Constraining the warm dark matter particle mass with Milky Way satellites

Rachel Kennedy,[1]★ Carlos Frenk,[1] Shaun Cole[1] and Andrew Benson[2]

[1]*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK*
[2]*Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA*

ABSTRACT

Particle physics theories predict the existence of particles (such as keV mass sterile neutrinos) which could behave as warm dark matter (WDM), producing a cutoff in the linear density power spectrum on the scale of dwarf galaxies. Thus, the abundance of Milky Way satellite galaxies depends on the mass of the warm particle and also scales with the mass of the host galactic halo. We use the GALFORM semi-analytic model of galaxy formation to compare predicted satellite luminosity functions to Milky Way data and determine a lower bound on the thermally produced WDM particle mass. This depends strongly on the Milky Way halo mass and, to some extent, on the baryonic physics assumed. For our fiducial model, we find that for a particle mass of 3.3 keV (the 2σ lower limit from an analysis of the Lyman α forest by Viel et al.) the Milky Way halo mass is required to be $>1.4 \times 10^{12}$ M$_\odot$. For this same fiducial model, we also find that all WDM particle masses are ruled out (at 95 per cent confidence) if the Milky Way halo mass is smaller than 1.1×10^{12} M$_\odot$, while if the mass of the Galactic halo is greater than 1.8×10^{12} M$_\odot$, only WDM particle masses larger than 2 keV are allowed.

Key words: galaxies: dwarf – galaxies: formation – dark matter.

1 INTRODUCTION

The nature of the dark matter that makes up most of the matter content of the Universe is still unknown. There are several particle candidates which could potentially serve as the dark matter. The prototype is generically known as a 'weakly interacting massive particle', or WIMP, which could be the lightest supersymmetric particle, and behaves as cold dark matter (CDM; see Frenk & White 2012 for a review). These particles have the property that they acquire negligible thermal velocities at early times, giving rise to a power spectrum of inflationary density perturbations at recombination that has power on all scales; this results in the well-known hierarchical build-up of cosmic structure.

But there are many other candidates which are also well motivated from particle physics. A class of them behave as warm dark matter (WDM). These particles acquire significant thermal velocities at early times and free-stream out of small wavelength perturbations creating a cutoff in the linear power spectrum at a wavelength that varies roughly inversely with the particle mass. In this case, structure formation on scales much larger than the cutoff wavelength proceeds in a very similar way to the CDM case, but the evolution on smaller scales is very different. Good examples of WDM candidates are the sterile neutrino (e.g. Dodelson & Widrow 1994; Shi &

Fuller 1999; Asaka, Blanchet & Shaposhnikov 2005; see Kusenko 2009 for a review), or the gravitino (the supersymmetric partner of the graviton; e.g. Pagels & Primack 1982; Moroi, Murayama & Yamaguchi 1993; Gorbunov, Khmelnitsky & Rubakov 2008). These particles could have a mass in the keV range, giving rise to a cutoff in the power spectrum on the mass scale corresponding to a dwarf galaxy. A mixture of CDM and WDM is also possible, for example if there is a population of resonantly produced sterile neutrinos (Boyarsky, Ruchayskiy & Shaposhnikov 2009).

Extensive efforts are underway to detect CDM particles either directly in the laboratory, indirectly through annihilation products of Majorana particles or at the Large Hadron Collider (see Strigari 2013 for a review). None of these searches have produced conclusive evidence. While we await developments on the experimental front, important conclusions regarding the identity of the dark matter may be obtained by confronting predictions for the growth of cosmic structure with astronomical data. The key scales to distinguish CDM from WDM candidates are subgalactic scales, where the effects of the cutoff in the WDM power spectrum are imprinted. Furthermore, since the cutoff wavelength depends on the particle mass, this approach leads to constraints on the WDM particle mass, m_{WDM}. At high redshift, the relevant scales are only mildly non-linear and so calculating the evolution of dark matter, and even gas, is relatively straightforward. Using high-resolution hydrodynamical simulations to interpret the small-scale clumpiness of the Lyman α flux power spectrum measured from high-resolution spectra of

★ E-mail: rachel.kennedy@durham.ac.uk

25 $z > 4$ quasars, Viel et al. (2013) have set a lower limit of $m_{WDM} \geq 3.3$ keV (2σ) for (thermally produced) WDM particles.

At the present day, the relevant scales are strongly non-linear and so N-body cosmological simulations (or analytical methods calibrated on them) are required to predict the evolution of the dark matter. The main differences between CDM and WDM are in the mass functions and internal structure of haloes and subhaloes of subgalactic mass. For CDM these mass functions increase steeply with decreasing mass (e.g. Jenkins et al. 2001; Diemand, Kuhlen & Madau 2007; Springel et al. 2008; Tinker et al. 2008). For WDM, the abundance of subgalactic mass haloes and subhaloes is much lower, and has a cutoff at small masses which scales inversely with m_{WDM} (Colín, Avila-Reese & Valenzuela 2000; Bode, Ostriker & Turok 2001; Schneider et al. 2012; Lovell et al. 2014). In CDM, haloes and subhaloes have cuspy 'NFW' dark matter density profiles (Navarro, Frenk & White 1996, 1997; Springel et al. 2008). In WDM, cores form but these are much too small to be astrophysically relevant (Macciò et al. 2012; Shao et al. 2013). In fact, over the relevant radial range, the profiles are also cuspy but have lower concentration than CDM haloes or subhaloes of the same mass. The central concentration, which reflects the formation time of the halo, decreases with decreasing m_{WDM} (Avila-Reese et al. 2001; Lovell et al. 2012, 2014; Schneider et al. 2012).

The differences between CDM and WDM haloes and, in the latter case the dependence of halo properties on m_{WDM}, suggest a number of astrophysical tests on subgalactic scales that might distinguish between the two types of dark matter or set constraints on m_{WDM}. One, based on the different degrees of central dark matter concentration between CDM and WDM subhaloes, takes advantage of recent kinematical data for Milky Way satellites which provide information about the distribution of dark matter within them. This test is related to the so-called too big to fail problem in CDM: an apparent discrepancy between the central dark matter concentration inferred for the brightest dwarf spheroidal satellites of the Milky Way and the most massive subhaloes found in CDM N-body simulations (Boylan-Kolchin, Bullock & Kaplinghat 2011, 2012) and in some gasdynamic simulations that follow the baryonic component of the galaxy including its satellites (Parry et al. 2012). Lovell et al. (2012) showed that the 'too big to fail' problem does not exist in simulations of WDM haloes with $m_{WDM} = 1.1$ keV,[1] and Lovell et al. (2014) showed that the Milky Way satellite data are not sufficiently precise to set an interesting upper limit on m_{WDM} using this test. Even in the case of CDM, the 'too big to fail' problem disappears if the mass of the Milky Way halo is less than about 1.5×10^{12} M$_\odot$ (Purcell & Zentner 2012; Wang et al. 2012).

The second test is based on the different number of subhaloes predicted to survive in CDM and WDM galactic haloes. In the case of CDM, there are many more subhaloes within galactic haloes than there are observed satellites in the Milky Way, a discrepancy often – and incorrectly – dubbed 'the satellite problem in CDM'. In fact, it has been known for many years that inevitable feedback processes, particularly the early reionization of gas by the first stars and winds generated by supernovae, prevent visible galaxies from forming in the vast majority of the small subhaloes that survive inside CDM haloes (Bullock, Kravtsov & Weinberg 2000; Benson et al. 2002; Somerville 2002).

A 'satellite problem', however, could exist in WDM because if m_{WDM} is too small, then there will be too few surviving substructures

to account for the observed number of satellites. A limited version of this test was recently applied to surviving dark matter subhaloes in high-resolution N-body simulations of WDM galactic haloes by Polisensky & Ricotti (2011), who found a limit of $m_{WDM} > 2.3$ keV, and by Lovell et al. (2014) who found a conservative lower limit of $m_{WDM} > 1.1$ keV. In this paper, we develop this theme further; however, we apply the test *not* to dark matter subhaloes but to visible satellites. This requires following the process of galaxy formation in galactic WDM haloes, which allows a more direct comparison with observations of the Milky Way satellites and leads to stronger limits on m_{WDM}. Since the number of surviving subhaloes scales with the parent halo mass (Gao et al. 2004), these limits will depend on the mass of the Milky Way halo. Unfortunately, this mass is still very uncertain, with estimates ranging from about 8×10^{11} to 2.5×10^{12} M$_\odot$ (e.g. Li & White 2008; Xue et al. 2008; Guo et al. 2010; Deason et al. 2012; Rashkov et al. 2013; Piffl et al. 2014).

In this study, we use the Durham semi-analytic model of galaxy formation, GALFORM, to follow galaxy formation in WDM models with different values of m_{WDM}. Nierenberg et al. (2013) used a different semi-analytic model to study the redshift evolution of satellite luminosity functions for hosts of different masses, finding that compared to CDM, a $m_{WDM} = 0.75$ keV particle captured better the observed evolution. Macciò & Fontanot (2010) also used a semi-analytic model, applied to N-body simulations of galactic haloes of mass 1.22×10^{12} M$_\odot$ to set a lower limit of $m_{WDM} > 1$ keV. This limit, however, is only valid for haloes of this particular mass. Here, we use a version of GALFORM in which galaxy merger trees are computed using Monte Carlo techniques (calibrated on WDM N-body simulations). In this way, we are able to explore models with a wide range of halo masses and thus set limits on m_{WDM} for different values of the, as yet poorly known, Milky Way halo mass. Another important advantage of our method is that it does not suffer from the problem of spurious halo fragmentation which is present in, and complicates the interpretation of, high-resolution N-body simulations of WDM models (Bode et al. 2001; Wang & White 2007; Lovell et al. 2014; but see also Angulo, Hahn & Abel 2013).

Not surprisingly, only a very minor adjustment to the galaxy formation model in CDM is required in WDM to obtain a good match to a variety of observed properties of the local galaxy population, such as galaxy luminosity functions in various passbands. We then apply this model to derive the expected luminosity function of satellites of galaxies like the Milky Way and thus set strong constraints on the value of m_{WDM} as a function of the Milky Way halo mass.

The rest of this paper is organized as follows: in Section 2, we introduce our methodology, including the computation of the fluctuation power spectrum, the construction of merger trees, and the adaptation of our semi-analytic model, GALFORM, to WDM. In Section 3, we predict satellite luminosity functions in galactic haloes of different mass as a function of m_{WDM}. In Section 4, we discuss the range of particle masses that are ruled out based upon various estimates of the Milky Way halo mass. A brief discussion of this limit in the context of other independent WDM constraints is presented, along with our conclusions, in Section 5.

2 METHODS

2.1 The WDM linear power spectrum

In the case where the WDM consists of thermal relics, the suppression of small-scale power in the linear power spectrum, P_{WDM},

---

[1] Some values of m_{WDM} quoted here differ slightly from those quoted in the original paper, to make them consistent with Viel et al. (2005).
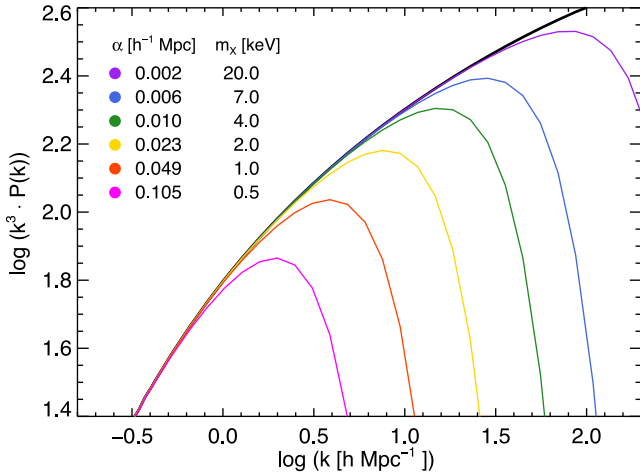
Figure 1. Linear power spectra (in arbitrary units) for WDM and CDM models. The thick black line shows CDM and the coloured lines various WDM models, labelled by their thermal relic mass and corresponding value of the damping scale, α, in the legend.

can be conveniently parametrized by reference to the CDM power spectrum, P_{CDM}. The WDM transfer function is then given by

$$T(k) = \left[\frac{P_{\text{WDM}}}{P_{\text{CDM}}}\right]^{1/2} = [1 + (\alpha k)^{2\nu}]^{-5/\nu} \tag{1}$$

(Bode et al. 2001). Here, k is the wavenumber, and following Viel et al. (2005), we take the constant $\nu = 1.12$; the parameter α can be related to the mass of the particle, m_{WDM} by

$$\alpha = 0.049 \left(\frac{\Omega_{\text{WDM}}}{0.25}\right)^{0.11} \left(\frac{h}{0.7}\right)^{1.22} \left(\frac{\text{keV}}{m_{\text{WDM}}}\right)^{1.11} h^{-1}\text{Mpc} \tag{2}$$

(Viel et al. 2005), in terms of the matter density parameter, Ω_{WDM}, and Hubble parameter, $h = H_0/(100\,\text{km s}^{-1}\,\text{Mpc}^{-1})$.

In the case where the WDM particle is a non-resonantly produced sterile neutrino, its mass m_{sterile}, can be related to the mass of the equivalent thermal relic, m_{WDM}, by requiring that the shape of the transfer function, $T(k)$, be similar in the two cases. Viel et al. (2005) give

$$m_{\text{sterile}} = 4.43 \left(\frac{m_{\text{WDM}}}{\text{keV}}\right)^{4/3} \left(\frac{0.25(0.7)^2}{\Omega_{\text{WDM}}h^2}\right)^{1/3} \text{keV}. \tag{3}$$

This conversion depends on the specific particle production mechanism (for a review, see Kusenko 2009); in the rest of this paper, we will refer only to the thermal relic mass, m_{WDM}, unless stated otherwise. We consider particles with masses, m_{WDM}, ranging from 0.5 to 20 keV. Fig. 1 shows the linear power spectra for 6 of the 11 WDM models we have investigated, as well as for CDM.

We adopt values for the cosmological parameters that are consistent with the *Wilkinson Microwave Anisotropy Probe* 7 (*WMAP*7) results (Komatsu et al. 2011): $\Omega_m = 0.272$, $\Omega_b = 0.0455$, $\Omega_\Lambda = 0.728$, $h = 0.704$, $\sigma_8 = 0.81$, $n = 0.96$. 200 merger trees were generated for each main halo mass and for each WDM particle mass.

2.2 Galaxy formation models

We calculate the properties of the galaxy population in our WDM models using the Durham semi-analytic galaxy formation model, GALFORM (e.g. Cole et al. 2000; Benson et al. 2003; Bower et al. 2006). Rather than applying it to merger trees obtained from an *N*-body simulation, we instead construct Monte Carlo merger trees

using the Extended Press–Schechter (EPS) formalism (Press & Schechter 1974; Bond et al. 1991; Bower 1991; Lacey & Cole 1993; Parkinson, Cole & Helly 2008) to generate conditional mass functions for haloes of a given mass. The standard formulation of the EPS formalism (in which the density field is filtered with a top hat in real space) is not applicable in the presence of a cutoff in the power spectrum. Instead, using a sharp filter in *k*-space produces a halo mass function in good agreement with the results of *N*-body simulations. We adopt this prescription which is justified and described in detail in Benson et al. (2013). A similar procedure was adopted by Schneider et al. (2013) but other authors, such as Smith & Markovic (2011) and Menci, Fiore & Lamastra (2012), have used a top hat filter in real space and then multiplied the resulting mass function by an ad hoc suppression factor. We do not apply the correction for finite phase-space density derived by Benson et al. (2013) because the effect of thermal velocities is negligible in the models we consider (Macciò et al. 2012; Shao et al. 2013). Halo concentrations were set according to the NFW prescription (Navarro et al. 1996, 1997), as described in Cole et al. (2000), thus explicitly taking into account the later formation epoch of WDM haloes compared to CDM haloes of the same mass. These concentrations are broadly in agreement with the WDM simulations of Schneider et al. (2012).

We use the latest version of GALFORM (Lacey et al., in preparation) which includes several improvements to the model described by Bower et al. (2006). The standard GALFORM model is tuned to fit a set of observed properties of the local galaxy population assuming CDM. Thus, an adjustment is required in the WDM case. On scales larger than dwarf galaxies at $z = 0$, there is little difference between WDM and CDM models. On smaller scales, the most important processes that influence galaxy formation are the feedback effects produced by the early reionization of the intergalactic medium and supernova feedback.

In GALFORM, reionization is modelled by assuming that no gas is able to cool in galaxies of circular velocity less than v_{cut} at redshifts less than z_{cut}. For CDM, the values $v_{\text{cut}} = 30\,\text{km s}^{-1}$ and $z_{\text{cut}} = 10$ result in a good approximation to more advanced treatments of reionization (Okamoto, Gao & Theuns 2008; Font et al. 2011). Supernova feedback, on the other hand, is controlled by the parameter β, the ratio of the rate at which gas is ejected from the galaxy to the star formation rate. This ratio is assumed to depend on the circular velocity of the disc, v_{circ}, as

$$\beta = \left(\frac{v_{\text{circ}}}{v_{\text{hot}}}\right)^{-\alpha_{\text{hot}}}, \tag{4}$$

where v_{hot} and α_{hot} are adjustable parameters fixed primarily by the requirement that the model should match the local b_J- and *K*-band galaxy luminosity functions. In the Lacey et al. model, these parameters take on the values $v_{\text{hot}} = 300\,\text{km s}^{-1}$ and $\alpha_{\text{hot}} = 3.2$. Since v_{circ} depends on the concentration of the host halo, which is lower for a WDM halo than for a CDM halo of the same mass (Lovell et al. 2012), we expect that a small adjustment to the parameters in equation (4) will be required to preserve the good match to the local luminosity functions.

Fig. 2 shows the b_J-band field galaxy luminosity function for different values of α_{hot} for the case of a 2 keV particle. Here, v_{cut} and z_{cut} are set to the CDM values. (The reionization model mostly affects galaxies fainter than those included in estimates of the field luminosity function.) The figure shows that only a small change in the value of α_{hot} is required to achieve as good a fit to the measured b_J-band luminosity function as in the CDM case. The best fit for $m_{\text{WDM}} = 3\,\text{keV}$ is obtained for $\alpha_{\text{hot}} \sim 3.0$ (green line; assuming the same value of $v_{\text{hot}} = 300\,\text{km s}^{-1}$ as in CDM). In general, we find
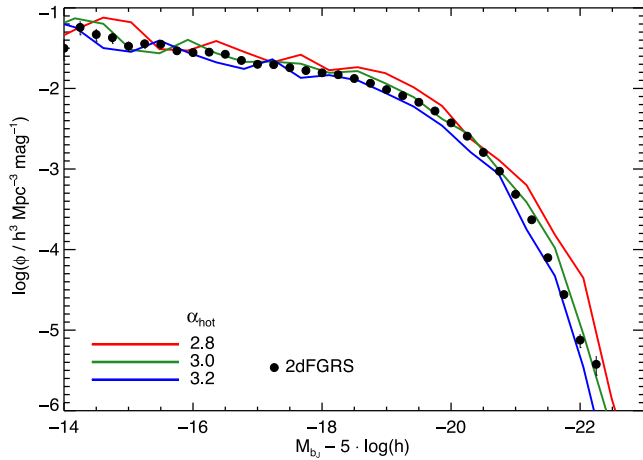
Figure 2. The b_J-band local galaxy luminosity function for $m_{WDM} = 3\,keV$ compared to the 2dFGRS determination (indicated by circles). Coloured curves show the effect of varying α_{hot}, as shown in the legend.

that the local galaxy luminosity function in WDM models is well reproduced for a wide range of values of m_{WDM} by setting

$$\alpha_{hot}(m_{WDM}) = 3.2 - 0.3 \left(\frac{m_{WDM}}{keV} \right)^{-1} \qquad (5)$$

(keeping the same values of v_{hot} and of v_{cut} and z_{cut} as above). This adjustment also results in acceptable matches to the K-band luminosity function, Tully–Fisher relation, size distribution and other observables. However, we find that for $m_{WDM} < 1.5\,keV$, we cannot obtain acceptable models using equation (5). Kang, Macciò & Dutton (2013) also found that it was not possible to find a consistent model of galaxy formation for such low-mass WDM particles. Since these masses are, in any case, ruled out by observations of the Lyman α forest, we restrict the rest of this analysis to the nine models with particle masses larger than 1.5 keV.

In Section 4.2, we vary the adjustable parameters in our models of reionization and supernova feedback to assess how they affect our inferred lower limits on the WDM particle mass. Throughout the remainder of this paper, we will refer to the model described here as the 'fiducial' model.

3 SATELLITE LUMINOSITY FUNCTIONS

We now consider satellite systems, first those predicted by GALFORM to exist in haloes of mass similar to that of the Milky Way's, and then the Milky Way's own system. We then describe the method we have adopted to compare the two.

3.1 The predicted satellite population

We use the models described in Section 2.1 with final halo masses ranging from 5×10^{10} to $1 \times 10^{13}\,M_\odot$, a significantly wider range than that covered by recent estimates of the Milky Way's halo mass. The mass resolution of the merger trees is set to $1 \times 10^6\,M_\odot$, which is below the free-streaming scale of our WDM models.

Fig. 3 shows the predicted cumulative V-band satellite luminosity functions for several examples. The three panels show results for $m_{WDM} = 2$, 3 and 20 keV and, within each panel, the effect of increasing the host halo mass from 8×10^{11} to $2.5 \times 10^{12}\,M_\odot$ is demonstrated. Increasing the host halo mass increases the number of satellites at all luminosities, and increasing the WDM particle mass increases the number of satellites particularly at fainter magnitudes.
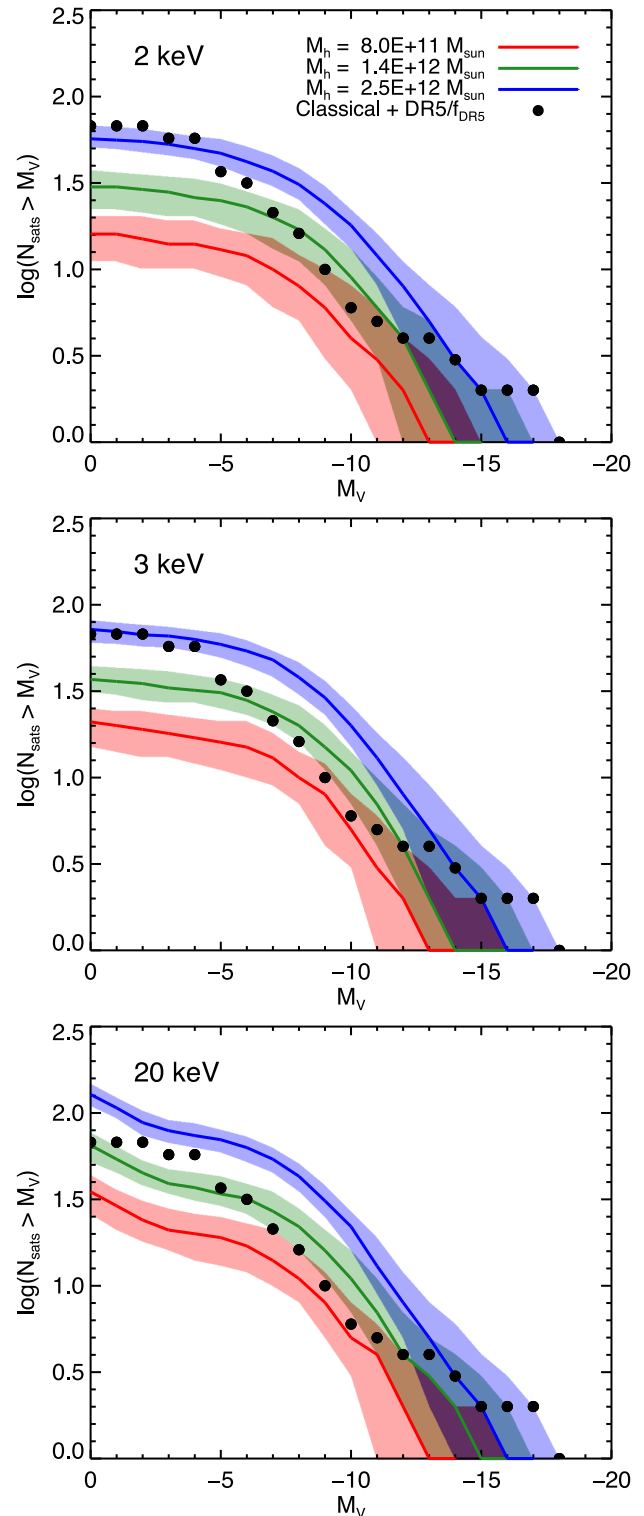

Figure 3. Satellite galaxy luminosity functions predicted by our fiducial semi-analytic model in galactic haloes of different mass, for WDM particle masses, m_{WDM}, of 2, 3 and 20 keV, as indicated in the legend. The different coloured curves correspond to different host halo mass. The solid line in each case is the median cumulative V-band satellite luminosity function and the edges of each band indicate the 10th and 90th percentiles. For reference, the luminosity function of the 11 observed classical satellites, plus the DR5 satellites (scaled for sky coverage assuming an isotropic distribution) is indicated by the black dots.

The number of bright satellites ($M_V \lesssim -12$) is insensitive to m_{WDM} because these satellites form in haloes with mass above the cutoff scale in the WDM power spectrum.

3.2 The observed satellite population

To determine whether a model produces a satisfactory number of satellites, we make use of observations of the satellites around the Milky Way. While there have been recent censuses of satellites around galaxies outside the Local Group (e.g. Guo et al. 2011; Lares, Lambas & Domínguez 2011; Liu et al. 2011; Strigari & Wechsler 2012; Wang & White 2012), these tend to be limited to the brightest few. Many faint satellites have been observed around M31 (e.g. Martin et al. 2006, 2009, 2013; Ibata et al. 2007; McConnachie et al. 2009), but in this analysis we limit ourselves to studies of the population in our own Galaxy.

There are 11 bright satellite galaxies around the Milky Way which were discovered in the previous century; these are dubbed the 'classical satellites'. In more recent years, the Sloan Digital Sky Survey (SDSS; e.g. Adelman-McCarthy et al. 2007) has revealed a number of fainter satellite galaxies. For this analysis we focus on 11 additional satellites found in the SDSS Data Release 5 (DR5; see summary in Tollerud et al. 2008), not double counting any classical satellites. This survey covers a fraction $f = 0.194$ of the sky, which is roughly 8000 square degrees, to a depth of around 22.2 in the g and r bands. We refer to these satellites here as the 'DR5 satellites'.

It is likely that there are yet more satellites in the DR5 region which have not been detected due to their faintness; at 260 kpc, the survey is only complete to $M_V \approx -6$ (Koposov et al. 2008). Attempts to correct for the detection limits of the survey by assuming a given radial profile of the satellites predict a total satellite population of hundreds (e.g. Koposov et al. 2008; Tollerud et al. 2008).

3.3 Assessing model population likelihoods

For the purposes of comparing our model predictions with satellite galaxy data, we will consider only those satellites brighter than $M_V = -2$, which is fainter than the magnitude of all the DR5 satellites. Since GALFORM only makes predictions for satellites which lie within the virial radius of the host halo, we limit our analysis of the real Milky Way satellites to those with a galactocentric distance less than the virial radius of a particular halo in the semi-analytic calculation. Here, the virial radius is defined as the boundary of the region enclosing an overdensity, Δ, with respect to the critical density, where, for the spherical collapse model, $\Delta \approx 93$ (Eke, Cole & Frenk 1996).

In order to estimate the total number of satellites brighter than $M_V = -2$ that we would expect around the Milky Way, it is necessary to make some assumptions about the underlying distribution since it is not fully sampled. First, we make the assumption that all the 'classical' satellites (those with apparent magnitudes brighter than $M_V \approx -8.5$) have been observed. This is probable, although our results would not change significantly even if one or two remained undetected behind the Milky Way disc.

Next, we assume that the underlying distribution of satellites is isotropic, so that the DR5 represents a geometrically unbiased sampling. This may be unrealistic because the eleven classical satellites of the Milky Way are known to lie in a 'pancake' structure oriented approximately perpendicular to the plane of the Milky Way disc (Lynden-Bell 1976, 1982; Majewski 1994; Libeskind et al. 2005). A large region of the DR5 footprint intersects this plane; if as yet undetected satellites also tend to lie in this disc, then the DR5 would

provide a biased sampling of the true satellite population, leading us to overpredict the number of satellites that are necessary to match the data. This would have the effect of weakening our lower limit on m_{WDM}. However, cosmological N-body simulations show that the preferentially flattened satellite distributions are restricted to the brightest satellites, and that as fainter and fainter populations are considered, their distribution tends to become increasingly isotropic (Wang, Frenk & Cooper 2013).

Finally, we make the extremely conservative assumption that every satellite in the DR5 footprint area has been detected, so that no more faint satellites are lurking below the detection threshold. Given the survey's radial completeness limits, this is unrealistic. This assumption works in the sense of making our inferred lower limits on m_{WDM} conservative. If future or current surveys, such as Pan-STARRS, were to reveal even more faint satellites, our lower mass limits would become correspondingly stronger.

To quantify whether the model satellite population is compatible with the MW data, we require that the model should produce at least as many satellites with $M_V < -2$ as are known to exist in the Milky Way. To find the likelihood of each model given the data, we calculate the probability that the predicted satellite population includes at least as many members falling within a region the size of the DR5 footprint, i.e. covering a fraction of the sky, $f = 0.194$, as the DR5 survey itself, which contains n_{DR5} satellites[2].

First, we define the number of classical Milky Way satellites (again within the virial radius of the model halo) to be n_{class}. This number is subtracted from the total number of predicted satellites, $n_{galform}$, to prevent double-counting in the DR5 region,

$$n_{pred} = n_{galform} - n_{class}. \tag{6}$$

Then, for this remaining population of n_{pred} satellites, we must find the likelihood that they are distributed such that at least as many satellites as are observed in DR5 fall in a region covering a fraction f of the total sky area. We find the probability, P, that a number between n_{DR5} and n_{pred} satellites lie in this region by assuming that a given satellite is equally likely to be found anywhere on the sky. Hence, P can be calculated from a binomial distribution,

$$P = \sum_{k=n_{DR5}}^{k=n_{pred}} \left(\frac{n_{pred}!}{k!(n_{pred} - k)!} \right) \cdot f^k \cdot (1 - f)^{n_{pred} - k}. \tag{7}$$

Equation (7) gives the probability that any given realization of a halo merger tree, for a particular value of m_{WDM}, within a given host halo mass, M_h, has produced enough satellites to be compatible with the Milky Way data. Since we have generated 200 merger trees for each WDM model at a given host halo mass, we take the average of the probabilities, P, computed for each individual host halo using equation (7).

If $\langle P \rangle$ is smaller than 0.05, we conclude that this model predicts too little substructure to account for the observations. Conversely, for each WDM particle mass, m_{WDM}, we find the minimum host halo mass, M_h, for which $\langle P \rangle$ is larger than 5 per cent. This value of m_{WDM} is therefore the limiting mass that cannot be excluded at 95 per cent confidence.

---

[2] The value of n_{DR5} used will depend upon the virial radius of the halo we compare to.
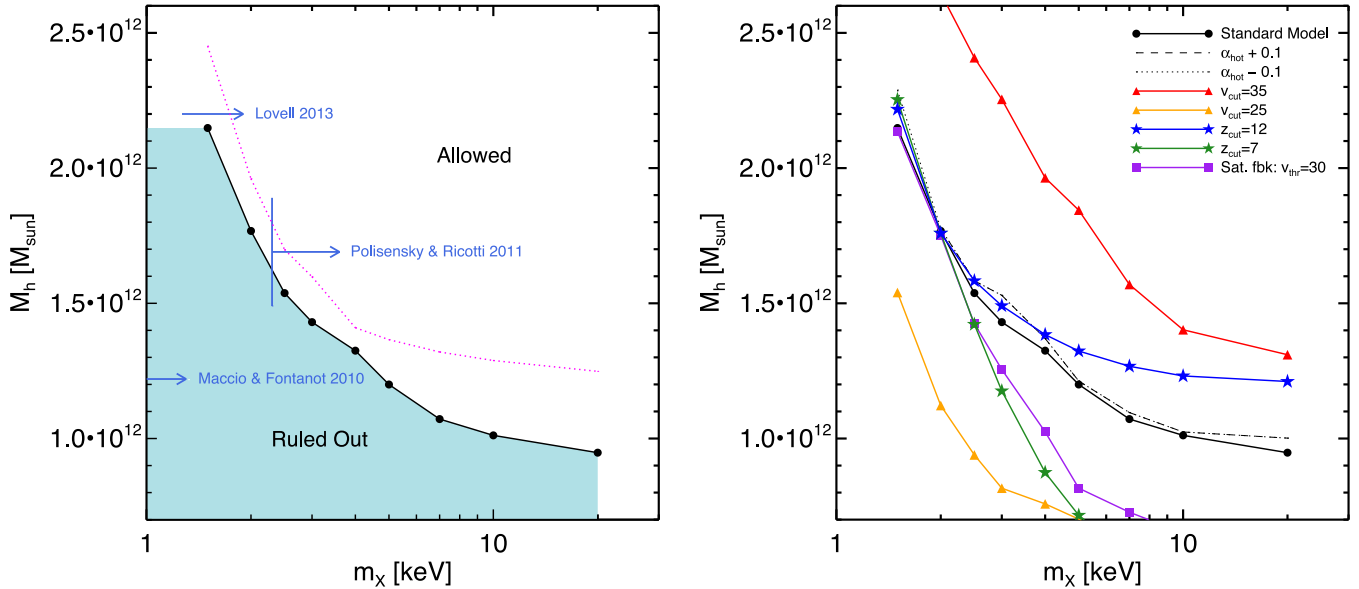
Figure 4. Left: exclusion diagram for thermal WDM particle masses, m_{WDM}, as a function of the Milky Way dark matter halo mass, M_h; the shaded region is excluded. The lower limits reported by other authors, as well as the host halo masses they considered, are indicated by the arrows. The dotted magenta line shows the limit if satellites not visible to an SDSS-type survey are excluded. Right: sensitivity of our constraints to variations in the parameters of our galaxy formation model; the lines show the envelope of the exclusion region.

4 RESULTS: LIMITS ON THE WDM PARTICLE MASS

In this section, we present the constraints[3] on the warm dark particle mass that follow from comparing our predictions for the satellite luminosity functions with the Milky Way data. We also discuss how our limits can be affected by uncertainties in our modelling of galaxy formation.

4.1 Fiducial model

The constraints on the WDM particle mass as a function of host halo mass set by the method described in Section 3.3 are shown in the exclusion diagram of Fig. 4(a). Each point in the plot gives the smallest Galactic halo mass that has at least a 5 per cent chance of hosting enough satellites to account for the observed number. Conversely, for a given Galactic halo mass, the minimum allowed WDM particle mass can be read off the *x*-axis. The shaded region shows the parameter space that is excluded. For example, if the Milky Way were found to have a mass of 1.5×10^{12} $M_\odot$, then the thermal relic dark matter particle must be more massive than 3 keV. The envelope of the exclusion region asymptotes to a value of 1.1×10^{12} $M_\odot$. Thus, for Milky Way halo masses below this value, all WDM particle masses are ruled out at 95 per cent confidence by our model.

As an additional test we have applied the SDSS visibility limits to our satellite populations, in order to discern which could actually be detected by the survey. Using equation (2) given in Tollerud et al. (2008), we find the threshold radius beyond which each of our satellites of a certain *V*-band magnitude would not be detected. Since the Monte Carlo based approach used here does not yield spatial information about the satellites, we use the radial distribution from Anderhalden et al. (2013, Fig. 4; very similar for CDM and WDM) in order to determine the probability that each satellite is inside this

completeness radius. Then, generating a random value between 0 and 1, we reject the satellite from our culled sample if this value is larger than the calculated probability. This yields the population of satellites which would be observable by SDSS. The result of this exercise is shown in Fig. 4(a) by the dotted magenta line. The limits become more stringent since this selection eliminates most of the faintest satellites from the sample.

An accurate measurement of the Milky Way's halo mass, M_h, could, in principle, rule out all astrophysically interesting thermally produced WDM particles. Unfortunately, this measurement is difficult and subject to systematic uncertainties. Several methods have been used to estimate M_h. (The values quoted below refer to different definitions of virial mass assuming different values of the limiting density contrast, Δ, as indicated by the subscript, M_Δ.) A traditional one is the timing argument of Kahn & Woltjer (1959) which employs the dynamics of the Local Group to estimate its mass. Calibrating this method with CDM *N*-body simulations, Li & White (2008) find $M_{200} \sim 2.43 \times 10^{12}$ $M_\odot$, with a lower limit of $M_{200} = 8.0 \times 10^{11}$ $M_\odot$ at 95 per cent confidence. A rather different method is based on matching the abundance of galaxies ranked by stellar mass to the abundance of dark matter haloes ranked by mass in a large CDM *N*-body simulation. This technique gives upper and lower 10 per cent confidence limits of $8 \times 10^{11} < M_{200} < 4.7 \times 10^{12}$ $M_\odot$ (Guo et al. 2010).

A third class of methods relies on the kinematics of tracer stars in the stellar halo to constrain the potential out to large distances. Using positions and line-of-sight velocities for 240 halo stars, Battaglia et al. (2005) find $6 \times 10^{11} < M_{100} < 3 \times 10^{12}$ $M_\odot$, depending on assumptions about the halo profile; using 2000 BHB stars out to 60 kpc, interpreted with the aid of simulations, Xue et al. (2008) find $8 \times 10^{11} < M_{102} < 1.3 \times 10^{12}$ $M_\odot$. Using a variety of tracers, Deason et al. (2012) find the mass within 150 kpc to be between 5×10^{11} and 1×10^{12} $M_\odot$. Most recently, Piffl et al. (2014) used a large sample of stars from the RAVE survey in conjunction with cosmological simulations to find $1.3 \times 10^{12} < M_{200} < 1.8 \times 10^{12}$ $M_\odot$.

[3] These data can be accessed by contacting the lead author.

4.2 Sensitivity to galaxy formation model parameters

Given an assumption about the nature of the dark matter, the abundance of galactic satellites depends primarily on two key astrophysical processes: the reionization of hydrogen after recombination and feedback from supernova explosions. The epoch during which the Universe became reionized is constrained by temperature anisotropies in the microwave background and their polarization to lie in the range $8 \lesssim z_{re} \lesssim 14$ (Planck Collaboration et al. 2014). Photoheating raises the entropy of the gas and suppresses cooling into haloes of low virial temperature.

In GALFORM, reionization is modelled by assuming that no gas cools in haloes of circular velocity smaller than v_{cut} at redshifts lower than z_{cut}. This simple prescription has been shown to be a good approximation to a more detailed semi-analytic model of reionization (Benson et al. 2002) and to full gasdynamic simulations (Okamoto et al. 2008). In our fiducial model, the parameters take the values $v_{cut} = 30\,\mathrm{km\,s^{-1}}$ and $z_{cut} = 10$. The simulations of Okamoto et al. (2008) suggest that v_{cut} is around $25\,\mathrm{km\,s^{-1}}$, but Font et al. (2011) conclude that a value of $v_{cut} = 34\,\mathrm{km\,s^{-1}}$ is required to match the results of the detailed semi-analytical calculation of the effects of reionization given by Benson et al. (2002). We explore the effect of varying both v_{cut} and z_{cut} within these bounds.

Supernova feedback is still poorly understood. In GALFORM, this process is modelled in terms of a simple parametrized power law of the disc circular velocity with exponent α_{hot} (equation 4). As discussed in Section 2.2, the parameter α_{hot} is constrained – as a function of m_{WDM} – by the strict requirement that the model should provide an acceptable fit to the observed local b_J-band galaxy luminosity function. This is a strong constraint which limits any possible variation of α_{hot} to less than ± 0.1. Our simple parametrization ignores, for example, environmental effects (Lagos, Lacey & Baugh 2013) but these are unlikely to make a significant difference to our conclusions so we do not consider them further. However, we do consider a model in which the effects of feedback saturate below $v_{circ} = 30\,\mathrm{km\,s^{-1}}$, similar to what Font et al. (2011) argue is required to explain the variation of metallicity with luminosity observed in the population of Milky Way satellites.

The effects of varying the galaxy formation model parameters (retaining agreement with the local field galaxy luminosity function) on our constraints on m_{WDM} as a function of M_h are shown in Fig. 4(b). Varying α_{hot} has a very small effect; varying z_{cut} affects, to some extent, the limits for WDM particle masses greater than 2–3 keV. The main sensitivity is to the parameter v_{cut} which has a strong effect on the number of small haloes which are able to form stars. At fixed halo mass, lower values of v_{cut} weaken the limits on m_{WDM} whereas larger values strengthen them. The range considered here, $25 < v_{cut}/\mathrm{km\,s^{-1}} < 35$, is realistic according to current understanding of the process of reionization.

5 DISCUSSION AND CONCLUSIONS

The cutoff in the linear power spectrum of density fluctuations produced by the free streaming of WDM particles in the early universe provides, in principle, the means to search for evidence of these particles. If the particle mass is in the keV range, the cutoff occurs on the scale of dwarf galaxies and no primordial fluctuations are present on smaller scales. Thus, establishing how smooth the universe is on these scales could reveal the existence of WDM or, since the cutoff length-scales inversely with the particle mass, set limits on its mass. The traditional method for testing the smoothness of the density field at early times is to measure the

flux power spectrum of the Lyman α forest in the spectra of high-redshift quasars. The most recent lower limit on the WDM particle mass using this method on data at redshifts $z \sim 2$–6 is that set by Viel et al. (2013), $m_{WDM} \geq 3.3\,\mathrm{keV}$ (2σ), for thermally produced WDM particles.

A different way to estimate the clumpiness of the matter density field on small scales, this time at the present day, is to count the number of substructures embedded in galactic haloes. The most direct way to do this is to count the satellites that survive in such haloes but these are so faint that sufficient numbers can only be found in our own Milky Way Galaxy and M31. Counting the Milky Way satellites thus provides a test of WDM which is independent from and complementary to the Lyman α forest constraint. There are several complications that need to be taken into account when carrying out this test. First, a suitable property to characterize the satellite population needs to be identified. The maximum of the circular velocity curve, v_{max}, is often used for this purpose, but this quantity is not directly measurable for the Milky Way's satellites. The luminosities of satellites, on the other hand, are accurately measured, but using this as a test of WDM requires the ability to predict the satellite luminosities and this, in turn, requires modelling galaxy formation. This is the approach we have adopted in this paper where we have made use of the semi-analytic model, GALFORM. This model has the virtue that it gives a good match to the field galaxy luminosity function in various bands and has been extensively tested against a variety of other observational data. The v_{max} test was carried out by Polisensky & Ricotti (2011) and by Lovell et al. (2014) but the uncertainty in the satellites' values of v_{max} introduces some uncertainty in the limits set.

The second complication is the requirement to understand the completeness of the satellite sample. The Milky Way has a population of 11 bright or 'classical' satellites which is thought to be complete (although one or two bright satellites could be lurking behind the Galactic plane, too small a number to affect our conclusions) and a population of faint and ultrafaint satellites that have been discovered in the fifth of the sky surveyed by the SDSS. While the classical satellites are known to be distributed on the thin plane, identified by Lynden-Bell (1976), it is not known if the SDSS sample is also anisotropic. Large N-body CDM simulations suggest that it is only the brightest satellites that lie on a plane whereas more abundant populations tend to be much less anisotropically distributed (Wang et al. 2013). Here, we assume that the spatial distribution of the Milky Way satellites other than the classical ones is isotropic. If this assumption were incorrect, we would overestimate the number of satellites which would cause us to overestimate the minimum WDM particle mass required to have enough satellites in a halo of a given mass. The simulations of Wang et al. (2013) suggest that this effect is unlikely to be large.

The third complication of our method is the difficulty in assessing possible systematic effects arising from uncertainties in our galaxy formation model. As we discussed in Section 4.2, the main source of uncertainty is our treatment of the inhibiting effect of the early reionization of the intergalactic medium on the cooling of gas in small haloes. We model this process in a relatively simple way which, however, has been validated both by realistic semi-analytic calculations (Benson et al. 2002) and by full cosmological hydrodynamic simulations (Okamoto et al. 2008). Another uncertainty arises from the fate of satellites prior to merging with the central galaxy: we do not currently consider tidal disruption effects in our model, meaning that all satellites survive until the point of merging. If tidal destruction is an important phenomenon, which may be especially true for WDM, then we would expect fewer surviving

satellites in our models. This would have the net effect of increasing further our lower limits on m_{WDM}.

Since the number of surviving subhaloes is a strong function of the parent halo mass, our limits on m_{WDM} depend on the mass of the Milky Way halo which, unfortunately, is still uncertain to within a factor of at least a few. For our fiducial model of galaxy formation, we find that if the halo mass is less than $1.1 \times 10^{12}\,M_\odot$, then *all* values of m_{WDM} are ruled out at 95 per cent confidence for the case of thermally produced WDM particles. If, however, the mass of the halo is greater than $1.3 \times 10^{12}\,M_\odot$, then, at the same confidence level, all masses greater than $m_{WDM} = 5\,keV$ are allowed, and if it is greater than $2 \times 10^{12}\,M_\odot$, then all masses greater than $m_{WDM} = 2\,keV$ are allowed. If the main parameter in our model of reionization, v_{cut}, had a value of $35\,km\,s^{-1}$, then most (thermal) masses of astrophysical interest would be ruled out even if the mass of the halo is $2 \times 10^{12}\,M_\odot$, but if this parameter is only $25\,km\,s^{-1}$, then only masses below $m_{WDM} = 2.5\,keV$ are ruled out for halo masses less than $1 \times 10^{12}\,M_\odot$. By contrast, using the abundance of dark matter subhaloes as a function of v_{max}, Lovell et al. (2014) were only be able to set a lower limit of $m_{WDM} = 1.3\,keV$ for dark matter haloes of mass $1.8 \times 10^{12}\,M_\odot$.

Our limits on the WDM particle mass from the abundance of satellites in the Milky Way are compatible with those set by the Lyman α forest constraints, except, of course, that they depend on the mass of the Milky Way halo. The value of the most recent lower limit ($m_{WDM} = 3.3\,keV$) derived from the Lyman α forest requires the halo mass to be $M_h > 1.4 \times 10^{12}\,M_\odot$ in order for there to be enough satellites in the Milky Way. All these limits apply only to thermally produced WDM and need not exclude specific warm candidates such as sterile neutrinos. In this case (and also for other types of WDM), there could also be additional resonantly produced particles that could behave as CDM, resulting in a different small scale behaviour of the linear density power spectrum, depending on the mass and formation epoch of these particles.

Sterile neutrinos can decay and emit a narrow X-ray line. The absence of such a line in the X-ray spectra of galaxy clusters can be used to set an *upper* limit to m_{WDM} but this depends in the sterile neutrino production mechanism. For example, for non-resonant production, Abazajian, Fuller & Tucker (2001) have set an *upper* limit of $m_{sterile} \lesssim 5\,keV$ which would correspond to a thermal mass of $\sim 1\,keV$.

The constraints presented in this study would become much tighter if the mass of the Milky Way halo could be measured accurately. While the recent RAVE results (Piffl et al. 2014) are encouraging, it is to be hoped that the forthcoming *GAIA* satellite mission will allow a better understanding of the systematic effects that complicate these kinds of measurements. In the meantime, gravitational lensing effects such as the flux ratio anomaly in multiply lensed quasar images may provide a direct measurement of the amount of substructure present in galactic dark matter haloes (Miranda & Macciò 2007; Xu et al. 2013). This is a powerful method that could, in principle, provide a conclusive test of whether the dark matter is cold or warm.

REFERENCES

Abazajian K., Fuller G. M., Tucker W. H., 2001, ApJ, 562, 593
Adelman-McCarthy J. K. et al., 2007, ApJS, 172, 634
Anderhalden D., Schneider A., Macciò A. V., Diemand J., Bertone G., 2013, J. Cosmol. Astropart. Phys., 3, 14
Angulo R. E., Hahn O., Abel T., 2013, MNRAS, 434, 3337
Asaka T., Blanchet S., Shaposhnikov M., 2005, Phys. Lett. B, 631, 151
Avila-Reese V., Colín P., Valenzuela O., D'Onghia E., Firmani C., 2001, ApJ, 559, 516
Battaglia G. et al., 2005, MNRAS, 364, 433
Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002, MNRAS, 333, 156
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38
Benson A. J. et al., 2013, MNRAS, 428, 1774
Bode P., Ostriker J. P., Turok N., 2001, ApJ, 556, 93
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bower R. G., 1991, MNRAS, 248, 332
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
Boyarsky A., Ruchayskiy O., Shaposhnikov M., 2009, Ann. Rev. Nucl. Part. Sci., 59, 191
Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, L40
Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2012, MNRAS, 422, 1203
Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Colín P., Avila-Reese V., Valenzuela O., 2000, ApJ, 542, 622
Deason A. J. et al., 2012, MNRAS, 425, 2840
Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859
Dodelson S., Widrow L. M., 1994, Phys. Rev. Lett., 72, 17
Eke V. R., Cole S., Frenk C. S., 1996, MNRAS, 282, 263
Font A. S. et al., 2011, MNRAS, 417, 1260
Frenk C. S., White S. D. M., 2012, Ann. Phys., 524, 507
Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004, MNRAS, 355, 819
Gorbunov D., Khmelnitsky A., Rubakov V., 2008, J. High Energy Phys., 12, 55
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Guo Q., Cole S., Eke V., Frenk C., 2011, MNRAS, 417, 370
Ibata R., Martin N. F., Irwin M., Chapman S., Ferguson A. M. N., Lewis G. F., McConnachie A. W., 2007, ApJ, 671, 1591
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Kahn F. D., Woltjer L., 1959, ApJ, 130, 705
Kang X., Macciò A. V., Dutton A. A., 2013, ApJ, 767, 22
Komatsu E. et al., 2011, ApJS, 192, 18
Koposov S. et al., 2008, ApJ, 686, 279
Kusenko A., 2009, Phys. Rep., 481, 1
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lagos C. d. P., Lacey C. G., Baugh C. M., 2013, MNRAS, 436, 1787
Lares M., Lambas D. G., Domínguez M. J., 2011, AJ, 142, 13
Libeskind N. I., Frenk C. S., Cole S., Helly J. C., Jenkins A., Navarro J. F., Power C., 2005, MNRAS, 363, 146
Li Y.-S., White S. D. M., 2008, MNRAS, 384, 1459
Liu L., Gerke B. F., Wechsler R. H., Behroozi P. S., Busha M. T., 2011, ApJ, 733, 62
Lovell M. R. et al., 2012, MNRAS, 420, 2318
Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, MNRAS, 439, 300
Lynden-Bell D., 1976, MNRAS, 174, 695
Lynden-Bell D., 1982, The Observatory, 102, 202

Macciò A. V., Fontanot F., 2010, MNRAS, 404, L16

Macciò A. V., Paduroiu S., Anderhalden D., Schneider A., Moore B., 2012, MNRAS, 424, 1105

McConnachie A. W. et al., 2009, Nature, 461, 66

Majewski S. R., 1994, ApJ, 431, L17

Martin N. F., Ibata R. A., Irwin M. J., Chapman S., Lewis G. F., Ferguson A. M. N., Tanvir N., McConnachie A. W., 2006, MNRAS, 371, 1983

Martin N. F. et al., 2009, ApJ, 705, 758

Martin N. F. et al., 2013, ApJ, 772, 15

Menci N., Fiore F., Lamastra A., 2012, MNRAS, 421, 2384

Miranda M., Macciò A. V., 2007, MNRAS, 382, 1225

Moroi T., Murayama H., Yamaguchi M., 1993, Phys. Lett. B, 303, 289

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Nierenberg A. M., Treu T., Menci N., Lu Y., Wang W., 2013, ApJ, 772, 146

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

Pagels H., Primack J. R., 1982, Phys. Rev. Lett., 48, 223

Parkinson H., Cole S., Helly J., 2008, MNRAS, 383, 557

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

Piffl T. et al., 2014, A&A, 562, A91

Planck Collaboration et al., 2014, A&A, preprint (arXiv:1303.5076)

Polisensky E., Ricotti M., 2011, Phys. Rev. D, 83, 043506

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

Purcell C. W., Zentner A. R., 2012, J. Cosmol. Astropart. Phys., 12, 7

Rashkov V., Pillepich A., Deason A. J., Madau P., Rockosi C. M., Guedes J., Mayer L., 2013, ApJ, 773, L32

Schneider A., Smith R. E., Macciò A. V., Moore B., 2012, MNRAS, 424, 684

Schneider A., Smith R. E., Reed D., 2013, MNRAS, 433, 1573

Shao S., Gao L., Theuns T., Frenk C. S., 2013, MNRAS, 430, 2346

Shi X., Fuller G. M., 1999, Phys. Rev. Lett., 82, 2832

Smith R. E., Markovic K., 2011, Phys. Rev. D, 84, 063507

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

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

Strigari L. E., 2013, Phys. Rep., 531, 1

Strigari L. E., Wechsler R. H., 2012, ApJ, 749, 75

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

Tollerud E. J., Bullock J. S., Strigari L. E., Willman B., 2008, ApJ, 688, 277

Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, Phys. Rev. D, 71, 063534

Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, Phys. Rev. D, 88, 043502

Wang J., White S. D. M., 2007, MNRAS, 380, 93

Wang W., White S. D. M., 2012, MNRAS, 424, 2574

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

Wang J., Frenk C. S., Cooper A. P., 2013, MNRAS, 429, 1502

Xu D. D., Sluse D., Gao L., Wang J., Frenk C., Mao S., Schneider P., 2013, preprint (arXiv:1307.4220)

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

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