

The spatial distribution of Milky Way satellites, gaps in streams, and the nature of dark matter

Mark R. Lovell ¹★, Marius Cautun ², Carlos S. Frenk,³ Wojciech A. Hellwing⁴ and Oliver Newton ⁵

¹Center for Astrophysics and Cosmology, Science Institute, University of Iceland, Dunhaga 5, 107 Reykjavík, Iceland

²Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

³Department of Physics, Institute for Computational Cosmology, University of Durham, South Road, Durham, DH1 3LE, UK

⁴Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, PL-02-668 Warsaw, Poland

⁵Univ. Lyon, Univ. Claude Bernard Lyon 1, CNRS, IP2I Lyon/IN2P3, IMR 5822, F-69622 Villeurbanne, France

Accepted 2021 August 20. Received 2021 July 21; in original form 2021 April 14

ABSTRACT

The spatial distribution of Milky Way (MW) subhaloes provides an important set of observables for testing cosmological models. These include the radial distribution of luminous satellites, planar configurations, and the abundance of dark subhaloes whose existence or absence is key to distinguishing among dark matter models. We use the COCO N -body simulations of cold dark matter (CDM) and 3.3 keV thermal relic warm dark matter (WDM) to predict the satellite spatial distribution in the limit that the impact of baryonic physics is minimal. We demonstrate that the radial distributions of CDM and 3.3 keV-WDM luminous satellites are identical if the minimum pre-infall halo mass to form a galaxy is $>10^{8.5} M_{\odot}$. The distribution of dark subhaloes is significantly more concentrated in WDM due to the absence of low mass, recently accreted substructures that typically inhabit the outer parts of a MW halo in CDM. We show that subhaloes of mass $[10^7, 10^8] M_{\odot}$ and within 30 kpc of the centre are the stripped remnants of larger haloes in both models. Therefore, their abundance in WDM is $3\times$ higher than one would anticipate from the overall WDM subhalo population. We estimate that differences between CDM and WDM concentration–mass relations can be probed for subhalo–stream impact parameters <2 kpc. Finally, we find that the impact of WDM on planes of satellites is likely negligible. Comprehensive comparisons with observations will require further work with high resolution, self-consistent hydrodynamical simulations.

Key words: Local Group – dark matter.

1 INTRODUCTION

The distribution of galaxies is a key observable that a viable cosmological model must reproduce in order to be an accurate description of our Universe. The success of the Λ cold dark matter (Λ CDM) model in explaining the properties of the large-scale Universe is remarkable, including the prediction of the existence of baryonic acoustic oscillations (Cole et al. 2005; Eisenstein et al. 2005). The predictions for galaxy distributions on small scales are controversial, with numerous claimed discrepancies between observations and theoretical predictions (see Bullock & Boylan-Kolchin 2017, for a review); these discrepancies could be removed by baryon physics, dark matter physics, or a combination of the two.

In this paper, we will consider three key observables related to the spatial distribution of satellite galaxies in the context of dark matter studies: the radial distribution of the Milky Way (MW) luminous satellite galaxies, the radial distribution of dark satellites, and the apparent distribution of some galaxies in planar structures.

First, we discuss the radial distribution of MW satellites. This distribution is known to be more concentrated than the distribution of massive subhaloes in many MW-analogue CDM simulations. The reasons for this discrepancy could include that the time-scale to

sink to the centre of the host is anticorrelated with the mass of the galaxy at infall, with more massive subhaloes more likely to host satellite galaxies; thus much of this discrepancy disappears when predictions are made for luminous satellites as opposed to simply massive subhaloes (Font & et al. 2011; Lovell et al. 2017; Newton et al. 2018; Bose et al. 2020; Samuel et al. 2020). It is likely that our census of massive satellites at distances >100 kpc from the MW is incomplete, in which case the satellite distribution would be less concentrated than currently thought; that the distribution of M31 satellites is consistent with both N -body and hydrodynamical simulation predictions supports this hypothesis (Yniguez et al. 2014; Font, McCarthy & Belokurov 2021). There is also the possibility that the accretion of the Magellanic Clouds has introduced an exceptionally large number of dwarf galaxies on orbits that run atypically close to the MW centre (Santos-Santos et al. 2021).

A more speculative, and certainly interesting problem for the purposes of understanding the particle nature of dark matter, is the distribution of dark satellites: subhaloes that have had all of their baryons evaporated by reionization radiation before they could form any stars (Bullock, Kravtsov & Weinberg 2000; Benson et al. 2002). The CDM model predicts the existence of such subhaloes down to scales of $1 M_{\odot}$ or even smaller (Schneider, Smith & Reed 2013). The presence of such subhaloes has been the subject of study in two contexts: the lensing of background galaxies and quasars by elliptical galaxies hosted in $\sim 10^{13} M_{\odot}$ haloes, where some detections have

* E-mail: m.r.lovell@durham.ac.uk

been reported (Vegetti et al. 2010; Hezaveh et al. 2016) and, more pertinently for this paper, stellar streams around the MW where interactions of dark subhaloes may have ripped gaps in stellar streams (Yoon, Johnston & Hogg 2011; Carlberg & Grillmair 2013; Erkal & Belokurov 2015; Bonaca et al. 2019). Detections of individual gaps have been reported, although it is not clear whether such gaps are indeed due to dark matter subhaloes or instead to some denser object (Amorisco et al. 2016; Bonaca et al. 2019).

Finally, the apparent location of some MW satellites in a planar structure is a particularly puzzling discrepancy. There is considerable evidence that some luminous satellites are located in a flat plane (Lynden-Bell 1976; Libeskind et al. 2005; Metz, Kroupa & Libeskind 2008; Metz, Kroupa & Jerjen 2009; Lux, Read & Lake 2010; Pawlowski & Kroupa 2020). Whether or not such planes are rotationally coherent is much more difficult to either confirm or refute given the challenges of obtaining proper motions for satellites with sufficient precision (Santos-Santos, Domínguez-Tenreiro & Pawlowski 2020) and due to the satellites following complex orbits that are affected by the non-spherical shape of the host halo and other massive satellites (Shao, Cautun & Frenk 2019). The likelihood of having a MW-like plane of satellites is still an open debate, ranging from 0.1 per cent (Pawlowski et al. 2014; Shao et al. 2019) to as high as 10 per cent when accounting for the diversity of potential Λ CDM planes of satellites and the look-elsewhere effect (Cautun et al. 2015). In CDM, planes of satellites arise from the anisotropic accretion of satellites, including group infall and filamentary accretion (Libeskind et al. 2005, 2007; Kang et al. 2005; Zentner et al. 2005; Li & White 2008; Lovell et al. 2011; Shao et al. 2018) and chance alignments (Cautun et al. 2015); the destruction of satellites on radial orbits by the disc may also play a role, although not to a sufficient degree to alleviate the tension with observations (Ahmed, Brooks & Christensen 2017).

To a first approximation, the radial distribution of satellites is determined by fluctuations on a ~ 1 Mpc scale, which are the same in both CDM and warm dark matter (WDM) with thermal relic masses > 2 keV. However, there are important secondary effects that can lead to differences, especially for subhaloes whose mass is close to the WDM cutoff in the power spectrum. The primary difference in halo properties between CDM and WDM models is that in the latter, the overall abundance of structure and therefore the number of dwarf haloes is suppressed. A second effect is the change in the internal properties of WDM haloes close to the mass cut-off scale, which are characterized by lower densities (Colín, Avila-Reese & Valenzuela 2000; Lovell et al. 2012) and thus can potentially experience faster mass loss and disruption when orbiting around a more massive host such as the MW. However, it has been shown that the switch from one model to the other also affects the radial distribution of low mass subhaloes, both in MW-analogue haloes (Lovell et al. 2014; Bose et al. 2017) and in $\sim 10^{13} M_{\odot}$ haloes that are used for lensing constraints (Despali et al. 2020). In both cases, the radial distribution of subhaloes with masses below the WDM characteristic half-mode mass, M_{hm} , is steeper in WDM than in CDM, which can be seen as a change in the radial distribution of satellite galaxies if the half-mode mass is high enough. As we discussed above, the analysis of gaps in streams is a popular test of WDM (e.g. Banik & Bovy 2019; Benito et al. 2020; Banik et al. 2021), and it is important to test whether any differences in radial distribution will either compound or compensate for the impact of the pure suppression of subhalo abundance in WDM.

Finally, one can envisage a situation in which the nature of the dark matter can affect the presence of satellite planes. The phase space distribution of satellites is influenced by group infall (Li &

White 2008; Shao et al. 2018) and filamentary accretion (Libeskind et al. 2005; Lovell et al. 2011). Haloes that form in low-density regions, such as lower mass filaments, are delayed in their collapse times, such that their gas is photo-evaporated by reionization before collapse and they can no longer form stars; they will also be less massive. Lovell et al. (2019) showed that the introduction of a power spectrum cutoff makes this delay longer than is the case for CDM, in a manner that is anticorrelated with the local density. If the WDM model is sufficiently extreme, the formation of galaxy-mass haloes in low-mass filaments may be inhibited, restricting infall of dwarf galaxies to a small number of massive filaments and thus a greater proportion of dwarf galaxies will be members of planes. The crucial question is whether the cut-off scale at which this effect happens is above or below the cut-off scale that is consistent with current bounds on WDM (Newton et al. 2020; Enzi et al. 2021; Nadler et al. 2021); this problem has not previously been addressed.

In this study, we use the COCO N -body simulations (Bose et al. 2016; Hellwing et al. 2016) to examine the spatial configurations of satellites in the WDM and CDM cosmologies. We will extract MW-analogue haloes that are produced in both simulations, and analyse the infall history of their subhaloes. We will briefly discuss the radial distribution of subhaloes as pertains to dwarf galaxies, devote the bulk of our analysis to the distribution and concentration of dark subhaloes, and present a brief coda on the question of planes of satellites.

A definitive study would also account for the contraction of the host halo and the presence of the stellar disc in destroying subhaloes on radial orbits (Blumenthal et al. 1986; Gnedin et al. 2004; Ahmed et al. 2017; Garrison-Kimmel et al. 2017; Sawala et al. 2017; Richings et al. 2020). This would require numerous hydrodynamical simulations and is beyond the scope of this paper: we will therefore highlight the limitations of our N -body approach throughout the text as necessary, including in a dedicated subsection (Section 3.4).

This paper is organized as follows. We describe our simulations in Section 2, present our results in Section 3, and draw conclusions in Section 4.

2 SIMULATIONS

The simulations used in this paper are the COCO N -body cosmological simulations. These are cosmological zoomed simulations of a high resolution sphere, ~ 24.7 Mpc in radius, embedded within a periodic cube of side-length 100 Mpc. The high resolution simulation particle mass is $1.6 \times 10^5 M_{\odot}$, the gravitational softening length is $\epsilon = 0.33$ kpc, and the cosmological parameters are consistent with the WMAP7 cosmology (Komatsu et al. 2011): Hubble parameter $h = 0.704$, matter density $\Omega_0 = 0.272$, dark energy density, $\Omega_{\Lambda} = 0.728$; spectral index, $n_s = 0.967$; and power spectrum normalization $\sigma_8 = 0.81$. The simulations were performed with the P-GADGET3 galaxy formation code, which is an updated version of the publicly available GADGET2 code (Springel et al. 2005). Haloes are identified with the friends-of-friends algorithm, and are deconstructed into subhaloes with the SUBFIND halo finder (Springel et al. 2001). The minimum number of particles required to identify a subhalo is 20, which is a mass of $3.2 \times 10^6 M_{\odot}$.

Two copies of the COCO volume were simulated: a CDM version (Hellwing et al. 2016) followed by a WDM counterpart (Bose et al. 2016). The WDM model is for a thermal relic with particle mass 3.3 keV, and has been implemented using the transfer function of Viel et al. (2005). The model can also be described by the half-mode mass, M_{hm} , which is defined as the mass corresponding to the sharp k -space filter where the amplitude of the WDM transfer function falls

to half of the CDM transfer function. For this 3.3 keV thermal relic, $M_{\text{hm}} = 2.8 \times 10^8 M_{\odot}$, which is approximately the same M_{hm} as for a resonantly produced sterile neutrino (Asaka, Shaposhnikov & Laine 2007; Laine & Shaposhnikov 2008; Lovell et al. 2016), of mass of 7 keV and lepton asymmetry, $L_6 = 10$ (Lovell 2020). WDM subhaloes that originate from spurious fragmentation of filaments are removed according to the method presented in Lovell et al. (2014). Briefly, we identify the shapes of haloes in the initial conditions – otherwise known as the shapes of the ‘protohaloes’ – and exclude those for which the protohalo sphericity is lower than 0.165. We also remove haloes for which the peak mass over the history of the simulation is lower than half the limiting mass, M_{lim} , which is a function of the WDM power spectrum and the simulation resolution: for the COCO volume this mass cutoff takes the value $\sim 3.8 \times 10^7 M_{\odot}$.

We now discuss our strategy and procedure for selecting host haloes. Our goal is to select haloes that are of a similar mass to the MW, that are relaxed systems, and that are isolated. The first criterion is based on virial mass, for which we use M_{200} . This is the mass enclosed inside a radius within which the density is 200 times the critical density required for collapse; this radius is labeled r_{200} . For our halo relaxation criterion, we use the ratio of the mass of the central smooth halo to the total mass gravitationally bound to the halo as measured by SUBFIND; see Neto et al. (2007) for a discussion on halo relaxation criteria. Third, we introduce an isolation criterion by requiring that a candidate halo is a minimum distance from any other halo of $M_{200} > 1 \times 10^{12} M_{\odot}$. We identify haloes in the CDM halo catalog that have a virial mass, $M_{200} = [1.0, 1.5] \times 10^{12} M_{\odot}$, a smooth-to-total mass ratio > 0.8 , and a minimum separation from other $M_{200} > 1 \times 10^{12} M_{\odot}$ haloes of 2 Mpc. Note that this last criterion excludes Local Group-like systems given the MW-M31 distance is ~ 0.75 Mpc. We then use the halo position and M_{200} to find matches to these haloes in the WDM simulation; our final sample consists of 24 hosts. We have checked these matches by comparing the density profiles of WDM and CDM hosts, and find that they agree to better than 5 per cent. We therefore highlight that massive haloes have the same profiles in WDM as in CDM.

3 RESULTS

We present our results in five subsections. We first discuss the radial distribution of satellites in general (Section 3.1), and then in a simple application to gaps-in-streams physics (Section 3.2). We examine the mass-concentration relations of subhaloes in Section 3.3, and in Section 3.4 discuss the potential impact of baryonic processes and the caveats these impose on our predictions for subhalo properties in the host inner regions. We end our results presentation with a discussion on planes of satellites in CDM and WDM in Section 3.5.

3.1 Radial distributions

We begin by computing the radial distribution of our MW analogue satellites. We separate our satellites into four bins by $z = 0$ subhalo mass, which we define as the mass gravitationally bound to the satellite as computed by SUBFIND; we denote this mass as M_0 . We then stack the subhaloes from all hosts into a single set of radial distributions. We present our results for CDM and WDM in Fig. 1.

The radial distributions of satellites more massive than $10^{8.5} M_{\odot}$ are largely the same in the two models. At lower masses – which correspond to subhaloes less massive than $3 \times 10^8 M_{\odot}$, i.e. M_{hm} – a strong discrepancy appears. For example, the normalized radial density of $M_0 = [10^{7.5}, 10^{8.0}] M_{\odot}$ WDM subhaloes at 30 kpc is twice that in CDM at radii < 70 kpc, and in the $[10^{7.0}, 10^{7.5}]$ bin the

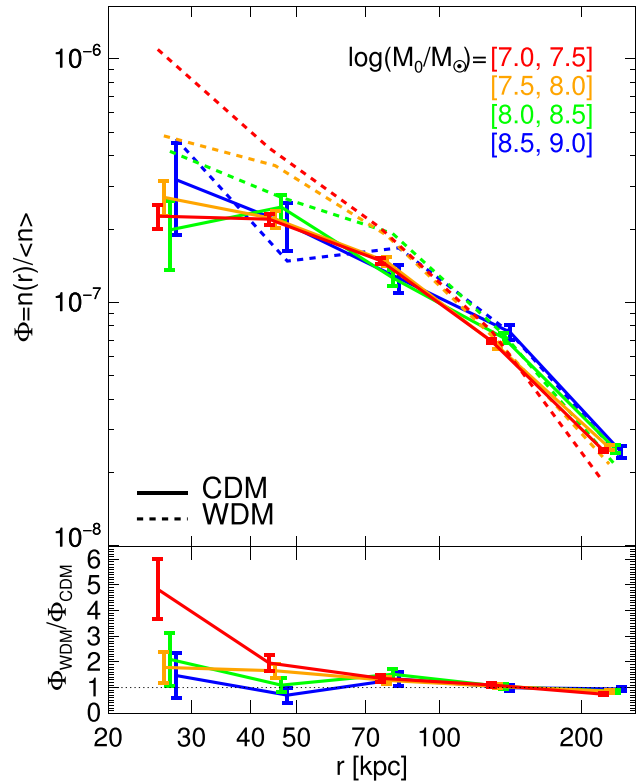


Figure 1. Top panel: the radial distribution of subhaloes in MW halo analogues, normalized by the number of subhaloes in each mass bin. CDM data are represented with solid curves and WDM data with dashed curves. The curve colour associated with each mass bin is given in the figure legend. Error bars are shown for CDM only. Bottom panel, the ratio of the WDM and CDM radial distributions. The error bars are the propagation of Poisson errors on both data sets.

normalized abundance is a factor of five times the CDM equivalent.¹ We thus recover the intriguing results found in Bose et al. (2017) and Despali et al. (2020).

Crucially, the difference between CDM and WDM correlates with halo mass, with bigger discrepancies between the two for mass bins $M_0 < M_{\text{hm}}$. The population of subhaloes in a particular mass bin is composed of two broad classes of objects. First, there are low mass haloes that were accreted and have experienced very little, if any, stripping. The second class are haloes that were much more massive at accretion and have since lost mass to tidal stripping. The relative number of these haloes will change between CDM and WDM, but it is less clear whether the suppression in WDM subhalo densities influences the process of dynamical friction. We first check for differences in subhalo processing, and then consider the difference between the contributions of stripped high mass and unstripped low mass subhaloes to each M_0 bin.

Chandrasekhar (1943), Lacey & Cole (1993), and Simha & Cole (2017) showed that the stripping time of the satellite becomes smaller as the mass ratio between host and satellite tends to 1. If the host halo density profiles were the same in the two models, we would expect that CDM and WDM subhaloes of the same infall mass should have

¹Note that the *total* number of WDM subhaloes in this mass bracket and radius range is still smaller than the number of CDM subhaloes in these same brackets: it is only the density normalized to the number of all WDM subhaloes that is larger.

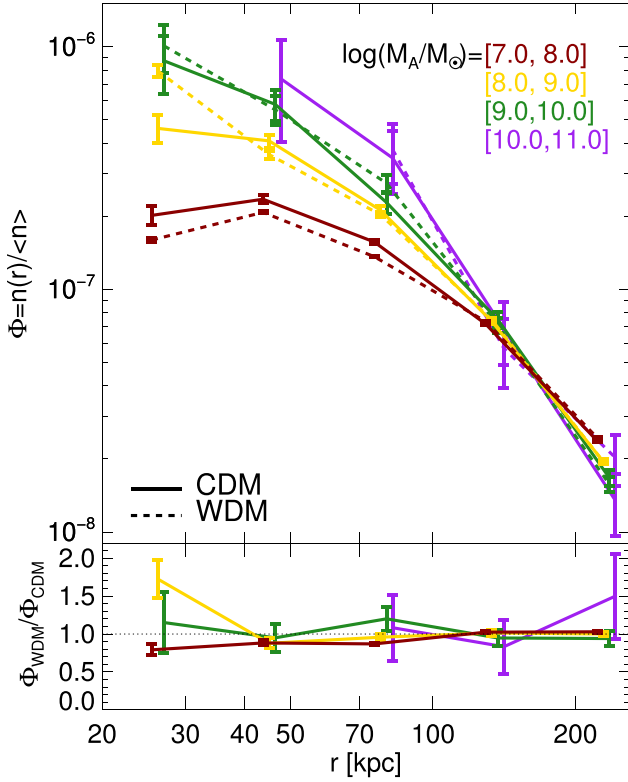


Figure 2. Top panel: the radial density profile of subhaloes binned by accretion mass, M_A . CDM curves are shown as solid lines and WDM curves as dashed lines. The relationship between curve colour and M_A bin is given in the figure legend. Error bars are Poisson. Bottom panel: the ratio of the two data sets.

the same dynamical friction time and therefore not contribute to the discrepancy in radial profiles.

We first check the assertion that the model does not impact the host density profile. We computed density profiles for each of the 24 pairs of hosts, and calculated the ratio of WDM-to-CDM densities: we found that the median WDM-to-CDM profile enhancement at any radius is 3 per cent, and 84 per cent of hosts show a WDM-to-CDM enhancement of less than 10 per cent at $r > 20$ kpc. We then check the hypothesis that dynamical friction is similar for CDM and WDM subhaloes by computing the radial distributions of subhaloes binned by the gravitationally bound mass at accretion, M_A , instead of M_0 . We define the time of accretion as the first time that the subhalo passes within r_{200} of its $z = 0$ host; for subhaloes that are within 300 kpc but never enter r_{200} , $M_A = M_0$. We plot the result in Fig. 2. Note that we use a different set of mass bins to Fig. 1.

There is a sharp distinction between the different M_A bins. The $M_A = [10^{10}, 10^{11}] M_\odot$ subhalo radial distribution is steeper than the $M_0 = [10^{9.5}, 10^{10}] M_\odot$ curve, and decreasing the bin mass results in a progressively shallower radial distribution (see Han et al. 2016). By contrast, there is no clear evidence for a systematic deviation between the CDM and WDM cases, which implies that satellites of a given M_A experience a similar degree of dynamical friction regardless of whether they are WDM or CDM. There is an apparent difference at radii < 40 kpc in the $[10^8, 10^9] M_\odot$ bin, with WDM subhaloes 60 per cent more abundant than in CDM, but the effect shrinks to 20 per cent at the lowest mass bin. We have repeated this process with the peak mass, M_P , which we define as the peak mass obtained by each subhalo across its merger tree main branch, and we find the

results are practically identical as for M_A . We note that subhaloes can survive more easily in the absence of a stellar disc: if a disc were to be included, the distributions would become less concentrated. Therefore our results can be considered to be an upper bound on the concentration of the satellite distribution.

We emphasize this result with an image of one of our haloes, devised to show the distribution of subhaloes with massive progenitors. For each of the subhaloes located within 300 kpc of the host halo centre, we identify its $z = 0$ particles and label those particles with the subhalo progenitor’s M_A . We then generate images such that particles whose progenitor subhaloes are more massive than $\sim 10^8 M_\odot$ are highlighted in red/yellow while all other particles – those that are members of the smooth halo, those with progenitors less massive than $10^8 M_\odot$, and those outside 300 kpc, are coloured in blue. We present the result in Fig. 3.

The CDM simulation shows a large number of subhaloes of various masses in the host’s central region, many of which are coloured as having a massive progenitor. The outskirts of the halo are occupied by many haloes of a similar apparent size that do not have a massive progenitor (i.e. are not coloured red/yellow); therefore the $[10^{7.5}, 10^{8.5}] M_\odot$ mass bracket is a combination of high-accretion mass–small radius haloes and low-accretion mass–large radius haloes. By contrast, the WDM version exhibits almost solely high infall mass subhaloes – as indicated by the colour – in the host centre and a paucity of similarly-sized subhaloes elsewhere; as a result the radial distribution is much steeper than for CDM.

We have argued that, to first order, dynamical friction proceeds in approximately the same manner in WDM as in CDM, i.e. that the radial distributions of CDM and WDM satellites binned by M_A show little, if any, difference, and therefore turn to the second model feature that influences satellite distributions: the relative contributions of different infall masses to each final $z = 0$ mass bin. We select subhaloes in two M_0 bins – $[10^{7.5}, 10^{8.5}] M_\odot$ and $[10^{8.5}, 10^{9.5}] M_\odot$ – and compute those subhaloes’ cumulative M_A mass functions. We present the results in Fig. 4.

In this variable, we identify clear differences between the two models. The median M_A of WDM haloes in the lower M_0 bin is a factor of two larger than CDM haloes, whereas for the higher mass bin it is instead ~ 20 per cent larger. 30 per cent of WDM satellites with $M_0 = [10^{7.5}, 10^{8.5}] M_\odot$ have been stripped into the bin and 5 per cent are the stripped remnants of $M_A > 10^9 M_\odot$ progenitors, compared to only 10 per cent and 1 per cent, respectively, of CDM haloes in this mass bracket. This compares to 15 per cent (10 per cent) of WDM (CDM) subhaloes stripped into the $M_0 = [10^{8.5}, 10^{9.5}] M_\odot$ bin.

In addition to the more rapid stripping time of massive subhaloes in general relative to lower mass subhaloes, we also investigate the phenomenon that WDM subhaloes are stripped more easily than their CDM counterparts with the same accretion mass due to their lower concentrations (see also Bose et al. 2017). We will consider the difference in concentration in the next subsection; here we instead do a simple test of how much the halo mass changes. We compute the ratio of present day mass, M_0 , to accretion mass M_A to obtain median stripping rates; note that we only include subhaloes that survive to $z = 0$. These rates are strongly influenced by infall time: we therefore use subhaloes in two infall time bins: $t_A = [3, 4]$ and $t_A = [7, 8]$ Gyr. We demonstrate the differences in mass loss caused by infall time and dark matter model in Fig. 5.

The median WDM subhalo loses approximately 15 per cent more of its initial mass than the CDM equivalents of the same mass. This figure is roughly consistent across all masses and at both infall time bins: the median amount of mass lost in WDM (CDM) is 70 per cent

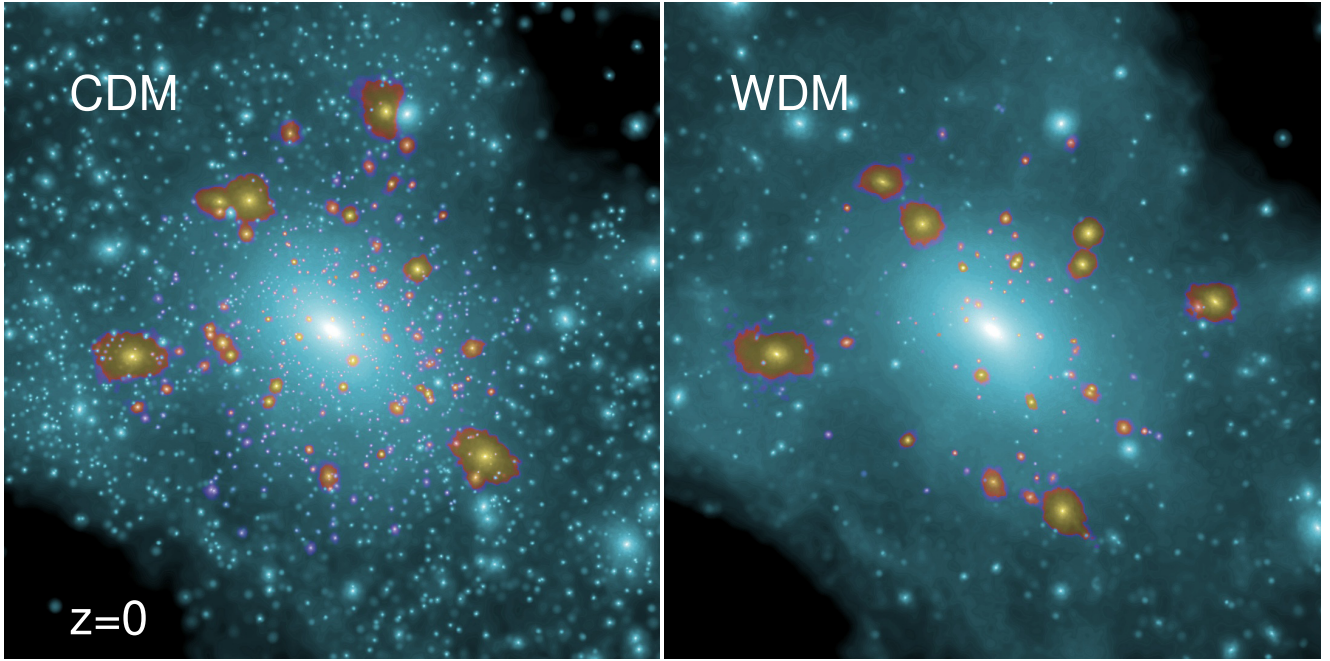


Figure 3. Images of the CDM (left) and WDM (right) realizations of one of our haloes at $z = 0$. Each image is ~ 450 kpc on a side. Subhaloes within 300 kpc of the host halo centre whose progenitor was more massive at $\sim 10^8 M_\odot$ are highlighted in red/yellow and all other particles are coloured in blue.

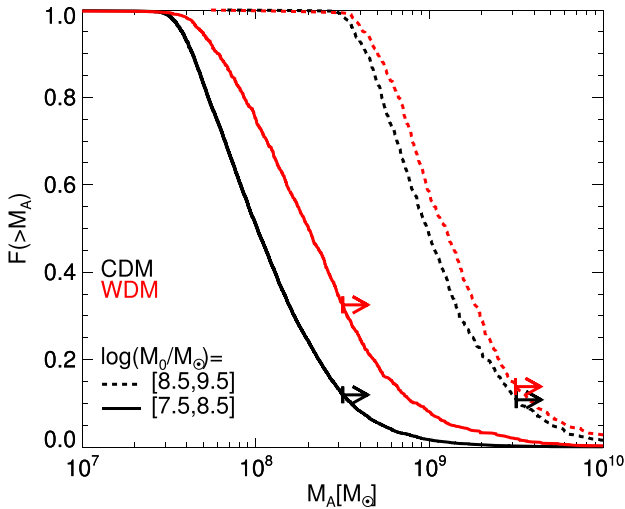


Figure 4. Normalized cumulative accretion mass functions for fixed $z = 0$ mass bins. CDM data are shown in black and WDM in red. The bin with subhaloes of $M_0 = [10^{8.5}, 10^{9.5}] M_\odot$ is shown with dashed lines and that for $M_0 = [10^{7.5}, 10^{8.5}] M_\odot$ with solid lines. The arrows mark the upper edges of each M_0 bin, and therefore subhaloes to the right of the arrow have to have been stripped.

(60 per cent) for the early infall bin and 40 per cent (25 per cent) in the late infall bin.

In conclusion, we recovered the result that the CDM and WDM models predict significantly different $\sim 10^8 M_\odot$ subhalo radial distributions. This difference arises from the sum of two contributing phenomena. First, the absence in WDM of haloes with low accretion mass and thus long dynamical friction time-scales, and secondly, the WDM haloes experience enhanced stripping, which shifts massive,

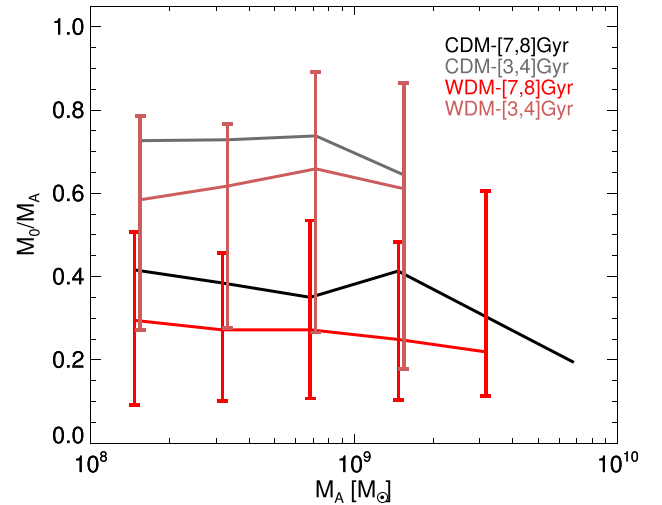


Figure 5. Ratio of present-day mass to accretion mass as a function of accretion mass. Curves are median relations, and the error bars denote the data 68 per cent region. Results in the [7,8] Gyr lookback time bin are shown in black for CDM and red for WDM, and in the [3,4] Gyr lookback time bin are shown in grey for CDM and brown for WDM. We select from all subhaloes within 300 kpc of the host centre. Error bars indicate the 68 per cent data region, and are shown only for the WDM data sets.

rapid-merger time subhaloes into mass bins that are lower than would be the case for CDM.

3.2 Radial distribution: implications for the number of stream perturbors

A detailed analysis of the impact of subhaloes on streams requires that we take into account adiabatic contraction of the host halo (Blumenthal et al. 1986; Gnedin et al. 2004), the destruction of

subhaloes by the MW stellar disc (Ahmed et al. 2017; Garrison-Kimmel et al. 2017; Sawala et al. 2017; Richings et al. 2020), and mass loss due to evaporation of baryons from subhaloes by supernova feedback and reionization (Sawala et al. 2016), all of which are beyond the scope of this study based on N -body simulations. We therefore restrict ourselves to a qualitative comparison to pure N -body work.

Initial published constraints using stream gaps have made the assumption that the spatial distribution of subhaloes is the same in WDM as in CDM. While this assumption appears to work well for massive, luminous satellites (Newton et al. 2020), it breaks down for dark, lower mass haloes as shown above and in Despali et al. (2020). We investigate the implications of this result by comparing the CDM and WDM subhalo populations in the halo region where stellar stream gaps are analysed versus the population of all subhaloes out to 300 kpc.

Banik et al. (2018) estimate that the GD-1 stream has a pericentre of ~ 14 kpc and an apocentre of ~ 30 kpc, and for Pal 5 (Banik & Bovy 2019) estimate these parameters to be >4 and ~ 14 kpc, respectively. Our goal is to compare the population of subhaloes in this small region of the host halo with the population of subhaloes across the host at large. For this purpose, we generate subhalo samples according to distance to the host centre and present-day mass. Given the desire to obtain a reasonably sized subhalo sample, we select for our ‘stream-generating’ subhaloes that are within 40 kpc of the centre of the host and for the host at large we select all subhaloes within 300 kpc, including those already present in the 40 kpc sample. In each radius selection, we count the number of subhaloes in two M_0 mass bins – $[10^7, 10^8] M_\odot$ and $[10^8, 10^9] M_\odot$ – and compute the ratio of the number of WDM subhaloes to the number of CDM subhaloes for each CDM–WDM host pair. We also compute the total mass in subhaloes in these two radial bins and the WDM/CDM ratios for each host halo.

This calculation reflects contributions from the two differences between CDM and WDM that we have explored: enhanced stripping in WDM as shown in Fig. 5 and also the difference in the halo mass function. We analyse the contribution of these two properties to the gaps-in-streams halo population by generating a ‘hybrid’ data set that factors out the difference in the halo mass function as follows.

For each CDM subhalo, we select at random a WDM counterpart that has the same M_A to within 50 per cent, the same $z = 0$ distance to the host centre to within 50 per cent, and the same infall time to within 0.5 Gyr. We then regard the M_0 of the WDM counterpart as a ‘WDM-equivalent stripped mass’. This quantity could be thought of as the present-day mass that the CDM subhalo would have had if it had a (low-concentration) WDM density profile instead of a (high-concentration) CDM density profile. For our hybrid data set, we have the total number of CDM subhaloes but with WDM stripping rates. We therefore repeat the process discussed in the previous paragraph but replace the WDM data with our hybrid data set; however, given the small number statistics involved in the $[10^8, 10^9] M_\odot$ – <40 kpc bin, we do not include hybrid data for that bin. We present the results in Fig. 6. Note that the CDM–WDM comparison contains contributions from both excess stripping and the halo mass function difference, whereas the CDM–hybrid only contain the contribution from the excess stripping.²

²We have chosen to select counterpart subhaloes purely by accretion mass. The predictions for individual subhaloes could be made more accurate if we selected by further variables such as eccentricity and the time of accretion;

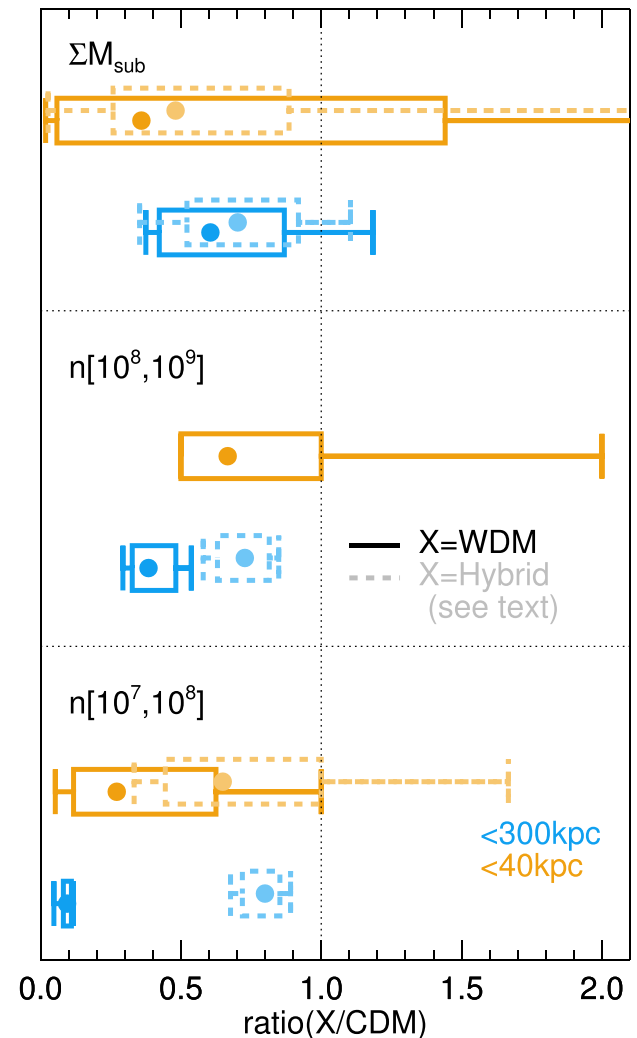


Figure 6. The ratios of substructure mass and subhalo abundance in WDM with respect to CDM for the 24 hosts. Subhalo abundance ratios in the $[10^7, 10^8]$ bin are indicated in the bottom panel, in the $[10^8, 10^9]$ bin in the middle panel, and total mass in substructure in the top panel. Results for these ratios measured for subhaloes within 300 kpc of the host centre are shown in blue, and within the inner 40 kpc in orange. For each data set, the box marks 68 per cent of hosts and the error bars mark the minimum and maximum ratio values. Median values are shown as dots. In addition to the ratios of WDM-to-CDM, shown as solid lines, we include the hybrid calculation as described in the text with faded, dashed-line boxes (top and bottom panels only).

For the full data set, the number of WDM subhaloes in the $[10^8, 10^9]$ ($[10^7, 10^8]$) mass bins within 300 kpc of the host halo centre is 40 per cent (10 per cent) of the CDM number, and the scatter between haloes is a factor of 10 per cent (<5 per cent). Within the 40 kpc sphere the median suppression in the $[10^7, 10^8]$ bin is 30 per cent of CDM, so the CDM-relative subhalo abundance is three times higher than in the halo as a whole. The scatter between hosts is large, but the vast majority have higher ratios than for the <300 kpc bin. The $[10^8, 10^9]$ bin contains ≤ 4 subhaloes per host; therefore while we can report that of the host haloes that have at least one WDM and one CDM subhalo in this bin, the median number ratio is ~ 0.7 and thus

we assume that because we are assembling each host’s subhaloes into a single figure that these details will not affect our result.

0.4 points higher than the whole halo sample, the statistics here are poor. For the subhalo mass ratio, we obtain the opposite result, with 60 per cent WDM-to-CDM for the halo as a whole and 40 per cent for the inner 40 kpc.

We can determine the relative contribution of the enhanced halo stripping in WDM to the suppression of the halo mass function by comparing these results to the hybrid model calculations. In the <300 kpc population, the halo abundance suppression in the hybrid model result is significantly weaker than in the full WDM case – the abundance of subhaloes is 80 per cent of the CDM figure at $[10^7, 10^8]$ and 70 per cent at $[10^8, 10^9]$. The median hybrid-to-CDM ratios in the $[10^7, 10^8]$ mass bin is 68 per cent. Therefore, the suppression in the abundance ratios for the stream-generating subhaloes receives a stronger contribution from stripping relative to mass function suppression than is the case for subhaloes in general. Finally, we note that the similar results are obtained for the total mass in subhaloes in the two radial populations, although this measurement is strongly influenced by the massive subhaloes and therefore the difference is much less pronounced than our two halo abundance mass bins.

We caution that in the inner region the limitations of the subhalo finder may play a role. For example, Onions et al. (2012) demonstrated that SUBFIND may potentially underestimate the number of subhaloes near the halo centres by a factor of two. The relative abundance of CDM subhaloes in these mass bins would increase if either: (i) the subhalo masses are significantly underestimated and so the cutoff in the WDM mass function effectively shifts to higher masses, or (ii) there are many undetected subhaloes in the host centre that, if detected at higher resolution, would exist in larger numbers in CDM than in WDM. On the other hand, if the subhalo finder is effectively reliant on the subhalo density to make an identification, it may underdetect low-concentration WDM haloes compared to their CDM counterparts. We have investigated these possibilities in Appendix A using the Aquarius N -body simulations of a MW-analogue halo at simulation particle masses up to a factor of 100 times smaller than is available for COCO, and we estimate from those data that the abundance of subhaloes in the $[10^7, 10^8] M_\odot$ halo mass range may plausibly be underestimated by a factor of ~ 2 , in agreement with (Onions et al. 2012).

One alternative approach to addressing resolution issues is to label subhaloes by a proxy for mass that is less susceptible to the uncertainties of the halo finder than is the case for the mass itself, such as the maximum of the subhalo circular velocity curve, V_{\max} , or the mass enclosed within the radius of V_{\max} , $M_{V_{\max}}$. However, these properties are also influenced by the halo density profile and its concentration, and are therefore difficult to compare between CDM and WDM. We illustrate these differences in Appendix B.

We now compare our results to observations. Banik et al. (2021) used the structure of Pal 5 and GD-1 streams to infer the abundance of subhaloes in the $[10^7, 10^8]$ and $[10^8, 10^9]$ bins. In both mass bins, they calculated that the measured value is ~ 20 per cent of the abundance measured in CDM simulations, and ascribe the discrepancy to gas removal and destruction by the MW disc. Our work has shown that the N -body prediction for 3.3 keV WDM haloes within 40 kpc is larger than this, therefore the 3.3 keV model can explain their results provided that the impact of baryonic physics processes is less than required for CDM to match the data. We do not find an enhancement in the substructure mass fraction from <300 to <40 kpc, which they measure estimate to be 14 per cent; however, the scatter in this quantity is large and likely driven by the presence of a small number of massive subhaloes.

In conclusion, we have demonstrated that our simulations predict that the difference between the abundances of CDM and WDM

subhaloes is much smaller for the gap-generating population than for subhaloes across the rest of the halo. We subsequently infer that both CDM and WDM models with $M_{\text{hm}} \leq 2.8 \times 10^8 M_\odot$ are consistent with observations. However, resolution and baryonic effects are important confounding issues in the inner halo, and higher resolution hydrodynamic simulations will be required to ascertain the true ratio of CDM-to-WDM gap perturber abundances.

3.3 Subhalo density and radial distribution

We now shift to investigating the internal properties of subhaloes found today in different radial bins, with a view to estimating the distributions of dark subhalo masses and densities. We select subhaloes with $M_0 = [10^{7.5}, 10^{8.5}]$ and split them into two groups by $z = 0$ radius: an inner halo sample (<50 kpc from the host centre) and a halo periphery sample ($[100, 120]$ kpc) which corresponds to the mean distance to classical dwarf spheroidal satellite galaxies (see McConnachie 2012, and references therein). Note that here we use 50 kpc rather than 40 kpc for our central region: this is because we need better statistics to split the subhalo population by infall time as well as mass compare to mass alone. We plot M_A as a function of accretion lookback time, t_A , for these two sets of subsamples in Fig. 7.

The accretion time-scales for both models are very similar. At both radii there is a very small bias towards earlier accretion times in WDM relative to CDM. The outer shell contains a population of late infalling subhaloes that is absent in the <50 kpc bin, and the relative size of that contribution is the same in the two models. The largest distinction between WDM and CDM is in the distribution of subhaloes accreted with masses $<10^8 M_\odot$ as anticipated above. The proportion of subhaloes in this mass range in WDM is suppressed by a factor of two relative to CDM in the larger radius shell but by a factor of 10 in the <50 kpc radius bin. There is also a strong suppression in the relative fraction of $10^8 M_\odot$ WDM haloes within 50 kpc compared to the relative fraction within 100 kpc. Overall, the population of resolved <50 kpc WDM subhaloes is more biased towards high mass progenitors than either outer-radius WDM subhaloes or inner-radius CDM subhaloes.

The difference in accretion masses and accretion times between the two radial bins leads to a difference in the mass–concentration relations, because formation time and mass both correlate with concentration: concentration decreases with mass and increases with redshift (Navarro, Frenk & White 1996, 1997; Neto et al. 2007). Concentration defined through the Navarro–Frenk–White profile (NFW; Navarro et al. 1996, 1997) can be defined as the ratio of the profile characteristic radius to r_{200} or with the characteristic overdensity, δ_c . Another definition of characteristic radius is the half-mass radius of the gravitationally bound dark matter, r_{half} , as this radius is relevant in the context of gaps-in-streams impact parameters.

We first compare the r_{half} of CDM and WDM haloes at infall. We select an initial sample of CDM and WDM subhaloes that fall in to the host at lookback times up to $t_A < 10$ Gyr. We split our CDM and WDM subhaloes into two distributions based on the accretion time, one at early lookback times ($[7, 9]$ Gyr), and one at very low lookback times (<2 Gyr). Note that we use all accreted haloes in this figure, including those that are disrupted before $z = 0$ or that are outside 300 kpc at $z = 0$. We present the relationship between r_{half} and M_A in Fig. 8.

We recover the expected result that concentrations are higher – i.e. r_{half} are smaller – for early accretion times than late accretion times. For $M_A > 2 \times 10^9 M_\odot$, the results for CDM and WDM are nearly identical, and both models exhibit 30 per cent larger r_{half} at late infall times. Below this threshold in M_A , the CDM M_A - r_{half}

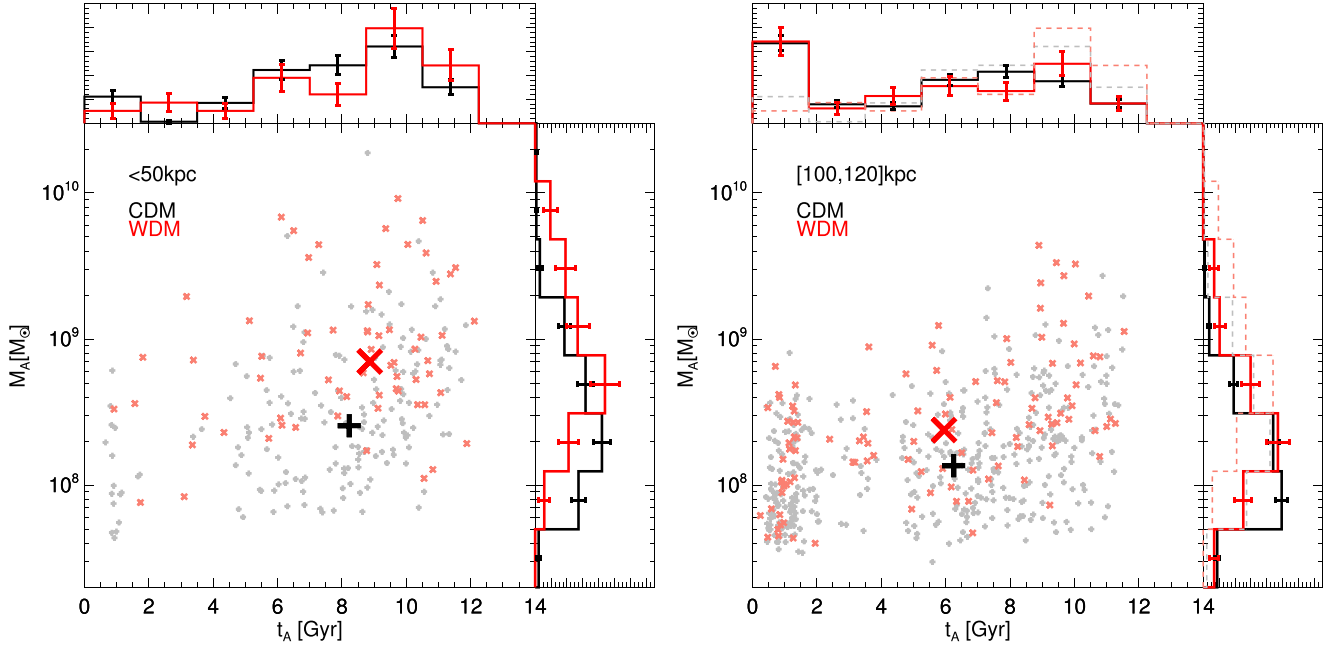


Figure 7. The distribution of accretion lookback times and accretion masses for subhaloes found today in two radial bins: within 50 kpc of the host centre (left) and the shell at [100, 120] kpc (right). CDM data are shown in black and WDM data in red. The data set medians are shown as large symbols. We include PDFs of each data set in the additional panels. In the right-hand set of panels, we also reproduce the 50 kpc data PDFs as dashed lines.

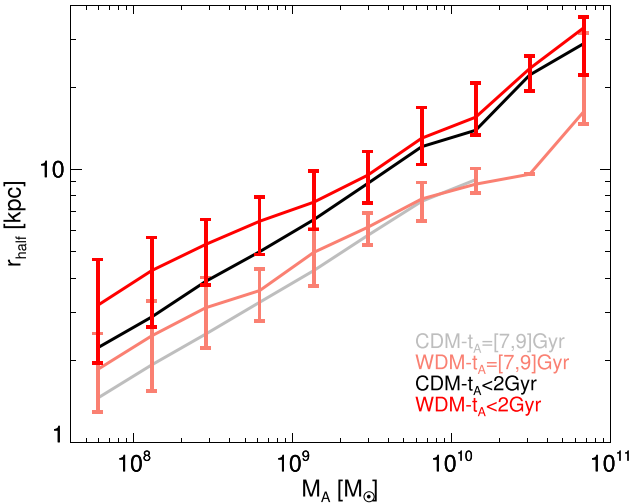


Figure 8. The distributions of half-mass radii for subhaloes at accretion, as a function of accretion mass. We show CDM and WDM data in two bins of infall time. CDM subhaloes with accretion times in the $t_A < 2$ Gyr ($[7, 9]$ Gyr) range are shown in black (grey); the WDM data in the same two lookback time bins are shown in red and pink, respectively. Curves denote median relations. For the WDM data sets, error bars indicate the 68 per cent regions; the errors on the CDM data are of a similar size but are omitted for clarity purposes.

distributions maintain the same power-law distribution as at higher masses. The WDM counterparts instead peel off towards larger r_{half} at fixed M_A . This is a result of the inversion of the halo concentration–mass relation at low masses in Bose et al. (2016). In the mass range $M_A = [10^8, 10^9] M_\odot$, the early accretion WDM subhaloes exhibit a smaller r_{half} / higher concentration than the late accreted CDM subhaloes in the same mass range.

We conclude this subsection by outlining possible ranges for radii of subhaloes that are relevant for gaps-in-streams studies, and in so doing demonstrate at which subhalo-stream impact parameters the difference in the mass–concentration relation between CDM and WDM subhaloes could be detected. Precise predictions would require very high resolution simulations with high time resolution output and likely 6D-phase space subhalo identification or even use a re-simulation technique similar to that of Lowing et al. (2011). In the absence of these particular requirements, we restrict ourselves to an approach that arguably brackets the possible range of mass–radius distributions. Our primary interest is in the subhaloes that generate gaps in streams. These are the subhaloes located within ~ 20 kpc of the host centre; in order to increase the sample of subhaloes, we will instead select subhaloes that are as far as 50 kpc from the host centre. We will then compare these ‘gap-generating’ subhaloes to all the subhaloes within 300 kpc.

It is in principle possible to measure the bound-mass half-light radius, r_{half} , of any subhalo, including those within 50 kpc. However, the halo finder can assign subhalo mass to the host halo or vice versa, and poor resolution in very small, stripped objects may affect the stripping rates. We therefore employ a method to estimate a post-tidal stripping halo mass and radius for subhaloes based on three factors: accretion mass, infall time, and a simple model of halo stripping. For each subhalo located within < 50 kpc of the host centre, we identify its accretion mass and infall time. We then compute the tidal radius, r_t , that the subhalo would have if it were situated at a distance of 20 kpc from the centre of our own MW without having its mass profile changed since infall. We use the formula for the tidal radius, r_t , presented in Springel et al. (2008):

$$r_t = \left[\frac{M_{\text{sub}}}{[2 - d \ln M_{\text{host}} / d \ln R] \times M_{\text{host}}(< R)} \right]^{1/3} \times R, \quad (1)$$

where M_{sub} is the gravitationally bound mass, for which we will use M_A , R is the distance from the subhalo centre to the host centre, for

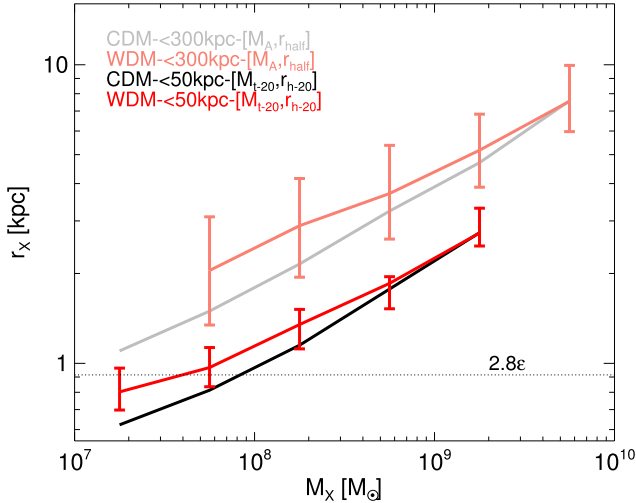


Figure 9. Approximate estimates for subhalo radii as a function of mass for CDM and WDM. We include subhaloes in two radial bins, the first within 50 kpc of the halo centre, and the second including all subhaloes out to 300 kpc. In the 50 kpc bin, we use the mass and half-mass radius definitions associated with the 20 kpc tidal stripping calculation described in the text; the CDM and WDM data for this radial bin are shown in black and red, respectively. In the larger radial bin, we show the half-mass radius computed at accretion as a function of the measured accretion mass, M_A ; here the CDM and WDM data are shown in grey and pink, respectively. Error bars mark 68 per cent of the data, and are only shown for WDM: the CDM error bars are of roughly the same size and are omitted for clarity. The horizontal dotted line marks $2.8 \times$ the gravitational softening length. The lowest-mass bin is lower than the WDM spurious subhalo mass limit $0.5M_{\text{lim}}$, and the whole-halo $-M_A$ bin at this mass is empty.

which we adopt 20 kpc, and $M_{\text{host}}(< R)$ is the spherically averaged mass profile of the host halo. For $M_{\text{host}}(< R)$, we use the total MW mass profile – including dark matter, stars, and gas – determined from observations by Cautun et al. (2020); therefore our result accounts for stripping due to the steepening of the dark matter halo by adiabatic contraction and also due to the presence of the stellar disc, although a full hydrodynamical treatment would be required to account for satellite destruction. We then compute the mass within r_t to be the mass of our gaps-subhaloes, which we label M_{t-20} . We then compute the half mass radius for material located within r_t , and label this $r_{\text{half}-20}$. We have therefore constructed an estimate of the mass and half-mass radius of a halo shorn of its outer layers by tidal effects at 20 kpc. Note that we do not attempt to model further evolution of the halo profile (see Errani & Navarro 2021, for a recent discussion on stripping to small fractions of the infall mass.)

Having generated a model for the gap-generating subhaloes, we turn to the whole subhalo population. For each of the subhaloes within 300 kpc, we identify its r_{half} at infall. The r_{half} represents approximately a maximum size for each halo. We thus compute two mass–radius distributions – $M_{t-20}-r_{\text{half}-20}$ for the gap-generating population and M_A-r_{half} for the general population. In so doing, we compare the mass–radius distributions of unprocessed subhaloes to an estimate of the mass–radius distributions of subhaloes that generate gaps. We present these in Fig. 9.

The curves of the general population subhaloes mirror the results of Fig. 8 closely, with both models tracking one another above a characteristic mass and diverging below it. The distribution that uses infall parameters predicts that differences between the stripping capabilities of WDM and CDM subhaloes could emerge at impact

parameters as large as 4 kpc, since 16 per cent of the WDM subhaloes with $M_A \sim 10^8 M_\odot$ have a value of r_{half} greater than 4 kpc. This threshold occurs for CDM at 2 kpc, at 80 per cent of the WDM r_{half} . By contrast, the <50 kpc population shows much more compact subhaloes at the same mass scale. The two dark matter models start to diverge from one another at tidal masses of $6 \times 10^{10} M_\odot$, where the tidal median r_{half} is estimated to be <2 kpc. We caution that this length scale is very close to the softening limit of 2.8ϵ .

In conclusion, we have argued that the combination of subhalo formation time, infall time, and tidal friction leads to massive, dense haloes in the regions of the MW halo where stellar stream gaps can be measured, and impact parameters of <2 kpc are required to discern whether perturber subhaloes are better described by CDM or WDM.

3.4 The impact of baryonic physics: potential caveats to the N -body approach in the inner 50 kpc

In the preceding sections we made predictions for the subhalo abundances in the CDM and WDM cosmologies using N -body simulations. Here we speculate about how baryonic physics make affect the results.

The first possible change is to the subhalo mass profiles during formation. Reionization photons from the first stars and galaxies can evaporate gas from galaxies before stars form, which removes mass and makes the potential well shallower. This process can inhibit the accretion of more dark matter, and can lower the masses of subhaloes by an average of 30 per cent (Sawala et al. 2013).

The interaction between this mass loss and the mass of WDM-to-CDM subhaloes in a given mass bin is complex. On the one hand, the CDM mass function is steeper than the WDM version: therefore, the relative number of subhaloes shifted out of the bin to the lower masses compared to into the bin from lower masses is larger in CDM and so the difference between the models will close slightly. On the other hand, the formation time of low mass subhaloes is delayed in WDM relative to CDM (Lovell et al. 2012, 2019) and potentially well beyond the end of the Epoch of Reionization, in which case a greater fraction of WDM subhaloes will have had their gas removed and so the gap between the models would widen. The lower concentrations of WDM subhaloes will also make them more susceptible to gas evacuation. Answering this question may well require bespoke simulations of reionization to answer fully. Moreover, in satellites that form sufficiently large numbers of stars, supernova feedback could conceivably drive larger changes in the lower density WDM haloes than in CDM.

A second consideration is the matter profile of the host halo, which we have already outlined above and approximated using the Cautun et al. (2020) MW halo profile estimate. Gas cooling results in adiabatic contraction of the dark matter halo (Blumenthal et al. 1986; Gnedin et al. 2004), thus steepening the density gradient, shrinking the tidal radius (Springel et al. 2008), and enhancing the stripping of each satellite. The stellar disc can also play an important role in destroying satellite galaxies (Garrison-Kimmel et al. 2017; Richings et al. 2020). The process of contraction and disc formation occurs at scales where CDM and WDM are indistinguishable, and therefore changes to the model predictions will result from differences in the subhalo mass functions and subhalo mass–concentration relations. The impact of contraction and will be strongest in the halo centre where the most massive haloes reside. It is unclear how the halo destruction rate changes between massive haloes – which sink faster under dynamical friction and are the least concentrated – and less massive haloes that have lower central densities over all even though their concentrations are higher. If it turns out that higher mass

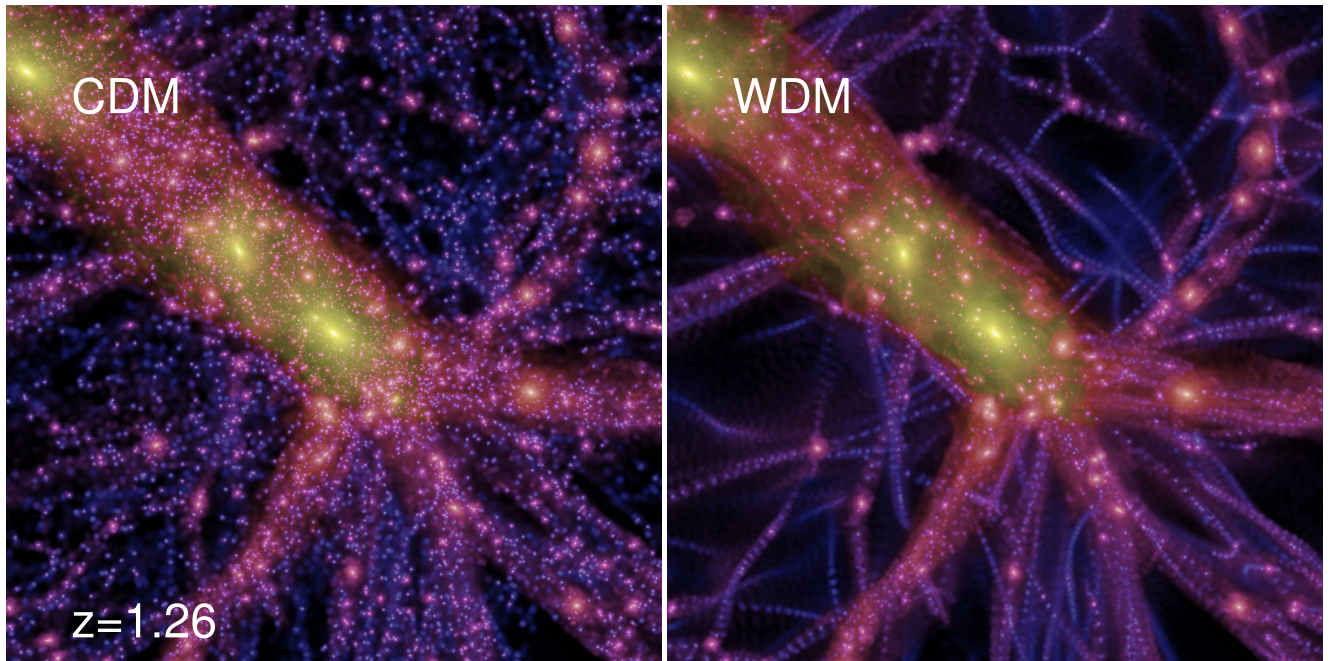


Figure 10. Images of one host halo at $z = 1.26$. CDM is shown in the left-hand panel and WDM in the right-hand panel. Each panel is 2 Mpc (comoving) on a side. The image intensity indicates the column density and the colour encodes the velocity dispersion.

subhaloes are more strongly affected than low mass haloes, this will drive the difference between the models to be larger than in the N -body prediction. If low mass haloes are suppressed equally in CDM and WDM then the differences will instead shrink. Perhaps the most complicated outcome is the case that both low mass subhaloes are suppressed more than high mass subhaloes and low concentration (WDM) subhaloes are suppressed more than high concentration (CDM) subhaloes. It would then be down to the relative contribution of both mass and concentration to subhalo disruption as to how the final result will differ from our N -body predictions.

3.5 Planes of satellites

We end our presentation of the results with a discussion of the prevalence – or lack thereof – of planes of satellites in CDM and WDM. It has been shown that extremely warm models, such as the 1 keV thermal relic used in Libeskind et al. (2013), can drastically reduce the number of filaments from which subhaloes can be accreted on to the host. Since filamentary accretion is an important factor in the creation of planes (Libeskind et al. 2005; Lovell et al. 2011), it is possible that WDM models may generate more distinct planes than CDM and therefore explain the observed plane of satellites around the MW and other galaxies. It has been shown that the local density in which subhaloes form can affect their mass, such that WDM haloes that form in voids will be less massive than those that form in filaments (Lovell et al. 2019). The structure of filaments also changes from one model to the other, such that the spines of WDM filaments are dense enough to form stars (Gao & Theuns 2007); we do not consider this possibility in this work.

In practice, the WDM cosmology includes a free parameter, M_{hm} , and not all observables will be affected equally at a given mass scale. For example, even though the number of dwarf haloes is strongly suppressed in WDM models with $M_{\text{hm}} \sim 10^8 M_{\odot}$ (Bose et al. 2016), the large scale structure at $z = 0$ is preserved (Lovell 2020). We will therefore investigate whether the scale of changes to the abundance of

filaments in the 3.3 keV thermal relic model is sufficient to influence the existence of planes of satellites. We begin this analysis with images of the CDM and WDM counterparts of one of our haloes at $z = 1.26$, shown in Fig. 10. We will use these images to develop an initial, visual intuition about the distribution of filaments in the two models.

The comparison of these images shows striking differences on small scales, with many more low-mass CDM haloes than WDM haloes. The main supply of subhaloes comes from a thick filament to the top left of the image plus four thinner filaments in the bottom right. Crucially, all these filaments are reproduced in the WDM version of the halo, implying that a half-mode mass $M_{\text{hm}} \leq 2.8 \times 10^8 M_{\odot}$ does not produce any striking effects on the direction of infalling haloes compared to CDM. Models with M_{hm} larger than this are frequently in strong tension with observations (Newton et al. 2020; Enzi et al. 2021; Nadler et al. 2021). Therefore it is unlikely that any viable WDM model could have a dramatic impact on the spatial distribution of satellites.

Attempts to turn this type of data into quantitative predictions for planes of satellites are numerous (Lovell et al. 2011; Cautun et al. 2015; Shao et al. 2018). In this study, we will adopt a modified version of the Shao et al. (2018) algorithm for identifying filaments and assigning subhaloes to filament membership. First, we identify the infall position of all subhaloes with peak mass (M_{p}) above some threshold. This position is defined as the location at which the subhalo enters the r_{200} of the host halo; for subhaloes that are within 300 kpc at $z = 0$ but have never entered r_{200} we take the $z = 0$ position. We rank these subhaloes by M_{p} .³ We identify our first filament as the location of the first ranked subhalo, and assign membership of that filament to subhaloes whose infall location is within 30 deg of that first subhalo

³We use M_{p} rather than M_{A} in this case because it is more likely to set the lower boundary for which subhaloes can form luminous satellite galaxies; in practice the choice of M_{A} or M_{p} does not affect our results.

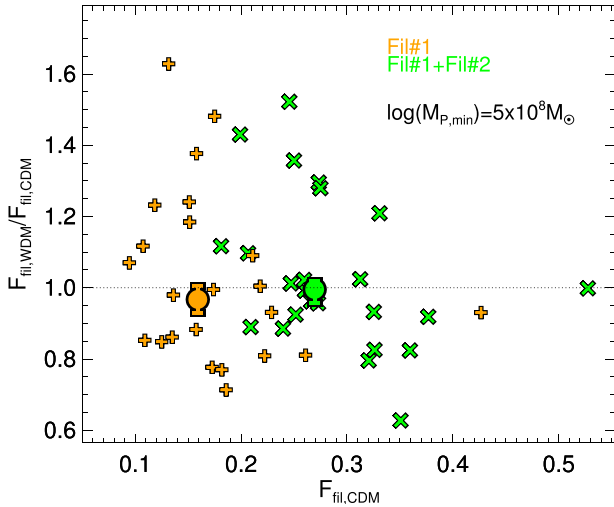


Figure 11. The fraction of subhaloes located in the first and second most massive filaments for each host as described in the text. We compute the ratio of the WDM and CDM filament fractions and plot the result as a function of the CDM counterpart fraction. We plot the filament fractions using the first filament alone as orange plus signs, and using the sum of the first and second filaments as green crosses. Filled circles denote the median values, the error bars on the circles denote bootstrap errors equivalent to 1σ . The threshold for selecting haloes is $M_A > 5 \times 10^8 M_\odot$.

infall position; we choose 30 deg because this is approximately the apparent size of filament at r_{200} as observed from the halo centre. The second filament is then defined with the next highest M_P -ranked halo that was not assigned to the first filament, and add subhaloes to the second filament that are within 30 deg but not included in the first filament. We repeat this process until all subhaloes have been associated to a filament. We then order the filaments by the number of member subhaloes, such that the ‘most massive filament’ is that which supplies the highest number of subhaloes, regardless of whether it contains the single most massive subhalo. Note that this method regards all subhaloes with M_P above a chosen threshold as being equally capable of forming a galaxy and contributing to a plane; Santos-Santos et al. (2020) argue that the mass / luminosity of a satellite does not appear to be related to whether or not it is a member of a plane.

The result of this process is a list of filaments for each host, ordered by the number of member subhaloes. We perform this for a threshold in M_A of $5 \times 10^8 M_\odot$, which we expect is the lowest halo mass capable of forming a luminous galaxy. We compute the fraction of subhaloes that are assigned to the first filament for each host, and also compute the fraction assigned to both the first and second filament. We compute the ratio of these fractions between the WDM and CDM counterparts of each host, and plot these ratios as a function of the CDM filament fraction in Fig. 11.

The first-filament subhalo fraction ranges from 0.1 to 0.42 in CDM. The WDM first-filament fraction varies from up to 60 per cent higher than the CDM counterpart to 30 per cent lower, with a median fraction of ~ 5 per cent lower than for CDM. A similar pattern occurs when the second filament is added to the first, with the highest subhalo fraction of 0.54; the median deviation between WDM and CDM reduces to < 5 per cent. In both cases, the bootstrap errors on the median are consistent with unity; implying that there is no statistical difference between CDM and WDM. We have repeated this process with an M_P threshold of $3 \times 10^8 M_\odot$ and find that the results are qualitatively and quantitatively very similar to the $5 \times 10^8 M_\odot$ threshold.

We have included all of the subhaloes that fall into the host halo, and thus have not attempted to remove subhaloes that would pass close enough to the Galactic disc as to be disrupted. It has been argued that the removal of such subhaloes on radial orbits would increase the likelihood of satellite planes (Ahmed et al. 2017). Given that these subhaloes would be the most massive – and therefore the most alike in abundance between CDM and WDM – there is perhaps some scope for the inclusion of baryon physics to increase the difference between the two models. A full hydrodynamical treatment will be required to check whether a disc would indeed drive the two models apart.

The main result from this analysis is that introducing the WDM cutoff changes the occupancy of filaments in a manner that is primarily stochastic rather than systematic. Any systematic effect that is present is of the order of < 10 per cent, and much smaller than the scatter between volumes. We have not found any evidence that is attempting to restrict our subhalo selection to those subhaloes more likely to host luminous galaxies will change this finding. We therefore conclude that the impact of WDM on the filamentary accretion of satellites phenomenon is at best marginal.

4 CONCLUSIONS

In this paper, we have considered the spatial distribution of satellites in CDM and WDM simulations, the regime where large scale filaments and the small-scale structure and abundance of dwarf haloes combine to set subhalo properties. Within this broad topic, we considered the potential for discerning between these two models using the distribution of luminous satellite galaxies, the distribution of dark subhaloes, and their implications for studies of gaps in stellar streams. We also considered the possibility that changes to filamentary accretion in WDM could increase the frequency of satellite planes.

We began by comparing the radial distributions of subhaloes in CDM and WDM, showing that the populations had different distributions when binned by mass at $z = 0$ but not when binned by accretion mass. We showed that the near-defining difference between CDM and WDM – the absence in the latter of low-mass, isolated haloes – is the source of the difference. Low mass subhaloes in WDM are invariably the stripped remnants of high mass progenitors. At fixed present-day mass, CDM contains more subhaloes in the outer regions – these are the largely unstripped versions of low-mass haloes either at first infall or on orbits far from the host centre. This difference is compounded by extra mass loss in WDM subhaloes due to their lower concentrations. Given that satellite galaxies will form only in the most massive subhaloes before accretion, we therefore expect that the distribution of luminous satellites will be very similar in CDM and WDM.

The less massive, dark subhaloes will have a very different distribution in the two models, with the radial distribution of WDM subhaloes being much more concentrated than that of CDM subhaloes. Therefore, the difference in abundance of WDM-to-CDM subhaloes is much reduced in the halo central regions than is the case for the halo at large. This result has important implications for the study of gaps in streams, which probe only the central ~ 30 kpc of the MW halo. For example, the abundance of WDM-to-CDM subhaloes in the mass range $[10^7, 10^8] M_\odot$ for the host at large is 10 per cent, whereas within the inner 40 kpc the abundance of WDM subhaloes rises to 30 per cent of CDM. We have also demonstrated that the differences within the inner 40 kpc are a combination of both extra stripping and mass function suppression, and the relative contribution

of stripping is higher in the inner 40 kpc than it is for the whole halo population.

We showed that in these central region subhaloes have earlier infall times than the subhalo population in general and have more massive progenitors, and therefore these subhaloes will have a different concentration–mass relation. We derived a simple estimate for the sizes of gap-generating subhaloes, demonstrating that they are on average less than half the size of subhaloes at infall of fixed subhalo mass, and argued that subhalo–stream impact parameters of <3 kpc are required in order to discern whether the perturber subhalo density profile is better described by CDM or by WDM.

Finally, we considered the possibility that evidence for a cutoff in the matter power spectrum could be imprinted on the abundance of planes of satellites, with planes being more common if WDM were to erase small-scale filaments. We computed the fraction of potentially luminous subhaloes that fell on to host haloes through the first and second richest filaments, and found that there was no statistically significant difference between the CDM and WDM models. There were large stochastic differences between CDM and WDM copies of each host, but these differences were smaller than between hosts. We conclude that, if there is indeed a mass scale at which the power spectrum cut-off favours planes of satellites, it is at a thermal relic mass scale <3.3 keV, which is already likely ruled out by a combination of Lyman- α forest observations, lensing limits, and MW satellite counts (Enzi et al. 2021; Nadler et al. 2021). We stress that our host halo selection explicitly excludes MW-analogues that have a nearby M31 companion, which may play a role in the large scale structure through which subhaloes are accreted; however, we expect that WDM will not change the likelihood that these systems generate planes of satellites.

We have thus argued that WDM models of interest have little impact on the radial distribution or the frequency of planes of luminous satellites. We have clarified the degree to which the CDM and WDM models differ in their predictions for gaps-in-streams analyses, and we have demonstrated that it is necessary to account for differences between the distribution of gap-generating subhaloes within ~ 30 kpc of the MW centre versus the MW subhalo population at large. In order to make more precise predictions for the properties of stream gaps in both models, simulations that include stellar discs, self-consistent contraction of the host halo, baryonic processes within subhaloes, evaporation of gas by reionization feedback, and high subhalo resolution are required.

ACKNOWLEDGEMENTS

MRL acknowledges support by a Grant of Excellence from the Icelandic Research Fund (grant number 206930). MC acknowledges support by the EU Horizon 2020 research and innovation programme under a Marie Skłodowska-Curie grant agreement 794474 (DancingGalaxies). CSF acknowledges support from European Research Council (ERC) Advanced Investigator grant DMIDAS (GA 786910). This work was also supported by the Consolidated Grant for Astronomy at Durham (ST/L00075X/1). WAH is supported by the Polish National Science Center grant nos. UMO-2018/30/E/ST9/00698 and UMO-2018/31/G/ST9/03388. ON acknowledges financial support from the Project IDEXLYON at the University of Lyon under the Investments for the Future Program (ANR-16-IDEX-0005) and supplementary financial support from La Région Auvergne-Rhône-Alpes. This work used the DiRACDurham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1,

ST/P002293/1, and ST/R002371/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

DATA AVAILABILITY

The data used in this paper were originally published in Hellwing et al. (2016) and Bose et al. (2016). Please the contact W. Hellwing for access.

REFERENCES

- Ahmed S. H., Brooks A. M., Christensen C. R., 2017, *MNRAS*, 466, 3119
 Amorisco N. C., Gómez F. A., Vegetti S., White S. D. M., 2016, *MNRAS*, 463, L17
 Asaka T., Shaposhnikov M., Laine M., 2007, *Journal of High Energy Physics*, 2007, 091
 Banik N., Bovy J., 2019, *MNRAS*, 484, 2009
 Banik N., Bertone G., Bovy J., Bozorgnia N., 2018, *J. Cosmol. Astropart. Phys.*, 2018, 061
 Banik N., Bovy J., Bertone G., Erkal D., de Boer T. J. L., 2021, *MNRAS*, 502, 2364
 Benito M., Criado J. C., Hütsi G., Raidal M., Veermäe H., 2020, *Phys. Rev. D*, 101, 103023
 Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002, *MNRAS*, 333, 177
 Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, *ApJ*, 301, 27
 Bonaca A., Hogg D. W., Price-Whelan A. M., Conroy C., 2019, *ApJ*, 880, 38
 Bose S. et al., 2017, *MNRAS*, 464, 4520
 Bose S., Hellwing W. A., Frenk C. S., Jenkins A., Lovell M. R., Helly J. C., Li B., 2016, *MNRAS*, 455, 318
 Bose S., Deason A. J., Belokurov V., Frenk C. S., 2020, *MNRAS*, 495, 743
 Bullock J. S., Boylan-Kolchin M., 2017, *ARA&A*, 55, 343
 Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, *ApJ*, 539, 517
 Carlberg R. G., Grillmair C. J., 2013, *ApJ*, 768, 171
 Cautun M. et al., 2020, *MNRAS*, 494, 4291
 Cautun M., Bose S., Frenk C. S., Guo Q., Han J., Hellwing W. A., Sawala T., Wang W., 2015, *MNRAS*, 452, 3838
 Chandrasekhar S., 1943, *ApJ*, 97, 255
 Cole S. et al., 2005, *MNRAS*, 362, 505
 Colín P., Avila-Reese V., Valenzuela O., 2000, *ApJ*, 542, 622
 Despali G., Lovell M., Vegetti S., Crain R. A., Oppenheimer B. D., 2020, *MNRAS*, 491, 1295
 Eisenstein D. J. et al., 2005, *ApJ*, 633, 560
 Enzi W. et al., 2021, *MNRAS*, 506, 5848
 Erkal D., Belokurov V., 2015, *MNRAS*, 454, 3542
 Errani R., Navarro J. F., 2021, *MNRAS*, 505, 18
 Font A. S., et al., 2011, *MNRAS*, 417, 1260
 Font A. S., McCarthy I. G., Belokurov V., 2021, *MNRAS*, 505, 783
 Gao L., Theuns T., 2007, *Science*, 317, 1527
 Garrison-Kimmel S. et al., 2017, *MNRAS*, 471, 1709
 Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, *ApJ*, 616, 16
 Han J., Cole S., Frenk C. S., Jing Y., 2016, *MNRAS*, 457, 1208
 Hellwing W. A., Frenk C. S., Cautun M., Bose S., Helly J., Jenkins A., Sawala T., Cytowski M., 2016, *MNRAS*, 457, 3492
 Hezaveh Y. D. et al., 2016, *ApJ*, 823, 37
 Kang X., Mao S., Gao L., Jing Y. P., 2005, *A&A*, 437, 383
 Komatsu E. et al., 2011, *ApJS*, 192, 18
 Lacey C., Cole S., 1993, *MNRAS*, 262, 627
 Laine M., Shaposhnikov M., 2008, *J. Cosmol. Astropart. Phys.*, 6, 31
 Li Y., White S. D. M., 2008, *MNRAS*, 384, 1459
 Libeskind N. I., Frenk C. S., Cole S., Helly J. C., Jenkins A., Navarro J. F., Power C., 2005, *MNRAS*, 363, 146
 Libeskind N. I., Cole S., Frenk C. S., Okamoto T., Jenkins A., 2007, *MNRAS*, 374, 16
 Libeskind N. I., Di Cintio A., Knebe A., Yepes G., Gottlöber S., Steinmetz M., Hoffman Y., Martinez-Vaquero L. A., 2013, *PASA*, 30, e039

Lovell M. R. et al., 2012, *MNRAS*, 420, 2318
 Lovell M. R. et al., 2016, *MNRAS*, 461, 60
 Lovell M. R. et al., 2017, *MNRAS*, 468, 4285
 Lovell M. R. et al., 2019, *MNRAS*, 485, 4071
 Lovell M. R., 2020, *MNRAS*, 493, L11
 Lovell M. R., Eke V. R., Frenk C. S., Jenkins A., 2011, *MNRAS*, 413, 3013
 Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, *MNRAS*, 439, 300
 Lowing B., Jenkins A., Eke V., Frenk C., 2011, *MNRAS*, 416, 2697
 Lux H., Read J. I., Lake G., 2010, *MNRAS*, 406, 2312
 Lynden-Bell D., 1976, *MNRAS*, 174, 695
 McConnachie A. W., 2012, *AJ*, 144, 4
 Metz M., Kroupa P., Libeskind N. I., 2008, *ApJ*, 680, 287
 Metz M., Kroupa P., Jerjen H., 2009, *MNRAS*, 394, 2223
 Nadler E. O., Birrer S., Gilman D., Wechsler R. H., Du X., Benson A., Nierenberg A. M., Treu T., 2021, *ApJ*, 917, 7
 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
 Neto A. F. et al., 2007, *MNRAS*, 381, 1450
 Newton O., Cautun M., Jenkins A., Frenk C. S., Helly J. C., 2018, *MNRAS*, 479, 2853
 Newton O., Leo M., Cautun M., Jenkins A., Frenk C. S., Lovell M. R., Helly J. C., Benson A. J., 2020, preprint (arXiv:2011.08865)
 Onions J. et al., 2012, *MNRAS*, 423, 1200
 Pawlowski M. S. et al., 2014, *MNRAS*, 442, 2362
 Pawlowski M. S., Kroupa P., 2020, *MNRAS*, 491, 3042
 Richings J. et al., 2020, *MNRAS*, 492, 5780
 Samuel J. et al., 2020, *MNRAS*, 491, 1471
 Santos-Santos I. M. E., Fattahi A., Sales L. V., Navarro J. F., 2021, *MNRAS*, 504, 4551
 Santos-Santos I. M., Domínguez-Tenreiro R., Pawlowski M. S., 2020, *MNRAS*, 499, 3755
 Sawala T. et al., 2016, *MNRAS*, 456, 85
 Sawala T., Frenk C. S., Crain R. A., Jenkins A., Schaye J., Theuns T., Zavala J., 2013, *MNRAS*, 431, 1366
 Sawala T., Pihajoki P., Johansson P. H., Frenk C. S., Navarro J. F., Oman K. A., White S. D. M., 2017, *MNRAS*, 467, 4383
 Schneider A., Smith R. E., Reed D., 2013, *MNRAS*, 433, 1573
 Shao S., Cautun M., Frenk C. S., Grand R. J. J., Gómez F. A., Marinacci F., Simpson C. M., 2018, *MNRAS*, 476, 1796
 Shao S., Cautun M., Frenk C. S., 2019, *MNRAS*, 488, 1166
 Simha V., Cole S., 2017, *MNRAS*, 472, 1392
 Spergel D. N. et al., 2003, *ApJS*, 148, 175
 Springel V. et al., 2005, *Nature*, 435, 629
 Springel V. et al., 2008, *MNRAS*, 391, 1685
 Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
 Vegetti S., Koopmans L. V. E., Bolton A., Treu T., Gavazzi R., 2010, *MNRAS*, 408, 1969
 Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, *Phys. Rev. D*, 71, 063534
 Yniguez B., Garrison-Kimmel S., Boylan-Kolchin M., Bullock J. S., 2014, *MNRAS*, 439, 73
 Yoon J. H., Johnston K. V., Hogg D. W., 2011, *ApJ*, 731, 58
 Zentner A. R., Kravtsov A. V., Gnedin O. Y., Klypin A. A., 2005, *ApJ*, 629, 219

APPENDIX A: RESOLUTION STUDY OF SUBHALO ABUNDANCE

As stated in the main text, the statistics of subhaloes close to the centre of the host are potentially affected by resolution due to the challenges of spurious disruption and the limitations of the halo finder. We check for the impact of resolution using the simulations of the Aquarius Aq-A halo (Springel et al. 2008). This halo was simulated using the WMAP-1 cosmology (Spergel et al. 2003) and has $M_{200} = 1.88 \times 10^{12} M_{\odot}$. There are five simulations of this

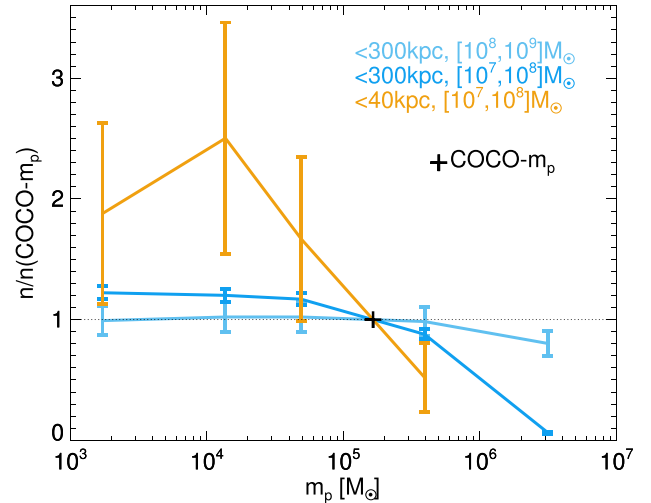


Figure A1. The subhalo abundance in mass and radial bins for the Aquarius Aq-A halo as a function of simulation particle mass. We interpolate a value for the abundance at the COCO resolution and normalize each curve by this value. Results for subhaloes <300 kpc from the centre and with mass $[10^8, 10^9] M_{\odot}$ are shown in light blue, <300 kpc with mass $[10^7, 10^8] M_{\odot}$ in dark blue, and <40 kpc with mass $[10^7, 10^8] M_{\odot}$ in orange. The error bars assume Poisson statistics. The COCO particle mass is shown as a black cross.

halo at different resolution levels, labeled from Aq-A-5 (particle mass $m_p = 3 \times 10^6 M_{\odot}$) to Aq-A-1 (particle mass $1.3 \times 10^3 M_{\odot}$). The Aq-A-1 remains among the highest resolution simulations of a MW halo-analogue to date; the resolution of COCO is between the resolutions of Aq-A-3 and Aq-A-4.

For each of these five simulations, we compute the halo abundance in radial and mass bins. For the whole halo (<300 kpc from the host centre), we use the mass bins $[10^7, 10^8] M_{\odot}$ and $[10^8, 10^9] M_{\odot}$, and for the gap-generating haloes, we use $[10^7, 10^8] M_{\odot}$; we do not use the $[10^8, 10^9] M_{\odot}$ bin due to low statistics for this halo. We interpolate a likely value at the COCO resolution for this halo, and normalize the halo abundances by this value. We present the results as a function of m_p in Fig. A1.

We find that the larger of the two mass bins, $[10^8, 10^9] M_{\odot}$, is converged to better than 10 per cent for $m_p < 4 \times 10^5 M_{\odot}$ across the whole halo. The number of $[10^7, 10^8] M_{\odot}$ subhaloes within 300 kpc is resolved with similar precision for $m_p < 5 \times 10^4 M_{\odot}$, and at COCO resolution, we likely underestimate the ‘true’ number of subhaloes by 20 per cent. At the <40 kpc aperture the discrepancy is clearly larger, with an average of a factor of two more subhaloes for $m_p < 10^4 M_{\odot}$ than is inferred for COCO although the uncertainties are very large. We conclude that the difficulty of resolving subhaloes in high background density environments could well play a role on estimates of gap-generating subhalo abundance, although as we state in the main text, we are not in a position to determine whether it is CDM (potentially more subhaloes to resolve) or WDM (low concentration peaks to resolve) that is more strongly affected.

APPENDIX B: COMPARING ACCRETION MASS WITH MASS PROXIES

In this appendix, we demonstrate the different behaviour of the mass– V_{\max} relation in CDM and WDM (3.3 keV thermal relic). For our sample of subhaloes at infall, we compute the circular velocity curves, extract V_{\max} and its corresponding radius, r_{\max} . We compute

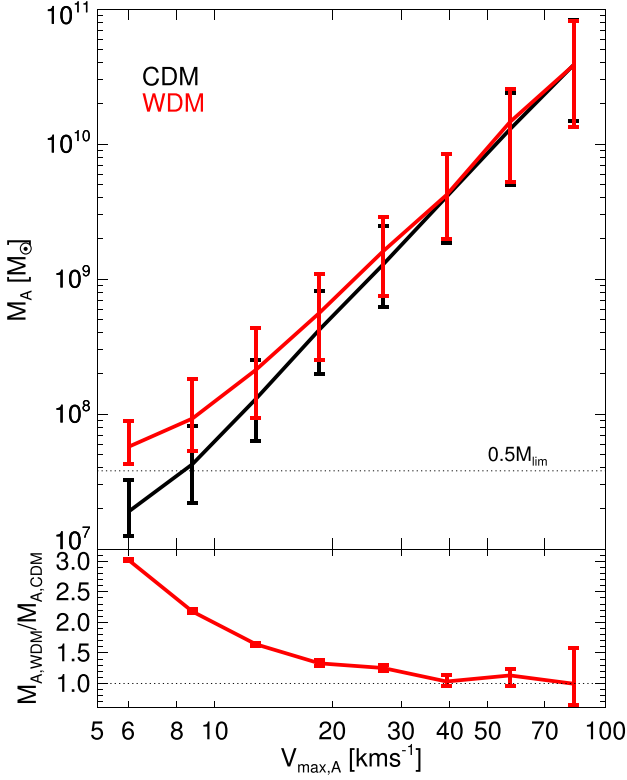


Figure B1. Top panel: the median accretion mass as a function of accretion V_{\max} . The error bars show the 68 per cent regions of the data; CDM is shown in black and WDM in red. The horizontal line marks the limit imposed for removing subhaloes as spurious, $0.5M_{\text{lim}}$. Bottom panel: the ratio of the WDM and CDM medians. Unlike the top panel, the bottom panel error bars are 68 per cent bootstrap errors on the median.

the median relation between the mass at infall, M_A and V_{\max} at infall, and present the results in Fig. B1.

The distributions are the same for the two models at $V_{\max} > 40 \text{ km s}^{-1}$. At smaller V_{\max} , CDM and WDM diverge. At $V_{\max} = 20 \text{ km s}^{-1}$, WDM haloes have M_A 40 per cent larger than in CDM, and at $V_{\max} = 10 \text{ km s}^{-1}$ the average difference is a full factor of 2, although this latter result is partly a consequence of our spurious subhalo removal parameter, $0.5 M_{\text{lim}}$. This difference in mass– V_{\max} relations arises from the difference in the halo profiles. Starting at $M_A \gtrsim 10 M_{\text{hm}}$, the mass–concentration relation for WDM is the same as CDM and thus the mass– V_{\max} relation is also the same. As M_A decreases, the halo concentrations of WDM haloes are systematically lower than their CDM counterparts, so r_{\max} increases and therefore V_{\max} must decrease. The inverse of this phenomenon is that for CDM and WDM haloes with the same V_{\max} , M_A must be higher in the WDM halo than in the CDM counterpart. We have therefore shown that care must be taken when comparing the V_{\max} values of subhaloes in different dark matter models.

In the discussion above, we have treated M_A as approximately equal to the mass within r_{\max} , given as $M_{\text{MaxV}} = V_{\max}^2 \times r_{\max} / G$,

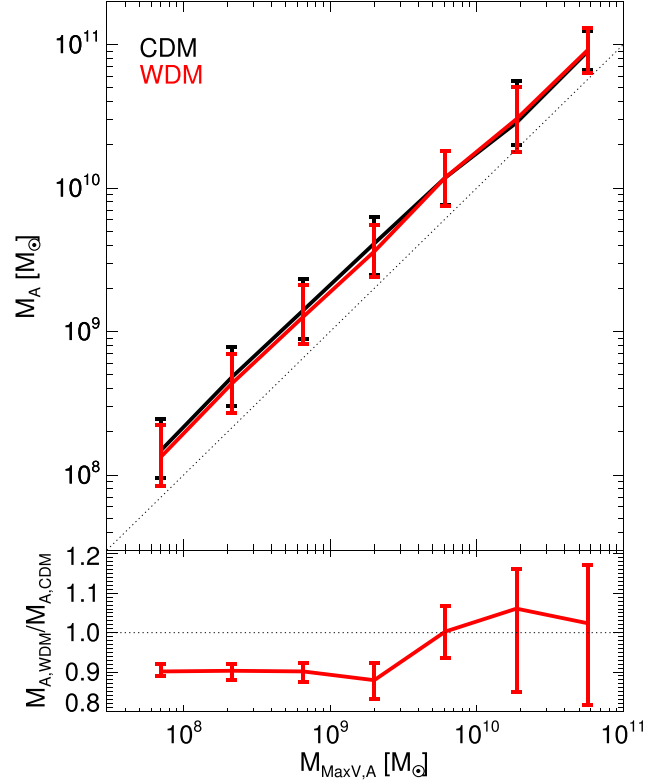


Figure B2. Top panel: the median accretion mass as a function of accretion mass within r_{\max} , M_{MaxV} . The error bars show the 68 per cent regions of the data; CDM is shown in black and WDM in red. The area below the bottom line is forbidden because M_{MaxV} must be less than M_A . Bottom panel: the ratio of the WDM and CDM medians. Unlike the top panel, the bottom panel error bars are 68 per cent bootstrap errors on the median.

where G is the gravitational constant. Although we have shown that the reason behind the change in V_{\max} – M_A relation with concentration is clear, we can also perform a direct comparison between M_A and the computed value of M_{MaxV} . We present the results in Fig. B2, using the same format as for Fig. B1. For this figure, we use all subhaloes that are accreted on to the 24 hosts, irrespective of whether they survive to $z = 0$.

At $M_{\text{MaxV}} \geq 6 \times 10^9 M_{\odot}$, the WDM and CDM median M_A are statistically consistent with each other and are on average a factor of 2 higher than M_{MaxV} . For masses $< 6 \times 10^9 M_{\odot}$, CDM and WDM masses diverge slightly such that at $M_{\text{MaxV}} \sim 10^8 M_{\odot}$, the WDM version is suppressed by 10 per cent relative to CDM. This difference is much smaller than we found for V_{\max} . We expect what difference comes from the larger WDM r_{\max} at fixed M_A : having a larger r_{\max} means that more of the total mass M_A is included in M_{MaxV} .

This paper has been typeset from a \LaTeX file prepared by the author.