

IFIC 21/29, FERMILAB-PUB-21-321-AE-T

e-Print: 2107.10293 [astro-ph.CO]

DOI: 10.1088/1475-7516/2021/12/026

(accepted)

Smallest Remnants of Early Matter Domination

Gabriela Barenboim,^a Nikita Blinov,^{b,c} Albert Stebbins^b

^aDepartament de Física Teòrica and IFIC, Universitat de València-CSIC, E-46100, Burjassot, Spain

^bFermi National Accelerator Laboratory, Batavia, IL 60510, USA

^cKavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

Abstract. The evolution of the universe prior to Big Bang Nucleosynthesis could have gone through a phase of early matter domination which enhanced the growth of small-scale dark matter structure. If this period was long enough, self-gravitating objects formed prior to reheating. We study the evolution of these dense early halos through reheating. At the end of early matter domination, the early halos undergo rapid expansion and eventually eject their matter. We find that this process washes out structure on scales much larger than naively expected from the size of the original halos. We compute the density profiles of the early halo remnants and use them to construct late-time power spectra that include these non-linear effects. We evolve the resulting power spectrum to estimate the properties of microhalos that would form after matter-radiation equality. Surprisingly, cosmologies with a short period of early matter domination lead to an earlier onset of microhalo formation compared to those with a long period. In either case, dark matter structure formation begins much earlier than in the standard cosmology, with most dark matter bound in microhalos in the late universe.

Contents

1	Introduction	1
2	Cosmological Evolution with Early Matter Domination	2
3	A Brief History of Early Halos	5
4	Evolution of an Isolated Halo	6
4.1	Ejection Velocity and Free-Streaming Length	7
4.2	Spherical Halos	8
4.3	Initially Circular Orbits	8
4.4	Spherical Halos in a Cosmological Context	10
4.5	Recollapse?	12
4.6	Halo Remnant Density Profile	13
4.7	Physical Parameters	17
5	Superposition of Halo Remnants	17
5.1	Remnant Overlap	18
5.2	One- and Two-Halo Power Spectra	19
6	Linear Evolution of Density Perturbations	20
6.1	Linear Evolution and Collapse Before Reheating	20
6.2	Linear Evolution and Collapse After Reheating	23
7	Power Spectrum and Formation of Microhalos	24
7.1	Comparison with Other Cutoffs	26
8	Conclusion	28
A	Adiabatic Homologous Expansion of Halos	29
B	Non-Circular Orbits and Shell Crossing	31
C	Linear Evolution	34

D Early Halo Model	35
D.1 An Analytic Example	37
D.2 Halo Exclusion	38

1 Introduction

The standard model of cosmology (Λ CDM) has proven to be very successful in explaining essentially all observations of the universe. It also includes extrapolations of assumptions to domains which have not been observationally tested. Observable relics from the early universe such as light element abundances [1], the cosmic microwave background (CMB) and its anisotropies (see, e.g., Refs. [2, 3]) and the large scale distribution of matter (both luminous and dark; see, e.g., Refs. [4–7]) give strong support for a hot big bang including an early stage of accelerated expansion (inflation). In order to accommodate both Big Bang Nucleosynthesis (BBN) [8–11] and the CMB the model requires an extended period of radiation domination (RD) where most of the mass/energy density is in the form of a nearly thermal distribution of photons and neutrinos as the universe cools by expansion from ~ 5 MeV to ~ 1 eV (1 sec to 10^5 yr). While it is usually assumed that radiation domination persisted throughout the period after inflation until just before BBN this is not required by observations. Given that the universe apparently has proceeded through several very different evolutionary phases (inflation, radiation domination, matter domination and currently dark energy domination) it is reasonable to suppose that there were additional intermediate phases which we are not currently aware of.

Recently much effort has been given to understanding the observational implications of an early matter domination (EMD) phase between inflation and BBN. Such a cosmology can be realized in a multitude of settings. For example, inflation can be followed by a long period where inflaton oscillations (which can have an equation of state identical to matter) dominate the energy content of the universe [12, 13]. An analogous situation can arise in ultra-violet completions of the Standard Model (SM) like string and supersymmetric theories where long-lived, non-relativistic particles (or scalar field oscillations) can drive the expansion for an extended period [14–19]; alternatively, a phase of EMD can result from the freeze-out of quasi-stable states in a dark sector [20–23]. Such a cosmology can have a profound impact of the production of dark matter (DM), favouring regions of theory space that are radically different from naive expectations [21, 22, 24–32]. An EMD phase would also produce potentially observable relics in the spatial distribution of dark matter (DM) on the very smallest scales which would be easily distinguishable from that expected in Λ CDM if one observes them [33–41]. These smallest structures are gravitationally bound clumps of DM (microhalos). Microhalos have shallow gravitational potential wells and are much smaller than the Jeans’ length of gas, so very little baryonic matter is bound to them. This would make them difficult to detect in any way except gravitationally. While microhalos are difficult to see we expect they will be detected eventually.

Matter domination allows for relatively rapid growth of inhomogeneities due to gravitational instability. An EMD epoch, with sufficient duration, will result in virialized (collapsed) structures in the very early universe which we call early halos (EHs). This paper explores the imprint of EHs on the smallest scale structure and how this is reflected in the microhalos present in the late universe we observe today. We find that, in contrast to the growth of small (linear) inhomogeneities during EMD, larger amplitude inhomogeneities which produce EHs suppress rather than enhance the late time inhomogeneities on the smallest scales. Such a suppression in the context of EMD was first pointed out in Ref. [42] though the suppression derived here is quantitatively different. This phenomenon which we call *explosive evaporation* has a simple intuitive explanation: the decay of early matter (EM) erases the gravitational potentials that bind EHs as the radiation decay products rapidly stream outward. The remaining stable particles, which constitute the DM of the current epoch, are also expelled outwards, but more slowly. This outward motion eventually dilutes the large

initially overdensity to a density below the cosmological mean at which point EH remnants overlap. If the decay were instantaneous the outward velocity would be comparable to the initial EH virial velocity, however a slower exponential decay leads to a much smaller outward velocity. The EHs become unbound well into the radiation, long after the decay time.

This evaporation and dilution of dense halos is generic in the sense that it is independent of how the EHs form. In the scenario considered here EMD occurs when a non-relativistic species (EM) comes to dominate the cosmic density and ends when this species decays into relativistic SM particles. We consider the case where nearly all the matter coalesces into gravitationally bound structures. We also require that DM be present and non-relativistic during EMD so that it will gravitationally cluster with the decaying species. In this scenario we find the remnant inhomogeneities after evaporation depend crucially on the EH mass function (distribution of masses) but is otherwise insensitive to their origin. We give analytic formula for the dark matter power spectrum in terms of this mass function.

The rest of this paper is organized as follows. In §2 we describe the background cosmology and briefly discuss the evolution of density perturbations during EMD. Enhanced growth during EMD ensures that most of the matter is bound up in EHs. We then summarize the evolution of EHs in §3. The evolution of a halo progresses through periods of 1) stable clustering during matter domination, 2) adiabatic expansion which describes their evolution well into the radiation era and 3) free expansion as successive outer layers become unbound during the radiation era. We use a simple spherical model of an isolated halo in §4 to study each of these steps quantitatively. Combining a population of EH remnants we construct the power spectrum of DM inhomogeneities which will eventually form microhalos in the late universe in §5. In §6 we specialize to primordial Gaussian adiabatic inhomogeneities which are an extrapolation to small scales of what we observe on large scales today. In this model we can estimate the mass distribution of EHs using linear perturbation theory and the Press-Schechter formalism. This distribution depends on the duration of the EMD era. Applying the previous results we predict the late time power spectrum for long and short durations of EMD in §7. From these spectra we deduce the mass spectra of microhalos which first collapse and discuss various implications of their existence. Our general conclusion is that non-linear collapse during EMD suppresses inhomogeneities on small scales; the suppression length scale, however, is much larger than the scales of sizes of the original EHs. Thus, “explosive evaporation” gives a power spectrum suppression that is independent of any DM microphysics. We outline avenues for further study and conclude in §8.

Our approach which starts with (nonlinear) bound halos complements linear theory estimates of the effect of EMD on the small scale power spectrum though the results are identical on much larger scales which do not collapse during EMD [33, 34, 39, 42]. Our results do not apply in cases where structures do not collapse during EMD but we show that linear theory predictions are qualitatively incorrect in the highly nonlinear case where they do. Somewhat surprisingly the maximal enhancement of small scale inhomogeneity at late times from EMD growth lies somewhere in between the linear and highly non-linear cases.

2 Cosmological Evolution with Early Matter Domination

We model the universe model as containing four components: “early matter” (EM - a particle species which dominates the density when non-relativistic during EMD and then decays with lifetime τ), “late matter” (LM - a stable species which is non-relativistic during EMD and is the dark matter (DM) today), late radiation (LR - the Standard Model decay products of EM assumed to rapidly thermalize into a relativistic gas of light species, e.g. photons, neutrinos, electrons, ...) and (optionally) early radiation (ER - remnant relativistic matter from the epoch which precedes EMD). The combination of LR and ER is denoted “rad” (radiation) and of LM and EM is denoted “mat” (matter). The energy

continuity equations for these homogeneous expanding fluids are

$$\dot{\bar{\rho}}_{\text{EM}} + 3H\bar{\rho}_{\text{EM}} = -\frac{1}{\tau}\bar{\rho}_{\text{EM}} \quad (2.1a)$$

$$\dot{\bar{\rho}}_{\text{rad}} + 4H\bar{\rho}_{\text{rad}} = +\frac{1}{\tau}\bar{\rho}_{\text{EM}} \quad (2.1b)$$

$$\dot{\bar{\rho}}_{\text{LM}} + 3H\bar{\rho}_{\text{LM}} = 0 \quad (2.1c)$$

where $\dot{} \equiv \partial/\partial t$, $H = \dot{a}/a$ gives the cosmological expansion in terms of the scale factor $a(t)$. The τ^{-1} terms transfer energy of decaying EM into LR. Note that we specialize to the case where LM is *not* produced in the decays of EM. The effect of gravity is given by the Friedmann equation for a spatially flat FLRW cosmology

$$H^2 = \frac{8\pi G}{3} (\bar{\rho}_{\text{EM}} + \bar{\rho}_{\text{rad}} + \bar{\rho}_{\text{LM}}) \quad (2.2)$$

which closes this system of equations. Solutions of Eq. 2.1 have the universe pass through four epochs which in chronological sequence are: early radiation domination (ERD), early matter domination (EMD), late radiation domination (LRD) and late matter domination (LMD). ERD is only present if $\bar{\rho}_{\text{ER}} \neq 0$. The transition between EMD and LRD occurs at $t \sim \tau$ (the *decay time*); we will refer to this epoch as reheating. LRD and LMD are identified with radiation and matter domination in our universe (later dark energy domination is not relevant to our analysis). Given τ and defining $t = 0$ by $H(0) = \infty$ the cosmological solution is fully specified by two parameters which set the duration of the EMD and LRD epochs. The LRD duration is determined by the LM to EM+LM density ratio at early times, $f \equiv \lim_{t \rightarrow 0} \bar{\rho}_{\text{LM}}/(\bar{\rho}_{\text{EM}} + \bar{\rho}_{\text{LM}})$: the universe expands by a factor $\sim f^{-1}$ during LRD. Given CMB determination of the end of radiation domination and BBN constraints on the minimum decay time (giving the radiation “reheating temperature” T_{RH} when $\bar{\rho}_{\text{rad}} = \bar{\rho}_{\text{mat}}$) one requires $f \lesssim 10^{-8}$. f could be much smaller. The duration of EMD is relevant here as it limits the amount of time for EHs to form.

One can solve equations 2.1–2.2 numerically. Two examples of such solutions are shown in Fig. 1. The left panel shows the cosmological evolution of the various fluids for the case where EMD is very long (i.e., the initial radiation density is negligible); we will refer to this scenario as “Long EMD”. An alternative possibility is that the period of EMD is brief and preceded by radiation domination; we will call this model “Short EMD”. We will use these two scenarios as benchmarks to study the evolution of density fluctuations. We will show that a longer period of EMD allows for the formation of more massive halos at early times. The initial densities of EM and ER and EM lifetime in the “Long EMD” (“Short EMD”) are chosen to give $a_{\text{eeq}}/a_{\text{RH}} = 0$ (1/2000), where a_{eeq} is the scale factor at “early equality” (i.e. when ERD ends and EMD begins) and a_{RH} is the scale factor at reheating.

It will be useful to have approximate analytic solutions to 2.1–2.2 in certain limits. When only one species dominates simple solutions exist: during EMD ($a \gg a_{\text{eeq}}$ and $t \ll \tau$, where a_{eeq} is the scale factor at “early equality” when the densities of ER and EM equal)

$$\bar{\rho}_{\text{LM}}(t) \simeq \frac{f}{6\pi G t^2} \quad \bar{\rho}_{\text{EM}}(t) \simeq \frac{1-f}{6\pi G t^2} \quad \bar{\rho}_{\text{LR}}(t) \simeq \frac{1}{10\pi G \tau t} \quad (2.3)$$

during LRD ($\tau \ll t \ll \tau/\sqrt{f}$)

$$\bar{\rho}_{\text{LM}}(t) \simeq \frac{f}{8.023 \pi G \sqrt{\tau} t^3} \quad \bar{\rho}_{\text{EM}}(t) \simeq \frac{(1-f)e^{-t/\tau}}{8.023 \pi G \sqrt{\tau} t^3} \quad \bar{\rho}_{\text{LR}}(t) \simeq \frac{3}{32\pi G t^2} \quad (2.4)$$

and during LMD ($t \gg \tau/\sqrt{f}$)

$$\bar{\rho}_{\text{LM}}(t) \simeq \frac{1}{6\pi G t^2} \quad \bar{\rho}_{\text{LR}}(t) \simeq \frac{3}{32\pi G t^2} \quad (2.5)$$

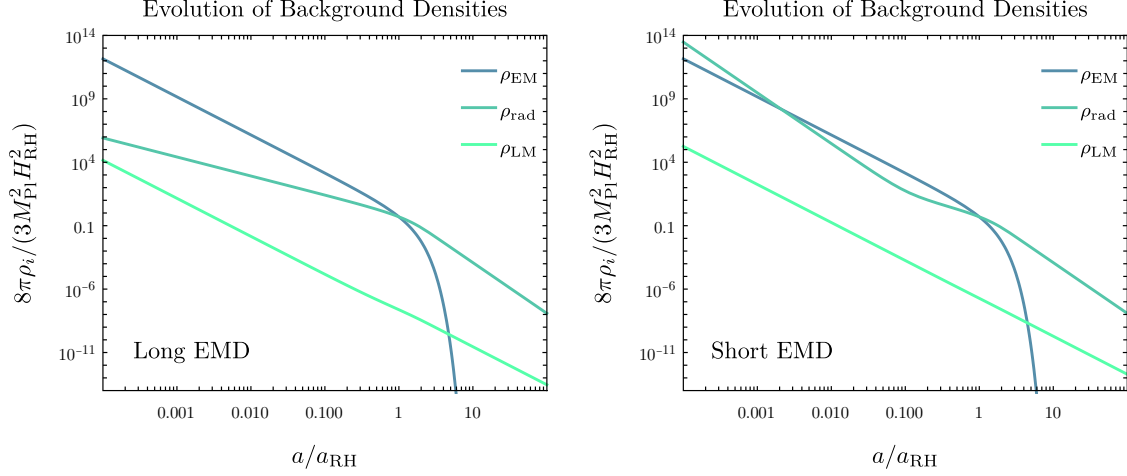


Figure 1. Evolution of background densities in a cosmology with a period of early matter domination as function of the scale factor normalized to its value at reheating. In the left panel the initial radiation density is small compared to EM, so all density fluctuations of interest enter the horizon during EMD. In contrast, in the right panel, the period of EMD is brief and is preceded by radiation domination. As a result, some modes enter the horizon before EMD starts. In both panels the densities are normalized to the total energy density at reheating.

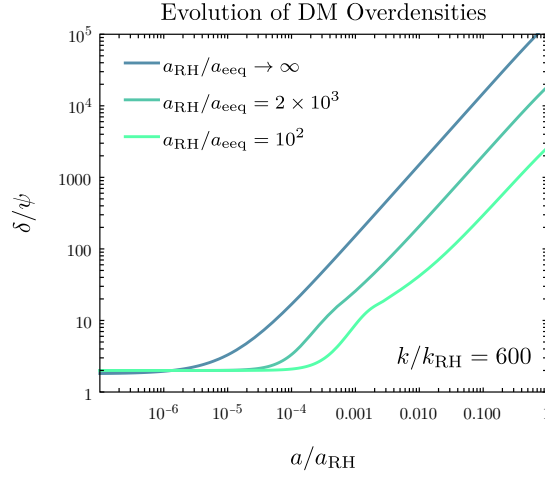


Figure 2. Evolution of the DM density contrast (normalized to the primordial value of the gravitational potential ψ) with $k/k_{\text{RH}} = 600$ in several cosmologies with early matter domination as a function of the scale factor normalized to its value at reheating. The upper line corresponds to a cosmology where the mode enters the horizon when the universe is already in the EMD phase. In the lower two lines, the background cosmology experiences only a brief period of EMD that lasts a factor of ~ 2000 (~ 100) in scale factor for the middle (lowest) line (a_{eeq} refers to the scale factor of the “early matter-radiation equality” when EMD began). The upper two lines correspond to the two background cosmologies shown in Fig. 1.

It is also useful to express the densities in terms of the cosmological scale factor, a . Defining a_{RH} such that during EMD $a = a_{\text{RH}}(t/\tau)^{2/3}$ then¹

$$\bar{\rho}_{\text{LM}}(a) = \frac{f}{6\pi G \tau^2} \left(\frac{a_{\text{RH}}}{a} \right)^3 \quad (2.6)$$

which is valid at all times. During LRD we find

$$a(t) \simeq 1.1 a_{\text{RH}} \sqrt{\frac{t}{\tau}} \quad (2.7)$$

so after nearly all the EM has decayed, i.e. during LRD and LMD

$$\bar{\rho}_{\text{LR}}(a) \simeq \frac{1}{7.2\pi G \tau^2} \left(\frac{a_{\text{RH}}}{a} \right)^4 \quad (2.8)$$

and the Hubble parameter is

$$H(a) \simeq \frac{1}{\tau} \left(\frac{a_{\text{RH}}}{a} \right)^2 \sqrt{0.37 + \frac{4}{9} f \frac{a}{a_{\text{RH}}}} = \frac{0.88 f^2}{\tau} \sqrt{\left(\frac{a_{\text{eq}}}{a} \right)^4 + \left(\frac{a_{\text{eq}}}{a} \right)^3} \quad (2.9)$$

where $a_{\text{eq}} = 0.83 a_{\text{RH}}/f$ is the scale factor when the LM and LR density are equal. This we identify with matter-radiation equality ($z_{\text{eq}} \simeq 3400$, $t_{\text{eq}} \simeq 51$ kyr) determined from late time observational cosmology [2]. Note that f and T_{RH} are not independent parameters, but are related by fixing the temperature and redshift of equality to the observed values, leading to $T_{\text{RH}} \sim \text{MeV} (10^{-8}/f)$.

Density perturbations experience enhanced growth during the EMD phase of cosmological evolution. Perturbations that enter the horizon at $a_{\text{hor}}(k)$ during EMD (i.e., $a_{\text{hor}} < a_{\text{RH}}$) are boosted by a factor of $a_{\text{RH}}/a_{\text{hor}}$. This effect was first studied in detail in Ref. [33]. The growth of density perturbations for several durations of EMD (i.e., various values of $a_{\text{RH}}/a_{\text{eeq}}$) is shown in Fig. 2. This evolution is obtained by numerically solving first order perturbation equations for EM, LM and radiation fluids coupled to gravity as in, e.g., Refs. [33, 39]. We will present partial analytic results for the scaling of the density contrast with wavenumber in §6 and numerical details in Appendix C. For now it is sufficient to observe that small-scale modes can be enhanced by orders of magnitude compared to the standard assumption of radiation-dominated growth of nearly-Harrison-Zeldovich perturbations. In particular, it is clear that certain perturbations can reach non-linearity and collapse before reheating, leading to the formation of EHs. The wavenumber in Fig. 2 is chosen to roughly correspond to the largest scales that can collapse – that is $\delta/\psi \sim 10^5$ at $a/a_{\text{RH}} \sim 1$ means that $\delta \sim 1$ for $\psi \sim 10^{-5}$, where ψ is the amplitude of the primordial gravitational potential. Smaller scales will collapse earlier, so if EMD lasts long enough most of the EM and LM will be bound in EHs of some size. The fraction of matter in EHs of a given mass and their distribution can be estimated by using linear perturbation theory in conjunction with the Press-Schechter formalism [43], as we describe in §6. In the following two sections we first focus on the evolution of individual EHs through reheating and beyond. Since these are non-linear structures, perturbation theory does not capture their dynamics, and we will instead study them using Newtonian equations of motion.

3 A Brief History of Early Halos

In what follows we show that in a broad range of scenarios the matter in the EHs progress through the following stages of evolution from early matter domination (EMD) to late matter domination (LMD):

¹There are several possibilities for defining a_{RH} , including via $t(a_{\text{RH}}) = \tau$, $\bar{\rho}_{\text{mat}}(a_{\text{RH}}) = \bar{\rho}_{\text{rad}}(a_{\text{RH}})$, $H(a_{\text{RH}}) = 1/\tau$, which give similar numerical values for a_{RH} up to $\mathcal{O}(1)$ factors. We will use the first definition for the semi-analytic results and the second in our numerics without introducing different notation for this characteristic scale factor.

- **stable clustering:** individual halo retain constant physical size during EMD, $a \lesssim 0.3 a_{\text{RH}}$. Self gravity dominates and EM has not yet undergone significant decay.
- **adiabatic expansion:** individual halos grow exponentially in physical size retaining their profile during the EMD/LRD transition, $0.3 a_{\text{RH}} \lesssim a \lesssim 10 a_{\text{RH}}$, while a significant fraction of EM has decayed but self gravity still dominates.
- **peeling:** successive outer layers of the halo end adiabatic expansion and begin free expansion soon after LRD begins, $2 a_{\text{RH}} \lesssim a \lesssim 10 a_{\text{RH}}$. The longer orbital timescale of outer layers become comparable to the expansion time invalidating the adiabatic approximation.
- **free expansion:** individual halos grow logarithmically in comoving size during LRD, $2 a_{\text{RH}} \lesssim a \lesssim a_{\text{eq}}$. Nearly all EMD has decayed and self-gravity is unimportant. The remnant LM moves ballistically in a radiation dominated universe.
- **halo overlap:** the ballistically expanding halo remnants will overlap with neighboring halos during LRD, $2 a_{\text{RH}} \lesssim a \lesssim a_{\text{eq}}$. Rapid expansion evacuates the LM from the initial halo position but this is mostly filled in by LM from neighboring expanding halos.
- **recollapse:** the inhomogeneous LM distribution created by the superposition of overlapping halo remnants will recollapse to form new structures during LMD, $a \gtrsim a_{\text{eq}}$, when self gravity of the LM dominates over the LR.

Many of these stages are very short in duration and some are concurrent. Different parts of the halo may exhibit different behaviour at the same time. The intervals during which different stages occur depend weakly on the overdensity of the EHs and may vary somewhat from those quoted. This qualitative scenario is valid for EH overdensities $\lesssim 10^{17}$. The first stage of “stable clustering” depends on the cosmology. In common scenarios where the early halos are formed from small initial density perturbations, this early structure formation proceeds via hierarchical assembly, so EHs undergo mergers with similarly-sized EHs or accrete smaller EHs. While these sub-halos undergo tidal stripping, they may survive for many dynamical times, maintaining roughly a constant physical size. Therefore in these models “stable clustering” is at best a coarse approximation. Despite this, we will assume initial stable clustering as it enables a simple analytical model; detailed N -body simulations are required to test the error made in this assumption.

In the next section we describe the evolution stages of an individual EH quantitatively, while the remainder of the paper is dedicated to studying the overlap of EH remnants at late times and the impact on the small-scale power spectrum.

4 Evolution of an Isolated Halo

Here we describe the evolution of matter, starting with an early halo (EH) which formed during EMD. We make three simplifying assumptions: 1) that EHs have collapsed sufficiently in advance of $t = \tau$ so that it has reached a quasi-equilibrium state of stable clustering before significant decay, 2) the EHs are sufficiently separated that tidal interactions and mergers are rare and 3) the EM and LM occupy the same phase space distribution in the halos. The first two assumptions are closely tied to each other, requiring mergers of halos to be episodic rather than continuous. One can imagine scenarios where 3) would not be valid: e.g. if one of the two components have sufficiently different velocity dispersions allowing one to collapse on scales where the other does not. These assumptions allow us to model the initial conditions purely in term of the density and velocity profiles of EHs along with their spatial correlations. As alluded to above, the first two assumptions may not hold in the strict sense in certain cosmologies. For example, if the primordial power spectrum is nearly flat, resulting continuous and hierarchical formation of EHs, it is not clear that a given EH is ever in the isolated or

stable clustering regime. In such a situation, estimates of the merger halo merger rate indicate that most of the mass gain of a halo is through minor mergers which are less likely to disrupt the “core” density profile of the halo [44]. Moreover, N -body simulations of small Λ CDM halos show stable-clustering-like evolution of the concentration parameter [45]. These qualitative observations suggest that the stable clustering and isolated ansatz might be a reasonable starting point even for these initial power spectra. In models with a different initial condition, such as a narrow power spectrum spike, the formation of isolated EHs may be explicitly realized and the first two assumptions would be clearly satisfied. Ultimately, the assumptions of stable clustering and rarity of disruptive tidal events enable a simple analytical study of the fate of EHs, but it will be important to thoroughly validate these in N -body simulations.

Upon collapse during EMD the matter overdensity, $\delta \equiv \rho_{\text{mat}}/\bar{\rho}_{\text{mat}} - 1$, increases to $\gtrsim 100$ and then increases as $\sim a^3$, becoming extremely large in a short number of expansion times. The halo dynamical timescale, $\tau_{\text{dyn}} \sim (H\sqrt{1+\delta})^{-1}$, becomes much less than the expansion timescale $\sim 1/H$, allowing a halo to rapidly approach stable clustering. So long as $\tau_{\text{dyn}} \ll \tau$ a halo will respond adiabatically to its decaying matter content, a process described in Appendix A and which we refer to as “adiabatic expansion”. During this phase a halo undergoes homologous expansion, with its physical size growing as $r \propto e^{+t/\tau}$ and its internal velocities shrinking as $v \propto e^{-t/\tau}$. During adiabatic expansion $\tau_{\text{dyn}} \sim r/v \propto e^{2t/\tau}$.

Adiabatic expansion ends when $\tau_{\text{dyn}} \gtrsim \tau$, i.e. when there is significant EM decay during a single orbit. As halos start with such a large overdensity this will take many decay times. The gravitation of the radiation also works to disrupt halos but most the radiation is between the halos and not within them; as a result, these forces are not significant before or during disruption. When disruption does occur the gravitational attraction binding the halos becomes ineffectual and the LM particles free-stream away from the initial halo center, acting as test particles in a radiation dominated universe. We call this stage “free expansion” since the gravitation of EM and LM are unimportant. Free expansion is slower than the exponential growth during adiabatic expansion. The total free streaming length during the radiation era is larger than the inter-halo separation erasing structures on and above the mass scale of the EHs. Individual halos freely expand into a much larger volume such that their density is below the cosmological mean. The final LM distribution is the superposition of these underdense halos, so that the locations of the EH remnants are not necessarily physically underdense once their overlap is taken into account.

Individual EHs start as non-linear overdensities, $\delta \gg 1$, and evolve into non-linear underdensities, $1 + \delta \ll 1$, so none of this evolution can be accurately described by linear theory. One can however use linear theory to track the evolution of the central position of halo remnants.

4.1 Ejection Velocity and Free-Streaming Length

First, we develop some intuition for the expected behaviour of EHs based on simple scaling arguments. The size of the free-streamed remnant will depend mostly on the velocity at which the constituent particles are ejected. If one characterizes a halo by a single mass and size, M and R , then $\tau_{\text{dyn}} \sim \sqrt{R^3/(GM)}$ and by the Virial theorem the internal velocities have magnitude $V \sim \sqrt{GM/R}$. During adiabatic expansion $M \simeq M_0 e^{-t/\tau}$, $R \simeq R_0 e^{t/\tau}$ and $V \simeq V_0 e^{-t/\tau}$ but this ends when $\tau_{\text{dyn}} \sim \tau$ or $t \equiv t_{\text{free}} \simeq \frac{1}{4}\tau \ln \delta_\tau$ where $\delta_\tau \sim GM_0 \tau^2/R_0^3$ is roughly the overdensity of the halo when $t = \tau$. The characteristic size and internal velocity during the transition to free expansion is

$$R_{\text{free}} = R_0 e^{t_{\text{free}}/\tau} \sim R_0 \delta_\tau^{1/4}, \quad V_{\text{free}} \sim \frac{R_{\text{free}}}{\tau_{\text{dyn}}(t_{\text{free}})} \sim \frac{R_0}{\tau} \delta_\tau^{1/4}. \quad (4.1)$$

Collapsed halos have $\delta_\tau \gg 1$ so t_{free} is significantly larger than τ and this transition occurs during the radiation era. After this transition the LM particles will free-stream outwards in a radiation

dominated universe, traversing a physical distance

$$\Delta r \sim V_{\text{free}} \sqrt{t t_{\text{free}}} \ln \left(\frac{t}{t_{\text{free}}} \right) \sim \frac{1}{2} R_0 \sqrt{\frac{t}{\tau} \sqrt{\delta_\tau} \ln \delta_\tau} \ln \left(\frac{4t}{\tau \ln \delta_\tau} \right) \quad (4.2)$$

at time t . $\Delta r \gg R_{\text{free}}$ soon after free expansion starts so Δr which excludes an initial offset gives an accurate estimate of the distance from the halo center.

If EHs are formed by hierarchical clustering during EMD they will not be well characterized by a single mass and size, M and R , but instead span a large range of densities $\propto \delta_\tau$, the central regions having much larger δ_τ while containing a small fraction of the mass. The time of transition from adiabatic to free expansion, $t_{\text{free}} \propto \ln \delta_\tau$ will happen somewhat later for orbits near the center of the halo than at the edge. More significantly $V_{\text{free}} \propto \sqrt[3]{M_0 R_0}$ will increase from the center to the edge for any halo profile. Thus the outer regions of the halo will be ejected first and at higher velocity than the inner regions and therefore free-stream to larger distances. Thus the inner parts of the halo will remain interior to the outer parts. The peeling dynamic determines the density profile of LM ejecta which in turn determines the remnants structure on the smallest scales. A quantitative description of peeling using a spherical halo model is given next.

4.2 Spherical Halos

In almost any formation scenario one expects collapsed structures (EHs) to be roughly spherical, similar to dark matter halos present today. We therefore use a spherical approximation for EHs. For spherical structures one need only follow the dynamics of spherical shells of dark matter rather than point particles; the shells contain all particles with the same radius, radial velocity and absolute value of angular momentum about the center. One also expects EHs to be non-relativistic and far smaller than the horizon, so we use the Newtonian equations of motion

$$\ddot{r}(t) = -\frac{G M_{<}(r(t), t)}{r(t)^2} + \frac{L^2}{r(t)^3} - \frac{8\pi G}{3} \bar{\rho}_{\text{rad}}(t) r(t) \quad (4.3)$$

where $\dot{} \equiv \partial/\partial t$, r is the physical (not comoving) radius of a particular shell, $M_{<}$ is the total mass of non-relativistic matter within radius r and $L = r v_\perp$ is the specific angular momentum which is conserved in spherically-symmetric potentials. The last term allows for the halo to be immersed in an expanding uniformly distributed bath of cosmological radiation of density $\bar{\rho}_{\text{rad}}$. Since non-relativistic halos have shallow gravitational potential wells ($-\Phi \ll c^2$) their presence does not alter the radiation distribution significantly. Recall that radiation contributes twice its mass density to its gravitationally attractive force since generally $\nabla^2 \Phi = 4\pi G (\rho + 3p/c^2)$ using $\bar{p}_{\text{rad}} = \bar{\rho}_{\text{rad}} c^2/3$ one has $\nabla^2 \bar{\Phi}_{\text{rad}} = 8\pi G \bar{\rho}_{\text{rad}}$.

The quantity $M_{<}$ contains both EM and LM which we assume have identical phase space distribution, so

$$M_{<}(r, t) = M_{\text{EM}}(r, t) + M_{\text{LM}}(r, t) \quad (4.4a)$$

$$M_{\text{EM}}(r, t) = \frac{1-f}{f} M_{\text{LM}}(r, t) e^{-t/\tau} . \quad (4.4b)$$

$M_{<}(r, t)$ will vary if different shells cross. We next consider a special case where they do not cross and one can evolve each shell independently.

4.3 Initially Circular Orbits

A simple initial condition, where all the particles start in initially circular orbits illustrates the general behavior. In this case $r(0) = r_0$, $\dot{r}(0) = 0$ and $L = \sqrt{G M_{<}(r_0, 0) r_0}$. Throughout the evolution of

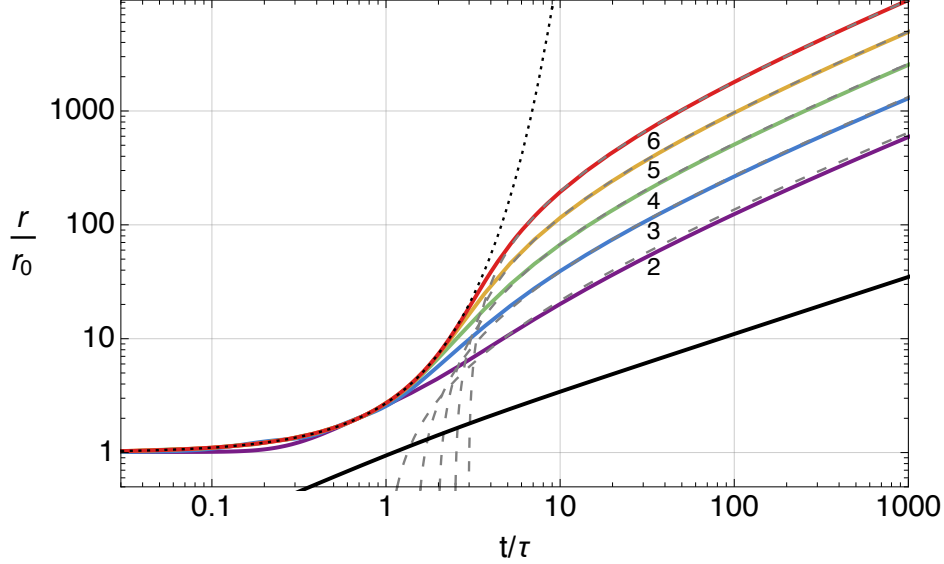


Figure 3. The colored curves give the evolution of the radius of shells in a spherical halo with initially circular orbits in units of the initial radius r_0 . The curves are labeled by $\log_{10}\delta_\tau$ ($\delta_\tau \equiv G M_0 \tau^2 / r_0^3$ where M_0 is the mass initially contained within the shell and τ is the EM lifetime). The dotted curve gives the predicted initial adiabatic exponential expansion and the long-dashed curves the approximation of Eqs. 4.6-4.8 to the asymptotic free expansion behavior. Shells successively “peel away” from exponential expansion: outer shells with smaller δ_τ first and inner shells with larger δ_τ later. The solid black curve shows the cosmological scale factor $a(t)/a(\tau)$. During adiabatic expansion the shells briefly expand much faster than $a(t)$ so their comoving size grows rapidly. During free expansion the shells grow only logarithmically faster than $a(t)$, their comoving size increasing further.

multiple shells the shells never cross: during adiabatic expansion $r(t) \simeq r_0 e^{t/\tau}$ while peeling retains the ordering of the shells. Thus M_{LM} is constant for each shell and one can replace $M_<(r(t), t) \rightarrow M(r_0) (f + (1-f) e^{-t/\tau})$. For initially circular orbits and $f \ll 1$ Eq. 4.3 becomes

$$\tau^2 \frac{\ddot{r}(t)}{r_0} = \delta_\tau(r_0) \left(\frac{r_0}{r(t)} \right)^3 \left(1 - e^{-t/\tau} \frac{r(t)}{r_0} \right) - \frac{8\pi G \bar{\rho}_{\text{rad}}(t) \tau^2}{3} \frac{r(t)}{r_0} \quad (4.5)$$

where $\delta_\tau(r_0) \equiv G M(r_0) \tau^2 / r_0^3$ as above. Assuming the mass density decreases from the center $\delta_\tau(r_0)$ will be a decreasing function of r_0 . From the form of Eq. 4.5 and initial conditions $E(0, \delta_\tau) = 1$ and $\dot{E}(0, \delta_\tau) = 0$ the solutions are of the form

$$r(t) = r_0 E\left(\frac{t}{\tau}, \delta_\tau(r_0)\right) \quad (4.6)$$

just as in Eq. 4.2. Initially ($t \lesssim \tau$) we expect adiabatic exponential expansion $E(T, \delta_\tau) \simeq e^T$. After adiabatic expansion ends in the radiation era at $T \sim T_{\text{free}}$, $8\pi G \bar{\rho}_{\text{rad}}/3 \rightarrow 1/(4t^2)$ and the first term on the right-hand side of Eq. 4.5 quickly becomes negligible so $\ddot{r} \simeq -r/(4t^2)$; the asymptotic solution is therefore

$$E(T, \delta_\tau) \rightarrow E_{\text{free}}(\delta_\tau) \sqrt{T} \ln\left(\frac{T}{T_{\text{free}}(\delta_\tau)}\right). \quad (4.7)$$

Only this asymptotic limit is relevant for late time inhomogeneities, and it is characterized by two functions:

$$E_{\text{free}}(\delta_\tau) \simeq 0.435 \sqrt{\sqrt{\delta_\tau} \ln \delta_\tau} \quad T_{\text{free}}(\delta_\tau) \simeq 0.215 \ln \delta_\tau \quad (4.8)$$

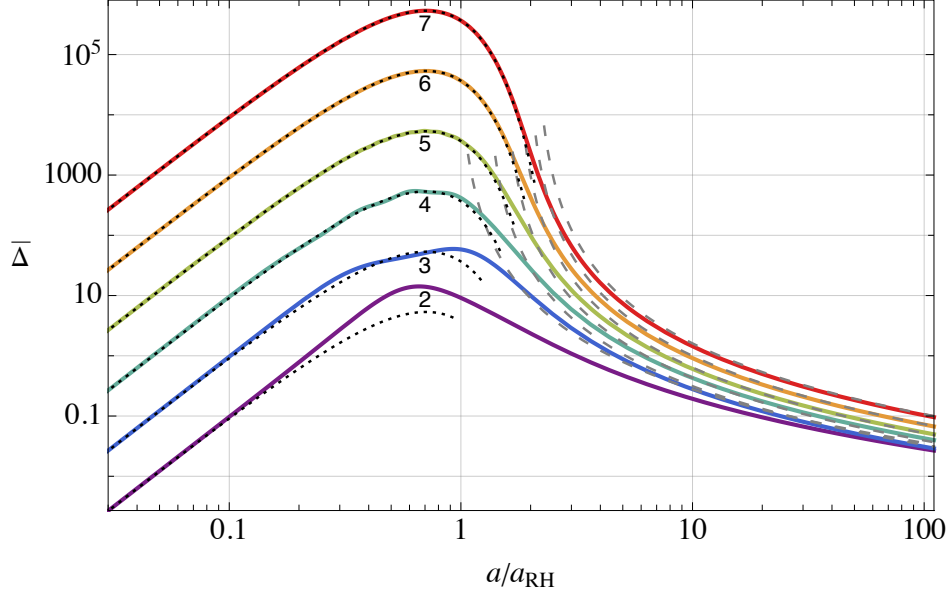


Figure 4. The colored curves show, as a function of the cosmological scale factor a , the ratio of the mean matter density within a shell to the mean cosmological density ($\bar{\Delta}$) for a halo with initially circular orbits. The curves are labeled by $\log_{10} \bar{\Delta}_{\tau}$ (see Eq. 4.13). The dotted curves give the predicted initial adiabatic exponential expansion and the long-dashed curves the asymptotic approximation of Eq. 4.15. $\bar{\Delta}$ does not include matter from overlapping neighboring halos.

The functional form is taken from the scaling derived in §4.1. T_{free} and E_{free} are determined by the cosmic time at the beginning of free expansion of a shell and by the velocity at this time:

$$\begin{aligned}
 t_{\text{free}}(\delta_{\tau}) &\equiv T_{\text{free}}(\delta_{\tau}) \tau \simeq 0.215 \tau \ln \delta_{\tau} = 0.215 \tau \ln \left(\frac{G M(r_0) \tau^2}{r_0^3} \right) \\
 v_{\text{free}}(\delta_{\tau}) &\equiv \frac{r_0}{\tau} \frac{E_{\text{free}}(\delta_{\tau})}{\sqrt{T_{\text{free}}(\delta_{\tau})}} \simeq 0.938 \frac{r_0}{\tau} \delta_{\tau}^{1/4} = 0.938 \left(\frac{G M(r_0) r_0}{\tau^2} \right)^{1/4}.
 \end{aligned} \tag{4.9}$$

For spherical halos with initially circular orbits individual shells accurately follow the δ_{τ} scaling derived in §4.1 for entire halos with small corrections to numerical factors: $1/4 \rightarrow 0.215$ and $1 \rightarrow 0.938$.

Numerical solutions of Eq. 4.7 and asymptotic fitting functions are shown in Fig. 3. Adiabatic exponential expansion very accurately describes the early evolution for $\delta_{\tau} \gtrsim 10^3$ as does the asymptotic fitting function. The accuracy of these approximations rapidly become worse for $\delta_{\tau} < 10^2$ because adiabaticity during EM decay is less closely realized.

4.4 Spherical Halos in a Cosmological Context

One can translate the physical radius of shells into a quantity more closely related to the background cosmology by using the cosmological scale factor a in place of t and in place of r using the “interior density ratio” defined by

$$\bar{\Delta}(a(t)) \equiv \frac{3M_{<}(r(t), t)}{4\pi r^3 \bar{\rho}_{\text{mat}}(t)} = \frac{9}{2} \frac{G M_{\text{LM}}(r(t)) \tau^2}{f r(t)^3} \left(\frac{a(t)}{a_{\text{RH}}} \right)^3, \tag{4.10}$$

where Eqs. 2.6 and 4.4 were used to obtain the second expression. This is the ratio of the matter mass within a spherical shell to the mass within the same sphere at the cosmological average density.

The last form emphasizes that the decay of EM does not in itself cause $\bar{\Delta}$ to change since $M_{\text{LM}}(a)$ only changes due to shell crossing. To transform the Newtonian Eq. 4.3 into an equation for $\bar{\Delta}$ define density parameters $\Omega_{\text{m}}(a) \equiv 8\pi G \bar{\rho}_{\text{mat}}/(3H^2)$ and $\Omega_{\text{r}}(a) \equiv 8\pi G \bar{\rho}_{\text{rad}}/(3H^2)$ and use Eqs. 2.1 and 2.2 to obtain the flatness condition $\Omega_{\text{m}} + \Omega_{\text{r}} = 1$, the deceleration parameter $q(a) \equiv -a\ddot{a}/\dot{a}^2 = 1 - \frac{1}{2}\Omega_{\text{m}}$ and find

$$a^2 \bar{\Delta}'' = \frac{3}{2} \Omega_{\text{m}} \bar{\Delta} (\bar{\Delta} - 1) - \left(1 + \frac{1}{2} \Omega_{\text{m}}\right) a \bar{\Delta}' + \frac{4}{3} \frac{(a \bar{\Delta}')^2}{\bar{\Delta}} - 3 \left(\frac{L}{H}\right)^2 \left(\frac{2f}{9G M_{\text{LM}} \tau^2}\right)^{4/3} \left(\frac{a_{\text{RH}}}{a}\right)^4 \bar{\Delta}^{7/3}. \quad (4.11)$$

where $' \equiv \partial/\partial a$ and M_{LM} is the LM mass within the shell which may vary due to shell crossing but not due to EM decay.

Overdensity, $\delta_{\text{m}} \equiv \rho_{\text{mat}}/\bar{\rho}_{\text{mat}} - 1$, is the quantity usually followed in cosmological perturbation analysis. For a spherical halo the mean matter overdensity within a spherical shell is $\bar{\delta}_{\text{m}} = \bar{\Delta} - 1$. One recovers linear perturbation theory of scalar (non-vortical) inhomogeneities by setting $L = 0$ and taking the limits $|\bar{\delta}_{\text{m}}| \ll 1$ and $(\bar{\delta}_{\text{m}}')^2 \ll |\bar{\delta}_{\text{m}}''|$ obtaining

$$\bar{\delta}_{\text{m}}'' + \left(1 + \frac{1}{2} \Omega_{\text{m}}\right) \frac{1}{a} \bar{\delta}_{\text{m}}' - \frac{3}{2} \frac{1}{a^2} \Omega_{\text{m}} \bar{\delta}_{\text{m}} = 0. \quad (4.12)$$

For standard evolution of the background energy densities (i.e., no EMD) this reduces to the Meszaros equation [46, 47]. Unfortunately, the conditions for the validity of this linear equation are not met for EHs and their subsequent evolution. During a brief period when $|\bar{\delta}_{\text{m}}| \ll 1$ the second condition on $\bar{\delta}_{\text{m}}''$ is not satisfied.

During EMD where $\Omega_{\text{m}} \rightarrow 1$ an exact circular orbit solution of Eq. 4.11 is

$$\bar{\Delta}(a) = \bar{\Delta}_{\tau} \left(\frac{a}{a_{\text{RH}}}\right)^3 \quad \bar{\Delta}_{\tau} \equiv \frac{9(GM_0)^4 \tau^2}{2L^6} = \frac{9GM_0 \tau^2}{2r_0^3} = \frac{9}{2} \delta_{\tau} \quad (4.13)$$

where $M_{\text{LM}} = M_0/f$ is constant since shells do not cross and $r_0 = L^2/(GM_0)$ is the constant physical radius of the orbiting shell. Note that $\bar{\Delta}_{\tau}$ here is a convenient short-hand, but the overdensity never reaches this value, since the decays of EM are already significant at $a = a_{\text{RH}}$; the quantity δ_{τ} was used in §4.3. Using Eq. 2.7 the approximate asymptotic LRD solution for initially circular orbits of Eqs. 4.6–4.8 & 2.7 in terms of a is

$$r(a) \rightarrow 0.450 \left(\frac{a}{f a_{\text{eq}}} r_0\right) \left(\bar{\Delta}_{\tau}^{1/4} \sqrt{l_{\text{ad}}(\bar{\Delta}_{\tau})}\right) l_{\text{free}}(\bar{\Delta}_{\tau}, a) \quad (4.14)$$

$$l_{\text{ad}}(\bar{\Delta}_{\tau}) \equiv \ln \frac{2}{9} \bar{\Delta}_{\tau} \quad l_{\text{free}}(\bar{\Delta}_{\tau}, a) = \ln \frac{1.62 a}{f a_{\text{eq}} \sqrt{\ln \frac{2}{9} \bar{\Delta}_{\tau}}}.$$

The first non-numerical term gives the cosmological expansion while the second and third give additional growth during adiabatic expansion and free expansion, respectively. Here $l_{\text{free}} \equiv \ln(a/a_{\text{free}})$ where $a_{\text{free}} \equiv a(t_{\text{free}})$; it is interesting to note that free expansion is the same physical effect that gives rise to the logarithmic growth of small-scale density fluctuations in Λ CDM - it generates a decay of density perturbations in our case due to the initial period of adiabatic expansion. Eq. 4.14 is valid for $a \lesssim a_{\text{eq}}$ so l_{free} will grow to be large since f is very small. Substituting Eq. 4.14 into Eq. 4.10 yields

$$\bar{\Delta}(a) \rightarrow \frac{5.33 \bar{\Delta}_{\tau}^{1/4}}{l_{\text{ad}}(\bar{\Delta}_{\tau})^{3/2} l_{\text{free}}(\bar{\Delta}_{\tau}, a)^3}. \quad (4.15)$$

From $\bar{\Delta}_{\tau}$ one can compute the density ratio $\Delta \equiv \rho_{\text{LM}}/\bar{\rho}_{\text{LM}} = \bar{\Delta} \left(1 + \frac{1}{3} \frac{d \ln \bar{\Delta}}{d \ln r}\right)$ which is asymptotically given by

$$\frac{\Delta(a)}{\bar{\Delta}(a)} \rightarrow \frac{\Delta_{\tau}}{\bar{\Delta}_{\tau} - \frac{3}{4} \left(1 + \frac{2}{l_{\text{ad}}(\bar{\Delta}_{\tau})} \left(1 - \frac{1}{l_{\text{free}}(\bar{\Delta}_{\tau}, a)}\right)\right) (\bar{\Delta}_{\tau} - \Delta_{\tau})}. \quad (4.16)$$

For a given shell the temporal dependence of $\bar{\Delta}$ and Δ is solely through $l_{\text{free}}(\bar{\Delta}_\tau, a)$ which is a large logarithm (since f is small) whose a dependence is slow. Thus $\bar{\Delta}(a)$ and $\Delta(a)$ decrease slowly during free expansion. If $\bar{\Delta}_\tau \gg 1$ then l_{ad} is large (l_{free} is usually even larger). In the “large logarithm approximation” $\Delta(a)/\bar{\Delta}(a) \approx 4 \Delta_\tau / (\bar{\Delta}_\tau + 3 \Delta_\tau)$.

In Fig. 4 numerical solutions for $\bar{\Delta}$ are plotted for initially circular orbit model of §4.3. Initial adiabatic exponential expansion and the asymptotic approximation are very accurate for $\bar{\Delta}_\tau \gtrsim 10^4$ and become more accurate for larger $\bar{\Delta}_\tau$. For $a \leq a_{\text{eq}}$ the errors are $< 10\%$ for $\bar{\Delta}_\tau \geq 10^4$ but $\leq 45\%$ and $\leq 63\%$ for $\bar{\Delta}_\tau = 10^3$ and 10^2 . The adiabatic approximation is poor for $\bar{\Delta}_\tau \lesssim 10^2$ which is not surprising since one sees that $\bar{\Delta}_\tau = 10^2$ only reaches $\max_a \bar{\Delta}(a) \simeq 10$ so it does not correspond to a large overdensity. For $\bar{\Delta}_\tau \gg 10^3$ one finds $\max_a \bar{\Delta}(a) \simeq 0.053 \bar{\Delta}_\tau$ at $a \simeq 0.7 a_{\text{RH}}$. Thus while δ_τ and $\bar{\Delta}_\tau$ are algebraically convenient quantities they overestimate the maximum overdensity by a factors ≈ 4 and 19 , respectively. For $\bar{\Delta}_\tau \lesssim 10^2$ the adiabatic approximation becomes inaccurate and at $\bar{\Delta}_\tau \leq 9/2$ ($\delta_\tau \leq 1$) it yields unphysical results since $l_{\text{ad}} < 0$ and l_{free} is complex. One should not use these formulae for small $\bar{\Delta}_\tau$; this approximation does not reduce to linear theory for $\bar{\Delta}_\tau$ near unity.

So far we have focused on spherical EHs with circular, non-crossing orbits. It is important to verify whether our results are robust to more physical initial conditions. While an N -body numerical study of EHs is beyond our scope, in Appendix B we used an N -shell code to confirm that the qualitative and quantitative evolution described here holds even for more generic choices of shell angular momenta, initial density profiles of EHs and allowing for shell crossing. The shell model does not account for the distribution of velocities/angular momenta at each radius which will serve to smear out the remnant in position space. We will argue below that these effects do not qualitatively change our findings.

4.5 Recollapse?

During EMD stable clustering the halo maintains constant physical density so the density ratio increases rapidly: $\bar{\Delta} \propto a^3$. During LRD the LM halo experiences adiabatic and free expansion, both of which are faster than the cosmological expansion rate, causing decreases in the density ratio: $\bar{\Delta} \propto a^3 e^{-3t/\tau}$ (adiabatic) and $\bar{\Delta} \propto (\ln a/a_{\text{free}})^{-3}$ (free). During LMD a shell may either recollapse or continue expanding depending on whether the shell is gravitationally bound or free. A bound shell has specific gravitational binding energy, $U = G M_{\text{LM}}/r$ larger than specific kinetic energy $K = (\dot{r}^2 + (L/r)^2)/2$, while free shells have $K > U$. During adiabatic expansion U and K retain their Virial ratio, $U \approx 2K$. During free expansion the trajectories become nearly radial so $K \rightarrow \dot{r}^2/2$ so we find using Eq. 4.10 to relate shell radius to the density ratio $\bar{\Delta}$

$$\frac{U}{K} = \frac{\bar{\rho}_{\text{mat}}}{(\bar{\rho}_{\text{mat}} + \bar{\rho}_{\text{rad}})} \frac{\bar{\Delta}^3}{(\bar{\Delta} - \frac{a}{3} \bar{\Delta}')^2} \xrightarrow{a \ll a_{\text{eq}} \ll a_{\text{eq}}} \frac{a}{(a_{\text{eq}} + a)} \left(\frac{l_{\text{free}}}{1 + l_{\text{free}}} \right)^2 \bar{\Delta} \quad (4.17)$$

The last expression, derived from Eq. 4.15 and valid during LRD free expansion, gives $\lim_{a \rightarrow a_{\text{eq}}} U/K \approx \bar{\Delta}_{\text{eq}}/2$ where $\bar{\Delta}_{\text{eq}} \equiv \bar{\Delta}(a_{\text{eq}})$. Thus if $\bar{\Delta}_{\text{eq}} \ll 1$ a shell is unbound and will continue to expand and never recollapse; while if $\bar{\Delta}_{\text{eq}} \gg 1$ the shell is bound entering the matter era and will recollapse. We show $\bar{\Delta}_{\text{eq}}$ as a function of $\bar{\Delta}_\tau$ in Fig. 5. One sees that the sign of the overdensity $\bar{\Delta}_{\text{eq}} - 1$ is a rough indicator of collapse just as with linear theory growing modes. The separatrix between these different behaviors we denote by $\bar{\Delta}_{\text{crit}}(a)$ which is the solution of Eq. 4.11 for which $\bar{\Delta}_{\text{crit}}(\infty) = 1$. This specifies a value $\bar{\Delta}_\tau = \bar{\Delta}_\tau^{\text{crit}}$. In Fig. 6 we plot $\bar{\Delta}_\tau^{\text{crit}}$ and $\bar{\Delta}_{\text{eq}}^{\text{crit}} \equiv \bar{\Delta}_{\text{crit}}(a_{\text{eq}})$ for various values of f . One sees that for allowed values of f that $\bar{\Delta}_\tau^{\text{crit}} \gtrsim 10^{20}$. One can understand this from Eq. 4.15 which estimates $\bar{\Delta}_{\text{eq}} \approx 5.33 \bar{\Delta}_\tau^{1/4} (\ln \frac{2}{9} \bar{\Delta}_\tau)^{-3/2} (\ln f^{-1})^{-3}$ so the condition for recollapse, $\bar{\Delta}_{\text{eq}} \gtrsim 1$, is $\bar{\Delta}_\tau \gg (\ln f^{-1})^{12}$ which is extremely large since $f \lesssim 10^{-8}$. Such large overdensities are not totally implausible: since $\bar{\Delta} \propto a^3$ during stable clustering and $\bar{\Delta}_\tau \sim 10^{21}$ could be attained when the universe expands by a factor $\gtrsim 10^7$ during EMD. Note that this means that these objects must have formed extremely early and therefore must be very light compared to the mass contained within the horizon at reheating.

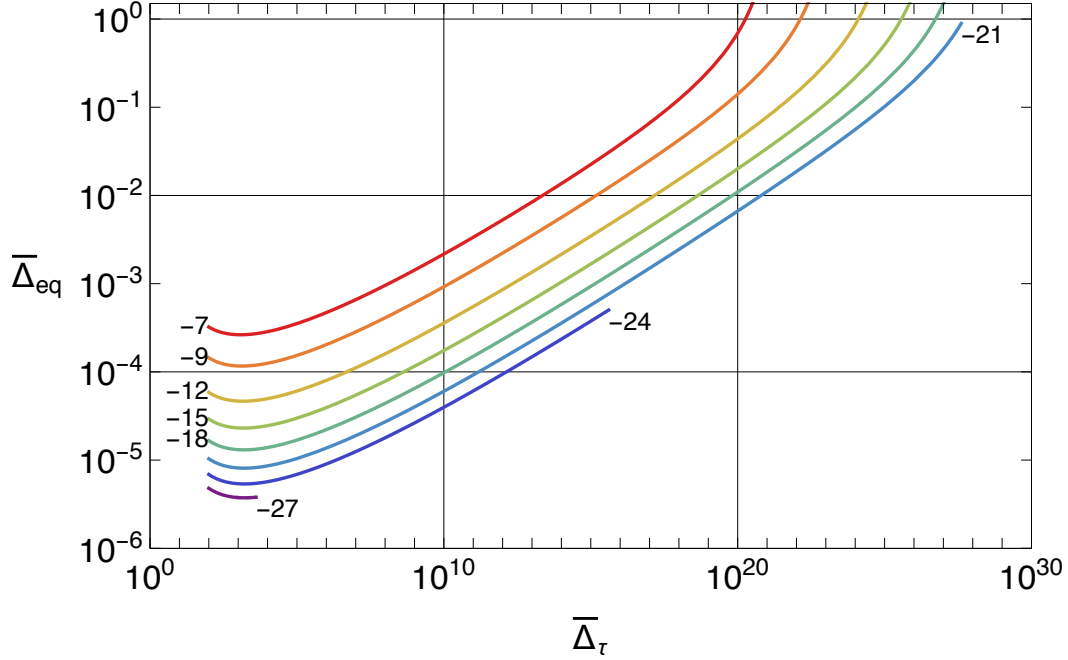


Figure 5. Density ratio of DM at equality $\bar{\Delta}_{\text{eq}}$ as a function of $\bar{\Delta}_{\tau}$ for different values f , the stable DM fraction which determines the duration of LRD. $\bar{\Delta}_{\tau}$ determines the EH density before reheating via Eq. 4.13. The numerical labels are values of $\log_{10} f$. These curves are computed for the model of §4.3 over the region of physical validity delineated in §4.7. Where visible the right endpoints correspond to shells with $\tau_{\text{dyn}} = t_{\text{pl}}$. Objects with $\bar{\Delta}_{\text{eq}} \gtrsim 1$ are expected to recollapse after matter-radiation equality. However, this requires enormous initial density contrasts before reheating.

For $\bar{\Delta}_{\tau} < \bar{\Delta}_{\tau}^{\text{crit}}$ an isolated halo becomes unbound and will not, by itself, recollapse to form a bound structure during LMD. In this case individual EHs which started as large overdensities during EMD are diluted below the cosmological mean density during LMD. It is only a superposition of these underdense remnants which will allow bound structures to form at late times.

4.6 Halo Remnant Density Profile

The nature of the bound structures which collapse during LMD will depend on the spatial distribution of LM given by the density structure of individual halo remnants and how these remnants are clustered. Here we again consider only individual halos treating them as isolated objects. As shown above the density profile of a halo remnant depends on the initial density profile of the EH. As we have not specified how EHs form there is a large degree of uncertainty in what these initial profiles are, though they clearly must satisfy the rules of gravitational clustering such as stability. Since we have already assumed sphericity this generally requires the density decreases with radius. Here we consider a variety of models of static stable spherical non-relativistic halos of gravitationally bound non-interacting particles using Eqs. 4.6-4.7, 4.15 to determine the asymptotic remnant profile. While this is formally correct for initially circular orbits Appendix B suggests it is a reasonable approximation more generally. Note that the density profile does not specify the phase space distribution and any density profile is consistent with purely circular orbits.

One characterizes a density profile by the radially-dependent logarithmic slope $\gamma(r) \equiv -\frac{\partial \ln \rho}{\partial \ln r} \geq 0$ (so $\rho \propto r^{-\gamma}$). Steepness/shallowness increases/decreases with $\gamma(r)$. Near the center EHs may have diverging (cusped $0 < \gamma(0) < 3$) or finite (cored $\gamma(0) = 0$) density. From Eq. 4.15 one finds the

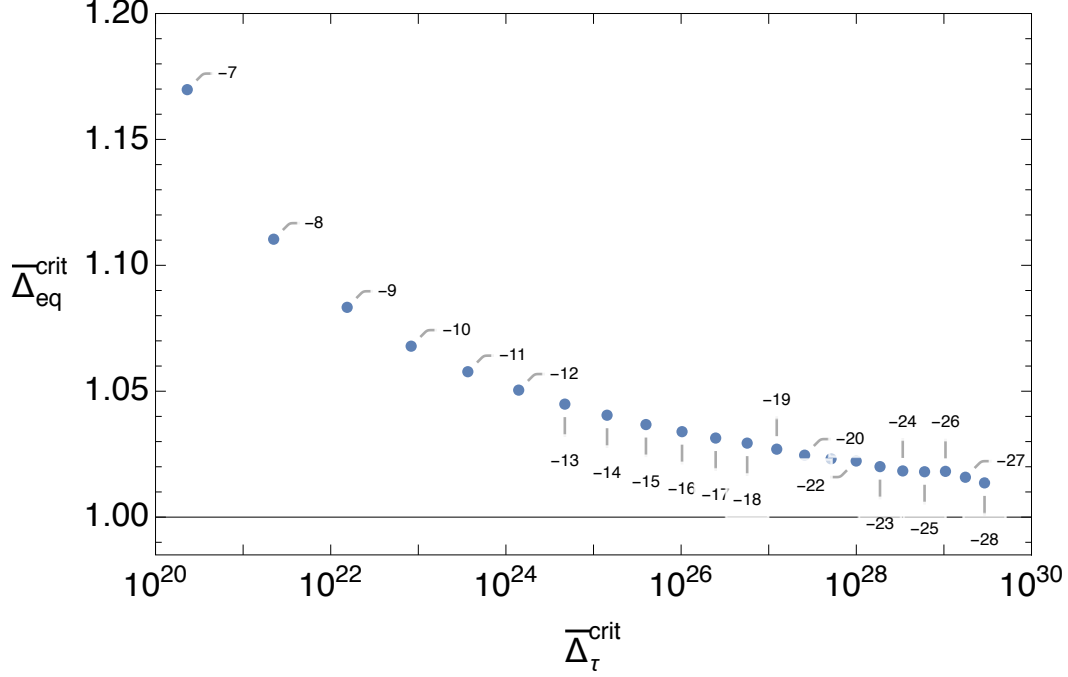


Figure 6. Each point gives $\bar{\Delta}$ for critical shells at two different epochs: $\bar{\Delta}_{\tau}^{\text{crit}} = 18.7 \bar{\Delta}(a_{\text{RH}})$ and $\bar{\Delta}_{\text{eq}}^{\text{crit}} = \bar{\Delta}(a_{\text{eq}})$. Critical shells evolve to $\bar{\Delta}(\infty) \rightarrow 1$. Shells with $\bar{\Delta}$ larger than these critical values at the corresponding time will recollapse while shells with smaller $\bar{\Delta}$ will evolve to $\bar{\Delta} \rightarrow 0$. The numerical labels give $\log_{10} f$ where f is the initial fraction of stable DM and determines the duration of late radiation domination (LRD). The values of f plotted are limited to the region of physical validity delineated in §4.7.

central slope of the asymptotic remnant is related to the central slope of the EH by

$$\gamma_{\text{rem}}(0) = \frac{\gamma_{\text{EH}}(0)}{4 - \gamma_{\text{EH}}(0)} \quad (4.18)$$

due to $\bar{\Delta} \sim \bar{\Delta}_{\tau}^{1/4}$ and $r \sim r_0 \bar{\Delta}_{\tau}^{1/4}$. Thus a cored EH evolves to a cored remnant and a cusped EH evolves to cusped remnant but with a shallower central slope. This indicates that the density is more uniform in the remnant than in the EH. The difference can be quite dramatic, e.g. $\rho_{\text{EH}} \propto r^{-1} \rightarrow \rho_{\text{rem}} \propto r^{-1/3}$.

Cosmological halos usually form by gravitational coalescence from a uniform medium and will grow in size and mass as more matter is accreted and becomes virialized. The outer edge of accretion, known as the virial radius, r_{vir} , is usually characterized by $\Delta_{\text{vir}} \equiv \frac{\rho_{\text{vir}}}{\rho_{\text{mat}}} = \frac{d(r^3 \bar{\Delta})}{dr^3} \big|_{r=r_{\text{vir}}} \approx 200$.² In contrast circular orbit halos have fixed physical size and do not allow for accretion since this involves radial not circular motion. The scenario we consider assumes most matter is virialized (described by stable clustering) by the end of EMD with only slow ongoing accretion / hierarchical clustering. To avoid small $\bar{\Delta}_{\tau}$ we need to define an edge to EHs. We do this by choosing an appropriate r_{vir} for the end of stable clustering and the beginning of adiabatic expansion at $t \sim \tau$. We choose this transition as the time when $\bar{\Delta}$ reaches its maximum value. As can be seen in Fig. 4 $\max \bar{\Delta} \approx \frac{1}{20} \bar{\Delta}_{\tau}$ so we define the edge of a halo, i.e. virial radius or virial shell, by $\bar{\Delta}_{\tau}^{\text{vir}} = 20 \times 200$. This model may not be accurate for the small fraction of matter with $r \gtrsim r_{\text{vir}}$.

Interior to a virial shell some but not all halos have an “extended atmosphere” which we define as an outer region which has large overdensity $\Delta_{\tau} \gg 1$ but contains a tiny fraction of the mass

²Note that here we distinguish the density ratio $\Delta = \rho/\bar{\rho}_{\text{mat}}$ from the averaged density ratio $\bar{\Delta}$ defined in Eq. 4.10.

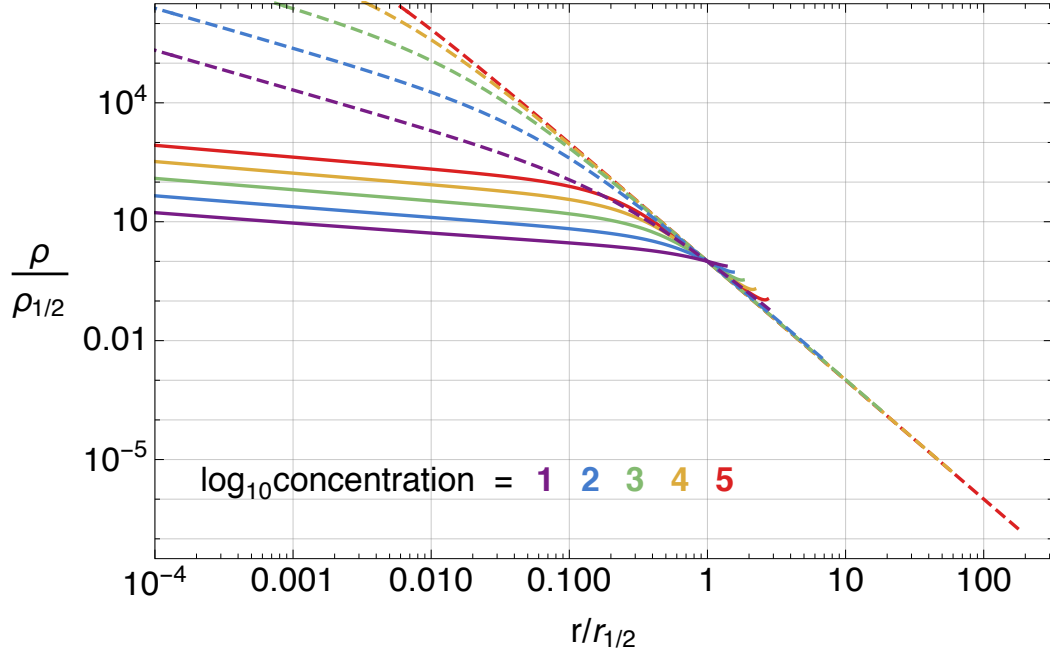


Figure 7. Density profiles of early halos relative to their half mass radius, $r_{1/2}$, and the density at this radius $\rho_{1/2}$. The dashed curves are the initial (NFW) density profile and the solid curve is the density profile of the asymptotic remnant from Eq. 4.15 for $a = 10^8 a_{\text{RH}}$ (the profile shape is insensitive to $a \gg a_{\text{RH}}$). The different curves have different concentrations = r_{vir}/r_s .

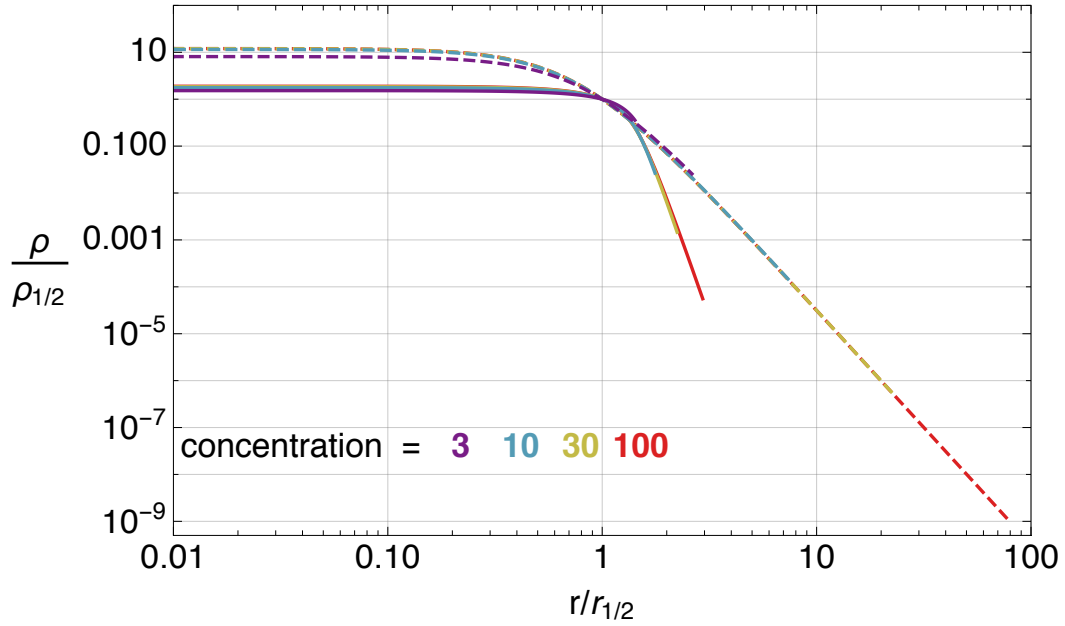


Figure 8. Same as Fig. 7 except for initial Plummer density profiles.

and $\Delta_\tau \ll \bar{\Delta}_\tau$. If the EH has a power law atmosphere $\Delta_\tau \propto r_0^{-\gamma_{\text{EH}}^{\text{atm}}}$ then using the large logarithm approximation and Eqs. 4.15-4.16 the remnant also has an atmosphere with

$$\gamma_{\text{rem}}^{\text{atm}} = 4\gamma_{\text{EH}}^{\text{atm}} - 9. \quad (4.19)$$

Since an extended atmosphere only exists for $\gamma_{\text{EH}}^{\text{atm}} > 3$ (otherwise the mass does not converge for large r_0) the slope of the remnant atmosphere is steeper than that of the EH: $\gamma_{\text{rem}}^{\text{atm}} > \gamma_{\text{EH}}^{\text{atm}}$.

Here we consider two contrasting models for EH density profile:

$$\rho_{\text{NFW}}(r) = \rho_{\text{vir}} \frac{r_{\text{vir}}(r_{\text{vir}} + r_s)^2}{r(r_s + r)^2} \quad \rho_{\text{P}}(r) = \rho_c \left(1 + \left(\frac{r}{r_s}\right)^2\right)^{-5/2}. \quad (4.20)$$

ρ_{NFW} is the Navarro-Frenk-White profile [48] which is currently the most commonly used model for dark matter halos in the late-time universe and ρ_{P} is the Plummer profile developed to model globular clusters [49]. ρ_{vir} is the virial density, ρ_c the central density and r_s a characteristic radius. We define the concentration $\mathcal{C} \equiv r_{\text{vir}}/r_s$ which is usually a large number. The contrasting properties are 1) NFW has a cusp and Plummer a core, 2) for Plummer most mass resides at $r \lesssim r_s \ll r_{\text{vir}}$ the total mass rapidly converging outside of this while for NFW the mass increases $\propto \ln r$ for $r \gg r_s$ and most resides at $r_s \ll r \ll r_{\text{vir}}$.

Given the EH profiles above the asymptotic solution in Eq. 4.15 can be used to derive the density profile of the remnant:

$$\rho_{\text{rem}}(r) = \frac{3M_{\text{LM}}(r)}{4\pi r^3} \left(1 + \frac{1}{3} \frac{d \ln \bar{\Delta}}{d \ln r}\right), \quad (4.21)$$

where the r -dependence of $\bar{\Delta}$ is through $\bar{\Delta}_\tau$ (see Eq. 4.13) and the original radius of a mass shell, r_0 (related to r via the asymptotic solutions in Eqs. 4.6 and 4.7). Note that since adiabatic expansion and peeling preserve shell order, the mass interior to r at late times, $M_{\text{LM}}(r, a \gg a_{\text{RH}})$ is equal to the mass interior to $r_0(r)$ at early times, $M_{\text{LM}}(r_0, a < a_{\text{RH}})$.

In Fig. 7 and 8 we show the density profiles of EHs and their remnants for various concentrations assuming the former are of the NFW and Plummer forms, respectively. The curves are normalized to the half mass radius and density. In both cases the remnant density profile is very shallow near the center and most of the mass resides at radii near the outer radius in the remnant which is a manifestation of the shallower density slope.

In summary remnants of EHs are not only larger in physical and comoving size relative to the original halo but their density profile significantly differs from the original halo. The inner profile is more shallow and the outer profile is steeper. In other words mass from an inner cusp is drawn outward being distributed more uniformly while the mass from any outer extended atmosphere is drawn inward leading to a sharper boundary to the remnant. The remnant mass is distributed among a smaller range of scales ($\ln r$) than its progenitor, closer to that of a Gaussian or top hat profile.

Our calculation of the remnant density profile relied on a shell model that does not fully account for the distribution of DM velocities in a bound halo. Particles that are in some common volume element in the halo at an initial instant in time will generally have different angular momenta and will therefore be on very different orbits that are not described by the same dynamical time (and for highly elliptical orbits the dynamical time is different in different phases of the orbit); as a result, these particles will exit adiabatic expansion at different times and in different phases of their orbits, which intuitively leads to a further flattening of the remnant density profile. In the following section we will show that treating the remnant as having a top hat profile gives a simple quantitative understanding of the power spectrum even when the original EH was NFW. Thus, a further flattening of the remnant profile due to the velocity dispersion should only serve to make the analogy with a top hat better.

4.7 Physical Parameters

In this section we collect various observational and theoretical constraints on physical quantities encountered so far.

Observational constraints limit viable values of the cosmological parameters $(\bar{\rho}_{\text{EM}}, \bar{\rho}_{\text{LM}}, \tau)$ and the derived parameter $f \equiv \lim_{t \rightarrow 0} \bar{\rho}_{\text{LM}}/(\bar{\rho}_{\text{EM}} + \bar{\rho}_{\text{LM}})$. The successful Λ CDM model of late time cosmology sets matter-radiation equality (LRD-LMD transition) at $t = t_{\text{eq}} \simeq 51$ kyr. In order for Big Bang Nucleosynthesis to yield the observed abundance of light elements requires LRD to begin with a reheating temperature, $T_{\text{RH}} \geq 5$ MeV ([8–11]) corresponding to $\tau \lesssim 0.03$ sec. Comparing 0.03 sec to 51 kyr one finds $f \lesssim 10^{-7}$.

Physical constraints limit the region of validity of the Newtonian analysis of this section, i.e. limit viable values of the EH parameters $(M_0$ and $r_0)$ and derived quantities $\delta_\tau = \frac{2}{9} \bar{\Delta}_\tau = G M_0 \tau^2 / r_0^3$. In order for the halo to have attained stable clustering before EM decay one requires $\tau \gg \tau_{\text{dyn}} = \sqrt{r_0^3 / (G M_0)}$ or $\delta_\tau \gg 1$.

A less certain constraint comes from shortest timescale on which our physical model is valid. If one takes the requirement that all timescales be much greater than the Planck time, t_{pl} , then $\tau \approx f^2 t_{\text{eq}} \gg t_{\text{pl}}$ and $\tau_{\text{dyn}} = \tau / \sqrt{\delta_\tau} \gg t_{\text{pl}}$ then one finds $10^{-28} \ll f \lesssim 10^{-7}$ and $1 \ll \delta_\tau = \frac{2}{9} \bar{\Delta}_\tau \ll 10^{83} f_{-7}^4$ where $f_{-7} \equiv f / 10^{-7}$. This does not make any allowance for a period of inflation which would further bound f from below.

Newtonian clustering requires non-relativistic EHs: $G M_0 / r_0 \ll c^2$ or $r_0 \ll 10^4 \text{ km } f_{-7}^2 \delta_\tau^{-1/2}$. The halos will become more non-relativistic during their subsequent expanding evolution. If $\bar{\Delta}_\tau < \bar{\Delta}_\tau^{\text{crit}}$ the halo remnant will expand by a factor given by Eqs. 4.7 and 4.8 until $a \sim a_{\text{eq}}$. Translating this limit on r_0 to the final expanded remnant size one finds $r_{\text{eq}} \ll 10^{11} \text{ km } f_{-7}^2 \sqrt{\frac{\ln \delta_\tau}{\delta_\tau}}$. EHs with $\bar{\Delta}_\tau \gg \bar{\Delta}_\tau^{\text{crit}}$ will collapse earlier during LRD and leave even smaller remnants.

If EHs coalesce from a uniform expanding cosmology then causality requires that their mass be much less than the horizon mass at $t = \tau$ or earlier. This requirement is $M_0 \lesssim c^3 \tau / G = 10^4 M_\odot f_{-7}^2$. The remnant LM mass must then satisfy $M_{\text{LM}} = f M_0 \lesssim 10^{-3} M_\odot f_{-7}^3$. The largest halo remnant in size and in mass is smaller than any dark matter structure yet observed.

Two physical quantities which do not enter our analysis are the masses of EM and LM particles. In order to form gravitationally bound structures both of their de Broglie wavelengths should be less than the size of the halo structures. This requirement may be written $m_{\text{min}} \equiv \min(m_{\text{EM}}, m_{\text{LM}}) \gg \hbar / \sqrt{G M_0 r_0} = \hbar (\sqrt{\delta_\tau} / (G M_0)^2 / t_{\text{eq}})^{1/3} f^{-2/3}$ or using the upper limits on M_0 one finds $m_{\text{min}} c^2 \gg 10^{-14} \text{ eV } \delta_\tau^{1/6} f_{-7}^{-2}$. This can be an interesting constraint in models of ultralight DM, such as axion-like particles or vector DM.

Thus the early halo model considered in this paper can only accommodate ratios of late matter to early matter in the range $10^{-28} \ll f \lesssim 10^{-7}$ producing remnant halo structures with mass $\ll 10^{-3} M_\odot$. This gives the characteristic mass scale of the smallest structures which will recollapse at late times (microhalos). Lower bounds on f and the microhalo mass scale are more uncertain, depending on assumptions for very early cosmology at the inflation or Planck scale and on the particle nature and phase space distributions of EM and LM.

5 Superposition of Halo Remnants

As halos remnants expand in comoving size they will begin to overlap. Overlapping spherical halos are necessarily aspherical, so the spherical halo model used so far cannot accommodate the gravitational

field of neighboring remnants. Fortunately this additional force is probably not important as we argue below that overlap occurs when $\bar{\Delta} \lesssim 1$ and we see from Fig. 5, that $\bar{\Delta} \gg 1$ at the start of free expansion. During free expansion the self-gravity of the matter becomes unimportant so when neighboring halo remnants do overlap there is no significant interaction between them. As a result, the matter distribution is simply the linear superposition of the spherical matter distribution of individual remnants:

$$\rho(\mathbf{x}) = \bar{\rho}_{\text{mat}} \sum_h \Delta_h(|\mathbf{x} - \mathbf{x}_h|) \quad (5.1)$$

where h labels individual remnants, $\mathbf{x}_h$ is the halo center, and $\Delta_h(r)$ gives its density profile. We first use the results from previous sections to show that the halo remnants overlap in §5.1. We then apply this (early) halo model description of the density field to express the matter power spectrum at late times in terms of the density profiles of the remnants in §5.2.

5.1 Remnant Overlap

One can quantify the amount of overlap of remnant h by the ratio³

$$\mathcal{N}_h \equiv \frac{\int d^3\mathbf{x} \Delta_h(|\mathbf{x} - \mathbf{x}_h|) \sum_{h' \neq h} \Delta_{h'}(|\mathbf{x} - \mathbf{x}_{h'}|)}{\int d^3\mathbf{x} \Delta_h(|\mathbf{x} - \mathbf{x}_h|)^2}, \quad (5.2)$$

where the denominator is the density weighted average of $\Delta_h(r)$ of remnant h and the numerator the same weighted average of the sum of $\Delta_{h'}$'s for all *other* halos. Roughly speaking the “overlap number”, $\mathcal{N}_h$, is the number of other remnants which overlap remnant h : iff there is no overlap then $\mathcal{N}_h = 0$ while $\mathcal{N}_h \gg 1$ indicates a great deal of overlap. One can place lower limits on the amount of overlap from the values Δ in halos though specific values depend on the clustering of the initial EHS: if EHS tend to be near to each other then they overlap more readily.

Let us consider the matter overdensity

$$\delta_{\text{m}}(\mathbf{x}) \equiv \frac{\rho(\mathbf{x})}{\bar{\rho}_{\text{mat}}} - 1 = -1 + \sum_h \Delta_h(|\mathbf{x} - \mathbf{x}_h|). \quad (5.3)$$

Denote spatial averages by $\langle f(\mathbf{x}) \rangle_{\mathbf{x}} \equiv V^{-1} \int d^3\mathbf{x} f(\mathbf{x})$ where the (usually infinite) cosmic volume is $V \equiv \int d^3\mathbf{x}$. Since the halos are assumed to contain all the mass it follows that $\langle \sum_h \Delta_h(|\mathbf{x} - \mathbf{x}_h|) \rangle_{\mathbf{x}} = 1$ or $\langle \delta_{\text{m}}(\mathbf{x}) \rangle_{\mathbf{x}} = 0$. One can quantify the amount of overlap by the decomposition

$$\begin{aligned} 1 + \langle \delta_{\text{m}}(\mathbf{x})^2 \rangle_{\mathbf{x}} &= \left\langle \sum_h \sum_{h'} \Delta_h(|\mathbf{x} - \mathbf{x}_h|) \Delta_{h'}(|\mathbf{x} - \mathbf{x}_{h'}|) \right\rangle_{\mathbf{x}} \\ &= \left\langle \sum_h \Delta_h(|\mathbf{x} - \mathbf{x}_h|)^2 \right\rangle_{\mathbf{x}} + \left\langle \sum_{h \neq h'} \Delta_h(|\mathbf{x} - \mathbf{x}_h|) \Delta_{h'}(|\mathbf{x} - \mathbf{x}_{h'}|) \right\rangle_{\mathbf{x}} \\ &= \sum_h (1 + \mathcal{N}_h) \langle \Delta_h(|\mathbf{x} - \mathbf{x}_h|)^2 \rangle_{\mathbf{x}}. \end{aligned} \quad (5.4)$$

For each remnant define the density weighted average Δ by

$$\langle \Delta \rangle_h \equiv \frac{\int d^3\mathbf{x} \Delta_h(|\mathbf{x} - \mathbf{x}_h|)^2}{\int d^3\mathbf{x} \Delta_h(|\mathbf{x} - \mathbf{x}_h|)} = \frac{\bar{\rho}_{\text{mat}}}{M_h} \int d^3\mathbf{x} \Delta_h(|\mathbf{x} - \mathbf{x}_h|)^2. \quad (5.5)$$

where $M_h = \bar{\rho}_{\text{mat}} \int d^3\mathbf{x} \Delta_h(|\mathbf{x} - \mathbf{x}_h|)$ is the mass of the remnant. Thus

$$1 + \langle \delta_{\text{m}}(\mathbf{x})^2 \rangle_{\mathbf{x}} = \sum_h \frac{M_h}{\bar{\rho}_{\text{mat}} V} (1 + \mathcal{N}_h) \langle \Delta \rangle_h = \langle (1 + \mathcal{N}_h) \langle \Delta \rangle_h \rangle_{\text{rem}}. \quad (5.6)$$

³The formulae of this section generalize to non-spherical remnants by $|\mathbf{x} - \mathbf{x}_h| \rightarrow \mathbf{x} - \mathbf{x}_h$.

where $\langle f_h \rangle_{\text{rem}} \equiv \sum_h \frac{M_h}{\bar{\rho}_{\text{mat}} V} f_h$ is a mass weighted average over halos since $\sum_h M_h = \bar{\rho}_{\text{mat}} V$. Defining a mean halo density ratio and a mean overlap number

$$\langle \Delta \rangle \equiv \langle \langle \Delta \rangle_h \rangle_{\text{rem}} \quad \langle \mathcal{N} \rangle \equiv \frac{\langle \langle \Delta \rangle_h \mathcal{N}_h \rangle_{\text{rem}}}{\langle \langle \Delta \rangle_h \rangle_{\text{rem}}} . \quad (5.7)$$

leads to the simple expression

$$\langle \delta_{\text{m}}(\mathbf{x})^2 \rangle_{\mathbf{x}} = (\langle \mathcal{N} \rangle + 1) \langle \Delta \rangle - 1 . \quad (5.8)$$

Since $\delta_{\text{m}}^2 \geq 0$ it follows that

$$\langle \mathcal{N} \rangle \geq \frac{1}{\langle \Delta \rangle} - 1 . \quad (5.9)$$

If $\langle \Delta \rangle \geq 1$ this is not a new constraint since $\langle \mathcal{N} \rangle \geq 0$ by construction. If $\langle \Delta \rangle \ll 1$ we see that $\langle \mathcal{N} \rangle \gg 1$ and there is a great deal of overlap.

The bound of Eq. 5.9 depends only on the density profiles and distribution of masses, M_h , of remnants and not how they are spatially distributed (clustering). The actual value of $\langle \mathcal{N} \rangle$ does depend on clustering. $\langle \mathcal{N} \rangle$ could be much larger than the bound of Eq. 5.9 if the remnant centers are strongly clustered. Our expectation is that clustering is modest so a baseline value is given by assuming the EH centers are uncorrelated in position in which case⁴

$$\langle \delta_{\text{m}}(\mathbf{x})^2 \rangle_{\mathbf{x}} \approx \langle \Delta \rangle \quad \langle \mathcal{N} \rangle \approx \frac{1}{\langle \Delta \rangle} , \quad (5.10)$$

which follows from Eqs. 5.4 and 5.8. This is only slightly larger than the bound if $\langle \Delta \rangle \ll 1$. Modest clustering would lead to similarly small fractional changes in the value of $\langle \mathcal{N} \rangle$.

One finds from Fig. 5 that $\Delta < \bar{\Delta} \lesssim 10^{-3}$ when $a \approx a_{\text{eq}}$ unless $\bar{\Delta}_\tau \gg 10^{10}$. Hence we expect $\langle \Delta \rangle \lesssim 10^{-3}$ and $\langle \mathcal{N} \rangle \gtrsim 10^3$ at the time of matter-radiation equality, i.e. a great deal of overlap.

5.2 One- and Two-Halo Power Spectra

The linear superposition of EH remnants in Eq. 5.1 is identical to the starting point of the halo model of DM density field as described in, e.g., Refs. [50, 51]. We review aspects of this formalism in Appendix D. We are ultimately interested in estimating the power spectrum at late times in order to infer structure formation history after matter-radiation equality. In the halo model, the power spectrum decomposes into one- and two-halo terms as

$$P(k) = P_{1\text{h}}(k) + P_{2\text{h}}(k), \quad (5.11)$$

where

$$P_{1\text{h}}(k) = \frac{1}{\bar{\rho}_{\text{LM}} a^3} \int dm \frac{df}{d \ln m} |\mathcal{F}(k|m)|^2 \quad (5.12\text{a})$$

$$P_{2\text{h}}(k) = \left| \int d \ln m \frac{df}{d \ln m} b(m) \mathcal{F}(k|m) \right|^2 \tilde{P}_{hh}(k), \quad (5.12\text{b})$$

where

$$\frac{dn}{d \ln m} = \frac{\bar{\rho}_{\text{LM}} a^3}{m} \frac{df}{d \ln m} \quad (5.13)$$

is the halo mass function; $\tilde{P}_{hh}(k)$ is a halo-halo power spectrum; $b(m)$ is a bias function that relates halo and mass power spectra⁵; $\mathcal{F}$ is the mass-normalized Fourier transform of the density profile in

⁴In the limit $V \rightarrow \infty$ assuming the rms M_h is finite.

⁵We will set $b \approx 1$ for simplicity but in principle the bias can be estimated analytically or from N -body simulations [51].

comoving coordinates

$$\mathcal{F}(k, t|m) = \frac{4\pi a^3}{m} \int_0^\infty dr_{\text{co}} r_{\text{co}}^2 j_0(kr_{\text{co}}) \rho_{\text{EH}}(r_{\text{co}}, t|m), \quad (5.14)$$

where $\rho_{\text{EH}} = \bar{\rho}_{\text{mat}} \Delta_h$ for some h (assuming all EHs have similar profiles and just differ in mass m). Note that $\mathcal{F} \rightarrow 1$ as $k \rightarrow 0$, which ensures that $P(k)$ reduces to the linear power spectrum on large enough scales (see below). The expressions in Eq. 5.12 are remarkable in that they nearly factorize effects from linear and non-linear scales; the integrals over the halo mass distributions depend on the halo profiles, while the halo-halo power spectrum is closely related to the linear power spectrum. The tilde in $\tilde{P}_{hh}(k)$, however, indicates that we must model the fact that on small scales prior to reheating halos did not overlap. This exclusion effect ensures that $\tilde{P}_{hh}(k) \approx P^{\text{lin}}(k)$ for $kr_e \ll 1$, and $\tilde{P}_{hh}(k) \approx C$ for $kr_e \gg 1$ with r_e the typical exclusion radius; the constant C is fixed such that the probability to find two halos with separation $r \lesssim r_e$ vanishes before EHs explosively evaporate. This prescription is described in detail in Appendix D.2.

In order to make use of Eq. 5.12 we will need to compute $\mathcal{F}$, P^{lin} and $df/d\ln m$ in specific cosmological models. The normalized density profile $\mathcal{F}$ can be calculated from the remnant profile derived in the previous section, Eq. 4.21, together with the definition in Eq. 5.14. The linear power spectrum P^{lin} and mass distribution $df/d\ln m$ can be obtained from linear perturbation theory and the Press-Schechter formalism which we discuss in the following section.

6 Linear Evolution of Density Perturbations

As shown in §4 and §5 even extremely overdense EHs disassociate and their remnants overlap with neighboring remnants producing an LDM distribution with rms overdensity $\langle \delta_{\text{m}}(\mathbf{x})^2 \rangle_{\mathbf{x}} \lesssim 10^{-3}$ (discussion following Eq. 5.10) by the beginning of LMD. Thus a very inhomogeneous (non-linear) matter distribution evolves into nearly homogeneous (linear) distribution. In this section we study the linear evolution of DM density perturbations both during EMD and after reheating. These results along with the Press-Schechter (PS) formalism [43] will enable us to estimate the properties of bound object of both early halos that eventually form during EMD and microhalos that form from density fluctuations that survive reheating or result from the overlap of EH remnants.

6.1 Linear Evolution and Collapse Before Reheating

The fluid equations describing the evolution of perturbations in LM, EM and radiation are given in, e.g., Refs. [33–35] and in Appendix C. During EMD these equations can be solved for the LM density contrast δ [33]

$$\delta = \delta_i - \frac{2}{3} \left(\frac{k}{k_{\text{RH}}} \right)^2 \left(\frac{a}{a_{\text{RH}}} \right) \psi, \quad (6.1)$$

where ψ and δ_i are the primordial (super-horizon) values of the gravitational potential and of the EM and LM fluctuations, respectively; if these were born as adiabatic fluctuations then $\delta_i \propto \psi$. This result is only valid for modes that enter the horizon during EMD; if a mode enters during an early period of RD (before the universe transitions to EMD) the evolution is more complicated. Numerical solutions exemplifying these different regimes are shown in Fig. 2 for a mode with $k/k_{\text{RH}} = 600$, where $k_{\text{RH}} = a_{\text{RH}} H(a_{\text{RH}})$ is the value of the conformal Hubble parameter at reheating. Equation 6.1 corresponds to the upper line of Fig. 2. For the background cosmologies in the lower lines, this mode enters during radiation domination, before transitioning to EMD; the impact of the gravitational driving effect, a hallmark of radiation domination [47], is evident as these modes enter the horizon. Note that in Fig. 2 modes with the same physical size at late times (after reheating), enter the horizon at different times and experience different amounts of growth. In the cosmologies with only a brief

period of EMD, the total energy density is greater at earlier times, implying that horizon entry of a mode with the same k/k_{RH} occurs later in scale factor evolution. As a result, these modes have a smaller amplitude at reheating, compared to models with a long period of EMD. Semi-analytic solutions for the density contrast evolution can be constructed in models with a brief period of EMD, e.g., by following the methods of Ref. [47]. For simplicity, we only present analytical expression for a period of long EMD; other models are treated numerically.

After a long period of EMD the initial condition in Eq. 6.1 becomes irrelevant and we have⁶

$$\langle \delta^2 \rangle \approx \left[\frac{2}{3} \left(\frac{k}{k_{\text{RH}}} \right)^2 \left(\frac{a}{a_{\text{RH}}} \right) \right]^2 \langle \psi^2 \rangle \quad (a_{\text{hor}} \ll a < a_{\text{RH}}), \quad (6.2)$$

which is valid well after the horizon entry of mode k but before the end of EMD. The autocorrelation function of the gravitational potential is related to the curvature power spectrum as

$$\langle \psi^2 \rangle = \left(\frac{3}{5} \right)^2 \times \frac{2\pi^2}{k^3} A_s \left(\frac{k}{k_0} \right)^{n_s-1}, \quad (6.3)$$

where the factor $3/5$ arises from conservation of the curvature perturbation for adiabatic fluctuations. While we do not know the spectrum of primordial curvature fluctuations on the small scales of interest, we will for simplicity take $A_s = 2.1 \times 10^{-9}$, $n_s = 0.965$ with $k_0 = 0.05/\text{Mpc}$ [2]. Note that this is an enormous extrapolation of the simple powerlaw ansatz from scales probed by the CMB, $\sim k_{\text{eq}}$, to $k > 10^7 k_{\text{eq}}$.

The above results can be used to estimate various properties of collapsed objects that form *during* EMD. First, it is useful to estimate the scale of the largest structures that can collapse during EMD. We can compute the mass and size of these non-linear structures using the Press-Schechter formalism by solving

$$\sigma(M_*, z_c) = \delta_c, \quad (6.4)$$

for the mass M_* of one-sigma overdensities of the density field collapsing at z_c ; in the above equation $\delta_c = 1.686$ is the collapse threshold and σ^2 is the density fluctuation variance

$$\sigma^2(M(R), z) = \int \frac{d^3k}{(2\pi)^3} \langle \delta^2 \rangle W(kR)^2, \quad (6.5)$$

where W is a window function. The density variance for the “Long EMD” and “Short EMD” cosmologies is shown in the left panel of Fig. 9 for $T_{\text{RH}} = 5 \text{ MeV}$. Since this quantity is computed in linear perturbation theory, it does not include the effects of adiabatic and free expansion of collapsed objects near the end of EMD; however, we can use it to estimate the properties of the largest structures formed just before/at reheating, i.e., those with a collapse redshift $z_c = z_{\text{RH}} \sim 10^{10}(T_{\text{RH}}/5 \text{ MeV})$. The mass scale of these structures can be estimated from Eq. 6.4 and corresponds to the intersection σ^2 with the gray dashed line equal to the collapse threshold δ_c^2 in the left panel of Fig. 9; numerically we find

$$\frac{M_*(z_{\text{RH}})}{M_{\text{RH}}} \sim \begin{cases} 8 \times 10^{-8} & \text{Long EMD} \\ 5 \times 10^{-12} & \text{Short EMD}, \end{cases} \quad (6.6)$$

where M_{RH} is the mass enclosed inside the horizon at reheating:

$$M_{\text{RH}} = 250 M_{\oplus} \left(\frac{5 \text{ MeV}}{T_{\text{RH}}} \right)^3 \left(\frac{10.75}{g_*(T_{\text{RH}})} \right)^{1/2}, \quad (6.7)$$

where $M_{\oplus} = 3 \times 10^{-6} M_{\odot}$ is the Earth mass. The tilde in the relation of Eq. 6.6 indicates that this is a window-function dependent quantity and therefore should be interpreted only as a order-of-magnitude characteristic mass of EHs forming near the end of EMD.

⁶We use $\langle \cdot \rangle$ as a shorthand power spectrum, or equivalently, the autocorrelation function with $(2\pi)^3 \delta^3(0)$ factored out.

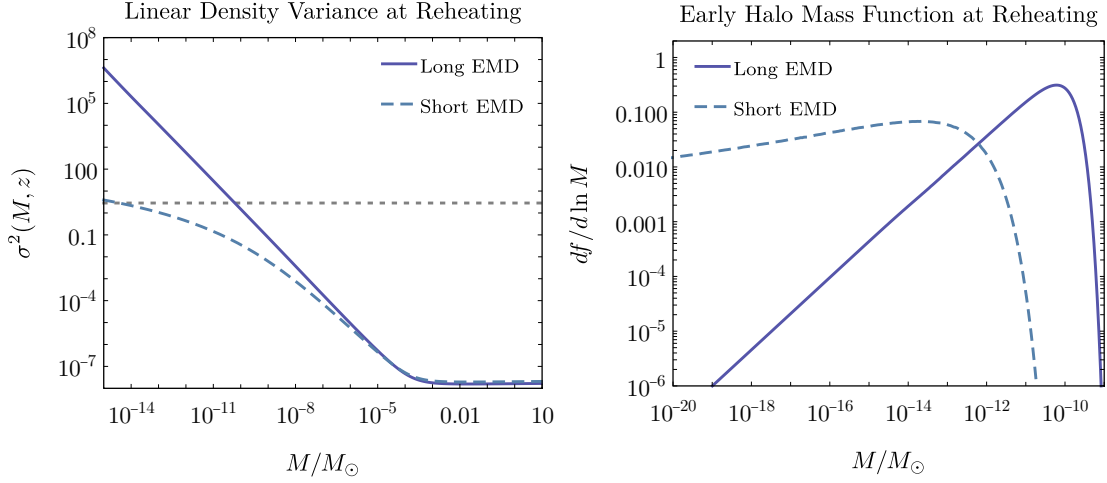


Figure 9. Linear theory density variance (left) and early halo mass function (right) computed using the PS formalism in Eq. 6.8 for the “Long EMD” and “Short EMD” cosmologies (solid and dashed lines in each panel, respectively). These quantities are evaluated at reheating (i.e., when the energy densities of EM and radiation are equal) and approximately describe the density field prior to adiabatic and free expansion of EHs, since these effects are not included in the linear theory. In the left panel, the gray dotted line indicates the EMD spherical collapse threshold of $\delta_c^2 \approx 1.686^2$.

We see that a shorter duration of EMD leads to much smaller EHs, since only the smallest scales have experienced enough growth to collapse. Using the Press-Schechter formalism [43] we can estimate the early halo distribution function from the density variance Eq. 6.5:

$$\frac{df}{d \ln m} = \sqrt{\frac{2}{\pi}} \frac{\delta_c}{\sigma} \left| \frac{d \ln \sigma}{d \ln m} \right| \exp \left(-\frac{\delta_c^2}{2\sigma^2} \right). \quad (6.8)$$

This distribution, along with the linear power spectrum above and individual halo profiles are needed to evaluate the matter power spectrum *after* reheating in Eq. 5.12. We show the distribution of EHs in the right panel of Fig. 9 for the “Long EMD” and “Short EMD” cosmologies. Note that in the latter example (dashed line) the distribution peaks at lower masses and has an extremely long tail as $M \rightarrow 0$; this is because modes that entered the horizon prior to the start of EMD grew only logarithmically with a for $a < a_{\text{eeq}}$, resulting in a nearly scale-invariant fluctuation amplitude on as can be seen in the left panel of Fig. 9.

To conclude the section we return to the evolution of linear perturbations described by Eq. 6.2. After the transition to RD at $a = a_{\text{RH}}$, linear perturbations continue to grow, albeit logarithmically with the scale factor. Matching the EMD and RD solutions at a_{RH} determines the power spectrum well after RH [33]

$$\langle \delta^2 \rangle \approx \left[\frac{2}{3} \left(\frac{k}{k_{\text{RH}}} \right)^2 \ln \left(\frac{ea}{a_{\text{RH}}} \right) \right]^2 \langle \psi^2 \rangle \quad (a_{\text{RH}} \ll a < a_{\text{eq}}). \quad (6.9)$$

This expression is the linear power spectrum of LM during radiation domination, P^{lin} . More explicitly, using Eq. 6.3, gives

$$\frac{k^3}{2\pi^2} P^{\text{lin}} = \left[\frac{2}{5} \left(\frac{k}{k_{\text{RH}}} \right)^2 \ln \left(\frac{ea}{a_{\text{RH}}} \right) \right]^2 A_s \left(\frac{k}{k_0} \right)^{n_s-1}. \quad (6.10)$$

In models with a finite duration of EMD, we evaluate $k^3 P / (2\pi^2)$ by using the numerical solution of $\langle \delta^2 \rangle$. Illustrative examples of these power spectra are presented in §7.

6.2 Linear Evolution and Collapse After Reheating

In this section we briefly describe how the power spectrum after reheating evolves through matter-radiation equality and beyond. The power spectrum in Eq. 6.10 applies only to the “Long EMD” cosmology, and it is valid only for modes with $k > k_{\text{RH}}$ during radiation domination. We therefore need a more general prescription to study the density field evolution for a wide range of scales and in different cosmologies. This evolution is captured by the Meszaros equation [46, 47], whose general solution is

$$\delta = [A_1 U_1(y) + A_2 U_2(y)], \quad (6.11)$$

where $U_{1,2}(y)$ are functions of $y = a/a_{\text{eq}}$ (given in Ref. [47]), and $A_{1,2}$ are coefficients that are obtained by matching Eq. 6.11 to numerics during radiation domination as described in Ref. [52]. The RD evolution can be written as (cf. Eq. 6.9)

$$\delta = I_1 \psi \ln \left(\frac{I_2 a}{a_{\text{hor}}} \right) \quad (a_{\text{RH}} \ll a < a_{\text{eq}}), \quad (6.12)$$

where $a_{\text{hor}}(k)$ is the scale factor when the mode k enters the horizon and $I_{1,2}$ are (typically) k -dependent coefficients that encode the pre-reheating evolution of small scales; $I_{1,2}$ are obtained by numerically solving for δ into RD and fitting the solution to Eq. 6.12. Matching Eq. 6.11 at $y \ll 1$ to Eq. 6.12, yields expressions for $A_{1,2}$ in terms of $I_{1,2}$:

$$A_1 \approx \frac{3}{2} \left[\ln \left(4 I_2 e^{-3 \frac{a_{\text{eq}}}{a_{\text{hor}}}} \right) \right] I_1 \psi. \quad (6.13)$$

$$A_2 \approx -\frac{4}{15} I_1 \psi. \quad (6.14)$$

For the modes of interest $k \gg k_{\text{eq}}$, $A_1 \gg A_2$, and for $a \gtrsim a_{\text{eq}}$ Eq. 6.11 simplifies to

$$\delta \approx A_1 U_1(y), \quad (6.15)$$

This is a useful result because the cosmology-dependent quantity A_1 factorizes from the scale factor evolution U_1 , meaning that we can write the linear power spectrum in *any* cosmology after MRE as

$$P^{\text{lin}}(k, a) = P_{\Lambda\text{CDM}}^{\text{lin}}(k, a) \left(\frac{\delta}{\delta_{\Lambda\text{CDM}}} \right)^2, \quad (6.16)$$

where $P_{\Lambda\text{CDM}}^{\text{lin}}$ is the ΛCDM power spectrum which can be computed using a Boltzmann code [53, 54], or using the transfer and growth functions of Ref. [55] (we use a combination of these methods as described in Appendix C of Ref. [39]). The density contrast ratio $(\delta/\delta_{\Lambda\text{CDM}})^2$ is independent of scale factor and therefore can be interpreted as a modification of the primordial power spectrum – it is an explicit realization of the “bump” functions considered in Ref. [52].

In summary, we have reduced the calculation of the linear power spectrum to the determination of $I_{1,2}$ which we obtain by numerically solving linear perturbation equations. In order to enable the evaluation of $P^{\text{lin}}(k, a)$ across a wide range of scales, we construct fitting functions for $I_{1,2}$ which have the correct asymptotics for $k \ll k_{\text{RH}}$ and $k \gg k_{\text{RH}}$. For example, in the “Long EMD” model $I_1 \approx 2(k/k_{\text{RH}})^2/3$ and $I_2 \approx (k_{\text{RH}}/k)^2$ for $k \gg k_{\text{RH}}$; modes that enter the horizon during RD ($k \ll k_{\text{RH}}$) modes that enter the horizon during RD have $I_1 \approx 9.11$ and $I_2 \approx 0.594$ [47]. The details of the numerical solutions and the fitting functions are presented in Appendix C. The resulting power spectra at matter-radiation equality are shown in Fig. 10 for the “Long EMD” and “Short EMD” cosmologies as dashed lines. In the following section, we will include the non-linear effects of EH explosive evaporation at small scales.

7 Power Spectrum and Formation of Microhalos

In this section we combine the results of Sections 4, 5 and 6 to compute the matter power spectra at small scales, and study early structure formation.

In Fig. 10 we show two benchmark DM power spectra evaluated at matter-radiation equality, corresponding to the “Long EMD” and “Short EMD” background cosmologies in Fig. 1 and described in §2. In each panel the gray line labeled Λ CDM indicates the expected spectrum in the absence of EMD based on the scale invariant primordial spectrum in Eq. 6.3; the slow growth with k is a consequence of logarithmic growth of perturbations during radiation domination. The dashed lines indicate the linear power spectrum that does *not* account for formation and eventual disruption of EHs during reheating. These power spectra monotonically increase at large k unless there is some small-scale cutoff due to the nature of the DM particle. These spectra first diverge from the Λ CDM expectation for wavenumbers comparable to the comoving horizon size at reheating [32, 33]:

$$\frac{k_{\text{RH}}}{k_{\text{eq}}} \approx 5.9 \times 10^6 \left(\frac{T_{\text{RH}}}{5 \text{ MeV}} \right) \left(\frac{g_*(T_{\text{RH}})}{10.75} \right)^{1/6}. \quad (7.1)$$

In the left panel, the EMD enhancement scales as $(k/k_{\text{RH}})^4$ as expected from the analytic arguments around Eq. 6.2. In contrast, in the “Short EMD” example in the right panel, the enhancement becomes logarithmic in $k \gtrsim k_{\text{eeq}} \sim \sqrt{a_{\text{RH}}/a_{\text{eeq}}} k_{\text{RH}}$ for modes that began their evolution before EMD in the early radiation era. The solid blue lines in both panels incorporate the effects of EH explosive evaporation as described in §5.2. We see that the impact of reheating on small scale power is significant: the maximum amplitude of the power spectrum can be significantly reduced compared to the naive linear theory expectation indicated by the dashed lines. The scale at which the full result begins to deviate from linear theory corresponds precisely to the size of EHs after they go through free expansion; this scale can be estimated by applying Eq. 4.14 to the virial radius of the largest EHs that have formed by the end of EMD; for EHs that formed at $a = a_{\text{RH}}$ this virial radius is

$$r_{\text{vir}} \approx a_{\text{RH}} \left(\frac{3M_{\text{LM}}}{800\pi\rho_{\text{cdm}}} \right)^{1/3} = 9 \times 10^{-4} \left(\frac{a_{\text{RH}}}{k_{\text{RH}}} \right) \left(\frac{M_{\text{LM}}/M_{\text{RH}}}{10^{-7}} \right)^{1/3}, \quad (7.2)$$

where ρ_{cdm} is the present DM density and M_{LM} is the LM mass in the typical EHs. These typical masses are given in Eq. 6.6. Using Eq. 4.14 we find that the *comoving* size of the remnant is approximately

$$r_{\text{vir}}^{\text{co}} \sim 200 \left(1 + 0.1 \ln \frac{a/a_{\text{RH}}}{10^8} \right) r_{\text{vir}}/a_{\text{RH}}. \quad (7.3)$$

This radius maps onto a characteristic wavenumber beyond which the Fourier transform of density profiles, Eq. 5.14, begins to fall off and suppress the power spectra (see Eq. 5.12). $r_{\text{vir}}^{\text{co}}$ is much larger than the comoving virial radius of the original halos due to the effects of adiabatic and free expansion. As a result we expect an imprint on the power spectrum at scales larger than the typical size of the early halos.

Our numerically-derived profiles in §4.6 do not have an analytic form if they originate from NFW or Plummer EHs. We note however, that the adiabatic and free expansion tend to produce nearly-flat interiors and a sharp cutoff at $r_{\text{vir}}^{\text{co}}$ (see, e.g., Fig. 7); as a result, the profiles resemble those of a spherical top hat, whose Fourier transform is just $3j_1(kr_{\text{vir}}^{\text{co}})/(kr_{\text{vir}}^{\text{co}})$. Using this approximate correspondence we see that $\mathcal{F}$ should significantly deviate from 1 for $k \gtrsim \sqrt{10}/r_{\text{vir}}^{\text{co}}$; combining Eqs. 7.3 and 7.1 this gives an estimate of the cut-off scale⁷

$$k_c/k_{\text{eq}} \sim \sqrt{10}(k_{\text{eq}}r_{\text{vir}}^{\text{co}})^{-1} \approx 20 \left(\frac{k_{\text{RH}}}{k_{\text{eq}}} \right) \left(\frac{10^{-7}}{M_{\text{LM}}/M_{\text{RH}}} \right)^{1/3} = 10^8 \left(\frac{T_{\text{RH}}}{5 \text{ MeV}} \right) \left(\frac{10^{-7}}{M_{\text{LM}}/M_{\text{RH}}} \right)^{1/3}, \quad (7.4)$$

⁷We dropped the logarithmic term from Eq. 7.3 and the g_* dependence in Eq. 7.1 for simplicity.

which is in excellent agreement with the two examples in Fig. 10 which use the NFW profile for EHs. The comparison with a top hat profile is also useful in explaining the wavenumber scaling for $k \gg 1/r_{\text{vir}}^{\text{co}}$: we find that the Fourier transform of exploded NFW profiles is also similar to the Fourier transform of a spherical top hat, so $\mathcal{F} \sim 1/k^2$ at large k . This means that power spectra are suppressed by $|\mathcal{F}|^2 \sim 1/k^4$ for $k \gg (r_{\text{vir}}^{\text{co}})^{-1}$. This suppression particularly evident in the left panel of Fig. 10 where this scaling nearly flattens k^4 growth of the power spectrum in the “Long EMD” cosmology. The k scaling for the “Short EMD” more difficult to derive because many different EH masses contribute to Eq. 5.12.

Equation 7.4 provides an approximate wavenumber beyond which modes are not enhanced by EMD. This model-independent cut-off allows for a very limited range of k for the power spectrum to grow. It is also interesting that because this cutoff depends on the mass of the largest EHs, cosmologies where this mass is lower have boosted power over a broader range of scales. This somewhat counter-intuitive point is illustrated in the right panel of Fig. 10, where we see that the “Short EMD” cosmology has more power at smaller scales than the “Long EMD” case, precisely because the EHs were much smaller than in “Long EMD”. Also, because the power is relatively flat at small scales in the “Short EMD” example (corresponding to the modes that entered the horizon in early radiation domination), the distribution of EHs has a long tail at small masses (see the right panel of Fig. 9), ensuring that the expulsion of matter from EHs during adiabatic and free expansion has little effect on larger scales. Thus, the two features of “Short EMD” – smaller characteristic mass of EHs and their relatively flat distribution – ensure that it has significantly more power at small scales than “Long EMD”.

The power spectra at matter-radiation equality shown in Fig. 10 can be evolved into matter domination as described in §6.2. We use this evolution to compute the density variance from Eq. 6.5, which enables us to estimate the formation history of microhalos of LM as a function of redshift via Eq. 6.4. The typical microhalo mass forming at z , shown in Fig. 11, highlights the physical importance of the explosive evaporation of EHs on late-time structure formation. The expulsion of matter from EHs at reheating erases small-scale power; in the “Long EMD” cosmology (solid blue line) this leads to a significant delay of the onset of microhalo formation compared to the naive expectation shown by the dashed blue line. As we discussed above, for “Short EMD” these non-linear effects are less severe on the scales of interest since the EHs were much smaller.

In §5.2 we decomposed the power spectrum into one- and two-halo terms. The one-halo term measures the correlations within the EH remnants, whose evolution is encoded in the time-dependence of the remnant density profile. Since the DM particles in EH remnants are travelling outwards as a result of explosive evaporation, their correlations do not grow as described in §6.2. This is explicitly illustrated by the numerical solutions in Fig. 4. In contrast, the halo-halo power spectrum, which is related to the two-halo term via Eq. 5.12b, continues to grow since it is proportional to the linear spectrum (see Eq. D.9). The physical interpretation here is that EH centers of mass continue to drift toward each other during RD even after the EHs have explosively evaporated. Well after reheating, the PS therefore becomes dominated by the two-halo terms.

The halo-halo power spectrum, however, must deviate from the linear matter power spectrum at small scales because of halo exclusion, as mentioned in §5.2: prior to evaporation EHs did not overlap by definition. A realistic implementation of exclusion effects is beyond the scope of the current work, so we opt for a simple prescription that is illustrated on a toy model in Appendix D.2: we set the PS $P(k)$ to a constant for $k > 1/r_e$, where r_e is the exclusion radius. r_e is in principle different for different EHs, but we roughly estimate it as the typical separation of the EHs forming at the end of EMD (it is of the same order as the virial radius of EHs forming at this time) – for more sophisticated models see, e.g., Ref. [56]. The scales affected by halo exclusion effects are evident in both panels of Fig. 10 at large k as a sharp kink. In the absence of the density profile factor in the definition of P_{2h} , Eq. 5.12b, dimensionless power spectra for both cosmologies would scale as k^3 . However, because the convolution of the EH mass function and the density profile $\mathcal{F}$ is different in the “Long EMD” and “Short EMD” examples, the behaviour of the PS also differs. For “Long EMD” the mass function

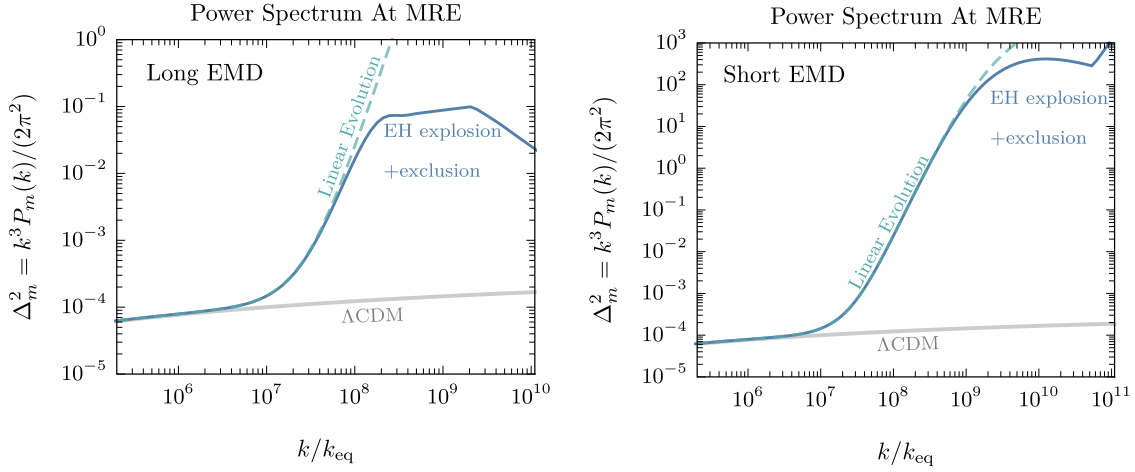


Figure 10. Matter power spectrum for a cosmology with $T_{\text{RH}} = 5$ MeV evaluated at matter-radiation equality. In the left panel we show a variant where EMD lasted for a long time and all modes with $k > k_{\text{RH}}$ entered the horizon during EMD. In the right panel EMD has a finite duration, and some modes entered the horizon during a period of early radiation domination which was then followed by EMD. In both panels the dashed lines are the linear power spectra without taking into account the disruption of halos that formed during EMD, while the solid line includes the effects of explosive evaporation of early halos at the end of EMD, which suppresses power at small scales. The gray band is the Λ CDM prediction. In the right panel the $\Delta_m^2 \gg 1$ region is shown for comparison with “Long EMD” only; linear perturbation theory used to evolve the PS to this point breaks down in this regime.

integral gives a $1/k^4$ suppression, so $k^3 P_{2h}(k) \sim 1/k$ decreases at large k ; in “Short EMD” the mass function integral is instead logarithmically growing (owing to the very flat mass function shown in the right panel of Fig. 9), so $k^3 P_{2h}(k)$ actually increases. It would be interesting to check whether this effect persists in other models of exclusion or in N -body simulations. However, it is likely irrelevant for the detectability of microhalos that form after equality since exclusion affects tiny length scales buried inside of microhalos.

In some cosmologies the linear theory logarithmic growth of density perturbations leads to overdensities $\gg 1$ at matter-radiation equality, as shown in the right panel of Fig. 10 at large k . We only show this for comparison with “Long EMD”, as linear perturbation theory is not valid in this regime; we do not use this highly-non-linear part of the power spectrum in Fig. 11. In reality, as the EH centers drift in the RD universe, they can pass through each other, leading to either a saturation or even a decrease of the halo-halo PS. Large matter overdensities can collapse into microhalos during RD if matter-radiation equality is attained locally. Since our formalism does not account for these effects, we focus on microhalo formation for $z < z_{\text{eq}}$ in Fig. 11 which is sensitive to $k^3 P(k)/(2\pi^2) \lesssim \mathcal{O}(1)$ in Fig. 10. We leave the study of these issues to future work.

7.1 Comparison with Other Cutoffs

The small-scale cut-off derived in the previous section is model-independent in the sense that it applies to any microphysical model of DM that is present during a period of EMD. In specific scenarios, however, other effects can lead to a more important suppression of power. For example, if DM is a thermal particle (that is, it is produced through contact with a plasma of relativistic particles), it can remain in kinetic equilibrium with the radiation bath long after chemical decoupling (freeze-out). This coupling to radiation provides pressure support to DM perturbations and prevents structure growth on scales that entered the horizon before kinetic decoupling. In fact, there are two different effects induced by DM-radiation coupling: free-streaming and acoustic oscillations, both of which suppress

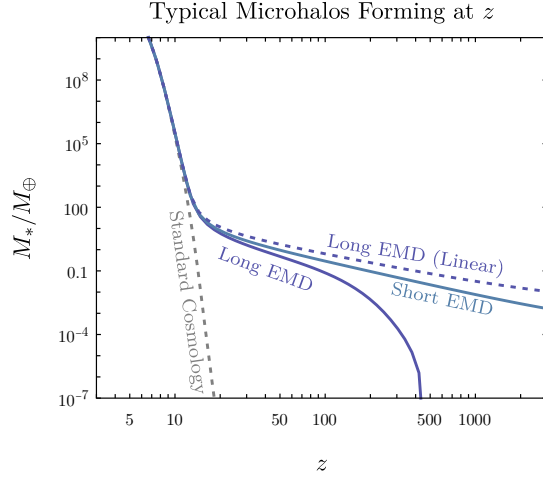


Figure 11. Mass of typical microhalos (normalized to Earth mass $M_\oplus$) forming at redshift z in cosmologies with Early Matter Domination. The solid lines labelled “Long EMD” and “Short EMD” correspond to background cosmologies illustrated in Fig. 1 and include the effects of explosive evaporation of EHs that formed during EMD. These effects wash out structure on small scales, resulting in smaller masses and later formation of microhalos compared to the naive expectation (blue dashed line) without accounting for the non-linear effects due to EHs.

power at small scales [57, 58]. The small scale cut-off is the smallest of the wavevectors

$$k_{\text{cut}} \approx \min(k_{\text{fs}}, k_{\text{ao}}), \quad (7.5)$$

where the free-streaming (fs) and acoustic oscillation scales are (ao)

$$k_{\text{fs}} = \left(\int_{a_{\text{kd}}}^{a_{\text{eq}}} \frac{da}{aH} \frac{v(T_{\text{kd}})}{a} \right)^{-1} \sim 300 \xi^{-1/2} \left(\frac{m}{100 \text{ GeV}} \right)^{1/2} \left(\frac{T_{\text{kd}}}{50 \text{ MeV}} \right)^{5/6} \left(\frac{5 \text{ MeV}}{T_{\text{RH}}} \right)^{4/3} k_{\text{RH}} \quad (7.6a)$$

$$k_{\text{ao}} = a_{\text{kd}} H(T_{\text{kd}}) \sim 20 \left(\frac{T_{\text{kd}}}{50 \text{ MeV}} \right)^{4/3} \left(\frac{5 \text{ MeV}}{T_{\text{RH}}} \right)^{4/3} k_{\text{RH}} \quad (7.6b)$$

where m is the DM mass, ξ is the ratio of DM and SM temperature at the time of kinetic decoupling. In evaluating the above expressions we assumed that the kinetic decoupling temperature is much larger than the reheat temperature, $T_{\text{kd}} \gg T_{\text{RH}}$, and we made use of the temperature scaling $T \sim a^{-3/8}$ during EMD. We see that for a large values of T_{kd} and m that k_{cut} can be comparable or larger to the model-independent estimate in Eq. 7.4.

Small-scale cut-offs also exist in models where DM is an axion-like particle (ALP), where the wave-like nature of the particle suppresses power for wavenumbers larger than the comoving horizon at the start of ALP oscillations [39]

$$k_{\text{osc}} = \left(\frac{m_a}{H(T_{\text{RH}})} \right)^{1/3} \sim 10^3 \left(\frac{m_a}{10^{-5} \text{ eV}} \right)^{1/3} \left(\frac{5 \text{ MeV}}{T_{\text{RH}}} \right)^{2/3} \left(\frac{10.75}{g_*(T_{\text{RH}})} \right)^{1/6} k_{\text{RH}}. \quad (7.7)$$

Depending on the ALP mass, the model-independent cut-off can be easily the dominant source of suppression in these scenarios as well.

Finally note that we have focused on the scenario where stable DM exists independently of the EM responsible for EMD. An alternative scenario is where the decays of EM also produce the majority of DM [25, 31, 59, 60]. If the DM particles are weakly-interacting and do not thermalize, we expect them to free-stream a distance of order the horizon size at reheating, $k_{\text{fs}} \sim k_{\text{RH}}$, precluding

any enhanced structure growth at all. If they do thermalize right after production, kinetic decoupling effects described above similarly wash out any potential enhancement from EMD [35]. Therefore our model-independent small-scale cut-off is accurate for models with EMD-enhanced structure only if other suppressions due to the microphysics of the specific DM candidate are subdominant.

8 Conclusion

Despite its enormous success the Standard Model of Cosmology has experimental support over a relatively modest range of scales and includes an aggressive extrapolation to scales where no observational evidence is accessible. The “desert”, between the end of inflation and the BBN period is usually assumed populated by a long and uneventful epoch of adiabatically expanding thermal radiation. While this is certainly the simplest extrapolation it provides little challenge for observational cosmology.

A similar desert exists in the particle physics Standard Model (SM) from the electroweak scale to the GUT or Planck scale, where uncertain gravitational physics will likely dominate. In particle physics a plethora of Beyond the Standard Model (BSM) physics scenarios have been developed as possible UV extensions of the SM. Differing BSM models populate the desert with a wide variety of new particles and interactions, giving numerous targets for experiments on the energy, luminosity and precision frontiers.

This paper studies models with an early period of matter domination (EMD) in place of part of the radiation desert. EMD is particularly interesting, not only because it naturally arises in several BSM models but also due to its intimate connection with dark matter (DM). EMD changes the preferred parameter space for DM models and modifies the present day cosmological inhomogeneities on very small scales that have not yet been observationally probed.

While EMD enhances the growth of linear inhomogeneities, we have shown that the non-linear evolution instead suppresses power on small scales. This surprising and perhaps counter-intuitive result can be easily understood. If EMD is sufficiently long, most of the matter ends up in gravitationally-bound early halos (EHs). EMD ends when the matter that dominates the energy budget of the universe decays into SM radiation, unbinding the EHs. As a result, the stable dark matter component present in the EHs free-streams and erases inhomogeneities on small scales. In other works it is assumed that the decay was instantaneous to enable an estimate of the free-streaming cut off scale. We have shown here that the evolution of EHs is better characterized by several stages: adiabatic expansion of the gravitationally bound object; peeling of successive outer layers as EH self-gravity becomes less important; and finally free-streaming of DM particles. We referred to the entirety of this process as explosive evaporation. The resulting free-streaming velocity, eq. 4.9, is significantly smaller than in the instantaneous decay approximation which would predict a velocity $\sim \sqrt{\frac{G M(r_0)}{r_0} \frac{\tau}{t_{\text{free}}}}$ at time t_{free} which is a factor $\sim 2 \sqrt{\frac{\sqrt{\delta_\tau}}{\ln \delta_\tau}}$ too large. The particle velocity distribution is also quite different.

The distribution of DM at late times is determined by a superposition of EH remnants. We use a halo model to construct the DM power spectrum that includes the non-linear effects described above, which cut off the EMD-driven enhancement. This cutoff depends only on the mass distribution of the EHs and their density profiles.

Our semi-analytic results for the evolution of individual EHs are based on the assumptions of isolated spherical halos and circular orbits. We have shown that these simple simulations are similar to more general orbital distributions using a simple shell code. We expect that aspherical triaxial or rotationally-supported halos to yield similar remnant DM distributions. It would be interesting to verify these results for more realistic halos and to study the impact of halo overlap in N -body simulations.

We applied our results to two benchmark models containing either a long or short period of EMD, which enhance a nearly scale-invariant primordial spectrum of density fluctuations (as extrapolated from that observed on much larger scales). In linear perturbation theory alone, one would expect the “Long EMD” cosmology to feature the greatest enhancement in small scale power. Interestingly, we find that the non-linear effects described above are more significant in this cosmology. As a result the “Short EMD” case tends to have larger inhomogeneity at small scales. As a final application of these results, we estimated the typical microhalo masses and formation times using the Press-Schechter approach. While both long and short EMD cosmologies lead to the formation of microhalos after matter-radiation equality, this occurs much earlier in “Short EMD” due to the relatively smaller impact of non-linear effects. We conclude that there is an “optimal” duration for EMD in order to generate the maximum enhancement of the power spectrum at late times, that minimizes the suppression due to explosive evaporation of EHs. This highlights the importance of the duration of EMD for the late-time matter distribution, an aspect of these cosmologies that has not yet been comprehensively studied.

While our calculation of explosive evaporation of EHs was agnostic to their formation history, we made several important assumptions about the nature of DM. We neglected any interactions between the dominant matter component responsible for EMD and the DM, and assumed that they have identical phase space distribution. It would be interesting to explore how relaxing some of these assumptions affects the late-time distribution of matter. We also specialized to EMD as a way to enhance density perturbations. Similar effects are generated in other cosmologies, such as an early period of kination. It is an open question as to whether non-linear physics plays a similarly important role here [61].

Finally, we stopped short of examining the observational signatures of microhalos forming at late times, such as those in pulsar timing arrays [62] or microlensing [63]. Given that enhancement of small-scale structure is so prevalent in many cosmological and particle physics models of DM, it is important and exciting to develop these and other techniques, as they can provide access to the earliest moments in the evolution of the universe.

Acknowledgements

We thank Arka Banerjee, Carlos Blanco and Patrick Draper for helpful discussions. This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. GB acknowledges support from the MEC Grant FPA2017-845438 and the Generalitat Valenciana under grant PROM-ETEOII/2017/033.

A Adiabatic Homologous Expansion of Halos

Here we model the evolution of early halos during EMD and the beginning of LRD. Assume that the early matter (EM) and late matter (LM) originally populate the orbits of a halo in the same way. When the EM decays we assume the decay products exit the halo on a timescale much smaller than the dynamical time of the halo so one may treat the decays as an instantaneous disappearance. This is accurate so long as the speed of the decay products is much larger than the escape velocity. The rate of disappearance is assumed uniform throughout the halo which would be accurate for particle decays if the halo is non-relativistic so the Lorentz factors are nearly unity.

Under these assumptions the phase space distribution of EM and LM are described by the same 1-particle distribution function $f[\mathbf{x}, \mathbf{v}, t]$ which for nonrelativistic halos obeys the Newton-Vlasov

equation

$$\frac{\mathbb{D}}{\mathbb{D}t}f \equiv \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{x}} - \left(\frac{\partial}{\partial \mathbf{x}} \Phi[\mathbf{x}, t] \right) \cdot \frac{\partial}{\partial \mathbf{v}} \right) f[\mathbf{x}, \mathbf{v}, t] = 0 \quad (\text{A.1})$$

where $\mathbf{x}$ and $\mathbf{v}$ are the physical (not comoving) Cartesian coordinate and velocity. This equation describes the trajectories of the particles which do not decay and requires no modification for decaying particles if one normalizes the distribution function as:

$$\begin{aligned} \rho_{\text{EM}}[\mathbf{x}, t] &= (1 - f_{\text{LM}}) e^{-t/\tau} \int d^3\mathbf{v} f[\mathbf{x}, \mathbf{v}, t] \\ \rho_{\text{LM}}[\mathbf{x}, t] &= f_{\text{LM}} \int d^3\mathbf{v} f[\mathbf{x}, \mathbf{v}, t] \end{aligned} \quad (\text{A.2})$$

where ρ_{EM} and ρ_{LM} are the EM and LM mass density; τ is the decay time; and f_{LM} the fraction of mass in LM particles. The Newtonian gravitational potential is given by

$$\begin{aligned} \Phi[\mathbf{x}, t] &= - \int d^3\mathbf{x}' \frac{G(\rho_{\text{EM}}[\mathbf{x}', t] + \rho_{\text{LM}}[\mathbf{x}', t])}{|\mathbf{x} - \mathbf{x}'|} \\ &= -d[t] \int d^3\mathbf{x}' \int d^3\mathbf{v} \frac{G f[\mathbf{x}', \mathbf{v}, t]}{|\mathbf{x} - \mathbf{x}'|}. \end{aligned} \quad (\text{A.3})$$

where $d[t] = (1 - f_{\text{LM}}) e^{-t/\tau} + f_{\text{LM}}$ in the decaying particle EMD scenario.

During a prolonged EMD epoch we first expect gravitational collapse into bound structures (halos) after which halos successively merge into larger and larger halos (hierarchical clustering). Well into hierarchical clustering most matter is in a halo environment much denser than the cosmological mean where the dynamical timescale is much shorter than the expansion (Hubble) time: $H\tau_{\text{dyn}} \ll 1$. Mergers are assumed episodic after which halos quickly relax into a nearly stationary equilibrium which is represented by $\frac{\partial}{\partial t}f = 0$ so

$$\begin{aligned} 0 &= \left(\mathbf{v} \cdot \frac{\partial}{\partial \mathbf{x}} - \left(\frac{\partial}{\partial \mathbf{x}} \Phi_{\text{eq}}[\mathbf{x}] \right) \cdot \frac{\partial}{\partial \mathbf{v}} \right) f_{\text{eq}}[\mathbf{x}, \mathbf{v}] \\ \Phi_{\text{eq}}[\mathbf{x}] &= - \int d^3\mathbf{x}' \int d^3\mathbf{v} \frac{G f_{\text{eq}}[\mathbf{x}', \mathbf{v}, t]}{|\mathbf{x} - \mathbf{x}'|}. \end{aligned} \quad (\text{A.4})$$

Such an equilibrium could persist indefinitely if not for the EM decay in the EMD scenario which causes Φ to vary introducing an unavoidable time dependence. So long as $\tau \gg \tau_{\text{dyn}}$ the rate of change is slow and one can follow the evolution of $f[\mathbf{x}, \mathbf{v}, t]$ in an adiabatic approximation.

When $\tau \gg \tau_{\text{dyn}}$ each particle will undergo many orbits during a decay time, and hence τ introduces no new length scale into the dynamics. One therefore expects homologous expansion where the halo slowly passes through a scaling of the initial equilibrium. Arbitrary homologous scaling of position, velocity and phase space density is of the form:

$$f[\mathbf{x}, \mathbf{v}, t] \cong \frac{1}{h_{\text{f}}[t]} f_{\text{eq}} \left[\frac{\mathbf{x}}{h[t]}, \frac{\mathbf{v}}{h_{\text{v}}[t]} \right] \quad \Phi[\mathbf{x}, t] \cong - \frac{h[t]^2 h_{\text{v}}[t]^3}{h_{\text{f}}[t]} \Phi_{\text{eq}} \left[\frac{\mathbf{x}}{h[t]} \right] \quad (\text{A.5})$$

where $h[t]$, $h_{\text{v}}[t]$ and $h_{\text{f}}[t]$ scale the size, velocity and phase space density. In order to preserve the total LM mass one requires $h_{\text{v}}[t] = h[t]^{-1}$. The Vlasov equation preserves the mass weighted distribution of phase space density so one also requires $h_{\text{f}}[t] = 1$. Thus adiabatic scaling must be of the form

$$f[\mathbf{x}, \mathbf{v}, t] \cong f_{\text{eq}}[\mathbf{x}_{\text{s}}, \mathbf{v}_{\text{s}}] \quad \Phi[\mathbf{x}, t] \cong \frac{1}{h[t]} \Phi_{\text{eq}}[\mathbf{x}_{\text{s}}]. \quad (\text{A.6})$$

where $\mathbf{x}_s \equiv \mathbf{x}/h[t]$ and $\mathbf{v}_s \equiv h[t]\mathbf{v}$. Substituting this ansatz into the Eq. A.2 and using Eq. A.4 one finds

$$\begin{aligned} \frac{\mathbb{D}}{\mathbb{D}t}f &= \left(d[t] - \frac{1}{h[t]}\right) \left(\frac{\partial}{\partial \mathbf{x}_s} \Phi_{\text{eq}}[\mathbf{x}_s]\right) \cdot \frac{\partial}{\partial \mathbf{v}_s} f_{\text{eq}}[\mathbf{x}_s, \mathbf{v}_s] \\ &\quad - \frac{\partial \ln h[t]}{\partial t} \left(\mathbf{x}_s \cdot \frac{\partial}{\partial \mathbf{x}_s} - \mathbf{v}_s \cdot \frac{\partial}{\partial \mathbf{v}_s}\right) f_{\text{eq}}[\mathbf{x}_s, \mathbf{v}_s] \cong 0. \end{aligned} \quad (\text{A.7})$$

In the adiabatic limit, $\left|\frac{\partial \ln h[t]}{\partial t}\right| \tau_{\text{dyn}} \ll 1$, so one neglects the last term and one finds that $h[t] = d[t]^{-1}$ gives the adiabatic scaling.

EMD halos will have differing mass, density profiles, angular momentum, etc. The distribution of halo properties will depend on the early universe scenario which gives rise to EMD. This adiabatic scaling is a general behavior and applies to each halo individually independent of halo properties. This scaling also works for any functional form $d[t]$.

In the EMD scenario the LM contributes negligibly to the mass density before LMD, i.e. $f_{\text{LM}} \ll 1$, so during EMD and early LRD the LM particles act as test particles in the gravitational field. In this case $d[t] \cong e^{-t/\tau}$ and $h[t] \cong e^{t/\tau}$: the halo size grows while the particle velocities shrink, both exponentially, during the adiabatic phase.

When a halo first collapses $\rho_{\text{halo}} \gg \bar{\rho}$ where ρ_{halo} and $\bar{\rho}$ are the halo and cosmological mean density. During EMD ρ_{halo} remains nearly constant while $\bar{\rho}$ decreases rapidly. The dynamical to expansion timescale ratio, $H\tau_{\text{dyn}} \sim (\bar{\rho}/\rho_{\text{halo}})^{1/2}$, decreases rapidly during EMD but adiabatic expansion gives $\rho_{\text{halo}} \propto e^{-4t/\tau}$ so $H\tau_{\text{dyn}}$ rises rapidly during LRD. Since for any bound object $\rho_{\text{halo}} \gg \bar{\rho}$ adiabatic expansion persists for a time significantly greater than τ and bound EMD halos exist well into the radiation era but grow larger and less bound. Adiabatic expansion eventually ends when $\rho_{\text{halo}} \lesssim \bar{\rho}$ giving rise to a period of free expansion (see text). As a result of the adiabatic expansion phase one finds free expansion rather than free contraction in the radiation era.

B Non-Circular Orbits and Shell Crossing

In this section we relax some of the assumptions that were used in §4 to describe the evolution of early halos through reheating. We use a simple shell code which retains spherical symmetry of an isolated halo but allows us to consider different angular momenta (i.e., non-circular orbits) and sensitivity to the initial density profile, as well as automatically including shell crossing. From the perspective of the late-time power spectrum, the most interesting quantity is the asymptotic velocity of (formerly) EH shells well after reheating, which determines the EH remnant profile. The physical velocity with respect to the comoving frame ($v = ad(r/a)/dt$) yields:

$$v(a) = -H \left(\frac{3M_{<}}{4\pi\bar{\rho}_{\text{LM}}} \right)^{1/3} \mathcal{V}, \quad (\text{B.1})$$

where we defined

$$\mathcal{V} = -\frac{a\bar{\Delta}'}{\bar{\Delta}^{4/3}} \quad (\text{B.2})$$

with $\bar{\Delta}$ defined in Eq. 4.10. Since $\bar{\Delta}' < 0$, $\mathcal{V} > 0$; $\mathcal{V}$ is a useful quantity because it encodes the dynamical effects due to EH explosive evaporation (the usual redshifting of velocity is captured by the first two factors in Eq. B.1). Moreover, $\mathcal{V}$ should approach a constant deep in the radiation era which can be seen from Eq. 4.11:

$$\Delta'' = -\frac{\Delta'}{a^2} + \frac{4}{3} \frac{(\Delta')^2}{\Delta} \Rightarrow \mathcal{V}' = 0. \quad (\text{B.3})$$

In principle the constant value of $\mathcal{V}$ depends on the parameters of the EH, such as the initial mass profile and average overdensity. However, using numerical solutions of Eq. 4.11 we have checked that $\mathcal{V} \approx 2$ over a wide range of “initial” conditions $\bar{\Delta}_\tau$. We will therefore use $\mathcal{V}$ as a simple diagnostic and check this result with an N -shell simulation allowing for non-circular orbits and shell crossings. We will find that $\mathcal{V} \approx 2$ even in this more general set-up, implying that the results in the main text are robust. We leave more detailed comparisons, the impact EH non-sphericity, substructure and halo overlap to future work.

We follow the evolution of N spherical, equal-mass shells, each with an equation of motion given by Eq. 4.3. Their evolution is coupled through the mass interior term

$$M_<(r_j, t) = \sum_{r_j < r_k} m \left[f + (1 - f)e^{-t/\tau} \right], \quad (\text{B.4})$$

where j and k are shell labels, m is the shell mass at early times and the two terms in the brackets represent LM and EM contributions to the mass.

We initialize the shell configuration similarly to Refs. [64, 65], which we summarize below. In order to specify the initial positions of the shells, we need postulate an initial mass profile for the overdensity:

$$1 + \delta(r_k, t) = \frac{3M_<(r_k, t)}{4\pi r_k^3 \bar{\rho}_{\text{mat}}(t)}, \quad (\text{B.5})$$

where $\bar{\rho}_{\text{mat}}$ is the background matter density. Following Ref. [64] we take at the initial time t_i

$$\delta(r_k, t_i) = \delta_{\text{avg}} \left(\frac{M_<(r_k, t_i)}{M_{\text{tot}}} \right)^{-\epsilon} = \delta_{\text{avg}} \left(\frac{k+1}{N} \right)^{-\epsilon}, \quad (\text{B.6})$$

where δ_{avg} and M_{tot} represent the average density contrast and mass of the whole shell configuration at early times, N is the number of shells and $\epsilon > 0$. The initial position of shell k is then [65]

$$r_k(t_i) = \left[\frac{3M_<(r_k, t_i)}{4\pi \bar{\rho}_{\text{mat}}(t_i)(1 + \delta(r_k, t_i))} \right]^{1/3}. \quad (\text{B.7})$$

If the initial mass distribution corresponds to δ in the linear regime (i.e., $\delta \ll 1$) then the initial velocity can be obtained from linear perturbation theory during matter domination. A standard result is that $\delta \propto a$, or equivalently $\dot{\delta}/\delta = H$. Differentiating Eq. B.5 with respect to time, using $\dot{\delta}/\delta$, $H = 2/(3t)$ during EMD and evaluating the result at $t = t_i$ gives

$$\frac{\dot{r}_k}{r_k} = H - \frac{\dot{\delta}_k}{3(1 + \delta_k)} = \frac{2}{3t_i} \left[1 - \frac{\delta_k}{3(1 + \delta_k)} \right], \quad (\text{B.8})$$

where $\delta_k = \delta(r_k, t_i)$, the first term is the usual Hubble expansion (i.e., if we only kept it the overdensity would remain constant in time) and the second term is responsible for the linear growth of δ . In practice we initialize our simulations with large overdensities $\delta_{\text{avg}} \gtrsim 1$; we find that initial density and velocity profiles do not impact the subsequent evolution since the shell configuration quickly virializes. Thus, as long as the initial time t_i is well before reheating, $t_i \ll \tau$, our results are not sensitive to the precise choice of shell position and velocity.

In order to assess the impact of shell angular momentum, we parametrize L similarly to Ref. [64]⁸

$$L^2 = 2\alpha GM_<(r_k, t_i)r_k, \quad (\text{B.9})$$

where α is a dimensionless parameter. The shells are gravitationally bound if

$$\alpha \leq 1 - \frac{(1 - \xi)^2}{1 + \delta}, \quad (\text{B.10})$$

⁸This differs slightly from Ref. [64] which took $L^2 = 2\alpha G \Delta M_<(r_k, t_i)r_k$, with ΔM the mass excess over background contained within r_k .

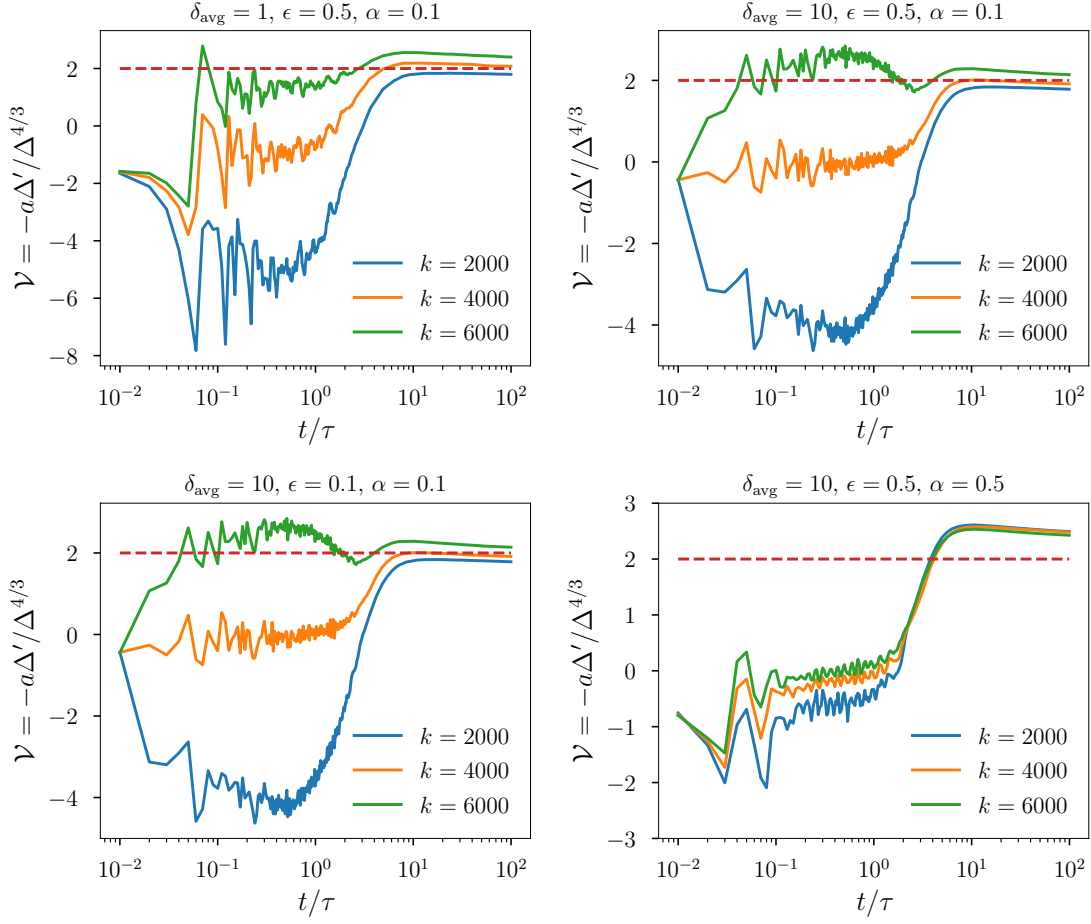


Figure 12. Evolution of velocity factor $\mathcal{V}$ (Eq. B.2) for three representative shells (out of a $N = 10000$ shell simulation) as a function of t/τ . Different panels correspond to different initial shell parameters, such as the initial average overdensity δ_{avg} , initial density profile slope ϵ (both defined in Eq. B.6) and the angular momentum parameter α (Eq. B.9). For a wide range of parameters $\mathcal{V}$ approaches ~ 2 (indicated by the dashed line) well after reheating, $t/\tau \gg 1$. The asymptotic value of $\mathcal{V}$ is mildly dependent on the initial angular momentum and overdensity.

where $\xi = \delta/[3(1 + \delta)]$ is the deviation from the Hubble expansion in the initial velocity, Eq. B.8. In this parametrization circular and radial orbits correspond to $\alpha = 1/2$ and $\alpha = 0$, respectively.

Following initialization, we evolve the shell ensemble forward in time using an adaptive leapfrog integration scheme well into radiation equality. In order to ensure stability of the solution we follow Ref. [65] to endow the shells with a finite thickness and constant density. This modification ensures that the inter-shell forces change continuously as the shells move past each other, leading to a more stable integration of the equations of motion.

In Fig. 12 we show the evolution of $\mathcal{V}$ for several choices of δ_{avg} , ϵ and α . These results confirm our initial finding that $\mathcal{V} \sim 2$ asymptotically for $t/\tau \gg 1$, although this value does have some mild dependence on the angular momentum of the shells.

C Linear Evolution

The Fourier-space density δ_i and velocity divergence $\theta_i = i\vec{k} \cdot \vec{v}_i$ perturbation equations for $i = \text{EM}$ (the EMD field), LM (the dark matter), and rad (SM radiation) are [33, 39]:

$$a^2 E \delta'_{\text{EM}} + \theta_{\text{EM}} + 3a^2 E \Phi' = \frac{a}{\tau} \Phi, \quad (\text{C.1a})$$

$$a^2 E \theta'_{\text{EM}} + a E \theta_{\text{EM}} + k^2 \Phi = 0, \quad (\text{C.1b})$$

$$a^2 E \delta'_{\text{rad}} + \frac{4}{3} \theta_{\text{rad}} + 4a^2 E \Phi' = \frac{a \bar{\rho}_{\text{EM}}}{\tau \bar{\rho}_{\text{rad}}} [\delta_{\text{EM}} - \delta_{\text{rad}} - \Phi], \quad (\text{C.1c})$$

$$a^2 E \theta'_{\text{rad}} + k^2 \Phi - k^2 \frac{\delta_{\text{rad}}}{4} = \frac{a \bar{\rho}_{\text{EM}}}{\tau \bar{\rho}_{\text{rad}}} \left[\frac{3}{4} \theta_{\text{EM}} - \theta_{\text{rad}} \right], \quad (\text{C.1d})$$

$$a^2 E \delta'_{\text{LM}} + \theta_{\text{LM}} + 3a^2 E \Phi' = 0, \quad (\text{C.1e})$$

$$a^2 E \theta'_{\text{LM}} + a E \theta_{\text{LM}} + k^2 \Phi = 0, \quad (\text{C.1f})$$

$$k^2 \Phi + 3a E^2 [a^2 \Phi' + a \Phi] = \frac{3}{2} a^2 [\bar{\rho}_{\text{EM}} \delta_{\text{EM}} + \bar{\rho}_{\text{rad}} \delta_{\text{rad}} + \bar{\rho}_{\text{LM}} \delta_{\text{LM}}], \quad (\text{C.1g})$$

where Φ is the gravitational potential ($\Phi = -\psi$ in the absence of anisotropic stress) and the prime denotes differentiation with respect to a , E is the dimensionless Hubble parameter

$$E^2 = \bar{\rho}_{\text{EM}} + \bar{\rho}_{\text{rad}} + \bar{\rho}_{\text{LM}}. \quad (\text{C.2})$$

The energy densities are normalized such that $E(a=1) = 1$, while $1/\tau$, k and θ_i are in units of H_1 , the Hubble parameter at $a=1$. These quantities can be obtained in physical units by, e.g., identifying the scale factor at reheating, a_{RH} with a physical temperature T_{RH} (or, equivalently, using Eq. 7.1).

The initial conditions for the above system of equations depends on whether modes enter during matter or radiation domination. In the “Long EMD” cosmology we assume all modes of begin their evolution during EMD, such that the initial conditions are the same as in Ref. [39]:

$$\delta_{\text{EM}} = \delta_{\text{LM}} = 2\delta_{\text{rad}} = 2\Phi_i \quad (\text{C.3})$$

$$\theta_{\text{EM}} = \theta_{\text{LM}} = \theta_{\text{rad}} = -\frac{2}{3} k^2 \sqrt{a} \Phi_i, \quad (\text{C.4})$$

where Φ_i is the superhorizon value of Φ . Since we specify Φ_i during EMD, its power spectrum is related to the primordial curvature power spectrum by

$$\langle \Phi_i^2 \rangle = \left(\frac{3}{5} \right)^2 \langle \mathcal{R}^2 \rangle. \quad (\text{C.5})$$

In the “Short EMD” cosmologies we assume that all modes begin their evolution during early RD, requiring a different set of initial conditions:

$$(4/3) \delta_{\text{EM}} = (4/3) \delta_{\text{LM}} = \delta_{\text{rad}} = 2\Phi_i \quad (\text{C.6})$$

$$\theta_{\text{EM}} = \theta_{\text{LM}} = \theta_{\text{rad}} = -\frac{1}{2\sqrt{\bar{\rho}_{\text{rad}}(a=1)}} k^2 a \Phi_i, \quad (\text{C.7})$$

which is in agreement with Newtonian gauge initial conditions in the standard cosmology [66]. Now Φ_i is fixed during early radiation domination, so its relationship to the curvature perturbation is

$$\langle \Phi_i^2 \rangle = \left(\frac{2}{3} \right)^2 \langle \mathcal{R}^2 \rangle. \quad (\text{C.8})$$

We solve the system of equations C.1 numerically with the initial conditions given in Eq. C.4 or Eq. C.7. Well after reheating, the solutions have the form given in Eq. 6.12, which we fit to numerics for a wide range of k . We present these fits below for the “Long EMD” and “Short EMD” cosmologies.

Long EMD We find that

$$I_1 \approx \frac{2}{3} \left(\frac{k}{k_{\text{RH}}} \right)^2 \left(0.757 + \frac{10.97}{(k/k_{\text{RH}})^{2/\alpha}} \right)^\alpha, \quad (\text{C.9})$$

where $\alpha = 1.13$;

$$I_2 \approx \frac{3.15}{(5.73 + (k/k_{\text{RH}})^{2/\beta})^\beta}, \quad (\text{C.10})$$

where $\beta = 0.923$;

$$a_{\text{hor}}/a_{\text{RH}} \approx 1.108 \left(\frac{1}{(k/k_{\text{RH}})^{1/\gamma} + 0.426(k/k_{\text{RH}})^{2/\gamma}} \right)^\gamma, \quad (\text{C.11})$$

where $\gamma \approx 0.302$. Note that for $k/k_{\text{RH}} \ll 1$, $(9/10)I_1 \approx 9$, $I_2 \approx 0.63$, and $a_{\text{hor}}/a_{\text{eq}} \approx k_{\text{eq}}/(\sqrt{2}k)$ reproduce the standard results for modes that enter the horizon during radiation domination well after reheating [47].⁹ In the opposite limit $k \gg k_{\text{RH}}$, these functions reproduce the approximate solution in Eq. 6.9.

Short EMD This cosmology has a finite period of EMD that is preceded by early RD; EMD starts when $a/a_{\text{RH}} = 4.5 \times 10^{-4}$, so in terms of scale factor EMD lasts a factor of ~ 2200 – see the right panel of Fig. 1. The mode that enters at the time of early equality is $k/k_{\text{RH}} \approx 90$. As before we solve the linear equations numerically for the DM density contrast and try to find simple analytic functions for I_1 and I_2 that fit the numerics well; these are more complicated because they need to capture 3 regimes now – early RD, EMD, and late RD. We find the following convenient fitting functions:

$$I_1 \approx \frac{2}{3} \left(\frac{k}{k_{\text{RH}}} \right)^2 \left(\frac{11.7}{(k/k_{\text{RH}})^{2/\alpha}} + \frac{1}{(1.15 + 0.00512(k/k_{\text{RH}})^{2/(\alpha\beta)})^\beta} \right)^\alpha \quad (\text{C.12})$$

where $\alpha \approx 1.06$, $\beta \approx 2.07$;

$$I_2 \approx \left(\frac{1}{745 + 24.2(k/k_{\text{RH}})^{1/\kappa}} + \frac{1}{1.68 + 0.289(k/k_{\text{RH}})^{2/\kappa}} \right)^\kappa, \quad (\text{C.13})$$

where $\kappa \approx 0.902$;

$$a_{\text{hor}}/a_{\text{RH}} \approx \left(\frac{0.005}{(k/k_{\text{RH}})^{1/\gamma}} + \frac{1}{(0.702(k/k_{\text{RH}})^{1/(\delta\gamma)} + 0.304(k/k_{\text{RH}})^{2/(\delta\gamma)})^\delta} \right)^\gamma, \quad (\text{C.14})$$

where $\delta \approx 0.406$, $\gamma \approx 0.686$.

The fitting functions given above match the numerical results to better than 5% for $10^{-2} \leq k/k_{\text{RH}} \leq 500$. They also have correct asymptotics where numerical solutions are impractical, enabling us to study the density field across an even wider range of scales.

D Early Halo Model

In this section we describe the halo model of the matter power spectrum making use of the results of Refs. [50, 51]. We start by writing the density field as a superposition of halos with different masses m_i :

$$\rho(\mathbf{r}) = \sum_i \rho_{\text{EH}}(\mathbf{r} - \mathbf{r}_i | m_i) = \sum_i \int dmd^3r' \delta(m - m_i) \delta^3(\mathbf{r}' - \mathbf{r}_i) m u(\mathbf{r} - \mathbf{r}' | m), \quad (\text{D.1})$$

⁹The additional factor of 9/10 is needed when comparing “Long EMD” to the standard cosmology because the relationship between the superhorizon gravitational potential (which sets the initial conditions for the density perturbations) and the curvature perturbation (whose power spectrum is usually specified) is different during EMD or RD – see Eqs. C.5 and C.8.

where we assume that all halos have the same functional form for the density profile ρ_{EH} that only depends on the halo mass. In the second equality we have written the sum over halos in terms of the mass-normalized density profile u . A realization of the halo model consists of a set of $\{\mathbf{r}_i, m_i\}$; an ensemble average over possible realizations is then defined by

$$\left\langle \sum_i \delta(m - m_i) \delta^3(\mathbf{r}' - \mathbf{r}_i) \right\rangle = \frac{dn}{dm} \equiv n(m), \quad (\text{D.2})$$

where $n(m)$ is the density of halos per unit mass (i.e., the halo mass function) and we have performed an average over space. It is easy to check that this leads to

$$\langle \rho(\mathbf{r}) \rangle = \int dm m n(m) = \bar{\rho}_{\text{LM}} a^3, \quad (\text{D.3})$$

using $\int d^3r u(r|m) = 1$. The two point function depends both on the one point function above, as well as on the halo-halo correlation function $\xi_{hh}(r|m_1, m_2)$:

$$\begin{aligned} \langle \rho(\mathbf{r}_1) \rho(\mathbf{r}_2) \rangle &= \int dm m^2 n(m) \int d^3y u(\mathbf{r}_1 - \mathbf{y}|m) u(\mathbf{r}_2 - \mathbf{y}|m) + \int dm_1 dm_2 m_1 n(m_1) m_2 n(m_2) \times \\ &\times \int d^3y_1 d^3y_2 [1 + \xi_{hh}(|\mathbf{y}_1 - \mathbf{y}_2||m_1, m_2)] u(\mathbf{r}_1 - \mathbf{y}_1|m_1) u(\mathbf{r}_2 - \mathbf{y}_2|m_2). \end{aligned} \quad (\text{D.4})$$

The first term is the one-halo contribution, encoding the correlation between elements of DM mass distribution inside a single halo; the second two-halo term describes the correlation between mass elements in different halos.

The above expression can be used to find the two-point correlation function of the density contrast. In Fourier space we have

$$\langle \delta(k) \delta(k') \rangle = [P_{1h}(k) + P_{2h}(k)] (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{k}') \equiv P(k) (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{k}'), \quad (\text{D.5})$$

where the one- and two-halo terms are

$$P_{1h}(k) = \frac{1}{\bar{\rho}_{\text{LM}}^2} \int dm m^2 n(m) |\tilde{u}(k|m)|^2 \quad (\text{D.6a})$$

$$P_{2h}(k) = \frac{1}{\bar{\rho}_{\text{LM}}^2} \int dm_1 dm_2 m_1 n(m_1) m_2 n(m_2) P_{hh}(k|m_1, m_2) \tilde{u}(k|m_1) \tilde{u}(k|m_2)^*. \quad (\text{D.6b})$$

The tildes denote Fourier transform (in comoving coordinates) and

$$P_{hh}(k|m_1, m_2) = \int d^3r \xi_{hh}(r|m_1, m_2) e^{i\mathbf{k} \cdot \mathbf{r}} \quad (\text{D.7})$$

is the halo-halo power spectrum. This result relates the non-linear mass density contrast power spectrum to the halo mass function and power spectrum of the halo centers. Note that from Eq. 5.14 we have

$$\tilde{u}(k|m) = \frac{1}{a^3} \mathcal{F}(k|m). \quad (\text{D.8})$$

The Press-Schechter formalism and its extensions provide a model of the halo mass function in terms of the density variance. Following Refs. [51, 67] we approximate

$$P_{hh}(k|m_1, m_2) \approx b(m_1) b(m_2) P(k)^{\text{lin}}, \quad (\text{D.9})$$

where $b(m)$ is bias parameter that can be estimated for a given mass function [51]. Note that this form cannot be valid up to arbitrarily large k : the halo-halo correlation function ξ_{hh} must turn over

since halos are mutually exclusive - equivalently P_{hh} must decrease at large k . We ignore this problem for now, but return to it in §D.2.

Combining Eqs. D.6, D.8 and D.9 yields the halo model for the matter power spectrum $P = P_{1h} + P_{2h}$, where

$$P_{1h}(k) = \frac{1}{\bar{\rho}_{\text{LM}} a^3} \int dm \frac{df}{d \ln m} |\mathcal{F}(k|m)|^2 \quad (\text{D.10a})$$

$$P_{2h}(k) = \left| \int d \ln m \frac{df}{d \ln m} b(m) \mathcal{F}(k|m) \right|^2 P(k)^{\text{lin}}, \quad (\text{D.10b})$$

and we used $n(m) = (\bar{\rho}_{\text{LM}} a^3 / m^2) df / d \ln m$. In the following section we apply these results to an analytic model.

D.1 An Analytic Example

In this section we apply some of the results from the previous subsection to a simple power law power spectrum of the form

$$P(k)^{\text{lin}} = A k^\alpha, \quad \Delta^2(k) = \frac{A}{2\pi^2} k^{3+\alpha} \quad (\text{D.11})$$

where A and α are constants; for EMD-motivated power spectra we expect $\alpha \approx 1$ (see Eq. 6.10), while in Λ CDM have $\alpha \approx 0$ for modes that enter the horizon during radiation domination (neglecting the k -dependence of the logarithmic growth during RD). Using the definition in Eq. 6.5 we can compute the density fluctuation variance for this power spectrum:

$$\sigma^2(R(M)) = \frac{A}{2\pi^2(3+\alpha)} \frac{1}{R^{3+\alpha}} = \delta_c^2 \left(\frac{M_*}{M} \right)^{1+\alpha/3}, \quad (\text{D.12})$$

where in the last equality we used the typical collapse mass M_* defined in Eq. 6.4 to simplify the expression. Note that we are focusing here on a particular instant in time; redshift dependence can be easily restored by introducing a growth function into σ^2 (or alternatively into the collapse threshold δ_c). The Press-Schechter halo mass function is naturally presented in terms of the variable

$$\nu = \frac{\delta_c^2}{\sigma^2} = \left(\frac{M}{M_*} \right)^{1+\alpha/3}, \quad (\text{D.13})$$

such that

$$\frac{m^2}{\bar{\rho}_{\text{LM}} a^3} \frac{dn}{dm} d \ln m = \frac{df}{d \ln m} d \ln m = \sqrt{\frac{\nu}{2\pi}} e^{-\nu/2} d \ln \nu. \quad (\text{D.14})$$

We can now estimate the non-linear power spectra of the halo model given in Eqs. D.10a and D.10b. We will do this in two cases: for the physically motivated late-time form of $\mathcal{F}$ computed in §4 which encodes the explosion of EHS, and for the scenario where no explosion occurs (and therefore $\mathcal{F}$ is just the Fourier transform of the density profiles). These cases can be handled simultaneously because $\mathcal{F}$ has a universal behaviour: $\mathcal{F} \rightarrow 1$ for $k \ll k_s$, and it decreases as $\mathcal{F} \sim (k_s/k)^\beta$ for large k , where $k_s(m)$ is a characteristic scale that depends on the mass of the EH and β is an m -independent constant. For example for the EMD case, $k_s = k_c$ given in Eq. 7.4 and $\beta = 2$; for unexploded NFW-like EHS, $k_s \propto r_s^{-1}$ and $\beta = 2$ follows from the Fourier transform of the NFW profile. In either case, the m -dependence of k_s can be expressed as a function of ν using Eq. D.13. The one-halo term then has the general form

$$P_{1h}(k) \sim \frac{m_*}{\bar{\rho}_{\text{LM}} a^3} \begin{cases} 1 & k \leq k_s(m_*) \\ (k_s(m_*)/k)^{2\beta} & k > k_s(m_*), \end{cases} \quad (\text{D.15})$$

where the integral over ν just gives $m_* \times \mathcal{O}(1)$, where the $\mathcal{O}(1)$ constant depends on α and β . Similarly we find for the two-halo power spectrum

$$P_{2h}(k) \sim P(k)^{\text{lin}} \begin{cases} 1 & k \leq k_s(m_*) \\ (k_s(m_*)/k)^{2\beta} & k > k_s(m_*), \end{cases} \quad (\text{D.16})$$

We can immediately see that both the one- and two-halo terms are expected to turn over at $k = k_c(m_*)$; for $T_{\text{RH}} = 5 \text{ MeV}$ $m_* \approx 6 \times 10^{-11} M_\odot$, which yields $k_c(m_*)/k_{\text{eq}} \approx 7 \times 10^7$ at $a = a_{\text{eq}}$. Note that for the EMD-inspired case with $\alpha = 1$ and $\beta = 2$, and the dimensionless two-halo powerspectrum $k^3 P_{2h}(k)/(2\pi)^3 \sim \text{const}$ for $k > k_c$.

It is well known that the simplest application of the halo model in Eqs. D.10a and D.10b results in some unphysical features of the matter power spectrum [51]. First, the one-halo term in Eq. D.10a contributes a constant to the power spectrum as $k \rightarrow 0$, which implies that the halo model power spectrum does not reduce to the linear one on large scales (alternatively, the integral constraint on the matter correlation function $\int_0^\infty r^2 \xi(r) dr = 0$ is not satisfied). Second, the above calculation ignores the fact that halos are non-overlapping objects by using the approximation in Eq. D.9. Ref. [68] found that these issues are, in fact, related, and addressing exclusion in a physical way results in the correct behaviour of the matter power spectrum as $k \rightarrow 0$. We therefore consider a simple implementation of halo exclusion effect in the context of a simplified model like Eq. D.11 in the following subsection.

D.2 Halo Exclusion

The approximate relationship between the halo-halo power spectrum and the linear matter power spectrum in Eq. D.9 is valid only on scales exceeding the typical halo size, i.e. on linear scales. We will attempt to understand the non-linear regime by starting with simple model of the PS, similar to the one in the previous section:

$$P(k)^{\text{lin}} = A k^\alpha e^{-\lambda k}, \quad \Delta^2(k) = \frac{A}{2\pi^2} k^{3+\alpha} e^{-\lambda k}, \quad (\text{D.17})$$

where we introduced an exponential cut-off to ensure the finiteness of the various integrals (the functional form of the cut-off is chosen to enable some analytic results); the cut-off wavenumber λ^{-1} is distinct from the one computed in the main text (k_c) – it is just a regulator that we will take to be much larger than other scales. We want to implement the fact that the probability of finding two halos separated by less than some exclusion radius r_e is 0; since the probability of finding two halos at separation r is $\propto 1 + \xi_{hh}(r)$, this requirement becomes

$$\xi_{hh}(r \leq r_e) + 1 = 0. \quad (\text{D.18})$$

We can mock up this behavior by modifying $\xi^{\text{lin}}(r)$, the Fourier transform of $P(k)^{\text{lin}}$, as follows. The exclusion condition, Eq. D.18, is enforced if we define

$$\xi_{hh}(r) = \begin{cases} \xi^{\text{lin}}(r) & r > r_e \\ -1 & r \leq r_e \end{cases}. \quad (\text{D.19})$$

This is a very coarse model for the halo-halo correlation because it uses the linear theory result in the mildly-non-linear regime $r \sim r_e$ and because it treats halos as hard spheres of equal sizes – see, e.g., Ref. [69] for possible improvements. The Fourier transform of this function is [69]

$$P_{hh}(k) = \int_{r_e}^\infty d^3r \xi^{\text{lin}}(r) e^{ik \cdot r} - \int_0^{r_e} d^3r e^{ik \cdot r} \quad (\text{D.20})$$

$$= P^{\text{lin}}(k) - \int_0^{r_e} d^3r \xi^{\text{lin}}(r) e^{ik \cdot r} - \int_0^{r_e} d^3r e^{ik \cdot r}. \quad (\text{D.21})$$

For the spectrum in Eq. D.17 some of the integrals can be performed analytically. Let us first discuss the explicit form of the linear correlation function. For $\alpha = 1$ (the EMD-like case) we have [70]

$$\xi(r)^{\text{lin}} = \frac{A}{\pi^2 \lambda^4} \frac{\sin(3 \arctan r/\lambda)}{(r/\lambda)(1 + r^2/\lambda^2)^{3/2}}. \quad (\text{D.22})$$

This correlation function approaches zero from below at large distances $r/\lambda \gg 1$

$$\xi(r)^{\text{lin}} \sim -\frac{A}{\pi^2 r^4}, \quad (\text{D.23})$$

and is nearly constant and positive at short distances $r/\lambda \ll 1$

$$\xi(r)^{\text{lin}} \sim \frac{3A}{\pi^2 \lambda^4}. \quad (\text{D.24})$$

This means that matter is *anti*-correlated at large distances and strongly correlated at small distances. The positive correlation, however, occurs for $r \lesssim \sqrt{3}\lambda$ which is sensitive to the precise nature of the cut-off. Note that this positive contribution is necessary to ensure that the integral constraint is satisfied [71]

$$\lim_{k \rightarrow 0} P^{\text{lin}}(k) = 0 = \int d^3 r \xi(r)^{\text{lin}}. \quad (\text{D.25})$$

The excision of the halo overlap, Eq. D.19, removes the region with $r < r_e$, including the region $r \lesssim \sqrt{3}\lambda$ where $\xi(r)^{\text{lin}}$ is positive. This means that the two-halo term computed from the halo-halo correlation function in Eq. D.19 cannot satisfy the integral constraint by itself; this is fine since only the total power spectrum (the sum of one- and two-halo terms) is an observable.

We therefore find that exclusion of overlap regions *reduces* the power spectrum, even on large $k \rightarrow 0$ scales, as discussed in, e.g., Ref. [69, 72]. This is just the statement that the last two terms in Eq. D.21 are negative. These two terms are constant in the $k \rightarrow 0$ limit and therefore represent a reduction in the halo shot noise (which is nominally part of the one-halo term). Ref. [68] argued that $k \rightarrow 0$ limit of these terms must nearly cancel against the one-halo shot-noise-like contribution ensuring the correct asymptotic behaviour of the entire power spectrum. This cancellation, however, is sensitive to the behaviour of $\xi(r)$ in the non-linear regime, which we do not model. We therefore opt for a simplified calculation of the power spectrum, focusing on the two-halo term from which we subtract off the shot-noise-like contribution (the $k \rightarrow 0$ limit); the small scale correlations encoded in the one-halo term (and the subtracted piece) become subdominant at late times.

The $k \rightarrow 0$ limit, C , of the exclusion terms (last two terms of Eq. D.21) can be computed analytically

$$C \equiv \int_0^{r_e} d^3 r \xi^{\text{lin}}(r) + \int_0^{r_e} d^3 r = \frac{4A r_e^3}{\pi \lambda^4 (1 + r_e^2/\lambda^2)^2} + \frac{4\pi r_e^3}{3} \rightarrow \frac{4A}{\pi r_e} + \frac{4\pi r_e^3}{3}, \quad (\text{D.26})$$

where the last step corresponds to $r_e/\lambda \gg 1$; note that the cut-off dependence drops out, so the result is insensitive to the precise nature of the regulator as long as r_e is well above λ .

We now define a “subtracted” halo-halo PS $\tilde{P}_{hh}$ via

$$P_{hh}(k) = \tilde{P}_{hh}(k) - C. \quad (\text{D.27})$$

The discussion in the previous paragraph implies that $\tilde{P}_{hh}$ has two nice properties: $\tilde{P}_{hh}(k) \approx P^{\text{lin}}(k)$ for $kr_e \ll 1$, and $\tilde{P}_{hh}(k) \approx C$ for $kr_e \gg 1$. The subtracted PS $\tilde{P}_{hh}(k)$ and the constant C are compared to the linear power spectrum in Fig. 13 for $r_e/\lambda = 10^2$ and $A = r_e^4$. We see that the main impact of halo exclusion is the flattening of $\tilde{P}_{hh}(k)$ for $kr_e \gtrsim 1$ (in other words, halo exclusion leads to a $1/k$ suppression at these scales). Because the $r_e/\lambda \gg 1$, these features are not sensitive to the cut-off λ .

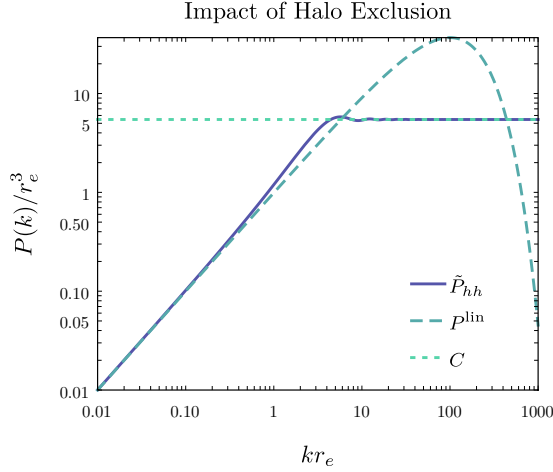


Figure 13. Impact of halo exclusion on the halo-halo power spectrum in a toy model with linear matter power spectrum in Eq. D.17. The wavenumber is given in units of inverse exclusion radius $1/r_e$.

The above results motivate the following definitions of the one- and two-halo power spectra

$$P_{1h}(k) = \frac{1}{\bar{\rho}_{\text{LM}} a^3} \int dm \frac{df}{d \ln m} |\mathcal{F}(k|m)|^2 - C \left| \int d \ln m \frac{df}{d \ln m} b(m) \mathcal{F}(k|m) \right|^2 \quad (\text{D.28a})$$

$$P_{2h}(k) = \left| \int d \ln m \frac{df}{d \ln m} b(m) \mathcal{F}(k|m) \right|^2 \tilde{P}_{hh}(k). \quad (\text{D.28b})$$

As mentioned above, we focus on the two-halo contribution that dominates at late times, thereby avoiding the difficulties of enforcing the proper cancellations in Eq. D.28a which is irrelevant at late times anyway. In Fig. 14 we show the result of evaluating Eq. D.28b using $\mathcal{F}$ for a spherical top hat density profile. This power spectrum roughly corresponds to $T_{\text{RH}} = 5$ MeV (i.e., we chose $A = 2\pi^2 A_s / k_{\text{RH}}^4$ with k_{RH} from Eq. 7.1), but does not include the logarithmic growth of perturbations on large scales. However, it illustrates the two main features of this result, which are apparent by comparing it with the original linear PS (dotted line) and the PS without halo exclusion (dashed line). First, we see that on scales smaller than $1/k_c$ (the left vertical dotted line) the expulsion of matter due to reheating flattens the power spectrum. On yet smaller scales, below the exclusion radius r_e (the right vertical dotted line), the PS decreases as $1/k$. The plateau feature arises because $r_e k_c < 1$, i.e. EHs grow to be much larger than their original size $\sim r_e$ during adiabatic expansion and subsequent free expansion.

References

- [1] B.D. Fields, K.A. Olive, T.-H. Yeh and C. Young, *Big-Bang Nucleosynthesis After Planck*, *JCAP* **03** (2020) 010 [[1912.01132](#)].
- [2] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [[1807.06209](#)].
- [3] ACT collaboration, *The Atacama Cosmology Telescope: DR4 Maps and Cosmological Parameters*, [2007.07288](#).
- [4] DES collaboration, *Cosmological Constraints from Multiple Probes in the Dark Energy Survey*, *Phys. Rev. Lett.* **122** (2019) 171301 [[1811.02375](#)].
- [5] DES collaboration, *Dark Energy Survey Year 1 Results: Cosmological constraints from cluster abundances and weak lensing*, *Phys. Rev. D* **102** (2020) 023509 [[2002.11124](#)].

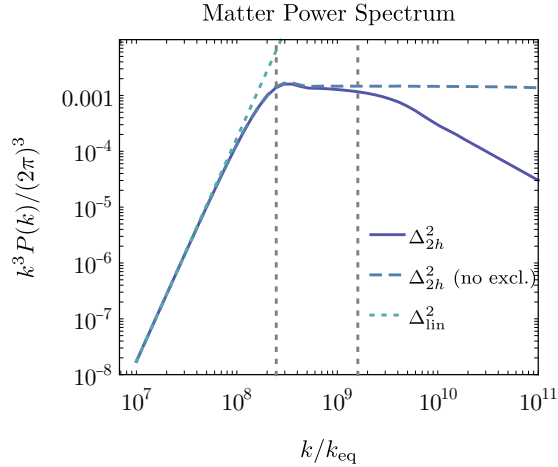


Figure 14. Matter power spectrum after the end of EMD in a toy model given by Eq. D.17 with the normalization A corresponding to $T_{\text{RH}} \approx 5$ MeV. The solid line shows the dimensionless power spectrum including both the explosive evaporation of early halos at the end of EMD, and the halo exclusion effect discussed in the text. The dashed line neglects the impact of halo exclusion, while the dotted line is the original linear power spectrum. The cutoff scale λ is taken to be a factor 10^{-3} smaller than the exclusion scale r_e , indicated by the right-most vertical dotted line. The left-most vertical dotted line indicates the expulsion cutoff k_c from Eq. 7.4. Only the two-halo contribution to the matter power spectrum is shown since the one-halo term becomes negligible well after reheating.

- [6] DES, SPT collaboration, *Cosmological Constraints from DES Y1 Cluster Abundances and SPT Multi-wavelength data*, [2010.13800](#).
- [7] KiDS collaboration, *KiDS-1000 Cosmology: constraints beyond flat Λ CDM*, [2010.16416](#).
- [8] M. Kawasaki, K. Kohri and N. Sugiyama, *MeV scale reheating temperature and thermalization of neutrino background*, *Phys. Rev. D* **62** (2000) 023506 [[astro-ph/0002127](#)].
- [9] S. Hannestad, *What is the lowest possible reheating temperature?*, *Phys. Rev. D* **70** (2004) 043506 [[astro-ph/0403291](#)].
- [10] P. de Salas, M. Lattanzi, G. Mangano, G. Miele, S. Pastor and O. Pisanti, *Bounds on very low reheating scenarios after Planck*, *Phys. Rev. D* **92** (2015) 123534 [[1511.00672](#)].
- [11] T. Hasegawa, N. Hiroshima, K. Kohri, R.S. Hansen, T. Tram and S. Hannestad, *MeV-scale reheating temperature and thermalization of oscillating neutrinos by radiative and hadronic decays of massive particles*, *JCAP* **12** (2019) 012 [[1908.10189](#)].
- [12] K. Jedamzik, M. Lemoine and J. Martin, *Collapse of Small-Scale Density Perturbations during Preheating in Single Field Inflation*, *JCAP* **09** (2010) 034 [[1002.3039](#)].
- [13] R. Easther, R. Flauger and J.B. Gilmore, *Delayed Reheating and the Breakdown of Coherent Oscillations*, *JCAP* **04** (2011) 027 [[1003.3011](#)].
- [14] O.K. Kalashnikov and M.Y. Khlopov, *ON THE POSSIBILITY OF CHECKING THE COSMOLOGY OF ASYMPTOTICALLY FREE $SU(5)$ THEORY*, *Phys. Lett. B* **127** (1983) 407.
- [15] G. Coughlan, W. Fischler, E.W. Kolb, S. Raby and G.G. Ross, *Cosmological Problems for the Polonyi Potential*, *Phys. Lett. B* **131** (1983) 59.
- [16] M. Khlopov, B.A. Malomed and I.B. Zeldovich, *Gravitational instability of scalar fields and formation of primordial black holes*, *Mon. Not. Roy. Astron. Soc.* **215** (1985) 575.
- [17] T. Banks, D.B. Kaplan and A.E. Nelson, *Cosmological implications of dynamical supersymmetry breaking*, *Phys. Rev. D* **49** (1994) 779 [[hep-ph/9308292](#)].
- [18] B. de Carlos, J. Casas, F. Quevedo and E. Roulet, *Model independent properties and cosmological*

- implications of the dilaton and moduli sectors of 4-d strings, *Phys. Lett. B* **318** (1993) 447 [[hep-ph/9308325](#)].
- [19] T. Banks, M. Dine and M. Graesser, *Supersymmetry, axions and cosmology*, *Phys. Rev. D* **68** (2003) 075011 [[hep-ph/0210256](#)].
 - [20] Y. Zhang, *Long-lived Light Mediator to Dark Matter and Primordial Small Scale Spectrum*, *JCAP* **05** (2015) 008 [[1502.06983](#)].
 - [21] A. Berlin, D. Hooper and G. Krnjaic, *PeV-Scale Dark Matter as a Thermal Relic of a Decoupled Sector*, *Phys. Lett. B* **760** (2016) 106 [[1602.08490](#)].
 - [22] A. Berlin, D. Hooper and G. Krnjaic, *Thermal Dark Matter From A Highly Decoupled Sector*, *Phys. Rev. D* **94** (2016) 095019 [[1609.02555](#)].
 - [23] J.A. Dror, E. Kuflik and W.H. Ng, *Codecaying Dark Matter*, *Phys. Rev. Lett.* **117** (2016) 211801 [[1607.03110](#)].
 - [24] T. Banks and M. Dine, *The Cosmology of string theoretic axions*, *Nucl. Phys. B* **505** (1997) 445 [[hep-th/9608197](#)].
 - [25] T. Moroi and L. Randall, *Wino cold dark matter from anomaly mediated SUSY breaking*, *Nucl. Phys. B* **570** (2000) 455 [[hep-ph/9906527](#)].
 - [26] G.F. Giudice, E.W. Kolb and A. Riotto, *Largest temperature of the radiation era and its cosmological implications*, *Phys. Rev. D* **64** (2001) 023508 [[hep-ph/0005123](#)].
 - [27] B.S. Acharya, P. Kumar, K. Bobkov, G. Kane, J. Shao and S. Watson, *Non-thermal Dark Matter and the Moduli Problem in String Frameworks*, *JHEP* **06** (2008) 064 [[0804.0863](#)].
 - [28] L. Visinelli and P. Gondolo, *Axion cold dark matter in non-standard cosmologies*, *Phys. Rev. D* **81** (2010) 063508 [[0912.0015](#)].
 - [29] M. Bose, M. Dine and P. Draper, *Moduli or Not*, *Phys. Rev. D* **88** (2013) 023533 [[1305.1066](#)].
 - [30] J. Fan and M. Reece, *In Wino Veritas? Indirect Searches Shed Light on Neutralino Dark Matter*, *JHEP* **10** (2013) 124 [[1307.4400](#)].
 - [31] N. Blinov, J. Kozaczuk, A. Menon and D.E. Morrissey, *Confronting the moduli-induced lightest-superpartner problem*, *Phys. Rev. D* **91** (2015) 035026 [[1409.1222](#)].
 - [32] N. Blinov, M.J. Dolan, P. Draper and J. Kozaczuk, *Dark matter targets for axionlike particle searches*, *Phys. Rev. D* **100** (2019) 015049 [[1905.06952](#)].
 - [33] A.L. Erickcek and K. Sigurdson, *Reheating Effects in the Matter Power Spectrum and Implications for Substructure*, *Phys. Rev. D* **84** (2011) 083503 [[1106.0536](#)].
 - [34] G. Barenboim and J. Rasero, *Structure Formation during an early period of matter domination*, *JHEP* **04** (2014) 138 [[1311.4034](#)].
 - [35] J. Fan, O. Özsoy and S. Watson, *Nonthermal histories and implications for structure formation*, *Phys. Rev. D* **90** (2014) 043536 [[1405.7373](#)].
 - [36] J.A. Dror, E. Kuflik, B. Melcher and S. Watson, *Concentrated dark matter: Enhanced small-scale structure from codecaying dark matter*, *Phys. Rev. D* **97** (2018) 063524 [[1711.04773](#)].
 - [37] A.E. Nelson and H. Xiao, *Axion Cosmology with Early Matter Domination*, *Phys. Rev. D* **98** (2018) 063516 [[1807.07176](#)].
 - [38] L. Visinelli and J. Redondo, *Axion Miniclusters in Modified Cosmological Histories*, *Phys. Rev. D* **101** (2020) 023008 [[1808.01879](#)].
 - [39] N. Blinov, M.J. Dolan and P. Draper, *Imprints of the Early Universe on Axion Dark Matter Substructure*, *Phys. Rev. D* **101** (2020) 035002 [[1911.07853](#)].
 - [40] A.L. Erickcek, P. Ralegankar and J. Shelton, *Cannibal domination and the matter power spectrum*, *Phys. Rev. D* **103** (2021) 103508 [[2008.04311](#)].
 - [41] A.L. Erickcek, P. Ralegankar and J. Shelton, *Cannibalism’s lingering imprint on the matter power spectrum*, **2106.09041**.

- [42] C. Blanco, M.S. Delos, A.L. Erickcek and D. Hooper, *Annihilation Signatures of Hidden Sector Dark Matter Within Early-Forming Microhalos*, [1906.00010](#).
- [43] W.H. Press and P. Schechter, *Formation of galaxies and clusters of galaxies by selfsimilar gravitational condensation*, *Astrophys. J.* **187** (1974) 425.
- [44] E. Neistein and A. Dekel, *Merger Rates of Dark-Matter Haloes*, *Mon. Not. Roy. Astron. Soc.* **388** (2008) 1792 [[0802.0198](#)].
- [45] B. Diemer and M. Joyce, *An accurate physical model for halo concentrations*, *Astrophys. J.* **871** (2019) 168 [[1809.07326](#)].
- [46] P. Meszaros, *The behaviour of point masses in an expanding cosmological substratum*, *Astron. Astrophys.* **37** (1974) 225.
- [47] W. Hu and N. Sugiyama, *Small scale cosmological perturbations: An Analytic approach*, *Astrophys. J.* **471** (1996) 542 [[astro-ph/9510117](#)].
- [48] J.F. Navarro, C.S. Frenk and S.D.M. White, *The structure of cold dark matter halos*, *The Astrophysical Journal* **462** (1996) 563.
- [49] H.C. Plummer, *On the problem of distribution in globular star clusters: (plate 8.)*, *Monthly Notices of the Royal Astronomical Society* **71** (1911) 460.
- [50] R.J. Scherrer and E. Bertschinger, *Statistics of Primordial Density Perturbations from Discrete Seed Masses*, *ApJ* **381** (1991) 349.
- [51] A. Cooray and R.K. Sheth, *Halo Models of Large Scale Structure*, *Phys. Rept.* **372** (2002) 1 [[astro-ph/0206508](#)].
- [52] N. Blinov, M.J. Dolan, P. Draper and J. Shelton, *Dark Matter Microhalos From Simplified Models*, [2102.05070](#).
- [53] A. Lewis, A. Challinor and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *ApJ* **538** (2000) 473 [[astro-ph/9911177](#)].
- [54] D. Blas, J. Lesgourgues and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes*, *J. Cosmology Astropart. Phys.* **2011** (2011) 034 [[1104.2933](#)].
- [55] D.J. Eisenstein and W. Hu, *Power spectra for cold dark matter and its variants*, *Astrophys. J.* **511** (1997) 5 [[astro-ph/9710252](#)].
- [56] R. Garcia, E. Rozo, M.R. Becker and S. More, *A Redefinition of the Halo Boundary Leads to a Simple yet Accurate Halo Model of Large Scale Structure*, [2006.12751](#).
- [57] A. Loeb and M. Zaldarriaga, *The Small-scale power spectrum of cold dark matter*, *Phys. Rev. D* **71** (2005) 103520 [[astro-ph/0504112](#)].
- [58] A.M. Green, S. Hofmann and D.J. Schwarz, *The First wimpy halos*, *JCAP* **08** (2005) 003 [[astro-ph/0503387](#)].
- [59] G.B. Gelmini and P. Gondolo, *Neutralino with the right cold dark matter abundance in (almost) any supersymmetric model*, *Phys. Rev. D* **74** (2006) 023510 [[hep-ph/0602230](#)].
- [60] G. Arcadi and P. Ullio, *Accurate estimate of the relic density and the kinetic decoupling in non-thermal dark matter models*, *Phys. Rev. D* **84** (2011) 043520 [[1104.3591](#)].
- [61] K. Redmond, A. Trezza and A.L. Erickcek, *Growth of Dark Matter Perturbations during Kination*, *Phys. Rev. D* **98** (2018) 063504 [[1807.01327](#)].
- [62] V.S.H. Lee, S.R. Taylor, T. Trickle and K.M. Zurek, *Bayesian Forecasts for Dark Matter Substructure Searches with Mock Pulsar Timing Data*, [2104.05717](#).
- [63] L. Dai and J. Miralda-Escudé, *Gravitational Lensing Signatures of Axion Dark Matter Minihalos in Highly Magnified Stars*, *Astron. J.* **159** (2020) 49 [[1908.01773](#)].
- [64] A. Nusser, *Selfsimilar spherical collapse with non-radial motions*, *Mon. Not. Roy. Astron. Soc.* **325** (2001) 1397 [[astro-ph/0008217](#)].
- [65] Y. Lu, H. Mo, N. Katz and M.D. Weinberg, *On the origin of cold dark matter halo density profiles*, *Mon. Not. Roy. Astron. Soc.* **368** (2006) 1931 [[astro-ph/0508624](#)].

- [66] C.-P. Ma and E. Bertschinger, *Cosmological perturbation theory in the synchronous and conformal Newtonian gauges*, *Astrophys. J.* **455** (1995) 7 [[astro-ph/9506072](#)].
- [67] U. Seljak, *Analytic model for galaxy and dark matter clustering*, *Mon. Not. Roy. Astron. Soc.* **318** (2000) 203 [[astro-ph/0001493](#)].
- [68] R.E. Smith, V. Desjacques and L. Marian, *Nonlinear clustering in models with primordial non-Gaussianity: The halo model approach*, *Phys. Rev. D* **83** (2011) 043526 [[1009.5085](#)].
- [69] T. Baldauf, U. Seljak, R.E. Smith, N. Hamaus and V. Desjacques, *Halo stochasticity from exclusion and nonlinear clustering*, *Phys. Rev. D* **88** (2013) 083507 [[1305.2917](#)].
- [70] P.J.E. Peebles, *The large-scale structure of the universe* (1980).
- [71] P.J.E. Peebles and E.J. Groth, *An integral constraint for the evolution of the galaxy two-point correlation function.*, *A&A* **53** (1976) 131.
- [72] R.E. Smith, R. Scoccimarro and R.K. Sheth, *The Scale Dependence of Halo and Galaxy Bias: Effects in Real Space*, *Phys. Rev. D* **75** (2007) 063512 [[astro-ph/0609547](#)].