Accurate mass and velocity functions of dark matter haloes

Johan Comparat,[1,2,3]⋆ Francisco Prada,[4] Gustavo Yepes[2] and Anatoly Klypin[5]

[1]*Instituto de Física Teórica UAM/CSIC, E-28049 Madrid, Spain*
[2]*Departamento de Física Teórica, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*
[3]*Max-Planck-Institut für extraterrestrische Physik (MPE), Giessenbachstrasse 1, D-85748 Garching bei München, Germany*
[4]*Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, E-18080 Granada, Spain*
[5]*Astronomy Department, New Mexico State University, Las Cruces, NM, USA*

ABSTRACT

N-body cosmological simulations are an essential tool to understand the observed distribution of galaxies. We use the MultiDark simulation suite, run with the Planck cosmological parameters, to revisit the mass and velocity functions. At redshift $z = 0$, the simulations cover four orders of magnitude in halo mass from $\sim 10^{11} M_{\odot}$ with 8783 874 distinct haloes and 532 533 subhaloes. The total volume used is ~ 515 Gpc^3, more than eight times larger than in previous studies. We measure and model the halo mass function, its covariance matrix w.r.t halo mass and the large-scale halo bias. With the formalism of the excursion-set mass function, we explicit the tight interconnection between the covariance matrix, bias and halo mass function. We obtain a very accurate (<2 per cent level) model of the distinct halo mass function. We also model the subhalo mass function and its relation to the distinct halo mass function. The set of models obtained provides a complete and precise framework for the description of haloes in the concordance Planck cosmology. Finally, we provide precise analytical fits of the V_{max} maximum velocity function up to redshift $z < 2.3$ to push for the development of halo occupation distribution using V_{max}. The data and the analysis code are made publicly available in the *Skies and Universes* data base.

Key words: dark matter – large-scale structure of Universe.

1 INTRODUCTION

N-body cosmological simulations are essential tools to understand the observed distribution of galaxies. In the last decades, development of numerical codes (Teyssier 2002; Springel 2005, 2010; Klypin, Trujillo-Gomez & Primack 2011; Habib et al. 2016) and the access to powerful supercomputers enabled the computation of high-resolution cosmological simulations over large volumes e.g. MultiDark (MD hereafter; Prada et al. 2012) and DarkSkies (DS hereafter; Skillman et al. 2014). Both simulations were run in the paradigm of the flat Λ cold dark matter cosmology (ΛCDM; Planck Collaboration XVI 2014). From MD emerged the most precise description to date of the dark matter (DM) halo (Klypin et al. 2016). While finding and describing the haloes formed by the DM is now well understood (Behroozi, Wechsler & Wu 2013; Knebe et al. 2013; Avila et al. 2014), connecting galaxies to haloes is a proven complicated subject. There are three main streams of galaxy assignment in simulations, we order them by decreasing computational needs and accuracy: (i) hydrodynamical simulations (HYDRO; Cen & Ostriker 1993; Springel & Hernquist 2003), (ii) semi-analytical models

of galaxy formation (SAMS; Cole et al. 2000; Baugh 2006), (iii) halo occupation distribution or subhalo abundance matching (HOD, SHAM; Cooray & Sheth 2002; Conroy, Wechsler & Kravtsov 2006, respectively). The existing methods will hopefully converge in the coming years (Knebe et al. 2015; Elahi et al. 2016; Guo et al. 2016).

The current and future cosmological galaxy and quasar surveys, e.g. BOSS, eBOSS, DES, DESI, 4MOST, *Euclid*, will cover gigantic volumes up to redshift 3.5 (The Dark Energy Survey Collaboration 2005; Laureijs et al. 2011; Dawson et al. 2013, 2016; DESI Collaboration et al. 2016). These volumes are too large to be entirely simulated with hydrodynamics. There is thus a need to improve the predictive power of the SAMS and HOD to the level of the expected two-point function measurements, i.e. around the per cent level. This challenge needs to be handled from both the hydrodynamical simulation point of view (Sawala et al. 2015; Chaves-Montero et al. 2016) and from the DM-only simulation perspective (Carretero et al. 2015; Favole et al. 2016; Rodríguez-Torres et al. 2016) to eventually join in an optimal semi-analytical model (Knebe et al. 2015). Lastly, Castro, Marra & Quartin (2016) argued that with such surveys, one would constrain directly the parameters of the mass function to the level that it is estimated in *N*-body simulations, enhancing again the need of a precise model for the halo mass function (HMF).

⋆ E-mail: comparat@mpe.mpg.de

From the DM-only simulation perspective, the most fundamental statistic is the HMF. Observational probes, such as weak lensing, galaxy clustering or galaxy clusters, also rely on the knowledge of the HMF. The mass function denotes, at a given redshift, the fraction of mass contained in collapsed haloes with a mass in the interval M and $M + \mathrm{d}M$. It was studied theoretically and numerically in various simulations and different cosmologies (Press & Schechter 1974; Sheth & Tormen 1999; Jenkins et al. 2001; Sheth, Mo & Tormen 2001; Sheth & Tormen 2002; Springel et al. 2005; Warren et al. 2006; Tinker et al. 2008; Bhattacharya et al. 2011; Angulo et al. 2012; Watson et al. 2013; Despali et al. 2016).

The theoretical formalism to describe the number density of haloes was initiated by Press & Schechter (1974). Its latest formulation by Sheth et al. (2001) and Sheth & Tormen (1999) includes the ellipsoidal collapse instead of spherical collapse. Heuristically, it corresponds to a diffusion across a 'moving' or across a mass-dependent boundary. The excursion-set formalism of the mass function constitutes today a good description of what is measured in N-body simulations. More precise predictions are actively being sought and eventually we might converge towards an ultimate universal mass function. The variety of existing and tested functional forms of the mass function are discussed and compared in Murray, Power & Robotham (2013). The description of the errors on the HMF is slightly less discussed subject. Nevertheless, Hu & Kravtsov (2003) and Bhattacharya et al. (2011) provided a solid background, used in this study, to model errors on the HMF and the large-scale halo bias.

Numerically, the HMF was extensively studied with a cosmology-independent (universal) model. The most recent measurements on N-body simulations enabled models to predict any HMF to about 10 per cent accuracy; see Despali et al. (2016). It is to date the latest HMF measurements in the Planck cosmology. We feel though, the lack of a per cent-level-accurate model for the HMF in the Planck cosmology.

The recent measurements of the cosmic microwave background indicate a significantly higher matter content than suggested by previous observations (*WMAP*; Komatsu et al. 2011). The matter content of the Universe is a parameter that strongly influences the HMF. We think it is thus necessary to revisit the parametrization of the mass function and understand to what accuracy the mass function is known in our best cosmological model. Previous works could not assess thoroughly the uncertainties on the measurement of the mass function due to the limited amount of N-body realizations available. With the MD and DS simulations, extracting covariance matrices becomes possible.

In this paper, we explore and model the HMF and its covariance matrix. We describe the model in Section 2. In Section 3, we describe the simulations used and we estimate the HMF, its covariance and the large-scale halo bias. The HMF results are presented in Section 4. Finally, in the Appendix we parametrize the redshift evolution of the distinct and satellite halo velocity function.

Data base

All the data and the results are available through the *Skies and Universes* data base.[1] The code is made public via GitHub.[2]

2 MODEL

2.1 Halo mass function

The formalism to describe the number density of haloes was initiated by Press & Schechter (1974). They assumed that the fraction of mass in haloes of mass greater than M at a time t, $F(>M, t)$, was equal to twice the probability, $\mathcal{P}$, for the smoothed density field, δ_s, to overcome the critical threshold for spherical collapse, δ_c i.e.

$$F(>M, t) = 2\mathcal{P}(\delta_s(t) > \delta_c(t)). \tag{1}$$

Assuming that δ_s is a Gaussian random field, they related the number density of haloes to F

$$n(M, t)\mathrm{d}M = \frac{\bar{\rho}}{M} \frac{\partial F(>M, t)}{\partial M}\mathrm{d}M. \tag{2}$$

The mass function depends on redshift and on halo mass. Rather than mass, it is physically more relevant to use the root mean square (rms) fluctuations of the linear density field smoothed with a filter encompassing this mass

$$\sigma^2(M, t) = 4\pi^2 \int_0^\infty P(k, t)W^2(k, M)k^2\mathrm{d}k, \tag{3}$$

where $P(k)$ is the linear power spectrum and W a top-hat filter.

Assuming that the initial Gaussian random density fluctuation field evolves and crosses via a random walk the spherical collapse barrier, these equations determine the number of regions in the simulation that underwent collapse at a given time

$$n(\sigma, t)\mathrm{d}M = f_{\mathrm{PS}}(\sigma)\frac{\bar{\rho}}{M^2}\frac{\mathrm{d}\ln\sigma}{\mathrm{d}\ln M}\mathrm{d}M, \tag{4}$$

where the function f, called the multiplicity function has the following expression:

$$f_{\mathrm{PS}}(\sigma) = \sqrt{\frac{2}{\pi}}\frac{\delta_c}{\sigma}\exp\left[-\frac{\delta_c^2}{2\sigma^2}\right]. \tag{5}$$

'PS' stands for 'Press–Schechter'. In other words, it is the fraction of mass associated with haloes in a unit range of $\mathrm{d}\ln\sigma$. Because the threshold δ_c increases with time, smaller haloes are formed first and then the larger ones (hierarchical clustering).

This model was revised using excursion-set theory by Bond et al. (1991). They argued that σ diffuses across the spherical collapse boundary or barrier, instead of crossing it via a random walk. This leads to a new multiplicity function

$$f_{\mathrm{EPS}}(\sigma) = f_{\mathrm{PS}}(\sigma)/(2\sqrt{\sigma}), \tag{6}$$

where 'EPS' stand for extended Press–Schechter.

Sheth & Tormen (1999) and Sheth et al. (2001) later explored the ellipsoidal collapse to replace the assumption of spherical collapse. Heuristically, it corresponds to a diffusion across a 'moving' barrier (or across a σ-dependent boundary). They found the following multiplicity function f_{ST}:

$$f_{\mathrm{ST}}(\sigma, A, a, p) = A\sqrt{\frac{2}{\pi}}\left[1 + \left(\frac{\sigma^2}{a\delta_c^2}\right)^p\right]\left(\frac{\sqrt{a}\delta_c}{\sigma}\right)\exp\left[-\frac{a}{2}\frac{\delta_c^2}{\sigma^2}\right], \tag{7}$$

where 'ST' stands for 'Sheth and Tormen'. It constitutes a further improvement compared to f_{EPS}.

The latter multiplicity function describes well the ΛCDM distinct HMF with the parameters (A, a, p)=(0.3222, 0.707, 0.3). These parameters were measured again by Despali et al. (2016) in the latest Planck-cosmology paradigm. They found (A, a, p)=(0.333,

0.794, 0.247). It remains a statistical scatter of the simulated data around this model of the order of 5–7 per cent at the high-mass end. More precise predictions are actively being sought (e.g. Pace, Batista & Del Popolo 2014, Rei; Del Popolo, Pace & Le Delliou 2017). Eventually, we will converge towards an ultimate physical model for the HMF.

Aside from the physical model of the mass function, exist a variety of functional forms created to best fit the mass function as measured in *N*-body simulations; see Murray et al. (2013) who compare and catalogue them. Among others, Bhattacharya et al. (2011) proposed a generalized form of the Sheth & Tormen (1999) function that we use here. Note that this generalization is not theoretically motivated by the excursion-set formalism.

The multiplicity function from Bhattacharya et al. (2011, equation 12–18) is

$$f_{Ba}(\sigma, z, \bar{A}, \bar{a}, \bar{p}, \bar{q}) = \bar{A}(z)\sqrt{\frac{2}{\pi}}\left[1 + \left(\frac{\sigma^2}{\bar{a}(z)\delta_c^2}\right)^{\bar{p}(z)}\right]$$
$$\times \left(\frac{\sqrt{\bar{a}(z)}\delta_c}{\sigma}\right)^{\bar{q}(z)} \exp\left[-\frac{\bar{a}(z)}{2}\frac{\delta_c^2}{\sigma^2}\right]. \quad (8)$$

In the case, $\bar{q} = 1$, the parameters of equation (8) are the same as that of equation (7) i.e. $\bar{A} = A$, $\bar{a} = a$, $\bar{p} = p$. The addition of the $\bar{q}$ parameter is strictly speaking not physically motivated, but provides a better fit to the data, see further down in the paper.

We then use the formalism of Hu & Kravtsov (2003) and Bhattacharya et al. (2011) to account for the large-scale halo bias and the mass function's covariance.

2.2 Large-scale halo bias

The large-scale halo bias function is written in terms of the conditional, the unconditional mass function and a Taylor expansion (Sheth & Tormen 1999; Bhattacharya et al. 2011). This allows its formulation with the same parameters as the mass function

$$b(\sigma, z, \bar{a}, \bar{p}, \bar{q}) = 1 + \frac{\bar{a}(z)(\delta_c^2/\sigma^2) - \bar{q}(z)}{\delta_c}$$
$$+ \frac{2\bar{p}(z)/\delta_c}{1 + (\bar{a}(z)(\delta_c^2/\sigma^2))^{\bar{p}(z)}}. \quad (9)$$

2.3 Covariance matrix

To model the covariance, we slightly adapt the notations from Hu & Kravtsov (2003) and Bhattacharya et al. (2011) as follows.

Let $\bar{\rho}$ be the average density of haloes. We assume the overdensity of haloes at a position $(z, \boldsymbol{x})$, denoted $\delta_{halo}(\sigma, z, \boldsymbol{x})$, to be related to the total mass density field $\delta_{DM}(\boldsymbol{x})$ by a biasing function, $b(\sigma, z)$. Note that, on large scales, this function is the bias mentioned in the previous section,

$$\delta_{halo}(\sigma, z, \boldsymbol{x}) = b(\sigma, z)\delta_{DM}(\boldsymbol{x}). \quad (10)$$

Then, within a window W_a, the average number density of haloes, n_a is given by

$$n_a(\sigma, z) = \bar{\rho}\int d\boldsymbol{x}\, W_a(\boldsymbol{x})\, b(\sigma_a, z_a)\delta_{DM}(\boldsymbol{x}). \quad (11)$$

The covariance between the number densities $n_a(\sigma_a, z_a)$ and $n_b(\sigma_b, z_b)$ within the windows W_a and W_b has two components: the shot-noise variance, proportional to the inverse of the density

times the volume $\sim(\bar{n}V)^{-1}$, and the sample variance:

$$\frac{\langle n_a n_b\rangle - \bar{n}_a\bar{n}_b}{\bar{n}_a\bar{n}_b} = b(\sigma_a, z_a)D(z_a)b(\sigma_b, z_b)D(z_b)$$
$$\times \int \frac{3d^3k}{(2\pi)^3}W_a(k\,R_{box,a})W_b^*(k\,R_{box,b})P(k), \quad (12)$$

where D is the growth factor, V the volume of the box, $R_{box} = (3V/4\pi)^{1/3}$ and $P(k)$ the DM power spectrum. We use top-hat window functions. The growth factor and the integral depend only on the cosmological model (and redshift) but not on the mass function model. The model of the bias function is directly related to the HMF model. Therefore once the mass function parameters are determined, the covariance matrix should be predictable. Also, we note how the large-scale structure makes number counts of haloes in distinct volumes covary. Our model of the covariance matrix is

$$C_{model}(\sigma_a, \sigma_b) = \frac{Q}{\sqrt{\bar{n}_a\bar{n}_b}(V_a + V_b)} + \left(\frac{\langle n_a n_b\rangle - \bar{n}_a\bar{n}_b}{\bar{n}_a\bar{n}_b}\right), \quad (13)$$

where the Q factor depends on the simulation size. This factor allows us to rescale small-subboxes estimates of the covariance to much larger computational simulations. We find the factor by observing how covariance scales with the box size. In the next section, we find that $Q = -3.62 + 4.89\log_{10}[L_{box}(h^{-1}\,Mpc)]$ accounts well for all of the estimated covariance matrices, see Fig. 7.

3 SIMULATIONS

The MD simulation suite[3] is currently the largest public data base of high-resolution large volume boxes with $\sim4000^3$ particles. The simulations were run in the Planck cosmology (Prada et al. 2012; Klypin et al. 2016) in a flat ΛCDM model with the $\Omega_m = 0.307$, $\Omega_\Lambda = 0.693$, $\Omega_b = 0.048$, $n_s = 0.96$, $h = 0.6777$ and $\sigma_8 = 0.8228$ (Planck Collaboration XVI 2014). They provide haloes plus subhaloes for all written outputs and for some boxes merger trees are also available. We found three other relevant simulation sets to be compared with our study. Despali et al. (2016) is the current state-of-the-art HMF in Planck cosmology. They ran a suite of 1024^3 particle simulations with different volumes and analysed the mass function up to redshift 1.25. The DS simulations discussed in Skillman et al. (2014), also run in Planck cosmology, used up to $10\,240^3$ particles and cover much larger volume, though the current data release only provides data at redshift 0. The exact cosmological parameters differ a little from the ones used in MD and Despali et al. (2016). Ishiyama et al. (2015) provide a new suite of simulation in Planck cosmology, the largest simulation (of interest for this analysis) is not yet publicly available, so we did not include their data in the analysis. Other simulations covering large volumes with large amount of particles exist Angulo et al. (e.g. 2012) and Heitmann et al. (e.g. 2015), but they were run in a different cosmology set-up and are not yet publicly available. For completeness, also we mention the P-Millennium $\sim4000^3$ simulation although it is not publicly documented and released yet. In this study, we therefore use only the MD simulations and the redshift 0 data produced by the DS simulation. These data sets constitute a non-negligible leap forward, for both resolution and volume, compared to the data used in Despali et al. (2016). Table 1 summarizes and compares the main parameters of each simulation: length of the boxes, number of particles, force resolution, particle mass and number of snapshots. We

[3] https://www.cosmosim.org

Table 1. Basic parameters of the simulations. L_{box} is the side length of the simulation cube. N_p is the number of particles in the simulation. ϵ is the force resolution at redshift $z = 0$. M_p is the mass of a particle. Ns is the number of snapshots available. The σ_8 column gives the input value and its measured deviation at redshift $z = 0$. The column 'cosmo' refers to the cosmology set-up used to run the simulation: (a) refers to Planck Collaboration XVI (2014) and (b) to Komatsu et al. (2011, *WMAP*). The column 'ref' gives the reference paper for each simulation: (1) stands for Klypin et al. (2016) $h=0.6777$, $\Omega_m = 0.307$, (2) for Skillman et al. (2014) $h=0.6846$, $\Omega_m = 0.299$, (3) for Despali et al. (2016) $h=0.677$, $\Omega_m = 0.307$, (4) for Heitmann et al. (2015) $h=0.71$, $\Omega_m = 0.27$, (5) for Angulo et al. (2012) $h=0.73$, $\Omega_m = 0.25$. (6) for Springel (2005) $h=0.73$, $\Omega_m = 0.25$. (7) for Ishiyama et al. (2015) $h=0.68$, $\Omega_m = 0.31$. A dash, '-', means information is the same as in the cell above. An empty space means the information is not available. The column nickname gives the naming convention used throughout the paper, figures and captions.

Box name	Set-up parameters				Ns	σ_8 input, measured	Cosmo	Ref	Nickname
	L_{box} Mpc	$N_p^{1/3}$	ϵ kpc	M_p $M_\odot$					
SMD	590.2	3840	2.2	1.4×10^8	88	0.8228, -2.8 per cent	(a)	(1)	M04
MDPL	1475.5	3840	7.3	2.2×10^9	128	$-$, $+0.2$ per cent	$-$	$-$	M10
BigMD	3688.9	3840	14.7	3.5×10^{10}	80	$-$, $+0.5$ per cent	$-$	$-$	M25
BigMDNW	3688.9	3840	14.7	3.5×10^{10}	1	$-$, $+0.5$ per cent	$-$	$-$	M25n
HMD	5902.3	4096	36.8	1.4×10^{11}	128	$-$, $+0.4$ per cent	$-$	$-$	M40
HMDNW	5902.3	4096	36.8	1.4×10^{11}	17	$-$, $+0.4$ per cent	$-$	$-$	M40n
DS	11 627.9	10 240	53.4	5.6×10^{10}	16	0.8355, $+0.0$ per cent	(a)	(2)	D80
$-$,	2325.5	4096	26.7	7.1×10^9	$-$	$-$	$-$	$-$	DS
$-$,	1162.7	4096	13.3	8.8×10^8	$-$	$-$	$-$	$-$	$-$
$-$,	290.7	2048	6.7	1.1×10^8	$-$	$-$	$-$	$-$	$-$
$-$,	145.3	$-$	3.3	1.3×10^7	$-$	$-$	$-$	$-$	$-$
Ada	92.3	1024	2.2	2.8×10^7	15	0.829,	(a)	(3)	De
Bice	184.6	$-$	4.4	2.2×10^8	15	$-$,	$-$	$-$	$-$
Cloe	369.2	$-$	8.8	1.8×10^9	15	$-$,	$-$	$-$	$-$
Dora	738.5	$-$	17.7	1.4×10^{10}	15	$-$,	$-$	$-$	$-$
Emma	1477.1	$-$	35.4	1.1×10^{11}	15	$-$,	$-$	$-$	$-$
Flora	2954.2	$-$	70.9	9.3×10^{11}	15	$-$,	$-$	$-$	$-$
ν^2GC-L	1647.0	8192		3.2×10^8		0.83	(a)	(7)	ν^2GC
ν^2GC-M	823.5	4096		3.2×10^8	4	$-$	(a)	(7)	$-$
ν^2GC-S	411.7	2048		3.2×10^8	4	$-$	(a)	(7)	$-$
ν^2GC-H1	205.8	2048		4.0×10^7	4	$-$	(a)	(7)	$-$
ν^2GC-H3	205.8	4096		5.0×10^6	2	$-$	(a)	(7)	$-$
ν^2GC-H2	102.9	2048		5.0×10^6	4	$-$	(a)	(7)	$-$
P-Millennium	800.0			1.5×10^8	271		(a)	In prep.	P-Mi
OuterRim	4225.3	10 240	7.0	2.6×10^9	34	0.84,	(b)	(4)	OR
QContinuum	1830.9	8192	2.8	2.1×10^8	$-$	$-$	$-$	$-$	QC
Millennium XXL	4109.6	6720	13.7	1.1×10^{10}		0.9	other	(5)	Mi-XXL
Millennium	684.9	2160		1.1×10^9		$-$	$-$	(6)	Mi

Table 2. More parameters for the MD simulation data used in this paper. The number of snapshots used in the analysis is the one that has a distinction between central and satellite haloes, which is a subsample of the complete simulations.

Box	Number of snapshots with parent IDs		
	all	$z < 3.5$	$z < 2.5$
M04	9	9	8
M10	11	11	10
M25	10	10	9
M25n	1	1	1
M40	128	67	56
M40n	17	15	13

note the latest advances in software enabling $20\,000^3$ particle simulations to converge in reasonable computing time (Potter, Stadel & Teyssier 2016).

We use a set of snapshots from each simulation to sparsely and regularly sample the redshift range $0 < z < 2.5$, i.e. to cover the extent of galaxy surveys. Table 2 provides the number of snapshots used per simulation in our analysis.

The rms amplitude of linear mass fluctuations in spheres of 8 h^{-1} Mpc comoving radius at redshift 0, denoted σ_8, holds a particular role when characterizing the abundance of haloes. To have a more accurate estimate of the actual σ_8 in the simulation, we compare the DM power spectrum at redshift 0 measured in each simulation with the predicted linear power spectrum in the same cosmology. The mean of the square-root of this ratio evaluated on scales where the linear regime dominates gives the relative variation of the value of σ_8. We find variation smaller than ~2 per cent; see Table 1. In the following, we compute the mass–$\sigma(M)$ relation using the measured value of σ_8 in each simulation. To compute these relations, we use the package HMFCALC[4] (Murray et al. 2013).

To visualize the challenges of bridging the gap between *N*-body simulations and galaxy survey, we designed Fig. 1. In this figure, we compare existing simulations with observed galaxy surveys in the resolved haloes mass versus comoving volume plane. We consider the resolved halo mass to be 300 times the particle mass of a simulation. The total comoving volume of our past light-cone within redshift 3.5 projected on two-third of the sky is $\sim10^{12}$ Mpc^3, the right boundary of the plot. We place the simulations enumerated
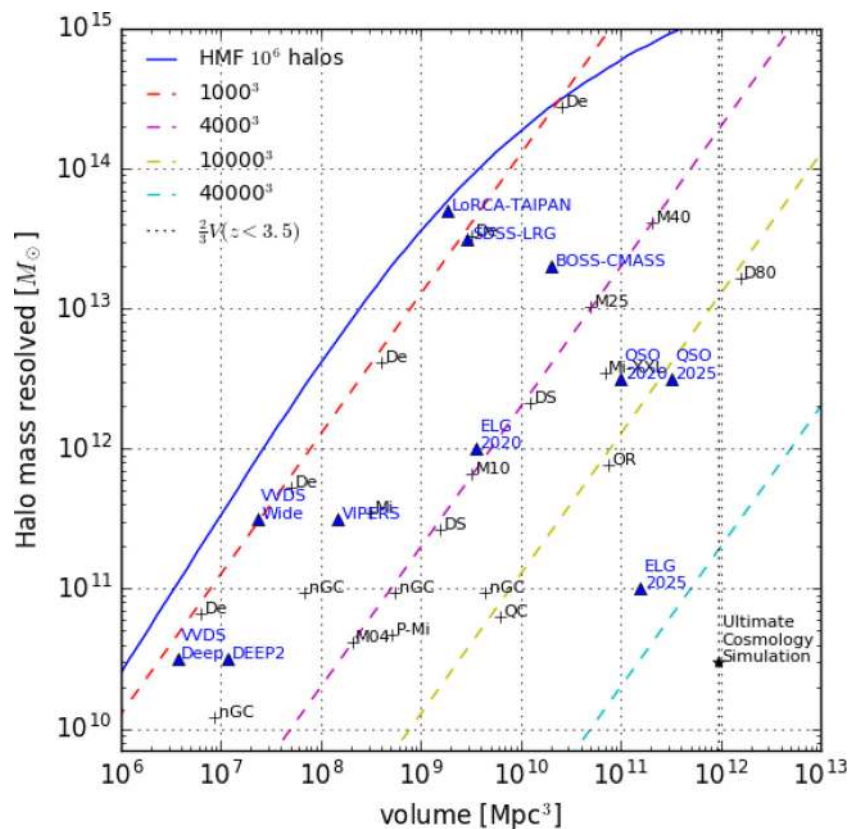
---

[4] http://hmf.icrar.org

Figure 1. Resolved halo mass versus volume. The resolved halo mass is taken as 300 times the particle mass. The set of simulations discussed in this paper (black crosses, De: Despali et al. 2016; M04, M10, M25, M40: MD; DS, D80: DarkSkies; OR: OuterRim; QC: QContinuum; Mi: Millennium; nGC: ν^2GC) are compared to current and future spectroscopic galaxy surveys (blue triangles). The galaxy surveys are tentatively placed according to halo mass values obtained with HOD models, the location is therefore not accurate but rather informative. Dashed diagonal lines relate the volume to the halo mass resolved assuming a constant number of particles 1000^3–$40\,000^3$. Assuming a halo abundance matching model, a simulation encompasses a galaxy sample located above and leftwards to its marker. We show a prediction of the redshift 0 cumulative HMF (blue curve). It is the mass of the least massive halo among the 1000 000 most massive haloes expected in a simulation of the volume given in the x-axis. The total comoving volume of our past light-cone within redshift 2.5 is $\sim 10^{12}$ Mpc^3, the right boundary of the plot.

in Table 1 according to their resolved halo mass and total volume (black crosses). We show with a set of dashed lines the relation between number of particles, volume and halo mass resolved.

It shows how simulations progressed and our future needs (black star on the bottom-right), from the top-left to the bottom-right. We show a prediction of the redshift 0 cumulative HMF. It is the mass of the least massive halo among the 1000 000 most massive haloes expected in a simulation of the volume given in the x-axis. For example, in a volume of 10^9 Mpc^3, there are a million haloes that have $M_{vir} > 4 \times 10^{13}$ M$_\odot$. The galaxy surveys (blue triangles) are tentatively placed according to halo mass values obtained with HOD models. Given the uncertainty on the HOD model parameters, the halo mass value used could shift around by say a factor of 2 or 3. The survey volumes are accurate. The galaxy surveys represented are VIPERS (Marulli et al. 2013), VVDS-Wide (Coupon et al. 2012), VVDS-Deep (Meneux et al. 2008), DEEP2 (Mostek et al. 2013), SDSS-LRG (Padmanabhan et al. 2009), BOSS-CMASS (Rodríguez-Torres et al. 2016), ELG 2020 (Comparat et al. 2013; Favole et al. 2016), ELG 2025 (DESI Collaboration et al. 2016), QSO 2020 (Rodríguez-Torres et al. 2017) and QSO 2025 (DESI Collaboration et al. 2016). If a simulation point is to the lower right of a data point, it means the simulation is sufficient to construct at least one realization of the observations (assuming a halo abundance matching model). We note the challenge to simulate upcoming ELG

samples to be observed by DESI, 4MOST and *Euclid*. Indeed a simulation with $L_{box} \sim 10\,000\,h^{-1}$ Mpc sampled with $\sim 20\,000$ cube particles is needed. It seems that such simulations should become available in the coming decade. However, we do not need to simulate in a single box the exact volume of the observations to extract the cosmological information, see Klypin & Prada (2017) for an extended discussion on the subject.

3.1 Halo catalogues

The halo finding process is a daunting task and in this analysis, we do not enter in this debate (see Knebe et al. 2011, 2013; Behroozi et al. 2015, for a review). For this analysis, we use the ROCKSTAR (Robust Over density Calculation using K-Space Topologically Adaptive Refinement) halo finder (Behroozi et al. 2013). Spherical DM haloes and subhaloes are identified using an adaptive hierarchical refinement of friends-of-friends (FoF) groups in six phase-space dimensions and one-time dimension.

ROCKSTAR computes halo mass using the spherical overdensities of a virial structure. Before calculating halo masses and circular velocities, the halo finder removes unbound particles from the final mass of the halo. We use haloes that have a minimum of a 1000 bound particles, a very conservative threshold for convergence (some analysis use haloes with 300 particles, or even down to only 30 particles

or so in the case of FoF haloes). We characterize the halo population with two properties, $M_{\rm vir}$ and $V_{\rm max}$ at present.

For the halo mass, we use $M_{\rm vir}$, defined relatively to the critical density ρ_c by

$$M_{\rm vir}(z) = \frac{4\pi}{3} \Delta_{\rm vir}(z) \Omega_{\rm m}(z) \rho_c(z) R_{\rm vir}^3. \qquad (14)$$

Indeed, the halo $M_{\rm vir}$ function was found to be closest to an eventual universal mass function (Despali et al. 2016). Throughout the analysis, we convert the mass variable to σ as defined in equation (3). To do so, we measure the DM power spectrum ($P_{\rm DM}$) on each simulation at redshift 0. Then, we take the mean of the ratio $P_{\rm DM}/P_{\rm lin}$ on large scales; where $P_{\rm lin}$ is the predicted linear power spectrum by CAMB using the cosmological parameters of the simulation. Finally, we rescale the M–σ relation accordingly to align all simulations to the input cosmological parameters. The value of the rescaling is given in the σ_8 column of Table 1.

The maximum of the circular velocity profile is a measure of the depth of the DM halo potential well. It is expected to correlate well with the baryonic component of galaxies such as the luminosity or stellar mass as followed from the Tully–Fisher relation (Tully & Fisher 1977). The maximum circular velocity is defined by equation (15). It has a very small dependence on radius and is therefore robustly determined,

$$V_{\rm max} = \max_r \left(\sqrt{\frac{GM(<r)}{r}}, \text{ over radius } r \right). \qquad (15)$$

3.2 Measurements

We divide each snapshot in 1000 subvolumes (on a grid of $10 \times 10 \times 10$). We compute the histogram of the halo mass in each subvolume. The bins start at 8 and run to 16 by steps of $\Delta \log_{10}^M = 0.05$. We denote, $N^{\rm bin\,i}$, the number count in a subvolume in a mass bin. Lukić et al. (2007) and Bhattacharya et al. (2011) corrected the mass assignment according to the force resolution of each simulation. We follow their corrections: $M_{\rm corrected} = [1 - 0.04(\epsilon/650\,{\rm kpc})] M_{\rm halo\,finder}$. The masses were overestimated by 0.3, 0.3, 0.1, 0.1, 0.05, 0.02 per cent in the M40, M40n, M25, M25n, M10, M04, respectively.

We estimate the uncertainty on the mass function using jackknife re-samplings by removing 10 per cent of the subvolumes. We obtain 10 mass function estimates based on 90 per cent of each volume. In each simulation snapshot, we select bins where the halo mass is greater than a 1000 times the particle mass and where the number of haloes is greater than 1000. We divide the number counts by the volume to obtain number densities

$$\mathrm{d}n(M) = \frac{N^{\rm bin\,i}\left(\log_{10}^{\rm bin}(M_i)\right)}{\rm Volume}, \qquad (16)$$

that we further divide by the natural logarithm of the bin width, to estimate the mass function, denoted interchangeably

$$n(\sigma, z) = \frac{\mathrm{d}n}{\mathrm{d}\ln M}. \qquad (17)$$

The resulting mass function estimation for distinct and satellite haloes at redshift 0 are presented in Fig. 2. The measurements span the range $11 < \log_{10}(M_{\rm vir}/M_\odot) < 15(13.5)$ for the distinct (satellites) haloes.

We find the DS HMF at redshift 0 to be 2 per cent lower than the combined MD mass function. This is due to the lower matter content in the DS simulation. Also due to its large volume, the resolution
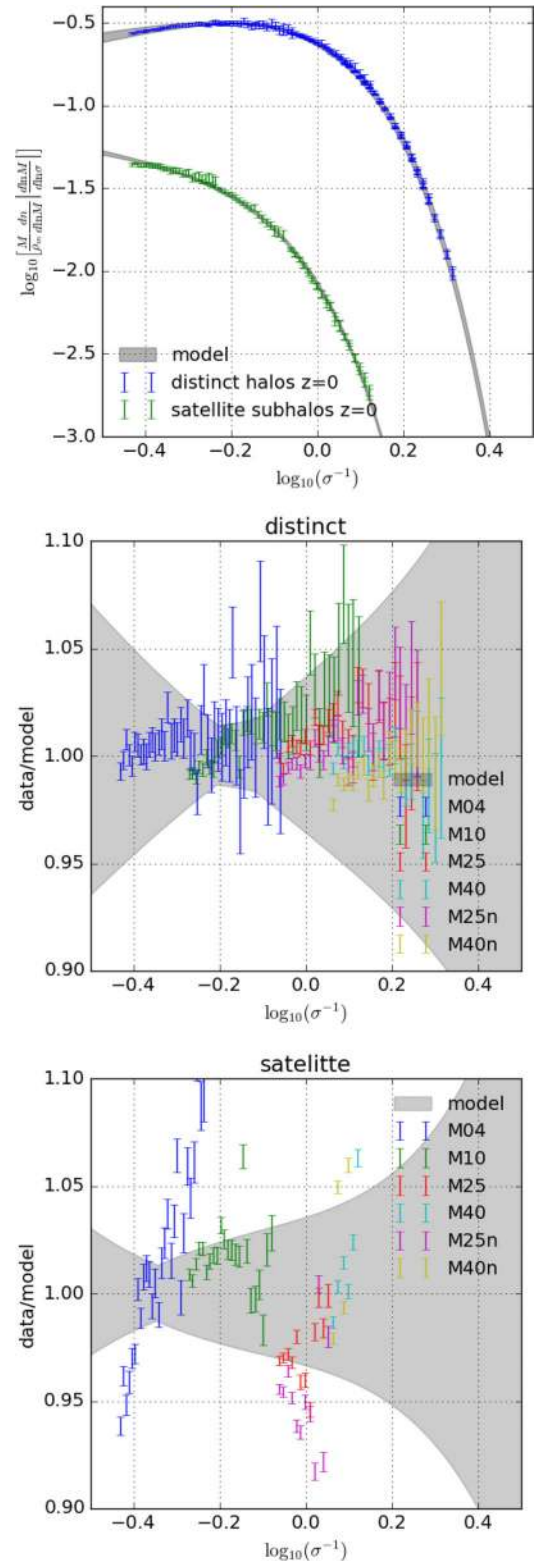

Figure 2. Measurements of the differential halo $M_{\rm vir}$ function for distinct haloes as a function of $\log_{10}(\sigma^{-1})$ for the MD simulations at redshift 0. The grey contours represent the best-fitting models discussed in Section 4. The mean of the residuals for the distinct (satellite) HMF is 0.8 per cent (0.4 per cent) and the standard deviation of the residuals is 1.6 per cent (4.2 per cent), which are shown in the middle (bottom) panel. It means the fit is very close to the data for the distinct haloes and a little further for the subhaloes.

does not enable to follow the mass function leftward of its knee, which prevents from fitting reliably the mass function models solely on the public DS data. The other DS simulations, which are smaller and complementary, are not provided to the public. Therefore, we do not push further the analysis with this simulation.

3.3 Covariance with mass

We construct two estimators of the uncertainty on the mass function measurements. We consider the redshift fixed. For both, we slice the simulations into 1000 subsamples of equal volume. The grid is $10 \times 10 \times 10$. Each subsample has a volume 1000 times smaller than the initial simulation. The first method goes as follows. On each subsample, we estimate the mass function to obtain $N_R = 1000$ of them. We denote by $f_i(\sigma)$ the multiplicity functions deduced. Then, we compute the covariance matrix C defined by

$$C(\sigma_a, \sigma_b) = \frac{\Sigma_i^{N_R}(f_i(\sigma_a) - \bar{f}(\sigma_a))(f_i(\sigma_b) - \bar{f}(\sigma_b))}{(N_R - 1)}, \quad (18)$$

where $\bar{f}$ is the mean multiplicity function. Because each subsample ends up being quite small, the matrices hereby obtained do not cover a large dynamic range in mass.

The second method is the jackknife. We group the subsamples by batches of 100 to obtain $N_R = 10$ realizations of the mass function using the complementary 900 subsamples. The mass functions obtained are not independent, but they cover a larger mass range. From this method, we only infer the diagonal error

$$C^{JK}(\sigma) = \frac{\Sigma_i^{N_R=10}(f_i(\sigma) - \bar{f}(\sigma))^2}{(N_R - 1)}. \quad (19)$$

We show the diagonal variances $C(\sigma, \sigma)$ and $C^{JK}(\sigma)$ on Fig. 3. There is one panel per simulation snapshot at redshift 0. We note that both methods are in agreement when estimating the errors in the low-mass regime. It is the regime where errors are dominated by sample variance. The jackknife method seems less sensitive to the shot-noise at the high-mass end. But this is simply a matter of the volume considered when estimating the uncertainty. Indeed in the jackknife method, we use 90 per cent of the volume whereas in the covariance, we only use 0.1 per cent of the volume. Therefore, a factor of $\sqrt{1000} \sim 30$ is expected between the two measurements. At the low-mass regime, the sample variance seems underestimated by the full covariance method. This discrepancy cannot be explained by the difference in volume covered, we therefore assume that there is a bias in the method.

The full covariance matrix varies smoothly with σ. The covariance matrix is not decreasing around its diagonal as the covariance matrix of the two-point correlation function does (see fig. 7 of Comparat et al. 2016). Indeed, there is a large amount of correlation between structures, i.e. the power spectrum of the DM is not zero. The model of the covariance matrix and its use in the analysis have been discussed in Section 2.

3.4 Covariance with redshift

The mass function at redshift 0 strongly depends on the mass function from previous redshifts i.e. on the complete formation history of the haloes. Therefore, fitting the redshift evolution of the parameters of the mass function is somewhat degenerate. The additional information between two redshift bins are the new (sub)haloes that formed, the mass increase of previous (sub)haloes and the cross-talk between the two functions (see Giocoli et al. 2010; van den

Bosch & Jiang 2016, for an exhaustive list of events occurring during the evolution of the mass function). Due to the limited number of N-body realizations (6 for MD), we cannot establish directly the redshift covariance of the mass function.

We run a set of approximate DM simulations to estimate the redshift covariance of the mass function to wisely choose the redshift sampling and avoid over-fitting in the later analysis. We run a set of Parallel Particle-Mesh GLAM simulations (PPM-GLAM; Klypin & Prada 2017) with lower resolutions and lower time-step resolution than a typical high-resolution N-body simulation to obtain a set of a 100 simulations with density field catalogues spanning the redshifts $0 \leq z < 3.2$ every 0.5 Gyr (23 time steps). With MD, the number of realizations available is 6, a rather small number to obtain variances. On each realization and at each time step, we estimate the density field with a Cloud-In-Cell estimator. Table 3 summarizes the PPM-GLAM runs.

We estimate the redshift covariance matrix, C_z, of the density field function, f^δ as

$$C_z^\delta(z_a, z_b) = \frac{\Sigma_i^{N_R}\left(f_i^\delta(z_a) - \bar{f}^\delta(z_a)\right)\left(f_i^\delta(z_b) - \bar{f}^\delta(z_b)\right)}{(N_R - 1)} \quad (20)$$

at fixed values of the density field δ. We deduce the Pearson product-moment correlation coefficients R defined by

$$R_z(z_a, z_b) = \frac{C_z^\delta(z_a, z_b)}{\sqrt{C_z^\delta(z_a, z_a)C_z^\delta(z_b, z_b)}}. \quad (21)$$

The DM density field function, f^δ, for $1 + \delta = \rho/\bar{\rho} > 10$ looks like a power law. At the highest densities, f^δ is cut-off exponentially (due to finite resolution of the PPM-GLAM simulations). In the cross-correlation matrix, we find two regimes; see Fig. 4. At the high-density field end, $1 + \delta = \rho/\bar{\rho} > 1000$, the cross-correlation coefficient is smaller than 20 per cent between redshifts 0 and 10. The off-diagonal cross-correlations coefficient are of the order of 10 per cent. Therefore, each snapshot brings significant information in this regime. At the lower end of the density field function, $\delta = \rho/\bar{\rho} < 200$, the cross-correlation coefficient is larger than 80 per cent. It means that using a single redshift gives most of the information available. In between the transition is quite sharp, it suggests we should retain for the analysis the $z = 0$ mass function measurements and the high-mass end of the $z > 0$ mass function measurements. A cut-off at $\sim$200 times the density field seems reasonable. It corresponds to $\sim 10^{12.9}\,\mathrm{M}_\odot$. For simplicity, in this analysis we only use the redshift $z = 0$ data and push back the question of accurate estimation of the redshift covariance for future studies.

These simulations give a sense of the redundancy of the information present in the data, but do not allow a robust estimation of the covariance matrix. With these simulations, we cannot weight each snapshot according to its information content. To do that, we would need a large amount of N-body simulations with halo finders run to estimate properly this covariance. Nevertheless, it allows rejection of data with high covariance.

Our understanding of the redshift covariance matrix is that the density field function at low overdensity is redundant with redshift. We agree that between a density field function and an HMF, there is a non-negligible step that is halo finding. Nevertheless, we think that adding all measured mass function points (in all written snapshots i.e. all the redshifts of the simulations) might lead to an incorrect statement as points cannot be considered to be strictly independent from one another.
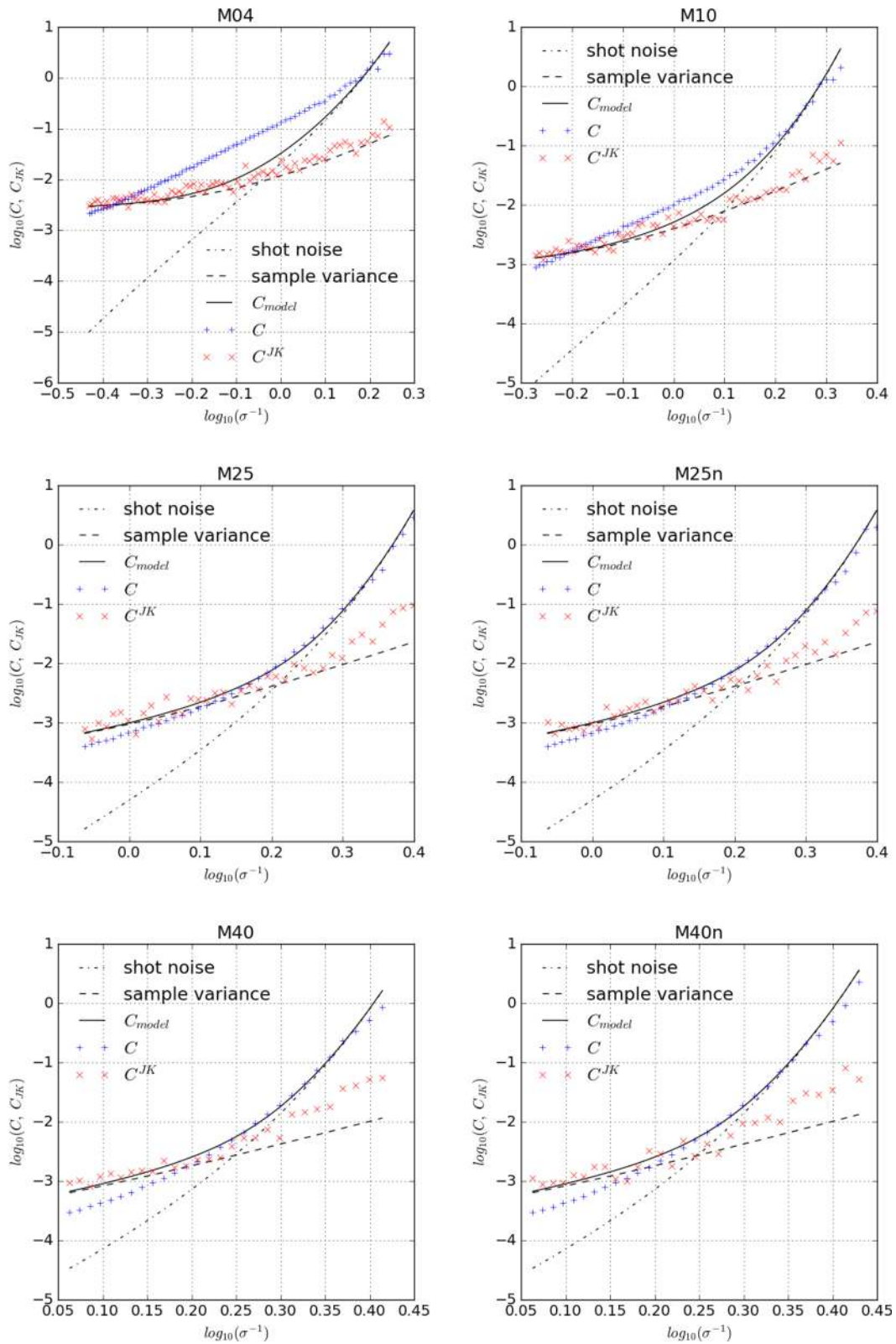
Figure 3. Diagonal component of the covariance matrix measured in each MD simulation (blue pluses) at redshift $z = 0$ compared to the errors obtained via the jackknife method (red crosses). The model is decomposed into shot-noise and sample variance.

Table 3. Parameters of the PPM-GLAM simulations run in Planck cosmology with $\sigma_8 = 0.8229$.

Name	L_{box} Mpc	$N_p^{1/3}$	M_p $M_\odot$	Grid	dt (Gyr)	da	N_R
pmA1	737.7	500	8.5×10^{10}	1000	0.5	0.0004	100
pmA2	737.7	500	8.5×10^{10}	1000	0.5	0.0002	100
pmA3	147.5	500	6.8×10^{8}	1000	0.5	0.0002	100
pmA4	1475.5	2000	1.1×10^{8}	2000	0.5	0.0002	10
pmB1	1475.5	1000	8.5×10^{10}	1000	0.5	0.0002	10
pmB2	147.5	1000	8.5×10^{7}	1000	0.5	0.0002	10
pmB3	14.7	1000	8.5×10^{4}	1000	0.5	0.0002	10
pmB4	1.4	1000	8.5×10^{1}	1000	0.5	0.0002	10
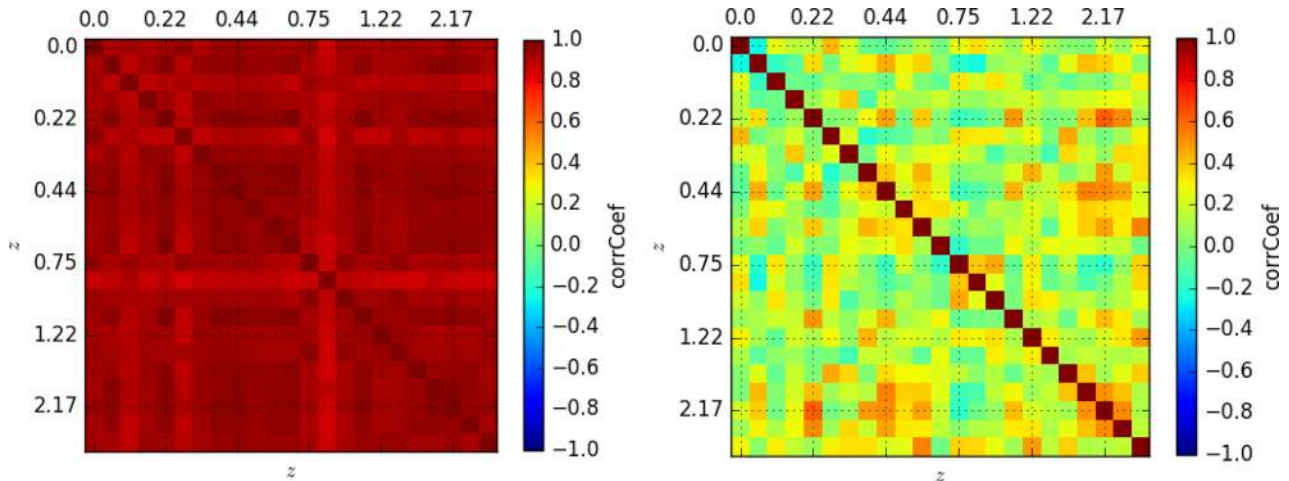
Figure 4. $R_z(z_a, z_b)$, redshift cross-correlation coefficient matrix of the number counts for density field values of $1 + \delta = 100$ (left) and 1000 (right).

It seems that to further improve the accuracy of the HMF and in particular its evolution with redshift, we need to properly work out its redshift covariance matrix, but this needs significantly more simulations to be run, so we leave it for future studies.

3.5 Large-scale halo bias

We compute the real-space two-point correlation function of the halo population in mass bins (identical as the ones used for the mass function) up separations to $r_{max} = 20\,h^{-1}$ Mpc. We follow a method described in Martinez & Saar (2002) that goes as follows.

We select all haloes in a mass bin $[M, dM]$. It constitutes the complete sample of haloes (H_C). Then, we select an 'inner' sample of haloes (H_I) that are located at least r_{max} away from any edge of the snapshot. We count all pairs between the H_C and H_I sample using the *scipy.spatial.ckdtree* python library (Jones et al. 2001). The histogram of the pair counts in bins of distance gives the number of pairs found at separation $r \pm dr/2$, denoted $N_{pairs}(r, dr)$. The real-space two-point correlation function, ξ, is then obtained by

$$1 + \xi(r, dr, M, dM) = \frac{N_{pairs}(r, dr)}{\#H_C \#H_I} \frac{3V_{snap}}{4\pi((r + dr)^3 - r^3)}, \quad (22)$$

where V_{snap} is the volume of the snapshot and the distance binning parameter $dr = 0.1\,h^{-1}$ Mpc. This is a fast and unbiased estimator of the two-point function in simulations.

We compute the redshift 0 linear correlation function, denoted ξ_{lin}^0, using CAMB and the Hankel transform (Szapudi et al. 2005; Challinor & Lewis 2011).[5]

For scales $8 < r < 20\,h^{-1}$ Mpc, we divide the correlation function measured by the linear one. We take the mean to estimate the large-scale halo bias

$$b_h^2(M_{vir}) = \frac{1}{N_i} \sum_i \frac{\xi(M_{vir}, r_i)}{\xi_{lin}^0(r_i)}. \quad (23)$$

We use the standard deviation of the latter ratio to estimate its uncertainty.

Fig. 5 shows the halo bias measured at redshift 0 and the best-fitting models. The agreement between the data and the model is very good; see the discussion in the next section.

4 RESULTS

The determination of the best-fitting model requires the assignment of errors on the data points. The covariance matrix discussed in the previous section is proportional to the product of the biases

$$C(\sigma_1, \sigma_2) \propto \frac{b(\sigma_1)b(\sigma_2)}{\sqrt{\bar{n}(\sigma_1)\bar{n}(\sigma_1)}}. \quad (24)$$

Thus, each line of the matrix is proportional to another lines of the matrix, making it singular. It prevents from estimating the χ^2 statistics for a given data-model pair, (D, M) via the inverse of the covariance matrix $\chi^2 = (D - M) \cdot C^{-1} \cdot (D - M)^T$.

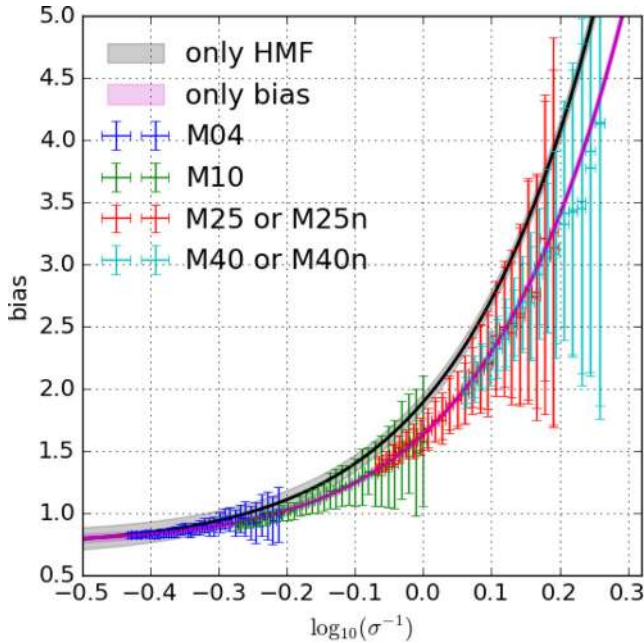[5] https://pypi.python.org/pypi/hankel

Figure 5. Large-scale halo bias versus halo mass. Error bars show the data from the MD simulations at redshift 0. The bias predicted using the best-fitting parameters obtained on the HMF is shown in grey and the bias model fitted on the bias data is shown in magenta.

We circumvent this issue as follows. First, in Section 4.1 we use the uncertainty estimated with the jackknife method on the mass function and fit only the mass function data. Then in Section 4.2, we fit the bias equation that involves the same parameters as the mass function to obtain another constraint on the parameters based on the covariance of the data. Finally in Section 4.3, we provide a relation to predict the covariance matrix for a given simulation.

4.1 Distinct HMF

To determine the best parameters for the mass function of distinct haloes, we use a χ^2 minimization algorithm[6] to obtain the set of best-fitting parameters. We fit the mass function model from equations (7) and (8) to the data at redshift 0. We thus constrain the two sets of parameters (A, a, p) and $(\bar{A}, \bar{a}, \bar{p}, \bar{q})$. We determined the parameters for different flavours of the data. 'MD D(S)MF' stands for the distinct (satellite) mass function from MD data. 'DS DMF´ stands for the distinct mass function from DS data. We use the Jack-knife diagonal errors. The fit of equation (7) on the MD DMF gives a reduced $\chi^2 = 1.28$. The model is not a satisfying statistical representation of the data as the probability of acceptance is 0.7 per cent. We find parameters somewhat discrepant to what was found in Despali et al. (2016). The fit of equation (8) to the MD DMF gives a reduced $\chi^2 \sim 0.75$, meaning it is an accurate description of the data. The probability of acceptance is >99 per cent. We find $(\bar{A}(0), \bar{a}(0), \bar{p}(0), \bar{q}(0))$=(0.280±0.002, 0.903±0.007, 0.640±0.026, 1.695±0.038). Table 4 hands out the best-fitting parameters obtained. We therefore think that adding the $\bar{q}$ parameter suggested by Bhattacharya et al. (2011) enhances significantly the quality of the fit to the DMF. The bottom panel of Fig. 2 shows the residuals after the fit of the model given in equation (8). The mean of the residuals for the distinct

HMF is 0.8 per cent and the standard deviation of the residuals is 1.6 per cent. It means the fit on average underestimates the HMF by less than 1 per cent. Furthermore, except for a few outliers the MD DMF is very well described by the model to the <2 per cent level.

We compare our fits to previous ones in Fig. 6. The mass function differs from up to a factor of 2 when compared to different cosmologies. Our fit agrees within <10 per cent with other analysis in Planck cosmology in the lower mass regime. At larger masses, the disagreement between our measurements and previous ones in Planck cosmology is due to the difference in the data used. In this paper, we use extremely large simulations whereas in previous analysis, the largest simulation were covering volumes 8–64 times smaller. The high-mass end being modelled by an exponential, it drives the fit to a different location in parameter space.

4.2 Large-scale halo bias

The fit of the model given in equation (9) suggests the following set of parameters $(\bar{a}, \bar{p}, \bar{q}) = (0.740 \pm 0.008, 0.61 \pm 0.02, 1.64 \pm 0.03)$. These are in slight tension with that of the HMF model (1σ contours do overlap); see Table 4 for a face-to-face comparison of the figures. It is slightly higher for large masses and slightly lower for low mass.

We are pleased to see that the excursion-set formalism works well to describe the mass function and the large-scale halo bias precisely. Such a low level of tension is worth the praise.

A joint fit to solve this issue is not straightforward. Indeed, the large-scale halo bias is related to the uncertainty on the mass function. We leave this for future studies.

4.3 Covariance matrix

In the comparison of the diagonal errors estimated, see Fig. 3, the two methods showed some disagreement: at the high-mass end where errors are dominated by the shot-noise and at the low-mass end where the errors are dominated by the sample variance. The difference in shot-noise is understood as the volumes used to differ in the two error-estimating methods. On the contrary, the difference in sample variance is puzzling. Indeed when using a larger volume, the sample variance estimated is higher than in the method using a smaller volume. This seems rather strange, as we expected the opposite. We take a conservative option. We consider the maximum of the two error estimates to fit the model: the *JK* estimates at the low-mass end and the covariance at the high-mass end.

According to the model, fitting all the coefficients of the covariance matrix is redundant. The shot-noise component is a scaling relative to the inverse of the density times the volume. The sample variance depends on the product of the biases and on the cosmology. Therefore, as soon as a single line of coefficient of the covariance matrix is reproduced by the model, other coefficients should be in line with the model. This is indeed what we observe. As data points, we simply use the diagonal of the covariance matrix. Note that the points are for $L_{box}[h^{-1} \text{Mpc}] = 40, 100, 250$ and 400, a factor of 10 smaller than the boxes used for the mass function estimate.

We fit a linear relation between the Q factor and the log of the side length of the simulations (i.e. the length of the simulations divided by 10 due to the subsampling). The uncertainty on the coefficients of the covariance matrix is unknown, so we perform a fit where the data points are equally weighted. Using the large-scale halo

---

[6] scipy.optimize.minimize: https://docs.scipy.org/doc/scipy-0.18.1/reference/generated/scipy.optimize.minimize.htmldocs.scipy.org

Table 4. Best-fitting parameters of the model at redshift 0. *D(S)MF* stands for distinct (satellite) mass function. B11: Bhattacharya et al. (2011). D16: Despali et al. (2016). A dash '-' means the entry is the same as above.

| A(0)	a(0)	p(0)		$\chi^2/n\,d.o.f$	$P(X > x, d.o.f)$	Data	Model Eq.	Ref
0.333 ± 0.001	0.794 ± 0.005	0.247 ± 0.009					(7)	D16
0.3170 ± 0.0008	0.818 ± 0.003	0.118 ± 0.006		$238.69/\,187 = 1.28$	0.7 per cent	MD DMF	–	This paper
0.0423 ± 0.0003	1.702 ± 0.010	0.83 ± 0.04		$31.03 / 84 = 0.37$	100 per cent	MD SMF	–	–
$\bar{A}(0)$	$\bar{a}(0)$	$\bar{p}(0)$	$\bar{q}(0)$	$\chi^2/n\,d.o.f$	$P(X > x, d.o.f)$	Data	Model Eq.	Ref
0.333	0.786	0.807	1.795				(8)	B11
0.280 ± 0.002	0.903 ± 0.007	0.640 ± 0.026	1.695 ± 0.038	$138.76 / 186 = 0.75$	99.6 per cent	MD DMF	–	This paper
0.27 ± 0.02	0.92 ± 0.03	0.36 ± 0.68	1.6 ± 0.6	$9.13 / 21 = 0.43$	98.9 per cent	DS DMF	–	This paper
Free	0.740 ± 0.008	0.61 ± 0.02	1.64 ± 0.03	$8.36/141 = 0.059$	100 per cent	Halo bias	–	This paper
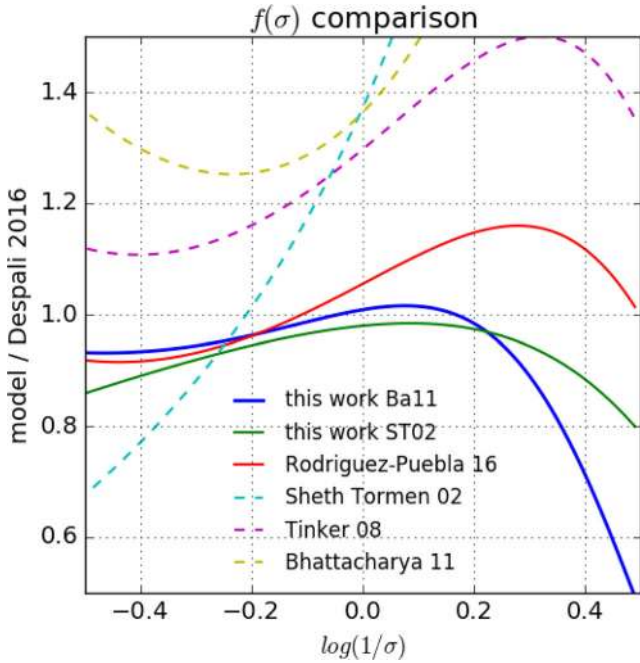
Figure 6. Comparison of mass functions with respect to the Despali et al. (2016) fit. The line 'this work Ba11' corresponds to the fits of equation (8) to the data and 'this work ST02' corresponds to the fits of equation (7) to the data. Studies done in the Planck cosmology have solid lines whereas studies in other cosmologies are shown with dashes. The difference at large masses is due to the difference in the simulation volumes.

Figure 7. *Q* covariance rescaling factor versus side length of the simulation and its linear fitting relation, see equation (25).

bias model from the previous subsection, we find that the following fitting relation,

$$Q = -3.62 + 4.89 \log_{10}[L_{\text{box}}(h^{-1}\,\text{Mpc})], \qquad (25)$$

produces a covariance matrix model very close to the MD data at redshift 0. Fig. 7 shows the *Q* versus the size of the simulation. We find the model to account well for the measured covariance, see Fig. 3 where the solid, dashed and dotted lines represent each component of the model. By combining equations (25), (13) and $\sqrt{C_{\text{model}}(\sigma, \sigma, L_{\text{box}})}$, one predicts a reliable uncertainty on the distinct HMF for any simulation in the Planck cosmology.

4.4 Subhalo and substructure mass function

In this analysis, we do not enter into the debate of the definition of subhaloes. We use the subhaloes as obtained by the ROCKSTAR halo finder at redshift 0. The substructure hierarchy in DM haloes was investigated in details by Giocoli et al. (2010) and van den Bosch & Jiang (2016). They argue two functions are needed to
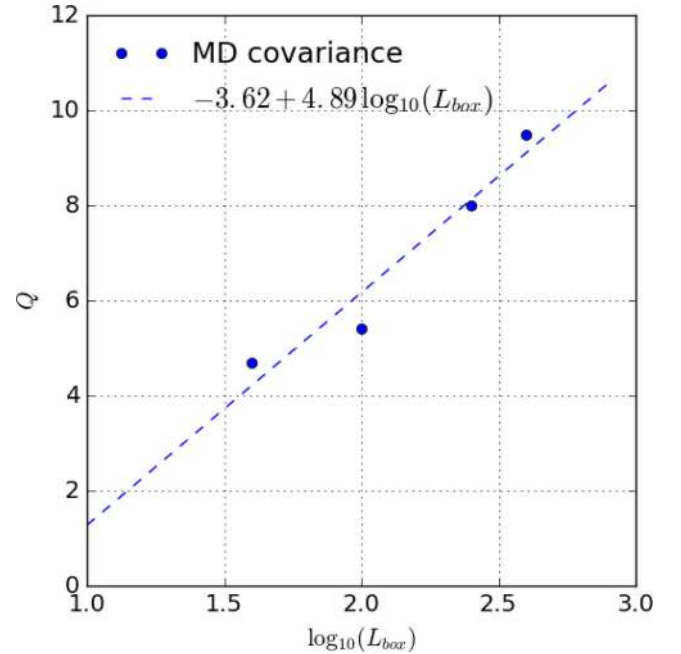
fully characterize in a statistical sense the subhalo population: the HMF and the substructure mass function. The convolution of the two gives the subhalo mass function.

We measure the subhalo mass function with the same method as for the distinct HMF; see Fig. 2. We fit the subhalo mass function (MD SMF in Table 4) with equation (7) and obtain a reduced $\chi^2 \sim 0.37$, meaning it is an accurate description of the data (probability of acceptance 100 per cent). We find $[A(0), a(0), p(0)]$=(0.0423±0.0003, 1.702±0.010, 0.83±0.04). Adding an additional parameter *q* is not necessary. The mean of the residuals for the subhalo mass function compared to this model is 0.4 per cent and the standard deviation of the residuals is 4.2 per cent. So the model is a little further away on average than for the MD DMF. To further refine the model, a complete discussion on what a subhalo is would be necessary. For the purpose of HOD, adding a subhalo mass function with a 4 per cent precision is a non-negligible advance. We warn the reader that the excursion-set formalism does not predict the subclumps within haloes. We simply use the function (7) as an analytical model to describe the data.

Then, for a subhalo of mass M_s we consider its relation to its host, a distinct halo of mass M_d, by studying the distribution of the ratio $Y = M_s/M_d$. In this aim, we measure the so-called substructure
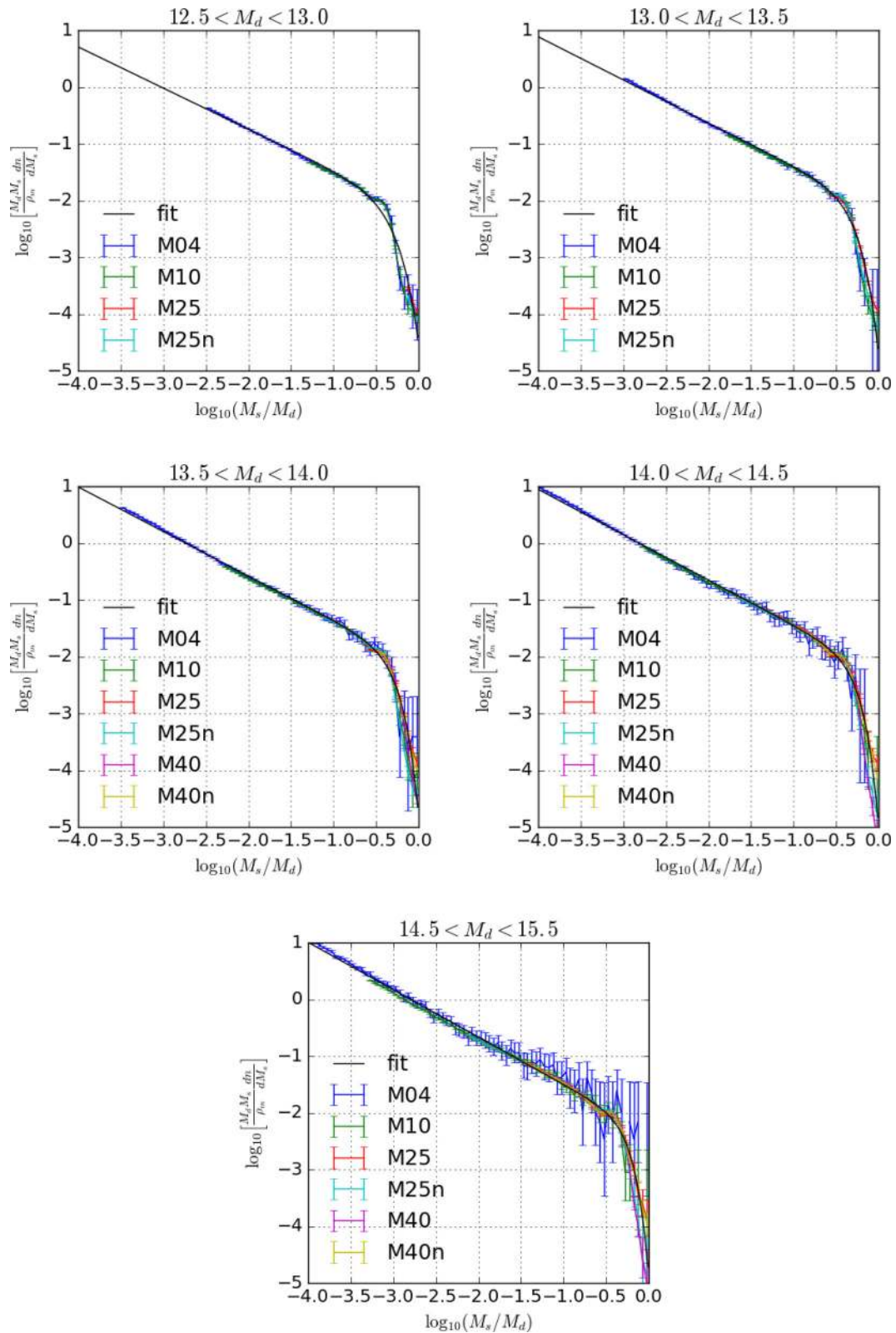
Figure 8. Substructure mass function for five distinct (host) halo mass bins. The model seems quite independent of the host halo mass bin.

mass function, defined by the left part of equation (26) and shown on Fig. 8,

$$\log_{10}\left[\frac{M_d^2}{\rho_m}\frac{\mathrm{d}n}{\mathrm{d}M_s}\right](Y) = N_{\mathrm{sub}}Y^{\alpha_{\mathrm{sub}}}\,\mathrm{e}^{-\beta_{\mathrm{sub}}Y^{\gamma_{\mathrm{sub}}}}. \tag{26}$$

Note that M_s is not the mass at the moment of accretion of the subhalo but the mass measured at redshift 0. We parametrize it similarly to van den Bosch & Jiang (2016) with four parameters: overall normalization, N_{sub}, power law at low-mass ratio, α_{sub}, and two parameters for the exponential drop: β_{sub} and γ_{sub}.

Table 5. Number of distinct haloes–subhalo pairs at redshift 0 split in distinct halo mass bins. Best-fitting parameters for equation (26) for each host halo mass bin are given below. The last column is the fit using all the data together.

Box	12.5–13	13–13.5	13.5–14	14–14.5	14.5–15.5	12.5–15.5
M04	515 922	504 923	441 228	284 992	144 352	1, 891 417
M10	938 628	879 394	729 358	480 041	200 699	3 228 120
M25	788 780	1 426 470	1 337 316	833 535	325 951	4 712 052
M25n	784 519	1 414 136	1 318 048	822 464	329 225	4 668 392
M40	19 793	7963 12	1 619 226	1 199 845	467 090	4 102 266
M40n	20 988	797 074	1 578 780	1 167 971	466 143	4 030 956
Total	3 068 630	5 818 309	7 023 956	4 788 848	1 933 460	22 633 203
Parameter			Best-fitting values			
$-\alpha_{sub}$	1.73 ± 0.03	1.76 ± 0.02	1.78 ± 0.01	1.799 ± 0.006	1.834 ± 0.004	1.804 ± 0.004
β_{sub}	5.34 ± 0.16	5.95 ± 0.18	6.12 ± 0.16	6.32 ± 0.21	5.87 ± 0.27	5.81 ± 0.09
$-\log_{10}N_{sub}$	2.19 ± 0.05	2.15 ± 0.03	2.15 ± 0.02	2.25 ± 0.01	2.33 ± 0.01	2.250 ± 0.008
γ_{sub}	1.95 ± 0.14	2.28 ± 0.11	2.46 ± 0.09	2.62 ± 0.09	2.92 ± 0.11	2.54 ± 0.05

The substructure mass function represents the abundance of subhaloes as a function of the mass ratio between the subhalo and its host distinct halo (in a distinct halo mass bin); Giocoli et al. (2010, see equation 2 and fig. 3) and van den Bosch & Jiang (2016, equation 6 and fig. 3). In these works, the authors consider a complete world model of how subhaloes evolve. In this analysis, we focus on the practical aspect of a relation that given a halo population, one can predict the characteristics of its subhalo population. Therefore, we do not apply the exact same formalism as in previous works, but rather something more practical, at fixed redshift. We use the mean density of the Universe to obtain a dimensionless measurement, therefore the normalization parameters have a different meaning than in previous studies. Subsequently, we adjust a four-parameter model, given in the right part of equation (26) to five host halo mass bins and to all the data simultaneously. Fig. 8 shows the substructure mass function measured at redshift 0 in the mass bins delimited by 12.5; 13; 13.5; 14; 14.5; 15.5. The parameters obtained are given in Table 5. The 22 633 203 subhaloes-halo pairs considered constitute a sample that is more than an order of magnitude larger than any previous study. The power law found is compatible with $-\alpha_{sub} = -1.804 \pm 0.004$ in every host mass bin. It confirms measurements from previous analysis, though with greater accuracy. The other parameters found are compatible between mass bins. To a good approximation, the parameters $-\alpha_{sub} = -1.8$, $\beta_{sub} = 5.8$ and $-\log_{10}N_{sub} = 2.25$, $\gamma_{sub} = 2.54$ provide a good description of the substructure mass function (whatever the host halo mass bin).

5 SUMMARY AND DISCUSSION

In this analysis, we measured at redshift 0 the mass function for distinct and satellite subhaloes and the substructure mass function to unprecedented accuracy thanks to the MD Planck simulation suite. Indeed, these simulations encompass eight times larger volumes than what was used in previous studies. We measured and modelled the large-scale halo bias of the distinct haloes. Then, we estimated for the first time the full covariance matrix of the distinct HMF with respect to mass. To refine our knowledge of the satellite subhalo population, we also estimated and modelled the substructure function.

We find that the Bhattacharya et al. (2011) model is a good description for the measurements related to the distinct halo population: its mass function, its large-scale bias and the covariance of the mass function. This new set of models for the mass function and for the velocity function should allow analytical HOD models

to reach better accuracy. We give practical fitting formula and their evolution with redshift of the V_{max} function in the Appendix.

Halo finding process

The halo finding is a difficult task, reason being that both the theoretical and the empirical definitions of, what a halo is, are not precise.

About the empirical definition of a halo, Knebe et al. (2011, 2013) and Behroozi et al. (2015) showed that when varying the halo finder on a single simulation, one should expect variations in the distinct HMF of the order of 10–20 per cent. This estimate, done on a rather small simulation (500 h^{-1} Mpc) with a small number of particles (1024^3), should be regarded today as an upper limit. Hopefully, such an exercise will be repeated with current and future simulations to reach a better empirical halo definition.

About the theoretical halo definition, it seems recent investigations on the extended spherical collapse models by Del Popolo et al. (2017) point towards a modification of the Sheth & Tormen (1999) along the lines of the modifications made by Bhattacharya et al. (2011). So, there might be a physical reason behind the fact that the Bhattacharya et al. (2011) is a better description of the data than Sheth & Tormen (1999).

Unlike distinct haloes, the satellite subhalo definition has not yet reached a consensus in the community. Theoretical advances are pushing towards a unified subhalo model so this uncertainty should hopefully vanish soon (van den Bosch & Jiang 2016). Nevertheless, we provided accurate fits of the statistics obtained with MD combined with ROCKSTAR.

Redshift evolution of the mass function

The redshift covariance of the density field function indicates that the debate about the universality of the mass function throughout redshift might be an ill-posed question.

Given the covariance between different redshift bins in the low-mass end of the density field function, it is hard to define properly how its evolution with redshift should be modelled. Simply using all the redshift outputs produced by the simulation is redundant. We therefore think the question of the universality needs to be approached with a slightly different theoretical background. Many more N-body simulations would need to be run to obtain deep insights on the redshift covariance of the HMF. But it does not seems reasonable to run a thousand MD of DS simulations? To save

Table 6. Ratio between the HMF with and without baryonic effect as a function of halo mass, $f_{\mathrm{hydro}}/f_{\mathrm{DM-only}}$. References are 1: Velliscig et al. ([2014](#)); 2: Vogelsberger et al. ([2014](#)); 3: Schaller et al. ([2015](#)); 4: Tenneti et al. ([2015](#)); 5: Bocquet et al. ([2016](#))

9–10	10–11	11–12	12–13	13–14	14–15	Simulation	Reference
		0.8	0.8	0.8	0.9	OWLS	1
0.8	0.8	1.1	1	0.9	0.9	Illustris	2
0.7	0.8	0.85	0.9	0.95	1	Eagle	3
0.8	0.85	0.85	0.9	0.95	1	Massive black 2	4
0.9	0.9	0.9	0.9	0.9	1	Magneticum	5

computation time, a possibility would be to study the evolution of the density field with the new PPM-GLAM method. In this paradigm, the number of realizations is not an issue and cosmological parameters are easily varied.

About the effects of baryons on the HMF

The baryons hosted by DM haloes influence the total mass enclosed in the halo. Supernovae and active galactic nuclei (AGNs) feedbacks expel gas from the halo to the intergalactic medium. The total mass enclosed in haloes where baryonic physics is accounted for is of the order of 20 per cent or lower. Therefore, the HMF estimated on DM-only simulations suffers a bias. It seems the number density of DM-only haloes is greater than that of DM+baryon haloes by a factor ∼20 per cent at $M \sim 10^9\, h^{-1}\, \mathrm{M}_\odot$. In clusters, the halo number densities seem in agreement. We summarize numbers obtained from various studies in Table [6](#). At redshift 0, it seems there is a consensus for clusters (impact negligible) and haloes with $\log_{10} M < 12$ (−20 per cent effect). The evolution of this effect with redshift is not clear. Vogelsberger et al. ([2014](#)) and Schaller et al. ([2015](#)) show an effect more or less constant with redshift. The most recent simulations (Bocquet et al. [2016](#)) advocate the effect is negligible at redshift 2 and starts around redshift 1. Recently, Despali & Vegetti ([2016](#)) tested these models by comparing with observed strong lensing events. With current statistics, it does not allow us to choose between feedback models, but with larger samples, the strong lensing probe should decide this problem. Note that the trend with mass vary from a simulation to another due to the differences in the AGN feedback or the supernovae model used. This result is indeed dependent on the recipe of AGN and supernovae feedback, so the true value could be larger (or smaller) but it is difficult to quantify by what amount.

Outlook

All in all, it seems assuming a few per cent statistical errors and of the order of tens of per cents systematical errors reasonably represents our current knowledge of the distinct HMF. To enable per cent precision with mass function cosmology, these results call for deeper investigations. First, about the redshift and mass covariances of the distinct HMF to be able to do proper statistical fits on the data. Secondly, about seeking a better empirical and theoretical definition of what a DM halo is. Lastly, about the remaining n-point functions that carry the next order of information about what haloes are and how they behave.

REFERENCES

Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Avila S. et al., 2014, MNRAS, 441, 3488
Baugh C. M., 2006, Rep. Prog. Phys., 69, 3101
Behroozi P., Wechsler R., Wu H.-Y., 2013, ApJ, 762, 109
Behroozi P. et al., 2015, MNRAS, 454, 3020
Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122
Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, MNRAS, 456, 2361
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, MNRAS, 447, 646
Castro T., Marra V., Quartin M., 2016, MNRAS, 463, 1666
Cen R., Ostriker J. P., 1993, ApJ, 417, 415
Challinor A., Lewis A., 2011, Phys. Rev. D, 84, 043516
Chaves-Montero J., Angulo R. E., Schaye J., Schaller M., Crain R. A., Furlong M., Theuns T., 2016, MNRAS, 460, 3100
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Comparat J. et al., 2013, MNRAS, 433, 1146
Comparat J. et al., 2016, MNRAS, 458, 2940
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Coupon J. et al., 2012, A&A, 542, A5
Dawson K. S. et al., 2013, AJ, 145, 10
Dawson K. S. et al., 2016, AJ, 151, 44
Del Popolo A., Pace F., Le Delliou M., 2017, J. Cosmol. Astropart. Phys., 3, 032
DESI Collaboration et al., 2016, preprint (arXiv:1611.00036)
Despali G., Vegetti S., 2016, preprint (arXiv:1608.06938)
Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, MNRAS, 456, 2486
Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859
Elahi P. J. et al., 2016, MNRAS, 458, 1096

Favole G. et al., 2016, MNRAS, 461, 3421

Giocoli C., Tormen G., Sheth R. K., van den Bosch F. C., 2010, MNRAS, 404, 502

Guo H. et al., 2016, MNRAS, 459, 3040

Habib S. et al., 2016, New Astron., 42, 49

Heitmann K. et al., 2015, ApJS, 219, 34

Hu W., Kravtsov A. V., 2003, ApJ, 584, 702

Ishiyama T., Enoki M., Kobayashi M. A. R., Makiya R., Nagashima M., Oogi T., 2015, PASJ, 67, 61

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Jones E. et al., 2001, SciPy: Open source scientific tools for Python. Available at: http://www.scipy.org/

Klypin A., Prada F., 2017, preprint (arXiv:1701.05690)

Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340

Knebe A. et al., 2011, MNRAS, 415, 2293

Knebe A. et al., 2013, MNRAS, 435, 1618

Knebe A. et al., 2015, MNRAS, 451, 4029

Komatsu E. et al., 2011, ApJS, 192, 18

Laureijs R. et al., 2011, preprint (arXiv:1110.3193)

Lukić Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, ApJ, 671, 1160

Martinez V. J., Saar E., 2002, Statistics of the Galaxy Distribution. Chapman & Hall/CRC, Boca Raton, FL

Marulli F. et al., 2013, A&A, 557, A17

Meneux B. et al., 2008, A&A, 478, 299

Mostek N., Coil A. L., Cooper M., Davis M., Newman J. A., Weiner B. J., 2013, ApJ, 767, 89

Murray S. G., Power C., Robotham A. S. G., 2013, Astron. Comput., 3, 23

Pace F., Batista R. C., Del Popolo A., 2014, MNRAS, 445, 648

Padmanabhan N., White M., Norberg P., Porciani C., 2009, MNRAS, 397, 1862

Planck Collaboration XVI, 2014, A&A, 571, A16

Potter D., Stadel J., Teyssier R., 2016, preprint (arXiv:1609.08621)

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018

Press W. H., Schechter P., 1974, ApJ, 187, 425

Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, 771, 30

Rodríguez-Puebla A., Behroozi P., Primack J., Klypin A., Lee C., Hellinger D., 2016, MNRAS, 462, 893

Rodríguez-Torres S. A. et al., 2016, MNRAS, 460, 1173

Rodríguez-Torres S. A. et al., 2017, MNRAS, 468, 728

Sawala T. et al., 2015, MNRAS, 448, 2941

Schaller M. et al., 2015, MNRAS, 451, 1247

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Sheth R. K., Tormen G., 2002, MNRAS, 329, 61

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Skillman S. W., Warren M. S., Turk M. J., Wechsler R. H., Holz D. E., Sutter P. M., 2014, preprint (arXiv:1407.2600)

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2010, MNRAS, 401, 791

Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V. et al., 2005, Nature, 435, 629

Szapudi I., Pan J., Prunet S., Budavári T., 2005, ApJ, 631, L1

Tenneti A., Mandelbaum R., Di Matteo T., Kiessling A., Khandai N., 2015, MNRAS, 453, 469

Teyssier R., 2002, A&A, 385, 337

The Dark Energy Survey Collaboration 2005, preprint (astro-ph/0510346)

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Tully R. B., Fisher J. R., 1977, A&A, 54, 661

van den Bosch F. C., Jiang F., 2016, MNRAS, 458, 2870

Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Dalla Vecchia C., 2014, MNRAS, 442, 2641

Vogelsberger M. et al., 2014, MNRAS, 444, 1518

Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, ApJ, 646, 881

Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230

APPENDIX A: $V_{\max}$ FUNCTION, MEASUREMENTS AND MODEL

The peak circular velocity was proven more efficient than the halo mass to map galaxies to haloes (Reddick et al. 2013; Guo et al. 2016; Rodríguez-Torres et al. 2016). The peak circular velocity ($V_{\max}$) is less affected than mass by tidal forces and it is thus better defined than halo mass. It traces best the assembly history of the halo and its potential well (Diemand, Kuhlen & Madau 2007). Thus exists an interest in formulating the halo model in terms of peak velocities instead of mass to obtain more accurate predictions with an analytical model. This section is aimed for a practical use in future exploration of the accuracy of the SHAM/HOD.

Using similar estimators as for the mass function, we measure the velocity function. Figs A1 and A2 show the differential velocity function for distinct and satellite subhaloes at redshifts below 2.3. We use jackknife as a proxy for errors to perform the fits. The analysis of errors is not as careful as done previously because we only pretend to provide fitting functions. The limits imposed on the $V_{\max}$ range are M04, [125, 450]; M10, [250, 800]; M25 and M25n, [600, 1100]; M40 and M40n, [900, 1400] km s^{-1}. We estimate a dimensionless velocity function, $V^3/H^3(z)\, \mathrm{d}n/\mathrm{d}lnV$, the left part of equation (A2). As in Rodríguez-Puebla et al. (2016), we model the measurements as the product of a power law and an exponential cut-off using four parameters

$$\log_{10}\left[\frac{V^3}{H^3(z)}\frac{\mathrm{d}n}{\mathrm{d}lnV}\right](V, A, V_{\mathrm{cut}}, \alpha, \beta) \qquad (A1)$$

$$= \log_{10}\left(10^A\left(1 + \frac{10^V}{10^{V_{\mathrm{cut}}}}\right)^{-\beta}\exp\left[\left(\frac{10^V}{10^{V_{\mathrm{cut}}}}\right)^{\alpha}\right]\right), \qquad (A2)$$

where A is the normalization, V_{cut} is the cut-off velocity, α the width of the cut-off and β the power-law index. We model the redshift trends using an expansion with redshift of each parameter, $p(z) = p_0 + p_1 z + p_2 z^2 + p_3 z^3 \cdots$.

We fit first the parameters at redshift 0. Then, we fit their redshift trends in the range $0 \leq z \leq 1$ and then in the range $1 \leq z \leq 2.3$. A model with four parameters is sufficient at redshift 0.6 parameters are used to describe the data in each further redshift ranges. At redshift 0, the fits converge with a reduced $\chi^2 = 1.43$ for the distinct haloes and $\chi^2 = 0.2$ for the subhaloes; see Fig. A3 that shows the residuals of the redshift 0 fits in greater details. Table A1 provides the parameters of the fits for both populations.

In the range redshift $0 \leq z \leq 1$, a linear evolution of the parameters A and V_{cut} is sufficient for the fits to converge with a reduced $\chi^2 = 1.56$ (0.54) for the distinct (satellite); see Fig. A1 (A2) left column row of panels that shows the data, the model and the residuals (from left to right). The parameters A and V_{cut} are compatible in the three redshift bins, while the parameters α and β are not. If we add an evolution term for α and β, the fits converge very slowly and the error on these parameters becomes very large i.e. current data do not allow us to constrain all the parameters at once. Among the parameters, V_{cut} and A are best constrained.
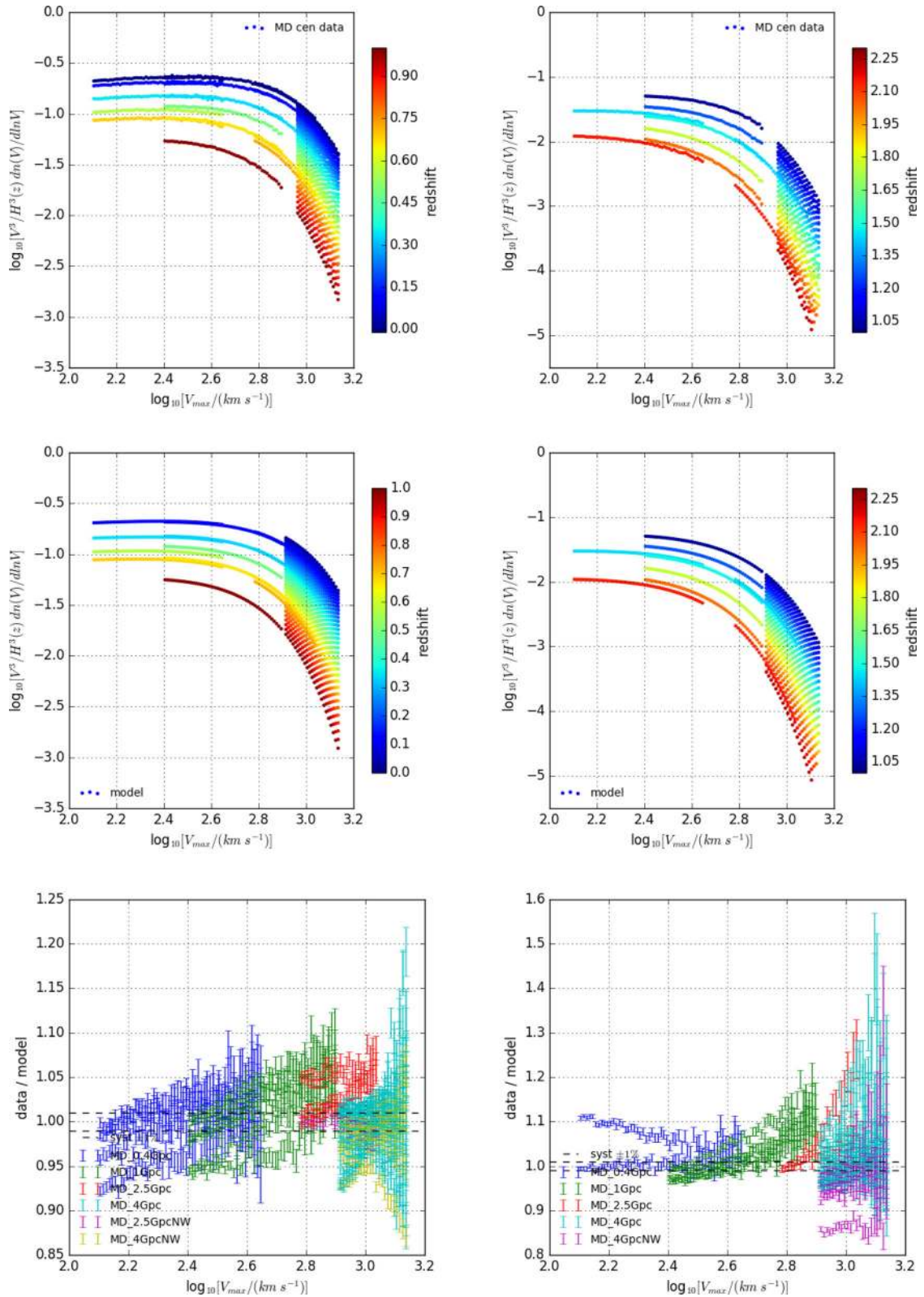
Figure A1. Measurements of the differential distinct halo V_{max} function versus V_{max} coloured with redshift (top row), its model (middle) and the residuals around the model (bottom row). The first column shows the range $0 \leq z \leq 1$ and the second column the $1 \leq z \leq 2.3$ range. Residual around the $0 \leq z \leq 1$ model are contained in ± 15 and ± 20 per cent for the high-redshift range.
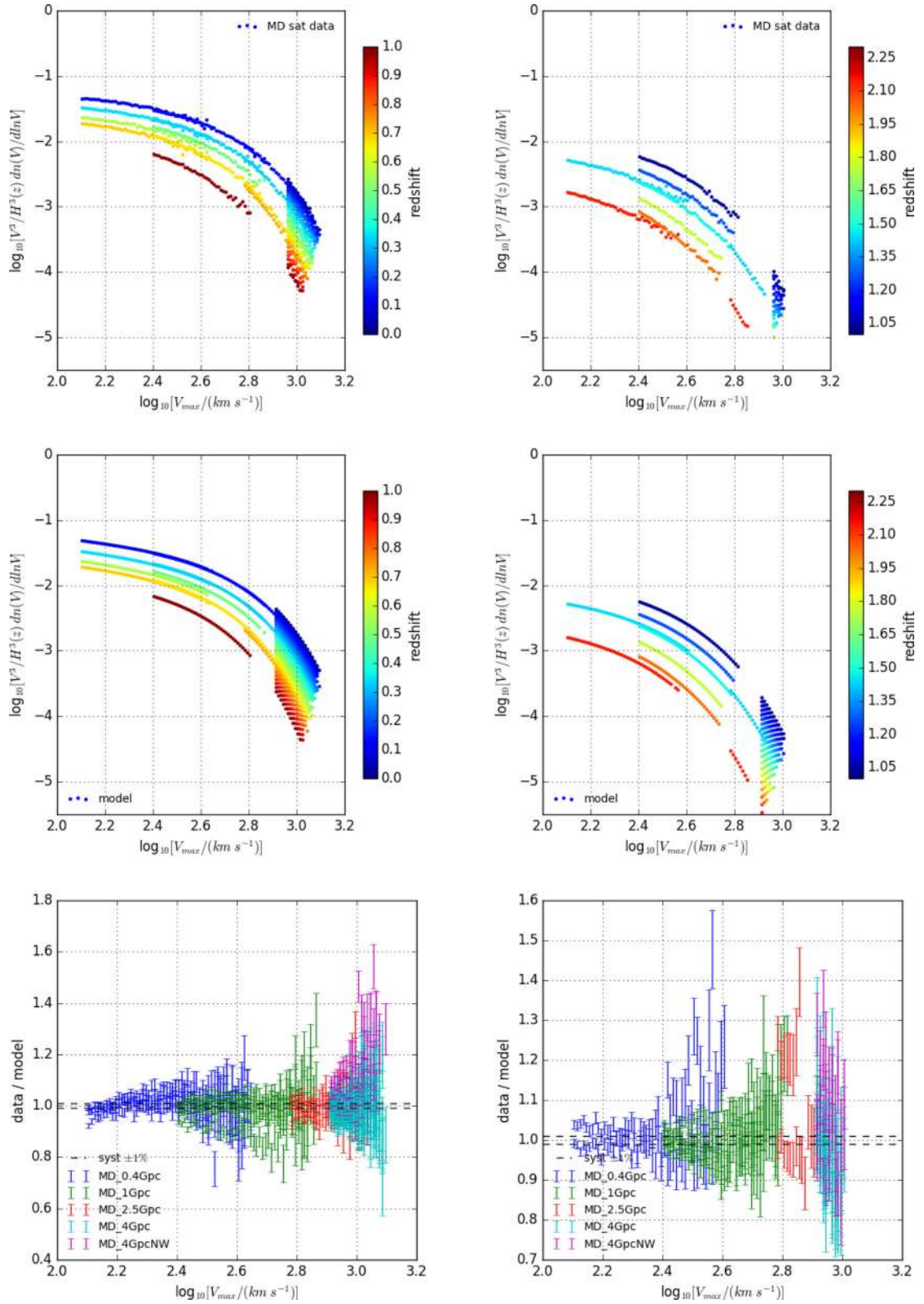
Figure A2. Continued Fig. A1 for the satellite subhaloes in the same redshift ranges. Residuals are of the same order of magnitude as for the distinct haloes.
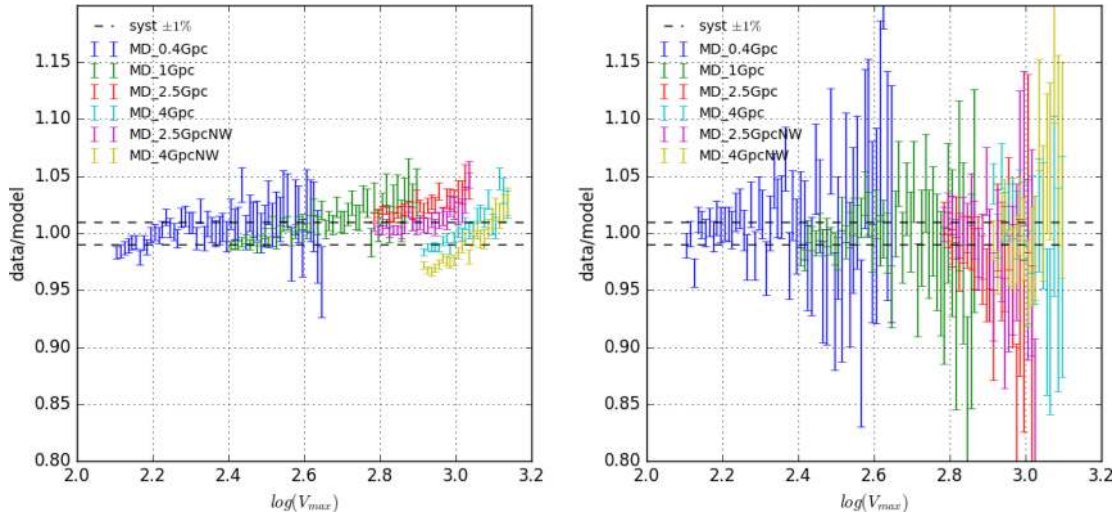
Figure A3. Residuals around the redshift 0 model are well ±5 per cent for the distinct haloes (left) and within ±10 per cent for the satellite haloes (right).

Table A1. Results of model fitting to the V_{max} differential function. Errors are the 1σ errors. Empty cells mean the parameter was not fitted.

		Distinct haloes	
z		p^0	p^1
0	A	-0.74 ± 0.04	
	V_{cut}	2.94 ± 0.02	
	α	2.02 ± 0.08	
	β	-0.79 ± 0.24	
	χ^2	$286.11/199 = 1.43$	
$0 \le z \le 1$	A	-0.71 ± 0.08	-0.62 ± 0.03
	V_{cut}	2.93 ± 0.09	-0.176 ± 0.001
	α	1.782 ± 0.07	
	β	-0.82 ± 0.07	
	χ^2	$2504.8/1599 = 1.56$	
$1 \le z \le 2.3$	A	-0.71 ± 0.14	-0.62 ± 0.05
	V_{cut}	2.85 ± 0.07	-0.15 ± 0.02
	α	1.58 ± 0.77	
	β	-0.77 ± 0.02	
	χ^2	$1555.6/1039 = 1.49$	
		Satellite haloes	
z		p^0	p^1
0	A	-1.66 ± 0.01	
	V_{cut}	2.69 ± 0.01	
	α	1.57 ± 0.02	
	β	0.36 ± 0.02	
	χ^2	$37.6/185 = 0.20$	
$0 \le z \le 1$	A	-1.67 ± 0.07	-0.62 ± 0.08
	V_{cut}	2.71 ± 0.05	-0.14 ± 1
	α	1.626 ± 0.08	
	β	-0.48 ± 0.01	
	χ^2	$591.8/1081 = 0.54$	
$1 \le z \le 2.3$	A	-1.45 ± 0.08	-0.63 ± 0.05
	V_{cut}	2.53 ± 0.05	-0.14 ± 0.03
	α	1.23 ± 0.12	
	β	0.03 ± 0.11	
	χ^2	$274.0/470 = 0.58$	

This paper has been typeset from a TEX/LATEX file prepared by the author.