

# Dark matter halo merger histories beyond cold dark matter – I. Methods and application to warm dark matter

Andrew J. Benson,<sup>1</sup>★ Arya Farahi,<sup>2</sup> Shaun Cole,<sup>3</sup> Leonidas A. Moustakas,<sup>4</sup> Adrian Jenkins,<sup>3</sup> Mark Lovell,<sup>3</sup> Rachel Kennedy,<sup>3</sup> John Helly<sup>3</sup> and Carlos Frenk<sup>3</sup>

<sup>1</sup>*Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA*

<sup>2</sup>*Michigan Center for Theoretical Physics, Randall Laboratory of Physics, The University of Michigan, Ann Arbor, MI 48109-1120, USA*

<sup>3</sup>*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE*

<sup>4</sup>*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, CA 91109, USA*

Accepted 2012 October 4. Received 2012 October 2; in original form 2012 September 13

## ABSTRACT

We describe a methodology to accurately compute halo mass functions, progenitor mass functions, merger rates and merger trees in non-cold dark matter universes using a self-consistent treatment of the generalized extended Press–Schechter formalism. Our approach permits rapid exploration of the subhalo population of galactic haloes in dark matter models with a variety of different particle properties or universes with rolling, truncated or more complicated power spectra. We make detailed comparisons of analytically derived mass functions and merger histories with recent warm dark matter cosmological  $N$ -body simulations, and find excellent agreement. We show that once the accretion of smoothly distributed matter is accounted for, coarse-grained statistics such as the mass accretion history of haloes can be almost indistinguishable between cold and warm dark matter cases. However, the halo mass function and progenitor mass functions differ significantly, with the warm dark matter cases being strongly suppressed below the free-streaming scale of the dark matter. We demonstrate the importance of using the correct solution for the excursion set barrier first-crossing distribution in warm dark matter – if the solution for a flat barrier is used instead, the truncation of the halo mass function is much slower, leading to an overestimate of the number of low-mass haloes.

**Key words:** galaxies: formation – galaxies: haloes – cosmology: theory – dark matter.

## 1 INTRODUCTION

The cold dark matter (CDM) paradigm (Peebles 1982) works extremely well on large scales (Seljak et al. 2005; Percival et al. 2007; Ferramacho, Blanchard & Zolnierowski 2009; Sánchez et al. 2009; Komatsu et al. 2011), but there remains the possibility of deviations from the CDM expectations on small scales – arising notably from the issue of cores versus cusps and the inner slope of dark matter density profiles (Salucci 2001; Donato, Gentile & Salucci 2004; Donato et al. 2009; Newman et al. 2009, 2011; de Blok 2010; Kuzio de Naray & Kaufmann 2011; Kuzio de Naray & Spekkens 2011; Salucci et al. 2012; Wolf & Bullock 2012) and the apparent paucity of bright satellites around the Milky Way (Boylan-Kolchin, Bullock & Kaplinghat 2011, 2012; although see Wang et al. 2012). Several extensions to dark matter phenomenology (Markevitch et al. 2004; Ahn & Shapiro 2005; Boehm & Schaeffer 2005; Miranda & Macciò 2007; Randall et al. 2008; Boyarsky et al. 2009; Lovell et al. 2012)

and several different particle physics candidates for dark matter (Raffelt 1990; Turner 1990; Jungman, Kamionkowski & Griest 1996; Hogan & Dalcanton 2000; Spergel & Steinhardt 2000; Abazajian, Fuller & Patel 2001; Cheng, Feng & Matchev 2002; Feng, Rajaraman & Takayama 2003; Sigurdson & Kamionkowski 2004; Hubisz & Meade 2005; Feng et al. 2009) have been put forward to explain these deviations, although it remains unclear if any of these proposals is able to fully explain observed phenomena (Kuzio de Naray et al. 2010) or if they are even necessary, with the observed phenomena simply being a consequence of galaxy formation physics in a CDM universe (Benson et al. 2002b; Libeskind et al. 2007; Li et al. 2009; Stringer, Cole & Frenk 2010; Font et al. 2011; Oh et al. 2011; Governato et al. 2012; Kuhlen et al. 2012; Pfrommer, Chang & Broderick 2012; Pontzen & Governato 2012; Starkenburg et al. 2012; Vera-Ciro et al. 2012). A variety of experimental measurements may be sensitive to the small-scale structure of dark matter haloes (Simon et al. 2005; Viel et al. 2008). The most promising are future lensing experiments, which have the potential to strongly constrain dark matter particle phenomenology (Keeton & Moustakas 2009; Vegetti & Koopmans

★E-mail: abenson@obs.carnegiescience.edu

2009a,b; Vegetti, Czoske & Koopmans 2010a; Vegetti et al. 2010b, 2012). A variety of previous works have shown that the subhalo mass function should depend sensitively on the dark matter physics and the shape of the power spectrum. To maximize the scientific return of future experiments therefore requires the ability to accurately and rapidly predict the distribution of dark matter substructure as a function of dark matter particle thermal or interaction properties, for arbitrary power spectra.

In this work, we develop techniques to follow the growth of non-linear structure in non-CDM universes using a fully consistent treatment of the extended Press–Schechter formalism. We demonstrate the performance of these techniques by applying them to a representative case of warm dark matter (WDM), which has the advantage of several pre-existing  $N$ -body simulations which we utilize to test the accuracy of our methods. WDM particles<sup>1</sup> are lighter than their CDM counterparts, allowing them to remain relativistic for longer in the early universe and to retain a non-negligible thermal velocity dispersion. This velocity dispersion allows them to free-stream out of density perturbations and so suppresses the growth of structure on small scales (Bond & Szalay 1983; Bardeen et al. 1986). While mass functions have been previously considered in this case (Barkana, Haiman & Ostriker 2001), we go one step further and develop methods to compute conditional mass functions and halo merger rates, and use these to construct merger trees in WDM universes. These merger trees are a key ingredient required to predict the distribution of substructure masses, positions and internal structure as is necessary to make detailed predictions for future lensing experiments.

The remainder of this paper is organized as follows. In Section 2, we describe the changes that we introduce to the extended Press–Schechter formalism to make it applicable to the case of non-CDM scenarios (including some specifics for the WDM case). In Section 3, we apply these methods to the case of WDM, first comparing their predictions to the available data from  $N$ -body simulations, then exploring the limitations of ignoring the effects of WDM velocity dispersion and presenting a comparison of key results between WDM and CDM. Finally, in Section 4, we discuss the consequences of this work and present our conclusions.

We also include two appendices. Appendix A gives a detailed derivation of our numerical procedure for solving the excursion set first-crossing problem for arbitrary barriers. Appendix B explores the numerical accuracy and robustness of the methods developed in this work.

When comparing our analytic theory with results from  $N$ -body simulations we will adopt the same cosmological parameters and dark matter particle properties as were used for the simulation. These values will be listed where relevant. For the rest of this work, specifically in Sections 2, 3.2–3.3 and Appendix B we adopt a canonical cosmological model with  $\Omega_M = 0.2725$ ,  $\Omega_\Lambda = 0.7275$ ,  $\Omega_b = 0.0455$  and  $H_0 = 70.2 \text{ km s}^{-1} \text{ Mpc}^{-1}$  (Komatsu et al. 2011) and a canonical WDM particle of mass,  $m_X = 1.5 \text{ keV}$  and effective number of degrees of freedom  $g_X = 1.5$  (the expected value for a fermionic spin- $\frac{1}{2}$  particle).

## 2 METHODS

Our approach makes use of the Press–Schechter formalism (Press & Schechter 1974; Bond et al. 1991; Bower 1991; Lacey & Cole 1993) which, after substantial development and tuning against  $N$ -body simulations, has proven to be extremely valuable in understanding the statistical properties of dark matter halo growth in CDM universes. The Press–Schechter formalism in its modern form is expressed in terms of excursion sets – the set of all possible random walks in density at a point as the density field is smoothed on ever smaller scales. Halo formation corresponds to a random walk making its first crossing of a barrier. The height of that barrier is determined from models of the non-linear collapse of simple overdensities.

The Press–Schechter algorithm requires three ingredients: (1) the power spectrum of fluctuations in the density field (characterized by  $\sigma(M)$ , the fractional root variance in the linear theory density field at  $z = 0$ ); (2) the critical threshold in linear theory corresponding to the gravitational collapse of a density perturbation  $\delta_c$ ; and (3) a solution for the statistics of excursion sets to cross this threshold. We will address each of these three ingredients below.

### 2.1 Power spectrum

We assume a power-law primordial power spectrum with  $n_s = 0.961$  (Komatsu et al. 2011) and adopt the transfer function of Eisenstein & Hu (1999). We include a modification for WDM using the fitting function of Bode, Ostriker & Turok (2001; as re-expressed by Barkana et al. 2001) to impose a cutoff below a specified length scale,  $\lambda_s$ :

$$T(k) \rightarrow T(k) [1 + (\epsilon k \lambda_s)^{2\nu}]^{-\eta/\nu}, \quad (1)$$

where  $\epsilon = 0.361$ ,  $\eta = 5$  and  $\nu = 1.2$  are parameters of the fitting function. For our canonical WDM particle, the smoothing scale<sup>2</sup> is  $\lambda_s = 0.124 \text{ Mpc}$  (Barkana et al. 2001; equation 4) corresponding to a mass of  $M_s = 4\pi\bar{\rho}\lambda_s^3/3 = 2.97 \times 10^8 M_\odot$ . The power spectrum is normalized to give the required  $\sigma_8 = 0.807$  (Komatsu et al. 2011) when integrated under a real-space top-hat filter of radius  $8 h^{-1} \text{ Mpc}$  (where  $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ).

#### 2.1.1 Window function

To derive the variance,  $S(M)$ , or, equivalently, the root variance,  $\sigma(M) \equiv \sqrt{S(M)}$ , from the power spectrum a window function,  $W(k|M)$ , must be adopted. Specifically,

$$S(M) \equiv \sigma^2(M) = \frac{1}{2\pi^2} \int_0^\infty 4\pi k^2 P(k) W^2(k|M) dk. \quad (2)$$

In the Bond et al. (1991) re-derivation of the Press & Schechter (1974) mass function and extended Press–Schechter conditional mass functions, one assumes a sharp  $k$ -space filtering,

$$W(k|M) = \begin{cases} 1 & \text{if } k \leq k_s(M) \\ 0 & \text{if } k > k_s(M), \end{cases} \quad (3)$$

of the linear density field so that the trajectories of density fluctuation,  $\delta$ , versus mass scale are true Brownian random walks. Despite

<sup>1</sup> The two usual candidates – both lying beyond the standard model of particle physics – are sterile neutrinos (Dodelson & Widrow 1994; Shaposhnikov & Tkachev 2006) and gravitinos (Ellis, Kim & Nanopoulos 1984; Moroi, Murayama & Yamaguchi 1993; Kawasaki, Sugiyama & Yanagida 1997; Gorbunov, Khmelnitsky & Rubakov 2008).

<sup>2</sup> This scale is usually approximated as being equal to the speed of the particles at the epoch of matter-radiation equality multiplied by the comoving horizon scale at that time; see Bode et al. (2001) for further discussion.

the use of this sharp  $k$ -space filtering in deriving the extended Press–Schechter solutions, it has been conventional to adopt the  $\sigma(M)$  that is given by a real-space top-hat window function,

$$W(k|M) = \frac{3[\sin(kR) - kR \cos(kR)]}{(kR)^3}, \quad (4)$$

when utilizing those solutions. The reason for this is that only in this case there is no ambiguity between the filtering scale,  $R$ , and the corresponding mass (see Lacey & Cole 1993),

$$R = \left( \frac{3M}{4\pi\bar{\rho}} \right)^{1/3}. \quad (5)$$

For CDM power spectra, which are essentially pseudo-power laws with a slowly varying slope, the shape of the  $\sigma(M)$  function varies little with the choice of filter function.

However, in the case of power spectra with a small-scale cutoff, such as those of WDM models and many other non-CDM candidates, the situation is very different. With a top-hat window function,  $\sigma(M)$  continues to increase with decreasing filter mass even for masses well below the cutoff scale. This happens because although there are no new intrinsic small-scale modes entering the filter, the longer wavelength modes are getting reweighted as the mass scale of the filter increases. In contrast, with a sharp  $k$  filter,  $\sigma(M)$  increases monotonically up to the cutoff scale and then becomes constant with the transition being determined by the abruptness of the cutoff in the input WDM power spectrum. Since the Press–Schechter mass functions depend directly on  $\sigma(M)$ , including a direct dependence on the logarithmic slope of  $\sigma(M)$ , its predictions for the case of WDM are sensitive to this choice. Hence, for non-CDM models, it is important to revisit the issue of filter choice and one would expect that the conventional top-hat choice will lead to an overestimate of the number of low-mass haloes [an effect apparent in the works of Barkana et al. (2001) and Menci, Fiore & Lamastra (2012)].

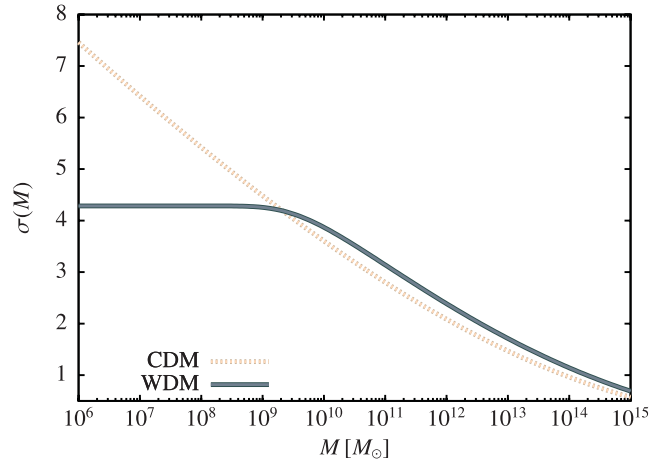
We begin by adopting the sharp  $k$ -space window function of equation (3) when working with truncated power spectra, such that the contribution from long-wavelength modes remains constant. In the case of a top-hat real-space window function, there is a natural relation between the top-hat radius and the smoothing mass scale, as given by equation (5). In the case of a sharp  $k$ -space window function, the choice of  $k_s(M)$  is less clear. Lacey & Cole (1993) suggest choosing  $\tilde{W}(r=0) = 1$  (where  $\tilde{W}(r)$  is the Fourier transform of the window function) and then choosing  $k_s(M)$  such that the integral of the mean density,  $\bar{\rho}$ , under  $\tilde{W}(r)$  equals the required smoothing mass,  $M$ . However, as noted by Lacey & Cole (1993), this choice for  $k_s(M)$  lacks any strong physical motivation.

Therefore, we advocate a different approach, namely we choose

$$k_s = a/R, \quad (6)$$

where  $R$  is computed from equation (5) and the parameter  $a$  is chosen such that the turnover in the halo mass function occurs in the same location as that seen in  $N$ -body simulations of WDM, as will be shown in Section 3.1. We find that  $a = 2.5$  is required to meet this condition.<sup>3</sup> This will mean that  $\sigma(M)$  differs from the usual result on large scales, resulting in a difference in the mass function for high-mass haloes that would not be expected. To remedy this, we note that the critical overdensity for collapse is usually motivated on the basis of spherical top-hat collapse models. Since we are no

<sup>3</sup> The normalization advocated by Lacey & Cole (1993) is equivalent to  $a = (9\pi/2)^{1/3} \approx 2.42$  which is not too different from our best-fitting value of 2.5.



**Figure 1.** The fractional root variance of the mass density field as a function of the mean mass contained within that filter. Both CDM and our canonical WDM cases are shown. A top-hat real-space window function is used for the CDM, while a sharp  $k$ -space filter is used for WDM. The WDM case is suppressed below the CDM value for  $M \lesssim M_s$ .

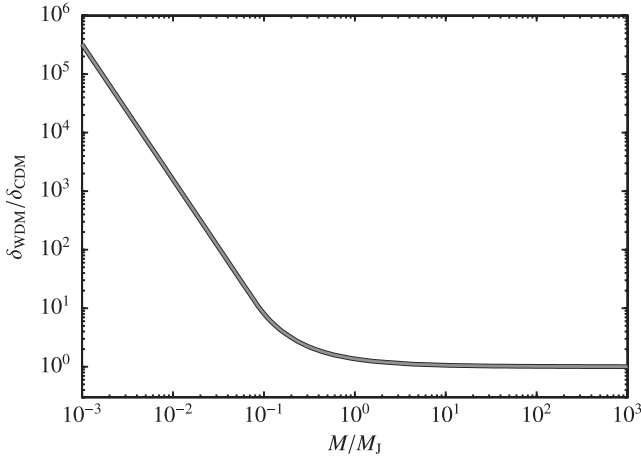
longer using a top-hat filter we choose to allow freedom in the choice of the critical overdensity such that the halo mass function is unchanged for high masses. This point will be discussed in greater detail in Section 2.2.2.

We note that fixing the  $k_s(M)$  relation in this way introduces a free parameter,  $a$ , into our method and introduces a dependence on the  $N$ -body simulation to which we will compare our results in Section 3.1.1. However, given that the CDM linear power spectrum is close to a power law on the scales that will be of interest for WDM models and that the modification to the transfer function due to WDM scales simply with the mass of the dark matter particle, we expect that the same choice of  $a$  should be valid for all WDM particle masses of interest.

Fig. 1 shows the resulting root variance in the density field (smoothed using a top-hat filter in real space) as a function of mass enclosed within the filter for both WDM and CDM models. As expected,  $\sigma(M)$  in our WDM model is suppressed below the CDM value for  $M \lesssim M_s$ .

## 2.2 Critical overdensity for collapse (excursion set barrier)

The Press–Schechter algorithm as formulated by Bond et al. (1991) associates the collapse of a dark matter halo with a random walk in  $\delta(M)$  making its first crossing of some barrier,  $B(S)$ . This barrier is usually associated with the critical linear theory overdensity for collapse of spherical top-hat perturbations (although see Section 2.2.1 for a refinement of this assumption). In the CDM case, that critical overdensity is independent of mass scale (since there are no preferred scales in the problem). In the non-CDM case, this is no longer true. For example, with WDM we expect collapse to become significantly more difficult on small scales, where the WDM particles can stream out of collapsing overdensities due to their non-zero random velocities. Barkana et al. (2001) addressed the question of collapse thresholds in WDM universes by performing a set of 1D hydrodynamical simulations, in which pressure acted as a proxy for the velocity dispersion of WDM particles. They find that the growth of collapsing overdensities is suppressed below a characteristic mass scale – i.e. the threshold for collapse increases rapidly with decreasing mass below the characteristic mass scale.



**Figure 2.** The critical linear theory overdensity for the collapse of a spherical top-hat density perturbation in a WDM universe normalized to that in a CDM universe as a function of halo mass,  $M$  (expressed in units of the effective Jeans mass of the WDM,  $M_J$ ). For  $M/M_J \gg 1$  the WDM and CDM cases coincide, but for  $M/M_J \lesssim 1$  WDM requires a much larger overdensity to undergo gravitational collapse.

We find that the results of Barkana et al. (2001) can be well fitted by the functional form:<sup>4</sup>

$$\delta_{\text{c,WDM}}(M, t) = \delta_{\text{c,CDM}}(t) \left\{ h(x) \frac{0.04}{\exp(2.3x)} + [1 - h(x)] \exp \left[ \frac{0.31687}{\exp(0.809x)} \right] \right\}, \quad (7)$$

where  $x = \log(M/M_J)$ ,  $M$  is the mass in question,  $M_J$  is the effective Jeans mass of the WDM as defined by Barkana et al. (2001; their equation 10):

$$M_J = 3.06 \times 10^8 \left( \frac{1+z_{\text{eq}}}{3000} \right)^{1.5} \left( \frac{\Omega_M h_0^2}{0.15} \right)^{1/2} \times \left( \frac{g_X}{1.5} \right)^{-1} \left( \frac{m_X}{1.0 \text{ keV}} \right)^{-4} M_\odot, \quad (8)$$

the redshift of matter-radiation equality is given by

$$z_{\text{eq}} = 3600 \left( \frac{\Omega_M h_0^2}{0.15} \right) - 1 \quad (9)$$

and

$$h(x) = 1/[1 + \exp[(x + 2.4)/0.1]]. \quad (10)$$

The ratio of the resulting critical overdensity to the CDM value with masses scaled to  $M_J$  is shown in Fig. 2. For  $M/M_J \lesssim 1$ , the critical overdensity for collapse in WDM universes is much higher than in CDM as a result of the non-zero velocity dispersion of WDM particles. A small-scale perturbation must have a much larger density for its self-gravity to overcome this velocity dispersion and cause collapse.

We emphasize that the calculations of Barkana et al. (2001) are approximate as they are based on a hydrodynamical approximation

<sup>4</sup> This fit is accurate for the regime where  $\delta_{\text{c,WDM}}/\delta_{\text{c,CDM}} < 600$ , but substantially overpredicts the results of Barkana et al. (2001) for smaller masses. This is a deliberate choice – our aim was to match the shape of the function through the region where it transitions away from the CDM value. On smaller mass scales the suppression is so dramatic that the precise value of  $\delta_{\text{c,WDM}}/\delta_{\text{c,CDM}}$  is unimportant.

which will not fully capture the collisionless dynamics of WDM. We expect that the approximation made by Barkana et al. (2001) could lead to a small (order unity) difference in the characteristic mass scale for suppression of overdensity collapse (e.g. in a similar way that the Toomre stability threshold for collisionless stars and gas differ by a factor of  $\pi/3.36$ ). There could plausibly be differences in detail in the shape of the collapse threshold as a function of mass, but we are unable to speculate what form those might take. A more realistic calculation using a Boltzmann solver should be carried out to improve upon these results and explore the dependence on cosmological parameters.

### 2.2.1 Barrier remapping

In the CDM case, the original Press–Schechter algorithm as formulated by Bond et al. (1991) adopted a barrier,  $B(S)$ , for excursion sets equal to the critical linear theory overdensity for the gravitational collapse spherical top-hat perturbations and which was independent of mass and equal to  $\delta_{\text{c}}(M[S], t) = \delta_{\text{c},0}/D(t)$ , where  $D(t)$  is the linear theory growth function<sup>5</sup> and  $\delta_{\text{c},0}$  is the collapse threshold [equal to  $(3/20)(12\pi)^{2/3} \approx 1.686$  in an Einstein–de Sitter universe; solutions for other cosmologies are given by Percival (2005), for example]. It is now well known that this constant barrier does not result in a halo mass function that agrees well with that measured from  $N$ -body simulations of CDM (e.g. Tinker et al. 2008). Motivated by this discrepancy, Sheth, Mo & Tormen (2001) introduced a remapping of the barrier<sup>6</sup> calibrated to improve the match:

$$B(S) \rightarrow B(S) \sqrt{A} \left( 1 + b \left[ \frac{S}{AB^2(S)} \right]^c \right), \quad (11)$$

where  $A = 0.707$ ,  $b = 0.5$  and  $c = 0.6$  are parameters that were tuned to obtain the best match. We retain this remapping of the barrier in this work<sup>7</sup> when computing halo mass functions but not when computing halo merger rates (see Section 2.4 for further discussion of this point).

### 2.2.2 Barrier scaling

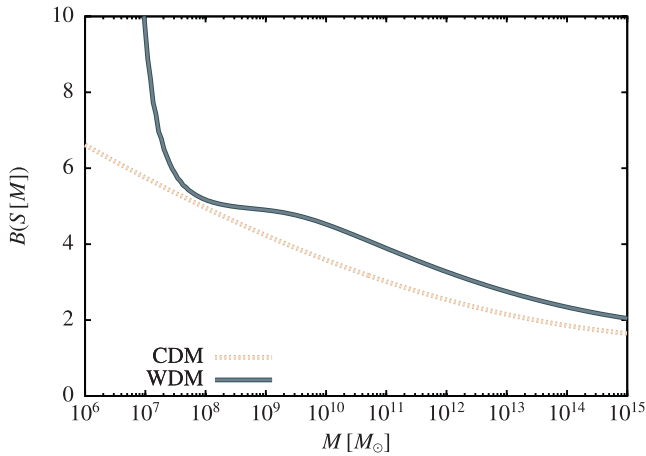
As noted in Section 2.1.1 our choice of normalization for the window function used to derived  $\sigma(M)$  from  $P(k)$  results in  $\sigma(M)$  for the

<sup>5</sup> Note that, as is conventional, we place this time dependence into the critical overdensity, such that we can always work with  $\sigma(M)$  at  $z = 0$ . The growth function used here is the usual growth function computed for CDM, consistent with the definitions of Barkana et al. (2001).

<sup>6</sup> This remapping was motivated by considerations of ellipsoidal collapse, the barrier for which differs from that for spherical collapse. In particular, Sheth et al. (2001) noted that the density perturbations leading to low-mass haloes are expected to be more ellipsoidal than those that give rise to the most massive haloes – effectively making the collapse barrier a function of mass scale.

<sup>7</sup> This mapping was derived strictly for the CDM case. We retain it here since for haloes with masses,  $M$ , much greater than  $M_s$ , we expect the WDM case to converge to the CDM solution. However, there is no guarantee that this mapping remains accurate for  $M \lesssim M_s$ . In particular, it is possible that low-mass haloes close to the cutoff scale could be more spherical in WDM than in CDM due to the isotropizing effects of the WDM velocity dispersion. Calibration of our methods against reliable  $N$ -body simulations of WDM universes would be required to evaluate the accuracy of the remapping in this regime. The  $N$ -body simulations currently available and examined later in this work do not address this particular issue as they do not include the effects of the non-zero velocity dispersion of WDM. However, it is understood how to include velocity dispersions in WDM simulations (Brandbyge et al. 2008), and this should be included in future simulations.





**Figure 3.** The barrier for excursion sets in both the CDM and our canonical WDM cases. For large masses the two are offset by the constant factor of 1.197 which accounts for the difference in definition (top-hat real-space versus sharp  $k$ -space window functions) of  $\sigma^2(M)$  in the two cases, but the WDM barrier becomes very large for  $M \ll M_s$ . Below  $M \sim M_s$  the WDM barrier rises rapidly due to the velocity dispersion of WDM particles.

WDM case lying above the CDM case for large masses. This will change the halo mass function for high-mass haloes – something which is not expected (i.e. WDM should behave just like CDM on sufficiently large scales). Therefore, we introduce a global rescaling of the barrier (applied after the remapping described in Section 2.2.1 – an important point since that remapping is non-linear), multiplying it by a factor of 1.197. This value is chosen to counteract the higher  $\sigma(M)$  on large scales<sup>8</sup> and ensure that the mass function of high-mass haloes remains unchanged (see Section 2.1.1 for discussion of this point). This rescaling of the barrier is always included, both when computing halo mass functions and when computing halo merger rates. We note that this factor is very insensitive to cosmological parameters and  $\sigma_8$  as it depends only upon the shape of the power spectrum (not the normalization). For example, if we compute the appropriate rescaling factor for *Wilkinson Microwave Anisotropy Probe 1* (WMAP1) cosmological parameters (rather than the WMAP7 cosmological parameters used throughout this work), we find that the factor changes by  $<0.1$  per cent (despite there being a difference of 11 per cent in  $\sigma_8$ ).

The resulting excursion set barriers for both CDM and our canonical WDM models (including both remapping and subsequent rescaling) are shown in Fig. 3. For large masses the two are offset by a factor of 1.197, but the WDM barrier becomes very large for  $M \ll M_s$  due to the velocity dispersion of the WDM particles.

### 2.3 Excursion sets/barrier crossing

Given a barrier,  $B(S)$ , and the variance of the density field,  $S(M)$ , the Press–Schechter algorithm proceeds by following random walks in  $\delta(S)$  beginning from  $(\delta, S) = (0, 0)$ . When a given random walk first crosses the barrier at some variance  $S$ , it is assumed that the point

<sup>8</sup> That is, the ratio of  $\sigma(M)$  computed using our sharp  $k$ -space filter to that computed using the top-hat filter is  $\sqrt{1.197}$ . Note that this scaling factor does not constitute a free parameter in our approach. Instead, it is fixed to offset the change in the  $S(M)$  relation on large scales resulting from the difference between top-hat and sharp  $k$ -space window functions. Therefore, given a power spectrum and a value of  $a$  (the coefficient in equation 6), the scaling factor is directly computable and uniquely determined.

in question has collapsed into a halo of mass  $M(S)$ . The fraction of random walks crossing between  $S$  and  $S + dS$ ,  $f(S)dS$ , corresponds to the fraction of mass in the universe in haloes of mass  $M$  to  $M + dM$ .

We therefore must compute the probability for random walks to cross our barrier between  $S$  and  $S + dS$ . We will assume a Gaussian distribution of density perturbations, motivated by the simplest inflation models. In the case of a constant (or linear in  $S$ ) barrier, there is a well-known analytic solution (Bond et al. 1991; Lacey & Cole 1993; Sheth 1998). However, for an arbitrary barrier, the solution must be computed numerically. Our approach, described in detail in Appendix A, is similar to that of Zhang & Hui (2006) but is numerically more robust. Briefly, the excursion set problem requires finding a solution to an integral equation:

$$1 = \int_0^S f(S') dS' + \int_{-\infty}^{B(S)} P(\delta, S) d\delta, \quad (12)$$

where  $P(\delta, S)d\delta$  is the probability for a trajectory to lie between  $\delta$  and  $\delta + d\delta$  at variance  $S$ . This equation expresses mass conservation, i.e. at any  $S$ , all trajectories must either have crossed the barrier at some smaller  $S$  (the first term in equation 12) or be below the barrier having never crossed at smaller  $S$  (the second term in equation 12). To numerically solve this equation we discretize the variance,  $S$ , using a grid that is uniform in  $S$  with a spacing of  $\Delta S$ . We then numerically solve equation (12) as described in detail in Appendix A. The solution is extended to the maximum value of  $S$ , if such exists (as is the case in a WDM model), or to sufficiently large  $S$  that smaller mass scales are of no interest for the problem in question. The accuracy of our numerical solver is explored in Appendix B1.

### 2.4 Merger rates and progenitor mass functions

The original Press–Schechter algorithm has been extended to compute conditional mass functions (i.e. the mass function of haloes which will all belong to a single halo of larger mass at some later time; Lacey & Cole 1993). Further, the form of the conditional mass function in the limit of small timestep has been used to estimate merger rates of dark matter haloes and so to construct merger trees (Cole et al. 2000). We therefore wish to compute the conditional mass function in the WDM case, specifically in the limit of small  $\delta t$ . In the excursion set approach, the conditional mass function is found by simply solving the first-crossing problem beginning from the  $(\delta, S)$  of the parent halo and using a barrier corresponding to an earlier time. This is equivalent to solving the original first-crossing problem with a modified barrier  $B'(S', t_1, t_0) = B(S' + S, t_1) - B(S, t_0)$ .

To compute merger rates we therefore solve the first-crossing problem using our numerical method but with an effective barrier  $B'(S', t_1, t_0)$ . We choose  $t_1 = (1 - \epsilon)t_0$  where  $\epsilon \ll 1$ . The rate of crossing this effective barrier can then be estimated as  $f(S)/\epsilon t_0$ . We will explore the sensitivity of our results to the value of the numerical parameter  $\epsilon$ .

In this case the choice of grid in  $S'$  is particularly important. Accurate numerical solution requires many grid points at small  $S'$  (since  $f(S')$  will peak very close to  $S' = 0$  for small values of  $\epsilon$ ), but also many points close to the maximum possible value of  $S'$ .  $S(M)$  becomes almost independent of  $M$  close to the maximum value of  $S$  – as a result, many points are required to resolve the cutoff of the halo mass function as a function of mass. Therefore, we adopt the following distribution for  $S_i$ :

$$S_0 = 0$$

$$S_{i>0} = \frac{1 + 1/r}{1/S_{i,\text{lin}} + 1/r S_{i,\text{log}}}, \quad (13)$$

where  $S_{i,\text{lin}}$  and  $S_{i,\text{log}}$  are sets of points spaced uniformly in  $S$  and the logarithm of  $S$  between the required minimum and maximum values, and  $r$  is a numerical parameter. For  $r > 1$  the resulting distribution of points is spaced uniformly in the logarithm of  $S$  for small  $S$  and transitions to being uniform in  $S$  for large  $S$  – as a result, the grid has good resolution both close to  $S = 0$  and close to the maximum value of  $S$ . The value of  $r$  controls the location of the transition between these two regimes. We find that a value of  $r = 10$  works well. The accuracy of our numerical solver for merger rates is explored in Appendix B2.

When solving for halo merger rates we *do not* include any remapping of the barrier function (as discussed in Section 2.2.1). The remapping function was chosen to result in good agreement between the Press–Schechter halo mass function and that measured in the Millennium  $N$ -body simulations (Springel et al. 2005). However, our merger rates will be used in merger tree construction algorithms which, in the CDM case, have been developed to work with a barrier with no remapping, but with their own empirical modification of merger rates designed to match progenitor mass functions found in  $N$ -body simulations (see Section 2.5 for further discussion). Scaling of the barrier (see Section 2.2.2) is included however.

## 2.5 Merger tree construction

To build merger trees, we follow the algorithm of Parkinson, Cole & Helly (2008). Briefly, at each point in a merger tree, this algorithm evaluates the probability per unit time of a binary merger occurring along the branch

$$\frac{df}{d\omega} = \int_{M_{\min}}^{M/2} \frac{M}{M'} \frac{df}{dM'} \frac{dS}{dM'} \left| \frac{dt}{d\omega} \right| G[\omega, \sigma(M), \sigma(M')] dM', \quad (14)$$

where  $\omega = \delta_{c,0}/D(t)$  and  $\delta_{c,0}$  is the critical overdensity for collapse in the CDM case and the integration is from the lowest mass halo to be resolved in the tree,  $M_{\min}$ , to haloes of half the mass of the current halo, corresponding to an equal mass merger. Cole et al. (2000) discuss the subtleties of why the integration is carried out over the lower half of the progenitor mass range. The rate of accretion of mass in haloes below the resolution limit of the merger tree is

$$\frac{dR}{d\omega} = \int_0^{M_{\min}} \frac{df}{dM'} \frac{dS}{dM'} \left| \frac{dt}{d\omega} \right| G[\omega, \sigma(M), \sigma(M')] dM', \quad (15)$$

where the integral is taken over all unresolved haloes (i.e. those less massive than  $M_{\min}$ ). In the above,  $G[\omega, \sigma(M), \sigma(M')]$  is an empirical modification of the progenitor mass function introduced to obtain results consistent with those from  $N$ -body simulations of CDM universes. Parkinson et al. (2008) show that the form

$$G[\omega, \sigma_1, \sigma_2] = G_0 \left( \frac{\sigma_2}{\sigma_1} \right)^{\gamma_1} \left( \frac{\omega}{\sigma_1} \right)^{\gamma_2}, \quad (16)$$

with  $G_0 = 0.57$ ,  $\gamma_1 = 0.38$  and  $\gamma_2 = -0.01$  provides a good match to progenitor mass functions measured from the Millennium Simulation (Springel et al. 2005), which assumes similar cosmological parameter values as are used in this work. The validity of this empirical modification on much smaller mass scales than are probed by the Millennium Simulation is explored in Appendix B3.

We retain the same empirical modification in this work since it should remain valid for WDM in the limit of high-mass haloes. For masses comparable to and below  $M_s$ , this empirical modification may no longer be accurate. Our justification for retaining this empirical modification is the good agreement achieved with progenitor

mass functions from  $N$ -body simulations of WDM as will be shown in Section 3.1.

A timestep is then chosen that is sufficiently small such that the probability of multiple branchings is small and the fractional change in mass due to subresolution accretion is small. This timestep is then taken, branching if a random deviate lies below the branching probability, and with mass removed at the rate of subresolution accretion. If branching does occur, the mass of one of the branched haloes is selected at random from the distribution  $d^2f/d\omega dM'$  and the other is chosen to ensure mass conservation.

### 2.5.1 Smooth accretion

In the CDM case, all mass is locked into haloes on some scale.<sup>9</sup> This implies that the progenitor mass function in CDM contains the entire mass of the parent halo, i.e.

$$\int_0^\infty f(S') dS' = 1. \quad (17)$$

For the WDM case, this is not true and we find

$$\int_0^{S_{\max}} f(S') dS' < 1, \quad (18)$$

i.e. not all trajectories cross the barrier. We identify these trajectories as corresponding to smooth accretion, i.e. accretion of mass which does not belong to any collapsed halo. This is physically distinct from the unresolved haloes accounted for by equation (15), but we can nevertheless account for this smooth accretion by boosting  $dR/d\omega$  by an amount

$$\frac{dR}{d\omega} \Big|_{\text{smooth}} = \frac{df_n}{d\omega} \left| \frac{dt}{d\omega} \right| G[\omega, \sigma(M), \sqrt{S_{\max}}], \quad (19)$$

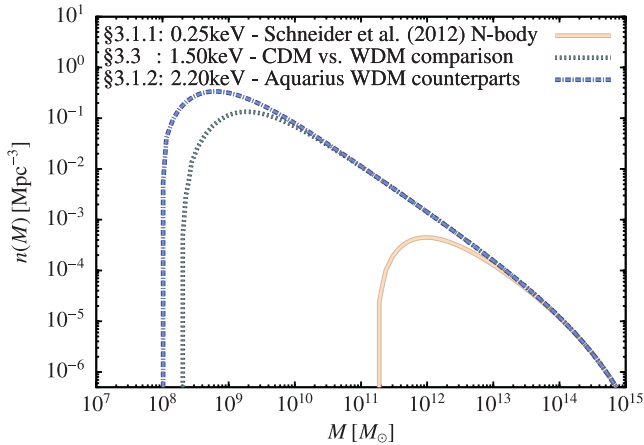
where  $f_n = 1 - \int_0^{S_{\max}} f(S') dS'$ . Note that we choose to include the empirical correction of Parkinson et al. (2008) here, using the maximum value of  $\sigma(M)$  for the  $\sigma_c$  argument. This ensures a treatment consistent with resolved haloes but, once again, should be tested and calibrated against WDM simulations. The importance of this smooth accretion is explored in Appendix B4.

## 3 RESULTS FOR WARM DARK MATTER

The methods described in Section 2 have been implemented within the open source semi-analytic galaxy formation code, GALACTICUS (Benson 2012). All results presented in this section were generated using GALACTICUS v0.9.1 r903. Control files and scripts to generate all results presented in this paper using GALACTICUS can be found at <http://users.obs.carnegiescience.edu/abenson/galacticus/parameters/dmMergingBeyondCDM1.tar.bz2>.

We consider three different WDM particle masses. In Section 3.1.1 we use  $m_X = 0.25$  keV to match the  $N$ -body simulations of Schneider et al. (2012), in Section 3.1.2 we use  $m_X = 2.2$  keV to match the Aquarius WDM counterpart simulations (described below) and in Section 3.3 we use  $m_X = 1.5$  keV in comparisons of WDM and CDM. For reference, Fig. 4 shows the halo mass function in these three cases, illustrating the expected increase in the cutoff mass,  $m_s$ , as the WDM particle mass is decreased.

<sup>9</sup> Assuming that  $\sigma(M)$  continues to rise monotonically to arbitrarily small scales. In practice, this is not true as even CDM will have some cutoff in its power spectrum on very small scales (Green, Hofmann & Schwarz 2004; Loeb & Zaldarriaga 2005). However, for our present purposes the assumption that all mass is locked into haloes in CDM is a good one.



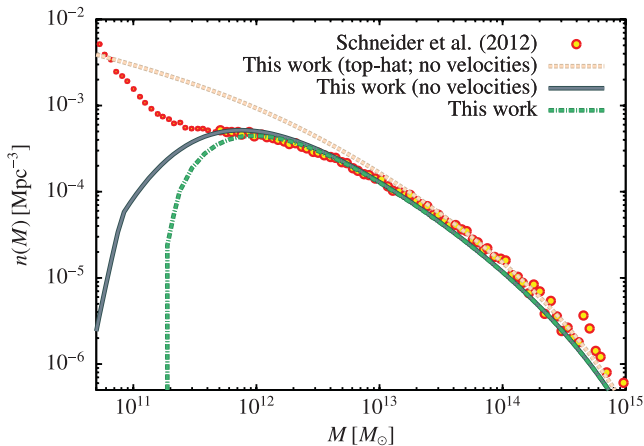
**Figure 4.** Dark matter halo mass functions, computed using the methods described in Section 2, for the three different WDM particle masses considered in this work. The labels give the particle mass and the section of this work in which that mass is used.

### 3.1 Comparison with $N$ -body simulations

#### 3.1.1 Halo mass function

$N$ -body simulations should, in principle, provide an accurate determination of the dark matter halo mass function in WDM cosmologies, provided that initial conditions are constructed carefully. The points in Fig. 5 show the mass function measured in an  $N$ -body simulation of 0.25 keV WDM carried out by Schneider et al. (2012). The upturn below  $2 \times 10^{11} M_{\odot}$  is a numerical artefact, arising from the fragmentation of filaments due to particle discreteness (Wang & White 2007). This is a challenging problem for  $N$ -body simulations of WDM as the mass scale at which the upturn appears decreases only as  $N^{-1/3}$  (where  $N$  is the particle number).

The solid line in Fig. 5 shows the mass function predicted by our calculation using the same cosmological and WDM particle parameters as in Schneider et al. (2012). Schneider et al. (2012)



**Figure 5.** The dark matter halo mass function for WDM in the case of a 0.25 keV particle. The points show the results of an  $N$ -body simulation of WDM from Schneider et al. (2012) (small points are used to indicate the region in which the  $N$ -body simulation results are unreliable as a result of being dominated by haloes formed through artificial fragmentation), while the line shows the result from this work with dark matter properties and cosmological parameters matched to those used by Schneider et al. (2012). Note that the upturn in the  $N$ -body mass function below  $2 \times 10^{11} M_{\odot}$  arises to artificial fragmentation of filaments (see Wang & White 2007).

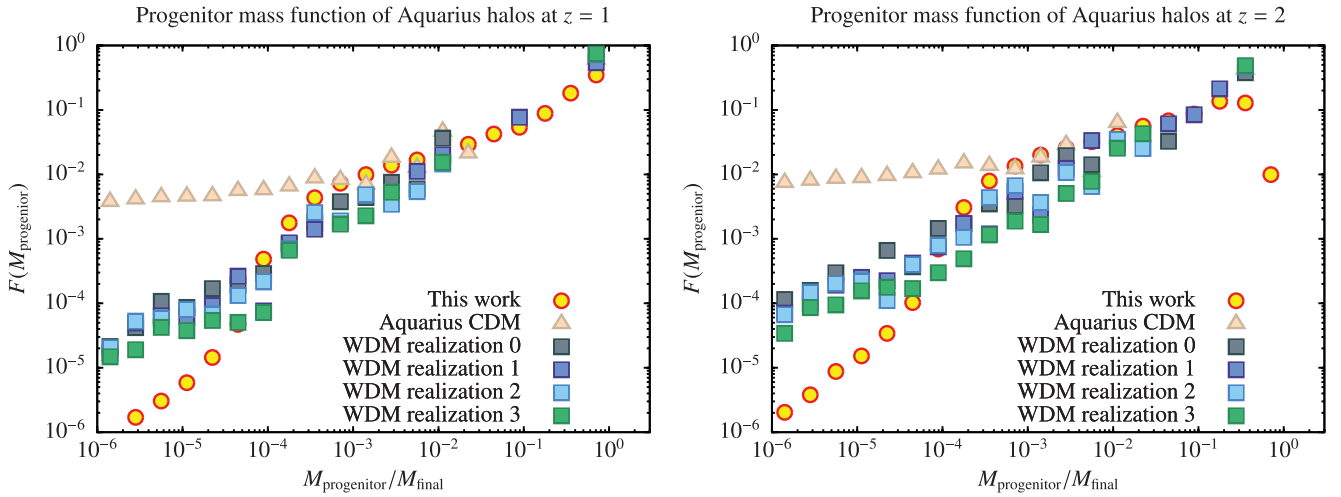
identified haloes using a friends-of-friends (FOF) algorithm with a linking length of  $b = 0.2$ , corresponding approximately to density contrasts of 200. In our model, the relevant density contrasts are those arising from the spherical top-hat collapse model (e.g. Percival 2005). We correct the masses reported by Schneider et al. (2012) for this difference by assuming that the haloes have NFW density profiles (Navarro, Frenk & White 1997) with concentrations given by the CDM results of Gao et al. (2008) modified by the WDM-to-CDM conversion factor reported by Schneider et al. (2012). Additionally, Schneider et al. (2012) do not include any velocity dispersion of dark matter particles in their initial conditions and so the fairest comparison is the one in which we do not modify the critical overdensity for collapse as described in Section 2.2.

Results of three different calculations from this work are shown in Fig. 5. Going from top to bottom, the first two lines do not include the effects of velocity dispersion (consistent with Schneider et al. 2012). The first line corresponds to a case where we use a top-hat real-space window function to determine  $\sigma(M)$ . It can be seen that this curve, while in good agreement with the  $N$ -body results for high masses, does not produce sufficient suppression of the mass function at lower masses – a similar discrepancy was seen by Barkana et al. (2001) when comparing their Press–Schechter-based model for WDM halo formation to the  $N$ -body simulations of Bode et al. (2001). The second line (blue) switches to using a sharp  $k$ -space window function as described in Section 2.1.1. For high-mass haloes, this results in only a small change in the mass function compared to the CDM case due to the difference in  $S(M)$  when computed with sharp  $k$ -space and top-hat window functions (even after scaling the barrier height to compensate for this difference as much as possible; see Section 2.2.2),<sup>10</sup> but also results in an excellent match to the suppression of the abundance of low-mass haloes, indicating that the discrepancies found in previous works were due to the artificial increasing in  $\sigma(M)$  below the cutoff which arises when a top-hat real-space filter is used. At the lowest masses the suppression of the mass function is masked in the  $N$ -body simulation as it is masked by the upturn due to artificial fragmentation of filaments. Finally, the green line adds in the effects of velocity dispersion (which are not included in the  $N$ -body simulation), illustrating the importance of this effect to accurately model the suppression of the lowest mass haloes.

#### 3.1.2 Progenitor mass functions/merger rates

A set of WDM Milky Way mass haloes have been simulated and analysed in Lovell et al. (in preparation). The haloes they simulated are WDM counterparts of the Aquarius CDM haloes presented in Springel et al. (2008), and so have masses of the order of  $10^{12} M_{\odot}$  – we will therefore refer to them as ‘Milky Way-mass WDM simulations’. Here we make use of a set of four (A–D) haloes simulated

<sup>10</sup> The slight difference between this curve and that computed using a top-hat window function is due to our inclusion of the Sheth et al. (2001) remapping. If this were not included, the factor of 1.197 increase in the barrier would precisely compensate for the difference in  $\sigma(M)$  between these two curves on large scales. With the Sheth et al. (2001) remapping included this constant factor cannot correct for the offset fully [due to the non-linear nature of the Sheth et al. (2001) remapping]. Future, high-precision work should consider recalibrating the Sheth et al. (2001) remapping to match a CDM halo mass function with  $\sigma(M)$  computed using our sharp  $k$ -space window function. This would obviate the need for a separate multiplicative increase in the barrier and could better capture the scale dependence of the required correction.



**Figure 6.** Progenitor mass functions of four realizations of the Milky Way-mass WDM dark matter haloes at  $z = 1$  and  $2$  (squares) compared with the predictions of this work (circles). For reference, the corresponding progenitor mass functions from one realization of the Aquarius CDM simulations are shown as triangles. Each panel shows the fraction of the halo mass contributed by progenitors in each mass bin. Masses are shown as a fraction of the final halo mass.

with approximately 40 million particles within their virial radii (resolution level 3 in the notation of the Aquarius project). One of the haloes (A) has also been run at higher (level 2) resolution, and we have used this to verify that the progenitor mass functions that we show in Fig. 6 have accurately converged over the range of masses plotted. The cosmological parameters for this set of simulations were chosen to match the *WMAP7* results of Komatsu et al. (2011); however, the CDM power spectrum was modified using the prescription of Bode et al. (2001) for WDM as given in our equation (1) with a smoothing scale of  $\lambda_s = 78.4$  kpc, corresponding to an approximately 2.2 keV thermal WDM particle. In each of the simulations, haloes and subhaloes were identified at each of the 128 output times using the FOF (Davis et al. 1985) and SUBFIND (Springel et al. 2001) algorithms, respectively. Merger trees of subhaloes were constructed by identifying the descendant of each subhalo and halo merger trees built from the subhalo membership of each halo as described in Jiang et al. (in preparation) and Merson et al. (2012). These halo merger trees can be thought of as FOF halo merger trees that have been cleaned up to avoid problems that occur when FOF haloes are essentially composed of two distinct haloes linked by a low-density bridge.

The issue of spurious haloes formed by the fragmentation of filaments due to numerical noise (Wang & White 2007) is investigated in Lovell et al. (in preparation). They find that most of such (sub)haloes can be identified by looking at the shape in the initial conditions of the region from which their particles originated. The spurious (sub)haloes typically originate from very flattened configurations. Lovell et al. determine a threshold on the axis ratio,  $c/a$ , of the inertia tensor of the initial particle distribution such that for a CDM simulation 99 per cent of haloes pass this cut, while in a WDM simulation most of the spurious haloes are excluded. Here we use this criterion to exclude complete subhalo branches from the  $N$ -body merger trees. In these  $N$ -body merger trees, the same subhalo is identified at subsequent epochs by tracing the fate of a fraction of its most bound particles. In this way we can identify the point at which a halo dissolves as a result of disruption within a larger (sub)halo and form a branch consisting of its main progenitor at all earlier times. We discard such a branch if at the point at which this subhalo had half its maximum mass, the subhalo fails the cut on the initial axis ratio.

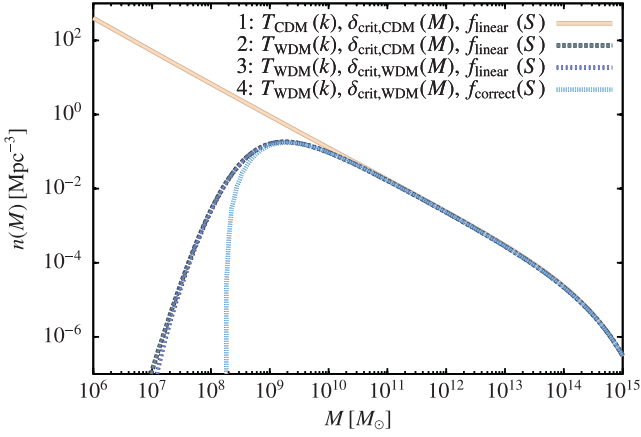
Fig. 6 compares the progenitor mass functions measured from a large number of merger tree realizations generated using the techniques described in this work (circles) compared with the progenitor mass functions measured from four Milky Way-mass WDM simulation haloes (squares). For reference, the equivalent CDM progenitor mass function from a single realization of the Aquarius simulations is shown (triangles). Given the halo-to-halo scatter, there is good agreement down to  $M_{\text{progenitor}}/M_{\text{final}} \approx 10^{-4}$ . Below this, the abundance of progenitors in the WDM  $N$ -body simulations exceeds that predicted by our techniques. At these mass scales, the artificial halo rejection algorithm reduces the number of progenitors by over an order of magnitude. This excess of low-mass progenitor haloes is consistent with the failure rate for the artificial halo rejection algorithm of slightly less than 10 per cent. We cannot rule out such a failure rate in the algorithm and so the true progenitor mass function could decline much more rapidly. The progenitor mass functions from our analytic calculations decline rapidly as they approach  $M_{\text{progenitor}}/M_{\text{final}} \approx 10^{-4}$ , but then transition to a slower, power-law decline at lower masses. As will be discussed in Section 3.3.3, this appears to be due to the finite difference approximation used to compute merger rates (see Section 2.4) and so can be suppressed as necessary by lowering the value of  $\epsilon$ .

In the CDM case, the mass function (expressed in this way) levels off to be almost constant below  $M_{\text{progenitor}}/M_{\text{final}} \sim 10^{-3}$ . In our calculation, which tracks the CDM case extremely well at high masses, this flattening begins (for  $M_{\text{progenitor}}/M_{\text{final}} \gtrsim 10^{-3}$ ) but is then interrupted by the cutoff due to WDM physics. The result is a ‘bump’ feature. The presence of such a bump is less clear in the WDM  $N$ -body progenitor mass functions (although there is a hint of it in the  $z = 1$  results) due to the noisiness of those results. This feature may have interesting implications for the expected number of surviving dwarf scale subhaloes, and so represents an interesting avenue for future investigation.

### 3.2 Limitations of only modifying power spectrum

Having demonstrated that our model is consistent with the available  $N$ -body simulation we now consider the effects of treating WDM incorrectly or incompletely in the extended Press–Schechter approach, as has previously happened in the literature (Menci et al.





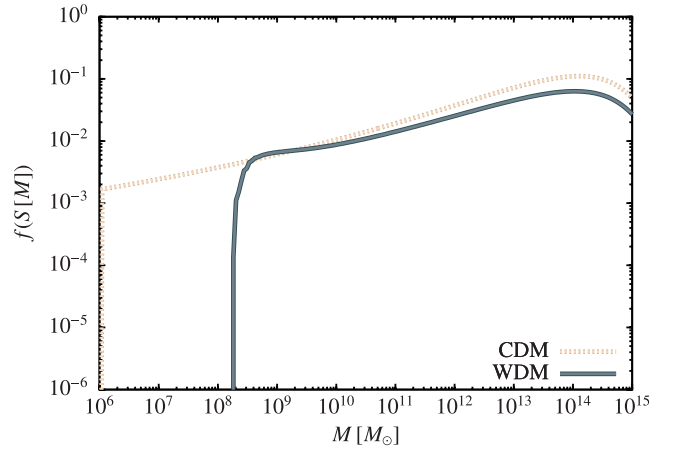
**Figure 7.** The dark matter halo mass function for CDM (line 1) and for various approximations to the 1.5 keV WDM solution. Line 2 shows the result of using the WDM transfer function, but retaining the CDM mass-independent critical overdensity and the functional form of  $f(S)$  appropriate to a linear barrier,  $f_{\text{linear}}(S)$ . Line 3 improves on this approximation by using the correct mass-dependent WDM critical overdensity, but still uses  $f_{\text{linear}}(S)$ . Finally, line 4 shows the result of using the correct functional form of  $f(S)$  for the WDM barrier. Note that, for large halo masses, all lines coincide precisely and so are hidden beneath line 4.

2012). Fig. 7 shows a series of results for the halo mass function for 1.5 keV WDM. For reference, line 1 shows the mass function for CDM derived assuming a mass-independent critical overdensity. Line 2 switches to using the WDM power spectrum, but retains the CDM mass-independent critical overdensity. There is a clear suppression of the mass function below  $M_s$ . Line 3 now uses the WDM power spectrum and adds mass dependence to the critical overdensity, but still uses the barrier crossing solution for a linear barrier<sup>11</sup> given in equation (B1). The suppression of the mass function is almost unchanged compared to the previous case where we used the CDM critical overdensity. Finally, line 4 shows the result of switching to using the correct, numerically determined, functional form for  $f(S)$ . The mass scale of the cutoff undergoes a dramatic shift of almost an order of magnitude to higher mass compared to the previous case. The reason for this is simple to understand. When using the WDM  $B(S)$  in  $f_{\text{linear}}(S)$  we are implicitly assuming that the barrier is flat at the same value of  $B(S)$  at all smaller  $S$  [this being the assumption used to derive  $f_{\text{linear}}(S)$ ]. This gives a certain crossing probability. When we numerically determine  $f_{\text{correct}}(S)$ , the full  $S$  dependence of the barrier is taken into account. In the first case, a random walk crossing  $B(S)$  at  $S$  must have remained below  $B(S)$  for all  $S' < S$ , while in the second case it must have remained below  $B(S')$  for all  $S' < S$ . Since  $B(S') < B(S)$  for all  $S' < S$ , this second condition is much more stringent, and so far fewer random walks will satisfy it. Thus, the crossing probability, and so the mass function, will be more strongly suppressed. This illustrates the importance of correctly solving the barrier crossing problem for WDM calculations.

### 3.3 Warm versus cold dark matter

In the following we compare example results for CDM and 1.5 keV WDM cases. This WDM particle mass differs from those used

<sup>11</sup> Specifically, we adopt a  $B(S)$  corresponding to the mass-dependent WDM critical overdensity, but simply use it in the solution for  $f(S)$  appropriate to a linear barrier.



**Figure 8.** The first-crossing probability,  $f(S)$ , is shown as a function of mass,  $M$ , for both CDM and WDM cases for  $z = 0$ .

in previous sections (where the choice was constrained to match those assumed in different  $N$ -body simulations). In all cases, we use the correct  $f(S)$  (determined numerically for the WDM case) and include remapping of the barrier for calculations of first-crossing probabilities and mass functions.

#### 3.3.1 Mass functions

Fig. 8 shows the first-crossing probability,  $f(S)$ , for both CDM and WDM as a function of mass,  $M$ . The two are offset by a constant multiplicative factor at large masses (small  $S$ ). This is expected – we have chosen to rescale the WDM barrier such that the *halo mass function* remains unchanged relative to the CDM expectation for large masses when we use the sharp  $k$ -space filter in our WDM calculations. The mapping from  $f(S)$  to the mass function,  $n(M)$ , is proportional to  $dS/dM$ :

$$n(M) = \frac{\bar{\rho}}{M} f(S[M]) \frac{dS}{dM}. \quad (20)$$

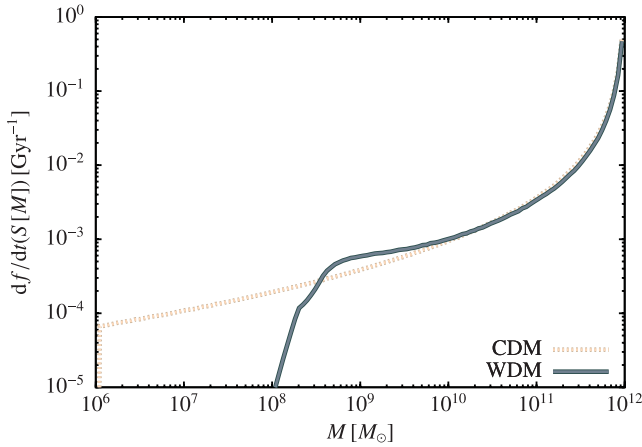
For large masses,  $dS/dM$  is larger at a fixed mass for large masses in our WDM calculations compared to the equivalent CDM calculation (a consequence of the different window functions adopted for the two cases). This difference in  $dS/dM$  offsets the difference in  $f(S)$  for large masses resulting in halo mass functions that agree between WDM and CDM.

Below around  $M_s$  the WDM first-crossing probability is suppressed due to the rapidly rising barrier  $B(S)$ . (There is a small region where the WDM  $f(S)$  exceeds that of CDM due to differences in the mapping from  $M$  to  $S$  in the two cases.)

These first-crossing distributions translate directly into halo mass functions, as shown by lines 1 and 4 of Fig. 7. As expected, the suppression in  $f(S)$  in WDM translates into a strong suppression in the mass function below around  $M_s$ . At larger masses, the two are indistinguishable.

#### 3.3.2 Merger rates

Fig. 9 shows the rate of first crossing for CDM and WDM barriers for a  $10^{12} M_\odot$  halo at  $z = 0$ . The two lines almost coincide at high mass, although the WDM line lies slightly below the CDM



**Figure 9.** The rate of first crossing,  $df/dt(S)$ , is shown as a function of mass,  $M$ , for both CDM and WDM cases, for the conditional barrier appropriate to a  $10^{12} M_{\odot}$  halo at  $z = 0$ .

line.<sup>12</sup> The merger rate in the WDM case is strongly suppressed below around  $M_s$  due to the lack of halo in that mass range. The slight enhancement in the rate of first crossing in WDM compared to CDM just above  $M_s$  is due to the different mapping between mass and variance in the two cases.

### 3.3.3 Merger tree statistics

Using merger rates computed as described in Section 2.4 we construct merger trees in both CDM and WDM cases using the algorithm described in Section 2.5, beginning with a halo of mass  $10^{12} M_{\odot}$  at  $z = 0$ . We generate 1743 trees in each case and construct the mean progenitor mass function and the mean mass accretion history.

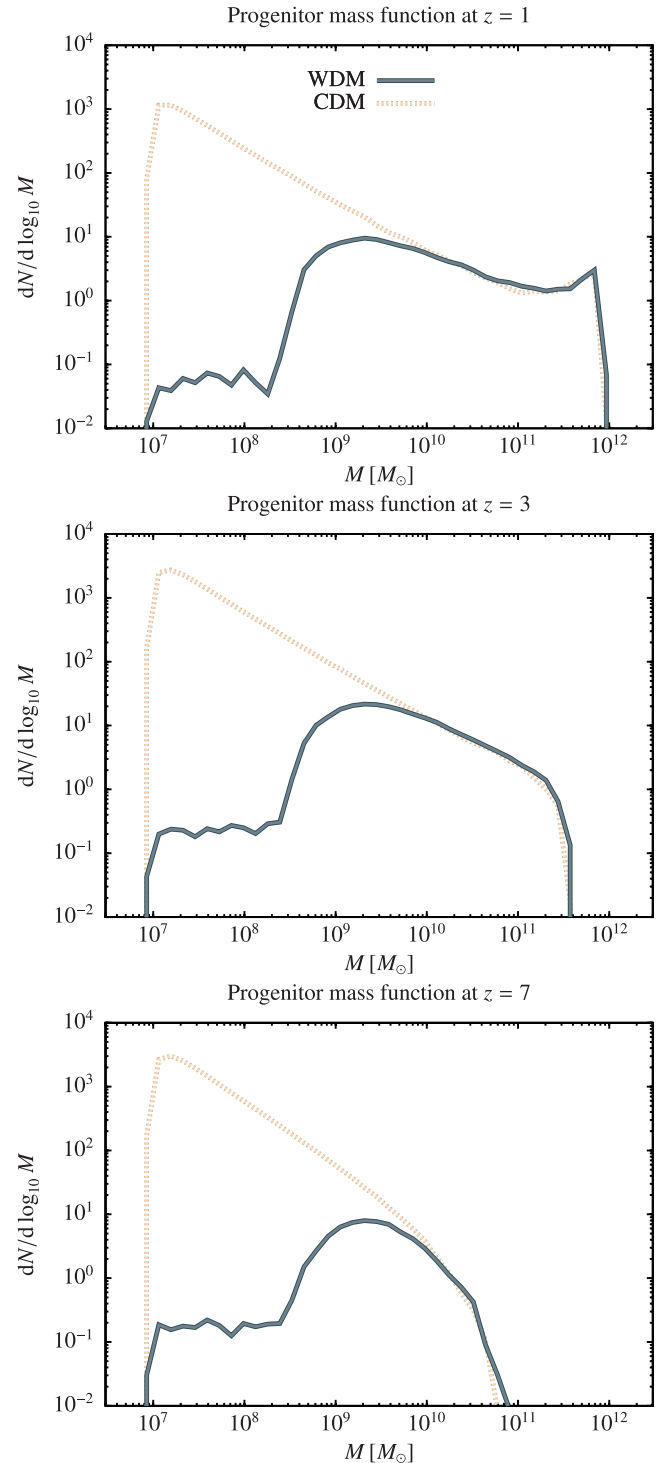
Fig. 10 shows progenitor mass functions at  $z = 1, 3$  and  $7$ . The sharp cutoff at  $10^7 M_{\odot}$  (present also in the CDM case) is due to the imposed resolution of our merger trees<sup>13</sup> (and so is unphysical). With smooth accretion included, the WDM progenitor mass function closely matches that of CDM above about  $3M_s$ , but is strongly suppressed below it at lower masses.

Well below the suppression scale a population of progenitor masses much less than  $M_s$  builds up. These are the result of smooth accretion – the first-crossing rate distribution (see Fig. 8) cuts off below  $M_s$ , so there is no way for these haloes to arise through branching of the merger tree – which gradually reduces the mass of the lowest mass haloes going back in time. The numerical robustness of our model in this regime is discussed in Appendix B4, in which we also demonstrate that the position of the peak in the progenitor mass function is numerically robust and well determined.

Fig. 11 shows the corresponding mean mass accretion histories of this halo (i.e. the mean mass of the most massive progenitor at each redshift). There is almost no difference between CDM and WDM. This is in agreement with the results of Knebe et al. (2002)

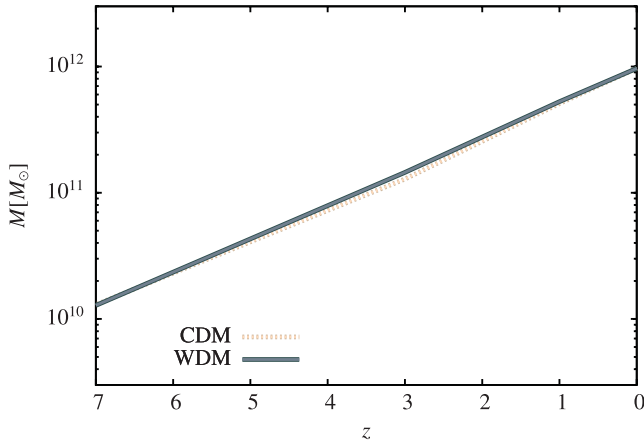
<sup>12</sup> For the same reason as in Fig. 8. Here the difference is less evident, partly because the two lines are closer to vertical, and because the largest mass scale showed here is lower than that in Fig. 8 such that  $S(M)$  differs less from WDM to CDM cases.

<sup>13</sup> The merger tree resolution is limited only by the available computational time and memory. We choose a resolution of  $10^7 M_{\odot}$  in this case to be sufficiently below  $M_s$  while keeping computing times tractable.



**Figure 10.** Progenitor mass functions derived via merger tree construction in CDM and WDM cases for a  $10^{12} M_{\odot}$  halo at  $z = 0$ .

who found almost no difference in the mass accretion histories of individual haloes in CDM and WDM  $N$ -body simulations. The two haloes presented by Knebe et al. (2002) were significantly more massive (almost  $10^{14} M_{\odot}$ ) than those considered here, but they report the same conclusion for lower mass haloes. They conclude that the number of mergers that contribute significant mass to the assembly of the haloes is unchanged between CDM and WDM case. Our results suggest that this is an accurate conclusion for



**Figure 11.** Mean mass accretion histories derived via merger tree construction in CDM and WDM cases for a  $10^{12} M_{\odot}$  halo at  $z = 0$ .

more massive haloes (the assembly of which will be dominated by haloes well above  $M_s$ ). However, for lower mass haloes, which gain a significant fraction of their mass from haloes close to  $M_s$ , our results clearly show that smooth accretion plays a crucial role in shaping the mass accretion history of WDM haloes – without it, substantial differences from the CDM case would occur.

#### 4 DISCUSSION

We have described algorithms for constructing halo mass functions and merger trees for dark matter haloes in WDM universes. Our methods improve upon previous treatments which did not correctly solve the barrier first-crossing problem (Menci et al. 2012) and which used a top-hat filter to compute  $\sigma(M)$  resulting in an overestimate of the abundance of low-mass haloes. Our results are in excellent agreement with the available  $N$ -body simulations. Illustrative results clearly demonstrate that the mass function and progenitor mass functions of WDM haloes are strongly suppressed relative to CDM below about  $M_s$ . Differences between CDM and WDM in coarse-grained statistics, such as the mass accretion history, are small for haloes well above the cutoff scale, providing that the accretion of smoothly distributed matter in WDM is accounted for.

The method that we describe has a single free parameter – the coefficient  $a$  appearing in equation (6). We have fixed the value of this parameter to match the location of the turnover in the  $N$ -body mass function reported by Schneider et al. (2012). This introduces a dependency in one of the simulations to which we compare our model. Nevertheless, as discussed in Section 2.1.1, we expect that the same value of  $a$  will be appropriate for all WDM particle masses of interest. The limited evidence available from our present work (i.e. that  $a$  chosen to fit the mass function for 0.25 keV WDM also successfully matches the progenitor mass functions for 2.20 keV WDM) certainly supports this claim.

Our approach has the advantage, compared to  $N$ -body simulations of WDM, of not being affected by numerical noise in the particle distribution which leads to the formation of large numbers of artificial low-mass haloes (Wang & White 2007) and of being substantially faster to evaluate mass functions and progenitor distributions. Using the techniques developed in this work, the procedure for applying them to dark matter with different phenomenology or to other physics that modifies the power spectrum or excursion set barrier is straightforward:

- (i) determine the linear theory power spectrum of density perturbations from the dark matter (or other) physics;
- (ii) determine, through analytic calculation or idealized  $N$ -body simulations, the critical linear theory overdensity for collapse, which will depend on the (thermal and interaction) physics of the dark matter particle.

Given these two inputs our techniques can be used to determine the resulting halo mass function and merger histories of dark matter haloes consistent with the input physics. The accuracy of our methods for phenomenology beyond that exhibited by WDM remains to be tested, but the success in this case leads us to expect that our methods will be generally applicable.

The nature of the dark matter distribution on small mass scales will be investigated by future lensing programmes (Keeton & Moustakas 2009; Vegetti & Koopmans 2009a). The techniques described in this work will allow detailed statistical predictions to be made for the expected numbers and masses of dark matter substructure as a function of dark matter particle properties.

To make accurate predictions for the dark matter subhalo distribution we have addressed only the first part of the problem, namely halo formation and merging. The second part, halo destruction by tidal forces must also be addressed. We plan to explore this process using the methods of Benson et al. (2002a; see also Taylor & Babul 2004), together with prescriptions for the internal structure of WDM haloes (which will, of course, differ from that of CDM haloes; Maccio' et al. 2012; Macciò et al. 2012).

The methods described in Section 2 have been implemented within the open source semi-analytic galaxy formation code, GALACTICUS (Benson 2012). All results presented in this work were generated using GALACTICUS v0.9.1 r903. Control files and scripts to generate all results presented in this paper using GALACTICUS can be found at <http://users.obs.carnegiescience.edu/abenson/galacticus/parameters/dmMergingBeyondCDM1.tar.bz2>.

#### ACKNOWLEDGMENTS

We thank Rennan Barkana for supplying results from his calculations of halo collapse in WDM universes, and Annika Peter for invaluable discussions and comments on an earlier draft of this paper. The work of LAM was carried out at Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. LAM acknowledges support by the NASA ATPF program.

#### REFERENCES

- Abazajian K., Fuller G. M., Patel M., 2001, *Phys. Rev. D*, 64, 23501  
Ahn K., Shapiro P. R., 2005, *MNRAS*, 363, 1092  
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, *ApJ*, 304, 15  
Barkana R., Haiman Z., Ostriker J. P., 2001, *ApJ*, 558, 482  
Benson A. J., 2012, *New Astron.*, 17, 175  
Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002a, *MNRAS*, 333, 156  
Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002b, *MNRAS*, 333, 177  
Bode P., Ostriker J. P., Turok N., 2001, *ApJ*, 556, 93  
Boehm C., Schaeffer R., 2005, *A&A*, 438, 419  
Bond J. R., Szalay A. S., 1983, *ApJ*, 274, 443  
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, *ApJ*, 379, 440  
Bower R. G., 1991, *MNRAS*, 248, 332  
Boyarsky A., Lesgourgues J., Ruchayskiy O., Viel M., 2009, *J. Cosmol. Astropart. Phys.*, 05, 012  
Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, *MNRAS*, 415, L40  
Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2012, *MNRAS*, 422, 1203

- Brandbyge J., Hannestad S., Haugbille T., Thomsen B., 2008, *J. Cosmol. Astropart. Phys.*, 08, 020
- Cheng H.-C., Feng J. L., Matchev K. T., 2002, *Phys. Rev. Lett.*, 89, 211301
- Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- de Blok W. J. G., 2010, *Adv. Astron.*, 2010, 5
- Dodelson S., Widrow L. M., 1994, *Phys. Rev. Lett.*, 72, 17
- Donato F., Gentile G., Salucci P., 2004, *MNRAS*, 353, L17
- Donato F. et al., 2009, *MNRAS*, 397, 1169
- Eisenstein D. J., Hu W., 1999, *ApJ*, 511, 5
- Ellis J., Kim J. E., Nanopoulos D. V., 1984, *Phys. Lett. B*, 145, 181
- Feng J. L., Rajaraman A., Takayama F., 2003, *Phys. Rev. Lett.*, 91, 11302
- Feng J. L., Kaplinghat M., Tu H., Yu H.-B., 2009, *J. Cosmol. Astropart. Phys.*, 07, 004
- Ferramacho L. D., Blanchard A., Zolnierowski Y., 2009, *A&A*, 499, 21
- Font A. S. et al., 2011, *MNRAS*, 417, 1260
- Gao L., Navarro J. F., Cole S., Frenk C. S., White S. D. M., Springel V., Jenkins A., Neto A. F., 2008, *MNRAS*, 387, 536
- Gorbunov D., Khmelnsky A., Rubakov V., 2008, *J. High Energy Phys.*, 12, 055
- Governato F. et al., 2012, *MNRAS*, 422, 1231
- Green A. M., Hofmann S., Schwarz D. J., 2004, *MNRAS*, 353, L23
- Hogan C. J., Dalcanton J. J., 2000, *Phys. Rev. D*, 62, 63511
- Hubisz J., Meade P., 2005, *Phys. Rev. D*, 71, 35016
- Jungman G., Kamionkowski M., Griest K., 1996, *Phys. Rep.*, 267, 195
- Kawasaki M., Sugiyama N., Yanagida T., 1997, *Mod. Phys. Lett. A*, 12, 1275
- Keeton C. R., Moustakas L. A., 2009, *ApJ*, 699, 1720
- Knebe A., Devriendt J. E. G., Mahmood A., Silk J., 2002, *MNRAS*, 329, 813
- Komatsu E. et al., 2011, *ApJS*, 192, 18
- Kuhlen M., Krumholz M. R., Madau P., Smith B. D., Wise J., 2012, *ApJ*, 749, 36
- Kuzio de Naray R., Kaufmann T., 2011, *MNRAS*, 414, 3617
- Kuzio de Naray R., Spekkens K., 2011, *ApJ*, 741, L29
- Kuzio de Naray R., Martinez G. D., Bullock J. S., Kaplinghat M., 2010, *ApJ*, 710, L161
- Lacey C., Cole S., 1993, *MNRAS*, 262, 627
- Li Y.-S., Helmi A., De Lucia G., Stoehr F., 2009, *MNRAS*, 397, L87
- Libeskind N. I., Cole S., Frenk C. S., Okamoto T., Jenkins A., 2007, *MNRAS*, 374, 16
- Loeb A., Zaldarriaga M., 2005, *Phys. Rev. D*, 71, 103520
- Lovell M. R. et al., 2012, *MNRAS*, 420, 2318
- Maccio' A. V., Ruchayskiy O., Boyarsky A., Munoz-Cuartas J. C., 2012, *arXiv:1202.2858*
- Macciò A. V., Paduoiu S., Anderhalden D., Schneider A., Moore B., 2012, *MNRAS*, 424, 1105
- Markevitch M., Gonzalez A., Clowe D., Vikhlinin A., 2004, in *Plionis M., ed., Multiwavelength Cosmology. Proceedings of the "Multiwavelength Cosmology" Conference, Astrophysics and Space Science Library, Vol. 301. Kluwer Academic Publishers, Dordrecht*, p. 263
- Menci N., Fiore F., Lamastra A., 2012, *MNRAS*, 421, 2384
- Merson A. I. et al., 2012, *arXiv:1206.4049*
- 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., 1997, *ApJ*, 490, 493
- Newman A. B., Treu T., Ellis R. S., Sand D. J., Richard J., Marshall P. J., Capak P., Miyazaki S., 2009, *ApJ*, 706, 1078
- Newman A. B., Treu T., Ellis R. S., Sand D. J., 2011, *ApJ*, 728, L39
- Oh S.-H., Brook C., Governato F., Brinks E., Mayer L., de Blok W. J. G., Brooks A., Walter F., 2011, *AJ*, 142, 24
- Parkinson H., Cole S., Helly J., 2008, *MNRAS*, 383, 557
- Peebles P. J. E., 1982, *ApJ*, 263, L1
- Percival W. J., 2005, *A&A*, 443, 819
- Percival W. J. et al., 2007, *ApJ*, 657, 645
- Pfrommer C., Chang P., Broderick A. E., 2012, *ApJ*, 752, 24
- Pontzen A., Governato F., 2012, *MNRAS*, 421, 3464
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Raffelt G. G., 1990, *Phys. Rep.*, 198, 1
- Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradac M., 2008, *ApJ*, 679, 1173
- Salucci P., 2001, *MNRAS*, 320, L1
- Salucci P., Wilkinson M. I., Walker M. G., Gilmore G. F., Grebel E. K., Koch A., Frigerio Martins C., Wyse R. F. G., 2012, *MNRAS*, 420, 2034
- Sánchez A. G., Crocce M., Cabré A., Baugh C. M., Gaztañaga E., 2009, *MNRAS*, 400, 1643
- Schneider A., Smith R. E., Macciò A. V., Moore B., 2012, *MNRAS*, 424, 684
- Seljak U. et al., 2005, *Phys. Rev. D*, 71, 103515
- Shaposhnikov M., Tkachev I., 2006, *Phys. Lett. B*, 639, 414
- Sheth R. K., 1998, *MNRAS*, 300, 1057
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Sigurdson K., Kamionkowski M., 2004, *Phys. Rev. Lett.*, 92, 171302
- Simon J. D., Bolatto A. D., Leroy A., Blitz L., Gates E. L., 2005, *ApJ*, 621, 757
- Spergel D. N., Steinhardt P. J., 2000, *Phys. Rev. Lett.*, 84, 3760
- 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
- Starkenburg E. et al., 2012, *arXiv:1206.0020*
- Stringer M., Cole S., Frenk C. S., 2010, *MNRAS*, 404, 1129
- Taylor J. E., Babul A., 2004, *MNRAS*, 348, 811
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
- Turner M. S., 1990, *Phys. Rep.*, 197, 67
- Vegetti S., Koopmans L. V. E., 2009a, *MNRAS*, 392, 945
- Vegetti S., Koopmans L. V. E., 2009b, *MNRAS*, 400, 1583
- Vegetti S., Czoske O., Koopmans L. V. E., 2010a, *MNRAS*, 407, 225
- Vegetti S., Koopmans L. V. E., Bolton A., Treu T., Gavazzi R., 2010b, *MNRAS*, 408, 1969
- Vegetti S., Lagattuta D. J., McKean J. P., Auger M. W., Fassnacht C. D., Koopmans L. V. E., 2012, *Nat*, 481, 341
- Vera-Ciro C. A., Helmi A., Starkenburg E., Breddels M. A., 2012, *arXiv:1202.6061*
- Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., Rauch M., Sargent W. L. W., 2008, *Phys. Rev. Lett.*, 100, 41304
- Wang J., White S. D. M., 2007, *MNRAS*, 380, 93
- Wang J., Frenk C. S., Navarro J. F., Gao L., Sawala T., 2012, *MNRAS*, 424, 2715
- Wolf J., Bullock J. S., 2012, *arXiv:1203.4240*
- Zhang J., Hui L., 2006, *ApJ*, 641, 641

## APPENDIX A: NUMERICAL METHOD

In this section, we describe our numerical method for solving the excursion set barrier first-crossing problem (equation 12). In the absence of a barrier,  $P(\delta, S)$  would be equal to  $P_0(\delta, S)$  which is simply a Gaussian distribution with variance  $S$ :

$$P_0(\delta, S) = \frac{1}{\sqrt{2\pi S}} \exp\left(-\frac{\delta^2}{2S}\right). \quad (\text{A1})$$



Since the barrier absorbs any random walks which cross at smaller  $S$ , the actual  $P(\delta, S)$  must therefore be given by

$$P(\delta, S) = P_0(\delta, S) - \int_0^S f(S') P_0[\delta - B(S'), S - S'] dS'. \quad (\text{A2})$$

The second term on the right-hand side of equation (A2) represents the distribution of random trajectories originating from the point  $(S, B(S))$ . The integral therefore gives the fraction of trajectories which crossed the barrier at  $S < S'$  and which can now be found at  $(S, \delta)$ .

Using this result, we can rewrite equation (12) as

$$1 = \int_0^S f(S') dS' + \int_{-\infty}^{B(S)} \left[ P_0(\delta, S) - \int_0^S f(S') P_0(\delta - B(S'), S - S') dS' \right] d\delta, \quad (\text{A3})$$

in general, and for the Gaussian distribution of equation (A1) as

$$1 = \int_0^S f(S') dS' + \int_{-\infty}^{B(S)} \left[ \frac{1}{\sqrt{2\pi S}} \exp\left(-\frac{\delta^2}{2S}\right) - \int_0^S f(S') \frac{1}{\sqrt{2\pi(S-S')}} \exp\left(-\frac{[\delta - B(S')]^2}{2(S-S')}\right) dS' \right] d\delta. \quad (\text{A4})$$

The integral over  $d\delta$  can be carried out analytically to give

$$1 = \int_0^S f(S') dS' + \operatorname{erf}\left[\frac{B(S)}{\sqrt{2S}}\right] - \int_0^S f(S') \operatorname{erf}\left[\frac{B(S) - B(S')}{\sqrt{2(S-S')}}\right] dS'. \quad (\text{A5})$$

We now discretize equation (A5). Specifically, we divide the  $S$  space into  $N$  intervals defined by the points:

$$S_i = \begin{cases} 0 & \text{if } i = 0 \\ \sum_{j=0}^{i-1} \Delta S_j & \text{if } i > 1. \end{cases} \quad (\text{A6})$$

Note that  $f(0) = 0$  by definition, so  $f(S_0) = 0$  always. We choose  $\Delta S_i = S_{\max}/N$  (i.e. uniform spacing in  $S$ ) when computing first-crossing distributions, and  $\Delta S_i \propto S_i$  (i.e. uniform spacing in  $\log(S)$ ) when computing first-crossing rates.

Discretizing the integrals in equation (A5) gives

$$\int_0^{S_j} f(S') dS' = \sum_{i=0}^{j-1} \frac{f(S_i) + f(S_{i+1})}{2} \Delta S_i \quad (\text{A7})$$

and

$$\int_0^{S_j} f(S') \operatorname{erf}\left[\frac{B(S) - B(S')}{\sqrt{2(S-S')}}\right] dS' = \sum_{i=0}^{j-1} \frac{1}{2} \left( f(S_i) \operatorname{erf}\left[\frac{B(S_j) - B(S_i)}{\sqrt{2(S_j - S_i)}}\right] + f(S_{i+1}) \operatorname{erf}\left[\frac{B(S_j) - B(S_{i+1})}{\sqrt{2(S_j - S_{i+1})}}\right] \right) \Delta S_i. \quad (\text{A8})$$

We can now rewrite equation (A5) in the discretized form as

$$1 = \sum_{i=0}^{j-1} \frac{f(S_i) + f(S_{i+1})}{2} \Delta S_i + \operatorname{erf}\left[\frac{B(S_j)}{\sqrt{2S_j}}\right] - \frac{1}{2} \sum_{i=0}^{j-1} \left( f(S_i) \operatorname{erf}\left[\frac{B(S_j) - B(S_i)}{\sqrt{2(S_j - S_i)}}\right] + f(S_{i+1}) \operatorname{erf}\left[\frac{B(S_j) - B(S_{i+1})}{\sqrt{2(S_j - S_{i+1})}}\right] \right) \Delta S_i. \quad (\text{A9})$$

Solving equation (A9) for  $f(S_j)$ ,

$$\begin{aligned} \left( \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left[\frac{B(S_j) - B(S_j)}{\sqrt{2(S_j - S_j)}}\right] \right) \Delta S_{j-1} f(S_j) &= 1 - \sum_{i=0}^{j-2} \frac{f(S_i) + f(S_{i+1})}{2} \Delta S_i - \frac{f(S_{j-1})}{2} \Delta S_{j-1} - \operatorname{erf}\left[\frac{B(S_j)}{\sqrt{2S_j}}\right] \\ &\quad + \frac{1}{2} \sum_{i=0}^{j-2} \left( f(S_i) \operatorname{erf}\left[\frac{B(S_j) - B(S_i)}{\sqrt{2(S_j - S_i)}}\right] + f(S_{i+1}) \operatorname{erf}\left[\frac{B(S_j) - B(S_{i+1})}{\sqrt{2(S_j - S_{i+1})}}\right] \right) \Delta S_i \\ &\quad + \frac{1}{2} f(S_{j-1}) \operatorname{erf}\left[\frac{B(S_j) - B(S_{j-1})}{\sqrt{2(S_j - S_{j-1})}}\right] \Delta S_{j-1}. \end{aligned} \quad (\text{A10})$$

For all barriers that we consider,

$$\operatorname{erf}\left[\frac{B(S_j) - B(S_j)}{\sqrt{2(S_j - S_j)}}\right] = 0. \quad (\text{A11})$$

We can then simplify equation (A10) as

$$\begin{aligned} f(S_j) &= \frac{2}{\Delta S_{j-1}} \left[ 1 - \sum_{i=0}^{j-2} \frac{f(S_i) + f(S_{i+1})}{2} \Delta S_i - \frac{f(S_{j-1})}{2} \Delta S_{j-1} - \operatorname{erf}\left[\frac{B(S_j)}{\sqrt{2S_j}}\right] + \frac{1}{2} \sum_{i=0}^{j-2} \left( f(S_i) \operatorname{erf}\left[\frac{B(S_j) - B(S_i)}{\sqrt{2(S_j - S_i)}}\right] \right. \right. \\ &\quad \left. \left. + f(S_{i+1}) \operatorname{erf}\left[\frac{B(S_j) - B(S_{i+1})}{\sqrt{2(S_j - S_{i+1})}}\right] \right) \Delta S_i + \frac{1}{2} f(S_{j-1}) \operatorname{erf}\left[\frac{B(S_j) - B(S_{j-1})}{\sqrt{2(S_j - S_{j-1})}}\right] \Delta S_{j-1} \right]. \end{aligned} \quad (\text{A12})$$

Consolidating terms in the summations:

$$f(S_j) = \frac{2}{\Delta S_{j-1}} \left[ 1 - \operatorname{erf} \left\{ \frac{B(S_j)}{\sqrt{2S_j}} \right\} - \sum_{i=0}^{j-1} \left\{ 1 - \operatorname{erf} \left[ \frac{B(S_j) - B(S_i)}{\sqrt{2(S_j - S_i)}} \right] \right\} f(S_i) \frac{\Delta S_{i-1} + \Delta S_i}{2} \right]. \quad (\text{A13})$$

In the case of constant  $\Delta S_i (= \Delta S)$  this can be simplified further as

$$f(S_j) = \frac{2}{\Delta S} \left[ 1 - \operatorname{erf} \left\{ \frac{B(S_j)}{\sqrt{2S_j}} \right\} \right] - 2 \sum_{i=0}^{j-1} \left\{ 1 - \operatorname{erf} \left[ \frac{B(S_j) - B(S_i)}{\sqrt{2(S_j - S_i)}} \right] \right\} f(S_i). \quad (\text{A14})$$

In either case (i.e. equations A13 and A14), solution proceeds recursively:  $f(S_0) = 0$  by definition,  $f(S_1)$  depends only on the known barrier and  $f(S_0)$ , and  $f(S_j)$  depends only on the known barrier and  $f(S_{<j})$ .

## APPENDIX B: NUMERICAL TESTS

In this appendix, we examine the numerical accuracy and robustness of our methods.

### B1 First-crossing probability solutions

To test the accuracy of our numerical solver for the first-crossing distribution we compare its results to the known analytic solution for a constant barrier. Specifically, we consider the CDM case with no remapping of the barrier such that  $B(S) = \delta_c = \text{constant}$  and

$$f(S) = \frac{\delta_c}{S\sqrt{2\pi S}} \exp \left[ -\frac{\delta_c^2}{2S} \right]. \quad (\text{B1})$$

Fig. B1 shows the fractional difference between results obtained using our numerical method and the analytic solution in this case. The numerical error is small except for at the highest masses (smallest variances) where the discreteness in our grid becomes an issue and the error reaches a few per cent. The accuracy achieved is sufficient for the present work (where we are mostly concerned with low-mass systems) and is easily improved by adopting a smaller value of  $\Delta S$ .

### B2 First-crossing rate solutions

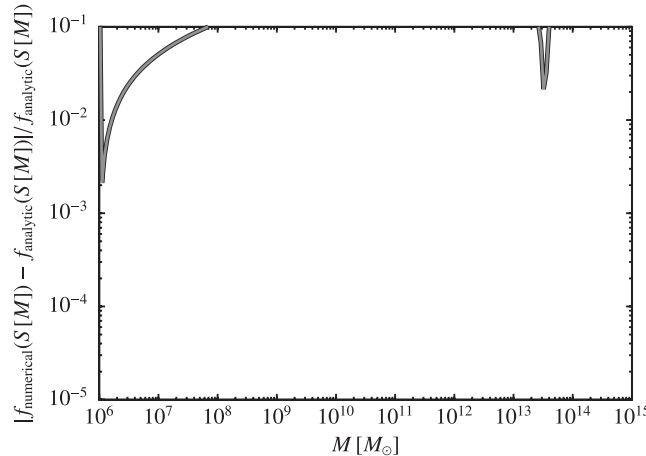
Once again, to test the accuracy of our numerical solver in the case of computing first-crossing rates we compare its results to the analytic solution for a constant barrier. Adopting the same constant barrier model as in Appendix B1, the crossing rate is

$$\frac{df}{dt} = \frac{1}{\sqrt{2\pi}} \frac{1}{(S - S_0)^{3/2}} \frac{d\delta_c}{dt}, \quad (\text{B2})$$

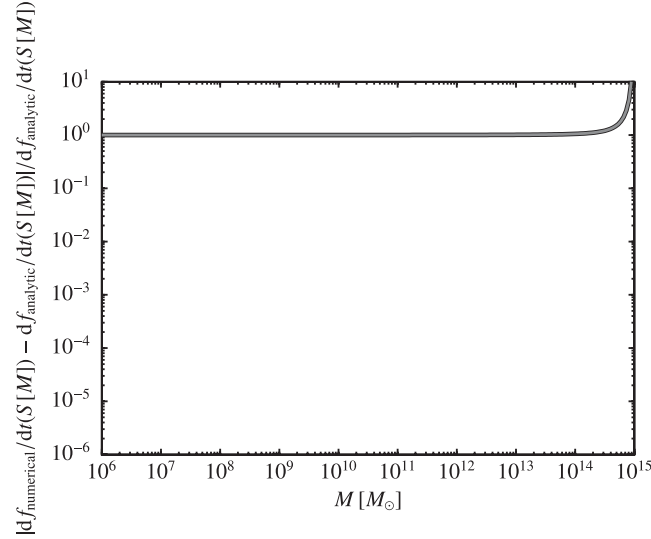
where  $S_0$  is the variance corresponding to the mass of the final halo (i.e. the halo formed through the merger event). Fig. B2 shows the fractional error in the numerically derived merger rate compared to this analytic solution for  $\epsilon = 0.01$ . The fractional error is constant across the range of masses shown and is limited by the value of  $\epsilon$  chosen.

### B3 Testing the Parkinson–Cole–Helly algorithm on smaller mass scales

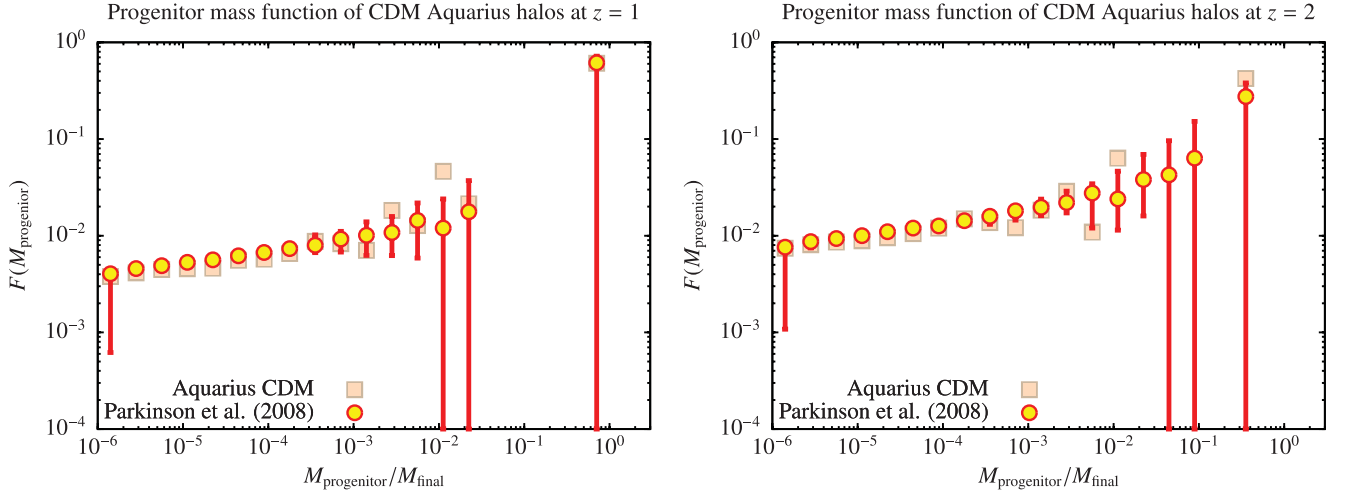
The Parkinson et al. (2008) empirical modification to the merger tree branching rate was calibrated against  $N$ -body merger trees drawn from the Millennium Simulation. As such, it has been tested for masses above approximately  $10^{10} M_\odot$ . Here we employ the same modification for much lower masses. Fig. B3 compares progenitor mass functions generated by the Parkinson et al. (2008) empirical modification with those extracted from the Aquarius CDM simulations of Springel et al. (2008) which resolve haloes of masses  $10^6 M_\odot$ . It can be seen that the



**Figure B1.** The absolute fractional error in the excursion set first-crossing distribution function,  $f(S)$ , computed using our numerical method compared to the analytic solution for a constant barrier.



**Figure B2.** The absolute fractional error in the excursion set first-crossing rate function,  $df(S)/dt$ , computed using our numerical method compared to the analytic solution for a constant barrier.



**Figure B3.** Progenitor mass functions of the Aquarius CDM dark matter haloes (Springel et al. 2008) at  $z = 1$  and  $2$  (histograms) compared with the predictions of the Parkinson et al. (2008) algorithm. Each panel shows the fraction of the halo mass contributed by progenitors in each mass bin. Masses are shown as a fraction of the final halo mass. The error bars on the predictions from the Parkinson et al. (2008) algorithm indicate the 20th and 80th percentiles of the distribution of progenitor mass functions. While this algorithm was calibrated on the much higher mass haloes found in the Millennium Simulation (Springel et al. 2005), it can be seen to work extremely well at these lower masses also.

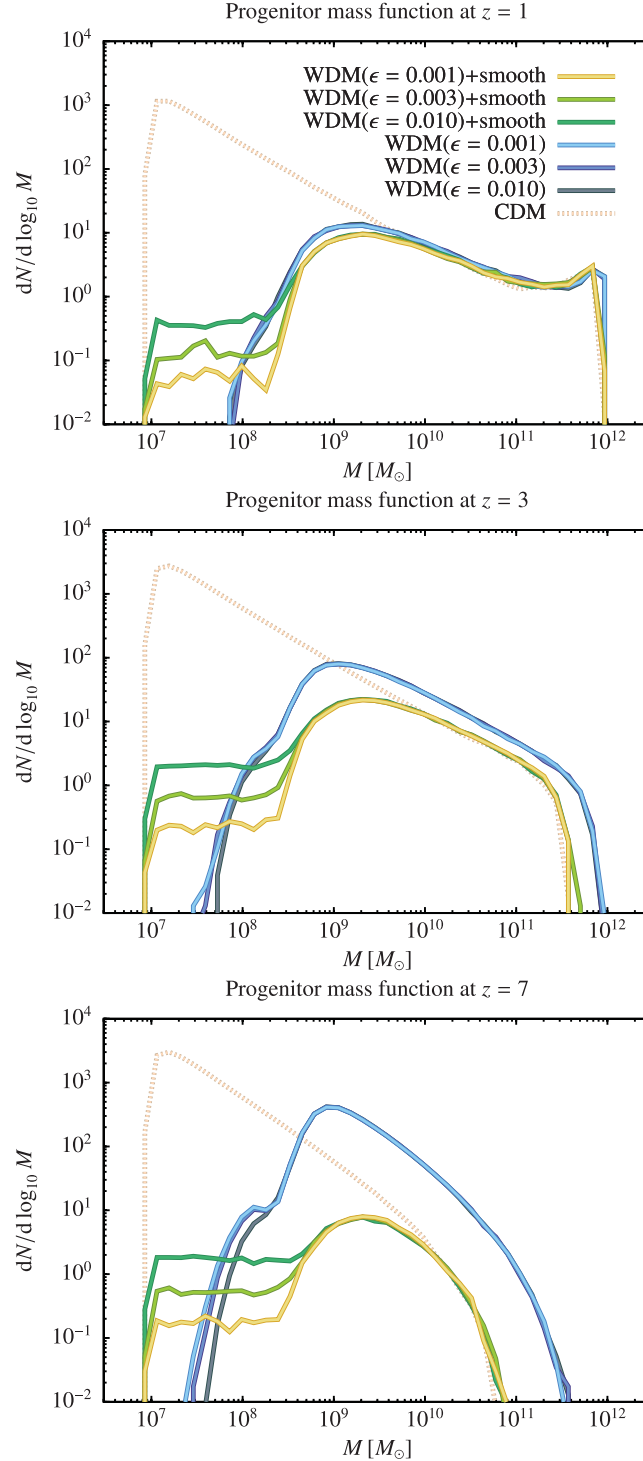
Parkinson et al. (2008) empirical modification performs very well (with unchanged parameter values) for these much lower mass haloes in the CDM case also.

#### B4 Merger tree accuracy and smooth accretion

To test the convergence of our merger trees with respect to the parameter  $\epsilon$  used in our numerical determination of the first-crossing rate distribution (see Section 2.4), we compute progenitor mass functions of  $10^{12} M_{\odot}$  haloes for  $\epsilon = 0.010, 0.003$  and  $0.001$ . Additionally, we perform these calculations both with and without the smooth accretion term of Section 2.5.1.

Fig. B4 shows progenitor mass functions at  $z = 1, 3$  and  $7$ . When smooth accretion is included, the WDM progenitor mass function closely matches that of CDM above about  $3M_s$ , but is strongly suppressed below it at lower masses. If smooth accretion is ignored the WDM progenitor mass function evolves much more slowly and diverges from the CDM case even at the highest masses. Ignoring this smooth accretion leads to significantly biased results.

It can clearly be seen that our WDM results are converged with respect to the  $\epsilon$  parameter except at masses well below the suppression scale in the progenitor mass function in the cases where smooth accretion is included. Here, a population of progenitor masses much less than  $M_s$  builds up. These are the result of smooth accretion – the first-crossing rate distribution (see Fig. 8) cuts off below  $M_s$ , so there is no way for these haloes to arise through branching of the merger tree – which gradually reduces the mass of the lowest mass haloes going back in time.



**Figure B4.** Progenitor mass functions derived via merger tree construction in CDM and WDM cases for a  $10^{12} M_{\odot}$  halo at  $z = 0$ . For the WDM case, results are shown for different values of  $\epsilon$  and also for cases with and without the accretion of smoothly distributed matter.

Our numerical determination of the first-crossing rate function is currently not robust in its determination of smooth accretion rates for these lowest mass (highest variance) haloes where the excursion set barrier is a very rapidly changing function of variance and our discretization of  $S$  used to obtain a numerical solution inevitably does a poor job of resolving the barrier. This could of course be mitigated by using a yet smaller value of  $\epsilon$  (requiring substantially increased precision in the numerical solutions) or a finer grid in  $S$  (requiring both increased precision and substantially more computing time). Nevertheless, the position of the peak in the progenitor mass function is well determined, and the differences in the abundance of low-mass progenitors have negligible effect on the mean mass accretion histories of haloes considered in this work.