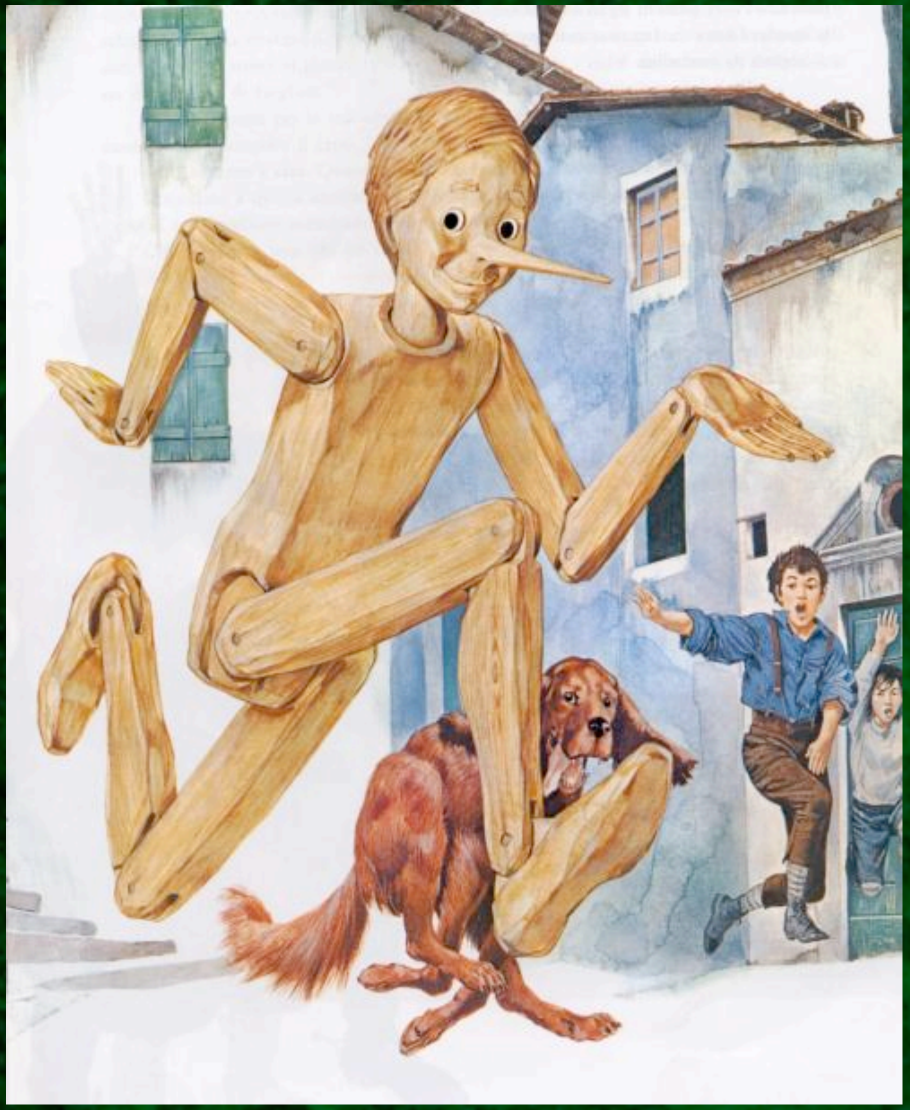
Approximate methods for the generation of dark matter halo catalogs in the age of precision cosmology

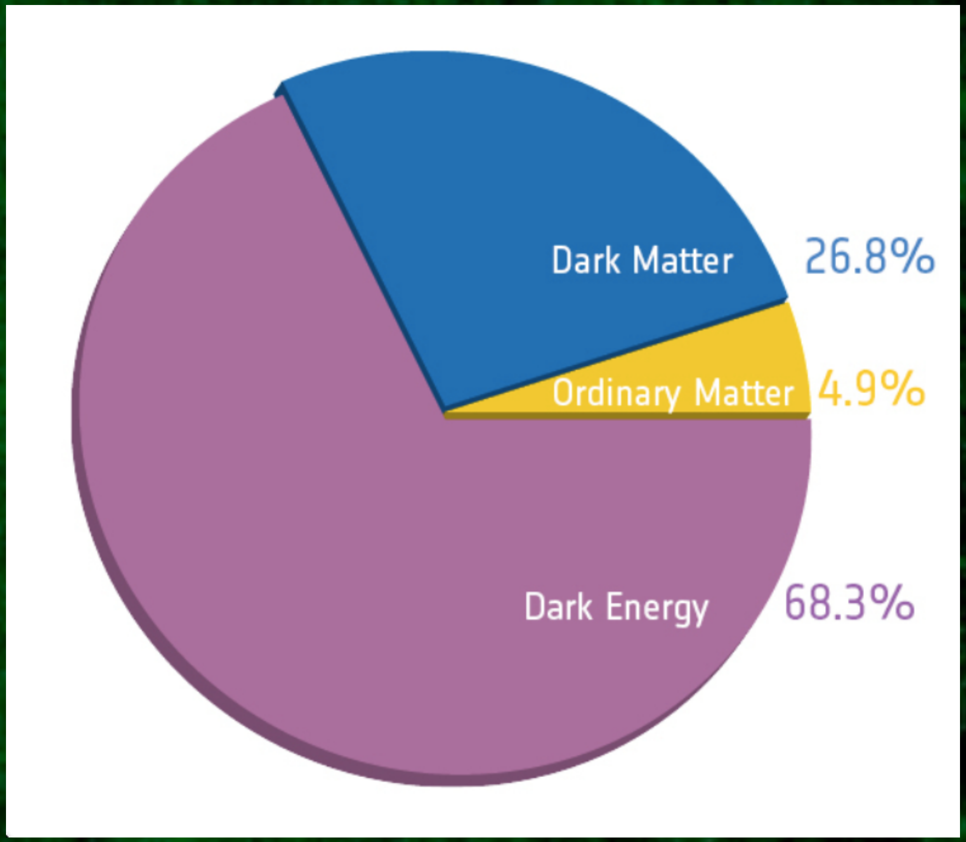
P. Monaco,
University of Trieste & INAF-OATs & INFN-TS

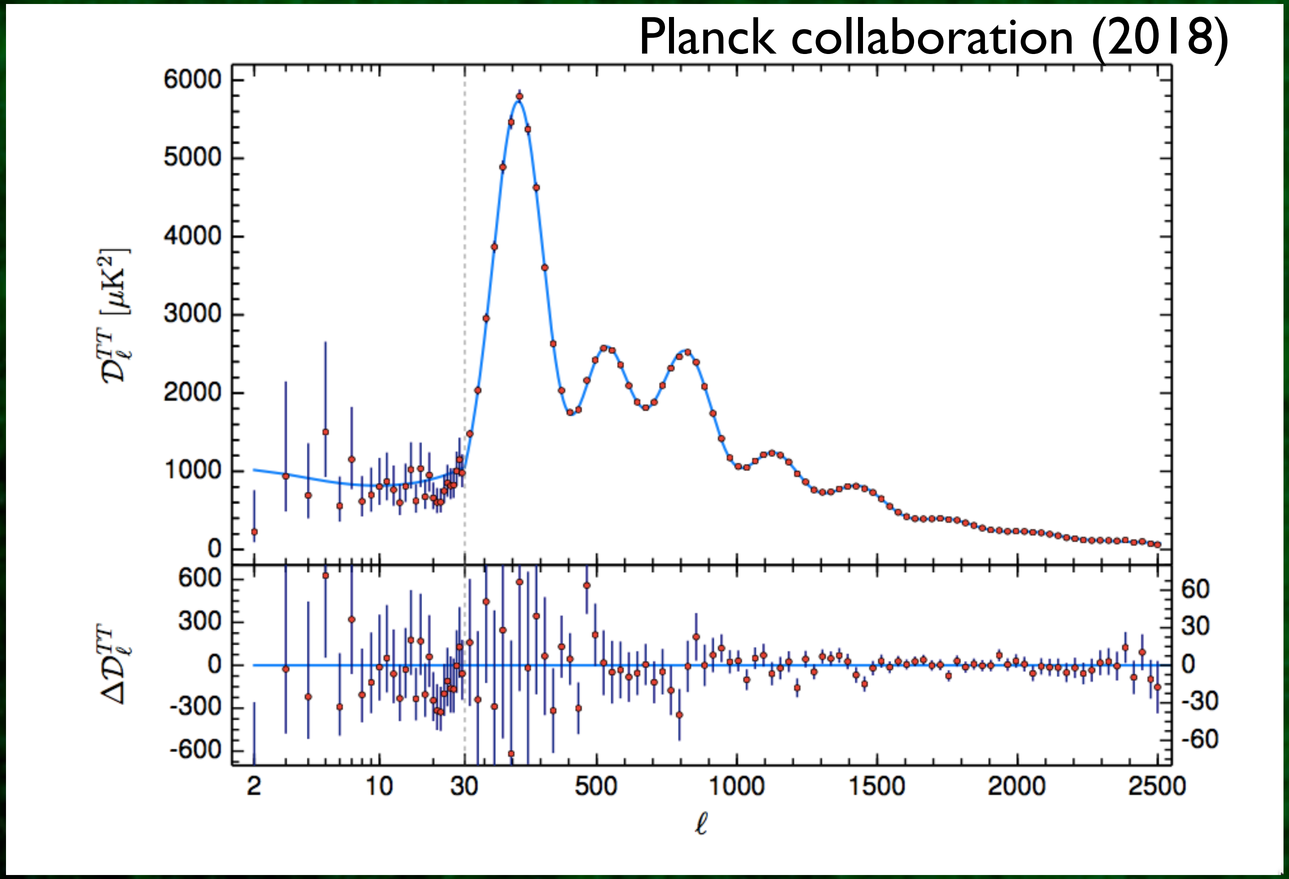


Background image credit: E. Munari

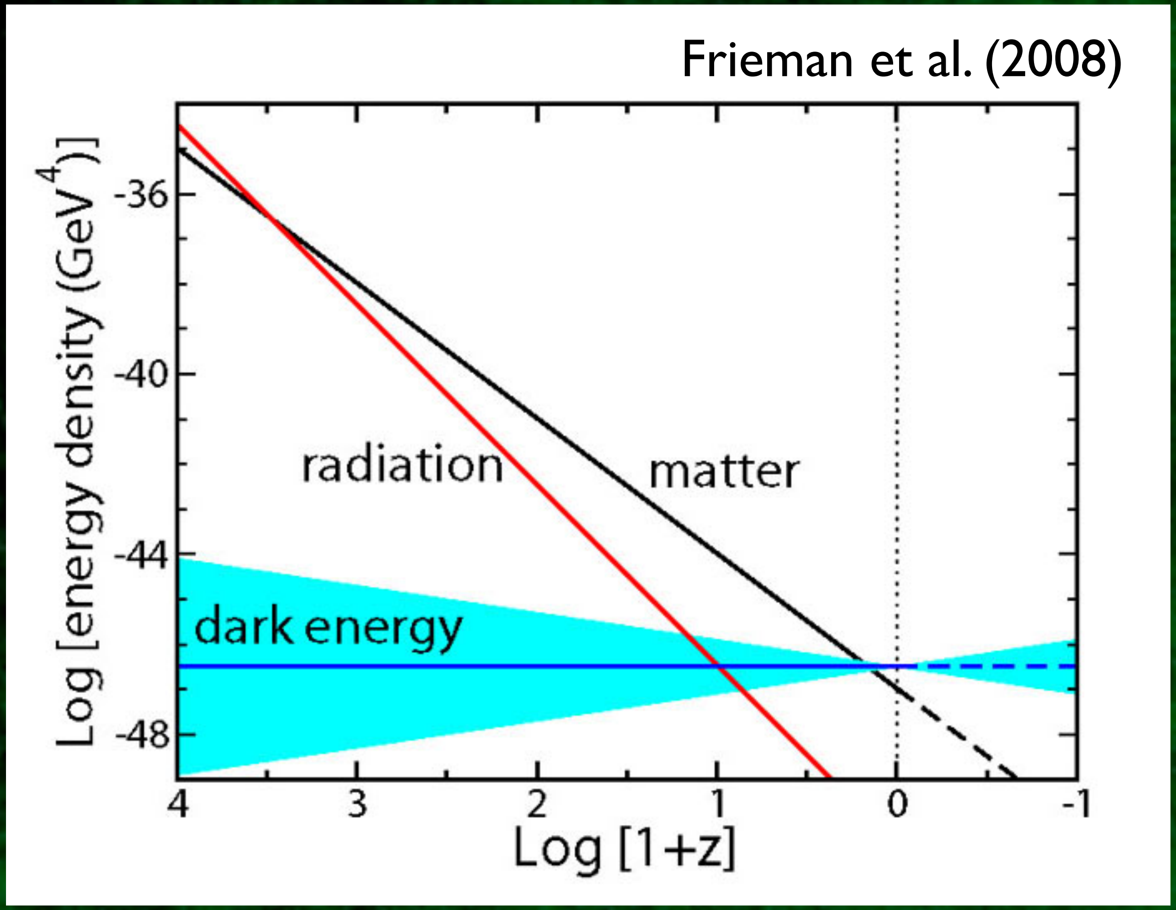
Cosmology in the next decade, Bangalore, January 2018

CMB observations give us an accurate snapshot of our Universe at z~1100

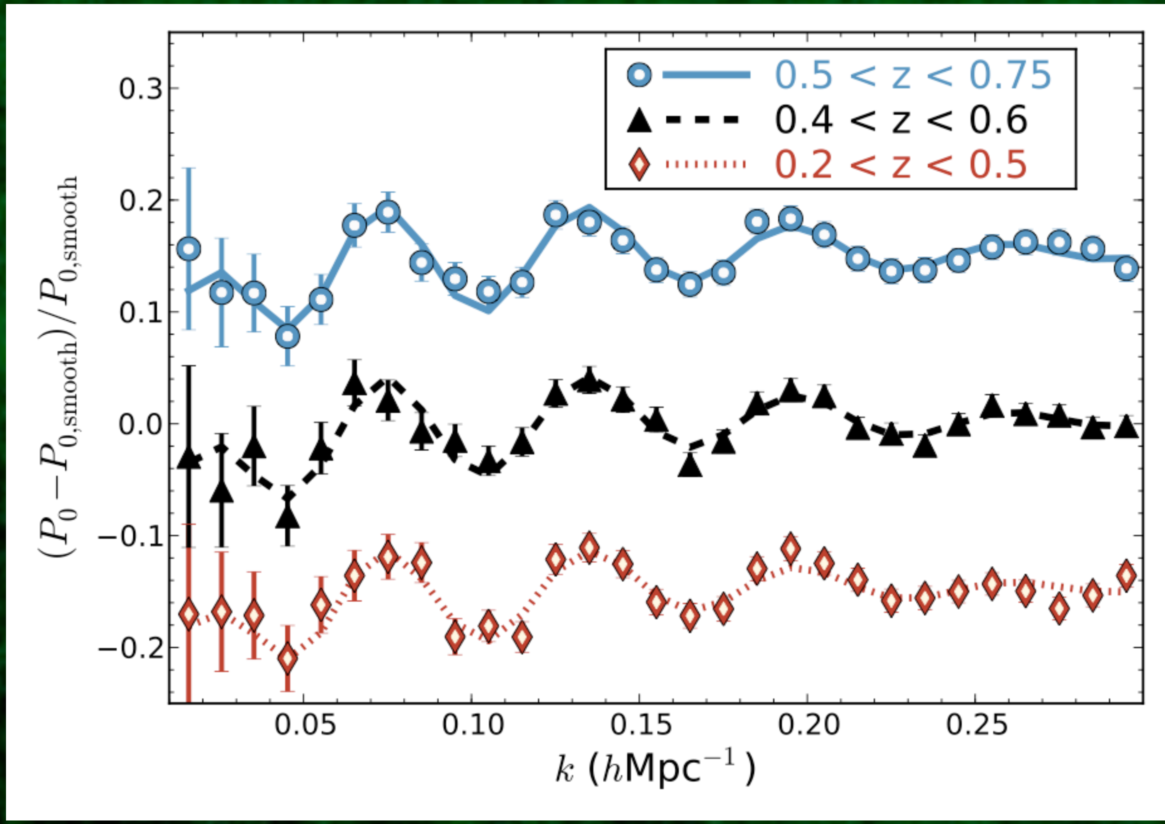




An assessment of the nature of dark energy requires <1% accurate measurements at z~1



SDSS BOSS collaboration, Alam et al. (2016)

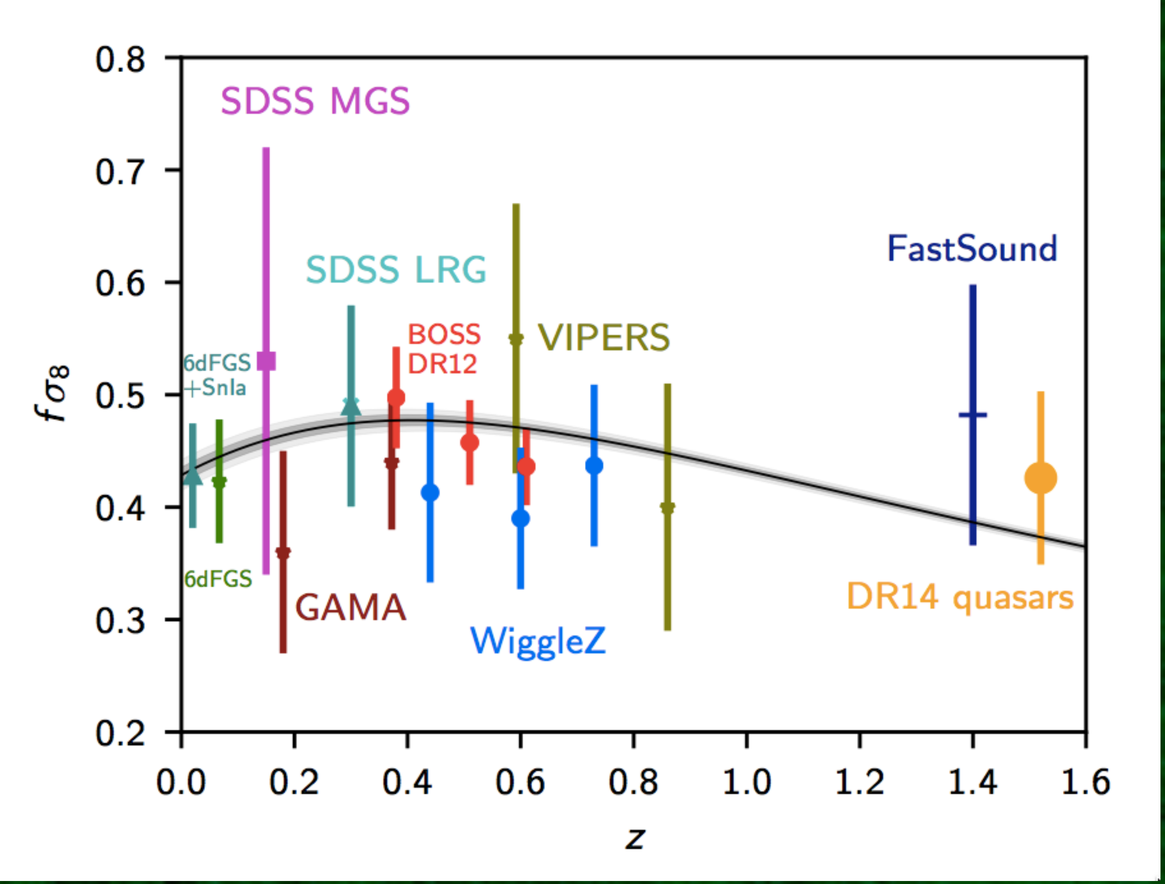


Clustering probes:

The Baryonic Acoustic Oscillation (BAO) feature of the matter power spectrum provides a standard ruler

The growth rate of structure is a specific prediction of the gravity theory

Planck collaboration (2018)



Euclid consortium

$w_0 = -0.95 \pm 0.1$ is not interesting

$w_0 = -0.95 \pm 0.01$ is interesting

Error budget becomes dominated by systematics

$$P(\theta|\mathbf{D}) = \frac{\mathcal{L}(\mathbf{D}|\theta)\, P(\theta)}{P(\mathbf{D})}$$

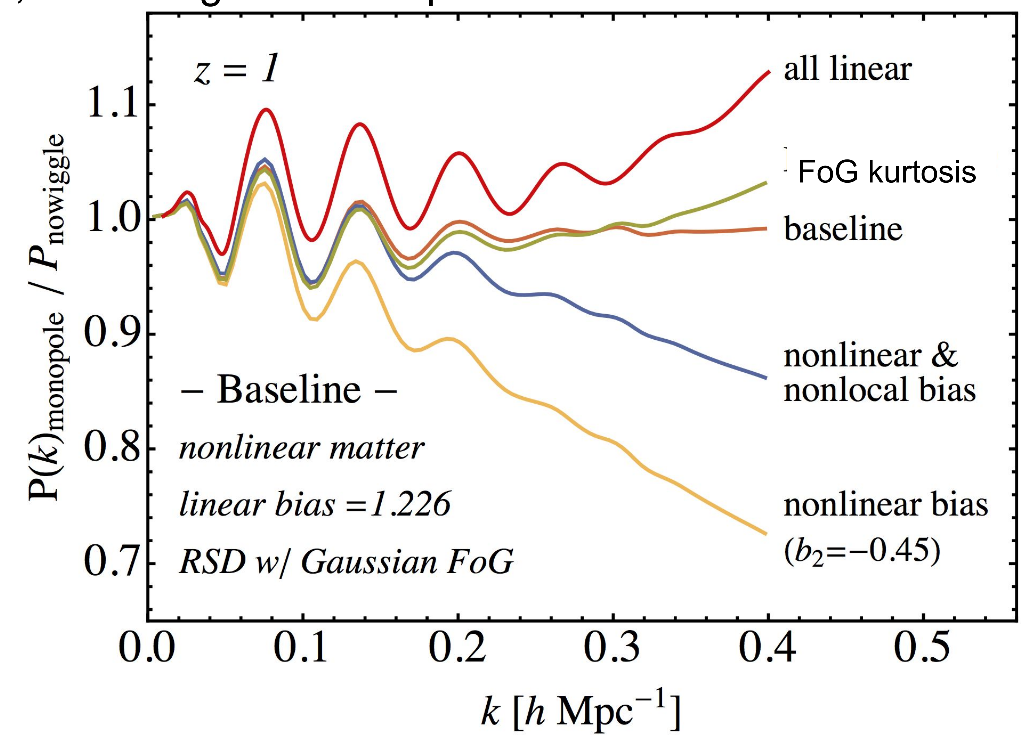
There are three possible sources of systematic errors:

- **Data**: all issues that concern the processing from raw data to measurements
- **Theory**: errors related to the models used to compute predictions as a function of cosmological parameters
- **Likelihood**: the assumptions made to link theory and observations

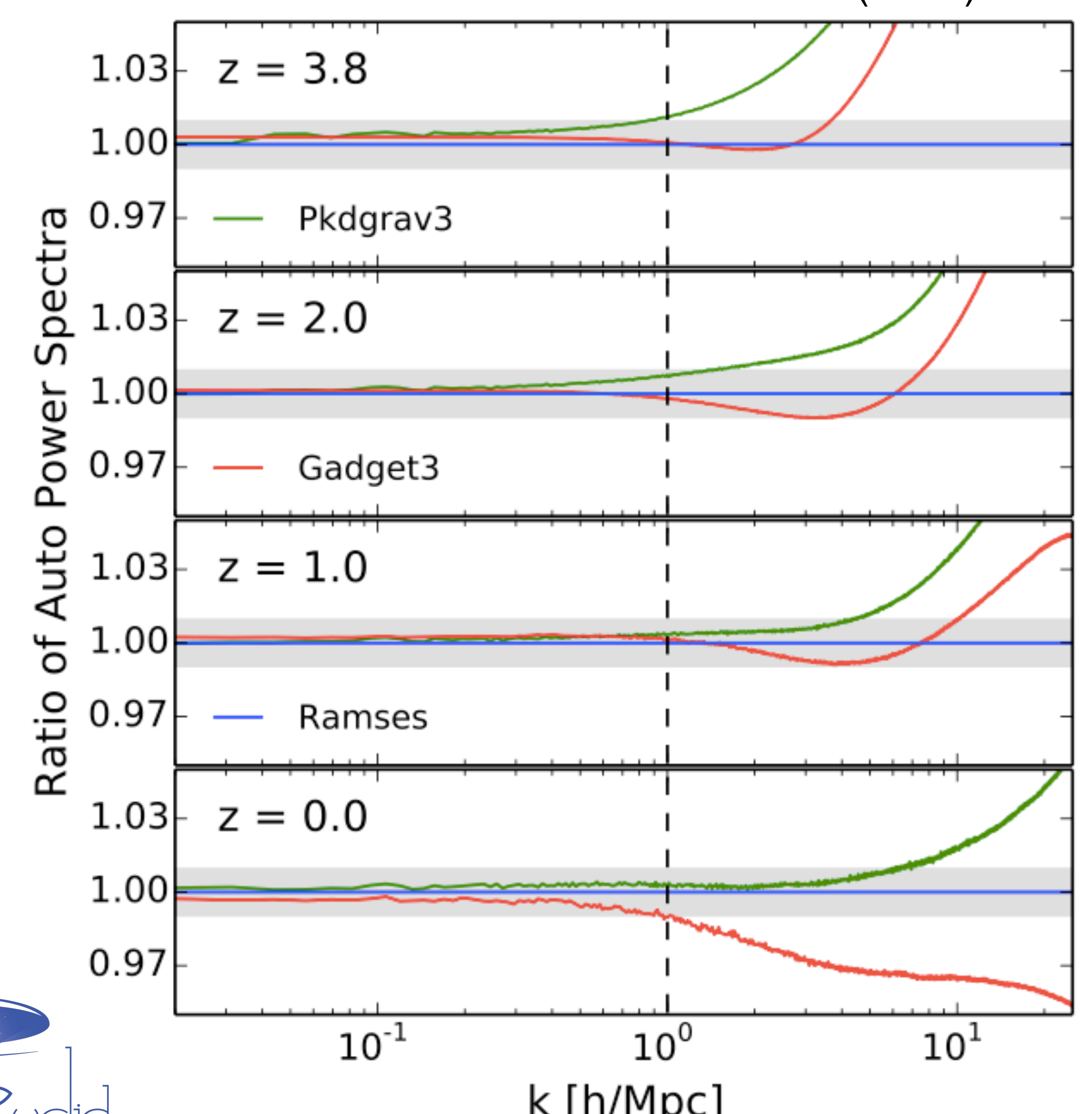
Theory systematics:

- Non-linear matter evolution
  - analytic theory: getting to k~0.3 h/Mpc
  - simulations: numerical convergence
- Reconstruction
- Galaxy bias
- Redshift-space distortions
- Baryon effects
- Massive neutrinos
- Light cone effects (including relativistic effects)

M. Crocce, Euclid Tiger Team report



Schneider et al. (2016)

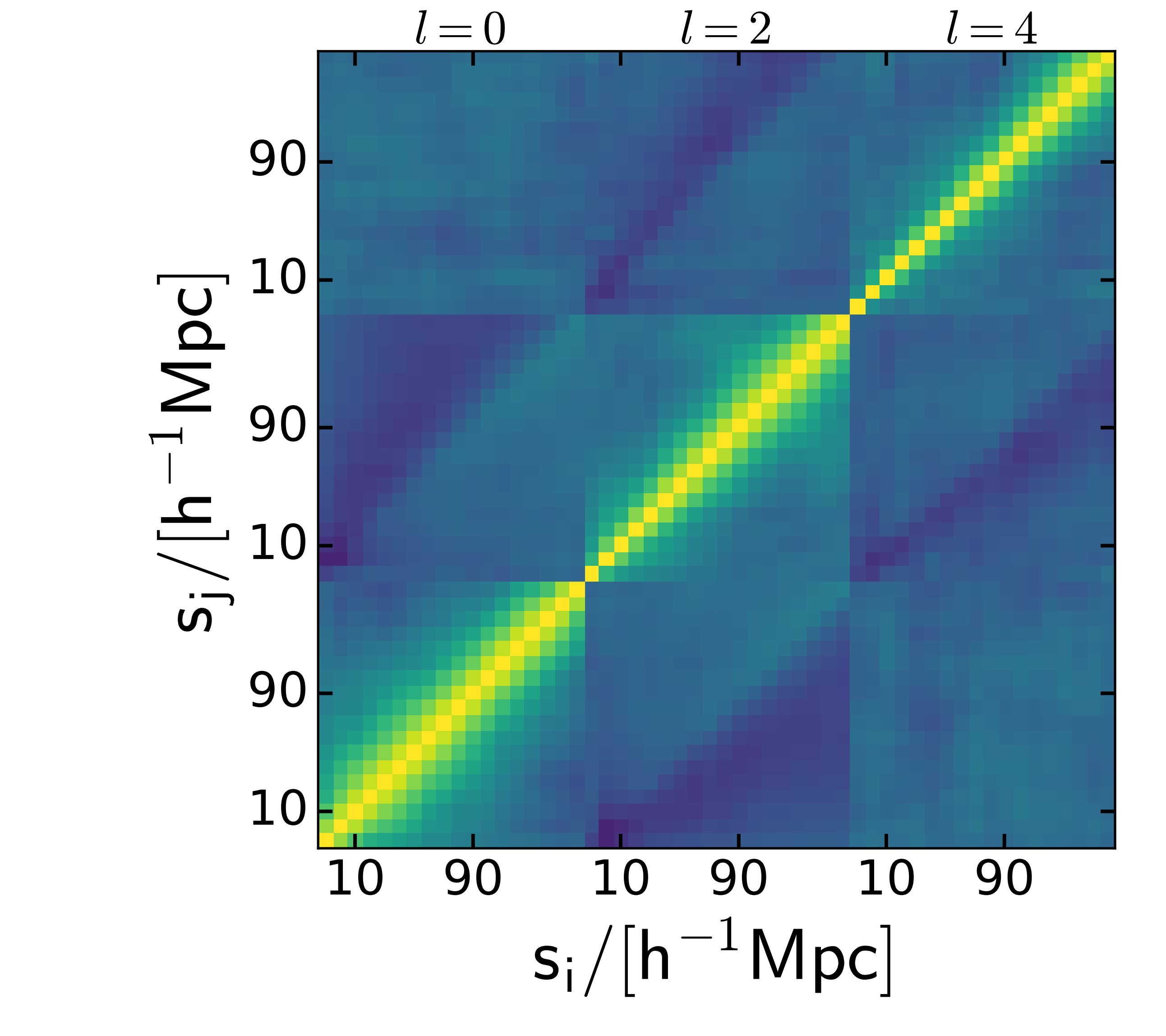


$$-2\ln\mathscr{L}(\boldsymbol{\xi}|\boldsymbol{\theta}) = (\boldsymbol{\xi} - \boldsymbol{\xi}_{\rm theo}(\boldsymbol{\theta}))^t \boldsymbol{\Psi}(\boldsymbol{\xi} - \boldsymbol{\xi}_{\rm theo}(\boldsymbol{\theta}))$$

$$C_{ij} = \frac{1}{N_s - 1}\sum_{k=1}^{N_s}(\xi_i^k - \bar{\xi}_i)(\xi_j^k - \bar{\xi}_j),$$

Likelihood systematics:

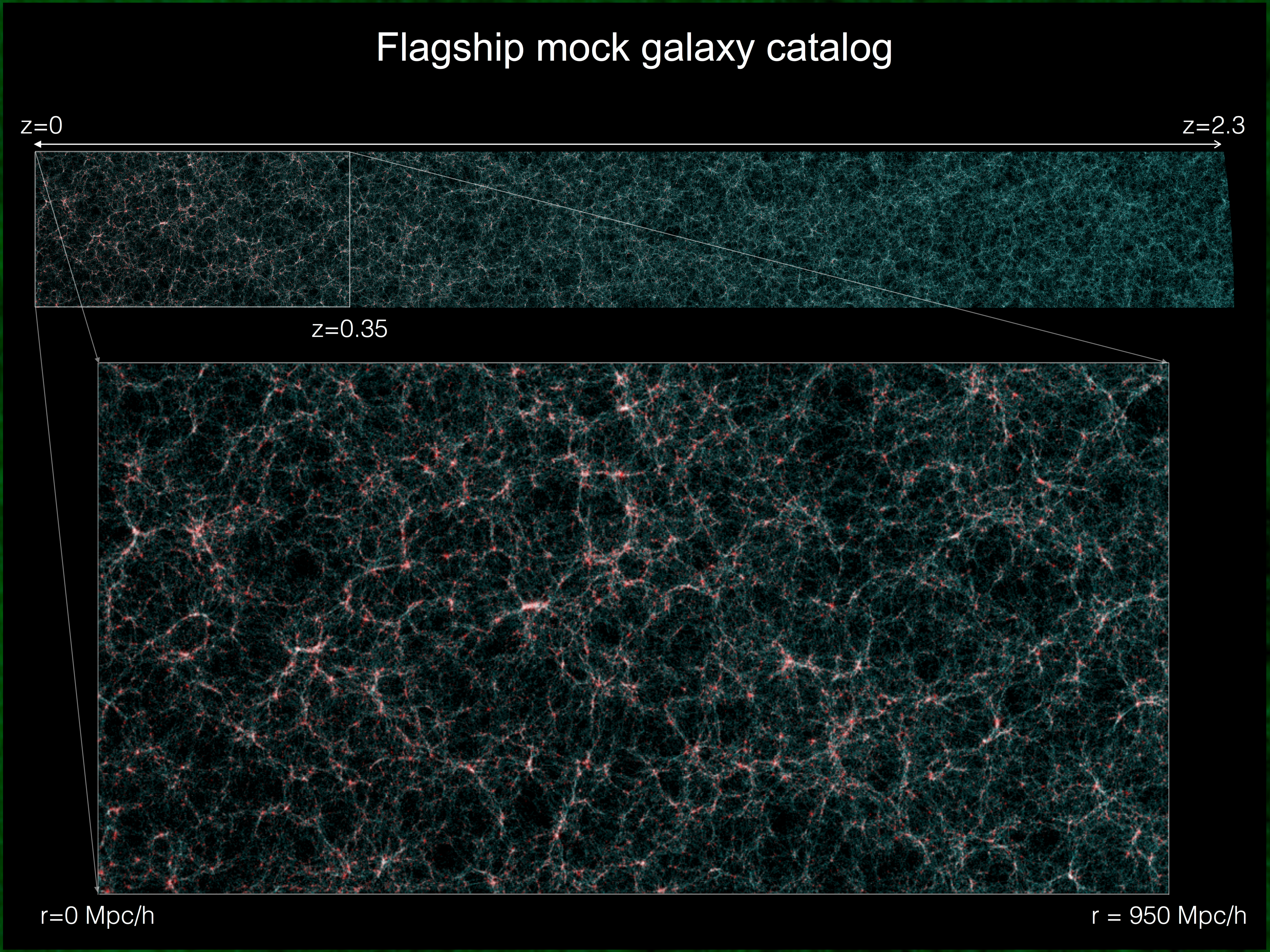
- Propagation of noise in C (covariance matrix)
- Biased estimate of C
- Cosmology dependence of C
- The shape of likelihood
- Combination of results from multiple statistics



Accurate errorbars and control of systematics require thousands of **huge** simulations

Contraints for a single simulation

- sample Baryonic Acoustic Oscillations
  → box size > 1 Gpc/h
- sample a past light cone (almost) without replications
  → box size ~ 4 Gpc/h
- resolve halos hosting the faintest galaxies
  → ~ 10^{11} $M_\odot$ if not smaller
- resolve halo substructure and sample merger trees for Semi-Analytic Models
  → ~ 10^{9} $M_\odot$ if not smaller
  → toward $16{,}000^3$ particles



Flagship 1:

$L=3780$ Mpc/h

$N_p=12600^3$

$M_p=2.4\ 10^9\ M_{sun}/h$

Flagship 2, wide:

$L=3600$ Mpc/h

$N_p=16000^3$

$M_p=1.0\ 10^9\ M_{sun}/h$

Flagship 2, deep:

$L=1000$ Mpc/h

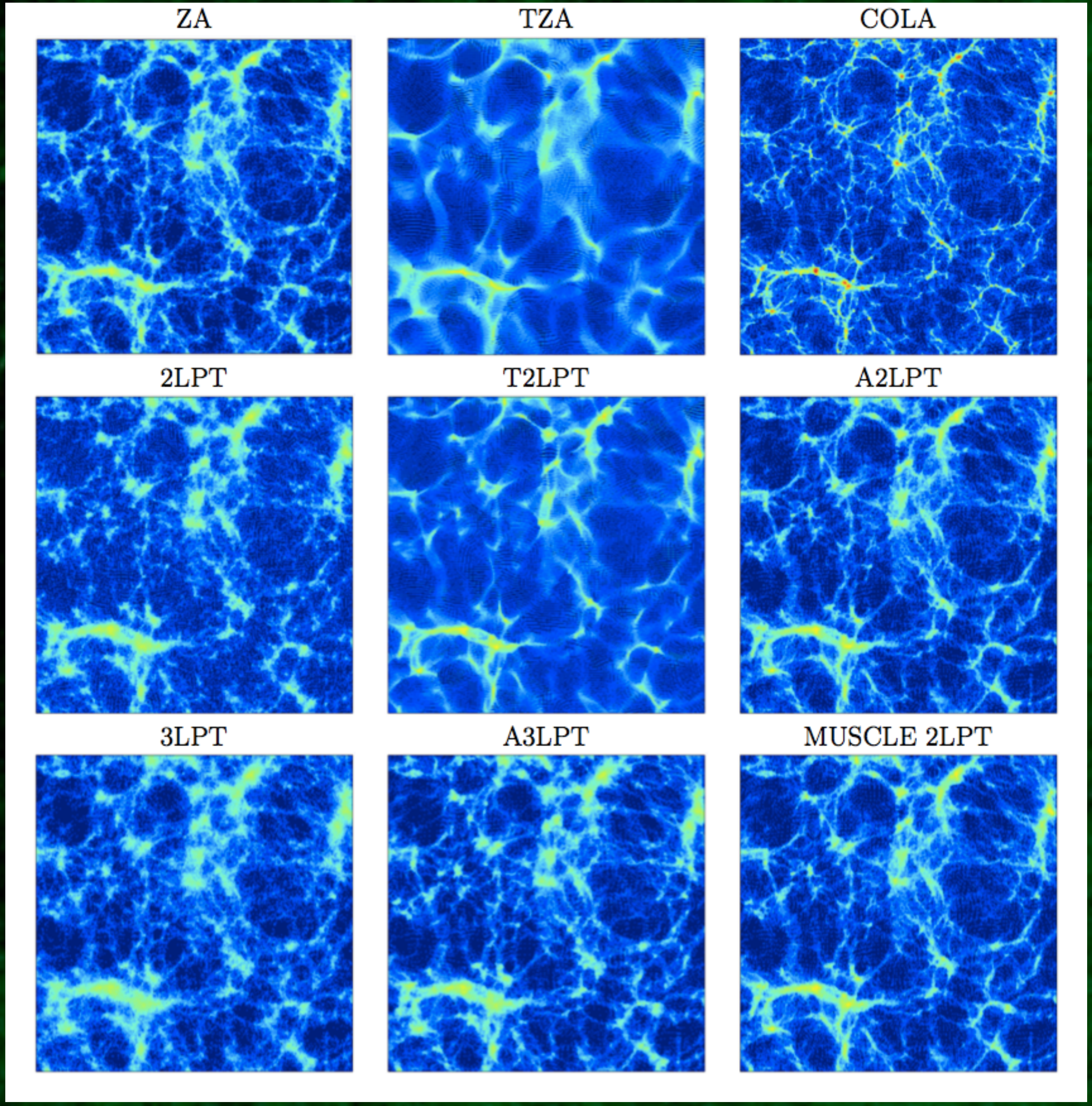
$N_p=9600^3$

$M_p=1.0\ 10^8\ M_{sun}/h$

This has prompted a new burst in the development of approximate methods to predict the clustering of dark matter and dark matter halos

P. M., 2016, *Galaxies* 4, 53

"The search for analytic and semi-analytic approximations to gravitational clustering, more than being pushed out of fashion, has been boosted by precision cosmology."



Two problems to solve:

- displace particles from Lagrangian to final position
- collect particles into halos

“Lagrangian” methods:

generate the ICs of a simulation and apply the method to find halos

- Peak-Patch (Bond & Myers 1996)
- Pinocchio (Monaco+ 2002, 2013, Munari 2016)
- PTHalos (Scoccimarro & Sheth 2002) as in Manera+ (2013)
- Quick PM integrators (White+ 2010, Feng+ 2014)
- COLA (Tassev+ 2013, Howlett+ 2015, Koda+ 2016, Izard+ 2016)

Pros: **predictive**, calibration is cosmology-independent

Cons: memory requirements can be high, resolution limits are as N-body

“Bias-based” methods:

generate a density field and populate it with halos using a bias scheme

- Patchy (Kitaura+ 2014, 2016)
- EZMocks (Chuang+ 2015)
- Halogen (Avila+ 2015)

Pros: very quick, resolution limits are much lighter

Cons: they need to be **calibrated** each time against a big simulation

P. M., 2016, *Galaxies* 4, 53

PINpointing Orbit Crossing-Collapsed HIerarchical Objects

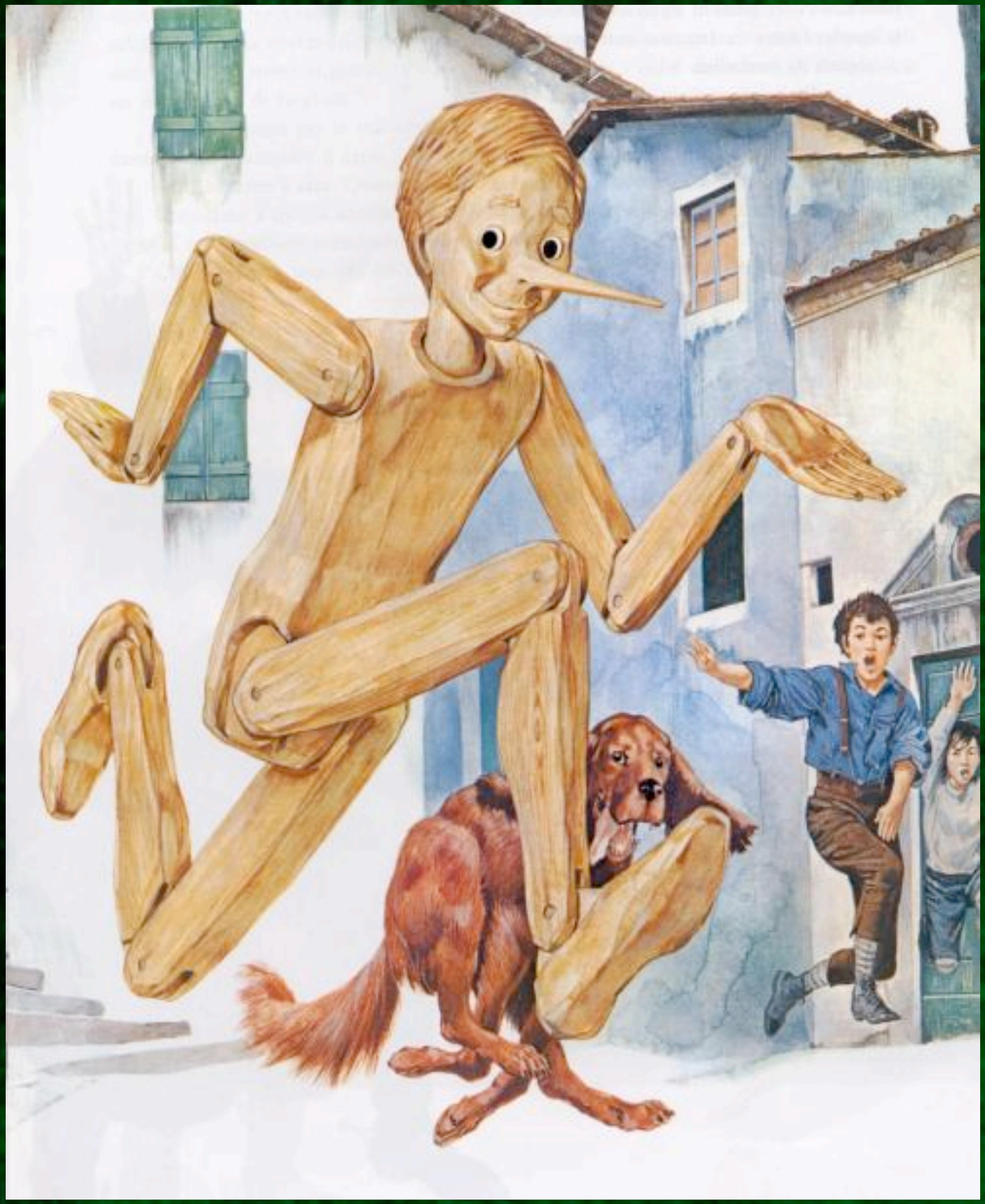
P.M., T. Theuns, G. Taffoni et al., 2002, ApJ, 564, 8

P.M., T. Theuns & G. Taffoni, 2002, MNRAS, 331, 587

G. Taffoni, P.M. & T. Theuns, 2002, MNRAS, 333, 623

P.M., E Sefusatti, S. Borgani, M. Crocce, P. Fosalba, R.K. Sheth, T. Theuns, 2013, MNRAS, 433, 2389

E. Munari, P.M., E. Sefusatti, E. Castorina, F. Mohammad, S. Anselmi, S. Borgani, 2017, MNRAS 465, 4658



He's cheating: he's not N-body, he's way too fast!

http://adlibitum.oats.inaf.it/monaco/pinocchio/
https://github.com/pigimonaco/Pinocchio

Three foundations:

+ Ellipsoidal collapse for mass elements

+ LPT to compute collapse times

+ Excursion sets theory to deal with smoothing scales

Computing the "collapse" (OC) time of a fluid element

- Taylor expansion of the gravitational potential:

$$\phi(\vec{q}_0) \simeq \not{\phi_0} + \phi_{,i}(\vec{q}_0)(\vec{q}-\vec{q}_0)_i + \phi_{,ij}(\vec{q}_0)(\vec{q}-\vec{q}_0)_{ij}$$

bulk flow — second-order term

- => evolution of a homogeneous ellipsoid
  - numerical solution, or
  - solution with 3LPT up to **ORBIT CROSSING**

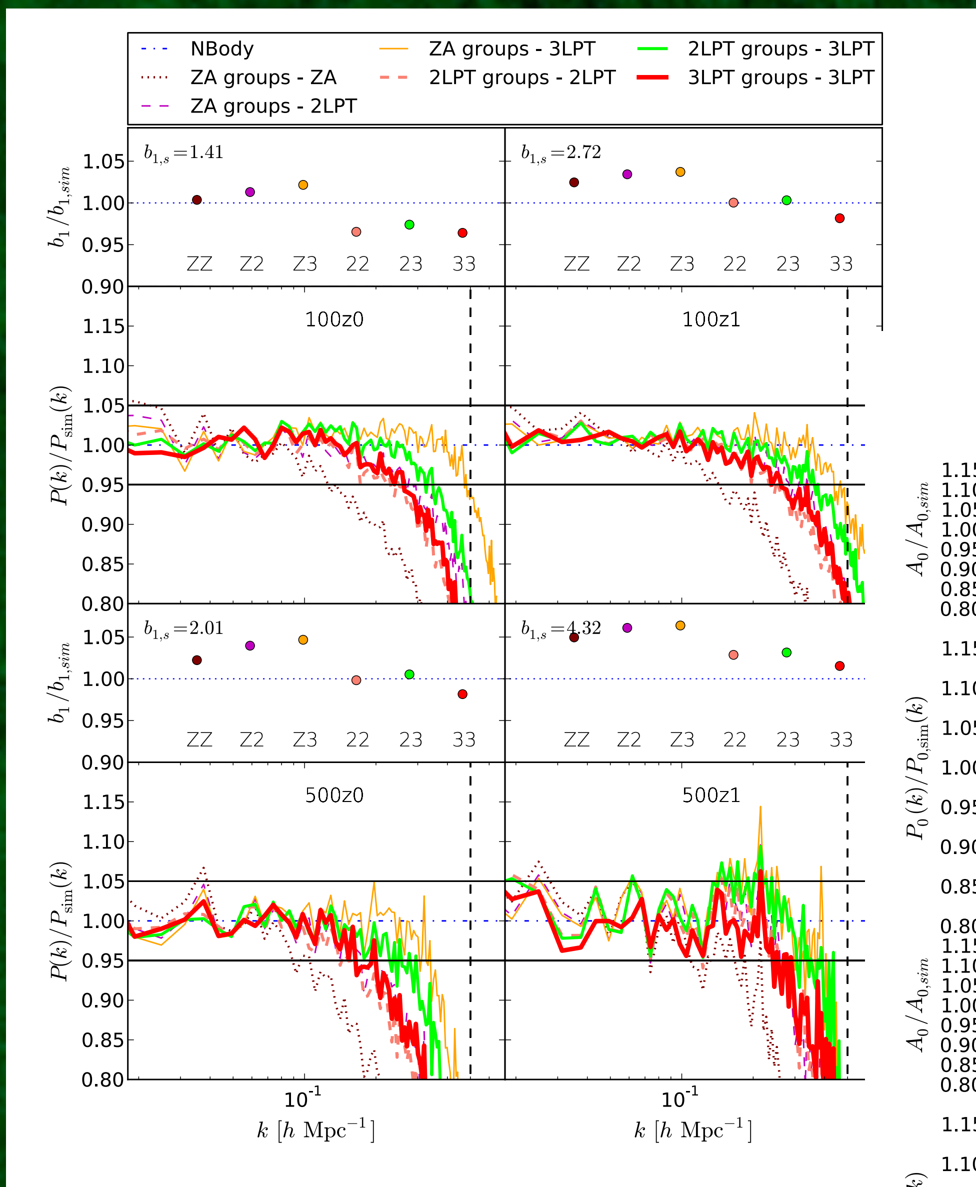
  + correction for quasi-spherical cases

(P.M. 1995, 1997a)

and an algorithm to mimic hierarchical assembly of particles into halos

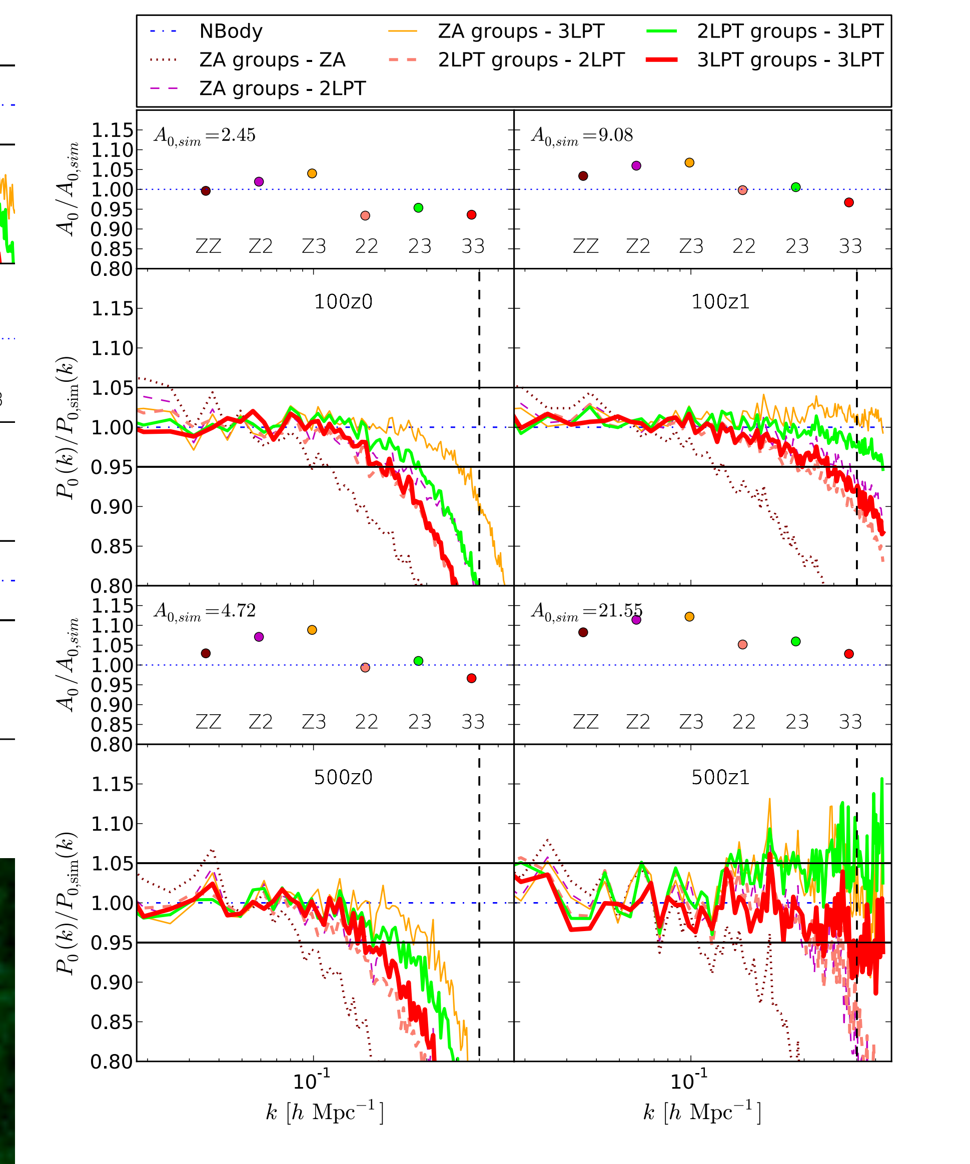
Clustering: P(k), pinocchio vs simulation, same phases

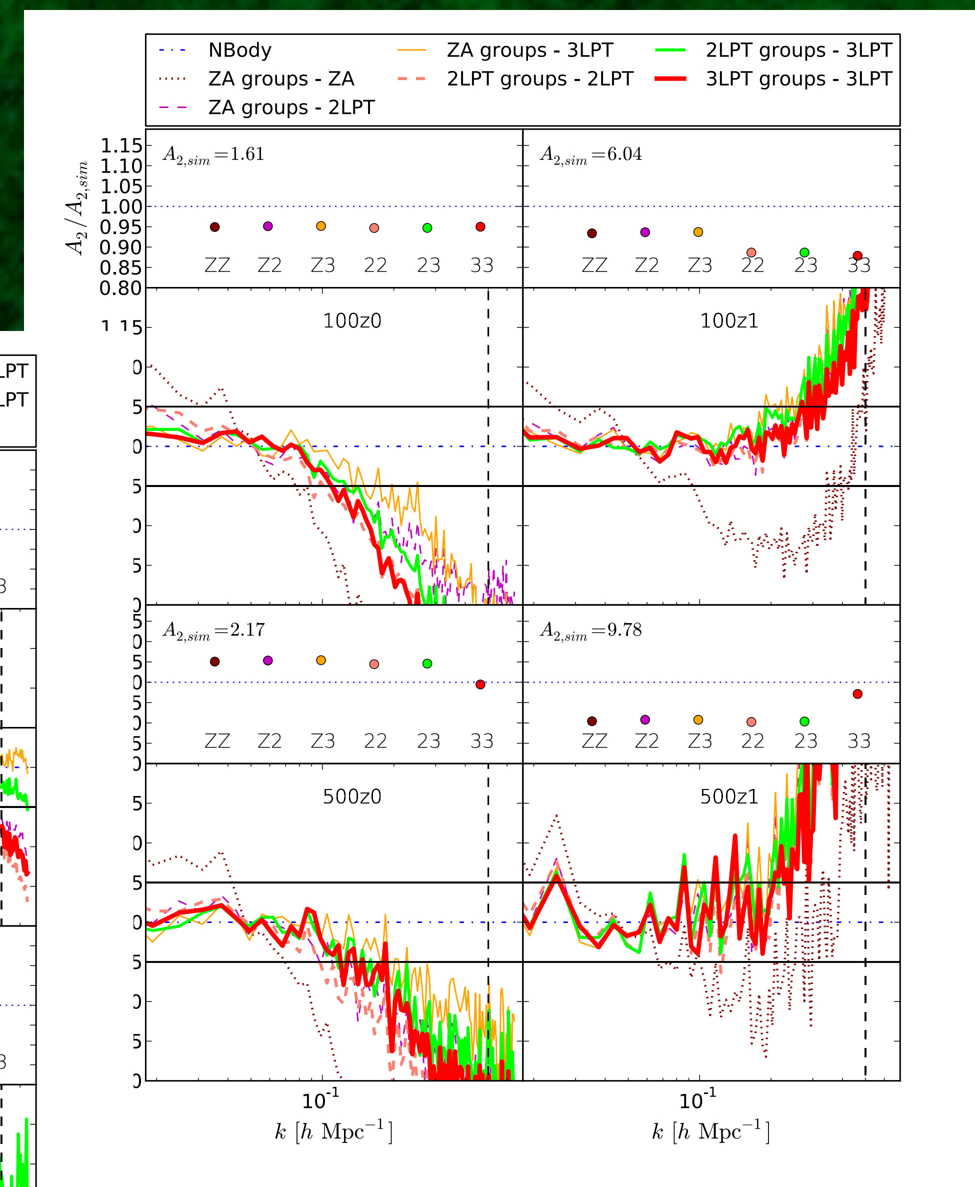
Munari et al. (2017)



real space

redshift space, monopole

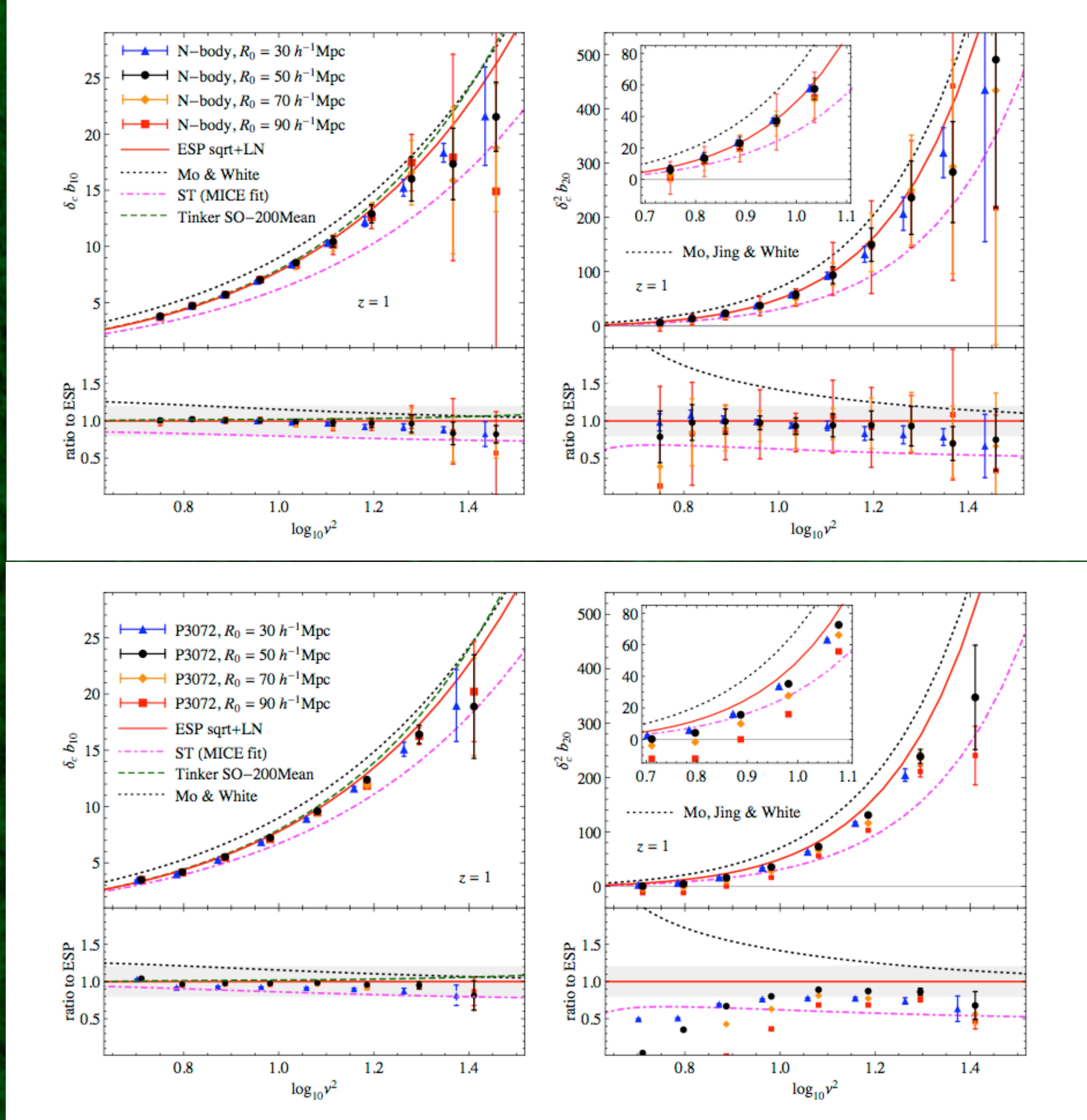




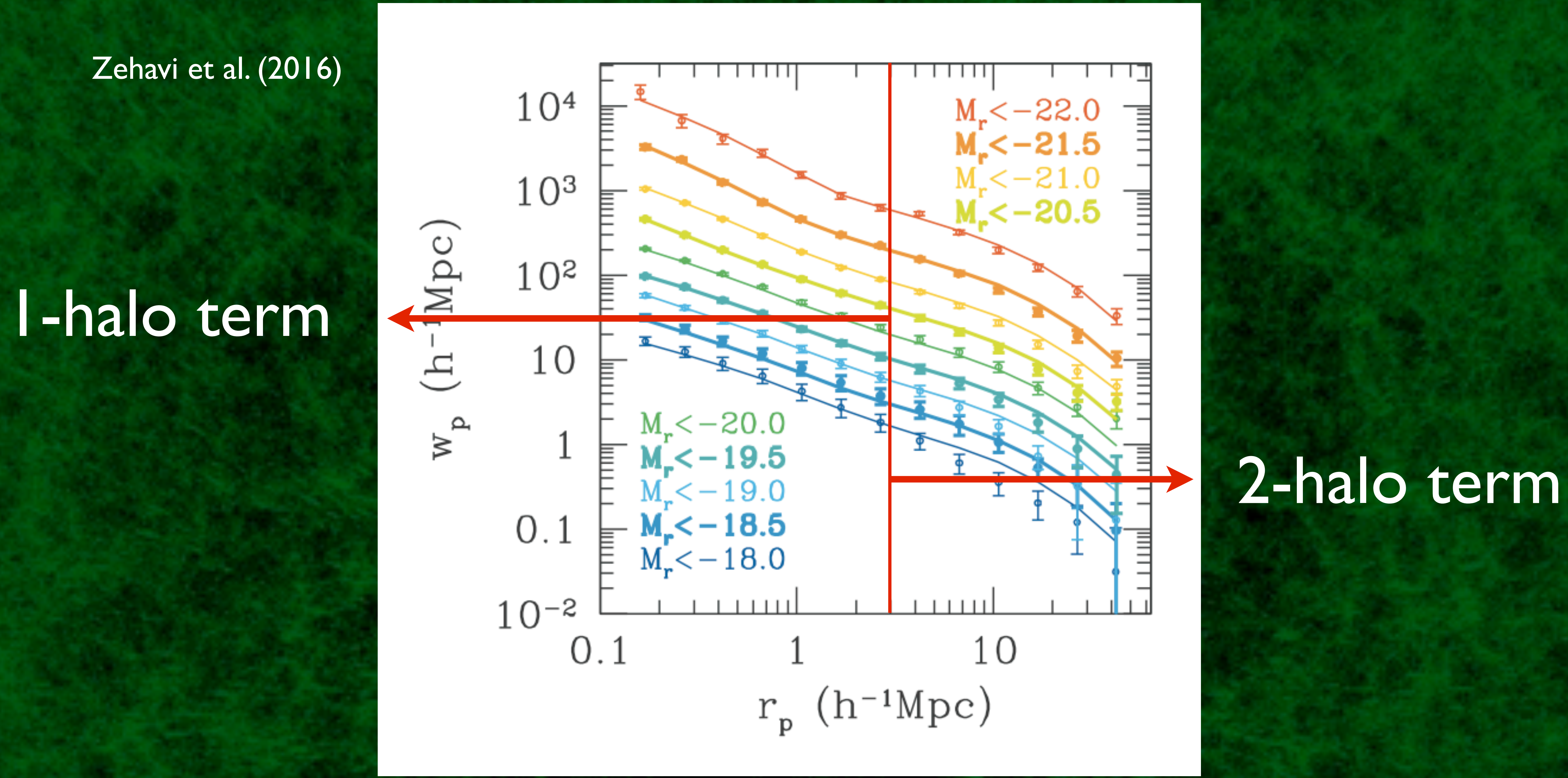
redshift space, quadrupole

Bias is a prediction

Paranjape, Sefusatti, Chan, Desjacques, P.M. & Sheth (2014)



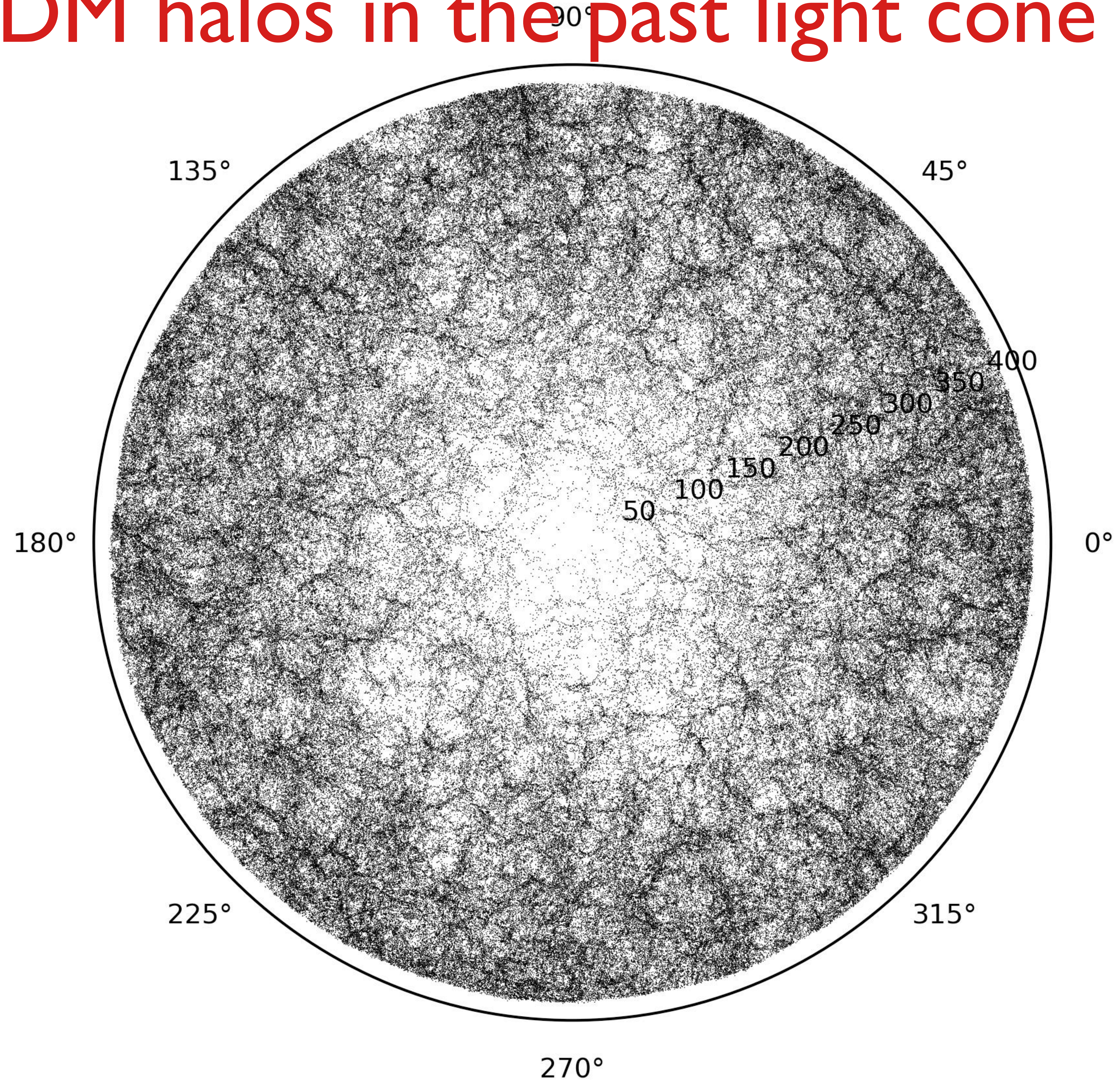
Zehavi et al. (2016)



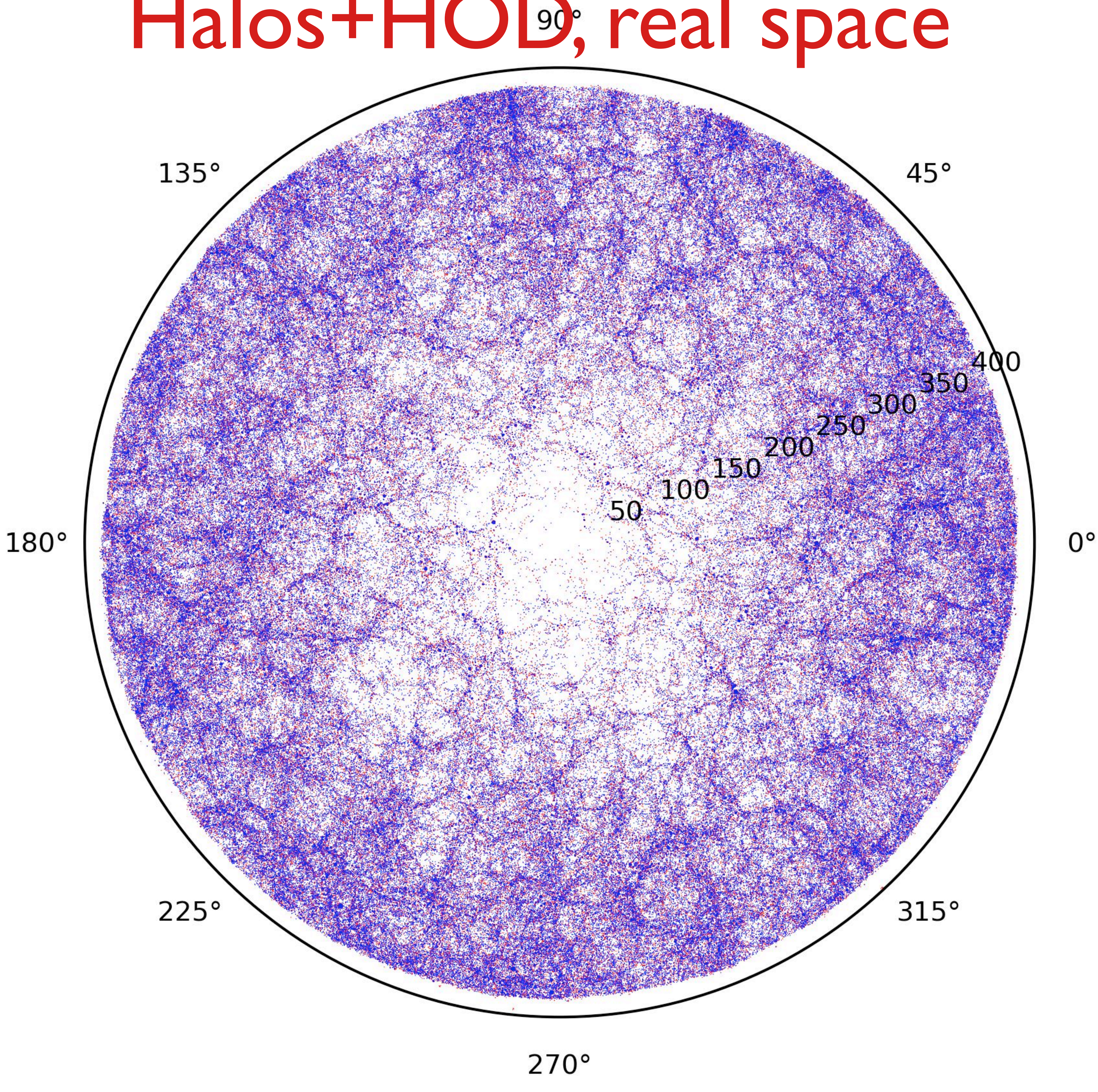
To populate dark matter halos with galaxies:

- Halo Occupation Distribution (HOD) models
- Sub-Halo Abundance Matching (SHAM) models
- Semi-Analytic Models (SAM) of galaxy formation

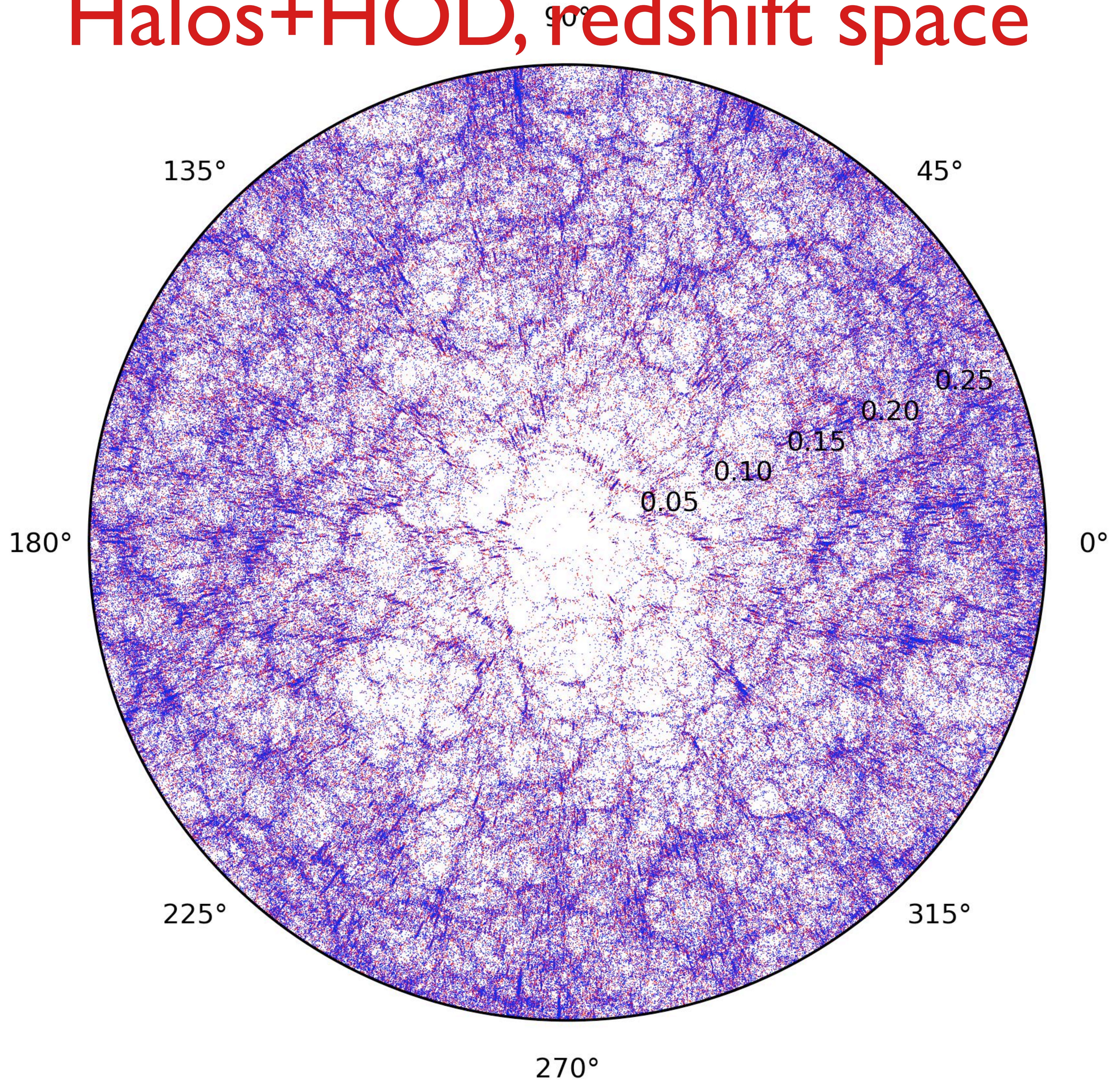
DM halos in the past light cone



Halos+HOD, real space



Halos+HOD, redshift space

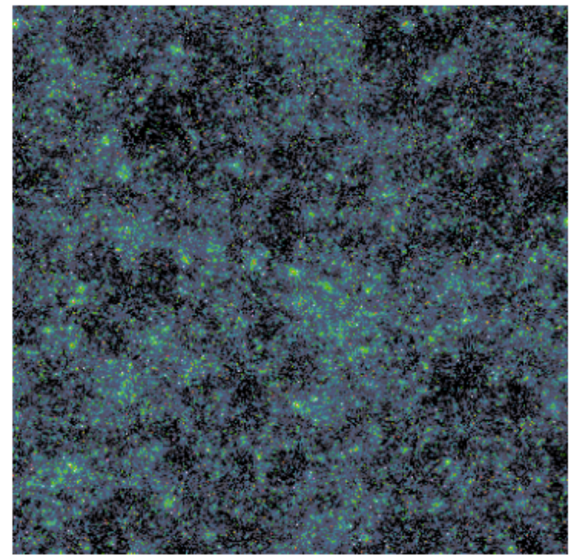


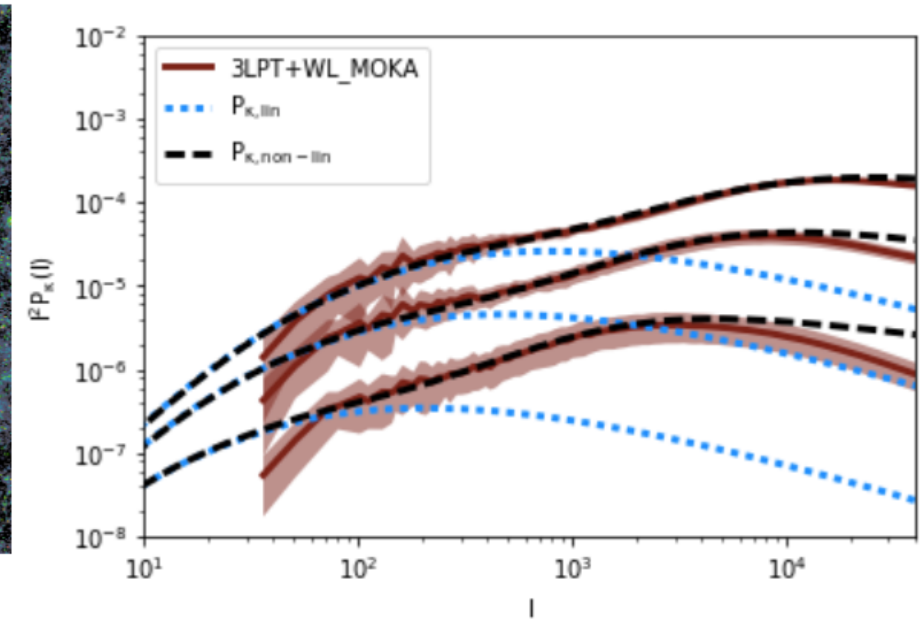
Development

Predicting lensing
(with Carlo Giocoli,
Tiago Castro)

Halo Convergence Power Spectra
64 light-cone realisations

modelling of the matter not resolved in
haloes as in Giocoli+17





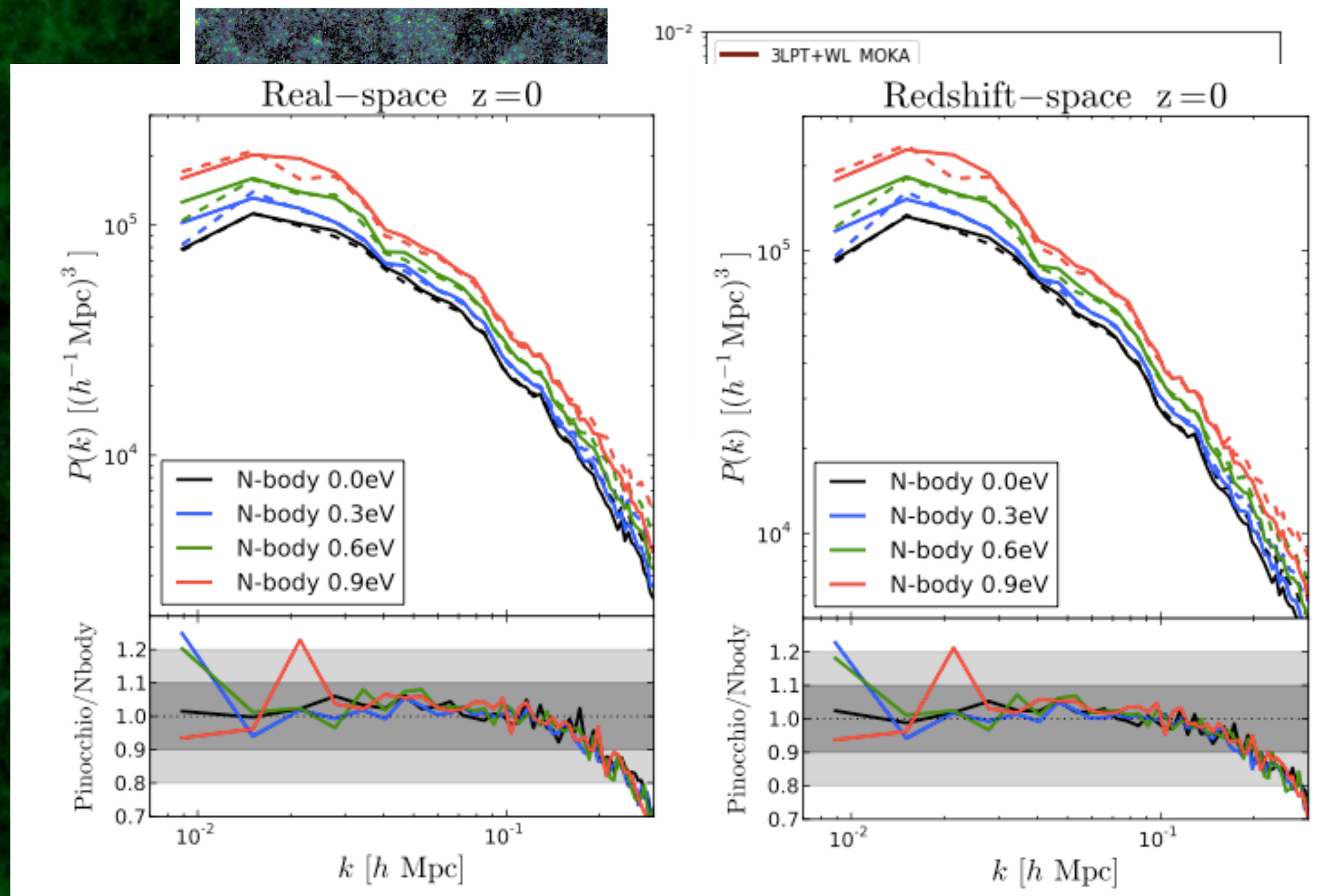
Development

Predicting lensing
(with Carlo Giocoli, Tiago Castro)

Halo Convergence Power Spectra
64 light-cone realisations

modelling of the matter not resolved in haloes as in Giocoli+17

Beyond ΛCDM: neutrinos
(with Luca A. Rizzo, Paco Villaescusa)

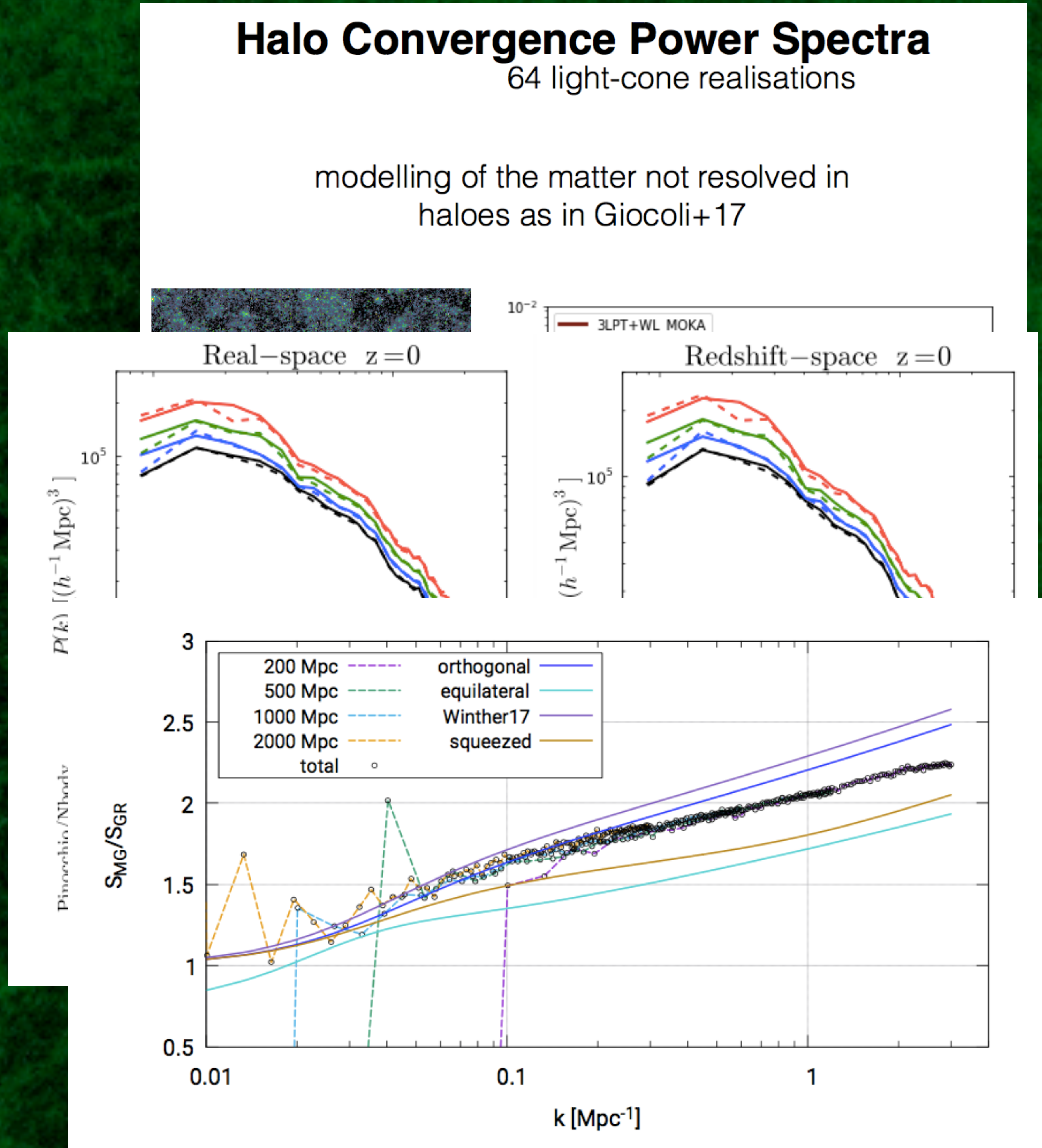


Development

Predicting lensing
(with Carlo Giocoli, Tiago Castro)

Beyond ΛCDM: neutrinos
(with Luca A. Rizzo, Paco Villaescusa)

Beyond ΛCDM: f(R) gravity
(with Chiara Moretti, Simone Mozzon, Marco Baldi, EFTCamb developers)



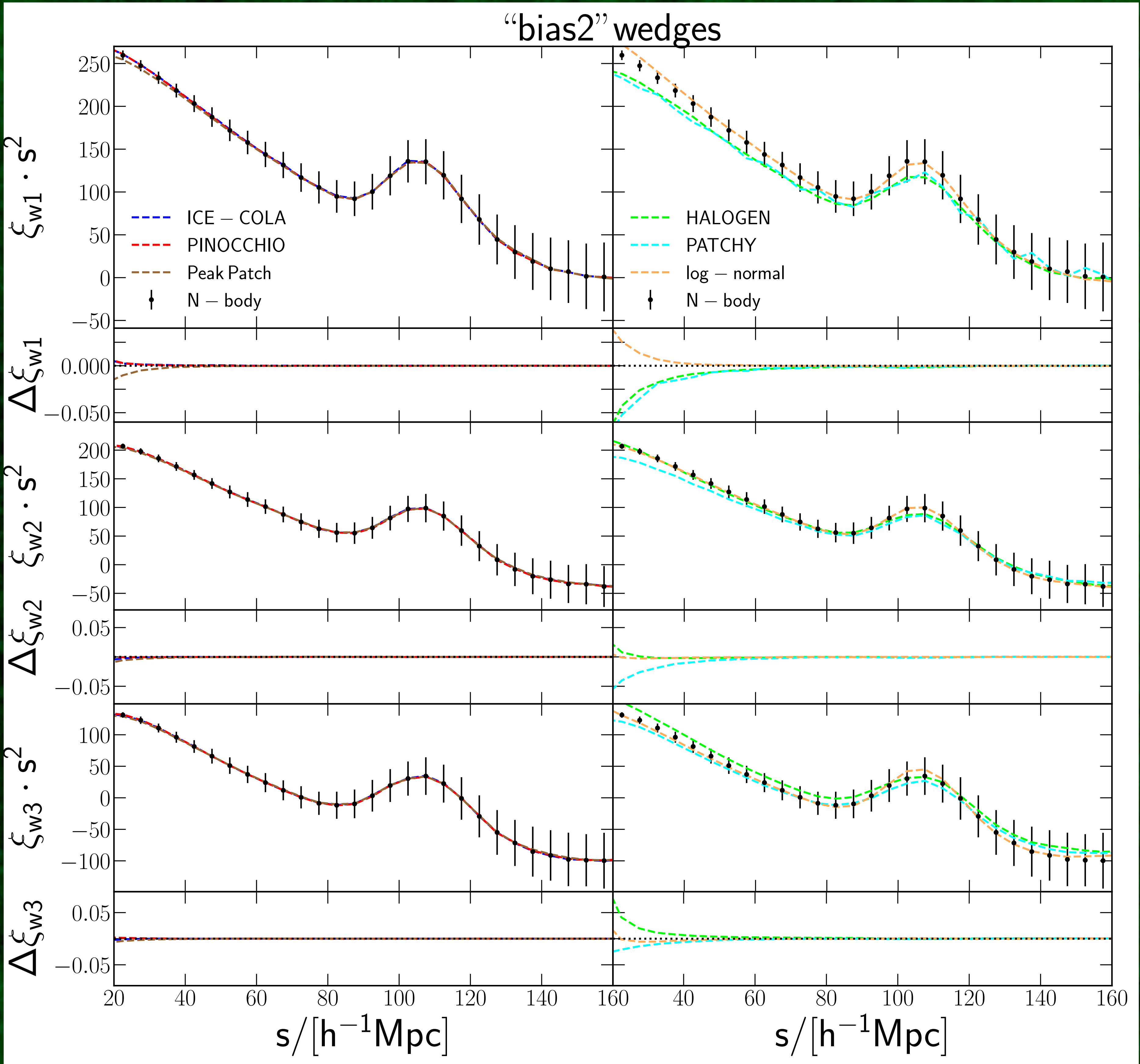
Question:
How accurate do these methods need to be?

+ Averages can be computed from simulations and theory, these methods are needed for covariances.

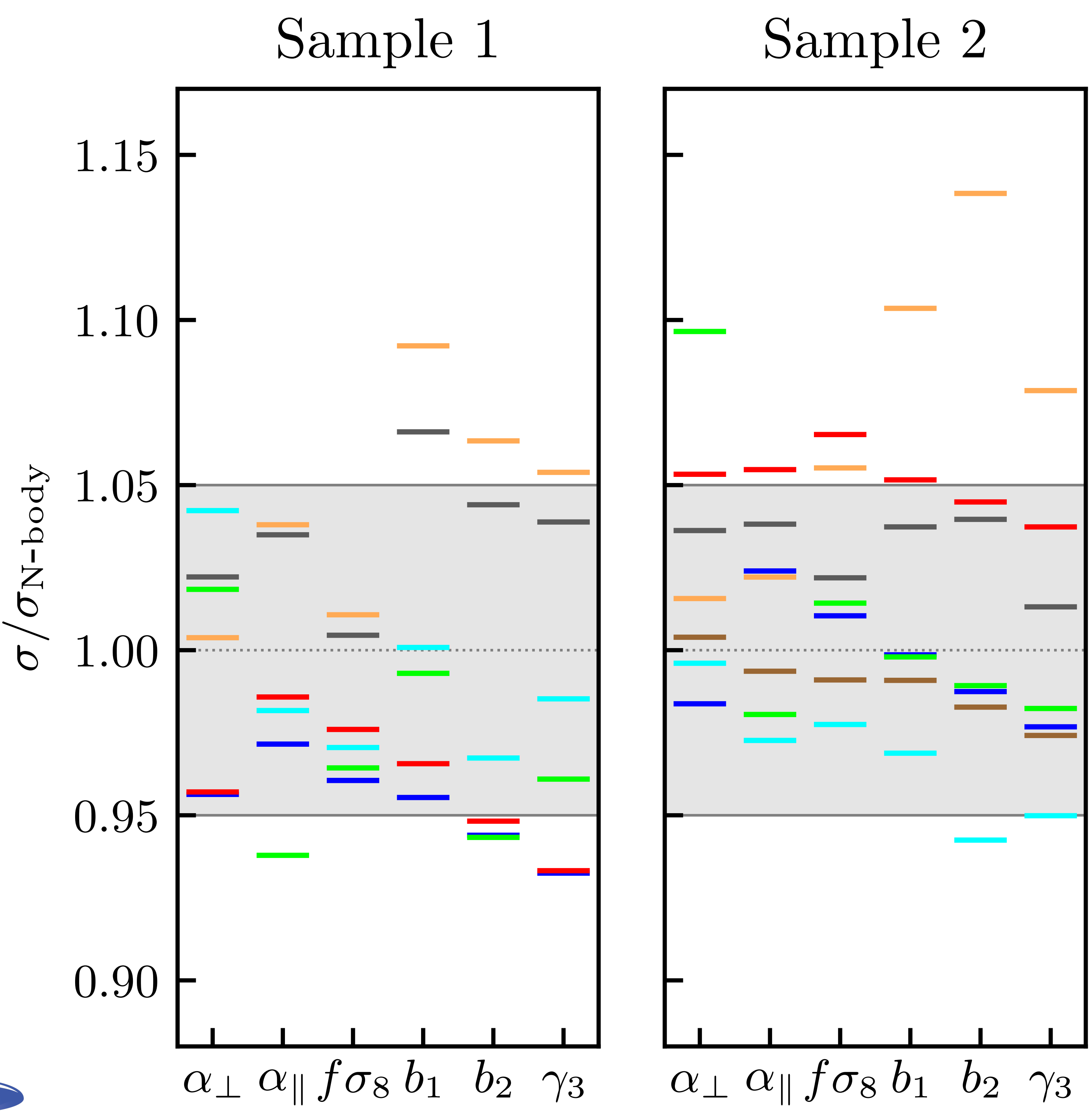
+ Small scales are dominated by the 1-halo term, dominated by galaxy formation effects.

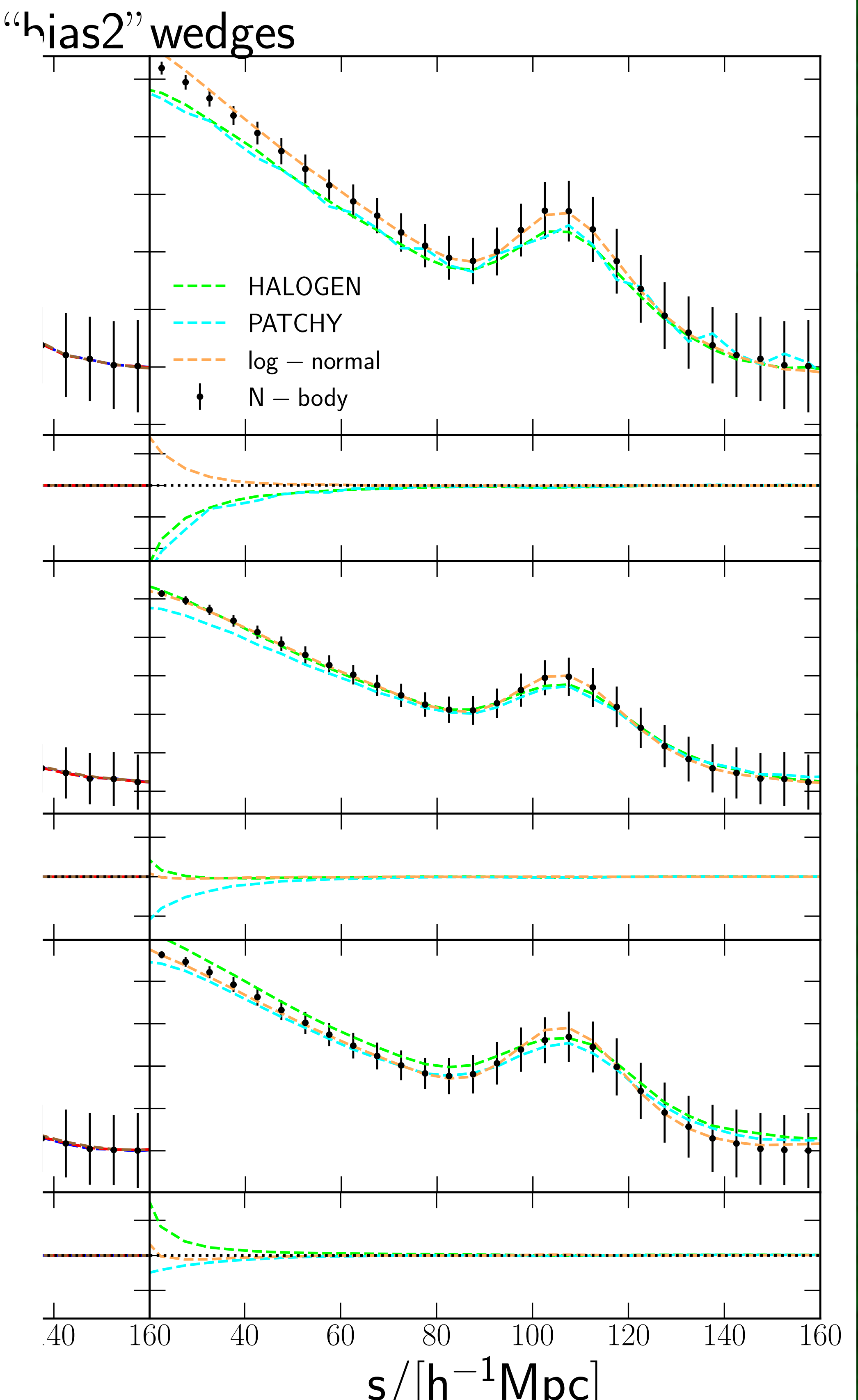
+ We want errorbars on cosmological parameters to be accurate by <10%

Two-point correlation function
Lippich et al. (2019)



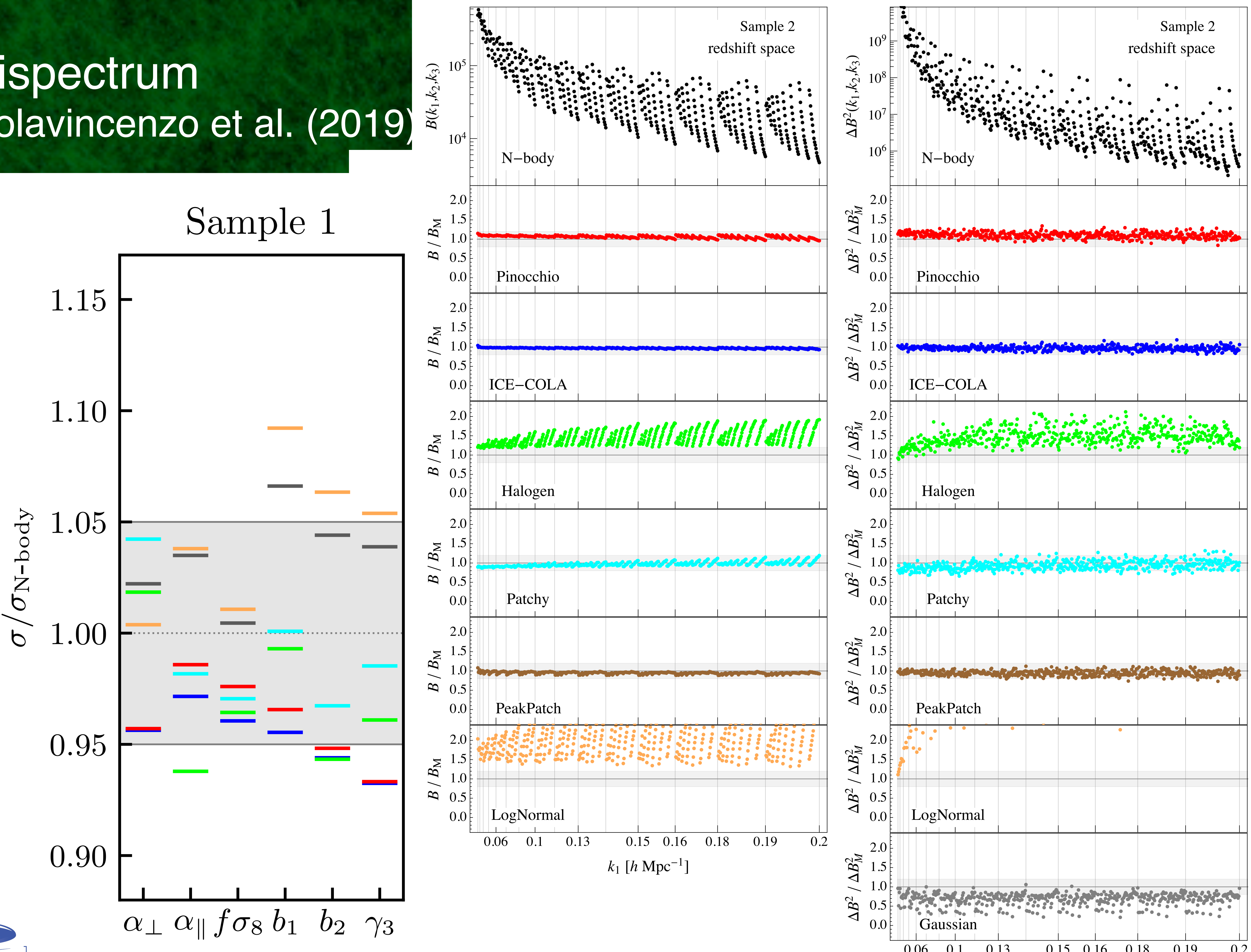
Power spectrum
Blot et al. (2019)





Bispectrum

Colavincenzo et al. (2019)



...tens of thousands of mock catalogues

Project	Number	Cosmology	Box size (Gpc/h)	Parts.	Part. mass (Msun/h)	PLC semi-aperture	PLC redshift range
Covariances	10,000	Minerva	1.5	1000^3	2.7e11	30 deg (1000 only)	0.9 - 1.8 (1000 only)
Systematics	20	Minerva	3.2	4096^3	3.8e10	60 deg	0.0 - 2.5
Systematics	10	Flagship	3.2	4096^3	4.2e10	60 deg	0.0 - 2.5
Bispectrum	300	Flagship	1.2	2160^3	1.5e10	20 deg	0.0 - 2.0
Clusters	100	Flagship	3.87	2160^3	4.9e11	60 deg	0.0 - 2.5

13

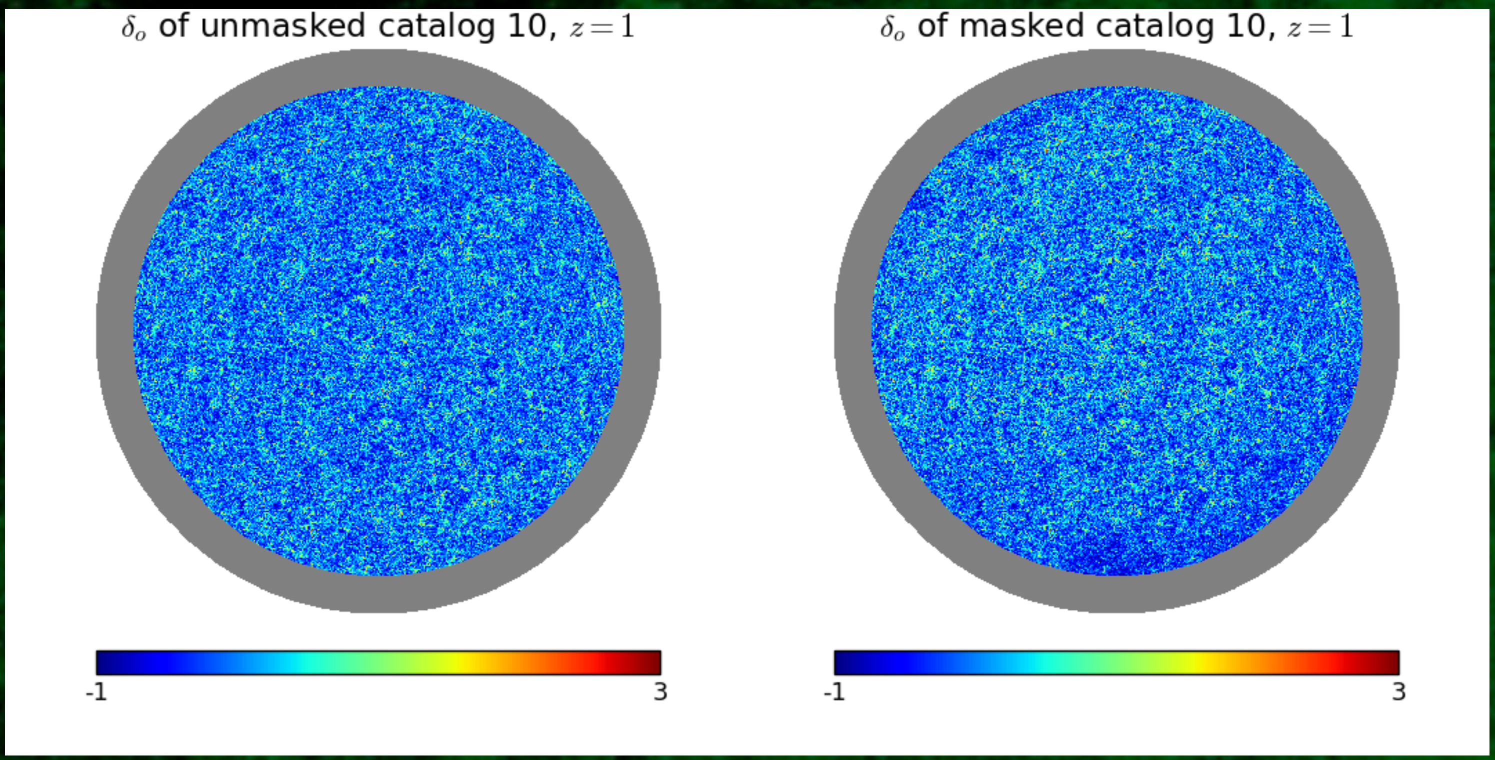
3.2 Gpc/h box sampled with 4096^3 particles

10^{12} M_{sun} well resolved

light cone since z=2.5

120 deg aperture

1/4 of the sky



Summary

- Control of systematics is a big challenge for near future cosmology surveys
- Thousands of simulated surveys are required to reduce noise in the covariance matrix, etc...
- Approximate methods are a precious tool for producing mocks, due to their speed
- Comparison with N-body simulations shows that some methods induce an acceptable perturbation in parameter confidence ellipsoids
- Predictivity pays!
- Lot of development to build complete mock machines