

Structure formation and microlensing with axion miniclusters

Malcolm Fairbairn,^{1,*} David J. E. Marsh,^{1,5,†} Jérémie Quevillon,^{1,2,‡} and Simon Rozier^{3,4,§}

¹King's College London, Strand, London WC2R 2LS, United Kingdom

²Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, 53 Avenue des Martyrs, F-38026 Grenoble, France

³École Polytechnique, Université Paris-Saclay, F-91128 Palaiseau, France

⁴Sorbonne Universités, UPMC Univ Paris 6, CNRS, UMR 7095, 75005 Paris, France and Institut d'Astrophysique de Paris—UMR 7095 du CNRS, Université Pierre & Marie Curie, 98^{bis} Boulevard Arago, F-75014 Paris, France

⁵Institut für Astrophysik, Georg-August Universität, Göttingen, Germany



(Received 4 August 2017; published 6 April 2018)

If the symmetry breaking responsible for axion dark matter production occurs during the radiation-dominated epoch in the early Universe, then this produces large amplitude perturbations that collapse into dense objects known as axion miniclusters. The characteristic minicluster mass, M_0 , is set by the mass inside the horizon when axion oscillations begin. For the QCD axion $M_0 \sim 10^{-10} M_\odot$, however, for an axionlike particle, M_0 can approach M_\odot or higher. Using the Press-Schechter formalism we compute the mass function of halos formed by hierarchical structure formation from these seeds. We compute the concentrations and collapse times of these halos and show that they can grow to be as massive as $10^6 M_0$. Within the halos, miniclusters likely remain tightly bound, and we compute their gravitational microlensing signal taking the fraction of axion dark matter collapsed into miniclusters, f_{MC} , as a free parameter. A large value of f_{MC} severely weakens constraints on axion scenarios from direct detection experiments. We take into account the non-Gaussian distribution of sizes of miniclusters and determine how this affects the number of microlensing events. We develop the tools to consider microlensing by an extended mass function of nonpointlike objects, and we use microlensing data to place the first observational constraints on f_{MC} . This opens a new window for the potential discovery of the axion.

DOI: 10.1103/PhysRevD.97.083502

I. INTRODUCTION

Models of particle dark matter (DM) can be broadly classified into two types: thermal and nonthermal. The prototypical thermal candidate is the weakly interacting massive particle [1,2]. The prototypical nonthermal candidate is the axion [3,4].

In this paper, we explore astrophysical implications and constraints on axion DM in the so-called “miniclusters” scenario [5]. In this scenario, dense lumps of axion DM form from the dynamics of symmetry breaking which leads to the axion production in the first place. Both the mass and the mass function of the miniclusters are determined by the axion particle mass, m_a , and in some cases miniclusters can be massive enough, dense enough, and abundant enough to impact astrophysical observables such as gravitational microlensing. Thus, searches for axion miniclusters are related to searches for nonparticle DM candidates including

massive compact halo objects (MACHOs) [6,7] and primordial black holes (PBHs) [8–10].

The axion is a pseudo-Nambu-Goldstone boson of a spontaneously broken global $U(1)$ symmetry, known as a Peccei-Quinn (PQ) symmetry [11]. PQ symmetry breaking occurs when the temperature of the Universe drops below the symmetry breaking scale f_a . The cosmology of the axion, and consequent phenomenology of axion DM, is determined by the cosmic epoch during which symmetry breaking occurs. Miniclusters are formed if the PQ symmetry is broken after smooth cosmic initial conditions are established (we refer from now on specifically to inflation [12], but the distinction is not important). The minicluster mass is determined by the axion Compton wavelength, with larger minicluster masses for lighter axions.

The cosmology and astrophysics of miniclusters comprised of the QCD axion [13–18] was explored in depth in a series of seminal papers by Kolb and Tkachev in the early 1990s [19–22]. At that time, microlensing observations were not yet precise enough to constrain QCD axion miniclusters. Modern observations are much improved, and here we perform the necessary improved theoretical calculations of

* malcolm.fairbairn@kcl.ac.uk

† david.marsh@kcl.ac.uk

‡ jeremie.quevillon@lpsc.in2p3.fr

§ rozier@iap.fr

the event rate to compare to the data. Furthermore, interest has been growing in the broader class of axion DM models, or “axionlike particles,” inspired in part by the theoretical observation that light ($m_a \ll 1$ eV) axions are abundant in the string theory landscape [23,24]. It is thus timely to reconsider the work of Kolb and Tkachev beyond the QCD axion. We take on this task, following also the work of Refs. [25–27].

The fraction of axion DM bound up in miniclusters is f_{MC} , and it is not known *a priori* from theoretical calculations. If $f_{\text{MC}} \approx 1$ then axion DM direct detection is severely limited, as encounters with a minicluster in our own Galaxy will be exceedingly rare. This has profound implications for axion direct detection. If $f_{\text{MC}} \approx 1$, then direct axion detection with ADMX [28], MADMAX [29,30], or the myriad other proposed experiments targeting the QCD axion in the $m_a \sim 1$ μeV mass range would be much more difficult. Null results could erroneously be interpreted as excluding the axion, when in fact it was just “hiding” in miniclusters. On the other hand experiments like ARIADNE [31], which detect axions via the forces they mediate [32], could still detect axions even if $f_{\text{MC}} = 1$ or axions are not the DM. In practice even if $f_{\text{MC}} \approx 1$ initially, tidal disruption will affect a few percent of miniclusters allowing for some prospect of DM direct detection, or even a measurement of f_{MC} in the laboratory [33,34]. In the present study we take f_{MC} as a free parameter that, if sufficiently constrained by observation, could be used to rule out entire classes of models for axion production. Our method proposes to measure the minicluster fraction using gravitational microlensing.

In the course of considering this signal, we address issues of structure formation with miniclusters and present a series of possible models. A new consequence of this investigation is the computation of the mass function of minicluster halos (which we term “MCHs”). Depending on the merging and tidal stripping of miniclusters, the MCH mass function may or may not be the appropriate mass function to consider for microlensing. There may be, however, other observational consequences of the existence of MCHs for which the mass function will be an important quantity.

We begin in Sec. II, where we present some introductory basics on miniclusters. There has been very little study in the literature on the subsequent gravitational evolution of axion miniclusters after their formation (although see [25]). In Sec. III we therefore present a new computation of the MCH mass function following the Press-Schechter [35] formalism. The form of the mass function is seen to arise simply from basic physical principles, and can be easily parametrized.

Miniclusters are extended objects and cannot be considered as pointlike lenses (radius larger than the Einstein radius). We discuss minicluster density profiles in Sec. IV. Section V presents tools to compute the lensing signal from a mass function of nonpointlike lenses.

We apply our minicluster lensing methodology to the historical EROS survey [36] and to the very recent Subaru Hyper Suprime Cam observations [37] in Sec. VI. We conclude in Sec. VII.

Since miniclusters can only be produced if the PQ symmetry is broken after inflation, the axion parameters are constrained by the precise prediction for the axion DM relic density in this scenario. In particular, the relic density can be too large if the axion domain wall number $N_{\text{DW}} > 1$, due to the late decay of domain walls. Thus, in the standard QCD axion model, miniclusters are a generic prediction only of the Kim-Shifman-Vainshtein-Zakharov (“KSVZ”)-type models, with the scenario in the “Dine-Fischler-Srednicki-Zhitnitsky”-type models being somewhat more complicated.¹ In this paper we consider only models with $N_{\text{DW}} = 1$, where the minicluster scenario is simplest. Appendix A collects results on the axion relic density, early time cosmology and thermal history, and we determine the range of axion masses and decay constants for which the minicluster scenario can occur.

Finding the relic density in general requires many complex computational steps: lattice QCD for the temperature dependence of the axion mass, and lattice field theory for the classical evolution of the axion field. These calculations have only been performed in the literature for the QCD axion. Our relic density computation uses analytic approximations to allow us to treat axion like particle (ALPs), and to account for uncertainties in the QCD calculation in a parametrized way. Our final constraints account for the uncertainties by allowing for a window in the relation between minicluster mass and axion mass for the QCD axion.

Appendix B presents the theoretical modeling of the MCH mass function, and some analytic results. Appendix C discusses how non-Gaussianity of the axion initial conditions affects the MCH mass function. Appendix D discusses how formation of “axion stars” might modify our lensing results in the case of axionlike particles with temperature-independent mass.

We use Planck (2015) [39] cosmological parameters $h = 0.67$, $\Omega_m = 0.32$, $\Omega_c h^2 = 0.12$, and $z_{\text{eq}} = 3402$, and for particle physics quantities we use natural units where $c = \hbar = 1$. Throughout this work we consider a standard thermal history of the Universe after the inflationary epoch.

II. AXION MINICLUSTERS

This section gives the briefest outline of the minicluster scenario to establish some important language and physical scales. The main results are given in Fig. 1 and Eq. (3).

There are two energy scales that define the cosmological evolution of the axion field: the decay constant, f_a , and the

¹A comprehensive discussion of the variations on these general themes can be found in Ref. [38].

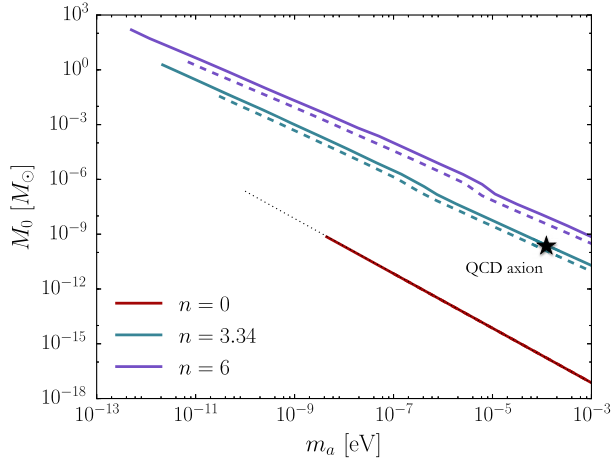


FIG. 1. The characteristic minicluster mass. We plot M_0 , as a function of the axion mass, m_a , for different temperature evolutions of the axion mass parametrized by index n . Solid lines show the most realistic assumptions about the relic density, while dashed lines relax those assumptions (see Appendix A). When the axion mass is temperature independent ($n = 0$), the two scenarios are equivalent for minicluster mass. The thin dotted line shows the approximation Eq. (3). Lines terminate at a lower bound on m_a set by the relic abundance.

mass, m_a . These two energy scales determine the two most important epochs in the life of a young axion.

The axion is the angular degree of freedom of a complex scalar field, φ , with a global $U(1)$ symmetry that is spontaneously broken by the potential:

$$V(\varphi) = \frac{\lambda}{4!} \left(|\varphi|^2 - \frac{f_a^2}{2} \right)^2. \quad (1)$$

Spontaneous symmetry breaking (SSB) occurs when the temperature of the Universe cools to $T \lesssim f_a$.² After SSB, the complex field is given by $\varphi = (f_a/\sqrt{2})e^{i\phi/f_a}$, with ϕ the (real) axion field.

If PQ symmetry is broken before or during inflation then the axion field takes on a uniform value across the observed Universe, with the addition of small isocurvature perturbations from the finite temperature during inflation, and density perturbations inherited from the adiabatic perturbations in the hot big bang plasma. On the other hand, if PQ symmetry is broken after inflation then topological defects and large amplitude axion field fluctuations are present on scales of order the horizon size at symmetry breaking [5,41,42]. In this case, since each horizon volume is causally disconnected, the axion field is uncorrelated across different horizon volumes and drawn from the distribution $\phi/f_a \in \mathcal{U}[-\pi, \pi]$.

²The precise critical temperature for the phase transition from thermal field theory is calculated in e.g. Ref. [40].

The Kibble mechanism [43] smoothes the axion field on the horizon scale until such a time that the axion mass becomes cosmically relevant, which defines the second important epoch in the life of the axion. At this epoch, the topological defects decay, and the axion field is left with large amplitude isocurvature fluctuations on the horizon scale.³ These fluctuations of order f_a between horizon volumes provide the initial conditions for the axion field on small scales. It is these fluctuations that subsequently evolve into gravitationally bound miniclusters. Thus, as long as f_a is low enough that SSB occurs during the normal thermal evolution of the Universe, then miniclusters are a logical possibility.

The time t_0 when the axion mass becomes significant is given by $3H(t_0) \approx m_a$, with H the Hubble rate. The axion acquires its mass due to nonperturbative effects such as instantons [45], which evolve with temperature as $m_a = m_{a,0}(T/\mu)^{-n}$. Therefore the critical time t_0 depends on the index n giving the temperature evolution. From this time onwards the axion field oscillates about its own quadratic potential minimum, and the equation of state for the background axion energy density becomes the same as that of pressureless matter [46–48]. The epoch when axion oscillations begin thus determines the axion relic density. Appendix A computes the range of (m_a, f_a) for various n for which axions provide the total DM relic density while having f_a small enough for SSB to occur after inflation.

The initial axion fluctuations laid down by SSB remain smooth up to scales of order the horizon size at t_0 . From this point on the density perturbations grow under gravity as usual, eventually collapsing into the gravitationally bound objects known as miniclusters. The total mass of axion DM contained within the horizon at time t_0 sets the characteristic minicluster mass, M_0 , given by

$$M_0 = \bar{\rho}_a \frac{4}{3} \pi \left(\frac{\pi}{\mathcal{H}(t_0)} \right)^3, \quad (2)$$

where $\bar{\rho}_a$ is the energy density in axions today, and we have used the fact that the comoving wave number associated with the horizon size at this time is $k_0 = \mathcal{H}(t_0)$ where $\mathcal{H} = aH$ is the conformal Hubble rate.

³It is important to realize that these minicluster isocurvature fluctuations seeded by the PQ phase transition are unrelated to the “axion isocurvature mode” constrained by large-scale CMB anisotropies [44]. The axion isocurvature mode is produced by vacuum fluctuations during inflation in the scenario when PQ symmetry is broken before inflation. The fluctuations are thus scale invariant and contribute to large angle anisotropies. The minicluster isocurvature mode is not scale invariant: the power dies off rapidly on scales larger than the horizon size at the QCD phase transition ($k < k_0$) and does not contribute to the observed CMB anisotropies.

In the case of a temperature-independent axion mass, and approximating the number of relativistic degrees of freedom as constant in T , a reasonable approximation to M_0 can be obtained by using $(H/H_0)^2 = \Omega_m(1+z_{\text{eq}})^{-1}a^{-4}$ to find $a(t_0)$. A “fudge factor” of two leads to a good agreement with the full numerical calculation using $g_*(T)$:

$$M_0(m_a, n=0) \approx 2.3 \times 10^{-7} M_\odot \left(\frac{m_a}{10^{-10} \text{ eV}} \right)^{-3/2} \times \left(\frac{\Omega_c h^2}{0.12} \right) \left(\frac{\Omega_m}{0.32} \right)^{-3/4} \left(\frac{1+z_{\text{eq}}}{3403} \right)^{3/4}. \quad (3)$$

This approximation is shown as the thin dotted line in Fig. 1.

In the general case, the characteristic minicluster mass M_0 is a function of the axion mass, the axion decay constant, and the index n determining the temperature evolution of the mass. Fixing the decay constant as a function of the mass from the relic density, we show $M_0(m_a, n)$ for various n in Fig. 1. The lines in Fig. 1 for axionlike particles terminate at a lower bound on m_a . For masses below this bound axions cannot produce the correct relic density while maintaining a low enough f_a to be of relevance for miniclusters. The lower bound on m_a translates into an upper bound on M_0 , which is around $10^3 M_\odot$ for $n=6$ temperature evolution.⁴

The QCD axion has a known temperature-dependent mass with $n=3.34$ from the “interacting instanton liquid” model for the QCD topological susceptibility [49], which is consistent with the results from lattice simulations ($n \approx 3.55 \pm 0.30$ [50,51]). The mass is given by $m_{a,\text{QCD}} = 6 \mu\text{eV}(10^{12} \text{ GeV}/f_a)$. For our modeling of the relic density the decay constant must be $f_a \approx 10^{10} \text{ GeV}$ leading to an axion mass $50 \mu\text{eV} \lesssim m_a \lesssim 200 \mu\text{eV}$. This implies a characteristic minicluster mass $M_0 \approx 1.8 \times 10^{-10} M_\odot$, in broad agreement with other estimates [20–22,26].

Our value of M_0 differs from the characteristic mass defined in e.g. Ref. [22], whose authors took the mass within a cubic volume of size the inverse horizon wave vector, rather than the spherical volume from the physical wavelength. This makes our definition of M_0 larger by a factor of $4\pi^4/3 \approx 130$. We believe our definition captures the symmetries of minicluster formation better, and also better represents their likely formation history (see Sec. IV). In practice the constraints on miniclusters from microlensing cover a broad range of masses and even a factor of 10^2 in M_0 does not have a large impact.

A far larger uncertainty is introduced if the axion string length parameter, ξ , rather than the horizon size, defines

⁴Constraints from the Lyman- α forest flux power spectrum [25] affect miniclusters with masses $M_0 \gtrsim 10^4 M_\odot$, which can be achieved in certain exotic scenarios for symmetry breaking [26].

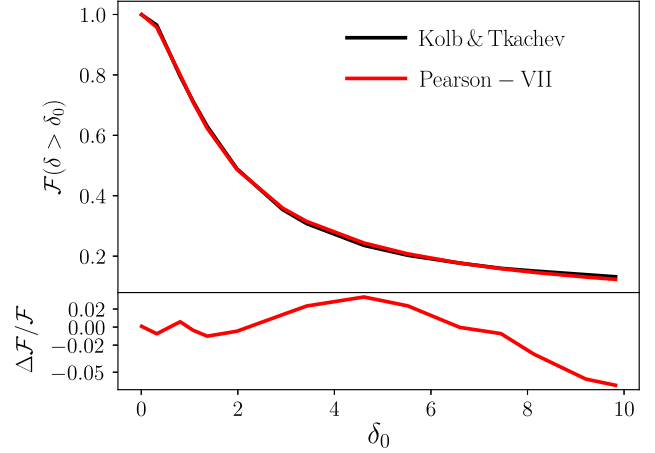


FIG. 2. Minicluster overdensity distribution. We show the cumulative mass fraction of miniclusters with overdensity parameter $\delta > \delta_0$. The black line shows the simulation results of Ref. [22], which we have fit using a Pearson-VII distribution. The overdensity distribution determines the halo concentration parameter (i.e. compactness) of miniclusters.

M_0 . This factor is historically uncertain in simulations: it is currently favored to be $\xi = 1.0 \pm 0.5$ [52], but could be much larger, with larger ξ leading to smaller M_0 . To an extent the effect of the uncertainty in ξ on M_0 is captured by the uncertainty in the relic density (see Appendix A).

The minicluster characteristic density, ρ_{MC} , is another important quantity, since it sets the typical radius of a minicluster, and thus its concentration. The characteristic density is given in terms of the initial overdensity parameter, δ , by [21]:

$$\rho_{\text{MC}} = 140\delta^3(1+\delta)\bar{\rho}_a(z_{\text{eq}}), \quad (4)$$

which can be derived using spherical collapse and understood by considering how regions of large δ collapse when $z \gg z_{\text{eq}}$. The initial condition simulations of Ref. [22] show that while the mass of miniclusters was approximately fixed to M_0 , the characteristic density showed a wide, non-Gaussian variation due to the anharmonicities in the axion potential.

Figure 2 shows the cumulative mass fraction, $\mathcal{F}(\delta > \delta_0)$, taken from digitizing Fig. 2 of Ref. [22]. The non-Gaussian distribution is well fit by a Pearson-VII-type distribution:

$$\mathcal{F}(\delta > \delta_0) = \frac{1}{[1 + (\delta_0/a_1)]^{a_2}}, \quad (5)$$

with $a_1 = 1.023$ and $a_2 = 0.462$ found by a nonlinear least squares fit.

We are not in possession of numerical simulations that would allow us to fully characterize the non-Gaussianity of the minicluster density field. In order to make some progress we assume that the non-Gaussianity in overdensity and minicluster size given by \mathcal{F} is a small-scale

phenomenon affecting only the stochastic distribution of minicluster sizes.

This approach to model the non-Gaussian effects on the density profile can be partially justified by considering that the non-Gaussianities that affect the density profile are caused by the axion self-interactions. The large-scale perturbations are of smaller amplitude (the dimensionless power spectrum falls as k^3 for small $k < k_0$). Small density perturbations mean small axion field perturbations. At small field values the axion potential is quadratic, giving free-field behavior that should be close to Gaussian. Appendix C discusses non-Gaussian corrections to the large-scale clustering of miniclusters and the minicluster halo mass function induced by the white-noise fluctuations of θ .

III. STRUCTURE FORMATION WITH MINICLUSTERS

The discussion in the previous section asserts that the initial conditions for the axion field on small scales caused by SSB lead to the collapse of objects of mass M_0 around matter-radiation equality. In the following section we consider how miniclusters, once initially formed, go on to merge into larger bound structures, which we term “minicluster halos,” or MCHs. This process has not been studied in great detail before. It deserves attention because the behavior will be quite different from cold dark matter: the initial conditions are isocurvature, structure formation begins much earlier, and the power spectrum is truncated. We address this situation by computing the standard halo mass function from linear growth of the minicluster initial conditions.

MCHs are small-scale structures; they are substructures within the larger-scale DM halos formed by the scale-invariant adiabatic initial conditions on large scales (see e.g. the combined power spectrum in Ref. [25]).

The logic of computing the mass function for the small-scale minicluster isocurvature initial conditions independently from the usual adiabatic large-scale cosmology is the following. As we will see, miniclusters collapse very early, at $z \approx z_{\text{eq}}$. Galactic halos like the Milky Way are formed from the dominant, but small amplitude, adiabatic perturbations on large scales and collapse at much lower redshifts. These galactic halos are, however, still formed of axion DM, and thus of the miniclusters formed early on. We treat these two periods of gravitational collapse independently, and we assume that the minicluster mass function established early on provides the substructure mass function on small scales within the larger galactic halos. In other words, the minicluster mass function is equivalent to the “field” mass function within the larger patches that collapse later on into galaxies. Since miniclusters collapse and “freeze out” from the expansion early on, they can be treated independently from the large scales.

In the explicit examples in this section we use a Heaviside initial power spectrum to model the effects of the Kibble mechanism and a Gaussian window function to define the mass variance. We consider other possibilities, and give analytic results, in Appendix B. The examples also use the $M_0(m_a)$ relation for an axionlike particle with T -independent mass.

We work in conformal time τ , with $dt = a d\tau$ and a the scale factor of the Friedmann-Lemaître-Robertson-Walker metric. For simplicity of presentation, we begin our numerical calculations in the matter-dominated era once all T -dependence of the axion mass and g_* can be neglected. We comment briefly on the transfer function in the radiation era.

A. Evolution of density perturbations

1. Initial conditions

We define the axion density perturbation as $\delta_a(\mathbf{r}, \tau) = [\rho_a(\mathbf{r}, \tau) - \bar{\rho}_a(\tau)]/\bar{\rho}_a(\tau)$. The power spectrum $P_{\mathbf{k}}$ is defined as the Fourier transform of the two point correlation function:

$$\xi(\mathbf{r}) = \langle \delta(\mathbf{x}) \delta(\mathbf{x} + \mathbf{r}) \rangle. \quad (6)$$

We wish to specify the value of the power spectrum at the initial time of our study τ_0 : the time when the axion field begins oscillating. At τ_0 , the Kibble mechanism assures us that the axion field is fixed to a constant value over each causal horizon, and randomly distributed (white noise) over the different horizons. Hence, we can approximate the power spectrum by a sharp- k function, cut at the typical (comoving) size of a horizon at τ_0 :

$$P_k(\tau_0) = P_0 \Theta(k_0 - k), \quad (7)$$

where $k_0 = \mathcal{H}(\tau_0)$ and $\Theta(x)$ is the Heaviside function.

We must now find the normalization, P_0 . From Eq. (6) it is easy to show that the integral of the power spectrum is (e.g. Ref. [53])

$$\int P_{\mathbf{k}}(\tau) \frac{d^3k}{(2\pi)^3} = \langle |\delta_a(\mathbf{x}, \tau)|^2 \rangle. \quad (8)$$

From our ansatz for $P_k(\tau_0)$, the left-hand side is equal to $\frac{4}{3} \pi k_0^3 P_0$. To find the value of $\int |\delta_a(\mathbf{x}, \tau_0)|^2 d^3x$, we note that $\rho_a \propto \theta^2$, where $\theta \in [-\pi, \pi]$ is the uniformly distributed axion field. Thus the mean value of $|\delta_a(\mathbf{x}, \tau_0)|^2$ is 4/5, and we find

$$P_0 = \frac{24}{5} \pi^2 k_0^{-3}. \quad (9)$$

This equation sets the initial condition on the modes of the perturbations. These initial conditions are of isocurvature type during the radiation-dominated epoch, and they

can be normalized to have $\delta_a(\tau_0) = 1$ with all other perturbations absent.⁵

2. The behavior in the radiation-dominated era

It is well known that subhorizon isocurvature perturbations undergo only a very small amount of logarithmic growth in the radiation-dominated era, and that the transfer function up to matter-radiation equality is approximately k -independent for large k [57]. The amplitude, i.e. the actual amount of growth in the radiation era, is somewhat cosmology dependent. For our cosmological parameters using the analytic transfer function of Ref. [57] we find

$$\frac{\delta(z_{\text{eq}})}{\delta_i} \approx 1. \quad (10)$$

This approach is also taken in e.g. Refs. [5,25,26] for the case of miniclusters. We have verified by full numerical solution of the Boltzmann equations that the transfer function indeed takes the form stated, and we use this result in the following.

3. The behavior in the matter-dominated era

We consider the subhorizon limit, assuming that most of the sizes we are interested in are subhorizon sized at matter-radiation equality. The equation of motion for the axion overdensity in an axion-dominated Universe at times $\tau \gg \tau_0$ is [4]

$$\delta_a'' + \frac{a'}{a} \delta_a' + (k^2 c_s^2 - 4\pi G a^2 \bar{\rho}_a) \delta_a = 0, \quad (11)$$

where primes denote derivatives with respect to conformal time, and $c_s^2 \approx k^2/4m_a^2 a^2$ is the axion effective sound speed. The sound speed leads to a Jeans scale, which balances pressure and gravity [58]:

$$k_J = (16\pi G a \rho_{a0})^{1/4} m_a^{1/2}, \\ = 66.5 \times 10^6 a^{1/4} \left(\frac{\Omega_m h^2}{0.12} \right)^{1/4} \left(\frac{m_a}{10^{-10} \text{ eV}} \right)^{1/2} \text{ Mpc}^{-1}. \quad (12)$$

The Jeans scale sorts the modes into three categories:

- (i) $k < k_J(\tau_{\text{eq}})$: These modes are already under the Jeans mode at matter-radiation equality. Their behavior is the usual growing/decaying as soon as matter-radiation equality is reached, and it matches the transfer function used in Ref. [25].

⁵This is sufficient accuracy for our purposes of computing the mass function in the matter era. For greater accuracy in the radiation era at $\tau > \tau_0$ for modes near the Jeans scale see the early-time expansion of Ref. [54]. For the complete treatment including $\tau < \tau_0$ see Refs. [55,56].

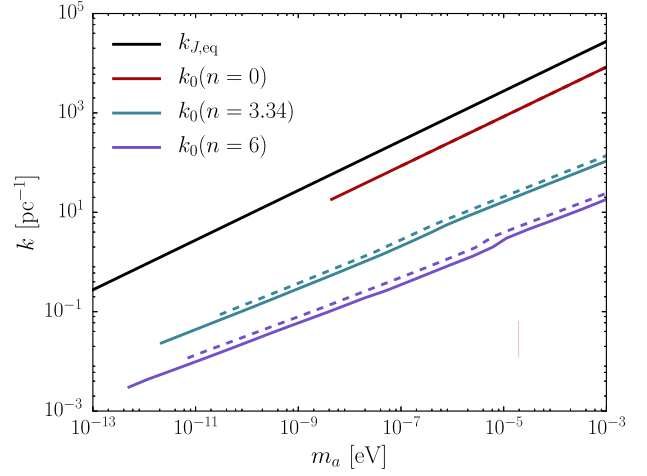


FIG. 3. Cutoffs in the power spectrum. We plot the wave numbers that cut off the axion power spectrum, k_0 and $k_{J,\text{eq}}$, as a function of axion mass. For k_0 , this depends on the temperature evolution of the mass as parametrized by n and on details of the relic density computation. The Jeans wave number is fixed only by the zero-temperature mass. We notice that we always have $k_{J,\text{eq}} > k_0$, which justifies that the Jeans scale can be ignored to a first approximation.

- (ii) $k_J(\tau_{\text{eq}}) < k < k_J(\tau_{\text{today}})$: These modes are bigger (physically smaller) than the Jeans mode at matter-radiation equality, and as k_J increases they cross the Jeans scale. The behavior of these modes is to oscillate at the beginning of the matter-dominated era, and then to follow the usual growing/decaying mode.
- (iii) $k > k_J(\tau_{\text{today}})$: These modes are still today physically smaller (have larger k) than the Jeans mode, and still follow the oscillating behavior.

The Jeans scale will affect the growth of linear perturbations when $k_J < k_0$. However, as can be seen in Fig. 3, we notice that we always have $k_{J,\text{eq}} > k_0$, which justifies that the Jeans scale can be ignored in our case to a first approximation when computing the variance and the evolution of the mass function.

4. The rms mass fluctuation

Let us define the variance of density fluctuations on a given scale by filtering the power spectrum with a window function of physical radius R :

$$\sigma^2(R) = \int \frac{dk}{k} \Delta^2(k) |W(kR)|^2, \\ = \int \frac{dk}{k} \frac{k^3 P(k)}{2\pi^2} |W(kR)|^2. \quad (13)$$

For ease of numerical integrals, we take the window function to be a real-space Gaussian, which in Fourier space gives

$$W(kR) = e^{-k^2 R^2/2}. \quad (14)$$

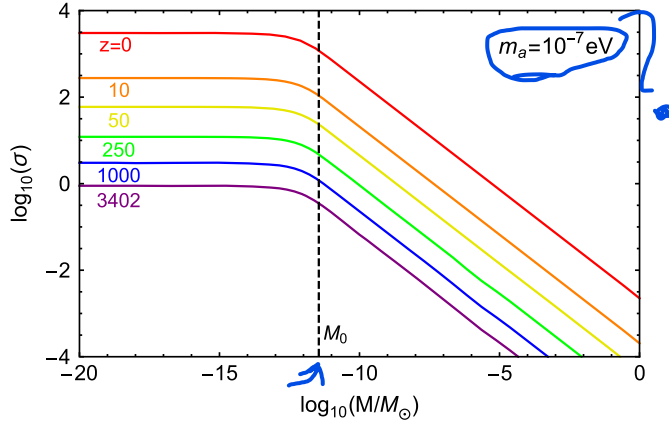


FIG. 4. The rms density fluctuation. The curves show a $\sigma(M) \propto M^{-1/2}$ above M_0 , which matches the rms of white noise, as in Ref. [5]. The constant behavior at low masses is due to the smoothed out axion field below the horizon at the phase transition, the effect of the Jeans scale cannot be seen in this diagram. The growth of the rms mass fluctuation through time is due to the linear growing mode: $\delta \propto a$ above the Jeans scale.

We define the mass to be that contained within the comoving volume of the radius R appropriate to the Gaussian window function, such that

$$M = (2\pi)^{3/2} \bar{\rho}_{a,0} R^3. \quad (15)$$

Thus we have the mass variance, $(\delta M/M)^2 \equiv \sigma^2(M)$ (see also Ref. [59]). The time dependence of $\sigma(M)$ comes from δ_a , so that

$$\begin{aligned} \sigma^2(M, \tau) &= \int \frac{k^3 P(k, \tau)}{2\pi^2} |W(k, M)|^2 \frac{dk}{k} \\ &= \int P(k, \tau_0) \mathcal{T}^2(k, \tau) |W(k, M)|^2 \frac{k^2 dk}{2\pi^2}, \end{aligned} \quad (16)$$

where we have defined the transfer function, $\mathcal{T}(k, \tau)$, to contain the evolution of the power spectrum. The above expression clearly separates the dependence of the mass fluctuation on the initial conditions, which depend on k_0 , the dynamics, given by the transfer function, and the window function which we use to define the mass scale.

Our normalization for the initial power spectrum sets the variance on the scale $R_0 = \pi/k_0$ as

$$\sigma^2(R_0) = \frac{P_0}{2\pi^2} \int_0^{k_0} k^2 \exp\left[-\frac{\pi^2 k^2}{k_0^2}\right] dk, \quad (17)$$

giving $\sigma(R_0) \approx 0.18$, which differs by a factor of two from the normalization of Ref. [25], whose authors considered quantum fluctuations of θ and found that the rms density fluctuation on the scale k_0 is $\sigma \approx 1/(2\sqrt{2}) \approx 0.35$.

Figure 4 shows the behavior of the rms mass fluctuation at several moments in the matter-dominated era.

It shows the white-noise behavior above M_0 , with $\sigma(M) \propto M^{-1/2}$, and a smooth convergence under M_0 , due to the sharp- k shape of the initial power spectrum. For this particular axion mass, the action of the Jeans scale is invisible, since the shape of $\sigma(M)$ is constant in time.

B. The MCH mass function

We consider the formation of gravitationally bound structures from linear density perturbations using the analytic Press-Schechter formalism [35]. We use the original formalism, rather than modern updates such as the Sheth-Tormen [60] as we are interested in only an approximate description of MCH formation and the mass function is subject to a number of theoretical uncertainties, which we discuss.

The quantity δ_c is the critical overdensity threshold for gravitational collapse and plays a key role in the Press-Schechter formalism. In spherical collapse of cold dark matter it is given by $\delta_c \approx 1.686$, and it is scale independent. For every point in space, the probability to have $\delta > \delta_c$ using the filtered version of δ by W_M is

$$p(\delta > \delta_c | W_M) = \frac{1}{2} \left(1 - \operatorname{erf} \left(\frac{\delta_c}{\sqrt{2}\sigma(M)} \right) \right) \quad (18)$$

where $\sigma(M)$ is the rms mass fluctuation that we defined. We take the point of view where δ_c is time independent, and we put the time dependence onto $\sigma(M)$. Defining $f(M)$ such that $f(M)dM$ is the comoving number density of mini-clusters in the range dM around M , the Press-Schechter formalism gives

$$\frac{M^2 f(M)}{\bar{\rho}_{a,0}} = \left| \frac{d \ln \sigma}{d \ln M} \right| \sqrt{\frac{2}{\pi}} \frac{\delta_c}{\sigma(M)} e^{-\frac{1}{2} \left(\frac{\delta_c}{\sigma(M)} \right)^2} \quad (19)$$

If $n(M)$ is the number density of structures of mass M , then the halo mass function (HMF), defined by $dn/d \ln M = M f(M)$, is the comoving number density of structures of mass M per logarithmic interval in masses.

The low mass end of the HMF is subject to large theoretical uncertainty. The density field on scales $M < M_0$ (wave numbers $k > k_0$) is non-Gaussian, and so the Press-Schechter formalism does not apply. The formalism can be applied perturbatively, or for certain special types of non-Gaussianity [61], though such a calculation is beyond the scope of this work. The Press-Schechter mass function we present for $M < M_0$ applies only to “the Gaussian part” of the small-scale density field.

Even for a Gaussian field, on small scales there are theoretical uncertainties in the mass function in the case of a truncated power spectrum such as the axion power spectrum. With a real-space window function the formula Eq. (19) predicts that structures are formed on all scales below the nonlinear scale even in the case of suppressed density perturbations. This is due to the asymptotic behavior of the variance. However, a number of considerations modify

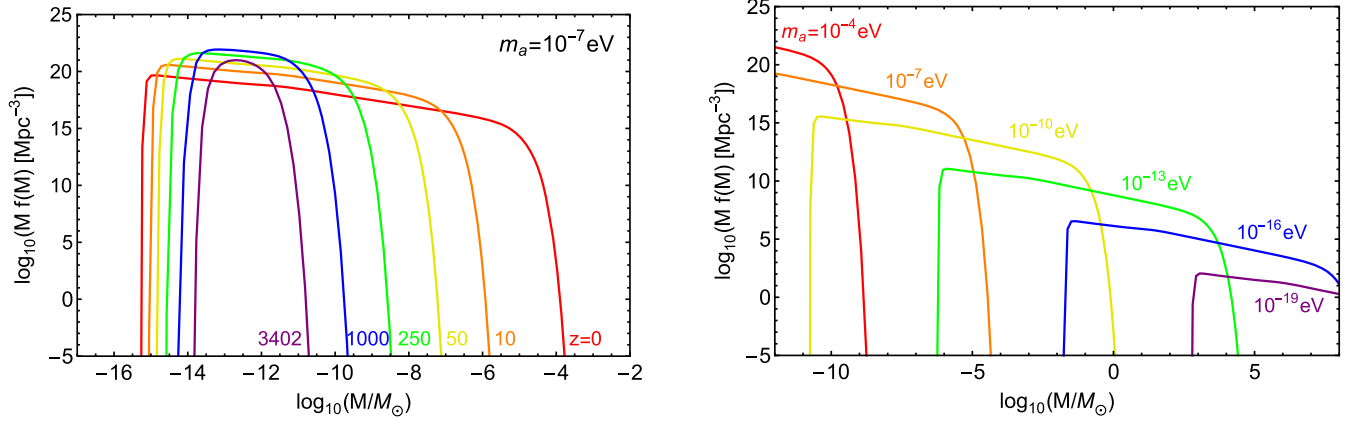


FIG. 5. The MCH mass function. Left panel: HMF as a function of time for fixed axion mass. The initial miniclusters at M_0 spread over time to form more massive objects due to hierarchical structure formation. Lighter objects are also formed as the late-time Jeans scale cutting off the mass function moves to scales smaller than M_0 . Right panel: Minicluster mass function today for various axion masses.

this prediction and predict a cutoff in the mass function for $M \lesssim M_0$. These can be summarized as

- (i) “Spurious structure”: Simulations with truncated power spectra display effects due to numerical discreteness on scales below the power spectrum cutoff [62].
- (ii) Filter dependence: With a truncated spectrum the sharp- k window function used to derive the Press-Schechter formula from the excursion set [63] predicts a cutoff in the mass function [64].
- (iii) Dynamical effects: The Jeans scale leads to a pressure on small scales that modifies the collapse barrier [65,66], the solution of the excursion set [67–69], and the formation of structure on small scales [70,71].

Appendix B presents three different models for the mass function for $M < M_0$: a fit to remove structure below M_0 , taken from the simulations of Ref. [72] (see also Ref. [73]); a sharp- k filter, with mass normalized to the Gaussian filter on large scales [64]; and a modified barrier at the Jeans scale, following the approximate implementation in Ref. [66]. We show later that, for the range of MCH masses relevant for microlensing, and in particular for the particle mass range of the QCD axion, the cutoff dependence of the mass function for $M < M_0$ has negligible effect on the observables.

For the purposes of illustration, the MCH mass function computed with the modified barrier is shown in Fig. 5. We show the mass function evolving over time for a fixed axion mass, and the mass function at $z = 0$ for a range of axion masses. The $M_0(m_a)$ relation used in these examples is for a temperature-independent axion mass ($n = 0$). The HMF is centered near M_0 , and is cut off at high and low masses. The mass function spreads over time as structure formation progresses. For lower axion masses, the HMF is centered around larger MCH masses.

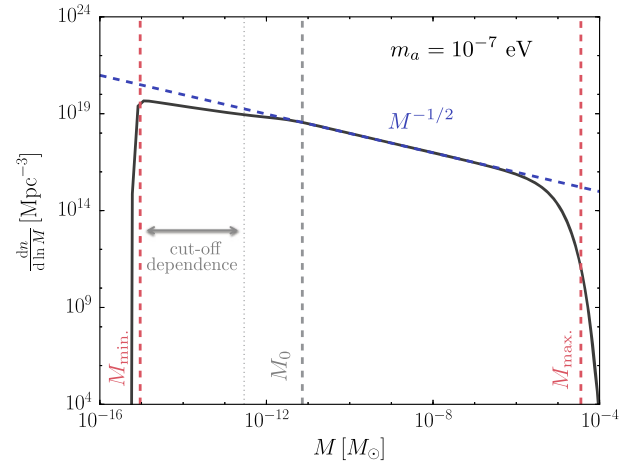


FIG. 6. Parametrization of the MCH mass function. The mass function can be well fit by two cutoffs and a single slope parameter, $M^{-1/2}$, derived from the white-noise initial conditions. For the numerical calculation in the previous subsection (solid line), the normalization is fixed to be per unit volume. For the substructure mass function, we normalize by f_{MC} (see text). For $M \lesssim M_0/25$ where the variance becomes flat, there is cutoff dependence from the barrier, window function, and non-Gaussianities. For illustration we take $m_a = 10^{-7} \text{ eV}$ and $n = 0$ to use the analytic formula for M_0 and show the cutoff from the Jeans scale only.

C. Parametrization of the mass function

The HMF for miniclusters that we have obtained can be parametrized by four quantities. We give an explicit parametrization of the HMF and show that (for arbitrary normalization) it matches our numerical calculations well. Since the parametric form is well understood, we can use this mass function to describe DM substructure within galactic halos, including the Milky Way. The substructure

mass function is normalized by the host galaxy mass and has units $[dn/dM] = [M]^{-1}$.

Our parametrized HMF is shown in Fig. 6 and compared to the numerical results from the previous subsection, for which the normalization per unit volume is fixed.

1. Cutoffs in the HMF

Cutoffs in the HMF are driven by the Gaussian term in Eq. (19), the argument of which depends on $\delta_c(M)/\sigma(M, z)$, with the possible addition of a fit to remove spurious structure at low masses discussed above. We parametrize $\sigma(M)$ for the Heaviside initial power spectrum and Gaussian window function following Fig. 4 as

$$\sigma(M < M_0/6.65) = \sigma_0 \approx 10^{3.4}, \quad (20)$$

$$\sigma(M > M_0/6.65) = \sigma_0 \left(\frac{25M}{M_0} \right)^{-1/2}. \quad (21)$$

The high mass cutoff in the mass function depends on the form of the initial power spectrum. At large $M \gg M_0$, cutoff dependence vanishes, giving a Gaussian density field suppressed as $-\delta_c^2/2\sigma^2 = -x^2/2$. Using $\delta_c = 1.686$ we solve for $x(M_{\max})$ when the HMF has dropped by 0.01 for the Heaviside cutoff initial power spectrum:

$$M_{\max}(z) \approx 4.9 \times 10^6 M_0 \times D(z)^2, \quad (22)$$

where $D(z)$ is the linear growth factor normalized to unity at $z = 0$. The coefficient of the Gaussian cutoff in the mass function is modified slightly when the nonzero bispectrum of the minicluster initial conditions is accounted for, leading to a slight increase in M_{\max} . This is a small effect, and depends on the model for non-Gaussian effects on the mass function. We do not account for it here. It is discussed in detail in Appendix C.

The maximum minicluster mass is fixed by M_0 and so depends strongly on the temperature evolution of the axion mass (see Fig. 1). MCHs of mass $M > M_0$ are formed by hierarchical structure formation from the seeds of mass $M \approx M_0$ present at matter-radiation equality. We discuss the mergers and density profiles of MCHs in Sec. IV.

We derive the low mass cutoff for the modified barrier (for the fit and Heaviside filtering cases the corresponding derivation is trivial). At small $M \ll M_0$, $\sigma(M) = \sigma_0$, and we use the asymptotic behavior of $\delta_c(M < M_J) \propto \exp[1.8(M/M_J)^{-1/2}]$, where M_J is the Jeans mass, to find when the Gaussian argument $x(M_{\min}) = 1$:

$$M_{\min}(z) \approx M_J \times \left[\frac{1.8}{7.5 + \log D(z)} \right]^2. \quad (23)$$

There is only a logarithmic dependence of the minimum minicluster mass on the growth rate, and so the minimum mass varies very little over time.

2. Slope

From the dependence of σ on the mass scale M , Eq. (21), we can deduce the dependence of the HMF on M far from either cutoff. Using Eq. (19) we have

$$\frac{dn}{d \ln M} = M f(M) \propto M^{-1/2}; \quad \text{for } (M_{\min} < M < M_{\max}). \quad (24)$$

Our parametrization sets the HMF to zero outside these boundaries.

We see in Fig. 6 that the slope of the mass function changes below $M_0/6.65$, where the variance becomes flat. The slope of the mass function in this regime is window function and cutoff dependent.

3. Normalization

The normalization of the substructure mass function is simply fixed by the total mass of the host galaxy, M_{host} . While we assume that all of the DM is composed of axions, we take the fraction collapsed into miniclusters, f_{MC} , as a free parameter. The normalization condition is

$$f_{\text{MC}} = \frac{1}{M_{\text{host}}} \int \frac{dn}{d \ln M} dM = \int \psi(M) dM, \quad (25)$$

where we have introduced the normalized mass function $\psi(M)$.

In principle, f_{MC} can be determined by numerical simulation of minicluster formation and the subsequent formation of galaxies from miniclusters. Such simulations will involve at least some modeling uncertainty. We take the alternate view that f_{MC} is a phenomenological quantity that can be used to constrain models of axion DM via astrophysical observations. Thus the mass function normalization, f_{MC} , is a free parameter in our constraints.

IV. MINICLUSTER AND MCH DENSITY PROFILES

A. Hierarchical structure formation

The scale of miniclusters is fixed by $k_0 \ll k_J$ and so we expect density profiles to be well described by cold dark matter (CDM). Density profiles formed by hierarchical structure formation of CDM are described by the famous Navarro-Frenk-White (NFW) profile [74]:

$$\rho_{\text{NFW}}(r) = \frac{\rho_{\text{crit}} \delta_{\text{char}}}{r/r_s (1 + r/r_s)^2}, \quad (26)$$

where r_s is the scale radius, which is specified in terms of the virial radius, r_{200} , and concentration, c , as $r_s = r_{200}/c$. The mass is defined as that contained within the virial radius, where the average density is 200 times the critical density. Thus the characteristic density is given in terms of the concentration as

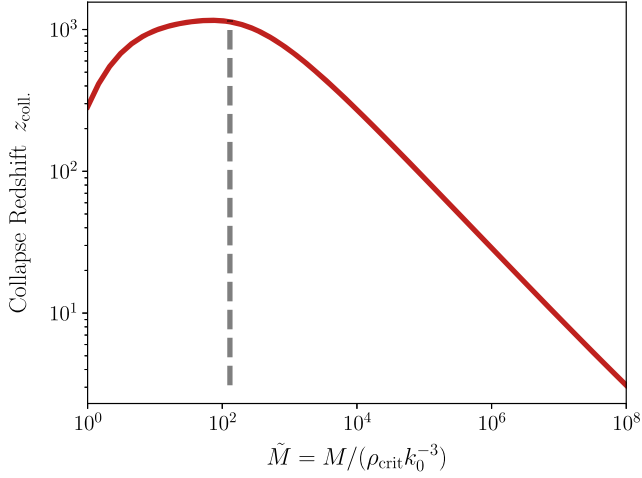


FIG. 7. MCH collapse redshift. The collapse redshift is computed following NFW (see main text). The earliest objects to collapse do so shortly after equality at $z_{\text{coll}} \approx 10^3$. The vertical line shows $M_0 = \rho_{\text{crit}}(4\pi/3)(\pi/k_0)^3$. Miniclusters at the characteristic mass M_0 are the first bound structures to form.

$$\delta_{\text{char}} = \frac{200}{3} \frac{c^3}{\ln(1+c) - c/(1+c)}. \quad (27)$$

Using our analytic result for the variance with a Heaviside filtering and Gaussian initial power spectrum, Eq. (B10), we first compute the collapse redshift, z_{coll} . Following NFW, this is defined using the extended Press-Schechter formalism as the redshift when half the mass of the halo was first contained in progenitors more massive than 1% of the final mass. The result is plotted in Fig. 7. The collapse redshift is maximized at $M \approx M_0$, indicating that objects of this characteristic mass are the first bound structures to form shortly after equality at $z \approx 10^3$. We note that this appearance of M_0 justifies our choice in Eq. (2). That is, the first objects to form are consistent with the spherical wavelength volume and are somewhat heavier than the cubic volume estimate.

It is these first bound structures which are “true” axion miniclusters, that is, MCHs with $M = M_0$. Due to the nonlinear dynamics of axion interactions in the instanton potential, these miniclusters display a spectrum of sizes and characteristic densities given by Eq. (4) and Fig. 2. The concentration is a stochastic function, which for the initially formed miniclusters has a large non-Gaussian tail.

We use Eq. (4) to define the characteristic density of miniclusters, $\rho_{\text{MC}}(\delta) = \rho_{\text{crit}}\delta_{\text{char}}(\delta)$, which defines a δ -dependent concentration for miniclusters, $c_{\text{MC}}(\delta)$. Due to the large spread in values of δ , there are miniclusters with concentrations far exceeding that expected from ordinary hierarchical structure formation. The physical reason for this is, as discussed, the nonlinear interactions in the axion potential, and the result is confirmed in numerical simulation.

The concentration-mass relation, $c(M)$, specifies the *spatially averaged* density profile of a MCH. Assuming

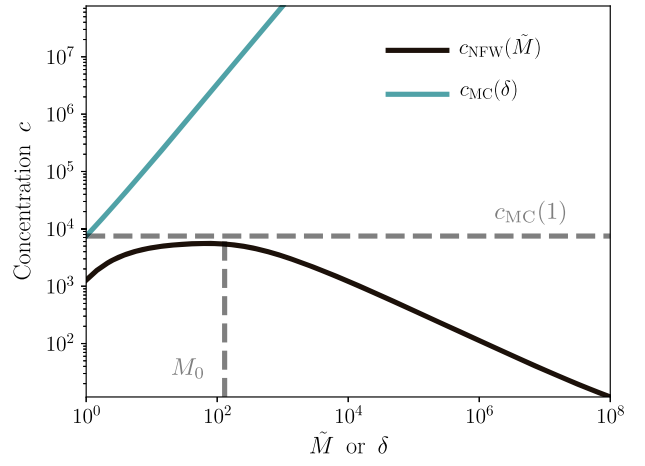


FIG. 8. Concentration of miniclusters and (diffuse) MCHs. For miniclusters we show $c_{\text{MC}}(\delta)$ derived for a NFW profile with characteristic density fixed by Eq. (4). For MCHs we show the concentration-mass relation assuming hierarchical structure formation. The vertical line shows the location of M_0 , which almost maximizes the concentration near $c_{\text{MC}}(1)$.

MCHs are formed hierarchically from the (assumedly) Gaussian large-scale white-noise density fluctuations of miniclusters in mergers that are described entirely by CDM we can adopt the analytic model for $c(M)$ proposed by NFW (related approaches include e.g. Ref. [75]):

$$\delta_{\text{char}} = 3000\Omega_m(1+z_{\text{coll}})^3. \quad (28)$$

The resulting concentration-mass relation, $c_{\text{NFW}}(M)$, is plotted alongside $c_{\text{MC}}(\delta)$ in Fig. 8. We notice that miniclusters of mass M_0 maximize $c_{\text{NFW}}(M)$ (since they have the largest z_{coll}) and that furthermore the NFW concentration agrees with the minicluster concentration for $\delta = 1$: $c_{\text{NFW}}(M_0) \approx c_{\text{MC}}(1)$. This is a pleasing coincidence that further validates our use of M_0 as the characteristic mass.

B. Minicluster mergers?

The NFW profile describes the average density profile of a halo. The Press-Schechter mass function (as we have used it) also simply groups all progenitors together into a single parent mass. Neither accounts for substructure. Using only the average density profile and ignoring the substructure is equivalent to the case where halo mergers completely disrupt the progenitors. Because observables, in particular microlensing, may be sensitive to the substructure of MCHs, and minicluster mergers may not be totally disrupting, we must be careful how the mass function and NFW density profile are used.

We will satisfy ourselves with some estimates. We begin with considering the Hill sphere:

$$r_{\text{Hill}} = a(1-e) \left(\frac{M_{\text{sat}}}{3M_{\text{host}}} \right)^{1/3}, \quad (29)$$

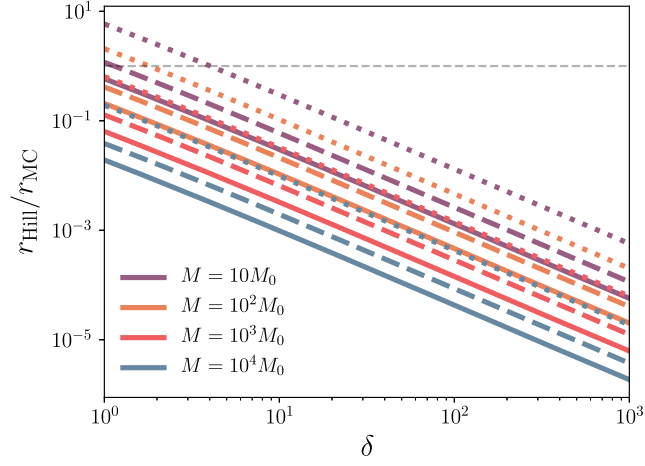


FIG. 9. The Hill radius. When the ratio of the Hill radius to the minicluster radius is larger than unity (horizontal dashed line), tidal stripping is significant. We consider minicluster satellites of mass M_0 in hosts of mass M located at the scale radius of the host. Solid, dashed, and dotted lines are for orbits of eccentricity $e = 0, 0.5, 0.9$ respectively.

where M_{sat} is the satellite mass, M_{host} is the host mass, and the satellite has an elliptical orbit of semimajor axis a and eccentricity e .

Consider a minicluster satellite of mass M_0 inside a MCH of mass M_{host} . The minicluster will be disrupted by tidal forces if the minicluster radius [calculated from $c_{\text{MC}}(\delta)$] is larger than the Hill radius. In other words stripping is significant if

$$\frac{r_{\text{Hill}}}{r_{\text{MC}}} > 1, \quad (30)$$

where we consider the minicluster radius to be given by the scale radius, $r_{\text{MC}}(\delta, M_0) = r_{200}(M_0)/c_{\text{MC}}(\delta)$.

For simplicity we consider the minicluster to be located at the scale radius of the host, where the majority of the mass is concentrated, with the concentration of the host given by $c_{\text{NFW}}(M)$. Figure 9 demonstrates that only the least concentrated miniclusters with the most eccentric orbits in relatively light MCHs are likely to undergo any significant tidal stripping as given by the Hill criterion.

A more complete analysis of stripping uses the tidal radius r_t for a satellite of mass M_{sat} orbiting at radius r_{sat} in a host halo of mass M_{host} [76]:

$$r_t = \left(\frac{GM_{\text{sat}}(< r_t)}{\omega^2 - \Phi''(r_{\text{sat}})} \right)^{1/3} \quad (31)$$

where Φ'' is the second derivative with respect to the radius of the gravitational potential in the host halo and ω is the angular velocity of the satellite. A full semianalytic merger tree calculation including tidal stripping (e.g. with GALACTICUS [77]) is beyond the scope of the present work. For simplicity we consider the satellite to be an initial

minicluster of mass M_0 located at the scale radius of the host (where most of the mass is) and take both satellite and host to have a concentration-mass relation given by $c_{\text{NFW}}(M)$ (equivalent to only the least dense minicluster satellites with $\delta = 1$). We find, similarly to the case with the Hill radius, that the tidal radius is always small compared to the minicluster radius, indicating that miniclusters do not undergo significant stripping within MCHs. We discuss this further in Sec. V C.

V. MICROLENSING WITH MINICLUSTERS

A. Microlensing basics

Microlensing is the fugitive amplification of a background star which occurs when a compact object passes close to the line of sight to that star [78].

The magnitude of the effect of microlensing by point objects is relatively simple to calculate and uses the normal equations for gravitational lensing [7]. The miniclusters that we are considering are not however point sources, and while they are very small, the deflection of light which they give rise to is also very small, so the extended nature of these objects is important to take into account.

For the configuration where observer, lens and source lie on the same line, the Einstein radius corresponds to the radius of closest approach of photons to the point mass lens as they pass by it:

$$R_E = \sqrt{\frac{4GM D_L D_{LS}}{c^2 D_S}} = 4 \times 10^{13} \sqrt{\frac{M}{M_\odot} \frac{D_S}{\text{kpc}}} \text{ cm} \quad (32)$$

where the enumeration assumes that $D_S = 2D_L = 2D_{LS}$ and the quantities D_L and D_S etc. are shown in Fig. 10. Taking benchmark values of $10^{-8} M_\odot$ mass halos at the distance of Andromeda, the Einstein radius will correspond to tens of nanoparsecs.⁶

A microlensing event occurs when a compact object passes through the microlensing “tube,” which has a radius of $u_T R_E$ where $u_T \approx 1$ is the minimum impact parameter for which the amplification of the background star is above the required threshold and R_E is the Einstein radius:

$$R_E(x) = 2 \left[\frac{GMx(1-x)D_S}{c^2} \right]^{1/2}, \quad (33)$$

where M is the lens mass, $D_S = L$ is the distance to the source, and $x = D_L/D_S$ is the distance of the lens from the observer [78]. In our case, the distance to the sources [Large Magellanic Cloud (LMC) or Andromeda (M31)] is much greater than their line of sight depth, so all of the

⁶This expression differs from that in Ref. [22] by a factor of 100 due to considering local rather than cosmological sources, and using the modern values of the cosmological densities. This leads us to find that larger values of δ are necessary for lensing.

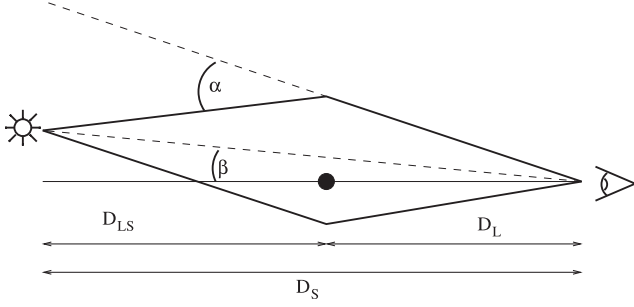


FIG. 10. The lensing of a source by a point mass at the origin. D_S is the distance from the observer (us) to the source, D_L is the distance to the lens and $D_{LS} = D_S - D_L$.

source stars can be assumed to be at the same distance (~ 50 kpc for LMC and ~ 770 kpc for M31) and the angular distribution of sources can be ignored.

The differential event rate for lenses of mass M in a halo with projected density profile $\rho(x)$ along the line of sight is given by [7]⁷

$$\frac{d\Gamma}{d\hat{t}} = \frac{32Lu_T^4}{\hat{t}^4 v_c^2} \frac{1}{M} \int_0^{x_h} \rho(x) R_E^4(x) e^{-Q(x)} dx, \quad (34)$$

where \hat{t} is the time taken by the lens to cross the Einstein diameter, x_h is the extent of the halo and $Q(x) = 4R_E^2(x)u_T^2/(\hat{t}^2 v_c^2)$, where $v_c = 220 \text{ km s}^{-1}$ is the local circular speed. The factor $u_T \approx 1$ defines the critical magnification for the lensing survey, which we take as $\mu = 1.34$ (see below). The factor e^{-Q} emerges by approximating the Bessel function in the lensing integral [7,79].

The expected number of events, N_{exp} , is given by

$$N_{\text{exp}} = E \int_0^\infty \frac{d\Gamma}{d\hat{t}} \epsilon(\hat{t}) d\hat{t}, \quad (35)$$

where E is the exposure in star years and $\epsilon(\hat{t})$ is the probability that a microlensing event with duration \hat{t} is detected (detection efficiency).

B. Miniclusters as nonpointlike objects

1. Lensing tube for nonpointlike objects

We are interested in the case where there is a magnification of 1.34, which is the threshold adopted by the EROS and HSC surveys. This magnification is not arbitrary, it corresponds precisely (in the pointlike mass case) to the outer ray passing the lens at a radius of $1.618 \times R_E$ where 1.618 is the golden ratio and R_E is the Einstein radius. The second image passes on the other side at a distance

$0.618 \times R_E$, i.e. the inverse golden ratio. The magnification of the two images is given in terms of the distance at which the light rays pass the lens $x = r/R_E$ and the magnifications μ are given by

$$\mu_{\pm} = \left[1 - \left(\frac{1}{x_{\pm}} \right)^4 \right]^{-1}. \quad (36)$$

The sum of the magnification of the two images is 1.34, the majority of which comes from the outer image which gives a magnification of 1.17.

We need to repeat the calculation for situations where the lens is potentially diffuse relative to the scales of interest, such that the enclosed mass is not a constant when δ is small.

We start by taking the characteristic density given by Eq. (4). We consider two functional forms for the minicluster density profiles. The NFW profile is a universal feature of CDM simulations emerging from hierarchical structure formation. However, miniclusters at the characteristic mass do not form from hierarchical structure formation, but probably from a more direct collapse. It has been suggested [34] that a more suitable profile for the initial seed miniclusters be given by that due to self-similar infall [80]:

$$\rho(r) = \rho_s \left(\frac{r_s}{r} \right)^{9/4}. \quad (37)$$

Such a power-law profile also appears in the minicluster N -body simulations of Ref. [25].

For both the NFW and self-similar profiles we need to make an identification with the density ρ_{MC} and the characteristic density. For the NFW profile we simply say that $\rho_{\text{MC}} = \rho_{\text{crit}} \delta_{\text{char}}$ as above, and rescale r_s until we obtain the correct mass of the halo at r_{max} . If we were to calculate the virial radius for these objects, they would be hugely larger than the scale radius by many orders of magnitude, but since the halos are within a galactic halo, we make the approximation that the NFW profile is cut off at a radius $r_{\text{max}} = 100r_s$. This simplifies the numerical lensing calculation by reducing the dynamic range.

The situation is slightly more complicated for the self-similar profile [Eq. (37)] because r_s is completely degenerate with ρ_s due to the scale invariance. The overall mass of such a halo which is truncated at a radius r_{max} is

$$M = \frac{16\pi}{3} \rho_s r_s^{9/4} r_{\text{max}}^{3/4} \quad (38)$$

and the average density of such a halo is

$$\rho_{\text{av}} = 4\rho_s \left(\frac{r_s}{r_{\text{max}}} \right)^{9/4} \quad (39)$$

or

⁷This expression assumes a spherical halo with an isotropic velocity distribution and ignores the transverse velocity of the microlensing tube, which has a small effect on the differential event rate [7].

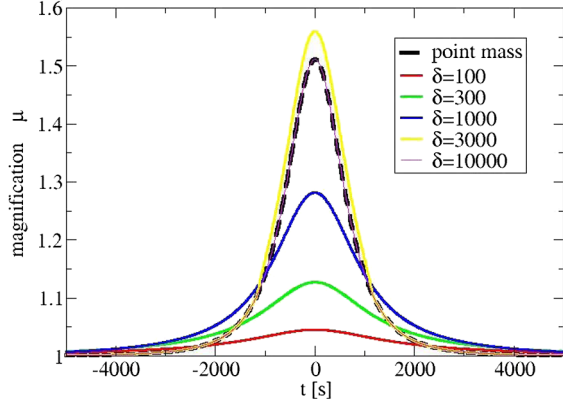


FIG. 11. Microlensing lightcurves. Magnification of source star as a minicluster passes through the lensing tube for various values of δ . The minicluster has a NFW profile, a mass of $10^{-9} M_{\odot}$, a tangential velocity of 200 km s^{-1} and an impact parameter with the line to the source of 1.6 milli-AU. The source star is assumed to be in the Andromeda Galaxy (M31), with $D_{LS} = D_S/2$. As we increase the value of δ the lensing curve approaches the value for a point mass, as described in the text.

$$r_{\text{max}} = \left(\frac{3M}{4\pi\rho_{\text{av}}} \right)^{1/3} \quad (40)$$

and then we make the identification $\rho_{\text{av}} = \rho_{\text{MC}}$.

We then want to turn the three-dimensional density into a surface density by collapsing it onto the lensing plane. We do this by integrating the three-dimensional density profile along the line of sight towards the center of the halo. In the situation where r_{max} is larger than the lensing radius we are probing, we only integrate the mass within the two cones

defined by the radius of interest and the distance between the lens and the source and the lens and the image.

Once we have surface mass as a function of radius we can calculate the magnification using the expression for an axisymmetric mass distribution (this can be derived from the equations in Ref. [81])

$$\mu = [(1-B)(1+B-C)]^{-1} \quad (41)$$

$$C = \frac{1}{\Sigma_c \pi r} \frac{dM(r)}{dr}; \quad B = \frac{M(r)}{\Sigma_c \pi r^2}; \quad \Sigma_c = \frac{c^2 D_S}{4\pi G D_L D_{LS}}. \quad (42)$$

What we need to do is to define the “lensing tube” [7] as being the tube within which a lens would create a magnification of at least 1.34.

We do this by starting at large r and zooming into the radius at which the magnification is the same as the outer image in the point mass case, i.e. $\mu = 1.17$. When a halo is diffuse, this will occur at a radius much less than the radius at which a point mass would give rise to the same lensing. The lightcurves for lensing computed as a function of δ for fixed minicluster mass are shown in Fig. 11. An interesting feature occurs as one increases the parameter δ in that the magnification rises above that for a point mass before settling down to the same value as the point mass. This is because there is an intermediate regime where there are multiple paths for photons from the same source to pass the lens at different radii and still arrive at the observer, adding to the overall magnification. This is captured by the gradient terms in the lens equations.

As δ increases and the halo becomes more compact it behaves increasingly like a point mass until r_{max} is well

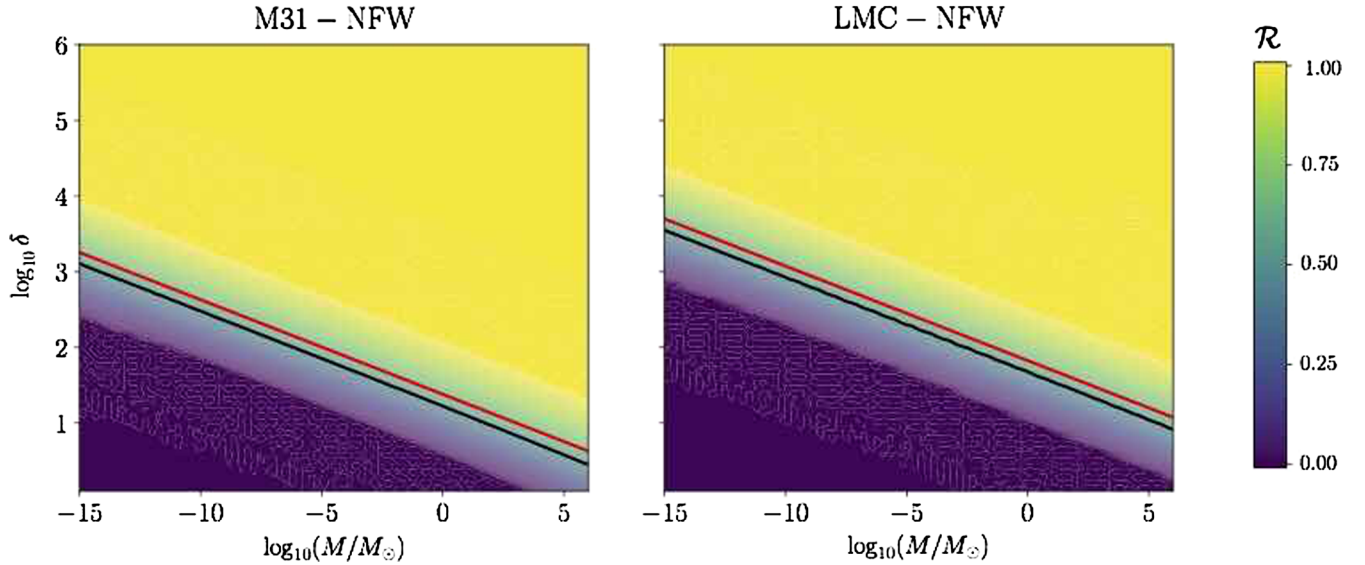


FIG. 12. Minicluster microlensing tube. Rescaling factor, \mathcal{R} , for a minicluster of mass M and overdensity δ with NFW radial profile. The black contour shows the location of $\mathcal{R} = 0.5$, while the red line shows the analytic result for $\delta_{\text{lens}}(M)$ defined by equating the point mass Einstein radius to the hard-sphere minicluster radius.

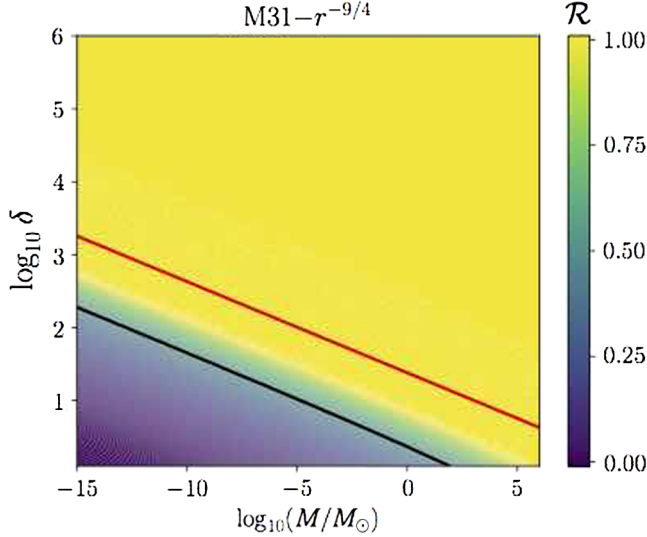


FIG. 13. As Fig. 12 for M31 in the case of the self-similar infall density profile, $\rho \propto r^{-9/4}$. This profile is more compact than the NFW profile (or a hard sphere) leading to a larger effective lensing tube at fixed M , δ .

below any of the scales of interest, the lensing tube is the same size as for the point mass case. One can check that for large values of δ the radius for the distributed mass and the radius for the point mass coincide.

For each value of M and δ we average the ratio between the radius of magnification 1.17 for ten values of x between 0 and 1 in order to obtain a numerical correction factor which we can then apply to the point mass lensing equation. For small values of δ which give rise to diffuse halos, there is no radius at which the magnification reaches 1.17 and such halos cannot contribute to the lensing integral.

We define the correction factor \mathcal{R} which corresponds to the average radius of the lensing tube with the specific magnification we are looking for divided by the average radius for a point mass. We plot values of \mathcal{R} for the NFW and self-similar profiles in Figs. 12 and 13. It is clear that for small values of δ the halos are diffuse and the corresponding lensing is suppressed. For larger values of δ the lensing increases and eventually when the vast majority of the mass is inside the lensing tube, the objects are indistinguishable from pointlike objects.

An estimate for the transition from $\mathcal{R} = 0$ to $\mathcal{R} = 1$ can be obtained using the hard-sphere minicluseter radius⁸:

$$r_{\text{mc}} = 4 \times 10^{16} \frac{1}{\delta((\delta+1)\Omega_m)^{1/3}} \left(\frac{M_0}{M_\odot}\right)^{1/3} \text{ cm.} \quad (43)$$

We equated this to the Einstein radius and solved for $\delta_{\text{lens}}(M)$; the value of δ when microlensing will be large.

⁸This differs by approximately a factor of three from the scale radius obtained by setting $\rho_{\text{MC}} = \rho_{\text{char}}$ and solving numerically for $c_{\text{MC}}(\delta)$.

The result is plotted in Figs. 12 and 13 as a red line. For the NFW case, the estimate is close to the contour for $\mathcal{R} = 0.5$ from the full numerical lensing calculation (black line) and accurately estimates the transition in the microlensing behavior.

It should be noted that we expect greater lensing if we adopt the self-similar profile rather than the NFW profile. The reason for this is that the majority of the mass contributing to the overall profile is greatest at the scale radius for the NFW profile, but reaches a maximum towards the center of the scale-invariant profile, so in the situation where only a fraction of the halo is within the normal lensing tube, the self-similar profile will give rise to a bigger lensing effect than the NFW profile. This can be seen in Fig. 13, where the transition in lensing behavior occurs at a smaller value of δ than for the NFW case, and the estimate for δ_{lens} from treating miniclusters as hard spheres is an overestimate.

2. Microlensing event rate for nonpointlike objects

From the previous numerical lensing calculations, we find that the shape of the microlensing tube is still reasonably well described by the Einstein radius, $R_E(x, M)$, but with a rescaling factor, \mathcal{R} , that depends on M and δ , such that the minicluseter lensing tube is given by

$$R_{\text{MC}}(x, M, \delta) = \mathcal{R}(\delta, M) R_E(x, M). \quad (44)$$

When a minicluseter/MCH is diffuse, the tube is smaller. There is a minimum value of δ below which there is no existing value of impact parameter ℓ for which $A \geq 1.34$, i.e. $\mathcal{R}(\delta < \delta_{\text{min}}) = 0$ with $\delta_{\text{min}} = \delta_{\text{min}}(M)$ given approximately by $r_{\text{mc}}/R_E > 1$. As we will see, this treatment reduces significantly the expected number of microlensing events for miniclusters compared to point masses (MACHOs, PBHs). For $\delta \gg \delta_{\text{min}}$ the limiting behavior is that of a point mass, $\mathcal{R} \rightarrow 1$.

Miniclusters of mass M_0 treated as nonpointlike objects have a rate of microlensing events of duration \hat{t} given by

$$\frac{d\Gamma}{d\hat{t}} = \frac{32Lu_T^4}{\hat{t}^4 v_c^2} \frac{1}{M_0} \int_0^\infty \frac{dn}{d\delta} \left[\int_0^1 \rho(x) R_E(x)^4 e^{-Q(R_E)} dx \right] d\delta, \quad (45)$$

which is a modified version of Eq. (34) where the Einstein radius has been replaced by the minicluseter lensing tube radius R_{MC} . The partition function $dn/d\delta$ is deduced from Fig. 2 of Ref. [22], which we showed earlier in Fig. 2 and the fit Eq. (4).

C. Mergers and the meaning of the mass function for microlensing

In the case of an extended mass function, Eq. (34) should be replaced by [82]

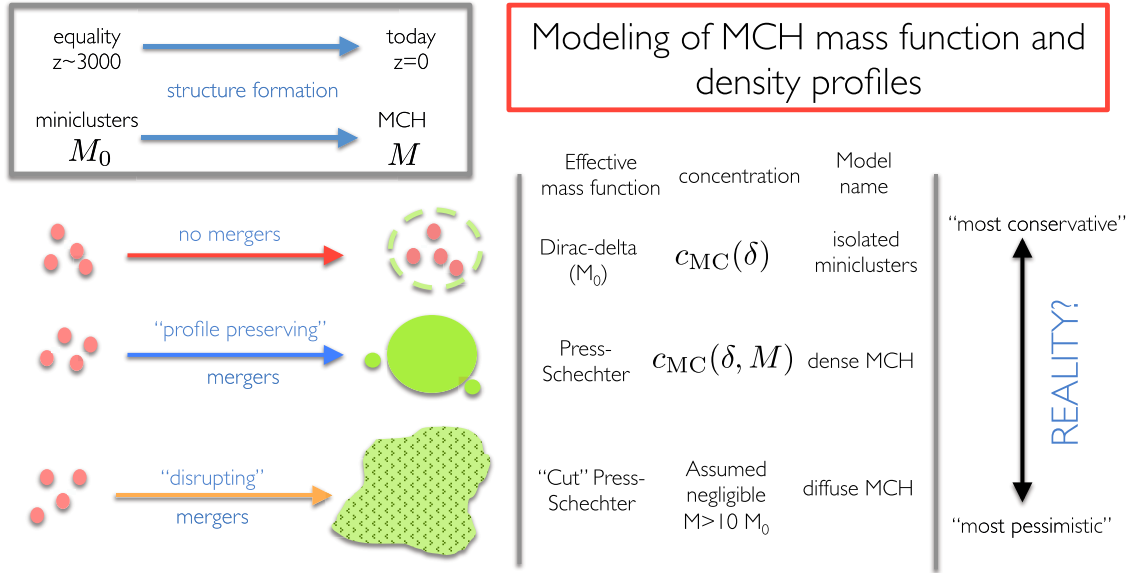


FIG. 14. Cartoon showing the modeling of the mass function and density profiles applied to the computation of the expected number of lensing events.

$$\frac{d\Gamma}{d\hat{t}} = \frac{32Lu_T^4}{\hat{t}^4 v_c^2} \int_0^\infty \frac{\psi(M)}{M} \left[\int_0^1 \rho R_{MC}^4 e^{-Q(R_{MC})} dx \right] dM, \quad (46)$$

where we have used the normalized mass function of Eq. (25).

The assumption in Eq. (46) is that an object of mass M in the mass function contributes to the microlensing only on a single time scale \hat{t} . The integral over the mass function ignores the possibility of microlensing events caused by substructure within an object of mass M . This is appropriate for pointlike objects such as PBHs and MACHOs [82]. For MCHs, which can have substructure, such an integral over the mass function assumes a smoothing of the lensing signal which is equivalent to assuming that miniclusters merge completely upon formation of a MCH.

The estimates in Sec. IV B suggest that miniclusters are dense enough to remain tightly bound within MCHs, like “plums in a pudding,” and do not undergo significant merging. In this case, there will always be lensing events with time scale $\hat{t}(M_0)$. A microlensing observation effectively smooths the lightcurve of a lensing event over a time scale given by the cadence [the lower limit of the efficiency $\epsilon(\hat{t})$]. If the microlensing observation in question is sensitive to $\hat{t}(M_0)$, and miniclusters do not merge when they form MCHs, then the effective mass function for microlensing is a Dirac-delta function, $\delta_D(M - M_0)$. We call this the “isolated miniclusters” scenario, and take it as our default model for the lensing constraints in Sec. VI.⁹

⁹In the isolated minicluster scenario MCHs can still play a role in modulating the arrival of miniclusters leading to correlation in the events of $\hat{t}(M_0)$ over the longer time scale $\hat{t}(M)$. Modeling such multi-time-scale microlensing would be challenging, but it

However, we may be wrong about minicluster mergers if scalar field dynamics plays an important part (i.e. if miniclusters cannot be considered as pure CDM). In such a case tunneling through the tidal radius can cause additional effects that need to be accounted for [83]. Scalar field dynamics also plays a significant role in axion star formation and core mergers (see Appendix D). To account for our inability to fully model these processes we consider two additional *ad hoc* possibilities for minicluster mergers into MCHs. Both scenarios assume complete merging of miniclusters (the “plums” dissolve), and use the Press-Schechter mass function to consider lensing by smooth MCHs.

- (i) “Dense MCHs”: Large non-Gaussian effects and interactions could cause MCHs to remain very dense. This scenario assumes that the concentration of a MCH is independent of mass and follows the same distribution with δ of miniclusters, $c_{MC}(\delta)$.
- (ii) “Diffuse MCHs”: Miniclusters could become completely stripped when they form MCHs. This scenario assumes MCHs follow $c_{NFW}(M)$ with no substructure.

In the dense MCH scenario, the MCHs themselves can still be dense enough to contribute to the microlensing signal, and so the lensing integral must account for both $dn/d\delta$ and the mass function $\psi(M)$, being¹⁰

could be used in the future to probe the MCH mass function and improve sensitivity to lower values of M_0 .

¹⁰For ease of notation in Ref. [27] we used dn/dM with implicit normalization by the host mass.

$$\frac{d\Gamma}{d\hat{t}} = \frac{32Lu_T^4}{\hat{t}^4 v_c^2} \int_0^\infty \left\{ \frac{dn}{d\delta} \int_0^\infty \left[\frac{\psi(M)}{M} \right. \right. \\ \left. \left. \times \int_0^1 \rho R_{\text{MC}}^4 e^{-Q} dx \right] dM \right\} d\delta. \quad (47)$$

The spread of masses away from M_0 can lead to a significant amount of mass moving in or out of the efficiency range of a given microlensing survey. In the range of scales relevant to the HSC microlensing survey, this effect ends up reducing the signal for the QCD axion minicluster masses and so (despite the vastly increased density of MCHs) it is *more pessimistic* than the isolated minicluster scenario (for lower minicluster masses than the QCD axion the microlensing event rate is increased and the dense MCH scenario is more optimistic).

In the diffuse MCH scenario we take only a small window of masses $M_0/10 < M < 10M_0$ to be described by $c_{\text{MC}}(\delta)$, i.e. those objects that have not undergone any mergers. We throw out all MCHs with $M > 10M_0$ from the mass function integral, since $c_{\text{NFW}}(M > M_0) < c_{\text{MC}}(\delta)$ for all M , and thus such diffuse MCHs will have $\mathcal{R} \ll 1$ and contribute negligibly to lensing. The diffuse MCH scenario is equivalent to a huge reduction in f_{MC} caused by structure formation and is the *most pessimistic* scenario for microlensing.

A cartoon depicting our three models for the microlensing signal of miniclusters and MCHs is shown in Fig. 14.

VI. RESULTS: MICROLENSING CONSTRAINTS ON AXIONS

A. The EROS microlensing survey

The EROS survey observed the LMC, at a distance $d_{\text{LMC}} = 50$ kpc, considering only lensing events of LMC stars by DM in the Milky Way (MW). EROS models the MW as a cored isothermal sphere:

$$\rho_{\text{MW,EROS}}(r) = \rho_0 \frac{R_c^2 + R_\oplus^2}{R_c^2 + r^2}, \quad (48)$$

where $R_\oplus = 8.5$ kpc is the radius of the Earth from the center of the MW and the MW halo parameters are $\rho_0 = 0.0079 M_\odot \text{pc}^{-3}$ and $R_c = 5$ kpc. A minicluster in the MW at distance d from Earth on the line of sight to the LMC has a radial coordinate in the MW halo $r_{\text{MW}}^2(d) = R_\oplus^2 - 2R_\oplus d \cos l_{\text{LMC}} \cos b_{\text{LMC}} + d^2$, where $(l, b) = (280^\circ, -33^\circ)$ are the measured Galactic coordinates.

For the EROS-2 survey $E = 3.68 \times 10^7$ star years. We extract the detection efficiency, in terms of Einstein radius crossing time, from Fig. 11 of Ref. [36].¹¹ The efficiency is

¹¹We follow Ref. [36] by multiplying the efficiency by an extra factor of 0.9 to take into account lensing by binary lenses.

large ($\epsilon \gtrsim 0.5$) for time periods of between 1 day and 1000 days.

B. Microlensing survey of Subaru HSC

The major limiting factor that prevents the EROS data from constraining MACHOs and miniclusters at low masses is the small-time efficiency, driven by the cadence of the observation. Very recently, observations in Ref. [37] used data from the Subaru Hyper Suprime Cam (HSC) to place constraints on low mass primordial black holes (PBHs) with $10^{-13} M_\odot < M_{\text{PBH}} < 10^{-6} M_\odot$. The HSC observations exclude a PBH fraction $f_{\text{PBH}} \gtrsim \mathcal{O}(\text{few}) \times 0.01$. This constraint was made by performing a microlensing survey with a cadence of 2 minutes over a single night of 7 hours.

For the microlensing survey, HSC observed Andromeda (M31). A major difference from the EROS survey, apart from the target galaxy, is that for microlensing in M31, due to the high DM density, one must account not only for lensing by DM in the MW, but also for lensing by DM in M31 itself. Thus the differential event rate is given by $d\Gamma = d\Gamma_{\text{MW}} + d\Gamma_{\text{M31}}$.

HSC adopted NFW radial density profiles for the DM in the MW and M31 with parameters from Ref. [84] of $r_s = 21.5$ kpc, $\rho_c = 4.88 \times 10^6 M_\odot \text{kpc}^{-3}$ giving $M_{\text{vir}} = 10^{12} M_\odot$ for the MW and $r_s = 25$ kpc, $\rho_c = 4.96 \times 10^6 M_\odot \text{kpc}^{-3}$ giving $M_{\text{vir}} = 1.6 \times 10^{12} M_\odot$ for M31. The different host masses for the MW and M31 normalize the minicluster mass function differently in each case.

The DM density profile along the line of sight is $\rho_{\text{DM}}(x) = \rho_{\text{MW}}(x) + \rho_{\text{M31}}(x)$, with $x = d/d_s$ for source distance $d_s = 770$ kpc. The line of sight distances are given in terms of the radial coordinate r as

$$r_{\text{MW}}(d) = \sqrt{R_\oplus^2 - 2R_\oplus \cos l \cos b + d^2}, \quad (49)$$

$$r_{\text{M31}}(d) = d_s - d, \quad (50)$$

with $R_\oplus = 8.5$ kpc the radial coordinate of Earth from the MW center and $(l, b) = (121.2^\circ, -21.6^\circ)$ the galactic coordinates of M31.

HSC has a microlensing efficiency of $\epsilon \sim 0.1$ – 0.8 for time periods 2 minutes $\lesssim \hat{t} \lesssim 7$ hours with a number of stars $N_s \sim 10^7$ – 10^9 . The advanced treatment of the efficiency and candidate selection employed in Ref. [37] is beyond the scope of the present work. In order to get a sense for the constraints that could be obtained with a dedicated analysis, we model the HSC microlensing efficiency as a step function with $\epsilon = 0.5$ in the given time scale (see Fig. 14 of Ref. [37]). To normalize the exposure we compute the expected number of events for PBHs, and we rescale E to approximately match the constraints in Fig. 29 of Ref. [37] without accounting for finite source

size and using the extrapolated number counts of source stars (this is the most optimistic constraint).

It is the short cadence that gives HSC access to low PBH masses, and for our purposes will allow constraints on the minicluster fraction for the QCD axion. This is because low mass objects create lensing events on shorter time scales due to the smaller radius of the microlensing tube. Thus detecting microlensing events by such objects requires large efficiency on small time scales, i.e. a short cadence.

HSC use pixel lensing and image subtraction to select microlensing candidates. Using this technique, they identify a large number of variable stars, eclipses, and other transient events. They find a single event with a light curve consistent with a PBH microlensing event, though the genuine nature is not confirmed. Thus, the Poisson statistics 95% C.L. upper limits on the expected number of microlensing events are

$$N_{\text{exp}} \leq \begin{cases} 3 & \text{with out the PBH candidate} \\ 4.74 & \text{with the PBH candidate} \end{cases} \quad (51)$$

We take $N_{\text{exp}} \leq 3$ as the conservative limit on the minicluster lensing events. A dedicated analysis of the HSC data with the minicluster light curve would be required to be more precise, and this is beyond the scope of the present work.

C. Results: Expected number of microlensing events

In Fig. 15 we show the expected number of microlensing events in various minicluster scenarios as a function of M_0 for HSC and EROS assuming $f_{\text{MC}} = 1$. The number of events in HSC is generally far larger than for EROS due to the huge volume of DM between the Earth and M31 leading to a larger optical depth to microlensing for HSC [37]. In order to understand the effects of our modeling we show four different calculations of N_{exp} for HSC.

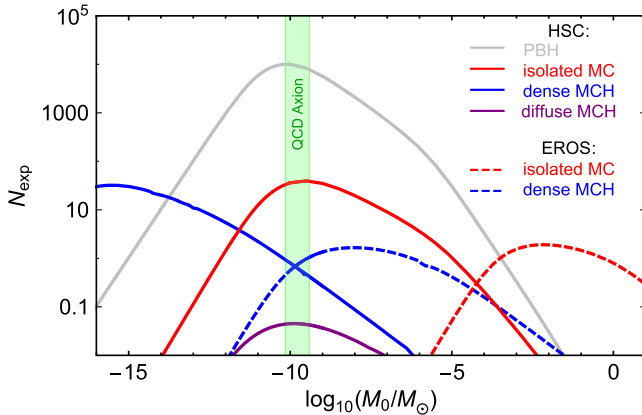


FIG. 15. Expected microlensing events. Here we assume that all the DM is composed of miniclusters on small scales. Lines show the effects of our modeling of the minicluster mass function and density profile for HSC and the EROS survey.

In the first (solid gray line in Fig. 15), we compute the event rate for pointlike objects (i.e. PBHs) of fixed mass M_0 (i.e. Dirac-delta-function mass distribution) to normalize the exposure and efficiency.

We then compute the case of isolated miniclusters (Dirac-delta-function mass distribution but nonpointlike objects), with density profiles determined by $dn/d\delta$ extracted from Fig. 2. This corresponds to the red solid/dashed line in Fig. 15 for the HSC/EROS survey. This additional treatment reduces the number of events by a factor of $\mathcal{O}(10^2)$ due to the requirement of large δ such that $\mathcal{R} > 0$. We consider this scenario as the most conservative: miniclusters are too dense to suffer much disruption on mergers, and MCHs are likely to be a “plum pudding” of objects of mass M_0 . In this case, the modulating role of the MCH mass function is not relevant for the HSC cadence and QCD axion.

The dense MCH case includes in addition the effects of dn/dM i.e. an extended mass function. A microlensing survey is sensitive to objects of fixed mass M . The mass function spreads the MCHs to $M > M_0$ (with more total mass at larger M), shifting the central M_0 to smaller values. The density profiles of the dense MCHs are also computed using $dn/d\delta$ i.e. mergers forming MCHs are assumed to preserve the distribution of halo concentrations. This treatment corresponds to the blue solid/dashed line in Fig. 15 for the HSC/EROS survey. This scenario is more conservative for the HSC survey and the QCD axion since it reduces the number of events by moving mass out of the central region of sensitivity.

Finally, the diffuse minicluster case uses dn/dM , but assumes that all MCHs with M outside the small window

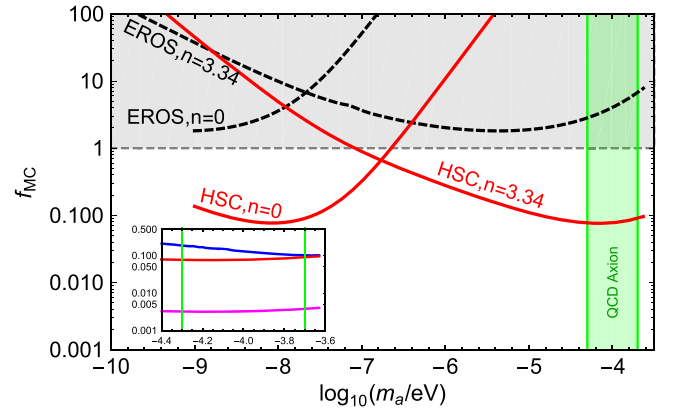


FIG. 16. Limits on the fraction of DM collapsed into miniclusters. The model adopted is for isolated miniclusters, which we consider the most realistic. The shaded region shows the allowed mass for the QCD axion with miniclusters. Where the $n = 3.34$ lines intersect this region, f_{MC} is constrained for the QCD axion. The inset shows a zoom-in. The magenta (blue) line in the inset shows a hypothetical improved observation by HSC of ten nights with an efficiency $\epsilon \sim 1$ in the case of isolated miniclusters (dense MCHs).

$M_0/10 \leq M \leq 10M_0$ have too low density for microlensing. Mergers are assumed to disrupt the miniclusters and the MCHs with $M > 10M_0$ are uniform with concentration $c_{\text{NFW}}(M)$, far too low to lens. The cut in dn/dM reduces significantly the number of events. This is the most pessimistic model, corresponding to an effective reduction in f_{MC} caused by mergers. This scenario corresponds to the purple line in Fig. 15 for the HSC survey.

Using the Poisson statistics 95% C.L. limit from above we find constraints on f_{MC} as a function of m_a , the axion mass, presented in Fig. 16 for the most realistic isolated miniclusters case. The dashed black lines correspond to the EROS limits in the $n = 0$ or $n = 3.34$ hypothesis. The solid red lines correspond to the HSC limits in the $n = 0$ or $n = 3.34$ hypothesis. We find that EROS does not place any bound on $f_{\text{MC}} < 1$; however HSC places very strong bounds on f_{MC} for an axionlike particle with $n = 3.34$, reaching as low as $f_{\text{MC}} \approx 8.0 \times 10^{-2}$ for $m_a \approx 50 \mu\text{eV}$.

The shaded green band shows the allowed mass for the QCD axion fixed by $m_a = 6.6 \mu\text{eV}(10^{12} \text{ GeV}/f_a)$ [13,14] and the relic density: $50 \mu\text{eV} \lesssim m_a \lesssim 200 \mu\text{eV}$ [40]. The $n = 3.34$ lines represent temperature evolution of the axion mass: where these lines intersect the shaded band, f_{MC} is bounded for the QCD axion, and we find $f_{\text{MC}} < 0.083(m_a/100 \mu\text{eV})^{0.12}$.

As shown in the inset of Fig. 16, these results could be improved. Indeed the magenta line shows a hypothetical improved observation by HSC, extending the current one night to ten nights with an efficiency $\epsilon \sim 1$, leading to a forecast bound of $f_{\text{MC}} \lesssim 0.004$ for the QCD axion in the isolated miniclusters case. The improved observation is also able to bound $f_{\text{MC}} \lesssim 0.1$ in the more pessimistic (for the QCD axion) dense MCH scenario. We advocate a dedicated analysis of the HSC microlensing data to place more rigorous bounds on f_{MC} than we have approximated

here, and for a longer microlensing survey in order to improve those bounds further. We are confident that a more thorough analysis by the observing teams will show that HSC, and microlensing in general, is now a powerful tool to constrain the QCD axion and, more generally, axionlike particles.

D. Theoretical uncertainties attached to our results

We now discuss various theoretical uncertainties and modeling that can give small shifts in the constraints. As we already discussed, an additional uncertainty comes from our simplified modeling of the lensing efficiency.

In Fig. 17, we study the impact of the modeling of the mass function on the computation of the expected number of lensing events. We used the analytical formulations of Appendix B. The solid blue line in Fig. 17 corresponds to our simple parametric mass function discussed earlier in Sec. III C.

For all combinations of initial power spectrum (Gaussian or Heaviside) and window function (Gaussian or Heaviside), we inject the variance σ_{XY}^2 of Appendix B in the usual Press-Schechter mass function also given in Appendix B in order to estimate the expected number of microlensing events. In the case of the Gaussian window function, we used the half-mode models for the low-mass cutoff. In order to isolate this source of theoretical uncertainty from other mass-dependent effects (such as the rescaling lensing tube factor) we considered in Fig. 17 the MCHs as pointlike objects, which is quite an unrealistic scenario. However we said previously that the additional treatment of considering MCHs as nonpointlike objects reduces the number of events by a factor of $\mathcal{O}(10^2)$ due to the requirement of large δ such that $\mathcal{R} > 0$. Our result is that our simple description of an extended mass function is conservative since for some characteristic mass, M_0 , a different choice of mass function could increase the expected number of lensing events by up to a factor of five.

The analytic results for the variance and associated MCH mass functions (GG, GH, HG, and HH) have generically a smaller upper limit compared to our simple parametric approximation (see Fig. 21). This small difference starts to matter when this upper limit approaches the critical value of the experiment. This critical value for HSC is around $M_0/M_\odot = 5 \times 10^{-5}$ and corresponds to the typical PBH mass when the lensing experiment loses all sensitivity due to the finite observing time (grey line in Fig. 15). At a particular moment, the parametric HMF will have MCH masses outside the sensitivity region of HSC; however for the other HMF all the masses would remain in the sensitive region and would consequently predict more lensing events. When every single HMF has MCHs heavier than that critical value then the discrepancy between them starts to shrink. This feature explains why the lines associated to the more refined HMF

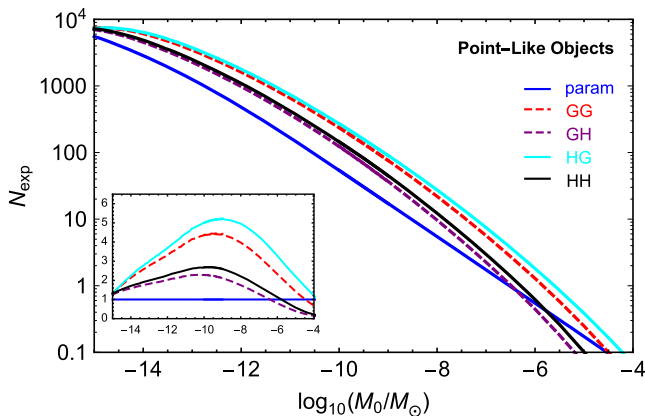


FIG. 17. Theoretical uncertainties in the mass function. Lines show the effects of our modeling concerning the mass function for the HSC survey. In order to isolate this effect from the effect of the density profile we considered the MCHs as pointlike objects (i.e. PBHs). The inset displays the various distributions normalized to our default simple parametric mass function (blue line).

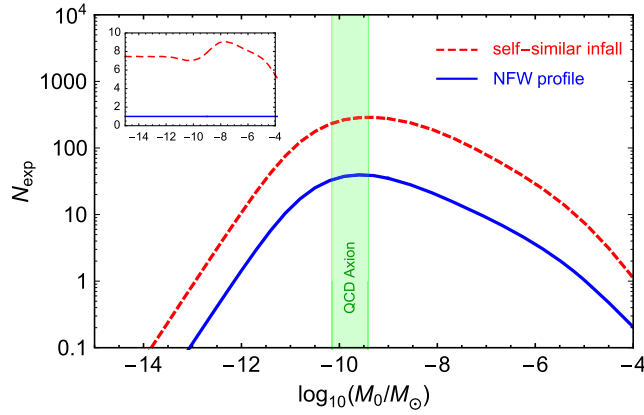


FIG. 18. Theoretical uncertainties in the density profile. Here we assume that all the DM is composed of miniclusters on small scales and take the isolated minicluster scenario. Lines show the effects of our choice concerning the density profile for the HSC survey. The dashed red line assumes a self-similar infall profile while the solid blue line assumes our default NFW profile. The inset displays the distributions normalized to the default NFW case (blue line).

are close to the parametric HMF for low and large values of M_0 and reach a maximum around $M_0/M_\odot \approx 10^{-9}$. The inset in Fig. 17 displays the various distributions normalized to our default simple parametric mass function (blue line).

In Fig. 18, we study the impact of the density profile on the computation of the expected number of lensing events for the HSC survey. The dashed red line assumes a self-similar infall profile (cf. Sec. IV) while the solid blue line assumes our default NFW profile. The inset displays the two distributions normalized to the default NFW line (blue line). As we can see, the self-similar infall profile predicts a higher expected number of lensing events by a factor of up to eight for M_0 masses below $10^{-6} M_\odot$. This is caused by the self-similar profile being more compact, thus requiring lower threshold values of δ for microlensing (see Fig. 13).

VII. SUMMARY AND CONCLUSIONS

The QCD axion remains one of the best-motivated dark matter candidates some 40 years after it was originally proposed. Unlike the case of thermal weakly interacting massive particle (WIMPs), the QCD axion parameter space remains wide open in the face of direct constraints due to the extraordinarily weak interactions between axions and the standard model. All of that is about to change, with a wide range of proposed experiments set to probe a large part of the parameter space in the coming decades. It is therefore timely to study more subtle aspects of axion DM that may affect direct detection signals.

If PQ symmetry breaking occurs during the radiation-dominated phase in the early Universe (or indeed during a putative matter-dominated phase) then the axion DM model

is severely constrained and makes rather precise predictions. One such prediction is the existence of miniclusters: gravitationally bound lumps of axions with masses of the order $M_0 \sim 10^{-10} M_\odot$. There is no theoretical prediction for the fraction of DM in miniclusters, f_{MC} , but naively we expect it to be of order unity. If f_{MC} is large, the direct detection signal for axions is reduced by $1 - f_{\text{MC}}$ (assuming that the probability of an encounter between the Earth and a minicluster over the course of an experiment is essentially zero).

Despite in some sense representing “half” of axion parameter space, the minicluster scenario has attracted relatively little attention over the years. In the present work we have attempted to revive interest in it and to observationally bound f_{MC} .

We began by computing the mass function of miniclusters. We predict the slope of the mass function should be $M^{-1/2}$ arising from the white-noise initial conditions. This is the mass function on large scales, where we have assumed Gaussianity of fluctuations. The minicluster halos that we predict are composed of many individual miniclusters (up to 10^6) and would not be seen in simulations such as Ref. [25] except in very large boxes with total mass many times larger than M_0 . We outlined uncertainties in the MCH mass function due to the initial power spectrum and window functions.

Our modeling of minicluster and MCH formation used the simplest possible implementation of Press-Schechter in an analytic form. More advanced semianalytic formulations of extended Press-Schechter can be used to answer deeper questions about the merger history and tidal effects and will provide more accurate mass functions [85]. The public code GALACTICUS is suitable for such a study [77]. Such approaches can account for gravitational N -body-like dynamics, but will not address issues associated to the unique minicluster initial conditions, which can only be answered by field theory simulations of the axion string network.

By treating the hierarchical structure formation as described by CDM on large scales, we computed the concentrations and formation times of MCHs. We used this to provide a simple estimate of the effect of tidal stripping on the “seed” miniclusters, and we have concluded that it is likely a minor effect. Thus, miniclusters should be present as plums in a pudding in MCHs, a scenario we dubbed isolated miniclusters. We also considered two other, less realistic, scenarios for mergers and their effects on observations. In the isolated minicluster case we concluded that the MCH mass function was irrelevant for microlensing constraints.

In the case of the QCD axion we expect the individual miniclusters to have masses very roughly comparable to larger asteroids like Vesta or Pallas with radius of order 1 AU for $\delta = 1$. After structure formation, these miniclusters find themselves gravitationally bound up into MCHs with typical masses up to that of Saturn and radii

of approximately 10^4 AU (0.1 light years). We are left wondering if there are other consequences of MCHs that can be tested observationally.

We have not discussed the formation of miniclusters themselves, where the assembly history will be very different. Our mass function on scales smaller than M_0 is also subject to large uncertainties. The shape of the mass function for $M < M_0$ (describing fragmentation and formation of miniclusters) could be drastically different from our estimates. We have tried to quantify this uncertainty using different window functions and cutoff models. We predict a cutoff in the mass function at least at the axion Jeans scale, and probably not much below M_0 itself.

The largest uncertainties relate to the non-Gaussianity of the minicluster density field. The non-Gaussianity will also affect the mass function on small scales. We have treated the non-Gaussianity on scales of order M_0 as giving the distribution of minicluster concentrations, using a fit from simulations [22] giving the distribution of overdensities $\mathcal{F}(\delta > \delta_0)$. This predicts that miniclusters come in a very wide range of different sizes for a fixed initial mass. As we have assumed a particular form for \mathcal{F} , our results could rather be interpreted as a constraint on the integrated tails of this distribution with $\delta_0 > \delta_{\text{lens}}$, where δ_{lens} is the critical density parameter for efficient microlensing.

Numerically verifying the distribution \mathcal{F} for large δ is a necessary future project, since our later results rest on an extrapolation of this function over a wide range.¹² Simulations of the minicluster field in the absence of gravity are not sufficient for this purpose: our results rely on the size distribution being preserved to the present day for bound objects of mass M_0 , which take some time to assemble gravitationally. The numerical verification should be done by performing a large number of small-box simulations containing a mass of just a few M_0 .

We computed the microlensing “tube size” and event rate for miniclusters. The distribution \mathcal{F} was key to this computation, since the densest miniclusters in the tail of the distribution have the largest effective tube radius and contribute dominantly to the expected number of events. Using the recent HSC limit on the number of microlensing events $N_{\text{exp}} \leq 3$ over event time scales between 2 minutes and 7 hours we were able to place the first observational bound on miniclusters of $f_{\text{MC}} < 0.083(m_a/100 \mu\text{eV})^{0.12}$ over the range relevant to the QCD axion. We also presented bounds for other axionlike particles.

In the event that observations are made which suggest microlensing events caused by DM, we would be faced with the pleasurable task of identifying the precise nature of these microlenses. The formation of miniclusters is in some

sense one of the more conservative scenarios for compact object formation, relying only on ordinary SSB leading to standard axion production. There are various other possible compact objects which could make up a significant fraction of the dark matter. We have already pointed out that with enough statistics we could tell the difference between pointlike objects such as PBHs and extended density profiles such as axion miniclusters. An interesting question is whether we could tell the difference between axion miniclusters and “ultracompact minihalos” that emerge as a result, for example, of features in the inflationary spectrum or a period of matter domination [87]. Just like in the case of miniclusters, there is significant theoretical uncertainty relating to the density profiles of such objects [88], and thus microlensing alone might not distinguish these types of compact object.

Ultracompact minihalo (UCMHs) can be composed of any type of DM, but they are typically thought of as being composed of WIMPs. Clearly, interactions distinguish axions and WIMPs and could also be used to distinguish miniclusters and UCMHs. A clear consequence of the high UCMH density is in boosting the WIMP annihilation rate, causing UCMHs to appear as luminous gamma ray sources [89]. On the other hand, axion-photon conversion leads to miniclusters acting as radio sources, with a possible relation to fast radio bursts [90–92]. Finally, on distinguishing different DM microlensing candidates, the early Universe physics responsible for each (minicluster, PBH, and UCMH) is in each case quite different (curvature perturbations for PBHs and UCMHs, isocurvature and SSB for miniclusters) and could lead to other indirect evidence supporting either case.

An $\mathcal{O}(1)$ uncertainty on the microlensing constraints comes from the minicluster density profile. We treated miniclusters as a NFW profile, which is close to a hard sphere with mass located predominantly at a single radius. In the case of a self-similar density profile, miniclusters are somewhat denser, and we found tighter constraints on the density fraction. Again, simulations are necessary to confirm which density profile is the correct one.

An observational bound on f_{MC} can be seen two ways. Firstly, observationally it is good to know that f_{MC} is small, since in the absence of other information this allows direct detection constraints to be interpreted with more confidence. Secondly, it allows the possibility to exclude or discover evidence for axions if theoretical predictions can be made. The same simulations that will provide us with more information about the distribution of overdensities \mathcal{F} could also narrow down predictions on the possible range of f_{MC} and help us to investigate the minicluster density profiles [86]. We hope that our preliminary investigations motivate the need for such further study.

If the eventual theoretical predictions are in violation of microlensing bounds, then microlensing constraints have excluded the QCD axion with late-time symmetry breaking.

¹²The simulations of Niemeyer [86] will be able to investigate this extrapolation. In the initial (no gravity) conditions some extrapolation seems justified to at least $\delta \gtrsim 100$. We thank Javier Redondo for discussion on this point.

On the other hand, if theoretical predictions are within an order of magnitude or so below our constraints, then future microlensing surveys could expect to see evidence of a subdominant population of miniclusters in the galactic DM. Upcoming observational efforts such as the Zwicky Transient Facility [93] and eventually the Large Synoptic Survey Telescope [94] are expected to provide much more sensitivity to microlensing events. We are excited to see what these telescopes observe and the implications of those observations for the nature of dark matter.

ACKNOWLEDGMENTS

The work of J. Q. and M. F. is supported by the UK Science and technology facilities council (STFC) Grant No. ST/L000326/1. M. F. is also funded by the European Research Council under the European Unions Horizon 2020 program (ERC Grant No. 648680 DARKHORIZONS). The work of D. J. E. M. is supported by a Royal Astronomical Society Postdoctoral Fellowship. We acknowledge useful discussions with Alberto Diez-Tejedor, Anne Green, Edward Hardy, Jens Niemeyer, Javier Redondo, and Paul Shellard. We also thank Torsten Bringmann, August Geelmuyden and Andrzej Jan Hryczuk for useful communications.

APPENDIX A: THE AXION RELIC DENSITY

The following section uses known results concerning the axion relic density to find the range of axion masses where the minicluster scenario is possible. We reproduce and extend the treatment of the relic density in Ref. [26], giving more details on the computation.

In the minicluster scenario the breaking of the PQ symmetry giving rise to axions must occur after any period of cosmic inflation (or indeed any other smoothing mechanism for the initial conditions). The axion relic density is produced by the coherent oscillations of the highly inhomogeneous classical axion condensate in a thermal background. In principle, the calculation of the relic density involves two computationally challenging pieces: the thermal evolution of the axion mass in finite temperature quantum field theory, and the classical evolution of the inhomogeneous condensate. We do not perform such calculations here, and in the literature results are only generally available for the QCD axion. However, both of these complexities can be parametrized into a much simpler computation, allowing us to treat the more general ALP models, and account for uncertainties in the numerical results in the literature.

In order to avoid having to compute the inhomogeneous field evolution, we make the standard assumption that there are two sources of axion relic density: production of cold axion particles via the decay of topological defects, and vacuum misalignment production from the homogeneous axion condensate. This separation of axion populations is

in fact artificial, the entire relic density being contained in coherent oscillations of the same classical field. It is a useful terminology, since the inhomogeneous calculation has the same parametric scaling as the homogeneous calculation, allowing a single fudge factor to be introduced, and the parameter space can then be explored in the much simpler homogeneous case.

In order to account for the finite temperature effects, we parametrize the mass evolution as a power law. The power law index can be computed either using an instanton model or a lattice computation, and it depends on the particle content (which is of course fixed only in the case of QCD).

Of the total axion relic density, some fraction of f_{MC} ends up bound in miniclusters. In principle this fraction can be determined by proper simulation of PQ symmetry breaking and axion production including the effects of gravity up to matter-radiation equality. Instead, we take f_{MC} as a phenomenological free parameter to be constrained by the data, and so our constraints on f_{MC} can be used to constrain models of PQ symmetry breaking.

The relic density of axions of mass m_a is determined by the decay constant, f_a , which for the minicluster scenario must satisfy $f_a < 8.2 \times 10^{12}$ GeV. The bound applies in an inflationary scenario by imposing the observational constraint on the cosmic microwave background tensor-to-scalar ratio $r < 0.07$ [95], which in turn constrains the Hubble scale during inflation. The constraint arises by imposing that the Gibbons-Hawking temperature, $T_{\text{GH}} = H_I/2\pi$, must be higher than the PQ scale in order that the PQ symmetry remain unbroken during inflation, and we have assumed for simplicity that the phase transition occurs at $T = f_a$. In a noninflationary scenario, if reheating is allowed to occur to temperatures $T_{\text{reh.}} > 8.2 \times 10^{12}$ GeV, then the minicluster scenario could be extended to larger values of f_a . For definiteness, we consider only the inflationary scenario from now on.

We must find the range of axion masses for which the minicluster scenario is relevant. To do this we impose the relic density constraint, $\Omega_a h^2 = 0.12$ [39], with $\Omega_a h^2$ computed in the minicluster scenario with the necessary requirement that $f_a < 8.2 \times 10^{12}$ GeV.

Consider the homogeneous misalignment production of axions. The equation of motion for the homogeneous axion field, ϕ , is

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0, \quad (\text{A1})$$

where dots denote derivatives with respect to physical time t . The axion potential is $V(\phi)$, and the prime denotes derivative with respect to ϕ . The precise form of $V(\phi)$ is not important for the present treatment of the relic density. An additional parameter N_{DW} sets the periodicity of the axion field and then determines the number of domain walls. For the QCD axion, N_{DW} is determined by the color anomaly of the quarks carrying PQ charge and is $N_{\text{DW}} = 1$ for the

KSVZ axion [15,16]. For simplicity of presentation we set $N_{\text{DW}} = 1$, the $N_{\text{DW}} > 1$ case having quite different cosmology [96].

Axion oscillations begin at temperature T_{osc} when the potential term dominates over the friction provided by the Hubble expansion. Thanks to the Kibble mechanism, the axion field remains homogeneous on scales up to the horizon size at this time. Once axion oscillations begin, the axion number density $n_a(T) = \rho_a(T)/m_a(T)$ becomes conserved (for slow m_a variation). The axion relic density at a later time, when the temperature is T_0 is then given by

$$\begin{aligned} \rho_a^{\text{mis}} &= m_a(T_0) n(T_{\text{osc}}) \left(\frac{a(T_{\text{osc}})}{a(T_0)} \right)^3 \\ &= \frac{1}{2} m_a(T_0) m_a(T_{\text{osc}}) f_a^2 \theta_i^2 \left(\frac{a(T_{\text{osc}})}{a(T_0)} \right)^3, \end{aligned} \quad (\text{A2})$$

where $\theta_i = \phi_i/f_a \in [-\pi, \pi]$ is the “initial misalignment angle.” The fraction of the critical density is given by $\Omega_a = \rho_a/(3H^2 M_{\text{pl}}^2)$. The scale factor is related to the temperature by the condition of constant entropy:

$$a(T) \propto g_{\star, S}(T)^{-1/3} T^{-1}, \quad (\text{A3})$$

with the proportionality normalized by fixing matter radiation equality at redshift $z_{\text{eq}} = 3402$ [39]. We use the fit for the entropic degrees of freedom, $g_{\star, S}(T)$, from Ref. [49].

In a harmonic potential, T_{osc} is given by

$$3H(T_{\text{osc}}) = m_a(T_{\text{osc}}). \quad (\text{A4})$$

We allow temperature variation of the mass parametrized as

$$m_a(T) = m_{a,0} \left(\frac{T}{\mu} \right)^{-n}, \quad (\text{A5})$$

for $T > \mu \propto \sqrt{m_{a,0} f_a}$ and $m_a(T < \mu) = m_{a,0} \equiv m_a$. Low-temperature variation of the mass occurs for axions acquiring their potential from a strongly coupled gauge theory such as QCD, and the index n can be computed given the particle content. As a representative of QCD we take $n \approx 3.34$ from the interacting instanton liquid model [49], which is close to the results from lattice simulations ($n \approx 3.55 \pm 0.30$ [50,51]) and the canonical dilute instanton gas ($n = 4$ [45]). For simplicity we neglect the case $\mu \neq \sqrt{m_a f_a}$. The QCD axion has $\mu \approx \Lambda_{\text{QCD}} \approx 2.5 \sqrt{m_a f_a}$. The very small effect of this factor on the relic density and M_0 is unimportant at our level of accuracy given the other associated uncertainties that we include below. The case $\mu \gg \sqrt{m_a f_a}$ is approximated by $n = 0$ and occurs for some string axions and “accidental axions” [97,98].

Assuming radiation domination, the Hubble rate is given by the Friedmann equation:

$$3H^2 M_{\text{pl}}^2 = \frac{\pi^2}{30} g_{\star, R}(T) T^4, \quad (\text{A6})$$

where $g_{\star, R}(T)$ is the number of relativistic degrees of freedom at temperature T . The epoch when axion oscillations begin is found by solving Eqs. (A4) and (A6), which we do numerically using the fit for $g_{\star, R}(T)$ from Ref. [49], which includes all the standard model degrees of freedom and the QCD phase transition.

Anharmonicities in the potential lead to flattening away from the origin and delay the onset of oscillations for large θ . The time that oscillations begin grows as [99]

$$t_{\text{osc}} = m_a^{-1}(t_{\text{osc}}) \ln[e/p(x)], \quad (\text{A7})$$

where $p(x)$ is a polynomial function of $x = \theta_i/\pi$ that is found from fitting to a numerical solution of Eq. (A1) with a specific potential [100], and it does not depend on m_a or f_a as each can be absorbed in an appropriate choice of units. For the potential $V(\theta) = m_a^2 f_a^2 (1 - \cos \theta)$ we find that $p(x) = 1 - x^4$ gives a good fit to our own numerical solutions for axionlike particles with T -independent mass and $g_{\star} = \text{const.}$

Using that $t \propto T^{-2}$ from the solution to the Friedmann equation, Eq. (A7) can be solved for the correction to $T_{\text{osc}}(\theta_i)$ that is then substituted into Eq. (A2) to find the relic density.

With the assumption of constant g_{\star} during the epoch over which anharmonic corrections affect T_{osc} , this solution can be found analytically, leading to an anharmonic correction to the relic density:

$$\rho_a^{\text{mis}} \rightarrow f_{\text{an}}(\theta_i) \rho_a^{\text{mis}}, \quad (\text{A8})$$

with the anharmonic correction function $f_{\text{an}}(\theta_i) = \{\ln[e/p(x)]\}^q$, and the power $q = 3/2 - n/(2n + 4)$. With non-constant g_{\star} , $f_{\text{an}}(\theta_i)$ cannot be found analytically, and furthermore it will in general depend on m_a , f_a , and n , though this is not usually stated.

In the minicluster scenario the vacuum misalignment relic density must be averaged over θ_i , reflecting the fact that the current observable Universe is many times larger than the horizon size when axion oscillations begin. One must replace $\theta_i^2 f_{\text{an}}(\theta_i)$ by

$$\langle \theta_i^2 f_{\text{an}}(\theta_i) \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \theta^2 f_{\text{an}}(\theta) d\theta \equiv c_{\text{an}} \frac{\pi^2}{3}. \quad (\text{A9})$$

The number c_{an} comes from the anharmonic corrections. Assuming one can treat g_{\star} as constant over the period in which anharmonic corrections act, with $p(x) = (1 - x^4)$ for the cosine potential and $n = (0, 3.34, 6)$ we find $c_{\text{an}} = (2.7, 2.1, 2.0)$, which we term c_n .

Now consider the population of axions from the topological defect decay. For $N_{\text{DW}} = 1$, the population of axions produced by decay of the string-wall network at T_{osc} can be parametrized by writing the total relic density as

$$\Omega_a h^2 = (1 + \alpha_{\text{dec}}) \Omega_a^{\text{mis}} h^2. \quad (\text{A10})$$

The parameter α_{dec} is computed from the numerical solution of the decay of the axion string-wall network [42,52,101–103]. The results can be expressed as a constant of proportionality, since the scalings are fixed by T_{osc} , with only weak dependence from nontrivial $g_*(T)$. The simulations of Ref. [96] find that $\alpha_{\text{dec}} = 2.48$.

For $N_{\text{DW}} > 1$ the relic density cannot be expressed in such a simple form. The domain wall network is long lived and the relic density of axions produced by its decay is too large (“overclosure”) unless some additional fine-tuning on CP -violation is allowed [96]. We do not consider the $N_{\text{DW}} > 1$ case any further. It would be interesting to consider in the future whether miniclusters or similar dense DM objects are produced in this scenario.

Thus, our final expression for the relic density in the minicluster scenario is

$$\Omega_a^{\text{total}} = \frac{1}{6H_0^2 M_{\text{pl}}^2} (1 + \alpha_{\text{dec}}) \frac{c_{\text{an}} \pi^2}{3} m_a(T_{\text{CMB}}) m_a(T_{\text{osc}}) f_a^2 \times \left(\frac{a(T_{\text{osc}})}{a(T_{\text{CMB}})} \right)^3, \quad (\text{A11})$$

where T_{osc} is found from solving Eq. (A4) for oscillations in a harmonic potential; $T_{\text{CMB}} = 2.725$ K is the CMB temperature; $a(T_{\text{CMB}}) = 1$; and we vary the parameters n , α_{dec} and c_{an} . We have not allowed for any additional entropy production or degrees of freedom beyond the standard model, which could suppress the relic density with an additional factor of $\gamma_{\text{ent}} < 1$ [104]. Allowing for $\gamma_{\text{ent}} < 1$ would raise the lower bound on m_a for axion miniclusters.

The results of our relic density computation are shown in Fig. 19. If axions are to provide the total DM relic density in the minicluster scenario we find a lower bound on the axion mass of $m_a \gtrsim 10^{-13}$ eV. This bound could be lowered somewhat for larger values of $n \gg 6$, which could be lowered slightly more for more extreme large values of n . For $n = 0$ we find $m_a \gtrsim 10^{-8}$ eV. We find that the QCD axion is where the dotted line $m_{a,\text{QCD}} \approx 6.6 \mu\text{eV} (10^{12} \text{ GeV}/f_a)$ intersects $n \approx 3.34$. Thus, when the PQ symmetry is broken post-inflation, the QCD axion must have $m_a \approx 100 \mu\text{eV}$, consistent with other estimates in the literature.

The relation between (m_a, f_a) found from the relic density affects the $M_0(m_a)$ relationship via the weak dependence on f_a for $n \neq 0$. For $n = 0$, the relic density only affects this by imposing a lower limit on n . We have verified the accuracy of the misalignment contribution to Eq. (A11) using a direct numerical solution of the temperature-dependent equations of motion following the method of Ref. [50], confirming both the use of the estimate for T_{osc} and the computation of c_{an} for the single cosine and chiral axion potentials. The approximation only breaks down for oscillation temperatures near Λ_{QCD} , where the effects of g_* on the dynamics are nontrivial. For $n = 0$ using the analytic

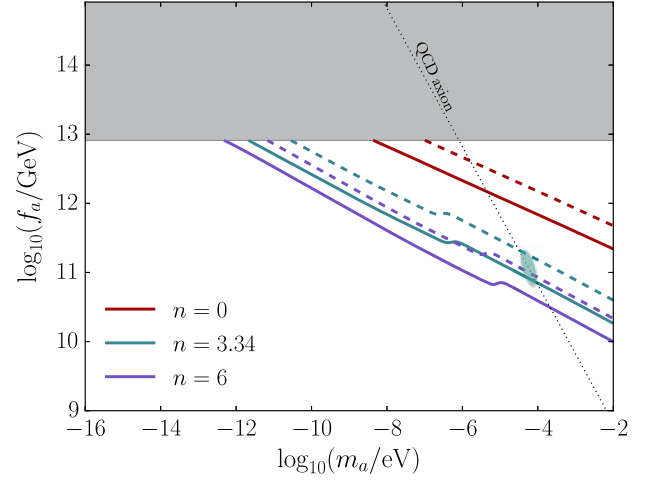


FIG. 19. The axion relic density. Contours $\Omega_a h^2 = 0.12$ are shown for a variety of models. Solid (dashed) lines have $\alpha_{\text{dec}} = 2.48(1)$ and $c_{\text{an}} = c_n(1)$ to account for uncertainty in the relic density from the decay of topological defects and anharmonicities in the axion potential. The shaded area indicates $f_a \geq 8.2 \times 10^{12}$ GeV, where inflationary constraints on the tensor-to-scalar ratio imply that miniclusters cannot be formed. The QCD axion is indicated by the ellipse giving $m_a \approx 100 \mu\text{eV}$.

result does not affect M_0 , and thus does not affect our microlensing results. For $n \neq 0$ the use of the analytic results has a smaller effect than the uncertainty in α_{dec} already accounted for.

The scenario discussed above does not apply to all possible axion models. For example, the relationship between f_a and the scale of SSB can be drastically modified in the “clockwork axion” theories [105,106]. Here a large hierarchy can be generated between f_a , which sets the periodicity of the axion field, and the symmetry breaking scale. For example, one could lower the symmetry breaking scale to $v_{\text{symm}} \approx 1$ TeV while keeping $f_a \approx 10^{16}$ GeV. Thus, lighter axions that require large $f_a > 8.2 \times 10^{12}$ GeV could still give rise to miniclusters. In the clockwork scenario one must be careful that the additional axions involved in realizing the clockwork do not produce dangerous relics themselves. The additional axions also significantly complicate the phase transition and subsequent production and decay of topological defects [107].

Geometric string theory (and supergravity) axions cannot give rise to miniclusters, as there is no notion of spontaneous PQ symmetry breaking in $3+1$ dimensions that could produce the necessary large field fluctuations as initial conditions. Accidental axions, where the $U(1)$ symmetry is explicitly broken by Planck suppressed operators, and which also arise in string theory, can undergo SSB and produce miniclusters.¹³

¹³We are grateful to Joseph Conlon, Matthew McCullough, Andreas Ringwald, and Fuminobu Takahashi for discussion on clockwork axions and string axions.

APPENDIX B: THEORETICAL UNCERTAINTIES IN THE MASS FUNCTION AND ANALYTIC RESULTS

There are four main theoretical uncertainties in our modeling of the MCH mass function:

- (i) Non-Gaussianity of the minicluster density field.
- (ii) The initial power spectrum, $P(k, \tau_0)$.
- (iii) The filtering function, $W(kR)$.
- (iv) The cutoff of the mass function.

Non-Gaussianity is addressed in detail in Appendix C. In this appendix we address the other three sources of uncertainty. We first compute the variance, $\sigma^2(M)$, at the initial time, τ_0 . We consider two simplified models for the initial power spectrum:

$$P_G(k) = P_{0,G} \exp \left[-\frac{1}{2} \left(\frac{k}{k_0} \right)^2 \right], \quad (\text{B1})$$

$$P_H(k) = P_{0,H} \Theta(k_0 - k), \quad (\text{B2})$$

with G for Gaussian and H for Heaviside. The Gaussian smoothing at k_0 to model the effects of the Kibble mechanism was used in Ref. [25]. We note that the normalizations, P_0 , of each power spectrum must be computed separately, and they are given by

$$P_{0,H} = \frac{24}{5} \pi^2 k_0^{-3}, \quad (\text{B3})$$

$$P_{0,G} = \frac{8\sqrt{2}}{5} \pi^{3/2} k_0^{-3}. \quad (\text{B4})$$

In addition to the power spectrum, we also consider two choices of window function (see e.g. Ref. [108]):

$$W_G^2(kR) = e^{-k^2 R^2}, \quad (\text{B5})$$

$$W_H^2(kR) = \Theta(1 - kR). \quad (\text{B6})$$

Since for all axion models considered we have $k_0 < k_{J,\text{eq}}$, the variance at $z = 0$ is simply given by the variance at $z = z_{\text{eq}}$ (i.e. computed with the initial power) divided by the CDM linear growth factor squared, $\sigma^2(z = 0) = \sigma^2(z_{\text{eq}})/D^2(z_{\text{eq}})$. The CDM linear growth factor is given by the integral (assuming flatness, $\Omega_\Lambda = 1 - \Omega_m$)

$$\tilde{D}(z) = \Omega_m \frac{5}{2} \frac{H(z)}{H_0} \int_z^\infty dz' \left[\frac{H(z')}{(1+z')H_0} \right]^{-3}. \quad (\text{B7})$$

The growth factor is normalized to unity at $z = 0$ such that $D(z) = \tilde{D}(z)/\tilde{D}(0)$.

For all combinations of initial power and window function the variance can be expressed analytically, in terms of error functions where necessary. We write the variance as σ_{XY}^2 , where X and Y take on either of the values

G and H, with X labeling the window function and Y labeling the initial power spectrum. The closed-form expressions in terms of the dimensionless radius $\tilde{R} = Rk_0$ are

$$\sigma_{GG}^2(\tilde{R}) = \frac{P_{0,G} k_0^3}{D^2(z_{\text{eq}})(2\pi^2)} \frac{\sqrt{\pi/2}}{(1 + 2\tilde{R}^2)^{3/2}}, \quad (\text{B8})$$

$$\sigma_{GH}^2(\tilde{R}) = \frac{P_{0,H} k_0^3}{D^2(z_{\text{eq}})(2\pi^2)} \frac{1}{4\tilde{R}^3} [-2\tilde{R}e^{-\tilde{R}^2} + \sqrt{\pi}\text{erf}(\tilde{R})], \quad (\text{B9})$$

$$\sigma_{HG}^2(\tilde{R}) = \frac{P_{0,G} k_0^3}{D^2(z_{\text{eq}})(2\pi^2)} \left\{ -\frac{e^{-1/(2\tilde{R}^2)}}{\tilde{R}} + \sqrt{\frac{\pi}{2}} \text{erf}[(\sqrt{2}\tilde{R})^{-1}] \right\}, \quad (\text{B10})$$

$$\sigma_{HH}^2(\tilde{R}) = \frac{P_{0,H} k_0^3}{D^2(z_{\text{eq}})(2\pi^2)} \frac{1}{3} \left[\frac{\Theta(\tilde{R} - 1)}{\tilde{R}^3} + \Theta(1 - \tilde{R}) \right]. \quad (\text{B11})$$

The mapping from radius to mass for the Gaussian window function is $M_G = (2\pi)^{3/2} \bar{\rho}_a R^3$. For the Heaviside window function the mass is not well defined. We choose a mass-radius relation that gives the same variance at large M as the Gaussian case, which is easily shown to be $M_H = (4/3)M_G$. We express the variance in terms of the dimensionless mass, $\tilde{M} = M/(\bar{\rho}_a k_0^{-3})$ (the dimensionless characteristic minicluster mass is $\tilde{M}_0 \approx 130$). The variance $\sigma_{HH}^2(\tilde{M})$ has a particularly simple closed-form expression:

$$\sigma_{HH}(\tilde{M}) = \sigma_0 \left[\Theta(\tilde{M} - c_m) \left(\frac{\tilde{M}}{c_m} \right)^{-1/2} + \Theta(c_m - \tilde{M}) \right], \quad (\text{B12})$$

with $c_m = (4/3)(2\pi)^{3/2}$ and $\sigma_0 \approx 10^{3.4}$. In terms of M_0 defined in Eq. (2) this gives the cutoff in the variance at the physical mass scale $M_{\text{cut}} = 2^{3/2} \pi^{-5/2} M_0 \approx M_0/6.2$.

The variances at $z = 0$ as a function of \tilde{M} for all four combinations of initial power spectrum and window function are plotted in Fig. 20. Despite the different normalizations of the power, P_0 , the variance has the same normalization in every case, as it should. We observe that the Gaussian initial power has a cutoff at lower \tilde{M} than for the Heaviside initial power, and consequently the variance on larger mass scales is smaller for the Gaussian case. The effect of the window function is to give a flatter variance as $\tilde{M} \rightarrow 0$ for the Heaviside window compared to the Gaussian window, a fact which has important consequences for the cutoff in the mass function, which we now turn to.

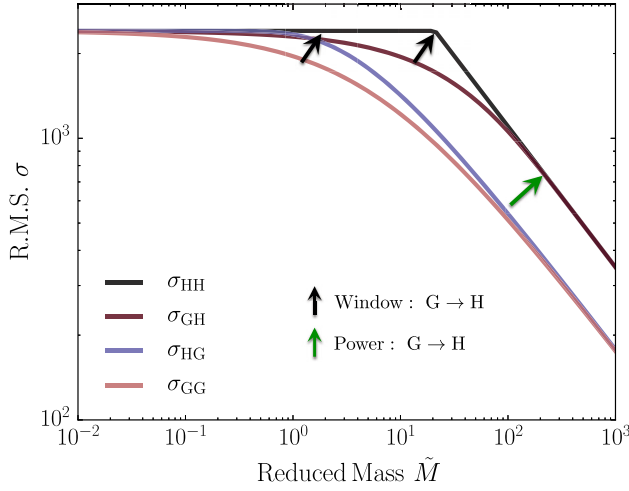


FIG. 20. Theoretical modeling of the rms. The mass fluctuation, σ , is shown as a function of mass for four different combinations of initial power spectrum and window function. The Gaussian initial power cuts off power earlier than the step function. The Heaviside window function leads to a more pronounced flattening of σ at low masses. The reduced mass $\tilde{M} = Mk_0^3/\bar{\rho}_a$.

The Press-Schechter mass function is given by

$$\frac{dn}{d \ln M} = \frac{1}{2} \frac{\bar{\rho}_a}{M} \left| \frac{d \ln \sigma^2}{d \ln M} \right| \sqrt{\frac{2}{\pi}} \frac{\delta_c}{\sigma} \exp \left[-\frac{1}{2} \left(\frac{\delta_c}{\sigma} \right)^2 \right]. \quad (\text{B13})$$

When the logarithmic derivative of the variance goes to zero fast enough, this cuts off the mass function at low M . This occurs for the Heaviside window function and was advocated by the authors of Ref. [64] to explain the downturn in the mass function seen in N -body simulations with truncated initial power spectra for warm DM. In our case, the flat variance with the Heaviside window also leads to a cutoff at low M for the HMF using σ_{HH} and σ_{HG} , with the cutoff for σ_{HH} being a step function at $\tilde{M} = c_m$.

For the Gaussian window, the logarithmic derivative of the variance does not go to zero fast enough to cut off the mass function. This is in conflict with simulations. We expect the Jeans scale to physically cut the power off in a simulation of the full axion field [71]. Even N -body simulations that simply use a truncated initial power spectrum see a downturn in the HMF after numerical artifacts have been removed [72,73].

Therefore, for the Gaussian window HMF we consider alternative cutoff procedures. The cutoff we adopt, as described in the main text, simply replaces δ_c with $\mathcal{G}(M)\delta_c$ in the Press-Schechter mass function. This is the approximation to the action of the Jeans scale considered in Refs. [66,109,110], which qualitatively agrees with the excursion set calculation using the same barrier of Ref. [69]. We refer to this cutoff as the “Jeans cutoff”: it is physically motivated, and qualitatively matches

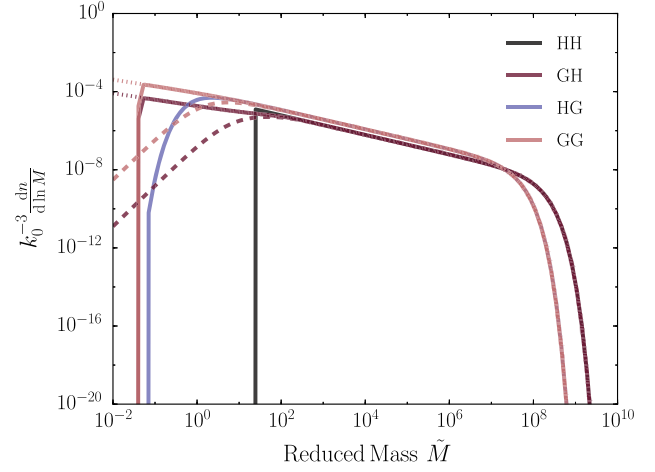


FIG. 21. Theoretical modeling of the mass function. The minicluster mass function is shown as a function of mass for four different combinations of initial power spectrum and window function. In this case of the Gaussian window function, we also show the uncut (dotted), half-mode (dashed), and Jeans (solid) models for the low-mass cutoff. The reduced mass $\tilde{M} = Mk_0^3/\bar{\rho}_a$.

simulations. However, this cutoff requires one to know the axion mass, as well as k_0 , and therefore it depends on the $M_0(m_a)$ relation. In the examples we use the fit for an $n = 0$ ALP.

Another possibility is the fit to the cutoff seen in the simulations of Ref. [72] (see also Ref. [73]):

$$\frac{dn}{d \ln M} \rightarrow \left[1 + \left(\frac{2.4M}{M_{1/2}} \right)^{-1.1} \right]^{-2.2} \frac{dn}{d \ln M}, \quad (\text{B14})$$

where we have rescaled the cutoff to be given by the “half-mode mass,” $M_{1/2}$, defined by the initial conditions cut at k_0 [72]. The half-mode masses are

$$\tilde{M}_{1/2,H} = (2\pi)^{3/2}, \quad (\text{B15})$$

$$\tilde{M}_{1/2,G} \approx 0.2(2\pi)^{3/2}. \quad (\text{B16})$$

The effects of all the modeling on the minicluster mass function are shown in Fig. 21, where we show all four combinations of initial power spectrum and window function, and for the Gaussian window we show the uncut, half-mode cutoff, and Jeans cutoffs.

The largest difference between different cutoff models with the same initial power happens in the case of the Heaviside initial power. The Heaviside window gives a sharp cutoff at \tilde{M}/c_m where the derivative of the variance goes exactly to zero. The half-mode cutoff comes in shortly after this, since we must identify the half-mode with k_0 in the absence of any other scale. The Jeans cutoff in this

case gives considerably more low-mass halos, a number of orders of magnitude below.

In the case of the Gaussian initial power, the different cutoffs make far less difference to the mass function. The Heaviside window leads to a cutoff around $\tilde{M} = 1$, due to the shallower flattening of the variance. The half-mode is $\tilde{M}_{1/2,G} \approx 3.1$, and the Jeans scale is $\tilde{M}_J \approx 0.3$, and thus the cutoff dependence occurs over less than one order of magnitude in \tilde{M} . For $n \neq 0$ we generically have $k_0 < k_{J,\text{eq}}$, and the Jeans mass can be much smaller than the characteristic minicluster mass. In this case the cutoff caused by the Kibble mechanism should dominate over the cutoff caused by the Jeans scale.

In the case of the barrier cutoff, the relation between $M_0(m_a, n)$ must be specified to derive an approximation to the cutoff, since the Jeans mass is given in terms of m_a (see Fig. 3). Factoring out the linear growth, the argument of the Gaussian $e^{-x^2/2}$ is $x(M) = 1.686\mathcal{G}(M)/\sigma_0(M)$, where $\mathcal{G}(M)$ is the fitting function of Ref. [110]. The barrier $\mathcal{G}(M)$ depends on the Jeans mass, M_J , which itself depends on the axion mass as

$$M_J = 5.1 \times 10^{-10} M_\odot \left(\frac{m_a}{10^{-10} \text{ eV}} \right)^{-3/2} \times \left(\frac{\Omega_m}{0.32} \right)^{1/4} \left(\frac{h}{0.67} \right)^{1/2}. \quad (\text{B17})$$

Figure 22 shows $x(M/M_J, M_0/M_J)$ for $0 < \mathcal{R} < 10^6$. Cutoffs in the HMF occur at small and large masses, when

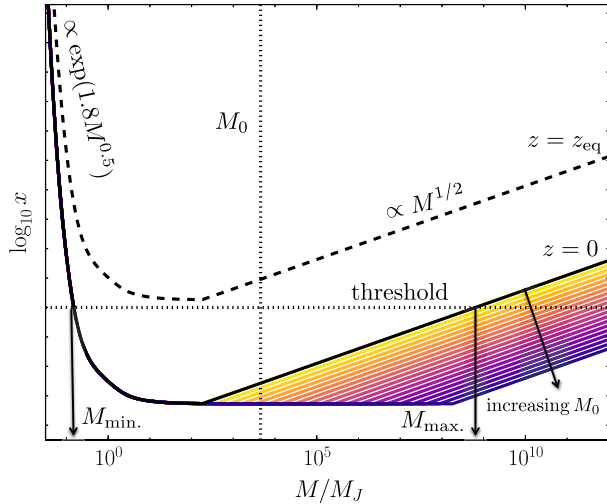


FIG. 22. Jeans scale and cutoffs in the mass function. These are driven by the argument of the Press-Schechter Gaussian, $e^{-x^2/2}$ and the threshold for a 1σ cutoff is $x = 1$. We plot x as a function of M/M_J with M_J the Jeans mass. As the ratio M_0/M_J increases from 1 to 10^6 , the mass function remains centered near M_0 but continues to have support at $M_J \ll M_0$. The spread of minicluster masses increases from a narrow distribution near M_0 at z_{eq} to a much wider distribution today.

$x(M)$ exceeds a particular threshold. The threshold for a 1σ cutoff is $x = 1$. The scale M_0 is the largest scale to cross the threshold shortly after z_{eq} , and thus sets the characteristic mass of miniclusters.

The Jeans mass depends only on the $T = 0$ axion mass, and so the mass function cutoff at low M is always at approximately the same value for fixed m_a regardless of the temperature dependence of the axion mass. However, as we saw in Fig. 1, as n is increased and the axion mass switches on more sharply, the value M_0 of the characteristic minicluster mass is increased for fixed m_a . When M_0/M_J is large, the spread of the minicluster mass function is increased. This demonstrates that the barrier cutoff has the largest difference in the low mass behavior of the mass function for different values of n . Halos/miniclusters at the Jeans mass are formed by “monolithic collapse” and in simulations are shown to be composed of isolated “axion stars” [70].

APPENDIX C: NON-GAUSSIAN EFFECTS ON THE MASS FUNCTION

We have already said something about the non-Gaussianity in the main text. We implicitly assume that on large scales $k < k_0$ the fluctuations are close to Gaussian and that Press-Schechter can be applied to compute the MCH mass function for $M > M_0$. In the main analysis we account for non-Gaussianity in the distribution of density profiles on small scales at $M \sim M_0$ using the distribution $\mathcal{F}(\delta > \delta_0)$.

In the following we discuss how non-Gaussianity in the white noise for $k_0 < k_0$ can alter the mass function. In analogy to Eqs. (6), (8) we consider the three-point correlation function of the overdensity:

$$\xi_3(r_1, r_2) = \langle \delta(\mathbf{x}) \delta(\mathbf{x} + \mathbf{r}_1) \delta(\mathbf{x} + \mathbf{r}_2) \rangle, \quad (\text{C1})$$

from which it can be shown that the bispectrum, $B(k_1, k_2)$, satisfies

$$\int \frac{d^3 k_1}{(2\pi)^3} \frac{d^3 k_2}{(2\pi)^3} B(k_1, k_2) = \langle \delta(\mathbf{x})^3 \rangle. \quad (\text{C2})$$

Another similar identity holds for the trispectrum, $T(k_1, k_2, k_3)$.

We model the axion initial field distribution as a random variable, θ , with uniform distribution on $[-\pi, \pi]$. At early times when the axion interactions can be neglected (the potential has not switched on, $T \ll \mu$ and $n > 0$), the uniform distribution is enough to allow us to calculate $\langle \delta(\mathbf{x})^3 \rangle = 16/35$. We note that we cannot use the numerically derived distribution $\mathcal{F}(\delta > \delta_0)$ to compute the cumulants since it is not shown for negative δ , and the Pearson distribution fit is not valid in this regime either.

Next we assume, just like for the power spectrum, that the non-Gaussianity is constant on large scales but that the

Kibble mechanism erases all correlations on small scales. Our ansatz for the bispectrum is thus

$$B(k_1, k_2) = B_0 \Theta(k_1 - k_0) \Theta(k_2 - k_0). \quad (\text{C3})$$

Plugging this ansatz into Eq. (C2) we can perform the integrals and compute the normalization:

$$B_0 = \frac{576}{35} \pi^4 k_0^{-6}. \quad (\text{C4})$$

A number of authors have considered the effect of a nonzero bispectrum on the mass function. For example, the authors of Ref. [111] performed an exact calculation using the path integral and excursion set theory. A simple formula to model non-Gaussian effects on the mass function is given in Ref. [112]:

$$\Pi_{\text{NL}}(M) = \exp \left[\frac{\delta_c^3 \mathcal{S}_3}{6\sigma^2} \right] \left(\frac{1}{6} \frac{\delta_c}{\sqrt{1 - \delta_c \mathcal{S}_3/3}} \frac{d\mathcal{S}_3}{d \ln \sigma} + \sqrt{1 - \delta_c \mathcal{S}_3/3} \right), \quad (\text{C5})$$

where Π_{NL} is the ratio of the non-Gaussian to Gaussian mass function. We do not use the related formula given by Ref. [113] because, for the large values of σ^2 on scales in our model, it gives rise to the unphysical values $\Pi_{\text{NL}} < 0$.

The non-Gaussian corrections depend on the *normalized skewness*:

$$\mathcal{S}_3 = \frac{\langle \delta^3 \rangle}{[\langle \delta^2 \rangle]^2} \equiv \frac{s_3(M)}{[\sigma^2(M)]^2}. \quad (\text{C6})$$

We are interested in the MCH mass dependence of this quantity. The denominator is simply the variance squared, computed under filtering by the appropriate window function. The bispectrum in the numerator must also be filtered to compute $s_3(M)$. The filtered bispectrum is

$$s_3(R) = \int \frac{d^3 k_1}{(2\pi)^3} \frac{d^3 k_2}{(2\pi)^3} B(k_1, k_2) W^2(k_1 R) W^2(k_2 R). \quad (\text{C7})$$

In the following we work with the Gaussian filter to be explicit, but the results do not depend on this choice. Using the results for P_0 and B_0 we find

$$\sigma^2(\tilde{R}) = \frac{12}{5} \int_1^\infty dy y^2 e^{-y^2 \tilde{R}}, \quad (\text{C8})$$

$$s_3(\tilde{R}) = \frac{144}{35} \left[\int_1^\infty dy y^2 e^{-y^2 \tilde{R}} \right]^2, \quad (\text{C9})$$

and the normalized skewness is

$$\mathcal{S}_3 = \frac{5}{7}. \quad (\text{C10})$$

It is important to note that the normalized skewness thus computed is *independent of scale*.

Plugging this result into Eq. (C5) we find that Π_{NL} diverges exponentially on large scales when $\sigma^2 \rightarrow 0$. However, this is not a fundamental problem, since Π_{NL} multiplies the mass function, which is cut off exponentially on large scales. On such large scales, the skewness-corrected nonlinear mass function goes like

$$\frac{dn}{d \ln M} \propto \exp \left[\frac{\delta_c^2}{\sigma^2} \left(\frac{\delta_c \mathcal{S}_3}{6} - \frac{1}{2} \right) \right] \approx \exp \left[-0.3 \frac{\delta_c^2}{\sigma^2} \right] \quad (\text{C11})$$

and thus the mass function remains cut off on large scales and gives sensible results in the presence of this type of non-Gaussianity. The change in the mass function cutoff leads to a small increase in the value of M_{max} , the maximum MCH mass.

On intermediate and small scales Π_{NL} is approximately constant and equal to 0.77. Thus we find that the bispectrum correction to the mass function leads to a small $\mathcal{O}(20\%)$ suppression of the number density of intermediate mass MCHs. The number of large MCHs is increased by a much larger fraction, but their number density remains exponentially small.

Under our approximations, we can also compute the normalized excess kurtosis from the trispectrum:

$$\mathcal{K}_4 = \frac{\langle \delta^4 \rangle - 3[\langle \delta^2 \rangle]^2}{[\langle \delta^2 \rangle]^3} = -\frac{15}{14}, \quad (\text{C12})$$

which is also mass independent. A negative excess kurtosis implies that the minicluster density field has a wider peak than a Gaussian. We are not aware of any study of the effect of kurtosis on the halo mass function.

We emphasize that the results presented here treat the axion field as a free random field and model the large-scale effects of the non-Gaussianity due to white noise on the formation of MCHs. These results do not model the effects of the potential or the gradient energy, and only account for the purely random large-scale effects. On large scales the gradients can be neglected and their effect on smoothing the perturbations on small scales via the Kibble mechanism is included in the ansatz for $B(k_1, k_2)$. On large scales the axion field values vary little, and the potential can be neglected. On smaller scales and at times $T \approx \mu$, the potential cannot be neglected and is important in establishing the exact form of the distribution $\mathcal{F}(\delta > \delta_0)$ and the formation of miniclusters and axion stars on scales $M \lesssim M_0$. Such effects can only be addressed by a combination of lattice field theory and N -simulations beyond the scope of the present work. For us, these effects fall under the umbrella of “cutoff dependence.”

APPENDIX D: AXION STARS

We have treated miniclust density profiles as if axions are entirely cold. On small scales, however, scalar field dynamics becomes important [70,83,114–117] and the final ingredient in any axion DM halo is the central solitonic core, or axion star. The axion star is formed when gradient pressure in the axion field is sufficient to halt gravitational collapse, leading to a stable (in the nonrelativistic limit) ground state solution (for axion star formation in halos, see the simulations of Ref. [114]; for relativistic corrections and stability see Refs. [118–121]; for interactions of axion stars with nuclear matter see Ref. [122]). We now briefly assess the role of axion stars in miniclusts, and whether they might modify any of our conclusions. We treat only the case where axion self-interactions can be neglected: the stable branch of dilute axion stars where collapse is driven by gravity.

The soliton is the ground state solution of the Schrödinger-Poisson equations. For a given axion mass, the soliton solution is specified entirely by the soliton mass. The mass-radius relation is (e.g. Ref. [114])

$$r_{\text{sol}} \simeq 1.54 \times 10^5 \left(\frac{M_{\odot}}{M_{\text{sol}}} \right) \left(\frac{10^{-10} \text{ eV}}{m_a} \right)^2 \text{ cm}. \quad (\text{D1})$$

When an axion density perturbation becomes large enough, the axions inside will condense into the soliton solution under the influence of gravity. We estimate the value of δ for which a miniclust condenses to an axion star by equating r_{sol} to the hard-sphere miniclust radius, Eq. (43).

We set $M_{\text{sol}} = M_0(m_a)$ and find δ_{sol} such that $r_{\text{sol}}(m_a) = r_{\text{mc}}(\delta_{\text{sol}}, m_a)$. Using the numerical results for $M_0(m_a, n)$ we solve numerically for δ_{sol} and plot the results for different n in Fig. 23. For temperature-dependent axion masses with $n = 3.34$ (the QCD axion) and $n = 6$, we find that enormously large values of δ are required for soliton formation. For such axions, since $\delta_{\text{sol}} > \delta_{\text{lens}}$ giving the transition in the behavior of the lensing tube parameter \mathcal{R} , soliton formation can be safely neglected in the lensing. The solitons are deep inside the pointlike regime with $\mathcal{R} \approx 1$. This implies that numerical simulations of miniclusts for these values of n will be free from large effects due to scalar field dynamics.

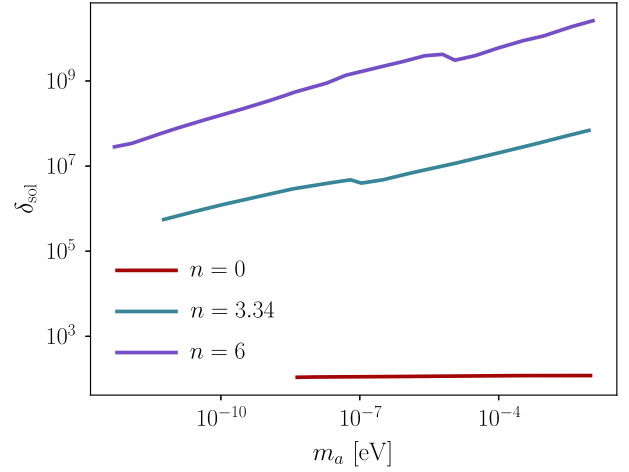


FIG. 23. Axion star formation. We find the critical value δ_{sol} such that the miniclust radius is equal to the soliton radius. For $\delta > \delta_{\text{sol}}$ miniclusts condense into axion stars.

The situation is quite different for axionlike particles with a temperature-independent mass, $n = 0$, where we find $\delta_{\text{sol}} \approx 10^2$ approximately independent of axion mass. This smaller value of δ compared to $n \neq 0$ arises because miniclusts with $n = 0$ are much lighter for fixed m_a . Such axions begin oscillating much earlier in cosmic history when the horizon is smaller, and thus the mass contained within the horizon is also much lower. The lighter miniclusts are closer to the soliton mass.

Using the analytic result for $M_0(m_a)$ for $n = 0$ [Eq. (3)] we can derive the value of δ_{sol} analytically. The axion mass dependence drops out of the ratio $r_{\text{mc}}/r_{\text{sol}}$ and we find $\delta_{\text{sol}} \approx 120$ independent of axion mass. This value of δ_{sol} is near the critical boundary for microlensing (Figs. 12 and 13) and implies that our results for axionlike particles with $n = 0$ could be altered by axion star formation. Miniclusts composed of such axions cannot become any denser without gaining mass, and this will reduce the expected number of microlensing events by effectively truncating the distribution \mathcal{F} for $\delta \gtrsim 120$. Furthermore, this implies that numerical simulations of miniclusts with $n = 0$ initial conditions cannot neglect scalar field dynamics.

-
- [1] G. Jungman, M. Kamionkowski, and K. Griest, *Phys. Rep.* **267**, 195 (1996).
 - [2] G. Bertone, D. Hooper, and J. Silk, *Phys. Rep.* **405**, 279 (2005).
 - [3] P. Sikivie, in *Axions*, Lecture Notes in Physics Vol. 741, edited by M. Kuster, G. Raffelt, and B. Beltrán (Springer Verlag, Berlin, 2008), p. 19.

- [4] D. J. E. Marsh, *Phys. Rep.* **643**, 1 (2016).
- [5] C. J. Hogan and M. J. Rees, *Phys. Lett. B* **205**, 228 (1988).
- [6] B. Paczynski, *Astrophys. J.* **304**, 1 (1986).
- [7] K. Griest, *Astrophys. J.* **366**, 412 (1991).
- [8] S. Hawking, *Mon. Not. R. Astron. Soc.* **152**, 75 (1971).
- [9] B. J. Carr and S. W. Hawking, *Mon. Not. R. Astron. Soc.* **168**, 399 (1974).

- [10] B. J. Carr, *Astrophys. J.* **201**, 1 (1975).
- [11] R. Peccei and H. R. Quinn, *Phys. Rev. Lett.* **38**, 1440 (1977).
- [12] Planck Collaboration, *Astron. Astrophys.* **594**, A20 (2016).
- [13] S. Weinberg, *Phys. Rev. Lett.* **40**, 223 (1978).
- [14] F. Wilczek, *Phys. Rev. Lett.* **40**, 279 (1978).
- [15] J. E. Kim, *Phys. Rev. Lett.* **43**, 103 (1979).
- [16] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, *Nucl. Phys.* **B166**, 493 (1980).
- [17] A. Zhitnitsky, *Sov. J. Nucl. Phys.* **31**, 260 (1980).
- [18] M. Dine, W. Fischler, and M. Srednicki, *Phys. Lett.* **104B**, 199 (1981).
- [19] E. W. Kolb and I. I. Tkachev, *Phys. Rev. Lett.* **71**, 3051 (1993).
- [20] E. W. Kolb and I. I. Tkachev, *Phys. Rev. D* **49**, 5040 (1994).
- [21] E. W. Kolb and I. I. Tkachev, *Phys. Rev. D* **50**, 769 (1994).
- [22] E. W. Kolb and I. I. Tkachev, *Astrophys. J.* **460**, L25 (1996).
- [23] P. Svrcek and E. Witten, *J. High Energy Phys.* **6** (2006) 051.
- [24] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, *Phys. Rev. D* **81**, 123530 (2010).
- [25] K. M. Zurek, C. J. Hogan, and T. R. Quinn, *Phys. Rev. D* **75**, 043511 (2007).
- [26] E. Hardy, *J. High Energy Phys.* **2** (2017) 46.
- [27] M. Fairbairn, D. J. E. Marsh, and J. Quevillon, *Phys. Rev. Lett.* **119**, 021101 (2017).
- [28] S. J. Asztalos *et al.*, *Phys. Rev. Lett.* **104**, 041301 (2010).
- [29] MADMAX Working Group, *arXiv:1611.04549*.
- [30] MADMAX Working Group, *Phys. Rev. Lett.* **118**, 091801 (2017).
- [31] A. Arvanitaki and A. A. Geraci, *Phys. Rev. Lett.* **113**, 161801 (2014).
- [32] J. Moody and F. Wilczek, *Phys. Rev. D* **30**, 130 (1984).
- [33] P. Tinyakov, I. Tkachev, and K. Zioutas, *J. Cosmol. Astropart. Phys.* **1** (2016) 035.
- [34] C. A. J. O'Hare and A. M. Green, *Phys. Rev. D* **95**, 063017 (2017).
- [35] W. H. Press and P. Schechter, *Astrophys. J.* **187**, 425 (1974).
- [36] P. Tisserand *et al.*, *Astron. Astrophys.* **469**, 387 (2007).
- [37] H. Niikura *et al.*, *arXiv:1701.02151*.
- [38] L. Di Luzio, F. Mescia, and E. Nardi, *Phys. Rev. Lett.* **118**, 031801 (2017).
- [39] Planck Collaboration, *Astron. Astrophys.* **594**, A13 (2016).
- [40] G. Ballesteros, J. Redondo, A. Ringwald, and C. Tamarit, *J. Cosmol. Astropart. Phys.* **08** (2017) 001.
- [41] P. Sikivie, *Phys. Rev. Lett.* **48**, 1156 (1982).
- [42] D. Harari and P. Sikivie, *Phys. Lett. B* **195**, 361 (1987).
- [43] T. W. B. Kibble, *J. Phys. A* **9**, 1387 (1976).
- [44] E. Komatsu *et al.*, *Astrophys. J. Suppl. Ser.* **180**, 330 (2009).
- [45] D. J. Gross, R. D. Pisarski, and L. G. Yaffe, *Rev. Mod. Phys.* **53**, 43 (1981).
- [46] J. Preskill, M. B. Wise, and F. Wilczek, *Phys. Lett.* **120B**, 127 (1983).
- [47] L. F. Abbott and P. Sikivie, *Phys. Lett.* **120B**, 133 (1983).
- [48] M. Dine and W. Fischler, *Phys. Lett.* **120B**, 137 (1983).
- [49] O. Wantz and E. Shellard, *Phys. Rev. D* **82**, 123508 (2010).
- [50] S. Borsanyi, M. Dierigl, Z. Fodor, S. D. Katz, S. W. Mages, D. Nogradi, J. Redondo, A. Ringwald, and K. K. Szabo, *Phys. Lett. B* **752**, 175 (2016).
- [51] S. Borsanyi *et al.*, *Nature (London)* **539**, 69 (2016).
- [52] T. Hiramatsu, M. Kawasaki, K. Saikawa, and T. Sekiguchi, *Phys. Rev. D* **85**, 105020 (2012).
- [53] J. A. Peacock, *Cosmological Physics* (Cambridge University Press, Cambridge, England, 1999).
- [54] D. B. Thomas, M. Kopp, and C. Skordis, *Astrophys. J.* **830**, 155 (2016).
- [55] D. J. E. Marsh, D. Grin, R. Hlozek, and P. G. Ferreira, *Phys. Rev. D* **87**, 121701 (2013).
- [56] R. Hlozek, D. Grin, D. J. E. Marsh, and P. G. Ferreira, *Phys. Rev. D* **91**, 103512 (2015).
- [57] G. Efstathiou and J. R. Bond, *Mon. Not. R. Astron. Soc.* **218**, 103 (1986).
- [58] M. Khlopov, B. Malomed, and I. Zeldovich, *Mon. Not. R. Astron. Soc.* **215**, 575 (1985).
- [59] E. W. Kolb and M. S. Turner, *The Early Universe*, Frontiers in Physics Vol. 69 (1990), p. 1–547.
- [60] R. K. Sheth and G. Tormen, *Mon. Not. R. Astron. Soc.* **308**, 119 (1999).
- [61] M. Musso and R. K. Sheth, *Mon. Not. R. Astron. Soc.* **439**, 3051 (2014).
- [62] J. Wang and S. D. M. White, *Mon. Not. R. Astron. Soc.* **380**, 93 (2007).
- [63] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, *Astrophys. J.* **304**, 15 (1986).
- [64] A. Schneider, R. E. Smith, and D. Reed, *Mon. Not. R. Astron. Soc.* **433**, 1573 (2013).
- [65] R. Barkana, Z. Haiman, and J. P. Ostriker, *Astrophys. J.* **558**, 482 (2001).
- [66] D. J. E. Marsh and J. Silk, *Mon. Not. R. Astron. Soc.* **437**, 2652 (2014).
- [67] A. J. Benson, A. Farahi, S. Cole, L. A. Moustakas, A. Jenkins, M. Lovell, R. Kennedy, J. Helly, and C. Frenk, *Mon. Not. R. Astron. Soc.* **428**, 1774 (2013).
- [68] A. R. Pullen, A. J. Benson, and L. A. Moustakas, *Astrophys. J.* **792**, 24 (2014).
- [69] X. Du, C. Behrens, and J. C. Niemeyer, *Mon. Not. R. Astron. Soc.* **465**, 941 (2017).
- [70] H.-Y. Schive, T. Chiueh, and T. Broadhurst, *Nat. Phys.* **10**, 496 (2014).
- [71] J. Veltmaat and J. C. Niemeyer, *Phys. Rev. D* **94**, 123523 (2016).
- [72] H.-Y. Schive, T. Chiueh, T. Broadhurst, and K.-W. Huang, *Astrophys. J.* **818**, 89 (2016).
- [73] P. S. Corasaniti, S. Agarwal, D. J. E. Marsh, and S. Das, *Phys. Rev. D* **95**, 083512 (2017).
- [74] J. F. Navarro, C. S. Frenk, and S. D. White, *Astrophys. J.* **490**, 493 (1997).
- [75] J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel, *Mon. Not. R. Astron. Soc.* **321**, 559 (2001).
- [76] I. King, *Astron. J.* **67**, 471 (1962).
- [77] A. J. Benson, *New Astron.* **17**, 175 (2012).
- [78] B. Paczynski, *Astrophys. J.* **304**, 1 (1986).
- [79] C. Alcock *et al.*, *Astrophys. J.* **461**, 84 (1996).
- [80] E. Bertschinger, *Astrophys. J. Suppl. Ser.* **58**, 39 (1985).

- [81] R. Narayan and M. Bartelmann, in 13th Jerusalem Winter School in Theoretical Physics: Formation of Structure in the Universe Jerusalem, Israel, 1996, 1996, [arXiv:astro-ph/9606001](#).
- [82] A. M. Green, *Phys. Rev. D* **94**, 063530 (2016).
- [83] L. Hui, J. P. Ostriker, S. Tremaine, and E. Witten, *Phys. Rev. D* **95**, 043541 (2017).
- [84] A. Klypin, H. Zhao, and R. S. Somerville, *Astrophys. J.* **573**, 597 (2002).
- [85] R. E. Angulo and S. D. M. White, *Mon. Not. R. Astron. Soc.* **401**, 1796 (2010).
- [86] J. Niemeyer, 13th Axion-WIMP Conference (Patras Workshop) (2017), <http://inspirehep.net/record/1476008>.
- [87] G. Aslanyan, L. C. Price, J. Adams, T. Bringmann, H. A. Clark, R. Easther, G. F. Lewis, and P. Scott, *Phys. Rev. Lett.* **117**, 141102 (2016).
- [88] M. Sten Delos, A. L. Erickcek, A. P. Bailey, and M. A. Alvarez, [arXiv:1712.05421](#).
- [89] T. Bringmann, P. Scott, and Y. Akrami, *Phys. Rev. D* **85**, 125027 (2012).
- [90] I. I. Tkachev, *JETP Lett.* **101**, 1 (2015).
- [91] A. Iwazaki, *Phys. Rev. D* **91**, 023008 (2015).
- [92] A. Iwazaki, [arXiv:1707.04827](#).
- [93] E. C. Bellm, [arXiv:1410.8185](#).
- [94] P. A. Abell *et al.* (LSST Science Collaboration), [arXiv:0912.0201](#).
- [95] BICEP2 Collaboration, *Phys. Rev. Lett.* **116**, 031302 (2016).
- [96] M. Kawasaki, K. Saikawa, and T. Sekiguchi, *Phys. Rev. D* **91**, 065014 (2015).
- [97] A. G. Dias, A. C. B. Machado, C. C. Nishi, A. Ringwald, and P. Vaudrevange, *J. High Energy Phys.* **6** (2014) 37.
- [98] J. E. Kim and D. J. E. Marsh, *Phys. Rev. D* **93**, 025027 (2016).
- [99] D. H. Lyth, *Phys. Rev. D* **45**, 3394 (1992).
- [100] M. S. Turner, *Phys. Rev. D* **33**, 889 (1986).
- [101] R. L. Davis, *Phys. Rev. D* **32**, 3172 (1985).
- [102] R. A. Battye and E. P. S. Shellard, *Phys. Rev. Lett.* **73**, 2954 (1994).
- [103] R. A. Battye and E. P. S. Shellard, *Phys. Rev. Lett.* **76**, 2203 (1996).
- [104] P. Fox, A. Pierce, and S. Thomas, [arXiv:hep-th/0409059](#).
- [105] D. E. Kaplan and R. Rattazzi, *Phys. Rev. D* **93**, 085007 (2016).
- [106] G. F. Giudice and M. McCullough, *J. High Energy Phys.* **02** (2017) 036.
- [107] T. Higaki, K. S. Jeong, N. Kitajima, T. Sekiguchi, and F. Takahashi, *J. High Energy Phys.* **8** (2016) 44.
- [108] A. R. Zentner, *Int. J. Mod. Phys. D* **16**, 763 (2007).
- [109] B. Bozek, D. J. E. Marsh, J. Silk, and R. F. G. Wyse, *Mon. Not. R. Astron. Soc.* **450**, 209 (2015).
- [110] D. J. E. Marsh, [arXiv:1605.05973](#).
- [111] M. Maggiore and A. Riotto, *Astrophys. J.* **717**, 526 (2010).
- [112] S. Matarrese, L. Verde, and R. Jimenez, *Astrophys. J.* **541**, 10 (2000).
- [113] M. LoVerde, A. Miller, S. Shandera, and L. Verde, *J. Cosmol. Astropart. Phys.* **4** (2008) 014.
- [114] H.-Y. Schive, M.-H. Liao, T.-P. Woo, S.-K. Wong, T. Chiueh, T. Broadhurst, and W.-Y. Pauchy Hwang, *Phys. Rev. Lett.* **113**, 261302 (2014).
- [115] D. J. E. Marsh and A.-R. Pop, *Mon. Not. R. Astron. Soc.* **451**, 2479 (2015).
- [116] B. Schwabe, J. C. Niemeyer, and J. F. Engels, *Phys. Rev. D* **94**, 043513 (2016).
- [117] P. Mocz, M. Vogelsberger, V. H. Robles, J. Zavala, M. Boylan-Kolchin, A. Fialkov, and L. Hernquist, *Mon. Not. R. Astron. Soc.* **471**, 4559 (2017).
- [118] J. Eby, P. Suranyi, and L. C. R. Wijewardhana, *Mod. Phys. Lett. A* **31**, 1650090 (2016).
- [119] D. G. Levkov, A. G. Panin, and I. I. Tkachev, *Phys. Rev. Lett.* **118**, 011301 (2017).
- [120] T. Helfer, D. J. E. Marsh, K. Clough, M. Fairbairn, E. A. Lim, and R. Becerril, *J. Cosmol. Astropart. Phys.* **3** (2017) 055.
- [121] J. Eby, M. Ma, P. Suranyi, and L. C. R. Wijewardhana, [arXiv:1705.05385](#).
- [122] Y. Bai, V. Barger, and J. Berger, *J. High Energy Phys.* **12** (2016) 127.