Signatures of the origin of the Universe from its emptiness

Leibniz Institute for Astrophysics in Potsdam

Francisco-Shu Kitaura

Karl-Schwarzschild fellow

INPA seminar 03/18/2016

How can we generate accurate mock galaxy catalogs for present large-scale structure surveys? (BOSS,eBOSS,J-PAS,DESI,EUCLID, 4MOST,WEAVE,DES,LSST,…)

4MOST: 4m VISTA 2021-2025
spectroscopic survey
~14k sq deg
15-20 M LRGs+eLGs+Quasars
z<1.4

JPAS: JST/T250 2016-2020
56 colors
1.2 Giga pixel camera
~8k sq deg
~200 M objects
z<1.4

Perform large N-body simulations:
MillenniumXXL, BigMultidark, MICE, DEUS, HORIZON
needed to produce reference catalogs.

However,
this effort can only be done a couple of times,
but not thousands of times!

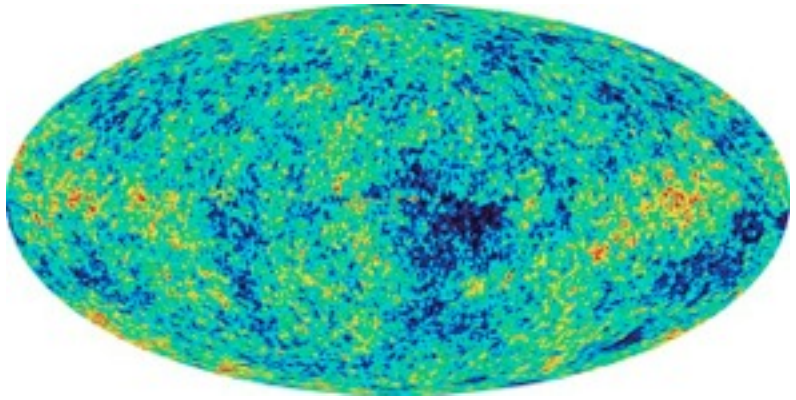
The ingredients

1) We start from the primordial fluctuations

   i) Gaussian

   ii) or non-Gaussian

Scoccimarro, Hui, Manera, Chang Chang 2012



Planck -8.9 < fNL < 14.3 (2 sigma)

(Suyama et al., 2013)

WMAP ~300 000 years after the Big Bang

2) We simulate structure formation

Why not simulate everything?

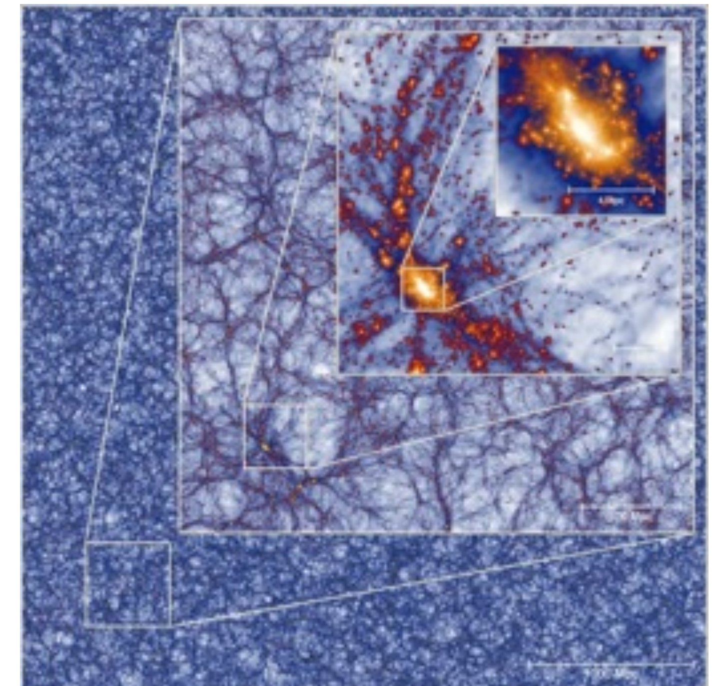
Full gravity and hydrodynamical solver in an expanding background

Vogelsberger et al 2014

about (70 Mpc/h)^3 !! too small volumes for our purposes!!

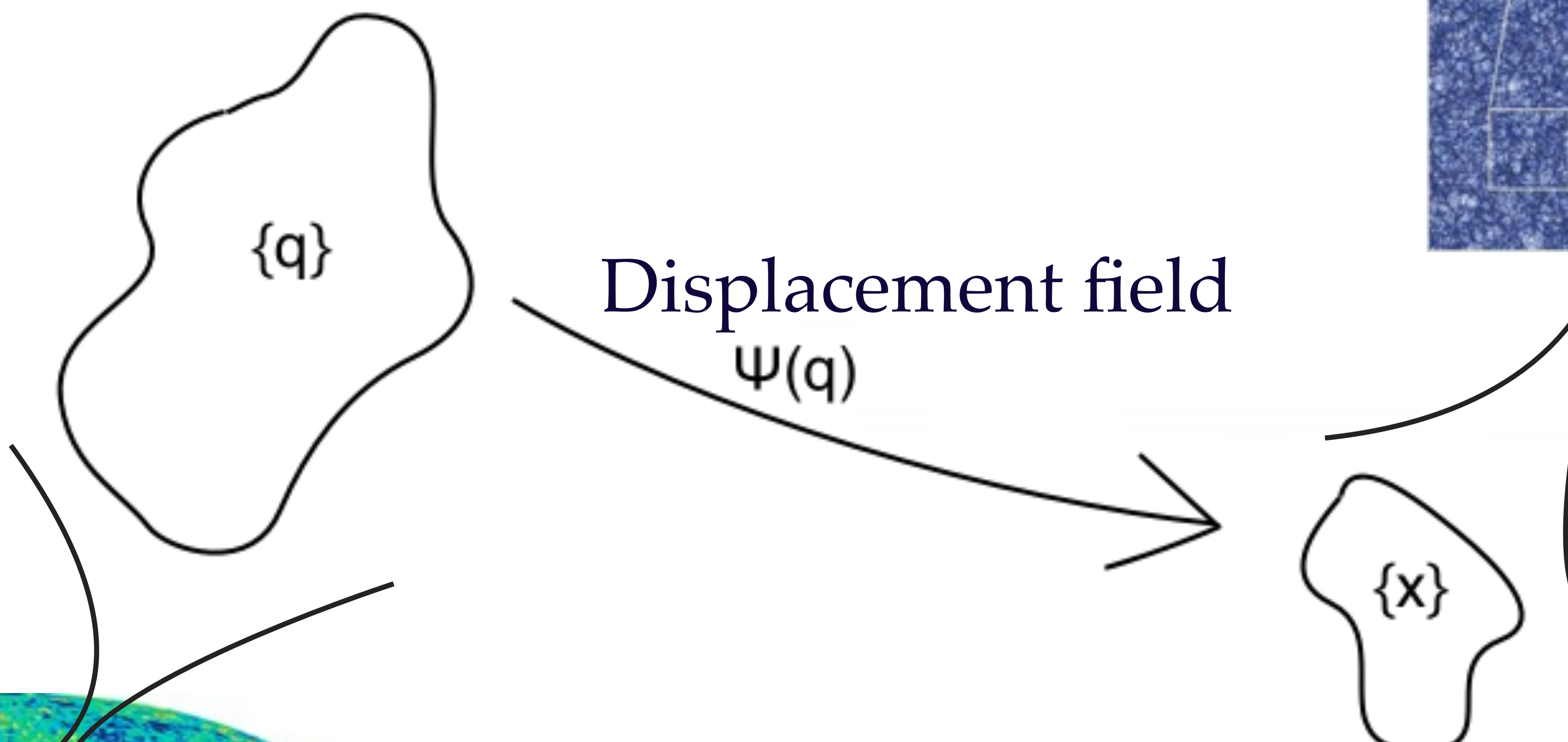
2) We simulate structure formation

Full gravity solver dark matter only in an expanding background, solution to the Vlasov equations
Monte Carlo approach: N-body code



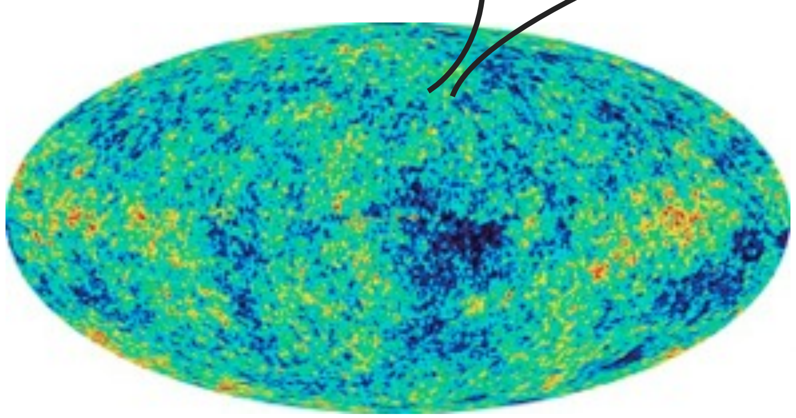
Millennium XXL
Angulo et al 2012
~(3 Gpc/h)^3

Lagrangian coordinates



Displacement field

Eulerian coordinates

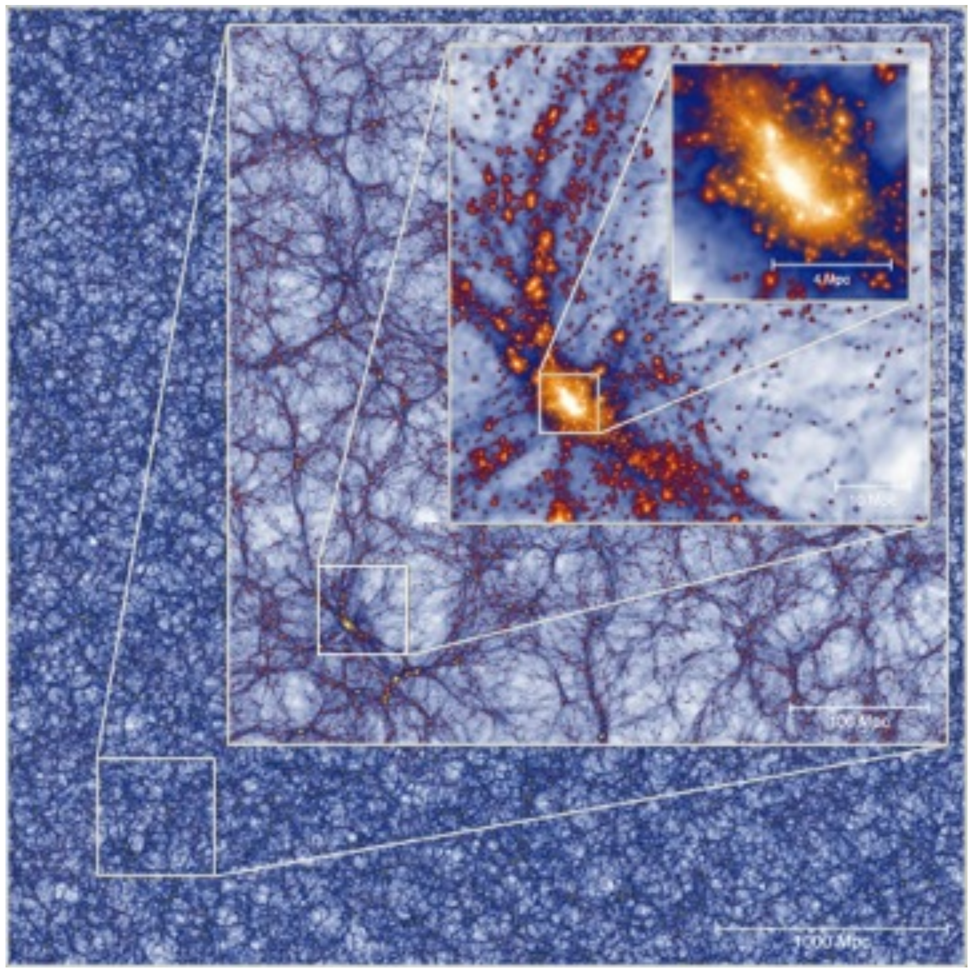


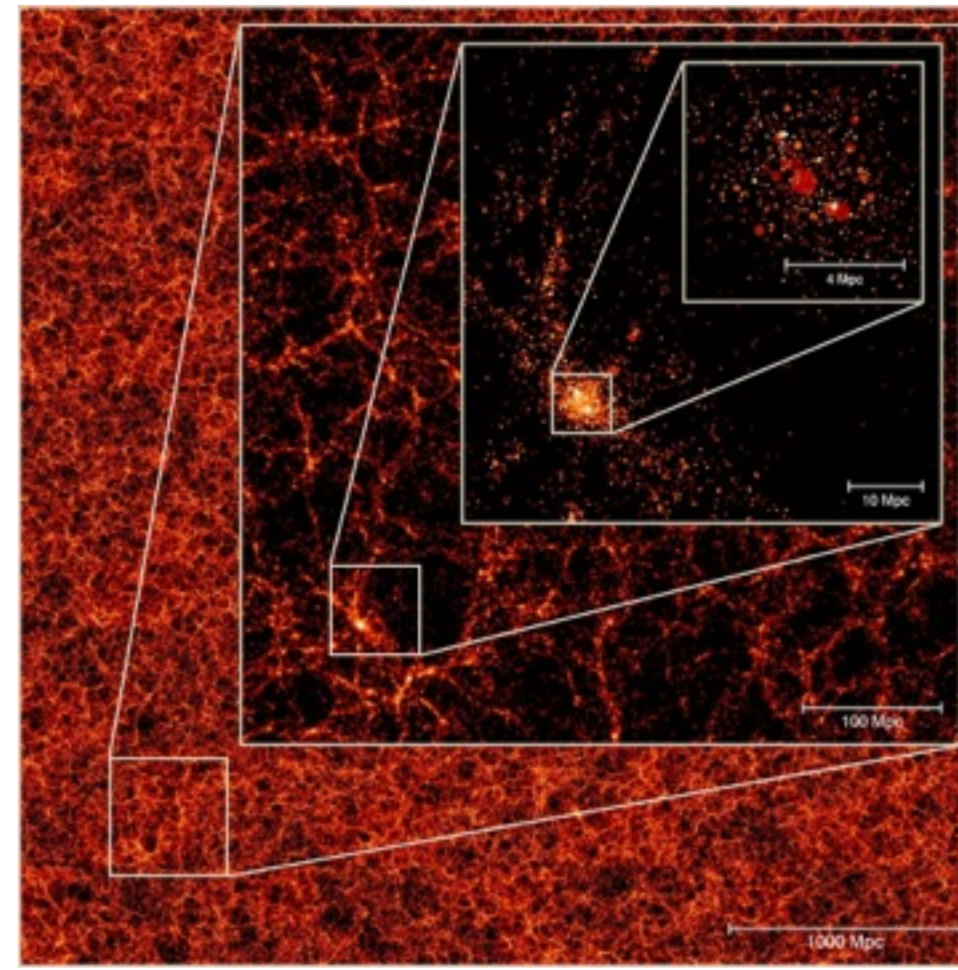
WMAP

~13.7 billion years after the Big Bang

3) We run a halo finder

Knebe et al. 2011





Millennium XXL
Angulo et al 2012

4) We produce a galaxy catalog based on either:

- semi-analytic model

  Cole 2000; Hatton 2003; Croton 2006; Bower 2006;
  Monaco 2007; Benson 2010; deLucia & Blaizot 2007; Baugh 2006

- halo occupation distribution

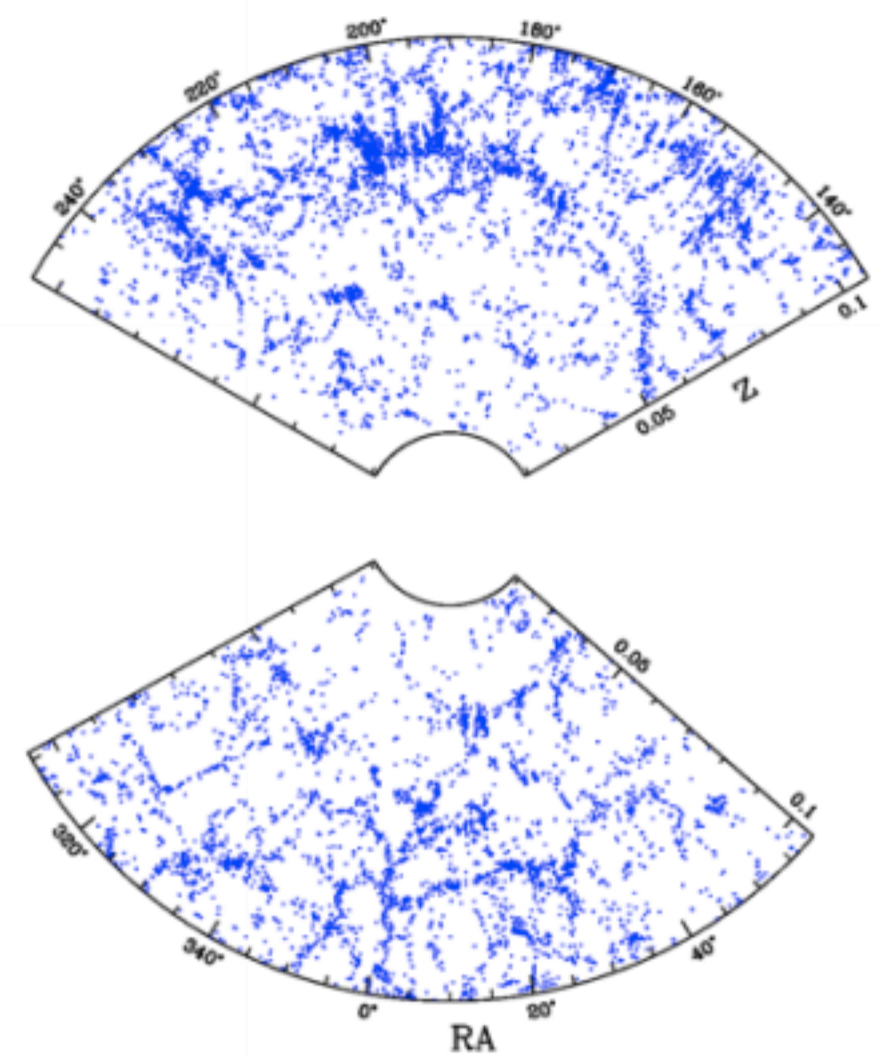
  Seljak 2000; Berlind & Weinberg 2002; Zheng 2005; Zehavi et al 2011

- abundance matching

  Kratsov 04, Tasitsiomi 04, Vale&Ostriker04, Conroy 06,
  Behroozi10, Trujillo-Gomez 11, Nuza 12

5) We apply observational effects

- selection function (magnitude cut)
- survey mask
- redshift-space distortions



Marc Manera SDSS

Perform large N-body simulations:
MillenniumXXL, BigMultidark, MICE, DEUS, HORIZON
needed to produce reference catalogs.

However,
this effort can only be done a couple of times,
but not thousands of times!

we need to scan different cosmologies,
different seeds (cosmic variance)

to estimate error bars/covariance matrices
to measurements based on our unique Universe;

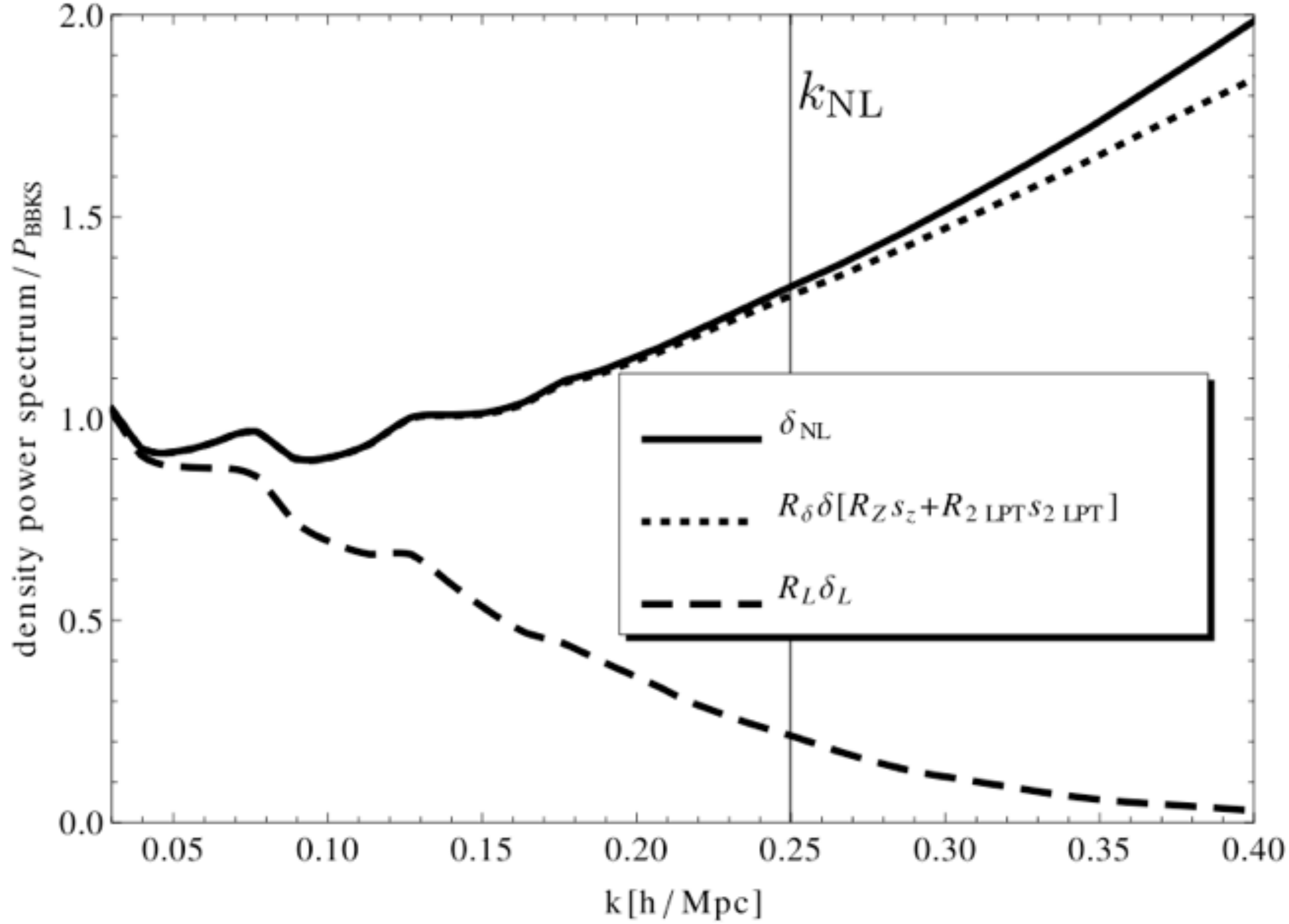
to test analysis tools with different realisations
in a statistical relevant way

What is the bottle-neck of the computations?

What is the bottle-neck of the computations?

1. the gravity solver

Perturbative approaches to model BAOs



$$\delta(\boldsymbol{k}) = R_\delta(k)\delta_1(\boldsymbol{k})$$

$$R_\delta(k,\eta) = \exp(0.58d)$$

Tassev & Zaldarriaga 2012

deviation in the power spectrum due to perturbative approach!

let us use approximate gravity solvers!

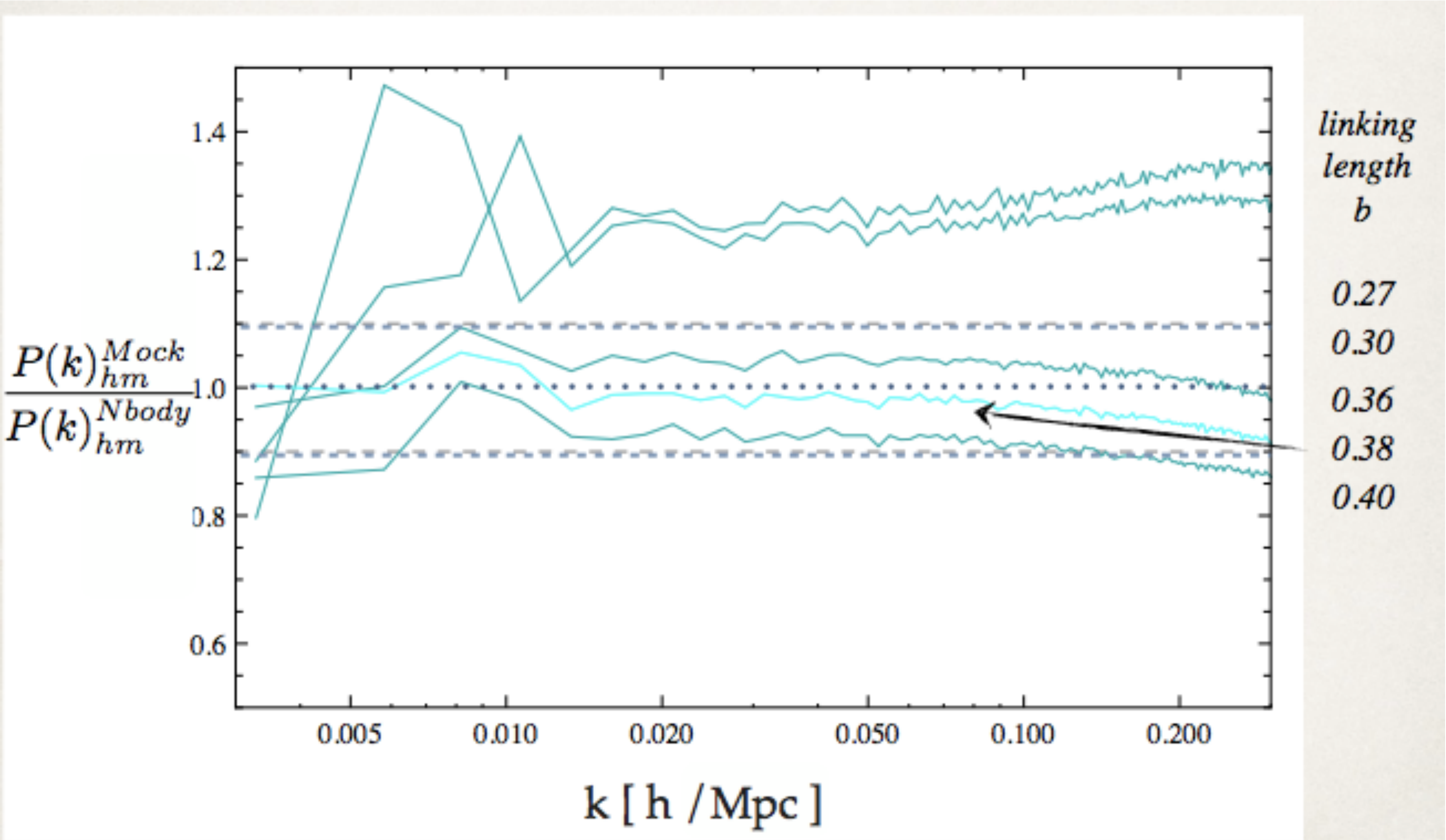
pioneering works by

1. Scoccimarro & Sheth 2002 (PThalos)
   2LPT
2. Monaco et al 2002 (Pinocchio)
   Zeldovich (being improved to 2LPT)
   includes merging histories

3. Tassev et al 2013 (COLA)
   make N-body code faster
4. White et al 2014 (QPM)
   use quasi PM solver

see FASTPM *Feng+16*

PThalos



Halos from N-Body using FoF with b=0.2 Mass = 3.0 10^{13} Mo/h

previous BOSS mocks
calibration with N-body
1000 mocks done!

about 10 % deviation in the power-spectrum at k~0.25 needs to resolve halos

see galaxy mocks from Manera M. & BOSS collaboration 2012 &2014, MNRAS

What is the bottle-neck of the computations?

1. the gravity solver
2. resolution (number of particles)

Let us ask a different question:

How can I generate a distribution of halos/galaxies in a statistical way?

(of a certain type within a certain mass range) the answer is:

we need all the higher order statistics from a reference sample
+ a way to draw from such a higher order PDF:

$$N_h \curvearrowleft \mathcal{P}(N_h | \xi_1^h, \xi_2^h, \xi_3^h, \xi_4^h, \dots)$$

Let us imagine we would know the halo/galaxy density field, i.e. the expected number of halos/galaxies per finite volume (cell).

Stochastic biasing

$$N_h \curvearrowleft \mathcal{P}(N_h|\rho_h)$$

caution! we still need to know the deviation from Poissonity!

over-dispersion modelled by the NB PDF:

$$P(N_i \mid \lambda_i, \beta) = \frac{\lambda_i^{N_i}}{N_i!} \frac{\Gamma(\beta + N_i)}{\Gamma(\beta)(\beta+\lambda)^{N_i}} \frac{1}{(1+\lambda/\beta)^{\beta}}$$

FSK, Yepes & Prada 2014, MNRAS

non-Poissonian PDFs:
W.C. Saslaw, A.J.S. Hamilton, 1984, ApJ, 276, 13
Sheth R. K., 1995, MNRAS, 274, 213

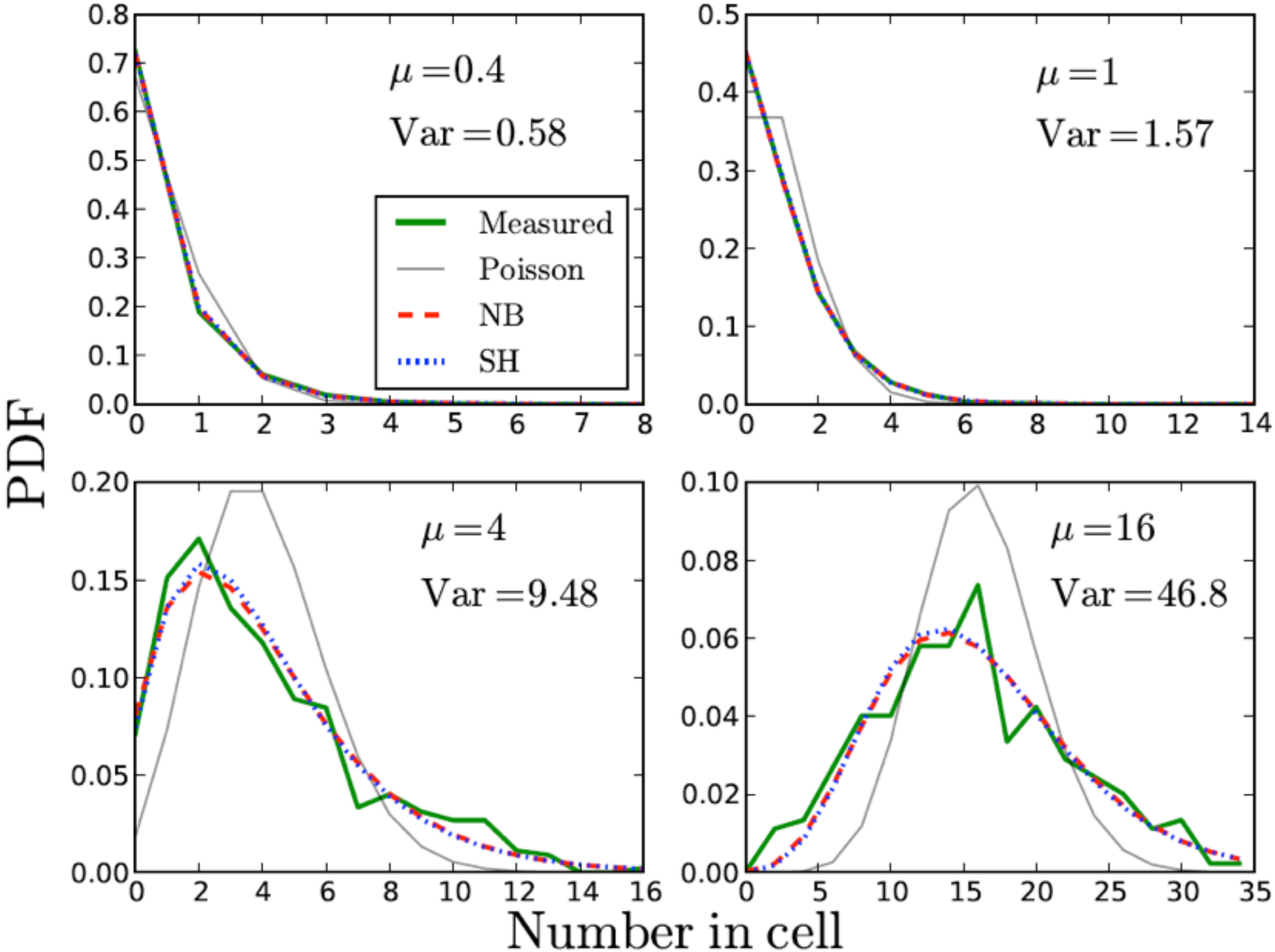
Over-dispersion 10% effect at BAO scales in Pk for LRGs, stronger effect 20-30% for eLGs

stochastic bias:
Dekel A., Lahav O., 1999, ApJ, 520, 24
Sheth R. K., Lemson G., 1999, MNRAS, 304, 767

and many more see references in e.g. Kitaura et al 2013

Somerville et al 2001, MNRAS, 320, 289
Casas-Miranda et al 2002, MNRAS, 333, 730

Neyrinck M et al 2013; Aragon-Calvo M. 2013



Deterministic biasing

we need to know the (deterministic) biasing,
but this implies knowing all the higher order correlation functions!

$$B(\rho_h|\rho_{\mathrm{M}}) = B(\rho_h|\rho_{\mathrm{M}}, \xi_1^h, \xi_2^h, \xi_3^h, \xi_4^h, \ldots)$$

$$N_h \curvearrowleft \mathcal{P}(N_h|B(\rho_h|\rho_{\mathrm{M}}, \xi_1^h, \xi_2^h, \xi_3^h, \xi_4^h, \ldots))$$

Deterministic biasing parametrization

Fry & Gaztañaga 1993

$$\rho_h = f_h^a \sum_i a_i \delta_{\mathrm{M}}^i$$

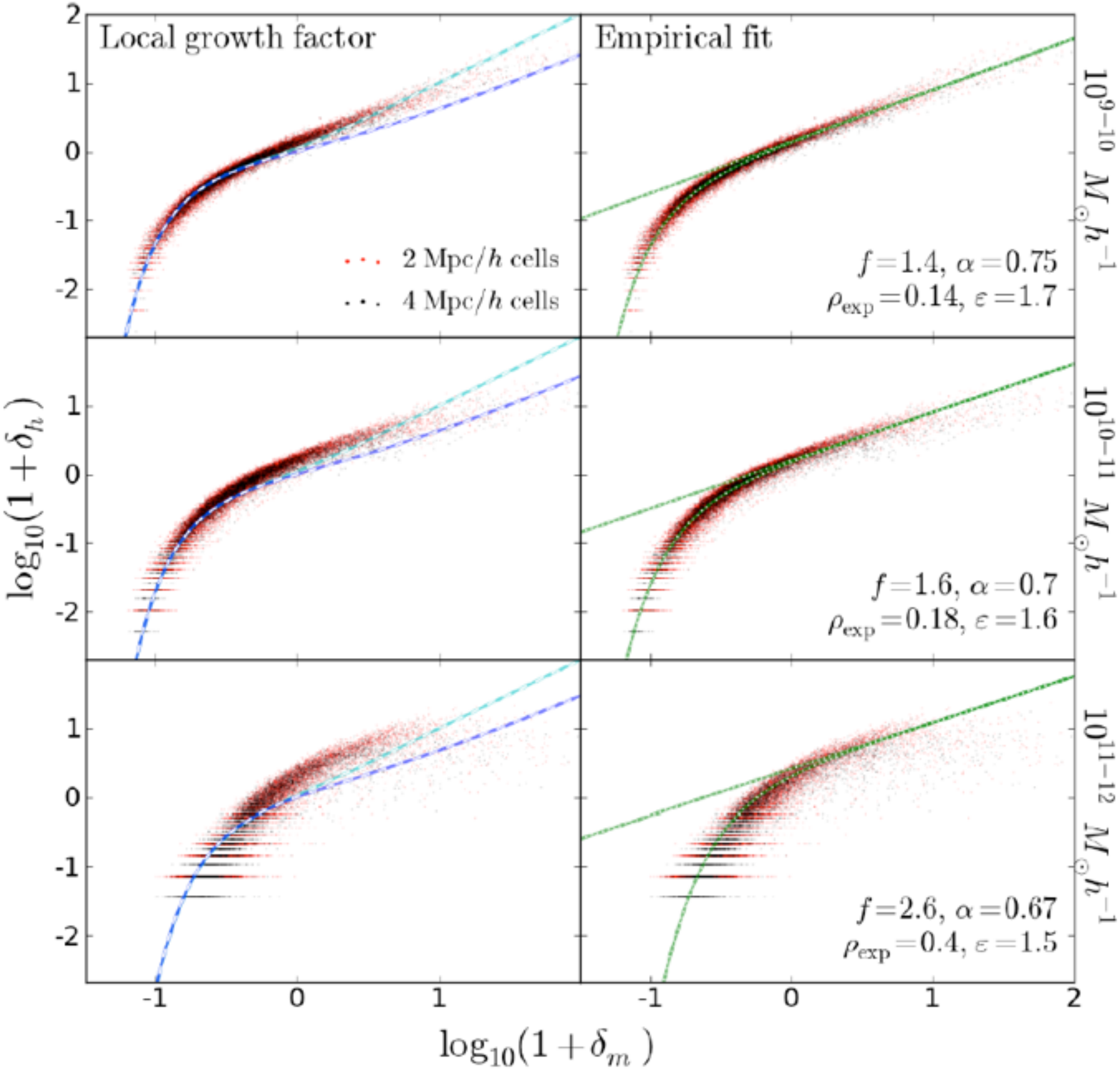
Cen & Ostriker 1993; de la Torre & Peacock

$$\rho_h = f_h^b \exp\left[\sum_i b_i \log\left(1+\delta_{\mathrm{M}}\right)^i\right]$$

FSK, Yepes & Prada 2014 + Neyrinck et al 2014

$$\rho_h = f_h\,\theta(\rho_{\mathrm{M}} - \rho_{\mathrm{th}})\,\rho_{\mathrm{M}}^{\alpha} \exp\left[-\left(\frac{\rho_{\mathrm{M}}}{\rho_\epsilon}\right)^{\epsilon}\right]$$

Neyrinck M et al 2013; Aragon-Calvo M. 2013



How can we do it?

given the dark matter field
(from low resolution N-body or approximate solver)
parametrise the bias
and constrain the bias parameters, in such a way that the higher order correlation functions are matched

expensive!

How can we do it?

FSK+14 arxiv:1407.1236

$$\bar{N}_h = \langle \rho_h \rangle \leftarrow \xi_1^h$$

$$P_h(k) \leftarrow \xi_2^h$$

$$\mathcal{P}^1(\rho) = \int_{-\sqrt{-1}\infty}^{\sqrt{-1}\infty} \frac{\mathrm{d}t}{2\pi\sqrt{-1}} \exp\left(t\rho + \mathcal{C}(t)\right)$$

$$\mathcal{P}_h^1\left(B\left(\rho_h|\rho_{\mathrm{M}}\right)\right) \leftarrow \{\xi_1^h, \xi_2^h, \xi_3^h, \xi_4^h, \ldots\}$$

$$N_h \curvearrowleft \mathcal{P}(N_h|B(\rho_h|\rho_{\mathrm{M}}, \bar{N}_h, P_h(k), \mathcal{P}_h^1))$$

our approach:

Let us use low resolution + approximate gravity solvers and augment all the missing halos!

- The approximate gravity solver accurately models the higher order statistics of the dark matter density field.

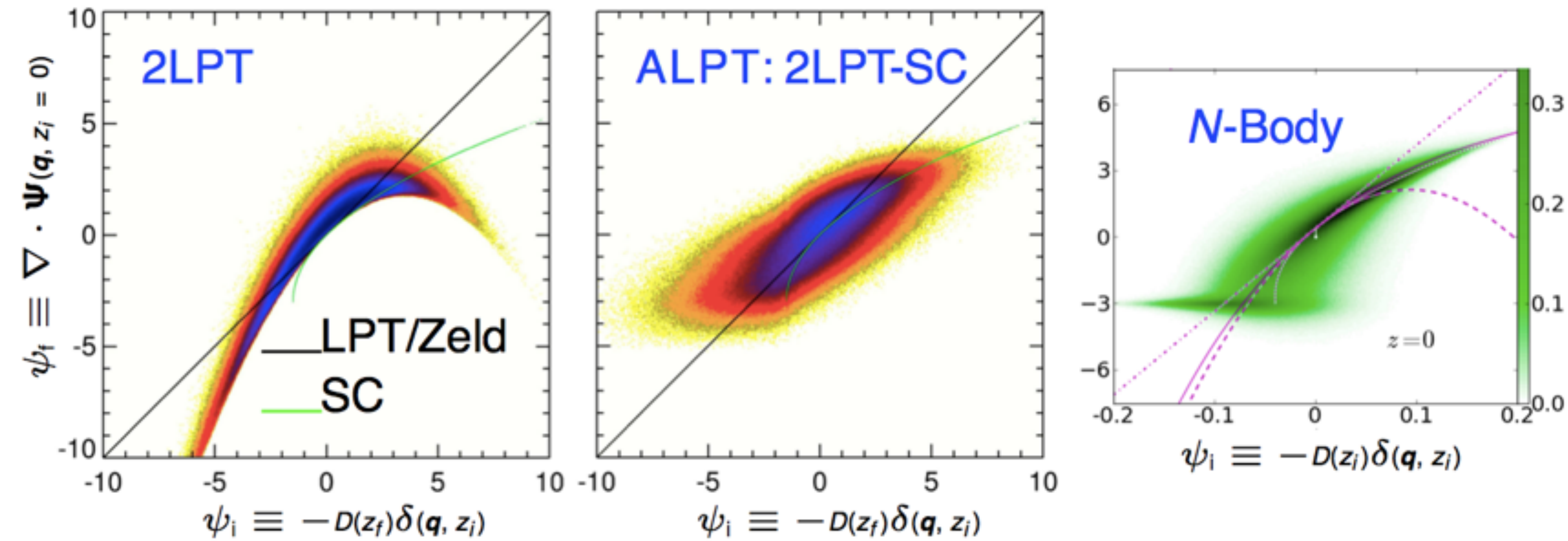
(low N-body resolution or perturbation theory based method)

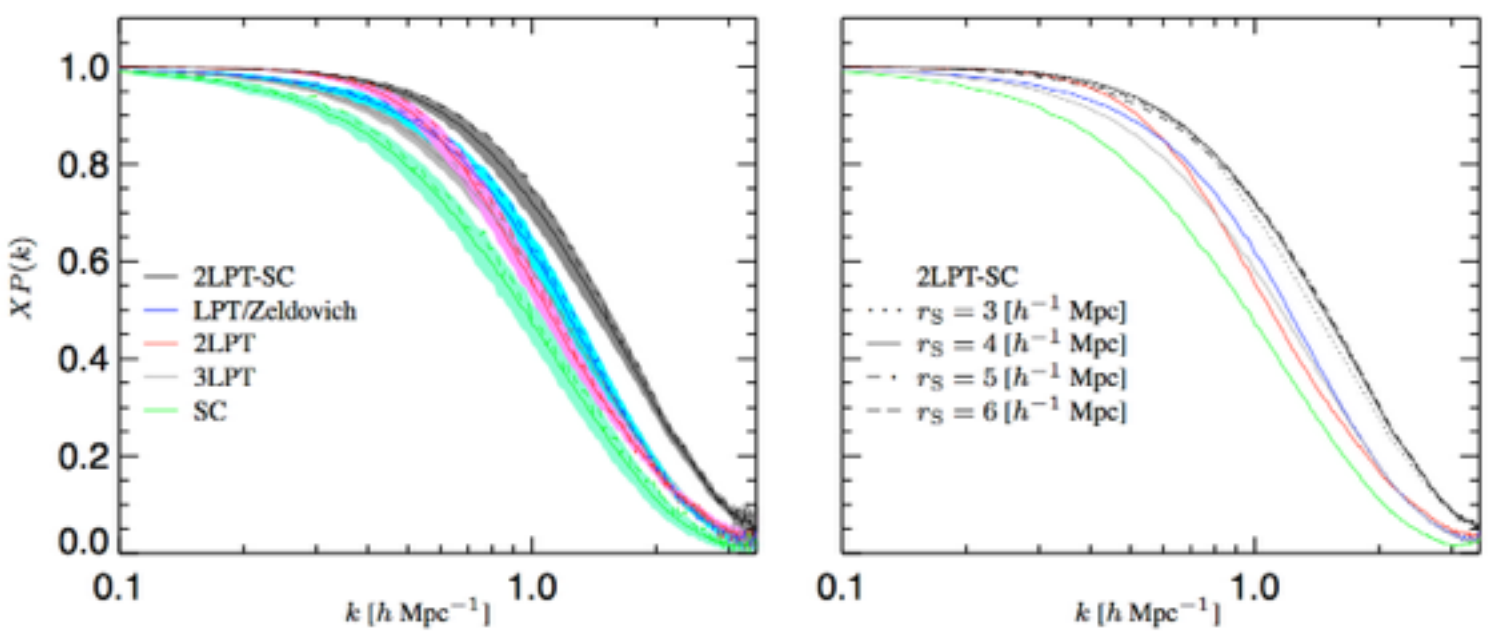
- The biasing model accurately connects the dark matter phase- space distribution with the halo distribution.

Simple efficient accurate one-step gravity solver…

ALPT: Augmented Lagrangian Perturbation Theory

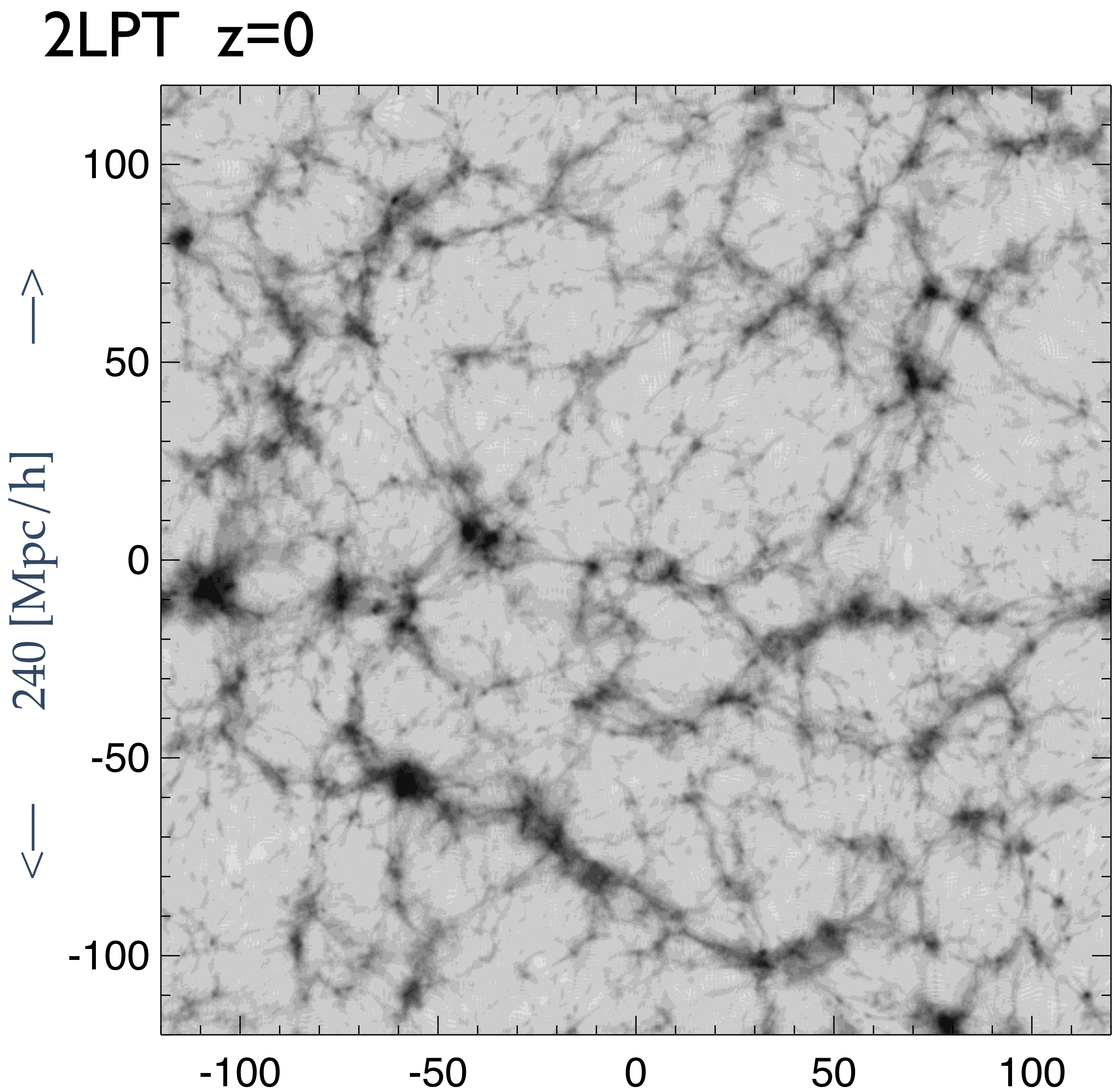
FSK & Heß 13, arxiv:1212.3514



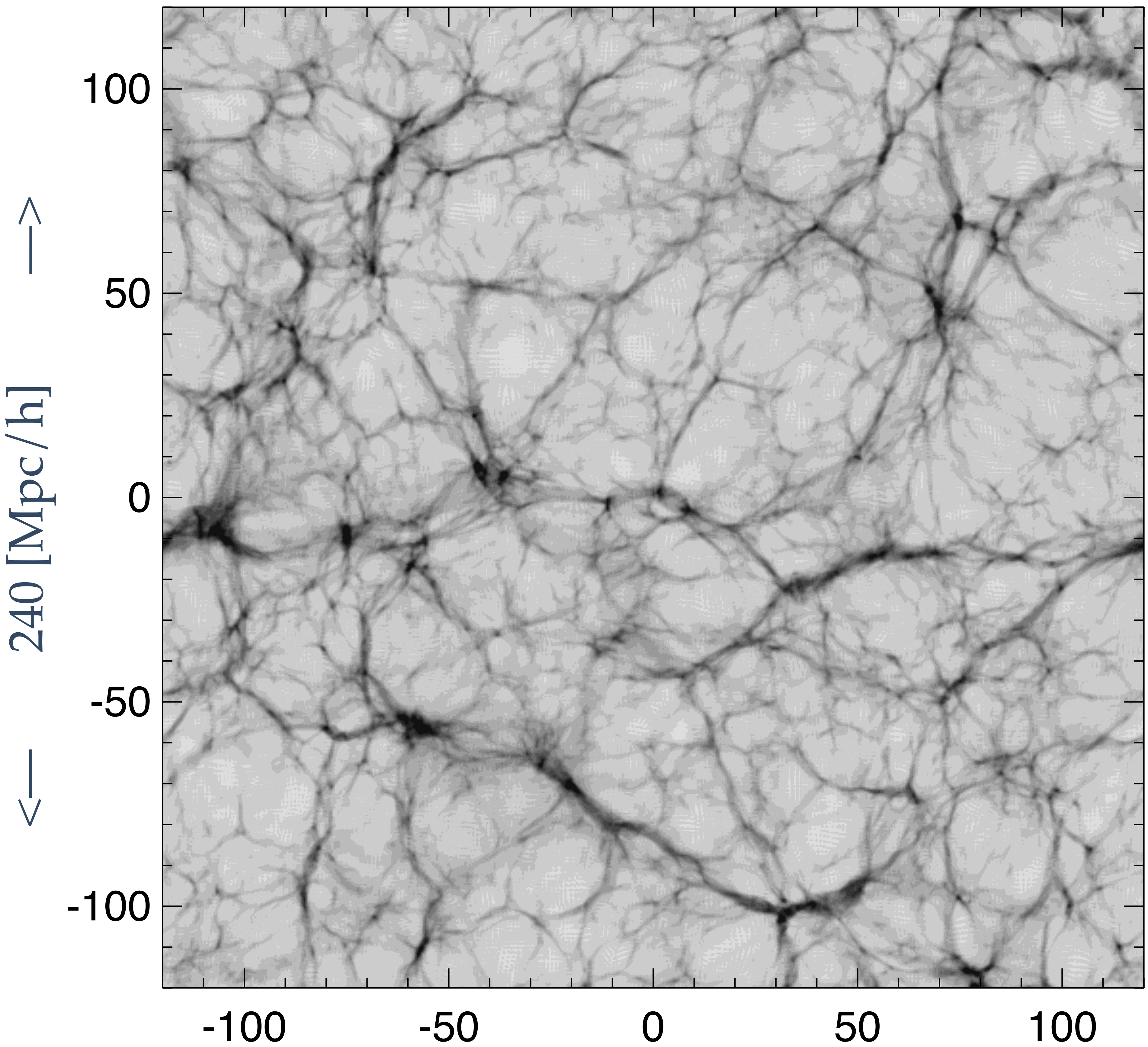


for N-body relation see Neyrinck13
for the spherical collapse relation see also Bernardeau94; Mohayee06
improvement for k>1 found in Neyrinck15
speed up N-body codes with LPT: COLA Tassev+13

ALPT is a fast one-step solver
75% more correlated with the full N-body solution than 2LPT at k~2 h/Mpc


2LPT z=0
240 [Mpc/h]
100
50
0
-50
-100
-100
-50
0
50
100

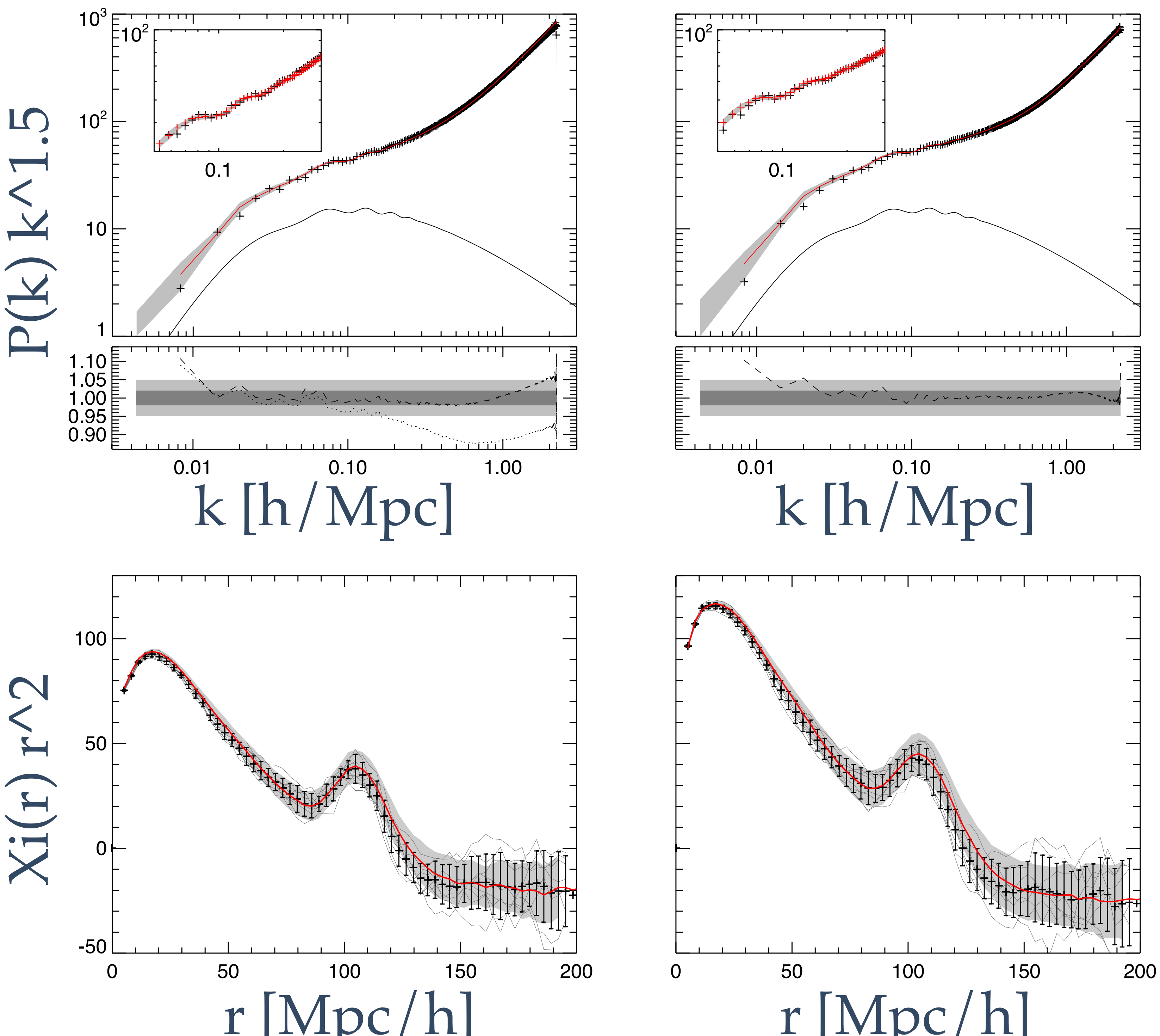
ALPT z=0



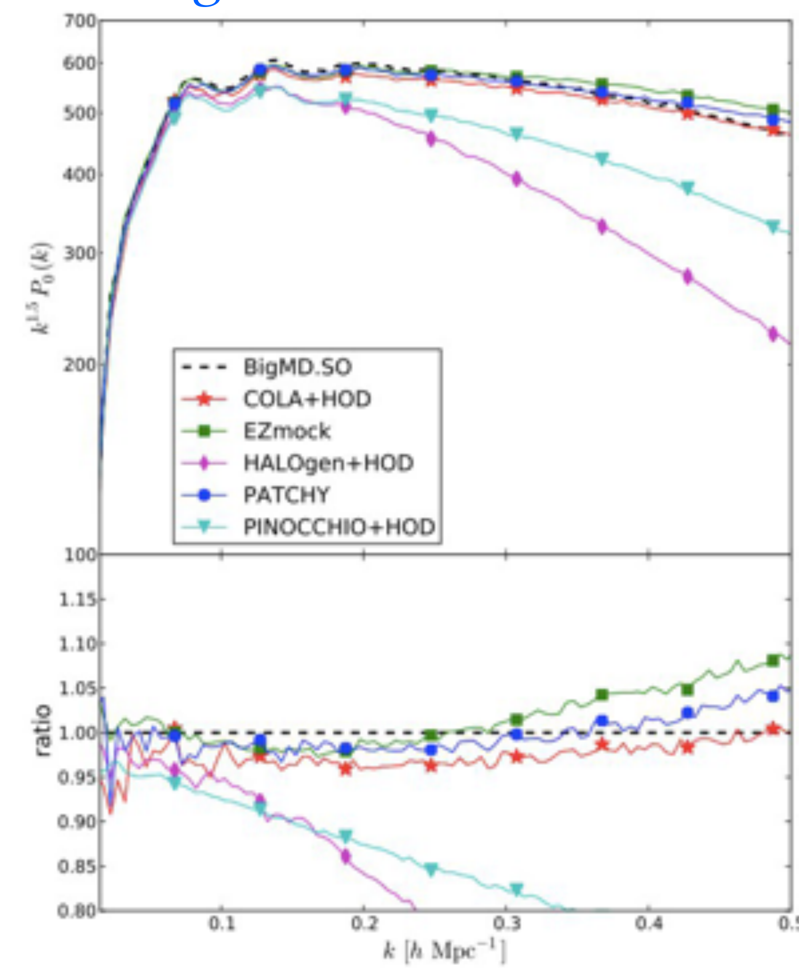
Calibration with N-body simulations

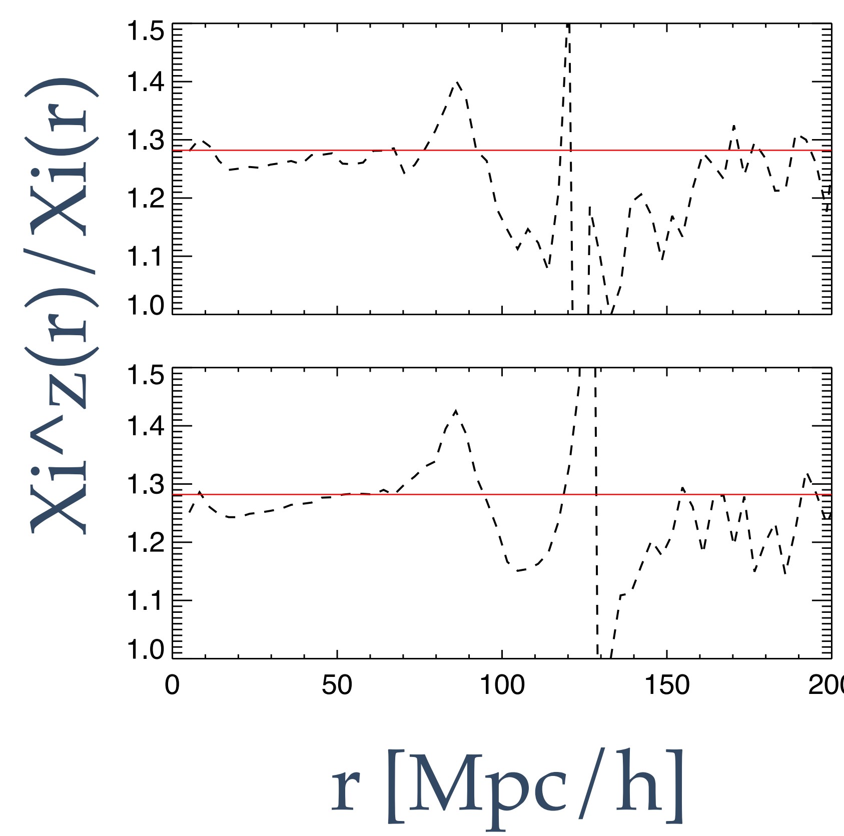
FSK, Yepes & Prada 14, arxiv:1307.3285

real-space redshift-space



Chuang et al.



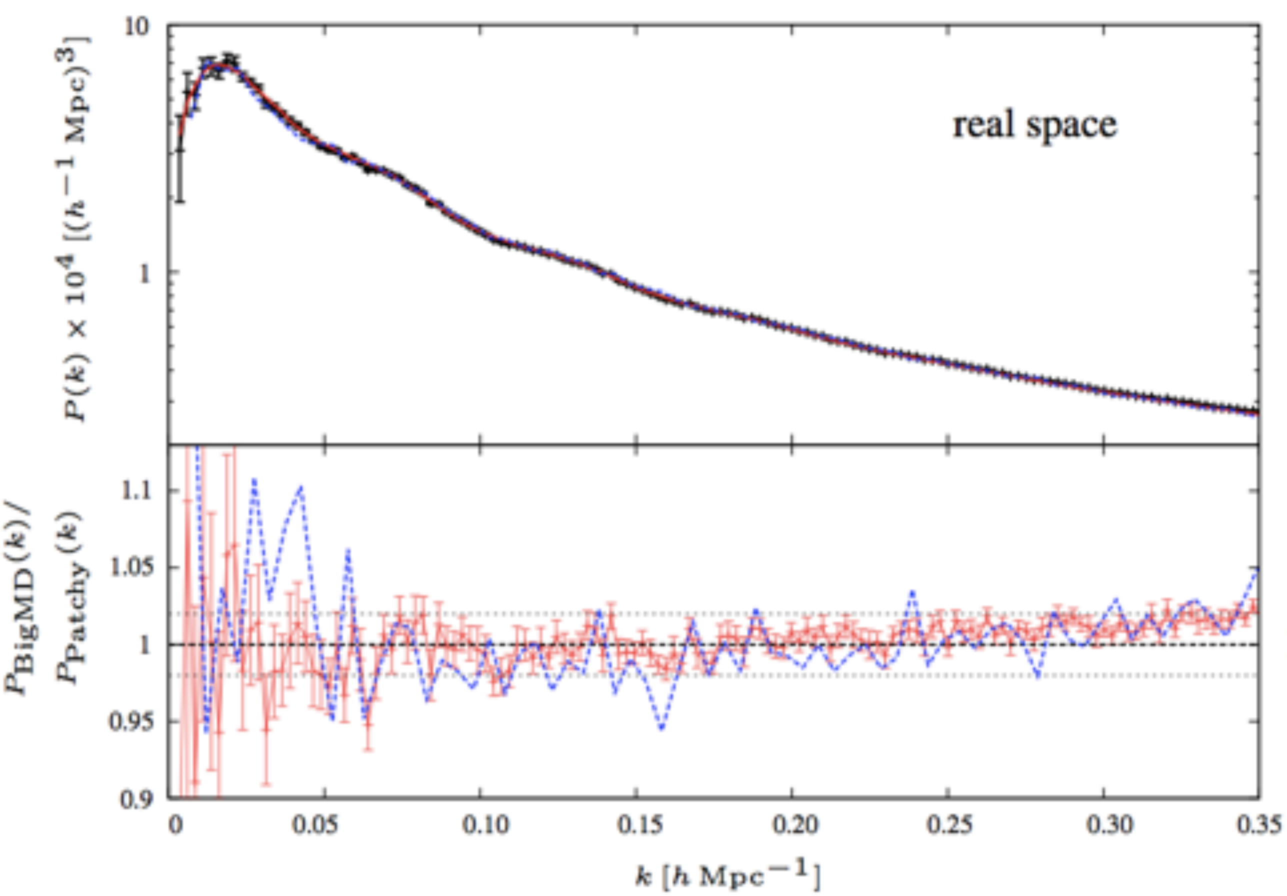


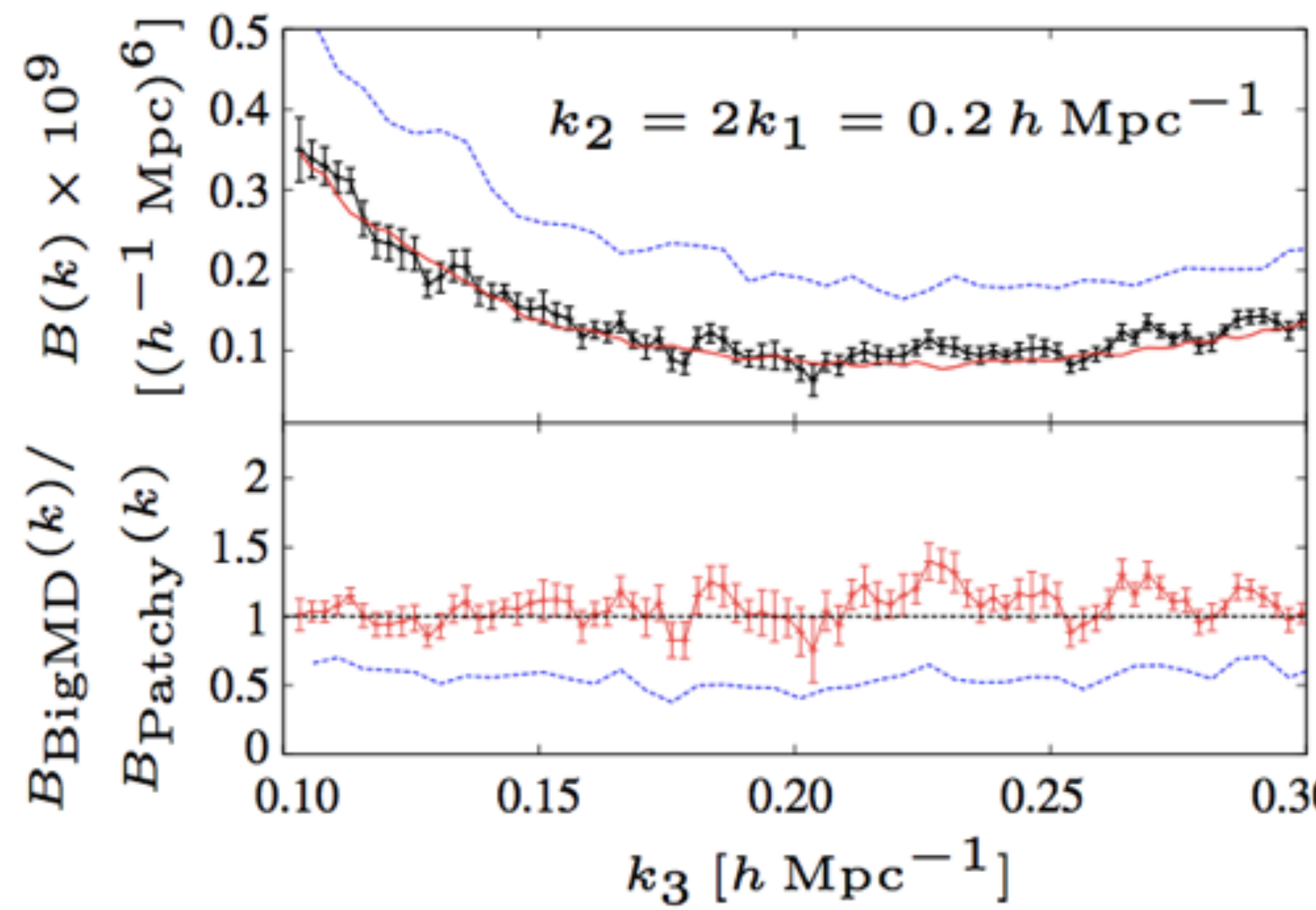
53 times lower resolution required!

PATCHY CODE: 3pt statistics bispectrum

FSK, Gil-Marín, Scóccola, Chuang+14

Same Pk different bias parameters: blue w/o threshold bias, red with threshold bias
with threshold bias reproduces bispectrum from N-body based catalogue





Threshold bias is an indispensable ingredient!

Kaiser84

Is this the end of the story?

Nonlocal biasing
(as part of deterministic biasing)

What is the impact of non-local bias?

see eg McDonald&Roy09
Sherwin&Zadarriaga12
Chan,Sheth&Schoccimarro12
Sheth,Chan&Schoccimarro13
Baldauf+12,13;
Saito+14

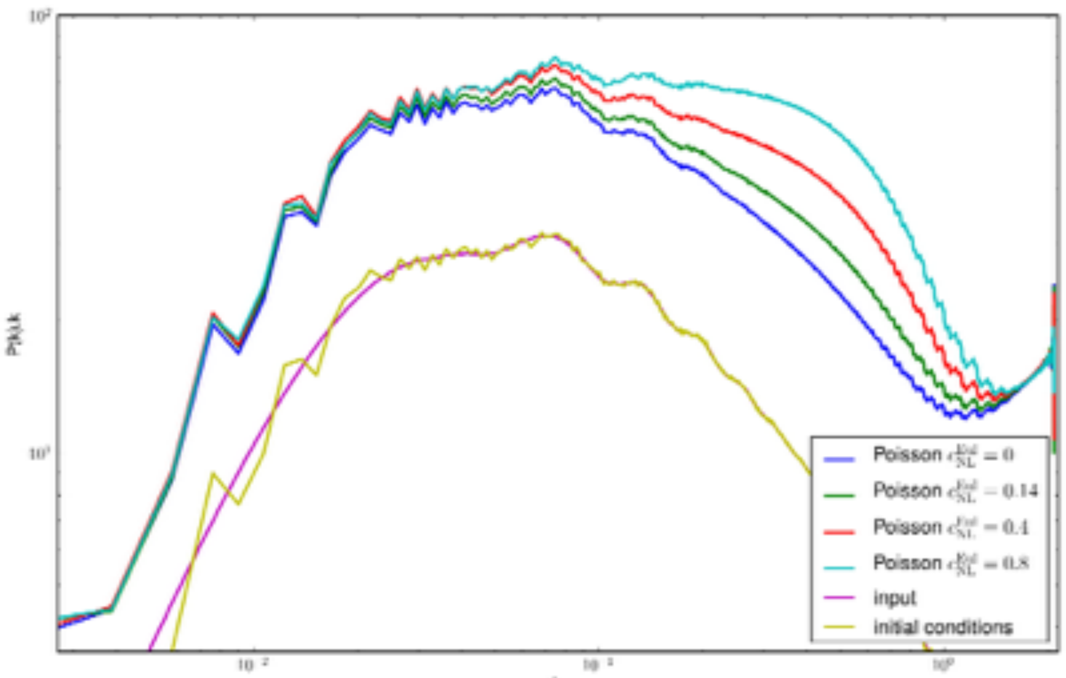
Can it be effectively modelled as part of the stochastic bias?

Nonlocal biasing

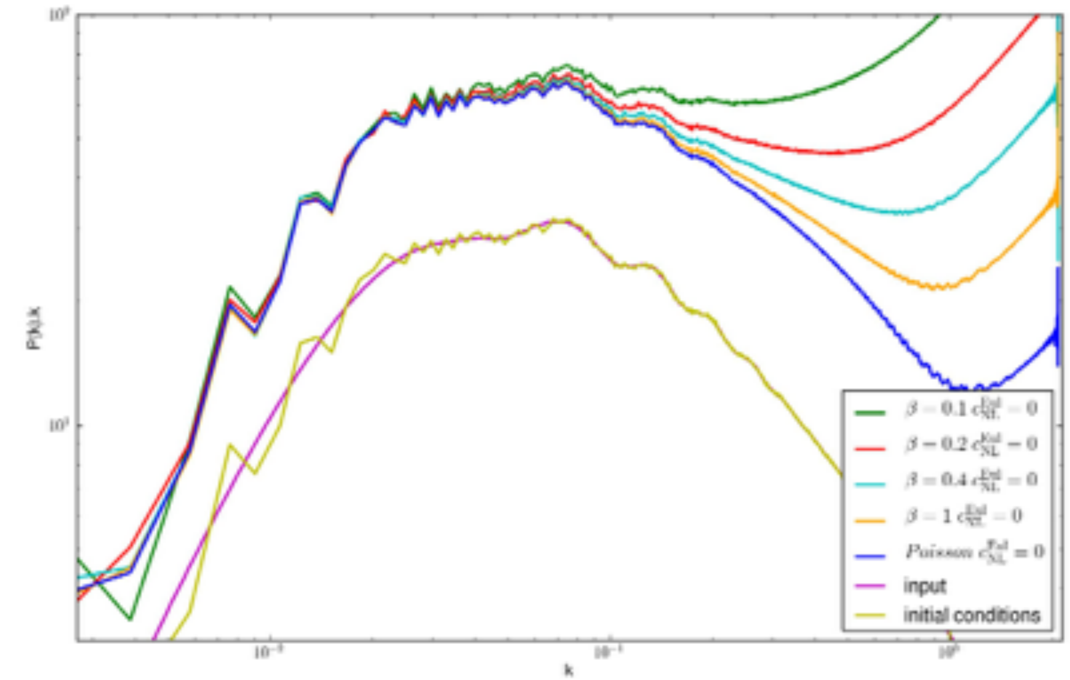
We include in PATCHY second order nonlocal bias:

$$B(\rho_h|\rho_{\mathrm{M}}) = B_{\mathrm{L}}(\rho_h|\rho_{\mathrm{M}}) + B_{\mathrm{NL}}(\rho_h|\rho_{\mathrm{M}})$$

Nonlocal bias

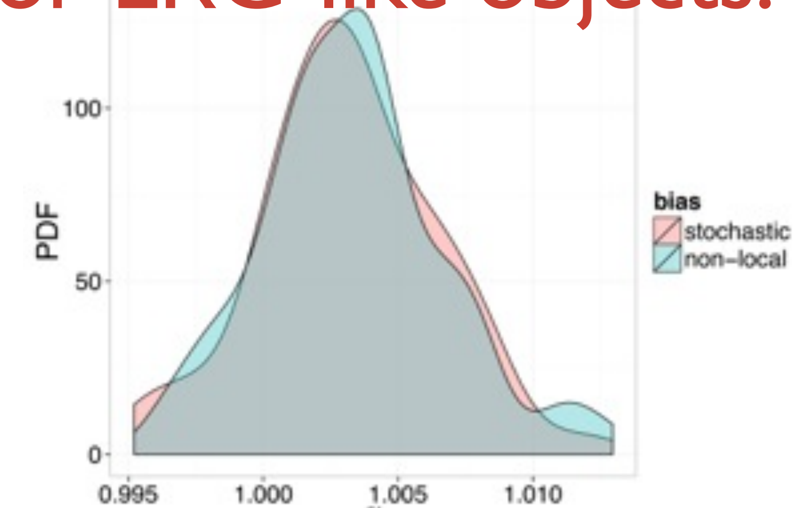


Stochastic bias



nonlocal bias is partially degenerate with stochastic bias!

not very relevant
for LRG-like objects:



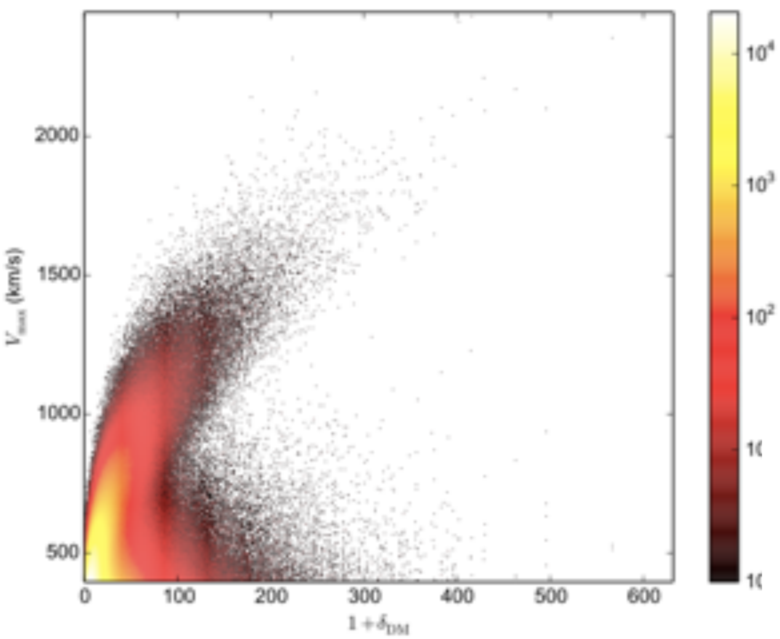
Mathieu Autefage, FSK, Christian Wagner & Raul Angulo in prep

Is this the end of the story?

How can we assign masses to the objects?

Cheng Zhao's master thesis

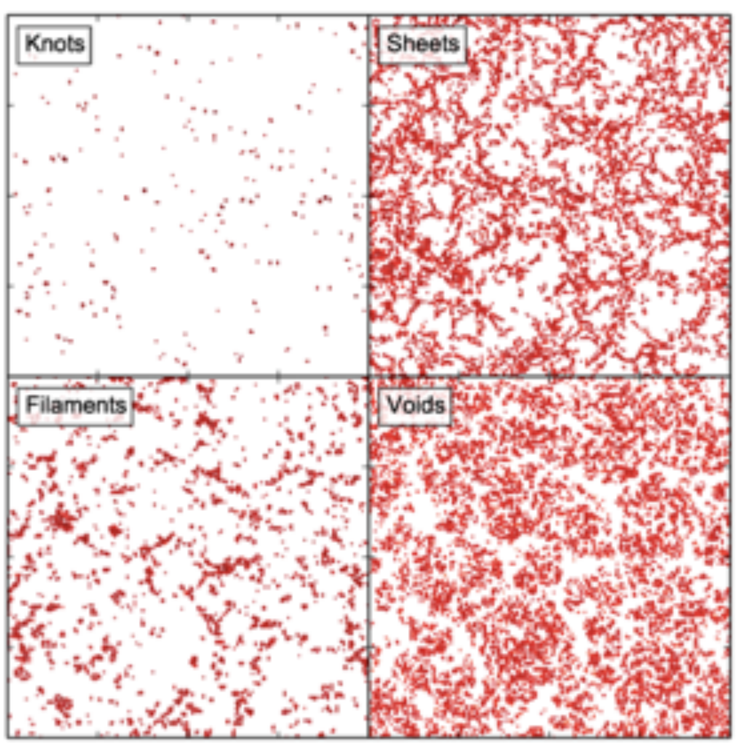
mass (Vmax) - density relation

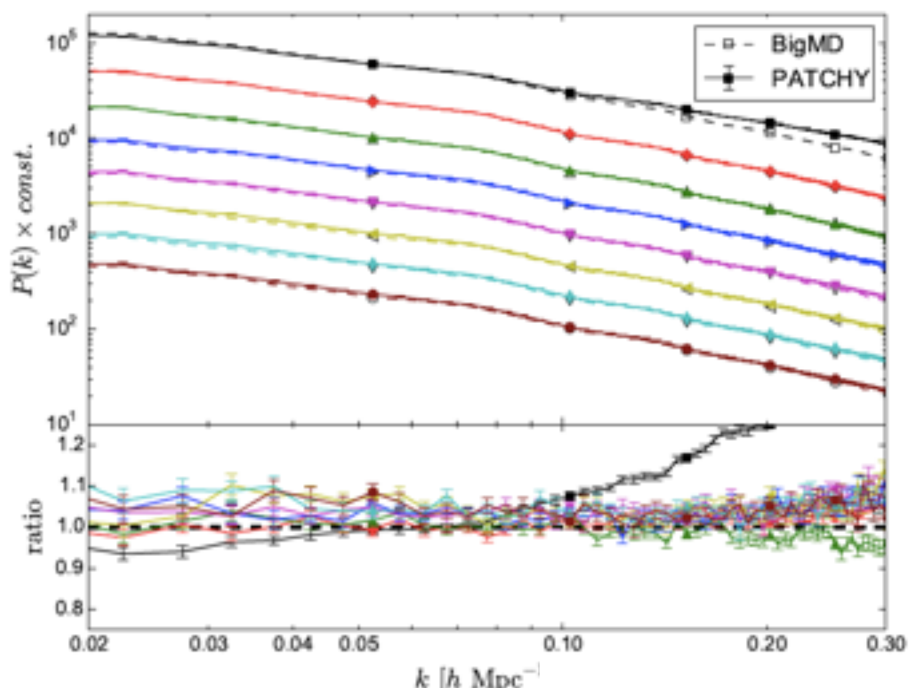


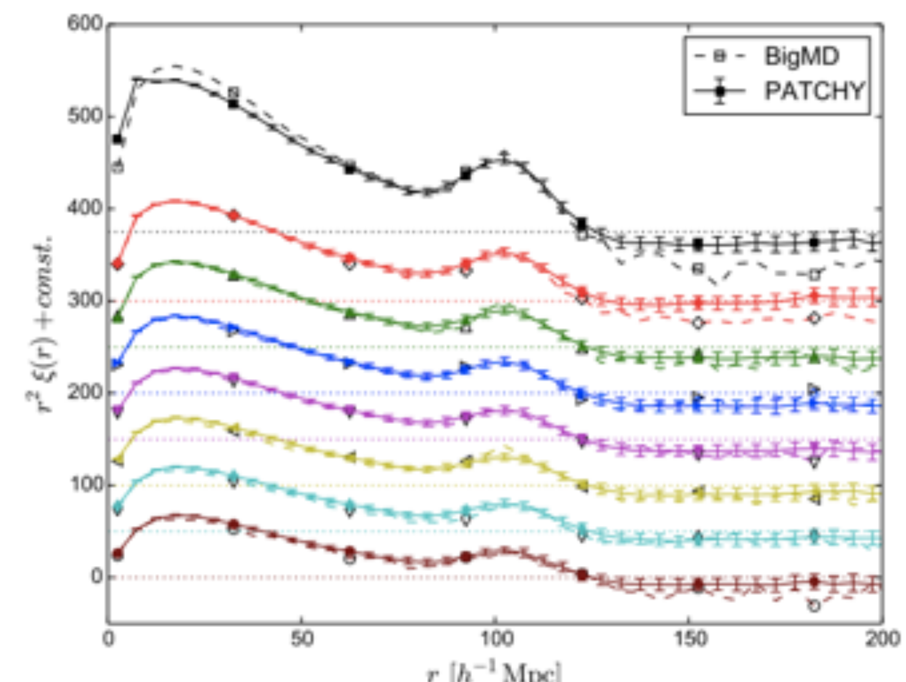
HADRON code: Cheng Zhao, FSK+15 arXiv:1501.05520

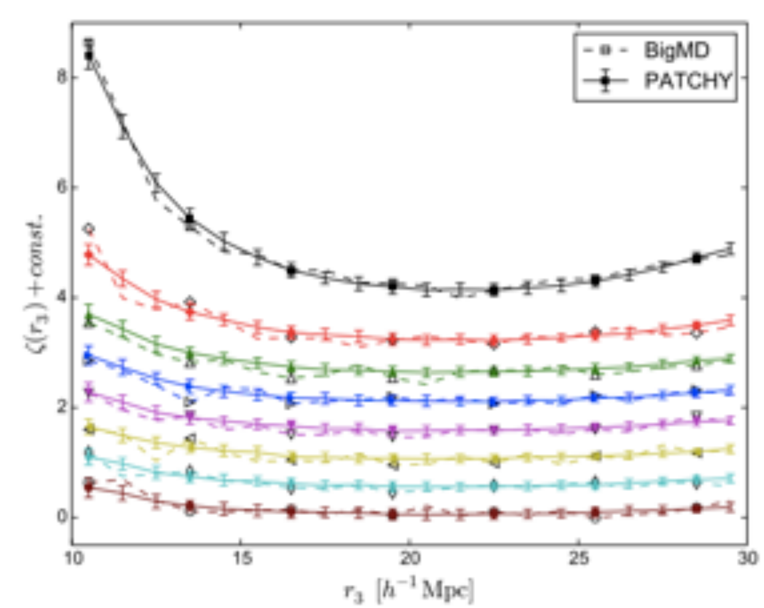
dependence on mass function, local density and cosmic web environment

$$P(M_{\rm h}^i|\{\boldsymbol{r}_{\rm h}\},\rho_{\rm M},T,\Delta r_{\rm min}^M,\{p_{\rm c}\},z)$$



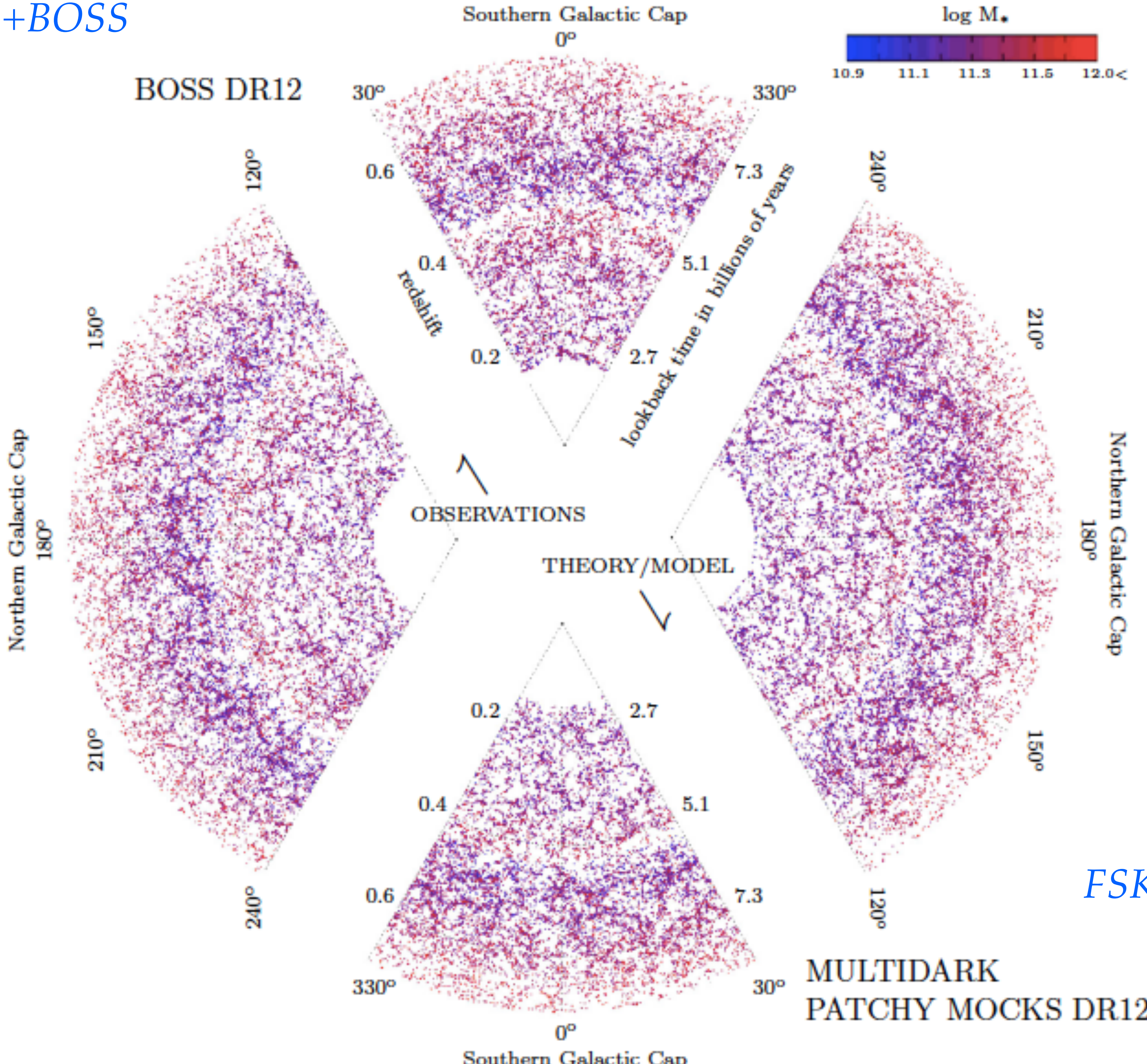






PATCHY MULTIDARK BOSS DR11/DR12

collaborators Chia-Hsun Chuang, Sergio Rodriguez-Torres, Cheng Zhao, F. Prada, G. Yepes, A. Klypin +BOSS



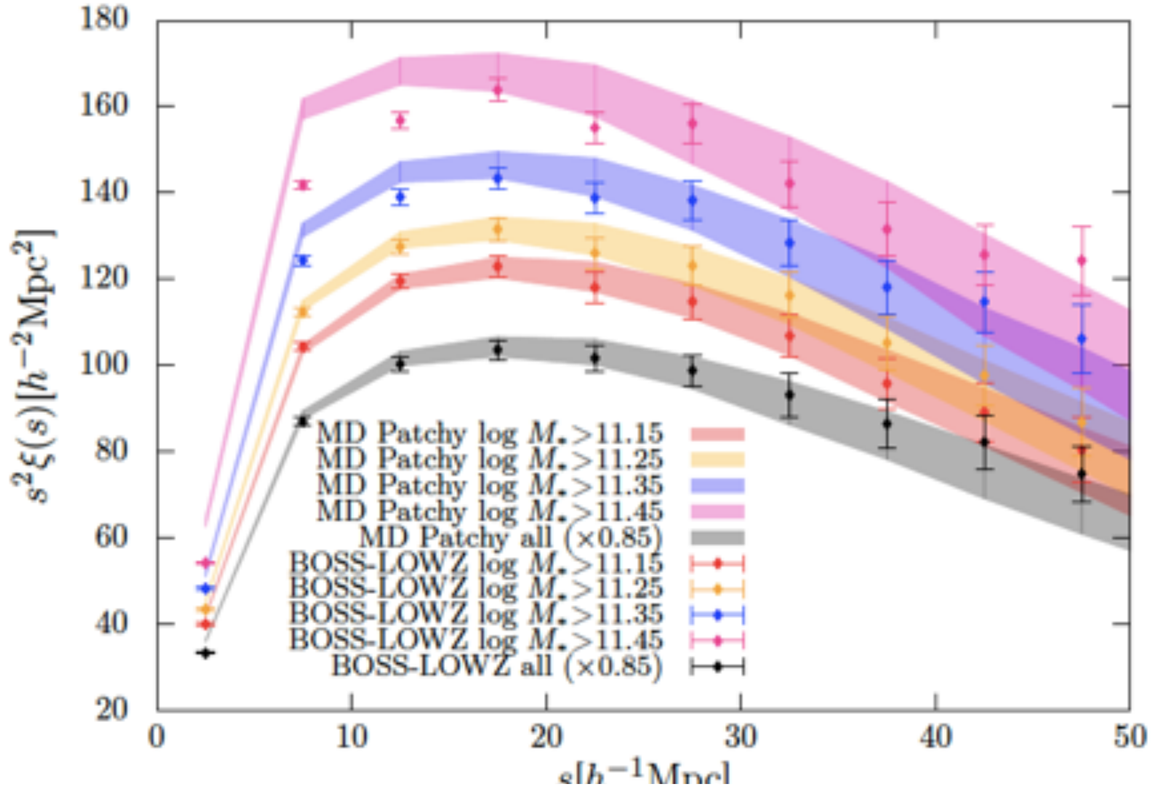
>12000 mocks including light cone and evolution effects (>4000 LOWZ, >4000 CMASS, >4000 COMBINED SAMPLE)

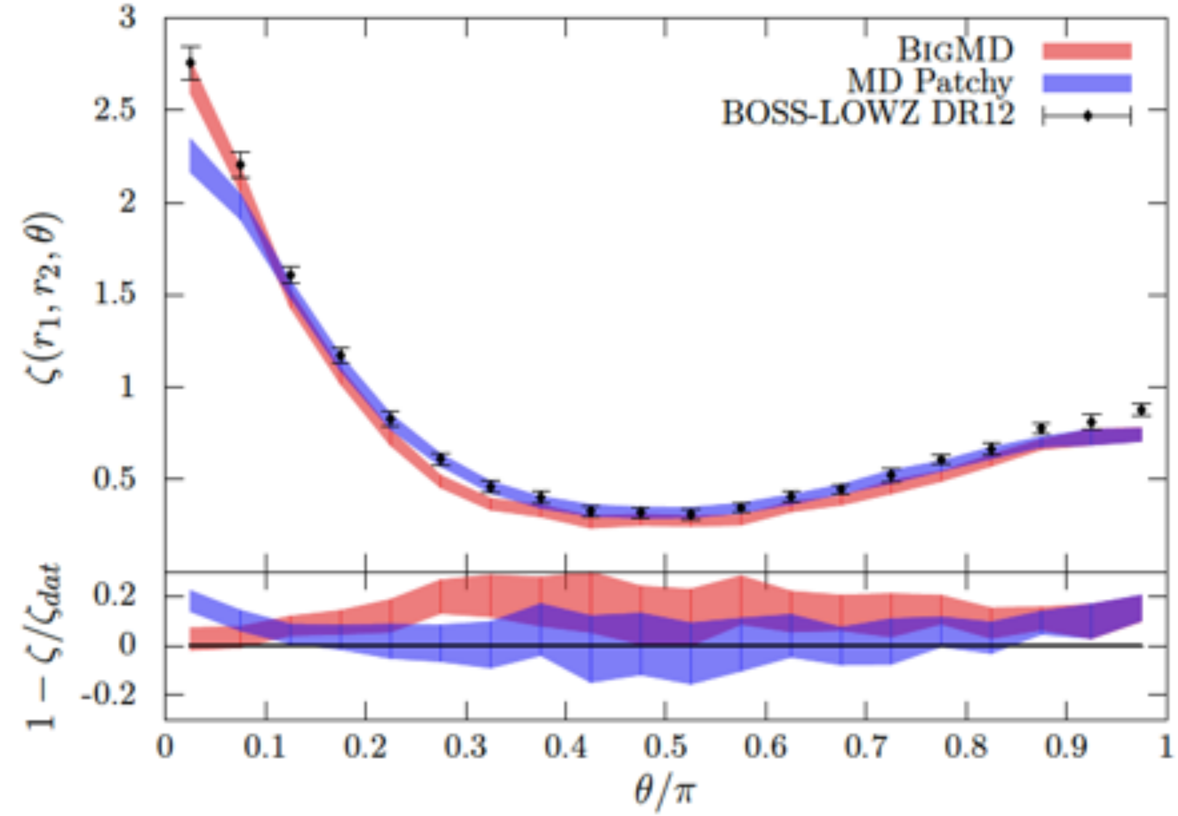
effective V ~ 192,000 [Gpc/h]^3
effective # p ~(61,440)^3

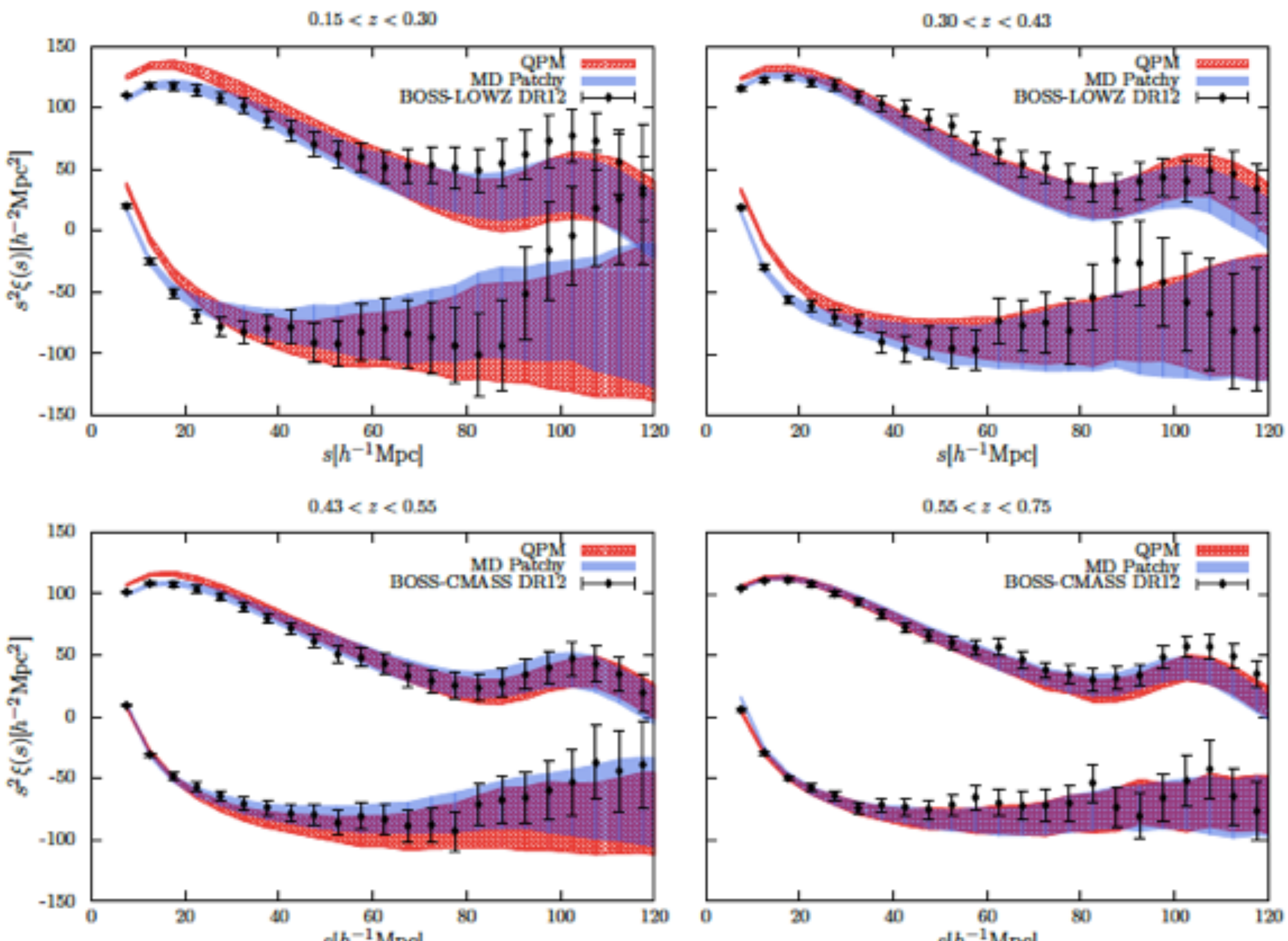
0.5M CPU hr vs 9billion CPU hr

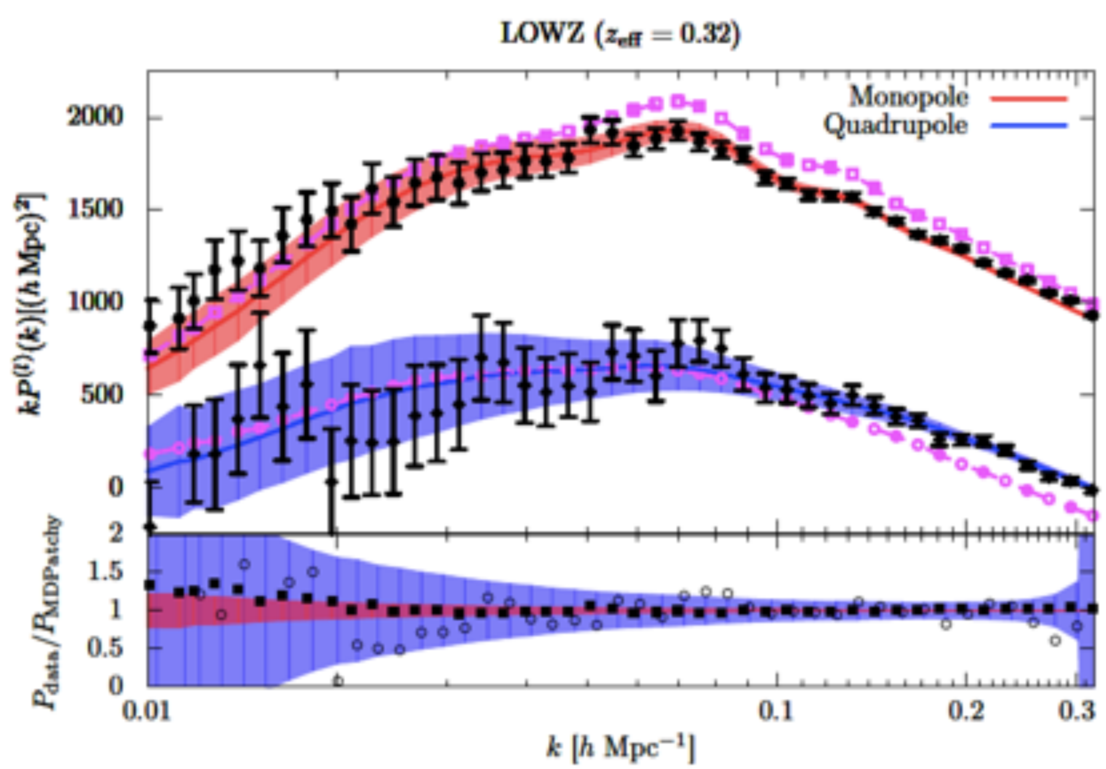
FSK+BOSS arXiv:1509.06400v1

At first glance: mocks are indistinguishable from observations!



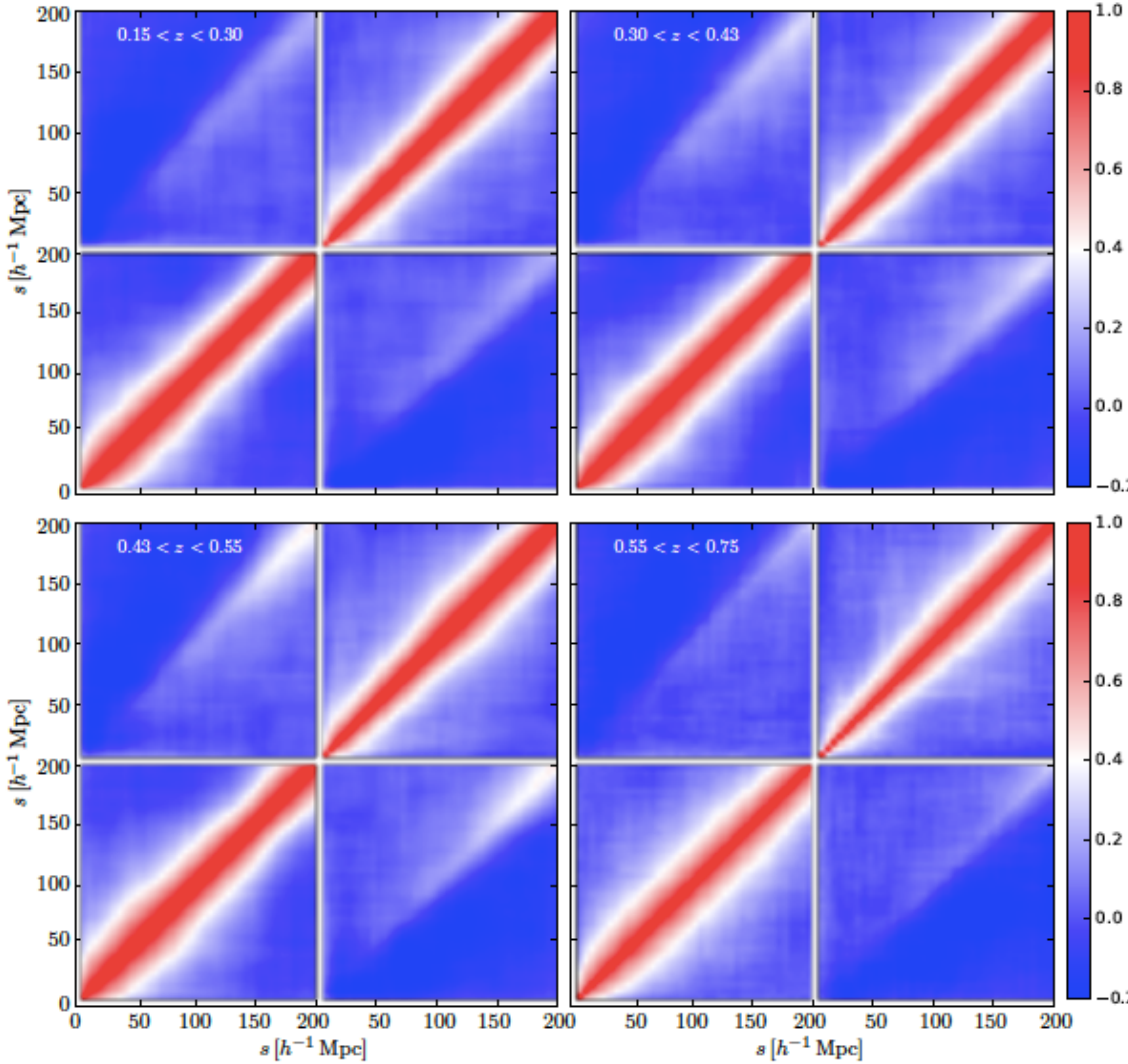






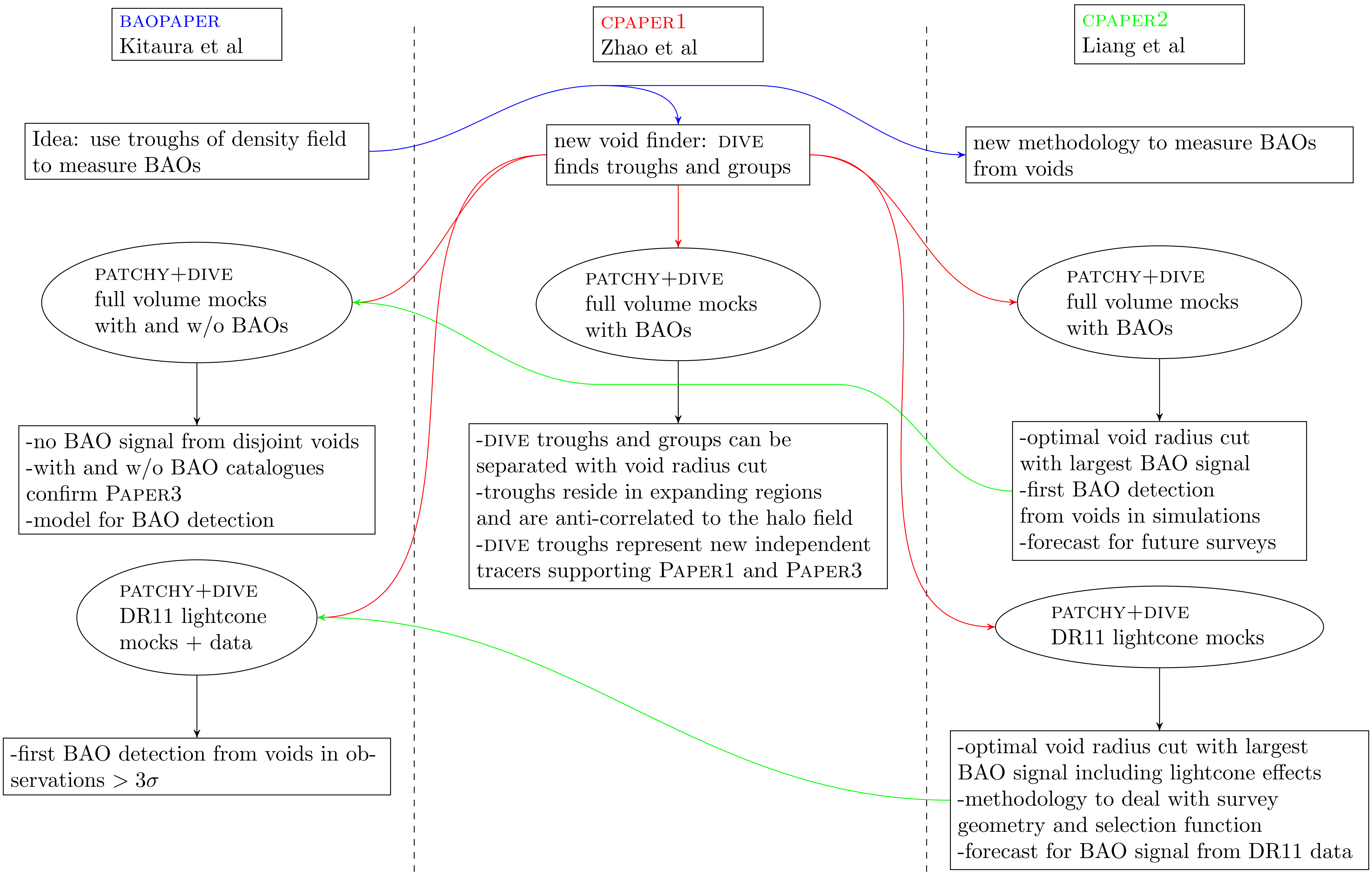
FSK+BOSS arXiv:1509.06400v1

Cosmic evolution of covariance matrices



we are also working on a comparison study of covariance matrices with different methods:
Ariel Sánchez, Pierluigi Monaco, Martín Crocce, Chia-Hsun Chuang, Claudio dalla Vecchia, FSK

BAOs from voids



in collaboration with Chia-Hsun Chuang, Yu Liang, Cheng Zhao, Charling Tao

+BOSS

Void finders

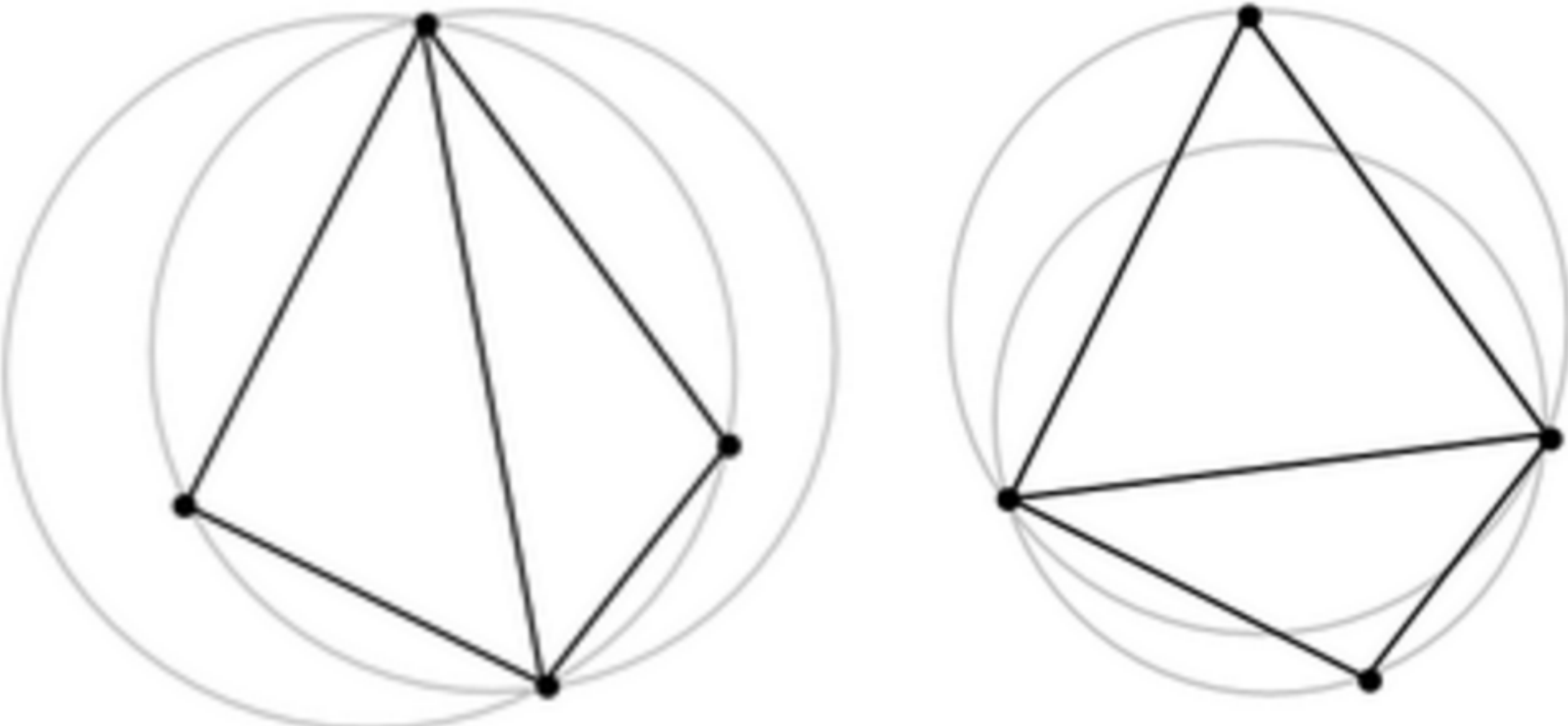
(i) under-dense regions based on the smoothed dark matter (or halo/galaxy) density field (e.g. Water-shed method) (*Colberg et al. 2005; Shandarin et al. 2006; Platen et al. 2007; Neyrinck 2008*);

(ii) gravitationally expanding regions based on the dynamics of the dark matter density field (*Hahn et al. 2007; Forero-Romero et al. 2009; Hoffman et al. 2012; Cautun et al. 2013*);

(iii) regions free of shell crossings based on phase-space tesselations of the particle distribution (*Abel et al. 2012; Falck et al. 2012; Shandarin et al. 2012*);

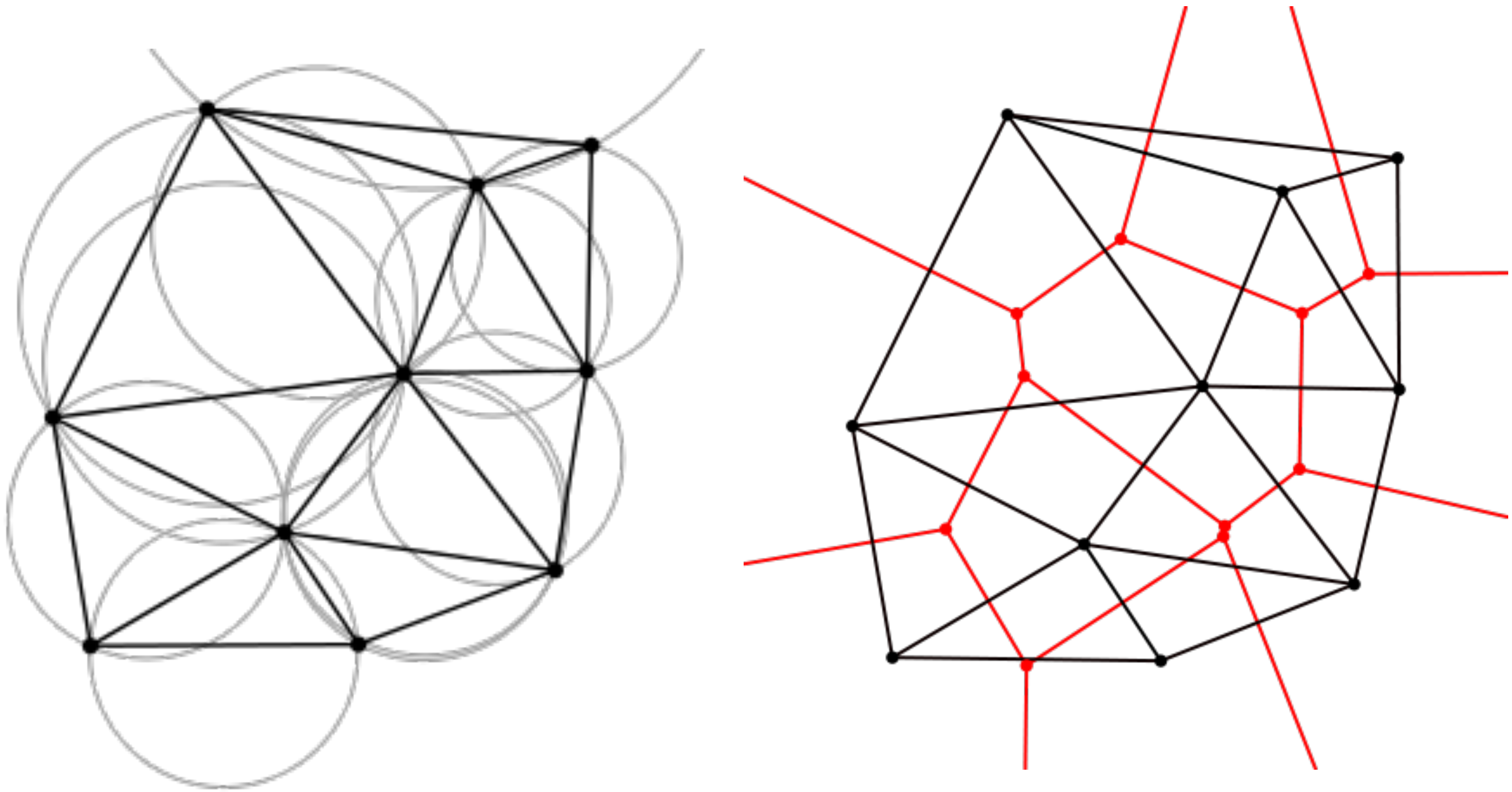
(iv) or empty spatial regions among discrete tracers (e.g. *El-Ad & Piran 1997; Aikio & Mähönen 1998; Hoyle & Vogeley 2002; Padilla et al. 2005; Patiri et al. 2006; Foster & Nelson 2009*).

Delaunay-triangulation in 2D: triangles and circum-circles

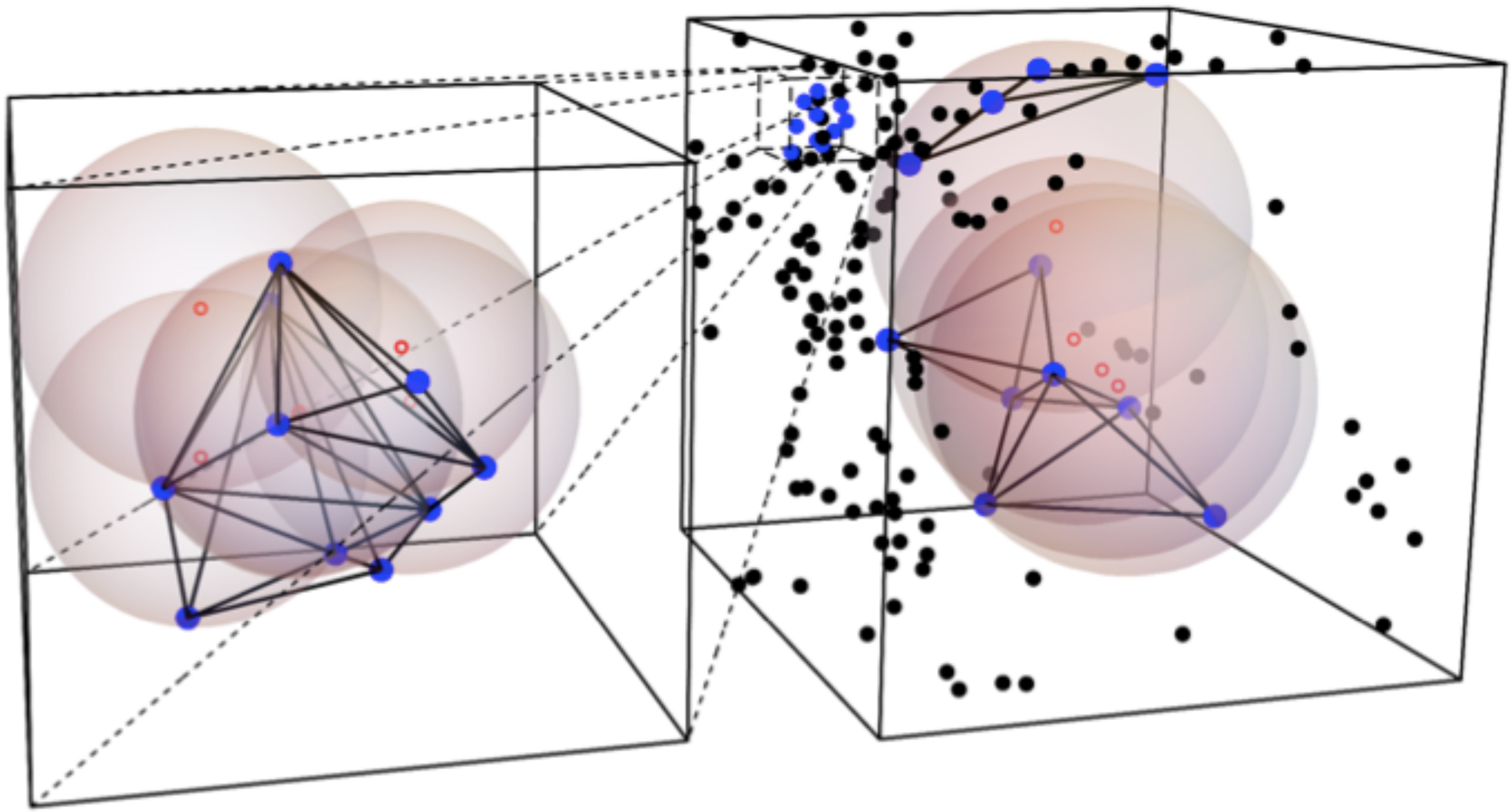


Incremental random and flip algorithms

Delaunay-Voronoi relation

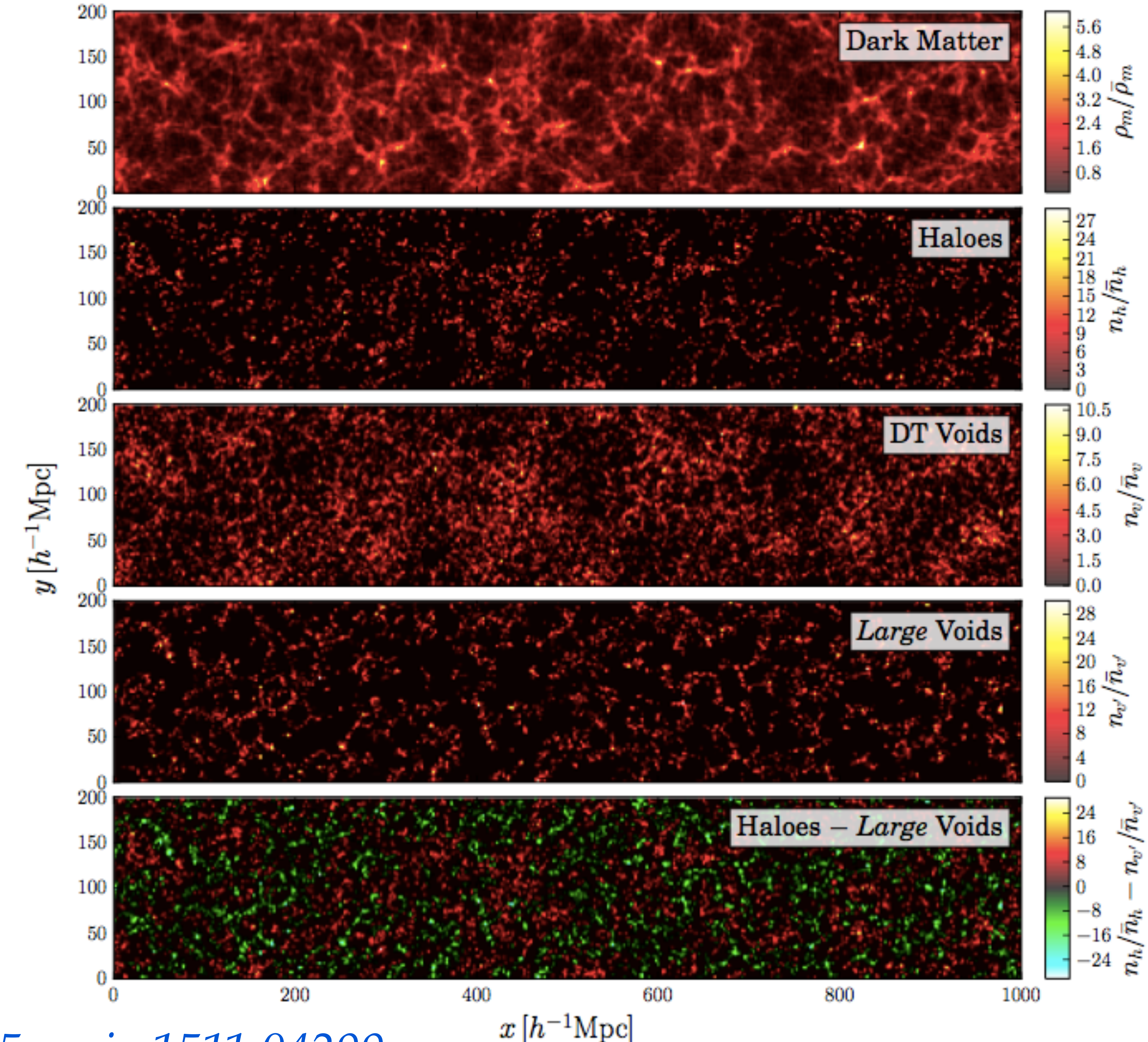


Delaunay-triangulation in 3D: tetrahedra and circum-spheres



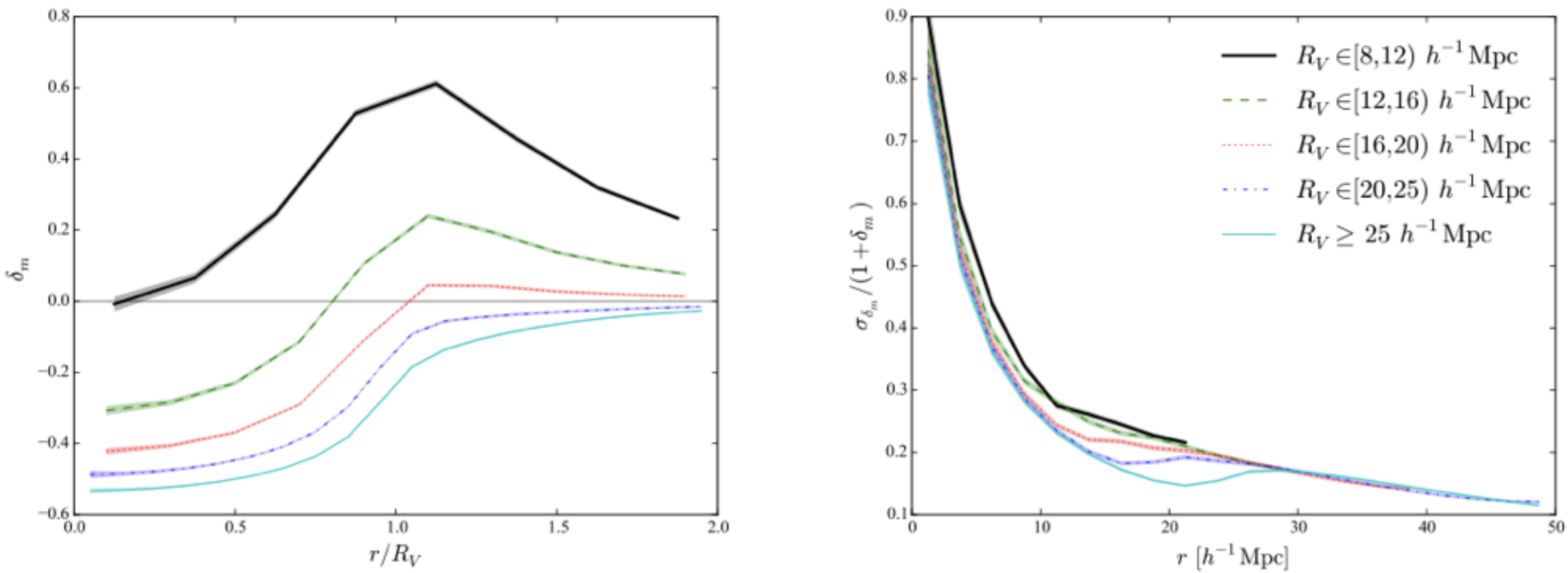
Zhao+15 arxiv:1511.04299

DIVE code applied



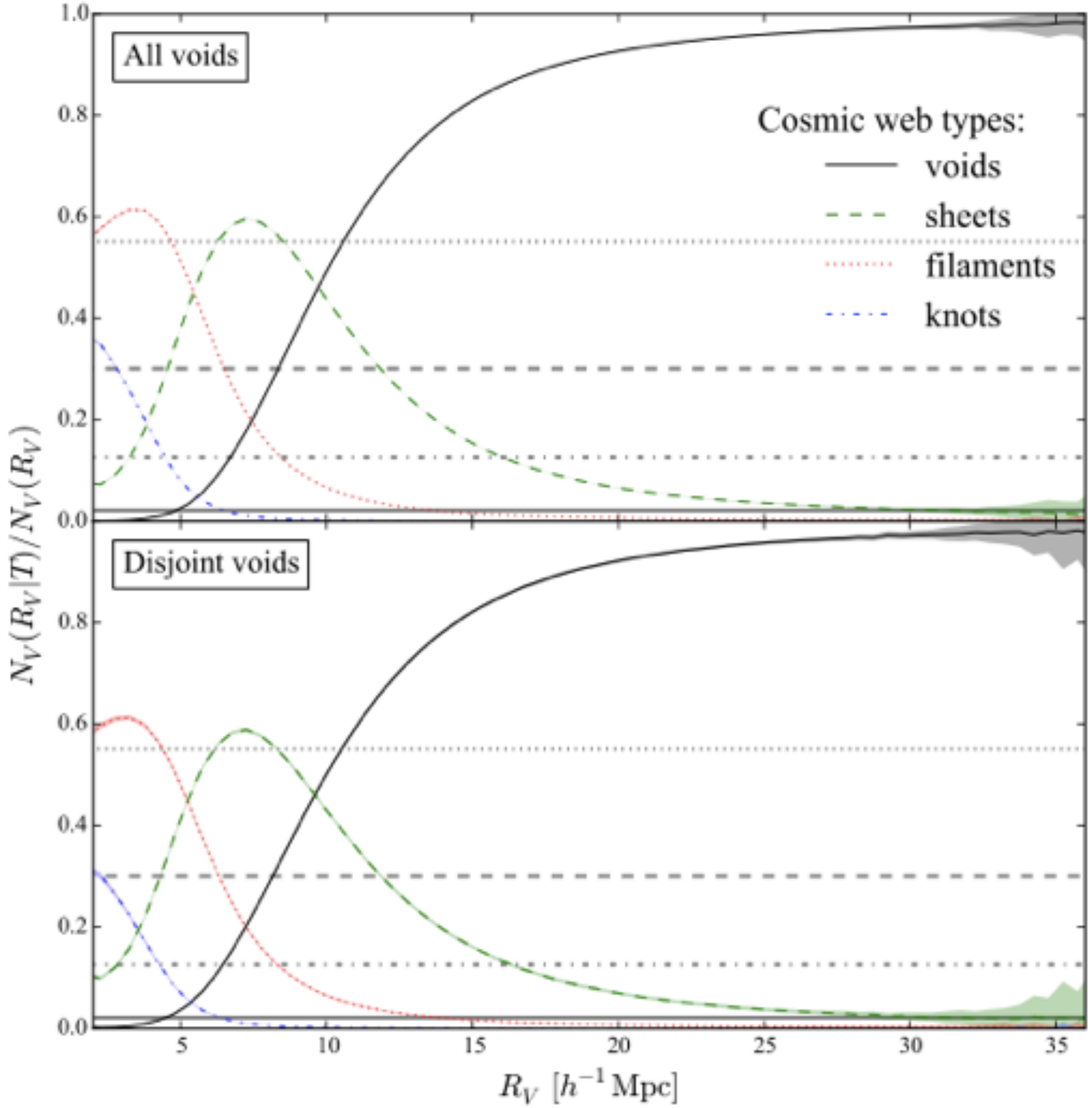
Zhao+15 arxiv:1511.04299

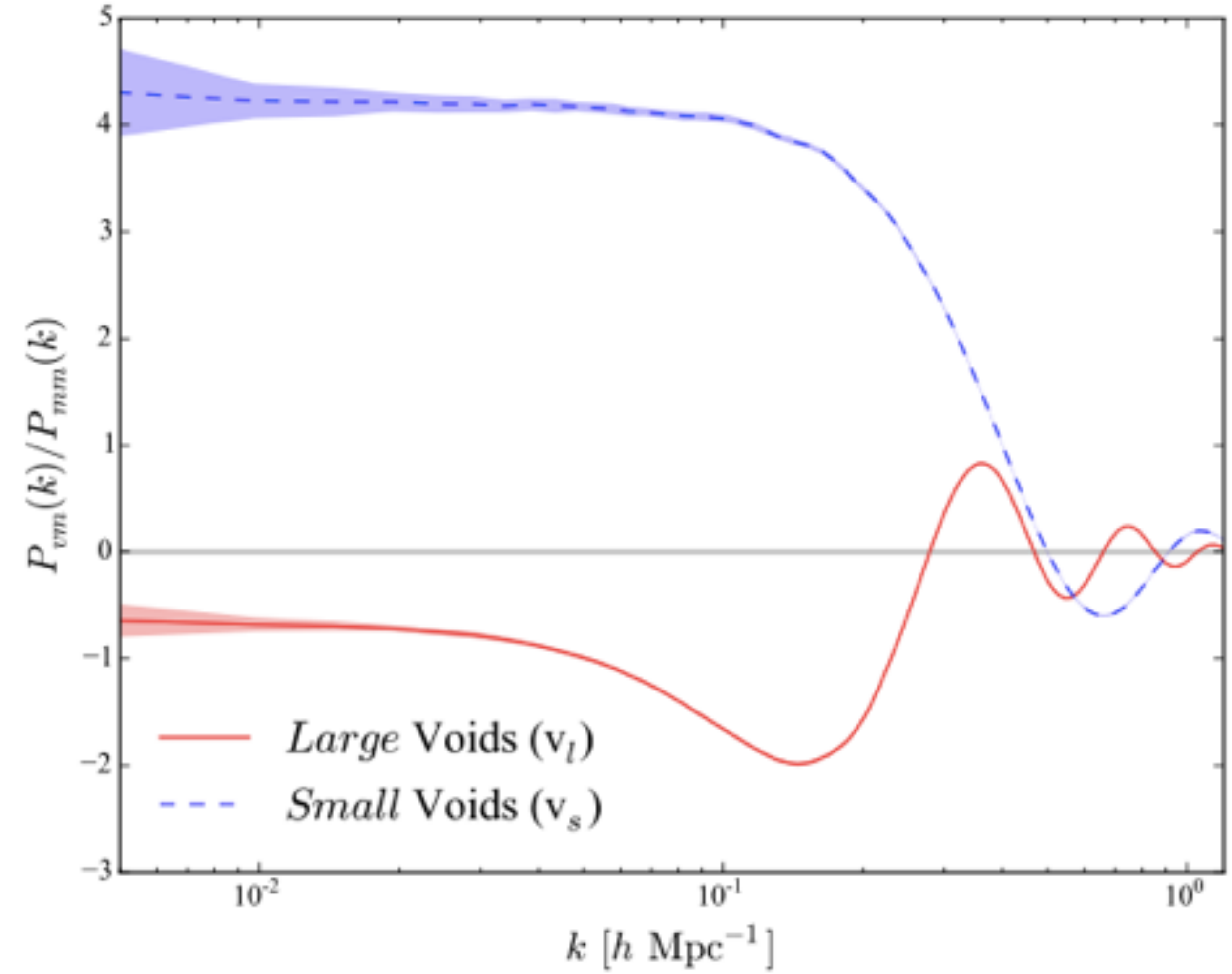
Tracing underdense regions



Zhao+15 arxiv:1511.04299

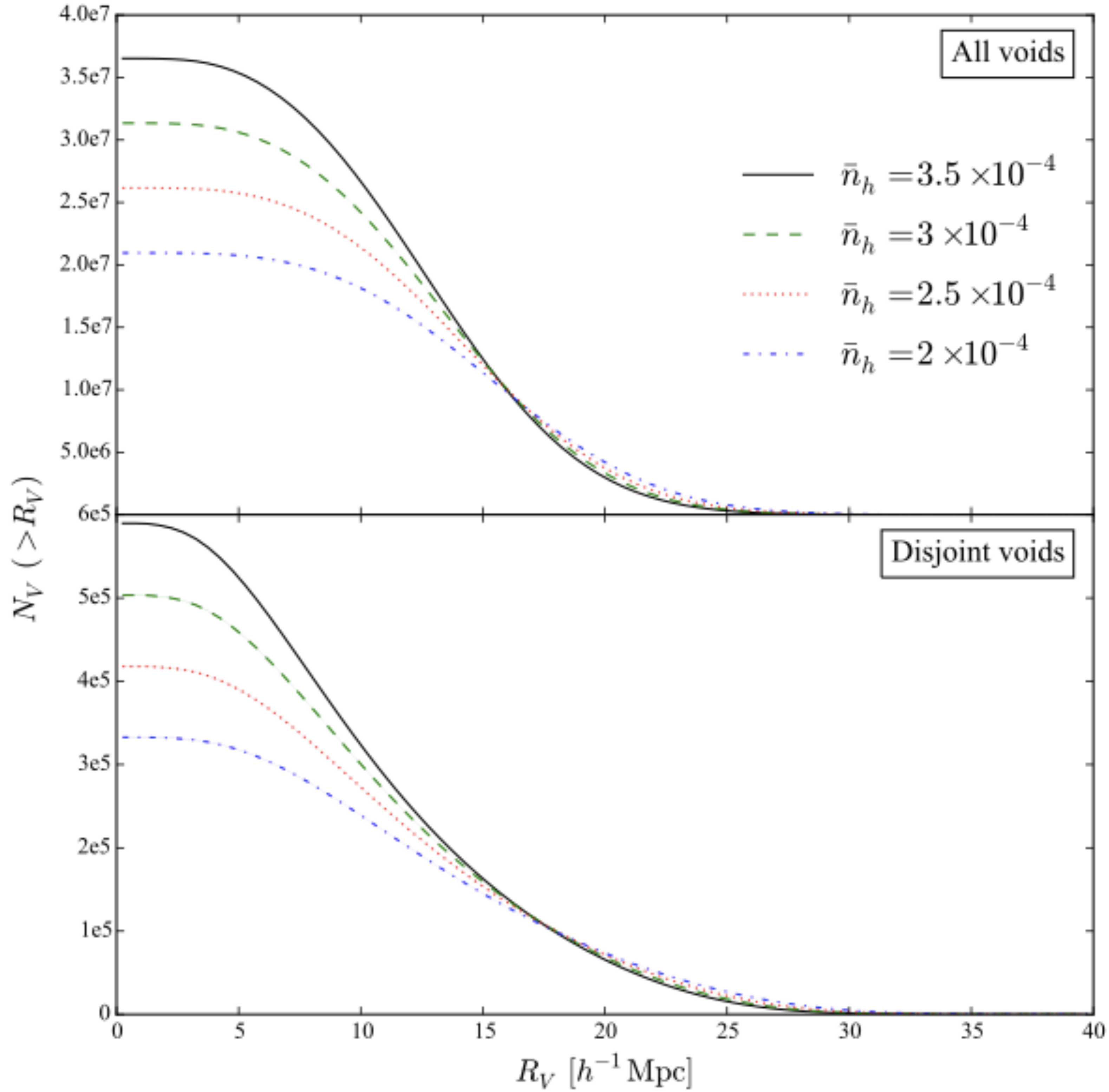
Tracing expanding regions





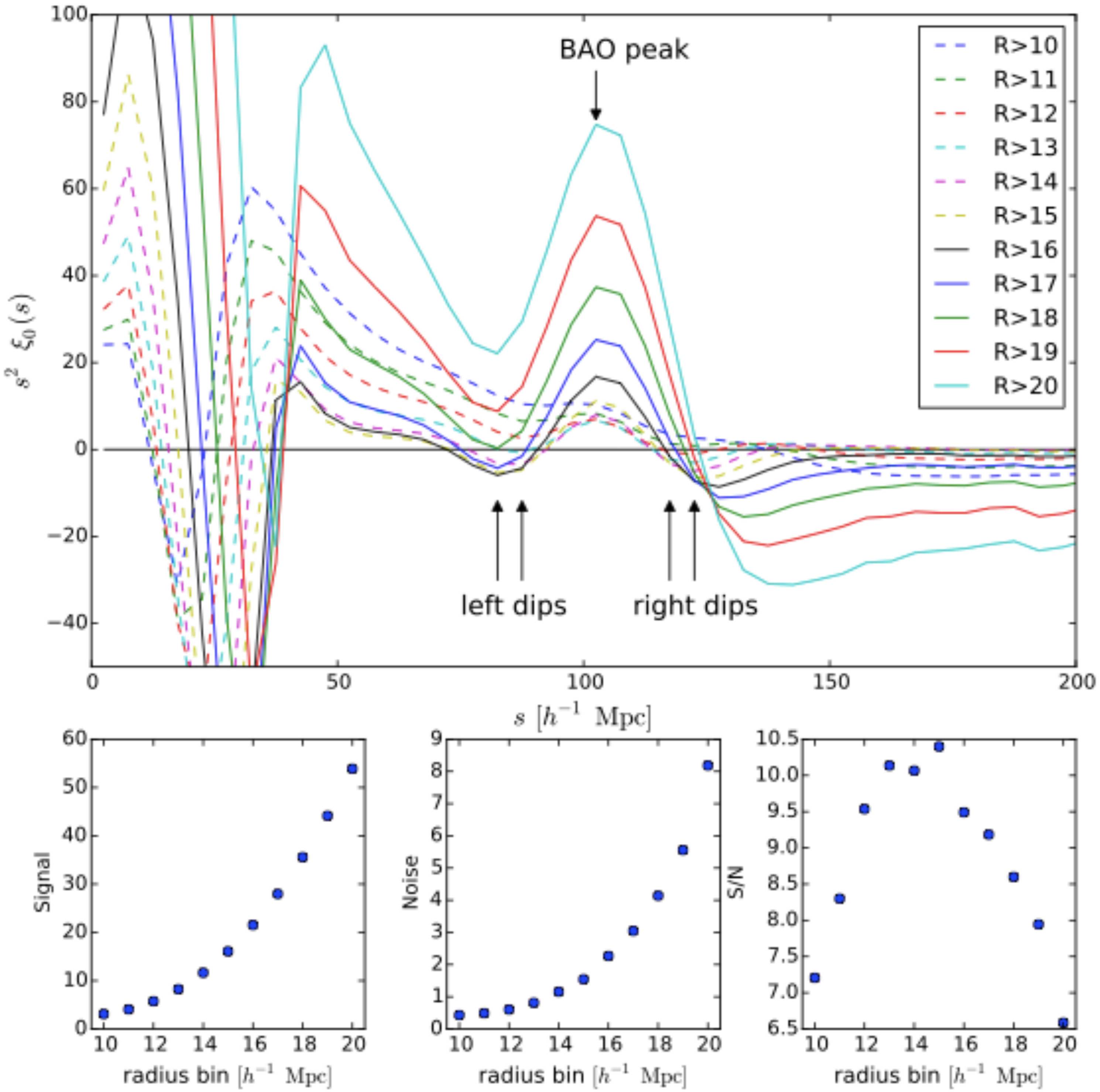
Zhao+15 arxiv:1511.04299

Optimal radius cut



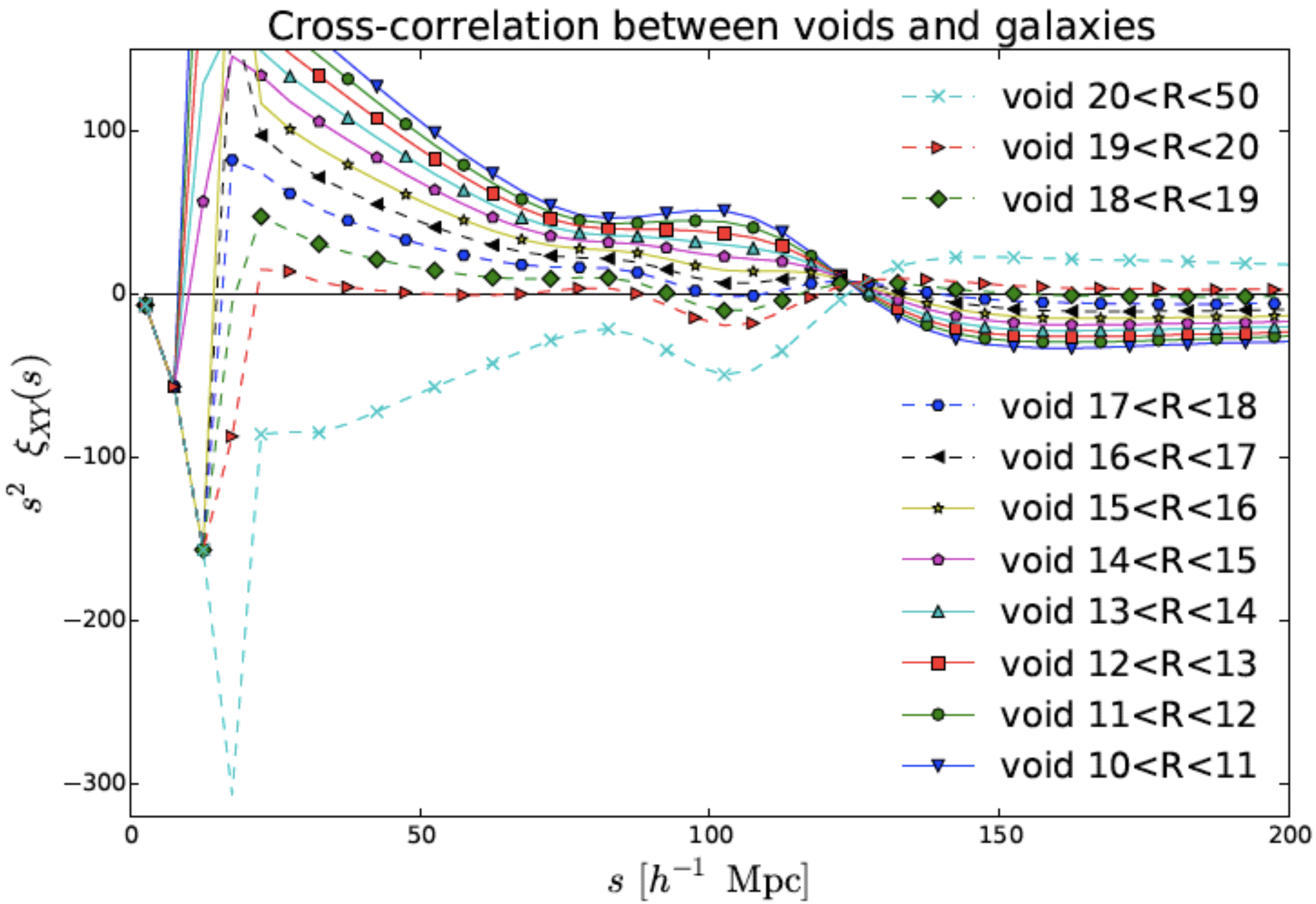
Zhao+15 arxiv:1511.04299

Optimal radius cut for BAO detection



Liang+15 arxiv:1511.04391

Optimal radius cut for BAO detection



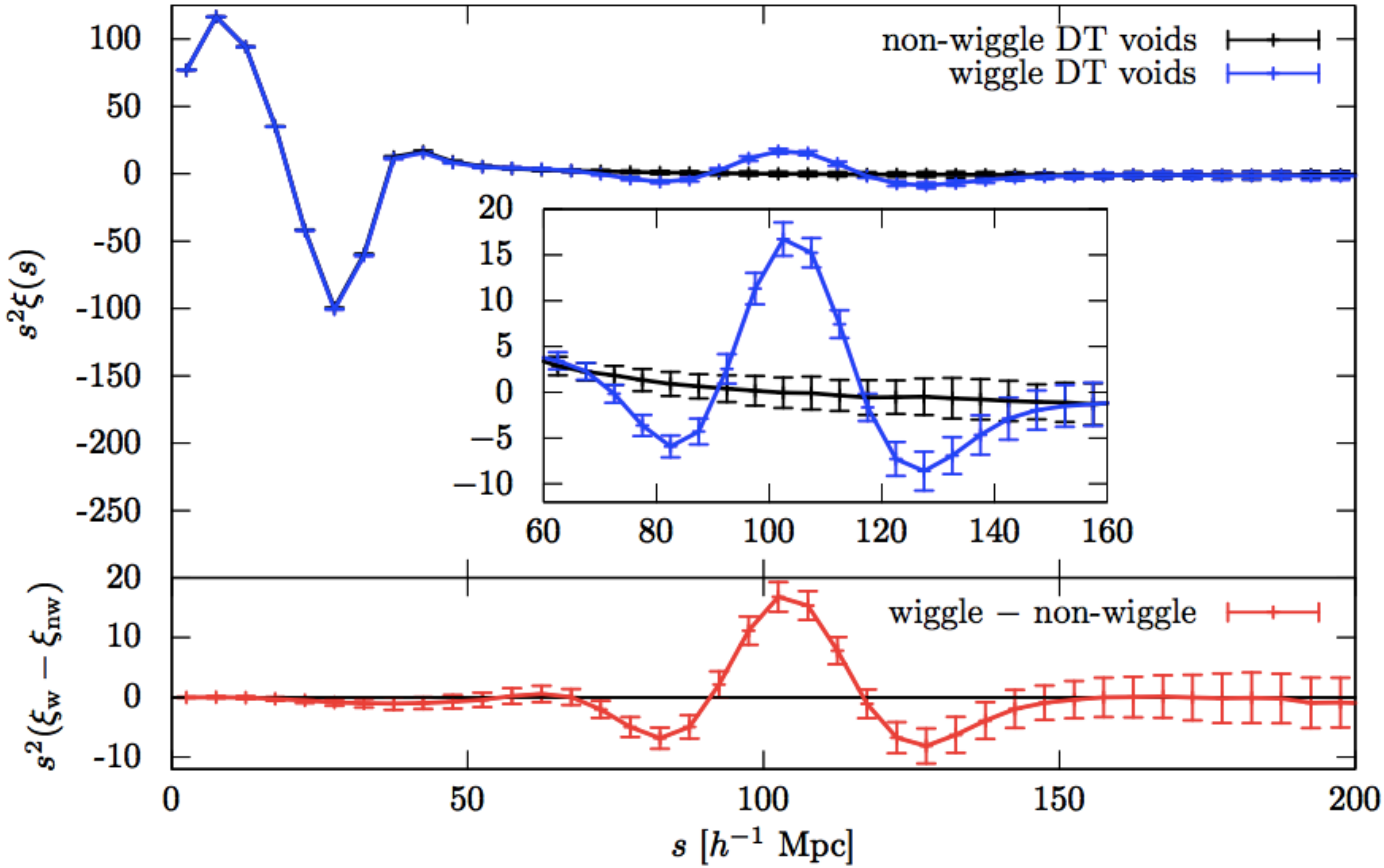
Liang+15 arxiv:1511.04391

BAO from voids in BOSS

Signatures of the primordial Universe from its emptiness

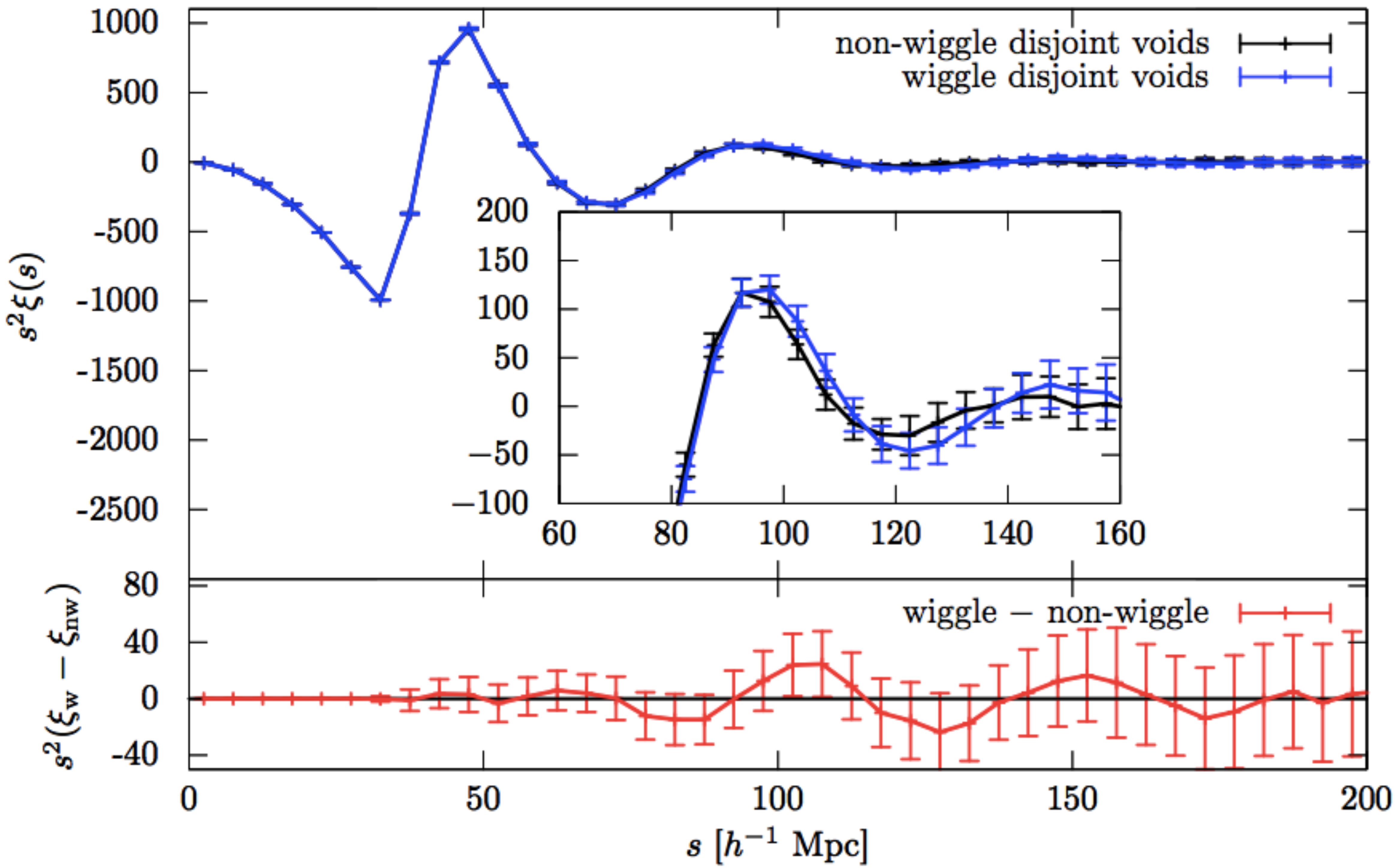
Francisco-Shu Kitaura[1*], Chia-Hsun Chuang[1], Yu Liang[2], Cheng Zhao[2], Charling Tao[3,2], Sergio Rodríguez-Torres[4,5,6], Daniel J. Eisenstein[7], Héctor Gil-Marín[8,9], Jean-Paul Kneib[10,11], Cameron McBride[7], Will J. Percival[12], Ashley J. Ross[12,13], Ariel G. Sánchez[14], Jeremy Tinker[15], Rita Tojeiro[16], Mariana Vargas-Magana[17], Gong-Bo Zhao[18,12]

BAO from voids in BOSS



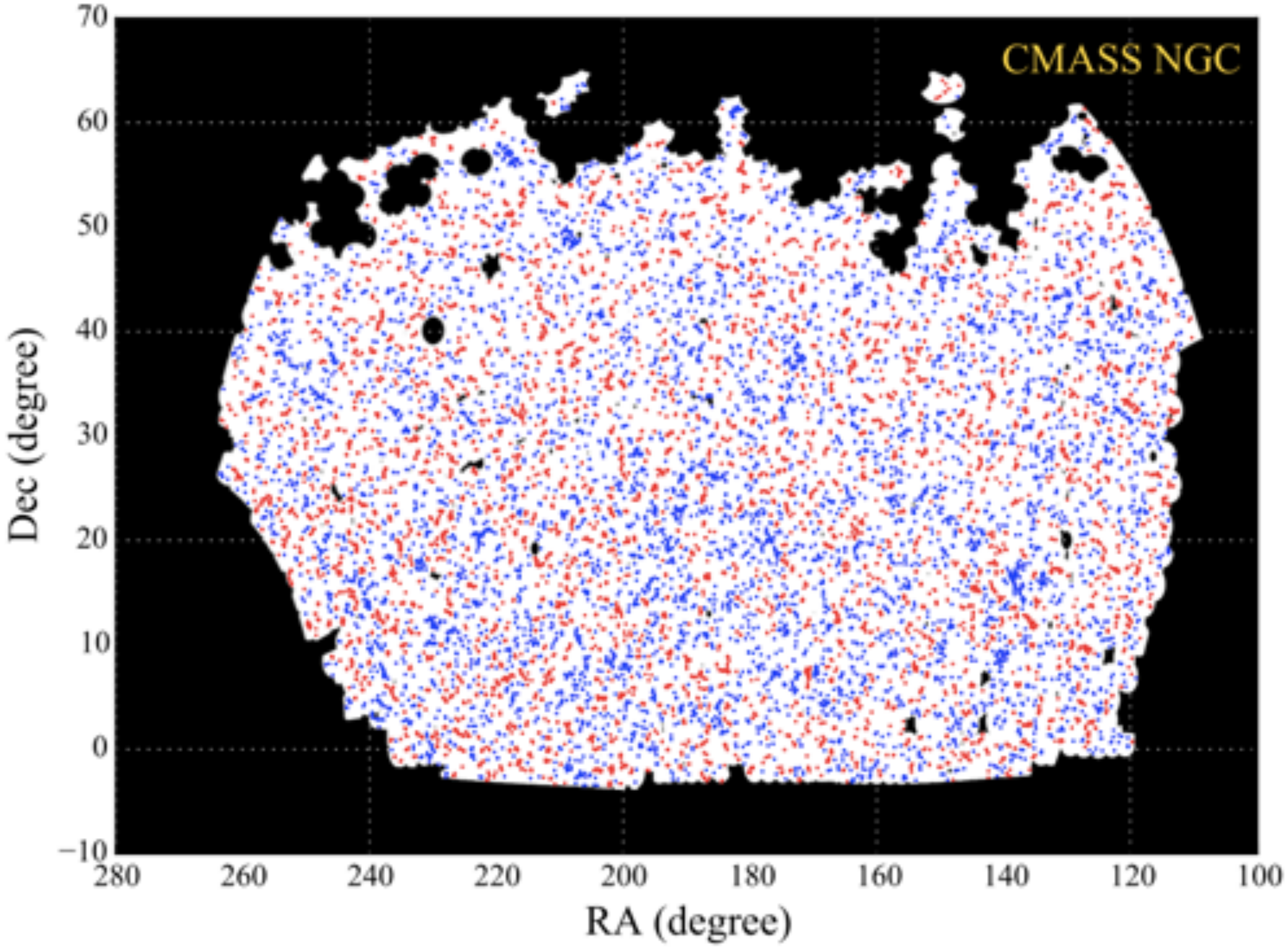
Kitaura+15 arxiv:1511.04405

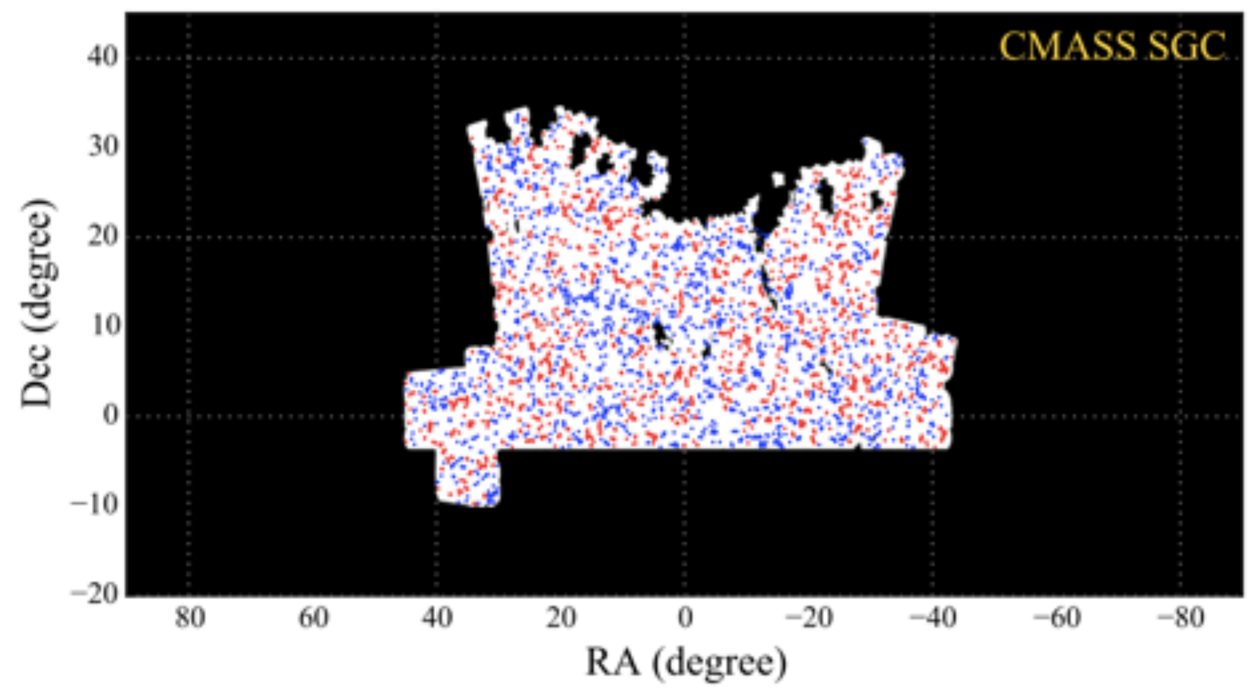
BAO from voids in BOSS



Kitaura+15 arxiv:1511.04405

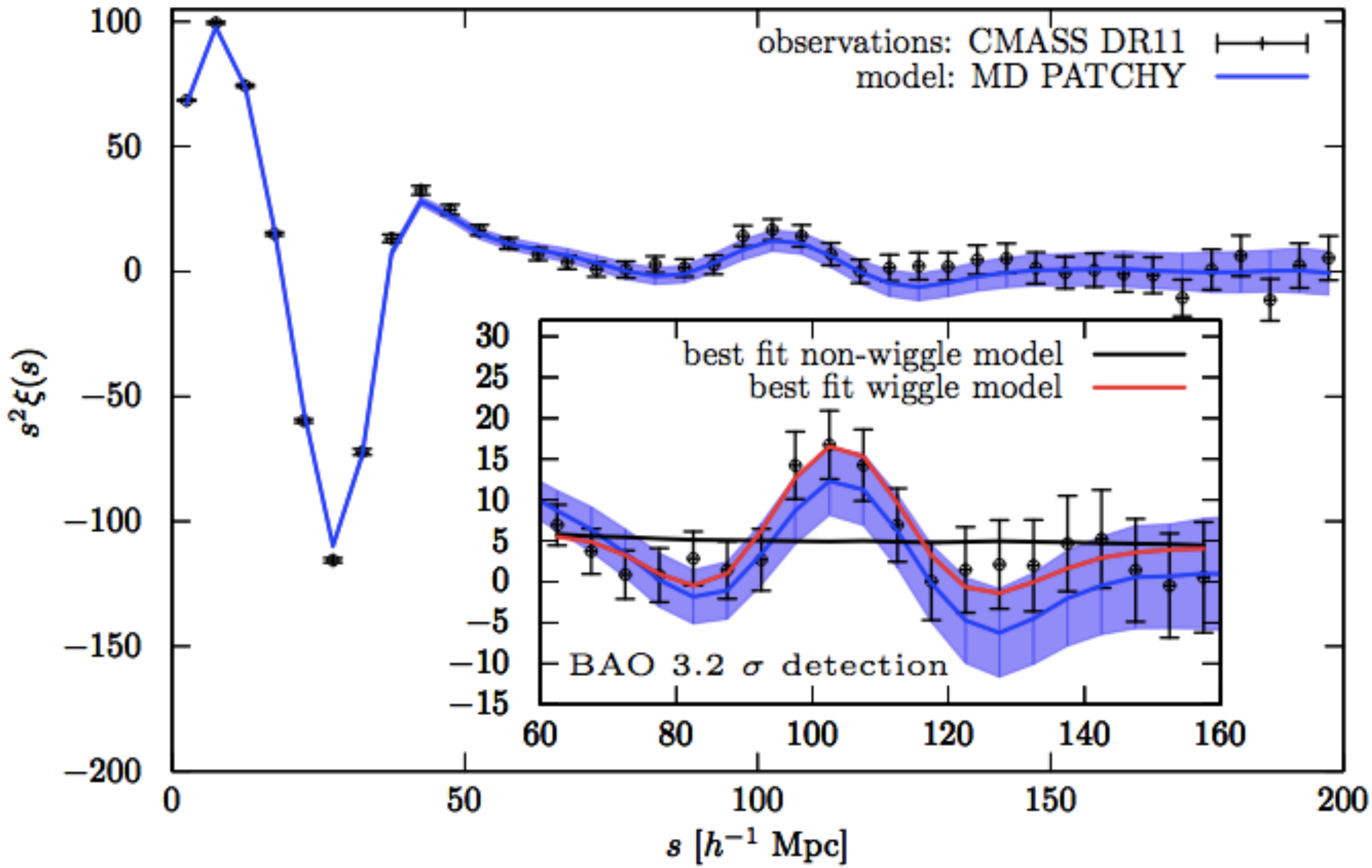
BAO from voids in BOSS



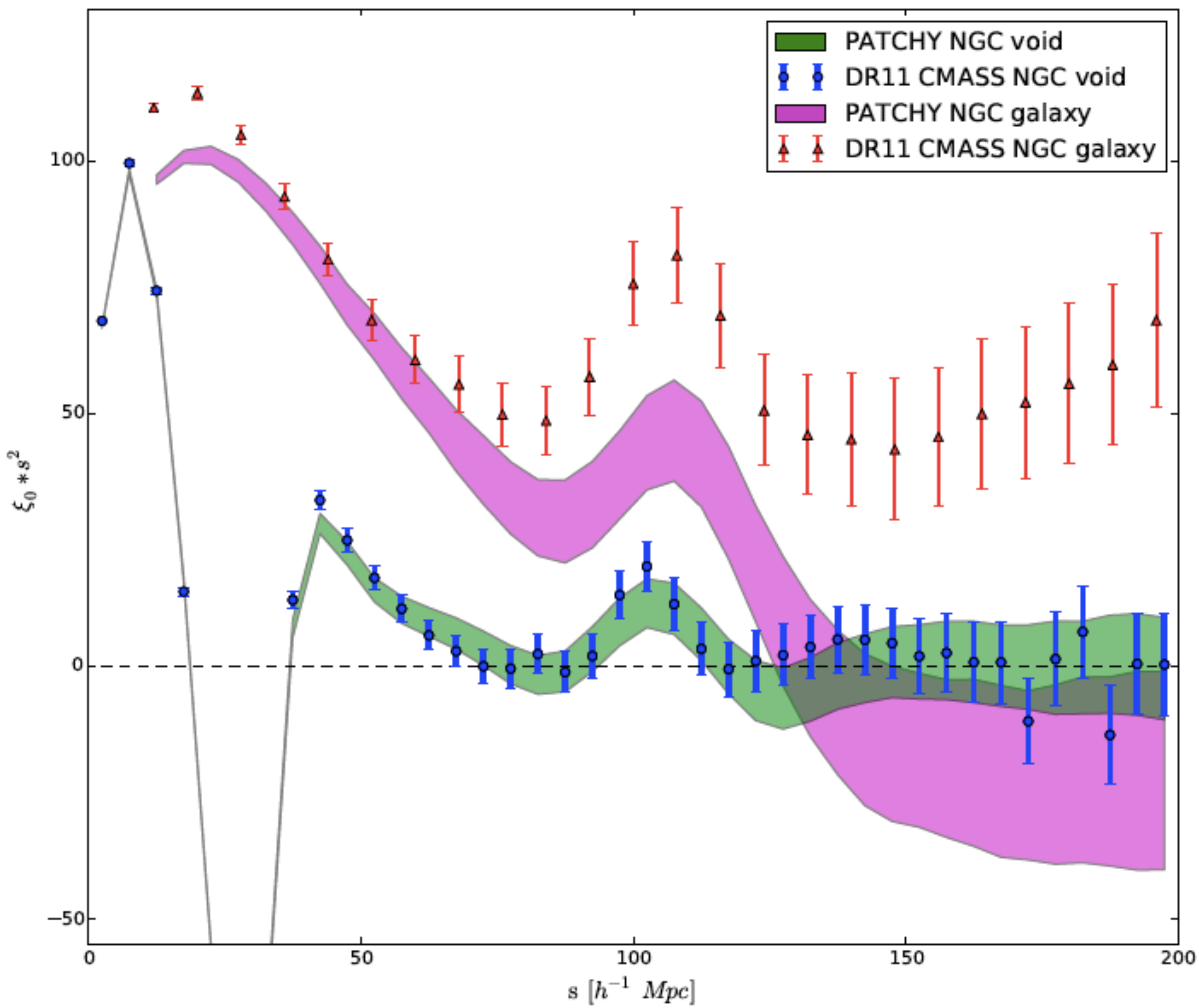


Kitaura+15 arxiv:1511.04405

BAO from voids in BOSS



Kitaura+15 arxiv:1511.04405



PATCHY NGC void
DR11 CMASS NGC void
PATCHY NGC galaxy
DR11 CMASS NGC galaxy
$\xi_0 * s^2$
$s\ [h^{-1}\ Mpc]$
100
50
0
−50
0
50
100
150
200

Conclusions

- We have presented ways of producing large sets of mock galaxy catalogues calibrated on N-body simulations.
- We have presented a measurement of the BAO from voids which matches accurately our models.
- The BAO from voids is essentially circumventing a more complex analysis, transferring information from the higher order statistics of galaxies to the two point correlation function analogously to BAO reconstruction (see Schmittfull +15) and the recent BAO detection from the 3-point statistics (Slepian+16)