Universidad de La Laguna

Connecting galaxies to the underlying dark matter field

Francisco-Shu Kitaura

IPMU 8/8/19

Team at the IAC

Andrés Balaguera-Antolínez
Sergio Rodríguez Torres
Andras Kovacs

PhD: Mónica Hernández Sánchez
Former PhD: Marcos Pellejero Ibanez

visitor: Metin Ata

collaborators: Chia-Hsun Chuang, Cheng Zhao, Gustavo Yepes, Raul Angulo, Claudio Dalla Vecchia, Kee-Ghan Lee, Ariel Sánchez, Rien van de Weygaert, Patrick Bos …

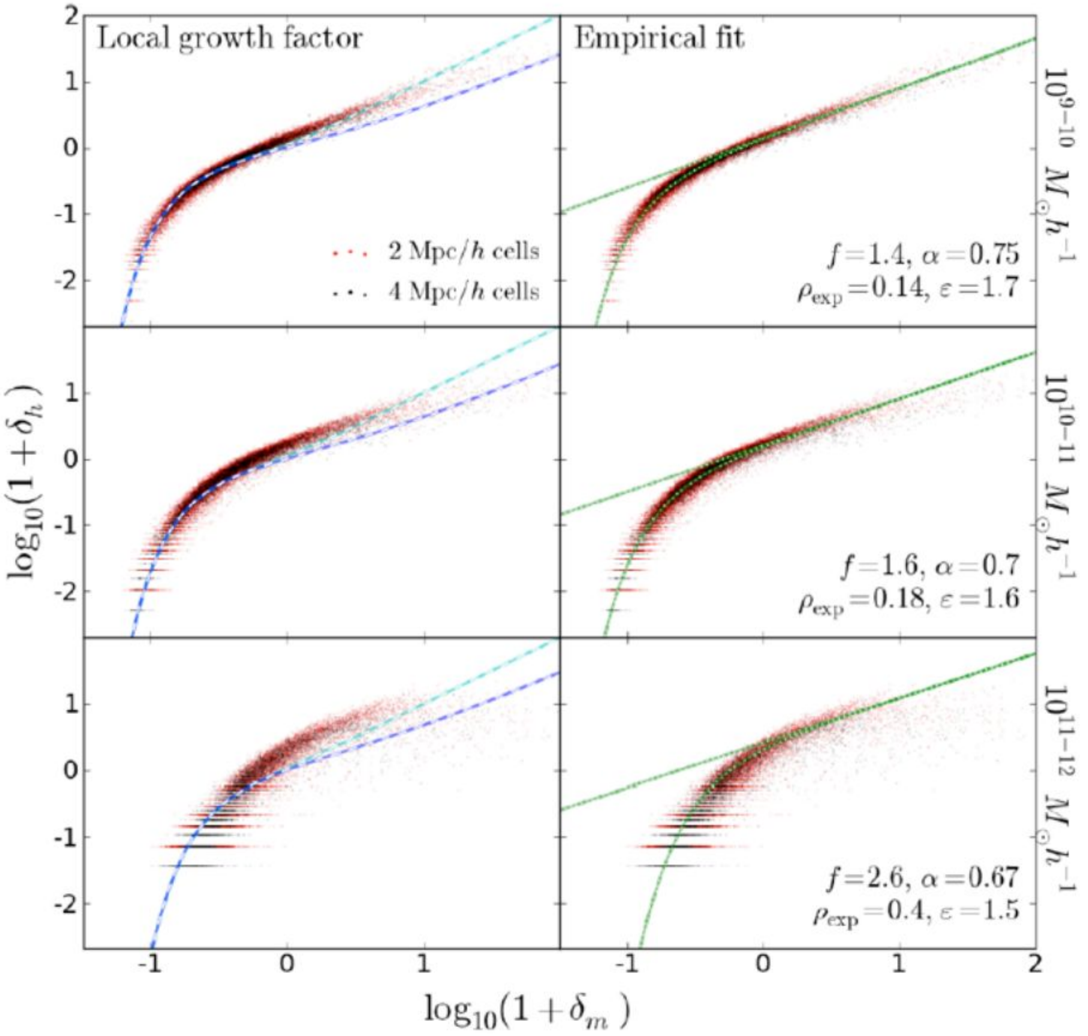
Excellence Severo Ochoa IAC Group Grant:	SEV-2015-0548
Personal Spanish National Grant:	RYC2015-18693
Spanish National LSS Project Grant:	AYA2017-89891-P

Effective Eulerian Bias

Kaiser 84
Fry & Gaztanaga 93
Cen & Ostriker 93
McDonald & Roy 09
de la Torre & Peacock 13
FSK et al 15

See review by Desjacques et al 2018

Neyrinck et al 14
Aragón-Calvo et al 14 (MIP sims)



Parametric bias model

Nonlinear deterministic bias

Kaiser 84
Fry & Gaztanaga 93
Cen & Ostriker 93
de la Torre & Peacock 13

FSK et al https://arxiv.org/abs/1407.1236

Stochastic component

Peebles 80
Saslaw 87
Sheth 95
Lahav & Lemson 99
FSK et al 14
Neyrinck et al 14

1st implementation of this bias model in a Bayesian framework:
Metin Ata, FSK & Mueller 2015: https://arxiv.org/abs/1408.2566

$$\langle\rho_{\mathrm{h}}\rangle_{\mathrm{d}V} = f_{\mathrm{h}}\, B(\rho_{\mathrm{h}}|\rho_{\mathrm{m}})\,,$$

$$f_{\mathrm{h}} = \frac{n_{\mathrm{h}}}{\langle B(\rho_{\mathrm{h}}|\rho_{\mathrm{m}})\rangle_V}\,,$$

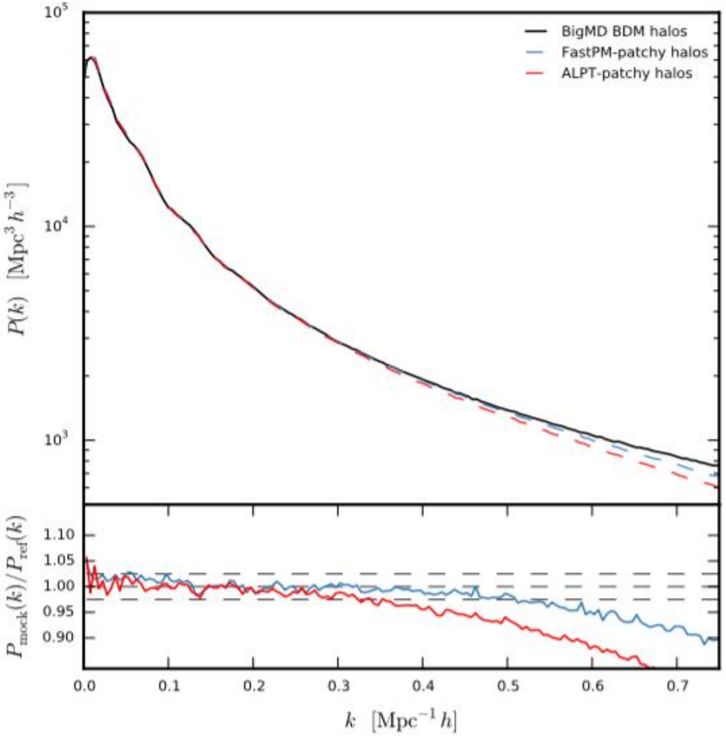
$$\begin{aligned} B(\rho_{\mathrm{h}}|\rho_{\mathrm{m}}) &= \underbrace{\rho_{\mathrm{m}}^{\alpha}}_{\text{nonlinear bias}} \\ &\times \underbrace{\theta(\rho_{\mathrm{m}}-\rho_{\mathrm{th}})}_{\text{threshold bias}} \times \underbrace{\exp\left(-(\rho_{\mathrm{m}}/\rho_{\epsilon})^{\epsilon}\right)}_{\text{exponential cutoff}}\,, \end{aligned}$$

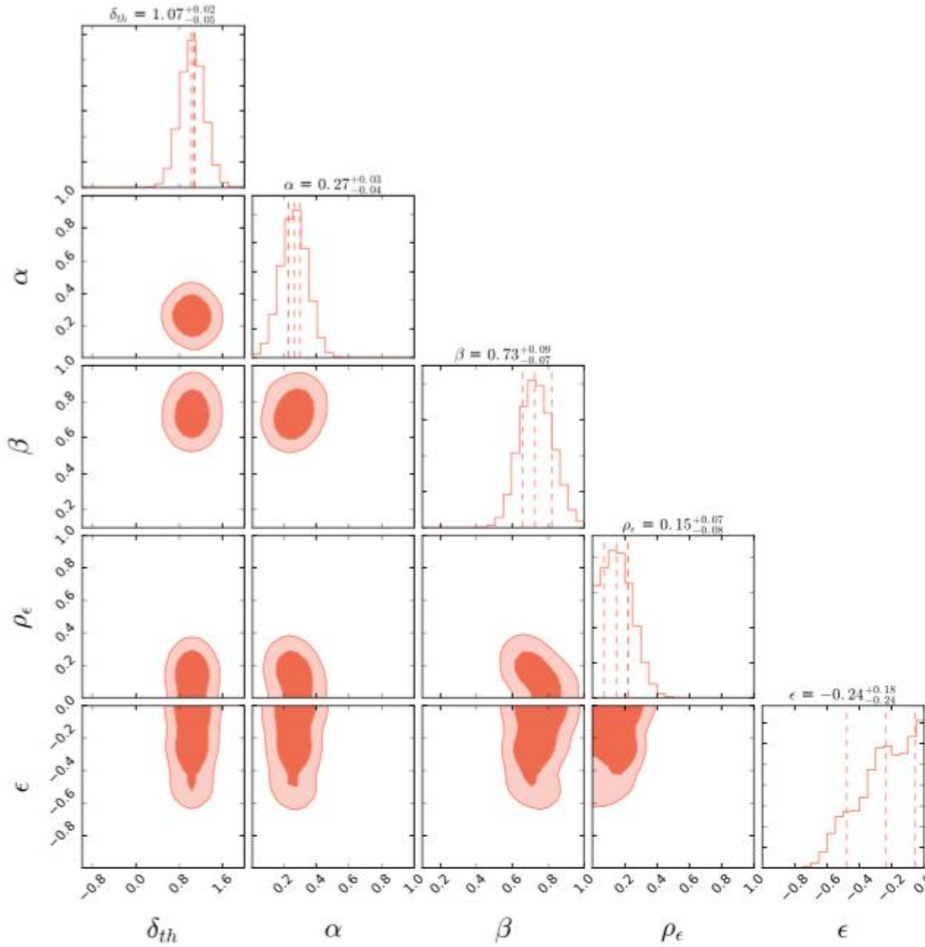
$$\begin{aligned} P(N_{\mathrm{h}}|\lambda_{\mathrm{h}},\beta) &= \underbrace{\frac{\lambda_{\mathrm{h}}^{N_{\mathrm{h}}}}{N_{\mathrm{h}}!}e^{-\lambda_{\mathrm{h}}}}_{\text{Poisson distribution}} \\ &\times \underbrace{\frac{\Gamma(\beta+N_{\mathrm{h}})}{\Gamma(\beta)(\beta+\lambda_{\mathrm{h}})^{N_{\mathrm{h}}}} \times \frac{e^{\lambda_{\mathrm{h}}}}{(1+\lambda_{\mathrm{h}}/\beta)^{\beta}}}_{\text{Deviation from Poissonity}}\,. \end{aligned}$$

Non-local bias is missing! McDonald & Roy 09

PATCHY code

FSK, Yepes & Prada 14
https://arxiv.org/abs/1307.3285

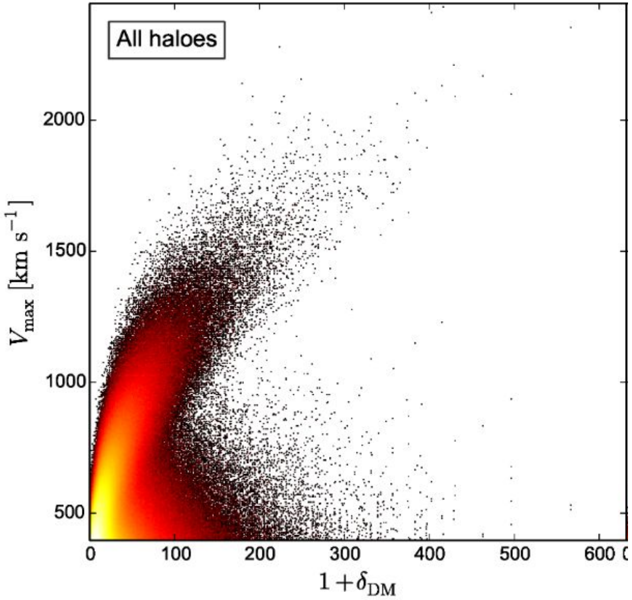




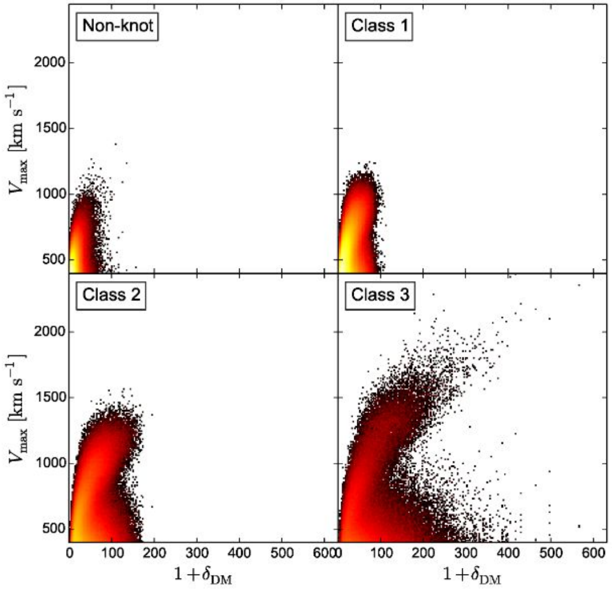
Mohammadjavad Vakili, FSK et al 17 ALPT, fastPM, MCMC automatic bias calibration

Halo mass dependence on the Cosmic Web

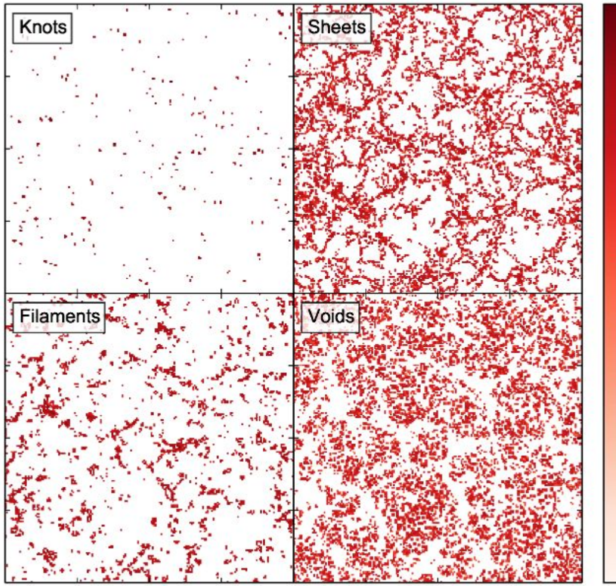
Local density



Mass of percolated knot region



Tidal field tensor



HADRON code Cheng Zhao, FSK et al https://arxiv.org/abs/1501.05520

Comparison project within EUCLID

special Credit to: Linda Blot (et al 19), Martha Lippich (et al 19), Manuel Colavincenzo (et al 19)
Ariel Sánchez, Martín Crocce, Emiliano Sefusatti, Pierluigi Monaco, Claudio Dalla Vecchia

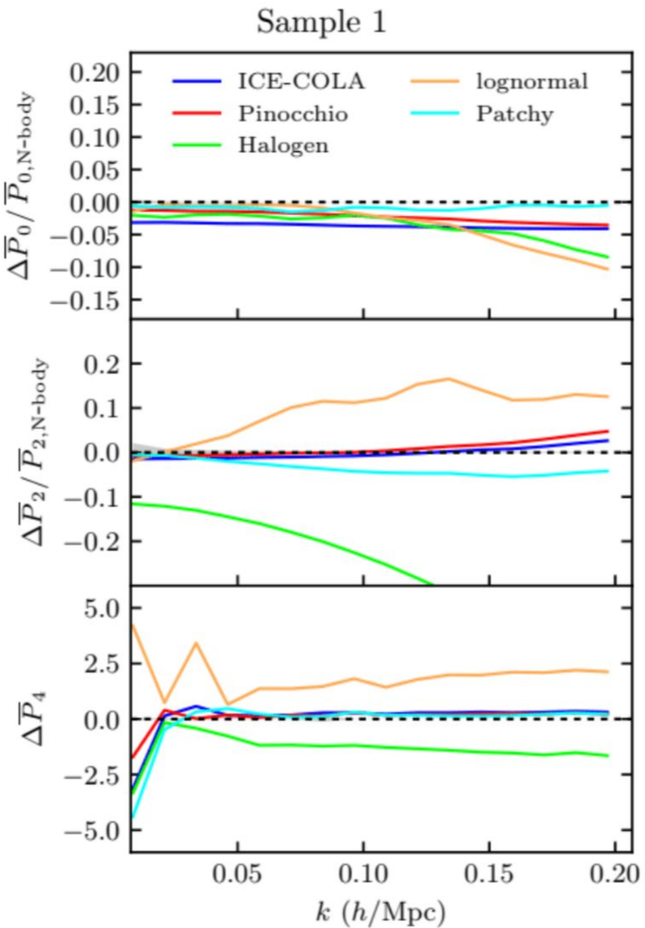
Method	Algorithm	Computational Requirements	Reference
Minerva	**N-body** Gadget-2 Halos : SubFind	CPU Time: 4500 hours Memory allocation: 660 Gb	Grieb et al. (2016) *https://wwwmpa.mpa-garching.mpg.de/gadget/*
ICE-COLA	Predictive 2LPT + PM solver Halos : FoF(0.2)	CPU Time: 66 hours Memory allocation: 340 Gb	Izard, Crocce & Fosalba (2016) Modified version of: *https://github.com/junkoda/cola_halo*
PINOCCHIO	Predictive 3LPT + ellipsoidal collapse Halos : ellipsoidal collapse	CPU Time: 6.4 hours Memory allocation: 265 Gb	Monaco et al. (2013); Munari et al. (2017b) *https://github.com/pigimonaco/Pinocchio*
PEAKPATCH	Predictive 2LPT + ellipsoidal collapse Halos : Spherical patches over initial overdensities	CPU Time: 1.72 hours* Memory allocation: 75 Gb*	Bond & Myers (1996a,b,c) Not public
HALOGEN	Calibrated 2LPT + biasing scheme Halos : exponential bias	CPU Time: 0.6 hours Memory allocation: 44 Gb Input: $\bar{n}$, 2-pt correlation function halo masses and velocity field	Avila et al. (2015). *https://github.com/savila/halogen*
PATCHY	Calibrated ALPT + biasing scheme Halos : non-linear, stochastic and scale-dependent bias	CPU Time: 0.2 hours Memory allocation: 15 Gb Input: $\bar{n}$, halo masses and environment	Kitaura, Yepes & Prada (2014) Not Public

Reference N-body simulations
4500 CPU hrs
660 Gb

approximate analytic solvers
2 orders of magnitude less CPU hrs
same order of magnitude memory requirements

PATCHY code
4 orders of magnitude less CPU hrs
2 orders of magnitude less memory requirements

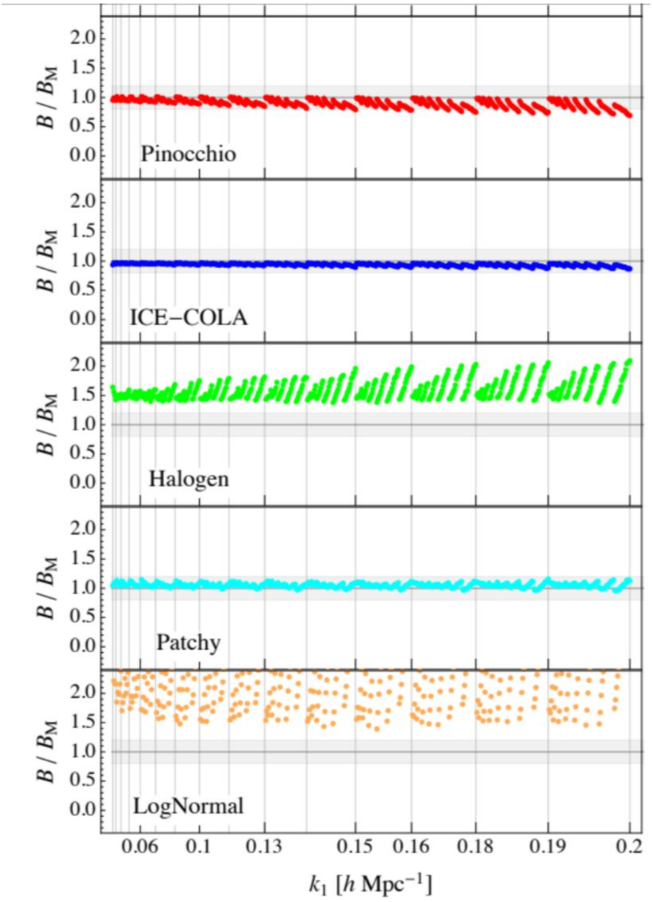
Comparison project within EUCLID



Most accurate method in 2-point statistics

Quadrupole can be improved tuning in the dispersion velocity term

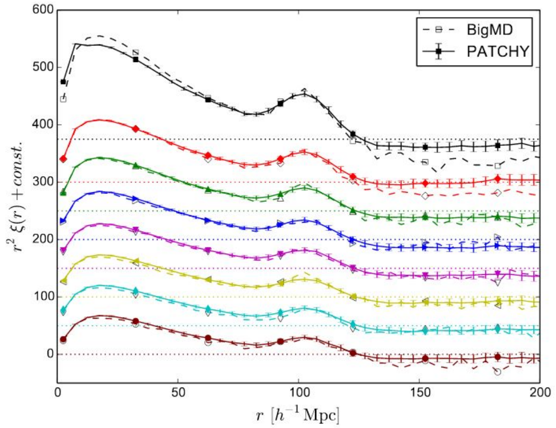
At the level of N-body solvers in 3-point statistics



Cosmic Web applied to mock halo and galaxy catalogs

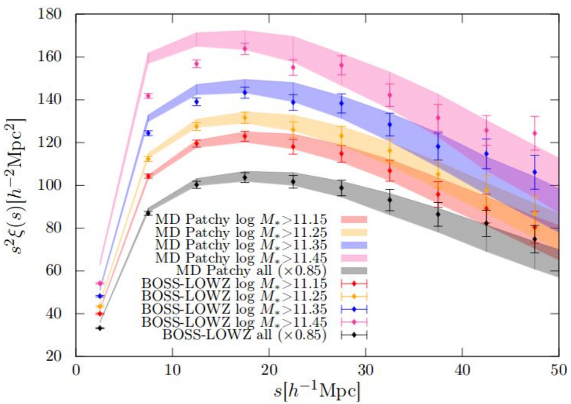
Including also exclusion effect

FSK, Rodriguez-Torres, Chuang et al
https://arxiv.org/abs/1509.06400

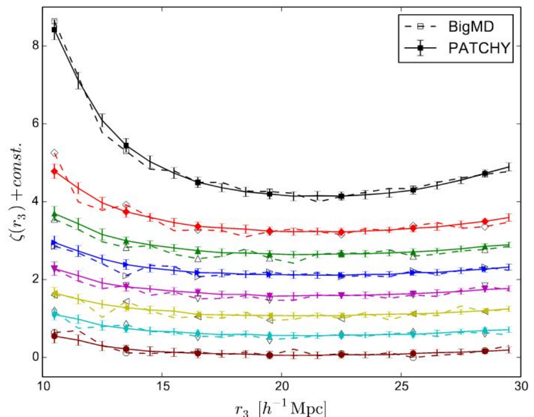


2-point statistics
in real-space
halo mass cuts

CMASS

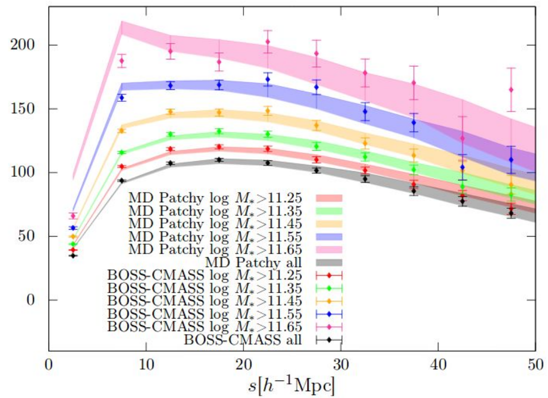


2-point statistics
in redshift-space
stellar mass cuts



3-point statistics
in real-space
halo mass cuts

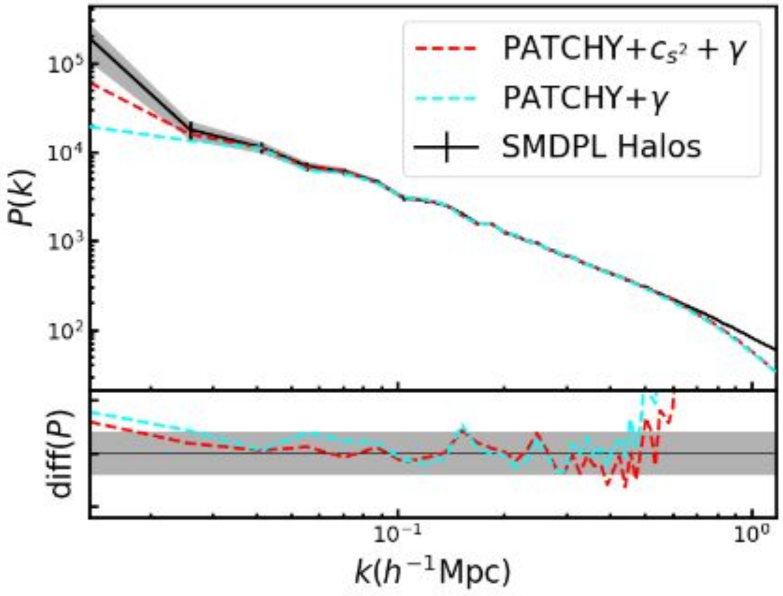
LOWZ

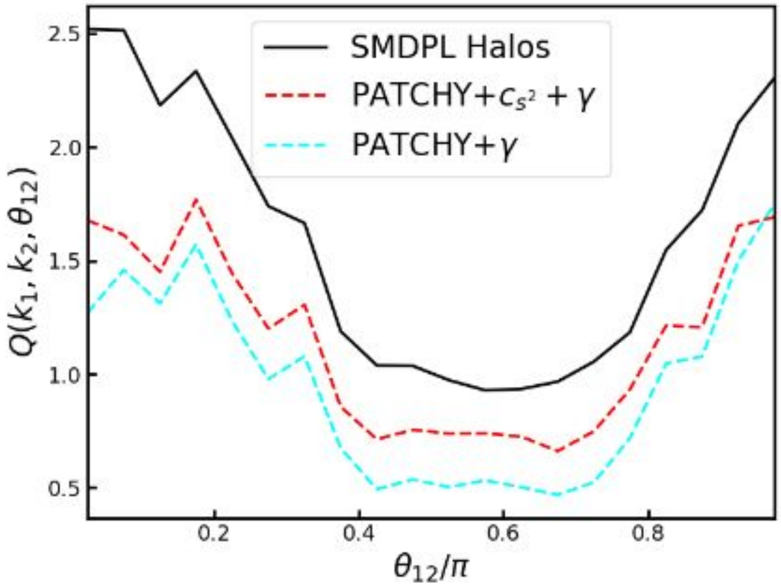


HADRON code Cheng Zhao, FSK et al
https://arxiv.org/abs/1501.05520

Going to lower halo masses for EUCLID and DESI

PATCHY including non-local bias, deviation from Poissonity etc fails!

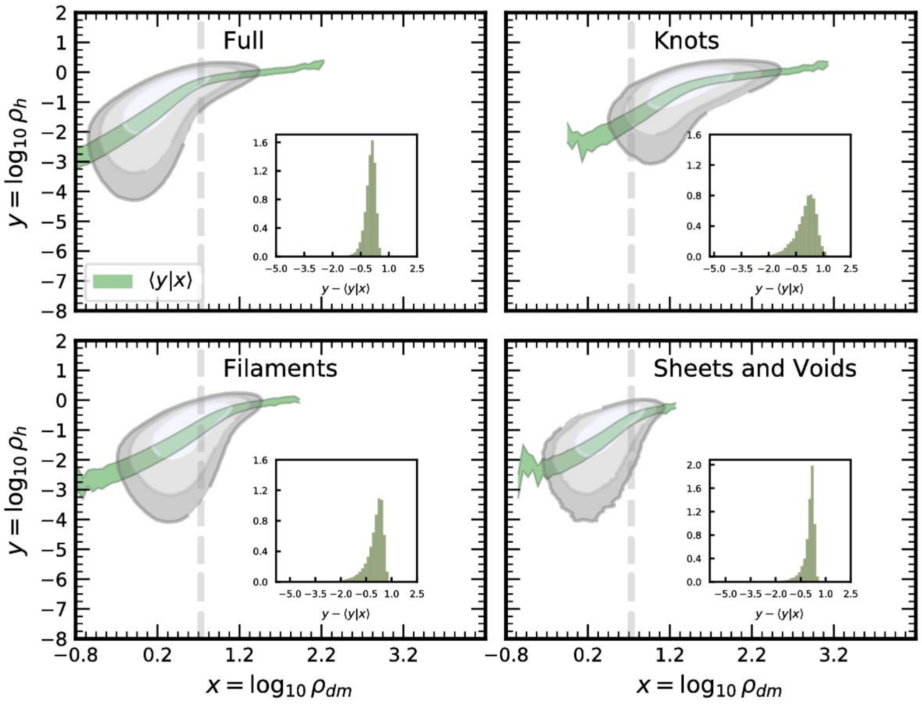


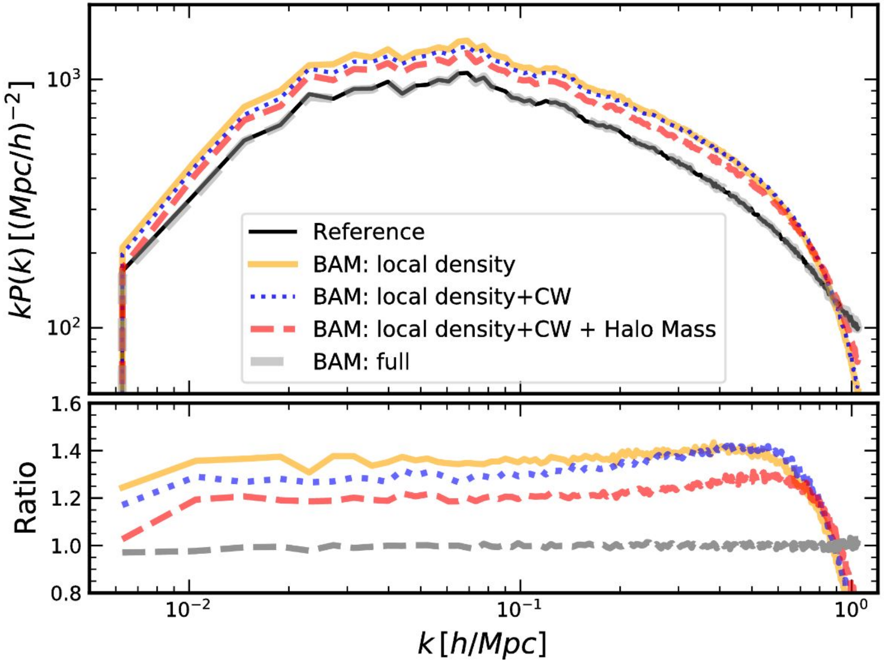


$M > 2.5 \times 10^{10} \mathrm{M}_{\odot} h^{-1}$

Marcos Pellejero-Ibanez et al to be submitted

Why not directly map the bias relation from accurate calculations?
Bias Assignment Method: BAM percentage accuracy up to the Nyquist frequency
non-parametric bias model

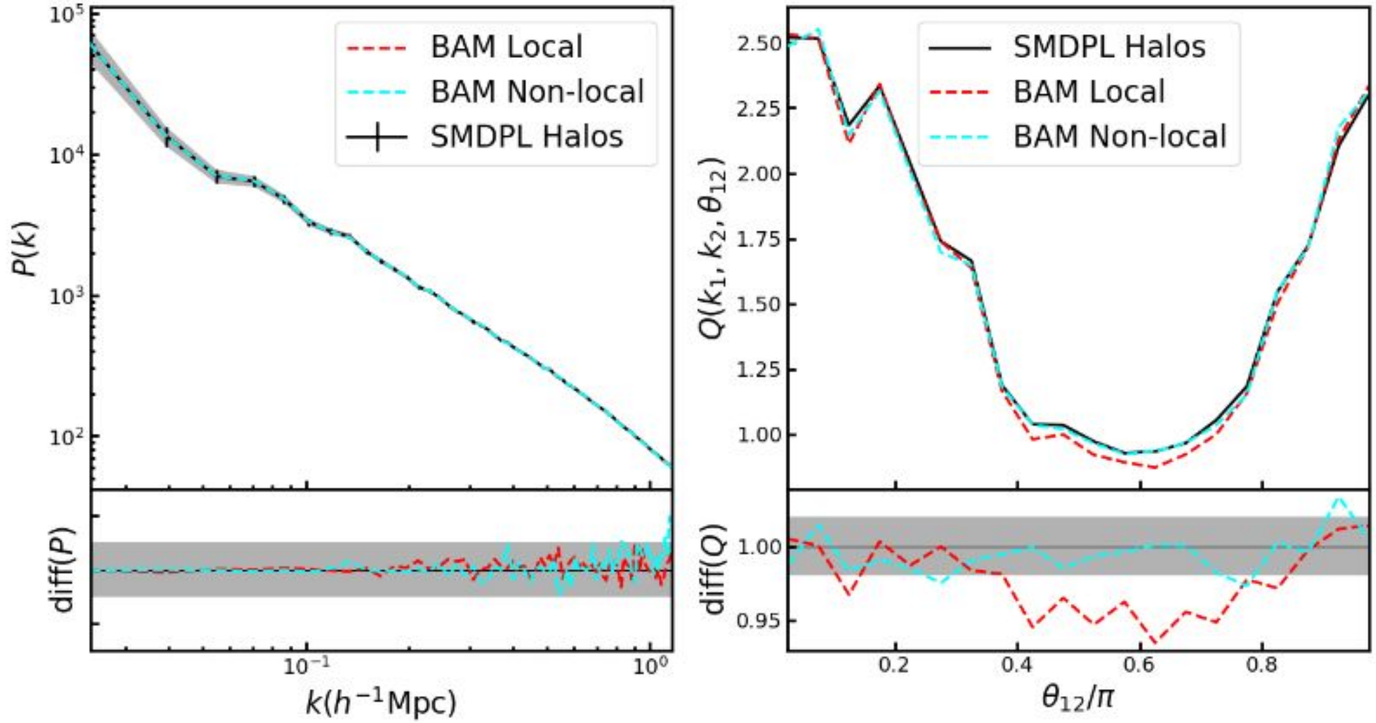




Learn the effective bias relation from N-body simulations including local and nonlocal dependencies.
Machine learning component with a kernel as a function of k

Andres Balaguera-Antolinez, FSK, Pellejero-Ibanez, Zhao, & Abel 19 arXiv:1806.05870

Performance of BAM using 160^3 particles instead of 3840^3 SMD sim
ALPT with phase space mapping (Shandarin 12; Abel et al 12; Hahn et al 13)

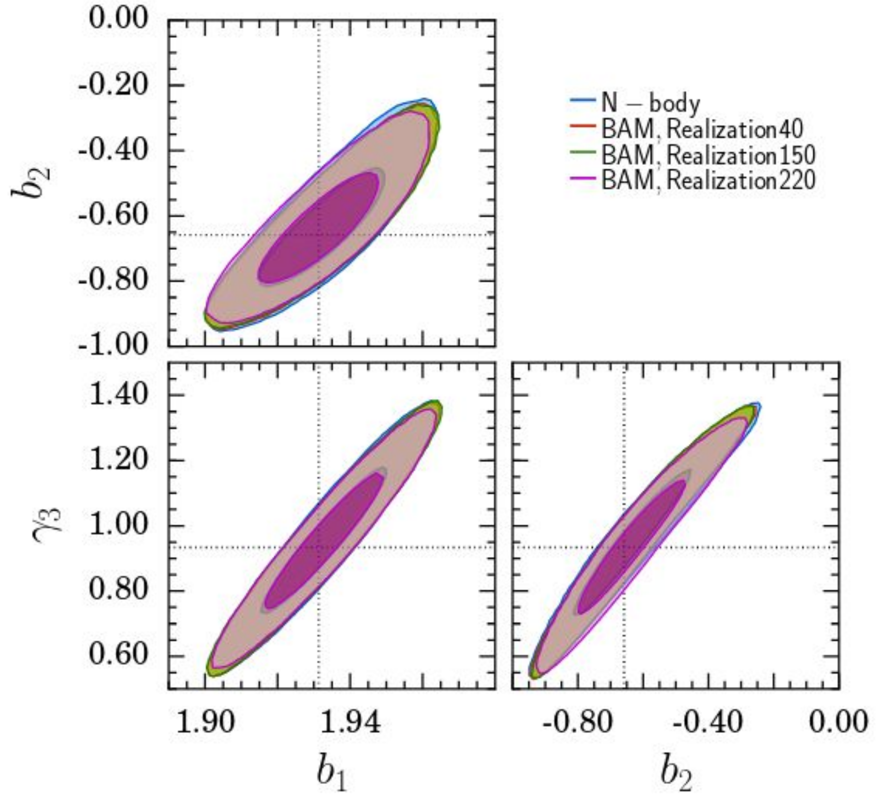


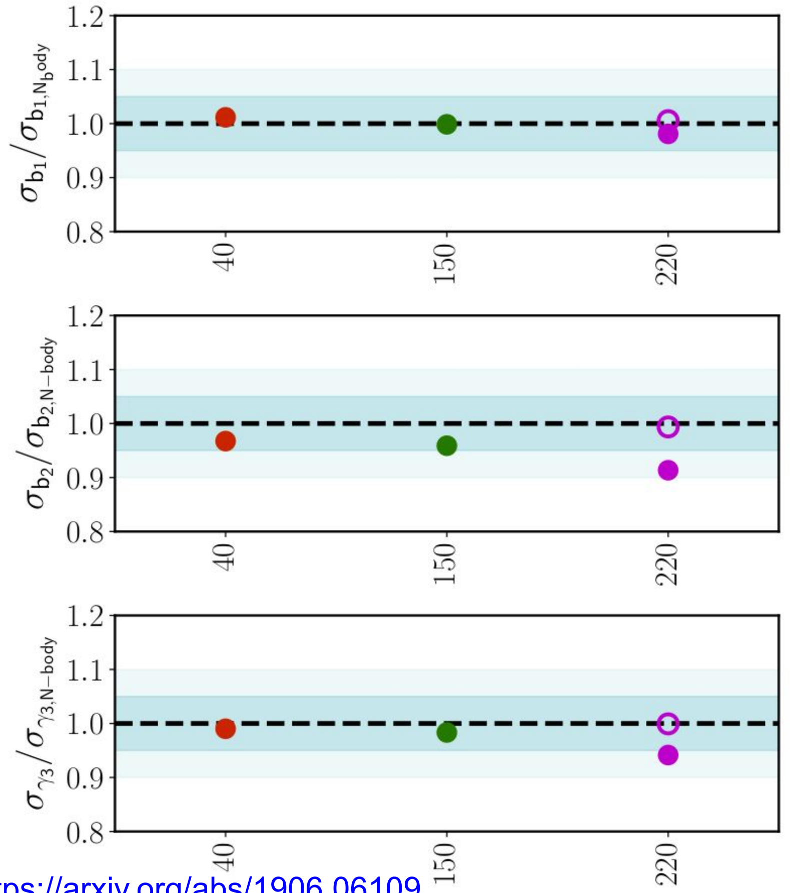
$M > 2.5 \times 10^{10} \mathrm{M}_{\odot} h^{-1}$

Marcos Pellejero-Ibanez et al to be submitted

BAM Covariance matrices

we will make mocks for EUCLID, DESI, JPAS, etc





Andres Balaguera-Antolinez, FSK, Pellejero-Ibanez, et al 19 https://arxiv.org/abs/1906.06109

Can we infer the dark matter field from the galaxy distribution?

Forward modelling with Bayesian approaches taking advantage of the simple statistics (Gaussian) of the primordial Universe.

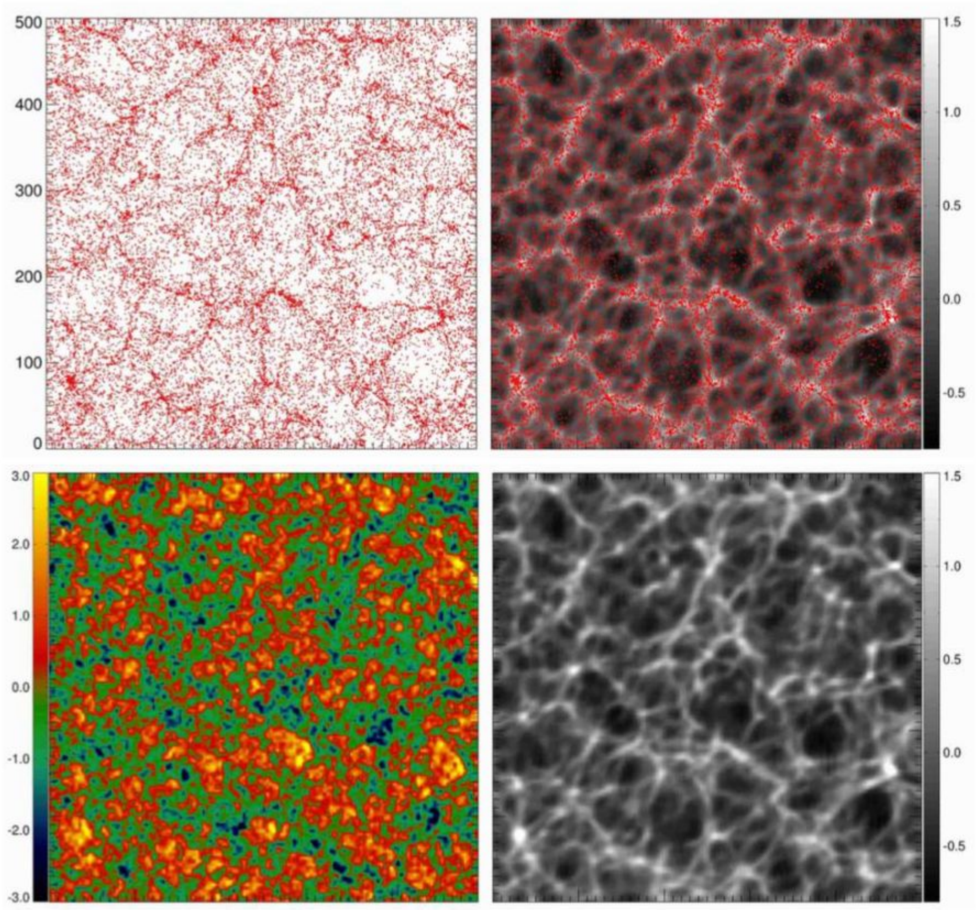
Solution to multi-streaming, as opposed to inverse approaches

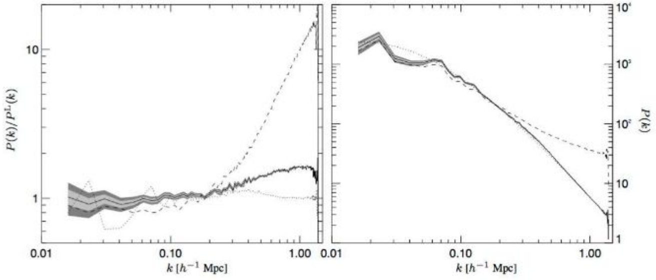
BARCODE	Bos, FSK, Weygaert	18
BORG	Jasche, Wandelt & co	13
ELUCID	Wang, Mo & co	13
KIGEN	FSK	13

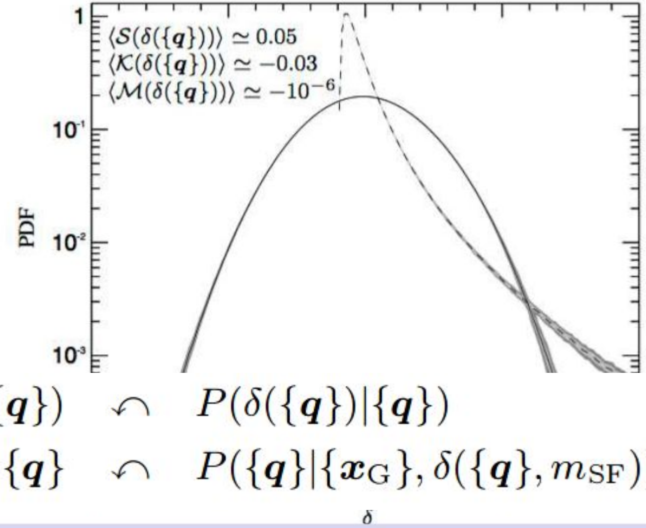
BORG and KIGEN appeared at the same time as independent works
ELUCID, BARCODE appeared later based on them
New BIRTH code FSK to be submitted

A forward modelling code KIGEN

FSK https://arxiv.org/abs/1203.4184



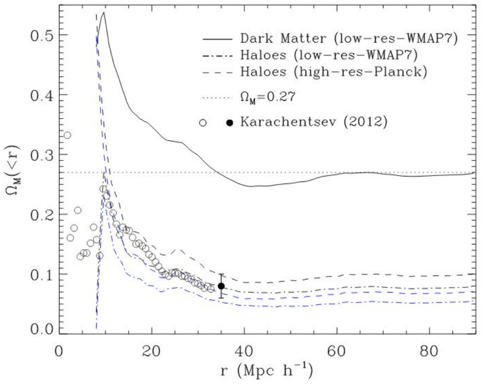




$$\delta(\{\boldsymbol{q}\}) \curvearrowleft P(\delta(\{\boldsymbol{q}\})|\{\boldsymbol{q}\})$$

$$\{\boldsymbol{q}\} \curvearrowleft P(\{\boldsymbol{q}\}|\{\boldsymbol{x}_\mathrm{G}\}, \delta(\{\boldsymbol{q}\}, m_\mathrm{SF}))$$

Application to the Local Universe 2MRS



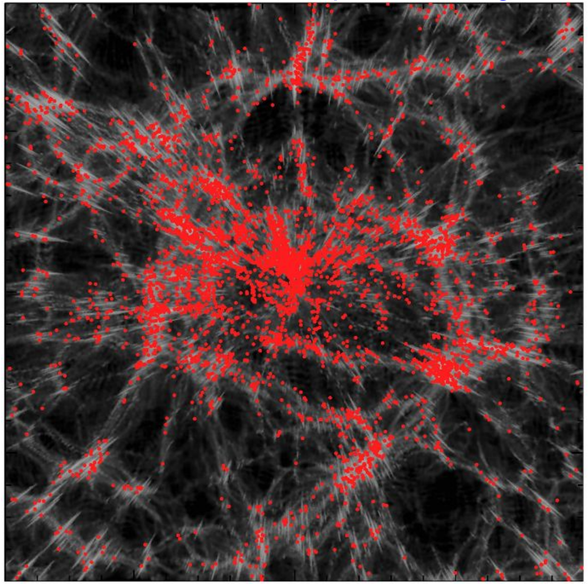
1st applications of forward model reconstructions to observations:

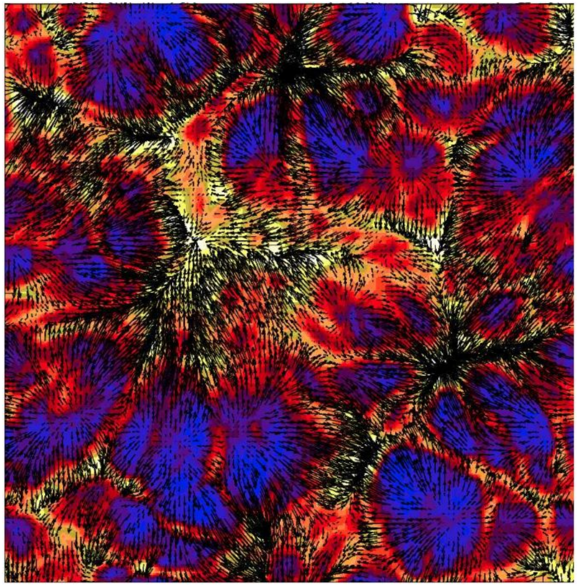
Including redshift space distortions correction also fogs!

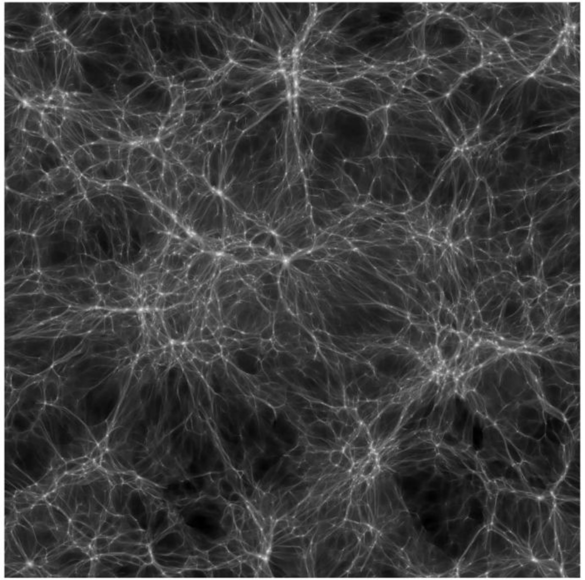
FSK et al 12 https://arxiv.org/abs/1205.5560

Hess, FSK et al 13 https://arxiv.org/abs/1304.6565

Nuza, FSK et al 14 https://arxiv.org/abs/1406.1004







See also Hess & FSK 16 https://arxiv.org/abs/1412.7310

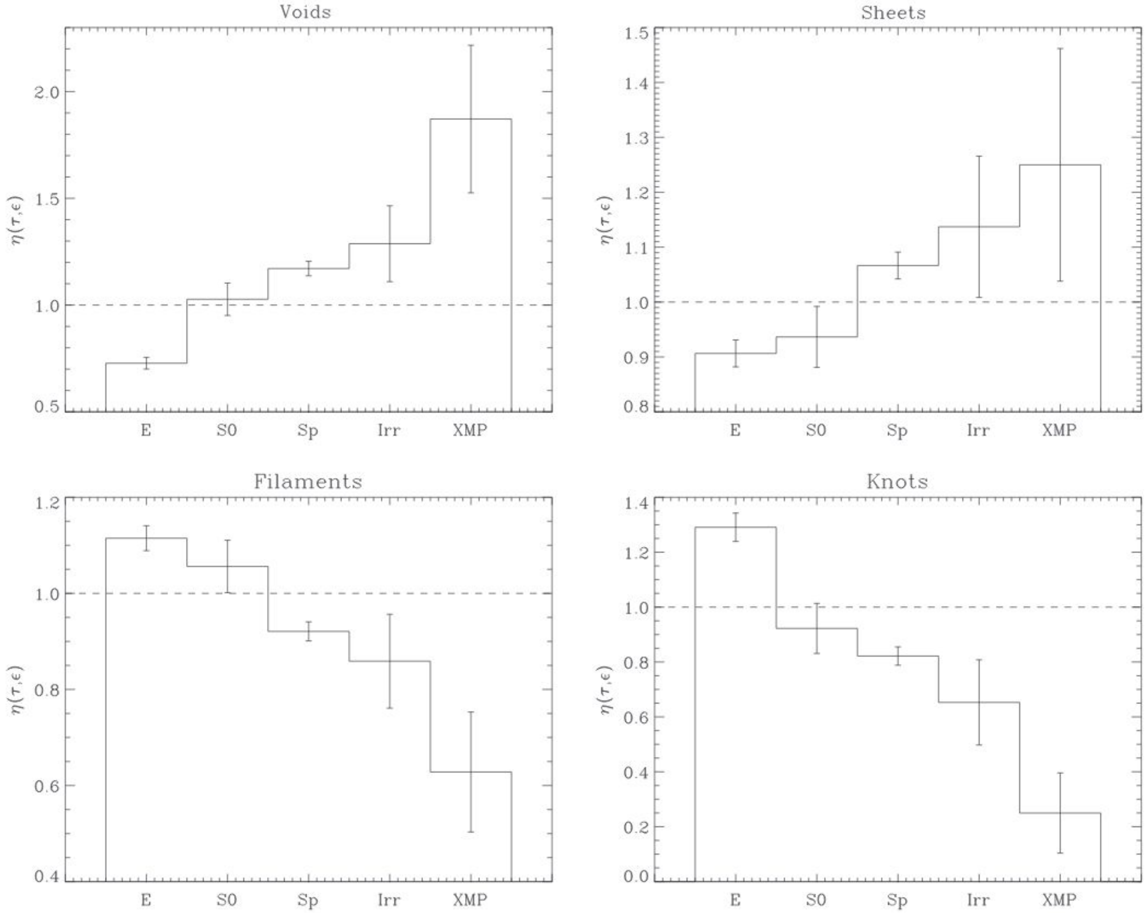
Abel, Kaehler, Hess &FSK 15 NatGeo

Environmental studies

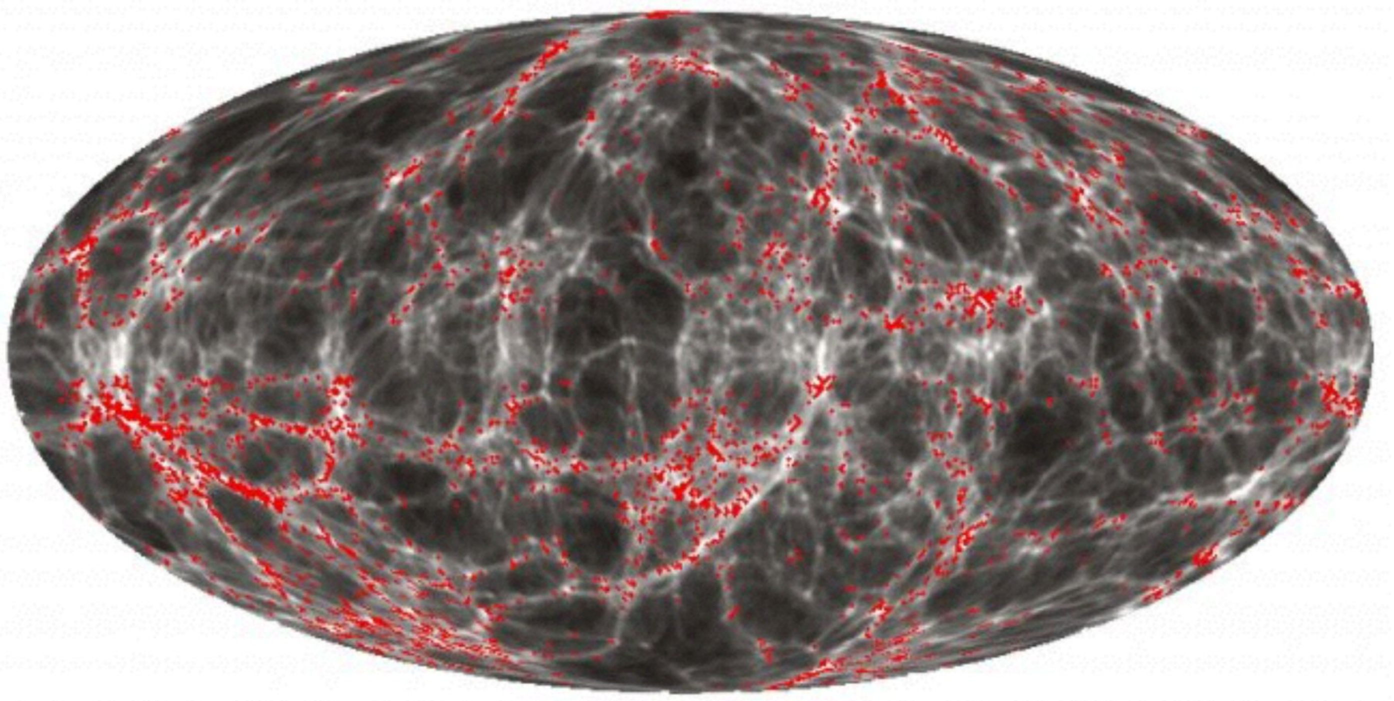
Nuza, FSK et al 14 https://arxiv.org/abs/1406.1004
Filho (FSK) et al https://arxiv.org/abs/1501.06709

See also J. Lee 08

Excess probability of finding certain galaxy types in different cosmic web environments



Application of KIGEN to the Local Universe (FSK et al 2012)



Problems:

1) bias including internal variables are chosen to fit an input power spectrum
2) survey mask is ignored gaps are filled with random mock galaxies
3) light-cone effects are ignored

Nuza, FSK et al 14 https://arxiv.org/abs/1406.1004

Getting ready for DESI, EUCLID ...

Cosmic BIRTH code

FSK et al to be submitted

Bayesian **I**nference for **R**eality vs **TH**eory

What is new?

- automatic nonlinear bias sampling
- reconstruction of the completeness at early cosmic times (survey geometry, selection function)
- light-cone reconstruction of bias, displacements, velocity fields, completeness

Towards Lagrangian bias

1) Observable clustering in redshift space and relation to large scale bias *b*

$$b^s(z) \equiv \sqrt{\frac{\xi_{\rm g}^s(z)}{\xi(z)}\Big|_{\rm LS}}$$

Kaiser 87
See also Metin Ata, FSK et al https://arxiv.org/abs/1605.09745

$$\begin{aligned}\xi_{\rm g}^s(z) &= K(z)\,\xi_{\rm G}(z) \\ &= K(z)\,b^2(z)\,\xi(z)\end{aligned}$$

$$b(z) = -\frac{1}{3}f_\Omega(z) + \sqrt{-\frac{4}{45}f_\Omega(z)^2 + (b^s(z))^2}$$

Towards Lagrangian bias

2) Time evolution of large scale bias:
Connecting Eulerian large scale bias to Lagrangian large scale bias

Fry 97

$$b(z_q) = (b(z) - 1)\,\frac{D(z)}{D(z_q)} + 1$$

A bias of 2 becomes a bias of about 60 at z=100!

At mesh resolutions of < order of 10 Mpc the over-density field becomes large enough to cause many cells with negative densities using this bias as a linear factor!

Towards Lagrangian bias

3) From large scale to non-linear Lagrangian large scale bias

$$\rho_g(\boldsymbol{q}) = \gamma(z_q)(1+\delta(\boldsymbol{q}))^{b(z_q)\, f_b(z_q)}$$

power-law bias

Simplest non-linear bias which yields positive definite density fields.
Not that simple! We need a correction factor *fb*, which depends on the resolution of the mesh: internal parameters!

Towards Lagrangian bias

4) renormalised perturbation theory

Attempt with renormalised perturbation theory
Truncated power law bias to third order:

Based on McDonald & Roy 09

$$\delta_{\rm g}(z) \equiv \frac{\rho_{\rm g}}{\bar{N}}(z) - 1 \simeq \tau(z)\Big[1 + b(z) f_b(z)\delta(z) + \frac{1}{2} b(z) f_b(z)(b(z) f_b(z) - 1)\left(\delta(z)\right)^2 + \frac{1}{3!} b(z) f_b(z)(b(z) f_b(z) - 1)(b(z) f_b(z) - 2)\left(\delta(z)\right)^3\Big] - 1\,, \quad (22)$$

$$b(z)\, f_b^3(z) + f_b(z)^2\left(\frac{5}{21} - b(z)\right) + \frac{f_b(z)}{b(z)}\left(\frac{2}{\sigma^2(z)} - \frac{26}{21} + b(z)\right) - \frac{2}{\sigma^2(z)\, b(z)} = 0$$

Beautiful cubic equation...

...but yields inaccurate power spectra due to truncation to third order! Due to high large scale bias higher orders become important too!

Towards Lagrangian bias

5) numerical computation to arbitrary accuracy

Can we compute the large scale bias from our model?

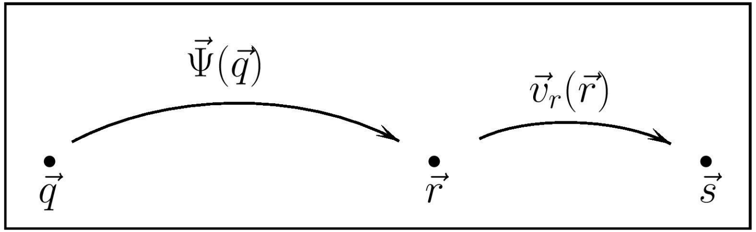
$$b(z) \equiv \sqrt{\frac{\sigma^2_{K\,g}(z)}{\sigma^2_K(z)}}$$

$$\sigma^2_K(z) = \langle (K \circ \delta(\boldsymbol{q}, z))^2 \rangle$$

$$\sigma^2_{K\,g}(z) = \langle (K \circ \delta_g(\boldsymbol{q}, z)[b_{\rm eff}])^2 \rangle$$

$$b_{\rm eff}(z) = b(z)\, f_b(z)$$

Lagrangian tracers

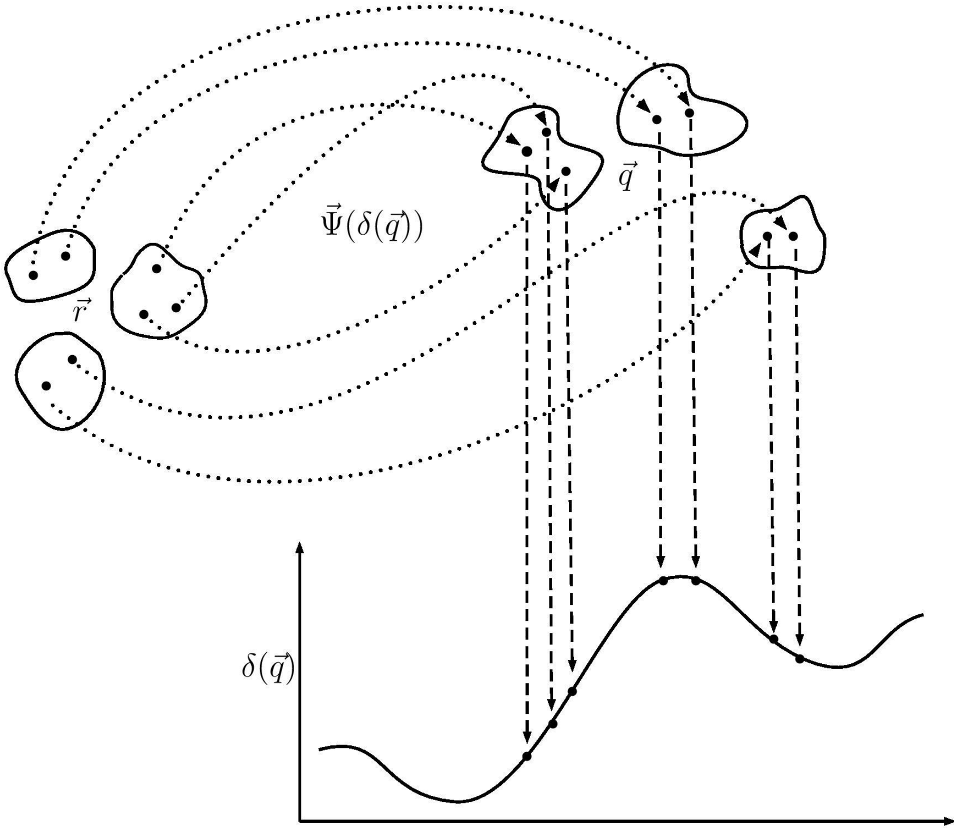


Galaxies are tracers of halos.

Multiple tracers can be used as in the KIGEN approach.

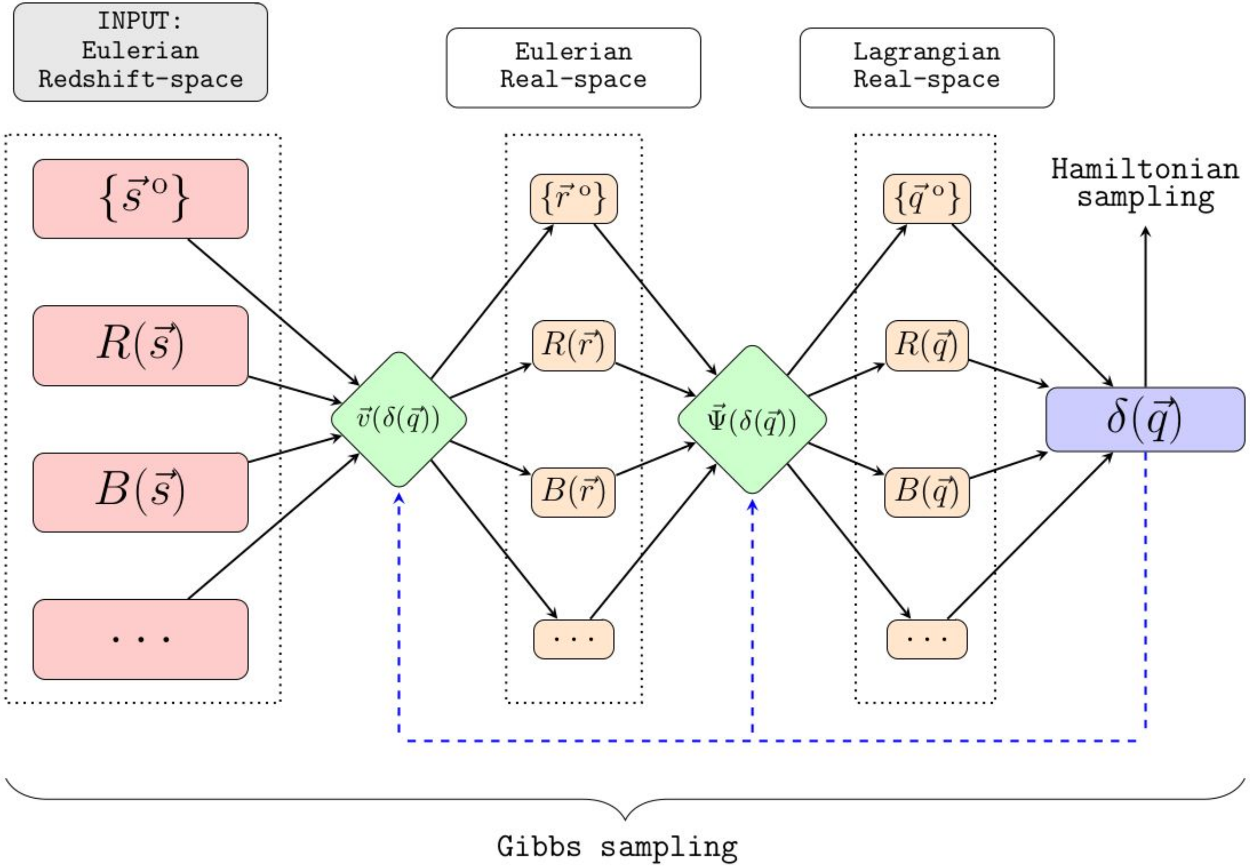
The tracers at Lagrangian positions are not defining spherical symmetric proto-halo regions.

Not including peak split-background: threshold bias, permits us to trace the whole regime of the density field!

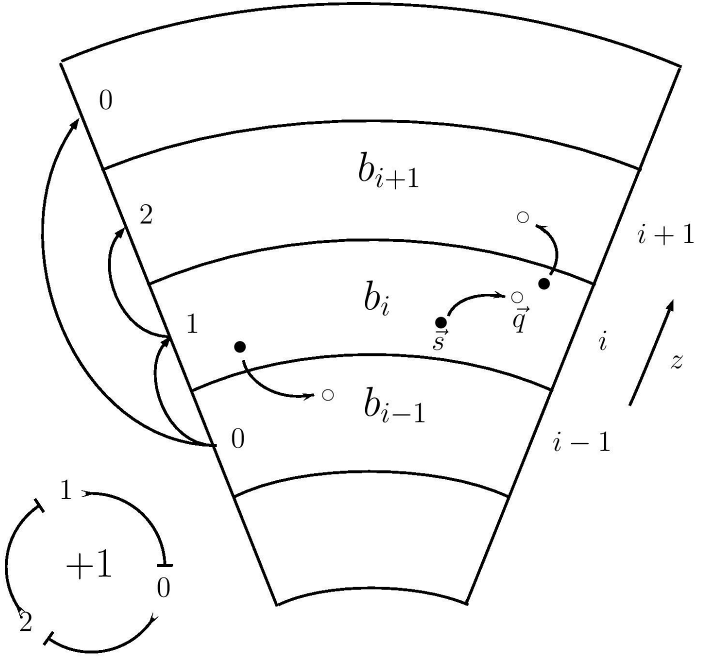


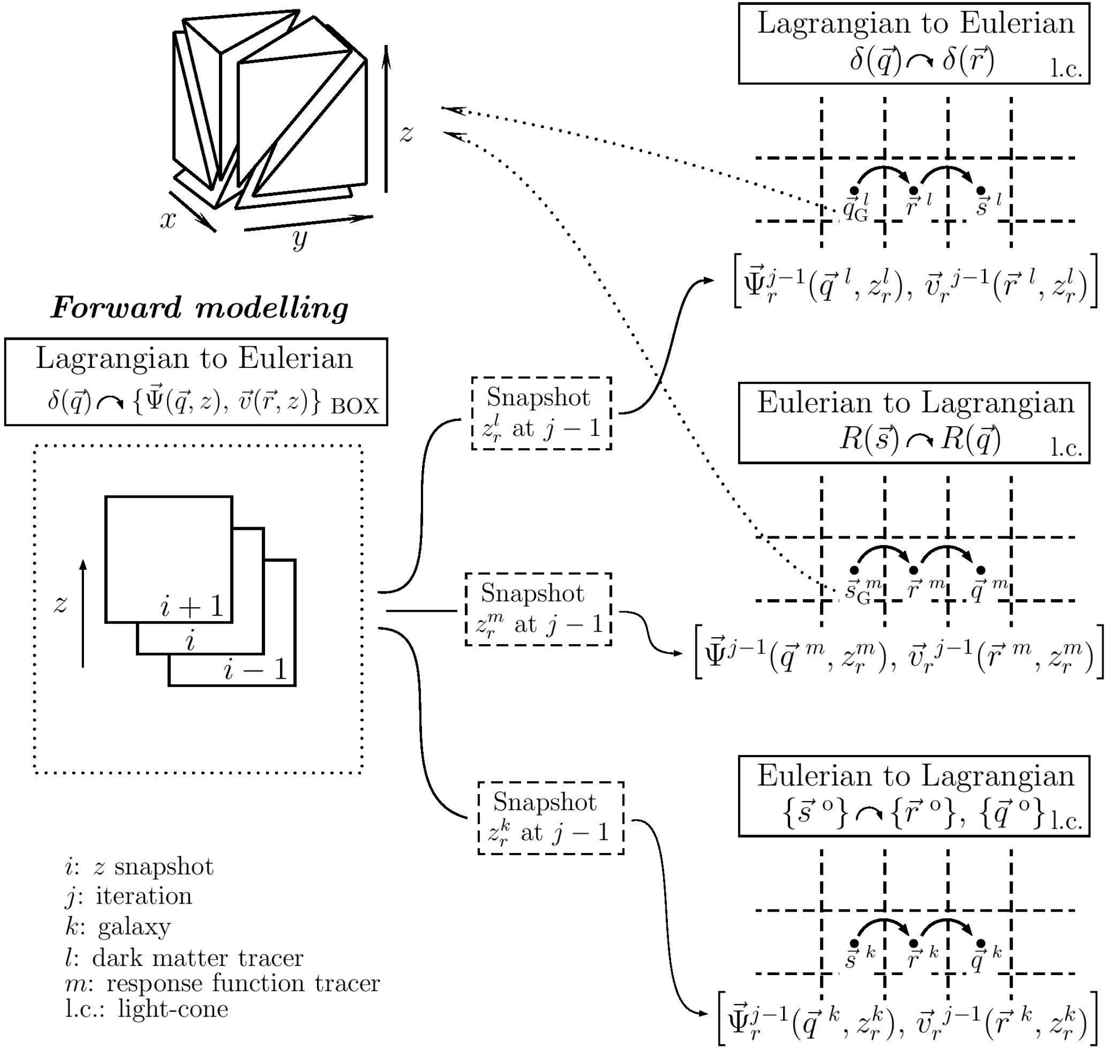
See Porciani et al 02 Sheth, Chan & Scoccimarro 13

Structure of the BIRTH code



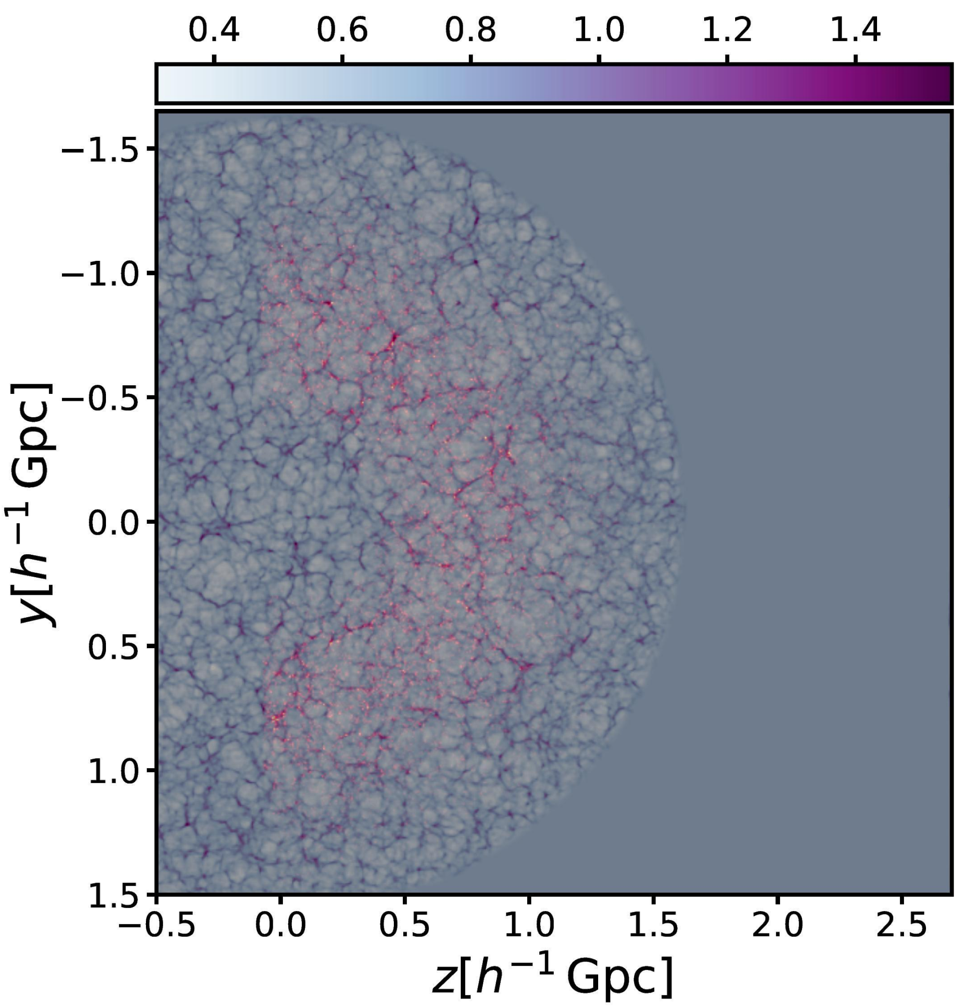
Going through z-snapshots



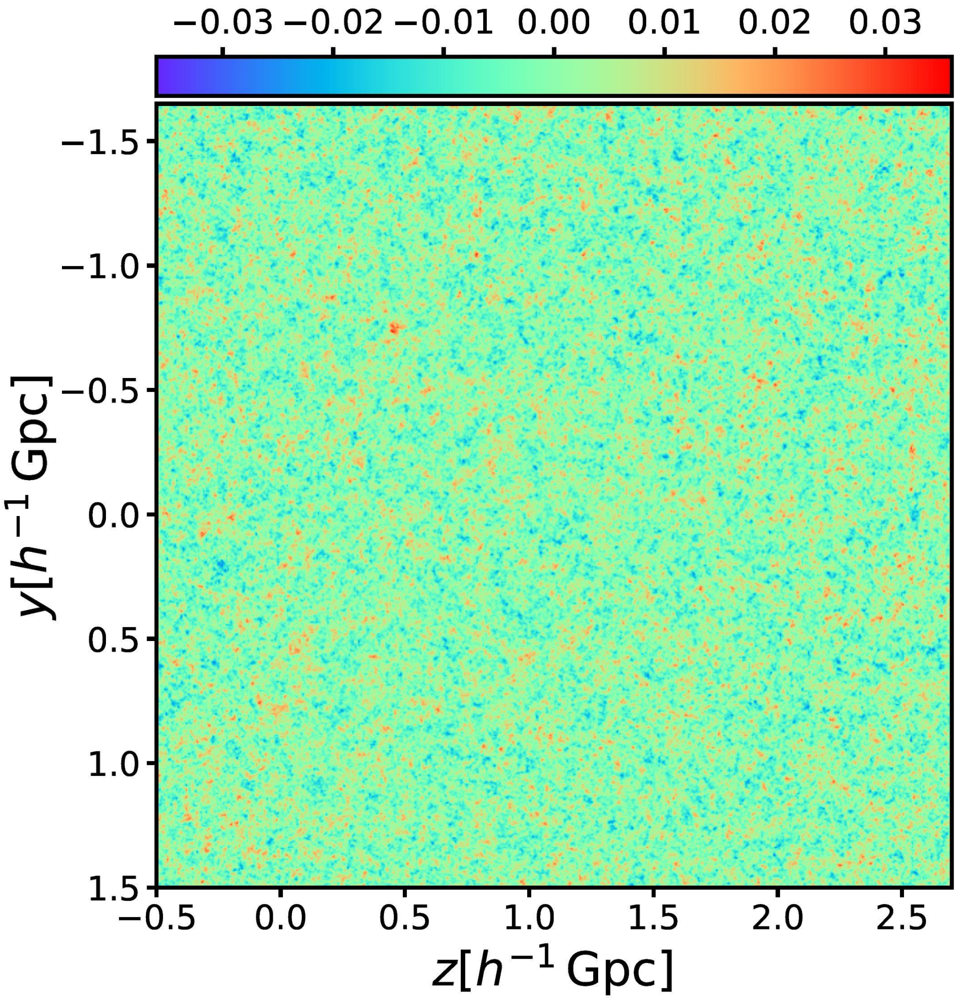


Study case: simulated CMASS galaxies

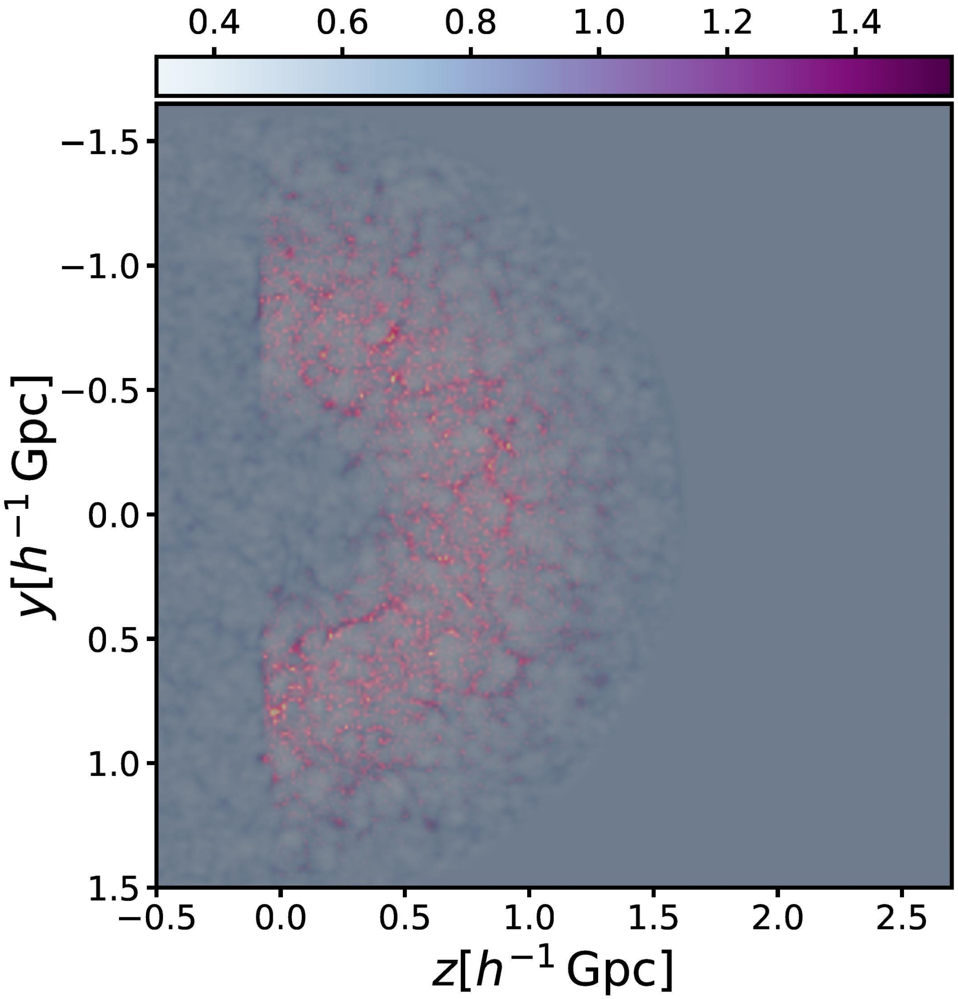
BigMD simulation Klypin et al 15
CMASS mock including evolution and SHAM Rodríguez-Torres et al 16

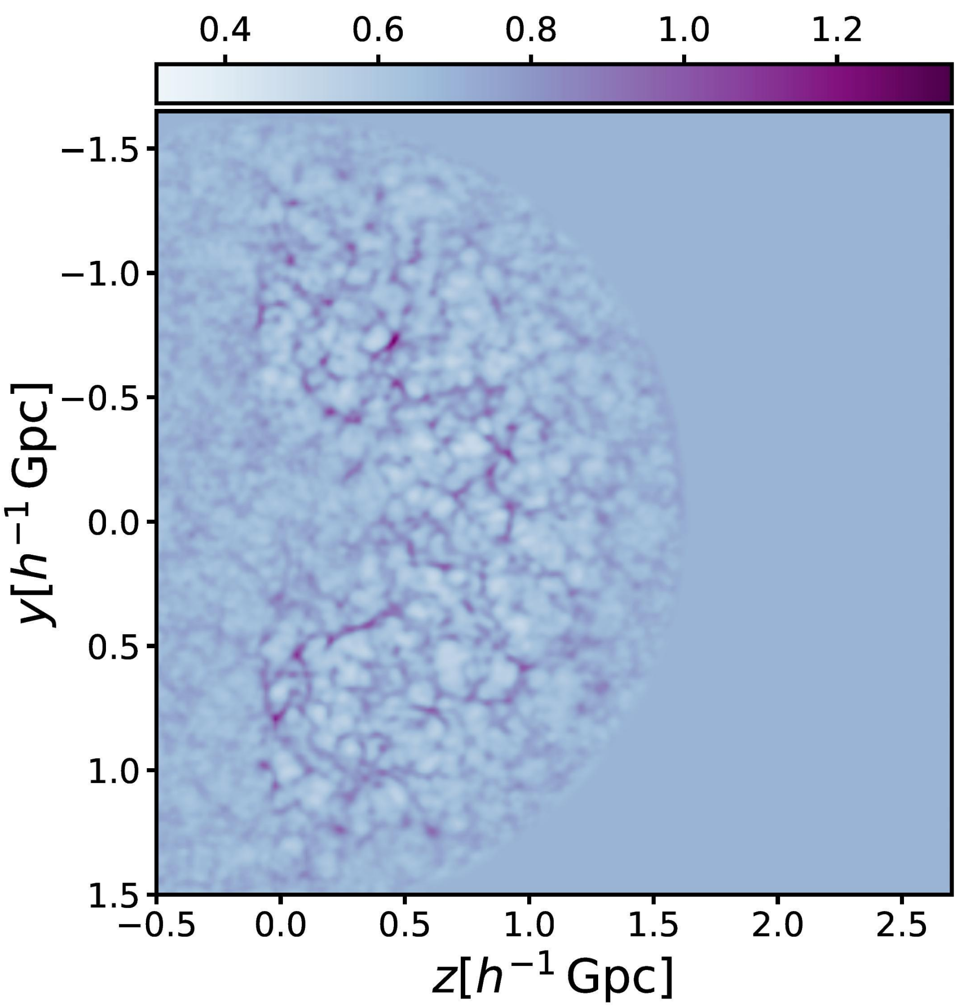


0.4
0.6
0.8
1.0
1.2
1.4
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
$y[h^{-1}\,\mathrm{Gpc}]$
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
$z[h^{-1}\,\mathrm{Gpc}]$

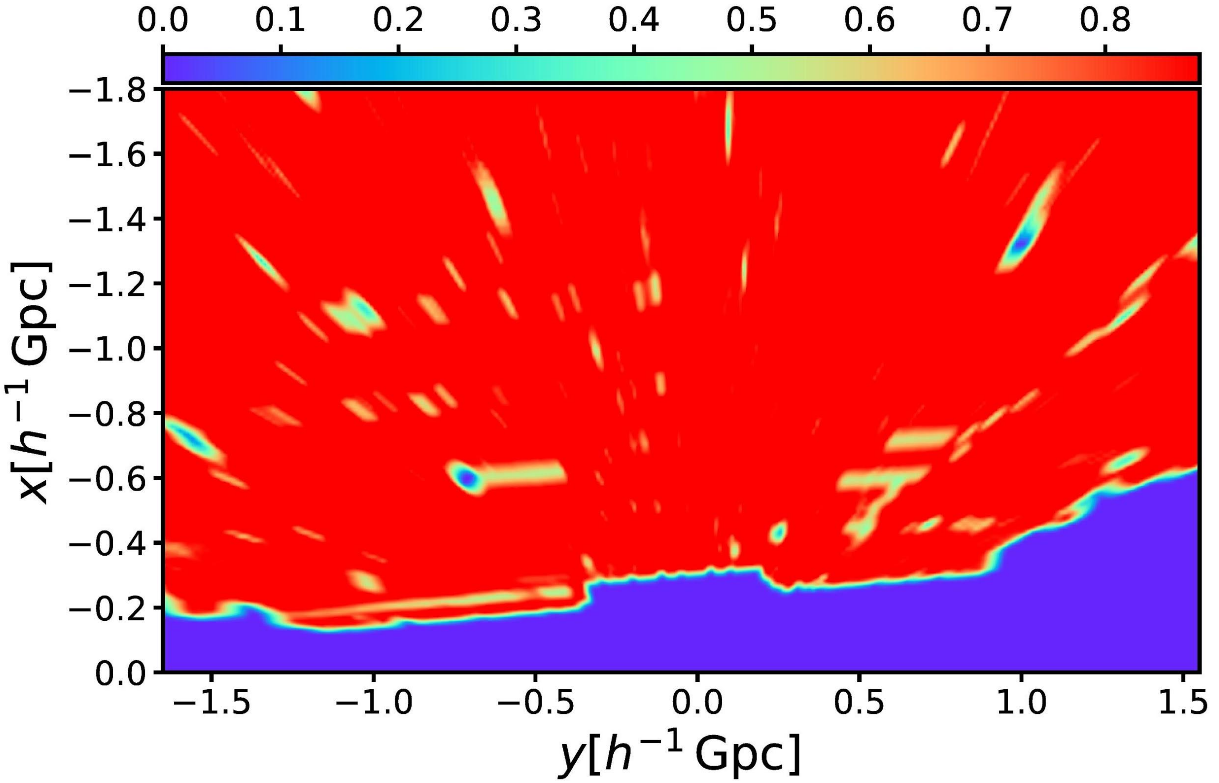


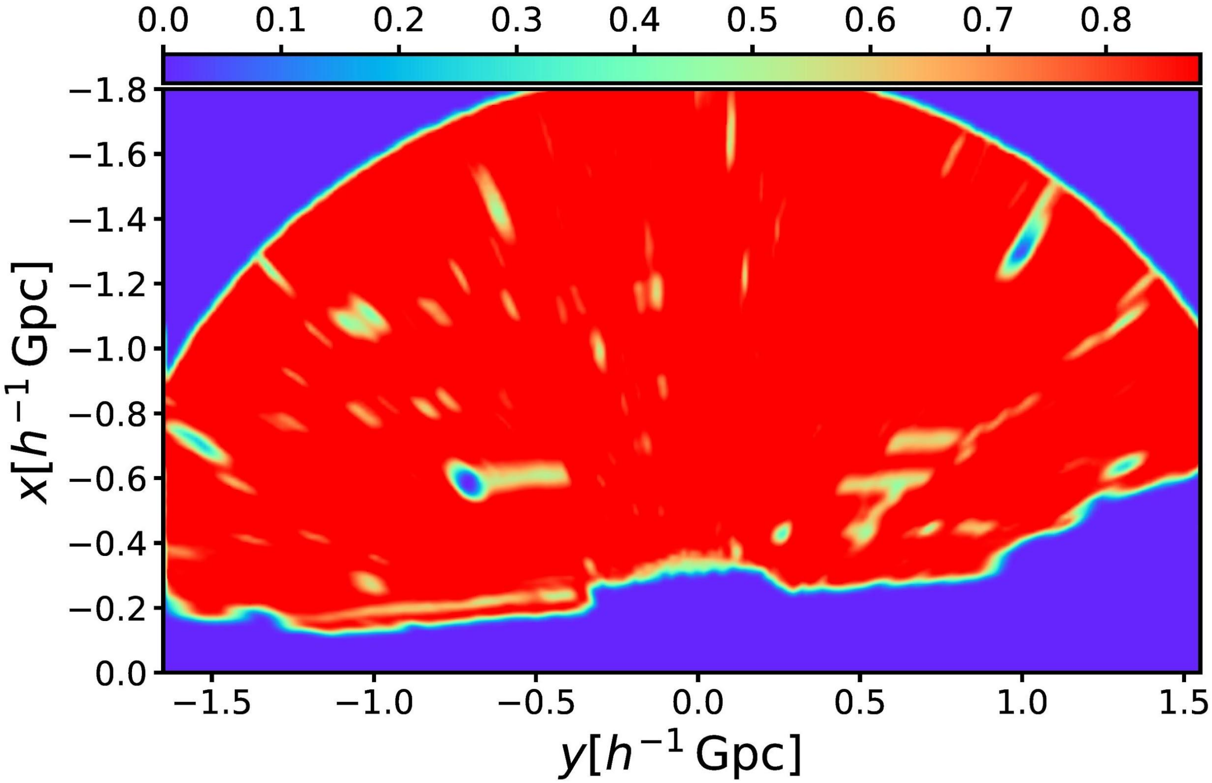
−0.03 −0.02 −0.01 0.00 0.01 0.02 0.03
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
$y[h^{-1}\,\mathrm{Gpc}]$
−0.5 0.0 0.5 1.0 1.5 2.0 2.5
$z[h^{-1}\,\mathrm{Gpc}]$


0.4
0.6
0.8
1.0
1.2
1.4
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
$y[h^{-1}\,\mathrm{Gpc}]$
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
$z[h^{-1}\,\mathrm{Gpc}]$



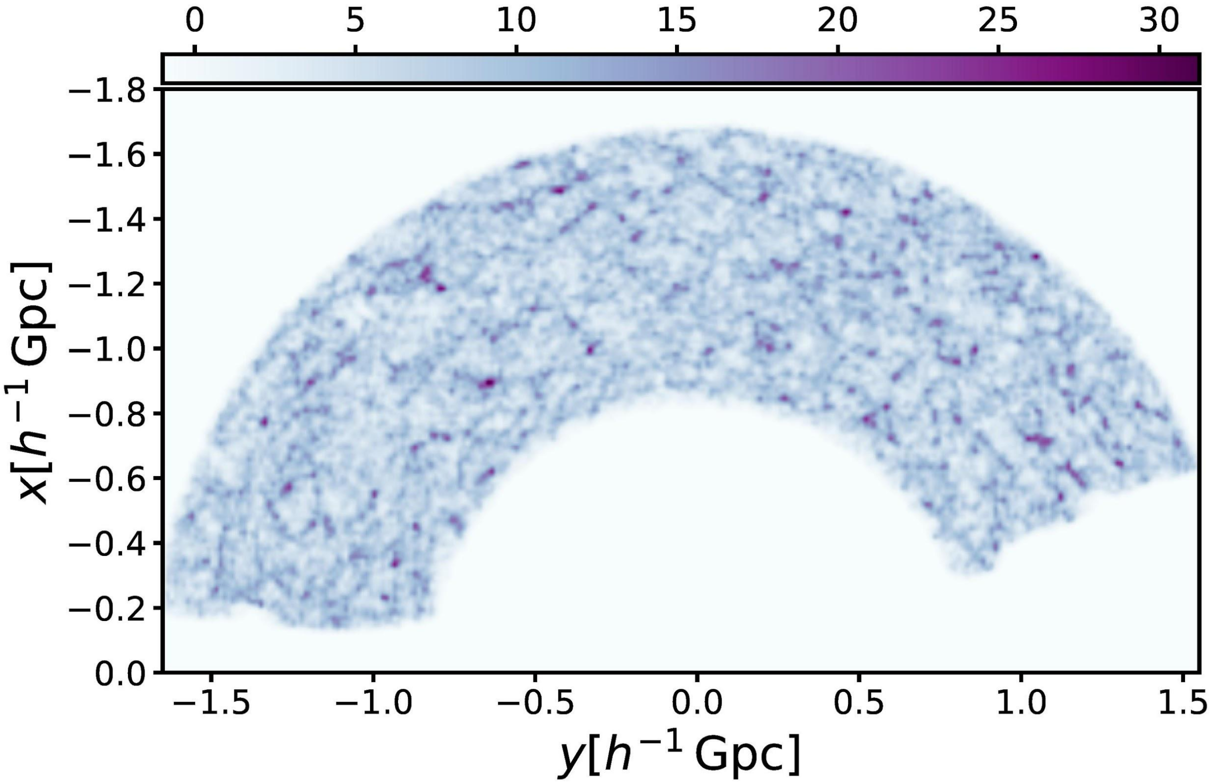
0.4
0.6
0.8
1.0
1.2
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
$y[h^{-1}\,\mathrm{Gpc}]$
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
$z[h^{-1}\,\mathrm{Gpc}]$


0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
−1.8
−1.6
−1.4
−1.2
−1.0
−0.8
−0.6
−0.4
−0.2
0.0
$x[h^{-1}\,\mathrm{Gpc}]$
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
$y[h^{-1}\,\mathrm{Gpc}]$

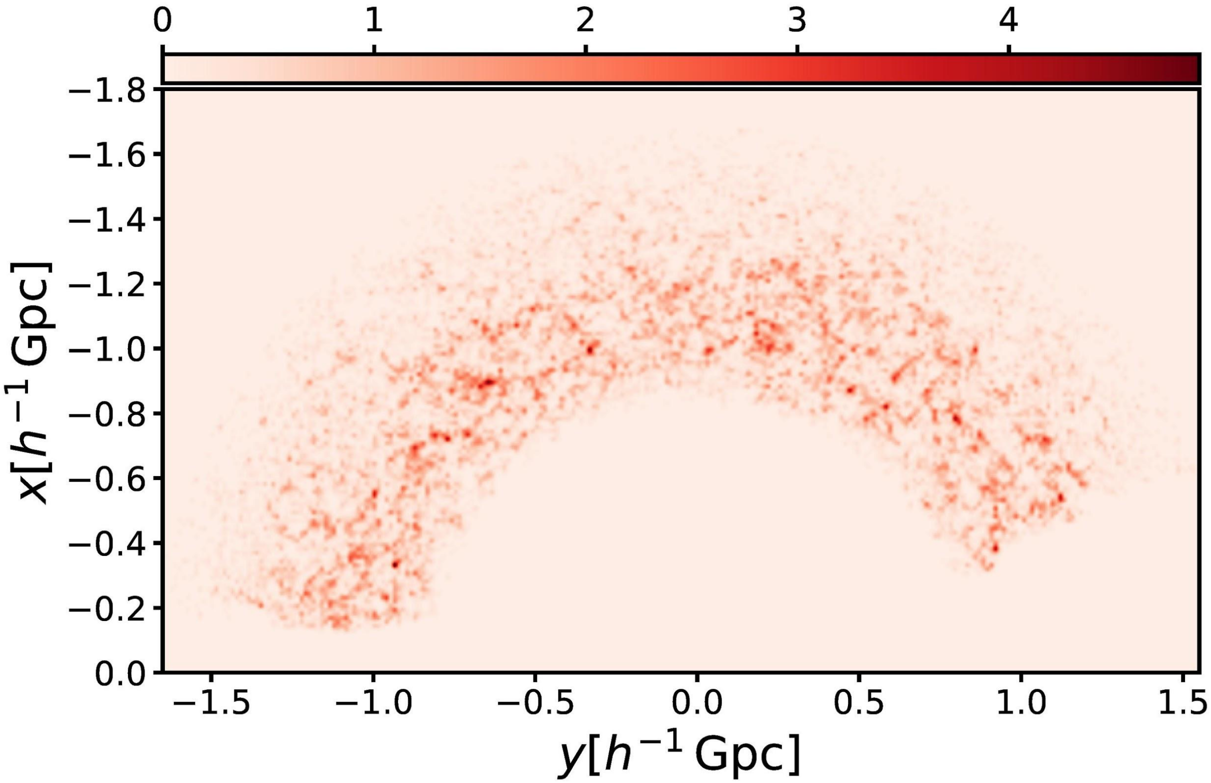


0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
−1.8
−1.6
−1.4
−1.2
−1.0
−0.8
−0.6
−0.4
−0.2
0.0
$x[h^{-1}\,\mathrm{Gpc}]$
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
$y[h^{-1}\,\mathrm{Gpc}]$

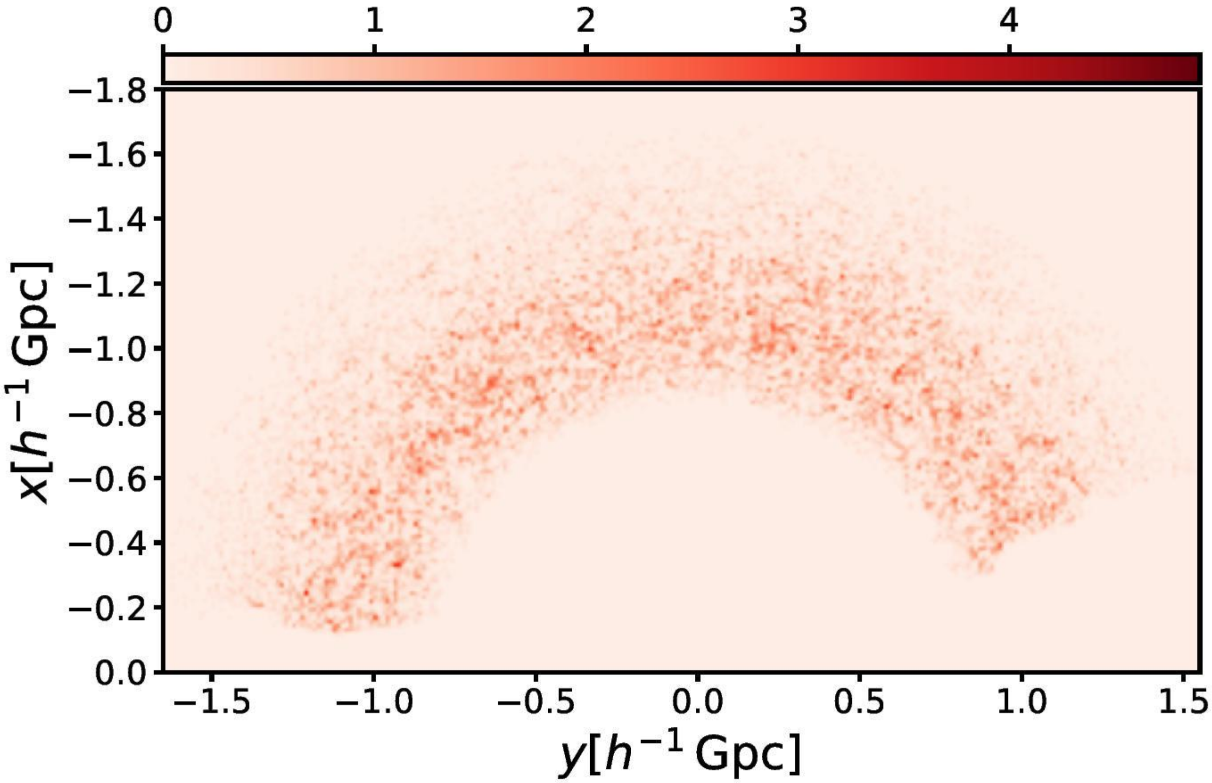
DM from the BigMD sim with light-cone evolution



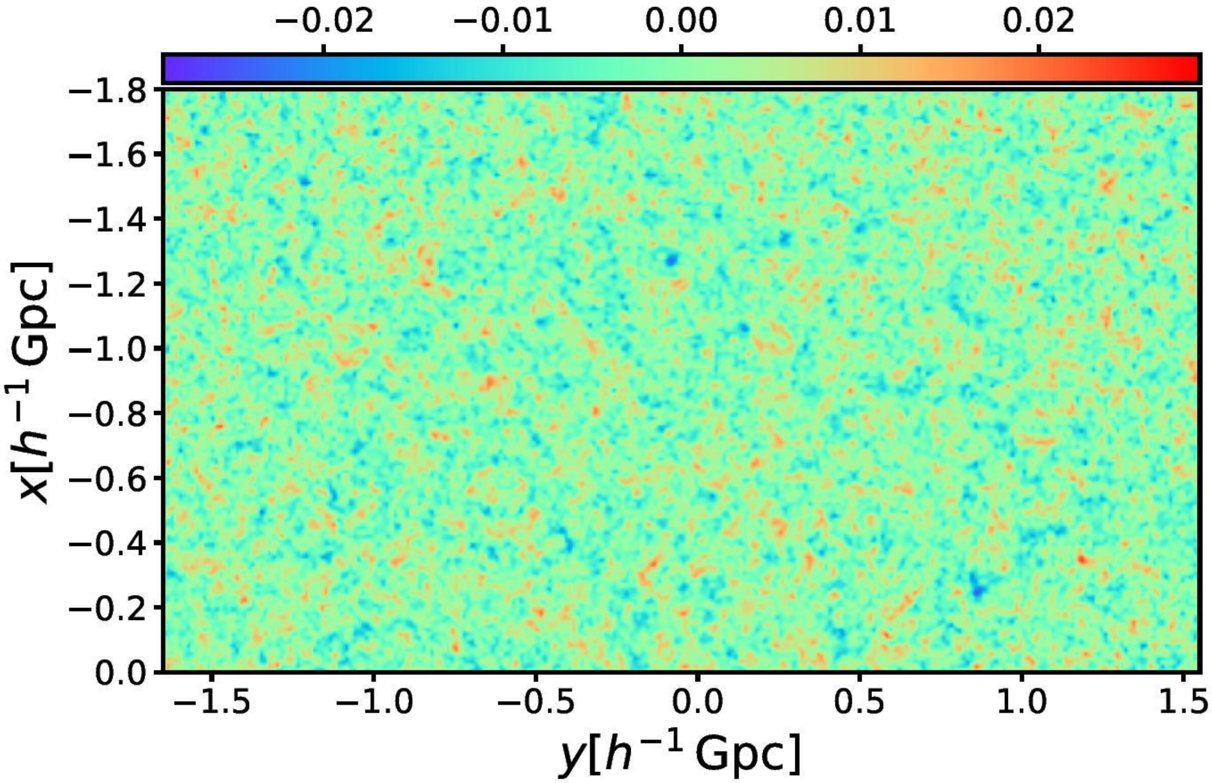
Mock galaxies using SHAM based on the BigMD sim



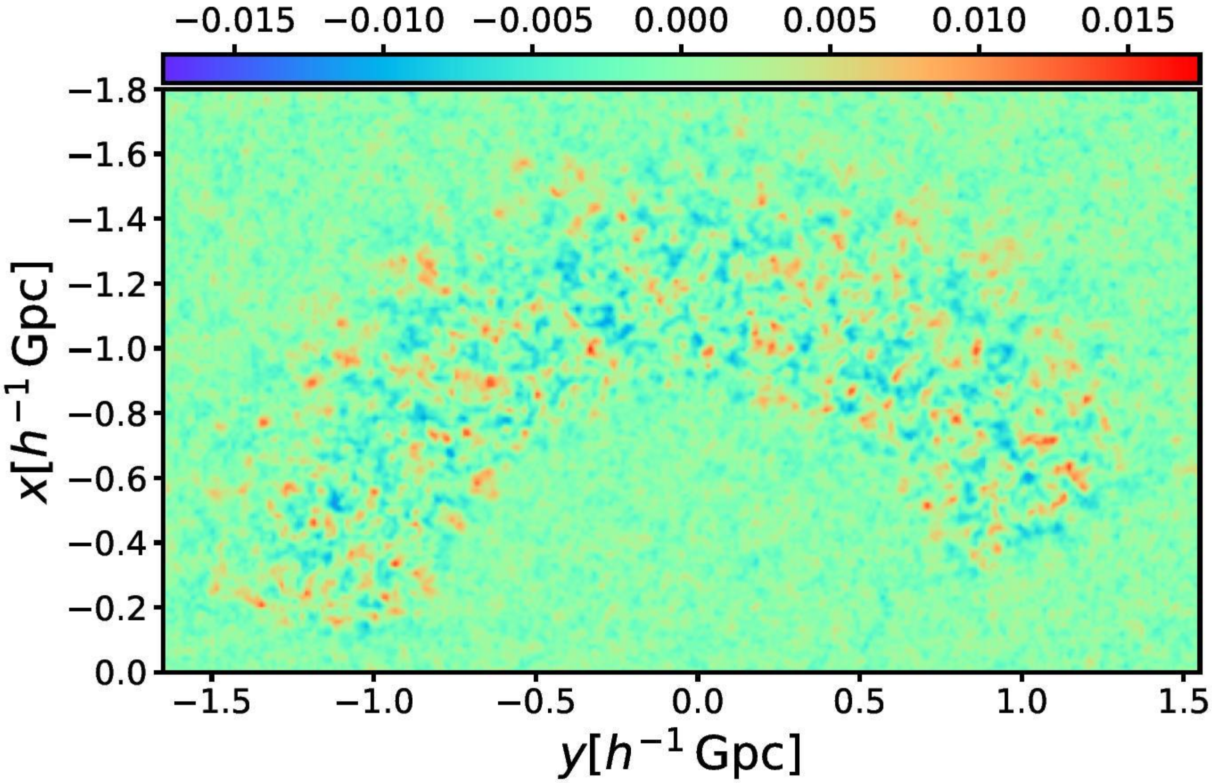
Mock galaxies at Lagrangian coordinates



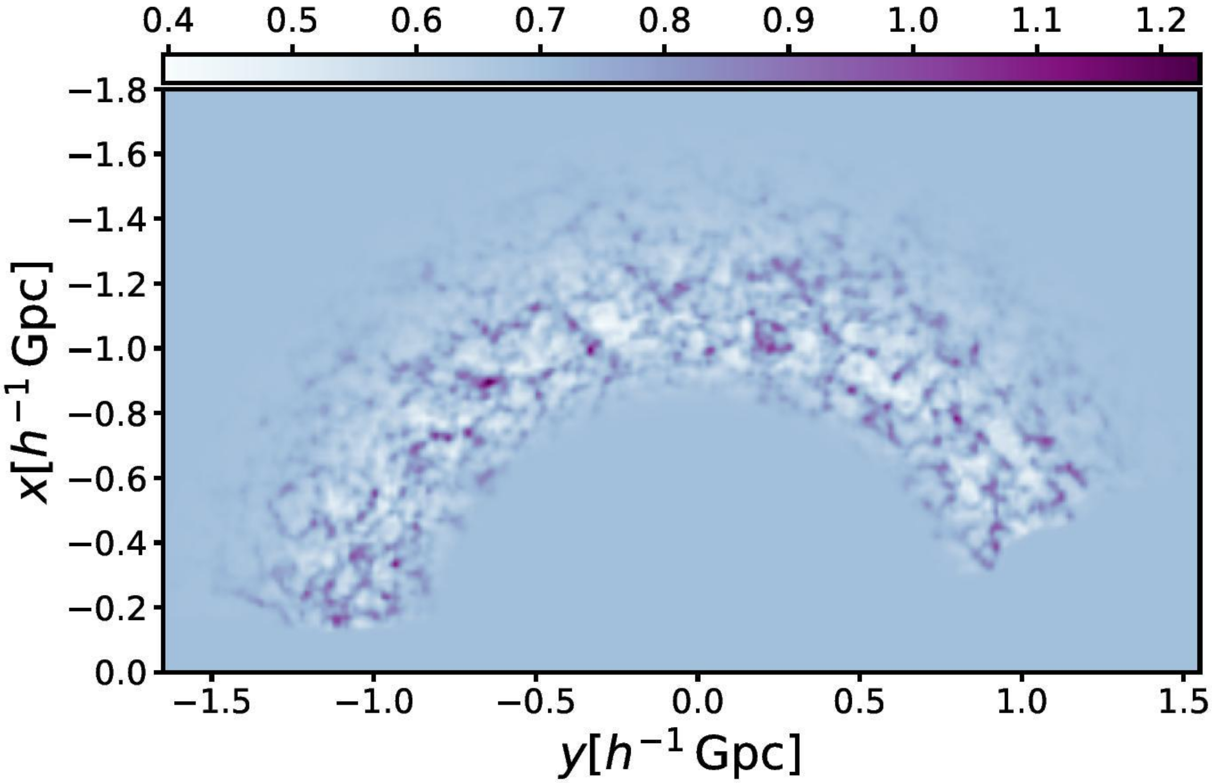
Sampled Gaussian field



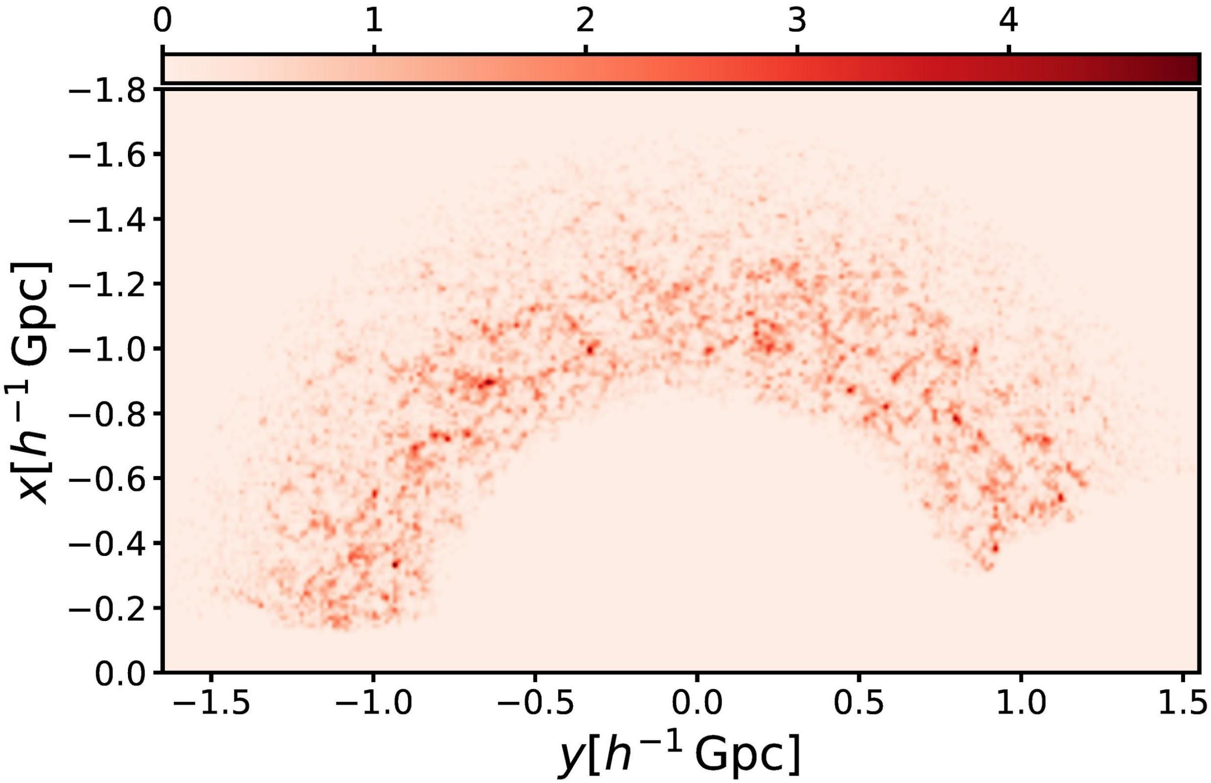
Ensemble average



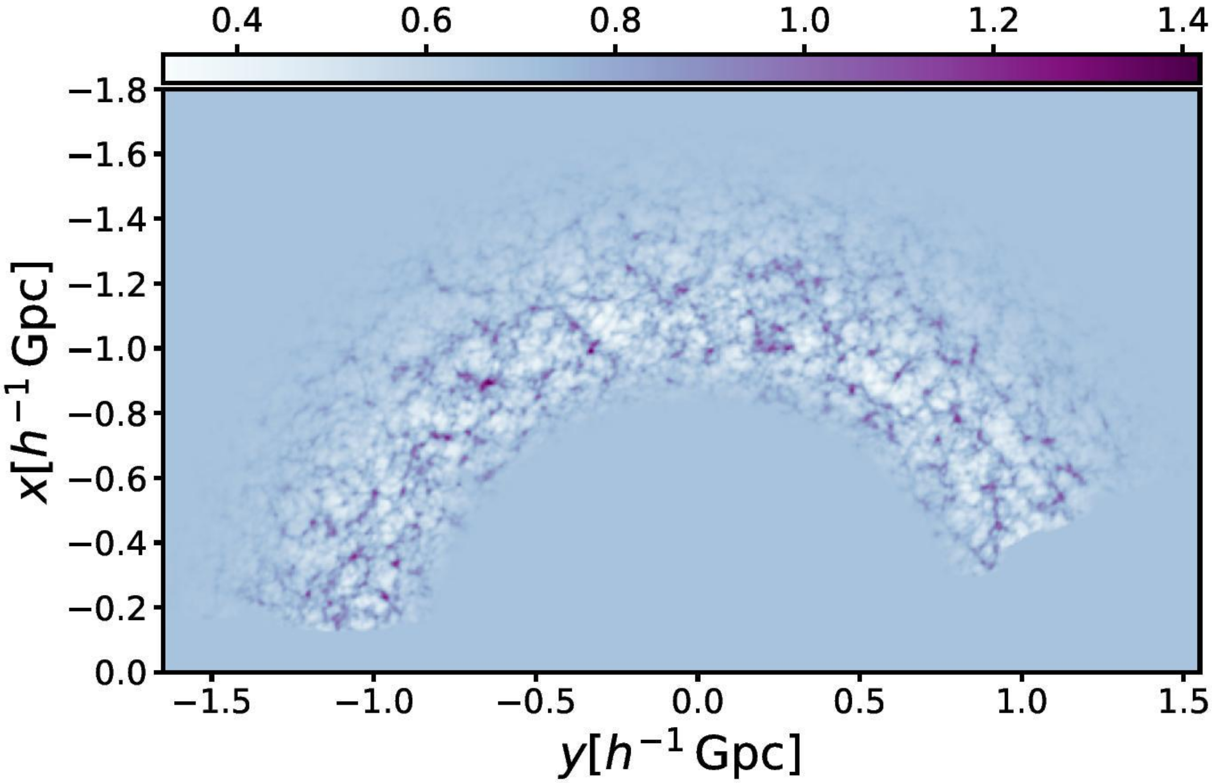
Ensemble average of density fields on the light-cone

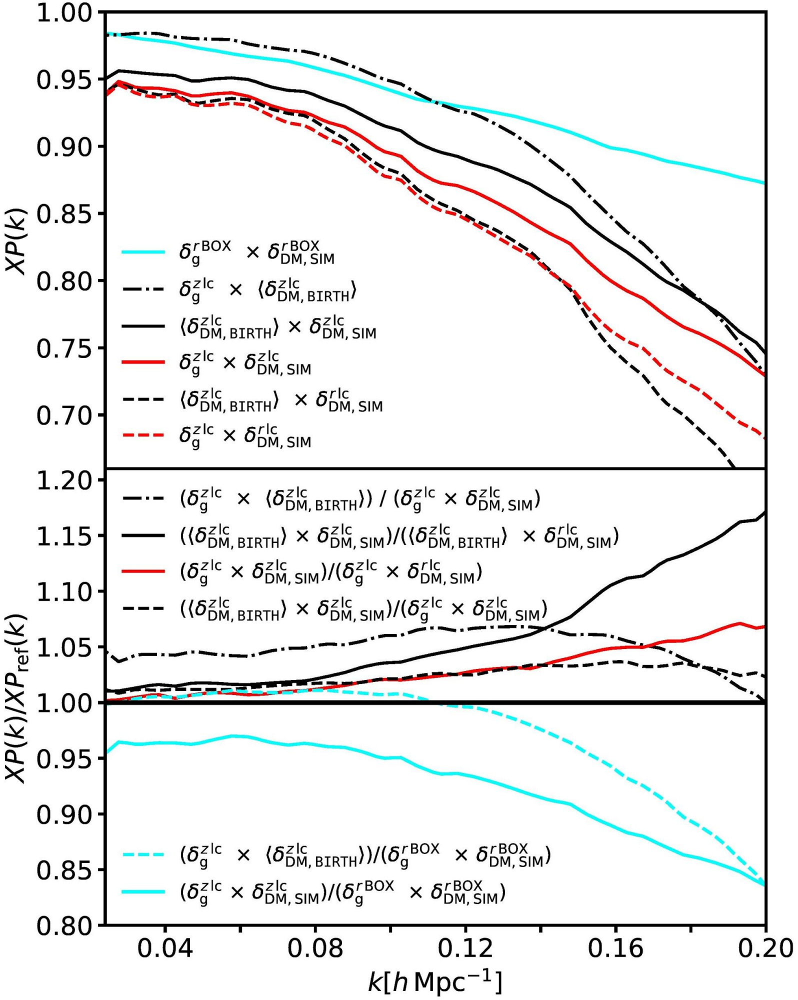


Mock galaxies using SHAM based on the BigMD sim



Ensemble average of density fields on the light-cone





$XP(k)$
1.00
0.95
0.90
0.85
0.80
0.75
0.70
$\delta_g^{r\mathrm{BOX}} \times \delta_{\mathrm{DM,SIM}}^{r\mathrm{BOX}}$
$\delta_g^{z\mathrm{lc}} \times \langle\delta_{\mathrm{DM,BIRTH}}^{z\mathrm{lc}}\rangle$
$\langle\delta_{\mathrm{DM,BIRTH}}^{z\mathrm{lc}}\rangle \times \delta_{\mathrm{DM,SIM}}^{z\mathrm{lc}}$
$\delta_g^{z\mathrm{lc}} \times \delta_{\mathrm{DM,SIM}}^{z\mathrm{lc}}$
$\langle\delta_{\mathrm{DM,BIRTH}}^{z\mathrm{lc}}\rangle \times \delta_{\mathrm{DM,SIM}}^{r\mathrm{lc}}$
$\delta_g^{z\mathrm{lc}} \times \delta_{\mathrm{DM,SIM}}^{r\mathrm{lc}}$
$XP(k)/XP_{\mathrm{ref}}(k)$
1.20
1.15
1.10
1.05
1.00
0.95
0.90
0.85
0.80
$(\delta_g^{z\mathrm{lc}} \times \langle\delta_{\mathrm{DM,BIRTH}}^{z\mathrm{lc}}\rangle) / (\delta_g^{z\mathrm{lc}} \times \delta_{\mathrm{DM,SIM}}^{z\mathrm{lc}})$
$(\langle\delta_{\mathrm{DM,BIRTH}}^{z\mathrm{lc}}\rangle \times \delta_{\mathrm{DM,SIM}}^{z\mathrm{lc}})/(\langle\delta_{\mathrm{DM,BIRTH}}^{z\mathrm{lc}}\rangle \times \delta_{\mathrm{DM,SIM}}^{r\mathrm{lc}})$
$(\delta_g^{z\mathrm{lc}} \times \delta_{\mathrm{DM,SIM}}^{z\mathrm{lc}})/(\delta_g^{z\mathrm{lc}} \times \delta_{\mathrm{DM,SIM}}^{r\mathrm{lc}})$
$(\langle\delta_{\mathrm{DM,BIRTH}}^{z\mathrm{lc}}\rangle \times \delta_{\mathrm{DM,SIM}}^{z\mathrm{lc}})/(\delta_g^{z\mathrm{lc}} \times \delta_{\mathrm{DM,SIM}}^{z\mathrm{lc}})$
$(\delta_g^{z\mathrm{lc}} \times \langle\delta_{\mathrm{DM,BIRTH}}^{z\mathrm{lc}}\rangle)/(\delta_g^{r\mathrm{BOX}} \times \delta_{\mathrm{DM,SIM}}^{r\mathrm{BOX}})$
$(\delta_g^{z\mathrm{lc}} \times \delta_{\mathrm{DM,SIM}}^{z\mathrm{lc}})/(\delta_g^{r\mathrm{BOX}} \times \delta_{\mathrm{DM,SIM}}^{r\mathrm{BOX}})$
0.04
0.08
0.12
0.16
0.20
$k[h\,\mathrm{Mpc}^{-1}]$

UNIT project: Universe N-body simulations for the Investigation of Theoretical models from galaxy surveys Chia-Hsun Chuang et al. arXiv:1811.02111

www.unitsims.org

Chia-Hsun Chuang (KIPAC, Stanford/SLAC, USA; Homepage)

Gustavo Yepes (UAM, Spain; Homepage)

Francisco-Shu Kitaura (IAC/ULL, Spain; Homepage)

Risa Wechsler (KIPAC, Stanford/SLAC, USA; Homepage)

Yu Feng (BCCP, USA; Homepage)

Shadab Alam (IfA, UK; Homepage)

Arka Banerjee (KIPAC, Stanford/SLAC, USA; Homepage)

Robert Benton Metcalf (DIFA/INAF, Italy)

Alexander Knebe (UAM, Spain; Homepage)

Carlo Giocoli (DIFA/FST/INAF/INFN, Italy; Homepage)

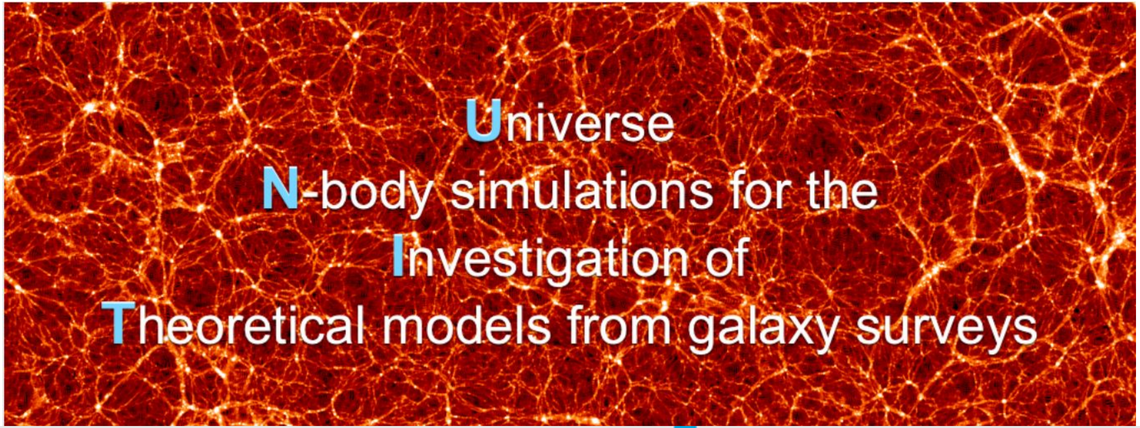
Marcos Pellejero-Ibanez (IAC/ULL, Spain)

Sergio Rodriguez-Torres (UAM, Spain)

Joe DeRose (KIPAC, Stanford/SLAC, USA)

Chun-Hao To (KIPAC, Stanford/SLAC, USA)

Cheng Zhao (EPFL, Switzerland; Homepage)



KIPAC KAVLI INSTITUTE FOR PARTICLE ASTROPHYSICS AND COSMOLOGY

UAM Universidad Autónoma de Madrid

STANFORD

SLAC

ULL Universidad de La Laguna

BERKELEY CENTER for COSMOLOGICAL PHYSICS

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

INAF ISTITUTO NAZIONALE DI ASTROFISICA

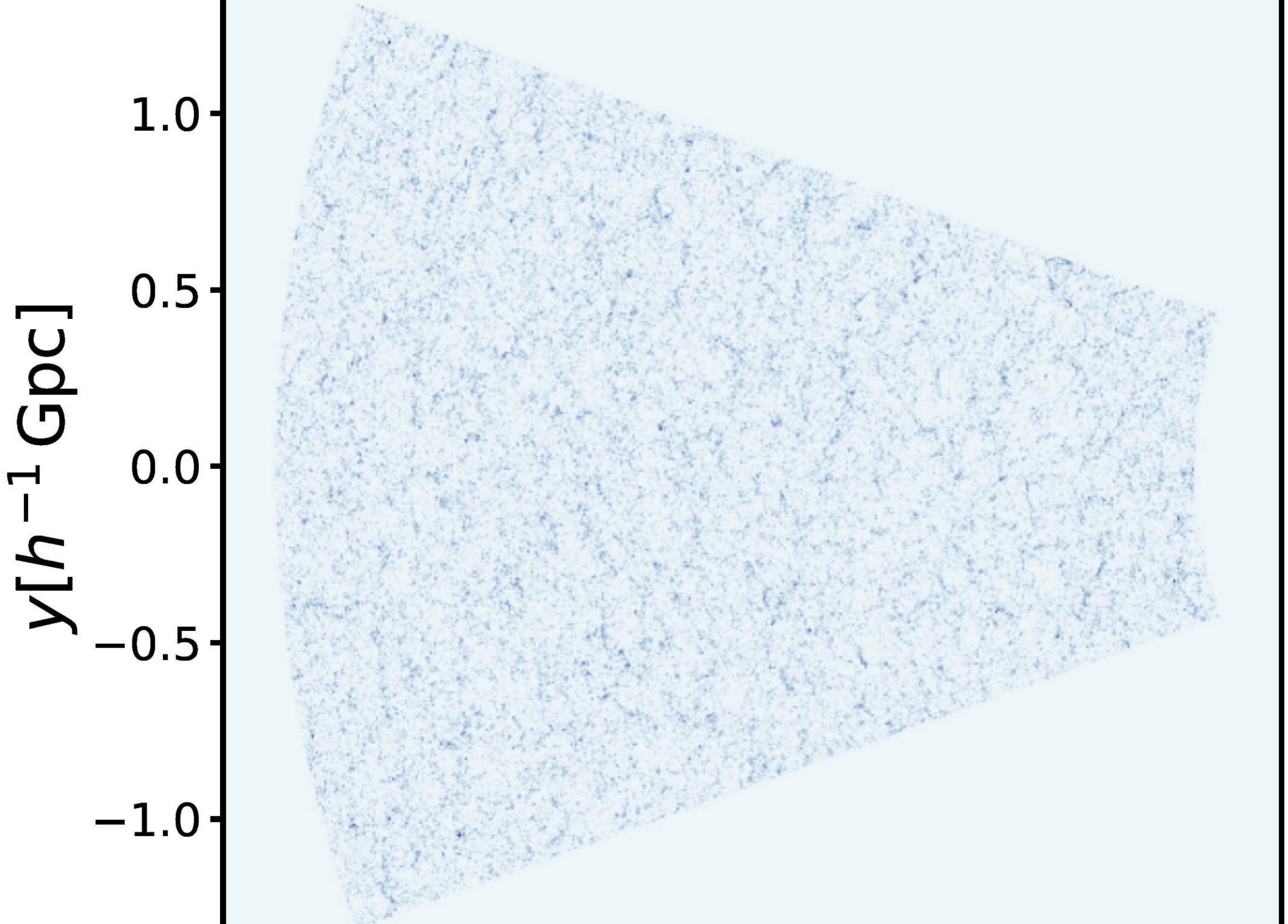
INFN Istituto Nazionale di Fisica Nucleare

Università degli Studi di Ferrara

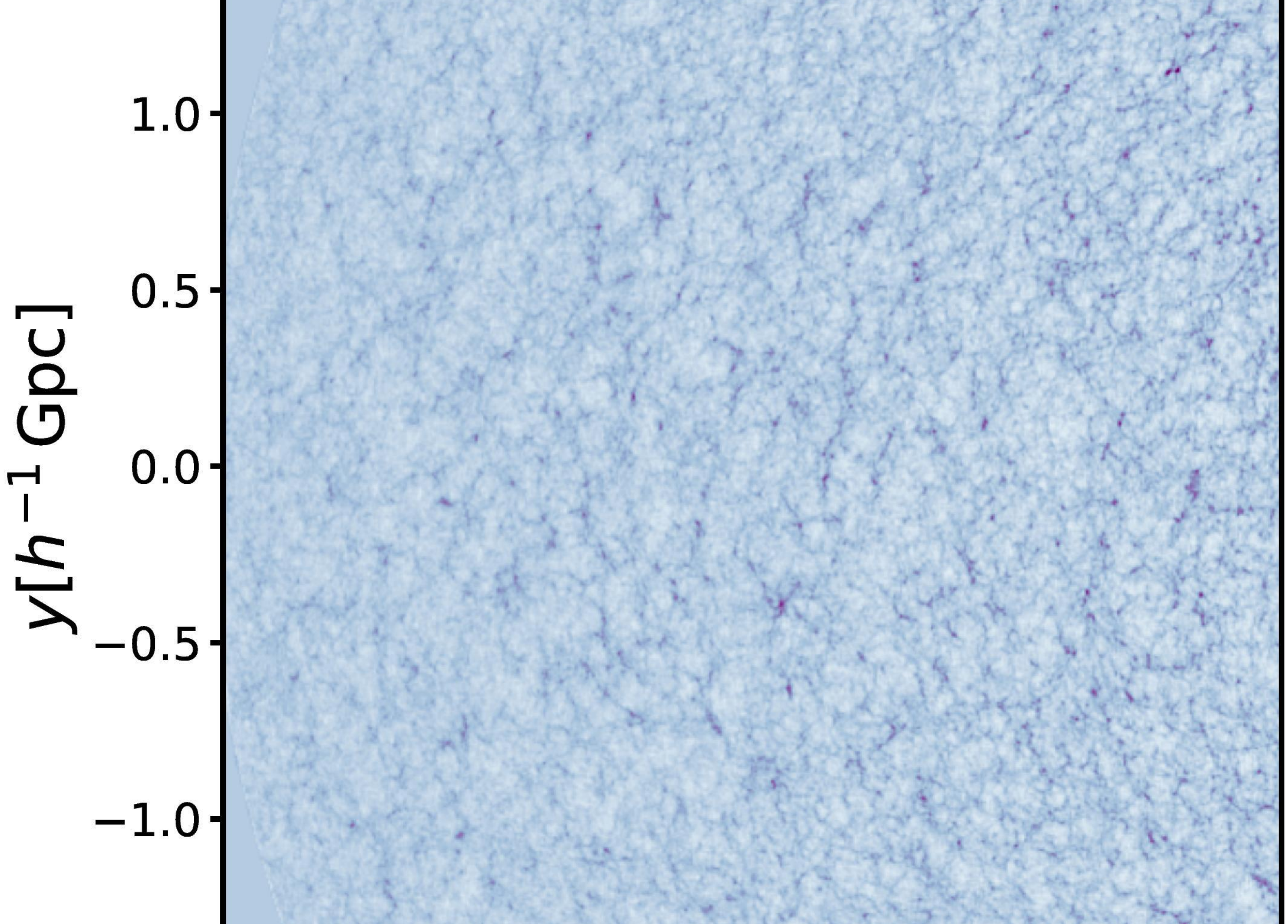
Institute for Astronomy

EPFL

PRACE



1.0
0.5
0.0
−0.5
−1.0
$y[h^{-1}\,\mathrm{Gpc}]$



$y[h^{-1}\,\mathrm{Gpc}]$
1.0
0.5
0.0
−0.5
−1.0

CONCLUSIONS

- Eulerian bias is complex especially going to eLG like galaxies (simple for LRGs)
- With a Bias Assignment Method we can learn the Eulerian bias in a non-parametric form from full calculation simulations and train algorithms to reproduce those results on coarse grids to high accuracy in the 2-, 3-, and 4- point statistics
- Bayesian forward modelling approaches help us to reconstruct the initial dark matter field from the galaxy distribution
- We have presented a framework in Lagrangian space able to deal with survey geometry radial selection functions, arbitrary structure formation models, multi-tracer analysis, and light-cone evolution