>2)

$$A_2(\mathbf{k}) = \frac{1}{k^2} \left\{ k_x^2 B_{xx}(\mathbf{k}) + k_y^2 B_{yy}(\mathbf{k}) + k_z^2 B_{zz}(\mathbf{k}) + 2 \left[k_x k_y B_{xy}(\mathbf{k}) + k_x k_z B_{xz}(\mathbf{k}) + k_y k_z B_{yz}(\mathbf{k}) \right] \right\}$$

- Where B_{ij} can be solved with FFTs

$$B_{ij}(\mathbf{k}) \equiv \int d\mathbf{r}\, \frac{r_i r_j}{r^2} N(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}$$

Bianci et al. 2015; arXiv:1505.05341; Scoccimarro; arXiv:1506.02729

Measuring BAO for SDSS DR12

- Split into two samples: LOWZ and CMASS at z=0.43
- Perform reconstruction
- Measure monopole and μ^2 moment using FFTs
- Fit each (different free parameters) with a function

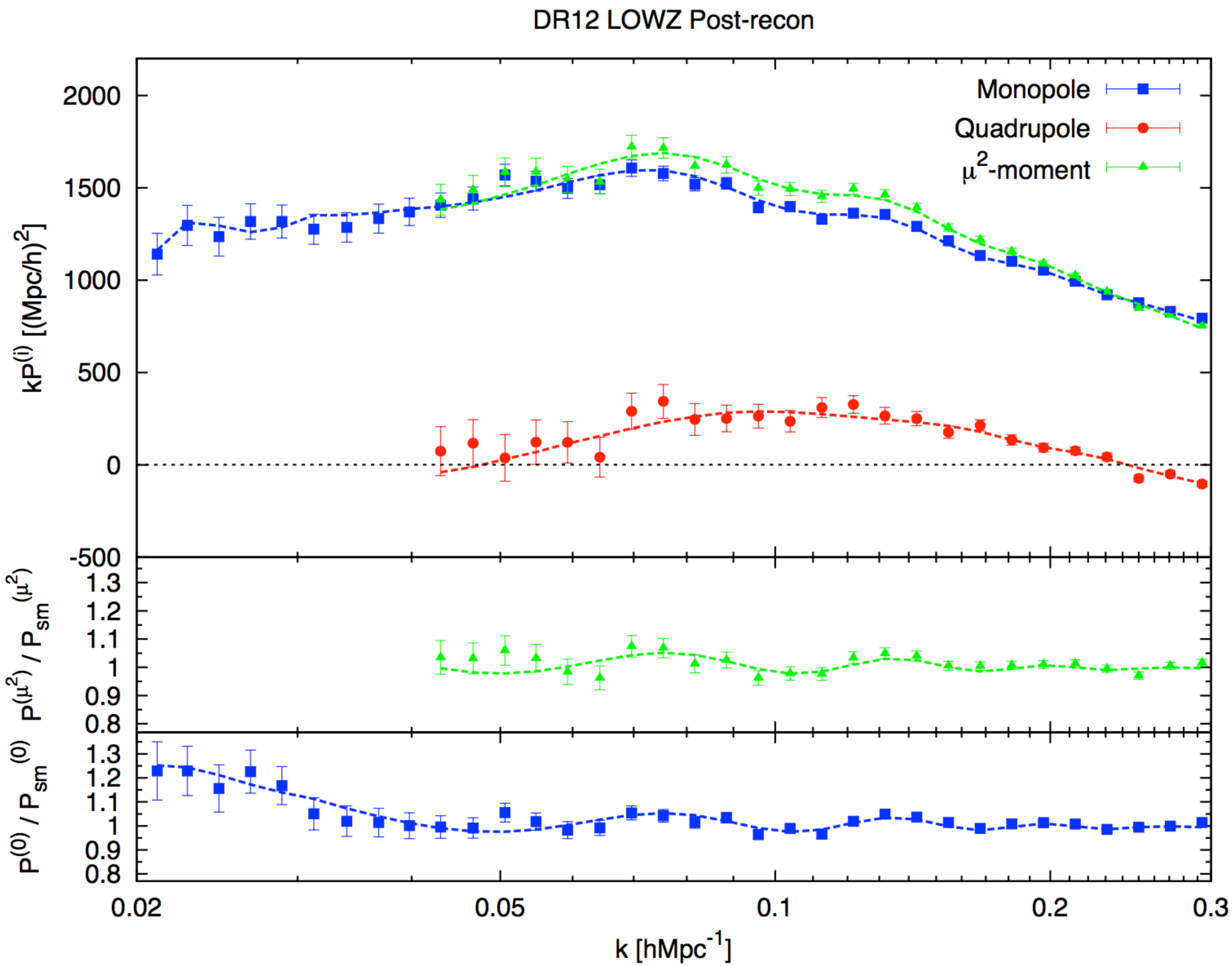
$$P_{\rm model}(k;\alpha) = P_{\rm model,sm}(k)\left\{1 + [O_{\rm lin}(k/\alpha) - 1]e^{-\frac{1}{2}k^2\Sigma_{\rm nl}^2}\right\}$$

$$P_{\rm model,sm}(k) = B^2 P_{\rm lin,sm}(k) + A_1 k + A_2 + \frac{A_3}{k} + \frac{A_4}{k^2} + \frac{A_5}{k^3}$$

- Convolve with the survey window function
- Allow varying BAO damping Σ_{nl} with a prior from mocks
- MCMC to find mean and variance for 16 parameters
- Interpret BAO peak position in monopole and μ^2-moment as power law combinations of $\alpha_\perp, \alpha_{||}$
- Test using different cosmologies to convert from angles & redshifts to distances
- Test using different BAO models to be fitted to the data

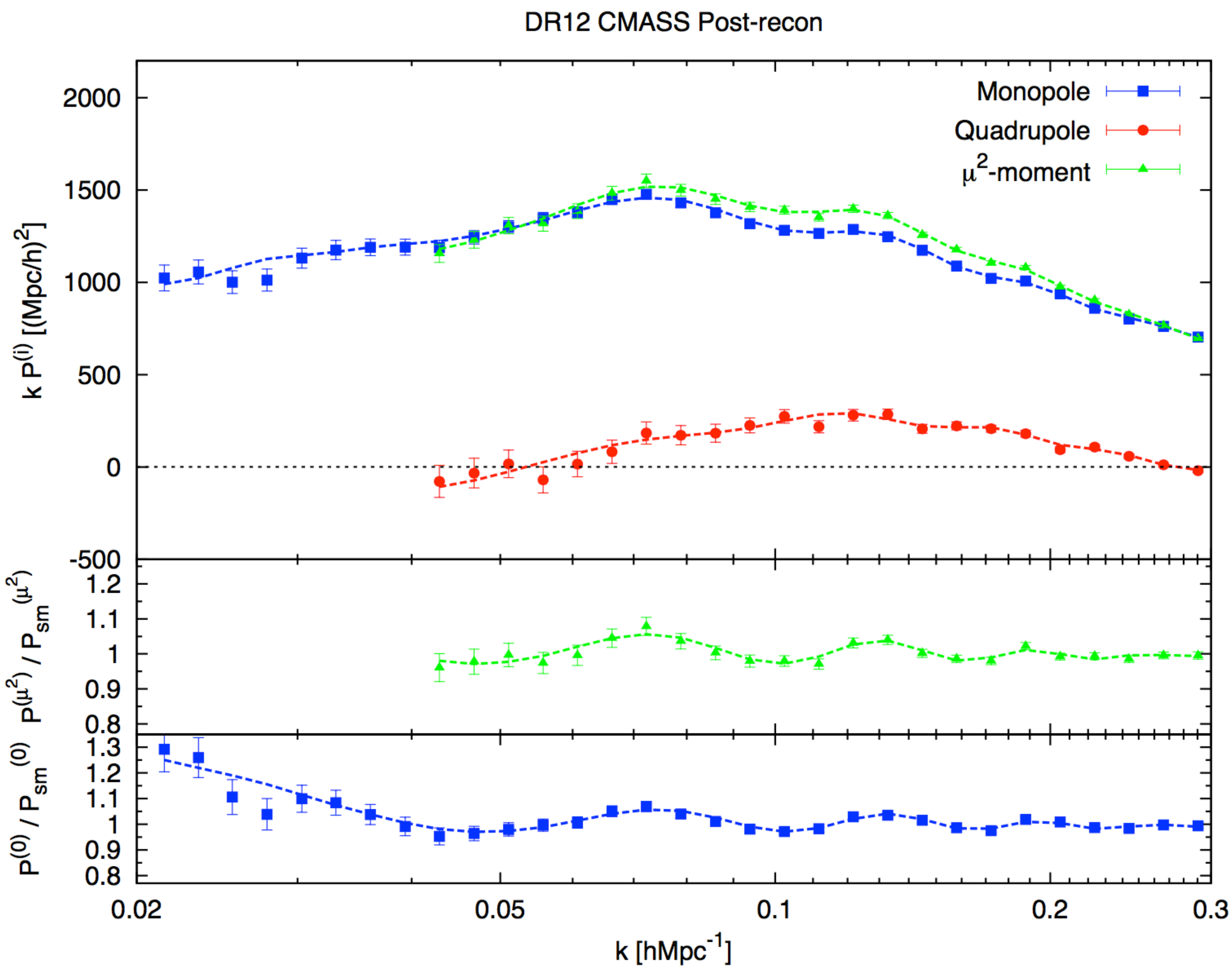
Gil-Marin et al. 2015; arXiv:1509.06386

Measuring BAO for SDSS DR12



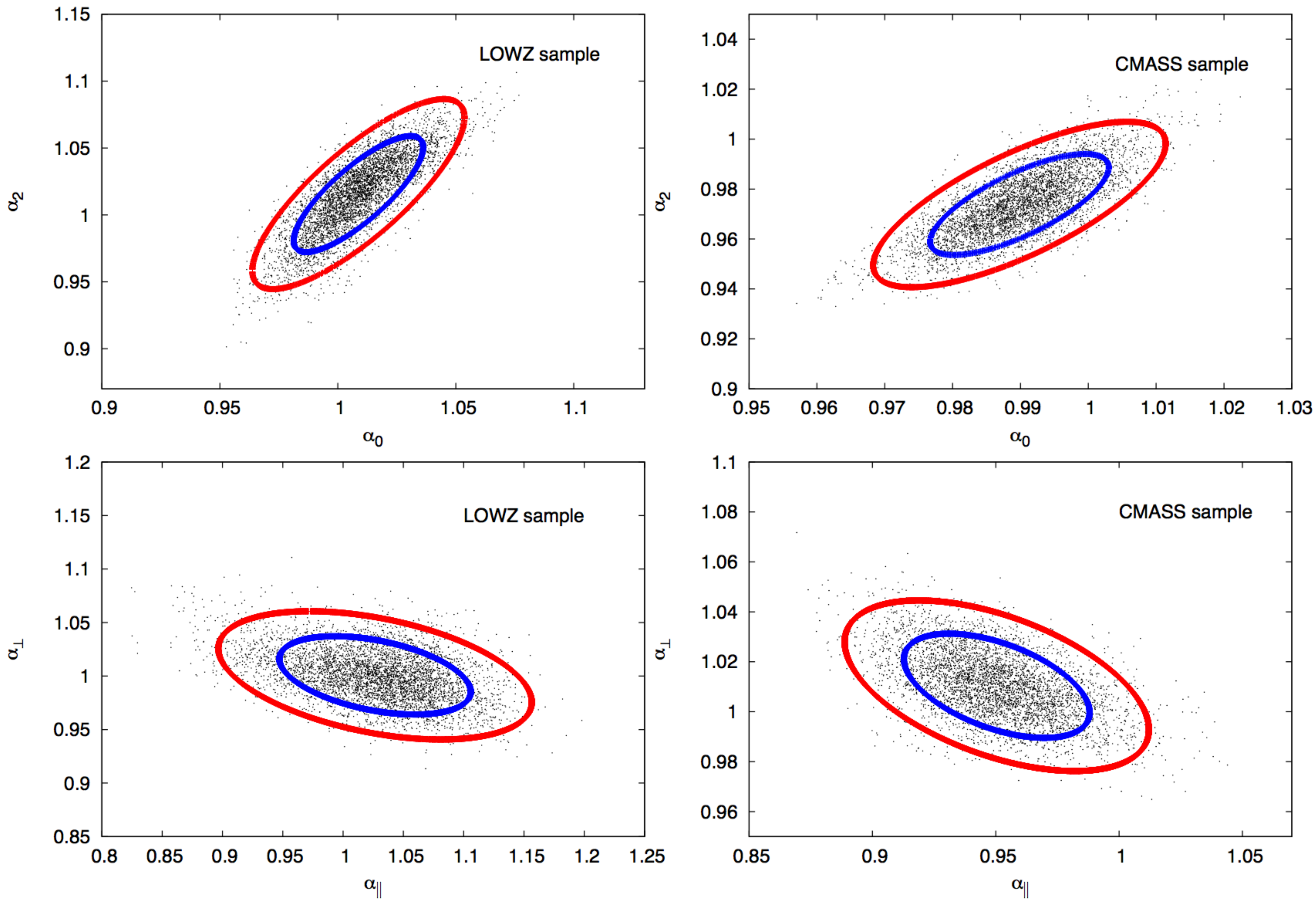
Gil-Marin et al. 2015; arXiv:1509.06386

Measuring BAO for SDSS DR12



Gil-Marin et al. 2015; arXiv:1509.06386

Measuring BAO for SDSS DR12



$$H(0.32)r_s(z_d) = (11.64 \pm 0.62)10^3\,\mathrm{kms}^{-1}, \quad D_A(0.32)/r_s(z_d) = 6.85 \pm 0.17, \quad r = 0.42$$

$$H(0.57)r_s(z_d) = (14.56 \pm 0.38)10^3\,\mathrm{kms}^{-1}, \quad D_A(0.57)/r_s(z_d) = 9.42 \pm 0.13, \quad r = 0.51$$

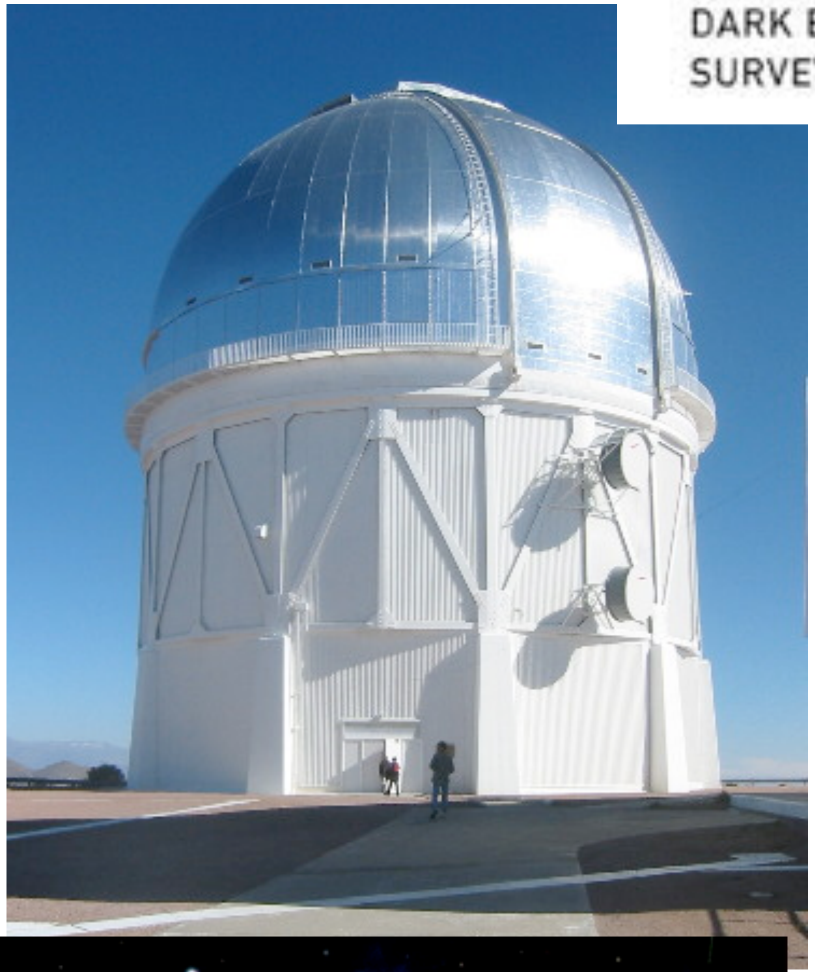
Gil-Marin et al. 2015; arXiv:1509.06386

Future surveys: next 4-6 years

Dark Energy Survey (DES)

- New wide-field camera on the 4m Blanco telescope
- Survey started, with first year of data in hand
- $\Omega = 5{,}000 \text{deg}^2$
- multi-colour optical imaging (g,r,i,z) with link to IR data from VISTA hemisphere survey
- 300,000,000 galaxies
- Aim is to constrain dark energy using 4 probes
  LSS/BAO, weak lensing, supernovae
  cluster number density
- Redshifts based on photometry
  weak radial measurements
  weak redshift-space distortions
- See also: Pan-STARRS, VST-VISTA, SkyMapper

DARK ENERGY SURVEY



eBOSS / SDSS-IV

- The new cosmology project with SDSS
- Use the Sloan telescope and MOS to observe to higher redshift
- Basic parameters
  - $\Omega = 1{,}500\text{deg}^2 - 7{,}500\text{deg}^2$
  - ~ 1,000,000 galaxies (direct BAO)
  - ~ 60,000 quasars (BAO from Ly-α forest)
- Distance measurements
  - 0.9% at z=0.8 (LRGs)
  - 1.8% at z=0.9 (ELGs)
  - 2.0% at z=1.5 (QSOs)
  - 1.1% at z=2.5 (Ly-α forest, inc. BOSS)
- Survey started 2014, lasting 6 years

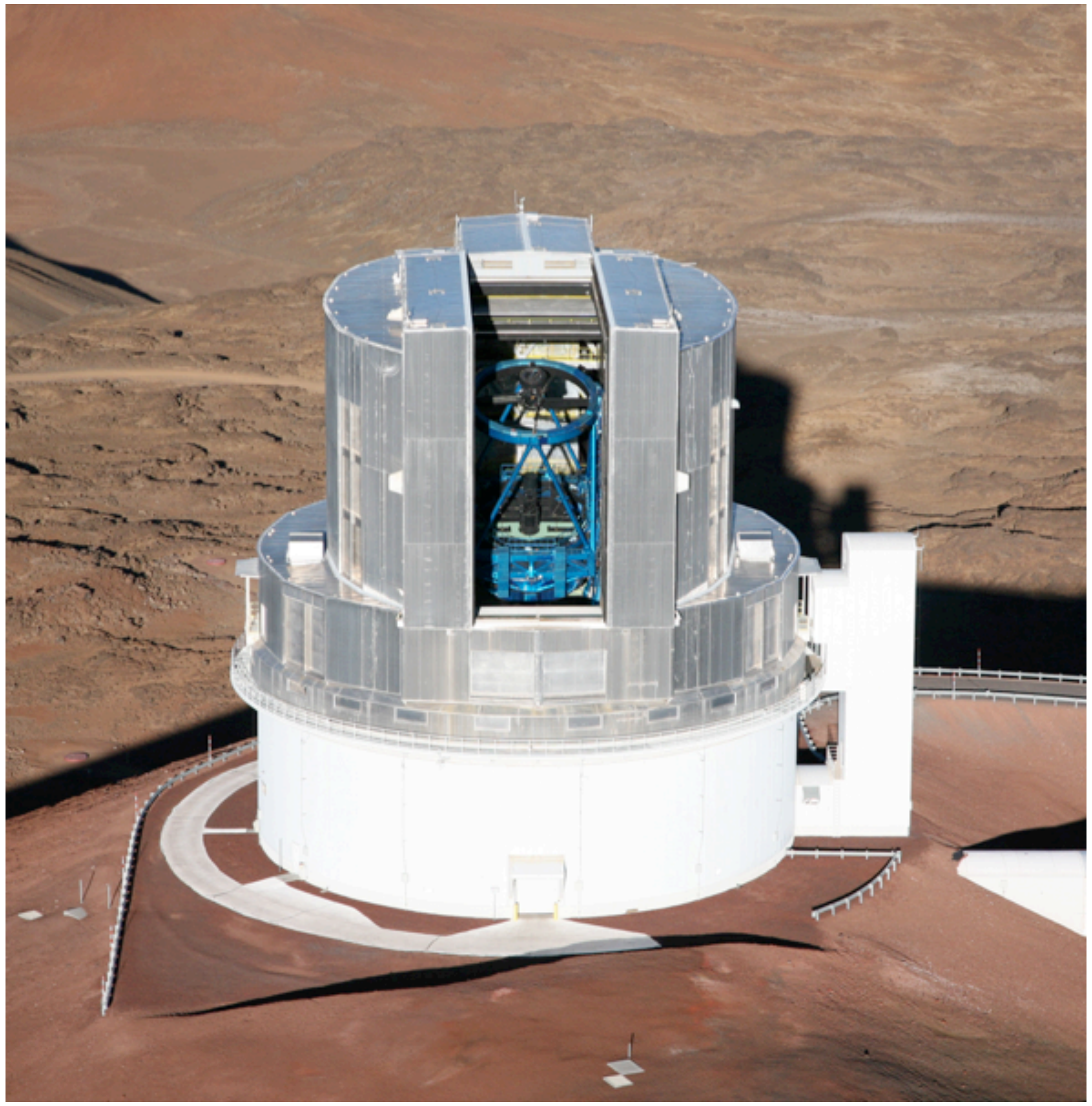
Future surveys: > 4 years

MOS on 4m-telescope

- New fibre-fed spectroscopes proposed for 4m telescopes
  - Mayall (BigBOSS) } DESI
  - Blanco (DESpec) } DESI
  - WHT (WEAVE)
  - VISTA (4MOST)
- Various stages of planning & funding
  - DESI at DOE CD-1, CD-2 soon, 2019 start
  - 4MOST chosen by ESO, 2020 start?
  - WEAVE, 2018 start
- All capable of observing
  - Ω =5--14,000deg^2
  - 2--40,000,000 galaxies (direct BAO)
  - 1--600,000 quasars (BAO from Ly-α forest)
  - Cosmic variance limited to z ~ 1.4

MOS on 10m-telescope

- New fibre-fed spectroscopes proposed for 10m telescopes
    - Hobby-Eberly (HETDEX)
    - Subaru (PFS)
- Different baseline strategies
- HETDEX
    - 420deg^2 Ly-alpha emitters
    - 800,000 galaxies $1.9<z<3.5$
    - Greig, Komatsu & Wyithe, 2012, arXiv:12120977
- PFS
    - 1400deg^2 ELGs
    - 3,000,000 galaxies $0.6<z<2.4$
    - Ellis et al., 2012, arXiv:1206.0737

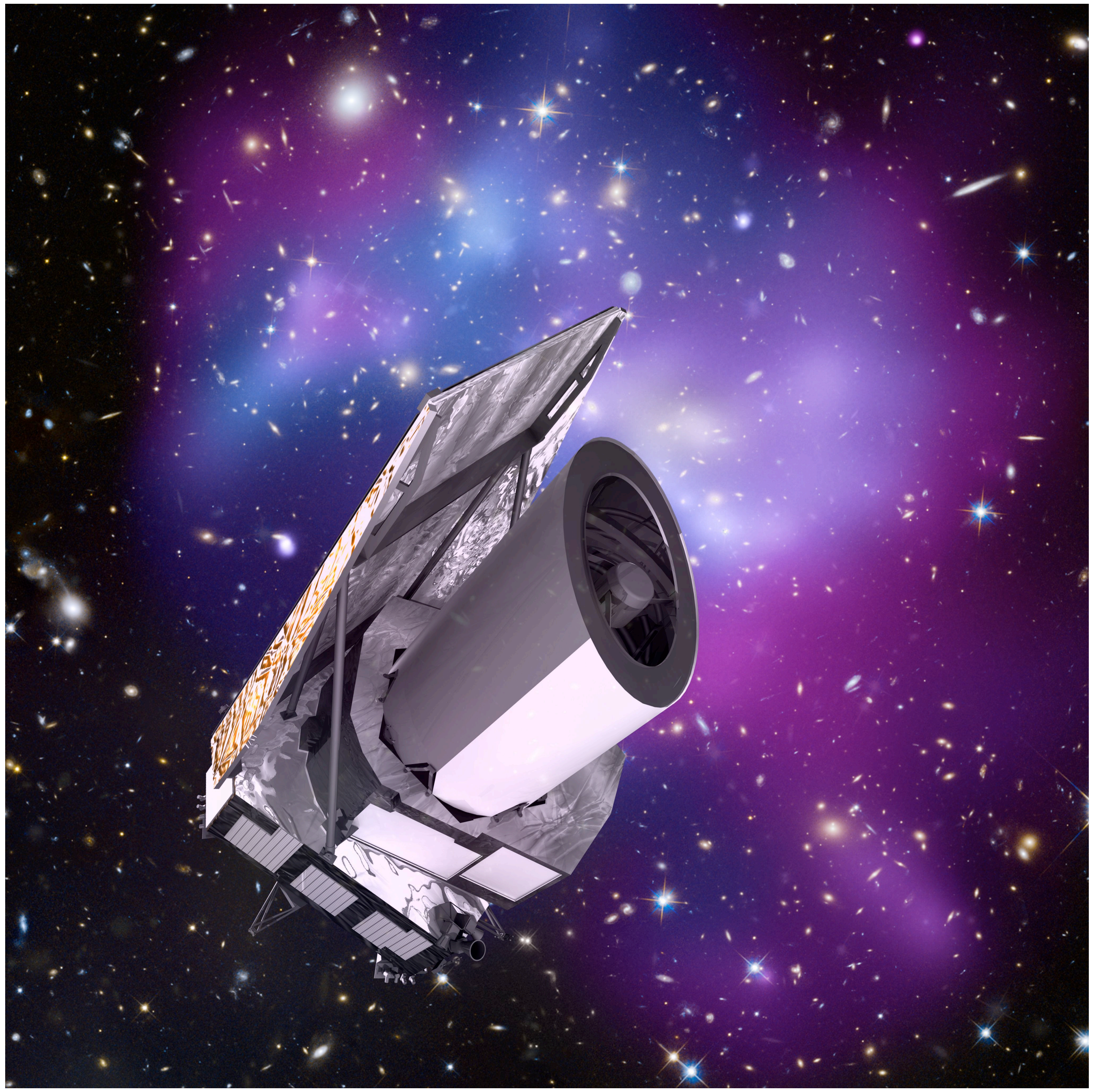


The ESA Euclid Mission

SURVEYS					
	Area (deg2)	Description			
Wide Survey	15,000 (required) 20,000 (goal)	Step and stare with 4 dither pointings per step.			
Deep Survey	40	In at least 2 patches of > 10 deg^2 2 magnitudes deeper than wide survey			
PAYLOAD					
Telescope	1.2 m Korsch, 3 mirror anastigmat, f=24.5 m				
Instrument	VIS	NISP			
Field-of-View	0.787×0.709 deg^2	0.763×0.722 deg^2			
Capability	Visual Imaging	NIR Imaging Photometry			NIR Spectroscopy
Wavelength range	550– 900 nm	Y (920-1146nm),	J (1146-1372 nm)	H (1372-2000nm)	1100-2000 nm
Sensitivity	24.5 mag 10σ extended source	24 mag 5σ point source	24 mag 5σ point source	24 mag 5σ point source	3 10^{-16} erg cm-2 s-1 3.5σ unresolved line flux
Detector Technology	36 arrays 4k×4k CCD	16 arrays 2k×2k NIR sensitive HgCdTe detectors			
Pixel Size Spectral resolution	0.1 arcsec	0.3 arcsec			0.3 arcsec R=250
SPACECRAFT					
Launcher	Soyuz ST-2.1 B from Kourou				
Orbit	Large Sun-Earth Lagrange point 2 (SEL2), free insertion orbit				
Pointing	25 mas relative pointing error over one dither duration 30 arcsec absolute pointing error				
Observation mode	Step and stare, 4 dither frames per field, VIS and NISP common FoV = 0.54 deg^2				
Lifetime	7 years				
Operations	4 hours per day contact, more than one ground station to cope with seasonal visibility variations;				
Communications	maximum science data rate of 850 Gbit/day downlink in K band (26GHz), steerable HGA				

http://www.euclid-ec.org/

The Euclid spectroscopic survey

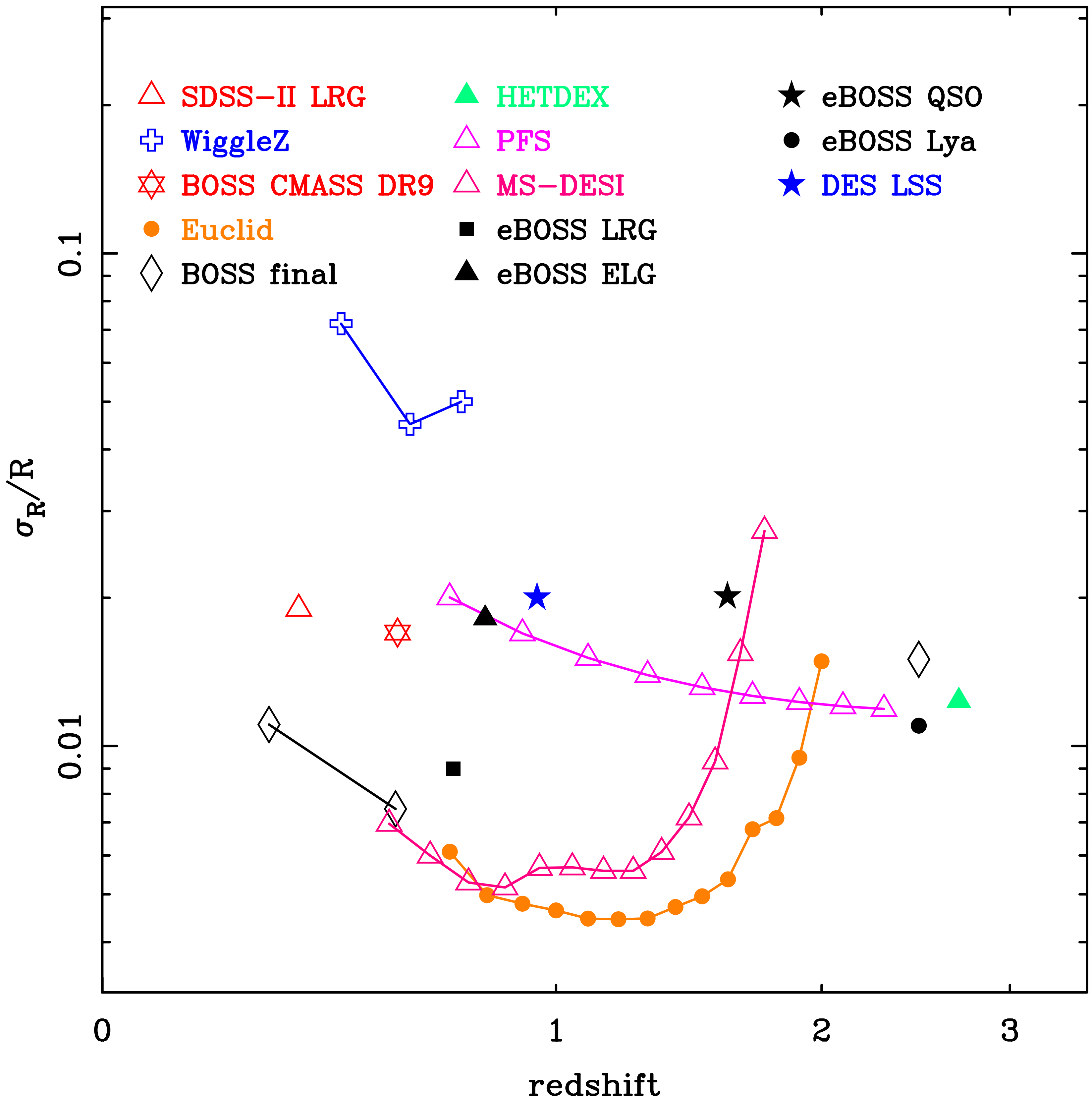


- Wide survey
  - 15,000deg^2
  - 4 dithers
  - NIR Photometry
    - Y, J, H
    - 24mag, 5σ point source
  - NIR slitless spectroscopy
    - red: 1.25-1.85μm,
    - 2×10^{-16}ergcm^{-2}s^{-1} 3.5σ line flux
    - 3 dispersion directions
    - 1 broad waveband 0.9<z<1.8
    - ~25M galaxies

- Deep survey
  - 40deg^2
  - 48 dithers
  - 12 passes, as for wide survey
  - additional blue spectra: 0.92-1.25μm
  - dispersion directions for 12 passes >10deg apart

http://www.euclid-ec.org/

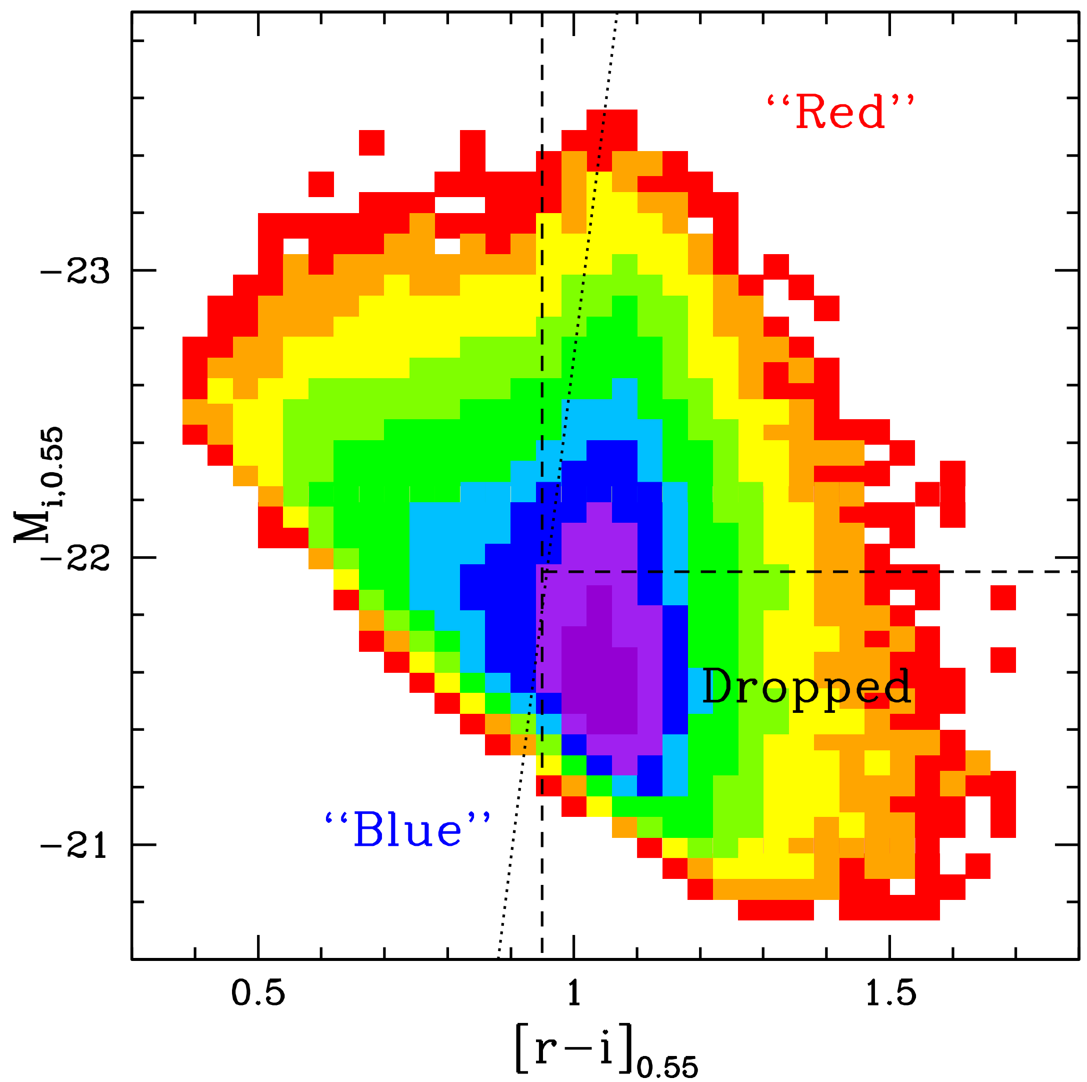
BAO measurements for future surveys

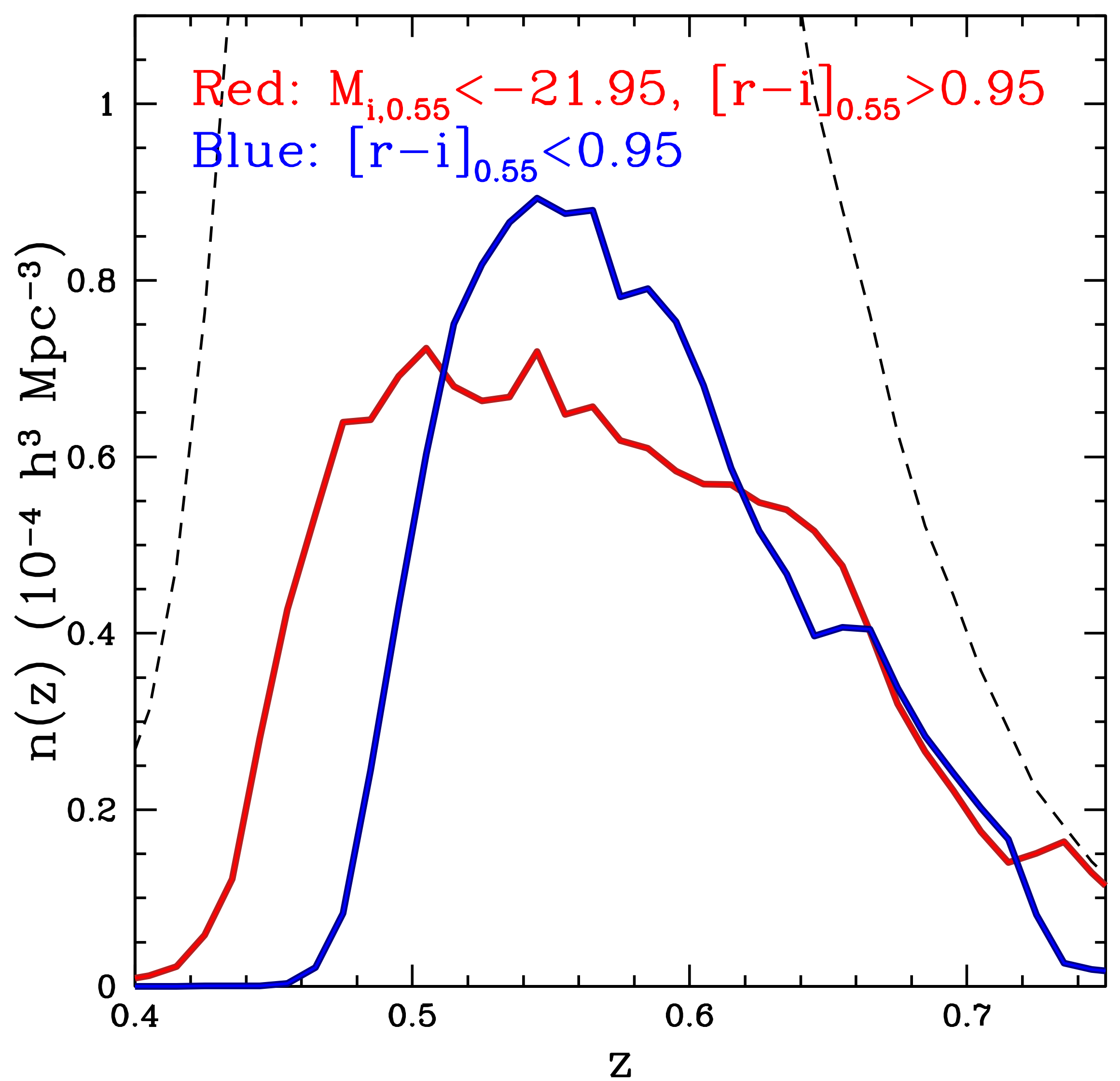


using the code of Seo & Eisenstein 2007, arXiv:0701079

Worrying about astrophysics

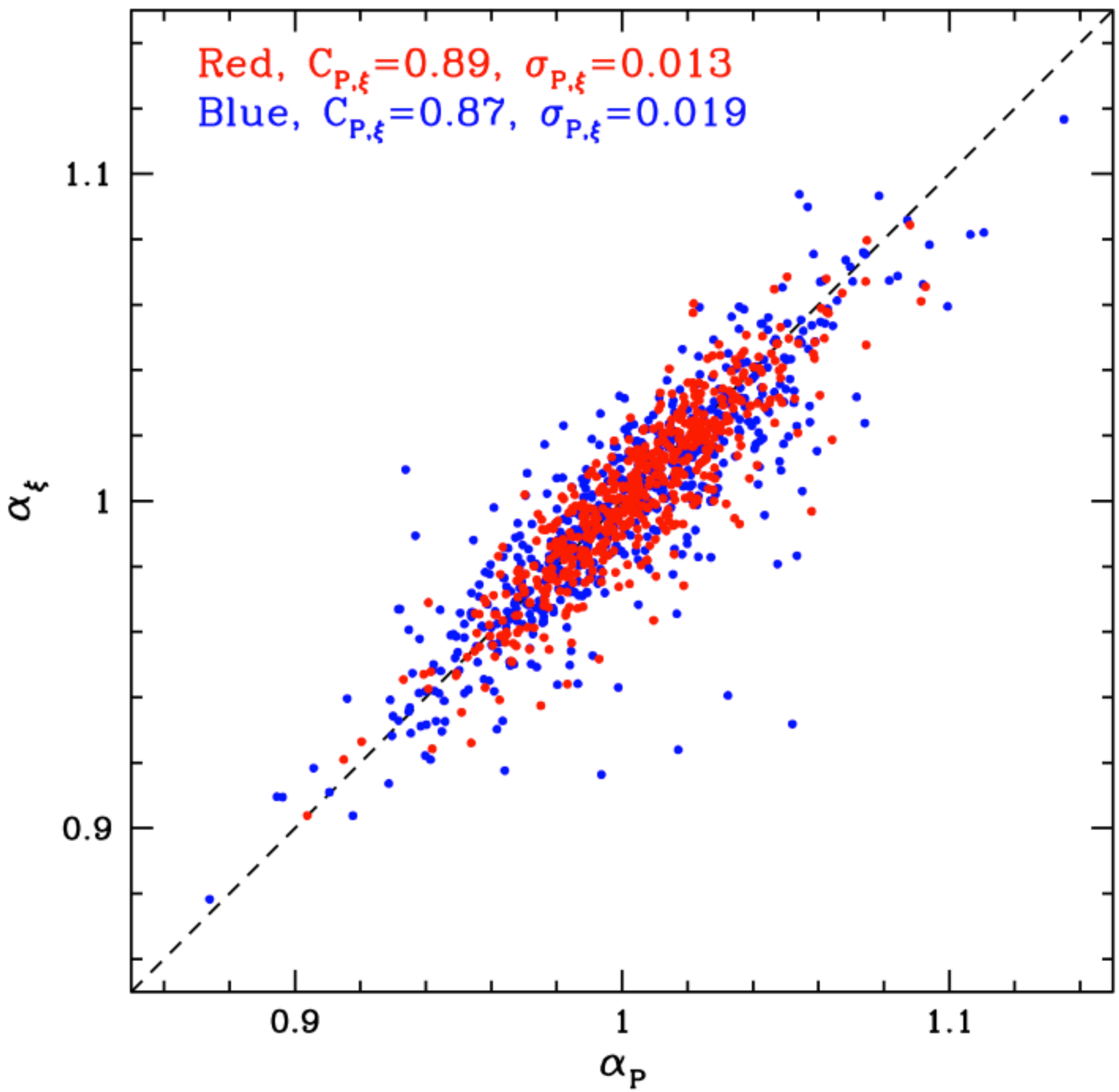
Testing with blue / red subsamples





Ross et al. 2014, MNRAS 437, 1109

Testing with blue / red subsamples



Ross et al. 2014, MNRAS 437, 1109

Worrying about statistics

Getting the likelihood calculation correct

The Likelihood under the standard assumption of a set of data drawn from a multi-variate Gaussian distribution is given by

$$\mathcal{L}(\mathbf{x}|\mathbf{p},\Psi^t) = \frac{|\Psi^t|}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\chi^2(\mathbf{x},\mathbf{p},\Psi^t)\right],$$

where $$\chi^2(\mathbf{x},\mathbf{p},\Psi^t) \equiv \sum_{ij} \left[x_i^d - x_i(\mathbf{p})\right] \Psi^t_{ij} \left[x_j^d - x_j(\mathbf{p})\right].$$

now suppose that the covariance matrix (size n_b x n_b) has been calculated from n_s simulations

$$\mu_i = \frac{1}{n_s}\sum_s x_i^s \qquad C_{ij} = \frac{1}{n_s - 1}\sum_s (x_i^s - \mu_i)(x_j^s - \mu_j)$$

then an unbiased estimator of the inverse covariance matrix is

$$\Psi = \frac{n_s - n_b - 2}{n_s - 1} C^{-1}$$

[compare with $\langle 1/x \rangle \neq 1/\langle x \rangle$]

Hartlap J., Simon P., Schneider P., 2007; A&A, 464, 399

Errors in the covariance matrix

Simply providing an unbiased estimator of the inverse covariance matrix is not enough

The inverse covariance matrix also has its own error

$$\langle \Delta\Psi_{ij} \Delta\Psi_{i'j'} \rangle = A \Psi_{ij} \Psi_{i'j'} + B(\Psi_{ii'} \Psi_{jj'} + \Psi_{ij'} \Psi_{ji'}),$$

$$A = \frac{2}{(n_s - n_b - 1)(n_s - n_b - 4)}$$

$$B = \frac{(n_s - n_b - 2)}{(n_s - n_b - 1)(n_s - n_b - 4)}$$

Strictly, we should form a joint likelihood

$$\mathcal{L}(\mathbf{x}, \Psi | \mathbf{p}, \Psi^t) = \mathcal{L}(\mathbf{x} | \mathbf{p}, \Psi) \mathcal{L}(\Psi | \Psi^t),$$

If we don’t, this leads to an additional error on the $n_{\mathbf{p}}$ parameters being fitted

$$\langle p_\alpha p_\beta \rangle |_{s.o.} = B(n_b - n_p) F^{-1}_{\alpha\beta},$$

Taylor et al., 2012, arXiv:1212.4359; Dodelson & Schneider 2007, arXiv:1212.4359

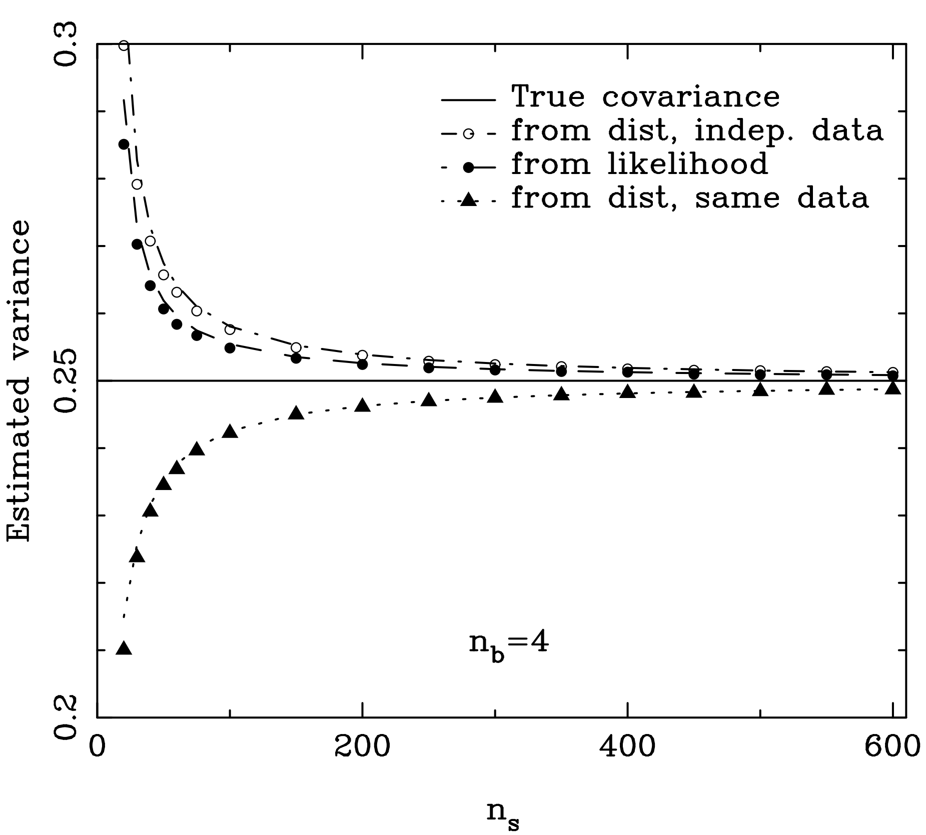
Errors in likelihood calculations

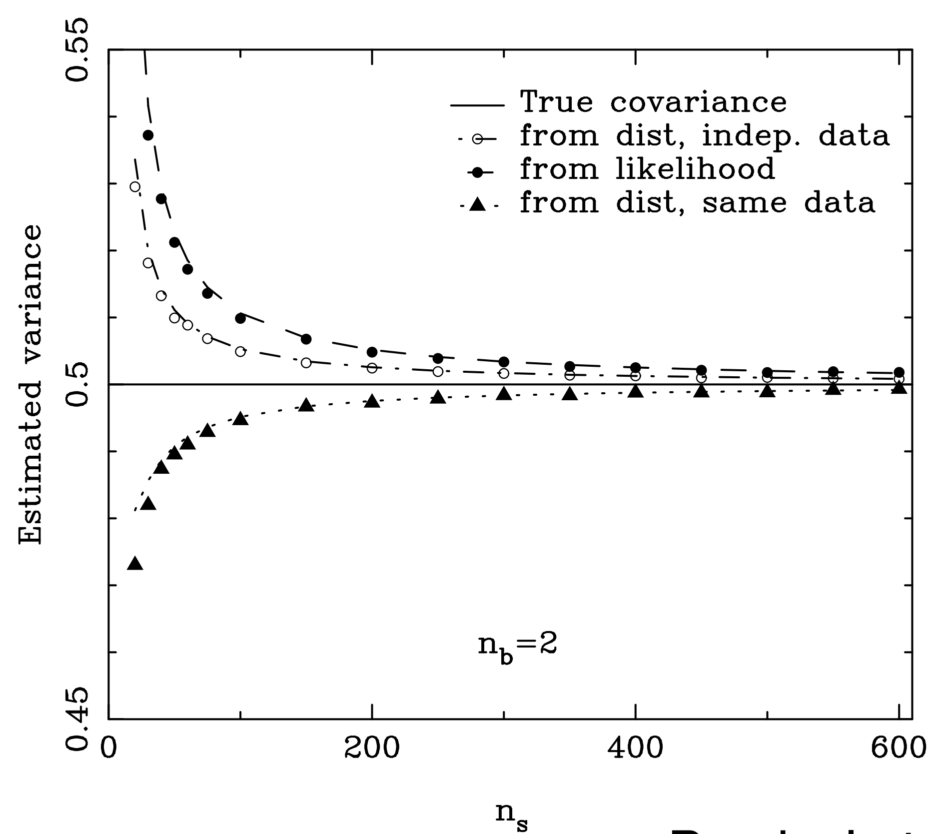
Given a set of mocks, we can form three possible estimates of the errors:

1. From the individual likelihood surface from each mock
2. From the distribution of recovered measurements from the set of mocks
3. From the distribution of a set of mocks not used to calcaulte covariance matrix

These should all agree!

The estimates from each are biased in subtly different ways given errors in the covariance matrix

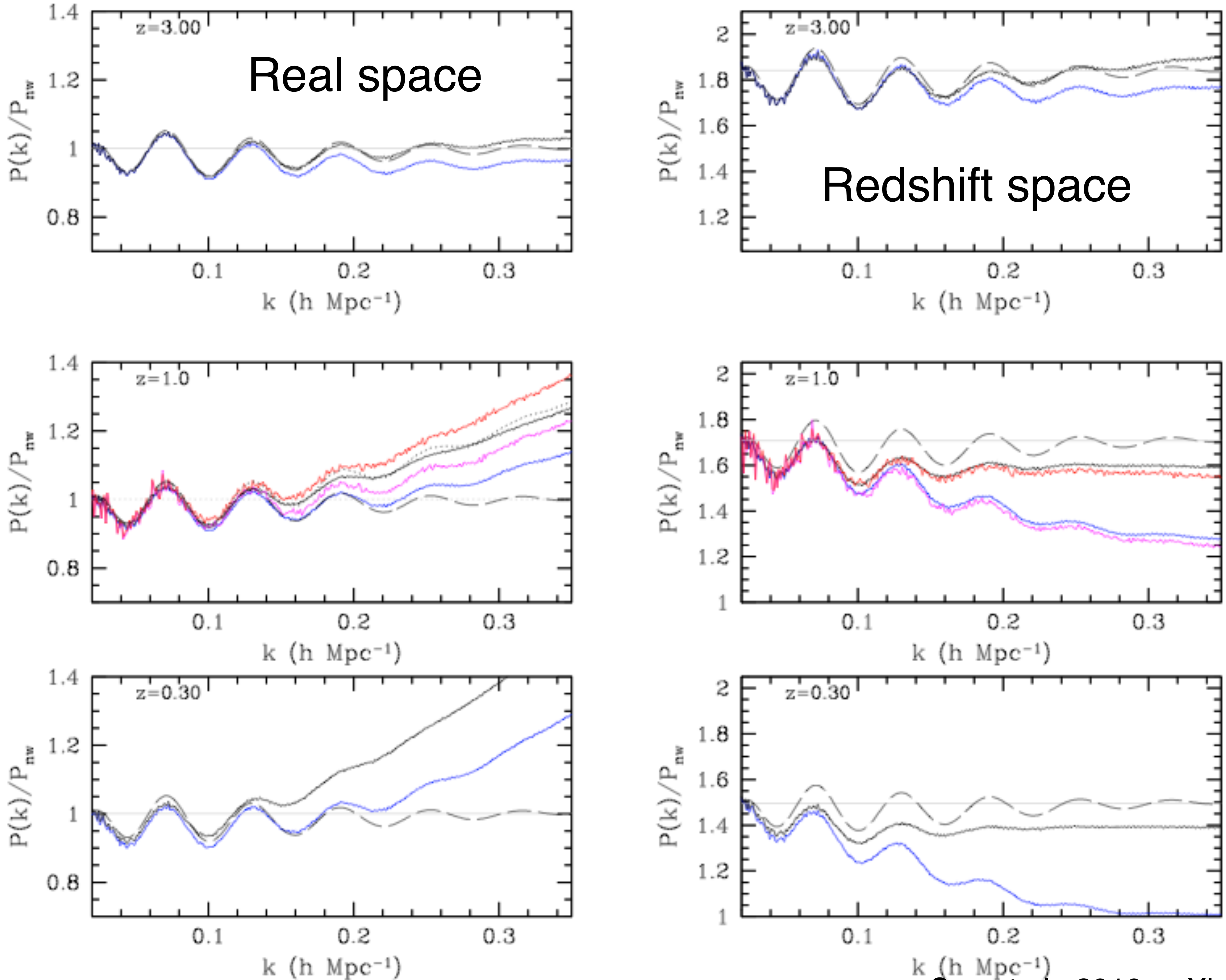




Percival et al., 2013: arXiv:1312.4841

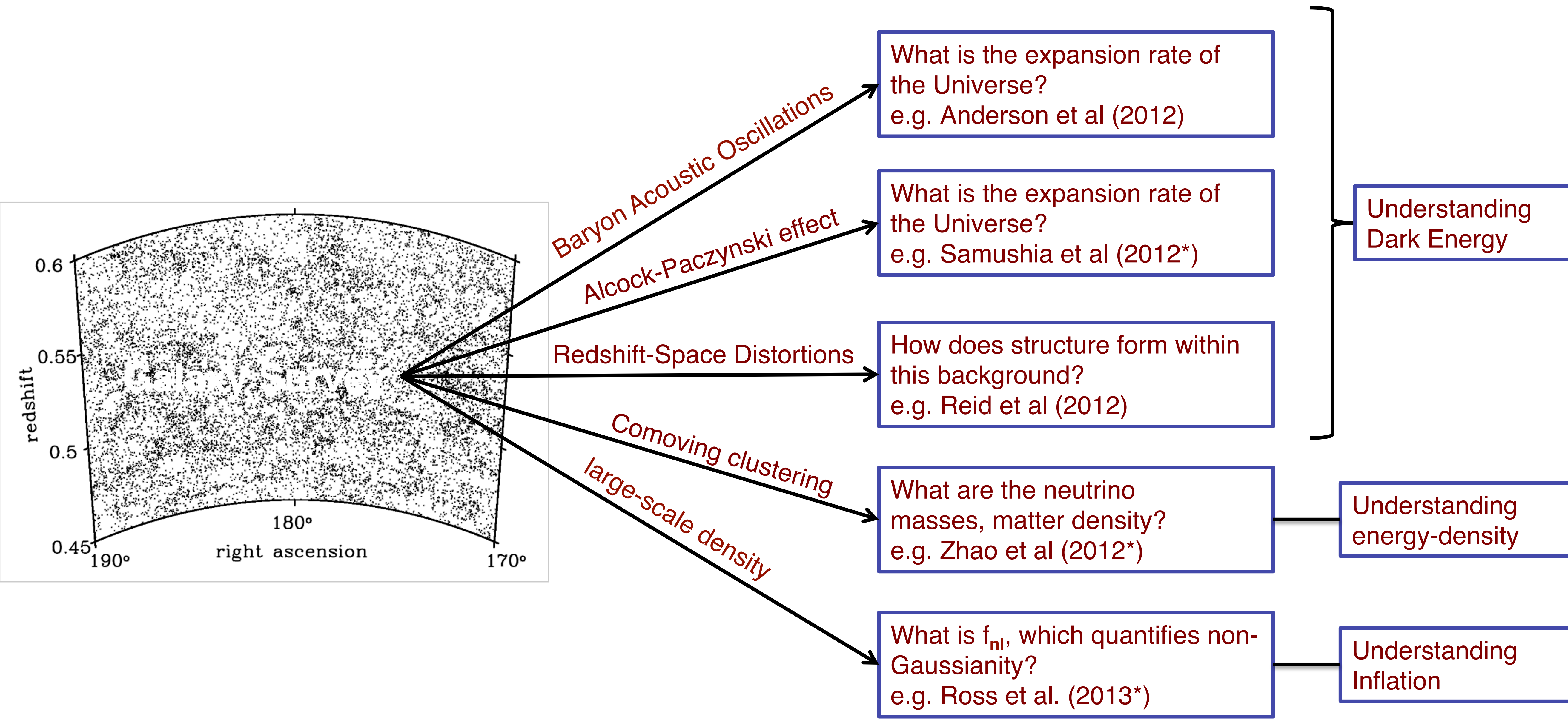
Worrying about non-linear physics

BAO from simulations



Seo et al., 2010, arXiv:0910.5005

Cosmology from galaxy surveys



Forthcoming surveys extremely exciting, but will require methodology & simulation development to reach statistical limit