Modeling Biased Tracers at the Field Level

Marko Simonović, CERN

based on 1811.10640

with Valentin Assassi, Marcel Schmittfull and Matias Zaldarriaga

South American Workshop on Cosmology in the LSST Era
December 17, 2018

Motivation

Perturbative bias model: how well does it work?

What models should we use for galaxy surveys?

What are the properties of the noise (stochastic part)?

The goal is to get unbiased cosmological parameters

Motivation

These questions have been extensively explored in the past

Desjacques, Jeong, Schmit: Large-Scale Galaxy Bias

Most of the analyses use n-point functions. Disadvantages:

- — Cosmic variance, compromise on resolution/size of the box
- — At high k hard to disentangle the nonlinearities
- — Overfitting (smooth curves, many parameters)
- — Only a few lowest n-point functions used
- — Difficult to isolate and study the noise

Motivation

These problems can be solved using fields rather than summary statistics

Baldauf, Schaan, Zaldarriaga (2015)
Lazeyras, Schmit (2017)
Abidi, Baldauf (2018)
McQuinn, D'Aloisio (2018)

Advantages:

- No cosmic variance, small boxes with high resolution are enough
- High S/N at low k, no need to go to the nonlinear regime
- No overfitting, each Fourier mode (amplitude and phase) is fitted
- "All" n-point functions measured simultaneously
- It is easy to isolate and study the noise

Motivation

Generate realizations using PT and simulations form the same ICs

$$\begin{tikzcd} & \text{Same initial conditions} \arrow[dl, "\text{Sim.}"'] \arrow[dr, "\text{PT}"] & \\ \delta_h^{\mathrm{truth}} & & \delta_h \end{tikzcd}$$

If PT was perfect, the two fields would be the same (all Fourier modes the same)

Outline:

Part 1: How to do PT on the field level?

Part 2: Comparison to simulations

Conclusions

Part 1: PT on the filed level

PT on the filed level

Given ICs, one can calculate the nonlinear field using standard PT kernels

nth order solution is a convolution of *n* initial Fourier modes

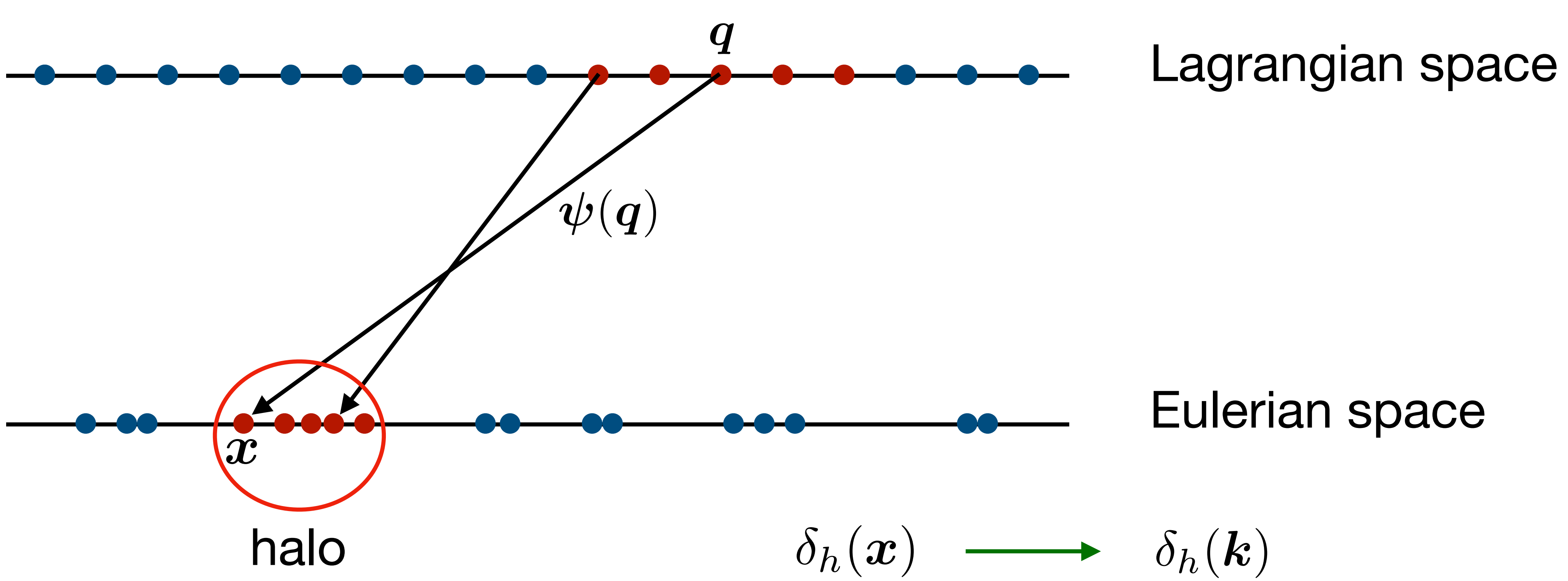
The problem are large displacements
(which do not cancel like for the *n*-point functions)

The standard Eulerian PT does not work, we need “IR resummation”

Lagrangian PT does not have this problem, but it gives only displacement…

PT on the filed level

We want a hybrid scheme



$$\boldsymbol{\psi}_1(\boldsymbol{q}) = \int_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{q}} \frac{i\boldsymbol{k}}{k^2} \delta_1(\boldsymbol{k})$$

linear displacement is large

Bias expansion on the field level

$$\delta_h^{\mathrm{L}}(\boldsymbol{q}) = b_1^{\mathrm{L}}\,\delta_1(\boldsymbol{q}) \,+\, b_2^{\mathrm{L}}\,(\delta_1^2(\boldsymbol{q}) - \sigma_1^2) \,+\, b_{\mathcal{G}_2}^{\mathrm{L}}\mathcal{G}_2(\boldsymbol{q}) + \cdots$$

$$\sigma_1^2 = \langle \delta_1^2(\boldsymbol{q}) \rangle = \int_0^\infty \frac{\mathrm{d}k}{2\pi^2}\, k^2 P_{11}(k)$$

$$\downarrow$$

$$\delta_h(\boldsymbol{k}) \equiv \int \mathrm{d}^3\boldsymbol{x}\,(1+\delta_h(\boldsymbol{x}))\,e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} = \int \mathrm{d}^3\boldsymbol{q}\,(1+\delta_h(\boldsymbol{q}))\,e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}(\boldsymbol{q}))}$$

$$\downarrow$$

$$\delta_h(\boldsymbol{k}) = \int \mathrm{d}^3\boldsymbol{q}\,\Big(1 + b_1^{\mathrm{L}}\,\delta_1(\boldsymbol{q}) \,+\, b_2^{\mathrm{L}}\,(\delta_1^2(\boldsymbol{q}) - \sigma_1^2) \,+\, b_{\mathcal{G}_2}^{\mathrm{L}}\mathcal{G}_2(\boldsymbol{q}) + \cdots$$
$$- i\boldsymbol{k}\cdot\boldsymbol{\psi}_2(\boldsymbol{q}) + \cdots\Big)\,e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}_1(\boldsymbol{q}))}$$

The usual approximation in LPT   for example: Vlah, Castorina, White (2016)

Bias expansion on the field level

This motivates us to write the bias expansion using the “shifted” operators

$$\tilde{\mathcal{O}}(\boldsymbol{k}) \equiv \int \mathrm{d}^3\boldsymbol{q}\; \mathcal{O}(\boldsymbol{q})\, e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}_1(\boldsymbol{q}))}$$

PT prediction

$$\delta_h(\boldsymbol{k}) = b_1\,\tilde{\delta}_1(\boldsymbol{k}) + b_2\,\tilde{\delta}_2(\boldsymbol{k}) + b_{\mathcal{G}_2}\,\tilde{\mathcal{G}}_2(\boldsymbol{k}) + \cdots \quad \text{+ noise}$$

Everything written in Eulerian space, easy comparison to simulations

IR resummation included, correct positions of halos, spread of the BAO peak…

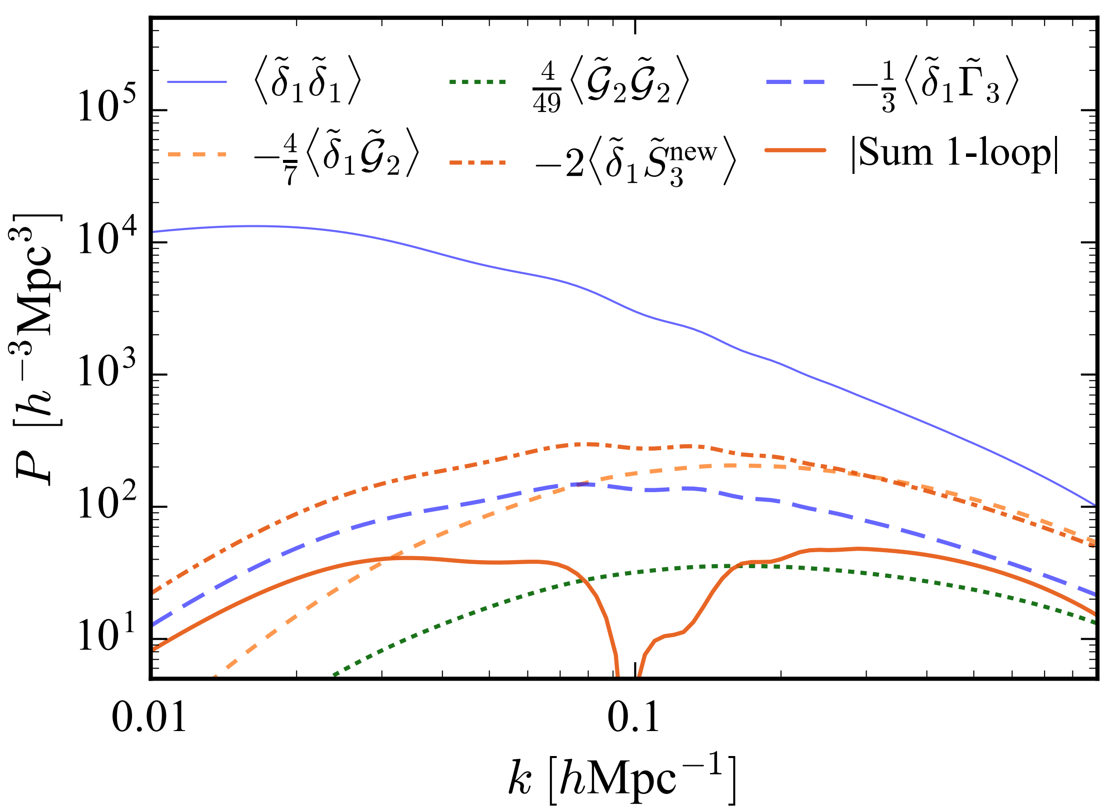
Shifted operators easy to generate, analytical calculations straightforward

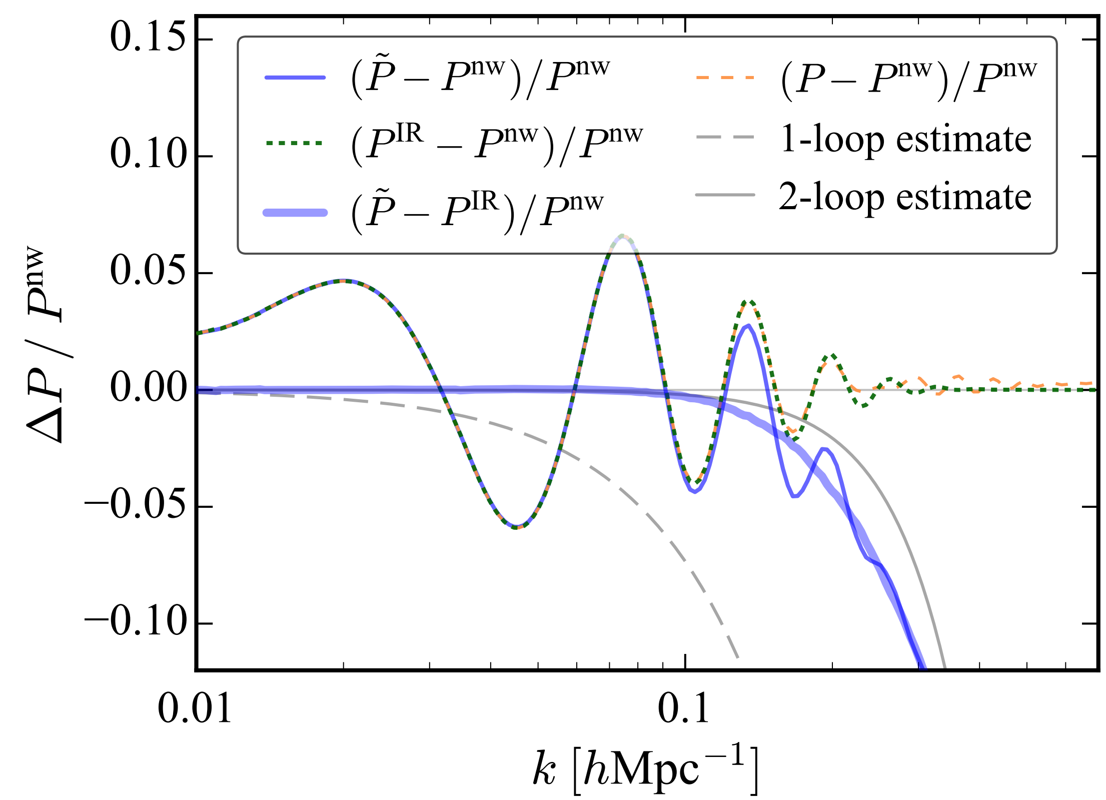
Only linear fields used in the construction

Bias expansion on the field level

Example of DM

$$\tilde{\delta} \;=\; \tilde{\delta}_1 \;+\; \frac{2}{7}\,\tilde{\mathcal{G}}_2 \;-\; \frac{3}{14}\,[\widetilde{\mathcal{G}_2\delta}] \;-\; \frac{2}{9}\,\tilde{\mathcal{G}}_3 \;+\; \frac{1}{6}\,\tilde{\Gamma}_3 \;-\; \tilde{\mathcal{S}}_3$$





The same results as in the standard PT approach with IR resummation

Bias expansion on the field level

What are the operators that we need for the one-loop prediction?

$$\delta_h(\boldsymbol{k}) = b_1\,\tilde{\delta}_1(\boldsymbol{k}) + b_2\,\tilde{\delta}_2(\boldsymbol{k}) + b_{\mathcal{G}_2}\,\tilde{\mathcal{G}}_2(\boldsymbol{k}) + \sum_i b_3^i\,\tilde{\mathcal{O}}_3^i$$

$$\tilde{\mathcal{O}}_3^i = \frac{\langle\tilde{\delta}_1\tilde{\mathcal{O}}_3^i\rangle}{\langle\tilde{\delta}_1\tilde{\delta}_1\rangle}\,\tilde{\delta}_1 + \left(\tilde{\mathcal{O}}_3^i - \frac{\langle\tilde{\delta}_1\tilde{\mathcal{O}}_3^i\rangle}{\langle\tilde{\delta}_1\tilde{\delta}_1\rangle}\,\tilde{\delta}_1\right) \equiv \frac{\langle\tilde{\delta}_1\tilde{\mathcal{O}}_3^i\rangle}{\langle\tilde{\delta}_1\tilde{\delta}_1\rangle}\,\tilde{\delta}_1 + \tilde{\mathcal{O}}_3^{i\perp}$$

No contribution
at 1-loop

Keep the second order fields, promote biases to k-dependent functions

Bias expansion on the field level

What are the operators that we need for the one-loop prediction?

$$\delta_h(\boldsymbol{k}) = b_1(k)\,\tilde{\delta}_1(\boldsymbol{k}) + b_2(k)\,\tilde{\delta}_2(\boldsymbol{k}) + b_{\mathcal{G}_2}(k)\,\tilde{\mathcal{G}}_2(\boldsymbol{k}) + \cdots$$

The final step is to make these fields orthogonal (simplifies the analysis)

$$\delta_h(\boldsymbol{k}) = \beta_1(k)\ \tilde{\delta}_1(\boldsymbol{k}) + \beta_2(k)\ \tilde{\delta}_2^{\perp}(\boldsymbol{k}) + \beta_{\mathcal{G}_2}(k)\ \tilde{\mathcal{G}}_2^{\perp}(\boldsymbol{k}) + \cdots$$

transfer functions

This is the model that we compare against simulations

Bias expansion on the field level

How much of the true halo density field correlates with these operators?

$$\delta_h(\boldsymbol{k}) = \beta_1(k)\ \tilde{\delta}_1(\boldsymbol{k}) + \beta_2(k)\ \tilde{\delta}_2^{\perp}(\boldsymbol{k}) + \beta_{\mathcal{G}_2}(k)\ \tilde{\mathcal{G}}_2^{\perp}(\boldsymbol{k}) + \cdots$$

$$\begin{aligned}
\beta_1(k) &= b_1 + c_s^2 k^2 + b_2 \frac{\langle \tilde{\delta}_1 \tilde{\delta}_2 \rangle}{\langle \tilde{\delta}_1 \tilde{\delta}_1 \rangle} + b_{\mathcal{G}_2} \frac{\langle \tilde{\delta}_1 \tilde{\mathcal{G}}_2 \rangle}{\langle \tilde{\delta}_1 \tilde{\delta}_1 \rangle} + b_{\Gamma_3} \frac{\langle \tilde{\delta}_1 \tilde{\Gamma}_3 \rangle}{\langle \tilde{\delta}_1 \tilde{\delta}_1 \rangle} - b_1 \frac{\langle \tilde{\delta}_1 \tilde{\mathcal{S}}_3 \rangle}{\langle \tilde{\delta}_1 \tilde{\delta}_1 \rangle} \\
\beta_2(k) &= b_2 + b_{\mathcal{G}_2} \frac{\langle \tilde{\delta}_2 \tilde{\mathcal{G}}_2 \rangle}{\langle \tilde{\delta}_2 \tilde{\delta}_2 \rangle} \\
\beta_{\mathcal{G}_2}(k) &= b_{\mathcal{G}_2}
\end{aligned}$$

The number of parameters is the same as in the standard 1-loop power spectrum

Part 2: Comparison to simulations

Comparison to simulations

How do we choose transfer functions (bias parameters)?

Minimize the difference of the model wrt. the true halo field

$$\sum_{\boldsymbol{k}, |\boldsymbol{k}| \approx k} |\delta_h^{\mathrm{truth}}(\boldsymbol{k}) - \delta_h^{\mathrm{model}}(\boldsymbol{k})|^2$$

Fitting the bias model on the entire field, instead *n*-point functions

An example:

$$\delta_h^{\mathrm{truth}} = b_1 \delta + \epsilon \longrightarrow b_1(k) = \frac{\langle \delta_h^{\mathrm{truth}}(\boldsymbol{k}) \delta^*(\boldsymbol{k}) \rangle}{\langle |\delta(\boldsymbol{k})|^2 \rangle}$$

More generally, for orthogonal fields: $\beta_i(k) = \dfrac{\langle \delta_h^{\mathrm{truth}} \tilde{\mathcal{O}}_i^{\perp} \rangle}{\langle \tilde{\mathcal{O}}_i^{\perp} \tilde{\mathcal{O}}_i^{\perp} \rangle}$

Comparison to simulations

What is the measure of success?

Cross-correlation coefficient: $r_{cc}(k) \equiv \frac{\langle \delta_h^{\rm model}(\boldsymbol{k})[\delta_h^{\rm truth}(\boldsymbol{k})]^* \rangle}{\left(\langle |\delta_h^{\rm model}(\boldsymbol{k})|^2 \rangle \langle |\delta_h^{\rm truth}(\boldsymbol{k})|^2 \rangle\right)^{1/2}}$

The power spectrum of the model error $\hat{\epsilon} \equiv \delta_h^{\rm truth} - \delta_h^{\rm model}$

$$P_{\rm err}(k) \equiv \langle |\hat{\epsilon}(\boldsymbol{k})|^2 \rangle$$

For the best-fit model $P_{\rm err}(k) = P_{\rm truth}(1 - r_{cc}^2)$

$$(P_{\rm model}/P_{\rm truth})^{1/2} = r_{cc}$$

Comparison to simulations

Common expectations:

The noise is always close to Poisson and scale-independent

Linear bias model is good enough, if DM from simulation is used

PT completely beaks down at the nonlinear scale (r_{cc} goes to zero)

Comparison to simulations

5 boxes, L = 500 Mpc/*h*, N = 1536^3, m = $3*10^9$ M_{sun}/h, z = 0.6

Halos identified using the standard FOF algorithm

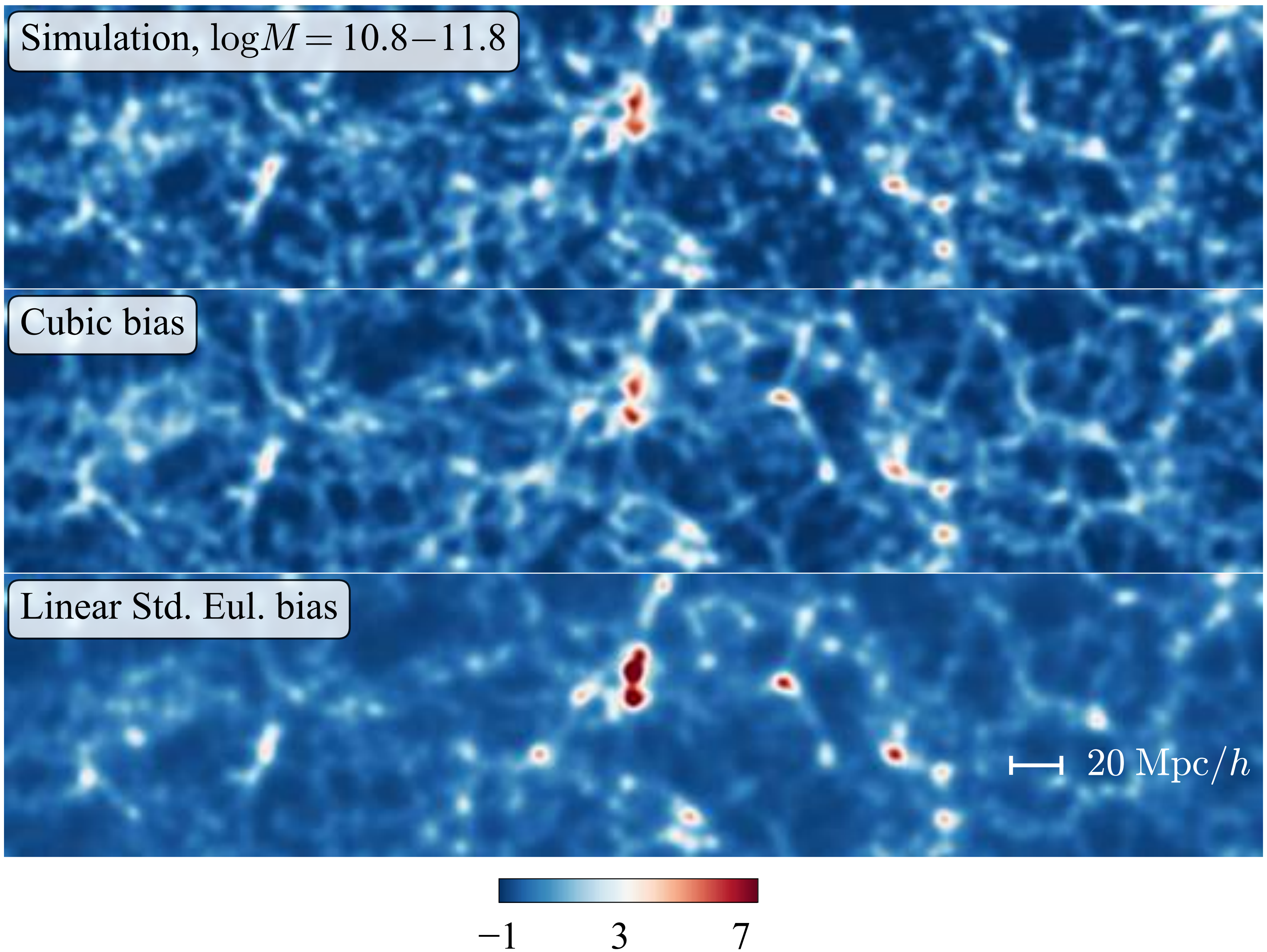
4 mass bins

$\log M[h^{-1}\mathrm{M}_\odot]$	$\bar{n}\,[(h^{-1}\mathrm{Mpc})^{-3}]$	$\bar{n}$ is comparable to
$10.8 - 11.8$	4.3×10^{-2}	LSST [80, 81], Billion Object Apparatus [82]
$11.8 - 12.8$	5.7×10^{-3}	SPHEREx [83, 84]
$12.8 - 13.8$	5.6×10^{-4}	BOSS CMASS [85], DESI [86, 87], Euclid [88–90]
$13.8 - 15.2$	2.6×10^{-5}	Cluster catalogs

Table I. Simulated halo populations at $z = 0.6$.

Comparison to simulations

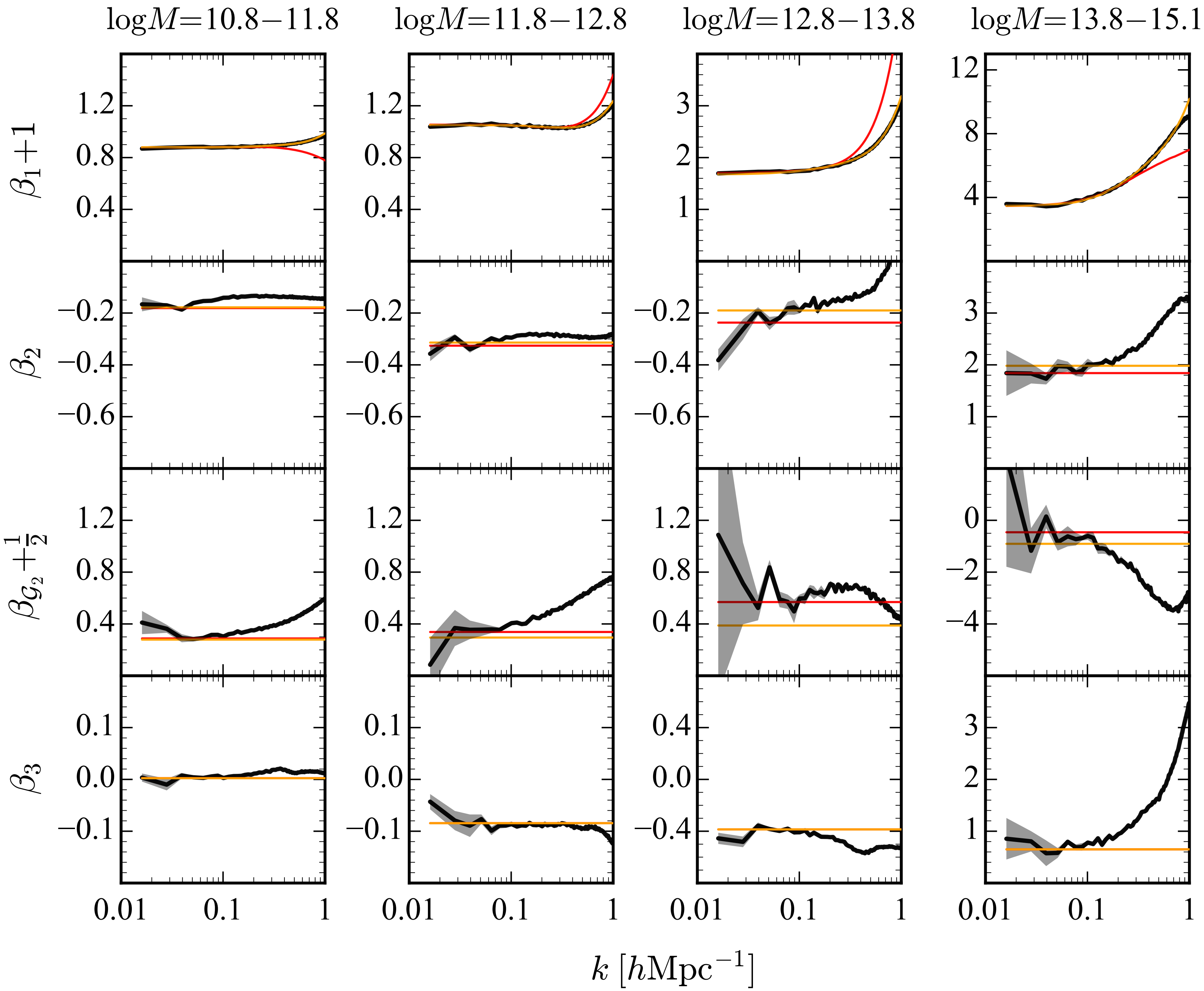
In real space things look decent



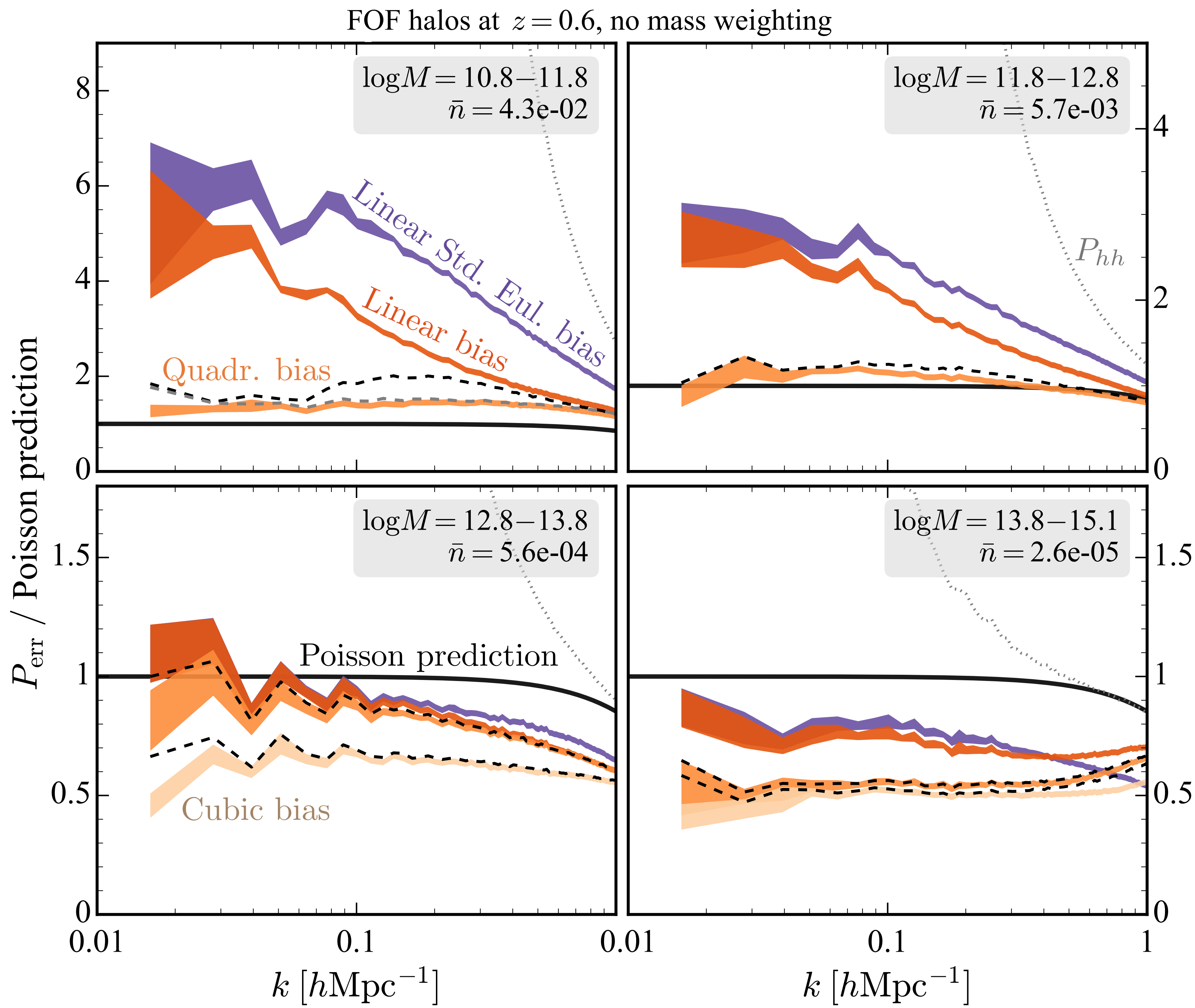
$\delta_h(\boldsymbol{x})$

Comparison to simulations

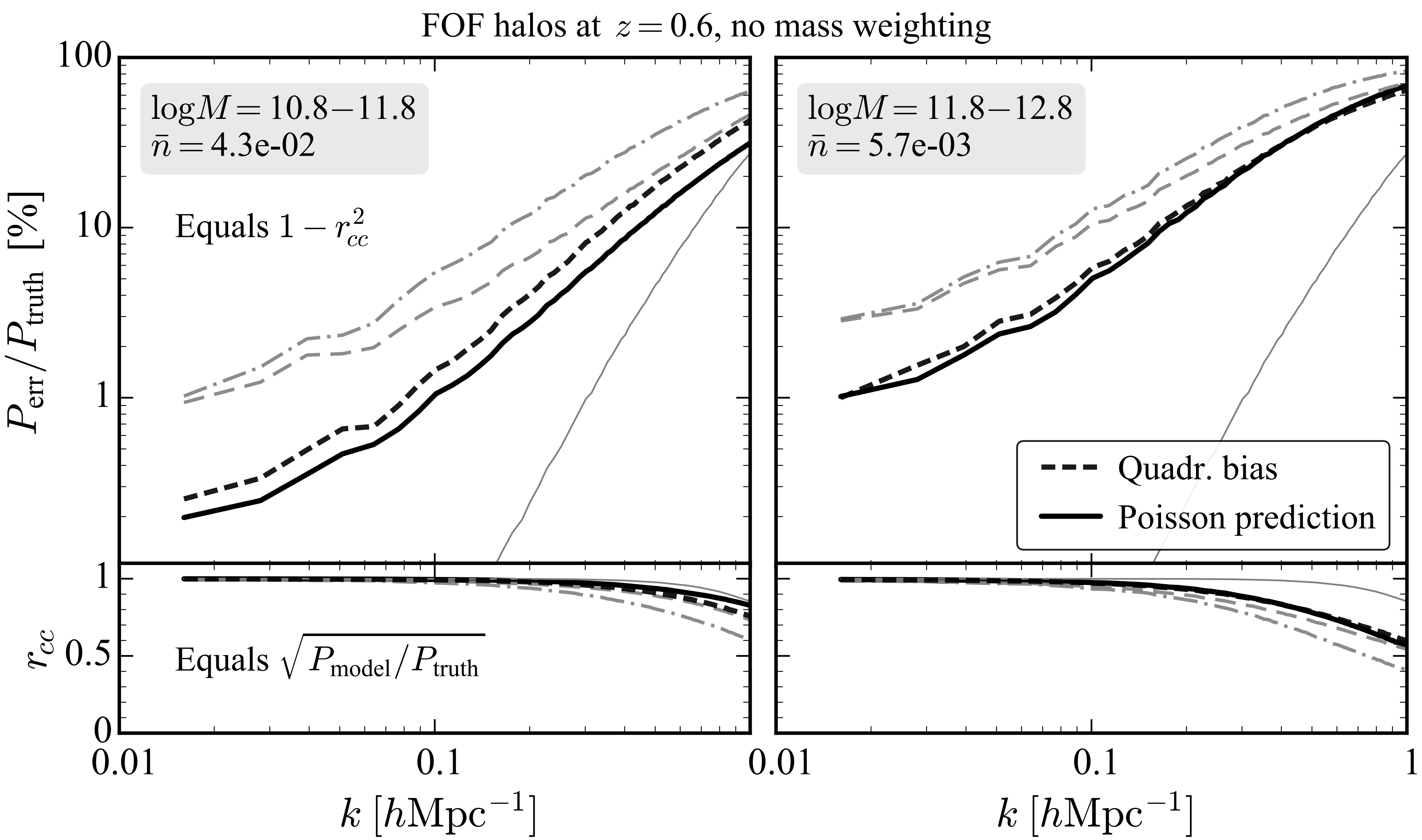
Transfer functions



Comparison to simulations

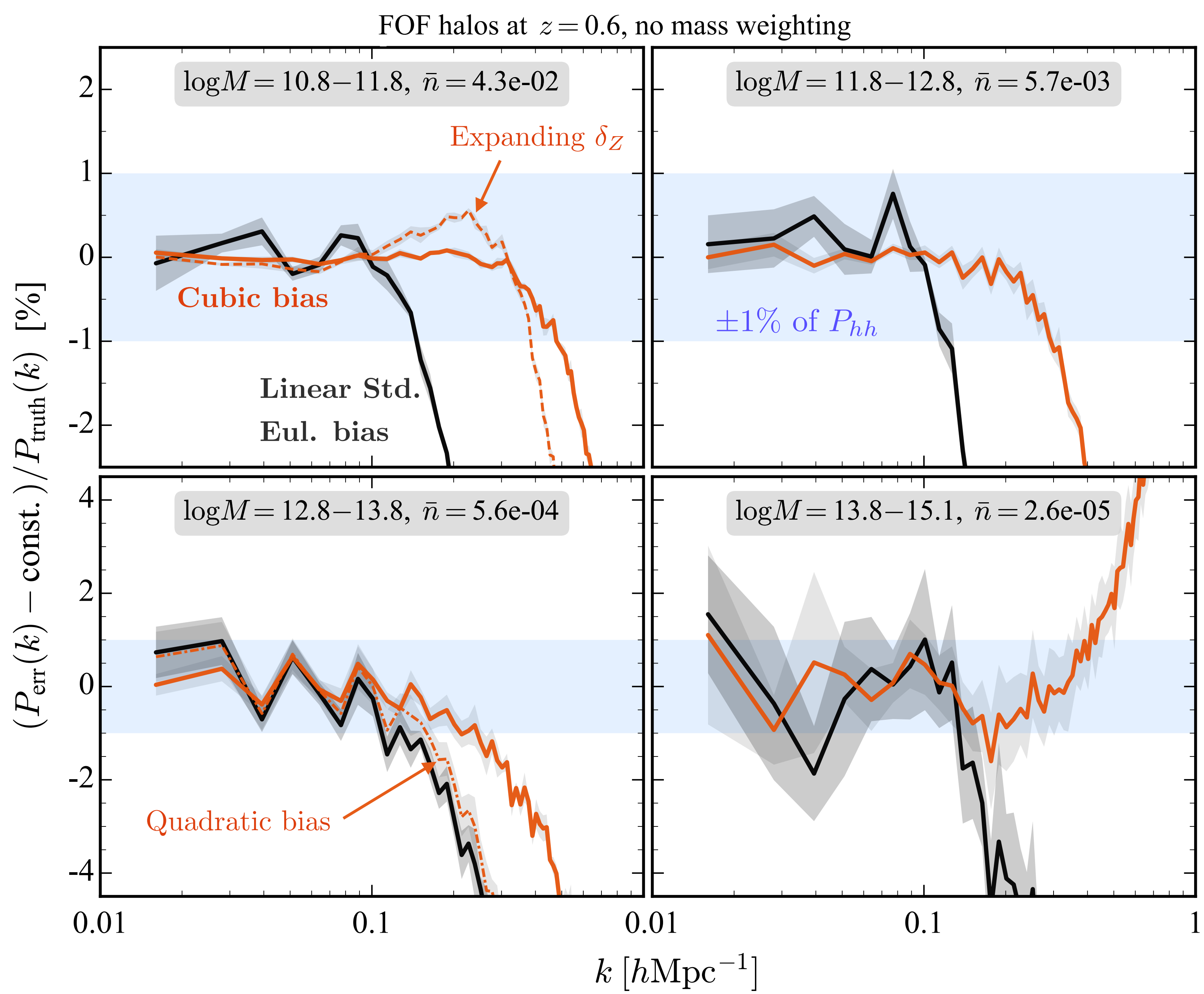


Comparison to simulations



Comparison to simulations

The scale-dependence of the noise is relevant for data analysis

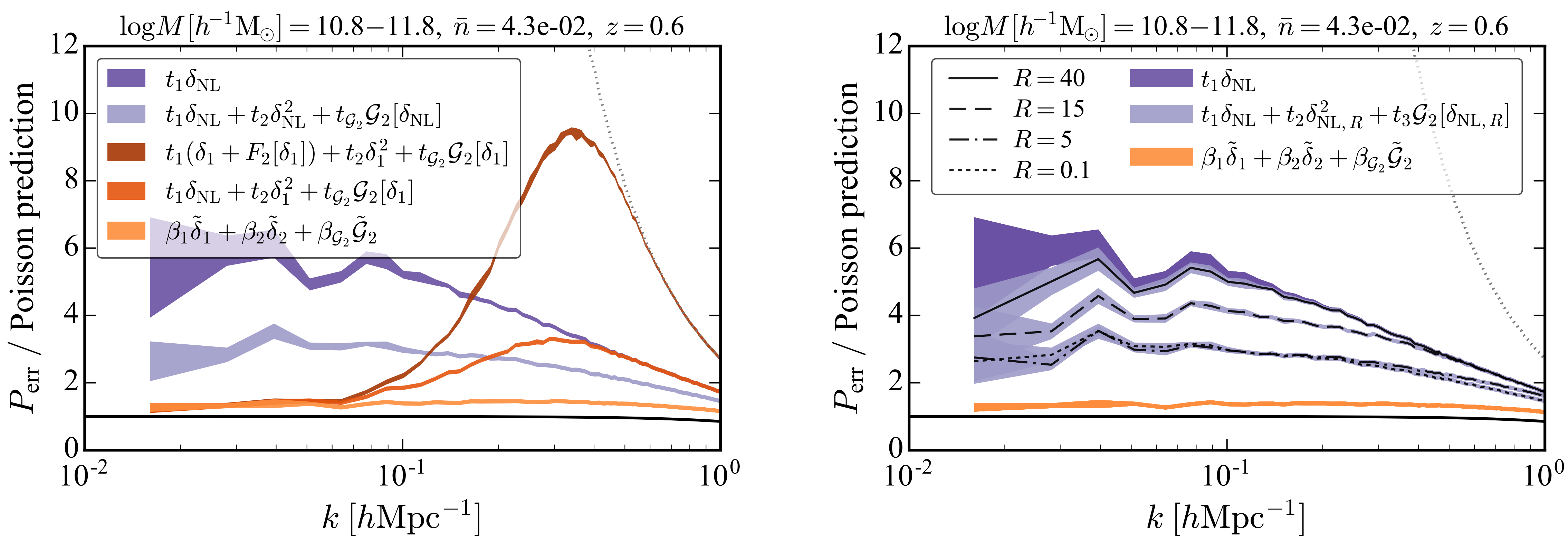


Comparison to simulations

Why not using the nonlinear DM field form simulations?

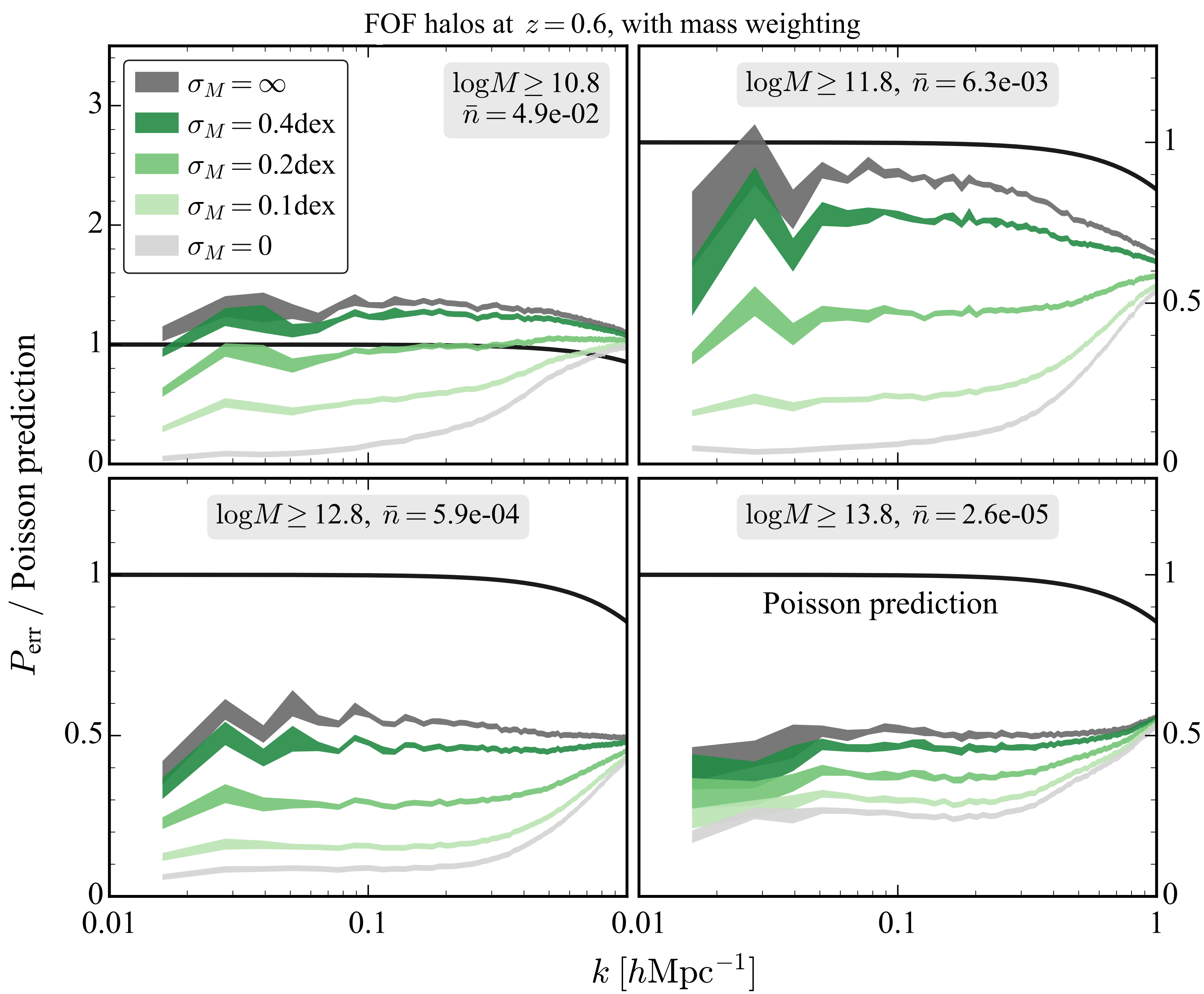
Small scales produce a lot of power if nonlinear DM field is used

Keeping the small scales is crucial to minimize the model error



Comparison to simulations

Mass-weighting reduces the noise



Conclusions

Comparison on the field level is very useful, no CV, all n-point functions

No bias model (linear or nonlinear) with simulation DM field is optimal

Good news is that the simple quadratic bias works quite well

The PT bias model has the noise close to Poisson and scale-independent

Small scale-dependence of the noise (~1%) relevant around the nonlinear scale

How well the perturbative bias model works for galaxies in redshift space?

Applications to the reconstruction of the initial conditions

A smooth forward model, useful for the likelihood on the field level

Backup slides

Open questions

All results presented so far are UV-dependent

Low-k limits of the transfer functions are not renormalized bias parameters

$$\beta_i(k) = \frac{\langle \delta_h^{\rm truth} \tilde{\mathcal{O}}_i^{\perp} \rangle}{\langle \tilde{\mathcal{O}}_i^{\perp} \tilde{\mathcal{O}}_i^{\perp} \rangle}$$

sensitive to high k and
definition of operators

Relation to similar methods to measure physical biases

Lazeyras, Schmit (2017)
Abidi, Baldauf (2018)

The key difference is the smoothing (not including the short modes)

R ~ 10-20 Mpc

Open questions

Standard bias expansion is well-defined and rigorous on large scales

It relays on integrating out the short scales, this leads to higher noise

Minimization on the field level is another way to define biases

The results depend on the UV completion

The noise is smaller and more well-behaved

Which procedure leads to tighter constraints on cosmological parameters?

Open questions

Power spectrum analysis:

The constant low-k part of the power spectrum is treated as noise

b_2 is measured only from the k dependence of the power spectrum

Large fraction of the constant low-k part of the power spectrum is treated as signal

b_2 is measured from the k dependence of the power spectrum and the constant part