

Inner cusps of the first dark matter haloes: formation and survival in a cosmological context

M. Sten Delos   and Simon D. M. White 

Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, D-85748 Garching, Germany

Accepted 2022 November 15. Received 2022 November 8; in original form 2022 July 14

ABSTRACT

We use very high resolution cosmological zoom simulations to follow the early evolution of 12 first-generation haloes formed from gaussian initial conditions with scale-free power spectra truncated on small scales by a gaussian in wavenumber. Initial collapse occurs with a diverse range of sheet- or filament-like caustic morphologies, but in almost all cases it gives rise to a numerically converged density cusp with $\rho = Ar^{-3/2}$ and total mass comparable to that of the corresponding peak in the initial linear density field. The constant A can be estimated to within about 10 per cent from the properties of this peak. This outcome agrees with earlier work on the first haloes in cold and warm dark matter universes. Within central cusps, the velocity dispersion is close to isotropic, and the equidensity surfaces tend to align with those of the main body of the halo at larger radii. As haloes grow, their cusps are often (but not always) overlaid with additional material at intermediate radii to produce profiles more similar to the Einasto or Navarro–Frenk–White forms typical of more massive haloes. Nevertheless, to the extent that we can resolve them, cusps survive at the smallest radii. Major mergers can disturb them, but the effect is quite weak in the cases that we study. The cusps extend down to the resolution limits of our simulations, which are typically a factor of several larger than the cores that would be produced by phase-space conservation if the initial power spectrum cutoff arises from free streaming.

Key words: methods: numerical – cosmology: theory – dark matter.

1 INTRODUCTION

In the idealized cold dark matter (CDM) picture, density variations are present at all scales within the initial mass distribution. These density variations seed gravitationally collapsed haloes that form hierarchically through accretion and mergers of smaller haloes. Numerical simulations of this scenario have found that the spherically averaged density profiles $\rho(r)$ of these haloes are remarkably universal as a function of the radius r . Navarro, Frenk & White (1996, 1997) showed that all CDM haloes are well described by a functional form – now called the Navarro–Frenk–White (NFW) profile – that approaches $\rho \propto r^{-1}$ at small radii and steepens towards $\rho \propto r^{-3}$ as the outer ‘virial’ radius of the halo is approached. Navarro et al. (2004) later found that the Einasto density profile (Einasto 1965), in which the logarithmic density slope varies with radius as $d \log \rho / d \log r \propto -r^{\alpha_E}$ with $\alpha_E \sim 0.17$, supplies a slightly better fit. This profile’s density slope becomes even shallower at small radii than the NFW asymptotic limit of -1 (Navarro et al. 2010).

This picture is incomplete, however. Most models of particle dark matter possess some thermal motion in the initial conditions. This motion smooths out the primordial density field on a characteristic, free-streaming scale. For example, if the dark matter is the lightest neutralino in a supersymmetric extension of the Standard Model with a mass of order 100 GeV, then the free-streaming scale corresponds to roughly an Earth mass, and the first haloes form at about this mass (e.g. Green, Hofmann & Schwarz 2005). For warm dark matter

(WDM) models, the free-streaming mass scale could be as large as $\sim 10^8 M_\odot$ (e.g. Enzi et al. 2021). The first haloes form via the direct monolithic collapse of smooth density peaks at this scale. Compared to the NFW or Einasto density profiles associated with idealized CDM haloes, the density profiles of these first haloes tend to be steeper at small radii, with slopes approaching $d \log \rho / d \log r = -3/2$ (Diemand, Moore & Stadel 2005; Ishiyama, Makino & Ebisuzaki 2010; Anderhalden & Diemand 2013; Ishiyama 2014; Polisensky & Ricotti 2015; Angulo et al. 2017; Delos et al. 2018a, b; Ogiya & Hahn 2018; Ishiyama & Ando 2020; Colombi 2021).¹

Yet the relevance of these $\rho \propto r^{-3/2}$ cusps to cosmological haloes remains uncertain. The central density slope tends to be shallower for more massive haloes (Ishiyama 2014; Angulo et al. 2017), which suggests that steep inner cusps may relax into shallower NFW or Einasto forms over time as haloes grow. Moreover, a separate body of work (Huss, Jain & Steinmetz 1999; Avila-Reese et al. 2001; Busha, Evrard & Adams 2007; Wang & White 2009; Lovell et al. 2014; Wang et al. 2020) continues to find no evidence of steep central cusps in even the smallest haloes, instead concluding that these haloes possess the same NFW or Einasto density profiles as their idealized CDM counterparts. It is also noteworthy that simulations that resolve the free-streaming scale are polluted by smaller-scale structures that arise artificially from discreteness noise (e.g. Wang & White 2007),

¹Dark matter halo density profiles that approach $\rho \propto r^{-3/2}$ at small radii were suggested earlier by Moore et al. (1999) in the context of idealized CDM, but later works (e.g. Navarro et al. 2004) found that this behaviour did not hold as simulation resolution improved.

* E-mail: sten@mpa-garching.mpg.de

an issue that could cast doubt on the numerical robustness of the density profiles in these simulations.

The present study is aimed at clarifying the extent to which $\rho \propto r^{-3/2}$ central density cusps arise and persist in a cosmological context. We carry out very high resolution zoom simulations of 12 first-generation haloes, grown from three different initial linear power spectra, and we track the structure and the density profiles of these haloes through their formation and early evolution. The details of the initial collapse differ dramatically in the 12 cases, as does the degree of artificial structuring present during the formation process. Our first objective is to test whether these diverse formation sequences all result in haloes with $\rho \propto r^{-3/2}$ in their inner regions.

Our second goal is to explore the survival of these cusps. Appealing to the results of idealized simulations, Ogiya, Nagai & Ishiyama (2016) and Angulo et al. (2017) argued that major mergers drive the transition of $\rho \propto r^{-3/2}$ cusps toward the shallower NFW or Einasto forms found in larger haloes. On the other hand, studies of CDM haloes by Dalal, Lithwick & Kuhlen (2010) and Ludlow et al. (2013), among others, found the overall mass accretion history of a halo, which is usually dominated by minor mergers and diffuse accretion rather than by major mergers (e.g. Angulo & White 2010), to be a good predictor of its density profile. Our halo sample exhibits a variety of accretion and merger histories, the range of which is enhanced by our sampling of very different initial power spectra. We explore how the later-time density profiles of our haloes vary with accretion and merger history, and we investigate how this affects their initial central cusps.

If they survive, $\rho \propto r^{-3/2}$ central cusps might have significant observational implications. Subhaloes with steeper central cusps are substantially more resistant to tidal disruption (Peñarrubia et al. 2010; Stücker et al. 2022a), so $\rho \propto r^{-3/2}$ cusps could boost the abundance of the lowest-mass subhaloes in the haloes of the Milky Way and other large galaxies. Steeper cusps are also associated directly with stronger observational signatures. For instance, prospects for the detection of dark subhaloes using annihilation radiation are enhanced because the annihilation luminosity diverges logarithmically towards $r = 0$ for a profile with $\rho \propto r^{-3/2}$ (for velocity-independent cross sections). This behaviour illustrates the importance of understanding the range of radii encompassed by the cusps of the first haloes. Our final aim is thus to explore this question. Our simulations can determine the maximum radii of the cusps, but they do not resolve the minimum radius. However, if the initial density field was smoothed by free streaming of the dark matter particles, then these same motions must tame the central cusp by imposing a core at the maximum phase-space density allowed by Liouville's theorem (Tremaine & Gunn 1979). We study the phase-space structure of the inner cusp analytically to determine the radius at which it must give way to a core.

Our main results support the robustness of $\rho \propto r^{-3/2}$ cusps in the first haloes. We find that such cusps arise for haloes with wildly different formation sequences and are well converged with respect to simulation resolution. As our haloes grow, rapid accretion is the main effect leading to shallower density profiles, but this process works by building up the density at intermediate radii without substantially altering the inner cusp. Consequently, haloes can develop apparently shallow Einasto-like density profiles at large and intermediate radii while retaining the initial $\rho \propto r^{-3/2}$ cusp at the smallest radii. Major merger events can disturb this central cusp, but their impact is relatively minor in the cases that we have studied. Finally, the radius at which phase-space constraints force a cusp to give way to a core is sufficiently small in realistic models that the $\rho \propto r^{-3/2}$ cusp can span several decades in radius. These findings suggest that $\rho \propto r^{-3/2}$

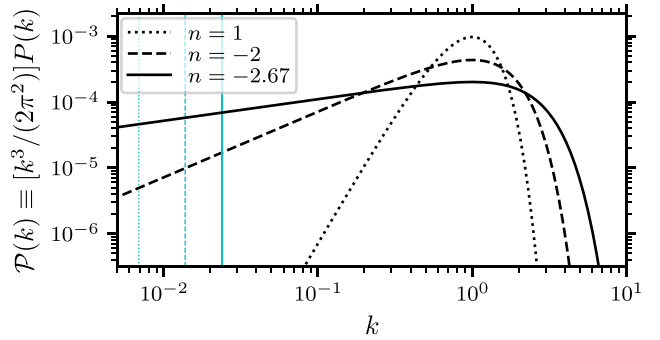


Figure 1. The power spectra from which we draw our simulation initial conditions. We show the dimensionless form $\mathcal{P}(k) \equiv [k^3/(2\pi^2)]P(k)$. The length units are fixed such that $\langle k^2 \rangle = 1$, so the gaussian cutoff scale is $k_{\text{cut}} = \sqrt{2/(3+n)}$. These spectra are normalized so that the rms density contrast is $\sigma = 0.03$. The vertical cyan lines indicate the fundamental frequency associated with the box employed for simulations with each power spectrum.

central cusps may survive long enough to be observationally relevant, although studies involving larger halo samples and longer time spans (and hence larger growth factors) are needed to be certain.

This article is organized as follows. In Section 2, we discuss our simulation setup and the properties of our simulated haloes. In Section 3, we explore halo formation mechanisms and the density cusps that emerge. In Section 4, we explore how and why halo density profiles evolve over time, and we study the extent to which the $\rho \propto r^{-3/2}$ cusps survive this process. In Section 5, we use analytic arguments to evaluate the radius at which an inner density cusp must give way to a core with finite central density. We summarize our results in Section 6. Finally, Appendices A and B validate the numerical robustness of our simulation results, while Appendix C presents mathematical details relating to the calculation of the core size in Section 5.

2 SIMULATION SETUP

2.1 Power spectra and units

To explore scenarios in which density variations are cut off at small scale, we consider Einstein–de Sitter cosmologies with initial power spectra taking the form

$$P(k) \propto k^n \exp[-(k/k_{\text{cut}})^2], \quad (1)$$

i.e. power laws multiplied by a gaussian of scale $k = k_{\text{cut}}$.² Specifically, we consider the three power-law indices $n = 1$, $n = -2$, and $n = -2.67$. These power spectra are depicted in Fig. 1; we show the dimensionless form $\mathcal{P}(k) \equiv [k^3/(2\pi^2)]P(k)$, indicating the typical (squared) relative amplitude of density variations as a function of their scale. Generally, power spectra with n close to -3 are associated with the presence of significant density variations over a broad range of scales. In contrast, the $n = 1$ power spectrum has significant support only within a narrow range. In standard cosmological scenarios, CDM is associated with $n \sim -3$ at scales close to the cutoff while warmer dark matter is associated with n farther from -3 . For instance, we will see later that the power spectrum associated with 3.5-keV WDM (e.g. Bode, Ostriker & Turok 2001) lies somewhere between the $n = -2.67$ and $n = -2$

²A cutoff of exactly this form arises from free streaming if the thermal velocity distribution is Maxwell–Boltzmann (e.g. Bertschinger 2006).

Table 1. The units that we employ in this study. We specify comoving units here; for physical units the scale factor a is defined such that $a = \sigma$ in linear theory.

Quantity	Unit
Length	$\langle k^2 \rangle^{-1/2}$
Density	$\bar{\rho}$
Mass	$\bar{\rho} \langle k^2 \rangle^{-3/2}$

power laws near the cutoff. Substantially larger power-law indices are not observationally viable in this context, although they arose in the original hot dark matter models (e.g. White, Frenk & Davis 1983) and are also found for CDM for some non-standard cosmological initial condition scenarios (e.g. Erickcek & Sigurdson 2011). In this paper, we do not aim to reproduce any particular cosmology but rather to explore the formation of first-generation objects using idealized initial power spectra.

The power spectra in Fig. 1 are normalized so that the (unfiltered) rms density contrast is $\sigma = 0.03$, where

$$\sigma^2 = \int_0^\infty \frac{dk}{k} \mathcal{P}(k). \quad (2)$$

We also fix our length units so that $\langle k^2 \rangle = 1$, where

$$\langle k^2 \rangle = \sigma^{-2} \int_0^\infty \frac{dk}{k} k^2 \mathcal{P}(k), \quad (3)$$

which leads to $k_{\text{cut}} = \sqrt{2/(3+n)}$ and puts the peak of \mathcal{P} at $k = 1$. The choice to fix σ and $\langle k^2 \rangle$ is convenient because it means that density peaks arising from different power spectra have comparable amplitudes and sizes. Within the linear density field $\delta(\mathbf{x})$, peak amplitudes δ are proportional to σ , while peak sizes, characterized by $|\delta/\nabla^2\delta|^{1/2}$, cluster around $\langle k^2 \rangle^{-1/2}$ in the high-peak limit (Bardeen et al. 1986). As we will see in Section 3.2, the inner structures of the first haloes are sensitive almost entirely to these quantities.³

In our simulations we will take the initial scale factor to be $a = 0.03$, so that $a = \sigma$ in linear theory. We also define mass units such that the mean comoving cosmological density is 1. The units and conventions are summarized in Table 1.

2.2 Simulation procedure

We use the GADGET-4 simulation code (Springel et al. 2021) to execute our simulations. First, we carry out one primary simulation for each of the three power spectra in Fig. 1. In each case we use second-order Lagrangian perturbation theory to produce a perturbed grid of 1024^3 particles inside a periodic box of width $643k_{\text{cut}}^{-1}$. This choice means that the Nyquist frequency associated with the initial particle grid is $5k_{\text{cut}}$. We then simulate each box up to some scale factor $a > 2$ under the assumption of matter domination. The simulation force softening length here is set to 0.03 times the initial

³If we were interested in halo *counts* instead of structures, the more natural length scale would instead be $(k^4)^{-1/2} \langle k^2 \rangle^{1/2} = \sqrt{2/(5+n)} k_{\text{cut}}^{-1}$, which sets the number density of peaks (Bardeen et al. 1986). On the other hand, the ‘half-mode’ scale commonly considered in studies of WDM (e.g. Enzi et al. 2021), where the cutoff suppresses the amplitudes of density variations by half, has no clear meaning for haloes. The half-mode mass in our units is $M_{\text{hm}} = 1800$, $M_{\text{hm}} = 225$, and $M_{\text{hm}} = 42.7$ for the equation (1) power spectrum with $n = 1$, $n = -2$, and $n = -2.67$, respectively; compare to Fig. 3.

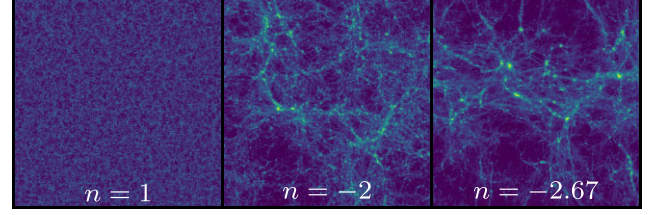


Figure 2. Projected density field within a representative region of each simulation at $a = 2$ (when the rms density contrast $\sigma = 2$ in linear theory). The region is a cube of width $261 \langle k^2 \rangle^{-1/2}$; for $n = -2.67$ this is the entire box. The vastly higher amplitude of large-scale power in the $n = -2$ and $n = -2.67$ simulations, compared to the $n = 1$ simulation, is immediately evident. The latter produces an essentially uniform distribution of haloes, while the former produce structures that approach the size of the displayed region.

interparticle spacing. A portion of each resulting simulation box at $a = 2$ is shown in Fig. 2; in particular we show a cube of width $261 \langle k^2 \rangle^{-1/2}$, which for $n = -2.67$ is the entire box. The imprint of the power spectrum is evident here. The $n = 1$ simulation’s dearth of large-scale power yields an essentially uniform distribution of tiny haloes, whereas the substantial amplitude of large-scale power in the $n \leq -2$ simulations produces much larger haloes and filaments; in the $n = -2.67$ case these structures begin to approach the box size.

From each of these primary simulations, we select a range of haloes to resimulate at much higher resolution. The selection is mostly arbitrary, but we favor haloes that formed by $a \sim 0.6$. We aim to ensure that there is a reasonable sense in which each halo initially formed by direct collapse. To boost the resolution, we identify each low-resolution particle that resides within the halo’s spherical-overdensity radius R_{178} (enclosing an average density 178 times the cosmological mean) at some late time. We then locate that particle within the initial conditions and replace it with a glass consisting of between 4.0×10^3 and 2.6×10^5 high-resolution particles (depending on the halo), whose centre of mass lies at the original particle’s location. We also remove low-resolution ‘holes’ in the high-resolution patch by iteratively ensuring that for any particle selected to be replaced by a glass, we also select its nearest neighbor along the cardinal direction that most closely points toward the high-resolution region’s centre of mass.

For each such high-resolution halo, we then carry out another simulation that includes that halo’s high-resolution particles along with the remaining original low-resolution particles (which retain their original mass and force-softening length). Note that due to the short duration of our simulations and our focus on inner density profiles, it is not necessary to ensure that low-resolution particles never enter the high-resolution halo. We discuss in Appendix A the extent to which low-resolution particles do enter our haloes, showing that they remain at radii too large to affect our conclusions.

The numerical parameters associated with the high-resolution halo simulations are listed in Table 2. Since the force-softening length represents a major resolution bottleneck (see Appendix B), we set the softening length of the high-resolution particles to be only about 0.016 their initial spacing. We simulate 12 haloes in this way, which we name H1–H3 (from the $n = 1$ power spectrum), W1–W5 (from the $n = -2$ power spectrum), and C1–C4 (from the $n = -2.67$ power spectrum). The mass accretion histories of these haloes are depicted in Fig. 3 over the period that we simulate them at high resolution. Note that this figure already demonstrates the value of our choice of units; by fixing the characteristic sizes $|\delta/\nabla^2\delta|^{1/2} \sim 1$ of peaks in the linear density field, we ensure that all haloes have masses of order unity at the time they collapse.

Table 2. Numerical parameters associated with our high-resolution halo simulations. For each halo (and associated power spectral index n), we list the number N_{high} of high-resolution particles, their mass m_p , the softening length ϵ , and the final expansion factor and halo mass, a_f and M_f .

n	halo	N_{high}	m_p	ϵ	a_f	M_f
1	H1	2.4×10^8	2.1×10^{-5}	4.3×10^{-4}	4.72	1.4×10^3
1	H2	1.5×10^8	2.1×10^{-5}	4.3×10^{-4}	5.06	2.1×10^3
1	H3	5.6×10^7	2.1×10^{-5}	4.3×10^{-4}	6.51	9.5×10^2
-2	W1	3.0×10^8	2.1×10^{-5}	4.3×10^{-4}	1.50	3.1×10^3
-2	W2	1.9×10^8	2.1×10^{-5}	4.3×10^{-4}	1.28	1.8×10^3
-2	W3	1.4×10^8	2.1×10^{-5}	4.3×10^{-4}	1.69	9.8×10^2
-2	W4	1.6×10^8	2.1×10^{-5}	4.3×10^{-4}	1.68	1.4×10^3
-2	W5	5.9×10^8	3.3×10^{-7}	1.1×10^{-4}	1.52	1.0×10^2
-2.67	C1	8.6×10^8	4.1×10^{-6}	2.5×10^{-4}	1.00	2.3×10^3
-2.67	C2	5.6×10^8	4.1×10^{-6}	2.5×10^{-4}	1.19	2.0×10^3
-2.67	C3	2.2×10^8	4.1×10^{-6}	2.5×10^{-4}	1.03	8.4×10^2
-2.67	C4	4.7×10^8	5.1×10^{-7}	1.2×10^{-4}	1.41	1.6×10^2

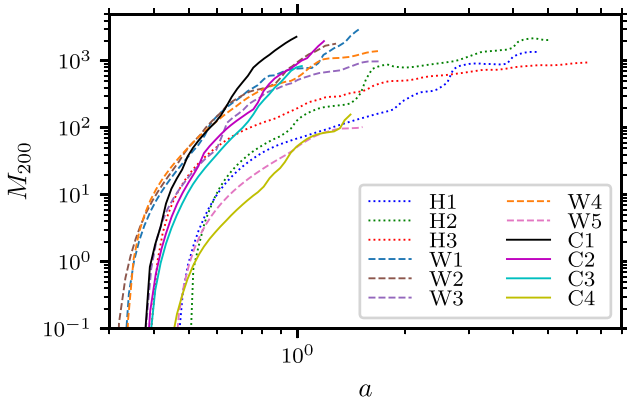


Figure 3. The spherical-overdensity mass M_{200} (within a spherical region enclosing a mean density 200 times the cosmic average) of each halo that we study, as a function of the scale factor a . The curve associated with each halo terminates where we end the simulation.

In later sections, we plot halo density profiles down to a minimum radius that is determined by the simulation particle count and the force softening length ϵ . Specifically, we consider the two radii r_{rel} , at which the age of the universe is equal to the two-body relaxation time-scale t_{rel} (equation B1), and $r_{\text{soft}} = 5\epsilon$. We then plot density profiles only at radii larger than both r_{rel} and r_{soft} . We test a density profile for numerical convergence in Appendix B and find that these choices determine reasonably well the minimum radius that is converged with respect to simulation resolution.

3 FORMATION OF THE INNER CUSP

We now explore the formation of the central cusp during the initial collapse of the high-resolution haloes presented above. A major goal of this section is to show that a wide variety of collapse sequences all give rise to the same power-law $\rho \propto r^{-3/2}$ inner density profile. We also discuss the connection between this feature and the initial density field.

3.1 The formation process

We first focus on haloes arising from the $n = 1$ power spectrum. Fig. 4 shows the density profile of the H1 halo during and shortly after its formation. This halo forms around $a \simeq 0.6$, and its inner density

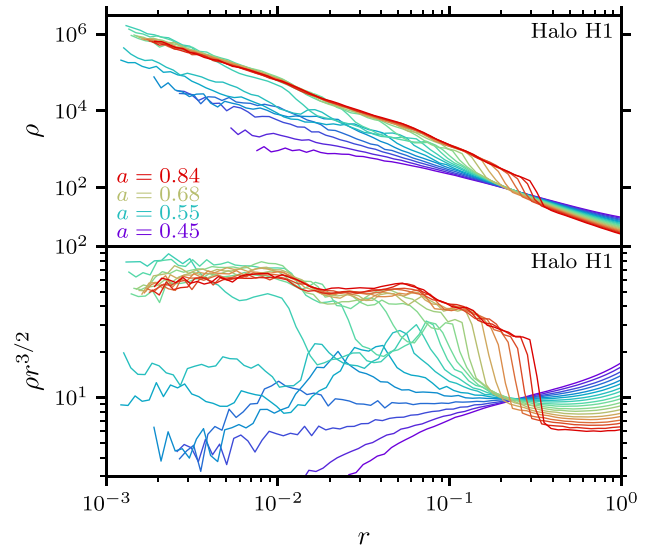


Figure 4. Density profile of halo H1 during its collapse. The upper panel shows density ρ as a function of radius r , while the lower panel scales the vertical axis by $r^{3/2}$. We plot physical, not comoving, quantities. Different colours represent different times. The density profile begins to stabilize around $a = 0.55$, after which it is essentially fixed and only grows outward in radius. The power law of the inner profile is about $\rho \propto r^{-3/2}$ (horizontal on the lower panel).

profile immediately settles at $\rho \simeq 60r^{-3/2}$ (in physical coordinates). From this point on, the density profile essentially remains stable at small radii, only growing outward due to new accretion. This picture closely resembles that in Delos, Bruff & Erickcek (2019), even with the present study's significantly improved spatial resolution: for $a \gtrsim 0.65$, there are 3×10^5 simulation particles at the radii $r < 0.06$ corresponding to the $\rho \propto r^{-3/2}$ part of the density profile.

Fig. 5 shows the halo's surrounding density field during its formation. We show three orthogonal projections at each time, and to clarify the picture, we align these projections with the principal axes of the tidal tensor at the halo's Lagrangian position within the initial density field. Evidently, the halo's precursor overdensity first collapses into a one-dimensional filament (second column in Fig. 5) aligned in the z direction with very little extent in either the x - or the y -direction. This filament begins to break up into regular fragments, easily visible in the third column, which are a well-known

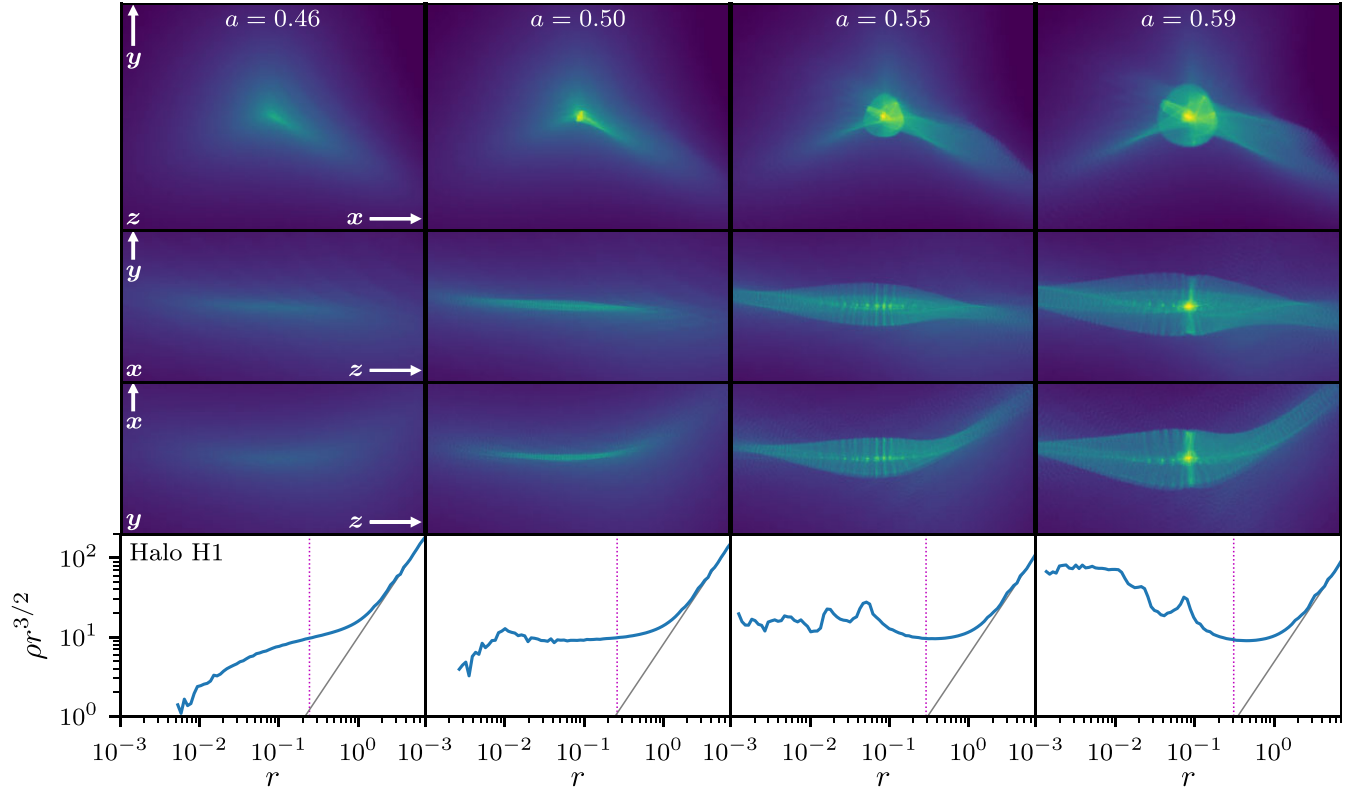


Figure 5. Formation of the halo H1. We plot three different orthogonal projections of the density field (upper three rows) at four different times (columns). These projections are aligned with the principal axes of the initial tidal tensor. Material initially collapses to form a filament (first three columns), which is evident because two projections show a linearly extended object while the third suggests a point-like object. Within this filament the collapsed halo eventually forms (last column). Fragmentation of the filament visible in the third column is a numerical artefact. In the bottom row, we show the spherically averaged density profile corresponding to the pictures above. Evidently, by $a = 0.59$ (last column) the halo’s stable $\rho \simeq 60r^{-3/2}$ density profile is already present. Within the lower panels, the vertical line marks half the width of the displayed box (i.e. the radius from the centre to the left or right edge), while the faint grey diagonal line indicates the cosmological mean density. In the upper panels of the left column we include directional references to clarify the relationship between the three projections; for instance the upper panels are projected in the z -direction, and the x - and y -directions therein are indicated by arrows.

numerical artefact in simulations of this type (e.g. Wang & White 2007; Angulo, Hahn & Abel 2013; Lovell et al. 2014). Finally, a collapsed halo forms near the densest point of the filament. For convenience we also plot the spherically averaged density profile at each time. By $a = 0.59$ the inner $\rho \simeq 60r^{-3/2}$ cusp is already present. The halo’s outer density caustic is clearly visible in the profile at this time around $r \simeq 2 \times 10^{-2}$, and the filament’s outer caustic can also be seen around $r \simeq 8 \times 10^{-2}$ (the latter will soon be swallowed by the growing halo; see Fig. 4).

In Fig. 5, the inner cusp could be interpreted as forming from a collection of artificial filament fragments, a prospect that might cast doubt on the validity of the central profile of the resulting halo. These fragments arise because the discrete simulation particles that comprise the filament tend to cluster (a process that may be viewed as the triggering of a one-dimensional Jeans instability by discreteness noise).⁴ However, two lines of evidence suggest that the $\rho \propto r^{-3/2}$ cusp is robust.

(i) In Appendix B, we explicitly test the density profile of the W5 halo for numerical convergence. This halo also exhibits a clear $\rho \propto r^{-3/2}$ power-law density profile at small radii, and its profile is

stable with respect to changes in the simulation’s spatial and mass resolution. In contrast, the mass and number of artificial filament fragments scale with a simulation’s linear resolution (Wang & White 2007).

(ii) As we will see next, a halo that collapses from a two-dimensional sheet (instead of a one-dimensional filament) also develops a clear $\rho \propto r^{-3/2}$ density cusp, even though the sheet exhibits far weaker artificial structuring.

Fig. 6 depicts the density field around the H3 halo during its collapse. We show again the field projected along the tidal tensor’s principal axes. The material in this case collapses into a sheet (first two columns), as indicated by two of the projections showing a linearly extended object while the third shows no apparent collapsed object at all.⁵ In the second column artificial small-scale structure is visible in the sheet and is amplified as it collapses along a second axis, as seen in the third column. It is, however, far weaker than the fragmentation in Fig. 5. Finally, the halo forms in the rightmost column. The lower panels show the spherically averaged density profile, and by $a = 0.45$ (rightmost column), the halo has developed a $\rho \simeq 90r^{-3/2}$ inner density cusp.

⁴Simulation methods exist that can eliminate these artefacts (Stücker et al. 2020, 2022b), although we do not employ them here.

⁵The principal-axis projections are particularly valuable in the case of sheet collapse, as most projection angles would show no clear evidence of collapse.

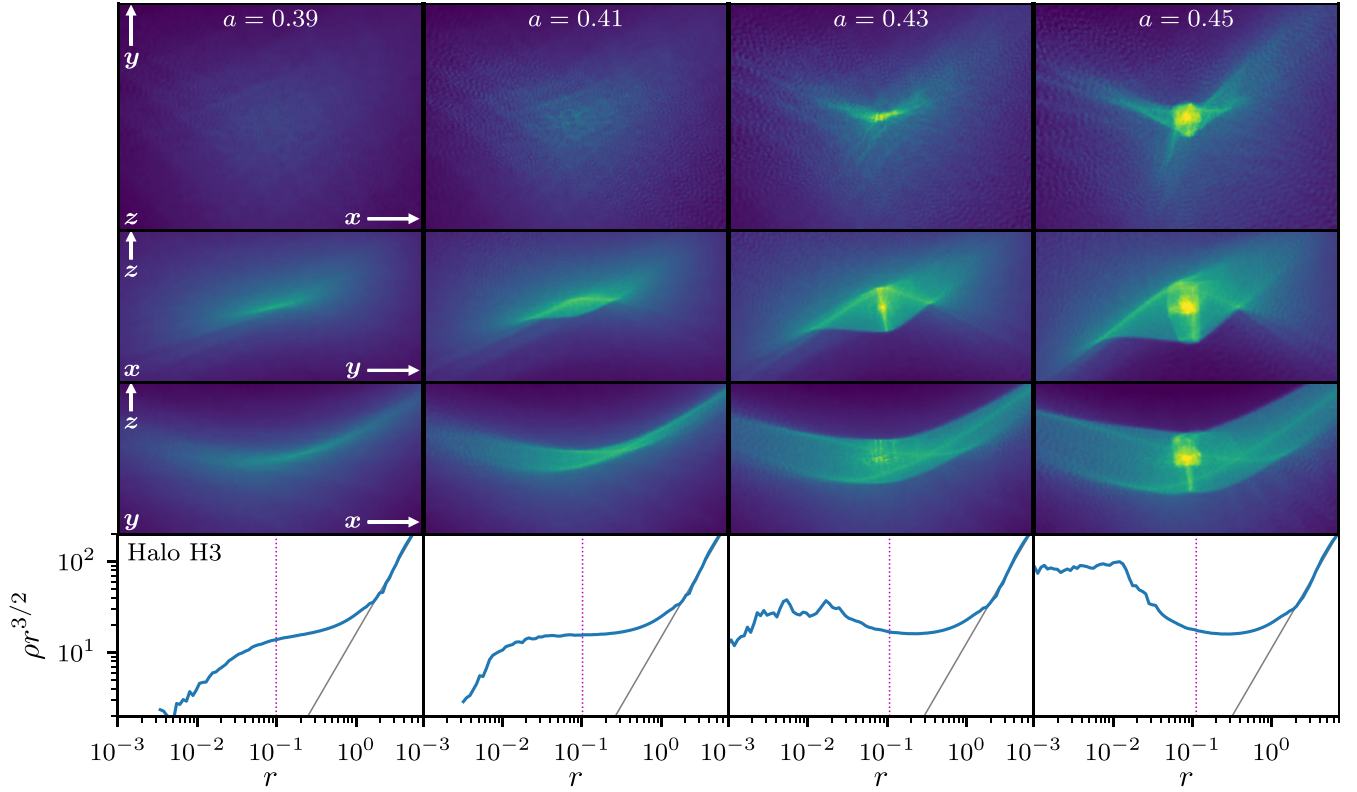


Figure 6. Formation of the halo H3. As in Fig. 5, we plot the density field projected along the principal axes of the initial tidal tensor (upper three panels) along with the spherically averaged density profile (lower panels) at four different times (columns); see that figure’s caption for further explanation. The material in this case collapses into a sheet (first two columns), which is evidenced by a linearly extended object visible along two of the principal axes and no collapsed object visible along the third. Artificial small-scale structure is visible in the sheet in the top image of the second and third columns and in the lower image (‘y’ projection) of the third column, but it is far less pronounced than the filament fragmentation seen in Fig. 5. The halo’s stable inner $\rho \simeq 90r^{-3/2}$ density profile has already developed by $a = 0.45$ (last column).

So far we have only studied haloes arising from the $n = 1$ power spectrum. We now show examples of haloes arising from each of the $n = -2$ and $n = -2.67$ power spectra. Figs 7 and 8 show the density fields surrounding haloes W1 and C1, respectively, during their collapse. Each exhibits its own unique behaviour. The material of the W1 halo (Fig. 7) initially collapses into a filament, but the halo forms at the end of this filament, so unlike H1 (Fig. 5), it is not clear that W1 can be viewed as arising from the collapse of the filament itself. Meanwhile, the material of the C1 halo (Fig. 8) exhibits a range of behaviour during its collapse. It first collapses into a very pronounced sheet; this sheet subsequently collapses into a filament followed almost immediately by formation of the halo, which promptly accretes another collapsed filament edge-on. Despite these variations, both haloes develop an inner density profile close to $\rho \simeq 80r^{-3/2}$ by the final snapshot shown.

Fig. 9 shows density profiles during the collapse phase and shortly afterwards for all 12 of our high-resolution haloes. All of these haloes develop density profiles with $\rho \sim r^{-3/2}$ that remain effectively stable over a factor of at least 1.3 in scale factor, extending to larger radii as the haloes accrete new material. Some of these density profiles, such as those of C1 and C2, do grow moderately in amplitude over time, an effect we explore further in Section 4. The density profiles of our highest-resolution haloes – W5 and C4 – show some tendency toward shallowing at the smallest radii. However, it is unclear whether this behaviour is converged with respect to simulation resolution, as we discuss in Appendix B, and studies with yet higher resolution are

needed to test this. Aside from the uncertainty in these two cases, the $\rho \propto r^{-3/2}$ density cusp appears to be a robust consequence of the direct collapse of a smooth initial density peak, independent of the details of the collapse.

3.2 Connection to the precursor density field

Despite the varied details of the collapse process, the central cusps of the first haloes develop from particularly simple initial conditions: the inner regions of smooth peaks in the primordial density field. Delos et al. (2019) noted that since the $\rho \propto r^{-3/2}$ cusp stabilizes so quickly after collapse, its properties must be sensitive only to the immediate neighborhood of the precursor density peak. Such a peak can be locally characterized by its height δ and its characteristic comoving radius $R \equiv |\delta/\nabla^2\delta|^{1/2}$. Its collapse dynamics are also sensitive to the tidal tensor $\partial_i\partial_j\phi$ (where ϕ is the peculiar potential), which can be described by its ellipticity e and prolateness p . Specifically, if $\lambda_1 > \lambda_2 > \lambda_3$ are the eigenvalues of $-\partial_i\partial_j\phi$, then $e \equiv \frac{\lambda_1 - \lambda_3}{2(\lambda_1 + \lambda_2 + \lambda_3)}$ and $p \equiv \frac{\lambda_1 + \lambda_3 - 2\lambda_2}{2(\lambda_1 + \lambda_2 + \lambda_3)}$. Note that since $\lambda_1 + \lambda_2 + \lambda_3 \propto \delta$, the three quantities δ , e , and p fully specify the tidal tensor up to rotation.

From dimensional considerations, we expect that the amplitude A of the $\rho = Ar^{-3/2}$ cusp should be proportional to $\bar{\rho}(a_c)(a_c R)^{3/2}$, where a_c is the expansion factor at which the peak collapses and $\bar{\rho}(a_c) = \bar{\rho}a_c^{-3}$ is the cosmological density at a_c . As above, $\bar{\rho}$ is the comoving cosmological density (equal to 1 in our simulations). Consequently,

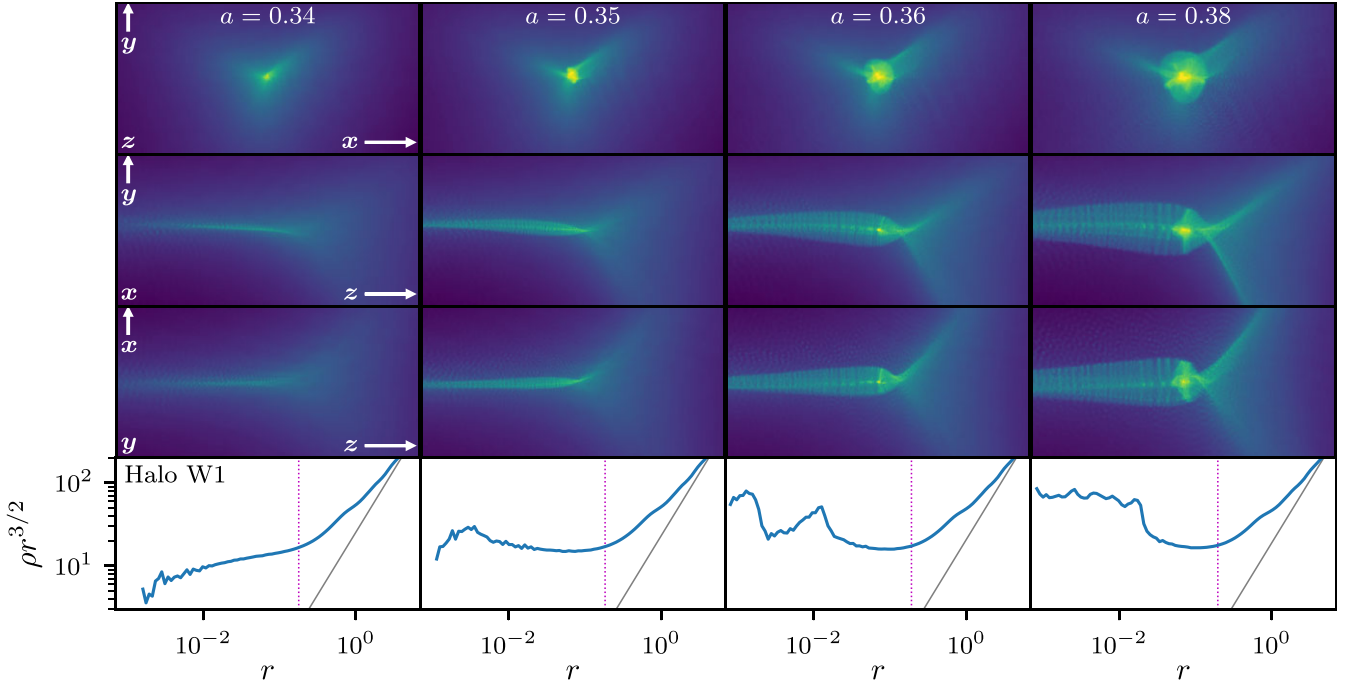


Figure 7. Formation of the halo W1. We plot the density field projected along the principal axes of the initial tidal tensor (upper three rows of panels) along with the spherically averaged density profile (lower panels) at four different times (columns); see the caption of Fig. 5 for further explanation. The halo’s material initially collapses into a filament (first two columns), similarly to the case of the H1 halo (Fig. 5). Unlike H1, however, W1 forms at the end of its filament (third column). Moreover, while artificial fragmentation is visible at this time, the halo has clearly already formed, as evidenced by comparing the density profiles in the third and fourth columns at the innermost resolved radii. By $a = 0.38$ (last column) a $\rho \simeq 70r^{-3/2}$ density profile is already well developed.

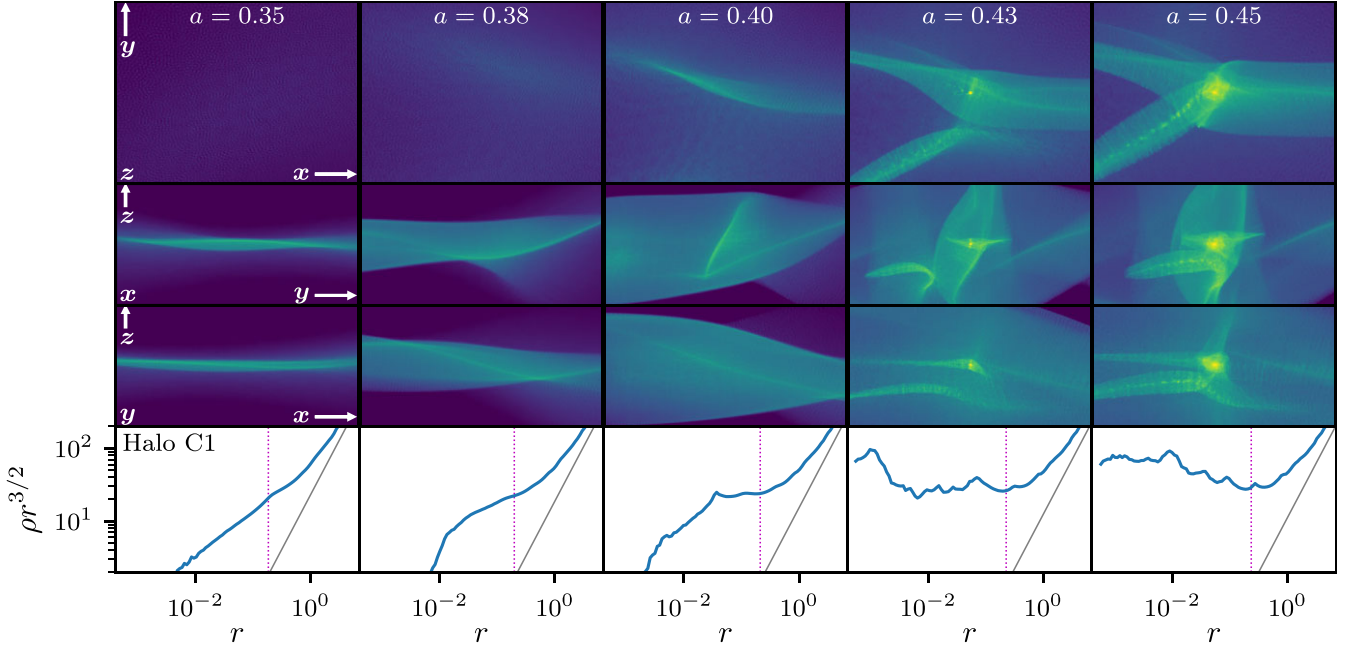


Figure 8. Formation of the halo C1. We plot the density field projected along the principal axes of the initial tidal tensor (upper three rows of panels) along with the spherically averaged density profile (lower panels) at five different times (columns); see the caption of Fig. 5 for further explanation. This halo’s material initially collapses into a sheet (first column), which is far more pronounced than the sheet associated with the H3 halo (Fig. 6). This sheet soon collapses into a filament (third column) followed almost immediately by the halo (fourth column). The halo’s $\rho \simeq 70r^{-3/2}$ density profile is well developed by $a = 0.45$ (last column); meanwhile, another filament is about to accrete edge-on, although its accretion will not change the density profile significantly (see Fig. 9).

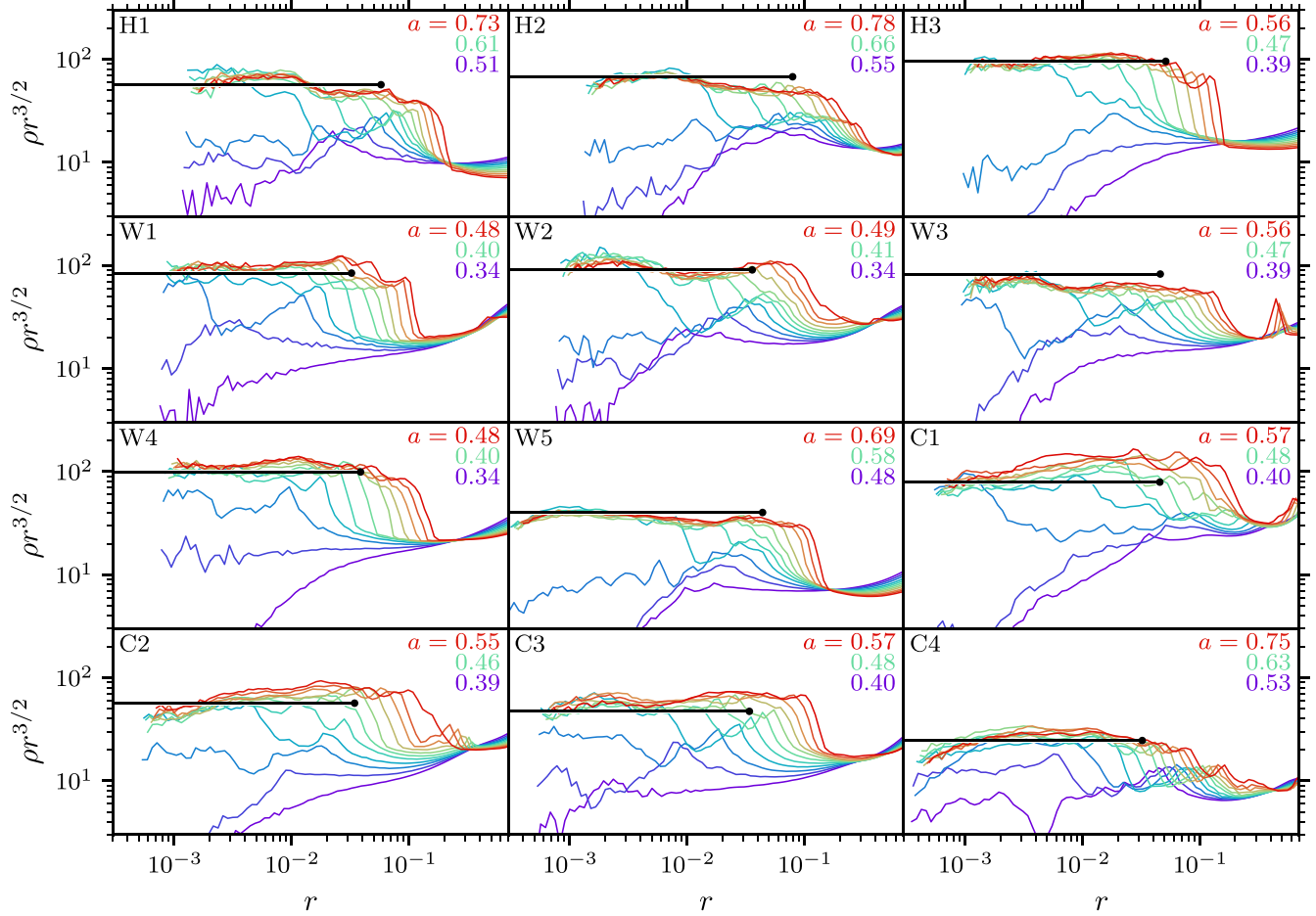


Figure 9. Density profiles of all of our high-resolution haloes during their collapse. As in Fig. 4, the vertical axis is scaled by $r^{3/2}$; we plot physical, not comoving, quantities; and different colours represent different times. The haloes all develop $\rho \propto r^{-3/2}$ density cusps very quickly after they collapse. In some cases there is a tendency for the profile to shallow at the smallest radii, but it is unclear whether this behaviour is numerically converged (see Appendix B). The horizontal black line indicates the cusp amplitude $A = \rho r^{3/2}$ predicted from each halo’s precursor density peak using equation (4) with $\alpha = 24$. This line terminates at the estimated initial cusp radius r_0 predicted from the same peak data using equation (6). There is some scatter, but overall these predictions match well the actual properties of the cusps.

we may write

$$A = \alpha \bar{\rho} a_c^{-3/2} R^{3/2} \quad (4)$$

for some universal proportionality constant α . Here we evaluate a_c as a function of δ , e , and p using the ellipsoidal collapse approximation in Sheth, Mo & Tormen (2001). This approximate a_c lies within about 5 percent of the scale factor at which the $\rho \propto r^{-3/2}$ cusp first appears in each simulation, and using the latter scale factor to predict A does not appreciably alter the degree to which it matches simulation results.

We can also estimate the size of the $\rho \propto r^{-3/2}$ central cusp shortly after collapse. Close to a peak in the linear density field, the mean density contrast within comoving radius q scales as $1 - (q/R)^2/10$.⁶ Spherical shells at radii $q > 0$ can be approximated as undergoing ellipsoidal collapse within the same tidal field with this lower enclosed density contrast. Now consider the shell that collapses at $a = 1.17a_c$. A factor of 1.17 in a corresponds to one dynamical time, in the sense of Binney & Tremaine (2008), for an

⁶This is the mean overdensity *within* q ; the mean overdensity *at* q scales as $1 - (q/R)^2/6$.

object whose density is 200 times the cosmological mean. This shell has Lagrangian radius q satisfying $1 - (q/R)^2/10 = 1.17^{-1}$, so it encloses the mass

$$M_0 = \frac{4\pi}{3} q^3 \bar{\rho} = \frac{4\pi}{3} 10^{3/2} \left(1 - \frac{1}{1.17}\right)^{3/2} R^3 \bar{\rho} \simeq 7.3 R^3 \bar{\rho}. \quad (5)$$

The collapsed halo therefore achieves a mass of approximately M_0 after one dynamical time interval. If its density profile is $\rho = Ar^{-3/2}$ out to the physical radius r_0 that encloses M_0 , then

$$r_0 = \frac{1}{4} \left(\frac{3}{\pi}\right)^{2/3} \left(\frac{M_0}{A}\right)^{2/3} \simeq 0.92 \alpha^{-2/3} a_c R. \quad (6)$$

If the $\rho \propto r^{-3/2}$ cusp is established over roughly the first dynamical time interval after collapse, then r_0 may be interpreted as an estimate of the cusp’s initial radius.

Within each of our haloes, we identify all of the particles that lie at a radius $r < 4.2 \times 10^{-3}$ at some time shortly after halo formation (the specific time makes no difference), and we locate the Lagrangian centre of mass of these particles in the initial conditions. We then follow the local density gradient to find the nearest peak in the initial density field. We record the peak height δ and use Fourier methods to evaluate R , e , and p . We use equations (4), (5), and (6) to evaluate

Table 3. Properties of our haloes’ precursor peaks in the linear density field $\delta(\mathbf{x})$. The ‘offset’ is the comoving distance between the central cusp’s mean Lagrangian position and the density peak. $R \equiv |\delta/\nabla^2\delta|$ is the characteristic size of the peak and $\delta(a)/a$ is its height (scaled by a^{-1} since $\delta(a) \propto a$). e and p are the ‘ellipticity’ and ‘prolateness’ of the tidal tensor $\partial_i\partial_j\phi$ at the location of the peak (see the text). Finally, a_c is the peak’s predicted collapse time (using the approximation in Sheth et al. 2001), A is the resulting halo’s predicted cusp amplitude (equation 4 with $\alpha = 24$), and M_0 and r_0 are the predicted mass and radius of the cusp after one dynamical time interval, respectively (equations 5 and 6).

Halo	Offset	R	$\frac{\delta(a)}{a}$	e	p	a_c	A	M_0	r_0
H1	0.31	0.96	3.40	0.09	-0.06	0.54	57	6.6	0.058
H2	0.54	1.20	3.32	0.18	0.00	0.60	68	13	0.079
H3	0.31	1.08	4.09	0.06	0.00	0.43	96	9.3	0.051
W1	0.25	0.82	4.86	0.04	-0.02	0.36	84	4.1	0.032
W2	0.19	0.89	5.00	0.10	-0.01	0.37	91	5.2	0.036
W3	0.00	0.98	4.20	0.09	0.01	0.43	83	6.8	0.046
W4	0.11	0.95	4.77	0.06	-0.01	0.37	99	6.3	0.039
W5	0.11	0.75	3.48	0.12	-0.01	0.53	40	3.1	0.044
C1	0.24	0.96	4.30	0.12	-0.00	0.43	79	6.5	0.046
C2	0.11	0.74	4.22	0.07	0.01	0.42	57	3.0	0.034
C3	0.00	0.70	4.11	0.10	-0.02	0.44	47	2.5	0.034
C4	0.00	0.55	3.66	0.15	-0.07	0.54	25	1.2	0.032

the predicted values of A , M_0 , and r_0 , respectively, for the associated halo. We assume the proportionality constant $\alpha = 24$.⁷ Table 3 lists these data. In Fig. 9, we mark the predicted A for each halo with a horizontal black line, and we terminate that line with a point at the radius r_0 . We find that with $\alpha = 24$, the predicted values of A match the haloes’ density profiles mostly very well, albeit with minor scatter. The radius r_0 also reasonably estimates the initial size of the $\rho \propto r^{-3/2}$ cusp, although this size cannot be precisely defined.

We also list in Table 3 the offset between the halo cusp’s Lagrangian centre and the density peak. In the ideal case where only δ , R , e , and p determine collapse, this offset should be zero because the halo should form precisely at the density peak. The presence of an offset thus implies that the broader density field around the peak does have a small influence on the collapse process (beyond determining the tidal tensor). The influence of the broader density field could also explain some of the scatter in the cusp amplitudes $\rho r^{3/2}$ in Fig. 9 with respect to the predicted values A from equation (4). However, we note that since this offset is always significantly smaller than R , there is no ambiguity as to whether a halo actually arose from the peak to which we associate it. We also tried evaluating the predicted value of A using the values of δ , R , e , and p at the location of the Lagrangian centre instead of the peak, but the outcome did not change significantly.

4 PERSISTENCE OF THE INNER CUSP

We saw in Section 3 that a halo’s $\rho \propto r^{-3/2}$ cusp forms immediately after collapse. In this section, we explore the extent to which these cusps persist over cosmic time. Fig. 10 shows the density profile of every halo at the end of its simulation (see Fig. 3). Specifically, we average the density profile over the final factor of 1.17 in the scale

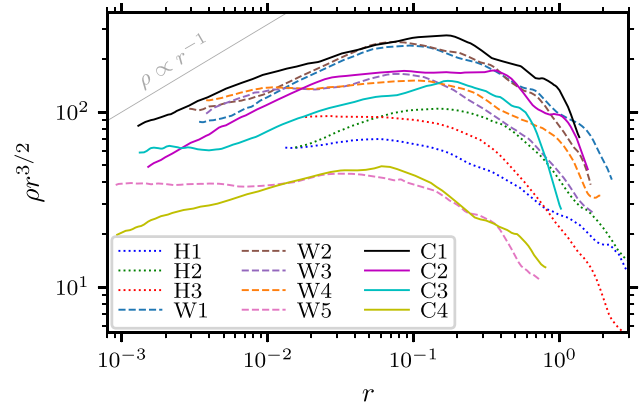


Figure 10. Density profiles of all 12 high-resolution haloes at the end of their simulations (averaged over the final factor of 1.17 in scale factor). We plot these profiles out to the radius R_{200} enclosing an average density 200 times the cosmological mean. While haloes H1, H3, and W5 still retain their $\rho \propto r^{-3/2}$ inner density profiles, most of the profiles have shallowed by this late time. We include a reference $\rho \propto r^{-1}$ line, corresponding to the innermost slope of the NFW profile. Some of the density profiles are beginning to approach this slope.

factor (corresponding again to one dynamical time for something 200 times denser than the cosmological mean). Some of these haloes – H1, H3, and W5 – evidently maintain their $\rho \propto r^{-3/2}$ density cusps. However, most of the late-time density profiles exhibit some degree of shallowing, with some, like C2, even approaching the $\rho \propto r^{-1}$ asymptote of the NFW form.

4.1 The impact of accretion rate

The shallowing of the inner profile can largely be understood as a consequence of high accretion rates. Due to its high energy and (typically) high angular momentum, newly accreted material tends to contribute significant density to a halo only at large and intermediate radii, close to the orbital apocentres of the newly accreted particles.⁸ Consequently, rapid accretion can shallow a halo’s profile by boosting the density at intermediate radii.

Fig. 11 shows how this process plays out. We plot halo mass (upper panel) and mass accretion rate $\Gamma \equiv \Delta \log M_{200} / \Delta \log a$ (centre panel; e.g. Diemer & Kravtsov 2014) for W1 as a function of scale factor a , where both $\Delta \log M_{200}$ and $\Delta \log a$ are evaluated over the preceding factor of 1.17 in a . The same figure also shows (lower panel) the evolution of the halo density profile over the same period; specifically we time average the profile over successive factors of 1.17 in a . The colour of each density profile in the lower panel matches the time range of the same colour marked in the upper panels. While the halo initially has $\rho \propto r^{-3/2}$ in the inner regions, it continues to accrete rapidly for some time after its formation (blue and green), which causes the halo’s density to grow at radii $r \gtrsim 10^{-2}$. This density growth is sufficient to produce a shallow density slope, $d \log \rho / d \log r > -3/2$, in the inner regions. The accretion rate later drops (yellow), leading to a period in which the density profile remains stable except

⁷The value of α that we assume is about 10 per cent smaller than the value that Delos et al. (2019) obtained. It is likely that the proportionality constant obtained in that work was biased upward by haloes’ later evolution (see Section 4), but since our own halo sample is small and not necessarily representative, the value that we use here might not be more accurate.

⁸New material tends to contribute almost uniform density at radii much smaller than the apocentres of its particle orbits (e.g. Dalal et al. 2010). However, this contribution is dwarfed by the much higher density already present at those radii. Only close to the apocentres can the new material significantly boost the halo’s density.

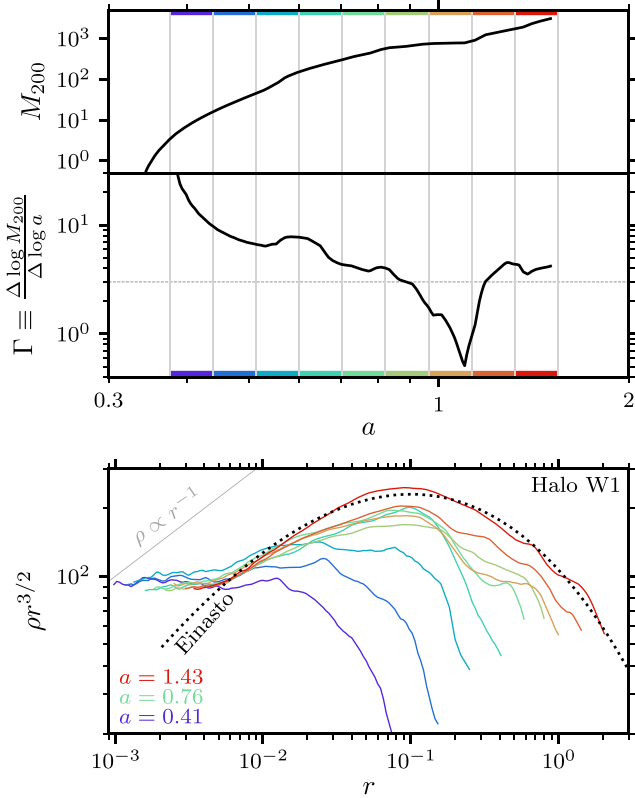


Figure 11. Mass (top), accretion rate Γ evaluated at each time over the preceding factor of 1.17 in a (centre), and density profile (bottom) for the W1 halo. We plot the density profile in physical coordinates at a range of different times, indicated by colours; each profile is averaged over the time interval in the upper panels that is marked with the same colour (we also separate these intervals with vertical lines). The halo begins with a $\rho \propto r^{-3/2}$ inner density profile (purple), but its consistently high accretion rate (blue and green) causes the density profile to trend shallower as the halo grows. The accretion rate later drops below $\Gamma \sim 3$ (yellow), and the density profile remains stable for a period. Finally, towards the end of the simulation the accretion rate rises above $\Gamma \sim 3$ again (red), inducing renewed shallowing. At the final time the density profile is well fit by an Einasto profile with $\alpha_E = 0.17$ (equation 8; dotted curve) except at the smallest radii where the natal $\rho \propto r^{-3/2}$ cusp appears to persist. New accretion has mostly built on top of the cusp rather than disrupting it (but see Section 4.2). A horizontal line marks $\Gamma = 3$; the idea that $\Gamma > 3$ corresponds to newly accreted material forming a density profile shallower than $\rho \propto r^{-3/2}$ is approximately borne out. In the lower panel we include a $\rho \propto r^{-1}$ line as reference (upper left); $\rho \propto r^{-3/2}$ would correspond to a horizontal line.

near R_{200} . Finally, near the end of the simulation the accretion rate again rises, causing further deposition of mass at intermediate radii.

The connection between the density profile slope $d \log \rho / d \log r$ and the accretion rate $\Gamma \equiv d \log M / d \log a$ can be understood on an approximate level by means of a simple model. Due to its mass growth, the halo’s radius R grows at the rate $d \log R / d \log a = 1 + \Gamma/3$ in physical coordinates. Now suppose that material accreted at the scale factor a' contributes significant density within the halo only near a particular radius $r = xR(a')$ that is proportional to the halo’s radius $R(a')$ at the accretion time (so x is some fixed number). Then the profile $M(r)$ of the halo’s enclosed mass obeys

$$\frac{d \log M(r)}{d \log r} = \frac{d \log M(a)}{d \log a} \left(\frac{d \log R(a)}{d \log a} \right)^{-1} = \frac{3\Gamma}{3 + \Gamma} \quad (7)$$

at $r = xR(a)$ if Γ is evaluated at a . Despite this model’s simplicity, close variants thereof have been shown to successfully predict trends in the density profiles within halo populations (Dalal et al. 2010; Ludlow et al. 2013; Delos et al. 2019).⁹ Equation (7) suggests that if a halo accretes material at a rate $\Gamma > 3$, then the mass profile resulting from this accretion should obey $d \log M / d \log r > 3/2$ at the corresponding radius, which requires a density profile shallower than $\rho \propto r^{-3/2}$. Thus, we expect that accretion at a rate $\Gamma \gtrsim 3$ should build up a density profile with a logarithmic slope shallower than $-3/2$. This expectation is approximately borne out in Fig. 11, where in the centre panel we mark $\Gamma = 3$ with a grey line. When $a \lesssim 0.9$ and when $a \gtrsim 1.2$, $\Gamma > 3$, which causes the density profile to gradually grow at intermediate radii $10^{-2} \lesssim r \lesssim 10^{-1}$ into a form with a slope shallower than $\rho \propto r^{-3/2}$. In contrast, when $0.9 \lesssim a \lesssim 1.2$, $\Gamma < 3$, and during this interval the inner and intermediate parts of the density profile remain stable.

Since the primary impact of rapid accretion is to boost the halo’s density, the shallowing of the density profile predominantly occurs at intermediate radii $10^{-2} \lesssim r \lesssim 10^{-1}$. The central $\rho \propto r^{-3/2}$ cusp largely survives the process, albeit with some mass loss that we will discuss further in Section 4.2 (where we attribute it to a major merger event). In fact, W1’s density profile in the final snapshot is well fit at almost all radii by the Einasto density profile (dotted curve)

$$\rho(r) = \rho_{-2} \exp \left\{ -\frac{2}{\alpha_E} \left[\left(\frac{r}{r_{-2}} \right)^{\alpha_E} - 1 \right] \right\} \quad (8)$$

(Einasto 1965), where ρ_{-2} and r_{-2} are scale parameters defined here as the density and radius at which the profile’s logarithmic slope crosses $d \log \rho / d \log r = -2$. Here we fix $\alpha_E = 0.17$ and find $r_{-2} = 0.55$ and $\rho_{-2} = 380$. Equation (8) is the same density profile seen in CDM simulations at much larger scales where the free-streaming cutoff is not resolved (e.g. Navarro et al. 2004, 2010), and such simulations also find that $\alpha_E \simeq 0.17$ supplies a good fit. Only at the smallest resolved scales, $r \lesssim 5 \times 10^{-3}$, does the density profile here break away from the Einasto form due to the persistence of the initial $\rho \propto r^{-3/2}$ central cusp.

Fig. 12 shows three more examples of how the accretion rate Γ relates to the shallowing of the halo’s density profile.

(i) Halo C2 (left) forms with a density profile that is initially close to $\rho \propto r^{-3/2}$, but its accretion rate Γ remains high, so this profile rapidly shallows. By the end of the simulation the inner density profile is close to $\rho \propto r^{-1}$.

(ii) In contrast, the accretion rate of halo H3 (centre) quickly drops after its formation, allowing its inner $\rho \propto r^{-3/2}$ density profile to remain stable over nearly an order of magnitude in scale factor. The accretion rate in this case becomes so low that even the outer, steeper part of the density profile stabilizes after $a \simeq 2$.

(iii) Halo H2’s accretion rate (right) also slows quickly early on, leading to an initially static inner density profile. However, around $a \simeq 1.6$ a brief period of rapid accretion builds up the density profile

⁹Dalal et al. (2010) and Delos et al. (2019) considered models framed in terms of a halo’s initial density peak instead of its accretion rate, but the idea is otherwise almost identical. These works also considered variants in which the radius at which accreted material settles is allowed to change in response to later accretion. The model in Ludlow et al. (2013) is equivalent to the claim that equation (7) holds if Γ is evaluated at the scale factor a such that $M(a) = M(r)$, where $M(a)$ is the halo’s mass at a while $M(r)$ is the enclosed mass profile at the final time. Note that these models cannot explain the $\rho \propto r^{-3/2}$ cusp.

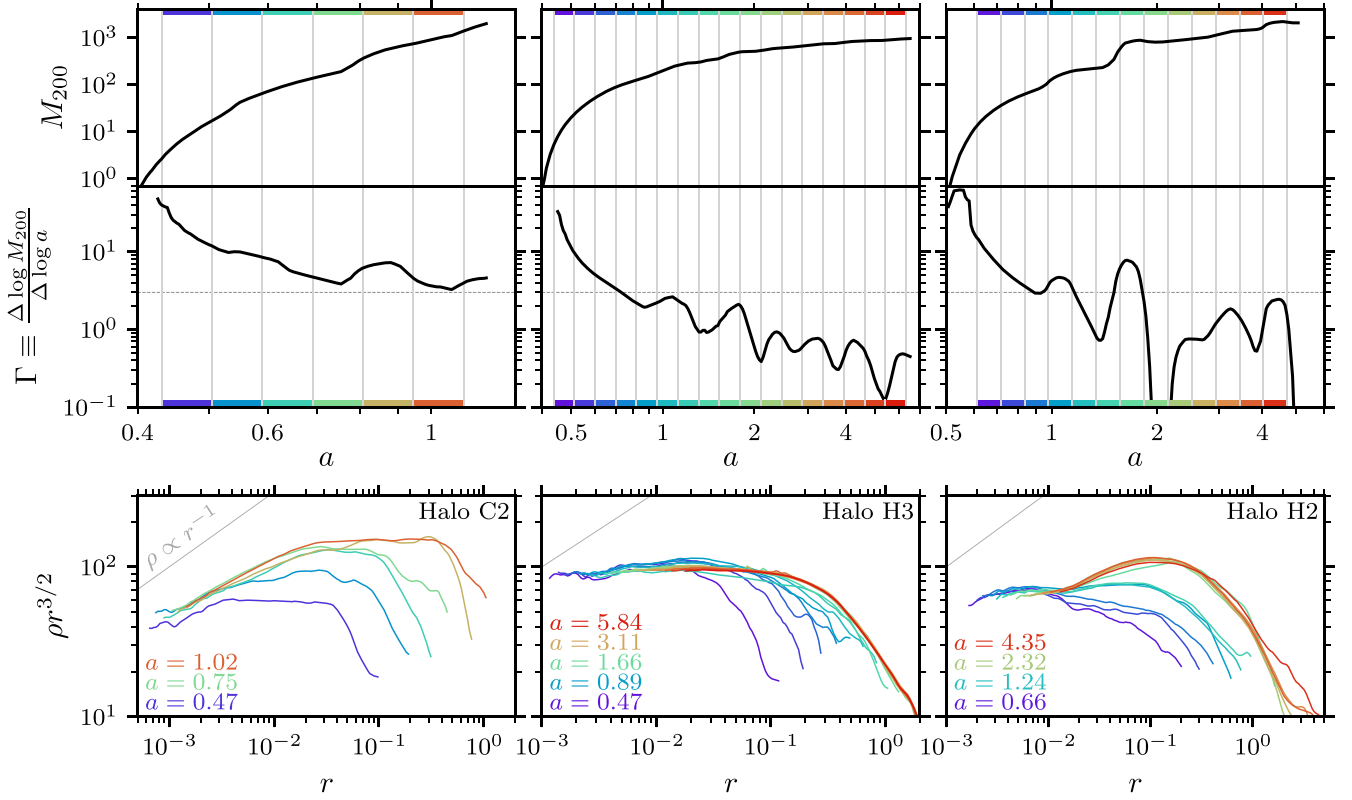


Figure 12. Masses (top), accretion rates Γ (centre) and density profiles (bottom) of the haloes C2 (left), H3 (centre), and H2 (right). As in Fig. 11 we use colours to match the density profiles in each lower panel to the corresponding scale factor band in the accretion-rate plot above it. (See that figure’s caption for further explanations.) *Left-hand panels:* Halo C2 maintains a high accretion rate Γ for some time after formation (like W1 in Fig. 11) and its density profile shallows significantly over time, almost reaching $\rho \propto r^{-1}$ by the end of the simulation. *Centre panels:* In contrast, halo H3’s accretion quickly slows below $\Gamma \sim 3$ and remains low. Its inner $\rho \propto r^{-3/2}$ density profile consequently remains stable over nearly an order of magnitude in scale factor. The accretion rate is so low at late times that even the outer, steeper part of the density profile remains stable from $a \simeq 2$ to $a \simeq 5$. *Right-hand panels:* Halo H2’s accretion rate initially drops below $\Gamma \sim 3$, causing the inner $\rho \propto r^{-3/2}$ profile to persist intact until $a \simeq 1.6$ (green), at which point a short burst of accretion builds up the inner density profile to a somewhat shallower slope. In fact this accretion episode was associated with a triple merger, which also marginally disrupted the central cusp (see Section 4.2). The accretion rate remains low thereafter, and the density profile correspondingly remains almost static.

to a shallower slope. Subsequently the halo’s accretion rate remains low, and its density profile correspondingly remains static again.

4.2 Impact of halo mergers

Ogiya et al. (2016) and Angulo et al. (2017) showed through controlled merger simulations that major mergers (between haloes of comparable masses) can disrupt $\rho \propto r^{-3/2}$ central cusps. We now present a couple of examples that confirm this effect in a cosmological setting.

(i) Halo H2 merged at $a \simeq 1.6$ with two other haloes of roughly 1/3 and 1/4 the main progenitor’s mass. This event is pictured in the upper row of panels in Fig. 13. The right-hand panels of Fig. 12 show that the amplitude of the central $\rho \propto r^{-3/2}$ cusp decreased by about 10 per cent at this time (and its slope may have also changed). The large influx of mass mostly settled at larger radii, boosting the density profile around $r \sim 10^{-1}$.

(ii) Halo W1 merged at $a \simeq 0.6$ with two other haloes of masses roughly 1/4 and 1/5 the main progenitor’s mass; this event is pictured in the central row of panels in Fig. 13. Fig. 11 shows that like H2’s merger, this event reduced the amplitude of the central density cusp around $r \sim 10^{-2}$ by about 10 per cent (although we do not have the resolution to determine whether the slope also changed). Here,

too, the mass primarily settled at larger radii, building up the density profile around $r \sim 10^{-1}$.

Major mergers have the capacity to alter a halo’s structure at small radii because dynamical friction causes the core of a large subhalo to sink rapidly to the centre of a host, shedding most of its mass along the way (e.g. Mo, van den Bosch & White 2010). The subhalo heats the host’s material in the process. This behaviour contrasts with that of smoothly accreted material and minor mergers, which do not sink in this way and therefore impact the growing halo’s structure only at intermediate to large radii (see Section 4.1). We illustrate this effect in Fig. 14. This figure shows a halo’s density profile at a particular time in the left-hand panels, and we mark radial bins within the density profile with different colours. The panels to the right of each density profile show the distribution of Lagrangian positions (within the initial conditions) of material residing within each coloured radial bin. The uppermost row of panels shows halo W1 at an early time, before any significant mergers have taken place. The Lagrangian regions associated with each radial bin of the density profile are simply connected and nested at this time, with shells of larger (final) radius associated with larger Lagrangian regions. This behaviour indicates that material that was initially at a larger radius within the halo’s precursor density peak – and therefore accreted later – contributes primarily at larger radii within the non-linear halo

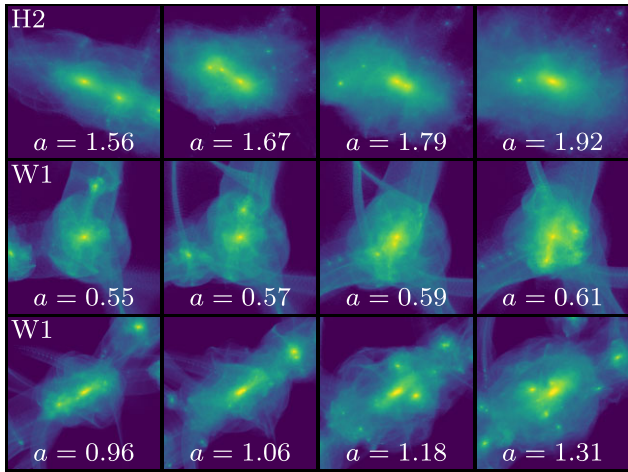


Figure 13. The significant merger events (each a different row) that we consider in Section 4.2. From left to right we advance in time. *Top row:* H2’s merger event at $a \sim 1.6$ is a three-body event where the two smaller haloes have masses of about $1/3$ and $1/4$ the main progenitor’s mass. This event disturbs the halo’s established $\rho \propto r^{-3/2}$ density profile by about 10 per cent (see Fig. 12). *Centre row:* W1’s merger event at $a \sim 0.6$ is also a three-body event; in this case the two smaller haloes have somewhat smaller masses of about $1/4$ and $1/5$ the main progenitor’s mass. This event disturbs the halo’s established $\rho \propto r^{-3/2}$ density profile to a similar degree (see Fig. 11). *Bottom row:* W1’s second major merger event at $a \sim 1.3$ has a mass ratio of about $1/3$ but does not appreciably disturb the central $\rho \propto r^{-3/2}$ cusp, at least within our resolution limits.

cusp. In contrast, the second row of panels shows halo W1 after the set of major mergers discussed above has taken place. Disconnected Lagrangian regions, which we label ‘A’ and ‘B’, now contribute even within the smallest radial bins (blue and red). The inner regions of W1 thus contain material from both merging haloes; both contribute to the final profile at all radii, even to the inner cusp. The amplitude of the density profile of W1 is decreased at small radii as a result of this merging (the density profile from the uppermost row is overplotted here as a dotted curve for comparison).

However, disruption of the central cusp is not a guaranteed outcome of a major merger. In fact, W1 underwent another merger event at $a \simeq 1.3$ with a mass ratio of about $1/3$, an event that is pictured in the lower panels of Fig. 13. The disconnected Lagrangian region associated with this merging halo is labeled ‘C’ in the third row of panels in Fig. 14. This halo’s material was able to contribute density as deep as the red radial bin within W1. However, this event did not appreciably disturb the innermost part of the density profile, at least within our resolution limits. As before, we overplot the previous row’s density profile as a dotted curve.

We can also use plots of this kind to show explicitly that major mergers are not necessary to produce a shallow density profile. In the lower two rows of Fig. 14, we show the density profile of halo C3 at two different times. As before, the right-hand panels plot the Lagrangian positions of the material that resides within each coloured radial bin at each time. This halo evidently develops a density profile at intermediate radii that is significantly shallower than $\rho \propto r^{-3/2}$. However, there is no significant disconnected Lagrangian region that would mark the impact of a major merger event.¹⁰ Instead, material

¹⁰The ring-like structure visible in the bottom-right panel is associated with coherence of the orbital phases of freshly accreted material. Material initially within the ring is close to the first pericentre of its orbit through the halo,

at large Lagrangian distances (which was accreted later) remains at large radii within the halo. Accordingly, the central $\rho \propto r^{-3/2}$ cusp remains undisturbed. Indeed, halo C3’s most major merger event was associated with a mass ratio of about $1/20$.

The degree to which mergers disrupt the natal cusps of our haloes appears at most moderate. Fig. 15 shows the density profile evolution for all 12 of our haloes. Despite the significant shallowing of many density profiles at intermediate radii – and sometimes even at the smallest resolved radii – due to deposition of accreted material, we find no case where the initial cusp suffered more than 10–20 per cent suppression. Resolution limitations mean that there are cases – particularly H2, W3, C1, C2, and C4 – where we cannot exclude that more major disruption occurred at the smallest radii. Nevertheless, in no case do we find positive evidence for substantial disruption of the central cusp. In cases where the central $\rho = Ar^{-3/2}$ cusp persists within our resolution limits, the properties of the halo’s precursor peak continue to predict its coefficient A with reasonable accuracy (horizontal black lines; see Section 3.2).

4.3 Mass and size of the inner cusp

We have now seen many examples of how a halo’s density profile evolves after formation of its initial $\rho \propto r^{-3/2}$ cusp. For instance, halo W1 (Fig. 11) rapidly built a shallow Einasto density profile on top of its initial cusp, so that the latter only persists at radii $r \lesssim R_{200}/400$. Halo C2 (Fig. 12) underwent similar evolution to the extent that its central cusp (if it persists) is not even resolved at the final snapshot. In contrast, halo H3’s $\rho \propto r^{-3/2}$ cusp remains a significant portion of the halo even at the end of the simulation. We now quantify more precisely how the sizes of our haloes’ natal cusps evolve over time.

We define the size and mass of the $\rho \propto r^{-3/2}$ cusp in the following manner. First, we evaluate $\text{dlog } \rho / \text{dlog } r$ at each radius r by carrying out a linear regression in log space over the surrounding factor of 1.5 in r (that is, radii r' satisfying $r/\sqrt{1.5} < r' < \sqrt{1.5}r$). Next, we search for the smallest radius at which $\text{dlog } \rho / \text{dlog } r$ deviates from -1.5 by more than 0.2, but since $\text{dlog } \rho / \text{dlog } r$ is noisy, we ignore deviations that are confined to less than a factor of 2 in radius. We define the deviation radius obtained in this way to be R_{cusp} , the radius of the natal cusp. We demonstrate the outcome of this process in Fig. 16 for the C3 halo. It does indeed place R_{cusp} where one might reasonably decide that the central cusp ends.

For each halo and each scale factor a , we average the density profile over the surrounding factor of 1.17 in a (that is, scale factors a' satisfying $a/\sqrt{1.17} < a' < \sqrt{1.17}a$) and then evaluate R_{cusp} by the above procedure.¹¹ The resulting evolution of R_{cusp} is plotted in Fig. 17. Note that R_{cusp} is not always resolved if the slope $\text{dlog } \rho / \text{dlog } r$ is already much shallower than -1.5 at the minimum

while material initially in the white region just inside the ring is close to its first apocentre subsequent to infall and currently lies outside the orange radial range in the profile panel. Note that halo C3 formed from a filament, and the Lagrangian ring is aligned along that filament’s axis.

¹¹We also make a further adjustment to clean up the evolution plots. Namely, we iteratively remove the data at a scale factor a if the enclosed mass M_{cusp} at that scale factor is larger by at least a factor of 3 than at both of its neighboring scale factors. We do the same for scale factors at which M_{cusp} is smaller by the same factor than both of its neighbors. Any data points removed in this way are instead interpolated between the neighboring points. This process suppresses a few instances in which R_{cusp} and M_{cusp} jump wildly. Note that we use all snapshot scale factors, which are separated by roughly a factor of 1.035 in a , and not only the scale factors shown in Fig. 16, so we do not remove a significant fraction of the data points.

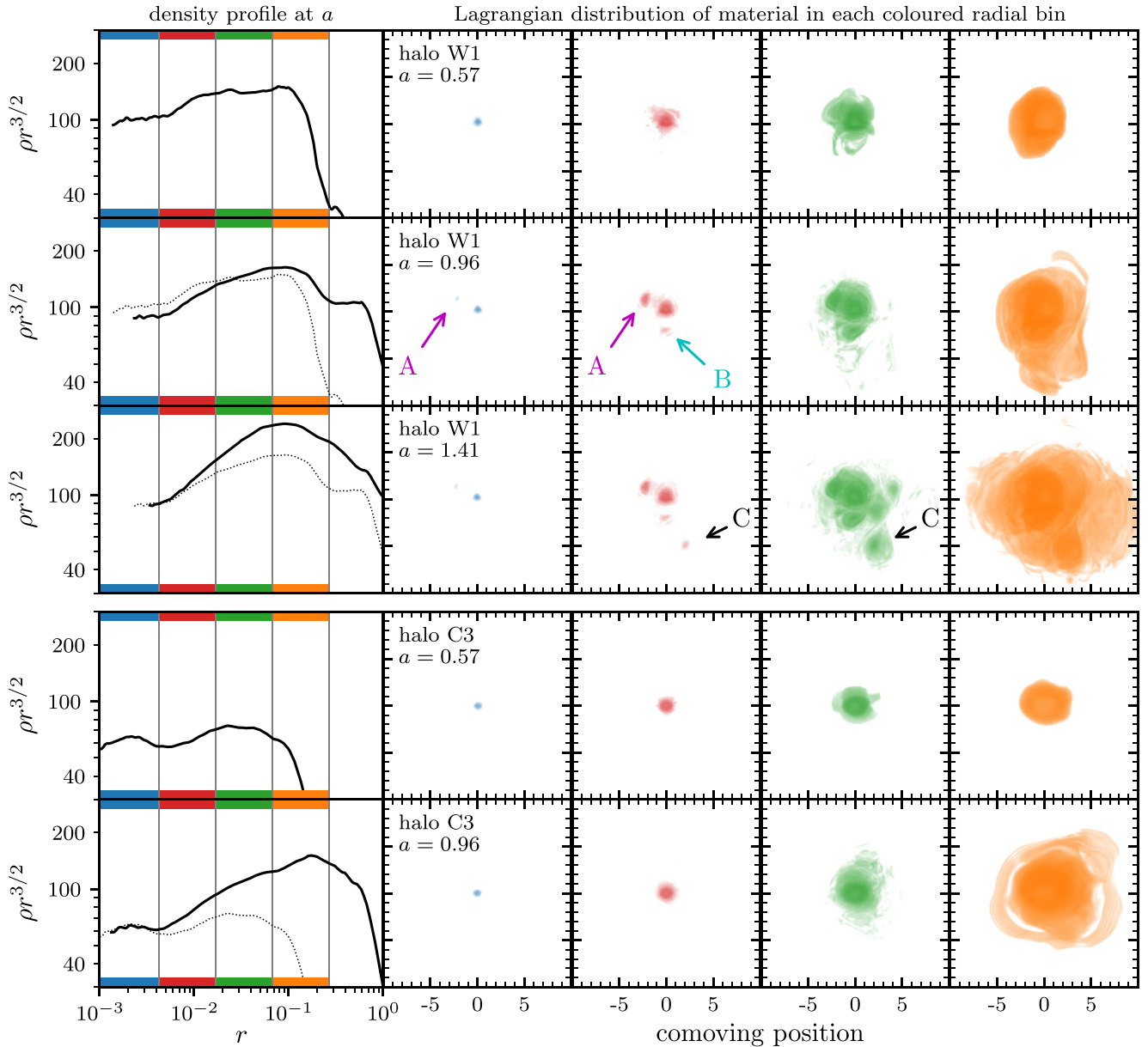


Figure 14. Lagrangian regions associated with halo material at set radii. Within each row, the left-hand panel shows a halo’s density profile at a particular time (averaged over a factor of 1.17 in a), and in some rows we also overplot the previous row’s density profile as a dotted curve for comparison. We mark four coloured radial bins (also separated by vertical lines). The panels to the right of each density profile show the projected distribution of Lagrangian positions (i.e. positions within the initial conditions) associated with the material that resides in the radial bin of the same colour (and we consider the same time interval, distributing each particle across multiple bins if necessary). The colour intensity scale is logarithmic between 10^{-3} and $1/4$ the total projected density (over a cubic region). The upper three panels show the W1 halo at different times. At first, the Lagrangian region associated with each radial bin is simply connected, and bins at larger radii are associated with larger Lagrangian regions. This behaviour indicates that material that accreted later (because it was initially farther from the peak) contributes at larger radii within the halo (note that the red, green and yellow regions in all panels are hollow, although this is difficult to see in projection). The second and third rows of panels for W1 show disconnected Lagrangian regions ‘A’, ‘B’, and ‘C’ contributing to the inner radial bins. Each additional region is associated with a separate halo that merged onto the main system; dynamical friction allowed these haloes to deposit material close to the centre of W1. The merger events associated with the ‘A’ and ‘B’ regions partially disrupted the halo’s central density profile, but the merger associated with the ‘C’ region had no such effect, depositing all its material at intermediate radii. In contrast, the lower two panels show the C3 halo. This halo did not undergo any major mergers, and the Lagrangian distribution of the material in its inner radial bins is simply connected and nearly spherically symmetric at both times. Nevertheless, rapid accretion of diffuse material caused this halo to develop a shallow density profile at all but the innermost radii, where the initial $\rho \propto r^{-3/2}$ cusp is preserved.

resolved radius. If R_{cusp} is not resolved at some time, we do not attempt to resolve it at any future time; attempting to do so may inappropriately identify the central cusp with, e.g. the portion of the Einasto density profile for which the density slope is close to -1.5 .

When R_{cusp} ceases to be resolved, we instead simply cut off its time evolution curve in Fig. 17.

In the upper panel of Fig. 17, we see that the physical cusp radius R_{cusp} often remains close to constant in time. This behaviour holds

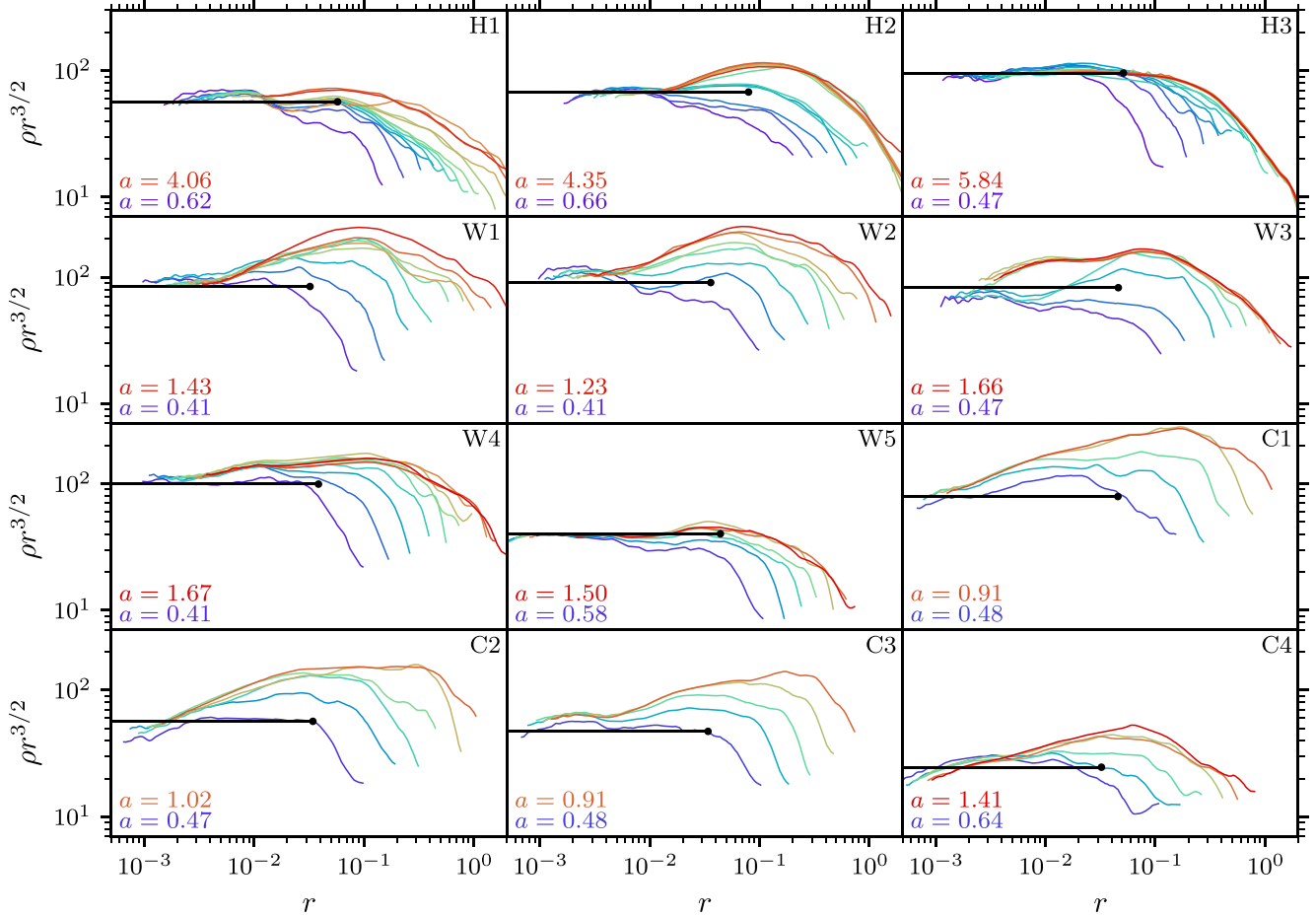


Figure 15. Density profile evolution for all 12 haloes. As in earlier figures, the density profiles are time averaged over successive factors of 1.17 in a (different colours). We find no evidence, within our resolution limits, for significant disruption of the natal cusps due to merger events. The effects are limited to a 10–20 per cent reduction in amplitude. We also repeat the cusp predictions (black lines) from Fig. 9, which are based on each halo’s precursor density peak. Even after halo growth, the initial peak continues to predict the inner cusp reasonably well in cases where the cusp persists within our resolution limits.

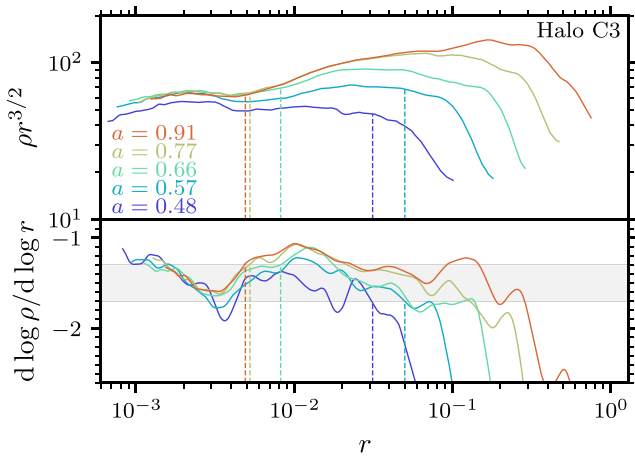


Figure 16. Demonstration of the procedure through which we identify the radius R_{cusp} of the central cusp. We plot the density profile (top) and logarithmic density slope $d \log \rho / d \log r$ (bottom) of the C3 halo at a number of different times (colours). At each time, we identify the cusp radius R_{cusp} (vertical line) as the largest radius at which $d \log \rho / d \log r$ remains within the range $(-1.7, -1.3)$ (grey band), but we ignore small excursions; see the text. The resulting radius values appear to reasonably mark the outer edge of the $\rho \propto r^{-3/2}$ cusp.

during periods of slow accretion but also during rapid accretion as long as accretion builds up the density profile primarily at radii outside of the central cusp. For instance, halo W1’s cusp radius remains essentially constant over the time interval $0.6 \lesssim a \lesssim 1.4$ even though accretion during that interval is sometimes rapid enough to build up a shallow density profile and sometimes not; compare Fig. 11. However, R_{cusp} also typically undergoes one (and only one) sudden drop when rapid accretion begins to build on top of an established $\rho \propto r^{-3/2}$ cusp. A prime example of this behaviour occurs for halo H2 at $a \simeq 1.8$; compare the right-hand panels of Fig. 12. Finally, at early times there can be growth in R_{cusp} as the initial collapse builds the density profile outward; compare for instance halo W1’s cusp growth during $0.4 \lesssim a \lesssim 0.55$ to the plot of its density in Fig. 9 over the same time period. In some cases, growth of R_{cusp} can proceed for some time. This outcome is likely attributable to an accretion rate that remains close to $\Gamma = 3$ for a while; for example, compare H3’s cusp growth during $0.7 \lesssim a \lesssim 1.2$ to the contemporaneous accretion rate and density profile in the centre panels of Fig. 12. In this case, the $\rho \propto r^{-3/2}$ cusp is not entirely set by the initial collapse but is further built outward by later accretion.

The lower panel of Fig. 17 plots the evolution of R_{cusp}/R_{200} for each halo. For the haloes arising from the $n = -2.67$ power spectrum,

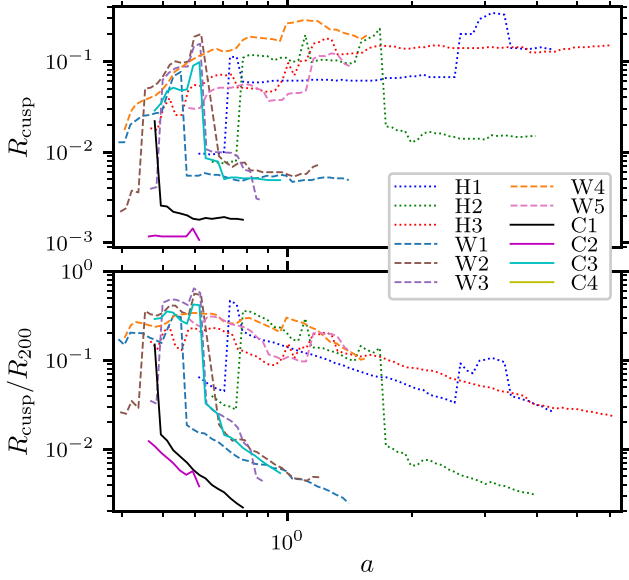


Figure 17. Evolution of the cusp radii R_{cusp} of our haloes as a function of the expansion factor a . The upper panel shows R_{cusp} in absolute units; generally we find that R_{cusp} grows initially, drops suddenly at one time as rapid accretion begins to build up a shallower density profile, and otherwise remains stable. The lower panel shows instead R_{cusp}/R_{200} , i.e. the cusp radius in units of the halo’s radius R_{200} (enclosing 200 times the cosmological mean density). Rapidly accreting haloes arising from the $n = -2.67$ and $n = -2$ power spectra quickly grow to the extent that $R_{\text{cusp}}/R_{200} < 10^{-2}$, but haloes W4, W5, H1, and H3 maintain slower accretion rates such that R_{cusp} remains close to $10^{-1}R_{200}$.

accretion is so rapid that the halo’s R_{200} radius quickly grows to more than 100 times the cusp radius. This picture demonstrates further why other studies did not find the natal $\rho \propto r^{-3/2}$ cusp inside haloes for which rapid accretion built up a shallower intermediate profile. For instance, Ishiyama (2014) and Angulo et al. (2017) concluded simply that larger haloes have shallower density profile slopes,¹² while Wang et al. (2020) found that all haloes are well described by the Einasto fitting form. These studies did not resolve small enough radii to find the natal cusps inside their rapidly accreting haloes.

In Fig. 18 we similarly plot the mass M_{cusp} enclosed inside R_{cusp} . For rapidly accreting haloes arising from the $n = -2$ and $n = -2.67$ power spectra, the natal cusp’s contribution quickly drops below 10^{-3} of the total halo mass. However, haloes W4 and W5 (arising from the $n = -2$ power spectrum) do not accrete rapidly enough to overwhelm the cusp in this way, so they maintain $M_{\text{cusp}}/M_{200} \sim 10^{-1}$ for the full duration of their respective simulations. The naturally slower accretion rates expected for the $n = 1$ power spectrum also cause haloes H1 and H3 to maintain M_{cusp}/M_{200} around 10^{-1} to 10^{-2} during an order-of-magnitude increase in the scale factor a . For reference, we show in Fig. 19 the accretion rates of our haloes; there is a clear correspondence between low accretion rates at late times and large fractional masses in the central cusps.

¹²Angulo et al. (2017) resolved radii below $10^{-2}R_{178}$ for some of their halo mass bins, and their plot of $\text{dlog } \rho / \text{dlog } r$ (fig. 6 of Angulo et al. 2017) does show signs of steepening at the smallest radii that they plot. Although this effect lies below the radii at which they expect the simulation to be numerically converged, numerical artefacts normally push the density profile towards a shallower rather than a steeper slope (see Appendix B).

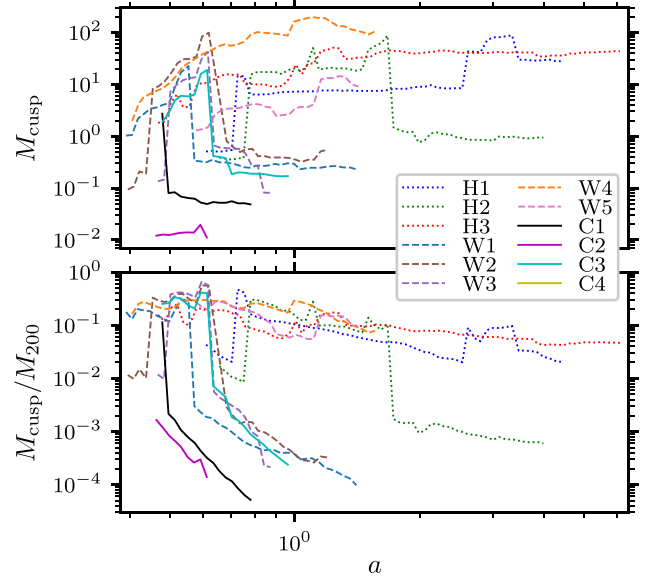


Figure 18. Evolution of the masses M_{cusp} enclosed within our haloes’ cusp radii R_{cusp} . The upper panel shows M_{cusp} in absolute units; behaviours similar to those in Fig. 17 are evident. The lower panel shows M_{cusp}/M_{200} , i.e. the cusp mass in units of the halo’s mass M_{200} (enclosed within R_{200}). We find that the rapidly accreting haloes C1–C4 and W1–W3 quickly grow to such extent that $M_{\text{cusp}}/M_{200} \lesssim 10^{-3}$. On the other hand, slowly accreting haloes like H1, H3, W4, and W5 are able to maintain $M_{\text{cusp}}/M_{200} \sim 10^{-1}$.

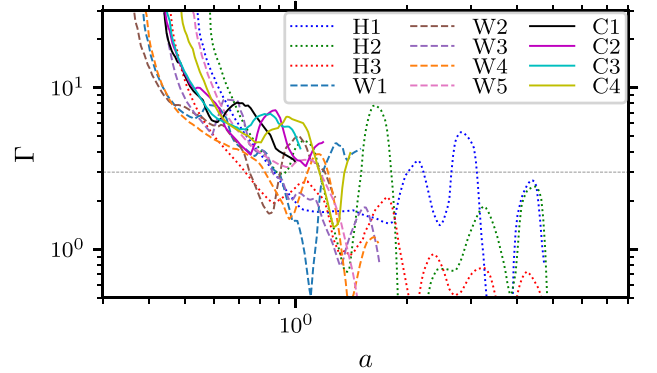


Figure 19. Accretion rates $\Gamma \equiv \Delta \log M_{200} / \Delta \log a$ for our haloes, evaluated as in Section 4.1. The horizontal reference line marks $\Gamma = 3$.

4.4 Three-dimensional structure

We briefly discuss the three-dimensional structures of our haloes and their evolution over time. We first explore anisotropy in the velocity structure. The centre panels of Fig. 20 show the velocity anisotropy parameter

$$\beta \equiv 1 - \sigma_r^2 / (2\sigma_t^2) \quad (9)$$

as a function of radius within each of haloes H3, W5, C3, and W1, where σ_r and σ_t are the radial and tangential velocity dispersions, respectively. We evaluate $\sigma_r^2 = \langle v_r^2 \rangle - \langle v_r \rangle^2$ and $\sigma_t^2 = \langle v^2 \rangle - \langle v_r \rangle^2$, where v_r is a particle’s radial velocity, v is its total velocity, and the angle brackets denote the average over particles in a narrow radial bin centred at radius r .

$\beta = 0$ represents an isotropic velocity dispersion, while $\beta = 1$ implies that all motion is radial. Generally, Fig. 20 shows that β

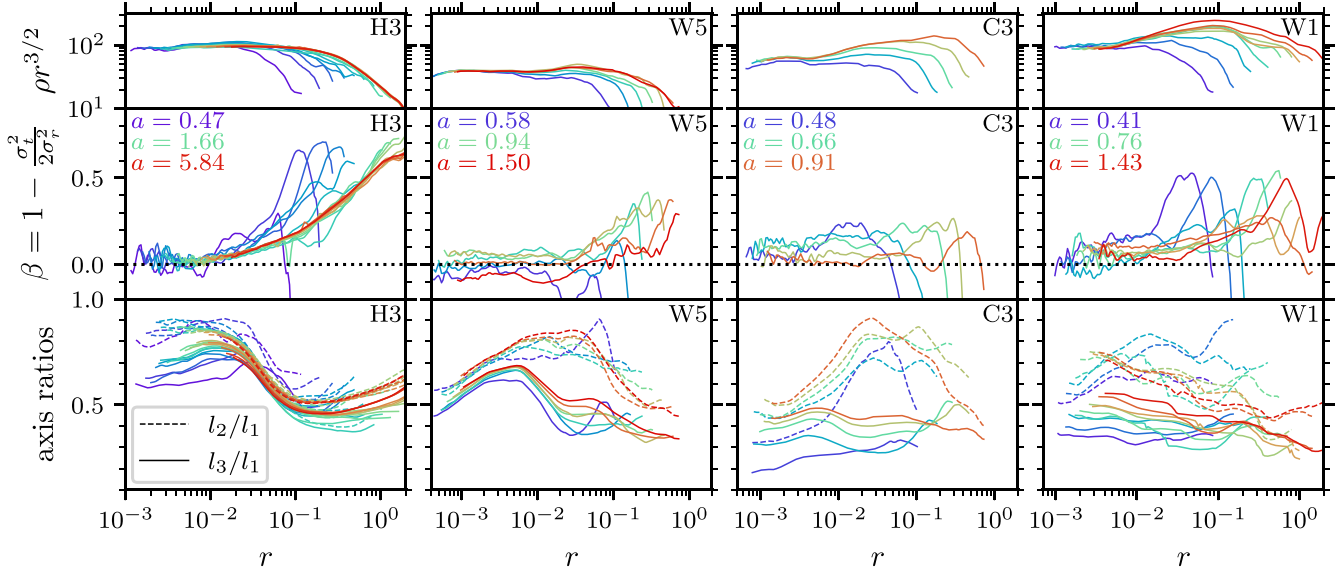


Figure 20. Velocity anisotropy parameter (centre panels) and ellipsoidal axis ratios (lower panels) for the haloes H3, W5, C3, and W1 (from left to right). We also show the density profile at the top, for reference, and all quantities are averaged over successive factors of 1.17 in a . The $\rho \propto r^{-3/2}$ central cusps of these haloes have nearly isotropic velocity distributions (β close to 0). However, there is far less uniformity in the three-dimensional shapes. H3 and W1 have triaxial cusps that evolve significantly toward a more spherical shape, with H1’s cusp becoming nearly spherical at late times. In contrast, W5 and C3 have very prolate central cusps that remain so throughout their evolution, at least at the smallest radii; W5’s cusp becomes more oblate at intermediate radii.

grows with radius at small and intermediate radii but drops again as we approach the halo’s outer edge. This pattern has been seen in CDM haloes as well (e.g. Navarro et al. 2010). In the context of the first haloes, the interesting feature is that the velocity structure is nearly isotropic ($\beta \simeq 0$) inside the central $\rho \propto r^{-3/2}$ cusps.

We also explore the three-dimensional shapes of our haloes. We use the algorithm of Allgood et al. (2006; see also Vera-Ciro et al. 2011), which involves regarding a halo (or concentric subset thereof) as an ellipsoid and evaluating the associated axis lengths $l_1 \geq l_2 \geq l_3$ via the halo’s inertia tensor. An ellipsoid’s axis lengths are proportional to the square roots of the eigenvalues of its inertia tensor. Specifically, this procedure employs the ‘reduced’ inertia tensor

$$I_{ij} = \sum_n m_n \frac{x_n^{(i)} x_n^{(j)}}{\tilde{r}_n^2}, \quad (10)$$

summed over particles n at ‘ellipsoidal’ radii \tilde{r}_n , which avoids weighting more distant particles more highly. If x_n, y_n , and z_n are coordinates in the ellipsoid’s principal-axes frame, then $\tilde{r}_n^2 \equiv (x_n/l_1)^2 + (y_n/l_2)^2 + (z_n/l_3)^2$. Beginning with a spherical ellipsoid with axes $l_1 = l_2 = l_3$, we evaluate I_{ij} for particles inside the ellipsoid, evaluate the resulting new axis lengths (keeping the longest axis length l_1 fixed), and rotate the system into the frame of these new axes. This procedure is iterated until the axis ratios l_2/l_1 and l_3/l_1 are converged to better than 10^{-3} .

The lower panels of Fig. 20 show the axis ratios l_3/l_1 and l_2/l_1 as a function of the radius $r = (l_1 l_2 l_3)^{1/3}$, which is the radius of the sphere that encloses the same volume as the ellipsoid. Generally these haloes are highly aspherical, although there is a weak tendency for them to become more spherical over time, as measured by the l_3/l_1 axis ratio. At late times, these haloes tend to be less spherical at large radii and more spherical at small radii, a trend that is interestingly the opposite of what Vera-Ciro et al. (2011) found for CDM haloes. Also, our haloes tend to be comparatively prolate at large radii, as evidenced by the ratios l_3/l_1 and l_2/l_1 both lying well below unity. With respect to the $\rho \propto r^{-3/2}$ central cusps, however, there is little uniformity. H3’s

cusp is initially triaxial but becomes nearly spherical at late times. W5 and C3 have highly prolate cusps that remain so through their evolution, although the portion of W5’s cusp at intermediate radii $r \gtrsim 10^{-2}$ becomes more oblate. W1’s cusp is triaxial, although it appears to be evolving rapidly toward a more spherical shape.

Despite the variety of three-dimensional shapes of haloes’ central cusps, we find one general feature: the cusp tends to align with the halo at larger radii. Fig. 21 shows slices of the equidensity surfaces of the same four haloes at radii $r \sim 0.01$ and $r \sim 0.2$. While some of the ellipsoidal axis ratios are clearly more extreme at larger radii, the axes themselves are closely aligned between the inner cusp at $r \lesssim 0.01$ and the broader halo at $r \sim 0.1$.

5 MINIMUM RADIUS OF THE INNER CUSP

We have found that the $\rho \propto r^{-3/2}$ density cusps of our haloes extend down to the resolution limits of our simulations, but it is unclear how much farther this power law could extend. If the dark matter were infinitely cold, but the initial power spectrum nevertheless had a small scale cutoff, then it is possible that the cusp could extend to arbitrarily small radii. However, any thermal motions of dark matter particles in the initial conditions would impose a minimum radius below which the $\rho \propto r^{-3/2}$ structure must give way to a core. We now explore what the radius of this core should be.

5.1 Maximum phase-space density

The thermal core can be understood as a consequence of Liouville’s theorem, which implies that the maximum coarse-grained phase-space density cannot increase with time (Tremaine & Gunn 1979). As a result, the (time-independent) maximum phase-space density f_{\max} of the dark matter at times prior to formation of any non-linear structure sets an upper limit on the value of the distribution function $f(\mathbf{x}, \mathbf{v})$ anywhere inside a halo at all later times (e.g. Dalcanton & Hogan 2001; Strigari et al. 2006; Villaescusa-Navarro & Dalal

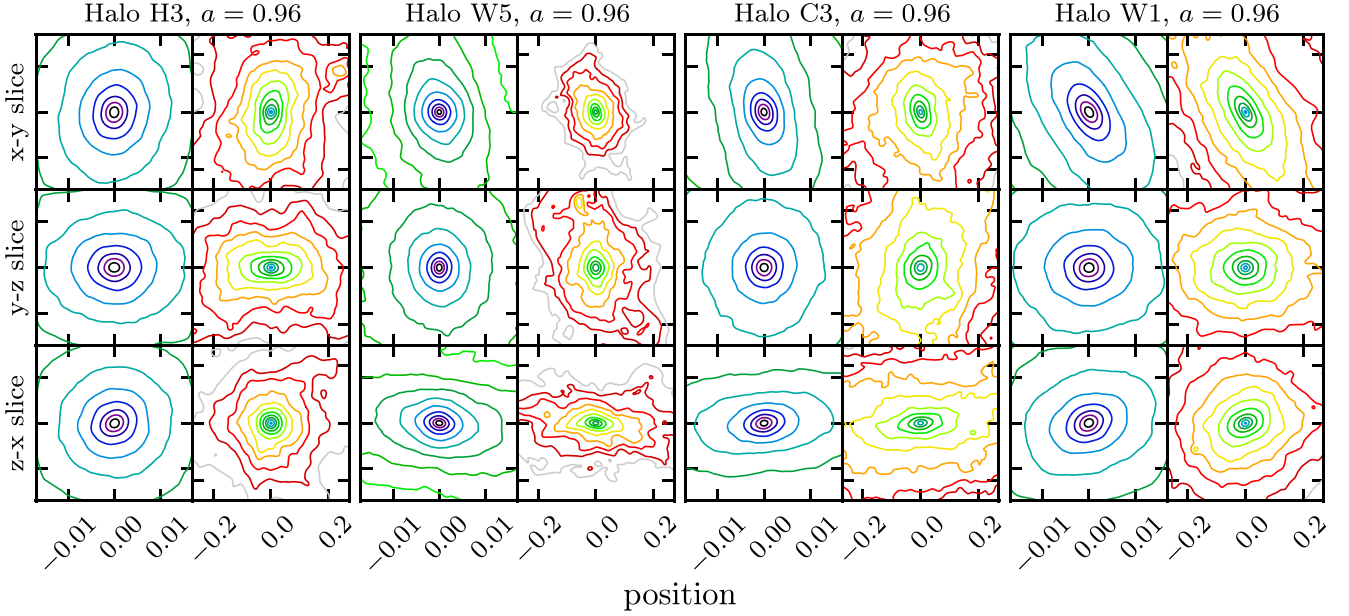


Figure 21. Slices of the equidensity surfaces of haloes H3, W5, C3, and W1 (from left to right) at $a = 0.96$. Specifically, we use the density field averaged over the surrounding factor of 1.17 in a . The upper, centre, and lower panels show different perpendicular slices, while for each halo, the left-hand and right-hand panels show two different length scales. The equidensity surfaces inside the central cusps (dark curves in each halo’s left-hand panels) tend to be aligned with those at larger radii (each halo’s right-hand panels).

2011; Macciò et al. 2012). Simulations suggest that halo centres approximately saturate this limit (Macciò et al. 2012).

We first estimate f_{\max} . In general, dark matter that was once coupled to the radiation bath is left with some momentum distribution $f_p(\mathbf{p}, a)$ that peaks at $\mathbf{p} = 0$ and has width proportional to some scale momentum $p_0 = p_{0d}(a_d/a)$ that declines as a^{-1} from its initial value p_{0d} , which was defined at the scale factor a_d when the dark matter kinetically decoupled. The maximum value of this distribution is given by $f_p(0) = Cp_0^{-3}$, where C is a normalization factor independent of a . Two specific examples are CDM and WDM:

(i) In the CDM scenario, dark matter decouples while it is non-relativistic. In this case it has a Maxwell–Boltzmann momentum distribution, $f_p(\mathbf{p}) = (2\pi)^{-3/2} p_0^{-3} \exp[-\mathbf{p}^2/(2p_0^2)]$, so that $C = (2\pi)^{-3/2}$ and $p_{0d} = (T_d m)^{1/2}$, where T_d is the temperature at decoupling and m is the particle mass.

(ii) In the WDM scenario, dark matter is a fermion that decouples while it is ultrarelativistic. In this case it has an ultrarelativistic Fermi–Dirac momentum distribution, $f_p(\mathbf{p}) = [6\pi\zeta(3)]^{-1} p_0^{-3} [\exp(p/p_0) + 1]^{-1}$ (where ζ is the Riemann zeta function), so that $C = [12\pi\zeta(3)]^{-1} \simeq 0.0221$ and $p_{0d} = T_d$.

Let us now consider a time when the dark matter is non-relativistic (regardless of whether it was relativistic at decoupling). Its spatial density is then $\rho_m a^{-3}$, where ρ_m is the present-day dark matter density (i.e. at $a = 1$). Consequently, the maximum density in position-momentum phase space is $(Cp_0^{-3})(\rho_m a^{-3}) = C\rho_m(p_{0d}a_d)^{-3}$. Within dark matter haloes we find it more convenient to work with the equivalent position-velocity phase-space density, the maximum value of which is therefore

$$f_{\max} = C\rho_m(p_{0d}a_d/m)^{-3}, \quad (11)$$

where m is the mass of the dark matter particle.

We may now estimate a halo’s thermal core radius r_c as the largest radius at which its phase-space density reaches the value f_{\max} . If we

also define $\rho_c \equiv \rho(r_c)$, then we show in Appendix C that

$$\rho_c r_c^6 \simeq 3 \times 10^{-5} G^{-3} f_{\max}^{-2} \quad (12)$$

for a halo with central cusp $\rho \propto r^{-\gamma}$ with $1 \lesssim \gamma \lesssim 1.5$. Thus, once a halo’s rescaled density profile ρr^6 approaches $\rho_c r_c^6$, its central cusp must give way to a thermal core.

5.2 WDM core sizes

Let us now explore the core sizes that would develop in our simulated haloes if they were to form in a WDM cosmology. Equation (A3) of Bode et al. (2001) gives the characteristic non-relativistic streaming velocity $(p_{0d}/m)(a_d/a)$ as a function of the present-day dark matter density Ω_χ (in units of the critical density), the Hubble parameter h , the dark matter particle’s effective number n_χ of degrees of freedom, and its mass m . We fix $m = 3.5$ keV, $g_\chi = 1.5$, $h = 0.68$, and $\Omega_\chi = 0.3$ (and for simplicity we assume that dark matter constitutes all of the matter). Consequently, $p_{0d}a_d/m = 7.8 \times 10^{-9}$ (taking the speed of light $c = 1$), and hence by equation (11), $f_{\max} = 1.8 \times 10^{33} \text{ M}_\odot \text{ Mpc}^{-3}$. Now equation (12) implies that the phase-space barrier lies at $\rho_c r_c^6 = 8.0 \times 10^{-14} \text{ M}_\odot \text{ Mpc}^3$.

The simulations, however, have been carried out with arbitrary units. To fix the units, we must compare our power spectra (equation 1) to the power spectrum that arises in the WDM scenario described above. We evaluate the matter power spectrum using the fitting form from Eisenstein & Hu (1998) and normalize it to $\sigma_8 = 0.81$ (Planck Collaboration 2020) at $a = 1$. Next, we use the transfer function described in Bode et al. (2001) to apply the small-scale cutoff corresponding to the WDM scenario above. In particular, we find that the free-streaming length α_{fs} of Bode et al. (2001) in this scenario is $\alpha_{\text{fs}} = 0.017$ Mpc. We now recall that our simulations begin when the rms density contrast is $\sigma = 0.03$. The WDM power spectrum described above must be scaled back to $z = 267$ (using the growth function) to have the same rms density contrast. Consequently, we interpret the starting redshift of our simulations to be $z = 267$, and

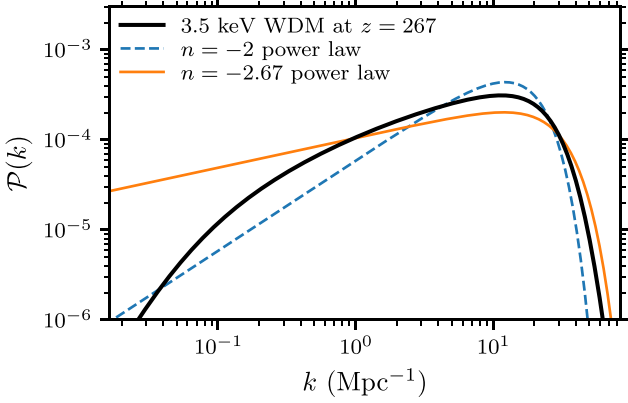


Figure 22. Matter power spectrum at redshift $z = 267$ in a 3.5 keV WDM scenario (solid black curve) compared to appropriately rescaled versions of our $n = -2$ and $n = -2.67$ power-law simulation power spectra. In particular we rescale to match σ , the rms density contrast, and $\langle k^2 \rangle$. The 3.5 keV WDM power spectrum appears to effectively lie somewhere between these two power laws.

their final redshifts range over $3.8 < z < 7.0$ for the $n = -2$ and $n = -2.67$ cases which are most relevant in this context. Fig. 22 shows the WDM power spectrum at the initial time (thick solid curve).

The remaining step is to fix the length scales in the simulations. Since the simulation power spectra have different forms than the WDM power spectrum, there is no unique way to make this association. We choose to fix $\langle k^2 \rangle$ (equation 3; see the considerations in Section 2), which evaluates to $\langle k^2 \rangle = 150 \text{ Mpc}^{-2}$ for the WDM power spectrum. Consequently we take the simulation length unit to be $\langle k^2 \rangle^{-1/2} = 8.2 \times 10^{-2} \text{ Mpc}$ (see Table 1), which implies that the exponential cutoff wavenumber is $k_{\text{cut}} \simeq 17 \text{ Mpc}^{-1}$ for $n = -2$ and $k_{\text{cut}} \simeq 29 \text{ Mpc}^{-1}$ for $n = -2.67$. Fig. 22 shows the rescaled $n = -2$ (dashed curve) and $n = -2.67$ (thin solid curve) power spectra.

The WDM power spectrum evidently lies effectively somewhere between the $n = -2$ and $n = -2.67$ power spectra, so we consider the nine simulated haloes that arose therefrom. After applying the scaling described above, we plot density profiles for these haloes in Fig. 23. We also plot the line $\rho r^6 = \rho_c r_c^6 = 8 \times 10^{-14} \text{ M}_\odot \text{ Mpc}^3$; the interpretation is that each halo’s inner cusp (as long as its slope lies between $\rho \propto r^{-3/2}$ and $\rho \propto r^{-1}$) must give way to a central core once it crosses this line. Evidently, each halo’s $\rho \propto r^{-3/2}$ cusp should extend over about two decades in radius shortly after formation and up to almost three decades in radius after halo growth in those cases where the $r^{-3/2}$ cusp persists. In general, haloes in this WDM scenario have core radii of about 3 pc. At early times, we resolve the inner cusps of our haloes down to about three times the predicted core radius, although this resolution worsens as time goes on.

5.3 CDM core sizes

We also explore core sizes that would arise in a CDM cosmology. In particular, we suppose that the dark matter has particle mass $m = 100 \text{ GeV}$ and kinetically decouples at the temperature $T_d = 30 \text{ MeV}$ (so $a_d \simeq 5.3 \times 10^{-12}$). In this case, $p_{\text{od}} a_d / m = 9.2 \times 10^{-14}$ (again taking $c = 1$) and hence $f_{\text{max}} = 3.1 \times 10^{48} \text{ M}_\odot \text{ Mpc}^{-3}$, which implies that the phase-space barrier lies at $\rho_c r_c^6 = 2.7 \times 10^{-44} \text{ M}_\odot \text{ Mpc}^3$.

To find the power spectrum, we scale the aforementioned Eisenstein & Hu (1998) fitting function using the CDM transfer function in Green et al. (2005). We then set $\sigma = 0.03$ by extrapolating this power spectrum back to $z = 1272$ using the growth function that is

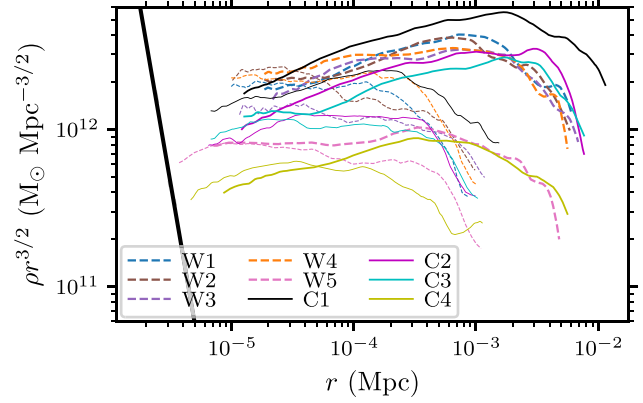


Figure 23. Predicted core sizes for our simulated haloes if they arise in a WDM scenario. In particular, we plot the density profiles $\rho(r)$ (scaled by $r^{3/2}$) for the haloes that arise from the $n = -2$ (dashed curves) and $n = -2.67$ (solid curves) power spectra. These haloes have been rescaled to correspond to a 3.5 keV WDM scenario (as described in the text). Each halo is plotted in a different colour, and for each halo we plot the density profile shortly after formation (thin curve) and a factor of about 1.9 in the scale factor later (thick curve). On the left, we show the $\rho r^6 = \rho_c r_c^6$ line (black) corresponding to equation (12). Each halo’s inner cusp is expected to give way to a central core where it crosses this line.

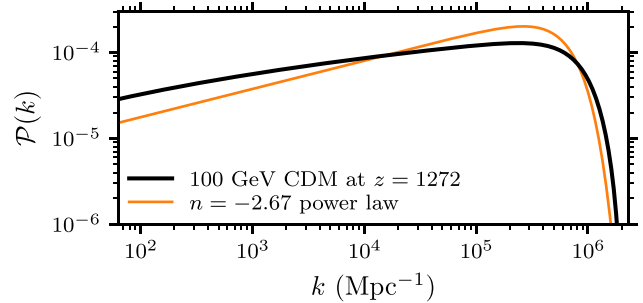


Figure 24. Matter power spectrum at redshift $z = 1272$ in a 100 GeV CDM scenario (black) compared to a version of our $n = -2.67$ power-law simulation power spectrum that we rescale to match σ and $\langle k^2 \rangle$ (orange).

valid in mixed matter-dark energy domination.¹³ The resulting power spectrum is shown as the black curve in Fig. 24. This power spectrum has $\langle k^2 \rangle = 7.4 \times 10^{10} \text{ Mpc}^{-2}$, so we take our simulation length unit to be $3.7 \times 10^{-6} \text{ Mpc}$.

Fig. 24 shows the appropriately rescaled version of our $n = -2.67$ power spectrum (orange curve). This power spectrum is somewhat too steep to faithfully represent the CDM scenario, but the mismatch is not serious. We now show in Fig. 25 the appropriately rescaled density profiles of the haloes arising from the $n = -2.67$ power spectrum along with the corresponding phase-space barrier $\rho r^6 = 2.7 \times 10^{-44} \text{ M}_\odot \text{ Mpc}^3$. Core radii in this scenario evidently lie around 10^{-5} pc , where the profiles would cross the phase-space barrier.

¹³At such early times, the radiation density also has a significant influence on the growth function, but since our simulations were carried out assuming matter domination, we neglect this effect in order to maintain consistency. Also, baryons do not cluster at the scale of the CDM cutoff, an effect that makes CDM density variations grow slightly more slowly on such scales (Hu & Sugiyama 1996). Since neither our simulations nor the Eisenstein & Hu (1998) power spectrum that we use include this effect, we also neglect it in the growth function.

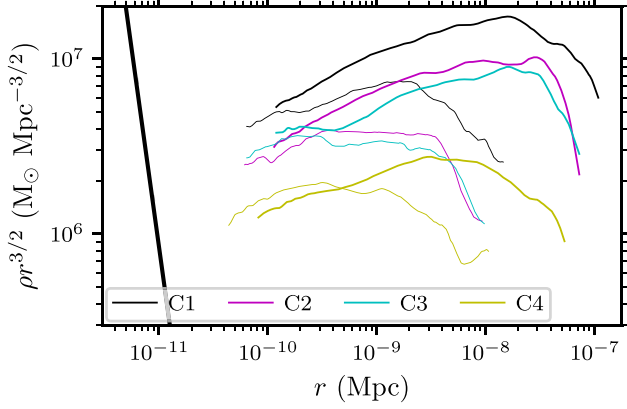


Figure 25. Predicted core sizes for our simulated haloes if they arise in a CDM scenario. We plot the density profiles $\rho(r)$ (scaled by $r^{3/2}$) for the haloes that arise from the $n = -2.67$ power spectrum, and these haloes have been rescaled to correspond to a 100 GeV CDM scenario with kinetic decoupling at 30 MeV. As in Fig. 23, each halo is represented by a different colour, and for each halo we plot the density profile shortly after formation (thin curve) and a factor of about 1.9 in the scale factor later (thick curve). We show on the left the $\rho r^6 = \rho_c r_c^6$ barrier (black) at which each halo’s inner cusp is expected to give way to a central core.

Interestingly, despite the huge difference in scales between the first haloes in these WDM and CDM scenarios, the core radii are very similar when considered as fractions of the halo (or cusp) radii. CDM core radii are only about three times smaller than WDM cores in this relative sense. Like in the WDM case, the $\rho \propto r^{-3/2}$ cusps of CDM haloes can extend across two to three decades in radius in cases where they persist.

5.4 Scaling behaviour

The observation that CDM and WDM core sizes are nearly the same, when considered in relation to the extent of the halo’s inner cusp, merits further discussion. In a general sense, it is expected that core sizes and cusp sizes scale together because the former are set by the dark matter thermal velocity distribution while the latter are set by the drift distances that result from these velocities. However, it is not obvious that their mutual scaling should be so tight.

Consider dark matter particles with mass m that kinetically decoupled at the scale factor a_d during the last radiation-dominated epoch with characteristic (possibly relativistic) momentum p_{0d} . Their characteristic comoving drift distance is

$$\lambda = \int_{a_d}^a da' \frac{\left[1 + \left(\frac{p_{0d} a_d}{m a'}\right)^2\right]^{-1/2}}{a'^2 H(a')} = \left(\frac{8\pi G}{3} \rho_m a_{\text{eq}}\right)^{-1/2} \frac{p_{0d} a_d}{m} \Lambda_{\text{fs}} \quad (13)$$

by the scale factor a , where

$$\Lambda_{\text{fs}} \simeq \ln \left[\frac{8}{1 + \sqrt{1 + (p_{0d}/m)^2}} \frac{a_{\text{eq}}}{a_d} \right] \quad (14)$$

is roughly the logarithm of the factor in a over which particles stream non-relativistically during radiation domination. Here $H(a)$ is the Hubble rate, ρ_m is the matter density at $a = 1$ today, and a_{eq} is the scale factor at matter-radiation equality. In equation (14) we assume the dark matter becomes non-relativistic long before a_{eq} and that we

are evaluating λ at a scale factor $a \gg a_{\text{eq}}$. We also neglect changes in the radiation content, which in general raise Λ_{fs} by a small amount.¹⁴

If R is the characteristic (comoving) size of a peak in the primordial density field smoothed by free streaming, we expect that $R \sim \lambda$. For a halo that collapses at the scale factor a_c from such a peak, we thus expect that the halo’s $\rho \propto r^{-3/2}$ cusp has coefficient $A = \rho r^{3/2} \sim \rho_m a_c^{-3/2} \lambda^{3/2}$ and initial radius $r_0 \sim a_c \lambda$ (see Section 3.2). Now through these relations and equations (11), (12), and (13) we find that

$$r_c/r_0 \sim (a_c/a_{\text{eq}})^{-2/3} \Lambda_{\text{fs}}^{-4/3}. \quad (15)$$

The mutual scaling between core radius r_c and cusp radius r_0 thus emerges, and we see that there are two main influences that can raise r_0 without raising r_c .¹⁵ The first possibility is to raise Λ_{fs} by increasing the duration of non-relativistic drift (or by raising the number of relativistic degrees of freedom in the early Universe). This change raises the free-streaming length without altering particle velocities and hence f_{max} . The second possibility is to form haloes later, which raises the physical free-streaming scale $a_c \lambda$ without altering f_{max} .

CDM and WDM have comparable values of Λ_{fs} (of order 10) and a_c/a_{eq} (of order 100), so they approximately preserve the mutual scaling between r_c and r_0 . However, non-standard early-universe physics can break this scaling to a more serious extent. For instance, certain inflationary (e.g. Ben-Dayan & Brustein 2010) or postinflationary (e.g. Erickcek & Sigurdson 2011) scenarios can greatly boost the power spectrum at small scales, potentially leading to $a_c \sim a_{\text{eq}}$ for the first haloes, or even $a_c < a_{\text{eq}}$ (e.g. Blanco et al. 2019; Delos & Silk 2022). Moreover, if the dark matter decoupled in the presence of a large number of beyond-Standard Model relativistic degrees of freedom, this effect could significantly raise Λ_{fs} . A large injection of entropy after dark matter decoupling, such as that arising from a period of early matter domination (e.g. Allahverdi et al. 2021), has a similar effect. Also, these considerations only apply to haloes that form at the free-streaming scale. Inflationary dynamics (e.g. Barnaby & Huang 2009) can give rise to features in the primordial power spectrum that are completely independent of the free-streaming scale, and similar features can arise in the matter power spectrum due to postinflationary physics (e.g. Hogan & Rees 1988; Erickcek, Ralegankar & Shelton 2021). If such a feature is sufficiently pronounced to form haloes by direct collapse, those haloes could develop $\rho \propto r^{-3/2}$ density profiles with arbitrarily small cores.

¹⁴ H is scaled by $(g_*/3.36)^{1/2} (g_{*s}/3.91)^{-2/3}$, where g_* and g_{*s} are the effective numbers of relativistic degrees of freedom for energy and entropy density, respectively. Assuming only Standard Model content, this factor is 1 at temperatures below about 1 MeV, roughly 0.9 at temperatures between 1 MeV and 100 MeV, and roughly 0.6 at temperatures above about 100 MeV (e.g. Husdal 2016). Since $\lambda \sim H^{-1}$, inclusion of this effect can boost λ by up to a factor of about 2, although typically the effect is far smaller because H^{-1} appears inside an integral. WDM, however, can require large g_* at decoupling to give dark matter the observed relic abundance.

¹⁵ r_c/r_0 is also sensitive to the shape of the initial momentum distribution, both through the normalization factor C (see Section 5.1) and the form of the power spectrum cutoff, which sets the connection between λ and the distribution of peak radii R . The R distribution also depends on the underlying initial power spectrum of density variations. These are order-unity considerations, however.

6 CONCLUSIONS

We have carried out high-resolution cosmological zoom simulations of 12 first-generation haloes of mass corresponding approximately to the small-scale cutoff in the fluctuation power spectrum. The three idealized linear power spectra we considered correspond roughly (but not exactly) to those associated with hot dark matter, WDM, and CDM. Our main findings are the following:

(i) Soon after collapse and for all three initial power spectra, almost all our haloes contained central cusps with $\rho \propto r^{-3/2}$ and with total mass comparable to that of the peak in the linear density field from which the halo originated. We could follow this cusp over as many as two decades in radius. Cusp formation is independent both of collapse morphology and of the prominence of small-scale N-body clumping artefacts in the pre-collapse structure. Collapses through a filament, through a sheet, and through more complex caustic networks all produce similar cusps. We explicitly verified in one case that the cusp profile is stable with respect to changes in the simulation resolution.

(ii) Since the cusp forms so quickly after collapse, its coefficient $A = \rho r^{3/2}$ is sensitive almost wholly to the immediate neighborhood of the halo's precursor density peak. This outcome confirms earlier findings by Delos et al. (2019). In particular, we find the peak's height and spherically averaged characteristic radius, along with the tidal tensor at the location of the peak, suffice to determine A to within 10–20 per cent.

(iii) Haloes that initially possess $\rho \propto r^{-3/2}$ central cusps tend to develop shallower density profiles as they grow. However, unlike previous works that attributed this process to major mergers, we find that it can be attributed to rapid accretion in general (which includes major mergers, minor mergers, and accretion of the smooth background). In fact, this shallowing is consistent with simple models of the sort presented by Dalal et al. (2010) and Ludlow et al. (2013) that connect halo density profiles to their accretion histories. But unlike mergers, rapid accretion tends to build up a halo's density profile at intermediate and large radii without disturbing its inner profile; it does not necessarily disrupt the initial $\rho \propto r^{-3/2}$ cusp.

(iv) Major mergers can disturb the central $\rho \propto r^{-3/2}$ cusp, as Ogiya et al. (2016) and Angulo et al. (2017) found. In two cases, we found that a major merger event reduced the amplitude of the central $\rho \propto r^{-3/2}$ cusp by about 10 per cent (although we do not have the resolution to assess whether the cusp's slope changed). In the absence of mergers, on the other hand, the central cusp remains undisturbed even when accretion is otherwise rapid. We also found a case where even a major merger had no impact on the central cusp. Within our halo sample and resolution limits, the extent to which mergers disrupt the $\rho \propto r^{-3/2}$ cusp remains small.

(v) For haloes that grow sufficiently slowly, the $\rho \propto r^{-3/2}$ central cusp can continue to comprise between 1 and 10 per cent of the total halo mass even long after the halo's formation. Two of the haloes arising from the hot dark matter-like power spectrum maintained a cusp mass fraction in this range over a factor of about 10 in the scale factor. On the other hand, a period of rapid accretion tends to swiftly reduce the cusp mass fraction below 10^{-3} .

(vi) The $\rho \propto r^{-3/2}$ cusp invariably has a nearly isotropic velocity dispersion. Conversely, it exhibits a wide range of three-dimensional shapes across different haloes and even for the same halo across different times. Equidensity surfaces within the cusp tend to align with those of the halo at larger radii.

(vii) It is still unclear how far down in radius the $\rho \propto r^{-3/2}$ cusp can extend, owing to the limits of our simulation resolution. However, if the dark matter was once in contact with the radiation bath, then its

residual thermal motion means that any density cusp must give way to a finite-density core at small enough radii. Some of our simulations probe radii down to as little as three times the expected core radius in a WDM scenario, which suggests that the question of how much farther the $\rho \propto r^{-3/2}$ cusp can extend may be largely immaterial, at least for conventional dark matter models in which the power spectrum's cutoff arises from thermal free streaming. Nevertheless, the core radius is small enough that the $\rho \propto r^{-3/2}$ density profile can be important if it survives.

Survival of the $\rho \propto r^{-3/2}$ central cusp through periods of rapid accretion can lead to density profiles resembling that in Fig. 11. There, a halo with an Einasto density profile at intermediate and large radii still possesses its natal $\rho \propto r^{-3/2}$ cusp at the smallest radii. In larger-scale cosmological simulations, this cusp might go unresolved, which could explain why some works (e.g. Wang et al. 2020) found no evidence of $\rho \propto r^{-3/2}$ cusps despite resolving the smallest haloes. Moreover, this picture demonstrates a new avenue by which the $\rho \propto r^{-3/2}$ cusps of the first haloes may transition into the Einasto density profiles seen at larger scales. Since this mechanism does not destroy the cusp, this discovery raises the prospect that a significant number of $\rho \propto r^{-3/2}$ cusps could survive up to the present day.

We studied only a small halo sample over a limited time range. More work is needed to determine the true extent to which the $\rho \propto r^{-3/2}$ cusps of the first haloes persist through cosmic time. If these cusps survive, they could have observational implications. The abundance of the smallest subhaloes could be larger than previously estimated, owing to the lower efficiency of tidal stripping for steep central cusps (Peñarrubia et al. 2010; Stücker et al. 2022a). For instance, Lee et al. (2021) concluded that the suppression of the smallest CDM subhaloes by tidal stripping pushes them outside the reach of pulsar timing searches. This conclusion may need to be reassessed if $\rho \propto r^{-3/2}$ density profiles persist. Separately from the question of subhalo survival, observable signatures of dark matter haloes, such as through dark matter annihilation (e.g. Ishiyama et al. 2010; Delos & White 2022) and astrometric microlensing (Erickcek & Law 2011), tend to be larger if they have steeper cusps. We also comment that these considerations are amplified for non-standard cosmological initial condition scenarios in which small-scale density variations are boosted (e.g. Erickcek & Sigurdson 2011). The haloes that arise at those boosted scales grow slowly, so their $\rho \propto r^{-3/2}$ cusps may remain a large fraction of the total mass in haloes.

Finally, we have not addressed the question of why the $\rho \propto r^{-3/2}$ density profile arises. This profile appears to be a general consequence of the collapse of a smooth density peak, independent of the details of the collapse process. It has been observed that the Einasto or NFW density profiles of idealized CDM haloes can largely be explained by appealing to their accretion histories (Dalal et al. 2010; Ludlow et al. 2013), an idea that is linked to the notion that material stratifies within the halo based on its accretion time (see Section 4.1). However, this idea cannot explain the $\rho \propto r^{-3/2}$ density profiles of the first haloes; these profiles form immediately upon collapse, before there is any meaningful accretion history. A new idea is needed (e.g. White 2022).

ACKNOWLEDGEMENTS

We thank Neal Dalal, Benedikt Diemer, and Go Ogiya for useful discussions. The simulations for this work were carried out on the Max Planck Institute for Astrophysics's Freya computing cluster at the Max Planck Computing and Data Facility.

DATA AVAILABILITY

Simulation data are available upon reasonable request to the corresponding author.

REFERENCES

- Allahverdi R. et al., 2021, *Open J. Astrophys.*, 4, 1
- Allgood B., Flores R. A., Primack J. R., Kravtsov A. V., Wechsler R. H., Faltenbacher A., Bullock J. S., 2006, *MNRAS*, 367, 1781
- Anderhalden D., Diemand J., 2013, *J. Cosmol. Astropart. Phys.*, 2013, 009
- Angulo R. E., White S. D. M., 2010, *MNRAS*, 401, 1796
- Angulo R. E., Hahn O., Abel T., 2013, *MNRAS*, 434, 3337
- Angulo R. E., Hahn O., Ludlow A. D., Bonoli S., 2017, *MNRAS*, 471, 4687
- Avila-Reese V., Colín P., Valenzuela O., D’Onghia E., Firmani C., 2001, *ApJ*, 559, 516
- Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, *ApJ*, 304, 15
- Barnaby N., Huang Z., 2009, *Phys. Rev. D*, 80, 126018
- Ben-Dayan I., Brustein R., 2010, *J. Cosmol. Astropart. Phys.*, 2010, 007
- Bertschinger E., 2006, *Phys. Rev. D*, 74, 063509
- Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*. Princeton University Press, 41 William Street, Princeton, New Jersey 08540
- Blanco C., Delos M. S., Erickcek A. L., Hooper D., 2019, *Phys. Rev. D*, 100, 103010
- Bode P., Ostriker J. P., Turok N., 2001, *ApJ*, 556, 93
- Busha M. T., Evrard A. E., Adams F. C., 2007, *ApJ*, 665, 1
- Colombi S., 2021, *A&A*, 647, A66
- Dalal N., Lithwick Y., Kuhlen M., 2010, preprint ([arXiv:1010.2539](https://arxiv.org/abs/1010.2539))
- Dalcanton J. J., Hogan C. J., 2001, *ApJ*, 561, 35
- Dehnen W., 2001, *MNRAS*, 324, 273
- Delos M. S., Silk J., 2022, preprint ([arXiv:2210.04904](https://arxiv.org/abs/2210.04904))
- Delos M. S., White S. D. M., 2022, preprint ([arXiv:2209.11237](https://arxiv.org/abs/2209.11237))
- Delos M. S., Erickcek A. L., Bailey A. P., Alvarez M. A., 2018a, *Phys. Rev. D*, 97, 041303
- Delos M. S., Erickcek A. L., Bailey A. P., Alvarez M. A., 2018b, *Phys. Rev. D*, 98, 063527
- Delos M. S., Bruff M., Erickcek A. L., 2019, *Phys. Rev. D*, 100, 023523
- Diemand J., Moore B., Stadel J., 2005, *Nature*, 433, 389
- Diemer B., Kravtsov A. V., 2014, *ApJ*, 789, 1
- Einasto J., 1965, *Tr. Astrofiz. Inst. Alma-Ata*, 5, 87
- Eisenstein D. J., Hu W., 1998, *ApJ*, 496, 605
- Enzi W. et al., 2021, *MNRAS*, 506, 5848
- Erickcek A. L., Law N. M., 2011, *ApJ*, 729, 49
- Erickcek A. L., Sigurdson K., 2011, *Phys. Rev. D*, 84, 083503
- Erickcek A. L., Ralegankar P., Shelton J., 2021, *Phys. Rev. D*, 103, 103508
- Green A. M., Hofmann S., Schwarz D. J., 2005, *J. Cosmol. Astropart. Phys.*, 2005, 003
- Hogan C. J., Rees M. J., 1988, *Phys. Lett. B*, 205, 228
- Hu W., Sugiyama N., 1996, *ApJ*, 471, 542
- Husdal L., 2016, *Galaxies*, 4, 78
- Huss A., Jain B., Steinmetz M., 1999, *MNRAS*, 308, 1011
- Ishiyama T., 2014, *ApJ*, 788, 27
- Ishiyama T., Ando S., 2020, *MNRAS*, 492, 3662
- Ishiyama T., Makino J., Ebisuzaki T., 2010, *ApJ*, 723, L195
- Lee V. S. H., Mitridate A., Trickle T., Zurek K. M., 2021, *J. High Energy Phys.*, 2021, 28
- Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, *MNRAS*, 439, 300
- Ludlow A. D. et al., 2013, *MNRAS*, 432, 1103
- Macciò A. V., Paduroiu S., Anderhalden D., Schneider A., Moore B., 2012, *MNRAS*, 424, 1105
- Mo H., van den Bosch F. C., White S., 2010, *Galaxy Formation and Evolution*. Cambridge University Press, Cambridge CB2 8RU, UK
- Moore B., Quinn T., Governato F., Stadel J., Lake G., 1999, *MNRAS*, 310, 1147
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Navarro J. F. et al., 2004, *MNRAS*, 349, 1039
- Navarro J. F. et al., 2010, *MNRAS*, 402, 21
- Ogiya G., Hahn O., 2018, *MNRAS*, 473, 4339
- Ogiya G., Nagai D., Ishiyama T., 2016, *MNRAS*, 461, 3385
- Peñarrubia J., Benson A. J., Walker M. G., Gilmore G., McConnachie A. W., Mayer L., 2010, *MNRAS*, 406, 1290
- Planck Collaboration, 2020, *A&A*, 641, A6
- Polisenky E., Ricotti M., 2015, *MNRAS*, 450, 2172
- Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, *MNRAS*, 338, 14
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Springel V., Pakmor R., Zier O., Reinecke M., 2021, *MNRAS*, 506, 2871
- Strigari L. E., Bullock J. S., Kaplinghat M., Kravtsov A. V., Gnedin O. Y., Abazajian K., Klypin A. A., 2006, *ApJ*, 652, 306
- Stücker J., Hahn O., Angulo R. E., White S. D. M., 2020, *MNRAS*, 495, 4943
- Stücker J., Ogiya G., Angulo R. E., Aguirre-Santaella A., Sánchez-Conde M. A., 2022a, preprint ([arXiv:2207.00604](https://arxiv.org/abs/2207.00604))
- Stücker J., Angulo R. E., Hahn O., White S. D. M., 2022b, *MNRAS*, 509, 1703
- Tremaine S., Gunn J. E., 1979, *Phys. Rev. Lett.*, 42, 407
- Vera-Ciro C. A., Sales L. V., Helmi A., Frenk C. S., Navarro J. F., Springel V., Vogelsberger M., White S. D. M., 2011, *MNRAS*, 416, 1377
- Villaescusa-Navarro F., Dalal N., 2011, *J. Cosmol. Astropart. Phys.*, 2011, 024
- Wang J., White S. D. M., 2007, *MNRAS*, 380, 93
- Wang J., White S. D. M., 2009, *MNRAS*, 396, 709
- Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020, *Nature*, 585, 39
- White S. D. M., 2022, *MNRAS*, 517, L46
- White S. D. M., Frenk C. S., Davis M., 1983, *ApJ*, 274, L1

APPENDIX A: MIXING OF LOW- AND HIGH-RESOLUTION PARTICLES

Section 2 describes how we prepare our high-resolution halo simulations. After choosing the halo in a low-resolution simulation, we identify the Lagrangian region spanned by halo particles and replace all of the low-resolution particles in that region with much more numerous high-resolution particles. To spare computational expense, however, we do not make the high-resolution region so large as to ensure that low-resolution particles never fall into the high-resolution halo. In this appendix, we show that any numerical artefacts that arise from this choice cannot impact our results.

Mixing of particles of different masses can accelerate two-body relaxation effects (e.g. Binney & Tremaine 2008). Artificial energy exchange between simulation particles causes the more massive particles to gradually sink to the centre of the system while the light particles are dynamically heated. However, we plot in Fig. A1 the total enclosed mass profile $M(r)$ of each of our haloes along with the mass profile contributed by low-resolution particles at the final scale factor a_f . Evidently, each halo’s low-resolution particles contribute 1 per cent or less of the total mass. Moreover, all such particles reside at radii so large that this small mass fraction cannot alter the conclusions of our study. Comparison with Fig. 10 shows that, for all haloes, the low-resolution particles lie outside both the central $\rho \propto r^{-3/2}$ cusp and any shallower portion of the density profile created by rapid accretion.

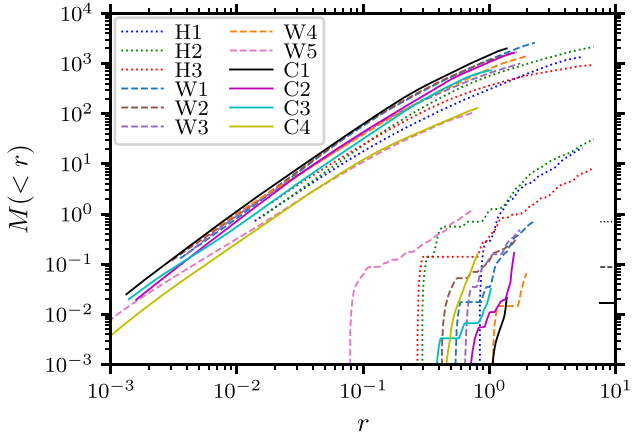


Figure A1. Total enclosed mass profiles of our haloes (diagonal curves through the centre) along with the mass profiles contributed by low-resolution particles only (lower right). We average these profiles over the final factor of 1.17 in the scale factor (i.e. from $a_f/1.17$ to a_f). Black tick marks in the lower right show for each primary (low-resolution) simulation the mass of one low-resolution particle (but since we time average it, the mass profile of low-resolution particles is not required to advance in integer multiples thereof). Evidently, low-resolution particles comprise $\lesssim 1$ per cent of each halo, and all such particles reside at radii too large to affect this study’s conclusions (compare Fig. 10).

APPENDIX B: NUMERICAL CONVERGENCE AT SMALL RADII

We simulated halo W5 at the highest spatial resolution of all of our haloes (see Table 2). In this section we test numerical convergence by simulating the same halo at worse resolution; specifically we raise the simulation particle mass m_p and the force softening length ϵ . We carried out two simulations in which we raise the particle mass by

factors of 8 and 64 and the softening length by factors of 2 and 4, respectively. Fig. B1 plots this halo’s density profile at four different times in the original simulation (cyan) and in both reduced-resolution simulations (red and black). All density profiles transition into a core at small radii, but this core is clearly artificial in that its radius scales with the spatial resolution. The $\rho \propto r^{-3/2}$ part of the density profile, on the other hand, appears to be converged with respect to simulation resolution at all four times.

We find that the radius above which the density profile is converged is reasonably well approximated by 5ϵ , where ϵ is the force-softening length. Fig. B1 depicts the density profile at radii below 5ϵ with a thin dotted line. The idea here is that force softening introduces a bias (e.g. Dehnen 2001) that makes systems less tightly bound, and the relative effect becomes larger as we consider systems closer to the softening scale. However, this criterion somewhat fails at the earliest times for reasons that are unclear. In particular, the solid red and cyan curves in the upper left-hand panel begin to shallow at small radii r even when $r > 5\epsilon$, but this shallowing occurs at different radii for different resolution levels. Consequently, tendencies of some of our haloes’ early density profiles to shallow at the smallest radii, e.g. in Fig. 9, should be interpreted with some skepticism.

Also, while the 5ϵ condition mostly suffices for W5, simulations carried out over a longer duration must also consider two-body relaxation between simulation particles. We evaluate the relaxation time-scale

$$t_{\text{rel}}(r) \simeq \frac{N}{8 \log(r/\epsilon)} \frac{2\pi r}{\sqrt{GM/r}} \quad (\text{B1})$$

(e.g. Power et al. 2003; Binney & Tremaine 2008) as a function of radius r , where M and N are the mass and particle count enclosed within r . The radius r_{rel} at which t_{rel} is equal to the age of the universe (a decent approximation for the age of any halo) then serves as an alternative minimum converged radius. In Fig. B1, radii with $5\epsilon < r < r_{\text{rel}}$ are shown in principle with a thin dashed line, but only at $a = 1.52$ (bottom right) for the lowest-resolution simulation (black) does

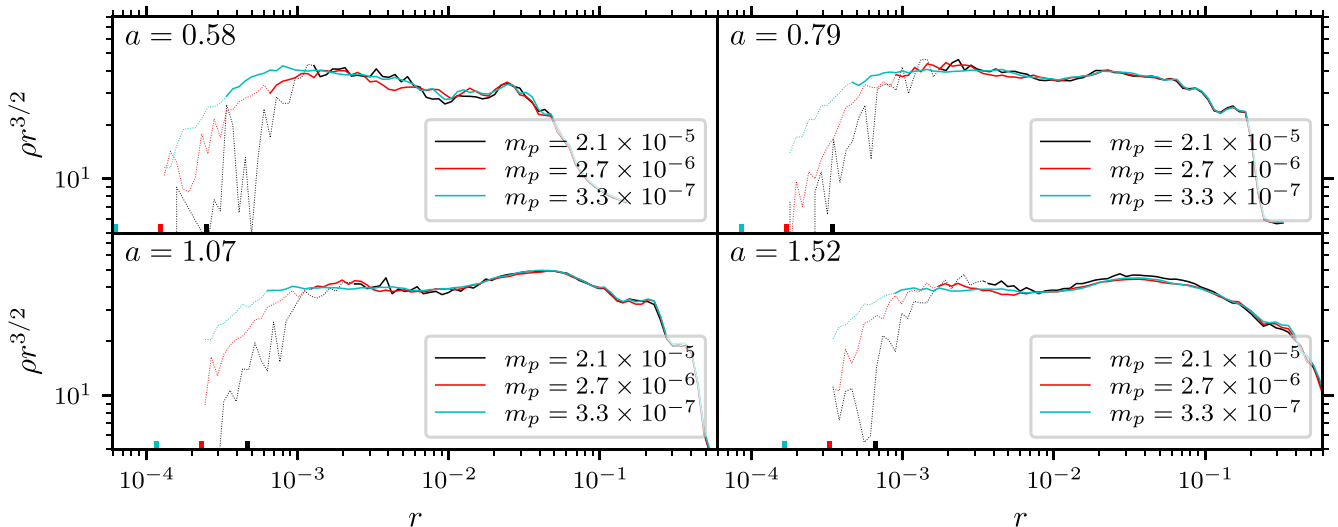


Figure B1. Density profile of the W5 halo at four different times (panels) and three different simulation resolutions (colours), indicated by the particle mass m_p . The force-softening length ϵ is scaled as $m_p^{1/3}$. While the density profiles all transition into a uniform-density core at small radii (indicated by a steeply rising curve in this scaled plot), this transition is evidently artificial because the transition radius scales with the simulation resolution. Meanwhile, the $\rho \propto r^{-3/2}$ behaviour (horizontal) is stable with respect to simulation resolution. We plot the density profile at radii below 5ϵ with a thin dotted line; ϵ itself is marked for each simulation by a coloured tick at the bottom of each panel. We also plot radii above 5ϵ but below the relaxation limit r_{rel} (where the relaxation time-scale becomes shorter than the age of the universe; see the text) with a thin dashed line, although for this halo the only example occurs for $a = 1.52$ (lower right-hand panel) at the lowest simulation resolution (black).

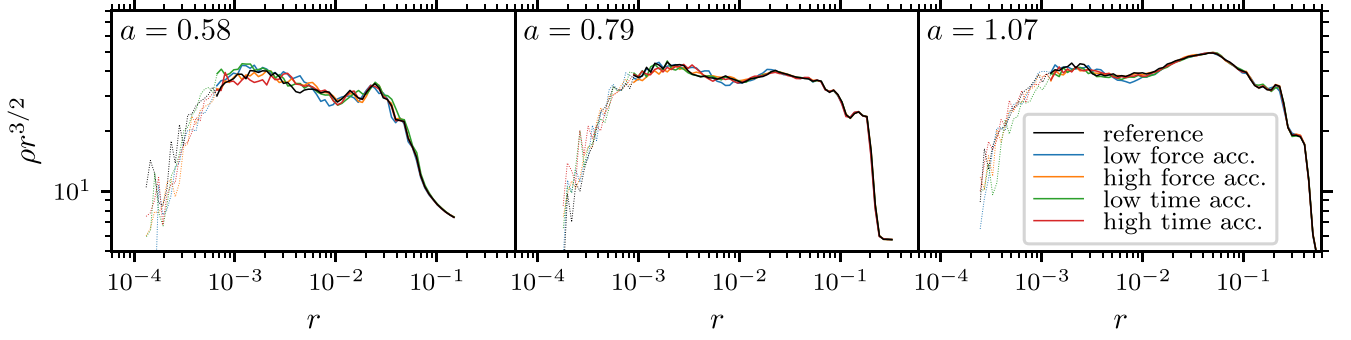


Figure B2. Density profile of the W5 halo at three different times (panels) and a variety of different GADGET-4 numerical parameters (colours). We vary the tree-force accuracy and time integration accuracy as described in Appendix B, but the impact on the density profile is marginal at early times (left-hand panel) and even smaller at later times (other panels).

an example of such a radius appear. This restriction becomes more relevant for haloes simulated for longer durations (see Fig. 3).

Since standard choices of numerical accuracy parameters are not widely validated in simulations of the first haloes, we also briefly explore the impact of GADGET-4’s time stepping and short-range force accuracy parameters (see also Delos et al. 2018b; Ogiya & Hahn 2018). For the simulations in the text, we set the tree-force accuracy parameter to be $\alpha_f = 0.005$, where tree nodes are only opened if the estimated fractional force error from truncating the calculation is more than α_f . We also set the adaptive time stepping parameter to be $\eta_t = 0.025$, which roughly ensures that a particle’s acceleration induces a displacement smaller than $\eta_t \epsilon$ over one time step, and we set the maximum time step to be $[\Delta \log a]_{\max} = 0.01$. In Fig. B2, we separately test $\alpha_f = 0.015$ (low force accuracy); $\alpha_f = 0.002$ (high force accuracy); $\eta_t = 0.075$ and $[\Delta \log a]_{\max} = 0.03$ (low time integration accuracy); and $\eta_t = 0.01$ and $[\Delta \log a]_{\max} = 0.004$ (high time integration accuracy). Evidently, the impact of these changes on the density profile of halo W5 is marginal at early times and essentially negligible at late times.

Finally, we remark that density cusps of the first haloes at least as steep as $\rho \propto r^{-3/2}$ have been confirmed to arise not only in N -body simulations but also in simulations that solve the Vlasov–Poisson equations directly (Colombi 2021). Consequently, they cannot be entirely an N -body artefact.

APPENDIX C: PHASE-SPACE STRUCTURE OF A DENSITY CUSP

In this appendix, we quantify the phase-space structure of a density cusp for the purpose of determining the size of the core that arises from phase-space conservation. Consider an equilibrium spherical dark matter system with the density profile

$$\rho(r) = Ar^{-\gamma}. \quad (\text{C1})$$

The coefficient A is the system’s sole scaling parameter with dimensions of density \times length $^\gamma$. If $\gamma < 3$, then the system’s enclosed mass profile is

$$M(r) = \frac{4\pi}{3-\gamma} Ar^{3-\gamma}. \quad (\text{C2})$$

If $\gamma < 2$, then the system’s potential is

$$\Phi(r) = \frac{4\pi}{(3-\gamma)(2-\gamma)} GA r^{2-\gamma}, \quad (\text{C3})$$

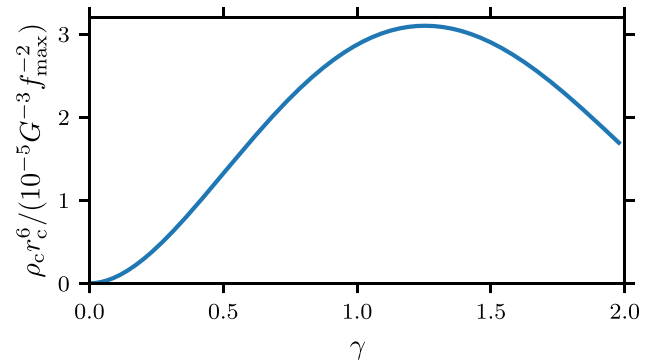


Figure C1. Sensitivity of the phase-space barrier $\rho_c r_c^6$ to the density profile slope γ . In particular, the maximum phase-space density within a $\rho \propto r^{-\gamma}$ density cusp begins to exceed f_{\max} where ρr^6 drops below $\rho_c r_c^6$. For the typical slopes $1 \lesssim \gamma \lesssim 1.5$ observed in this work and others, $\rho_c r_c^6 \simeq 3 \times 10^{-5} G^{-3} f_{\max}^{-2}$.

and dimensional considerations imply that the velocity dispersion has the same scaling, $\sigma^2 \propto r^{2-\gamma}$. Consequently, the distribution function scales as $f \propto \rho/\sigma^3 \propto r^{\gamma/2-3} \propto \Phi^{-1/2-2/(2-\gamma)}$. Let us assume that the distribution function is isotropic, so it depends on energy $E = \Phi + v^2/2$ alone, i.e. $f(r, v) \propto E^{-1/2-2/(2-\gamma)}$. Recall from Section 4.4 that the velocity dispersion inside the central cusp is indeed nearly isotropic. If we also assume that $\gamma > 0$, then we can normalize f by requiring that $\int d^3v f(r, v) = \rho(r)$. Then

$$f(r, v) = (2\pi)^{-3/2} \frac{\Gamma\left(\frac{\gamma}{2-\gamma} + \frac{3}{2}\right)}{\Gamma\left(\frac{\gamma}{2-\gamma}\right)} \frac{\rho(r)}{\Phi(r)^{3/2}} \left[1 + \frac{v^2}{2\Phi(r)}\right]^{-\frac{1}{2}-\frac{2}{2-\gamma}}, \quad (\text{C4})$$

where Γ is the gamma function.

At each radius r , the maximum phase-space density is obtained by setting $v = 0$,¹⁶

$$f(r, 0) = \frac{[(3-\gamma)(2-\gamma)]^{3/2} \Gamma\left(\frac{\gamma}{2-\gamma} + \frac{3}{2}\right)}{2^{9/2} \pi^3 \Gamma\left(\frac{\gamma}{2-\gamma}\right)} G^{-3/2} A^{-1/2} r^{\gamma/2-3} \quad (\text{C5})$$

(where we have substituted equations C1 and C3). If some maximum phase-space density f_{\max} is set by dark matter microphysics, as discussed in Section 5, then the radius r_c at which f_{\max} is reached is

¹⁶Interestingly, equation (C5) implies that spatially denser haloes (with larger A) have lower phase-space density, owing to their higher velocity dispersion.

thus

$$r_c = \left[\frac{(3-\gamma)(2-\gamma)}{8\pi^2} \right]^{\frac{3}{6-\gamma}} \left[\frac{\Gamma\left(\frac{\gamma}{2-\gamma} + \frac{3}{2}\right)}{\Gamma\left(\frac{\gamma}{2-\gamma}\right)} \right]^{\frac{2}{6-\gamma}} (G^3 A f_{\max}^2)^{-\frac{1}{6-\gamma}}. \quad (\text{C6})$$

If we define $\rho_c = \rho(r_c)$ and substitute $A = \rho_c r_c^\gamma$, then

$$\rho_c r_c^6 = \left[\frac{(3-\gamma)(2-\gamma)}{8\pi^2} \right]^3 \left[\frac{\Gamma\left(\frac{\gamma}{2-\gamma} + \frac{3}{2}\right)}{\Gamma\left(\frac{\gamma}{2-\gamma}\right)} \right]^2 G^{-3} f_{\max}^{-2} \quad (\text{C7})$$

defines the general barrier in ρ - r space at which f_{\max} is reached. The right-hand side of this equation depends only on f_{\max} and not on the parameters of the particular halo, other than γ . In fact, within the range $1 < \gamma < 1.5$, $\rho_c r_c^6$ is almost independent of γ as well. We plot its dependence on γ in Fig. C1; evidently,

$$\rho_c r_c^6 \simeq 3 \times 10^{-5} G^{-3} f_{\max}^{-2} \quad (\text{C8})$$

for any halo within this slope range.

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