The locations of halo formation and the peaks formalism

Oliver Hahn* and Aseem Paranjape*

Institute for Astronomy, Department of Physics, ETH Zürich, CH-8093 Zürich, Switzerland

Accepted 2013 November 19. Received 2013 October 30; in original form 2013 August 19

ABSTRACT

We investigate the problem of predicting the halo mass function from the properties of the Lagrangian density field. We focus on a perturbation spectrum with a small-scale cut-off (as in warm dark matter cosmologies). This cut-off results in a strong suppression of low-mass objects, providing additional leverage to rigorously test which perturbations collapse and to what mass. We find that all haloes are consistent with forming near peaks of the initial density field, with a strong correlation between protohalo density and ellipticity. We demonstrate that, while standard excursion set theory with correlated steps completely fails to reproduce the mass function, the inclusion of the peaks constraint leads to the correct number of haloes but significantly underpredicts the masses of low-mass objects (with the predicted halo mass function at low masses behaving like \(dn/d\ln m \sim m^{n_2}\)). This prediction is very robust and cannot be easily altered within the framework of a single collapse barrier. The nature of collapse in the presence of a small-scale cut-off thus reveals that excursion set calculations require a more detailed understanding of the collapse-time of a general ellipsoidal perturbation to predict the ultimate collapsed mass of a peak – a problem that has been hidden in the large abundance of small-scale structure in cold dark matter. We demonstrate how this problem can be resolved within the excursion set framework.

Key words: methods: analytical – methods: numerical – galaxies: formation – cosmology: theory – dark matter – large-scale structure of Universe.

1 INTRODUCTION

Where and when do dark matter haloes form? The problem of identifying the locations where gravitational collapse leads to bound haloes of dark matter, and predicting the cosmic time at which this will occur, is among the oldest problems of cosmic structure formation theory. The idea that small perturbations in the primordial matter density formed the seeds of the large-scale structure we observe at present is among the cornerstones of our current picture of the evolution of the Universe. An understanding of the relevant processes and a robust theoretical model enables us to map properties such as the abundance and clustering of dark matter haloes – which are directly tied to the corresponding observed properties of galaxies – to well-understood statistical properties of the initial dark matter density.

Although this problem can now be tackled directly using numerical simulations of large, cosmological volumes, it is still important to explore analytical approximations and identify the key physical features that decide the sites of halo formation. The main motivation behind this exercise is to gain a better understanding of the physical processes that affect structure formation in the Universe. From a practical viewpoint, however, this can also lead to useful, fast approximations to the halo mass function, clustering and predictions of collapse time for a given patch in the initial conditions. The latter especially could be useful from the point of view of ‘semi-analytic’ mock catalogue algorithms such as PThALOS (Scoccimarro & Sheth 2002), PINOCCHIO (Monaco, Theuns & Taffoni 2002; Monaco et al. 2013), COLA (Tassev, Zaldarriaga & Eisenstein 2013), ALPT (Kitaura & Heß 2013), etc., which are becoming increasingly popular in the construction of covariance matrices in current and upcoming surveys such as Baryon Oscillation Spectroscopic Survey (BOSS; Manera et al. 2013), WigleZ (Marín et al. 2013), Euclid (Laureijs et al. 2011) and others.

Our focus in this paper is on the mass function of dark matter haloes, which is the most basic diagnostic of the fully non-linear density field. Analytical descriptions of the halo mass function have traditionally used two parallel approaches: the excursion set approach (Press & Schechter 1974; Epstein 1983; Peacock & Heavens 1990; Bond et al. 1991; Lacey & Cole 1993; Sheth 1998; Sheth, Mo & Tormen 2001; Maggiore & Riotto 2010; Achitouv et al. 2012; Musso & Sheth 2012, 2013; Paranjape, Lam & Sheth 2012) and the peaks formalism (Bardeen et al. 1986; Bond 1989; Appel & Jones 1990; Manrique et al. 1998; Hanami 2001), both of which aim to characterize the locations of collapse in the initial conditions using some criteria. The former relies on counting sufficiently overdense regions in the initial conditions, which it maps to collapsed haloes in the final, gravitationally evolved density field, while the latter

*E-mail: hahn@phys.ethz.ch (OH); aseemp@phys.ethz.ch (AP)
associates haloes specifically to peaks in the initial matter density. In other words, while both approaches rely on the statistical properties of the initial conditions to predict final halo abundances, the excursion set approach does this by treating all locations in the initial conditions on the same footing, while the peaks formalism treats density peaks as being special.

The key aspect of the excursion set approach (Bond et al. 1991), which is missing in the traditional peaks approach (Bardeen et al. 1986), is that it explicitly accounts for the so-called ‘cloud-in-cloud’ problem which avoids overcounting overdense regions embedded in larger overdense regions as individual objects. The ‘peak-patch’ approach of Bond & Myers (1996) is a numerical prescription for unifying the two approaches to solve the cloud-in-cloud problem for peaks, or, equivalently, to study excursion sets for a special subset of initial locations, namely peaks. Recent work (Musso & Sheth 2012; Paranjape & Sheth 2012) has shown that this can also be achieved analytically by making some simple but accurate approximations (see also Bond 1989). There are several motivations for doing so (Sheth, Mo & Tormen 2001; Paranjape & Sheth 2012), not least the fact that N-body simulations of cold dark matter (CDM) show that a large fraction of haloes do, in fact, originate from initial density peaks (Ludlow & Porciani 2011a). Further, Paranjape, Sheth & Desjacques (2013) showed that this unified analytical formalism of excursion set peaks (ESP) gives a self-consistent description of the CDM halo mass function as well as clustering which is accurate at the ∼10 per cent level.

It is worth asking whether this formalism has correctly captured all the relevant aspects of structure formation that affect the mass function. One way of addressing this issue is to apply the same formalism in an ‘extreme’ situation which it was not explicitly built to describe. Structure formation from an initial matter power spectrum with highly suppressed small-scale power, as found in warm dark matter (WDM) cosmologies, offers the perfect playground. The reason is that, apart from having a truncated initial power spectrum, simulations of WDM in fact solve exactly the same problem as those of CDM: the evolution of a cold, collisionless, self-gravitating fluid.

Analytically, one then expects that the same ESP expressions, which correctly describe the CDM mass function and clustering, should work for the WDM case as well, with the simple replacement of the CDM initial power spectrum with that of WDM. In this regard, as we describe in detail below, the ‘out-of-the-box’ ESP calculation does considerably better than traditional top-hat-filtered excursion sets: it correctly predicts a turnover in dn/dln m at the correct scale whereas the latter predicts a monotonic rise at low masses. We will see, however, that ESP predicts a power law decrease at low masses dn/dln m ∼ m^1/3 which is incompatible with the results of simulations. This analytical prediction is very robust and hints at a missing physical ingredient in the excursion set logic.1

Our goal in this paper is to characterize the collapsed objects identified in a WDM simulation in terms of the properties of the initial density field. This will allow us to understand the reasons behind the mismatch of the measured mass function and the ESP prediction. The paper is structured as follows.

In Section 2, we describe the numerical simulation and halo finding algorithm, which are the same as presented by Angulo, Hahn & Abel (2013). We then compare the resulting halo mass function with theoretical expectations based on the ESP formalism, and discuss possible reasons for the differences we see between the theory and numerics. To better understand where haloes form, in Sections 3 and 4 we turn to an in-depth analysis of the initial conditions of the simulation. In Section 3 we analyse the initial density field at the ‘Lagrangian patches’ of the haloes (i.e. the initial locations of groups of particles that will eventually be identified as haloes) and demonstrate that all haloes in the simulation are consistent with forming near peaks of the initial density. We also explore correlations between the initial overdensity and shape of the Lagrangian patches, and use these results to motivate the construction of an empirical catalogue of ‘ESPeaks’, which we describe in Section 4. These ESPeaks are a numerical realization of what the ESP calculation aims to accomplish, and we compare their Lagrangian properties with those of the haloes.

Our main conclusion from this exercise is that, while the ESP calculation on average correctly identifies the locations of halo formation, it systematically underpredicts the mass of the resulting object, and that this effect is especially enhanced at halo masses that are small compared to the characteristic mass scale where the WDM mass function turns around. In Section 5 we argue that this mismatch is related to a systematic overprediction of the time of collapse of a given perturbation, and propose a modification to the ESP calculation to re-assign masses by correcting for this effect. We show that the resulting mass function not only agrees very well with the WDM result, but also describes the CDM mass function accurately with the simple replacement of the WDM power spectrum with that of CDM.

We close with a summary and discussion in Section 6. The appendices collect technical details and arguments used to motivate some of the results in the main text.

2 THE HALO MASS FUNCTION: CONFRONTING SIMULATIONS AND THEORY

Matter power spectra with an initial small-scale truncation arise naturally in warm and hot dark matter cosmologies where density fluctuations on small scales are suppressed due to the late transition to the non-relativistic regime of the respective dark matter particle. Such a power spectrum leads to a corresponding turnover in the late time halo mass function. The numerical determination of such mass functions has, however, proven extremely challenging due to the presence of low-mass objects that arise – completely unphysically – from the fragmentation of filaments (see e.g. Avila-Reese et al. 2001; Bode, Ostriker & Turok 2001; Melott 2007; Wang & White 2007; Hahn, Abel & Kaehler 2013). In the presence of artificial fragmentation, mass functions can only be measured indirectly after filtering or correcting for the spurious haloes (see e.g. Lovell et al. 2012, 2013; Schneider et al. 2012). Only more recently has the behaviour of the halo mass function around and below the turnover scale been explicitly demonstrated by Angulo et al. (2013, hereafter AHA13).

While such WDM cosmologies are of course of genuine physical interest in their own right, we are mainly concerned with a different aspect here: the suppression of low mass haloes provides powerful additional leverage to test models of structure formation in such
cosmologies. The exponential fall in the halo mass function at large masses – whose sensitivity to cosmological parameters has been exploited for decades – is replicated here at the small-mass end. Any theoretical model must now describe both of these strong features in the mass function.

We begin by briefly discussing the numerical simulations of AHA13 and the WDM halo mass function they measure. We will see that these numerical results do not meet theoretical expectations based on the ESP formalism. We discuss possible reasons for this, which will motivate our subsequent analysis.

2.1 Numerical simulation

The numerical simulation discussed by AHA13 employs the novel \( \tau_{4Pm} \) method (Hahn et al. 2013), which completely suppresses artificial fragmentation and allows the determination of the halo mass function at and below the turnover scale in the absence of numerical artefacts.

Specifically, this simulation resolves a \( 80 h^{-1} \) Mpc cosmological volume with 1024\(^3 \) particles with cosmological parameters \( \Omega_{m} = 0.276, \Omega_{\Lambda} = 0.724, \Omega_b = 0.045, h = 0.703, \sigma_8 = 0.811 \) and \( n_s = 0.96 \), consistent with the 7-year Wilkinson Microwave Anisotropy Probe (WMAP7) data release (Komatsu et al. 2011). The normalization of the power spectrum using \( \sigma_8 \) was set using a CDM simulation, so that the amplitude of fluctuations on large scales is independent of the truncation scale of the power spectrum.

The truncation of power at small scales is done by assuming a toy model cosmology with a 0.25 keV thermally produced WDM particle. Such a particle is, of course, completely ruled out by observations as the dominant component of dark matter (e.g. Viel et al. 2013, who derive a current lower bound of 3.3 keV). However, it allows resolving the entire power spectrum up to the truncation scale with sufficient particles and, as we have already argued above, is studied in this paper for the main purpose of testing analytical predictions. In particular, AHA13 used the fitting formula of Bode et al. (2001) to modify the CDM transfer function:

\[
T_{\text{WDM}}(k) = T_{\text{CDM}}(k) \left[ 1 + (\alpha k)^{-5.0} \right],
\]

with

\[
\alpha = 0.05 \left( \frac{\Omega_m}{0.4} \right)^{0.15} \left( \frac{h}{0.65} \right)^{1.3} \left( \frac{m_{\text{dm}}}{1 \text{ keV}} \right)^{-1.15} h^{-1} \text{Mpc},
\]

where \( m_{\text{dm}}(= 0.25 \text{ keV}) \) is the dark matter particle mass. This results in \( \alpha = 0.26 h^{-1} \) Mpc, equivalent to a free-streaming mass-scale:

\[
M_{\text{fs}} = \frac{4\pi}{3} \rho (\alpha/2)^3 \simeq 7 \times 10^8 h^{-1} M_{\odot},
\]

and a ‘half-mode’ mass-scale (cf. e.g. Schneider et al. 2012):

\[
M_{\text{lm}} \simeq 4.3 \times 10^5 M_{\odot} \simeq 3.0 \times 10^{12} h^{-1} M_{\odot}.
\]

Note that, as discussed in more detail in AHA13, these simulations do not include the (small) thermal velocity dispersion that a real WDM fluid would possess, so that the collisionless dark matter fluid is in fact treated in the perfectly cold limit, after perturbations have been suppressed below the maximum free-streaming scale in linear perturbation theory. A thermal velocity however is expected to have little effect on the abundance of collapsed structures at late times, which is the main topic of our interest here.

Adopting the fit of Eisenstein & Hu (1999) as the fiducial CDM transfer function, initial conditions were generated using the MUSIC code (Hahn & Abel 2011) at an initial redshift of \( z = 63 \) using the Zel’dovich approximation. We note that a simulation initialized at such a rather low redshift using first-order Lagrangian Perturbation Theory is to some degree affected by transients from the initial conditions (e.g. Crocce, Pueblas & Scoccimarro 2006). The high-mass end of the halo mass function is thus expected to deviate from the true one. The small volume of the simulation adds further to a systematic deviation. Furthermore, it is possible that the detailed behaviour of the mass function around the half-mode mass is also affected by transients. This possibility needs to be considered for precision determinations of the halo mass function but we do not expect it to alter the qualitative behaviour with which we are mostly concerned here (the results of AHA13 are roughly consistent with e.g. the predictions of Schneider et al. 2013 who use second-order Lagrangian Perturbation Theory). We note that the half-mode mass is resolved with almost 100 000 particles in the simulation of AHA13.

2.2 Halo identification

AHA13 found that the suppression of artificial fragmentation leads to a failure of the Friends-of-Friends (FoF) halo finder. The dense cores of filaments (in the absence of artificial fragmentation) lead to a percolation of large regions of several haloes when the standard linking parameter \( b = 0.2 \) is used. Instead, they first adopted a linking parameter of \( b = 0.05 \) times the mean interparticle separation and then determined the spherical-overdensity (SO) mass centred on the centre-of-mass of the parent FoF group. A halo was defined as the sphere of radius \( R_{200} \), which has a mean density of 200 times the critical density, \( \rho_{\text{crit}} \). This corresponds to a halo mass \( M_{200} = (4\pi/3) R_{200}^3 \rho_{\text{crit}} \).

Further, by analysing all haloes individually, AHA13 found that the halo sample could be divided into various subsamples or ‘types’. ‘Type-1’ objects are virialized haloes, while ‘type-2’ include haloes in late stages of formation; the latter do not show an isotropic density structure and instead contain larger scale caustics that are remnants of their formation. The remaining objects were haloes in early stages of formation that have e.g. just started collapsing along the third axis.

In what follows, we only consider the ‘type-1’ objects clearly identified as haloes. The red histogram in Fig. 1 is identical to the line labelled ‘haloes’ in fig. 7 of AHA13 and shows the mass function of these ‘type-1’ haloes, which has a sharp cut-off between \( 10^{12} \) and \( 10^{12} h^{-1} M_{\odot} \). (The other two histograms will be discussed in Section 4.) It is important to note here that the classification was performed visually and we thus expect that, on an object-by-object basis, the distinction between types 1 and 2 is likely not perfectly robust.

2.3 Theoretical expectations

As discussed earlier, as far as late time structure formation is concerned, the only difference between the CDM and WDM cosmologies is the lack of initial small-scale power in the latter. So one might expect that a physically motivated description of CDM structure should apply equally well to WDM – at least at scales much larger than the free-streaming scale – with the simple replacement of the CDM transfer function with the one in equation (1). One can already anticipate that the standard hierarchical excursion set calculation would have difficulty in describing the low-mass end of the WDM mass function where the effective logarithmic slope of the power spectrum \( n_{\text{eff}} \equiv d \ln P(k)/d \ln k \) becomes steeper than \(-3\). This is the boundary beyond which hierarchical prescriptions are
known to fail, with predicted mass accretion rates becoming ill-defined (Lacey & Cole 1993, 1994). The ESP formalism, however, introduces a new ingredient into the picture – the peaks constraint – and since we know that it works well for CDM, we can ask what it predicts for the WDM case.

In what follows, we will frequently use integrals over the power spectrum of the filtered initial overdensity field $\delta_k$ and its spatial derivatives, all linearly extrapolated to the present epoch:

$$\sigma^2(R) \equiv \int d \ln k \Delta(k) k^2 \tilde{W}_g(k)^2,$$

where $\Delta(k) \equiv k^3 P(k, z = 0)/(2\pi^2)$ is the dimensionless matter power spectrum in linear theory and $\tilde{W}_g(k)$ is the Fourier transform of the smoothing filter, for which we will use a spherical top-hat $\tilde{W}_g(k) = (3(\sin kr - kr \cos kr))/(kr)^3$ in our numerical analysis and later also a Gaussian $\tilde{W}_g(k) = \exp(-k^2 R^2/2)$ in our analytical modelling. The above definitions correspond to setting $\sigma^2_0(R) = \langle \delta_k^2 \rangle$, $\sigma^2_1(R) = \langle (\nabla \delta_k)^2 \rangle$ and $\sigma^2_2(R) = \langle (\nabla^2 \delta_k)^2 \rangle$ and appear in peaks formalism calculations. Another quantity, which is relevant for excursion set models of the mass function, is the derivative of the smoothed density field with respect to smoothing scale, $d\delta_k/dR$, which is in general different from the spatial derivatives of $\delta_k$. A special case is that of Gaussian filtering, for which $d\delta_k/dR = R \nabla^2 \delta_k$, a result which will be useful later in our analytical modelling. For top-hat filtering, $d\delta_k/dR$ and $\nabla^2 \delta_k$, although different, are strongly correlated (Paranjape et al. 2013).

To get an idea about the scale of the problem, consider that simply counting all peaks in the unsmoothed initial density field of the WDM simulation gives us 6713 objects, where ‘unsmoothed’ refers to the density on a 512$^3$ grid and a grid cell is labelled a peak if its density is higher than all its 26 neighbours (see Section 3 for more details). This grid size just about resolves the cut-off scale $\alpha$ (equation 2) below which no initial fluctuations exist; we have verified that a 1024$^3$ grid (the resolution at which the initial conditions of AHA13 were generated) leads to a consistent result (6822 peaks). This matches very well with the theoretical prediction for this number in the simulation volume $V_{\text{box}}$ (equation 4.11b of Bardeen et al. 1986, hereafter BBKS):

$$N_{pk} = n_{pk} V_{\text{box}} = 3.12 \times 10^{-3}(\sigma_2/\sigma_1)^3 V_{\text{box}} = 6608,$$

in which we evaluated $\sigma_1$ and $\sigma_2$ using equation (5) with $R = 0$ and the transfer function (1). Comparing this with the number of ‘type-1’ objects in the simulation – 1522 – we clearly see that not every individual peak forms a halo. This is fully expected within the analytical framework – e.g. there is nothing special about a peak of height $\delta \equiv -10\sigma_0$ and we will show later that the ESP calculation does lead to a number close to the measured number of haloes.

The ESP halo mass function can be written as (Paranjape & Sheth 2012; Paranjape et al. 2013)

$$dN_{\text{ESP}}/d\ln m = N_{\text{ESP}} \, d\nu/d\ln m,$$

where $\nu \equiv \delta_k(z)/\sigma_0(m)$ with $\delta_k(z)$ being the critical linear overdensity or ‘barrier’ for spherical collapse in a $\Lambda$CDM background, and where $N_{\text{ESP}}$ has the structure

$$N_{\text{ESP}} \sim 1/V_{c} \g_{\text{ESP}}(\nu, \gamma).$$

Here $g_{\text{ESP}}(\nu, \gamma)$ is a dimensionless function of its arguments (details in Section 5), and $\nu$ and $V_c$ are spectral quantities that define the distribution of peaks (BBKS):

$$\gamma \equiv \sigma_1^2/\langle \sigma_0 \sigma_2 \rangle; \quad V_c \equiv (6\pi)^{1/2} \langle \sigma_1/\sigma_2 \rangle^3,$$

where $\gamma$ is a dimensionless measure of the width of the power spectrum, $V_c$ sets the mean number density of all peaks on scale $R$ (equation 4.11b of BBKS).

It is easy to see that the spectral integrals $\sigma_1$ and, consequently also $\nu$, $\gamma$ and $V_c$, will each approach a constant value for WDM as $m \to 0$. The behaviour of $\nu$ and $V_c$ in this respect is very different from that in CDM in the same limit, where $\nu$, $V_c \to 0$ as $m \to 0$. This reflects the fact that peaks can only form on scales large enough to be inhomogeneous; reducing the smoothing scale cannot wipe out existing peaks for any power spectrum and, for a truncated spectrum, cannot introduce new peaks. This can be seen in the top panel of Fig. 2: for each choice of $m_{\text{min}}$, $V_c$ at large $m$ is close

\footnote{All these integrals remain finite at all scales, including the unsmoothed limit $R \to 0$, since the WDM free-streaming scale itself acts as a smoothing filter. In contrast, ultraviolet power in CDM causes $\sigma_0(R)$ and $\sigma_1(R)$ to diverge as $R \to 0$, while the top-hat smoothed $\sigma_2(R)$ always diverges, meaning that any analysis of small-scale CDM peaks would be limited by effects at the spatial resolution limit of the simulation.}

\footnote{The redshift dependence of $\delta_k(z)$ in a flat $\Lambda$CDM universe is slightly different from that in an Einstein–deSitter (EdS) background (see e.g. Eke, Cole & Frenk 1996), and can be approximated by $\delta_k(z) D(z)/D(0) = \delta_{k, \text{EdS}}(1 - 0.0123 \log_{10}(1 + x^2))$, where $x \equiv (\Omega_m - 1)^{1/2}/(1 + z)$ and $\delta_{k, \text{EdS}} = 1.686$ (Henry 2000). In our case, requiring collapse at present epoch gives $\delta_k(z = 0) = 1.674$.}
to the CDM value (and hence approximately proportional to $V = m/\bar{\rho}$, where $\bar{\rho}$ is the comoving mean density), but deviates as $m$ approaches the half-mode mass $M_{\text{hm}}$, finally approaching a constant value at $m \ll M_{\text{hm}}$, this value being close to $M_{\text{hm}}/\bar{\rho}$. This last aspect gives us an interesting physical interpretation of the half-mode mass as being essentially the same as the asymptotic peaks scale.

Finally, the Jacobian appearing in equation (6) behaves, as $m \to 0$ for WDM, like

$$\left| \frac{d\nu}{d \ln m} \right| = \frac{\nu^3 R^3}{3\delta^2(z)^2} \int d \ln k \Delta(k) W(k R) \left| \frac{dW}{d R} \right| \to \frac{R^2 \nu^3}{\#\delta(z)^2} \int d \ln k \Delta(k) k^2 \sim R^2 \sim m^{1/3},$$

where $\# = 15$ for a top-hat filter and 3 for a Gaussian filter, so that the ESP mass function (equation 6) for WDM at small masses behaves like

$$dn_{\text{ESP}}/d \ln m \sim |d\nu/d \ln m| \sim m^{2/3}.$$  \hspace{1cm} (10)

This power-law behaviour at small masses will be true for any single-barrier excursion set model of peaks in which the barrier depends on halo mass only through the spectral integrals; in particular, this behaviour is independent of details such as barrier shape, stochasticity, etc.

The behaviour of $V_\nu$ discussed above shows that the turnover occurs at around $m \sim M_{\text{hm}}$. The solid curve in Fig. 1 shows the ESP calculation of Paranjape et al. (2013) using the WDM transfer function (1). The dashed curve shows the ESP result with a somewhat different, convenient choice of barrier for the random walks, which we will discuss in detail in Section 5.2. The main point is that both of these curves show the turnover and the asymptotic scaling of $\sim m^{2/3}$. As discussed above, the latter is a very robust prediction that cannot be easily altered by technical modifications within the framework of a single barrier.

It is interesting to contrast this result with the corresponding one for the traditional excursion set approach, as this emphasizes the key role played by the behaviour of $V_\nu$. For traditional excursion sets (Bond et al. 1991; Musso & Sheth 2012), one has

$$dn_{\text{trad}}/d \ln m = (\bar{\rho}/m) f(\nu) |d\nu/d \ln m|,$$  \hspace{1cm} (11)

where $f(\nu)$ is a dimensionless function analogous to $g_{\text{ESP}}(\nu, \gamma)$ discussed earlier. More importantly, $V_\nu$ is replaced by the Lagrangian volume $V = m/\bar{\rho}$ of the halo, so that for WDM at small masses, this mass function behaves like

$$dn_{\text{trad}}/d \ln m \sim m^{-1} |d\nu/d \ln m| \sim m^{-1/3}.$$  \hspace{1cm} (12)

This asymptotic behaviour is unphysical because the hierarchical excursion set calculation should not predict objects at small masses where no hierarchical formation can occur in the absence of small-scale power (see the discussion above, equation 5). The dotted curve in Fig. 1 shows the result of using equation (1) to compute the relation $\sigma_0(m)$ in the CDM fit provided by Tinker et al. (2008). For completeness, the bottom panel of Fig. 2 compares the $\sigma_0(m)$ relation for top-hat filtering when using equation (1) with the corresponding relation for CDM. For WDM we clearly see a ‘freezing-out’ of $\sigma_0(m)$ at small masses. The other spectral integrals also show similar behaviour.

2.4 Theory versus simulation – what could be going wrong?

Although the physical requirement of being a density peak naturally accounts for a turnover in the mass function at the correct mass scale, the asymptotic scaling (very robustly) predicted by ESP is clearly wrong. There are several issues which could in principle affect this result.

**Dynamics.** A dramatic possibility is that, since small-mass haloes in WDM do not form hierarchically (e.g. at some point in time the first object forms, with no virialized progenitor), the peaks calculation might simply not be applicable. This would lead to the interesting question of just what it is that characterizes the locations and dynamics of the collapse of small-mass objects. The cut-off scale in the initial spectrum could in principle allow for higher order catastrophes (cf. Arnold, Shardarin & Zeldovich 1982) to become relevant, and these need not necessarily appear as peaks when filtered on the protohalo scale. In CDM, the situation is quite different in this respect, since fluctuations persist down to very small scales and so every protohalo has a progenitor at a smaller scale.

**Barrier shape.** It has been argued that the collapse barrier appropriate for WDM haloes is very different from the corresponding CDM one due to thermal effects in WDM, and that this can introduce a sharp cut-off in the mass function (Benson et al. 2013). However, since WDM simulations see a cut-off despite ignoring thermal effects (e.g. AHA13; Schneider et al. 2013), the origin of
the cut-off must be rooted in the suppression of initial small-scale power, and must therefore be a generic feature of cold collisionless dynamics in such conditions. The failure of excursion set (peaks) models to reproduce the correct mass function hence indicates quite clearly that these models are still not accounting for some important physical processes. One of the primary goals of this work is to investigate the cause of this behaviour.

Patch shape. Traditional excursion sets, as well as the ESP calculation of Paranjape et al. (2013), use spherical filters when assigning masses to objects, and it could be that asphericity of the Lagrangian patches affects the mass assignment significantly at small masses. For example, recently Despali, Tormen & Sheth (2013) have demonstrated using CDM simulations that accounting for halo asphericity using an ellipsoidal halo finder can lead to small increases in mass for low-mass haloes (see also Ludlow & Porciani 2011b).

Stochasticity. Regardless of the importance of thermal effects, the specific details of the barrier, e.g. those related to stochasticity in the barrier height, are in fact somewhat uncertain (even in the CDM case). The ESP calculation for CDM is self-consistent but not fully predictive, and needs some inputs from simulations (Paranjape et al. 2013). In particular, the barrier used in that calculation was adjusted to match measurements by Robertson et al. (2009) of protohalo overdensity in CDM simulations, and the same results might not apply in the case of WDM.

Peak-in-peak. Another possible source of error is that the ESP framework treats the peak-in-peak problem approximately, by introducing the effects of the peaks constraint as an extra weight in the mass function, rather than by explicitly accounting for spatial correlations between walks centred at different locations in space (see e.g. Scannapieco & Barkana 2002), and this approximation needs testing.

We address these issues in the next two sections by exploring the properties of the initial conditions of the simulation in greater detail.4

3 LAGRANGIAN PROPERTIES OF HALOES

In this section, we turn to the initial conditions of the simulation and perform an in-depth study of the Lagrangian properties of regions that will eventually form haloes; we call such regions protohaloes and give a precise definition below. Several authors have performed such studies in CDM simulations (e.g. Bond & Myers 1996; White 1996; Sheth et al. 2001; Porciani, Dekel & Hoffman 2002; Robertson et al. 2009; Ludlow & Porciani 2011a, b; Elia, Ludlow & Porciani 2012; Despali et al. 2013). To our knowledge, the current work is the first to extend these studies to the case of WDM, and is interesting for the reasons discussed in Section 2.

An advantage of using a WDM model with m_th = 0.25 keV is that the number of objects is reasonably small. A disadvantage is that the half-mode mass is close to being unit significance, v(M_th, z = 0) ≈ 0.9, which does not allow us to explore low-significance objects with sufficient statistical precision. This could also potentially confuse non-linear assembly-bias-like effects with the peculiarities of halo formation at and below the half-mode mass scale. Nevertheless, this simulation provides us with an invaluable testing ground for several ideas in the peaks framework.

We focus on the overdensity of the protohalo patch (which is indicative of the collapse threshold), its curvature, velocity shear (ellipticity and prolateness) and moment of inertia. We will demonstrate two important features of the protohalo patches: (a) that they are all consistent with forming at initial density peaks and (b) their overdensities are strongly correlated with their ellipticities but not their prolateness.

3.1 Lagrangian density and shear fields

The initial conditions code MUSIC allows us to output the density field that was used to generate the simulation initial conditions as three-dimensional grid data. We used this function to re-generate the density field directly on a 5123 mesh with the same Fourier modes as the original simulation. We refer to this as the unsmoothed field. Using the particle IDs that encode the three-dimensional Lagrangian coordinate q on the unperturbed initial particle lattice (i.e. before applying the Zel’dovich approximation), we can directly evaluate the density at q without interpolating the perturbed particle position back on a grid. We linearly scale the density field to z = 0.

Using the unsmoothed density field on a mesh, we compute various derived fields using the fast Fourier transform (FFT). We compute the gradient and the Hessian of the density field:

$$\nabla \delta = \mathcal{F}^{-1}\left\{i k \tilde{\delta}(k)\right\}, \quad \partial_{ij} \delta = \mathcal{F}^{-1}\left\{-k^j k_i \tilde{\delta}(k)\right\},$$

where the tilde indicates the Fourier transformed field. Additionally, we compute the velocity potential as well as its Hessian (the so-called tidal tensor which reflects the velocity shear),

$$\psi = \mathcal{F}^{-1}\left\{-k^2 \tilde{\delta}(k)\right\}, \quad \partial_{ij} \psi = \mathcal{F}^{-1}\left\{\frac{k_i k_j}{k^2} \tilde{\delta}(k)\right\},$$

so that the velocity field is $u \propto -\nabla \psi$. When computing filtered fields, we replace $\tilde{\delta}(k)$ with $\tilde{\psi}(k) = \tilde{\psi}(k) \tilde{W}_k(k)$.

We define the ordered eigenvalues of $\partial_{ij} \psi$ as $\lambda_1 \leq \lambda_2 \leq \lambda_3$ and those of the velocity shear $\partial_{ij} u$ as $\lambda_1 \leq \lambda_2 \leq \lambda_3$. The normalized negative trace of the density Hessian gives us the dimensionless peak curvature,

$$x \equiv -(\lambda_1 + \lambda_2 + \lambda_3) / \sigma_2,$$

while the trace of the velocity shear gives back the density,

$$\delta = \text{Tr} \partial_{ij} \psi = \lambda_1 + \lambda_2 + \lambda_3 \propto -\text{div} \, u.$$

Peaks in $\delta$ are thus equivalent to regions of maximum convergence in the Lagrangian flow. We will also need the ellipticity $e_e$ and prolateness $p_e$ associated with the tidal tensor:

$$e_e \equiv (\lambda_3 - \lambda_1) / 2 \equiv Y,$$

$$p_e \equiv (\lambda_3 - 2\lambda_2 + \lambda_1) / 2 \equiv Z,$$

where we have defined $Y$ and $Z$ to be the corresponding un-normalized quantities which will be useful below. Similarly, we can define

$$e_{pk} x \equiv -(\lambda_1 - \lambda_3) / (2\sigma_2) \equiv y,$$

$$p_{pk} x \equiv -(\lambda_1 - 2\lambda_2 + \lambda_3) / (2\sigma_2) \equiv z,$$

so that $e_{pk}$ and $p_{pk}$ describe the shape of the peak (BBKS).
3.2 Protohaloes and their properties

For each halo, we recorded the particle IDs and recovered their respective Lagrange coordinates $q$. We call the set of $N_l$ Lagrangian particles comprising halo $k$ its Lagrangian patch or protohalo, denoted $L_k \equiv \{ q_i \mid i = 1 \ldots N_l \}$.

We compute the patch average of a Lagrangian field $f(q)$ by evaluating

$$
\langle f \rangle_k = \frac{1}{N_l} \sum_{q_i \in L_k} f(q_i),
$$

and the spherical average by first determining the Lagrange radius $R_l = (3m/4\pi \bar{\rho})^{1/3}$, where $m$ is the halo mass, and then evaluating

$$
\langle f \rangle_k = \frac{1}{V_l} \int_{q_\text{med}} f(q) dq,
$$

Here $W_{Rh}$ is the top-hat filter at scale $R_l$ and $q_\text{med}$ is the median Lagrange coordinate of the Lagrangian patch, where the median is taken of each separate Cartesian component. Using the median instead of the mean coordinate reduces the influence of outliers in the Lagrangian patch.

Fig. 3 shows three examples of protohaloes with masses $\sim 10^{13} h^{-1} M_\odot$. The top panel shows a well behaved protohalo. We notice two disconnected shells surrounding the connected interior of this patch. This is a beautiful example of the mapping between Lagrangian and Eulerian space. The gaps between the shells appear because the outer caustics of the halo are not inside the virial radius and are thus cut off. The two shells correspond to material on first and second infall. The other two examples show evidence of mixing due to large-scale interactions.

In addition to the ellipticity and prolateness associated with the density Hessian, we can also characterize the shape of the Lagrangian patch $L_k$ through the dimensionless reduced moment of inertia tensor:

$$
I_{ij} = \sum_{q_i \in L_k} (q_i^2 \delta_{ij} - q_i q_j) / q^2,
$$

which we define to be centred on the centre-of-mass of the object (rather than its median location), since this minimizes its values.

The eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$ of $I_{ij}$ give the corresponding axes of the homogeneous ellipsoid $a \leq b \leq c$:

$$
a = \sqrt{2N_k/(\lambda_1 + \lambda_2 - \lambda_3)},$$
$$
b = \sqrt{2N_k/(\lambda_2 - \lambda_1 + \lambda_3)},$$
$$
c = \sqrt{2N_k/(\lambda_3 + \lambda_1 - \lambda_2)},
$$

and the sphericity

$$
S \equiv a/c.
$$

3.3 Haloes form at peaks

We start by verifying statistically that haloes in cosmologies with truncated small-scale power do indeed form from peaks. Being a peak requires the overdensity field $\delta$ to be locally extremal on the scale of the protohalo, i.e.

$$
\langle \nabla^2 \delta \rangle = 0 \quad \text{and} \quad \langle \zeta_i \rangle < 0, \quad i = 1, 2, 3.
$$

We find that all protohalo patches have $\langle \zeta_i \rangle^{(3)} < 0$ and $\langle \zeta_i \rangle^{(9)} < 0$ and thus the total peak curvature $x$ is positive in both cases. Note that to define averaged eigenvalues of a tensor, we diagonalize after computing the average of the tensor.

In Fig. 4, we show the distribution of the total curvature $x$ (equation 15; top panel) and the magnitude of the gradient $\eta \equiv \sqrt{3} |\nabla \delta|/\sigma_1$ (bottom panel) averaged over the halo patches. For computing $\sigma_1$ and $\sigma_2$ which make these quantities dimensionless, we used a top-hat filter at the Lagrangian scale $R_l$ of each object.

The distribution of $x$ for a Gaussian random field would be a Gaussian with zero mean and unit variance (dotted black curve in the top panel). The measured distribution on the other hand has only positive values as mentioned above, and its shape is very similar to the analytical prediction using ESP with a deterministic barrier (dashed black curve, see Section 5, equation 36), although the measured mean value for $x$ is lower than the predicted mean by about 0.4.

The distribution of $\eta$ for a Gaussian random field would be $p(\eta) = \sqrt{2/\pi \eta} e^{-\eta^2/2}$ (because in this case $\eta^2$ is chi-squared distributed with 3 degrees of freedom.) This is shown as the dotted black curve in the bottom panel; the measured values clearly
The formation of haloes in WDM, which does not proceed hierarchically below the mass scale of interest, has been the subject of various studies. We have compared our results with those of Tinker et al. (2008), who found similar results in their CDM simulations. There is no clear reason for the discrepancy.

Figure 4. (Top panel: the distribution of peak curvatures \( x \) (equation 15) at protohalo locations, averaged over the Lagrangian patch (blue histogram) and using a spherical aperture of the same mass (green histogram). The dashed curve shows the theoretically expected distribution (equation 36), while the dotted curve shows the distribution at random locations (a Gaussian with zero mean and unit variance). Bottom panel: the distribution of density gradient \( \eta \equiv \sqrt{3} |\nabla \delta| / \sigma_1 \) at protohalo locations, averaged over the Lagrangian patch (blue histogram) and using a spherical aperture of the same mass (green histogram). The dotted line is the distribution at random locations; in this case \( \eta \) is distributed as chi-squared with 3 degrees of freedom. These plots show that all haloes in our sample are consistent with having formed near initial density peaks.

populate the low tail of this distribution. (Ideally all the values would be zero.) We also see that the patch-averaged values of \( \eta \) have a significantly lower scatter than the spherically averaged ones. This is not surprising since the requirement \( \eta = 0 \) is quite unstable to choices of filtering, and the spherical filter is known to introduce an additional randomization as compared with the actual Lagrangian patch (BBKS; Despali et al. 2013).

We therefore conclude that all haloes in our sample are consistent with having formed near initial density peaks. In principle, we should also have explicitly checked for the presence of local density maxima at or near the protohalo locations, e.g. along the lines discussed by Ludlow & Porciani (2011a). This, however, would involve making a specific choice regarding the smoothing scale. We defer such a calculation to Section 4.1, where we implement an algorithm that makes this choice while simultaneously centring the smoothing filter at locations that are most likely to collapse according to the excursion set formalism.

3.4 Overdensity of Lagrangian patches

Fig. 5 shows the patch-averaged (left-hand panel) and spherically averaged (middle panel) overdensities of the protohaloes as a function of their mass. We find that the spherical overdensities are strongly correlated with the corresponding spherically averaged values of \( Y \) (equation 17). This is evident in the middle panel where we have coloured the points using \( \langle Y \rangle^{(s)} \). (The patch-averaged overdensities show a similar strong correlation with the patch-averaged \( Y \); we omitted the colouring in the left-hand panel for clarity.)

The right-hand panel of the figure shows the difference \( \langle \delta \rangle^{(s)} - \langle Y \rangle^{(s)} \) coloured by prolateness \( \langle Z \rangle^{(s)} \) (equation 18). We see that the scatter in this difference is significantly smaller than that in \( \langle \delta \rangle^{(s)} \), and its distribution is curiously similar to that of \( \langle \delta \rangle^{(p)} \) with a mean close to the standard spherical collapse value \( \delta_c \) (horizontal dashed line in all panels). More importantly, we have found no correlation of \( \langle \delta \rangle^{(s)} \) with prolateness \( \langle Z \rangle^{(s)} \). In other words, the spherical overdensities of the protohaloes are well approximated by the relation

\[
\langle \delta \rangle^s = \delta_c + \langle Y \rangle^s, \tag{27}
\]

with a residual scatter that does not correlate with prolateness. This is consistent with the CDM results of Ludlow & Porciani (2011b). We have also checked that the overdensity does not correlate with the other shape parameters \( y \) and \( z \) defined in equations (19) and (20).

Robertson et al. (2009) performed similar spherically averaged measurements in the initial conditions of the CDM simulations presented by Tinker et al. (2008). The left-hand panel of Fig. 5 shows the mean and standard deviation of the distribution of overdensities reported by Robertson et al. (2009), but using the WDM transfer function. We see that their spherically averaged measurements, extrapolated to WDM, are in reasonable agreement with our patch-averaged overdensities. Our spherical overdensities, on the other hand, have a higher mean and scatter than theirs (see also Elia et al. 2012, who found similar results in their CDM simulations). There is no clear reason for this discrepancy.

These results for the spherically averaged overdensity and shear ellipticity will form the basis of the empirical walks that we describe in the next section. In general, however, we note that our measurements of spherically averaged quantities tend to have larger scatter than the corresponding patch-averaged ones. For completeness, in Appendix C we also show the distributions of ellipticity and prolateness defined using the tidal tensor and the Hessian of the density.

In summary, the results of this section show us that haloes form at peaks and have Lagrangian (spherically averaged) overdensities that are consistent with equation (27), with a residual scatter that is uncorrelated with other properties such as shear prolateness or peak shapes. One aspect we have not explored here is the relative (mis)alignment between the velocity shear, density Hessian and moment of inertia tensors, which can be an important ingredient in any recipe for predicting collapse time based on dynamical arguments. This is especially interesting given previous results from CDM simulations (Porciani et al. 2002; Despali et al. 2013) which suggest that, contrary to expectations based on Gaussian statistics (e.g. van de Weygaert & Bertschinger 1996), the direction of maximum initial compression is on average well aligned with the longest geometrical axis of the protohalo, rather than the shortest. We will return to an analysis of tensor alignments and their dynamical consequences in future work.

4 EMPIRICAL EXCURSION SET PEAK WALKS

The most serious issue raised in Section 2.4 was whether or not the excursion set formalism can capture at all the formation of haloes in WDM, which does not proceed hierarchically below the
Figure 5. Patch overdensity as a function of protohalo mass. Left panel: the overdensity \( \delta \) (equation 16) at protohalo locations, averaged over the Lagrangian patch (points). The thick blue line and the blue band show, respectively, the mean and standard deviation of the spherically averaged protohalo overdensities measured by Robertson et al. (2009) in CDM simulations, extrapolated to our WDM initial power spectrum (i.e. we used our WDM transfer function in the fit in their equation D2 with the \( \Delta = 200 \) parameters from their table 1). Middle panel: protohalo overdensity \( \delta \) averaged over spherical apertures of the same mass, coloured by the spherically averaged ellipticity \( Y = c_\text{ec} \delta \) (equation 17). We see a strong trend of \( \delta \) with \( Y \), as expected from ellipsoidal collapse arguments (Sheth et al. 2001). Right panel: the quantity \( \delta - Y \), coloured by the spherically averaged prolateness \( Z = p_\text{ec} \delta \) (equation 18). The residual scatter shows no trend with \( Z \), which is at odds with the predictions of ellipsoidal collapse models.

Figure 4. \( \delta \) of actual Lagrangian patch, \( \delta \) of spherical patch, and \( \delta - Y \) of spherical patch as a function of the protohalo mass. The thick blue line and the blue band show, respectively, the mean and standard deviation of the spherically averaged protohalo overdensities measured by Robertson et al. (2009) in CDM simulations, extrapolated to our WDM initial power spectrum (i.e. we used our WDM transfer function in the fit in their equation D2 with the \( \Delta = 200 \) parameters from their table 1). Middle panel: protohalo overdensity \( \delta \) averaged over spherical apertures of the same mass, coloured by the spherically averaged ellipticity \( Y = c_\text{ec} \delta \) (equation 17). We see a strong trend of \( \delta \) with \( Y \), as expected from ellipsoidal collapse arguments (Sheth et al. 2001). Right panel: the quantity \( \delta - Y \), coloured by the spherically averaged prolateness \( Z = p_\text{ec} \delta \) (equation 18). The residual scatter shows no trend with \( Z \), which is at odds with the predictions of ellipsoidal collapse models.

Our algorithm is essentially a more accurate version of what the original peak-patch algorithm implemented by Bond & Myers (1996), since we are not interested in the initial locations, profiles and velocity dispersions of the haloes, but only in their mass.

We consider a hierarchy of \( N_i = 100 \) smoothing scales \( R_i \), logarithmically spaced in the top-hat spherical mass contained in \( R_i \), between \( M_0 = 10^{10} h^{-1} M_\odot \) and \( M_N = 10^{15} h^{-1} M_\odot \). We then proceed as follows, starting with the largest smoothing scale.

(i) We determine the coordinates \( x_i \) of all peaks in \( \delta_{R_i} \). Being a peak requires that \( \delta_{R_i} \) is larger at \( x_i \) than in all 26 surrounding cells.

(ii) We discard all peaks for which \( \delta_{R_i} \) is below the barrier, i.e. where \( \delta_{R_i}(x_i) < B(x_i; R_i) \); \( B \) is given by equation (28).

(iii) Additionally, we discard all peaks that are within the Lagrangian radius of a peak that has been identified before. This explicitly solves the cloud-in-cloud problem.

(iv) Finally, we also discard all peaks where the density was above threshold on a larger scale, i.e. \( \delta_{R_{i+1}}(x_i) > B(x_i; R_i) \).

This step improves numerical stability but is otherwise redundant.

(v) We proceed to the next smaller scale \( i \to i - 1 \) and start over at step 1.

A small fraction of objects have partially overlapping Lagrangian volumes. We flag the smaller of such pairs as ‘subpeaks’ and, for the current analysis, do not include them in the sample of protohaloes. (These form about 10 per cent of the total sample.) At the end, we arrive at a catalogue of ESPeaks whose Lagrangian properties we can analyse in exactly the same way as for the actual protohalo patches.

4.2 The mass function of ESPeaks

The solid orange histogram in Fig. 1 corresponds to the mass function of ESPeaks obtained using the algorithm described above. The dashed blue histogram shows the result of the same algorithm, but now using a deterministic barrier equation (33) which, as we argue in the next section, is a useful approximation to equation (28). Indeed, we see that these two histograms agree quite well, indicating that stochasticity in the barrier arising from statistical fluctuations in the initial conditions does not lead to dramatic effects in the mass function.

The most important feature of the orange histograms is that they show a low-mass tail consistent with \( d\nu/d\ln m \propto m^{-2/3} \) as discussed earlier. In fact, the histograms are well described by the WDM version of the ESP calculation (solid black) of Paranjape et al. (2013) who used a stochastic barrier adjusted to match CDM simulations, as well as a similar ESP calculation with the deterministic barrier (equation 33) which we discuss below.
The overall number of ESPeaks identified by our algorithm (1261 when using equation 28 and 1335 when using equation 33) is reasonably close to the total number of protohaloes, which is 1522. The lower numbers of ESPeaks could partially be because we stop our algorithm at the lower mass limit of $10^{10} h^{-1} M_\odot$. For comparison, integrating the ESP prediction using equation (33) (dashed line in Fig. 1) above the free-streaming scale $M_{fs}$ gives a prediction of 1380 objects in the simulation volume.

4.3 Matching ESPeaks and haloes

If the excursion set picture is valid, the ESPeaks we identify should be correlated with the actual protohaloes. Could the mass function mismatch simply be because our low-mass ESPeaks are not associated with protohaloes? To assess this, we match the protohalo catalogue and the ESPeaks catalogue as follows.

For every protohalo, we find the ESPeaks contained inside of a sphere of its Lagrangian radius, and associate the ESPeak of highest mass with the halo. This procedure matches 978 (i.e. 64 per cent) of the protohaloes to ESPeaks. We then repeat the same procedure matching ESPeaks to haloes using the filter radius on which the ESPeak was identified, and in this case we can match 935 (i.e. 74 per cent) of the ESPeaks to protohaloes. (Including the subpeaks in the analysis makes these numbers 68 and 72 per cent, respectively.) We discuss possible reasons for the relatively large fraction of mismatched objects below (see also Ludlow & Porciani 2011a). In Appendix C, we also discuss the effect of repeating the exercise with the ellipsoidal collapse barrier (equation 29).

As a visual example, we note that the well-behaved protohalo in the top panel in Fig. 3 was assigned a matching ESPeak while the other two distorted objects were not. While the majority of the protohaloes we can match to ESPeaks look like the object in the top panel and the majority of unmatched protohaloes are distorted, there are also a number of examples of well-behaved protohaloes that are not matched, as well as distorted protohaloes that are.

The top panel of Fig. 6 shows the masses of the protohaloes that we could match to ESPeaks compared to the corresponding ESPeak masses. The points are coloured by the spherically averaged protohalo sphericity $S$ (equation 25). There is a strong trend of $S$ with halo mass: low-mass haloes are decidedly aspherical. This is consistent with the CDM results of Ludlow & Porciani (2011b). Additionally, the scatter in mass assignment also correlates strongly with $S$, with low-mass, aspherical haloes having the largest scatter. However, at a given halo mass, the mass mismatch on average does not seem to correlate strongly with halo shape.

The histograms in the bottom panel of the figure show the mass function of matched (orange) and unmatched (blue) protohaloes, with the grey histogram showing the total halo mass function (same as the red histogram in Fig. 1). The matched fraction is quite large at the highest masses (reaching 100 per cent for $m \gtrsim 3 \times 10^{13} h^{-1} M_\odot$), remains approximately constant at intermediate masses $m \sim M_{fs}$ and falls significantly at low masses $m \lesssim 10^{12} h^{-1} M_\odot$.

In the top panel of Fig. 7, we show the ratio of ESPeak mass to protohalo mass as a function of ESPeak mass, for ESPeaks that we could match to haloes, coloured in this case by the spherically averaged protohalo curvature $x$. We see a weak trend of curvature with mass mismatch: ESPeaks matching shallower protohaloes appear to have larger mass mismatches.

The histograms in the bottom panel of the figure show the mass function of matched (orange) and unmatched (blue) ESPeaks, while the grey histogram is the same as in Fig. 6 and shows the mass function of all haloes. We clearly see that most of the unmatched ESPeaks were assigned dramatically lower masses than any protohalo in the sample. This could indicate that the ESP picture is, in fact, not appropriate for these objects. However, the presence of a small fraction of low-mass ESPeaks that do have matching protohaloes, which in turn have larger true masses and low curvatures, suggests that the explanation of this trend could be more subtle. We therefore explore the properties of the mismatched objects in more detail in the next subsection. Recall that the sum of the mass functions of matched and unmatched ESPeaks is consistent with analytical ESP predictions (compare the smooth black curves and solid orange histogram in Fig. 1).

4.4 Failures of the protohalo ↔ ESPeak matching

We have already seen above that the masses of the ESPeaks that cannot be matched to protohaloes are too low compared to the mass function of haloes. As illustrated in Fig. 8, these unmatched protohaloes tend to have significantly lower peak curvatures and higher densities (solid blue curves in the top and bottom panels,
respective) than those that could be matched to ESPeaks (solid orange curves in the respective panels). Similar trends are also seen, albeit to a lesser extent, in the densities and curvatures of ESPeaks that (do not) have matching protohaloes, as seen in the dashed orange (blue) curves in the two panels. The inset in the top panel of the figure shows the joint distribution of curvature and mass for the matched (orange) and mismatched (blue) protohaloes. The absence of a significant mass dependence of $\chi$ indicates that the difference in peak curvature is not a result of the two populations occupying different mass ranges.

Fig. 9 shows the distribution of mass and sphericity $S$ for the matched (orange) and unmatched (blue) protohaloes. There is a strong trend of sphericity with mass, as noted in Fig. 6. Apart from this, however, there is no significant difference between the sphericities of matched and unmatched protohaloes. The dashed line, which is somewhat shallower than the measured mass trend, shows the linear fit to corresponding measurements in CDM simulations presented by Ludlow & Porciani (2011b).
4.5 Discussion

Our comparison of the outcomes of simulation and the empirical walks suggests that, at least at masses $m \gtrsim M_{\text{min}}$, the empirical approach more or less correctly predicts both the locations and masses of collapsed haloes. At smaller masses, however, the algorithm is able to predict the location of a collapsed object only for a relatively small fraction ($\sim 55$ percent of protohaloes below $10^{12} h^{-1} M_\odot$ have an associated ESPeak, and the fraction falls below $50$ percent quickly below this mass scale) and almost always predicts too small a mass in these cases. The large fraction of ESPeaks that cannot be matched to protohaloes are also predicted to have dramatically lower masses than any protohalo in the simulation.

We have also seen that, on average, the mass and position mismatches seem to be uncorrelated with the shapes of the protohaloes. That is to say, although smaller protohaloes are decidedly aspherical (with a scatter in mass mismatch that correlates with sphericity), there seems to be no trend between the average sphericity and the ratio $m_{\text{ESPeak}}/m_{\text{ halo}}$ in the matched cases (Fig. 6) and no significant difference between sphericities of matched and unmatched protohaloes with $m_{\text{ halo}} < M_{\text{min}}$ (Fig. 9).

The unmatched protohaloes do have smaller curvatures than the matched ones. In principle, this could simply be because of their lower masses. However, the absence of a significant mass dependence of $x$ in the inset in the top panel of Fig. 8 indicates that this is not the case. Additionally, in the case of ESPeaks matched to protohaloes, the protohalo curvature correlates with the mass mismatch (Fig. 7).

Note that the objects identified in the simulation of AHA13 were split into different types (see Section 2.2), the most important being ‘type-1’ (virialized haloes) and ‘type-2’ (objects in late stages of formation). The results above are for the 1522 ‘type-1’ objects, while we have ignored the 438 ‘type-2’ objects. If the latter also form from peaks, our choice of ‘type-1’ could be a cause for concern since a peaks-based analysis such as ESP might simply not be able to distinguish between them. Indeed, we do not find a significant difference between the protohalo regions of ‘type-2’ and ‘type-1’ objects. Moreover, we find that including the ‘type-2’ objects in the analysis on the same footing as ‘type-1’ leads to a larger number (1097) of matched ESPeaks, meaning that 87 percent of our ESPeaks can be matched to some object that is either about to or has completely virialized. (The fraction of unmatched protohalo patches are now $\sim 44$ percent, as compared with 33 percent when using only ‘type-1’. This is largely simply because the combined number of ‘type-1’ and ‘type-2’ objects (160) is significantly larger than that of the ESPeaks, which can only be accommodated in the ESP calculation by lowering the collapse threshold.)

Interestingly, AHA13 found that the transition between ‘type-2’ and ‘type-1’ occurs fast and is associated with a rapid mass growth, bringing a ‘type-2’ object to a mass around or above the half-mode mass by the time it has virialized and thus turned into a ‘type-1’ object. This is consistent with the picture that power spectra with steeper (effective) slopes show enhanced accretion rates (Lacey & Cole 1993, 1994). These observations suggest that the excursion set calculation could be failing because it is unable to capture the quick mass growth that ‘type-1’ objects experience around the half-mode mass scale, possibly due to an incorrect prediction of collapse-time for a given peak patch. The rapid transition between these two types of objects means that even small errors in predicting the collapse time could dramatically alter the predicted locations and masses of fully virialized haloes. As we discuss in the next section, such an error can also account for the correlations we find between protohalo curvature and mass/location mismatches.

5 ANALYTICAL RESULTS

In this section we use the results of our numerical study to motivate an analytical approximation which captures the sharp cut-off in the mass function better than the standard ESP calculation.

5.1 A possible explanation for the mismatch between ESPeaks and protohaloes

The behaviour discussed above might be explained if, at small masses, the algorithm systematically overpredicts the time at which a given peak patch should collapse. This is because a patch that collapses earlier than predicted will have time to accrete mass by the time of interest, and will consequently have a larger mass than predicted. As noted by AHA13, low-mass WDM haloes tend to grow much more rapidly than their high-mass counterparts, so even a small error in collapse time could have a dramatic impact on the predicted mass function. This is further corroborated by our observation in Section 4.3 that the assignment of peaks to either virialized haloes or objects in the late stages of formation is somewhat uncertain.

Let us suppose that there is in fact such a systematic uncertainty in collapse time. This is not an unreasonable assumption; similar effects have been noticed and discussed by other authors (Monaco 1999; Giocoli et al. 2007) in the case of CDM. Such effects could arise due to simplifying choices made in models such as ellipsoidal collapse (see Appendix B for a justification), as well as due to other physical mechanisms such as assembly bias. Can this explain the orders-of-magnitude mass increases that are required to go from the ESPeaks mass function to the halo mass function? To see why this is indeed the case, consider the following.

A collapse-time uncertainty can be interpreted as introducing a second barrier in the problem, in the sense that if a patch has been identified at mass $m$ using the barrier $B$ (in this case the one in equation 28) it must then be allowed to accrete until mass $M$...
where the integral is over all relevant variables (e.g. peak density, curvature, shear, etc.) and the function $\mathcal{N}_{ab}(X)$ incorporates the intrinsic (Gaussian) probability of these variables, as well as the peaks constraint (equation 26) and the excursion set constraint which require first crossing of the chosen barrier. The latter means that the integration variables also include $\sigma_0$, and the Dirac delta $\delta_\beta(\ln m - \ln \bar{m}(\sigma_0))$ then assigns the mass according to the first-crossing scale $\sigma_0$, where $\bar{m}(\sigma_0)$ is the inverse function of $\sigma_0(m)$ (see Appendix A for details).

In the calculation of Paranjape et al. (2013), the barrier was assumed to be $B = \delta_c + \beta \sigma_0$, where $\beta$ is a stochastic variable whose distribution was motivated by the CDM measurements of Robertson et al. (2009) and was assumed independent of the tensors $\delta_i \psi$ and $\delta_i \delta$. In particular, $p(\beta)$ was taken to be lognormal with mean 0.5 and variance 0.25. The resulting mass function is then (equations 12 and 13 of Paranjape et al. 2013)

$$\frac{d n_{ESPStd}}{d \ln m} = v \frac{N_{ESPStd}(v)}{|d \ln \sigma_0/d \ln m|},$$

with $v \equiv \frac{\delta_c(z)}{\sigma_0(\delta)}$ and

$$v \frac{N_{ESPStd}(v)}{d \ln m} = \int d\beta \ p(\beta) \frac{e^{-\nu + \nu \beta^2/2}}{\sqrt{2\pi} \nu^2} \int_{\nu \beta}^{\infty} dx \ (x/\nu - \beta) F(x) \times p_G(x - \beta \nu - \nu; \nu; 1 - \nu^2),$$

where $F(x)$ is the BBKS curvature function (equation A11) and $p_G(y - \mu; \Sigma^2)$ is a Gaussian in the variable $y$ with mean $\mu$ and variance $\Sigma^2$. The solid black curve in Fig. 1 shows this expression, using the WDM transfer function (equation 1).

In order to implement the barrier (equation 28), we must account for the correlation between the eigenvalues of the velocity shear $\delta_i \psi$ and those of the density Hessian $\delta_i \delta$. Although this is straightforward in principle, in practice the misalignment between these tensors turns out to be cumbersome to deal with (see Appendix A). We therefore explore a simpler, albeit approximate, solution. Following Sheth et al. (2001), we look for the value at which the distribution $p(Y|\delta)$ has its maximum, ignoring the peaks constraint. (This distribution can be obtained by integrating equation A3 of Sheth et al. 2001, over the prolateness, and is different from $p(Y, Z = 0|\delta)$ which is what those authors worked with.) This happens at $Y_{\max} = 0.5020 \sigma_0 \approx 0.50 \sigma_0$. We therefore look for the first crossing of the "deterministic" barrier:

$$B = \delta_c + \beta \sigma_0 = \bar{\delta}_c + 0.5 \beta \sigma_0.$$  

The dashed black curve in Fig. 1 shows this expression, which amounts to replacing the integral over $\int d\beta \ p(\beta)$ in equation (32) with the single value $\beta = \bar{\beta} = 0.5$.

$$\frac{d n_{ESPDet}}{d \ln m} = v \frac{N_{ESPDet}(v)}{|d \ln \sigma_0/d \ln m|}.$$  

Since Paranjape et al. (2013) were interested in a CDM mass function, they used a top-hat filter to compute $\sigma_0$ but a Gaussian filter for $\sigma_1$ and $\sigma_2$. (Recall $\sigma_2$ diverges for a top-hat filtered CDM spectrum.) The smoothing scale $R_\theta$ for the latter was set by demanding $\langle \delta_i \delta_i \theta \rangle = \delta_\Theta$, i.e. $\langle \delta_i \delta_i \theta \rangle = \sigma_0(R_\Theta)^2$. Consequently, $V_i$ was computed using the Gaussian filter and $\gamma$ was defined using mixed filtering. We used this prescription for the solid black curve in Fig. 1. To keep things simple in the present work, however, we will define $a/d$ quantities in the analytical calculation using Gaussian filtering, with the filtering scale matched to the mass using the relation mentioned above. We have checked that switching to top-hat filtering for defining $\sigma_0$ has little impact on our results. Additionally, using Gaussian filtering throughout guarantees self-consistency; e.g. the relation $d\beta/d\sigma_0 = x/\nu$, which we use below, is exact in this case.
with
\[ v\mathcal{N}_{\text{ESPdet}}(v) = \int_{\beta_1}^{\infty} dx \mathcal{N}_{\text{ESPdet}}(v, x) \]
\[ = \frac{e^{-\frac{1}{2}(v^2 + \bar{\beta}^2)}}{\sqrt{2\pi} V_\nu} \int_{\beta_1}^{\infty} dx \left( \frac{x/\gamma - \bar{\beta}}{\gamma^2} \right) F(x) \]
\[ \times p\delta(x - \bar{\beta}) \gamma^2 - 1 - \gamma^2). \]  
(35)

We see that this describes the dashed blue histogram quite well (this was the result of our empirical walks algorithm for the barrier, equation 33, see Section 4.2). Consequently, it is also not very different from the solid orange histogram, which was the result of the empirical walks using the stochastic barrier (equation 28), as well as the standard ESP calculation (solid black curve).

A similar calculation gives the predicted distribution \( p(x|\text{ESP}) \) of ESPeak curvature. For the deterministic barrier (equation 33) this is
\[ p(x|\text{ESPdet}) = \frac{\int dx \ln \sigma_0 v\mathcal{N}_{\text{ESPdet}}(v, x)}{\int dx \ln \sigma_0 \int_{\beta_1}^{\infty} dx v\mathcal{N}_{\text{ESPdet}}(v, x)}, \]
(36)
where \( \mathcal{N}_{\text{ESPdet}}(v, x) \) was defined in equation (35). The dashed black curve in Fig. 4 shows the result; this is very similar in shape to the measured protohalo curvature distribution but has a higher mean value.

5.3 Re-assigning mass

To implement the mass re-assignment, we modify equation (30) by writing
\[ \frac{dn}{d\ln M} = \int D\mathcal{X} \mathcal{N}_{\text{pk}}(\mathcal{X}) p(\ln M|\mathcal{X}), \]
(37)
where the probability distribution \( p(\ln M|\mathcal{X}) \) accounts for the mass re-assignment. If the re-assignment were deterministic, this distribution would be a Dirac delta centred on the appropriate re-assigned mass value. Indeed, this is precisely what equation (30) does, except that it gets the mass wrong. In practice, in addition to changing the mass, we allow for some scatter, which is more realistic and also improves the numerical stability of our calculation.

Suppose the standard calculation identifies an ESPeak and assigns it a mass \( M \). If the collapse-time uncertainty discussed earlier leads to a barrier shift \( -\Delta B \), then the strongly correlated nature of the filtered density contrasts \( \delta \) at different smoothing scales means that the mass scale \( M \) at the new barrier satisfies
\[ B_M - \Delta B = B_m + \frac{x_m}{\gamma_m} (\sigma_M - \sigma_m), \]
(38)
where the subscript indicates smoothing scale, and we approximated the walk in density \( \delta(\sigma) \) as a straight line with slope \( d\delta/d\sigma = x/\gamma \). If we assume that \( \Delta B \) is Gaussian distributed with mean \( \Delta B \) and variance \( \sigma_{\Delta B}^2 \), and that \( B \) is given by equation (33), then we have
\[ p(\sigma_M|X) = \int d\Delta B p_G(\Delta B - \Delta B; \sigma_{\Delta B}^2) \theta_H \left( \sigma_m - \frac{\Delta B}{x_m/\gamma_m - \bar{\beta}} \right) \]
\[ \times \delta_D \left( \sigma_M - \sigma_m + \frac{\Delta B}{x_m/\gamma_m - \bar{\beta}} \right) \]
\[ = \theta_H (\sigma_M) \left( \frac{x_m}{\gamma_m - \bar{\beta}} \right) \]
\[ \times p_G \left( \frac{x_m}{\gamma_m - \bar{\beta}} \right) \left( \sigma_m - \sigma_M - \Delta B; \sigma_{\Delta B}^2 \right), \]
(39)
where the Heaviside function \( \theta_H(\cdot) \) ensures that \( \sigma_M > 0 \). The cumulative probability \( p(\ln M|\mathcal{X}) \) satisfies
\[ p(\ln M|\mathcal{X}) = p(< \sigma_M|X) \]
\[ = \frac{1}{2} \left[ \text{erf} \left( \frac{(x_m/\gamma_m - \bar{\beta})(\sigma_m - \sigma_M - \Delta B)}{\sqrt{2} \sigma_{\Delta B}} \right) \right. \]
\[ - \text{erf} \left( \frac{(x_m/\gamma_m - \bar{\beta})(\sigma_m - \Delta B)}{\sqrt{2} \sigma_{\Delta B}} \right) \right]. \]
(40)
The mass function with re-assigned masses then becomes
\[ \frac{dn}{d\ln M} = \int d\ln \sigma_0 e^{-\frac{1}{2} \frac{B_m^2}{2\sigma_{\text{pk}}}} \int_{\beta_1}^{\infty} dx \left( \frac{x/\gamma_m - \bar{\beta}}{\gamma^2} \right) F(x) p_G \left( x - \frac{\gamma_m B_m}{\sigma_m}; 1 - \gamma^2 \right) p(\ln M|\mathcal{X}). \]
(41)

We were to account for the full stochasticity of \( B \) using equation (28), the expression for the mass function in equation (41) would have an additional integral over \( Y \), and the integrand, including the distribution \( p(\ln M|\mathcal{X}) \), would be more complicated. However, since the mass function without mass re-assignment and with \( \beta = \bar{\beta} = 0.5 \) describes the results of the empirical walks with barrier (equation 28) quite well (cf. discussion below equation 33), we will continue to ignore this inherent stochasticity due to the ellipticity \( Y \).

5.4 An explicit example

The distribution of \( \Delta B \) will in general depend on the scale \( m \) at which the ESPeak is originally identified; we know that at large \( m \) the mass assignment is essentially correct, with small scatter, while there is a trend towards underpredicting masses at small \( m \). Since there is little guidance from theory for the actual values of \( \Delta B \) and \( \sigma_{\Delta B} \), we have left these as free parameters, except for requiring that they become numerically small for large \( m \) or small \( \sigma_{\text{pk}} \). The following is intended as a proof of principle, and we leave a more detailed analysis and estimate of \( \Delta B \) to future work.

We compare \( \sigma_{\text{pk}} \) with the scale \( \sigma_{\text{pk}} \) which we define as the scale at which the Jacobian between \( \sigma \) and \( m \) becomes small. In particular, we set
\[ \frac{d\ln \sigma_0}{d\ln m}|_{\sigma_{\text{pk}}} = 0.1. \]
(42)
In practice, for \( m_{\text{pk}} = 0.25 \text{keV} \) this occurs at \( m \approx 3.2 \times 10^{12} h^{-1} \text{M}_\odot \) which is close to the half-mode mass scale \( M_{\text{hm}} \approx 3 \times 10^{12} h^{-1} \text{M}_\odot \). We have chosen this definition since it remains well behaved in the CDM limit as well, whereas the half-mode mass goes to zero in that case. The choice of 0.1 as the threshold in equation (42) is, however, arbitrary.

The solid blue curve in Fig. 11 shows the result of using equation (41) after setting
\[ \Delta B = 5\sigma_{\Delta B} = 0.175(\sigma_{\text{pk}}/\sigma_{\text{pk}})^3. \]
(43)
The shape of the turnover is quite sensitive to the value of \( \Delta B \), less so to \( \sigma_{\Delta B} \). The numerical values of the amplitude and exponent in the expression on the right are quite degenerate. With these settings, the ESP mass function with re-assigned masses gives a fairly good description of the halo masses in the simulation (histogram). For comparison, the figure also shows the ESP mass function using the barrier (equation 33) but before mass re-assignment (dashed black; this is the same as in Fig. 1), and the Tinker et al. (2008) fitting function (dotted black). Additionally, the dot-dashed blue curve shows the sharp-\( \kappa \) excursion set fit proposed by Schneider et al.
Halo mass functions: the solid red histogram and dashed black curve are the same as in Fig. 1 and show, respectively, the mass function of haloes in the simulation and the ESP prediction (equation 34) using the deterministic barrier (equation 33). The solid blue curve shows the result of re-assigning masses using equation (41), with parameter values from equation (43). For comparison, the dot–dashed blue curve shows the fit based on sharp-$k$ filtering presented by Schneider et al. (2013), their spherical collapse fit with $q = 1$ and $c = 2.7$, which is a good description of the mass function measured in their simulation. The dotted black curve shows the Tinker et al. (2008) fitting form.

The top panel of Fig. 12 shows our analytical mass function at $z = 0$ with masses re-assigned using the same parameter values as in equation (43) for three different values of dark matter particle mass $m_{dm}$ (solid curves). We also show the corresponding curves fit by Schneider et al. (2013), again using their spherical collapse fit (dot–dashed curves). In the bottom panel we show our analytical prediction for $m_{dm} = 0.25$ keV, now for three different redshifts. The solid curves marked 0.25 keV in both panels are identical, and the same as the solid blue curve in Fig. 11.

The predictions of the ellipsoidal collapse model, augmented by a systematic uncertainty in mass assignment, accurately describe the WDM mass function. By the logic discussed earlier, the same expressions with the CDM transfer function should describe the CDM mass function. In particular, the half-mode mass scale for CDM is small enough that, in practice, every halo has a virialized progenitor. This means that the effects of collapse-time uncertainty – which were very pronounced around the half-mode mass of WDM due to the rapid growth of those objects – are now essentially an

\[ f_{coll} = \int \frac{dM}{\bar{\rho}} \frac{d\ln M}{d\ln M} \]  

(44)

is not held fixed during the re-assignment, which is obvious since the calculation allows haloes to accrete more mass than is predicted in the standard ESP treatment. The values for $f_{coll}$ returned by the analytical calculation before and after mass re-assignment are, respectively, 0.19 and 0.23. In comparison, the mass fraction in actual haloes is 0.18, but note that this number can fluctuate due to sample variance effects at the high-mass end.

5.5 Consequences for CDM

The predictions of the ellipsoidal collapse model, augmented by a systematic uncertainty in mass assignment, accurately describe the WDM mass function. By the logic discussed earlier, the same expressions with the CDM transfer function should describe the CDM mass function. In particular, the half-mode mass scale for CDM is small enough that, in practice, every halo has a virialized progenitor. This means that the effects of collapse-time uncertainty – which were very pronounced around the half-mode mass of WDM due to the rapid growth of those objects – are now essentially an
In other words, our proposed modifications to the ESP calculation not only correctly describe the sharp turn in the WDM mass function, but also describe the CDM mass function with the same accuracy as the standard ESP calculation of Paranjape et al. (2013). For CDM we have also checked that the linear Lagrangian halo bias predicted by this model (not shown) matches measurements in CDM simulations (Tinker et al. 2010) with the same accuracy as the Paranjape et al. (2013) calculation. However, as noted earlier, low mass in WDM does not mean low significance, and, in principle, low significance CDM haloes could be different from low-mass WDM haloes. Testing this would need high-resolution CDM simulations, or WDM simulations with a slightly larger mass such as \( m_{\text{dm}} \sim 0.5 \text{ keV} \).

6 DISCUSSION AND CONCLUSIONS

Can we predict where and when haloes form? In this paper, we have thoroughly evaluated our ability to predict the abundance of collapsed objects by performing an in-depth analysis of the properties of the initial density field at the locations where collapse occurs in numerical simulations. To accomplish this, we used a perturbation spectrum with a small-scale cut-off such as those arising in WDM cosmologies. As discussed in Section 2, the resulting suppression of low-mass haloes provides powerful additional leverage which is absent in the CDM case.

Numerical simulations have traditionally had great difficulty in making a prediction for the abundance of haloes in such a scenario due to the artificial fragmentation of filaments – a problem that has only recently been overcome by AHA13. As a consequence, we were in the unique situation of being able to perform a thorough comparison between this numerical experiment and the mass function predicted from excursion set theory, by analysing the properties of haloes on an object-by-object basis. We summarize our results below, and discuss some outstanding issues.

It is well known that the standard excursion set approach predicts a mass function that is completely inconsistent with the numerical results (see e.g. Schneider et al. 2012, and also our Fig. 1). We showed that the inclusion of the peaks constraint in excursion sets (Musso & Sheth 2012; Paranjape & Sheth 2012; Paranjape et al. 2013) leads to a turnover in the mass function as well as an overall number of collapsed objects that is consistent with the simulation results. However, it also predicts masses around and below the half-mode mass scale that are significantly smaller than those measured in the simulation, leading to a small-mass slope of the mass function (\( dn/d\ln m \sim m^{2/3} \)) that is inconsistent with that found from the simulation (cf. Fig. 1). This prediction is remarkably robust against changing details of the calculation such as the shape and stochasticity of the barrier.

We next investigated the origin of this discrepancy between simulation and theoretical predictions. In particular, we analysed the Lagrangian properties of ‘protohaloes’ (the initial locations of groups of particles which are eventually identified as haloes in the simulation), and also performed empirical excursion set peak walks in the initial density field used in the simulation. We can summarize our findings as follows.

(1) All haloes in the simulation are consistent with forming near peaks in the initial density field (Fig. 4).

(2) The overdensities of protohaloes are strongly correlated with their shear ellipticities, but show no correlation with the shear prolateness (Fig. 5). The former is expected from arguments based on ellipsoidal collapse dynamics, while the latter is not (compare

Figure 13. ESP mass function predictions for CDM using the deterministic barrier (equation 33) (dashed black) and after re-assigning masses (solid blue). The dotted black curve shows the Tinker et al. (2008) fitting form, which was calibrated over a range between \( 10^{10.5} \) and \( 10^{15.5} \ h^{-1} \ M_{\odot} \).

The solid red curve shows the prediction, after mass re-assignment, for a ‘realistic’ WDM particle candidate with \( m_{\text{dm}} = 3.3 \text{ keV} \). In all cases the mass re-assignment parameters are the same as in equation (43). The lower panel shows the ratio of the CDM predictions to the Tinker et al. fit.

uncertainty in the time of major mergers, and consequently the associated mass mismatch must be significantly smaller. We see in Fig. 13 that this is indeed the case for masses \( m \gtrsim 10^{11} h^{-1} \ M_{\odot} \). It is also reassuring to note that changing the value of \( \Delta_{\text{vir}} B \) in the CDM case has much less effect on the mass function at \( m \gtrsim 10^{11} h^{-1} \ M_{\odot} \) than in the WDM case. For example, we have checked that increasing the amplitude in equation (43) by a factor of 1.5 or changing the exponent from 3 to 2 both lead to \( \lesssim 3 \) per cent changes in the CDM mass function for \( m \gtrsim 10^{11} h^{-1} \ M_{\odot} \).

At lower masses our specific implementation of mass re-assignment predicts a factor of \( \sim 2 \) larger number of haloes than expected from the mass function fit by Tinker et al. (2008). This behaviour of the re-assigned mass function at low masses is not very robust; however, it is sensitive to the specific numerical choice in equation (42). It is possible to adjust this number (and those in equation 43) to simultaneously get a good match to the CDM and WDM simulations, although we have not pursued this exercise here. One must also keep in mind that the Tinker et al. (2008) fit was calibrated for masses between \( \sim 10^{0.5} \) and \( \sim 10^{15.5} \ h^{-1} \ M_{\odot} \).
equation 28 with equation 29). The fact that the protohalo overdensity has no correlation with its prolateness is intimately connected with the distribution of individual shear eigenvalues and hence with the dynamical ordering of the collapse times of each (Ludlow & Porciani 2011b; Despali et al. 2013). It will be interesting to find a dynamical model that is consistent with our results, perhaps along the lines presented by Ludlow & Porciani (2011b).

(3) The number of ‘ESPeaks’ (1261) identified by our empirical algorithm (Section 4.1) is reasonably close to the actual number of protohaloes (1522).

(4) A significant fraction (74 per cent) of ESPeaks can be matched to actual protohaloes, while 64 per cent of the protohaloes can be matched to ESPeaks (details in Section 4.3).

(5) The curvatures of these matched objects are significantly higher, and their overdensities significantly lower, than those of protohaloes and ESPeaks that could not be matched to each other (Fig. 8).

(6) Most strikingly, the masses of ESPeaks are systematically lower than the protohalo masses. This is true for both matched and unmatched objects (respectively, top and bottom panels of Fig. 7). Since the ESPpeak mass function is very well described by the ESP calculation (Fig. 1), this fully accounts for the discrepancy between the ESP halo mass function and that measured in the simulation.

(7) For matched objects, the mismatch in mass assignment correlates with protohalo curvature (Fig. 7), while the scatter in the mass mismatch correlates with protohalo shape (Fig. 6).

(8) We have checked that, apart from having larger overdensities and lower curvatures than their matched counterparts, the unmatched protohaloes do not appear to be special in any other property related to the velocity shear, density Hessian or moment of inertia.

(9) If we also include in the analysis objects in late stages of formation (‘type-2’ in AHA13), 87 per cent of the ESPeaks can be matched to protohaloes (although the fraction of unmatched protohaloes are now larger, largely because the total number of protohaloes increases). These ‘type-2’ objects are known to be undergoing rapid mass growth (AHA13).

Based on these results, we argued that the likely cause for the observed mass mismatch between ESPeaks and protohaloes is a systematic overprediction of the collapse time for a given perturbation. We then showed how such an uncertainty can be accounted for and corrected in the excursion set language (Section 5.1 and Fig. 10), and presented an explicit example of such a correction which describes the numerical WDM results very well (Figs 11 and 12). As an important consistency check, we also showed that the same model gives an accurate description of the CDM mass function, with the simple replacement of the WDM initial power spectrum with that of CDM (Fig. 13).

We emphasize that our solution works because it explicitly alters the mass assignment step of the ESP calculation, in our case by introducing a second barrier. Simply introducing new statistical variables defined by smoothing the initial density field in a single-barrier calculation (say, by setting \( B \to \delta_\ell + \gamma + \tau \sigma_\ell \)) would not work, because the predicted mass function in this case would still behave as \( \log m \sim m^{1/3} \) at low masses, as discussed in Section 2.3. At the heart of this issue is the difference between the physics of individual halo formation and the statistics of the initial density field: mismatches in collapse time predictions are primarily a physical, not statistical, problem. In CDM, since the \( \sigma_\ell(m) \) relation is always steep, errors in the physical collapse model can be accommodated by altering the statistical modelling (e.g. by changing the barrier shape as a function of \( \sigma_\ell \), or by introducing stochasticity in the barrier). WDM, on the other hand, presents us with a situation where such solutions no longer work since the \( \sigma_\ell(m) \) relation ‘freezes out’ at small masses (Fig. 2). A full solution of the problem would likely involve a single barrier with an explicit dependence on mass (rather than \( \sigma_\ell \)) which is fixed by an accurate model of collapse.

Our work can be extended in several directions which could yield clues towards building a more predictive model. Although we argued that uncertainties in collapse time can easily arise in toy models such as ellipsoidal collapse, or due to assembly-bias-like effects, we have not provided conclusive evidence pinpointing a specific physical mechanism. Furthermore, the protohalo curvatures we measure are significantly lower than the simplest ESP prediction (Fig. 4). While this might be partially accounted for by the mass re-assignment, it is not clear whether there is a deeper reason for this discrepancy. Also, the scatter in ESPpeak mass at fixed halo mass for matched objects correlates strongly with the protohalo shapes (Fig. 6). Our mass re-assignment currently incorporates only curvature, and it will be interesting to additionally account for protohalo shapes. The overdensities of unmatched protohaloes have a significantly higher tail than that of matched objects (Fig. 8). This could be indicating that low-mass WDM haloes can form near density peaks without necessarily satisfying the excursion set peaks constraints; e.g. there could be some other criterion for a sufficiently overdense patch to virialize. One way forward would be to analyse the local environments of these unmatched objects to look for peculiarities. Finally, while our chosen particle mass of \( m_{\text{gas}} = 0.25 \, \text{keV} \) is untypically warm, any realistic massive dark matter candidate would lead to very similar phenomenology around its half-mode mass scale, and our analysis would be relevant whenever this scale can be numerically resolved. As mentioned earlier, however, for a particle much colder than our present choice, low-mass haloes will also have low significance \( (\nu \ll 1) \) which was not possible in our case, and this could lead to additional effects. Clearly, it will be of great interest to investigate these aspects further in future work.

ACKNOWLEDGEMENTS

We thank Raul Angulo for providing us with the halo catalogues published in AHA13. We wish to thank Ravi Sheth, Tom Abel and Cristiano Porciani for insightful discussions and comments on the draft. OH acknowledges support from the Swiss National Science Foundation (SNSF) through the Ambizione fellowship.

REFERENCES

Achitouv I., Rasera Y., Sheth R. K., Corasaniti P. S., 2012, preprint (arXiv:1212.1166)

Angrick C., Bartelmann M., 2010, A&A, 518, A38

Angulo R. E., Hahn O., Abel T., 2013, MNRAS, 434, 3337 (AHA13)

Appel L., Jones B. J. T., 1990, MNRAS, 245, 522

Arnold V. I., Shandarin S. F., Zeldovich I. B., 1982, Geophys. Astrophys. Fluid Dyn., 20, 111

Avila-Reese V., Colin P., Valenzuela O., D’Onghia E., Firmani C., 2001, ApJ, 559, 516

Bardes J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Bode P., Ostriker J. P., Turok N., 2001, ApJ, 556, 93

Bond J. R., 1989, in Ashbury A., Campbell B., Israel W., Kamal A., Khanne F., eds, Proceedings of the Lake Louise Winter Institute, Frontiers in Physics – From Colliders to Cosmology. World Scientific Press, Singapore, p. 182
van de Weygaert R., Bertschinger E., 1996, MNRAS, 281, 84
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
White S. D. M., 1996, in Schaeffer R., Silk J., Spriro M., Zinn-Justin J., eds, Cosmology and Large Scale Structure. Elsevier, Amsterdam, p. 349
White S. D. M., Silk J., 1979, ApJ, 231, 1

APPENDIX A: FULLY STOCHASTIC EXCURSION SET PEAKS

Here we present some of the formal details of the excursion set peaks calculation. This will highlight the difficulty in working with the full stochasticity that must be dealt with when using a barrier such as equation (28), and motivate the simpler, deterministic approximation (equation 33) used in the main text.

A1 Formal expression for mass function

In what follows, an overdot denotes a derivative with respect to \( \sigma_0 \). All variables are assumed to be Gaussian filtered on a scale \( R \) which is related to the mass \( m \) through \( m = 4(\pi/3)V R^3 \), where \( V R^3 \) satisfies 

\[
\mathcal{N}(R) = \mathcal{N}(R_{TH}) = \left( \frac{\delta_{TH}(R_{TH})}{\delta_{TH}} \right)^3 \]

with the subscripts ‘G’ and ‘TH’ on \( \delta \) denoting Gaussian and top-hat filtering, respectively (Paranjape et al. 2013, see also Footnote 7). In practice this gives \( R \approx 0.46 R_{TH} \) with a slow variation. We will drop the filter subscripts below.

The ESP mass function assuming a barrier \( B \) (which can be stochastic) can be written as equation (30) where, in full glory, we have

\[
\int D\xi \equiv \int d\ln \sigma_0 \, d^3 \psi \, d^3 \xi \, d^3 \eta \tag{A1}
\]

and

\[
\mathcal{N}_{\sigma_0}(X) \equiv p(\partial_i \psi, \partial_j \delta, \nabla \delta) \, Pk(\partial_i \delta, \nabla \delta) \, ES(\sigma_0, [\delta, B]). \tag{A2}
\]

The expression (A1) involves a total of 16 integration variables: the smoothing scale represented by \( \sigma_0 \), the six independent components of the shear tensor \( \partial_i \psi \) represented by \( d^3 \psi \), similarly the components of the Hessian of the density \( \partial_i \delta \) represented by \( d^3 \xi \), and finally the three components of the density gradient \( \nabla \delta \) represented by \( d^3 \eta \). Strictly speaking, we must also include the scale derivative of density \( \delta \) (which will appear in the excursion set constraint) as a separate variable, but Gaussian filtering ensures that \( \delta = x/y \) where \( x \) was defined in equation (15) and \( y \) in equation (8), so this is included in \( \partial_i \delta \).

The raw number density of peaks \( \mathcal{N}(X) \) defined in equation (A2) consists of the following quantities: the joint (Gaussian) distribution function \( p(\partial_i \psi, \partial_j \delta, \nabla \delta) \) of the shear, density gradient and density Hessian smoothed on scale \( R \); the peaks constraint \( Pk(\partial_i \delta, \nabla \delta) \) which enforces \( \nabla \delta = 0 \) and \( \zeta_i < 0 \), where \( \zeta_i \) are the eigenvalues of \( \partial_i \delta \); and the excursion set constraint \( ES(\sigma_0, [\delta, B]) \) which enforces up-crossing of the barrier by the random walk at the scale \( \sigma_0(R) \). The compact notation in ES(...) hides the fact that the up-crossing condition will introduce a dependence on \( \delta = x/y \) and possibly other stochastic quantities through \( B \). The Dirac delta in equation (30) then sets the mass to be \( m = m(\sigma_0) \), where \( m(\sigma_0) \) is the inverse function of \( \sigma_0(R(m)) \) as discussed above, which is straightforward to compute numerically.

---

\(^8\)Musso & Sheth (2012, 2013) discuss why up-crossing is a sufficiently accurate approximation to first-crossing for walks with correlated steps.
In detail, following BBKS, we have
\[
P_k(\partial_{ij} \delta, \nabla \delta) = \delta_0(\nabla \delta) \mid_{\xi_1, \xi_2, \xi_3}[\partial_t(-\xi_1)],
\] (A3)
where we have assumed the ordering $\xi_1 \leq \xi_2 \leq \xi_3$, while the excursion set constraint can be written as
\[
\mathcal{E}(\sigma_0, \{ \delta, B \}) = (x/y - B) \partial_t(x/y - B) \delta_0(\mu - B/\sigma_0),
\] (A4)
where we defined $\mu \equiv \delta/\sigma_0$. The Dirac delta in equation (A4) enforces barrier-crossing, while the terms involving $x/y = \delta$ ensure that this is an $\nu$-crossing.

The intrinsic Gaussian distribution of these variables couples the tensors $\partial_{ij} \psi$ and $\partial_{ij} \delta$ through the correlation structure (van de Weygaert & Bertschinger 1996)
\[
\begin{align*}
\langle \partial_{ij} \psi \partial_{kl} \psi \rangle &= \frac{\sigma_z^2}{15} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) , \\
\langle \partial_{ij} \delta \partial_{kl} \delta \rangle &= \frac{\sigma_z^2}{15} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) , \\
\langle -\partial_{ij} \delta \partial_{kl} \psi \rangle &= \frac{\sigma_z^2}{15} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) ,
\end{align*}
\] (A5)
where the $\delta_{ij}$ are Kronecker deltas, whereas the vector $\nabla \delta$ is uncorrelated with the tensors and satisfies
\[
\begin{align*}
\langle \nabla_i \delta \nabla_j \delta \rangle &= \frac{\sigma_z^2}{3} \delta_{ij} , \\
\langle \nabla_i \delta \partial_{ij} \psi \rangle &= 0 = \langle \nabla_i \delta \partial_{ij} \delta \rangle .
\end{align*}
\] (A6)

**A2 Consequences of misalignment**

The 12 degrees of freedom in the two tensors $\partial_{ij} \psi$ and $\partial_{ij} \delta$ are most conveniently organized as the three eigenvalues of each, the three relative Euler angles between their respective eigenvectors and three additional Euler angles that fix the orientation of one of them with respect to the chosen basis.

The correlation structure between the tensors $\partial_{ij} \psi$ and $\partial_{ij} \delta$ implies that, in a generic realization of the field, they will be misaligned. That is, although they are strongly correlated (with correlation coefficient $\gamma$), their respective eigenvectors will not be parallel (see e.g. Desjacques 2008; Lavaux & Wandelt 2010; Despali et al. 2013). The misalignment is captured by the three relative Euler angles, and these must be marginalized over (the other three angles never appear in the distribution due to statistical isotropy, and can be trivially marginalized over). The effect of this marginalization is to introduce a non-trivial coupling between the anisotropic combinations of the two tensors. So if we define $y$ and $z$ as in equations (19) and (20) then, as shown in an elegant calculation by Desjacques (2008), the marginalization over Euler angles will introduce terms involving products between at least one of $\{y + z\}$, $\{3y - z\}$ and at least one of $\{Y + Z\}/\sigma_0$, $\{3Y - Z\}/\sigma_0$, where $Y$ and $Z$ were defined in equations (17) and (18), respectively. This is quite different from the coupling between the isotropic variables $x$ and $\mu$ through the familiar Gaussian term $\sim -\gamma$ that appears in the BBKS calculation.

This anisotropic coupling is irrelevant if we were to ignore the excursion set constraint $\mathcal{E}(\cdot)$ in equation (A2) or if the barrier $B$ in $\mathcal{E}(\cdot)$ did not depend on any of the anisotropic variables $\{y, z, Y, Z\}$ (an example would be the barrier used by Paranjape et al. 2013), since in these cases one could marginalize over, e.g. $Y$ and $Z$ to recover expressions similar to those analysed by BBKS (note that ignoring the excursion set constraint altogether would give back the BBKS calculation exactly). In the case of the barrier (equation 28), however, the calculation will involve integration of a term like $\sim p(\lambda | \xi e^{-\gamma \delta^2/2(1-\gamma^2)}$ over $Y$ and $Z$, where the distribution $p(\lambda | \xi)$ of the eigenvalues of $\partial_{ij} \psi$ given the eigenvalues of $\partial_{ij} \delta$ is given by equation (22) of Desjacques (2008).

This is not all, however. The excursion set constraint also involves the derivative $B$, which in this case would lead to a term $\hat{Y}$. For Gaussian filtering, it is easy to check that the individual components of $\partial_{ij} \psi$ and $\partial_{ij} \delta$ are related by
\[
d(\partial_{ij} \psi)/d\sigma_0 = -(\partial_{ij} \delta)/(\gamma \sigma_2),
\] (A7)
whose trace gives the relation $\delta = x/y$ quoted earlier. If the tensors were perfectly aligned, this would also imply $\hat{Y} = y/y'$. In the general case, however, $\hat{Y}$ depends on the eigenvalues of $\partial_{ij} \delta$ and the relative Euler angles between the two tensors in a highly non-linear way. Using the ellipsoidal collapse barrier (equation 29) would make things even more complicated.

As an example, consider using the barrier (equation 28) with the approximation $\hat{Y} = y/y'$ but otherwise assuming a generic misalignment between the tensors. In this case, after a BBKS-like calculation, the mass function (with or without non-standard mass assignment) becomes
\[
\frac{dn}{d \ln m} = \int d \ln \sigma_0 \frac{N}{V_c} \frac{1}{y'} \left(1 - \frac{\gamma}{y'}\right)^4
\times \int d \bar{x} d \bar{y} d \bar{z} \chi(\bar{x}, \bar{y}, \bar{z}) \int d \bar{y} d \bar{Z} \chi(\bar{y} + \bar{Y}, \bar{Y}, \bar{Z})
\times (\bar{x} - \bar{y}) \bar{F}(\bar{x}, \bar{y}, \bar{z}) (\bar{Y}^2 - \bar{Z}^2) W(\bar{y}, \bar{z}, \bar{Y}, \bar{Z})
\times e^{-\frac{1}{2}(\bar{Y} - \bar{Z})^2} e^{-\frac{1}{2}(\bar{Y} + \bar{Z})^2} e^{-\frac{1}{2}(\bar{X} - \bar{Y})^2}
\times e^{\frac{1}{2}x^2} p(\ln m | \bar{X}),
\] (A8)
where
\[
N \equiv S^3(2\pi)^2, \quad \chi(\bar{x}, \bar{y}, \bar{z}) = \{x, y, z\}/\sqrt{1 - \gamma^2},
\]
\[
\{\bar{y}, \bar{Y}, \bar{Z}\} = \{\bar{y}, \bar{Y}, \bar{Z}\}/\sigma_0 \sqrt{1 - \gamma^2},
\]
\[
\bar{F}(\bar{x}, \bar{y}, \bar{z}) = \frac{\bar{y} (\bar{z}^2 - \bar{x}^2)}{\sqrt{\bar{y}^2 - 2 \bar{x}^2 (\bar{z}^2 - \bar{x}^2)}},
\]
the function $\chi(x, t, u)$ is unity when $-t \leq u \leq t, t \geq 0, s \geq 3t - u$ and is zero otherwise, and the function $W(\bar{y}, \bar{z}, \bar{Y}, \bar{Z})$ which captures the effect of marginalizing over the relative Euler angles $\hat{Y}$ has the Taylor expansion (Desjacques 2008):
\[
W(\bar{y}, \bar{z}, \bar{Y}, \bar{Z}) = 1 + \frac{\kappa^2}{10} W_2 + \frac{\kappa^3}{45} W_3 + \frac{\kappa^4}{105} W_4 + \frac{\kappa^5}{280} W_5
\]
\[
+ \frac{\kappa^6}{2310} W_6 + O(\kappa^7),
\] (A9)
where $\kappa = S^2/4$ and
\[
W_2 = 16(\bar{y}^2 + \bar{z}^2)(\bar{Y}^2 + \bar{Z}^2),
\]
\[
W_6 = 64z \bar{Z}(9\bar{y}^2 - \bar{z}^2)(9\bar{Y}^2 - \bar{Z}^2).
\] (A10)

The integrals over $\bar{y}$, $\bar{z}$ and $\bar{Z}$ are tedious but can be expressed in closed form. The remaining integrals over $\bar{x}$, $\bar{y}$ and $\sigma_0$ must be done numerically. Fig. A1 shows the result of the zeroth-order calculation and that of successively including higher powers of $\kappa$, with standard mass assignment. We see that the odd powers do not contribute significantly. The even powers, however, have not converged since the $\kappa^4$ term gives a significant contribution compared to the second-order calculation. Presumably one would have to continue the calculation at least to order $\kappa^6$ (if not $\kappa^8$), to get reasonable convergence.
Figure A1. Accounting for the effects of stochasticity: ESP mass function predictions for WDM using the stochastic barrier (equation 28) and the approximation $\bar{y} = y/y'$ (see text for details). From bottom to top, the solid curves show the result of keeping the zeroth-, second- and fourth-order terms in equation (A9), while the dashed curves (almost indistinguishable from the upper solid curves) include the third- and fifth-order terms. The upper short-dashed black curve is the same as in Fig. 1 and shows the ESP calculation using the deterministic barrier (equation 33), which gives a very good description of the ESPeaks mass function (solid orange histogram in Fig. 1) obtained with the stochastic barrier (equation 28). While the odd order terms do not contribute significantly to the stochastic calculation, the even order terms have clearly not converged – the fourth-order term gives a significant contribution compared to the second-order calculation.

**A3 A simpler way out**

As the results of the previous section show, it rapidly becomes very complicated to deal with the full stochasticity inherent in even a simple barrier prescription like equation (28). Luckily, approximating this barrier with the deterministic expression (33) leads to an excellent description of the mass function of ESPeaks found using the stochastic barrier (equation 28) (compare the dashed black curve in Fig. 1 with the solid orange histogram).

The calculation proceeds by setting $B = b_0 + 0.5\sigma_0$ and $\dot{B} = 0.5$ in equation (A4), which means that the excursion set constraint is independent of $Y$ and $Z$. These variables, together with the problematic relative Euler angles between the tensors, can then be trivially marginalized over, without having to Taylor expand as in equation (A9). The remaining variables can then be dealt with exactly like in BBKS; the only integral that cannot be done analytically is the one over $x$, since this involves the BBKS curvature function $F(x)$ given by

$$F(x) = \frac{1}{2} \left( x^3 - 3x \right) \left\{ \text{erf} \left( x \sqrt{\frac{3}{2}} \right) + \text{erf} \left( x \sqrt{\frac{3}{8}} \right) \right\} + \left[ \frac{2}{5\pi} \left( \frac{3x^2}{4} + \frac{8}{5} \right) e^{-5x^2/8} + \left( \frac{x^2}{2} - \frac{8}{5} \right) e^{-5x^2/2} \right]$$

(equations A14–A19 in BBKS). This is, of course, exactly the calculation performed by Paranjape & Sheth (2012) and later used by Paranjape et al. (2013).

The main message here is that the barrier stochasticity induced by ellipsoidal effects is a technical detail that does not address the main problem – that of mass re-assignment – we are faced with. Rather, the simplicity of the deterministic barrier solution allows us to tackle this problem in a computationally straightforward way, as discussed in the main text. In principle, one could imagine having a physically better motivated model of mass re-assignment; provided it can be expressed in the formal language of equation (37), the hurdles in using it to make mass function predictions would be purely technical.

**APPENDIX B: ELLIPSOIDAL DYNAMICS AND COLLAPSE-TIME UNCERTAINTY**

To understand why one might expect a small systematic uncertainty in the predicted collapse time in ellipsoidal dynamics, it will help to first recapitulate some of the basic features of spherical collapse.

Consider the simplest case of an EdS ($\Omega_{\text{tot}} = \Omega_{\text{m}} = 1$) background. Recall that the comoving Eulerian radius $R$ and overdensity $\Delta = \rho/\rho_{\text{crit}}$ of a collapsing homogeneous sphere with initial overdensity $\delta_i = (5/3)\delta_\text{th}/(1 + z) > 0$ and comoving Lagrangian radius $R_0$ are given by (Gunn & Gott 1972)

$$\Delta = \left( \frac{R_0}{R} \right)^3 = \frac{9}{2} \left( \frac{\theta - \sin \theta}{1 - \cos \theta} \right)^2,$$

$$\delta_0 = \delta_i \left( \frac{t}{t_0} \right)^{2\pi} = \delta_l = \frac{3}{2} \left( \frac{3}{4} \right)^{2\pi} \left( \theta - \sin \theta \right)^{2/3},$$

where $t_0$ is the present epoch and $\delta_l$ is the linear overdensity extrapolated to redshift $z$. The radius turns around at $\theta = \tau$ and reaches zero at $\theta = 2\pi$. In the standard approach, instead of evolving the solution to zero radius, one argues that the radius will become essentially constant once the object virializes. While this is not captured by the spherical model, a simple energy conservation argument says that the physical ‘virial radius’ must be one half of the physical radius at turnaround.

The key conceptual point here is that the model itself reaches the virial radius at $t_{\text{vir}} = t(\theta = 3\pi/2)$ which occurs slightly before the collapse time $t_{\text{coll}} = t(\theta = 2\pi)$; in particular, $t_{\text{vir}} \approx 0.91 t_{\text{coll}}$. However, one can argue that virialization actually occurs over a dynamical time-scale $t_{\text{dyn}}$ (the free-fall time at virialization; see e.g. Padmanabhan 1993) which, coincidentally, is of order 10 per cent of the collapse time. So in practice it is quite reasonable to use $t_{\text{coll}} \approx t_{\text{vir}} + t_{\text{phys}}$ in place of $t_{\text{vir}}$ when computing the linearly extrapolated overdensity, which gives the well-known result $\delta_l = 1.686$. (Had we instead evaluated $\delta_l$ at $t_{\text{vir}}$, we would get $\delta_l = 1.583$, a ∼6 per cent decrement from the traditional value; see e.g. Bond & Myers 1996, hereafter BM96.)

The non-linear overdensity at virialization follows from energy conservation and matching to the turnaround time $t_{\text{ta}}$:

$$\Delta_{\text{vir}} = \Delta_{\text{vir}}(t_{\text{vir}}/t_{\text{ta}})^2(R_{\text{phys,vir}}/R_{\text{phys,ta}})^3$$

$$= (9\pi^2/2)(t_{\text{vir}}/t_{\text{ta}})^2$$

$$\approx (9\pi^2/2)(t_{\text{coll}}/t_{\text{ta}})^2 = 18\pi^2.$$

Notice that the traditional value of $18\pi^2 \approx 178$ arises after approximating $t_{\text{vir}} \approx t_{\text{coll}}$: the actual non-linear density at $t_{\text{vir}}$ would be ∼20 per cent smaller.
In the ellipsoidal model discussed by BM96 (see also White & Silk 1979; Eisenstein & Loeb 1995; Monaco 1999), it is more difficult to identify conserved quantities since there is no spherical symmetry and non-linear tides can have a complicated influence. The definition of virialization is therefore somewhat ambiguous. BM96 settled on a simple prescription in which each principle axis is evolved non-linearly until it shrinks to a predetermined fraction $f_c$ of the global scale factor, after which it is assumed to remain fixed at that value. Collapse is defined as the time at which the longest axis (i.e. the smallest shear eigenvalue) satisfies this condition. Choosing $f_c = 0.178$ ensures that the overdensity at virialization for a spherical configuration is $f_c^{-1/3} \approx 178$, the traditional value. Since there is no compensation for dynamical time-scales, it seems reasonable that this prescription underpredicts collapse times. In the spherical limit, this would be a $\sim 10$ per cent effect as discussed above.

The fit (equation 29) by Sheth et al. (2001, hereafter SMT01) to the resulting linear overdensity values at collapse as a function of ellipticity and prolateness reduces to the traditional spherical collapse result when $Z = 0 = Y$, meaning that it rescales the BM96 prescription for the barrier by a factor of $1.686/1.583 = 1.06$. This would cause a corresponding $\sim 10$ per cent increase in collapse time. While this increase is of the correct magnitude in the spherical limit (see above), it is less clear whether this is also true for significantly triaxial configurations. In the EdS case where the growth factor is proportional to the scale factor, one can check that a barrier rescaling as above does, in fact, rescale the collapse time by a constant at any prolateness. However, the situation is different e.g. in flat $\Lambda$CDM with $\Omega_m < 1$. In this case, the growth factor is different from the scale factor and one can show that a barrier rescaling leads to a slightly larger effect on the collapse time for $e_\psi > 0$ than it does for $e_\psi = 0$. In other words, if one were aiming for a $\sim 10$ per cent increase in collapse time for all ellipticities, then rescaling the barrier by a constant will tend to overcompensate at large ellipticities, and would go in the direction of explaining the results of Section 5.3.

Of course, the specific prescription for virialization itself is somewhat ad hoc, and BM96 discuss several modifications, all of which lead to few per cent changes in collapse time (see also Angrick & Bartelmann 2010; Ludlow & Porciani 2011b). Additionally, the specific choice of halo finder in the simulation will also lead to (probably unquantifiable) systematics. It is therefore not unreasonable that uncertainties in dynamical modelling lead to a small (but, in the WDM case, important) systematic overprediction of collapse times.

**APPENDIX C: PEAK SHAPES AND THE COLLAPSE THRESHOLD**

Fig. C1 shows the distributions of ellipticity and prolateness defined using the tidal tensor (top panel) and using the Hessian of the density (bottom panel), coloured by the halo mass in each case. These quantities were defined in equations (17)–(20) above.

In the top panel we see that low-mass haloes have a weak preference for larger values of ellipticity $e_\psi$, while there is no trend of prolateness $p_\psi$ with mass. This is consistent with the results of Ludlow & Porciani (2011a) and Despali et al. (2013) for CDM haloes. In the bottom panel (note the difference in axes scales) we see that there is a large scatter in values of $e_\psi$ at any mass, while there is a weak trend for $p_\psi > 0$ at low masses. The latter is consistent with the BBKS results for peak shapes.

We noted earlier that the barrier (equation 29) associated with ellipsoidal dynamics does not describe the measured densities of the prot halo patches; in particular, we found no correlation between the measured proto halo density and prolateness. We have repeated the exercise of finding ESPeaks and matching them with protohaloes using the full ellipsoidal collapse barrier (equation 29). While we find that a similar fraction ($\sim 62$ per cent) of protohaloes have matching ESPeaks, the scatter in the assigned masses, especially at low masses, is somewhat larger in this case as compared to using equation (28). Fig. C2 compares the mass distributions of matched objects in these two cases.
Figure C2. Comparing ESPeak masses assigned by the algorithm described in Section 4.1 when using the simple stochastic barrier (equation 28; blue points) and when using the fully stochastic SMT01 barrier (equation 29; orange points). The points show the masses $M_{\text{halo}}$ of protohaloes that could be matched to ESPeaks, against ESPeak mass $M_{\text{walks}}$. In both cases the matched fraction was comparable (64 per cent when using equation 28 and 62 per cent when using equation 29); however, the scatter when using the SMT01 barrier is somewhat larger than when using equation (28), especially at low masses, which can be traced to the fact that the SMT01 barrier introduces a dependence of the peak-centred overdensity on prolateness which is not present in the case of the actual protohaloes.

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