Simulating the Universe with MICE: the abundance of massive clusters

Martín Crocce, Pablo Fosalba, Francisco J. Castander and Enrique Gaztañaga

Institut de Ciències de l’Espai, IEEC-CSIC, Campus UAB, Facultat de Ciències, Torre C5 par-2, Barcelona 08193, Spain

ABSTRACT

We introduce a new set of large N-body runs, the Marenostrum Institut de Ciències de l’Espai (MICE) simulations, that provide a unique combination of very large cosmological volumes with good mass resolution. They follow the gravitational evolution of \( \sim 8.5 \) billion particles \((2048^3)\) in volumes covering up to \( \sim 15 \) Hubble volumes \( (450 \, h^{-3} \, \text{Gpc}^3) \), and sample over five decades in spatial resolution. Our main goal is to accurately model and calibrate basic cosmological probes that will be used by upcoming astronomical surveys of unprecedented volume. Here, we take advantage of the very large volumes of MICE to make a robust sampling of the high-mass tail of the halo mass function (MF). We discuss and avoid possible systematic effects in our study, and do a detailed analysis of different error estimators. We find that available fits to the local abundance of haloes match well the abundance estimated in the large volume of MICE up to \( M \sim 10^{14} \, h^{-1} \, M_\odot \), but significantly deviate for larger masses, underestimating the MF by 10 per cent (30 per cent) at \( M = 3.16 \times 10^{14} \, h^{-1} \, M_\odot \) \((10^{15} \, h^{-1} \, M_\odot)\). Similarly, the widely used Sheth & Tormen fit, if extrapolated to high redshift assuming universality, leads to an underestimation of the cluster abundance by 30, 20 and 15 per cent at \( z = 0, 0.5, 1 \) for fixed \( v = \delta_c/\sigma \approx 3 \) (corresponding to \( M \sim [7 - 2.5 - 0.8] \times 10^{14} \, h^{-1} \, M_\odot \), respectively). We provide a recalibration of the MF over five orders of magnitude in mass \([10^{10} < M/\, h^{-1} \, M_\odot < 10^{15}]\), that accurately describes its redshift evolution up to \( z = 1 \). We explore the impact of this recalibration on the determination of the dark energy equation of state \( w \), and conclude that using available fits that assume universal evolution for the cluster MF may systematically bias the estimate of \( w \) by as much as 50 per cent for medium-depth \((z \lesssim 1)\) surveys. The halo catalogues used in this analysis are publicly available at the MICE webpage, http://www.ice.cat/mice.

Key words: methods: N-body simulations – galaxies: clusters: general – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Near future extragalactic surveys will sample unprecedentedly large cosmological volumes, of the order of tens of cubic gigaparsecs, by combining wide fields with deep spectroscopy or photometry, typically reaching \( z \sim 1 \) \( \text{(e.g. DES, PAU, BOSS, PanSTARRS, WiggleZ)}\). In addition, they will be able to capture very faint objects and lower their shot-noise level to become close to sampling variance limited. Optimizing the preparation and scientific exploitation of these upcoming large surveys requires accurate modelling and simulation of the expected huge volume of high-quality data. This is quite a non-trivial task, because it involves simulating a wide dynamic range of cosmological distances in order to accurately sample the large-scale structure, and well resolved dark matter haloes as a proxy to galaxy clusters, to model the physics of galaxy formation and other non-linear physics.

Over the past decades numerical simulations have provided one of the most valuable tools to address these issues, and their relevance will certainly increase in the near future. They allow us to follow the growth of cosmological structure, shed light on the process of galaxy formation, model non-linear effects entering different clustering measures, lensing and redshift distortions, track the impact of a dark energy component and more. Among projects related to the development of very large simulations are those carried out by the Virgo consortium \( \text{(Frenk et al. 2000)} \), the Millennium I and II \( \text{(Springel et al. 2005; Boylan-Kolchin et al. 2009)} \), the Horizon run \( \text{(Kim et al. 2009) and the Horizon project (Teyssier et al. 2009).} \) They have all benefited from the vast computational power and hardware developed over the past years.

In this paper, we present a new effort to tackle the demand of large simulations and mock catalogues, the MareNostrum Institut de Ciències de l’Espai (MICE) simulations, that aims at the
development of very large and comprehensive \( N \)-body runs to deliver an unprecedented combination of large simulated volumes with good mass resolution.

As a first step, we have developed two \( N \)-body simulations including more than eight billion particles (\( 2048^3 \)) each, in volumes similar and well beyond the one corresponding to the Hubble length (\( \sim 30 \, h^{-1} \text{Gpc}^3 \)), in addition to several other large runs of typically smaller volume and corresponding higher mass resolution that are complementary to the large volume runs.

Some of the MICE simulations used in this paper have already been used to develop the first full-sky weak-lensing maps in the light cone (Fosalba et al. 2008), or study the clustering of luminous red galaxies (LRG) with multiple-band photometric surveys such as PAU (Benitez et al. 2009). More recently, using the largest volume simulations, a series of papers has studied the large-scale clustering in the spectroscopic LRG Sloan Digital Sky Survey (SDSS) sample, through the redshift space distortions (Cabrè & Gaztañaga 2009a,b), the baryon oscillations in the three-point function (Gaztañaga et al. 2008a), and in the radial direction (Gaztañaga, Cabre & Hui 2008b).

In our paper, we will focus on the mass function (MF) of the most massive objects formed through hierarchical clustering, since their low abundance makes the need of large sampling regions crucial. In turn, a precise description of this regime is of paramount importance since the abundance of clusters, to which it corresponds, is very sensitive to cosmological parameters (particularly the matter density), the normalization of the matter power spectrum and the expansion history of the universe, characterized by the dark energy density and its equation of state (e.g. see Mantz et al. 2008; Cunha 2009; Henry et al. 2009; Rozo et al. 2009; Vikhlinin et al. 2009, and references therein). This regime is also one of the best probes to search for primordial non-Gaussianities (e.g. Matarrese, Verde & Jimenez 2000; Grossi et al. 2007; Maggiore & Riotto 2009; Pillepich, Porciani & Hahn 2010).

The halo MF and related topics have been extensively studied in the literature. Analytic models predicting not only the abundance as a function of mass but also the evolution were developed as far back as the 1970s by Press & Schechter (1974) and followed by Bond et al. (1991), Lee & Shandarin (1998) and Sheth, Mo & Tormen (2001). However, the development of \( N \)-body simulations showed that these predictions were in general not sufficiently accurate for cosmological applications, and demanded the need for calibrations against numerical results (see also Robertson et al. 2009). Although \( N \)-body studies were likewise developed early (e.g. two 1000 particle simulations by Press & Schechter 1974), the reference work in this direction was set by Sheth & Tormen (1999) and Jenkins et al. (2001). More recent recalibrations of the MF to within few per cent were put forward by Warren et al. (2006) and Tinker et al. (2008). In addition, these or other papers have focused their attention on the redshift evolution of the MF (Reed et al. 2003, 2007; Lukic et al. 2007; Cohn & White 2007), different definitions of halo and the redshift evolution of the MF (Reed et al. 2003, 2007; Lukic et al. 2007; Tinker et al. 2008), with some dependence on the linking length (Jenkins et al. 2001; Tinker et al. 2008). Given the complicated non-linear processes that drive halo formation, having a universal MF (i.e. independent of redshift and cosmology) at this level could be very convenient, alleviating the need to simulate individual cosmologies. In contrast, the abundance of SO haloes is considerably less universal, showing evolution in the halo MF amplitude at the level of 20–50 per cent, depending on cosmology, but also evidence for redshift dependence of its overall shape (Tinker et al. 2008). The current error budget of cluster science is dominated by the unknowns in the calibration of mean and scatter of the mass–observable relations (e.g. Voit 2005; Rozo et al. 2009; Stanek et al. 2009b). Only after understanding these better will we be able to state what is an acceptable degree of non-universality in future precision cosmology (but see also Stanek et al. 2009a).

For concreteness, we will next focus on FoF haloes and leave SO for a follow-up study, although we have tested to some extent that our main results are robust in front of this choice, as we further discuss in our Conclusions.

This paper is organized as follows: in Section 2, we describe the MICE simulations. Section 3 recapitulates known theoretical predictions and fits to the halo MF and concludes with a comparison between MICE and results from previous simulations. In Section 4, we discuss systematic effects that are most relevant in the measurement of the high-mass end of the MF, such as transients from initial conditions, finite sampling of the mass distribution and mass resolution effects. A detailed error analysis including different estimators is provided in Section 6, whereas, in Section 7, we derive a new fitting function to account for the high-mass tail of the halo MF. The higher redshift evolution, including results regarding MF universality, is the subject of Section 8. We discuss the implications of our results for dark energy constraints in Section 9.
and we finish by summarizing and discussing our main findings in Section 10.

2 THE MICE SIMULATIONS

The set of large N-body simulations described in this paper were carried out on the Marenostrum supercomputer at the Barcelona Supercomputing Center (http://www.bsc.es), hence their acronym MICE (Marenostrum Institut de Ciències de l’Espai).

All simulations were run with the GADGET-2 code (Springel 2005) assuming the same flat concordance Λ cold dark matter (ΛCDM) model with parameters Ω_m = 0.25, Ω_b = 0.75, Ω_k = 0.044 and h = 0.7. The linear power spectrum had spectral tilt n_s = 0.95 and was normalized to yield σ_8 = 0.8 at z = 0. Special care was taken in order to avoid spurious artefacts from the initial conditions (transients). Thus, the initial particle distributions were laid down using either the Zeldovich approximation (ZA) with a high starting redshift or second order Lagrangian Perturbation Theory (2LPT) (Scoccimarro 1998; Crocce, Pueblas & Scoccimarro 2006) (see Section 4.1 for details).

The main goal of the MICE set is to study the formation and evolution of structure at very large scales, with the aim of simulating with enough mass resolution the size of future large extragalactic surveys, such as DES (Annis et al. 2005) or PAU (Benitez et al. 2009), and test robustly statistical and possible systematic errors. Fig. 1 shows the set of MICE simulations in the mass resolution–volume plane. They sample cosmological volumes comparable to the SDSS main sample (0.1 h^−3 Gpc^3), the SDSS-LRG survey (1 h^−3 Gpc^3), PAU or DES (9 h^−3 Gpc^3), and those of huge future surveys such as EUCLID (~100 h^−3 Gpc^3), in combination with mass resolutions from 3 × 10^12 h^−1 M⊙ down to 3 × 10^9 h^−1 M⊙. In turn, the largest volume simulations (squares) map the MF at the high-mass end, ~10^{15} h^−1 M⊙, whereas the test simulations (triangles) extend the dynamic range down to haloes of 10^{10} h^−1 M⊙. Table 1 summarizes the identifying parameters of the main MICE simulations.

Note that for one particular case (MICE1200), we implemented a set of 20 independent realizations, in order to compare statistical errors on different quantities obtained from a strictly ‘ensemble error’ approach from other internal or external error estimates.

In addition to the production of comoving outputs at several redshifts, we have constructed projected density and weak lensing maps as well as light-cone outputs from the main MICE runs. The mass projected and lensing catalogues were discussed in Fosalba et al. (2008), while the light-cone catalogue will be presented in future work. Note that both represent a huge compression factor (~1000) that may turn out to be essential in dealing with very large numbers of particles as in our case. Further details and publicly available data can be found at http://www.ice.cat/mice.

3 THE HALO MASS FUNCTION

The very large simulated volume spanned by the MICE set allows us to study accurately not only Milky Way size haloes, but specially the most massive and rarest haloes formed by hierarchical clustering. To this end, we built dark matter halo catalogues at each snapshot of interest according to the FoF algorithm (Davis et al. 1985) with linking length parameter b set in units of the mean interparticle distance in each simulation. We will refer to haloes defined in this way as FoF(b). For the most part, we will deal with the b = 0.2 catalogues, although we have also implemented b = 0.164 for a first validation of our simulations against the Hubble volume simulation (HVS) (Jenkins et al. 2001; Evrard et al. 2002). The HVS is one of the very few publicly available halo catalogues comparable in simulated volume to MICE.

The halo finder algorithm was implemented using the FoF code publicly available at the N-body Shop (http://www-hpcc.astro.washington.edu/), with some additional modifications needed in order to handle the large number of particles in a reasonable amount of time. The resulting halo catalogues contain not only the mass, position and velocity of the centre of mass, and virial velocity, but also information on all the particles forming each halo.

As an example of the size of our outputs, we mention that MICE3072 contains at z = 0 a total of about 25 million haloes more massive than 3.9 × 10^{12} h^−1 M⊙ if the minimum number of particles per halo is set to 20. The most massive object weighs 5.27 × 10^{15} h^−1 M⊙ and is made of 22,561 particles. In turn, MICE7680 contains about 15 million haloes with mass greater than 7.3 × 10^{13} h^−1 M⊙, with the biggest reaching 8.4 × 10^{13} h^−1 M⊙.

3.1 Theoretical predictions

Let us start by recalling some well-known results regarding the abundance of haloes. The differential MF is defined as

\[ f(\sigma, z) = \frac{M}{\rho_0} \frac{d n(M, z)}{d \ln \sigma^{-1}(M, z)} \]

(1)

where n(M, z) is the comoving number density of haloes with mass M and \( \sigma(M, z) \) is the variance of the linear density field smoothed with a top hat filter of radius R and enclosing an average mass \( M = \rho_0 R^3 / 3 \),

\[ \sigma^2(M, z) = \frac{D^2(z)}{2\pi^2} \int k^2 P(k)W^2(kR) dk, \]

(2)
Table 1. Description of the MICE N-body simulations.

| Run       | \(N_{\text{part}}\) | \(L_{\text{box}}/h^{-1}\text{Mpc}\) | \(m_p/h^{-1}\text{M}_\odot\) | \(l_{\text{soft}}/h^{-1}\text{kpc}\) | IC | \(z_i\) |
|-----------|----------------------|----------------------------------|-------------------------------|----------------------|----|-------|
| MICE7680  | 2048\(^3\)          | 7680                             | \(3.66 \times 10^{12}\)       | 50                    | ZA | 150   |
| MICE3072  | 2048\(^3\)          | 3072                             | \(2.34 \times 10^{11}\)       | 50                    | ZA | 50    |
| MICE4500  | 1200\(^3\)          | 4500                             | \(3.66 \times 10^{12}\)       | 100                   | 2LPT | 50   |
| MICE3072LR\(^\star\) | 1024\(^3\)    | 3072                             | \(1.87 \times 10^{12}\)       | 50                    | ZA | 50    |
| MICE768\(^\star\)  | 1024\(^3\)         | 768                             | \(2.93 \times 10^{10}\)       | 50                    | 2LPT | 50   |
| MICE384\(^\star\)  | 1024\(^3\)         | 384                             | \(3.66 \times 10^{9}\)        | 50                    | 2LPT | 50   |
| MICE179\(^\star\)  | 1024\(^3\)         | 179                             | \(3.70 \times 10^{8}\)        | 50                    | 2LPT | 50   |
| MICE1200\(^\star\) (\(\times 20\)) | 800\(^3\)       | 1200                             | \(2.34 \times 10^{11}\)       | 50                    | ZA | 50    |

Note: \(N_{\text{part}}\) denotes number of particles, \(L_{\text{box}}\) is the box-size, \(m_p\) gives the particle mass, \(l_{\text{soft}}\) is the softening length, IC is the type of initial conditions (ZA or 2LPT), and \(z_i\) is the initial redshift of the simulation. Their cosmological parameters were kept constant throughout the runs (see text for details), the initial global time-step is of the order 1 per cent of the Hubble time (i.e. \(d \log a = 0.01\), being \(a\) the scale factor), and the number of global time-steps to complete the run \(N_{\text{steps}}\) \(\gtrsim\) \(2000\) in all cases. We ran an ensemble of 20 different realizations with the parameters of MICE1200 primarily to calibrate error estimators. We mark with \(\star\) those runs that were done for completeness or testing as main purpose.

with

\[ W(x) = \frac{3}{(x^3)}[\sin(x) - x \cos(x)] \]

In equation (2), \(D(z)\) is the linear growth factor between \(z = 0\) and the redshift of interest, and \(P(k)\) the linear power spectrum of fluctuations at \(z = 0\).

In equation (1), we have explicitly assumed that all the cosmology dependence of the differential MF enters through the amplitude of linear fluctuations, equation (2), at the mass scale \(M\). If the redshift dependence also satisfies this condition the halo abundance as defined by equation (1) is said to be universal (Press & Schechter 1974; Sheth & Tormen 1999; Jenkins et al. 2001).

Several analytical derivations (Press & Schechter 1974; Bond et al. 1991; Sheth et al. 2001) or fits (Sheth & Tormen 1999; Jenkins et al. 2001; Evrard et al. 2002) have been provided in the literature over the past years, starting from the original Press–Schechter formalism in 1974 (Press & Schechter 1974). In our work, we refer only to the Sheth and Tormen (ST) fit given by Sheth & Tormen (1999).

\[
\frac{dM}{d\ln M} = \frac{M}{L_{\text{box}}} \frac{\Delta N}{\Delta M},
\]

that is then related to the differential MF in equation (1) after multiplying by the prefactor \(-\sigma/\sigma' \rho_b\). However, for the most part, we will directly compare the number density of haloes in a given mass bin with the prediction binned in the same way as the measurements.

That is, from equation (1), we obtain the theoretical number density of objects per unit mass \(n_M\), which is then integrated as

\[
M_{\text{bin}} = \int_{M_1}^{M_2} \frac{dM}{d\ln M} M dM = \int_{M_1}^{M_2} \frac{\rho_b}{M} \frac{d\sigma}{dM} f(M, z) dM
\]

3.2 The binned mass function

In order to compare the predictions for the differential MF in equation (1) with the observed halo abundance in a simulation of volume \(L_{\text{box}}^3\), one would measure the number of haloes \(\Delta N\) in a given mass bin \([M_1 - M_2]\) of width \(\Delta M\) and characteristic mass \(M\), and define

\[
\frac{dn}{d\ln M} = \frac{M}{L_{\text{box}}} \frac{\Delta N}{\Delta M},
\]

with \(A = 0.7234, a = 1.625, b = 0.2538, c = 1.1982\), obtained from a fit to the mass range \((10^{10} - 10^{15}) h^{-1}\text{M}_\odot\) at \(z = 0\). We will use equation (5) as our benchmark reference fit.

3.3 Comparison with previous work

As a first validation of our set of large volume simulations, we compared the halo abundance in MICE3072 to that in the HVS (Jenkins et al. 2001; Evrard et al. 2002), since they both simulate almost the same total volume. The HVS is among the biggest...
simulated volumes with halo data publicly available,\(^2\) at http://www.mpa-garching.mpg.de/Virgo/hubble.html.

We used the catalogue corresponding to a \(\Lambda\)CDM cosmology and FoF haloes with linking length parameter \(b = 0.164\) (Jenkins et al. 2001). Thus, in what follows, we will refer to the MICE catalogues for this value of \(b\). Finally, for this comparison, we employed the low-resolution run of MICE3072 (MICE3072LR) in Table 1 that has a similar mass resolution to that in the HVS.

Fig. 2 shows the ratio of the MFs measured in the HVS and MICE3072LR at \(z = 0\). The higher abundance of massive haloes in the HVS is due mainly to its larger value of \(\sigma_s\) equal to 0.9 against 0.8 in MICE. Therefore, we include in the figure the expected value for this ratio as predicted by the Jenkins fit in equation (4). The difference between symbols and the prediction are within the claimed accuracy for the Jenkins fit. In all cases, we show only haloes with no less than 50 particles, and Poisson errors.

\(4.1\) Transients from initial conditions

Several potential sources of systematic errors must be considered and controlled when implementing an \(N\)-body run, with their relevance sometimes dictated by the regime in which one is interested (see Lukic et al. 2007 for a detailed analysis). We have performed convergence tests regarding force and mass resolution, initial time steeping, finite volume effects and more. But of particular relevance to the abundance of the largest haloes at a given output is the initial redshift and the approximate dynamics used to set the initial conditions and start the run (Scoccimarro 1998; Crocce et al. 2006; Tinker et al. 2008; Knebe et al. 2009).

The generally adopted way to render the initial mass distribution is to displace the particles from a regular grid or a glass mesh, using the linear order solution to the equations of motion in Lagrangian Space. This is known as ZA. Particle trajectories within the ZA follow straight lines towards the regions of high initial overdensity. The ZA correctly describes the linear growth of density and velocity fields in Eulerian Space but, failing to account for tidal gravitational forces that bend trajectories, underestimates the formation of non-linear structure. To leading order this can be incorporated using the second order solution in 2LPT effectively reducing the time it takes for the correct gravitational evolution to establish itself (known as transients) once the \(N\)-body started. During the period where transients are present, the abundance of the most massive objects, that originate from the highest density peaks, is systematically underestimated.

In Crocce et al. (2006), it is shown that transients affect the \(z = 0\) mass function reducing it by 5 per cent at \(10^{15} h^{-1} M_{\odot}\) if ZA, as opposed to 2LPT, is used to start at \(z_i = 49\). This value rises to 10 per cent for \(M > 2 \times 10^{14} h^{-1} M_{\odot}\) at \(z = 1\). Also Tinker et al. (2008) finds evidence for transients in the Hased-Oct-Tree (HOT) runs introduced in Warren et al. (2006) and the HVS (Jenkins et al. 2001). These runs were started in the redshift range \(z = 24\) to 35 using ZA. However, their own run with \(z_i = 60\) is in good agreement with the 2LPT predicted abundance from Crocce et al. (2006) by \(\sim 1.25\). The impact of the starting redshift in the high-redshift MF has been investigated in Lukic et al. (2007) and Reed et al. (2003, 2007).

To test the significance of transients ourselves, we implemented two runs of MICE1200 (\(L_{\text{box}} = 1200 h^{-1}\) Mpc and 800\(^\prime\) particles) using ZA and 2LPT, both with \(z_i = 50\), and the same initial random phases (not listed in Table 1). Fig. 3 shows the ratio of the measured MFs, \(n_{ZA}/n_{2\text{LPT}}\), at \(z = 0\) (top panel), \(z = 0.5\) (middle) and \(z = 1\) (bottom). Top and bottom panels show in addition the results obtained by Crocce et al. (2006) for the same combination of \(\{z, z_i\}\) with a solid line. The dash line in the middle panel corresponds to a simple second order polynomial fit to the ratio at \(z = 0.5\). Our results agree very well with those in Crocce et al. (2006) despite the difference in cosmology of the \(N\)-body runs (most notably \(\sigma_s\)), confirming an underestimation of the halo abundance by \(\sim 5\) per cent at \(M \sim 10^{15}, 3.16 \times 10^{14}\) and \(10^{14} h^{-1} M_{\odot}\) at \(z = 0, 0.5\) and 1, respectively (and larger for larger masses, at fixed redshift), if ZA \(z_i = 50\) is used instead of starting at higher redshift or using 2LPT.

In line with the results above, almost all our runs in Table 1 were started using 2LPT at \(z_i = 50\) to avoid transients in the low-redshift outputs. The convergence of 2LPT with \(z_i \sim 50\) is discussed in detail in Croce et al. (2006). For MICE7680, we implemented ZA at high starting redshift (\(z_i = 150\)) to minimize transients. In this case, the convergence is assured by the results in Fig. 6, where its halo abundance is compared with the one in MICE4500, that was started with 2LPT at \(z_i = 50\) with completely different random phases.
In what follows, we will therefore correct the MF measured in MICE3072 by a simple fit to the ratios shown in Fig. 3.

### 4.2 FoF mass correction

As noted by Warren et al. (2006), the mass of haloes determined using the FoF algorithm suffers from a systematic overestimation due to the statistical noise associated with sampling the mass density field of each halo with a finite number of particles. By systematically subsampling an $N$-body simulation and studying the associated FoF($0.2$) halo abundance (keeping the linking length parameter fixed), Warren et al. (2006) determined an empirical correction of the mass bias that depends solely on the number of particles $n_p$ composing the halo through the simple expression

$$n_p^{corr} = n_p \left(1 - n_p^{-0.6}\right).$$

However, as remarked by Lukic et al. (2007), the correction should be checked on a case-by-case basis since it is not the result of a general derivation (see also Tinker et al. 2008). While it is true that for well sampled haloes the correction is relatively small (e.g., $2.5$ per cent for haloes with $500$ particles), the impact that a few per cent correction to the mass has in the halo abundance can be large if one refers to the most massive haloes living in the rapidly changing high-mass tail of the MF, as we are investigating in this paper.

For this reason, we have carried out an independent check of the correction in equation (9), with particular emphasis on the regime $M > 10^{13 - 14} h^{-1} M_\odot$, where the MF is exponentially suppressed.

We randomly subsampled every simulation in the MICE set to several degrees ($1$ every $2$, $4$ and $8$ particles) and ran the FoF algorithm afterwards keeping the linking length parameter $b = 0.2$ fixed (i.e., with the link length $n^{1/3}$ larger in each case). Results are shown in Fig. 4 for two representative cases, MICE768 and MICE7680, but they extrapolate to all others in Table 1.

The correction in equation (9) is able to bring the subsampled MFs into agreement with the original fully sampled one over the whole dynamic range (up to $4 \times 10^{15} h^{-1} M_\odot$). Most notably, in the case of MICE7680, whose MF is very sensitive to such halo mass corrections because it samples the exponential tail with low Poisson shot-noise but with haloes of no more than $\sim 2300$ particles. Finally, we have also tested that varying the factor $0.6$ leads to worse matching.

Hence, in what follows, we will refer to the abundance of mass corrected FoF haloes, unless otherwise stated.

### 4.3 Mass resolution effects

For a first glimpse of the abundance of massive objects in MICE, we display in Fig. 5 the MF of FoF($0.164$) haloes obtained from our largest runs (in terms of simulated volume), including the corrections for mass and abundance discussed above in Sections 4.1 and 4.2.

MICE3072 (blue squares) is in good agreement with the Jenkins prediction over more than two decades in mass, overlapping with the results from the larger MICE7680 (green circles) and MICE4500 (red triangles) for $M > 10^{14 - 15} h^{-1} M_\odot$.

However, as we transit towards the high-mass end ($M \gtrsim 10^{15} h^{-1} M_\odot$), the abundance in the grand sampling volume of MICE7680 rises over the one in MICE3072 (and HVS) reaching a $20$ per cent difference. In addition, measurements in MICE4500 (red triangles) are in very good agreement with those in MICE7680.
even though these runs correspond to completely different initial conditions, softening length, box-size, etc. (see Table 1).

We recall that MICE7680 and MICE4500 have roughly the same mass resolution as that in the HVS, but a volume 16.7 and 3.4 times larger, respectively. In turn, MICE3072 has roughly the box-size of the HVS, but eight times better mass resolution.

To check that the ‘excess’ abundance at large masses is not an artefact due to poor mass resolution, we have included in Fig. 5 the MF measured in MICE3072LR and in a very low mass resolution run not listed in Table 1 ($L_{\text{box}} = 3072 h^{-1} \text{Mpc}$, $N_p = 512$ and $m_p = 1.5 \times 10^{13} h^{-1} \text{M}_\odot$). They both agree remarkably well with MICE3072 at $M \gtrsim 10^{15} h^{-1} \text{M}_\odot$, showing that the abundance we found using MICE7680 and MICE4500 is robust to mass resolution effects, once the Warren et al. (2006) correction is taken into account.

In summary, we have tested the robustness of our results in front of several possible systematic effects. Early starting redshifts or inappropriate initial dynamics can affect the high-end MF at the several per cent level. Hence, the MICE simulations used either the 2LPT dynamics or ZA with high starting redshift ($z \sim 150$), with the exception of MICE3072 whose MF we none the less correct as described in Section 4.1. In addition, halo masses were computed using the correction of Warren et al. (2006), that we independently tested focusing on the regime of very massive haloes, in order to avoid a bias towards higher masses at low particle number. Lastly, we studied whether the rather high particle mass of MICE7680 and MICE4500 can impact the abundance of massive haloes by implementing a set of runs with fixed (large) volume and decreasing mass resolution. The abundances of massive haloes in these runs are in very good agreement once the Warren et al. (2006) correction is considered, reinforcing the robustness of our MF measurements in front of mass resolution effects.

## 5 The abundance of FoF(0.2) haloes

Let us now turn to the MF measurements in our catalogues of FoF(0.2) haloes. Fig. 6 shows the measured MF in the MICE simulations tabulated in Table 1. We display the ratios to the Sheth & Tormen fit in equation (3), binned in the same way as the measurements. The top panel corresponds to masses corrected for the}

![Figure 5](https://example.com/fig5.png)

**Figure 5.** Mass resolution effects on the high-mass end: we show the abundance of FoF(0.164) haloes at $z = 0$ in MICE3072 in blue squares, MICE4500 in red triangles and MICE7680 in green circles. We find a systematic rise over the Jenkins prediction for $M \gtrsim 10^{15} h^{-1} \text{M}_\odot$. The low-resolution simulation MICE3072LR (solid line) and the very-low resolution simulation MICE3072CR (dashed line) are evidence that our results are robust to mass resolution effects at large masses.

![Figure 6](https://example.com/fig6.png)

**Figure 6.** The MICE MF at $z = 0$, measured after combining data from the set of MICE simulations with box-sizes $L_{\text{box}} = 179, 384, 768, 3072, 4500$ and $7680 h^{-1} \text{Mpc}$ (in red, green blue, magenta, sea-green and black, respectively) and varying mass resolutions (see Table 1 for further details). The top panel corresponds to the measured MF after applying the correction to the FoF mass as described in Warren et al. (2006). In the bottom panel, we do not include this correction. In both figures, we display the ratio to the Sheth & Tormen (1999) prediction and include the corresponding Warren & Jenkins fits for reference (solid and dashed lines). In each case, the low-mass end is set by requiring a minimum of 100 particles per halo while the high-mass end is set by requiring a relative error below 5 per cent (displayed error bars correspond to Poisson shot-noise).

FoF(0.2) bias as described in Warren et al. (2006) and discussed in Section 4.2 (equation 9). The bottom panel contains uncorrected MFs. In both panels, the solid line represents the Warren fit given in equation (5), while the dashed line corresponds to the Jenkins fit in equation (4). The corrected MFs agree very well with the Warren fit, but only up to $10^{14} h^{-1} \text{M}_\odot$. After that mass, there is a systematic underestimation of the halo abundance in MICE768 and MICE3072 that reaches 20 per cent at $M \sim 5 \times 10^{14} h^{-1} \text{M}_\odot$ (note that we show only points with relative Poisson error $\leq 5$ per cent). Part of this effect can be attributed to transients in the simulations used by Warren et al. (2006) to calibrate the high-mass end, as discussed in Crocce et al. (2006) and Tinker et al. (2008).

For larger masses, $M \gtrsim M^{15} h^{-1} \text{M}_\odot$, the underestimation of the Warren fit is even more severe, and grows monotonically with $M$. This in part might be due to volume effects: the abundance of haloes at the high-mass end is expected to be extremely low, of the order of $n_{\text{halo}}/h^{-3} \text{Mpc}^3 \lesssim 10^{-7}$ ($z = 0$), $10^{-8}$ ($z = 0.5$) and $10^{-9}$ ($z = 1$) at $10^{15} h^{-1} \text{M}_\odot$ (integrated over mass bins of $\Delta \log_{10} M = 0.1$). This means that already at moderate redshifts, $z = 0.5$, a simulation of $L_{\text{box}} = 3 h^{-1} \text{Gpc}$ will contain only about 300 haloes and therefore measuring the abundance of haloes will be subjected to large uncertainties, i.e. the expected (shot-noise) error will be already of the order of 6 per cent. By including larger volume simulations, such as MICE4500 ($L_{\text{box}} = 4.5 h^{-1} \text{Gpc}$) and MICE7680 ($L_{\text{box}} = 7.68 h^{-1} \text{Gpc}$), we are able to increase the number of haloes by up to a factor of $\sim 16$, thus decreasing the associated halo abundance uncertainties by a factor of $\sim 4$. As shown in Fig. 6 (top panel), results from both large volume simulations (MICE4500 & MICE7680) agree very well in the high-mass end ($M \gtrsim 3 \times 10^{14} h^{-1} \text{M}_\odot$). This agreement serves as a validation test for the implementation of each
of them as well as for the high-mass end result, given that these two simulations share the same particle mass but have different initial dynamics (2LPT versus ZA) and random phases.

6 ERROR ESTIMATION

In a rather general sense, the most common source of statistical error considered in theoretical studies of halo abundance is solely the shot-noise contribution (e.g. Jenkins et al. 2001; Warren et al. 2006; Reed et al. 2007; Lukic et al. 2007). The importance of considering sample variance in addition to Poisson shot-noise had been highlighted in Hu & Kravtsov (2003), where it is shown that it cannot be neglected in front of shot-noise for deriving precise cosmological constraints. Following this criterion, Tinker et al. (2008) recently used jackknife (JK) errors with the intention to account for both sampling variance at low mass and Poisson shot-noise at high ones.

To explore these considerations more deeply, we will dedicate this section to perform a detailed study of different methods to estimate the error or variance in MF measurements. One particular goal is to obtain well calibrated errors in order to implement an accurate fit that could improve the high-mass description of equation (5).

We will pay particular attention in comparing how internal errors (i.e. those derived using only the N-body for which the mean MFs are measured, such as JK) perform against external ones and theoretical predictions, depending on the mass regime and total simulated volume under consideration.

One of the internal methods that we implemented is JK resampling (Zehavi et al. 2005). For this, we divided the simulation volume under consideration into \( N_{\text{JK}} \) non-overlapping regions, and computed the halo number density in the full volume omitting one of these regions at a time. The variance (defined as the relative error squared) in the i-bin of the number density is then obtained as

\[
\sigma_{\text{JK}}^2 = \frac{1}{N_{\text{JK}}} \sum_{j=1}^{N_{\text{JK}}} \left( n_j - \bar{n}^{\text{JK}} \right)^2,
\]

where \( \bar{n}^{\text{JK}} \) is the mean number density of halos for that bin. In what follows, we will show results using \( N_{\text{JK}} = 5^3 \), but we have checked that the estimates have already converged with varying \( N_{\text{JK}} \).

Another internal method we considered was to assume that the haloes are randomly sampled and form a Poisson realization of the underlying number density field. In this case, it is given by

\[
\sigma_{\text{Poisson}}^2 = \frac{1}{N_i},
\]

where \( N_i \) is the number of haloes in the i mass bin.

For estimating the variance externally in a volume \( V \), we used an N-body of volume \( V_L \), with \( V_L \gg V \). We then divide \( V_L \) into several non-overlapping regions of volume \( V \) and measure the number density in each subvolume. This method, which we refer to as subvolumes, is similar in spirit to boot-strap sampling except that the subvolumes are not thrown at random and do not overlap. Thus, this method has the advantage of incorporating the effect of long-wavelength modes which are absent in the volume \( V \).

For example, for MF errors in MICE179, we divided MICE384 into eight subvolumes and MICE768 into 80 subvolumes. In this way, the best statistics for the error is achieved at the mid-to-high mass regime of MICE179 because the mass resolution of MICE768 does not allow us to test all the way down to \( M \sim 3.16 \times 10^{11} h^{-1} M_{\odot} \) (although MICE384 does). None the less both MICE384 and MICE768 leads to a very consistent error estimation in MICE179 for its whole dynamic range, showing no dependence on mass resolution. For MICE384 we divided MICE768 into eight subvolumes and MICE3072 into 512 subvolumes. The rest of the box-sizes follow this same logic, that is, their variance in the mean number density was obtained by analysing the next-in-volume runs as listed in Table 1.

Our last external method is ensemble average. This we can only apply to one box-size, \( L_{\text{box}} = 1200 h^{-1} \text{Mpc} \), using the ensemble of 20 independent realizations of MICE1200 as listed in Table 1.

Finally, to derive a theoretical estimate of the variance in the measured MF, consider fluctuations in the mean number density of haloes of a given mass, \( \bar{n}_h(M) \), as coming from two different sources (see Hu & Kravtsov 2003 for the original derivation). First, a term arising from fluctuations in the underlying mass density field \( \delta_m \), if we consider the halo number density to be a tracer of the mass. If this relation is simply linear and local, then \( \delta n_h(M, x)/\bar{n}_h = b(M \delta_m(x)) \), where \( b \) is the halo bias.

Secondly, a shot-noise contribution \( \delta n_{\text{sn}} \) due to the imperfection of sampling these fluctuations with a finite number of objects. This noise satisfies \( \langle \delta n_{\text{sn}} \rangle = 0 \) and is assumed to be uncorrelated with \( \delta_m \). Furthermore, if we assume the halo sample to be a Poisson realization of the true number density, this error becomes a simple Poisson white-noise with variance \( \delta n_{\text{sn}}^2 / \bar{n}_h^2 = 1/\bar{n}_h V = 1/N \), where \( N \) is the total number of objects sampled within the volume \( V \). Within these assumptions we then have

\[
\delta n_h(M, x) = b(M \bar{n}_h(M)) \delta_m(x) + \delta n_{\text{sn}}, \tag{12}
\]

The number density of objects of this mass within the simulation is estimated by

\[
n = \int_V d^3x W(x) \langle \bar{n}_h + \delta n_h \rangle, \tag{13}
\]

where \( W(x) \) is the simulation window function, normalized such that \( \int_V d^3x W(x) = 1 \). The variance of the number density measurements in the simulation is then given by

\[
\langle n^2 \rangle - \bar{n}_h^2 = \bar{n}_h^2 + b^2 \bar{n}_h^2 \int d^3x_i \int d^3x_j W(x_i) W(x_j) \times \langle \delta_m(x_i) \delta_m(x_j) \rangle, \tag{14}
\]

and can be cast as

\[
\sigma_n^2 = \frac{\langle n^2 \rangle - \bar{n}_h^2}{\bar{n}_h^2} = \frac{1}{\bar{n}_h V} + b_h^2 \int \frac{d^3k}{(2\pi)^3} |W(k)R|^2 P(k), \tag{15}
\]

where \( P(k) \) is the linear power spectrum of mass. For simplicity, we will assume the simulation window function \( W \) to be top-hat in real space (equation 3), with smoothing radius such that the volume equals the simulated one, i.e., \( R = (3V/4\pi)^{1/3} \).

The first term in equation (14) is the usual shot-noise contribution to the variance, that we introduced in a rather ad hoc way in equation (12), but it can also be derived in the context of the halo model as the contribution from the one-halo term (Takada & Bridle 2007). The second term, known as sampling variance, is the error introduced by trying to estimate the true number density using a finite volume.

As discussed in Section 3.2, one is in practice interested in the halo abundance within bins of mass range \( [M_1 - M_2] \) and characteristic mass \( M \). Thus, in equation (15), we will compute \( \bar{M} \) and \( \bar{n}_h \) from equations (7) and (8) and the bias as

\[
b_h = \int_{M_1}^{M_2} b_{\text{ST}}(M) dM, \tag{16}
\]

where \( b_{\text{ST}} \) is the prediction for the linear bias dependence on halo mass from Sheth & Tormen (1999),

\[
b_{\text{ST}}(M) = 1 + \frac{q^2/\sigma^2 - 1}{\delta_c} + \frac{2p/\delta_c}{1 + (q^2/\sigma^2)^p}, \tag{17}
\]

© 2010 The Authors. Journal compilation © 2010 RAS, MNRAS 403, 1353–1367
The abundance of massive clusters

with parameters \( q = 0.707 \) and \( p = 0.3 \) and \( \sigma = \sigma(M) \) given in equation (2). Equation (17) follows from considering variations of the unconditional MF in equation (3) with respect to the critical overdensity for collapse \( \delta_c \). Thus, strictly speaking this bias expression should be weighted by \( dn/dm \) from equation (3) when integrating equation (16) (Sheth & Tormen 1999; Manera, Sheth & Scoccimarro 2009). However, we have found that using a fit to the MICE MF instead leads to better agreement with measurements of clustering in the simulations. Accordingly, we will use this fit also to estimate \( M \) and \( h_n \), entering in equation (15).

Fig. 7 shows the result of our external, internal and theoretical error study. Clearly, the Poisson shot-noise dominates the error budget for \( M \gtrsim 10^{14} h^{-1} \, M_\odot \) corresponding to \( \sigma_{MF}/MF > 5 \) per cent (and \( \log \sigma^{-1} \lesssim 0.06 \)) at smaller masses, the sampling variance becomes increasingly important, rapidly dominating the total error (this is more so for smaller box-sizes). JK resampling does capture this trend but only partially, in particular for the smaller box-sizes (\( \lesssim 500 h^{-1} \, \text{Mpc} \)) where sampling variance from the absence of long-wavelength modes is more significant. This seems to indicate that the JK regions must have a minimum volume (e.g. while JK works well at \( 10^3 h^{-1} \, M_\odot \) in MICE768, it does not for MICE179 at the same mass). This is an important result to consider in further studies where JK resampling is used to improve upon Poisson shot-noise.

![Figure 7. Sampling variance versus shot-noise: the panels show different estimates for the variance in the halo MF within each of the six box-sizes used throughout this paper. Dot symbols correspond to Poisson shot-noise and are only shown for haloes with a minimum of 200 particles and until the relative error reaches 10 per cent (as used in Section 7). The dot-dashed blue line is the result of implementing JK resampling in the given box-size. Solid and dashed black lines are the subvolumes method described in the text that uses independent subdivisions of larger size boxes. For this method, we include cases in which two larger volume runs are used (panels of MICE179, MICE384 and MICE1200). Empty magenta squares are the theoretical estimate in equation (15) that includes both sampling and shot-noise variance. The variance is shot-noise dominated roughly for \( M > 10^{14} h^{-1} \, M_\odot \) (depending on the box-size). While JK does capture some sampling variance at smaller masses, it is not fully satisfactory. It can underestimate the error by a factor of 2 to 3 at \( 10^{12} h^{-1} \, M_\odot \). Notably, the subvolumes and theoretical estimates are in very good agreement across all box-sizes and full mass range.](https://academic.oup.com/mnras/article-abstract/403/3/1353/1048411)

The total error can be underestimated by a factor of a few, e.g. 3 (1.5) at \( 10^{12} h^{-1} \, M_\odot \) in MICE179.

On the other hand, the theoretical error from equation (15) is in remarkably good agreement with the subvolumes method described above across all box-sizes (not shown in MICE7680 and MICE3072 panels of Fig. 7 because it is indistinguishable from the other estimates). This can be taken as a cross-validation of these two methods.

In addition, we have tested how these different methods compare with the ensemble error in the MF obtained from the 20 independent runs of MICE1200. For the subvolumes estimation, we divided MICE7680 into 264 regions of volume almost identical to that of MICE1200 (as well as MICE3072 into 18 regions for consistency checks). The result is that, for \( L_{\text{box}} = 1200 h^{-1} \, \text{Mpc} \), the subvolume error is larger than the ensemble one by a factor of about 20 per cent at \( M \sim (10^{13} - 3.16 \times 10^{14}) h^{-1} \, M_\odot \). The reason is that each subvolume region ‘suffers’ fluctuations in the mean density caused by the long-wavelength modes present in the larger box-size from which they have been obtained (MICE7680 or MICE3072 in this case). These modes, that introduce an extra-variance, are absent in each of the ensemble members that satisfy periodic boundary conditions at the scale \( L_{\text{box}} = 1200 h^{-1} \, \text{Mpc} \). Thus, the ensemble error (from running different simulations) does not always work well because it can suffer from volume effects.

This conclusion can be nicely reinforced using equation (15), which performs very well here also. One can mimic the absence of long-wavelength modes by setting the low-\( k \) limit in the sampling variance integral in equation (15) to be the fundamental mode of MICE1200 (\( k_f = 2 \pi / 1200 h \, \text{Mpc}^{-1} \)). In this case, the theoretical model agrees with the ensemble error (in fact, \( \sigma_\delta \) turns to be mostly dominated by shot-noise). Instead, by setting it to \( k_f \approx 0 \) one recovers the subvolumes estimate.

In summary, the subvolume method should be considered as the one comprising all statistical uncertainties: shot-noise, sampling variance and volume effects (the fact that there are fluctuations in scales larger than the sample size). The theoretical estimate in equation (15) is consistent with it to a remarkably good level, for the masses and box-sizes tested in this paper. Hence, it is a powerful tool for studies involving the abundance of massive haloes (Rozo et al. 2009; Vikhlinin et al. 2009).

7 THE FITTING FUNCTION FOR THE ABUNDANCE OF MASSIVE HALOES

The accurate sampling of the MF requires a demanding combination of very big volumes and good mass resolution. As shown in Fig. 1, using the MICE set of simulations, we have sampled volumes up to \( 450 h^{-3} \, \text{Gpc}^3 \) with a wide range of mass resolutions yielding a dynamic range \( 10^8 < M/(h^{-1} \, M_\odot) < 10^{12} \) in particle mass (see also Table 1).

As shown in Fig. 6, the ST and Warren fits underpredict the abundance of the most massive haloes found in our \( N \)-body simulations for \( M > 10^{14} h^{-1} \, M_\odot \), although Warren gives accurate results for lower masses.

In this section, we derive new fits based on the MICE set of simulations, sampling the MF over more than five orders of magnitude in mass, and covering the redshift evolution up to \( z = 1 \). For this purpose, we use a set of MICE simulations with increasing volume and corresponding decreasing mass resolution, in order to sample the mass range from the power-law behaviour at low masses, \( M \sim 10^{10} h^{-1} \, M_\odot \), and up to the exponential cutoff at the high-mass end. As discussed in Section 5, the abundance of haloes
at the high-mass end is expected to be extremely low, of the order of $n_{\text{halo}} h^{-3} \text{Mpc}^{-3} \lesssim 10^{-7}$ ($z = 0$), $10^{-8}$ ($z = 0.5$) and $10^{-9}$ ($z = 1$) at $10^{15} h^{-1} \text{M}_{\odot}$ (integrated over mass bins of $\Delta \log_{10} M = 0.1$), and thus we shall use MICE4500 ($L_{\text{box}} = 4.5 h^{-1} \text{Gpc}$) and MICE7680 ($L_{\text{box}} = 7.68 h^{-1} \text{Gpc}$) in order to get a more accurate measurement of the halo abundance in this regime. In practice, we shall combine cluster counts from both simulations to get a more robust estimate. As it will be shown below, our fitting functions recover the measured MF over the entire dynamic range and its redshift evolution with $\sim 2$ per cent accuracy.

For our fitting procedure, we use the following simulations: MICE179, 384, 768, 3072, 4500 and MICE7680 (see Fig. 1 and Table 1) and match them so that in overlapping mass bins we shall adopt the abundance estimated from the simulation with the lower associated error, provided haloes include a minimum of 200 particles, except for the smallest box-size simulation, MICE179, for which we use down to 50 particle haloes. This is done in order to sample small enough haloes (as small as $10^{10} h^{-1} \text{M}_{\odot}$) whose abundance has been accurately measured in previous analyses (see e.g. Warren et al. 2006; Tinker et al. 2008) and that we aim at recovering as well as with our fitting functions.

On the other hand, for $M \gtrsim 3 \times 10^{14}$, we average results from both MICE4500 and MICE7680. As shown above, for $z = 0$, results for $M > 10^{15} h^{-1} \text{M}_{\odot}$ from MICE7680 are found to be in full agreement with those of MICE4500 despite the different initial conditions and time-step size used, which provides a robustness test to our measurements at the highest mass bins.

The fit to the MF is then determined using a diagonal $\chi^2$ analysis,

$$
\chi^2 = \sum_{i=1}^{N} \frac{(n_i^{(0)} - n_i^{(\text{bbody})})^2}{\sigma_i^{(0)^2}},
$$

where $n_i^{(0)} (n_i^{(\text{bbody})})$ is the theoretical ($N$-body) MF integrated over the $i$th logarithmic mass bin of width $\Delta \log_{10} M / (h^{-1} \text{M}_{\odot}) = 0.1$, and the errors $\sigma_i^{(0)}$ are computed using the JK estimator. The JK estimator is consistent with theoretical errors and subvolume dispersion (see Fig. 7), except for the smallest mass bins (sampled by MICE179, see lower right-hand panel in Fig. 7) for which the JK errors significantly underestimate other error estimates. By using JK errors for those bins, we just give them a larger statistical weight in the $\chi^2$ analysis, thus making sure the fit recovers the expected low-mass behaviour (i.e. we assume something similar to a low-mass prior).

We summarize the fitting results for the MICE MFs at $z = 0$ and 0.5 in Table 2. The best fit to the $N$-body measurements at $z = 0$ (‘MICE’ fit, solid line) is given by the Warren-like MF (equation 5), with parameters $A = 0.58$, $a = 1.37$, $b = 0.30$, $c = 1.036$ with $\chi^2 / \nu = 1.25$. As shown in Fig. 8, the fit recovers our $N$-body data to 2 per cent accuracy in practically all the dynamic range (i.e. for $10^{15} > M / h^{-1} \text{M}_{\odot} > 2 \times 10^{10}$). The Warren fit (dashed line) matches the $N$-body data to 3 per cent accuracy in the low-mass end $M / h^{-1} \text{M}_{\odot} < 10^{14}$, but it significantly underestimates the abundance of the most massive haloes: we find a 10 per cent (25 per cent) underestimate at $M / h^{-1} \text{M}_{\odot} \sim 3 \times 10^{14}$ (10^{15}$), and larger biases for more massive objects.

Table 2. MF best-fitting parameters for $f(\sigma, z)$, equation (5) and goodness of fit.

| $z$ | $A$ | $a$ | $b$ | $c$ | $\chi^2 / \nu$ |
|-----|-----|-----|-----|-----|---------------|
| 0   | 0.58| 1.37| 0.30| 1.036| 1.25          |
| 0.5 | 0.55| 1.29| 0.29| 1.026| 1.20          |

can be attributed in part to transients in the Warren et al. (2006) $N$-body data, as discussed in Section 5.

At higher redshifts, the MF deviates from the universal form in equation (5), and the fitting parameters change accordingly. In particular, for $z = 0.5$, we find the best-fitting values $A = 0.55$, $a = 1.29$, $b = 0.29$, $c = 1.026$, yielding a $\chi^2 / \nu = 1.20$. As seen in Fig. 9 (left-hand panel), this fit recovers the $N$-body measurements to 2 per cent accuracy in all the dynamic range (i.e. for $10^{15} > M / h^{-1} \text{M}_{\odot} > 2 \times 10^{10}$). The MICE fit at $z = 0$ extrapolated with the linear growth to $z = 0.5 (z = 1)$ overestimates the measurements by 3–6 per cent (10 per cent) as the halo mass increases. This in turn shows to what extent the FoF(0.2) MF deviates from universality.

8 HALO GROWTH FUNCTION

In the previous section, we found that the MF deviates significantly from universality (or self-similarity). Here, we investigate in detail the halo growth function, i.e. the evolution of the halo abundance with redshift, using the scaling evolution of the best-fitting parameters as a starting point.

The evolution of the fitting function parameters with redshift, as shown in Fig. 9, indicates that the FoF(0.2) MF, at least for the linking length $l = 0.2$, is non-universal at the 5–10 per cent level in the redshift range tested. This is in agreement with previous studies (Jenkins et al. 2001; Reed et al. 2003, 2007; Lukic et al. 2007; Tinker et al. 2008), but is consistently extended here to the high-mass regime, hard to sample robustly, particularly at higher redshifts.
To account for this evolution, we next try to fit the MF growth with a simple ansatz. If we follow Tinker et al. (2008) and assume that the fitting parameters are a simple function of the scale factor, \( a = 1/(1 + z) \), we can model the evolution as

\[
P(z) = P(0)(1 + z)^{-\alpha_i}; \quad P = [A, a, b, c]; \quad \alpha_i = [\alpha_1, \ldots, \alpha_4],
\]

where \( P(0) \) are the fitting parameters at \( z = 0 \), as given by Table 2. Therefore, we can use the lowest redshift measurements at \( z = 0 \) and 0.5 to determine the slope parameters \( \alpha_i \):

\[
\alpha_1 = 0.13, \quad \alpha_2 = 0.15, \quad \alpha_3 = 0.084, \quad \alpha_4 = 0.024.
\]

If the ansatz is correct, i.e. the growth of the MF can be modelled to a good approximation with equation (19), one should be able to predict the measured cluster evolution at higher redshifts. Using the values of \( \alpha_i \) as given above, we predict the following fitting parameters at \( z = 1 \): \( A = 0.53, a = 1.24, b = 0.28, c = 1.019 \), which gives a good match to simulations. \( \chi^2 / \nu = 1.92 \). As shown in both panels of Fig. 9, the scaling ansatz recovers the measured MF to 3 per cent accuracy in most of the dynamic range (i.e. for \( 3.16 \times 10^{14} > M / (h^{-1} \text{M}_\odot) > 2 \times 10^{13} \)).

We conclude from this that the ansatz equation (19) can be safely used to make predictions about the abundance of the most massive haloes at intermediate redshifts.

It has been argued by Tinker et al. (2008) that the non-universality of the MF is basically a consequence of the evolution of the halo concentrations, which in turn is mostly due to the change of the matter density \( \Omega_m \) with redshift and thus \( f(\sigma) \) should be rather modelled as a function of the linear growth rate of density perturbations, \( D(z) \). In order to test this hypothesis, we have repeated the analysis where the scaling of \( f(\sigma) \) is parametrized as follows:

\[
P(z) = P(0)(D(z)/D(0))^{\beta_i}; \quad \beta_i = [\beta_1, \ldots, \beta_4].
\]

In this case, the slope parameters are found to be \( \beta_1 = 0.22, \beta_2 = 0.25, \beta_3 = 0.14, \beta_4 = 0.04 \). Using this model, we estimate \( f(\sigma) \) parameters at \( z = 1 \) to be \( A = 0.52, a = 1.22, b = 0.28, c = 1.017 \), which provides a slightly worse fit to the N-body measurements, with \( \chi^2 / \nu = 2.85 \). Therefore, we find some evidence in favour of a scaling ansatz based on the scale factor with respect to that based on the growth rate. We note that our analysis is of limited validity since we have only considered one cosmology and one should explore a wider parameter space to draw stronger conclusions on this point.

In summary, the fit to the FoF(0.2) halo MF measured in the MICE simulations between redshift 0 and 1 is given by

\[
f_{\text{MICE}}(\sigma, z) = A(z) [\sigma^{-\alpha_i(z)} + b(z)] \exp \left[ -\frac{c(z)}{\sigma^2} \right] \quad (22)
\]

with \( A(z) = 0.58(1 + z)^{-0.13}, \alpha(z) = 1.37(1 + z)^{-0.15}, b(z) = 0.3(1 + z)^{-0.08}, c(z) = 1.036(1 + z)^{-0.021} \).

We can now explore how the halo growth function evolves in more detail. For this purpose, we study the halo MF integrated in wider logarithmic bins \( \Delta \log M / (h^{-1} \text{M}_\odot) = 0.5 \), and concentrate on the highest mass bins where we find the largest deviations between our N-body measurements and available fits. Fig. 10 (left-hand panel) shows the halo growth factor as measured in several comoving redshifts in the MICE simulations.

Our simulations show that the abundance of massive haloes drops by half (one) order of magnitude for the log \( M / (h^{-1} \text{M}_\odot) = 13.5 - 14.5 \) from \( z = 0 \) to \( 1 \), in rough agreement with analytic fits. However, as displayed in the right-hand panels of Fig. 10, the (self-similar) ST and Warren fitting functions (see short- and long-dashed lines, respectively) only match the high-mass end of the measured MF at \( z = 0.5 \) to 15 per cent accuracy. The Warren fit at \( z = 1 \) underestimates simulation data by up to 30 per cent. On the other hand, using the predicted halo abundance growth from the MICE fits at low redshift (red solid line) recovers the measured abundance at \( z = 1 \) to better than 1 per cent. We have used the scaling functions given by equation (19), however, we have checked that our results.
MF can be mistaken by the right cluster abundance for a different cosmology, the bias on $w$, for a given redshift, will be determined by the relative sensitivity on $w$ of the two cluster count probes of dark energy: the MF growth and the survey volume up to a given depth.

We perform a $\chi^2$ analysis to determine the bias as a function of survey depth by comparing halo counts in redshift shells as follows:

$$\chi^2 = \sum_{z_i} \frac{(n(w)^{\mu} - n(z)^{\nu}_{\text{Nbody}})^2}{\sigma^{\mu 2}},$$

where $n(w)^{\mu}$ are the counts from the assumed self-similar MF, for a cosmology with a given value of $w$, integrated from a minimum mass $M_{\text{min}}$ up to some maximum mass $M_{\text{max}}$, for the redshift shell $z_i$ of width $\Delta z_i$ within which we can safely consider the MF to be independent of $z$. In turn, $n(z)^{\nu}_{\text{Nbody}}$ are the corresponding counts measured in simulations that use the fiducial cosmology with $w = -1$, and that are accurately described by the scaling ansatz (equation 22). The associated error is assumed to be pure shot-noise given by the $N$-body counts, $\sigma = \sqrt{n_{\text{Nbody}}}$. Here, we make estimates for full-sky surveys (i.e. we will draw optimistic forecasts for a given survey depth), although smaller areas can be easily incorporated in our analyses by scaling the shot-noise error accordingly. We shall consider surveys with Sunyaev–Zeldovich (SZ) detected clusters, i.e. with a redshift independent mass threshold $M \sim 10^{14} h^{-1} M_\odot$, and constant photo-$z$ error $\Delta z = 0.1$. The upper limit in the mass is taken to be $10^{15.5} h^{-1} M_\odot$ to avoid possible systematic departures of the $N$-body fit used beyond this mass cut with respect to the simulation measurements. However, we have checked that our results do not change if we take larger mass cuts.

Fig. 11 shows the bias on $w$ as a function of survey depth $z$ for two different priors on self-similar MFs, the ST fit and the MICE fit (i.e. assuming the fit at $z = 0$ extrapolated to higher-$z$). We find that the estimated bias is robust to better than 20 per cent to changes in

9 COSMOLOGICAL IMPLICATIONS: BIAS ON DARK ENERGY CONSTRAINTS

The cluster MF is one of the standard cosmological probes used by current and proposed surveys to constrain cosmological parameters. In particular, the cluster abundance as a function of cosmic time is a powerful probe to determine the nature of dark energy. However, usage of the cluster abundance as a cosmological probe is limited by systematics in the mass–observable relations and the potential impact of priors (see e.g. Battye & Weller 2003a; Weller & Battye 2003b).

Here, we concentrate on the impact of priors in the MF on extracting the dark energy equation of state, $w$. As shown in Section 8, the halo MF measured in simulations deviates from self-similarity by as much as 15 per cent, depending on redshift and halo mass. We shall estimate how this systematic departure from universality can potentially bias estimates of $w$.

As a working case, we will consider a tomographic survey with photometric accuracy $\Delta z$, and estimate the shift in the recovered value of $w$, by including cluster counts in redshift shells up to a given depth, $z$. For simplicity, we consider a constant dark energy equation of state, although this same analysis could be easily extended to a time varying $w(z)$. Since the wrong prior on the halo

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure10.pdf}
\caption{Halo growth function: evolution of the halo abundance with redshift. We show the three most massive logarithmic bins (with $\Delta \log M/(h^{-1} M_\odot) = 0.5$), where we find the most significant deviations between N-body and available fits. Left-hand panel: the abundance of massive haloes drops by half (one order of magnitude for the log $M/(h^{-1} M_\odot) = 13.5-14$) from $z = 0$ to $z = 1$, in rough agreement with analytic fits. Right-hand panel: residuals between analytic fits (lines) to N-body measurements (symbols) for the same mass bins (increasing mass from top to bottom panels). ST and Warren fitting functions (short- and long-dashed lines, respectively) only match the high-mass end of the measured MF at $z = 0.5$ to 15 per cent accuracy. The extrapolated Warren fit to $z = 1$ under-estimates MICE data by up to 30 per cent. The predicted halo abundance growth from the MICE fits at low redshift (red solid line; see text for details) recovers the measured abundance at $z = 1$ to better than 1 per cent.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure11.pdf}
\caption{Bias on $w$ induced by self-similar prior on the MF. Upper panel: assuming the self-similar ST (MICE) fit induces up to 50 per cent (20 per cent) shift in the recovered value of $w$ for shallow or medium depth surveys $z \lesssim 1$, whereas the effect tends to be negligible for deep surveys. Bottom panel: the relative shot-noise error decreases with redshift, which makes high-$z$ cluster counts, which are dominated by the strong variation of the MF with $w$, down-weight biases induced by the low-$z$ counts.}
\end{figure}
the SZ mass threshold $M_{\text{min}}$, from $10^{13.8}$ to $10^{14.2}$ (see dashed lines in Fig. 11). For the ST fit, the bias can be as large as 50 per cent for survey depth $z < 1$, whereas for the MICE self-similar fit it reduces to 20 per cent at most for the same depth. The bias at low-redshift results from the relatively small and comparable sensitivities to changes in $w$ of the geometric (volume) and the shape of the MF growth. On the other hand, the strong bias at low-redshift for the ST prior is due to the poor fit it gives to $N$-body for the relevant masses $M \gtrsim 10^{14} h^{-1} \text{M}_\odot$. This systematic effect tends to decrease for increasing depth as the MF growth becomes a much stronger function of the dark energy equation of state than the redshift shell volume (for constant $\Delta z$), and thus it determines the cluster counts irrespective of the prior on the MF. This trend is strengthened by the fact that the shot-noise error per $z$-shell drops with depth as well (see bottom panel of Fig. 11), so the high-$z$ counts down-weight the observed low-$z$ bias on $w$. Therefore, for deep surveys $z \gtrsim 1$ any bias associated with the MF tends to be washed out.

10 DISCUSSION AND CONCLUSIONS

The abundance of clusters as a function of cosmic time is a powerful probe of the growth of large-scale structure. In particular, clusters provide us with a test of the sensitivity of the growth of structure to cosmological parameters such as the dark matter and dark energy density content, its equation of state or the amplitude of matter density fluctuations. More importantly, the abundance and clustering of clusters are complementary to other large-scale structure probes such as the clustering of galaxies, weak lensing and supernovae (see e.g. Hu & Kravtsov 2003; Cunha 2009; Cunha, Huterer & Frieman 2009; Estrada, Sefusatti & Frieman 2009).

In order to plan and optimally exploit the scientific return of upcoming astronomical surveys (e.g. DES, PAU, PanSTARRS, BOSS, WiggleZ, EUCLID) one needs to make accurate forecasts of the sensitivity of these probes to cosmological parameters in the presence of systematic effects. Accurate determination of the abundance of clusters is limited by our knowledge of the relation between cluster mass and observable (see e.g. Hu 2003). One key ingredient in this mass–observable calibration is the precise determination of the halo MF.

In this paper, we use a new set of large simulations, including up to 2048$^3 \approx 10^{10}$ particles, called MICE (see Table 1), to accurately determine the abundance of massive haloes over five orders of magnitude in mass, $M = (10^{10} - 10^{15.5}) h^{-1} \text{M}_\odot$, and describe its evolution with redshift. Note that the MICE simulations sample cosmological volumes up to $7.68^3 \approx 450 h^{-3} \text{Gpc}^3$. Armed with these simulations, we estimate, for the first time, the MF over a volume more than one order of magnitude beyond the HVS size, $3^3 = 27 h^{-3} \text{Gpc}^3$ (Jenkins et al. 2001; Evrard et al. 2002). Our results have been tested against mass resolution, choice of initial conditions and simulation global time-step size.

Our findings can be summarized as follows:

(i) We confirm previous studies (see Crocce et al. 2006) showing that an accurate determination of the MF is sensitive to the way initial conditions are laid out. For a fixed starting redshift, the usage of 2LPT instead of the ZA to set up the initial conditions helps to avoid the effects of transients that tend to artificially decrease the abundance of large-mass haloes. Alternatively, one can use ZA with a high enough starting redshift. In particular, we find an underestimate of the halo abundance by ~5 per cent at $M \sim 10^{15}, 3.1 \times 10^{14}$ and $10^{14}$ h$^{-1} \text{M}_\odot$ at $z = 0, 0.5$ and 1, respectively (and larger for larger masses, at fixed redshift), if ZA $z_i = 50$ is used instead of 2LPT. All our MFs were safe from transients or corrected accordingly for those simulations that need it (i.e. MICE3072, see Table 1).

(ii) As highlighted by Warren et al. (2006), the mass of haloes determined using the FoF algorithm suffers from a systematic overestimation due to the statistical noise associated with sampling the halo density field with a finite number of particles. The precise form of this correction has to be checked on a case by case basis. We confirm the results of Warren et al. (2006), particularly testing them for large halo masses, that the mass correction for the FoF(0.2) haloes follows equation (9).

(iii) For masses $M > 10^{14} h^{-1} \text{M}_\odot$, we find an excess in the abundance of massive haloes with respect to those from the HVS (Jenkins et al. 2001; Evrard et al. 2002), once corrected for the different cosmology used, that is a few times the Poisson error (see Fig. 5). As argued in Section 4.1, we conclude that this difference can be largely explained by systematics due to transients affecting the HVS.

(iv) From an extensive study of error estimates we conclude: (1) Sample variance is significant only for haloes of $M < 10^{14} h^{-1} \text{M}_\odot$, and dominates over shot-noise for box-sizes $L_{\text{box}} < 1 h^{-1} \text{Gpc}$. Consistently, Poisson shot-noise errors underestimate the total error budget by factors 2–4 at $10^{12} - 10^{13} h^{-1} \text{M}_\odot$ in these volumes. (2) JK resampling is in general consistent with external estimators such as ensemble averages, and theoretical errors. However, it underestimates the total error budget for small box-sizes ($<500 h^{-1} \text{Mpc}$) where sampling variance is more important due to the lack of long-wavelength modes and to the fact that JK regions are not independent. (3) For all our runs the theoretical error estimate in equation (15) is in remarkably good agreement with our external subvolumes estimator that incorporates fluctuations due to long-wavelength modes in addition to sampling and shot-noise variance at the scale of the given box-size.

(v) Existing analytic fits (Sheth & Tormen 1999; Warren et al. 2006) accurately reproduce our N-body measurements at up to $10^{14} h^{-1} \text{M}_\odot$, but fail to reproduce the abundance of the most massive haloes, underestimating the MF by 10 per cent (30 per cent) at $M = 3.16 \times 10^{14} h^{-1} \text{M}_\odot$ ($10^{15} h^{-1} \text{M}_\odot$).

(vi) The FoF halo MF deviates from universality (or self-similarity). In particular, the Sheth & Tormen (1999) fit, if extrapolated to $z \geq 0$ assuming universality, leads to an underestimation of the abundance of 30, 20 and 15 per cent at $z = 0, 0.5, 1$ for fixed $v = N_{\Delta/\sigma} \approx 3$ (corresponding to $M \sim 7 \times 10^{14} h^{-1} \text{M}_\odot, 2.5 \times 10^{14} h^{-1} \text{M}_\odot$ and $8 \times 10^{13} h^{-1} \text{M}_\odot$, respectively, see Fig. 9). This is due to some extent to the ST not being a very good fit at $z = 0$. If we instead extrapolate our $z = 0$ MICE best fit, the level of evolution in the amplitude of the halo MF is ~5 per cent (10 per cent) at $z = 0.5 (z = 1)$.

(vii) We provide a new analytic fit, equation (22), that reproduces N-body measurements over more than five orders of magnitude in mass, and follows its redshift evolution up to $z = 1$, that is accurate to 2 per cent. The new fit has the functional form of Warren et al. (2006), but with different parameters to account for the excess in the high-mass tail (see also Table 2).

(viii) Systematic effects in the abundance of clusters can strongly bias dark energy estimates. We estimate that medium depth surveys $z \lesssim 1$, using SZ cluster detection, could potentially bias the estimated value of the dark energy equation of state, $\omega$, by as much as 50 per cent. This effect is, however, an upper limit to the amplitude of this systematic effect, and it drops quickly with depth. For deep surveys $z \sim 1.5$, such as DES, the estimated bias is largely negligible.
As discussed in Section 1, there are several ways to define haloes and their masses, but the most commonly employed algorithms are FoF and SO. From an observer viewpoint, it is not clear that a single definition is optimal for all kinds of detection techniques (see Voit 2005 for a review). X-ray observations are easier and more robust in the inner centre of clusters that have higher density contrasts and are more relaxed. Hence, it seems clear that spherical apertures are more appropriate in this case, although with threshold density contrasts as high as 500 or more (Vikhlinin et al. 2009). Optical richness is measured within regions of fixed projected physical radius, resembling a modified FoF method. Hence, they demand a proper calibration of the mass-richness relation (Bode et al. 2001) if one does not want to rely on the correlation with the cluster X-ray properties. Cluster detections through the SZ effect are relatively recent in statistical terms and suffer from several sources of error (e.g. Carlstrom, Holder & Reese 2002). Yet, both halo definitions have been employed in the literature. For instance, mock SZ maps have been built and exploited using FoF (Schulz & White 2003), and correlations between X-ray properties and the SZ signal were studied starting from SO catalogues (Stanek et al. 2009b).

Our particular work focused on haloes defined through the FoF percolation method, leaving the SO for a follow-up study. None the less, we have tested that the trend seen towards high masses when comparing different MICE runs persists for decreasing linking length, that is, when probing the innermost regions of the haloes. We also leave for future work a more thorough investigation of the cosmological implications of our results. In particular, it remains to be seen what is the impact of the bias on cosmological parameters induced by MF priors in the presence of other systematic effects such as the uncertainties in the cluster mass–observable relations and the cosmological parameter degeneracies present in cluster abundance measurements.

ACKNOWLEDGMENTS

We would like to thank Jorge Carretero, Gus Evrard, Marc Merenda, Sebastián Pueblas, Román Scoccimarro, Volker Springel and Masahiro Takada for very useful help and discussions at different parts of this project. We acknowledge support from the MareNostrum supercomputer (BSC-CNS, www.bsc.es), through grants AECT-2008-1-0009, AECT-2008-2-0011, AECT-2008-3-0010 and Port d’Informació Científica (PIC, www.pic.es) where the simulations were run and stored, the use of the Friends-of-Friends code from the University of Washington N-body Shop (www-hpccastro.washington.edu) and the 2LPT code for initial conditions (cosmo.nyu.edu/roman/2LPT). The MICE simulations were implemented using the gadget-2 code (www.mpargarching.mpg.de/gadget). Funding for this project was partially provided by the Spanish Ministerio de Ciencia e Innovacion (MICINN), projects 2008SAO795, 2008SAO3176, 2008SAO6341, Consolider-Ingenio CSD2007-00060, research project 2005SGR00728 from Generalitat de Catalunya and the Juan de la Cierva MEC programme.

REFERENCES

Annis J. et al., 2005, submitted (arXiv:astro-ph/0510195)
Battye R. A., Weller J., 2003a, Phys. Rev. D, 68, 083506
Battye R. A., Weller J., 2003b, New Astron. Rev., 47, 775
Benítez N. et al., 2009, ApJ, 691, 241
Bode P., Bahcall N. A., Ford E. B., Ostriker J. P., 2001, ApJ, 551, 15
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150
Cabré A., Gaztañaga E., 2009a, MNRAS, 393, 1183
Cabré A., Gaztañaga E., 2009b, MNRAS, 396, 1119
Carlstrom J. E., Holder G. P., Reese E. D., 2002, ARA&A, 40, 643
Cohn J. D., White M., 2007, MNRAS, 385, 2025
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Cunha C., 2009, Phys. Rev. D, 79, 063009
Cunha C., Huterer D., Frieman J. A., 2009, Phys. Rev. D, 80, 063532
Davis M., Efstathiou G., Frenk C., White S. D. M., 1985, ApJ, 292, 371
Eke V. R., Baugh C. M., Cole S., Frenk C. S., Navarro J. F., 2006, MNRAS, 370, 1147
Estrada J., Sefusatti E., Frieman J. A., 2009, ApJ, 692, 265
Evrard A. E. et al., 2002, ApJ, 574, 7
Fosalba P., Gaztañaga E., Castander F. J., Manera M., 2008, MNRAS, 391, 435
Frenk C. S. et al., 2000, preprint (arXiv:astro-ph/0007362)
Gaztañaga E., Cabré A., Castander F., Crocce M., Fosalba P., MA09a, MNRAS, 399, 801
Gaztañaga E., Cabré A., Hui L., 2009b, MNRAS, 399, 1663
Grossi M., Dolag K., Branchini E., Matarrese S., Moscardini L., 2007, MNRAS, 382, 1261
Heitmann K., Lukic Z., Habib S., Ricker P. M., 2006, ApJ, 642, L85
Henry J. P., Evrard A. E., Hoekstra H., Babul A., Mahdavi A., 2009, ApJ, 691, 1307
Hu W., 2003, Phys. Rev. D, 67, 081304
Hu W., Kravtsov A. V., 2003, ApJ, 584, 702
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
Kim J., Park C., Gott R., Dubinski J., 2009, ApJ, 701, 1547
Knebe A., Wagner C., Knollmann S., Dieckershoff T., Krause F., 2009, MNRAS, 698, 266
Lacey C. G., Cole S., 1993, MNRAS, 262, 627
Lacey C. G., Cole S., 1994, MNRAS, 271, 676
Lee J., Shandarin S. F., 1998, ApJ, 500, 14
Lukic Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, ApJ, 671, 1160
Lukic Z., Reed D., Habib S., Heitmann K., 2009, ApJ, 692, 217
Magni M., Riotta A., 2009, preprint (arXiv:0903.1251)
Manera M., Sheth R. K., Scoccimarro R., 2009, MNRAS, preprint (arXiv:0906.1314)
Mantz A., Allen S. W., Ebeling H., Rapetti D., 2008, MNRAS, 387, 1179
Matarrese S., Verde L., Jimenez R., 2000, ApJ, 541, 10
Pillepich A., Porciani C., Hahn O., 2009, MNRAS, 402, 191
Press W. H., Schechter P., 1974, ApJ, 187, 425
Reed D. S., Gardner J., Quinn T., Stadel J., Fardal J., Lake G., Governato F., 2003, MNRAS, 346, 565
Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2007, MNRAS, 374, 2
Robson B. E., Kravtsov A. V., Tinker J., Zentner A. R., 2009, ApJ, 696, 636
Rozo E. et al., 2009, preprint (arXiv:0902.3702)
Schulz A. E., White M., 2003, ApJ, 586, 723
Scoccimarro R., 1998, MNRAS, 299, 1097
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Springel V., 2005, MNRAS, 364, 1105
Springel V. et al., 2005, Nat, 435, 629
Stanek R., Rudd D., Evrard A. E., 2009a, MNRAS, 394, L11
Stanek R., Rasia E., Evrard A. E., Pearce F., Gazolla L., 2009b, ApJ, submitted (arXiv:0910.1599)
Tago E., Einasto J., Saar E., Tempel E., Einasto M., Vennik J., Muller V., 2008, A&A, 479, 927
Takada M., Bridle S., 2007, New J. Phys., 9, 446
Teyssier R. et al., 2009, A&A, 497, 335

© 2010 The Authors. Journal compilation © 2010 RAS, MNRAS 403, 1353–1367
The abundance of massive clusters

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlober S., Holz D. E., 2008, ApJ, 688, 709
Vikhlinin A. et al., 2009, ApJ, 692, 1060
Voit G. M., 2005, Rev. Mod. Phys., 77, 207
Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, ApJ, 646, 881
Wen Z. L., Han J. L., Liu F. S., 2009, ApJS, 183, 197

White M., 2001, A&A, 367, 27
White M., 2002, ApJS, 143, 241
Zehavi I. et al., 2005, ApJ, 630, 1

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.