Cosmology with Marked Power Spectra

Elena Massara

UNIVERSITY OF WATERLOO | WATERLOO CENTRE FOR ASTROPHYSICS

Cosmology from Home 2022

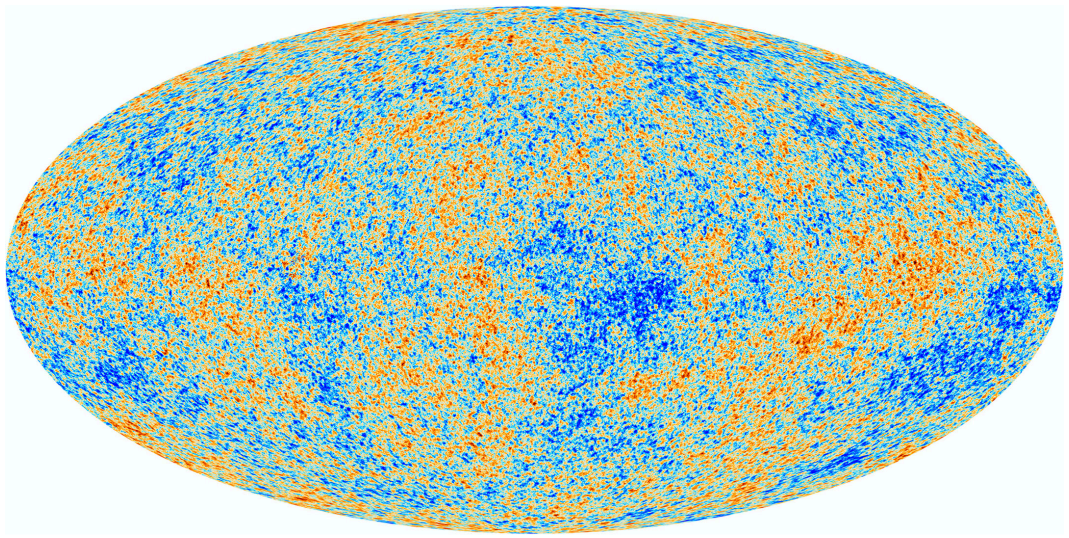
THE LARGE SCALE STRUCTURE

The distribution of matter in the Universe is sensitive to:

- properties of dark matter
- nature of dark energy
- neutrino mass scale
- initial condition of the Universe

$\longrightarrow \Omega_m, \Omega_b, h, \sigma_8, n_s, M_\nu$

NON-GAUSSIAN density field



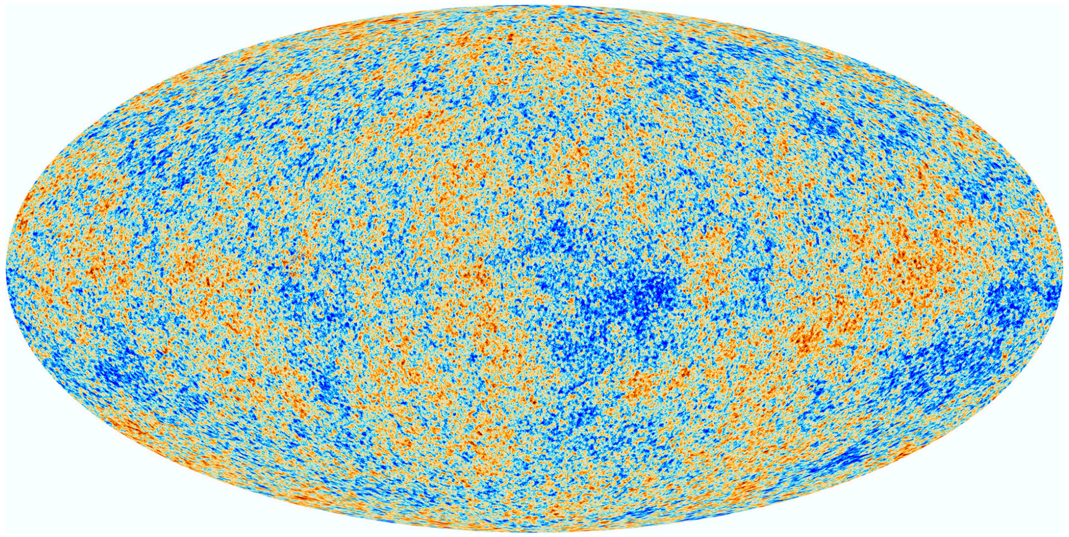
$$\delta(\mathbf{k}) \sim N(0, P(\mathbf{k}))$$

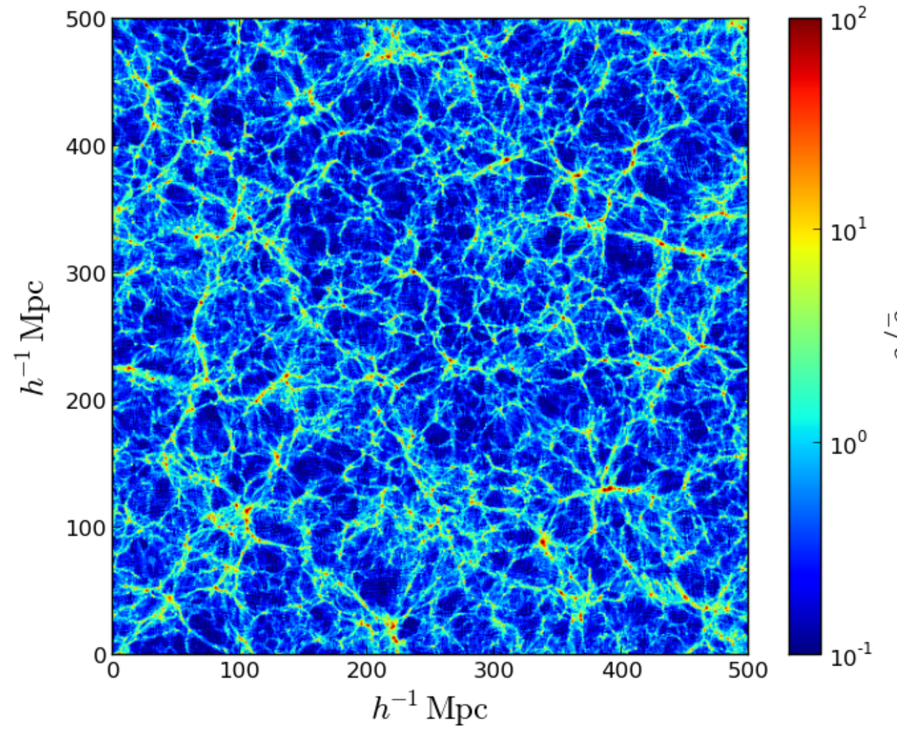
All information contained in 2-pt statistics:

- correlation function
- power spectrum

Higher order statistics are not needed to describe the field

NON-GAUSSIAN density field





$$\delta(\mathbf{k}) \sim N(0, P(\mathbf{k}))$$

$$\delta(\mathbf{k}) \nsim N(0, P(\mathbf{k}))$$

All information contained in 2-pt statistics:

- correlation function
- power spectrum

Higher order statistics are not needed to describe the field

NOT all information contained in 2-pt statistics

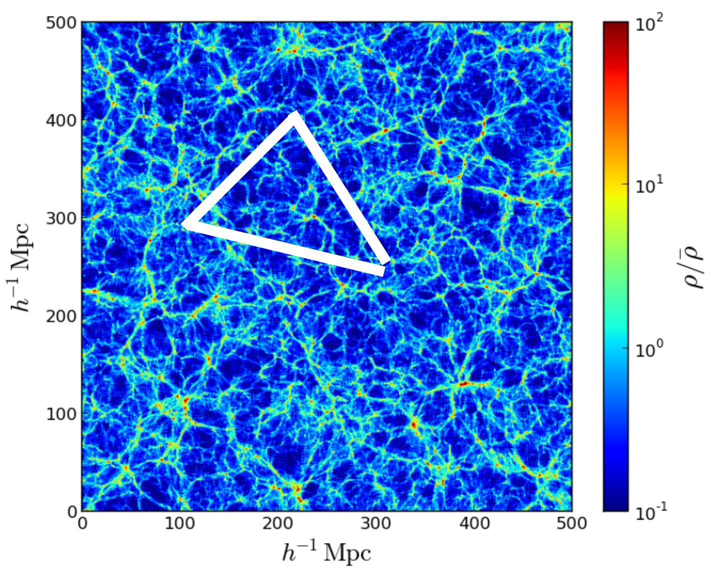
Higher order statistics contain information to describe the field

NON-GAUSSIAN statistics

A variety of statistics have been proposed to retrieve the cosmological information beyond the two point functions

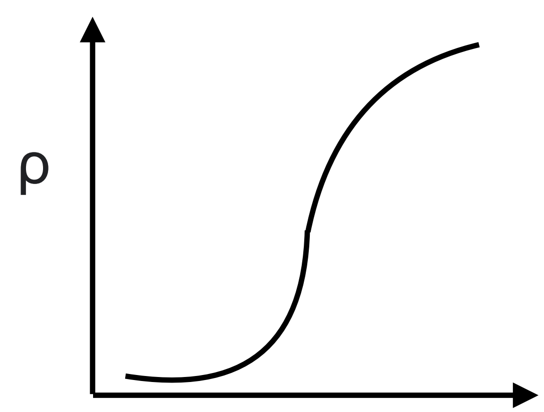
Higher-order statistics:

- bispectrum
- trispectrum
- …



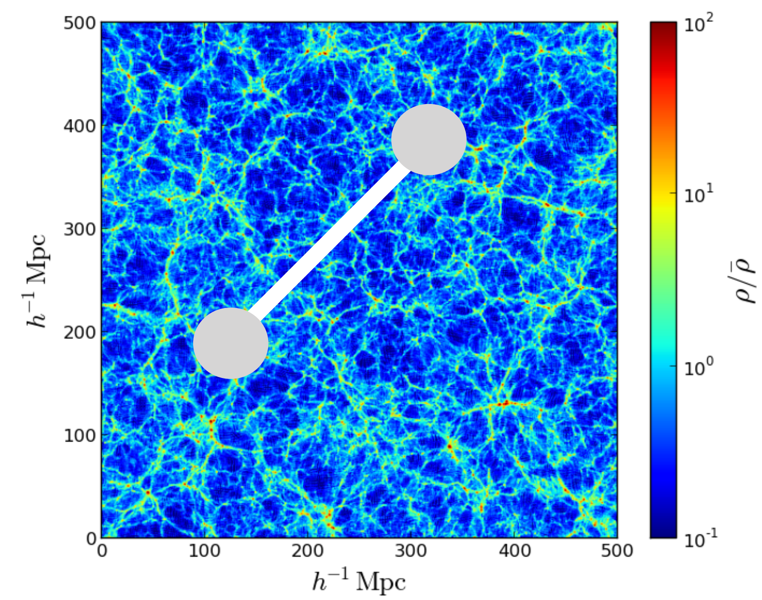
Different summary statistics:

- peaks
- voids
- scattering transforms
- minimum spanning tree
- ….



Non-linear transformations of the field:

- log-transformations
- marked power spectra



And many others …

COSMOLOGY with VOIDS

Low-density regions are good laboratories to study cosmology because

- They are unvirialized, thus they are expected to retain most of their initial cosmological information

COSMOLOGY with LOW-DENSITY regions

Low-density regions are good laboratories to study cosmology

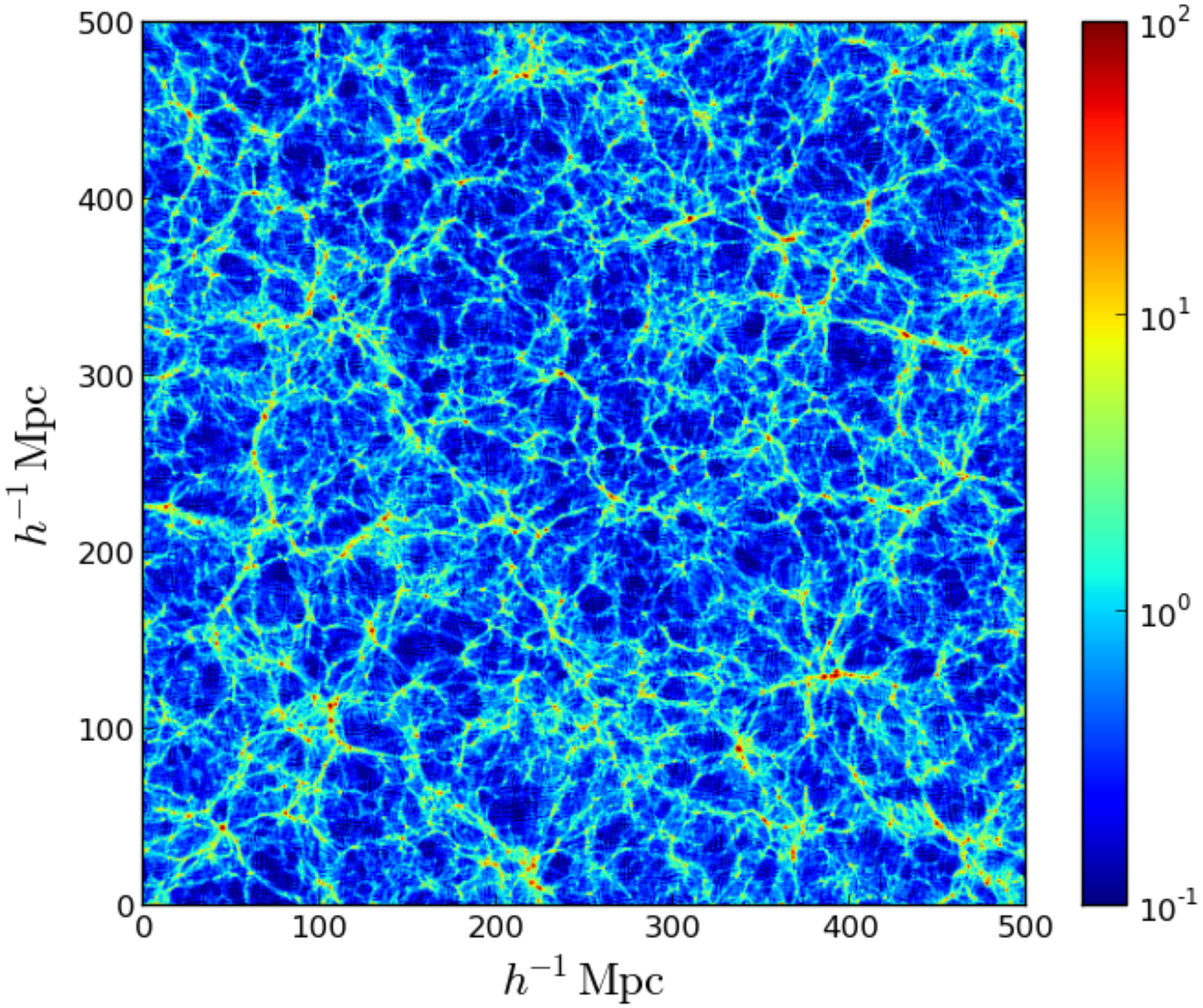
because

- They are unvirialized, thus they are expected to retain most of their initial cosmological information

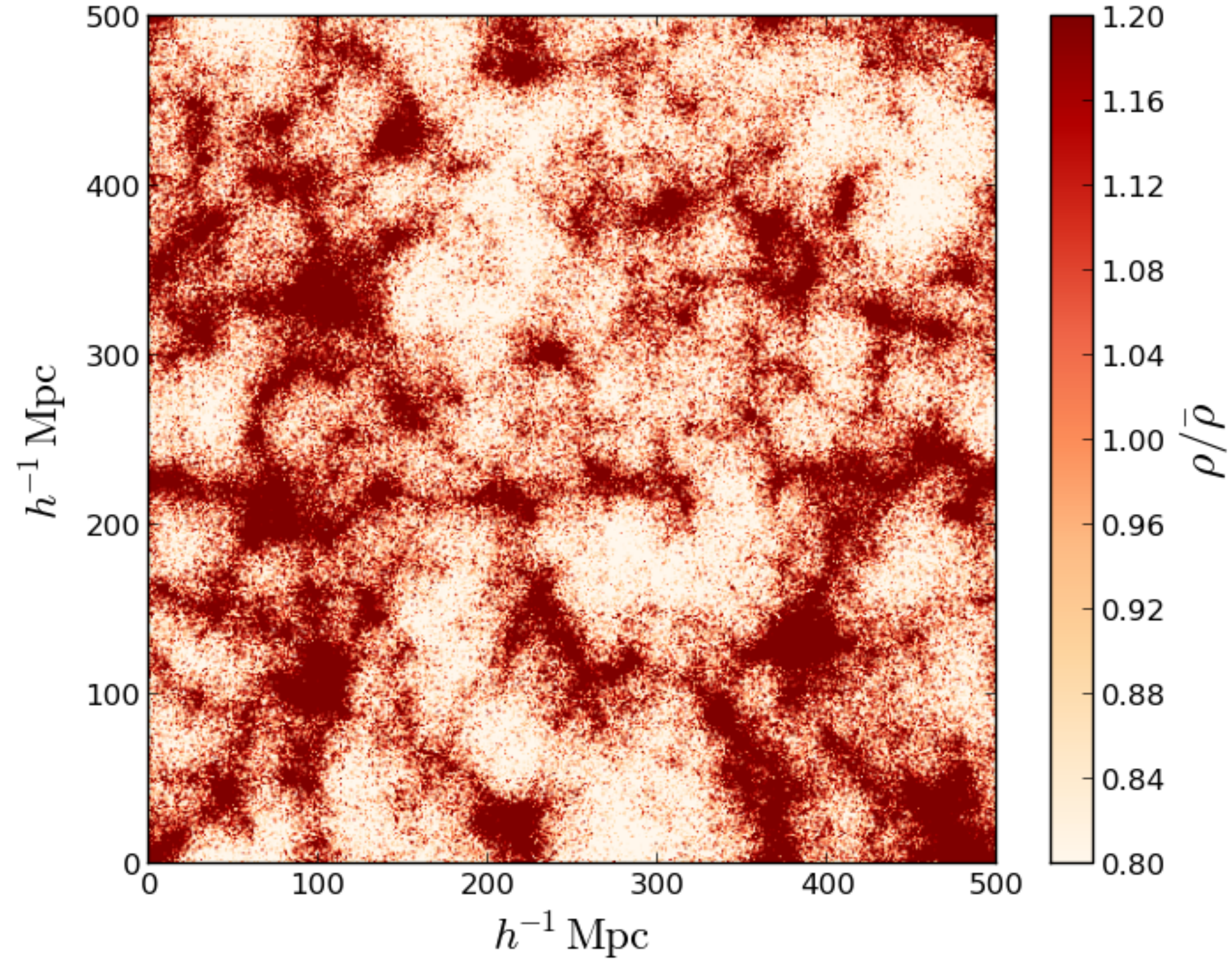
- They are sensitive to diffuse components such as
    - neutrinos
    - dark energy

COSMOLOGY with LOW-DENSITY regions

CDM density field



Neutrino density field



COSMOLOGY with LOW-DENSITY regions

Low-density regions are good laboratories to study cosmology because

- They are unvirialized, thus they are expected to retain most of their initial cosmological information
- They are sensitive to diffuse components such as
  - neutrinos
  - dark energy
- Screening mechanism are inefficient in them

LOW DENSITY REGIONS

Low-density regions are good probe to study cosmology

↓

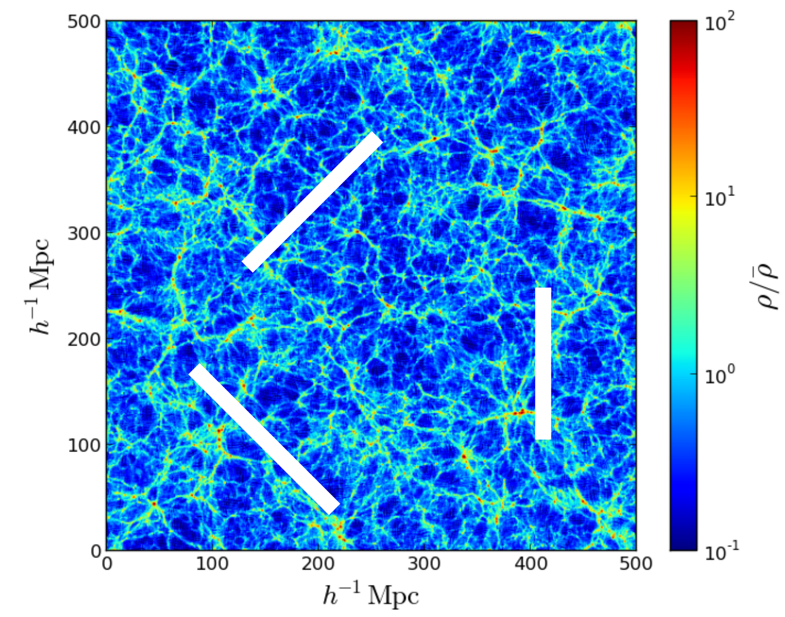
1. Do 2-pt functions depend on low-density regions?

2. Can we modify standard 2-pt functions to incorporate more information from low-density regions?

CORRELATION FUNCTION

Correlation function

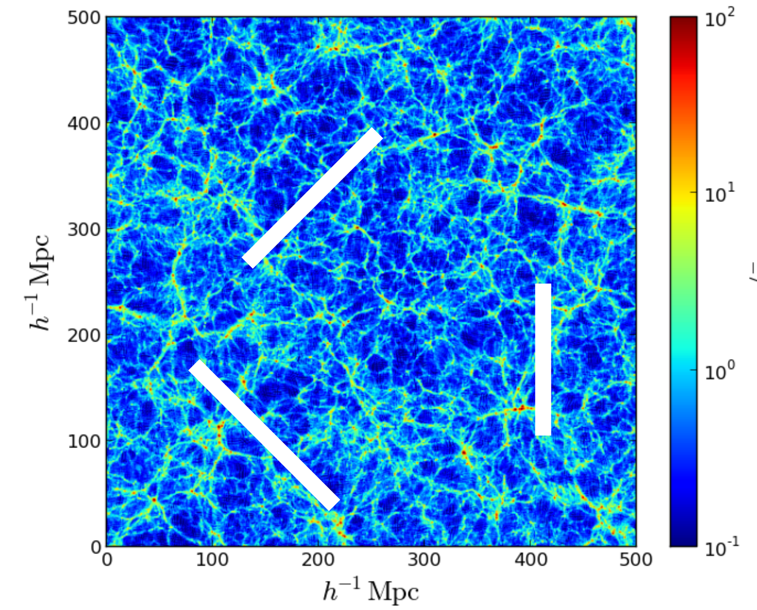
$$1 + \xi(r) = \frac{V}{N^2} \sum_{i,j=1}^{N} \delta_D(|\vec{x_i} - \vec{x_j}| - r)$$



CORRELATION FUNCTION

Correlation function

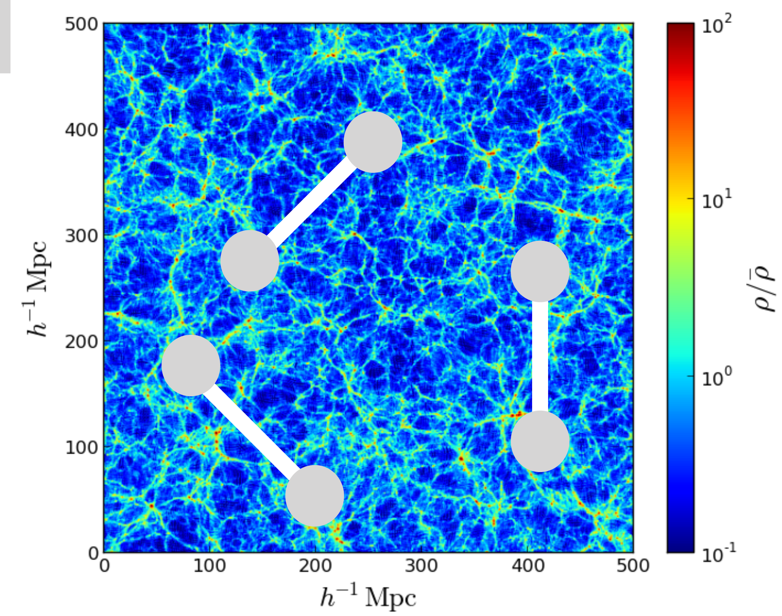
$$1 + \xi(r) = \frac{V}{N^2} \sum_{i,j=1}^{N} \delta_D(|\vec{x_i} - \vec{x_j}| - r)$$



Marked correlation function

$$1 + M(r, \phi) = \frac{V}{N^2} \sum_{i,j=1}^{N} \frac{\delta_D(|\vec{x_i} - \vec{x_j}| - r)\, m(\vec{x_i}, \phi)\, m(\vec{x_j}, \phi)}{\bar{m}^2}$$

1. m depends on the local density around each point
2. m up-weights low-density regions and down-weights high-density regions



MARKED CORRELATION FUNCTION

$$1 + M(r,\phi) = \frac{V}{N^2} \sum_{i,j=1}^{N} \frac{\delta_D(|\vec{x}_i - \vec{x}_j| - r)\, m(\vec{x}_i,\phi)\, m(\vec{x}_j,\phi)}{\bar{m}^2}$$

$$\boxed{m(\vec{x},\phi = R, p, \delta_s) = \left[\frac{1+\delta_s}{1+\delta_s+\delta_R(\vec{x})}\right]^p}$$

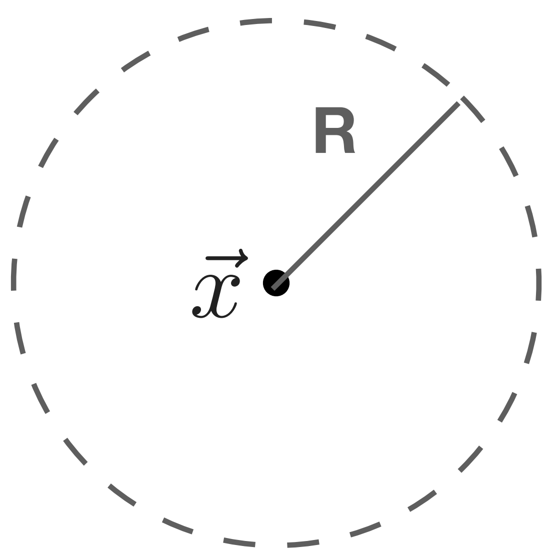
(M. White 2016)

MARKED CORRELATION FUNCTION

$$1 + M(r,\phi) = \frac{V}{N^2} \sum_{i,j=1}^{N} \frac{\delta_D(|\vec{x_i} - \vec{x_j}| - r)\, m(\vec{x_i},\phi)\, m(\vec{x_j},\phi)}{\bar{m}^2}$$

$$\boxed{m(\vec{x},\phi = R, p, \delta_s) = \left[\frac{1+\delta_s}{1+\delta_s+\delta_R(\vec{x})}\right]^p}$$

(M. White 2016)



$$\delta_R(\vec{x}) = \frac{1}{V_R}\int_{V_R} d^3y\, \delta(\vec{y})$$

MARKED CORRELATION FUNCTION

$$1 + M(r, \phi) = \frac{V}{N^2} \sum_{i,j=1}^{N} \frac{\delta_D(|\vec{x_i} - \vec{x_j}| - r)\, m(\vec{x_i}, \phi)\, m(\vec{x_j}, \phi)}{\bar{m}^2}$$

$$m(\vec{x}, \phi = R, p, \delta_s) = \left[\frac{1 + \delta_s}{1 + \delta_s + \delta_R(\vec{x})} \right]^p$$

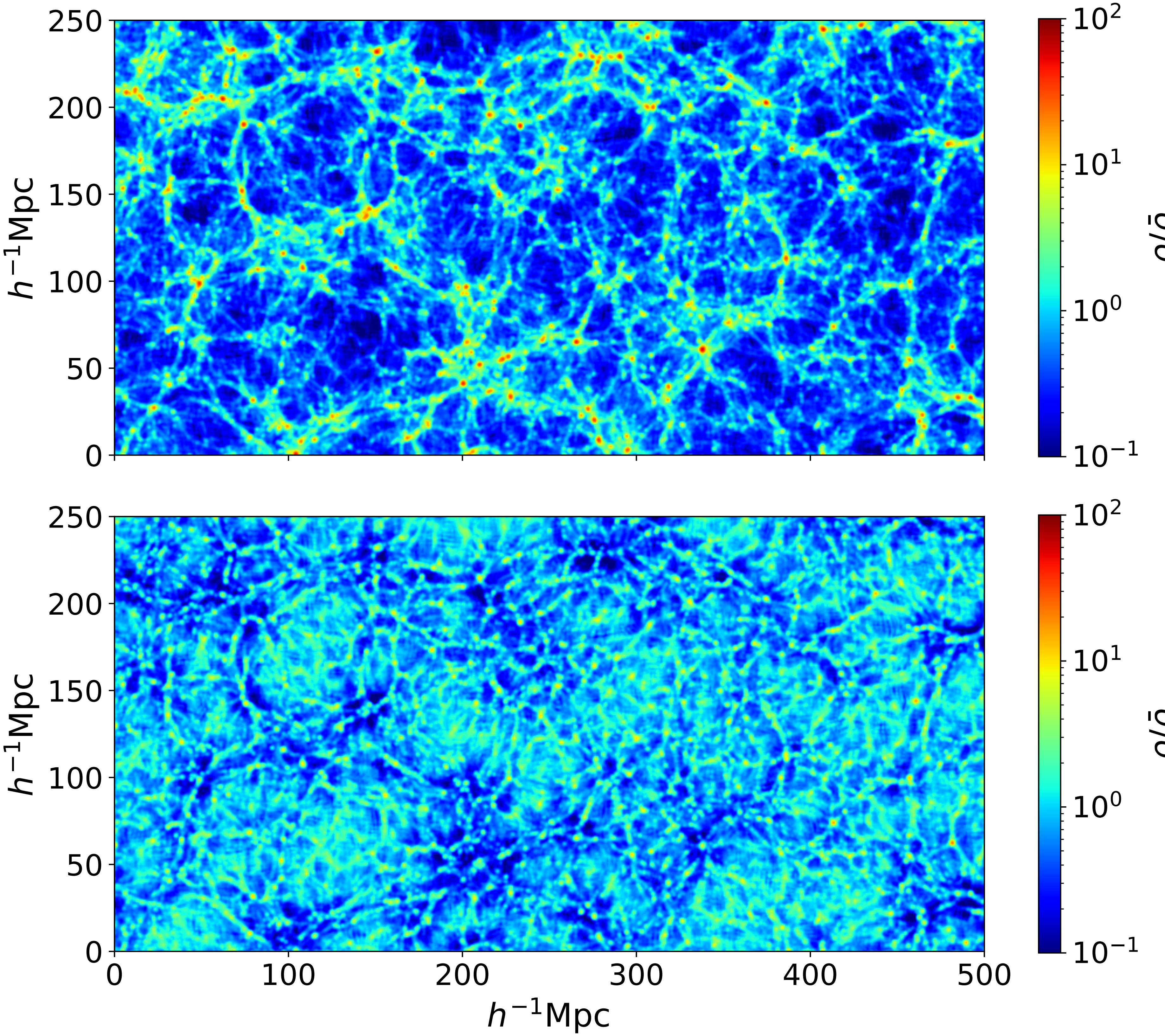
(M. White 2016)

$p > 0$ up-weight galaxies in **low** density regions

$p < 0$ up-weight galaxies in **high** density regions

MARKED DENSITY FIELD

EM et al. 2020



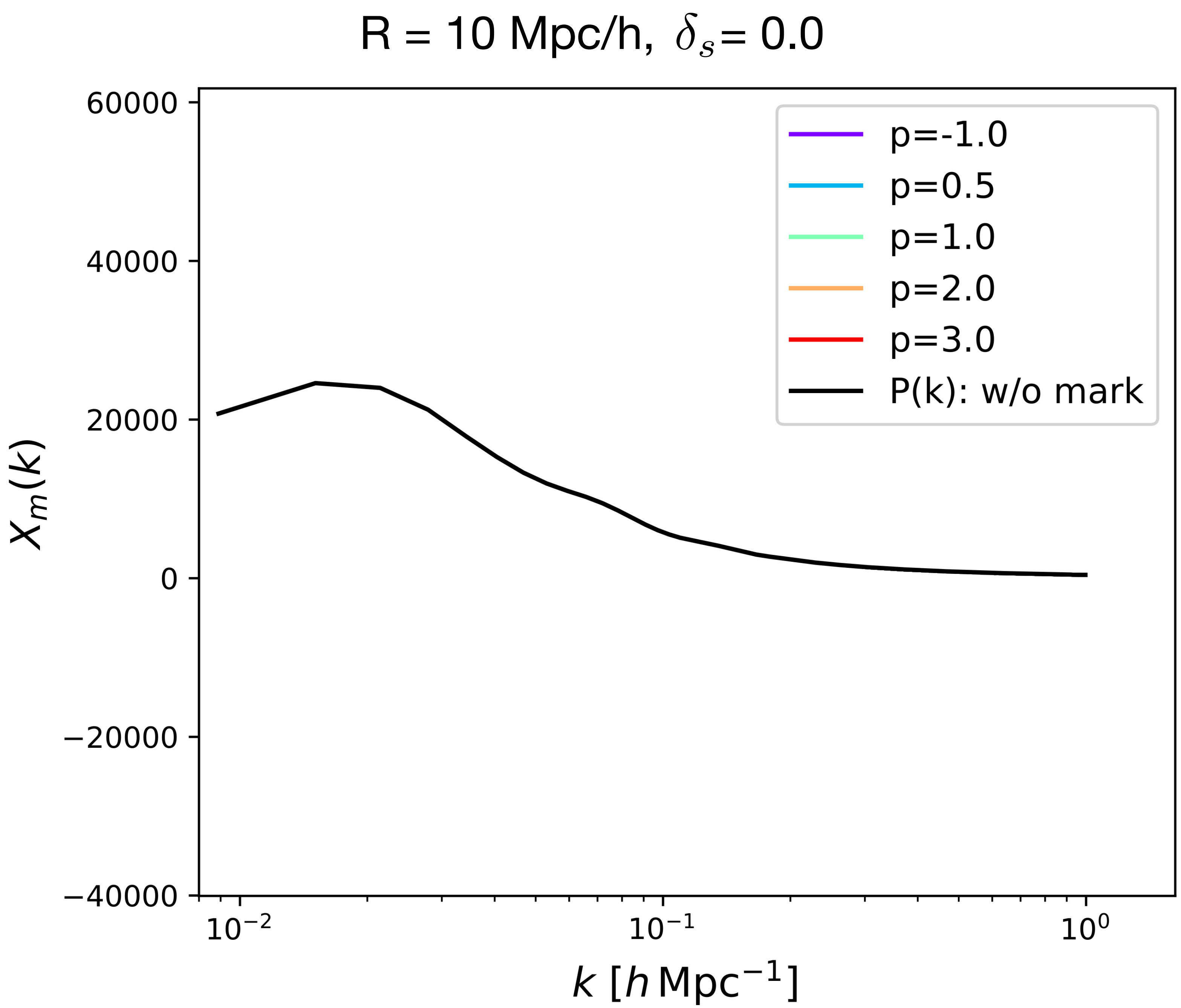
Density field

Marked density field with p>0

MARKED POWER SPECTRUM

EM et al. 2020

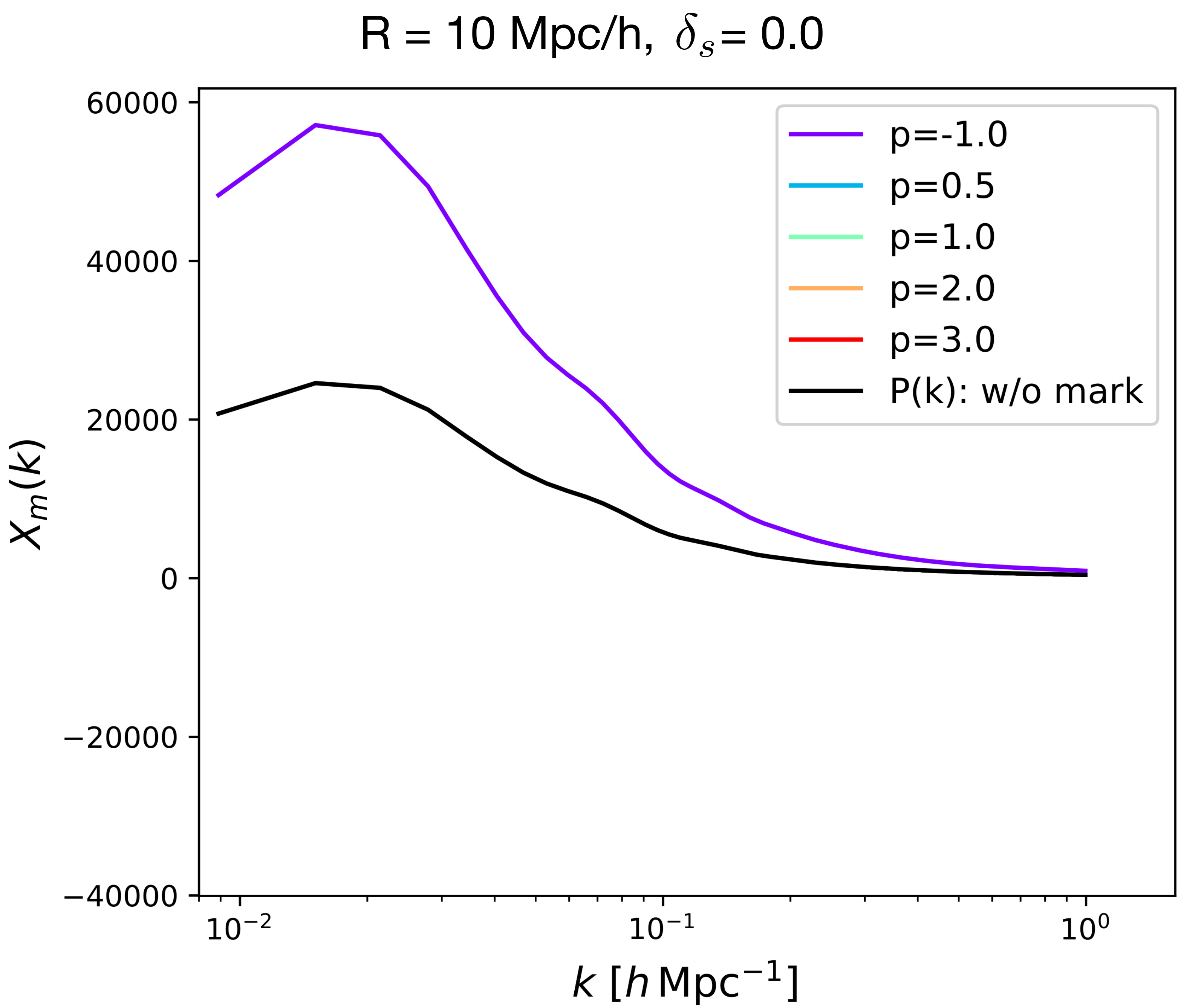
Marked-standard density cross-power spectrum



MARKED POWER SPECTRUM

EM et al. 2020

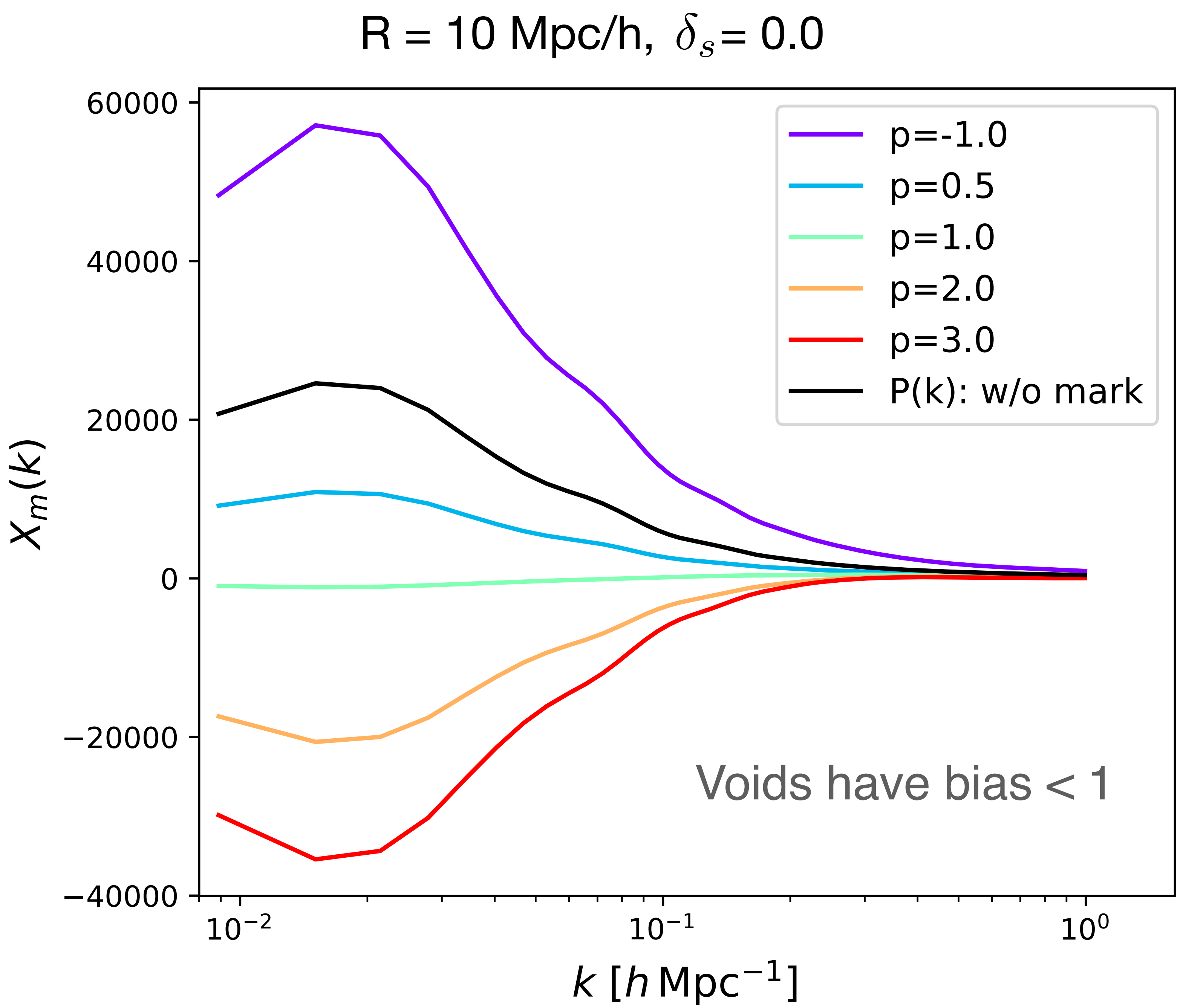
Marked-standard density cross-power spectrum



MARKED POWER SPECTRUM

EM et al. 2020

Marked-standard density cross-power spectrum



INFORMATION CONTENT in MARKED POWER SPECTRA of the MATTER FIELD

FISHER ANALYSIS

Cosmological parameters: $\vec{\theta} = \{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}$

Data vector (observables): $\vec{d} = \{P(k_1), P(k_2), .., P(k_n)\}$

Error on each parameter:

$$\sigma(\theta_\alpha) \leq \sqrt{(F^{-1})_{\alpha\alpha}}$$

Fisher matrix:

$$F_{\alpha,\beta} = \frac{\partial \vec{d}}{\partial \theta_\alpha} C^{-1} \frac{\partial \vec{d}}{\partial \theta_\beta}$$

QUIJOTE SIMULATIONS

Villaescusa-Navarro, Hanh, **EM** et al 2019

- https://github.com/franciscovillaescusa/Quijote-simulations
- Set of 43,100 full N-body simulations
- **1 Gpc/h** box size, **512^3 CDM** particles

  (512^3 neutrinos)
- More than 7000 models with different
  $$\Omega_m, \Omega_b, h, \sigma_8, n_s, M_\nu, \omega$$
- 1 Pb of publicly available data

QUIJOTE SIMULATIONS

Villaescusa-Navarro, Hanh, **EM** et al 2019

Boxes to compute the covariances

Name	Ω_m	Ω_b	h	n_s	σ_8	M_ν [eV]	realizations	ICs
Fiducial	0.3175	0.049	0.6711	0.9624	0.834	0	15,000	2LPT

Table 1. Description of the N-body simulations used in the Fisher analysis.

QUIJOTE SIMULATIONS

Villaescusa-Navarro, Hanh, **EM** et al 2019

Boxes to compute the numerical derivatives

Name	Ω_m	Ω_b	h	n_s	σ_8	M_ν [eV]	realizations	ICs
Fiducial ZA	0.3175	0.049	0.6711	0.9624	0.834	0	500	Zel'dovich
Ω_m^+	**0.3275**	0.049	0.6711	0.9624	0.834	0	500	2LPT
Ω_m^-	**0.3075**	0.049	0.6711	0.9624	0.834	0	500	2LPT
Ω_b^{++}	0.3175	**0.051**	0.6711	0.9624	0.834	0	500	2LPT
Ω_b^{--}	0.3175	**0.047**	0.6711	0.9624	0.834	0	500	2LPT
h^+	0.3175	0.049	**0.6911**	0.9624	0.834	0	500	2LPT
h^-	0.3175	0.049	**0.6511**	0.9624	0.834	0	500	2LPT
n_s^+	0.3175	0.049	0.6711	**0.9824**	0.834	0	500	2LPT
n_s^-	0.3175	0.049	0.6711	**0.9424**	0.834	0	500	2LPT
σ_8^+	0.3175	0.049	0.6711	0.9624	**0.849**	0	500	2LPT
σ_8^-	0.3175	0.049	0.6711	0.9624	**0.819**	0	500	2LPT
M_ν^+	0.3175	0.049	0.6711	0.9624	0.834	**0.1**	500	Zel'dovich
M_ν^{++}	0.3175	0.049	0.6711	0.9624	0.834	**0.2**	500	Zel'dovich
M_ν^{+++}	0.3175	0.049	0.6711	0.9624	0.834	**0.4**	500	Zel'dovich

Table 1. Description of the N-body simulations used in the Fisher analysis.

MARKED POWER SPECTRA

EM et al. 2020

The Mark

$$m(\vec{x}, \phi = R, p, \delta_s) = \left[\frac{1+\delta_s}{1+\delta_s+\delta_R(\vec{x})}\right]^p$$

Considered values for the mark parameters

R = [5, 10, 15, 20, 30] Mpc/h

p = [-1, 0.5, 1, 2, 3]

δ_s = [0, 0.25, 0.5, 0.75, 1]

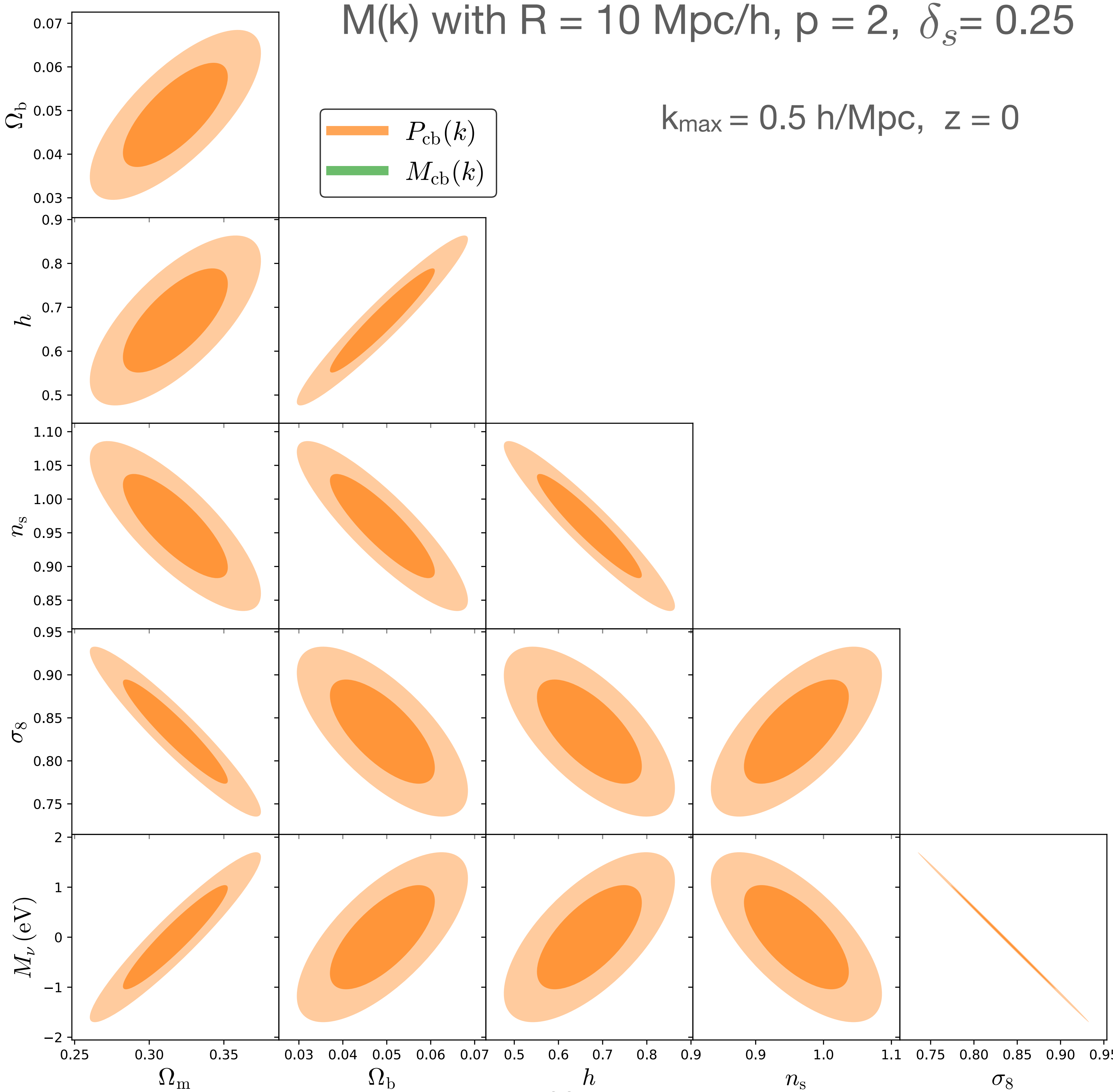
125 marked power spectra compute on the matter fields
cb (cdm) and m (cdm+neutrinos)

Forecast for statistics of the cold dark matter

EM et al. 2020

M(k) with R = 10 Mpc/h, p = 2, δ_s= 0.25

k_{max} = 0.5 h/Mpc, z = 0

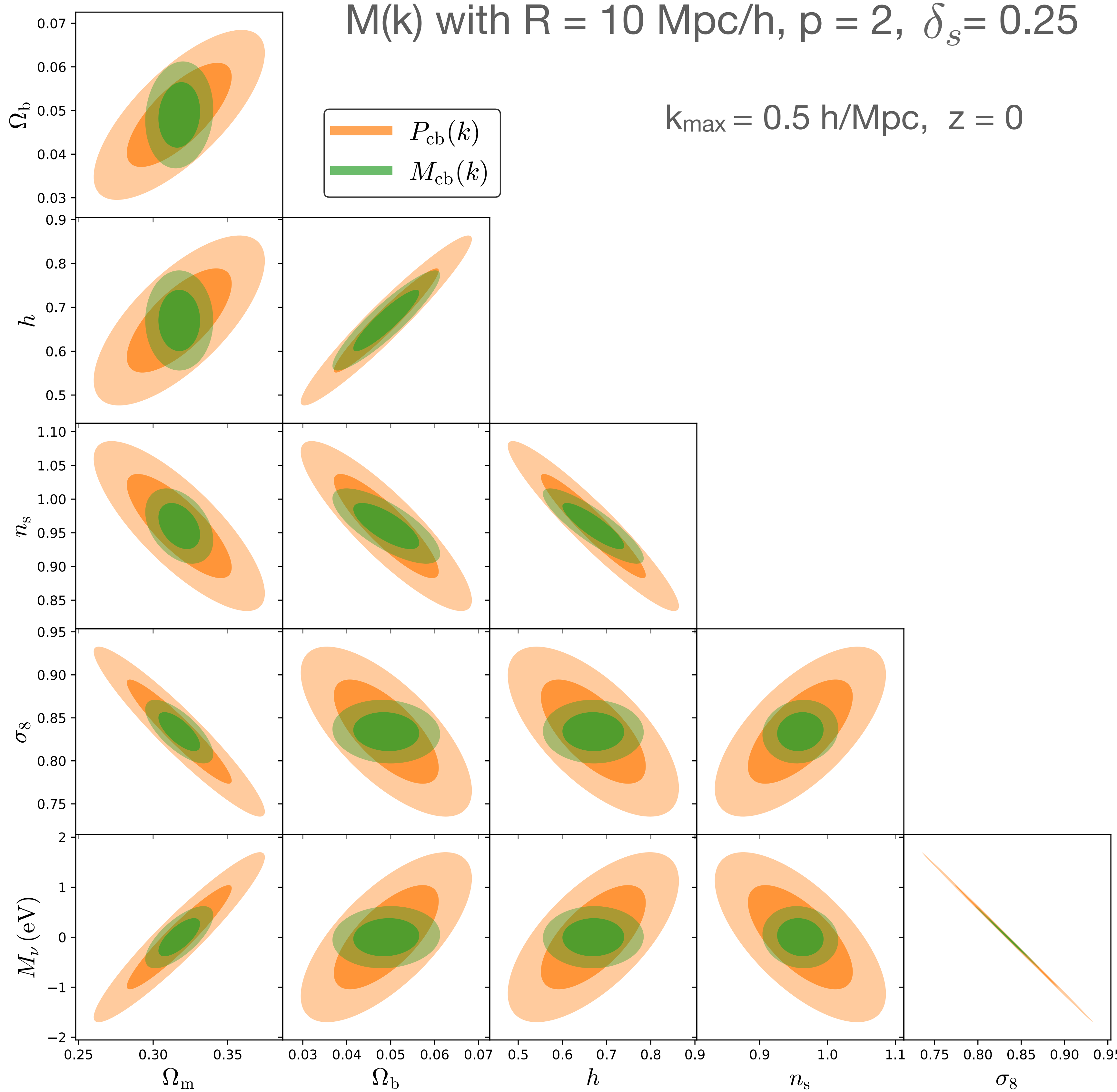


Forecast for statistics of the cold dark matter

EM et al. 2020

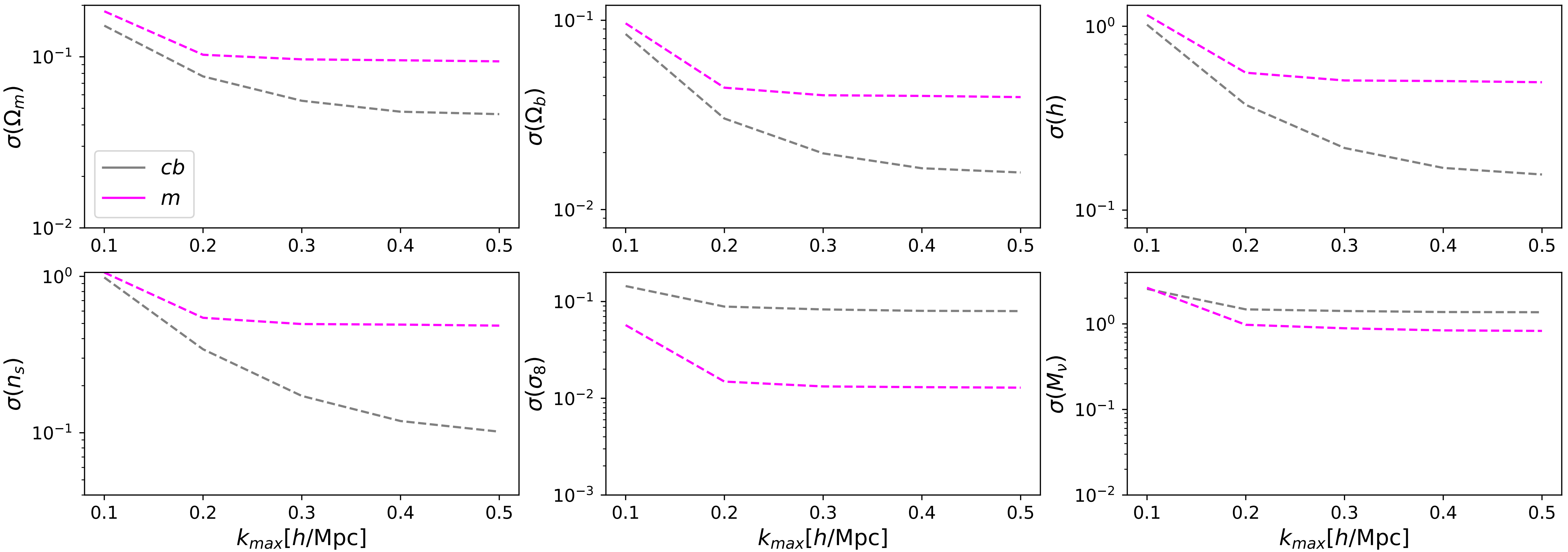
M(k) with R = 10 Mpc/h, p = 2, δ_s= 0.25

k_{max} = 0.5 h/Mpc, z = 0



Marginalized errors

EM et al. 2020



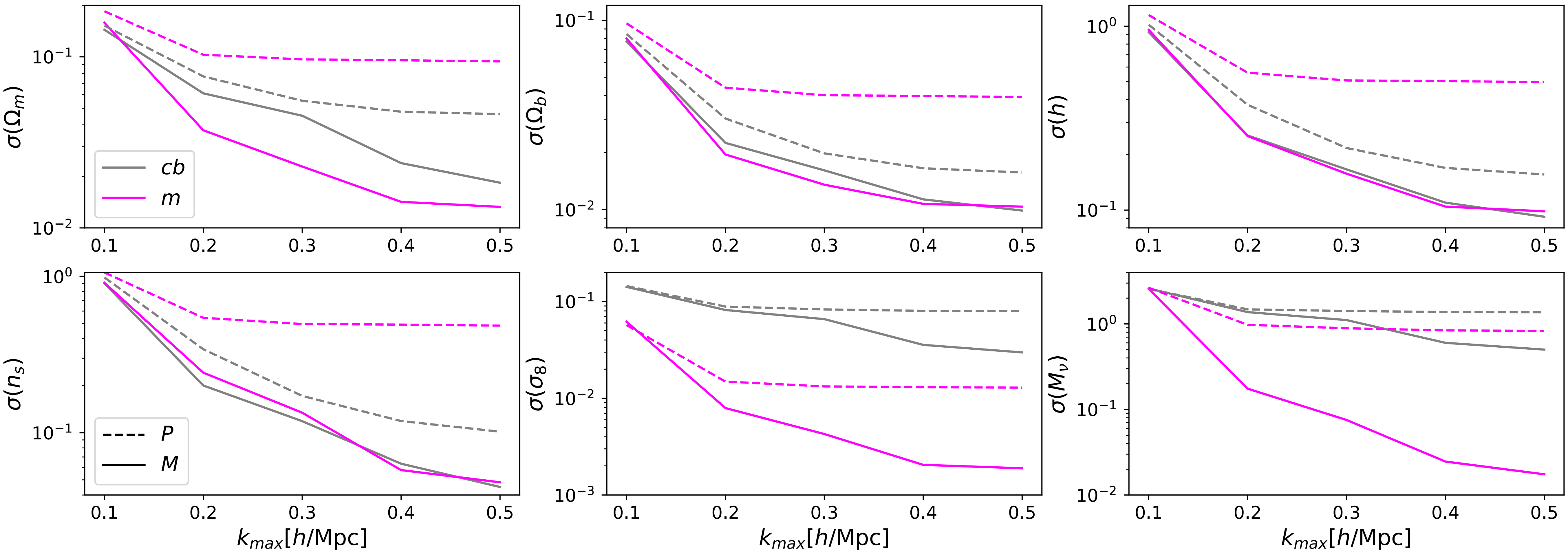
Marginalized errors for k_{max} = 0.5 h/Mpc

Parameter	P_{cb}	M_{cb}	P_{cb} / M_{cb}	P_m	M_m	P_m / M_m
Ω_m	0.046			0.094		
Ω_b	0.016			0.039		
h	0.16			0.50		
n_s	0.10			0.48		
σ_8	0.080			0.013		
M_ν	1.4			0.83		

Marginalized errors

EM et al. 2020

M(k) with R = 10 Mpc/h, p = 2, δ_s= 0.25



Marginalized errors for k_{max} = 0.5 h/Mpc

Parameter	P_{cb}	M_{cb}	P_{cb} / M_{cb}	P_m	M_m	P_m / M_m
Ω_m	0.046	0.018	**2.5**	0.094	0.013	**7.2**
Ω_b	0.016	0.0099	**1.6**	0.039	0.010	**3.9**
h	0.16	0.092	**1.7**	0.50	0.098	**5.1**
n_s	0.10	0.045	**2.2**	0.48	0.048	**10**
σ_8	0.080	0.030	**2.7**	0.013	0.0019	**6.8**
M_ν	1.4	0.50	**2.8**	0.83	0.017	**48**

INFORMATION CONTENT in MARKED POWER SPECTRA of the GALAXY FIELD

Molino galaxy catalogs

Hahn et al. 2021

Built upon the Quijote simulations using Halo Occupation Distribution (HOD) framework from Zheng et al. (2007):

Mean central galaxy occupation

$$\langle N_c \rangle = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{\log M_h - \log M_{\min}}{\sigma_{\log M}}\right)\right]$$

Mean satellite galaxy occupation

$$\langle N_s \rangle = \langle N_c \rangle \left(\frac{M_h - M_0}{M_1}\right)^{\alpha}$$

5 additional parameters to describe the BIAS scheme of GALAXIES

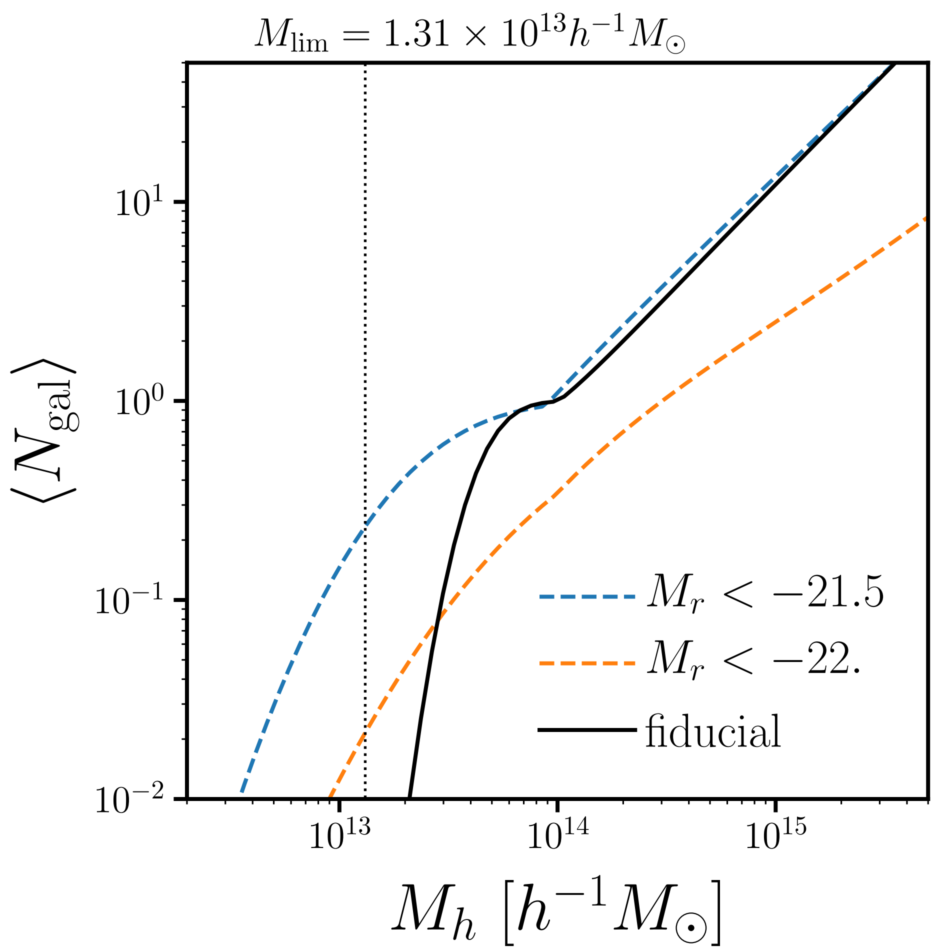
This prescription allow us to compute the **redshift-space** multiples (**monopole** and **quadrupole**) of the **marked power spectrum** of the galaxy field

Halo Occupation Distribution (HOD)

Hahn et al. 2021

Molino mock catalog

Fiducial: $\{\log M_{\rm min}, \sigma_{\log M}, \log M_0, \alpha, \log M_1\} = \{13.65,\ 0.2,\ 14.0,\ 1.1,\ 14.0\}.$



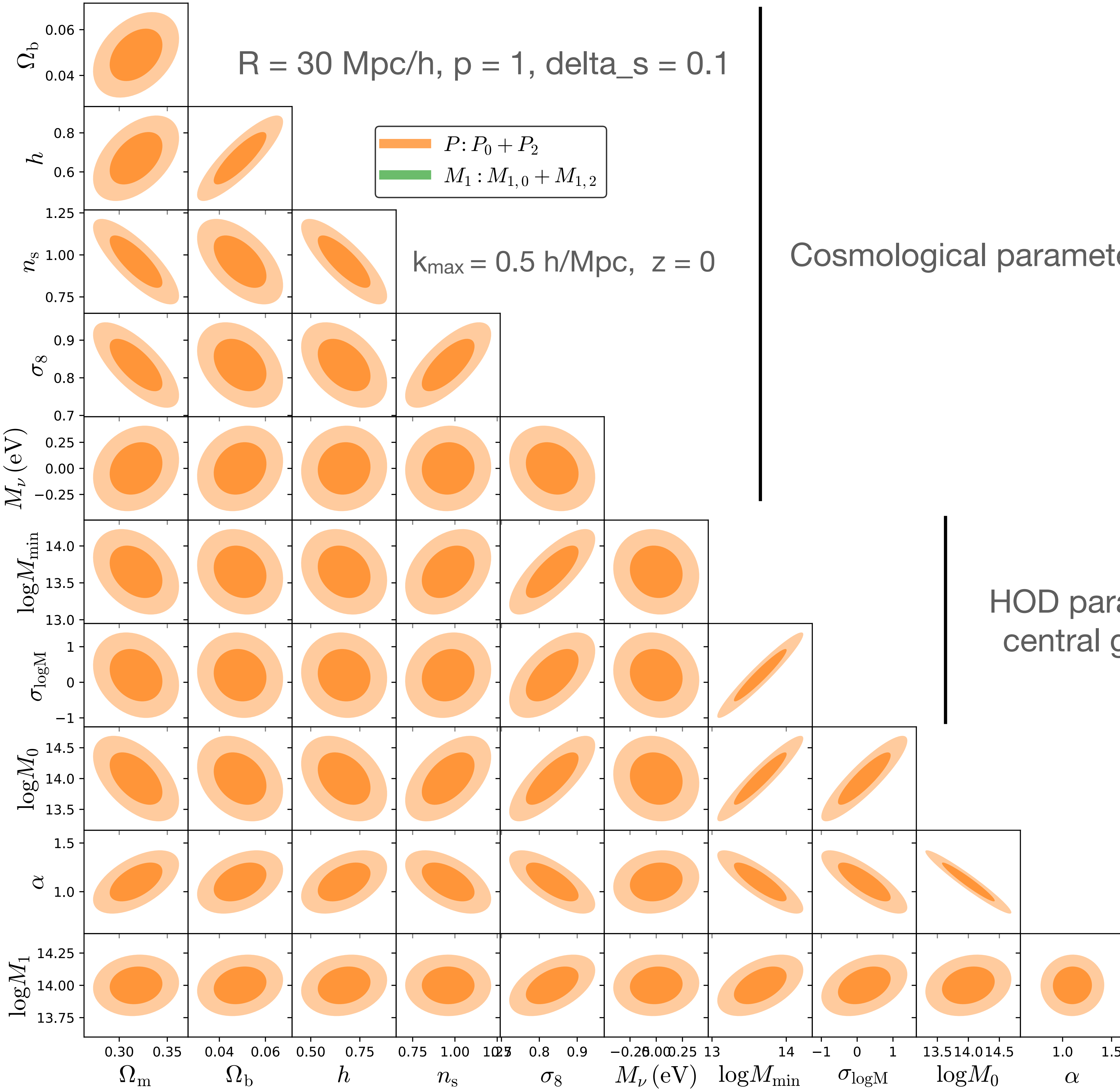
$\bar{n}_g \sim 1.63 \times 10^{-4}\ h^3\ {\rm Mpc}^{-3}$

$b_g \sim 2.55$

Variation to compute derivatives: $\{\Delta \log M_{\rm min}, \Delta\sigma_{\log M}, \Delta \log M_0, \Delta\alpha, \Delta \log M_1\} = \{0.05,\ 0.2,\ 0.2,\ 0.2,\ 0.2\}$

Forecast for the redshift-space Galaxy field

EM et al. (2022)



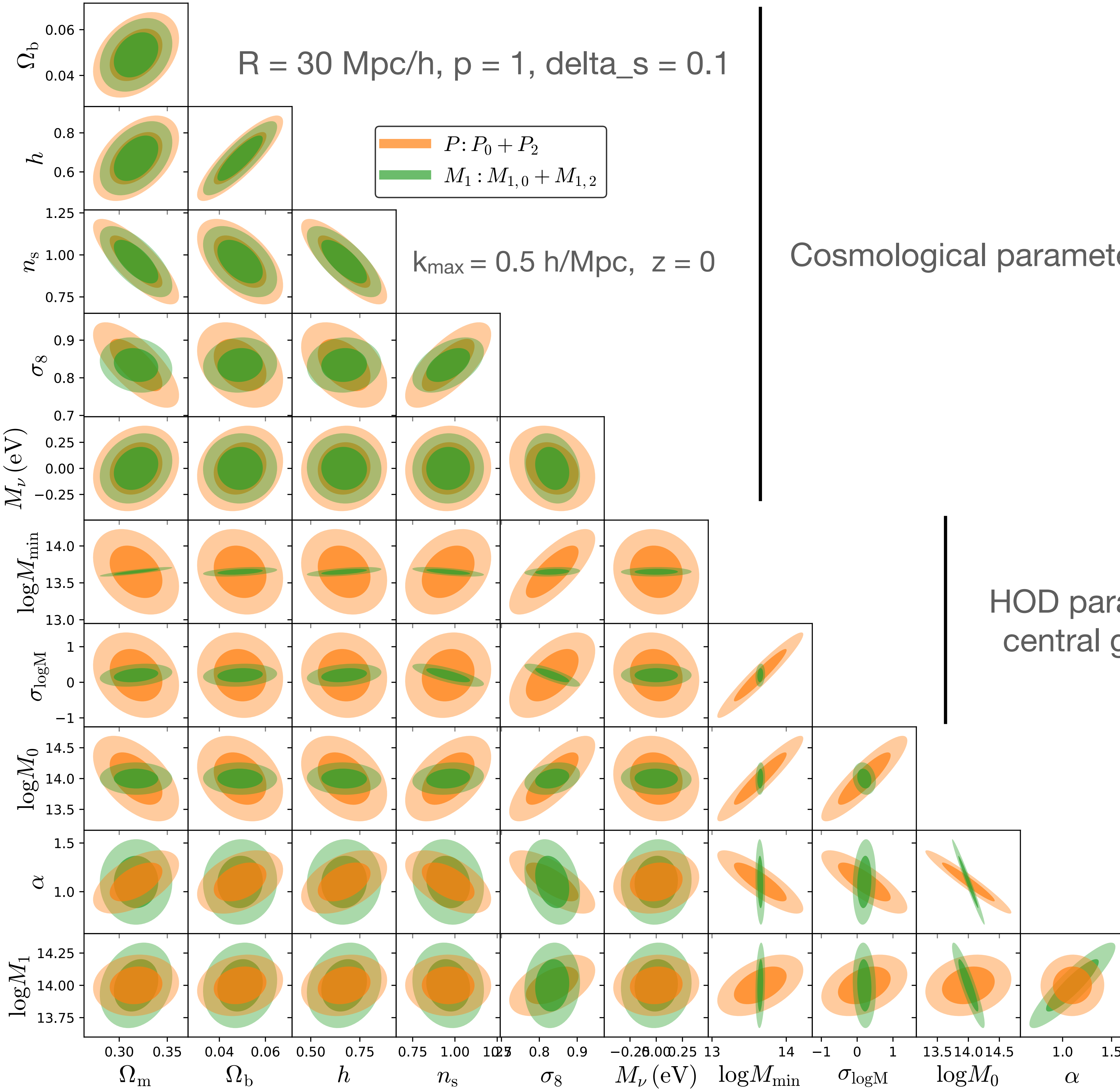
Cosmological parameters

HOD parameters, central galaxies

HOD parameters, satellite galaxies

Forecast for the redshift-space Galaxy field

EM et al. (2022)



Cosmological parameters

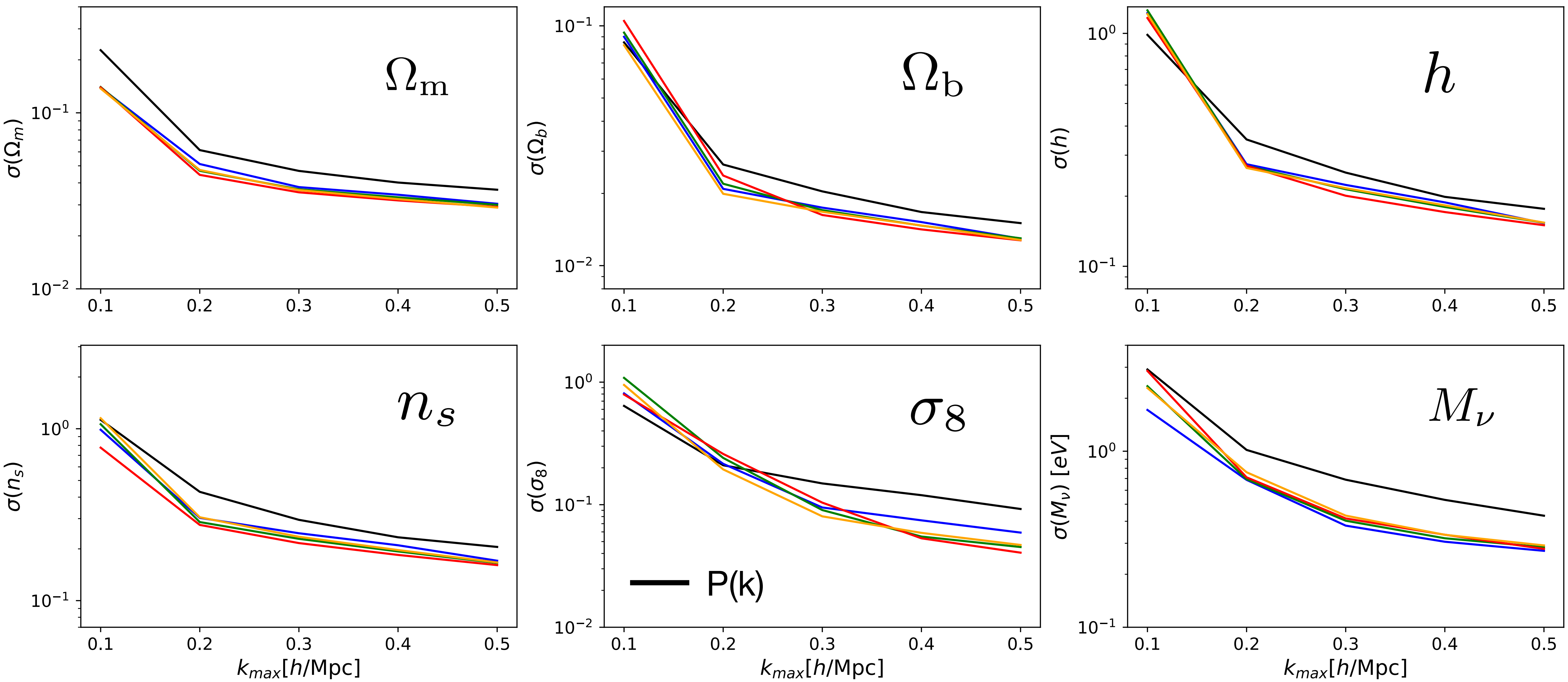
HOD parameters, central galaxies

HOD parameters, satellite galaxies

Marginalized errors - galaxy field

EM et al. (2022)

Constraints using monopole and quadrupole of different statistics up to k = 0.5 h/Mpc



M(k)	$R[h^{-1}\text{Mpc}]$	p	δ_s
M_1	30	1	0.1
M_2	25	1	0.25
M_3	20	1	0.5
M_4	30	1	0.5

Table 2. Values for the mark parameters (R, p, δ_s) of selected marked power spectra M_1, M_2, M_3, and M_4.

Marginalized errors - galaxy field

EM et al. (2022)

The combination of **4 marked power spectra and the standard power spectrum** can largely improve the cosmological constraints coming from the power spectrum alone.

Θ	$P/(P+\Sigma_i M_i)$ $0+2$
Ω_m	2.4
Ω_b	2.5
h	2.6
n_s	3.3
σ_8	6.1
M_ν	3.0

Future surveys (DESI, EUCLID, Roman) will observe a **LARGER number density** of galaxies that will allow them to **trace the inner part of voids** better than the Molino catalogs. Thus, they will be able to exploit the potential the M(k) even further!

WHY DO MARKED POWE SPECTRA CONTAIN INFORMATION BEYOND THE STANDARD POWER SPECTRUM?

MARKED POWER SPECTRA

Philcox, **EM** et al. 2020

The marked density field:

$$\delta_M(\mathbf{x}) \equiv \frac{1}{\bar{m}} m(\mathbf{x}) \left[1 + \delta(\mathbf{x})\right] - 1$$

with mark:

$$m(\mathbf{x}) = \left(\frac{1 + \delta_s}{1 + \delta_s + \delta_R(\mathbf{x})} \right)^p \equiv \left(1 + \frac{\delta_R(\mathbf{x})}{1 + \delta_s} \right)^{-p}$$

MARKED POWER SPECTRA

Philcox, **EM** et al. 2020

The marked density field:

$$\delta_M(\mathbf{x}) \equiv \frac{1}{\bar{m}} m(\mathbf{x}) \left[1 + \delta(\mathbf{x})\right] - 1$$

with mark:

$$m(\mathbf{x}) = \left(\frac{1 + \delta_s}{1 + \delta_s + \delta_R(\mathbf{x})}\right)^p \equiv \left(1 + \frac{\delta_R(\mathbf{x})}{1 + \delta_s}\right)^{-p}$$

Taylor expanding the mark:

$$\delta_M(\mathbf{x}) = \frac{1}{\bar{m}} \left[1 + \delta(\mathbf{x})\right] \left[1 - C_1 \delta_R(\mathbf{x}) + C_2 \delta_R^2(\mathbf{x}) - C_3 \delta_R^3(\mathbf{x})\right] - 1 + \mathcal{O}\left(\delta^4\right).$$

MARKED POWER SPECTRA

Philcox, **EM** et al. 2020

The marked density field:

$$\delta_M(\mathbf{x}) \equiv \frac{1}{\bar{m}} m(\mathbf{x}) \left[1 + \delta(\mathbf{x})\right] - 1$$

with mark:

$$m(\mathbf{x}) = \left(\frac{1 + \delta_s}{1 + \delta_s + \delta_R(\mathbf{x})} \right)^p \equiv \left(1 + \frac{\delta_R(\mathbf{x})}{1 + \delta_s} \right)^{-p}$$

Taylor expanding the mark:

$$\delta_M(\mathbf{x}) = \frac{1}{\bar{m}} \left[1 + \delta(\mathbf{x})\right] \left[1 - C_1 \delta_R(\mathbf{x}) + C_2 \delta_R^2(\mathbf{x}) - C_3 \delta_R^3(\mathbf{x})\right] - 1 + \mathcal{O}\left(\delta^4\right) .$$

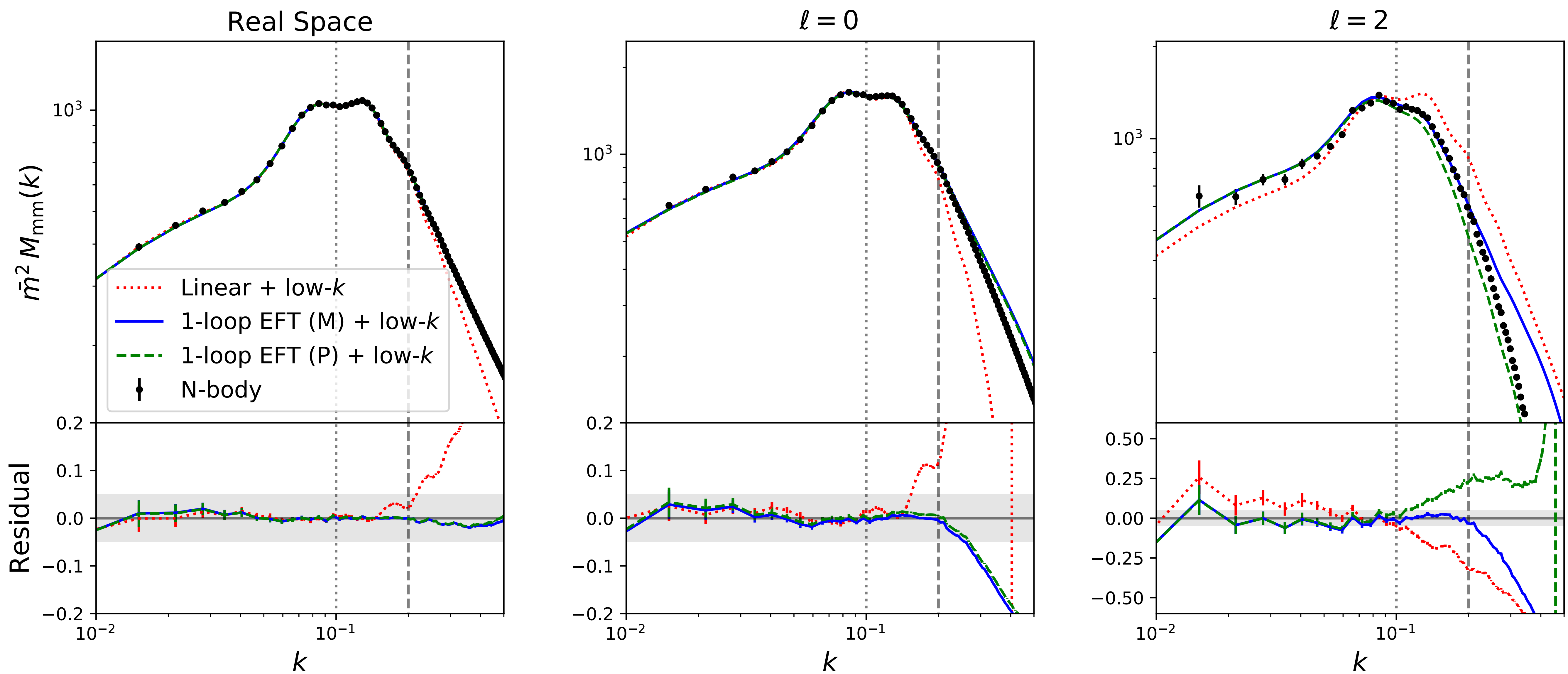
2-point function of the marked field contains higher order statistics of the original field

EFT for MARKED POWER SPECTRA

Philcox, **EM** et al. 2020, Philcox, Aviles, **EM** et al. 2021

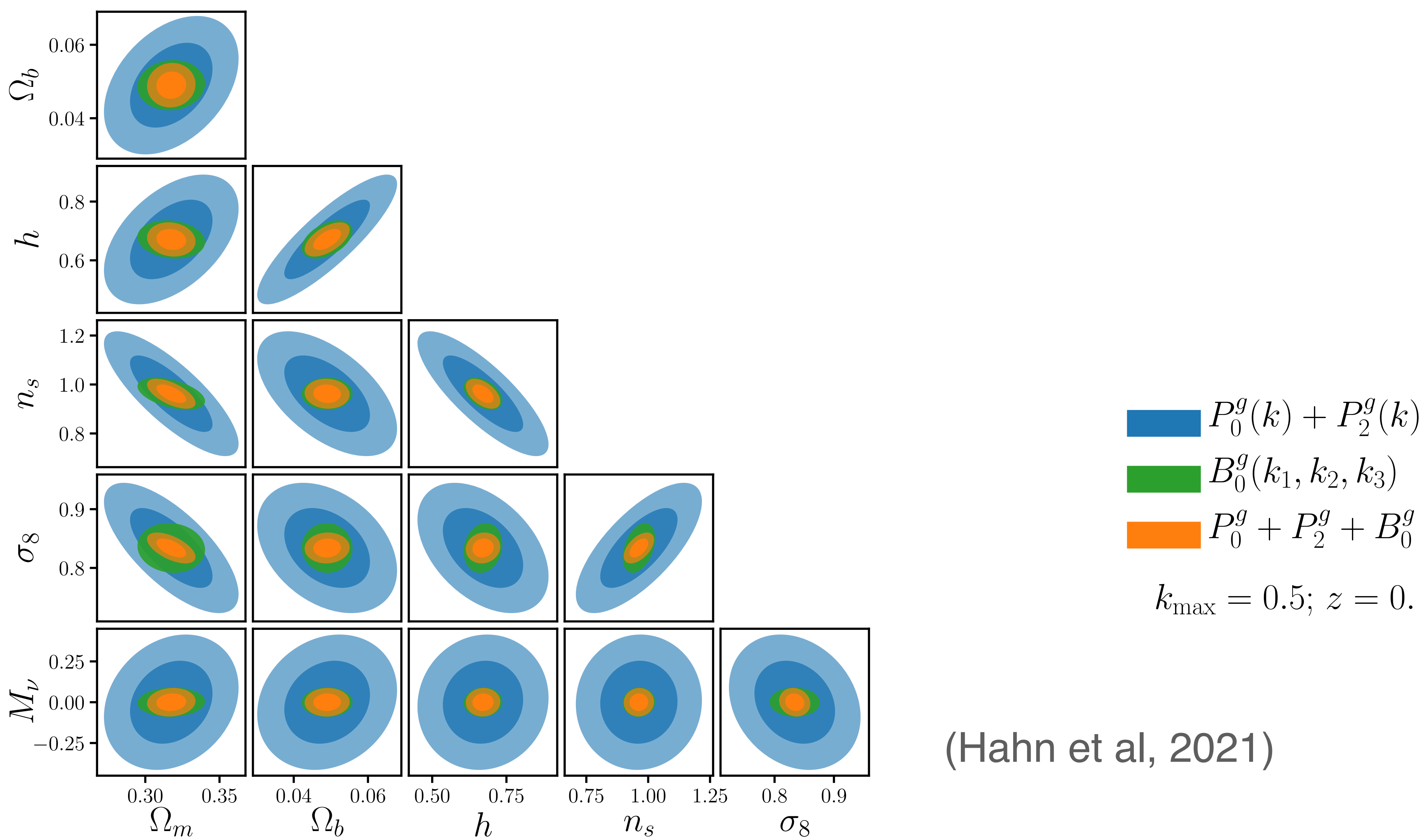
MATTER FIELD

R = 15 Mpc/h, p = 1, delta_s = 0.25



THE INFORMATION CONTENT

- Marked power spectra contain **higher order statistics** of the density field

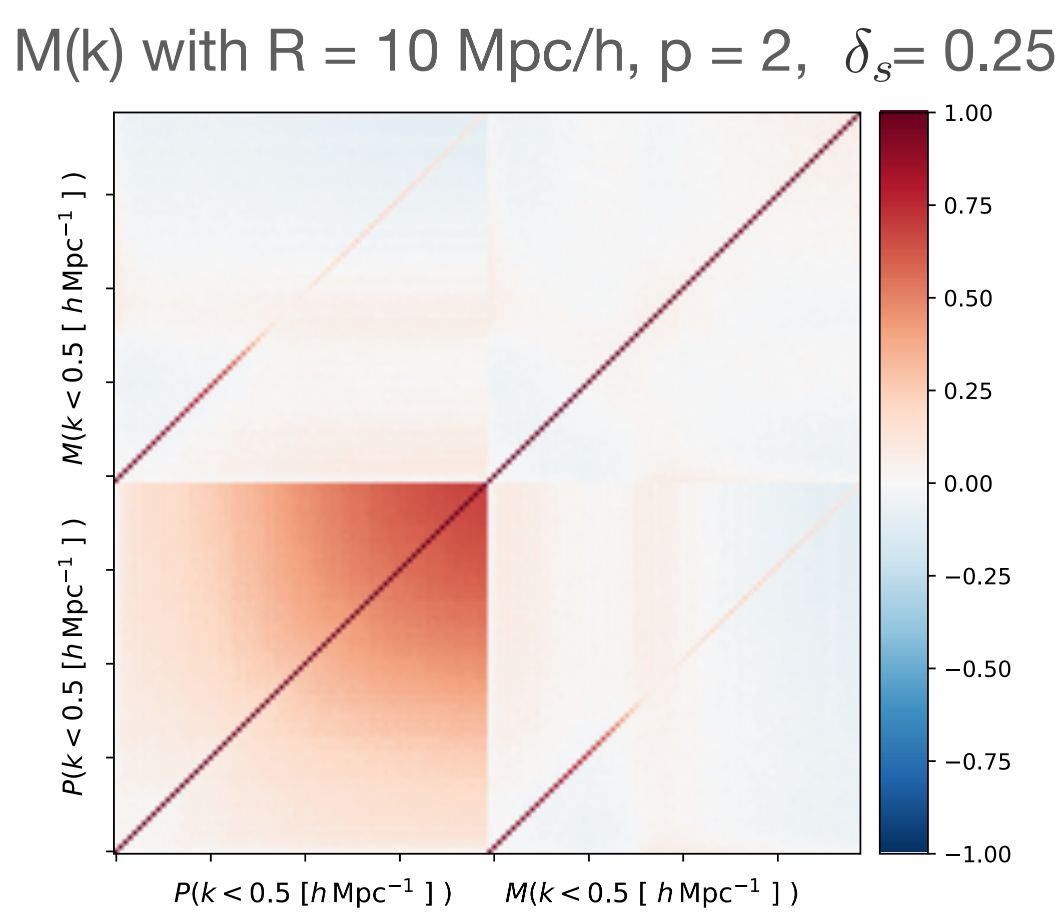


(Hahn et al, 2021)

THE INFORMATION CONTENT

- Marked power spectra contain **higher order statistics** of the density field

- The **covariance** matrix of some marked power spectra M(k) that up-weight low-density regions is almost **diagonal**



Other nonlinear transformations, such as the log-transformation, have shown to make the field more Gaussian (Neyrinck et al, 2009, 2010, 2011**)**

THE INFORMATION CONTENT

- Marked power spectra contain **higher order statistics** of the density field

- The **covariance** matrix of some marked power spectra M(k) that up-weight low-density regions is almost **diagonal**

- Marked power spectra that up-weight low-density regions incorporate **information from voids**

CONCLUSIONS

- Results from Fisher analyses: **marked power spectra** that **up-weight low density regions** improve parameter constraints beyond the standard power spectrum.
- **6x** tighter constraints for sigma8 and **2-3x** for the other cosmological parameters when considering combinations of marked and standard power spectra of the galaxy field.
- Upcoming surveys (DESI, EUCLID, Roman) will probe **larger volumes** and **higher galaxy number density**, that will allow them to better explore low-density regions and improve the performance of marked power spectra.
- Next step: cosmological analysis with marked power spectra in available surveys. We are building a **simulation-based inference** framework that will allow us to forward modeling survey systematics and geometry.