Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **431,** 1366–1382 (2013)
Advance Access publication 2013 March 8

doi:10.1093/mnras/stt259

The abundance of (not just) dark matter haloes

Till Sawala,[1]★ Carlos S. Frenk,[1] Robert A. Crain,[2] Adrian Jenkins,[1] Joop Schaye,[2] Tom Theuns[1,3] and Jesus Zavala[4,5]

[1]*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK*
[2]*Leiden Observatory, Leiden University, Postbus 9513, NL-2300 RA Leiden, the Netherlands*
[3]*Department of Physics, University of Antwerp, Campus Groenenborger, Groenenborgerlaan 171, B-2020 Antwerp, Belgium*
[4]*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada*
[5]*Perimeter Institute for Theoretical Physics, 31 Caroline St. N., Waterloo, ON N2L 2Y5, Canada*

Accepted 2013 February 8. Received 2013 January 25; in original form 2012 June 27

ABSTRACT

We study the effect of baryons on the abundance of structures and substructures in a Λ cold dark matter (CDM) cosmology, using a pair of high-resolution cosmological simulations from the Galaxies-Intergalactic Medium Interaction Calculation project. Both simulations use identical initial conditions, but while one contains only dark matter, the other also includes baryons. We find that gas pressure, reionization, supernova feedback, stripping and truncated accretion systematically reduce the total mass and the abundance of structures below $\sim 10^{12}\,\mathrm{M}_\odot$ compared to the pure dark matter simulation. Taking this into account and adopting an appropriate detection threshold lower the abundance of observed galaxies with maximum circular velocities $v_{\rm max} < 100\,\mathrm{km\,s^{-1}}$, significantly reducing the reported discrepancy between ΛCDM and the measured H I velocity function of the Arecibo Legacy Fast ALFA survey. We also show that the stellar-to-total mass ratios of galaxies with stellar masses of $\sim 10^5$–$10^7\,\mathrm{M}_\odot$ inferred from abundance matching of the (sub)halo mass function to the observed galaxy mass function increase by a factor of ∼2. In addition, we find that an important fraction of low-mass subhaloes are completely devoid of stars. Accounting for the presence of *dark* subhaloes below $10^{10}\,\mathrm{M}_\odot$ further reduces the abundance of *observable* objects and leads to an additional increase in the inferred stellar-to-total mass ratio by factors of 2–10 for galaxies in haloes of 10^9–$10^{10}\,\mathrm{M}_\odot$. This largely reconciles the abundance matching results with the kinematics of individual dwarf galaxies in ΛCDM. We propose approximate corrections to the masses of objects derived from pure dark matter calculations to account for baryonic effects.

Key words: galaxies: evolution – galaxies: formation – galaxies: luminosity function, mass function – cosmology: theory.

1 INTRODUCTION

The large-scale evolution of the Universe is determined by gravity which drives hierarchical structure formation in an expanding space. While the nature of dark matter (DM) remains an open question at a fundamental level, gravity makes no distinction between baryonic and DM. Consequently, numerical simulations of cosmic structure formation are generally purely gravitational, or 'dark matter only' (DMO), which greatly reduces the computational cost and complexity compared to gas-dynamical simulations. Over the past decades, DMO simulations have played a key part in establishing the currently preferred Λ cold dark matter (CDM) paradigm (e.g. Davis et al. 1985; Frenk et al. 1988) and the hierarchical formation scenario (e.g. White & Rees 1978; White & Frenk 1991), explaining the observed abundance and correlation of structures at different redshifts and over many mass and length scales.

On dwarf galaxy scales (haloes up to $10^{11}\,\mathrm{M}_\odot$ or $v_{\rm max} < 100\,\mathrm{km\,s^{-1}}$, or galaxies with stellar masses up to $\sim 10^9\,\mathrm{M}_\odot$), however, CDM simulations appear to overpredict the amount of structure, to a degree that appears incompatible with observations: the abundance of haloes with a maximum circular velocity of $v_{\rm max} \sim 35\,\mathrm{km\,s^{-1}}$ exceeds the abundance measured in the Arecibo Legacy Fast ALFA (ALFALFA) H I survey by a factor of ∼10 (e.g. Zavala et al. 2009; Papastergis et al. 2011); the simulations fail to match the dynamics of satellites (e.g. Parry, Eke & Frenk 2009; Strigari, Frenk & White 2010; Boylan-Kolchin, Bullock & Kaplinghat 2011, 2012, but see Wang et al. 2012); and the stellar-to-total mass ratios inferred from the abundances do not match those of individual dwarf galaxies (e.g. Ferrero et al. 2012; Sawala et al. 2011). In this work,

★E-mail: till.sawala@durham.ac.uk

we show that the inclusion of baryons and their relevant astrophysical processes changes the abundance of (sub)haloes sufficiently to reconcile CDM with observations.

While the observable matter in the Universe traces the underlying total mass distribution, galaxies are 'biased' with respect to the DM (Kaiser 1984; Davis et al. 1985), and the stellar-to-total mass ratio varies greatly with galaxy mass, evolutionary stage and environment. Linking simulated structures to observations of galaxies must therefore be done indirectly, via methods such as (sub)halo abundance matching (e.g. Vale & Ostriker 2006; Conroy & Wechsler 2009; Guo et al. 2010; Moster et al. 2010), halo occupation distribution models (e.g. Benson et al. 2000; Seljak 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2002) or semi-analytical models (e.g. White & Frenk 1991; Kauffmann, White & Guiderdoni 1993; Cole et al. 1994; Somerville & Primack 1999; Bower et al. 2006; Croton et al. 2006). While these approaches vary greatly in complexity, aim and predictive power, they share the fundamental assumption that the total mass distribution in the Universe is governed by gravity alone, and can be represented by DM.

In this work, we investigate the impact of baryon physics on structure formation on smaller scales. We find that baryons systematically affect the masses and abundances of objects in three ways.

(i) Interstellar gas is expelled from haloes in simulations that include supernova feedback, while intergalactic gas is prevented from accreting in the presence of ionizing radiation.

(ii) Ram-pressure stripping preferentially removes gas from low-mass satellite galaxies, and satellites are more easily disrupted.

(iii) The shallower potential well caused by the loss of baryons subsequently leads to diminished accretion, both of baryons and of DM.

Of these, the first two effects can substantially reduce the mass of baryons in small haloes, and lower the total mass of an object by up to the universal baryon fraction, compared to the corresponding DMO simulation. The third effect can reduce an object's mass even further, as haloes with a lower mass and shallower potential accrete less material. All baryon effects show a strong mass dependence, and stripping also depends on environment.

To quantify the influence of baryons, we use a set of simulations of structure formation in identical cosmological volumes: a pure DM simulation, and one which includes a gas physics model that incorporates hydrodynamics, radiative heating and cooling, a cosmic UV background, star formation, chemical evolution and supernova feedback. This simulation, part of the 'Galaxies-Intergalactic Medium Interaction Calculation' ('GIMIC'; Crain et al. 2009), has sufficient dynamic range and volume to resolve objects with total mass between 10^9 and $10^{14}\,\mathrm{M}_\odot$, from dwarf galaxies to massive groups and clusters. It has been used previously to study many aspects of galaxy formation, such as environmental dependences (Crain et al. 2009), disc galaxies (Sales et al. 2012) and their X-ray coronae (Crain et al. 2010), the origin of the baryonic Tully–Fisher relation (McCarthy et al. 2012a), and the assembly of stellar haloes (Font et al. 2011; McCarthy et al. 2012b).

Here, our focus is not on the formation of galaxies, but on the formation of their DM hosts, and in particular, on the abundance of haloes and subhaloes of different masses and circular velocities, in different environments, and over time. Matching individual objects across both simulations also enables us to understand the mechanisms by which baryons affect the abundance of structures, and to derive an approximate expression that allows us to correct for baryonic effects in generic DMO simulations.

The effect of baryons on the structure of haloes has been studied extensively, both using analytical models (e.g. Blumenthal et al. 1986; Sellwood & McGaugh 2005) and in numerical simulations (e.g. Abadi et al. 2010; Duffy et al. 2010; Bryan et al. 2012). Whereas dissipation universally increases the concentration of haloes (e.g. Mo, van den Bosch & White 2010; Gnedin et al. 2011), a non-adiabatic response to rapid outflows induced by supernovae can erase the central cusp in the profiles of low-mass haloes (Navarro, Eke & Frenk 1996; Governato et al. 2010; Pontzen & Governato 2012), potentially resolving the discrepancy between cusped profiles predicted by CDM and the cored profiles indicated by some observations (e.g. Walker & Peñarrubia 2011, but see Strigari et al. 2010; Wolf & Bullock 2012).

Comparisons of the masses and abundances of haloes in DMO and gas simulations have also been performed previously, by Weinberg et al. (2008), Rudd, Zentner & Kravtsov (2008), Dolag et al. (2009) and Cui et al. (2012), but these have focused on scales of $10^{12}\,\mathrm{M}_\odot$ and above. Semboloni et al. (2011) and van Daalen et al. (2011) have investigated the effects of baryons on estimates of cosmological parameters, for $k = 10\,h$/Mpc and larger. On these scales, Dolag et al. (2009) concluded that baryons only have a minor effect on halo masses, with adiabatic contraction balancing feedback, while Rudd et al. (2008) and Cui et al. (2012) both found an increase in the mass of haloes, with Cui et al. (2012) reporting a net increase of 1–2 per cent in M_{200} for haloes of $10^{13.5}\,\mathrm{M}_\odot$ and above, and Rudd et al. (2008) report an increase in the cumulative halo mass function of $\sim$10 per cent above $10^{12}\,h^{-1}\,\mathrm{M}_\odot$. An opposite effect of a reduction in the matter power spectrum was found by van Daalen et al. (2011), in simulations which include AGN feedback. Simha et al. (2012) have compared hydrodynamic simulations to results of (sub)halo abundance matching, and found the results to be consistent.

By comparison to these works, our study is based on much higher resolution simulations, which allows us to investigate the effect of baryons in lower mass haloes. In showing that the effect of baryons is highly mass dependent, we find that while our results are consistent with previous results for objects of $10^{12}\,\mathrm{M}_\odot$ and above,[1] on smaller scales, baryons significantly reduce the mass of haloes and subhaloes, with consequences for abundance matching.

This paper is organized as follows: in Section 2, we describe the simulations, and outline the procedures employed to identify and link substructures across the two simulations. Section 3 contains our main results: in Section 3.1, we compare individual structures in both simulations, and we give an analytic expression for the change in the median subhalo mass in Section 3.2. In Section 3.3, we describe the effect of baryons on the mass functions of haloes and subhaloes, while in Section 3.4, we examine the v_{max} function, and compare our results to observational H I data. In Section 3.5, we show the effect of baryons on the satellite mass functions of groups and clusters. Section 3.6 focuses on *dark* subhaloes, which do not contain any gas or stars. In Section 4, we describe how the baryonic effects change the results of abundance matching, and compare the inferred stellar-to-total mass ratios to observations. We summarize our results in Section 5. Parameters for the analytical mass correction are given in Appendix A, and numerical convergence is discussed in Appendix B.

[1] All masses and distances in our own work are expressed in physical units.

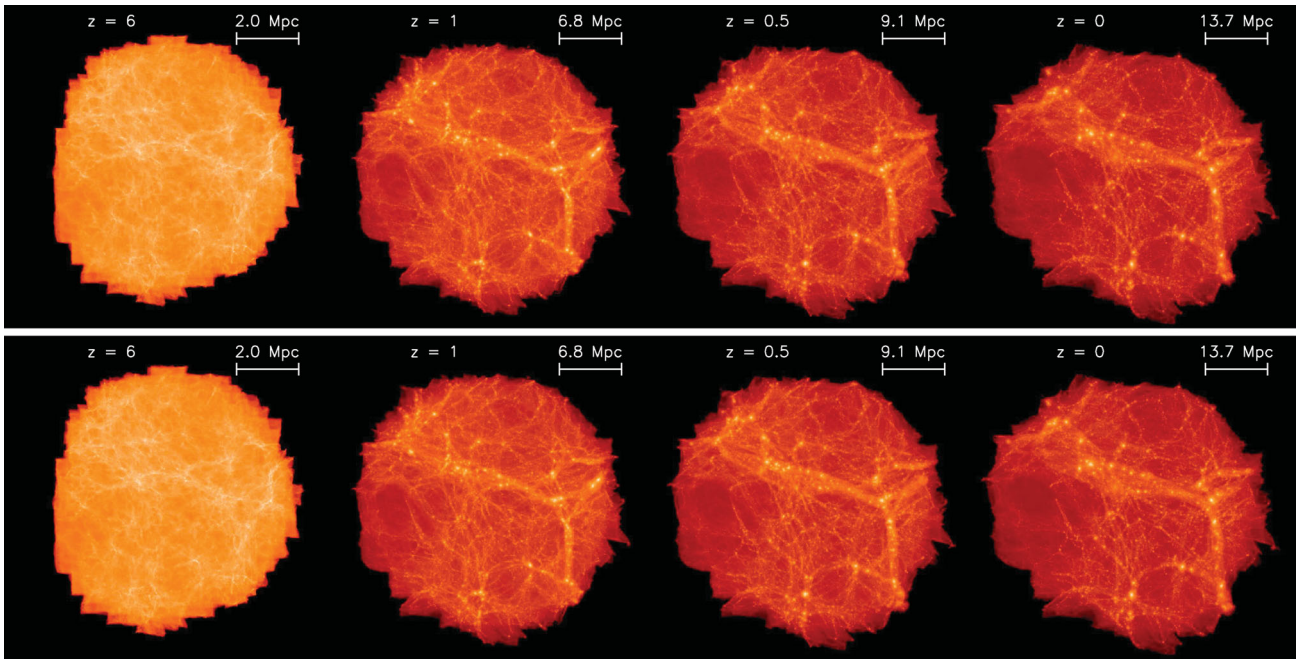

Figure 1. Evolution of a slice of $16\,h^{-1}$ Mpc in depth through the total mass distribution within the high-resolution regions of the DMO simulation (top) and GIMIC simulation (bottom), at $z = 6$, 1, 0.5 and 0 (left to right). Shown in each panel is a scale bar of a constant *comoving* length of $10\,h^{-1}$ Mpc, and its corresponding length in physical coordinates. On large scales, the mass distributions in the DMO and GIMIC simulations appear identical.

2 METHODS

2.1 The simulations

Our comparison is based on a set of *N*-body simulations of the same cosmological volume: a hydrodynamical simulation ('GIMIC'), performed by the Virgo Consortium and described in detail by Crain et al. (2009), and a matching DMO simulation where the total mass is composed of DM. The simulations refine a spherical Lagrangian volume (labelled 0σ in Crain et al. 2009) of radius $\sim 18\,h^{-1}$ Mpc (a volume of ~ 63 Gpc^3 at $z = 0$) and median density, selected from the Millennium Simulation (Springel et al. 2005). Within the zoom region, additional small-scale modes are added using the zoom method described by Power et al. (2003), while the external large-scale density field is represented by lower resolution DM particles.

The simulations were performed using the TreePM smoothed particle hydrodynamics (SPH) code P-GADGET-3, an extension of the publicly available code GADGET-2 (Springel 2005). The cosmological parameters are identical to those of the Millennium Simulations: density parameters $\Omega_\mathrm{m} = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_\mathrm{b} = 0.045$, Hubble parameter $h = 0.73$, power spectrum normalization $\sigma_8 = 0.9$ and spectral index $n = 1$. The evolution of the DM distribution in the full, high-resolution region in both simulations is shown in Fig. 1. By construction, the simulated volume is approximately spherical at $z = 1.5$, and becomes more irregular at both higher and lower redshifts. On scales of megaparsecs and above, baryons do not significantly affect structure formation, and no differences can be seen in the density distribution shown in Fig. 1.

2.1.1 *Initial conditions*

Both simulations evolve the same number, 2.75×10^8, of DM particles, from $z = 127$ to 0, and are analysed at 58 co-temporal snapshots. In the GIMIC simulation, the high-resolution region additionally contains an equal number of gas particles that can be converted to star particles.

The initial particle masses are $m_\mathrm{g} = 1.98 \times 10^6\,\mathrm{M}_\odot$ and $m_\mathrm{DM} = 9.05 \times 10^6\,\mathrm{M}_\odot$ for gas and dark matter, respectively, in the GIMIC simulation, and $m_\mathrm{DM} = 1.98 + 9.05 = 11.03 \times 10^6\,\mathrm{M}_\odot$ in the DMO simulation. In the GIMIC simulation, gas particles can be converted into star particles of equal mass, and mass can be exchanged between gas and star particles in the form of stellar winds and supernovae. The gravitational softening scale for all particle types in the high-resolution region is initially fixed in comoving coordinates, and chosen so that it reaches a value of $0.5\,h^{-1}$ kpc Plummer equivalent in physical coordinates at $z = 3$, whereupon it remains fixed in physical coordinates, contracting in comoving coordinates as $1/a$. To test convergence, both simulations were also repeated at lower resolution, with all particle masses increased by a factor of 8, and the softening increased by a factor of 2. Unless we specifically refer to the lower resolution simulations for the purpose of comparison, all results are based on the high-resolution simulations. We discuss convergence in Appendix B, and consider the abundance of (sub)haloes to be unaffected by resolution down to $10^9\,\mathrm{M}_\odot$.

2.1.2 *Physical model*

While the DMO simulation includes only gravitational interactions, the GIMIC simulation also includes astrophysical processes, such as gas dynamics, photoionization, radiative element-by-element cooling, star formation and supernova feedback. Here, we give a brief outline of the most important aspects of the physics model; a more complete description can be found in Crain et al. (2009) and references therein.

(i) Radiative cooling is computed assuming the optically thin limit and ionization equilibrium, depending on redshift, local temperature and density, and chemical composition. The

intergalactic medium (IGM) is heated uniformly, with a redshift-dependent UV/X-ray background as given by Haardt & Madau (2001), with hydrogen and helium-II being reionized at $z \sim 9$ and $z \sim 3.5$, respectively.

(ii) Following Schaye & Dalla Vecchia (2008), star formation is implemented by imposing an effective equation of state for gas particles above a density threshold of $n_{\rm H} = 0.1\,{\rm cm}^{-3}$, that makes the Jeans mass independent of the gas density. In this regime, gas particles can be transformed into star particles at a pressure-dependent rate that reproduces a local Kennicutt–Schmidt law, with a star formation rate column density of $\dot{\Sigma} = 1.5 \times 10^4\,{\rm M}_\odot\,{\rm yr}^{-1}\,{\rm kpc}^{-2}\,\Sigma_{\rm g}^{1.4}$, where $\Sigma_{\rm g}$ is the gas column density in ${\rm M}_\odot\,{\rm pc}^{-2}$. Each star particle is assumed to represent a single stellar population, with an initial mass function based on Chabrier (2003). As described in Wiersma et al. (2009), stellar evolution follows the abundances of 11 metals, which are released over time to the surrounding gas particles by Types II and Ia supernovae and asymptotic giant branch stars.

(iii) Following Dalla Vecchia & Schaye (2008), supernova feedback is imparted on the surrounding gas particles in the form of kinetic energy, with each star in the mass range 6–100 $\rm M_\odot$ providing 10^{51} erg, sufficient to accelerate an average of four neighbours to a wind velocity of 600 km s^{-1}.

Whilst some parts of the astrophysical model are still subject to considerable uncertainty, both in terms of the underlying physics and its numerical implementation, it has been shown to reproduce the cosmic star formation rate (Crain et al. 2009; Schaye et al. 2010) and the properties of individual galaxies in some detail (Font et al. 2011; McCarthy et al. 2012a). By comparison, the results presented in this paper are quite general.

The GIMIC simulation does not include AGN. This omission is the most likely cause of a too high stellar and baryon fraction in the most massive haloes. Because the total abundance of these massive haloes is low, the effect on the cumulative mass function is negligible for less massive haloes.

2.2 Analysis

We compare the populations of haloes and subhaloes in the two simulations, as well as individual objects which are matched across both. In Section 2.2.1, we describe the identification of structures and substructures, and the definition of different subsets used in the analysis. In Section 2.2.3, we describe the procedure and criteria used for matching substructures across the two simulations.

2.2.1 *Identification of substructures*

In each snapshot, structures are identified in a two-step process. First, DM *haloes* are found using the friend-of-friend (FoF) algorithm, with a linking length of 0.2 (Davis et al. 1985). In the GIMIC simulation, star particles and gas particles are then attached to the nearest dark matter particle, and inherit its FoF association. In a second step, self-bound substructures within haloes are identified using the SUBFIND algorithm of Springel et al. (2001), with the extension of Dolag et al. (2009) to account for baryons. SUBFIND computes the density field within each FoF halo using an SPH interpolation. A set of $N > 20$ particles (of any type) is considered as a potential substructure if it exceeds the local smoothed density estimate. It is then subjected to unbinding, whereby all particles whose combined kinetic and internal energy exceeds their gravitational binding energy to the substructure are iteratively removed, until the substructure either vanishes, falling below the 20 particle threshold, or is identified as a genuine self-bound subhalo.

2.2.2 *Mass estimators*

Having identified haloes and subhaloes, their masses can be measured in several ways. We define the FoF mass, $M_{\rm FoF}$, of a halo as the sum of the masses of all particles belonging to the FoF group. In structure formation theory, it is customary to characterize a halo by M_{200}, which describes an object with an overdensity of 200, near the threshold for a DM halo to collapse and virialize. Different definitions exist for M_{200} and r_{200}, the radius which encloses a spherical volume of corresponding overdensity: $M_{200,\,\rm crit}$ is defined such that the mean overdensity within r_{200} is 200 times the critical density; $M_{200,\,\rm mean}$ such that the overdensity within r_{200} is 200 times the mean background density; $M_{\rm top\text{-}hat}$ is the mass in a spherical volume whose overdensity equals the value at virialization in the top-hat collapse model. In Fig. 2, we show a comparison of the three mass estimators relative to the FoF mass for all haloes at $z = 0$. We find that $M_{\rm top\text{-}hat}$ most closely matches $M_{\rm FoF}$, and that the spread between the three estimators increases with halo mass in the GIMIC simulation, from ∼20 per cent at $10^9\,{\rm M}_\odot$ to ∼30 per cent at $10^{14}\,{\rm M}_\odot$. Importantly for our purpose, below $10^{12}\,{\rm M}_\odot$ the ratio between each of the three halo mass estimators and the FoF mass is very similar in both simulations, and much smaller than the measured baryonic effect on the halo mass. We use $M_{\rm FoF}$ as our halo mass definition throughout the rest of the paper, as it is independent of the concentration of the halo. The fact that our results are very similar for the different mass estimators indicates that a measured halo mass difference is a true physical effect.

For subhaloes, the mass $M_{\rm SH}$ is defined as the total mass of particles bound to the subhalo. For *isolated* subhaloes, i.e. central subhaloes with no satellites, we find that the subhalo mass is typically very close to the halo mass, indicating that most particles belonging to the halo are also gravitationally bound to it. We have also examined individual subhaloes, and compared density profiles of the gravitationally bound particles to those of all particles in the vicinity. We find no noticeable difference between the two

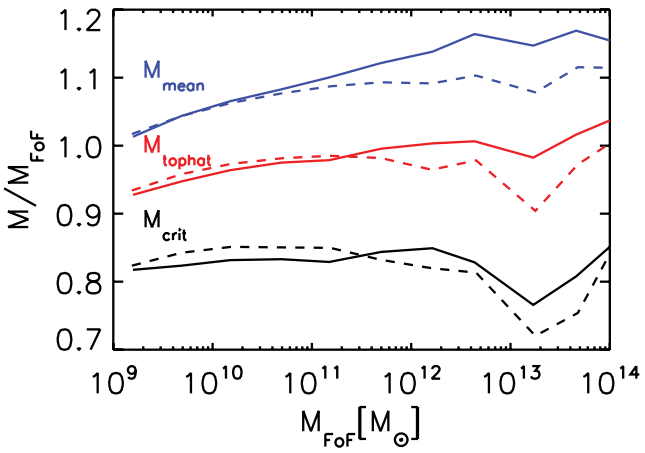

Figure 2. Comparison of halo mass estimators $M_{200,\,\rm mean}$ (blue), $M_{\rm top\text{-}hat}$ (red) and $M_{200,\,\rm crit}$ (black), as a function of $M_{\rm FoF}$ for all haloes at $z = 0$. The solid lines show the results of the GIMIC simulation, and the dashed lines show results of the DMO simulation. As expected, for all objects, $M_{200,\,\rm mean}$ gives the highest mass per halo, followed by $M_{\rm top\text{-}hat}$ and $M_{200,\,\rm crit}$. Importantly, there is no systematic bias in mass estimators between the GIMIC and the DMO simulation.

simulations in the ratio of bound to unbound particles for each subhalo, indicating that any subhalo mass difference that we find between the two simulations is a genuine physical effect. Examples of density profiles, including a comparison of the bound and unbound particles, are shown in Fig. 5.

We discuss the convergence of the simulations in terms of the halo and the subhalo mass functions in Appendix B, and show that the principal effects of the baryon physics are independent of the definition of the objects, and are not affected by resolution.

2.2.3 *Matching substructures across simulations*

In order to understand how the differences between the DMO and the GIMIC simulations arise, we also compare the evolution of individual objects. We use the fact that DM particles of identical position in the initial conditions[2] have identical IDs in order to match substructures across both simulations. Specifically, at each snapshot and for every subhalo in the DMO simulation, we consider its five most bound particles, and require that at least three of them belong to a single subhalo in the GIMIC simulation. This ensures that each subhalo in the DMO simulation is matched to at most one subhalo in the GIMIC simulation. Because mergers can occur at different times in the two simulations (and in some cases occur only in one and not in the other), we also impose a threshold on the mass ratio, matching only objects whose mass ratio is less than 4:1.

As shown in Fig. 3, this method results in >90 per cent of subhaloes above $10^9\,\mathrm{M}_\odot$ in the DMO simulation being matched to unique counterparts in the GIMIC simulation, including >95 per cent of centrals and >80 per cent of satellites. The reduction of the matching fraction for present-day satellites compared to central subhaloes is largely due to their more chaotic orbits, and the associated increase in the likelihood of diverging merger histories of individual objects in the two simulations. In addition, low-mass satellites are more likely to be disrupted in the GIMIC simulation compared to the DMO simulation.

In total, at $z = 0$, the DMO simulation contains 214 676 haloes and 254 305 subhaloes; the GIMIC simulation contains 148 765 haloes and 162 866 subhaloes. Including only subhaloes with a total mass above $10^9\,\mathrm{M}_\odot$, the DMO simulation has 86 137 subhaloes, 69 699 centrals and 16 438 satellites, while the GIMIC simulation has 62 265 subhaloes, 50 077 centrals and 12 188 satellites. Of the central subhaloes in the DMO simulation with masses above $10^9\,\mathrm{M}_\odot$, 92.6 per cent are matched to subhaloes in GIMIC, 97 per cent of which are matched to centrals and 3 per cent to satellites. In the same mass range, 80.0 per cent of satellites in the DMO simulation are matched to subhaloes in GIMIC, 70 per cent of which are matched to satellites and 30 per cent to centrals.

We also compared the matching success rate for subhaloes in the central parts and near the boundary of the high-resolution region (see Section 2.1), but found no significant edge effects. While the criteria for matching subhaloes appear somewhat arbitrary, we note that the matching of objects across the two simulations only serves to understand the origin of the differences, but that statistical quantities, including mass functions and satellite mass functions, are independent of the matching rate and criteria.

[2] A DM particle in the DMO simulation is replaced by a pair of DM and gas particles with the same centre of mass in the initial conditions of the GIMIC simulation.

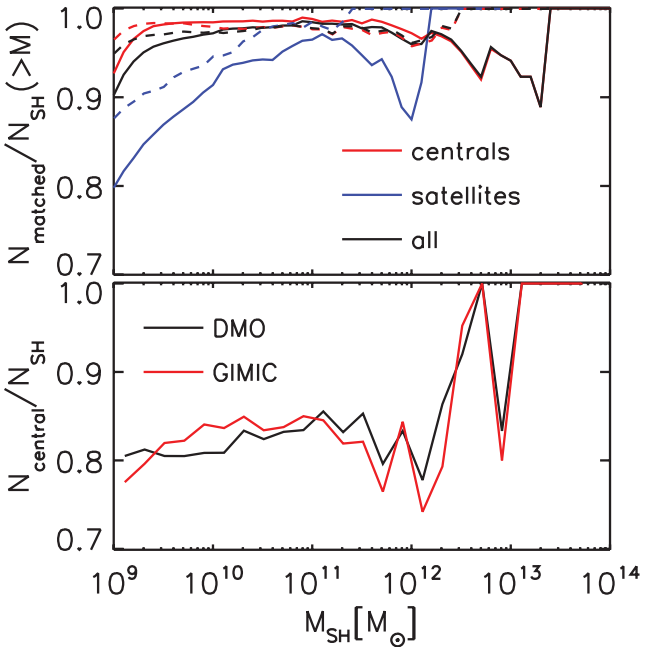

Figure 3. Top panel: cumulative fraction of subhaloes in the DMO simulation that are uniquely matched to subhaloes in the GIMIC simulation at $z = 0$ (solid) and $z = 1$ (dashed), for centrals (red), satellites (blue) and all subhaloes (black). Bottom panel: fraction of all subhaloes at $z = 0$ that are central in the DMO simulation (black), and in the GIMIC simulation (red). In general, the fraction of subhaloes that are matched increases with mass, and it is slightly lower at $z = 0$ compared to $z = 1$, particularly for satellites. The lower matching fraction among satellites and its decrease over time are attributable to a higher rate of mergers and a higher probability for divergent evolution between the two simulations. Overall, more than 90 per cent of subhaloes above $10^9\,\mathrm{M}_\odot$ are matched at $z = 0$. In both simulations, centrals account for ∼80 per cent of subhaloes in the mass range 10^9–$10^{12}\,\mathrm{M}_\odot$.

3 EFFECT OF BARYONS ON STRUCTURES

In this section, we describe the effects that the inclusion of baryons have on the formation and evolution of structure mainly in statistical terms. We compare the evolution of individual objects across the two simulations in Section 3.1. The effect on the overall abundance of haloes and subhaloes is discussed in Section 3.3. Section 3.4 shows the effect on the v_{max} function, and compares our results to measurements of the ALFALFA H I survey in the Local Volume. In Section 3.5, we focus on the mass functions of individual groups and clusters. We highlight the importance of *dark* subhaloes in Section 3.6, which will be relevant to the discussion of abundance matching in Section 4.

From the similarity in the overall mass distribution of the DMO simulation and its counterpart with baryons ('GIMIC') shown in Fig. 1, it is clear that the large-scale evolution of structures is nearly identical. On scales of several Mpc, gravity remains the only relevant force even in the hydrodynamical simulation, and the gas distribution closely follows that of the DM. Fig. 4 shows the projected density distribution of gas (left), stars (centre) and DM (right) in a region of 11.2^3 Mpc^3 from the GIMIC simulation at $z = 0$. On scales of several hundred kpc and below, it is apparent that the gas distribution is much smoother than the DM, a reflection of the fact that most of the interstellar medium of small haloes has been expelled, while the IGM is heated through the UV background, resulting in a large Jeans length (Theuns, Schaye & Haehnelt 2000).

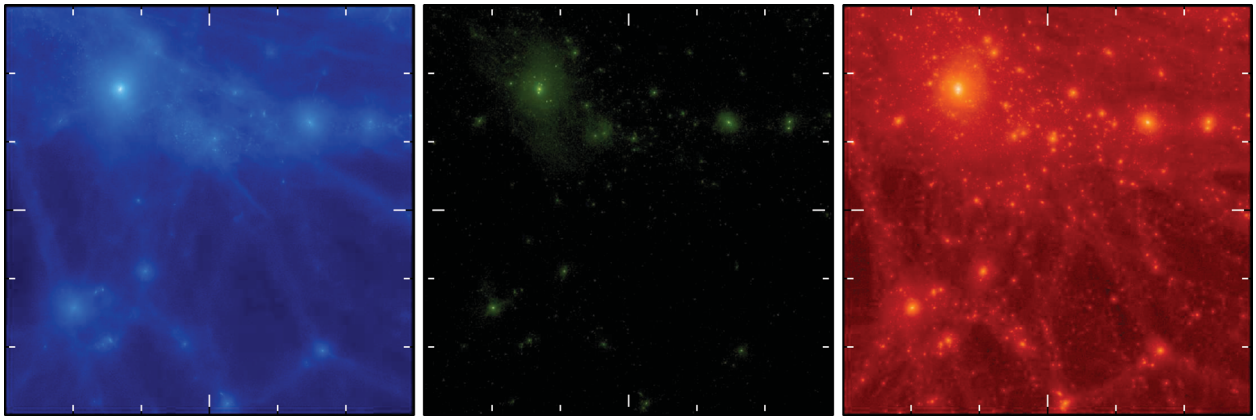

Figure 4. Projected density distributions of gas (left), stars (centre) and DM (right) in the GIMIC simulation, from a cubic region of side length $l = 11.2$ Mpc at $z = 0$. On large scales, stars and gas follow the DM distribution, but on galaxy scales, the distributions differ, leading to a difference in the total mass distributions between the GIMIC and the DMO simulations.

While the stellar mass is also highly clustered, it is only significant in the most massive objects.

3.1 Individual objects

We use pairs of individual matched subhaloes in the DMO and the GIMIC simulation to study the effect of baryons on the formation of haloes and subhaloes. In Fig. 5, we show the mass profiles in GIMIC at $z = 0$, for three typical matched central subhaloes of 10^{10}, 10^{11} and $10^{12}\,\mathrm{M}_{\odot}$. For each subhalo, the density in DM is shown in black, the density in stars in green and the density in gas in blue. As the subhalo mass increases from left to right, the baryon fraction also increases. Whereas the total density is dominated by DM throughout for 10^{10} and $10^{11}\,\mathrm{M}_{\odot}$ subhaloes, $10^{12}\,\mathrm{M}_{\odot}$ subhaloes contain a significant mass of stars in the centre. For subhalo masses of 10^{10} and $10^{11}\,\mathrm{M}_{\odot}$, the total mass density in GIMIC is systematically lower than the mass density of the corresponding object in the DMO simulation, which is plotted as a black dashed line. Note that for the $10^{12}\,\mathrm{M}_{\odot}$ subhaloes, the central mass profile is steeper in GIMIC compared to the DMO simulation as a result of cooling and adiabatic contraction; this is not seen in the smaller haloes.

Fig. 6 compares the masses of many individual subhaloes across the two simulations, as a function of their mass in the DMO simulation, subdivided into centrals (left-hand panel) and satellite (right-hand panel). For the purposes of this comparison, we only include

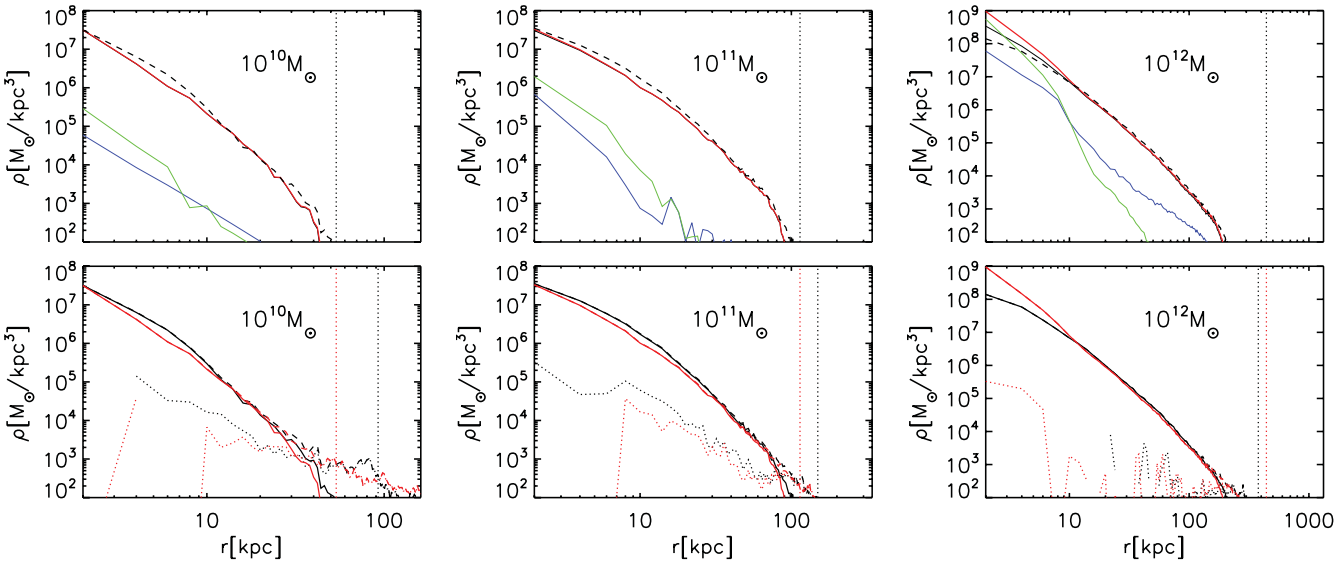

Figure 5. Spherical density profiles of three pairs of individual matched, central subhaloes. In the top row, solid lines show the different components in the GIMIC simulation: DM (black), stars (green), gas (blue) and total mass density (red), each including only particles bound to the subhalo. The dashed line shows the density of the same object in the DMO simulation, and the vertical lines indicate the radius of the outermost bound particle. The material within $\sim 10^{12}\,\mathrm{M}_{\odot}$ subhaloes has been rearranged appreciably in the GIMIC simulation relative to the DMO simulation, with baryons accumulating in the centre. The difference between the solid and dashed black lines in the inner part shows that the DM itself is also more concentrated in GIMIC, but the total subhalo mass does not change. By contrast, $\sim 10^{10}\,\mathrm{M}_{\odot}$ and $\sim 10^{11}\,\mathrm{M}_{\odot}$ subhaloes have lost mass in the GIMIC case. Here, the total density is dominated by DM throughout, and lower than in the DMO simulation. In the bottom row, solid lines denote the density profiles computed using bound particles, dotted lines show unbound particles in the vicinity (not counted towards the subhalo mass), and dashed lines show the sum of the densities of bound and unbound particles. In all cases, the density of unbound particles is low, and there is no significant difference between the two simulations in the ratio of bound to unbound particles.

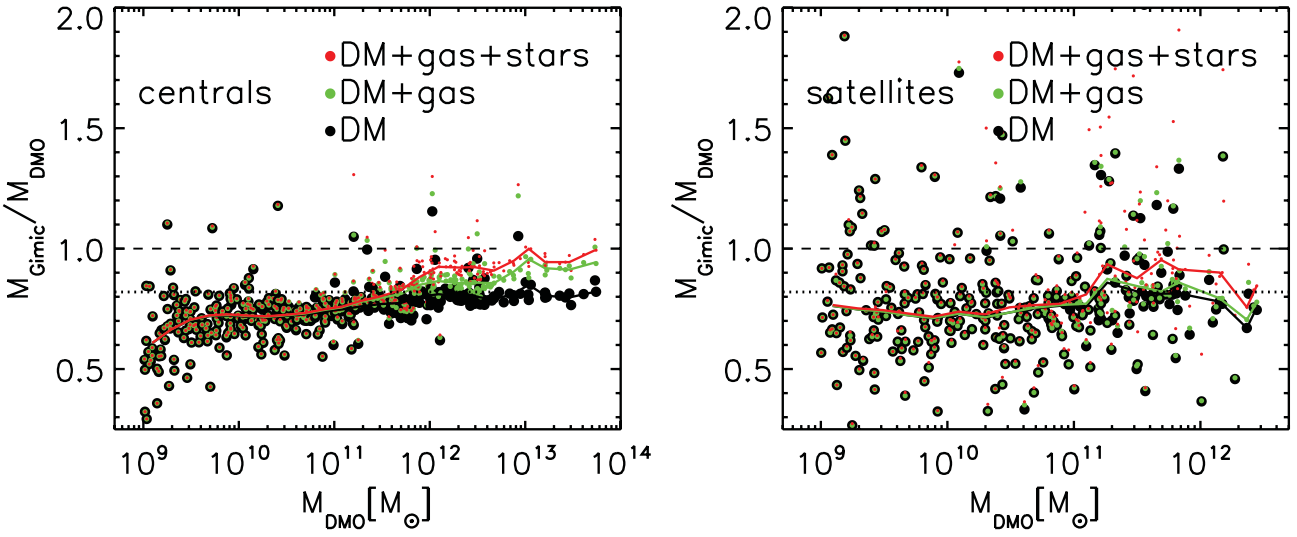

Figure 6. Ratio of subhalo mass in GIMIC relative to the DMO simulation for individual, matched subhaloes at $z = 0$, as a function of DMO mass, for central subhaloes (left) and satellite subhaloes (right). Every subhalo is represented by a set of three circles, with the y-values indicating its mass components in the GIMIC simulation: only the DM (black), DM and gas (green), and total mass including DM, gas and stars (red). The solid lines of corresponding colour indicate the median value within each bin, from which individual subhaloes are drawn at random, so as to show the same number of objects per logarithmic mass interval. The black dashed and dotted lines denote a 1:1 ratio, and a ratio of $\Omega_{\rm DM}/(\Omega_{\rm DM}+\Omega_{\rm b})$, respectively. For lower mass subhaloes, the circles lie closer together, indicating a decrease in baryon fraction with decreasing subhalo mass, reflected also in the convergence of the three curves. The decline of the red curve indicates a decrease in total subhalo mass in GIMIC relative to the DMO simulation. For satellites, there is a similar trend of decreasing baryon fraction and mass ratio for lower mass objects, although the scatter among individual subhaloes is much greater.

subhaloes that are of the same type in both simulations, i.e. satellites matched to satellites and centrals matched to centrals, using the criteria discussed in Section 2.2.3. We show a random subset of subhaloes, normalized so as to have equal numbers per logarithmic mass interval.

Each subhalo is represented by a set of three points that denote the different composition of the subhalo in the GIMIC simulation: black circles measure only the DM component, green circles also include the gas, and red circles represent the total mass, consisting of DM, gas and stars.

Overplotted in both panels are the median mass ratios of the corresponding components; black lines for DM, green also including gas, and red including DM, gas and stars. As the baryon fraction decreases with decreasing subhalo mass, circles which are well separated at high mass become closer, and for lower mass subhaloes, they mostly overlap. This trend is reflected in the convergence of the median mass ratios of the three components. At the high-mass end, on average, subhaloes reach the same total mass in the GIMIC simulation as in the DMO simulation because their baryon fraction is close to the universal value. The total mass (red line) approaches a 1:1 ratio, while the DM component (black line) approaches the value expected from subtracting the universal baryon fraction. At lower masses, the median total mass in the GIMIC simulation is dominated by the DM component. It is worth noting that the mass of low-mass subhaloes in the GIMIC simulation is not only below the mass in the DMO simulation, but even below the value expected from a simple loss of baryons. This points to the fact that objects which lose mass during an early stage in their formation acquire a shallower potential, which also diminishes their subsequent accretion. Due to the requirement for subhaloes to be self-bound, a subhalo whose gravitational binding energy is reduced due to outflows may also lose additional particles that are no longer bound. However, as we show in Fig. 9, we find a very similar reduction in the masses of FoF haloes, indicating that any difference in the attribution of masses by SUBFIND is only a secondary effect.

Comparing the central subhaloes in the left-hand panel of Fig. 6 to the satellite subhaloes in the right-hand panel, it can be seen that the scatter increases substantially for satellites, and also increases with decreasing mass. As we will show in Fig. 12, the overall gas fraction is lower for satellites than for centrals as a result of stripping.

The main cause for the increased scatter among satellites lies in their more chaotic orbits, meaning that matched pairs are more likely to follow divergent evolutionary paths in the two simulations. This effect is illustrated in Fig. 7, which shows the projected displacement between pairs of satellite subhaloes of $\sim 10^{10}\,{\rm M}_\odot$ in the GIMIC simulation relative to the DMO simulation. Tracing the evolution since $z = 1.5$, it can be seen that satellite pairs whose final mass ratios deviate from the median by more than 2σ (top panel) are much more likely to evolve along divergent paths compared to those whose final mass ratios are within 10 per cent of the median (bottom panel).

Qualitatively, the overall reduction in subhalo mass due to baryonic effects at $z = 0$ is similar for centrals and satellites. However, despite the considerable scatter, at fixed subhalo mass, the average mass-loss in the GIMIC simulation appears to be somewhat larger for central subhaloes than for satellites, as is also illustrated in Figs 8 and A1. The offset between the relations seen for satellites and for centrals may be partly attributed to the fact that while centrals only experience mass-loss due to outflows in the GIMIC simulation, the masses of satellites are further reduced by stripping in both the DMO and the GIMIC simulations.

It is worth noting that, despite the fact that every halo contains at most one central subhalo but potentially many satellites, over 80 per cent of all subhaloes in the simulation volume are central, as shown in Fig. 3. This is easily understood when recognizing that within each halo, by definition, the central subhalo is more massive than the satellites, so at a fixed subhalo mass, satellites

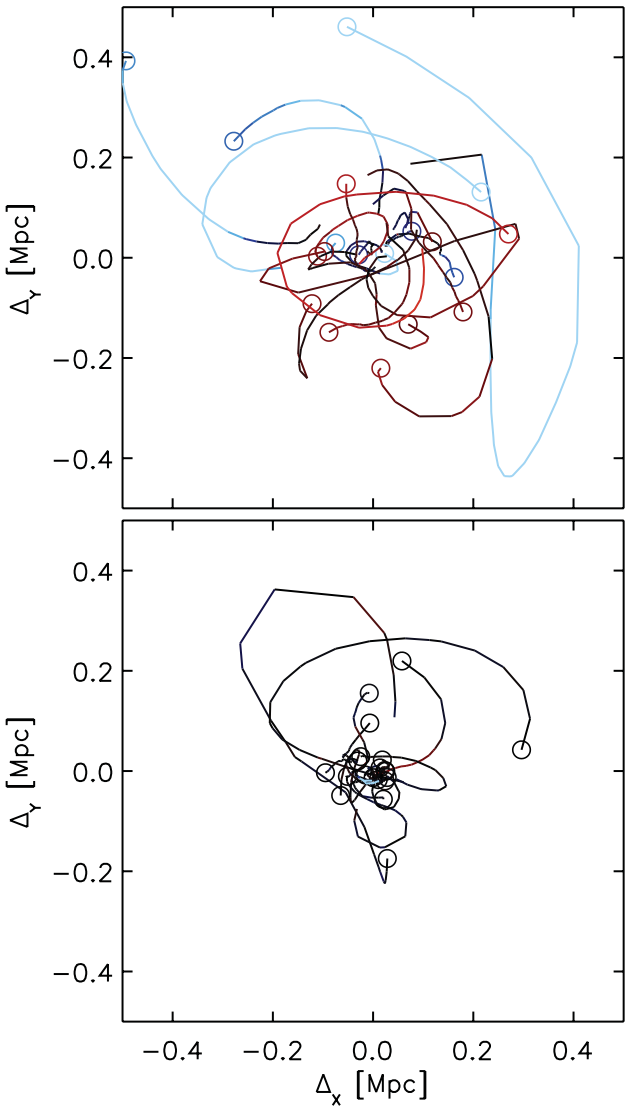

Figure 7. Relative displacement between matched pairs of satellite subhaloes with final masses of $\sim 10^{10}\,\mathrm{M}_\odot$, for pairs with mass ratios that deviate from the median by more than 2σ (top panel), or by less than 10 per cent (bottom panel). The circles show the present displacement, while the traces show the evolution from $z = 1.5$, coloured according to the mass ratio: black for mass ratios close to the median, red (blue) if the mass in GIMIC is higher (lower) than the median at the time. While most pairs of subhaloes in the bottom panel evolve along similar paths, those in the top panel show more divergent histories, linking the increased scatter among satellites in Fig. 6 to their more chaotic orbits.

belong to larger haloes than centrals. As a consequence of the steep halo mass function (see Fig. 9), the total subhalo population closely resembles that of centrals, with some additional scatter introduced by satellites.

3.2 Analytic correction

Noting from Fig. 6 that the change in mass ratio between the GIMIC and the DMO simulations saturates both at the high-mass and at the low-mass limit, we parametrize the mean ratio of the subhalo mass

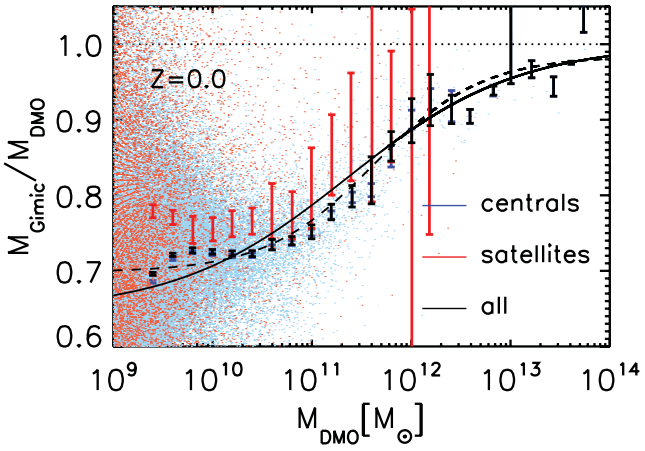

Figure 8. Ratio of subhalo masses between the GIMIC and DMO simulations for matched subhaloes at $z = 0$ compared to analytic fits. The blue and red points show the results for centrals and satellites, respectively, while the error bars show the estimate of the median ratio and its statistical error, for centrals (blue), satellites (red) and all subhaloes (black). The solid and dashed lines show the best fit to the median for all subhaloes, assuming four free parameters (dashed line, equation 1), or three free parameters, with the upper limit fixed at $b = 1$ (solid line, equation 2).

in the GIMIC simulation, $M_{\rm GIMIC}$, to that in the DMO simulation, $M_{\rm DMO}$, in the following way:

$$\frac{M_{\rm GIMIC}}{M_{\rm DMO}} = b \cdot \frac{a/b + (M_{\rm DMO}/M_{\rm t})^w}{1 + (M_{\rm DMO}/M_{\rm t})^w}. \quad (1)$$

With $w > 0$, equation (1) has two horizontal asymptotes,

$$\lim_{M_{\rm DMO} \ll M_{\rm t}} \left(\frac{M_{\rm GIMIC}}{M_{\rm DMO}} \right) = a$$

and

$$\lim_{M_{\rm DMO} \gg M_{\rm t}} \left(\frac{M_{\rm GIMIC}}{M_{\rm DMO}} \right) = b.$$

The mass ratio takes an intermediate value of $(a + b)/2$ at $M_{\rm DMO} = M_{\rm t}$, i.e. $M_{\rm t}$ sets the scale for the transition between the two limits, while the width of the transition is parametrized by the value of w. Assigning equal weight to the median ratio in each logarithmic mass interval, the best fit to all subhaloes at $z = 0$ yields values of $a = 0.69$, $b = 0.98$, $M_{\rm t} = 10^{11.6}\,\mathrm{M}_\odot$ and $w = 0.79$. It is shown by the dashed line in Fig. 8.

The asymptotic behaviour in the low-mass limit signifies an almost complete loss of baryons at very early times. The fact that the value of b at $z = 0$ is very close to unity indicates that the physical processes in GIMIC lead to no significant change in mass at the high-mass end. By setting $b = 1$, equation (1) simplifies to

$$\frac{M_{\rm GIMIC}}{M_{\rm DMO}} = \frac{a + (M_{\rm DMO}/M_{\rm t})^w}{1 + (M_{\rm DMO}/M_{\rm t})^w} \quad (2)$$

with horizontal asymptotes at a and 1, and an intermediate value of $(a + 1)/2$. A fit to this equation yields values of $a = 0.65$, $M_{\rm t} = 10^{11.4}\,\mathrm{M}_\odot$ and $w = 0.51$. The solid line in Fig. 8 shows the fit to equation (2), and can be compared to the dashed line obtained from equation (1).

Because the additional free parameter of equation (1) does not change the results significantly, we adopt the simpler equation (2) for the remainder of this work. It is possible that other physical processes, particularly the effect of AGN, could lead to a significant effect also at the high-mass end, in which case the more general

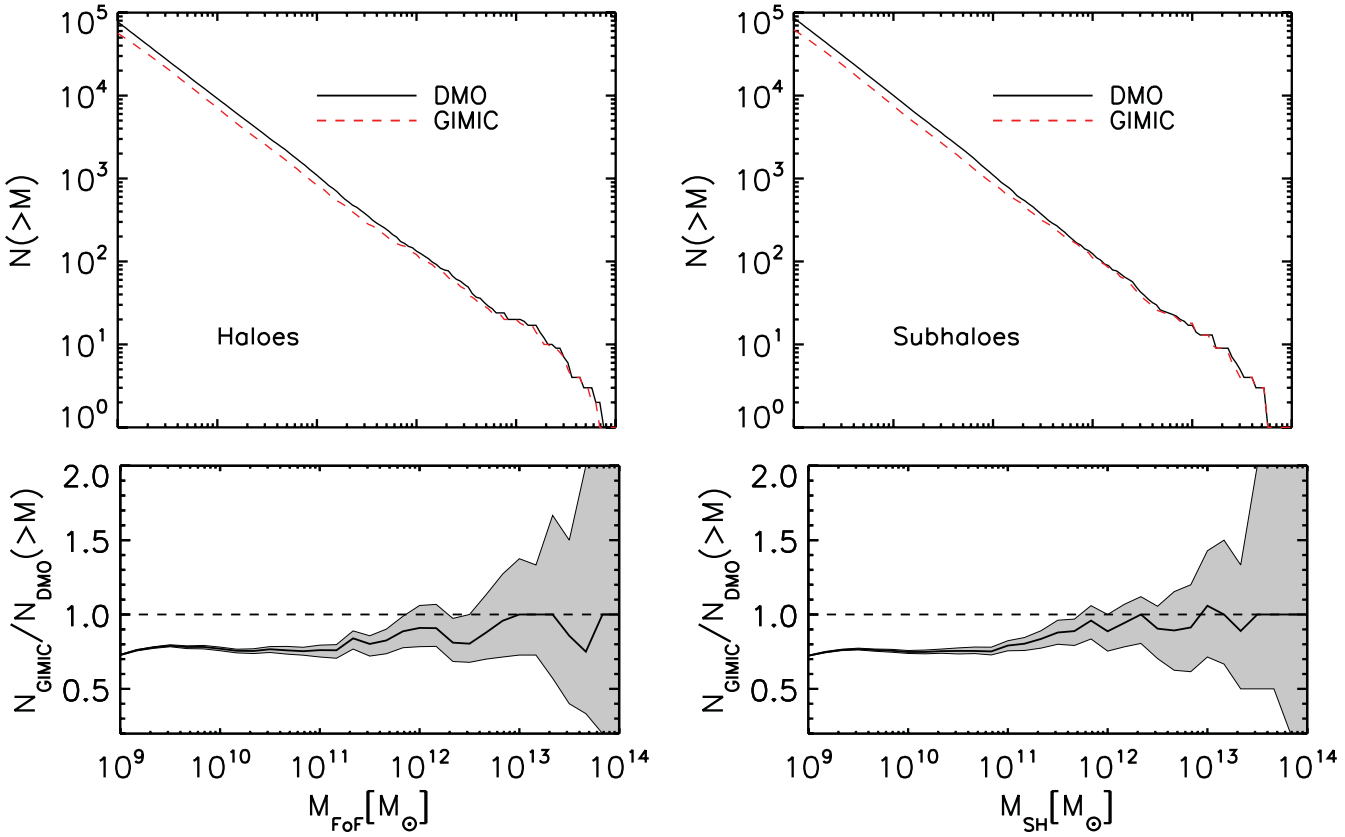

Figure 9. Comparison of the cumulative halo mass functions (left column) and subhalo mass functions (right column) of the GIMIC and the DMO simulations. The top row shows the cumulative mass functions at redshift $z = 0$, with the DMO results shown in black, and the GIMIC results overplotted in red. The bottom row shows the ratio of the cumulative abundances at $z = 0$, with the shaded area denoting the statistical uncertainty. The difference in abundance depends strongly on mass: above $\sim 10^{12}\,\mathrm{M}_\odot$, both the subhalo and halo abundances have ratios close to 1, but they drop to $\sim$0.75 from $\sim 10^{11}\,\mathrm{M}_\odot$ and below.

form of equation (1) would be more appropriate. Separate fits for satellites and centrals, as well as for different redshifts, are discussed in Appendix A.

3.3 Global mass functions

Mass or multiplicity functions of haloes or subhaloes constitute one of the most fundamental measures of structure formation. Because DM haloes (or more precisely: subhaloes) are believed to be the locations of galaxy formation, predicting their abundance also allows us to relate observations of galaxies to theories about the underlying mass distribution. We return to this point in Section 4.

In Fig. 9, we show the cumulative mass functions of haloes (left-hand panel) and subhaloes (right-hand panel), for the DMO simulation (shown in black) and the GIMIC simulation (red). In both cases, the mass includes the total mass, i.e. the mass in DM in the DMO simulation, and the combined mass of DM, gas and stars in the GIMIC simulation. The cumulative mass functions are nearly identical for the most massive objects, but below $\sim 10^{12}\,\mathrm{M}_\odot$, the abundance in the GIMIC simulation is significantly reduced relative to the DMO simulation, with the difference increasing at lower masses.

It should be noted that this statistical result is not reliant on the matching of haloes or subhaloes and, as we show in Appendix B, it is not affected by the resolution of our simulations. Above the resolution threshold, the decrease in numbers is only partly due to the erasure of structures, but mostly due to each object having a lower mass in the GIMIC simulation. The shape of the mass functions then results in this 'horizontal' shift in mass also manifesting itself as a 'vertical' shift in abundance.

3.4 Velocity functions

The circular velocity, $v_\mathrm{c}(r)$, at a given radius, r, is related to the enclosed mass $m(<r)$ via $v_\mathrm{c}(r) = \sqrt{Gm(<r)/r}$. Thus, measurements of the velocity profiles of H I gas in galaxies can provide a measurement of the underlying mass distribution. In Fig. 10, we compare the differential velocity function of subhaloes in the DMO simulation to results from the GIMIC simulation, and to measurements from the ALFALFA survey (Giovanelli et al. 2005, 2007). Following constrained simulations of the Local Volume by Zavala et al. (2009), we have normalized the simulated velocity functions to account for the overdensity in the survey volume, which is dominated by the Local Supercluster, and whose velocity function exceeds the universal value. A detailed description of the normalization can be found in Zavala et al. (2009), but we note that it has little effect on scales where the DMO results significantly exceed the observations.

When compared to the ALFALFA observations, it can be seen that the velocity function of the DMO simulation (black) overpredicts the abundance of subhaloes of $v_\mathrm{max} \sim 35\,\mathrm{km\,s^{-1}}$ by a factor of $\sim$10. This discrepancy, which has been pointed out by Zavala et al. (2009), Trujillo-Gomez et al. (2011) and Papastergis et al. (2011), persists when the observations are compared to the velocity function computed using all subhaloes in GIMIC (red). It has been

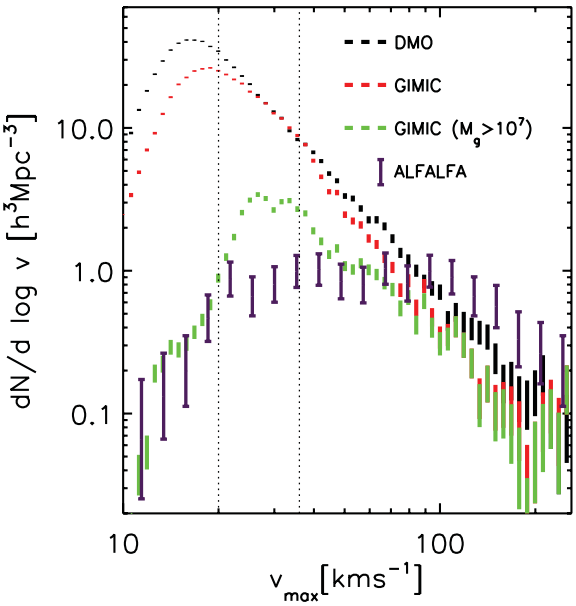

Figure 10. Differential velocity function of subhaloes in the DMO simulation and in the GIMIC simulation, compared to the H I velocity function measured in the ALFALFA survey for a detection limit of $M_{\rm H\,I} = 10^7\,h^{-1}\,{\rm M}_\odot$. The results of the DMO simulation are shown in black; for GIMIC, red shows the velocity function computed using all subhaloes, and green only includes those subhaloes with $M_{\rm g} > 10^7\,h^{-1}\,{\rm M}_\odot$. The simulated mass functions are corrected for the overdensity of the survey volume; their width denotes the Poisson errors within each bin. The vertical lines at 20 and 35 km s^{-1} indicate the resolution limit of our simulations and the completeness limit of the ALFALFA survey, respectively. While the DMO simulation significantly overpredicts the abundance of subhaloes with $v_{\rm max} = 35\,{\rm km\,s^{-1}}$, accounting for baryonic effects and adopting an appropriate detection threshold for the ALFALFA survey significantly reduce the reported discrepancies between CDM and the measured velocity function, down to the survey limit of 35 km s^{-1}.

considered a challenge to the CDM paradigm and evidence in favour of warm dark matter. However, allowing for the detection threshold of the survey for the sample of nearby galaxies analysed by Zavala et al. (2009), $M_{\rm H\,I} > 10^7\,h^{-1}\,{\rm M}_\odot$, and recomputing the GIMIC velocity function only for subhaloes with $M_{\rm g} > 10^7\,h^{-1}\,{\rm M}_\odot$, we find that the resulting velocity function (green symbols) agrees to within a factor of ∼3 with the ALFALFA observations, while also giving a better fit to the shape of the $v_{\rm max}$ function. Considering that the total gas mass, $M_{\rm g}$, is only an upper limit for $M_{\rm H\,I}$, and that we have not included other constraints such as the inclination limit of 30 for galaxies in the ALFALFA survey, the remaining difference may be yet be explained within the CDM framework.

3.5 Satellite mass functions

Local Group dwarf galaxies and Milky Way (MW) satellites constitute important test cases for CDM theory: they are among the smallest and most DM-dominated galaxies, and their proximity enables detailed studies of their abundance and their internal structure. While the apparent discrepancy between the predicted satellite mass function and the observed satellite *luminosity* function (e.g. Klypin et al. 1999; Moore et al. 1999) may be explained through inefficient star formation due to astrophysical mechanisms including photoionization and feedback (e.g. Bullock, Kravtsov & Weinberg 2000; Benson et al. 2002; Somerville 2002), the reported discrepancies between the predicted and measured satellite *mass* functions

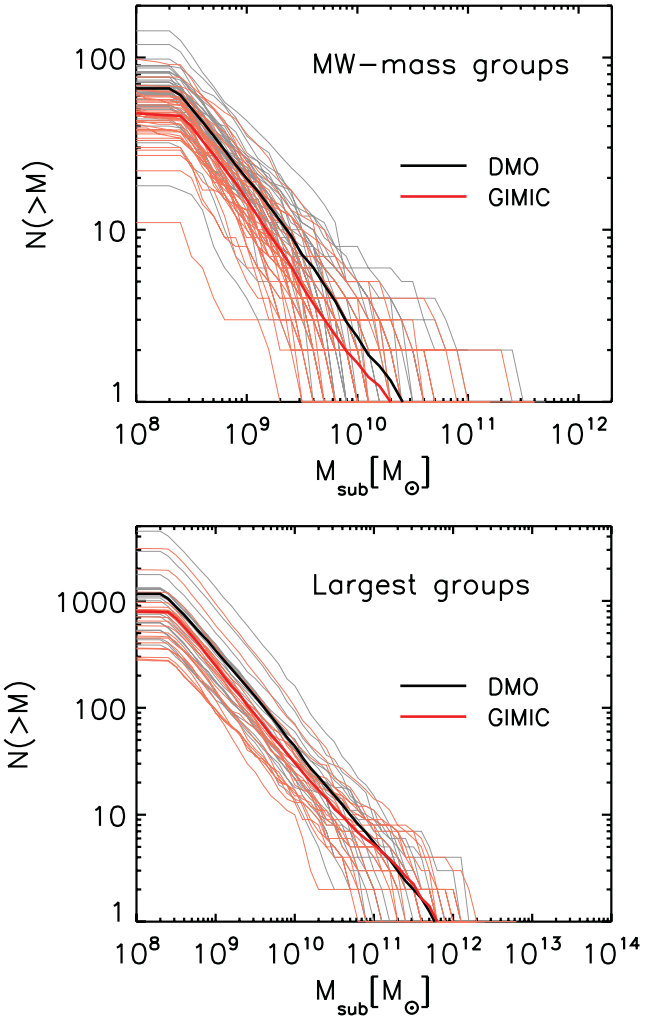

Figure 11. Cumulative satellite mass functions for ∼50 individual, MW-sized ($10^{12} < M_{200} < 2.5 \times 10^{12}\,{\rm M}_\odot$) haloes (top), and for the 20 most massive clusters (bottom), in the DMO simulation (black) and the GIMIC simulation (red), at $z = 0$. The thin lines show the mass functions of individual groups, and the thick lines show the average abundance for the groups in either simulation. For clusters and for MW groups, the number of members is reduced in the GIMIC simulation, as expected from the overall reduction in subhalo abundance. It is also worth noting that there is more than an order of magnitude variation in the number of satellites amongst individual groups.

are more challenging: ΛCDM simulations such as the Aquarius runs (Springel et al. 2008) produce a higher number of satellites with circular velocities above 50 km s^{-1} than is allowed by observations (Boylan-Kolchin et al. 2011, 2012; Parry et al. 2012). It has also been pointed out that the MW appears special in having two satellites as bright as the Magellanic Clouds (Busha et al. 2011; Tollerud et al. 2011; Guo et al. 2012).

In Fig. 11, we compare the satellite mass functions of groups of different masses in the DMO simulation (black lines) and the GIMIC simulation (red lines). In the top panel, in both simulations, we have selected groups with total mass in the range $1-2.5 \times 10^{12}\,{\rm M}_\odot$, compatible with current mass estimates of the MW (e.g. Li & White 2008; Xue et al. 2008; Gnedin et al. 2010). Each thin line shows the satellite mass function of an individual group, while the two thick lines represent the average cumulative subhalo abundance over all groups in the mass range in both simulations.

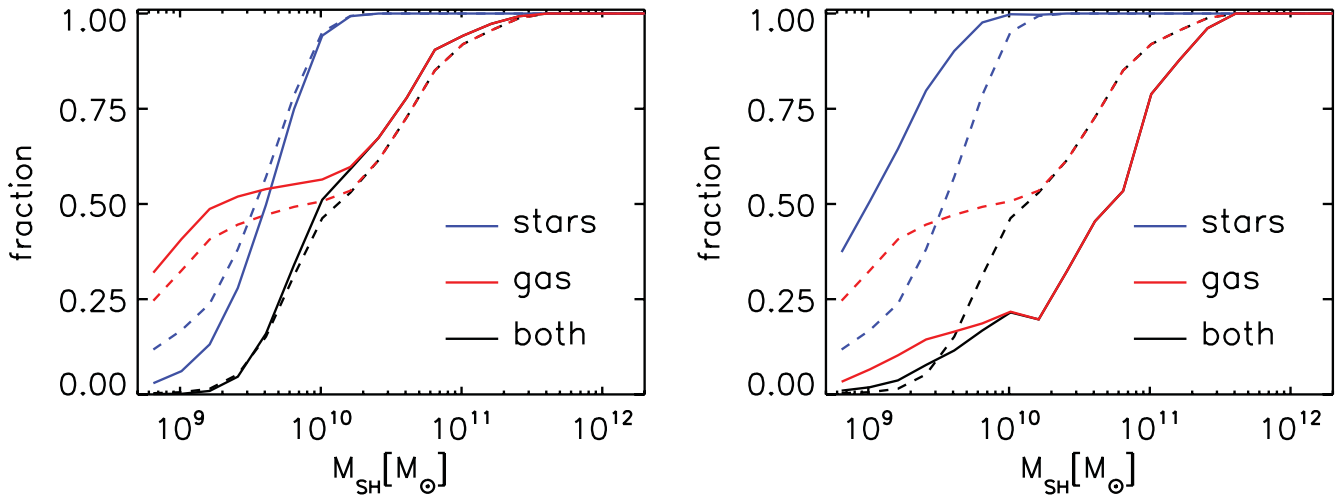

Figure 12. Fraction of subhaloes with stars (blue), with gas (red), or both (black), at $z = 0$. The solid lines show the results for centrals (left-hand panel) and satellites (right-hand panel); the dashed lines show the fraction of all subhaloes for reference and are identical in the two panels. For both centrals and satellites, the fraction of subhaloes with stars decreases for subhalo masses below $\sim 2 \times 10^{10}\,\mathrm{M}_\odot$, while the fraction of subhaloes with gas starts to decrease below $\sim 3 \times 10^{11}\,\mathrm{M}_\odot$. The main environmental dependence is in the gas content; while the fraction of central galaxies with gas remains above ∼50 per cent down to $10^9\,\mathrm{M}_\odot$, it drops below ∼25 per cent among satellites with a total mass of $\sim 2 \times 10^{10}\,\mathrm{M}_\odot$. The decrease in the fraction of subhaloes without stars is less steep for satellites than for centrals; this is because stars form in more bound objects, which are more likely to survive as satellites.

We find that, on average, the MW-mass groups contain fewer satellites in GIMIC than in the DMO simulation. The reduction in the number of satellites in haloes of a fixed mass can be understood because the mass-dependent baryonic effects are stronger in the lower mass satellites than in their more massive hosts.

In line with previous results by di Cintio et al. (2011), the average baryon effects we see in the GIMIC simulation alone are not enough to sufficiently reduce the number of *massive* satellites in order to explain the results of Boylan-Kolchin et al. (2011). However, in both the GIMIC and the DMO simulations, we also find a scatter of more than one order of magnitude in the satellite abundances of individual groups. In particular, among the sample of ∼50 MW-mass groups, we find several examples with satellite numbers almost an order of magnitude below the median. Confirming the results of Wang et al. (2012), we therefore find a considerable chance for a low number of massive satellites in a MW-mass halo, to which baryon effects contribute further.

In the bottom panel of Fig. 11, we repeat the same analysis for the satellite mass functions of the 20 most massive groups in both simulations; masses of the clusters are in the range $\sim 10^{13}$–$10^{14}\,\mathrm{M}_\odot$. As expected from the global mass functions shown in Fig. 9, at the high-mass end, the average satellite mass functions of clusters show no systematic difference between the two simulations, a result noted previously by Dolag et al. (2009). However, at lower masses, clusters also have fewer satellites in GIMIC than in the DMO simulation. Compared to the satellite mass function of MW-sized groups shown in the top panel of Fig. 11, the satellite mass functions of individual clusters show less scatter, even though the mass range of clusters shown is larger than that of MW-sized groups.

3.6 Dark subhaloes

At $z = 0$, the GIMIC simulation contains a substantial number of low-mass subhaloes which are completely free of stars or gas, i.e. that they do not contain a single (bound) star or gas particle of mass $1.98 \times 10^6\,\mathrm{M}_\odot$. *Dark* subhaloes are predicted to exist in ΛCDM, as they may explain the ratio of the predicted number of substructures to the observed number of dwarf galaxies (Moore et al. 1999), and because astrophysical processes including reionization prevent cooling and star formation in the smallest systems. Likewise, very low mass galaxies like dwarf spheroidals contain stars but no detectable gas, which is understood to have been lost either through supernova feedback, reionization (if stars formed earlier), stripping or a combination of these effects.

In Fig. 12, we show the fraction of low-mass subhaloes in the GIMIC simulation containing some amount of stars (blue lines), gas (red lines) or both (black lines), as a function of total subhalo mass. For both centrals (left-hand panel) and satellites (right-hand panel), we find that the fraction of *dark* subhaloes increases with decreasing mass below $\sim 10^{10}\,\mathrm{M}_\odot$ (i.e. $\sim 10^3$ particles). Intriguingly, the fraction of subhaloes with stars decreases less steeply with mass among satellites compared to centrals; while ∼50 per cent of satellites with a total mass of $10^9\,\mathrm{M}_\odot$ contain stars, the fraction is below 10 per cent for centrals of the same mass at $z = 0$. This difference can be understood from the effective threshold in binding energy required for star formation, apparent from Fig. 13: subhaloes that are more concentrated, and therefore have a higher maximum circular velocity v_{max}, are both more likely to cool gas efficiently to allow star formation and also less likely to be destroyed by stripping. As can be seen in Fig. 13, a v_{max} threshold of $\sim 30\,\mathrm{km\,s^{-1}}$, rather than a threshold in total mass, separates subhaloes that contain stars from those that do not. In preferentially destroying less bound objects, tidal disruption leads to a higher fraction of subhaloes with stars among the surviving satellites relative to the fraction among centrals, which suffer weaker tidal forces. Our result is in agreement with simulations of Crain et al. (2007), who found that photoheating can prevent the collapse of baryons in systems with circular velocities below $35\,\mathrm{km\,s^{-1}}$ (at $z = 0$), and with the high-resolution simulations of Okamoto, Gao & Theuns (2008), who found that haloes with a circular velocity of $25\,\mathrm{km\,s^{-1}}$ can lose half of their baryons due to photo-heating, and are unable to accrete and cool gas efficiently.

The fraction of subhaloes with gas decreases from $\sim 10^{11}\,\mathrm{M}_\odot$ downwards, and decreases much faster for satellites than for

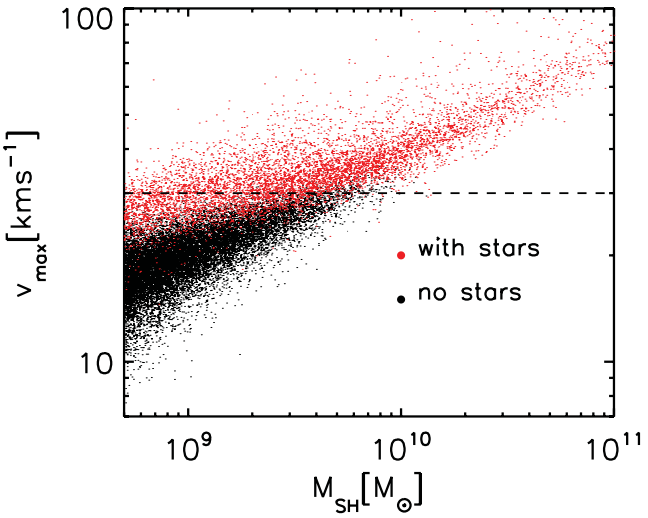

Figure 13. Distribution of maximum circular velocities, v_{max}, of a random sample of low-mass subhaloes in the GIMIC simulation, as a function of subhalo mass. Subhaloes that contain stars are shown in red, and those without stars are shown in black. At masses below $\sim 10^{10}$, it is apparent that a v_{max} of $\sim 30\,\mathrm{km\ s^{-1}}$ splits the two populations: subhaloes with $v_{max} > 30\,\mathrm{km\ s^{-1}}$ have typically been able to form stars, while most subhaloes with lower v_{max} have not.

centrals. While stars in satellites are concentrated towards the centre, gas is easily removed from satellites by tidal and ram-pressure stripping (e.g. McCarthy et al. 2008). At $10^{10}\,\mathrm{M}_\odot$, where most subhaloes contain stars, the fraction with gas is ∼55 per cent among centrals, and only ∼20 per cent among satellites.

We also find that the fraction of subhaloes without gas and the fraction without stars are correlated, but not proportional. The black lines in Fig. 12 show the fraction of subhaloes that contain both stars *and* gas. It can be seen that central galaxies with total mass below $\sim 3 \times 10^9\,\mathrm{M}_\odot$ typically only contain either gas *or* stars: subhaloes that undergo star formation lose their gas due to feedback, while those that never form stars due to a lack of cooling can typically retain their gas.

As subhaloes without stars are not directly observable, their presence has important consequences for abundance matching, that will be addressed in Section 4. Of course, the finite particle mass of $m_g = 1.98 \times 10^6\,\mathrm{M}_\odot$ sets a natural lower limit to the baryonic mass detectable in our simulation. However, the occurrence of *dark* subhaloes at relatively high subhalo masses, and the dependence on v_{max} and environment, rather than on subhalo mass, indicates the physical origin of their existence. As we will discuss in Appendix B, while the gas-free subhaloes are well converged, predictions for the exact fraction of *dark* subhaloes are still affected by resolution in the GIMIC simulation.

4 ABUNDANCE MATCHING

Abundance matching (Yang, Mo & van den Bosch 2003; Vale & Ostriker 2006; Guo et al. 2010; Moster et al. 2010) is a simple but powerful method to statistically link galaxies to DM haloes, without making detailed assumptions about the physics of galaxy formation. By assuming a monotonic relationship (with a limited amount of scatter) between an observed property such as stellar mass and a simulated property such as total mass, and equating their cumulative abundances, it allows a determination of quantities like the stellar-to-total mass ratio as a function of mass. In integral form, the relation (for stellar mass) can be expressed, in the absence of scatter, as follows:

$$\int_{M_{\mathrm{h,max}}}^{M_{\mathrm{h,min}}} N_h(m)\,\mathrm{d}m = \int_{M_{\star,\mathrm{max}}}^{M_{\star,\mathrm{min}}} N_\star(m)\,\mathrm{d}m, \tag{3}$$

where $N_h(m)$ and $N_\star(m)$ are the subhalo and stellar mass functions. The upper limits of the subhalo and stellar mass ranges, $M_{\mathrm{h,\,max}}$ and $M_{\star,\,\mathrm{max}}$, are fixed, while the respective lower limits, $M_{\mathrm{h,\,min}}$ and $M_{\star,\,\mathrm{min}}$, are chosen such that the relation holds over the entire interval. Equivalently, abundance matching corresponds to assigning the most massive observed galaxy to the most massive (sub)halo, and stepping down towards the minimum galaxy mass or subhalo mass that can still be matched, which determine $M_{\mathrm{h,\,min}}$ and $M_{\star,\,\mathrm{min}}$.

4.1 Abundance matching with baryons

The effect of baryons on the abundance of subhaloes shown in Fig. 9 has consequences for the results of abundance matching. Furthermore, as described in Section 3.6, an increasing fraction of low-mass subhaloes does not host any stars. Because these would not be observable, we propose to modify equation (3) further, by multiplying the (sub)halo mass function by a completeness term $0 \le f_\star(m) \le 1$, to give the *reduced* subhalo mass function (RMF) of *observable* objects:

$$\int_{M_{\mathrm{h,max}}}^{M_{\mathrm{h,min}}} f_\star(m) N_h(m)\,\mathrm{d}m = \int_{M_{\star,\mathrm{max}}}^{M_{\star,\mathrm{min}}} N_\star(m)\,\mathrm{d}m. \tag{4}$$

Unlike a simple scatter term, dropping the assumption that every simulated DM subhalo with $m > M_{\mathrm{h,\,min}}$ hosts a galaxy has the effect of extending the domain to lower mass subhaloes, and assigning higher stellar masses to subhaloes above the new limit.

For the purpose of abundance matching, the RMF differs depending on the observable: for a stellar mass survey, the RMF includes only subhaloes that contain stars, while for an H I survey (e.g. Zavala et al. 2009), it includes those galaxies that contain gas, or both stars and gas.

From the ratio of the subhalo mass functions, we compute the subhalo mass in GIMIC with the same cumulative abundance as a subhalo of a given mass in the DMO simulation. Fig. 14 shows the relations of equating the DMO subhalo mass function to the total subhalo mass function in GIMIC (blue line), and to the RMF in GIMIC (red line), where subhaloes without stars have been removed. These relations can be applied directly to abundance matching results based on a DMO simulation, to derive stellar-to-total mass ratios in a universe that contains baryons.

Applying our results to the abundance matching performed by Guo et al. (2010), we note that the GIMIC simulation uses the same cosmological parameters as the Millennium Simulations studied by Guo et al. However, while the total abundance of (sub)haloes depends on the slope and the normalization of the power spectrum (e.g. Vale & Ostriker 2006), within the CDM paradigm, the effect of baryons should be similar.

In line with other authors, Guo et al. have computed the subhalo mass function as a combination of the present-day mass for centrals and the mass before infall for satellites. In adopting their result for the DMO case which forms the baseline of all of our abundance matching results, we use the same prescription, but we consider the ratio of the abundance at $z = 0$ when we apply the baryon effects. As shown in Fig. A1, the average mass ratio shows little evolution between $z = 1$ and 0, suggesting that the difference between applying the correction at infall and at $z = 0$ would be negligible.

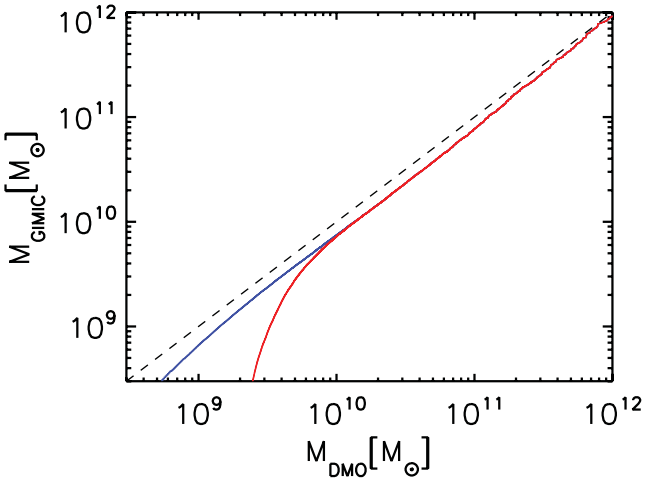

Figure 14. Mass of subhaloes in the GIMIC simulation and in the DMO simulation, at *equal cumulative subhalo abundance*. The blue line includes all subhaloes, while the red line assumes the *reduced* mass function, with *dark* subhaloes removed. The dashed line provides a reference for a 1:1 abundance ratio. Below $10^{12}\,M_\odot$, the mass attributed to subhaloes of equal abundance is considerably lower in the GIMIC simulation compared to the DMO simulation, and the occurrence of *dark* subhaloes becomes the dominant factor below $\sim 4 \times 10^9\,M_\odot$.

In Fig. 15, we show the effect of baryons on the results of abundance matching, in terms of the inferred stellar mass (left-hand panel) and the stellar-to-total mass ratio (right-hand panel). For reference, black lines show the DMO case; the solid line denotes the original results of Guo et al. (2010), while the dashed line reflects a change in the faint-end slope of the stellar mass function from $\alpha = -1.15$ (Li & White 2009) to $\alpha = -1.58$ (Baldry, Glazebrook & Driver 2008). The set of blue lines shows the corresponding relations, but taking into account the baryonic effects on the subhalo mass function obtained from the GIMIC simulation. The set of red lines uses the RMF, which includes baryonic effects, and additionally has the *dark* subhaloes removed from the subhalo mass function.

The effect of baryons on the attributed stellar mass strongly depends on the subhalo mass. It is insignificant at subhalo masses above $10^{12}\,M_\odot$, but appreciable at the faint end. Purely due to the inclusion of baryons, the stellarto-total mass ratio increases by a factor of $\sim$1.5 for subhaloes of $10^{11}\,M_\odot$, and by a factor of $\sim$2.2 for subhaloes of $10^{10}\,M_\odot$ and below. If the RMF is considered, stellar masses assigned to subhaloes below $10^{10}\,M_\odot$ increase further, and the mean stellar-to-total mass ratio reaches a finite minimum at $1 : 10^3$– $1 : 10^5$, depending on the slope of the stellar mass function.

The apparent discrepancy in the stellar-to-total mass ratios inferred from abundance matching and state-of-the-art simulations of dwarf galaxies (Sawala et al. 2011, and references therein), as well as kinematics of individual dwarf galaxies (Ferrero et al. 2012), has been described as a challenge to the CDM paradigm. Included in the left-hand panel of Fig. 15 are observational estimates for seven classical dwarf spheroidals, with stellar masses adopted from Misgeld & Hilker (2011), and virial masses derived by Peñarrubia et al. (2008).

It is worth pointing out that at such low masses, a direct comparison to observations should be taken with a grain of salt: mass measurements rely on assumptions about the shape and anisotropy of the DM halo; the halo mass is measured at present rather than at infall, and the limited sample may be biased towards the objects with higher stellar mass [indeed, Wolf et al. (2010) show that despite stellar masses extending to $\sim 10^4\,M_\odot$, most MW satellites are compatible with a formation in $\sim 10^9\,M_\odot$ haloes]. Nevertheless, it can be seen that when the CDM subhalo mass function is corrected to take into account baryonic effects, the stellar masses inferred from abundance matching become consistent with observations of individual dwarf galaxies.

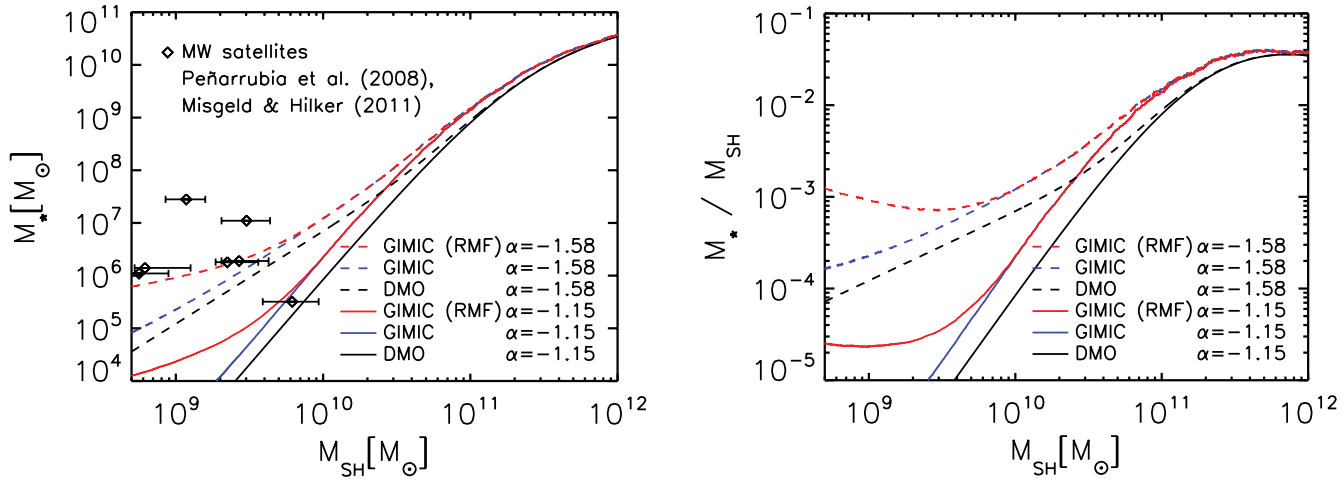

Figure 15. Stellar mass as a function of subhalo mass (left-hand panel), and ratio of stellar mass to subhalo mass (right-hand panel). The black lines are adapted from Guo et al. (2010), using the halo mass function of the Millennium I and II DMO simulations. The blue lines are the result of taking into account the baryonic effects on the subhalo mass function measured in GIMIC, as shown in Fig. 9. The red lines are obtained using the RMF of GIMIC, with *dark* subhaloes removed. The solid lines correspond to a faint-end slope of the stellar mass function of $\alpha = -1.15$ (Li & White 2009), while the dashed lines assume $\alpha = -1.58$. For comparison, we show stellar masses (Misgeld & Hilker 2011) and dynamical masses (Peñarrubia, McConnachie & Navarro 2008) of MW satellites. The inclusion of baryons doubles the inferred stellar-to-total mass ratio for subhaloes below $10^{11}\,M_\odot$, while the effect of the RMF further increases the inferred ratios by up to an order of magnitude for subhaloes in the range 10^9–$10^{10}\,M_\odot$. Combined with a steep faint-end slope for the stellar mass function, the reported discrepancy between abundance matching and measurements of individual galaxies is alleviated. The RMF also results in a finite minimum stellar-to-total mass ratio in subhaloes below $\sim 10^9\,M_\odot$.

Tikhonov & Klypin (2009) have likewise argued that the velocity function of voids in the Local Volume presents a challenge to CDM. The observed number of voids is so high that only haloes down to masses of $(6-8) \times 10^9$ or $(1-2) \times 10^{10}\,\mathrm{M}_\odot$ (for values of σ_8 of 0.75 and 0.9, respectively) can be populated by galaxies brighter than $M_B = -12$, but individual dwarf galaxies are observed with H I rotational velocities below $v_{\rm max} \sim 25\,\mathrm{km\,s^{-1}}$ indicative of halo masses closer to $10^9\,\mathrm{M}_\odot$. Our results suggest that this discrepancy may also be resolved by baryonic processes. As shown in Fig. 14, the mass limits derived by Tikhonov & Klypin (2009) based on the abundance of haloes in DMO simulations would have to be reduced in a universe where baryons affect the masses of low-mass haloes. In addition, the appearance of *dark* subhaloes discussed in Section 3.6 instead of a sharp threshold in (sub)halo mass can reconcile the observation of a population of galaxies in lower mass haloes with theoretical predictions.

5 DISCUSSION

We have compared a set of high-resolution cosmological *N*-body simulations with and without baryons to study the effect that baryonic processes have on structure formation in a ΛCDM universe. While both simulations agree well on large scales, objects below $\sim 10^{12}\,\mathrm{M}_\odot$ have systematically lower masses in the GIMIC simulation that includes baryons compared to its 'DMO' counterpart. Consequently, the cumulative abundance of haloes and subhaloes at every mass below $\sim 10^{12}\,\mathrm{M}_\odot$ is reduced in the GIMIC simulation. Given that the Universe includes baryonic processes such as gas pressure, cooling, reionization, star formation and supernova feedback, and assuming that they are represented at least approximately in the GIMIC simulation, it follows that DMO simulations *overpredict* the true abundance of substructures, by ∼10 per cent at $10^{11.5}\,\mathrm{M}_\odot$, ∼20 per cent at $10^{11}\,\mathrm{M}_\odot$ and ∼30 per cent at $10^{10}\,\mathrm{M}_\odot$.

The difference in halo and subhalo abundance has consequences for all analyses that take their mass functions as a starting point. For example, the stellar-to-total mass ratio of dwarf galaxies with stellar masses of $\sim 10^5$–$10^9\,\mathrm{M}_\odot$ inferred from subhalo abundance matching doubles after taking into account the lower mass of subhaloes, resulting from baryonic effects. If, in addition, the *reduced* mass function accounts for the fact that not all subhaloes contain stars, stellar-to-total mass ratios increase further, particularly for galaxies with stellar masses below $\sim 10^6\,\mathrm{M}_\odot$, i.e. the classical dwarf spheroidals. It appears that both effects are needed in order to reconcile abundance matching results with kinematic estimates of mass-to-light ratios of dwarf galaxies (Ferrero et al. 2012), or current high-resolution simulations (e.g. Sawala et al. 2011).

We conclude that the discrepancy between dwarf galaxy abundances and kinematics, and the results of CDM simulations paired with abundance matching, may be due to the inability of DMO simulations to capture the physical effects relevant on small scales, aggravated by the assumption that a visible galaxy forms inside every single DM (sub)halo. Likewise, taking into account baryonic effects and excluding subhaloes that fall below the observational detection threshold significantly reduce the reported discrepancy between the ΛCDM paradigm and the velocity function measured in the ALFALFA H I survey (Zavala et al. 2009; Papastergis et al. 2011).

For haloes more massive than $\sim 10^{12}\,\mathrm{M}_\odot$, our results corroborate those of previous works (e.g. Dolag et al. 2009; Cui et al. 2012), who concluded that the net effect of gas loss via stripping and feedback is largely counterbalanced by the increased concentration of the stellar component. We cannot exclude the possibility, however, that as suggested by van Daalen et al. (2011), efficient AGN feedback would also change the results on these scales; indeed McCarthy et al. (2012b) have reported overcooling in galaxies with stellar masses above $10^{11.5}\,\mathrm{M}_\odot$ in the GIMIC simulation. While there are also differences in the physics model, most notably the inclusion of photoionization in the GIMIC simulation that suppresses star formation in small objects, the most significant difference compared to previous works is the increased resolution, which has allowed us to study objects more than two orders of magnitude lower in mass. As we show, the baryonic results are strongly mass dependent. At low masses, re-ionization, feedback and stripping outweigh adiabatic contraction, leading to a net mass-loss. The GIMIC results also confirm the results of high-resolution resimulations of individual DM haloes with baryons (Sawala et al. 2011), which showed a similar reduction in mass due to efficient outflows from $\sim 10^{10}\,\mathrm{M}_\odot$ haloes.

Quantitatively, our results will likely be refined by future simulations at still higher resolution and with more complete physics models. In particular, our result for the fraction of *dark* subhaloes straddles the resolution limit of our simulations. However, the requirements of an RMF to match the faint end of observable galaxies, of taking into account the gas fraction when comparing to H I surveys and of including baryon physics on the underlying mass distribution of (sub)haloes appear to be largely model independent. To the precision achievable with our simulations, the failures previously attributed to the CDM paradigm appear to be largely attributable to a neglect of baryons in DMO simulations.

Overall, we believe that the GIMIC simulation, which has already been demonstrated to successfully reproduce many aspects of galaxy formation (Crain et al. 2010; Font et al. 2011; McCarthy et al. 2012a,b), is a much closer approximation to the observable Universe than a DMO model. With current and upcoming surveys such as ALFALFA, GAMA and GAIA pushing the observations to even fainter limits, simulations must not only increase in resolution, but will also have to take baryons into account, if they are to resolve the scales required to distinguish alternative DM models, or to model galaxy formation on sub-MW scales.

While it seems somewhat unsatisfactory to modify apparently 'assumption-free' and elegant methods like abundance matching and to replace DM simulations with complex, and in many ways uncertain astrophysical simulations, it was clear from the outset that a 'DMO universe' itself was merely an assumption of convenience. On small scales, it falls short.

ACKNOWLEDGEMENTS

We are indebted to Ian McCarthy, who developed and performed some of the original GIMIC simulations. We are very grateful to Q. Guo for providing the abundance matching data. TS acknowledges a fellowship by the European Commission's Framework Programme 7, through the Marie Curie Initial Training Network Cosmo-Comp (PITN-GA-2009-238356). CSF acknowledges a Royal Society Research Merit Award and ERC Advanced Investigator Grant COSMIWAY. Some simulations used in this paper were performed on the Cosmology Machine Supercomputer at the ICC, which is part of the DiRAC Facility jointly funded by STFC, the Large Facilities Capital Fund of BIS, and Durham University. This work was supported in part by an STFC rolling grant to the ICC and by the National Science Foundation under Grant No. NSF PHY11-25915. JZ is supported by the University of Waterloo and the Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada

and by the Province of Ontario through the Ministry of Research and Innovation. JZ acknowledges financial support by a CITA National Fellowship.

REFERENCES

Abadi M. G., Navarro J. F., Fardal M., Babul A., Steinmetz M., 2010, MNRAS, 407, 435

Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945

Benson A. J., Cole S., Frenk C. S., Baugh C. M., Lacey C. G., 2000, MNRAS, 311, 793

Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002, MNRAS, 333, 177

Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587

Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645

Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, L40

Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2012, MNRAS, 422, 1203

Bryan S. E., Mao S., Kay S. T., Schaye J., Dalla Vecchia C., Booth C. M., 2012, MNRAS, 422, 1863

Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517

Busha M. T., Wechsler R. H., Behroozi P. S., Gerke B. F., Klypin A. A., Primack J. R., 2011, ApJ, 743, 117

Chabrier G., 2003, PASP, 115, 763

Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781

Conroy C., Wechsler R. H., 2009, ApJ, 696, 620

Crain R. A., Eke V. R., Frenk C. S., Jenkins A., McCarthy I. G., Navarro J. F., Pearce F. R., 2007, MNRAS, 377, 41

Crain R. A. et al., 2009, MNRAS, 399, 1773

Crain R. A., McCarthy I. G., Frenk C. S., Theuns T., Schaye J., 2010, MNRAS, 407, 1403

Croton D. J. et al., 2006, MNRAS, 365, 11

Cui W., Borgani S., Dolag K., Murante G., Tornatore L., 2012, MNRAS, 423, 2279

Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

di Cintio A., Knebe A., Libeskind N. I., Yepes G., Gottlöber S., Hoffman Y., 2011, MNRAS, 417, L74

Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, MNRAS, 405, 2161

Ferrero I., Abadi M. G., Navarro J. F., Sales L. V., Gurovich S., 2012, MNRAS, 425, 2817

Font A. S., McCarthy I. G., Crain R. A., Theuns T., Schaye J., Wiersma R. P. C., Dalla Vecchia C., 2011, MNRAS, 416, 2802

Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, ApJ, 327, 507

Giovanelli R. et al., 2005, AJ, 130, 2613

Giovanelli R. et al., 2007, AJ, 133, 2569

Gnedin O. Y., Brown W. R., Geller M. J., Kenyon S. J., 2010, ApJ, 720, L108

Gnedin O. Y., Ceverino D., Gnedin N. Y., Klypin A. A., Kravtsov A. V., Levine R., Nagai D., Yepes G., 2011, arXiv:e-prints

Governato F. et al., 2010, Nat, 463, 203

Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111

Guo Q., Cole S., Eke V., Frenk C., 2012, MNRAS, 427, 428

Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays. Commissariat à l'Energie Atomique, France

Kaiser N., 1984, ApJ, 284, L9

Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201

Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82

Li Y.-S., White S. D. M., 2008, MNRAS, 384, 1459

Li C., White S. D. M., 2009, MNRAS, 398, 2177

McCarthy I. G., Frenk C. S., Font A. S., Lacey C. G., Bower R. G., Mitchell N. L., Balogh M. L., Theuns T., 2008, MNRAS, 383, 593

McCarthy I. G., Font A. S., Crain R. A., Deason A. J., Schaye J., Theuns T., 2012a, MNRAS, 420, 2245

McCarthy I. G., Schaye J., Font A. S., Theuns T., Frenk C. S., Crain R. A., Dalla Vecchia C., 2012b, 427, 379

Misgeld I., Hilker M., 2011, MNRAS, 414, 3699

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution. Cambridge Univ. Press, Cambridge

Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903

Navarro J. F., Eke V. R., Frenk C. S., 1996, MNRAS, 283, L72

Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920

Papastergis E., Martin A. M., Giovanelli R., Haynes M. P., 2011, ApJ, 739, 38

Parry O. H., Eke V. R., Frenk C. S., 2009, MNRAS, 396, 1972

Parry O. H., Eke V. R., Frenk C. S., Okamoto T., 2012, MNRAS, 419, 3304

Peñarrubia J., McConnachie A. W., Navarro J. F., 2008, ApJ, 672, 904

Pontzen A., Governato F., 2012, MNRAS, 421, 3464

Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14

Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, ApJ, 672, 19

Sales L. V., Navarro J. F., Theuns T., Schaye J., White S. D. M., Frenk C. S., Crain R. A., Dalla Vecchia C., 2012, MNRAS, 423, 1544

Sawala T., Guo Q., Scannapieco C., Jenkins A., White S., 2011, MNRAS, 413, 659

Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210

Schaye J. et al., 2010, MNRAS, 402, 1536

Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, ApJ, 546, 20

Seljak U., 2000, MNRAS, 318, 203

Sellwood J. A., McGaugh S. S., 2005, ApJ, 634, 70

Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020

Simha V., Weinberg D., Dave R., Fardal M., Katz N., Oppenheimer B. D., 2012, MNRAS, 423, 3458

Somerville R. S., 2002, ApJ, 572, L23

Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087

Springel V., 2005, MNRAS, 364, 1105

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

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

Springel V. et al., 2008, MNRAS, 391, 1685

Strigari L. E., Frenk C. S., White S. D. M., 2010, MNRAS, 408, 2364

Theuns T., Schaye J., Haehnelt M. G., 2000, MNRAS, 315, 600

Tikhonov A. V., Klypin A., 2009, MNRAS, 395, 1915

Tollerud E. J., Boylan-Kolchin M., Barton E. J., Bullock J. S., Trinh C. Q., 2011, ApJ, 738, 102

Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, ApJ, 742, 16

Vale A., Ostriker J. P., 2006, MNRAS, 371, 1173

van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649

Walker M. G., Peñarrubia J., 2011, ApJ, 742, 20

Wang J., Frenk C. S., Navarro J. F., Gao L., Sawala T., 2012, MNRAS, 424, 2715

Weinberg D. H., Colombi S., Davé R., Katz N., 2008, ApJ, 678, 6

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574

Wolf J., Bullock J. S., 2012, arXiv:e-prints

Wolf J., Martinez G. D., Bullock J. S., Kaplinghat M., Geha M., Muñoz R. R., Simon J. D., Avedo F. F., 2010, MNRAS, 406, 1220

Xue X. X. et al., 2008, ApJ, 684, 1143

Yang X., Mo H. J., van den Bosch F. C., 2003, MNRAS, 339, 1057

Zavala J., Jing Y. P., Faltenbacher A., Yepes G., Hoffman Y., Gottlöber S., Catinella B., 2009, ApJ, 700, 1779

APPENDIX A: MASS CORRECTION FOR CENTRALS AND SATELLITES

In Section 3.2, we parametrized the average relative change of a subhalo's mass from the DMO to the GIMIC simulation in equation (2):

$$\frac{M_{\rm GIMIC}}{M_{\rm DMO}} = \frac{a + (M_{\rm DMO}/M_{\rm t})^{w}}{1 + (M_{\rm DMO}/M_{\rm t})^{w}}.$$

In Fig. A1, we show the ratio of subhalo masses for matched pairs of satellites, centrals, and the combination of all subhaloes at four different redshifts, from $z = 6$ to 0. Allowing a, $M_{\rm t}$ and w to vary freely, we fit equation (2) to the different populations of subhaloes, giving equal weight to the median ratio within each logarithmic mass bin. It can be seen that for central subhaloes, the lower limit does not evolve strongly with redshift from $z = 1$, although the mass $M_{\rm t}$ at which the mass ratio reaches the intermediate value of $(a + 1)/2$ rises from $10^{11.2}\,{\rm M}_\odot$ at $z = 1$ to $10^{11.6}\,{\rm M}_\odot$ at $z = 0$. At $z = 6$, the reduction in subhalo mass is smaller.

We noted in Section 3.1 that at a fixed subhalo mass at $z = 0$, the average difference in mass between the GIMIC simulation and the DMO simulation is slightly less for satellites than for centrals, which may be attributed to the fact that satellites also experience tidal effects, which are similar in both simulations. A pair of isolated subhaloes that have evolved to a given mass ratio before infall, and whose mass is subsequently reduced in equal proportion by tidal effects in both simulations, would evolve to the same mass ratio, but for a lower total mass. Fig. A1 shows that a difference also exists at higher redshifts, where the decrease in mass ratio appears at higher masses for satellites than for

Table A1. Coefficients of equation (2) for fits to the median mass ratios of individual, matched subhaloes in the GIMIC relative to the DMO simulation. For each redshift, the rows show the result of different sets of subhaloes, as defined in Section 2.2.3. Note that at high redshift, scatter dominates the mass ratio for satellites, but the total population remains dominated by centrals.

z	Type	a	$\log_{10} M_t$	w
0.0	All	0.65	11.37	0.51
	Centrals	0.66	11.54	0.59
	Satellites	0.68	11.20	0.50
0.5	All	0.63	11.20	0.65
	Centrals	0.64	11.38	0.47
	Satellites	0.75	11.14	9
1.0	All	0.63	11.20	0.41
	Centrals	0.62	11.20	0.40
	Satellites	0.78	11.07	5
6.0	All	0.77	11.18	1.95
	Centrals	0.77	11.20	1.58
	Satellites	–	–	–

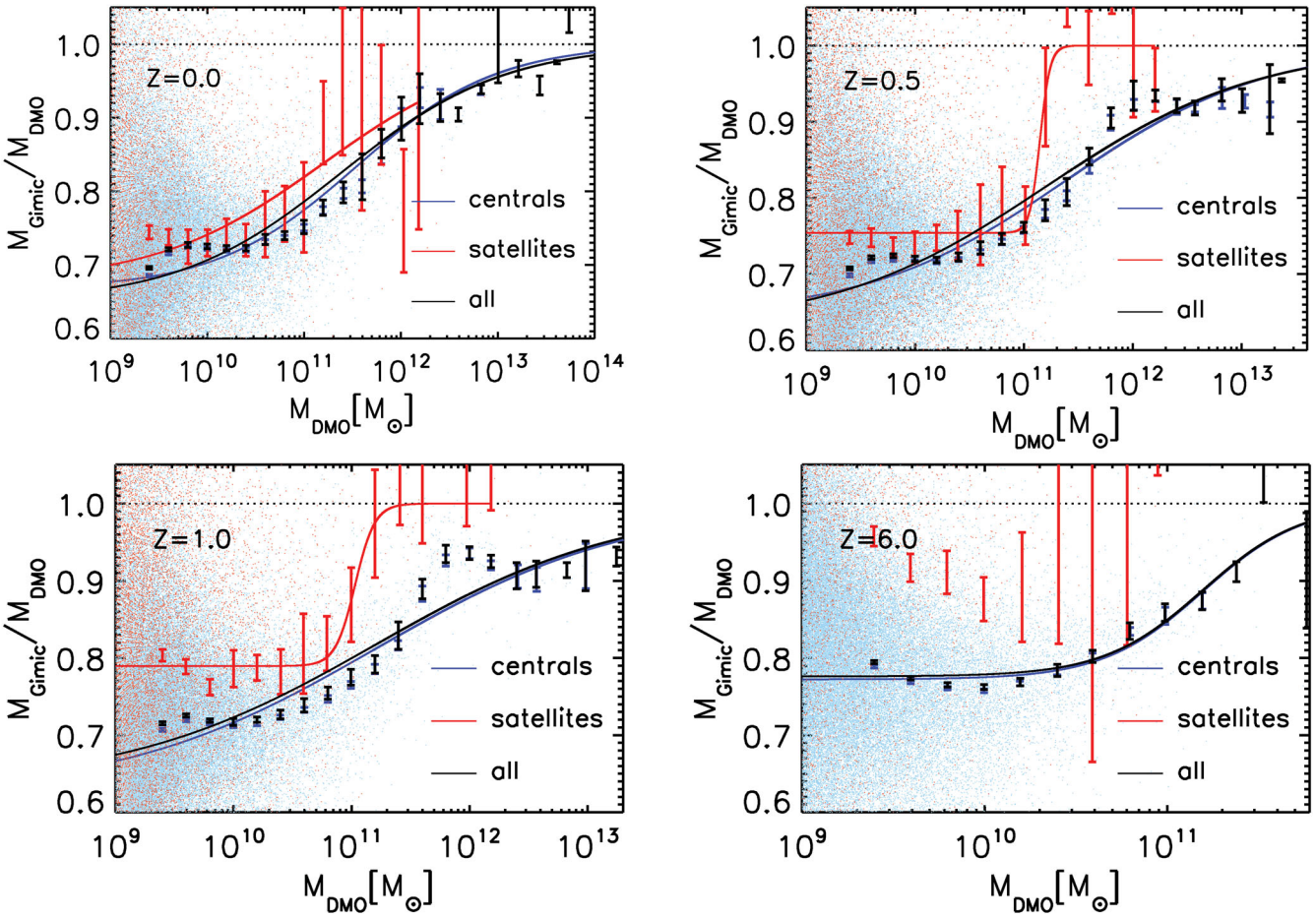

Figure A1. Ratio of subhalo masses between the GIMIC and DMO simulations for matched pairs of subhaloes at different redshift. The error bars indicate the estimate of the median and its error for centrals (blue), satellites (red) and all subhaloes (black), while the lines show the corresponding best fits to equation (2), with coefficients listed in Table A1.

centrals, although at high redshifts, the number of massive satellites is low, and the associated statistical uncertainty is large. It is worth noting that the fraction of satellites increases with time, and that more than half of the satellites at $z = 0$ were still (isolated) centrals up to $z = 0.5$, partly explaining the convergence of the two curves over time. At all times, the average reduction in mass for the total population of subhaloes closely resembles that for the centrals.

Table A1 contains the coefficients of the different fits, and may be useful for constructing subhalo mass functions based on a DMO catalogue.

APPENDIX B: CONVERGENCE

As described in Section 2, the GIMIC and DMO simulations are each performed at two different resolutions, with particle masses that differ by a factor of 8. In Fig. B1, we compare the cumulative mass functions in both sets of simulations, with thin and thick lines denoting the low- and high-resolution results, respectively. The left-hand panel shows the halo mass function, while the right-hand panel shows the subhalo mass function, both at $z = 0$. In both panels, black lines denote the DMO simulations and red lines the total mass functions of the GIMIC simulations. While the total number of both haloes and subhaloes decreases by an expected factor of $\sim$5 for the lower resolution runs, it can be seen that the mass functions are well converged, up to the resolution limit. Hence, the baryonic effects on the total number of haloes and subhaloes as a function of mass, as measured in the GIMIC simulation, are largely independent of resolution, at least for masses of $\sim 10^9\,\mathrm{M}_\odot$ and above.

In the right-hand panel of Fig. B1, we also plot the *reduced* cumulative mass functions of subhaloes that contain stars (green) or gas (blue). While the reduced mass functions for subhaloes with gas are well converged, the reduced mass functions of subhaloes with stars begin to diverge above the absolute resolution limit, indicating some resolution dependence of the star formation threshold. In addition to model dependence, this lack of convergence implies some uncertainty in the effect on stellar-to-total mass ratios discussed in Section 4, but does not change our finding that a significant fraction of *dark* subhaloes are present in a universe with baryons, which has a significant effect on abundance matching at the low-mass end.

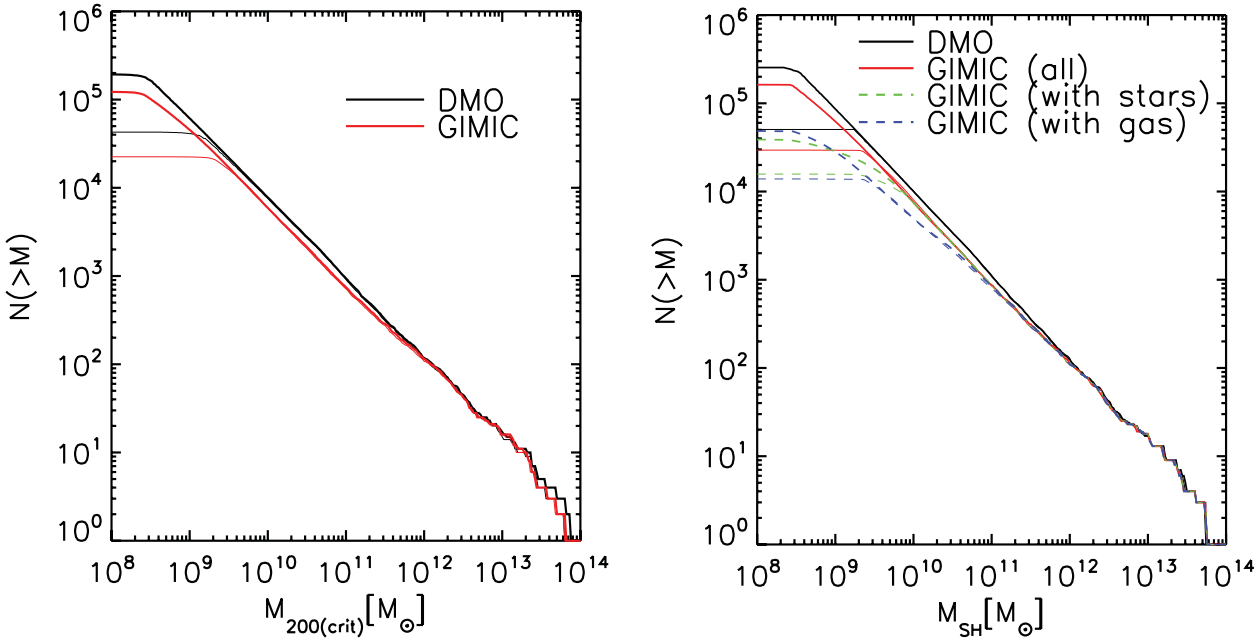

Figure B1. Cumulative mass function of haloes (left-hand panel) and subhaloes (right-hand panel) at $z = 0$ in the high-resolution (thick lines) and intermediate-resolution (thin lines) realizations. In both panels, results from the DMO simulations are shown in black, while those of the GIMIC simulations are shown in red. Also plotted in the right-hand panel are the RMFs for subhaloes with stars (green), and with gas (blue). It can be seen that the total number of haloes and subhaloes are converged to the resolution limit, both in the DMO and the GIMIC simulations. While the reduced mass functions for subhaloes with gas are also converged, the reduced mass functions of subhaloes with stars begin to diverge at higher masses, indicating some resolution dependence.

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