Imprints of the early Universe on axion dark matter substructure

Nikita Blinov,[1,2] Matthew J. Dolan,[3] and Patrick Draper[4]

[1]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[2]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA*
[3]*ARC Centre of Excellence for Particle Physics at the Terascale, School of Physics, University of Melbourne, 3010, Australia*
[4]*Department of Physics, University of Illinois, Urbana, Illinois 61801, USA*

(Received 29 November 2019; accepted 14 January 2020; published 5 February 2020)

Despite considerable experimental progress large parts of the axionlike particle (ALP) parameter space remain difficult to probe in terrestrial experiments. In some cases, however, small-scale structure of the ALP dark matter (DM) distribution is strongly enhanced, offering opportunities for astrophysical tests. Such an enhancement can be produced by a period of prenucleosynthesis early matter domination (EMD). This cosmology arises in many ultraviolet completions and generates the correct relic abundance for weak coupling $f_a \sim 10^{16}$ GeV, ALP masses in the range 10^{-13} eV $< m_a < 1$ eV, and without fine-tuning of the initial misalignment angle. This range includes the QCD axion around $10^{-9} - 10^{-8}$ eV. EMD enhances the growth of ALP small-scale structure, leading to the formation of dense ALP miniclusters which can contain nearly all of DM (depending on ALP mass and reheating temperature). We study the interplay between the initial ALP oscillation, reheating temperature, and effective pressure to provide analytic estimates of the minicluster abundance and properties. ALP miniclusters in the EMD cosmology are denser and more abundant than in ΛCDM. While enhanced substructure generically reduces the prospects of direct detection experiments, we show that pulsar timing and lensing observations can discover these minihalos over a large range of ALP masses and reheating temperatures.

DOI: 10.1103/PhysRevD.101.035002

I. INTRODUCTION

Axionlike particles (ALPs) provide a compelling and elegant explanation for the dark matter (DM) of the Universe [1–4]. These DM candidates arise in ultraviolet (UV) completions of the Standard Model (SM) as pseudo-Nambu-Goldstone bosons of spontaneously broken global symmetries, or as zero modes of higher-dimensional gauge fields [4–7]. As such, their masses can be naturally light. A relic density of ALPs can be produced via several mechanisms, including misalignment, thermal and inflationary production, and from the decays of heavier particles or topological defects. In many well-motivated cases several mechanisms can contribute. For reviews, see, e.g., Refs. [8,9].

The nonthermal production of ALPs suggests that both their abundance and late-time distribution are sensitive to physics in the UV. This is in contrast to thermally produced DM (such as weakly interacting massive particles, or WIMPs), where the abundance only depends on processes at energies similar to the DM mass. This UV sensitivity can open a window into the prenucleosynthesis universe, where few other probes are currently available.

One of the principal means for discovering and measuring the properties of ALPs is through terrestrial direct detection experiments. Previously we have studied prospects for the direct detection of ALPs in a variety of cosmological scenarios [10], providing a range of experimental targets which are free of fine-tuning; some of these targets are shown in Fig. 1. The most difficult scenario to probe via direct detection involves a period of early matter domination (EMD). In EMD scenarios, the value of the ALP decay constant required to achieve the correct DM relic density is large, corresponding to a small value of the ALP-photon coupling $g_{a\gamma\gamma} \sim 10^{-18}$ GeV^{-1}. This value is independent of the ALP mass over a large range of masses, including a range of QCD axion candidates around a nano-eV.

Although challenging for ALP direct detection experiments, a period of EMD is an interesting possibility for early Universe physics, motivated both by top-down model building and by bottom-up phenomenological considerations.

From the phenomenological perspective, the prenucleosynthesis expansion history has not been determined by observations, and it is natural to consider alternatives to the standard radiation-dominated assumption. Among these alternatives, a period of matter domination is perhaps the simplest possibility. Furthermore, as alluded to above, EMD is a natural way to achieve the correct relic abundance of axions with decay constants around the grand unified theory (GUT) scale *independent of the axion mass* [10,20]. Whereas in the standard cosmology the axion fraction of

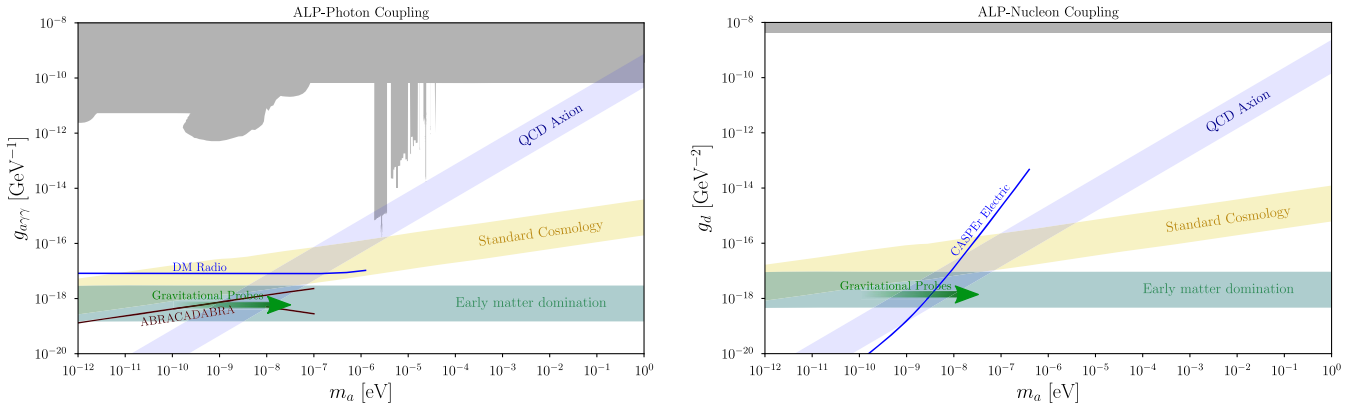

FIG. 1. Existing constraints (gray regions), target parameter space (colored bands) and sensitivity of future experiments (colored lines) in the ALP mass m_a and photon coupling $g_{a\gamma\gamma} \sim 1/f_a$ (left panel) and nucleon dipole moment coupling $g_d \sim 1/(f_a \Lambda_{\rm QCD})$ (right panel) planes. The QCD axion band is shown in blue and corresponds to a particular approximate mass-coupling relation, $m_a f_a \sim m_\pi f_\pi$. Relaxing this relationship but imposing saturation of the dark matter relic density results in other mass-coupling relations. These relations depend on the pre-BBN expansion history and initial misalignment angle θ_i. The green and yellow bands correspond to cosmologies with early matter domination (with a reheat temperature of 10 MeV) and standard radiation domination before nucleosynthesis and natural values of $\theta_i \sim 1$. Larger masses and lower couplings favored by EMD are challenging to probe in terrestrial experiments. In this paper we show that EMD enhances ALP DM small-scale structure, resulting in the formation of ALP minihalos. These minihalos can be probed with lensing and pulsar timing observations through gravitational interactions alone for $m_a \gtrsim 10^{-10}$ eV as indicated by the green arrow. We also show the far-future sensitivities of DM Radio [11,12] and ABRACADABRA [13–15] to the ALP-photon coupling and CASPER-electric to the ALP-nucleon coupling [16–18]. These experiments are projected to reach the EMD target region for a range of reheating temperatures (the upper and lower ABRACADABRA lines correspond to the broadband and resonant searches, respectively [15]). The left panel is adapted from Ref. [10] including results of Ref. [19].

the energy density grows linearly with temperature from $H \sim m_a$ until matter-radiation equality, generally leading to overclosure for high-scale decay constants unless the axion is extremely light, in EMD this energy fraction is frozen around $(f_a/M_{\rm Pl})^2$ until reheating. The correct relic abundance is then obtained for $T_{\rm eq}/T_{\rm RH} \sim (f_a/M_{\rm Pl})^2$, independent of m_a.

From the top-down perspective, EMD is thought to occur generically in string models of light axions [20], where axions typically have GUT-scale decay constants [5] and are accompanied by a saxion (a heavier scalar modulus partner of the axion) that comes to dominate the energy density at early times. Saxion moduli in these models typically have Planck-suppressed couplings, and there can also be other modulus fields which are similarly weakly coupled (see, for example, the textbook discussion of Refs. [21,22] or the models discussed in Refs. [5,23,24]). Moduli fields are thus naturally long-lived, and the energy stored in their oscillations can come to dominate the energy density of the early Universe before they decay, leading to a period of EMD. If the moduli decay happens too late, the energy injection can ruin the success of big bang nucleosynthesis (BBN): this is known as the cosmological moduli problem [25–27]. The moduli problem can be avoided if the moduli mass scale is above $\mathcal{O}(10)$ TeV, leading to reheating (RH) above 5 MeV and satisfying BBN constraints [28–31].

A period of EMD can modify dark matter physics in various ways. In WIMP scenarios, dark matter can be produced in the decay of the fields responsible for EMD, favoring different parts of the supersymmetric parameter space [32–36]. In ALP scenarios, axions that begin to oscillate during an EMD phase have modified relic densities and perturbation growth relative to the standard radiation dominated cosmology. The authors of Ref. [37] performed an early study of axion dark matter with a period of EMD; more recent studies include Refs. [38–41] and projections for a wide range of proposed experiments for this scenario were given in Ref. [10].

In this work we explore the impact of a stage of early matter domination on the growth of ALP density perturbations, which has also recently been considered in Refs. [39,40] for ALPs and in Refs. [42–46] for WIMPs. Density perturbations grow linearly with the scale factor during EMD, as opposed to logarithmically during radiation domination. Linear growth leads to enhanced structure on scales that enter the horizon before the end of EMD. These structures decouple from the Hubble flow and collapse at high redshifts, leading to the formation of ALP miniclusters or minihalos (we use these terms interchangeably) [47–52]. Galactic dark matter halos are then hierarchically assembled from these miniclusters. If the miniclusters survive the galactic assembly process, their presence can significantly alter the optimal DM search strategy. For example, the terrestrial minicluster encounter rate may be too low for effective direct detection searches [53,54]. On the other hand, such compact structures can be searched for via lensing and pulsar timing. Future lensing

[55–57] and pulsar-timing searches [58] will be able to probe compact DM substructure at an unprecedented level.

The formation of ALP miniclusters is known to occur for ALP initial conditions generated by postinflationary Pecci-Quinn (PQ) breaking, topological defect decay or when ALP self-interactions are important. In this work, we focus on the complementary case where the initial ALP perturbations are small $\mathcal{O}(10^{-4})$, in analogy to the preinflationary PQ breaking scenario for the QCD axion. We show that even in this case, miniclusters form due to the enhanced growth during EMD. Unlike the post-PQ breaking case, the growth of ALP perturbations can be treated analytically. Consequently the minicluster distributions can be characterized with mild assumptions about their survival probability. We generalize and extend previous analyses of Refs. [39,40] by considering non-QCD ALPs, treating growth before and after matter-radiation equality, and studying the impact of the ALP mass, sound speed effects, reheating temperature, and initial conditions on the growth and distribution of miniclusters.

Recent numerical work [59] suggests QCD axion miniclusters formed from postinflationary PQ breaking in the standard cosmology may be too light to be relevant for lensing and pulsar timing searches. In contrast, in the ALP parameter space where EMD provides a natural explanation for the relic density, we find that these searches, particularly photometric microlensing, offer strong sensitivity. In EMD scenarios it is plausible that most of the present-day relic density is bound in miniclusters, greatly weakening direct detection prospects. These conclusions affect a large range of weakly coupled ALP models, including a QCD axion with a mass around an neV.

This work is organized as follows. In Sec. II we discuss the ALP relic density and a model for the period of early matter domination. In Sec. III we assemble the Boltzmann equations for the evolution of the background energy densities and the ALP perturbations. The physics involves a number of different scales: the ALP mass (and hence scale of oscillation), the comoving horizon size, the ALP Jeans scale, and the reheating scale. We carefully assess the impact of each scale on the growth of perturbations, and we present numerical results tracking the perturbations from early matter domination, through reheating and into standard radiation domination, and through standard matter-radiation equality. In Sec. IV we use the Press-Schechter formalism to estimate the statistical distributions of the ALP miniclusters. We discuss our results in Sec. V, and study their implications for the minicluster survival rate, direct detection, and pulsar timing and lensing searches. We conclude in Sec. VI.

II. ALP RELIC ABUNDANCE

Our model consists of the Standard Model, an ALP field a which will constitute the dark matter, and a heavy scalar field ϕ. The energy density in coherent ϕ oscillations dominates the Universe at early times and redshifts like matter, leading to a period of EMD. The scalar field a is minimally coupled to gravity with action

$$S \supset \int d^4x\sqrt{-g}\left[\frac{1}{2}(\partial a)^2 - V(a)\right], \tag{1}$$

where for simplicity we consider a quadratic potential

$$V(a) = \frac{1}{2}m^2a^2. \tag{2}$$

If the field a is an axionlike particle then we expect its interactions with SM fields to be suppressed by factors of $1/f_a$, where f_a is the ALP decay constant. While critical for direct detection, these interactions are not relevant for the early Universe cosmology we study in this paper. We assume that the ALP relic density is set through the misalignment mechanism [1–3], and that the ALP mass is independent of temperature.

Higher-order terms may also be present in the potential of Eq. (2). These terms delay the onset of ALP oscillations for large initial values of a, and including them can have $\mathcal{O}(1)$ impact on the relic density. For $a/f_a \lesssim 1$ these effects are unimportant. Nonlinear terms in the ALP equation of motion can also lead to important effects on small scales, such as solitonlike configuration known as axitons or oscillons [48,59–62]. The precise effect of self-interactions depends on the interaction terms and temperature dependence of the ALP mass. On the whole, the presence of axitons could provide additional substructure within the miniclusters we identify below and would be interesting to pursue further through dedicated numerical studies.

In a Friedmann-Robertson-Walker cosmology, the equation of motion for the ALP background is

$$\frac{d^2a}{dt^2} + 3H\frac{da}{dt} + m_a^2a = 0, \tag{3}$$

where H is the Hubble parameter. The evolution of the ALP field depends on the distribution of initial conditions $\theta_i = a_i/f_a$. If the ALP exists prior to inflation then θ_i is uniform across the initially causally disconnected regions constituting the Universe at the present time. This is the scenario we study in this paper. Another possibility is that θ_i is stochastically distributed over the separate causal patches throughout the Universe. In models where the ALP is a pseudo-Nambu Goldstone boson, this corresponds to the scenario where associated global symmetry is broken after inflation. This can be studied analytically using cosmological perturbation theory by taking an effective average misalignment angle corresponding to $\theta_i = \pi/\sqrt{3}$ [63]. However, one also expects the formation of topological defects such as strings and domain walls at the boundaries of different causal patches which are not captured by this approach. The decay of these defects

leads to large fluctuations in the ALP field which later evolve into miniclusters. Structure formation in this scenario has recently been studied numerically in Refs. [59,60] and analytically in Refs. [51,52,64] using the Press-Schechter formalism, and leads to the formation of miniclusters [47–50].

It is well known that the preinflationary PQ-breaking axion scenario generates isocurvature perturbations, which are strongly constrained by cosmic microwave background (CMB) measurements [65–67]. Suppressing these modes either implies an upper bound on the scale of inflation H_I [68] or requires nontrivial axion-inflaton dynamics (see, e.g., Ref. [68]). In the former case we estimate in Appendix D that $H_I \lesssim 10^{9-10}$ GeV depending on the reheating temperature. This is less constraining than for the QCD axion in a standard cosmology [38]. For the rest of this paper we assume that inflation has taken place at a sufficiently low scale to satisfy the isocurvature constraint.

The scalar field ϕ comes to dominate the energy density in the early Universe before it decays. Reheating occurs when the Hubble parameter is approximately equal to the decay width of the scalar, $H \sim \Gamma_\phi$. During matter-domination $H \propto a^{-3/2}$, and the scale factor a at reheating is approximately $a_{\rm RH} \sim \Gamma_\phi^{-2/3}$.[1] We denote the temperature of the Universe when reheating occurs by $T_{\rm RH}$. This is constrained by big bang nucleosynthesis to be larger than $\mathcal{O}(\text{MeV})$. The lowest reheat temperature we consider in this work is 5 MeV [28–31].

In UV-complete models the field ϕ could correspond to a saxion or modulus field. We assume that the ϕ decays predominantly into Standard Model fields, corresponding to radiation in the early Universe. It is also possible that ϕ decays into ALPs. In that case the ALP relic density would be made up partly from a population due to misalignment and partly from a population due to ϕ decay. Whether this population behaves as matter or radiation depends on the relative mass of ϕ and a, and on $T_{\rm RH}$. However, for $m_a <$ eV and the low reheat temperatures we are interested in, ALPs produced from ϕ decays are still relativistic at matter-radiation equality (MRE). Therefore, they contribute to the total energy density of the Universe as dark radiation and their abundance is constrained by the concordance of standard cosmology with observations of the CMB and light element abundances. These constraints are conveniently expressed as bounds on the effective number of relativistic degrees of freedom, $N_{\rm eff}$. In the instantaneous decay approximation we estimate that ALPs from ϕ decays would contribute

$$\Delta N_{\rm eff} \lesssim \frac{4}{7}\left[\frac{\mathrm{BR}(\phi \to aa)}{\mathrm{BR}(\phi \to \mathrm{SM\,SM})}\right] g_*(T_{\rm RH}), \tag{4}$$

where the upper bound arises from assuming $T_{\rm RH} \lesssim 10$ MeV; higher $T_{\rm RH}$ is more weakly constrained due to additional SM entropy injections which dilute the relativistic ALP contribution relative to the SM. CMB and BBN limit $\Delta N_{\rm eff} \lesssim 0.5$ at 95% C.L. (see, e.g., Refs. [69–71]), which translates into $\mathrm{BR}(\phi \to aa) < 0.08$. In the absence of self-interactions, this ALP component does not contribute to the formation of ALP clumps, so we set $\mathrm{BR}(\phi \to aa) = 0$ throughout this work. We also note that topological defects such as domain walls and strings tend to be irrelevant in preinflationary ALP scenarios, since inflation smooths out inhomogeneities. Recent numerical studies suggest that the magnitude of the defect contribution is small even in the postinflationary ALP scenario [59,60,72] (although significant uncertainty remains—see Refs. [59,73,74]). We therefore assume that the relic density is determined entirely by production through misalignment.

In the early Universe while $H > m_a$ the field a is effectively frozen in its initial value. As the Universe expands and Hubble decreases and becomes comparable with the ALP mass $H \sim m_a$ the ALP field starts evolving at time $t_{\rm osc}$ and oscillates in its potential. After oscillations begin, the ALP energy density redshifts as matter,

$$\rho_a = \frac{1}{2} m_a^2 f_a^2 \theta_i^2 (a(t_{\rm osc})/a)^3 + \mathcal{O}(H^2/m_a^2), \tag{5}$$

where the a^{-3} is the redshifting of the energy density with the scale factor.

The ALP density in the current epoch can be shown to be approximately [10]

$$\Omega_a h^2 \simeq 0.12 \times \left(\frac{f_a \theta_i}{9 \times 10^{14}\ \text{GeV}}\right)^2 \times \left(\frac{T_{\rm RH}}{10\ \text{MeV}}\right), \tag{6}$$

where $\theta_i = a_i/f_a$ is the initial ALP misalignment angle. This equation holds for temperature-independent ALP masses, and also assumes that the reheating temperature $T_{\rm RH}$ is lower than the ALP oscillation temperature $T_{\rm osc}$. If this were not the case then ALP oscillations would commence during radiation domination, and EMD would not have any impact on dark matter physics and structure formation. Notably, Eq. (6) is independent of the ALP mass: EMD is an efficient mechanism for preventing heavier ALPs from overclosing the Universe without fine-tuning the misalignment angle [10,75]. Similarly, it achieves the correct relic abundance for weaker couplings than in the standard scenario. In the next section we study the evolution of inhomogeneities in the ALP field and the corresponding cosmological density and velocity perturbations.

III. GROWTH OF ALP PERTURBATIONS

In this section we trace the evolution of the ALP perturbations through EMD, reheating, and into standard

[1]We use a to refer to both and ALP field and the cosmological scale factor. It will be clear from context which one is in use.

radiation domination. The initial ALP perturbations are small for our choice of initial conditions, so their growth during these stages is well described by linear perturbation theory. We determine the evolution of the background densities, then find the Boltzmann equations and initial conditions describing the perturbations in the EMD field, ALPs and radiation. We solve the Boltzmann system numerically and discuss the growth of perturbations of different physical sizes. Finally, we use these results to construct approximate transfer functions, which will be employed in the following section to characterize mini-cluster formation and distribution in the nonlinear regime.

Our analysis closely follows that of Ref. [43] for WIMP-like DM with EMD, so it is useful to highlight some differences from the WIMP case. First, since the ALP masses are low, if they are to account for the cold DM of the Universe, they cannot be produced during reheating in the decays of the massive particles responsible for EMD. Second, ALPs only behave like matter after they begin to oscillate. This sets a characteristic minimum scale for enhanced structure growth, and it modifies the Boltzmann equations compared to the WIMP scenario. Finally, the ALP coupling to the SM is so weak that there are no annihilations or decays of the ALP during reheating [45].

A. Background

We model the background evolution of the Universe as a three fluid system adapted from [43,76]. This scenario is described by the following evolution equations for the energy densities of the scalar field ρ_ϕ, the ALP ρ_a, and SM radiation ρ_r:

$$\frac{d\rho_\phi}{dt} + 3H\rho_\phi = -\Gamma_\phi \rho_\phi \tag{7a}$$

$$\frac{d\rho_r}{dt} + 4H\rho_r = +\Gamma_\phi \rho_\phi \tag{7b}$$

$$\frac{d\rho_a}{dt} + 3H\rho_a = 0. \tag{7c}$$

We will follow the conventions of Ref. [43] and set the scale factor and Hubble parameter at an initial time t_0 to be $a(t_0) = 1$ and $H(t_0) = H_1$. We also define dimensionless variables

$$\tilde{\rho}_i = \rho_i / \rho_{\mathrm{crit},0} \tag{8}$$

$$\tilde{\Gamma}_\phi = \Gamma_\phi / H_1, \tag{9}$$

where $\rho_{\mathrm{crit},0} = 3M_{\mathrm{Pl}}^2 H_1^2/(8\pi)$. We work in these scaled units throughout and drop the tildes below for simplicity. Physical quantities are ratios of scales and the dependence on H_1 drops out.

Equations (7) are valid after the ALP starts to oscillate; before this era, the energy density must be obtained by solving the field equation (3). We obtain the initial conditions for Eq. (7) by assuming that ϕ dominates the energy density at early times, that the dominant component of radiation has been produced by ϕ decay alone (i.e., any primordial contribution has been diluted away), and that the Universe is flat. These assumptions give

$$\rho_r(t_0) \approx \frac{2}{5}\Gamma_\phi \rho_\phi(t_0), \tag{10}$$

$$\rho_a(t_0) \approx \rho_a(t_0), \tag{11}$$

$$\rho_\phi(t_0) + \rho_a(t_0) + \rho_r(t_0) \approx 1, \tag{12}$$

where the last condition can be solved for $\rho_\phi(t_0) \approx 1 - \mathcal{O}(\Gamma_\phi)$. Since we are considering times well before the standard matter-radiation equality, we have $\rho_a(t_0) \ll \rho_{\phi,r}(t_0)$.

The background evolution, Eq. (7), can be solved in the early-time limit:

$$\rho_\phi(t) \approx \rho_\phi(t_0) a^{-3}, \tag{13a}$$

$$\rho_a(t) \approx \rho_a(t_0) a^{-3}, \tag{13b}$$

$$\rho_r(t) \approx \rho_r(t_0) a^{-3/2}. \tag{13c}$$

The ALP and ϕ redshift as matter, but the radiation density redshifts slower than the usual a^{-4} due to the ϕ decays. These solutions are approximately valid until reheating when $\Gamma_\phi / H(a_{\mathrm{RH}}) \sim 1$, which occurs when

$$a_{\mathrm{RH}} \sim \Gamma_\phi^{-2/3}. \tag{14}$$

Compared to Ref. [43], we assume in Eq. (7) that the DM component does not arise from decays of ϕ [$f = \mathrm{BR}(\phi \to aa) = 0$ in their notation]. Instead, the ALP initial condition is set by requiring that we get the correct DM abundance by the time of (standard) matter-radiation equality. A representative numerical solution of the system in Eq. (7) is shown in Fig. 2. Before it commences oscillating, the axion energy density behaves as a component of dark energy and is constant. In this section we have presented the background evolution equations in terms of the time t. In all sections after this we will use either the scale factor a, or conformal time τ as our time variables.

B. Perturbations and initial conditions

Density and velocity perturbations in the ϕ and radiation fluids are governed by the Einstein and stress-energy conservation equations in the perturbed FRW space-time. The ALP fluid behaves as a cosmological constant before oscillations commence and as matter afterwards. Even deep

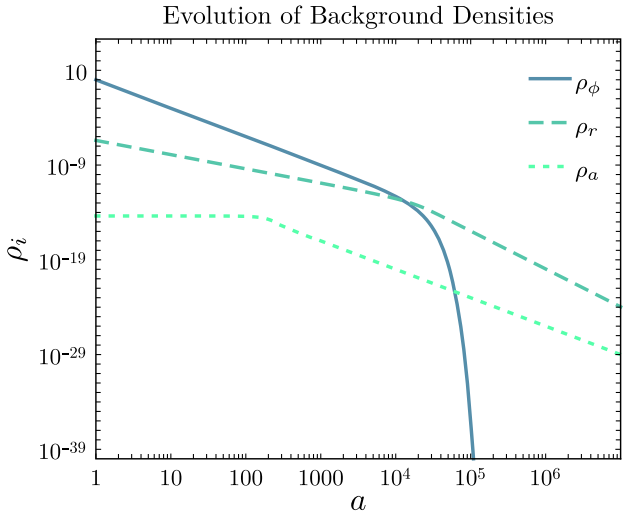

FIG. 2. Evolution of background densities for a cosmology with early matter domination. At early times, the energy budget of the Universe is dominated by a nonrelativistic field ϕ, which eventually decays into radiation (at $a = a_{\rm RH} \approx 10^4$), reheating the Universe. The DM of the Universe consists of an axionlike particle with density ρ_a. This density is constant before the ALP field commences oscillation at $a_{\rm osc} \approx 10^2$. We study the growth of ALP perturbations in this background cosmology.

in the matterlike regime, however, the ALP is not CDM-like at all scales. At small physical scales (large comoving wavevector k), the wavelike nature of the ALP introduces an effective pressure (and a corresponding sound speed) for the ALP fluid, preventing clustering below a characteristic scale, the Jeans wavenumber k_J, which is estimated below following Refs. [6,77,78]. We focus on times after oscillations have begun and take into account these effects by integrating out the fast oscillations of the ALP, giving rise to effective fluid equations. The derivation of these equations, including the ALP sound speed, is discussed in greater detail in the Appendixes A and B. Here we collect the results.

We work in the conformal Newtonian gauge and take the metric to be

$$ds^2 = a^2[(1+2\Psi)d\tau^2 - (1+2\Phi)dx^2], \tag{15}$$

where τ is the conformal time, x is the comoving coordinate, a is the scale factor, and Ψ and Φ are the metric perturbations. We neglect neutrino anisotropic stress and set $\Phi = -\Psi$; this is a good approximation for modes that entered the horizon before $T \gtrsim$ MeV, since Weak interactions are still in equilibrium, so neutrinos behave as a perfect fluid. The equation of motion for the linear ALP field perturbation a_1 is [79]

$$\ddot{a}_1 + 2\mathcal{H}\dot{a}_1 + (k^2 + m_a^2 a^2)a_1 - (\dot{\Psi} - 3\dot{\Phi})\dot{a}_0 + 2a^2 m_a^2 \Psi a_0 = 0, \tag{16}$$

where the dots denote derivatives with respect to conformal time, $\mathcal{H} = aH$ is the comoving Hubble parameter and a_0 is the background solution of Eq. (3). The oscillation-averaged equations for the energy density and velocity perturbations are obtained by constructing approximate solutions in the $\mathcal{H}/(m_a a)$ expansion as in Ref. [78].

We find the Fourier-space density δ_i and velocity divergence $\theta_i = i\vec{k}\cdot\vec{v}_i$ perturbation equations for $i = \phi$ (the EMD field), a (the ALP), and r (SM radiation) to be

$$\dot{\delta}_\phi + \theta_\phi + 3\dot{\Phi} = -a\Gamma_\phi \Psi, \tag{17a}$$

$$\dot{\theta}_\phi + \mathcal{H}\theta_\phi - k^2\Psi = 0, \tag{17b}$$

$$\dot{\delta}_r + \frac{4}{3}\theta_r + 4\dot{\Phi} = a\Gamma_\phi \frac{\rho_\phi}{\rho_r}[\delta_\phi - \delta_r + \Psi], \tag{17c}$$

$$\dot{\theta}_r - \frac{1}{4}k^2\delta_r - k^2\Psi = a\Gamma_\phi \frac{\rho_\phi}{\rho_r}\left[\frac{3}{4}\theta_\phi - \theta_r\right], \tag{17d}$$

$$\dot{\delta}_a + \theta_a + 3\dot{\Phi} = -3c_{\rm nad}^2\mathcal{H}\delta_a - 9c_{\rm nad}^2\mathcal{H}^2\theta_a/k^2 \tag{17e}$$

$$\dot{\theta}_a + \mathcal{H}\theta_a - k^2\Psi = +3c_{\rm nad}^2\mathcal{H}\theta_a + k^2 c_{\rm nad}^2\delta_a \tag{17f}$$

$$k^2\Phi - 3\mathcal{H}[\mathcal{H}\Psi - \dot{\Phi}] = 4\pi G a^2[\rho_\phi\delta_\phi + \rho_r\delta_r + \rho_a\delta_a], \tag{17g}$$

where the nonadiabatic sound speed $c_{\rm nad}$ is

$$c_{\rm nad}^2 = \frac{k^2}{k^2 + 4m_a^2 a^2}. \tag{18}$$

We take standard adiabatic initial conditions for all fluids. These are discussed in detail in Appendix A. Note that this assumption for the ALP is not trivial. Under standard assumptions and high-scale inflation, quantum fluctuations of the ALP will generate isocurvature initial conditions at large scales. Consistency with the CMB then implies either a low scale of inflation or nontrivial inflationary dynamics, as discussed in Sec. II. For example, if the scale of inflation is low, the isocurvature component in the ALP is tiny. The adiabatic component vanishes until the ALP begins to oscillate. After oscillations begin, superhorizon modes of the ALP develop adiabatic perturbations, as we show in Appendix B.

The system of Eqs. (17) is nearly identical to those derived in Ref. [43] for the WIMP, except for the appearance of the sound speed terms, the initial conditions, and the generation of the relic abundance through misalignment. As a result, modes that enter the horizon after oscillations and for which the sound speed is not important evolve as described in Ref. [43]. Likewise, the other fluids evolve as in Ref. [43].

Before numerically analyzing Eq. (17), it is useful to study the behavior of the ALP density perturbations analytically. First, let us focus on longer wavelength modes, which enter the horizon after the ALP begins to oscillate, and for which sound speed effects can be neglected. The growing mode of perturbations that enter the horizon before RH is

$$\delta_a(a,k) = 2\Phi_0 + \frac{2}{3}\left(\frac{k}{H_1}\right)^2 a\Phi_0 (a < a_{\rm RH}, k > k_{\rm RH}), \quad (19)$$

where Φ_0 is the initial value of the gravitational potential (prior to reheating), and $k_{\rm RH}$ is the comoving scale corresponding to the comoving horizon at reheating:

$$k_{\rm RH} = \mathcal{H}(a_{\rm RH}) \sim a_{\rm RH}^{-1/2}. \quad (20)$$

This wavenumber can be expressed in physical variables [43]:

$$\frac{k_{\rm RH}}{k_{\rm eq}} = \frac{k_{\rm RH}}{a_{\rm eq}H(T_{\rm eq})} \approx 5.9\times 10^6 \left(\frac{T_{\rm RH}}{5\ {\rm MeV}}\right)\left(\frac{g_*(T_{\rm RH})}{10.75}\right)^{1/6}, \quad (21)$$

where we used $a_{\rm eq} = 1/(1+z_{\rm eq})$, $z_{\rm eq} = 3365$ and $T_{\rm eq} = 0.79$ eV [80].

After RH, these modes continue to grow logarithmically; assuming horizon entry occurs well before instantaneous reheating at $a = a_{\rm RH}$ and matching solutions in the two regimes, one finds that for $a > a_{\rm RH}$

$$\delta_a(a,k) = \frac{2}{3}\left(\frac{k}{k_{\rm RH}}\right)^2 \Phi_0 \ln\frac{ea}{a_{\rm RH}} (a > a_{\rm RH}, k > k_{\rm RH}). \quad (22)$$

Here the prefactor $\propto (k/k_{\rm RH})^2$ encodes the period of enhanced growth between horizon entry and reheating; since $\mathcal{H} \sim a^{-1/2}$ during EMD, linear growth corresponds to

$$a_{\rm RH}/a_{\rm hor} = (k/k_{\rm RH})^2. \quad (23)$$

Modes that enter the horizon after RH evolve as in ΛCDM; the solution is well approximated by the fitting formula [81]

$$\delta(a,k) = \frac{10}{9} A\Phi_0 \ln\left(\frac{Ba}{a_{\rm hor}}\right) \quad (a > a_{\rm RH}, k < k_{\rm RH}) \quad (24)$$

after horizon entry ($a \gg a_{\rm hor}$), with $A \approx 9.11$ and $B \approx 0.594$; the prefactor $10/9$ accounts for the transition from EMD to RD.

How much growth can an axion perturbation undergo during EMD? Linear growth in the scale factor corresponds to $\delta_a(T_{\rm RH})/\delta_a(T_{\rm osc}) \sim (k_{\rm osc}/k_{\rm RH})^2$ [see Eq. (22)], where $k_{\rm osc}$ is the comoving horizon size when oscillations begin:

$$k_{\rm osc} = m_a a_{\rm osc}. \quad (25)$$

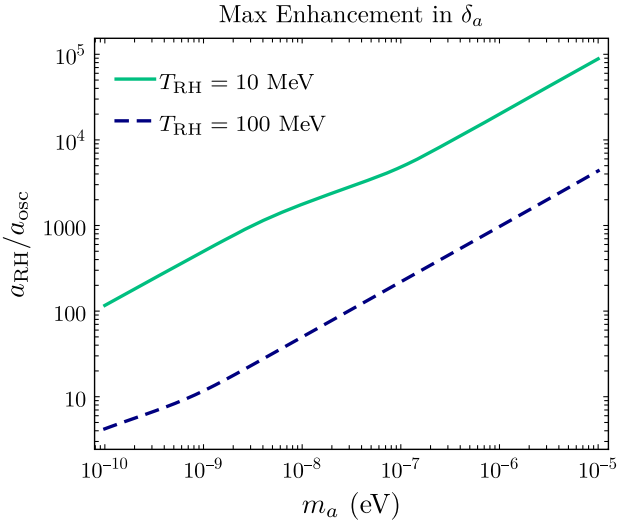

FIG. 3. The ratio of scale factors at the onset of ALP oscillation and at reheating $a_{\rm osc}/a_{\rm RH}$, corresponding to the maximum possible enhancement in the growth of the dark matter perturbation δ_a, for reheating temperatures $T_{\rm RH} = 10$ and 100 MeV.

Therefore

$$\left(\frac{k_{\rm osc}}{k_{\rm RH}}\right)^2 = \frac{a_{\rm RH}}{a_{\rm osc}} \approx \left(\frac{m_a}{H(T_{\rm RH})}\right)^{2/3} \sim 10^6 \left(\frac{m_a}{10^{-5}\ {\rm eV}}\right)^{2/3}\left(\frac{5\ {\rm MeV}}{T_{\rm RH}}\right)^{4/3}\left(\frac{10.75}{g_*(T_{\rm RH})}\right)^{1/3}. \quad (26)$$

Note also that during RD the temperature falls as $T \propto a^{-1}$, but during EMD, entropy release from the ϕ decay causes the temperature to fall as a smaller power, $T \propto a^{-3/8}$ [82]. The fact that axion miniclusters are produced in an EMD cosmology relies both on this fact and on the linear growth of perturbations during EMD.

The only modes that could possibly grow by the amount (26) must already be inside the horizon at $T_{\rm osc}$, as well as being nonrelativistic and below the Jeans scale (discussed below). This is an estimate that represents the maximum theoretically possible growth of a mode during EMD. We show this result in Fig. 3 for $T_{\rm RH} = 10$ and 100 MeV. We see that we can indeed achieve a duration of matter domination with $a_{\rm RH} \sim \mathcal{O}(10^4)$ after ALP oscillation (taking $a_{\rm osc} = 1$), but that this requires $m_a \gtrsim 10^{-6}$ eV and a low reheating temperature. For low masses, the enhancement is quite small, since the ALP starts to oscillate later and so does not benefit as much from EMD.

The previous discussion neglected the effective ALP pressure. This approximation turns out to be excellent for modes that enter the horizon well after oscillation and well before reheating, which will be the most important for the formation of miniclusters. To see this, and to understand the effect of the sound speed on other modes, we combine

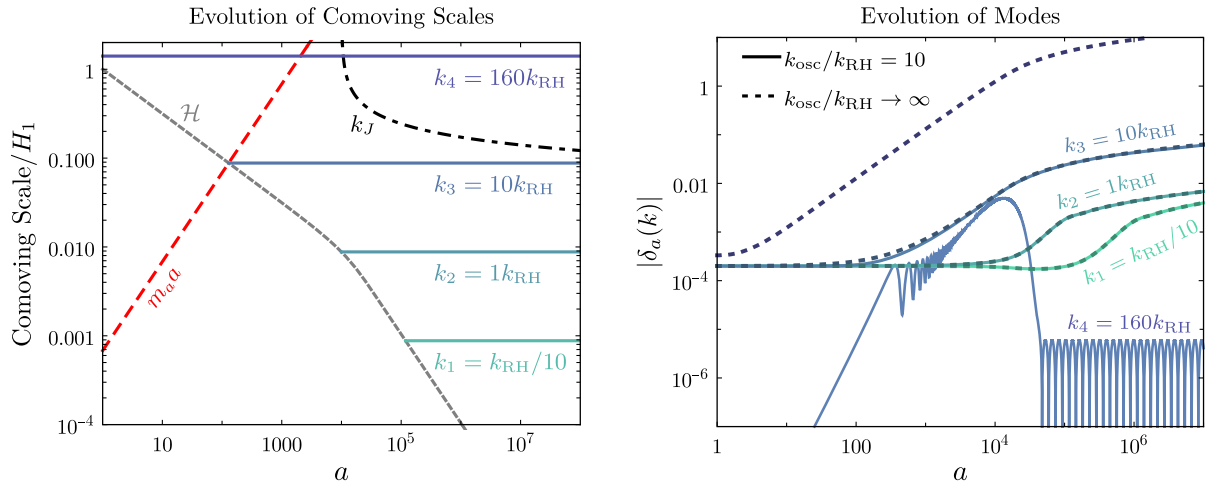

FIG. 4. Left: comoving wavevectors k_i of representative ALP perturbations (solid lines) compared to the evolution of comoving Hubble distance $\mathcal{H}$ (short dashed gray line), ALP mass m_a (dashed red line), and the Jeans scale k_J (dot-dashed black line). Each mode begins evolving when $k \approx \mathcal{H}$ and its subsequent behavior depends on its size relative to the other scales at a given time. The ϕ decay rate is such that reheating happens at $a_{\rm RH} \approx 10^4$ and the ALP mass is chosen such that $k_{\rm osc}/k_{\rm RH} = 10$, i.e., ALP oscillations begin before reheating. k_1 enters the horizon after reheating and evolves as in the standard cosmology; k_2 enters at reheating and corresponds to physical scales much larger than the ALP Jeans scale; k_3 and k_4 enter deep during the EMD; k_3 (k_4) is physically larger (smaller) than the Jeans scale. Right: evolution of modes with different wavevectors k (same as in the left panel) as a function of the scale factor (solid lines) compared to the CDM case (dotted lines), corresponding to $k_{\rm osc}/k_{\rm RH} \to \infty$ or, equivalently, large m_a. The growth of modes with $k > k_J \sim k_{\rm osc}$ is suppressed by the effective ALP pressure, while those with $k < k_J$ are identical to the CDM case. Modes that enter the horizon before RH experience enhanced growth due to EMD.

Eqs. (17e) and (17f) into a single second-order equation. The exact result is complicated, but if we take $\mathcal{H} \sim k \ll ma$ (i.e., $c_{\rm nad} \ll 1$) then it can be expanded in these small quantities, with the result

$$\ddot{\delta}_a + \mathcal{H}\dot{\delta}_a + c_{\rm nad}^2 k^2 \delta_a = -k^2\Psi - 3\mathcal{H}\dot{\Phi} - 3\ddot{\Phi}. \quad (27)$$

This equation was also obtained in Ref. [6]. It is clear that the sound speed term competes with the gravitational driving term, and it prevents growth of the perturbations with k greater than some Jeans scale k_J, which we estimate as follows. First we analyze the EMD era, where the gravitational potential is dominated by ϕ and is constant in time. When the ALP oscillations begin, the ALP density contrast is $\delta_a \approx 2\Phi_0$ (see Sec. III B and Appendix A). The sound speed term prevents growth until it becomes comparable to the gravitational source term $k^2\Psi = -k^2\Phi$, which occurs when $k < m_a a$. Thus during EMD, $k_J \sim m_a a \equiv k_{\rm Compt}$.[2] Moreover, modes that enter the horizon during EMD but well after oscillations satisfy $k \ll a_{\rm osc} m_a \ll a m_a \sim k_J$. On the other hand, since $k_J \sim k_{\rm Compt}$, modes with $k > k_J$ during EMD are not well described using the oscillation-averaged equations used here. It is still true, however, that during this period the density perturbation does not grow beyond the CDM adiabatic initial condition $2\Phi_0$, as illustrated in the right panel of Fig. 4.

After reheating, the gravitational potential rapidly decays and oscillates, averaging to zero over cosmological time scales for $k \gg \mathcal{H}$. In this limit, the source terms on the right-hand side of Eq. (27) can be dropped, leading to an approximate solution

$$\delta_a \sim \sin\left(\frac{k^2}{2m_a\sqrt{a_{\rm RH}}} \ln\frac{a}{a_{\rm RH}}\right). \quad (28)$$

The perturbation grows logarithmically as for a CDM fluid when the argument of the sine is small, and begins to oscillate (with a constant frequency in $\ln a$) when $k^2 \sim 2m_a\sqrt{a_{\rm RH}}/\ln(a/a_{\rm RH})$. Combining the EMD and RD regimes we find that the Jeans scale (expressed in physical units) is

$$k_J = \begin{cases} m_a a & (a \le a_{\rm RH}) \\ \dfrac{a_{\rm RH}\sqrt{2m_a H(a_{\rm RH})}}{\sqrt{\ln(a/a_{\rm RH})}} & (a > a_{\rm RH}). \end{cases} \quad (29)$$

C. Numerical solutions

In this section we solve the full linearized system of equations described above numerically and study the growth of ALP perturbations. We interpret the evolution

[2]Note that this is a different scaling with scale factor and ALP mass than found in, e.g., Ref. [6,52,77,83], due to the fact that the gravitational potential is dominated by the EMD field ϕ, rather than the ALP itself.

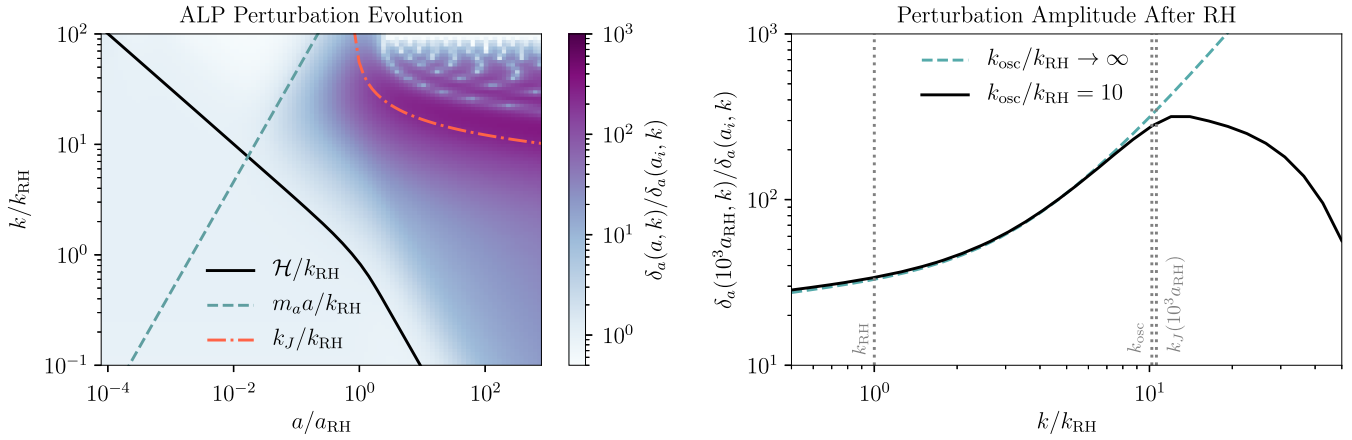

FIG. 5. Growth of ALP perturbations δ_a with different comoving wavenumbers. The left panel shows the evolution of δ_a (normalized to its initial value) in the plane of $a/a_{\rm RH}$ and $k/k_{\rm RH}$. Important comoving scales are indicated by lines (in analogy to Fig. 4): the comoving Hubble horizon $\mathcal{H}$ (solid black line), the comoving Compton scale $m_a a$ (teal dashed line) and the Jeans scale (red dot-dashed line). The right panel shows the amplitude of ALP perturbations well after reheating (at $a \sim 10^3 a_{\rm RH}$) with and without the ALP sound-speed effects, solid and dashed lines respectively. Key scales are shown by vertical dashed gray lines. In both panels, the ALP mass is chosen such that $k_{\rm osc}/k_{\rm RH} = 10$. The growth of modes with $k \gtrsim k_J \sim k_{\rm osc}$ is suppressed because of the effective ALP pressure. These modes also enter the horizon prior to ALP oscillations, when the ALP field still behaves like a cosmological constant. Modes with $k/k_{\rm RH} \lesssim 1$ enter after reheating and therefore evolve as in standard cosmology. Note that the $k > k_J$ part of the right panel is the envelope of the oscillating function shown in the left panel.

of perturbations with different k in terms of their relation to the key physical scales:

(i) the horizon size $\mathcal{H}$,
(ii) the Compton scale $k_{\rm Compt} = m_a a$,
(iii) the horizon size at RH, $k_{\rm RH}$, defined in Eq. (20),
(iv) the horizon size when ALP oscillations begin, $k_{\rm osc}$, defined in Eq. (25),
(v) the Jeans scale, k_J, defined in Eq. (29), which encodes the effects of the ALP effective pressure.

In Fig. 4 we illustrate the evolution for several values of k; the left panel shows the relation of these wavenumbers to the key scales above. The right panel shows the time evolution of the corresponding perturbations. The mode $k_1 < k_{\rm RH}$ enters during RD, and is always smaller than the Jeans and Compton scales, so its evolution is CDM-like with no enhancement from EMD or suppression from the ALP effective pressure.[3]

The mode $k_2 = k_{\rm RH}$ entered at reheating, while the Jeans scale is unimportant, so it evolves as a CDM perturbation. The mode $k_3 > k_{\rm RH}$ enters during EMD, while quasirelativistic, and it is slightly sensitive to the Jeans scale at the start of its evolution. This can be seen in the right-hand plot in the small suppression in the growth of the k_3 mode relative to the same mode in the CDM case as they both enter the horizon. Finally, the mode k_4 enters the horizon while relativistic. After it becomes nonrelativistic however ($a = 5000$), we see that its amplitude is suppressed and it undergoes rapid oscillations driven by gravity and ALP effective pressure. We note that ALP oscillation-averaged equations [Eqs. (17)] are not adequate in this case before the onset of oscillations; we therefore solve for its early evolution using the field equation (16). Its late-time evolution matches onto the solutions of the oscillation-averaged equations with adiabatic initial conditions discussed above. We therefore only use these equations in the remainder of this work.

The previous discussion can be used to understand Fig. 5, which shows the evolution of a wide range of k modes in the left panel; the right panel illustrates the suppression of growth due to the effective pressure well after reheating at $a = 10^3 a_{\rm RH}$. The most important features are evident in the left panel: modes that enter the horizon benefit from EMD, while those with $k \gg k_J \sim k_{\rm osc}$ are suppressed. The magnitude of the suppression is scale and somewhat time dependent [due to the scale factor dependence of k_J; see Eq. (29)]. In the following subsection we combine these numerical solutions with late-time growth of perturbations during standard radiation and matter domination. The goal is to evaluate the ALP perturbation power spectrum and smoothed density variance at various redshifts. In order to simplify these considerations we will model the small-scale suppression of power as a sharp cutoff at $k = k_{\rm osc}$. This is an approximation as the actual fall-off is much smoother; see the right panel of Fig. 5. However, this simplification will enable a fast exploration of the minihalo parameter space without having to solve the full Boltzmann system for each $(m_a, T_{\rm RH})$. A more sophisticated treatment of these small-scale effects would

[3]There is a brief kick of nonlogarithmic growth as modes cross the horizon due to gravitational driving [81]. This is particularly clear for the k_1 mode and it is unrelated to the period of EMD.

involve using the analytic solution during RD [Eq. (28)] to evaluate the density contrast and the corresponding transfer function at late times. This approximation is appropriate for the gravitational signals we will study in Sec. V, since they are sensitive to the largest surviving minihalos which originate as $k < k_{\rm osc}$ perturbations in models with $k_{\rm RH} \ll k_{\rm osc}$ (i.e., when there is an extended period between the start of ALP oscillations and end of reheating). In scenarios where there is only a mild hierarchy $k_{\rm RH} \lesssim k_{\rm osc}$, the sharp cutoff underestimates the abundance of minihalos and their survival probability, and therefore leads to conservative estimates of signal rates.

D. Growth and transfer function

The minicluster abundance depends on the evolution of ALP perturbations towards collapse. For our choice of adiabatic initial conditions this happens at or after the standard matter-radiation equality. In order to match onto this "standard" growth period we modify the semianalytic prescription of Ref. [43] to account for ALP mass effects. The quantity of interest is the time-dependent fluctuation variance smoothed over comoving length scales R

$$\sigma^2(a,R) = \int \frac{d^3k}{(2\pi)^3} \langle \delta_a(a,k)^2 \rangle W(kR)^2 = \frac{1}{2\pi^2} \int dk k^2 W(kR)^2 |TD(a,k)|^2 P(k). \quad (30)$$

The averaging is achieved through the use of a window function $W(kR)$, which we take to be a spherical top hat in real space. In the second line we wrote the time-dependent ALP perturbation as

$$\delta_a(a,k) \approx \frac{2k^2}{5\Omega_m H_0^2} TD(a,k)\mathcal{R}(k) \quad (a > a_{\rm eq}), \quad (31)$$

where $\mathcal{R}$ is the primordial scalar curvature fluctuation amplitude that determines the initial conditions for the evolution of δ_a; this equation defines $TD(a,k)$, the scale-dependent growth function which we discuss in more detail below. The primordial matter power spectrum $P(k)$ is related to the power spectrum of $\mathcal{R}$ by

$$P(k) = \left(\frac{2k^2}{5\Omega_m H_0^2}\right)^2 P_{\mathcal{R}}(k), \quad (32)$$

where

$$P_{\mathcal{R}}(k) = \frac{2\pi^2}{k^3} A_s \left(\frac{k}{k_0}\right)^{n_s - 1} \quad (33)$$

is set by inflation, with $k_0 = 0.05\ \mathrm{Mpc}^{-1}$ and Planck best-fit values of $\ln(10^{10} A_s) = 3.044 \pm 0.014$ and $n_s = 0.965 \pm 0.004$ [70]. We note that the scales we are interested in are far smaller than those probed by the CMB; thus the assumption of a power-law spectrum with a constant n_s amounts to a significant extrapolation. While the uncertainties in the n_s measurement do not qualitatively affect the results below, we emphasize that the $k \gg k_0$ part of the matter power spectrum we are studying has not been measured directly.

The scale-dependent growth function $TD(a,k)$ contains both the dynamics of the Boltzmann equations solved in the previous section and the post-reheating growth of ALP overdensities. For cosmologies without baryons or the cosmological constant, TD factorizes as $TD(a,k) = T(k)D(a)$, where T and D are the standard transfer and growth functions, respectively. The normalization factors in Eq. (31) are chosen such that $T(k) \approx 1$ for modes that entered the horizon after matter-radiation equality and $D(a) \approx a$ deep in the matter domination era. The factor of k^2 in Eq. (31) combines with $D(a) = a$ to become $k^2 a = a/a_{\rm hor}$, representing linear growth during the most recent era of matter domination. The transitions from EMD to RD at $T_{\rm RH}$, and from RD to MD at $T_{\rm eq}$ introduce two characteristic scales $k_{\rm eq}$ and $k_{\rm RH}$ into the transfer function. Schematically, T behaves as

$$T(k) \sim \begin{cases} 1 & k < k_{\rm eq} \\ (k/k_{\rm eq})^{-2} \ln(k/k_{\rm eq}) & k_{\rm RH} \geq k \geq k_{\rm eq} \\ (k_{\rm RH}/k_{\rm eq})^{-2} & k_{\rm osc} > k > k_{\rm RH} \\ 0 & k > k_{\rm osc} \end{cases}, \quad (34)$$

The first line reflects the fact that modes with $k < k_{\rm eq}$ enter the horizon after matter-radiation equality and their growth is entirely captured by $D(a)$. The second line applies to modes that enter the horizon during radiation domination (after reheating and before MRE); for these modes, the transfer function measures the deviation from linear growth. The factor of $(k/k_{\rm eq})^{-2}$ removes this linear growth from $k^2 D$ and the factor of $\ln(k/k_{\rm eq})$ restores the standard logarithmic growth during RD. The transfer function plateaus for modes that enter the horizon before reheating, simply removing linear growth between reheating and equality, but preserving it at higher scales. Finally, modes that enter before oscillations begin are suppressed due to the ALP effective pressure (note that this sharp cutoff is only a rough approximation; the actual falloff is k and a dependent as discussed in the previous section).

In a universe with a non-negligible baryon abundance and a late-time period of dark energy domination, TD does not factorize. The former ensures that the growth of modes below the baryonic Jeans scale is slower (since baryons are pressure supported on these scales), while the latter affects the late-time evolution of structure. In both cases the growth rate becomes scale dependent as can be seen by solving the Meszaros equation [43]. However, the

approximate scaling relations in Eq. (34) can still be used to qualitatively understand our results. The transfer function can also be altered through the presence of a nonzero velocity dispersion for dark matter, which washes out structure at low masses. ALPs produced through the misalignment mechanism have a momentum dispersion which is of order the Hubble scale at the time of oscillation, which is negligible [38].

In order to facilitate the exploration of parameter space, we follow Ref. [43] in defining a semianalytic approximation to $\delta_a(a,k)$ that allows us to propagate our results into standard matter domination. The linearity of the Boltzmann equations and decoupling of different k-modes allows us to approximate the full scale-dependent growth function as

$$TD(a,k) = TD_{\Lambda\mathrm{CDM}}(a,k)R(k), \tag{35}$$

where

$$R(k) = \frac{\delta_a(a,k)}{\delta_a^{(\Lambda\mathrm{CDM})}(a,k)} \sim \begin{cases} 1 & k < k_{\mathrm{RH}} \\ \frac{(k/k_{\mathrm{RH}})^2}{\ln(k^2/k_{\mathrm{RH}}^2)} & k_{\mathrm{osc}} \geq k \geq k_{\mathrm{RH}} \\ 0 & k > k_{\mathrm{osc}} \end{cases}. \tag{36}$$

The growth of perturbations after equality cancels in the ratio such that R does not depend on a. This parametrization is useful because $R(k)$ encodes the effects of EMD and the ALP mass relative to ΛCDM, while $TD_{\Lambda\mathrm{CDM}}(a,k)$ can be computed using a standard Boltzmann code [84,85] or using well-known fitting formulas [81]. The precise form of $R(k)$ and the computation of $TD_{\Lambda\mathrm{CDM}}$ are described in Appendix C. The scalings in Eq. (36) follow from the approximate solutions in Eqs. (22). Since $TD_{\Lambda\mathrm{CDM}} \sim 1/k^2$ at $k \gg k_{\mathrm{eq}}$ and $R \sim k^2$ for modes that entered during EMD, we see that the rescaling by R results in the "flattening" of the transfer function as indicated in the third line of Eq. (34).[4]

We can capture the EMD-enhanced growth in terms of an effective primordial power spectrum $P_{\mathcal{R}}^{\mathrm{eff}}(k) = P_{\mathcal{R}}(k) \times R(k)^2$ that can be used as input to N-body simulations. This effective power spectrum is shown in Fig. 6 as a function of k/k_{eq}. The gray band is for the standard cosmology as in Eq. (32). The blue lines correspond to our scenario for two different choices of $T_{\mathrm{RH}} = 5$ and 100 MeV. Modes with $k < k_{\mathrm{RH}}$ enter the horizon after reheating and evolve as in the standard cosmology. The wavenumbers of those modes can be estimated using Eq. (21). For modes $k_{\mathrm{osc}} \geq k \geq k_{\mathrm{RH}}$ the power spectrum grows as k^{3+n_s} as in Eq. (36). At large k power is cut off by the Jeans scale. The dashed blues lines correspond to the heavy limit $m_a \to \infty$, for which there is no Jeans scale cutoff, and the solid blue lines correspond to $m_a = 10^{-9}$ eV and 10^{-8} eV.

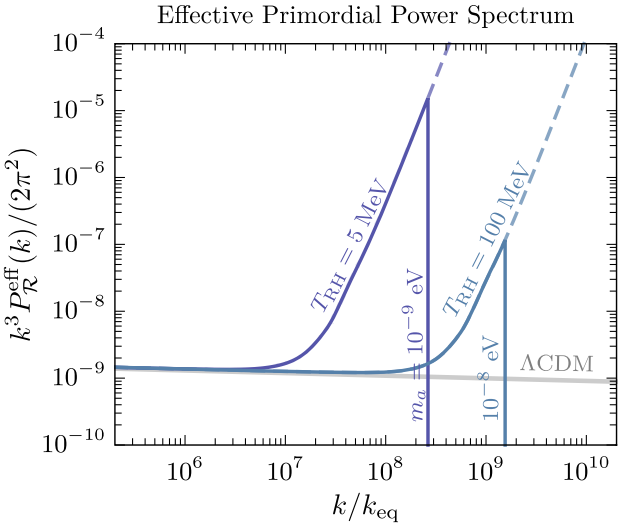

FIG. 6. The dimensionless effective primordial curvature power-spectrum $k^3 P_{\mathcal{R}}^{\mathrm{eff}}(k)/(2\pi^2)$ that encodes EMD structure growth at small scales. The gray band corresponds to the power spectrum in the standard cosmology as in Eq. (33). The solid blue lines correspond to $T_{\mathrm{RH}} = 5$ MeV, $m_a = 10^{-9}$ eV (upper) and $T_{\mathrm{RH}} = 100$ MeV, $m_a = 10^{-8}$ eV (lower). The sharp cutoffs at small scales in the small m_a cases roughly approximate the suppression of power for modes with $k > k_J \sim k_{\mathrm{osc}}$ due to the effective ALP sound speed. This suppression is absent in the CDM case (dashed lines), corresponding to $m_a \to \infty$.

In Fig. 7 we show the density standard deviation σ for $T_{\mathrm{RH}} = 5$ MeV, different choices of m_a and two different times, $z = 1000$ and $z = 100$, illustrating the linear growth of density perturbations during standard matter domination. Its value in the standard cosmology is indicated by the dotted gray lines. In this case there is no enhancement at large k, corresponding to small M. The dashed lines show the small-scale enhancement in models with a large ALP mass where the effective pressure is not important. In this regime,

$$\sigma \propto \frac{1}{1+z} M^{-2/3} k_{\mathrm{RH}}^{-2} \propto \frac{1}{1+z} M^{-2/3} T_{\mathrm{RH}}^{-2}. \tag{37}$$

In contrast, if m_a is small, the effective pressure introduces a small scale cutoff in the EMD mode "amplification," leading to a flattening of σ at small M. Larger T_{RH} reduce the largest mass scales that benefit from increased growth

[4]To qualitatively summarize the slightly ridiculous parametrization $\delta \sim k^2 TD_{\Lambda\mathrm{CDM}} R$: $k^2 D$ puts in linear growth for all modes for all times, $TD_{\Lambda\mathrm{CDM}}$ removes all linear growth prior to matter-radiation equality for modes entering during this time, and R restores it again between ALP oscillations and reheating for modes entering during this time, while setting to zero modes that enter the horizon before oscillations begin.

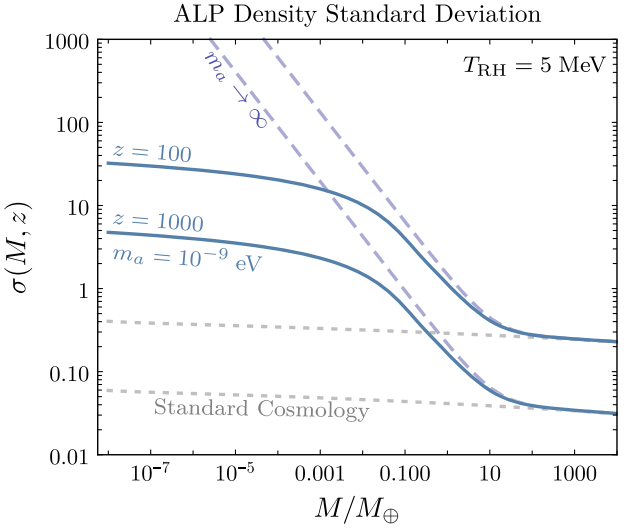

FIG. 7. The density fluctuation standard deviation for a standard cosmology (dotted lines) and EMD cosmology with reheating at $T_{\rm RH} = 5$ MeV with $m_a = 10^{-9}$ eV and $m_a \to \infty$ (solid and dashed lines, respectively). For each model, the lower and upper lines correspond to $z = 1000$ and $z = 100$.

during EMD. In the following section, we use these results to estimate minihalo size, density and distribution.

IV. MINICLUSTER ABUNDANCE

Early matter domination leads to enhanced structure growth for scales that entered the horizon before reheating. Given the bound on $T_{\rm RH} \gtrsim 5$ MeV, the *maximum* mass of fluctuations that benefit from this enhancement simply corresponds to the matter mass enclosed by the horizon when $T = T_{\rm RH}$:

$$M_{\rm RH} = 250 M_\oplus \left(\frac{5\ {\rm MeV}}{T_{\rm RH}}\right)^3 \left(\frac{10.75}{g_*(T_{\rm RH})}\right)^{1/2}. \quad (38)$$

Perturbations with $M > M_{\rm RH}$ evolve as in standard cosmology, while those with $M < M_{\rm RH}$ experience enhanced growth. However, perturbations in the ALP fluid only begin to grow like matter after the oscillations have begun, so there is also a *minimum* mass of fluctuations, corresponding to the mass inside the horizon at the start of oscillations, even in standard cosmology without EMD:

$$M_{\rm osc} \approx 2.7 \times 10^{-11}\ M_\odot \Omega_a \left(\frac{10^{-5}\ {\rm eV}}{m_a}\right)\left(\frac{5\ {\rm MeV}}{T_{\rm RH}}\right), \quad (39)$$

where we used Eqs. (26) and (21) and Ω_a is given in Eq. (6). It is these objects that are usually called miniclusters, especially in models with postinflationary PQ breaking where they collapse very early. The purpose of this section is to show that these smallest miniclusters are assembled into larger clumps, since EMD enhances the growth of density fluctuations over a range of scales.[5]

We estimate the statistics of collapsed ALP DM objects as a function of size and redshift using the Press-Schechter (PS) formalism [88,89] with the results from linear theory discussed in the previous sections. We wish to estimate the mass spectrum of miniclusters, their sizes and densities, and their assembly history.

A. Halo function

The central assumption of PS is that the fraction of mass in structures of size $\sim R$ is equal to the probability that the smoothed density contrast δ_R exceeds a threshold δ_c. The critical density contrast δ_c can be estimated from spherical collapse, with the result that overdensities with $\delta_R = \delta_c = 1.686$ (as derived in linear perturbation theory) should have collapsed; this number is insensitive to the precise cosmological model, i.e., variations in Ω_Λ and Ω_m, as long as the collapse occurs during matter domination [90].[6] Using this prescription, the fraction f of matter in objects of mass in the range $[M, M + dM]$ at redshift z is

$$df(M, z) = \sqrt{\frac{2}{\pi}} \frac{\delta_c}{M\sigma} \left|\frac{d\ln\sigma}{d\ln M}\right| \exp\left(-\frac{\delta_c^2}{2\sigma^2}\right) dM. \quad (40)$$

The halo function is the number density of collapsed objects in this mass range:

$$n(M, z) = \frac{\rho}{M}\frac{df}{dM}. \quad (41)$$

Equation (40) can be integrated to find the fraction of matter contained in collapsed objects with a mass in the range $[M_1, M_2]$

$$F(M_1, M_2) = {\rm erf}\left(\frac{\delta_c}{\sqrt{2}\sigma(M_2, z)}\right) - {\rm erf}\left(\frac{\delta_c}{\sqrt{2}\sigma(M_1, z)}\right). \quad (42)$$

It is useful to define a characteristic mass $M_*(z)$,

$$\sigma(M_*, z) = \delta_c. \quad (43)$$

[5]Coincidentally, in the standard ΛCDM scenario where the dark matter is a thermal relic whose density is set through freeze-out, the smallest gravitationally bound structures are also approximately Earth-mass microhalos [86]. The formation of these structures is determined by the time of kinetic decoupling from the Standard Model thermal bath. This occurs when the temperature has dropped a further factor of 10–1000 after freeze-out [87]. For a 100 GeV WIMP this corresponds to a temperature of order 10–100 MeV, similar to $T_{\rm RH}$ in our scenario.

[6]Collapse can also occur during RD following the end of EMD. In this regime the collapse criterion is different; see, e.g., Ref. [91]. Objects forming during this era would be very light and unimportant for the gravitational probes we consider in Sec. V. We therefore focus on objects that reach nonlinearity at or after standard matter-radiation equality.

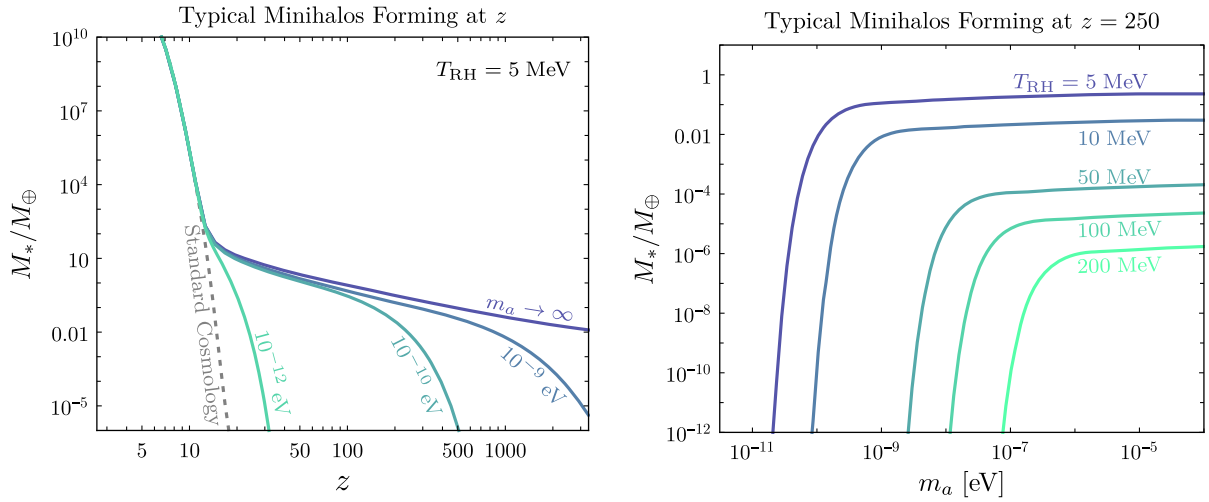

FIG. 8. The typical mass M_* (in units of the Earth mass) of minihalos forming at z, defined by $\sigma(M_*, z) = \delta_c$. Results are shown as a function of redshift at fixed reheat temperature $T_{\rm RH} = 5$ MeV for different ALP masses (left panel) and for fixed redshift $z = 250$ as a function of the ALP mass for different $T_{\rm RH}$ (right panel). We expect that minihalos forming before this redshift will survive tidal disruption through encounters with stars in the galaxy; see Sec. V A. In the right panel the curves remain flat for higher values of m_a.

At a given redshift z, M_* gives the mass of the typical structures that collapse at this time. The fact that M_* is a monotonically increasing function of redshift suggests that structures are formed hierarchically in this model. Using Eq. (30) we see that $\sigma \sim M^{-(n_s+3)/6}$ for $M < M_{\rm RH}$. Therefore we can approximately solve for $M_*(z)$:

$$M_*(z) \approx M_{\rm RH}\left(\frac{c}{1+z}\right)^{6/(n_s+3)} \tag{44}$$

where $c \sim 2.4$–2.8 for $T_{\rm RH} \sim [5, 100]$ MeV is a slowly varying function of $T_{\rm RH}$. This scaling is valid for $z \gtrsim 20$ and $m_a \to \infty$. In Fig. 8 we show the numerical solution for M_* as a function of the collapse redshift in the left panel and the ALP mass in the right panel for several choices of the reheating temperature. In the limit of a "heavy" ALP we find a result similar to that of Ref. [43] for CDM. For small ALP masses, however, there is a noticeable suppression of minicluster formation. As explained in Sec. III C, this is due to the delayed onset of oscillations of the ALP, which reduces the amount of EMD linear growth experienced by the perturbations. Since this linear growth factor is $\sim T_{\rm osc}/T_{\rm RH}$, the "cutoff" occurs at larger m_a for larger reheat temperatures, as is evident in the right panel of Fig. 8. The minimum m_a needed to form minihalos at z for a given $T_{\rm RH}$ may be estimated by requiring the total amount of growth between oscillations and z to be of order 10^5,

$$10^5 \approx \left(\frac{k_{\rm osc}}{k_{\rm RH}}\right)^2 \log\left(\frac{T_{\rm RH}}{T_{\rm eq}}\right)\left(\frac{1+z_{\rm eq}}{1+z}\right), \tag{45}$$

where the first factor is given by Eq. (26). For thermal DM, an analogous cutoff in the power on small scales occurs in the presence of nonzero DM velocities and the resulting free-streaming [43]. The choice of $z = 250$ in the right panel is motivated in Sec. V A: minihalos forming later than this are expected to undergo significant tidal disruption in stellar encounters.

The Press-Schechter formalism is based on the assumption of spherical collapse, and when compared with the results of numerical simulations overpredicts the amount of structure at the smallest scales and underpredicts the amount at larger scales [92]. Better agreement with simulations in ΛCDM can be achieved using formalisms such as Sheth-Tormen [93,94] which allow for ellipsoidal collapse. It would be interesting to study the implications of ellipsoidal collapse for ALP miniclusters in future work.

In Fig. 9 we show the fractional minicluster abundance at a redshift of $z = 250$ as function of the minicluster mass M and $T_{\rm RH}$. For finite ALP masses and at larger $T_{\rm RH}$, there is less time during EMD for perturbations to grow, and the abundance peaks at smaller M. If the ALP is too light and $T_{\rm RH}$ too high, oscillations begin after EMD ends, and the minicluster abundance is suppressed on all scales.

Another quantity of interest is the fraction of DM mass in minihalos, which is given in Eq. (42) and corresponds to the area under the constant-$T_{\rm RH}$ slices of the distributions in Fig. 9. This equation should be evaluated at a high enough redshift that standard ΛCDM-like halos have not started to form. In Fig. 10 we show the fraction of DM in minihalos in the mass range $M \in [M_{\rm osc}, M_{\rm RH}]$ at $z = 250$ as a function of $T_{\rm RH}$ (these masses are defined in Eqs. (39) and (38), respectively). There is no collapse in ΛCDM in this mass range at this time, so all minihalos form solely due to EMD. Lower $T_{\rm RH}$ implies a longer period of enhanced growth, while the effective pressure of lighter ALPs inhibits it. We see that in the regime where the small-scale cutoff of the power spectrum is not important (the $m_a = 10^{-6}$ eV curve

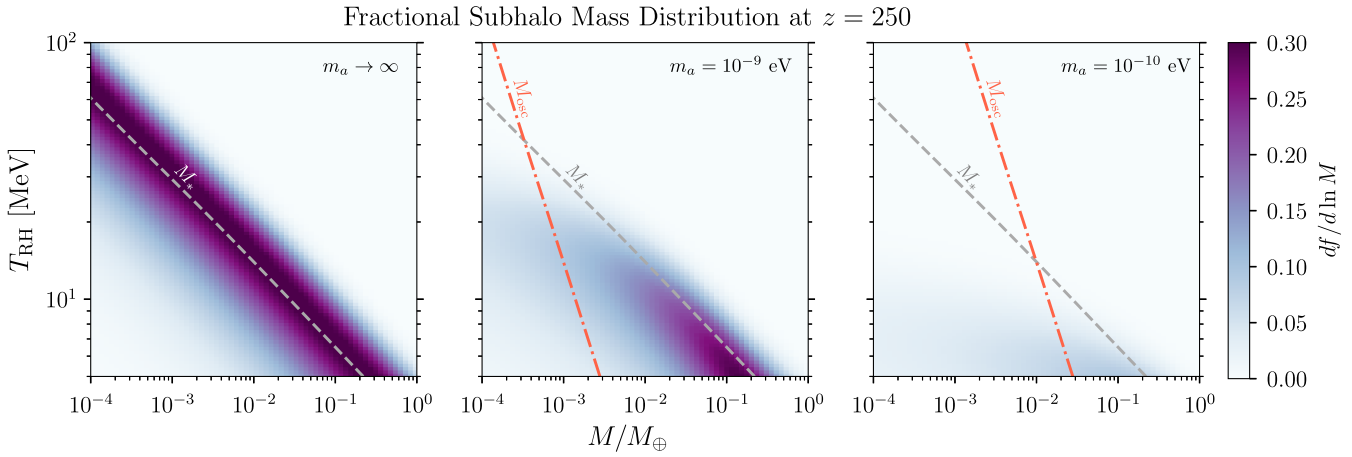

FIG. 9. Fractional minicluster mass distribution $df/d\ln M$ at $z = 250$ as a function of minicluster mass M and reheating temperature $T_{\rm RH}$ for $m_a \to \infty$ (left panel), $m_a = 10^{-9}$ eV (middle panel) and $m_a = 10^{-10}$ eV (right panel). Minicluster formation is suppressed for large $T_{\rm RH}$ and small ALP masses. In each panel the gray dashed line shows the approximate value of M_* [from Eq. (44)], the mass of a typical minicluster forming at this redshift in the CDM ($m_a \to \infty$) limit. The dot-dashed red line shows $M_{\rm osc}$, the DM mass within the horizon at the start of oscillations. Minihalos forming after $z = 250$ are expected to undergo significant tidal disruption in encounters with stars; see the discussion in Sec. V A.

and lower $T_{\rm RH}$), all of DM is already in minihalos at $z = 250$. Higher $T_{\rm RH}$ and lower ALP masses can result in smaller minihalo fractions. Note that these calculations do not take into account possible disruption of minihalos at later times, which may liberate some of the ALPs. Because the minihalos are built up hierarchically, the contents of the disrupted minihalos themselves are made of smaller miniclusters (depending on the ALP mass and the resulting small-scale cutoff). More detailed questions about the distribution of substructure in larger ALP minihalos and galactic halos can be studied using the extended Press-Schechter/excursion set formalism [89,95,96].

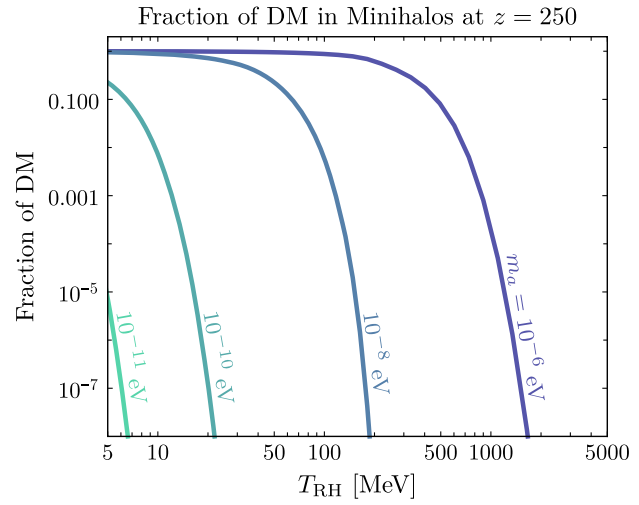

FIG. 10. Fraction of ALP DM bound in minihalos of mass $M \in [M_{\rm osc}, M_{\rm RH}]$ at $z = 250$ as a function of $T_{\rm RH}$ for different choices of the ALP mass. In models with heavier ALPs and lower reheating temperatures, nearly all of DM is bound into minihalos at early times. Estimates in Sec. V A indicate that minihalos forming at $z \gtrsim 250$ are resilient to tidal disruption, suggesting that the fraction of ALPs in minihalos evaluated at early times is robust, i.e., it remains roughly constant through the remaining evolution.

The formation of QCD axion miniclusters in the standard cosmological scenario with postinflationary PQ breaking has been the topic of recent numerical simulation [59,60]. Reference [59] finds that the differential minicluster mass function peaks at $10^{-14}\ M_\odot$ with a long tail to very small masses of around $10^{-17}\ M_\odot$ and a shorter tail to heavier miniclusters up to $10^{-12}\ M_\odot$, measured at matter-radiation equality. The characteristic size of the QCD axion miniclusters in Ref. [60] is slightly larger at $10^{-12}\ M_\odot$. Both of these studies find that the average overdensities in miniclusters are smaller than estimated in Refs. [48–50]. Analytic studies also based on the Press-Schechter formalism find results which are similar [51] or somewhat heavier [52,64]. In contrast, we find that ALP miniclusters which benefit from a period of EMD are heavier than these, with masses $10^{-(6-10)} M_\odot$ depending on the ALP mass, as in Fig. 9.

B. Minicluster density and size

Collapse and decoupling from Hubble flow occur when $\delta = \delta_c \sim 1$. If a minicluster collapses at redshift $z_{\rm c}$ during standard matter domination, the density of the final virialized object can be estimated using the spherical collapse model (see, e.g., Ref. [90]):

$$\rho(z_{\rm c}) \approx 178\bar{\rho}(z_{\rm c}) \approx 3500\ {\rm GeV/cm^3}\left(\frac{1+z_{\rm c}}{250}\right)^3. \quad (46)$$

This allows us to calculate a characteristic radius R_* of a minicluster with mass $M_*(z)$,

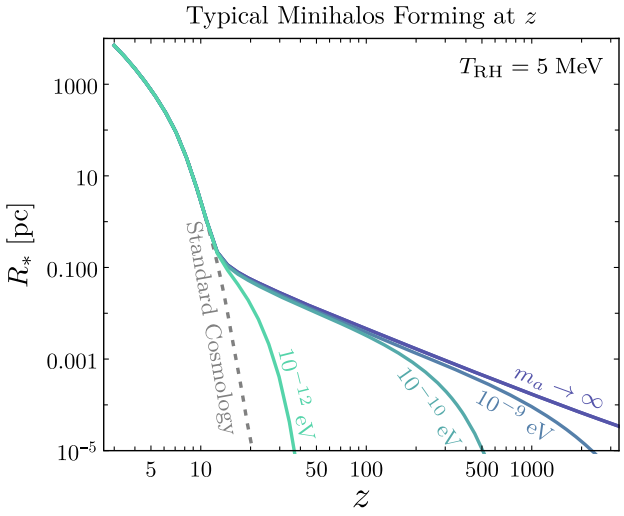

FIG. 11. The radius of a minicluster of mass M_* for $T_{\rm RH} = 5$ MeV and various ALP masses. Color coding is the same as in the left panel of Fig. 8.

$$R_*(z_c) = \left(\frac{3M_*(z_c)}{4\pi\rho(z_c)}\right)^{1/3} \tag{47}$$

$$\sim 10^{-3}\ \text{pc}\left(\frac{5\ \text{MeV}}{T_{\rm RH}}\right)\left(\frac{250}{1+z_c}\right)^{(5+n_s)/(3+n_s)}, \tag{48}$$

where the second line was estimated using Eqs. (44) and (46). R_* is shown in Fig. 11 as function of the redshift of collapse for different ALP masses.

Note that unlike Ref. [39], these calculations include the growth of ALP perturbations during standard matter domination after MRE. As a result, we find that the miniclusters are hierarchically assembled into much more massive minihalos at later redshifts.

In the following section we discuss the implication of ALP DM clustering in the EMD cosmology on direct and indirect probes.

V. ANALYSIS

Early matter domination enhances the growth of structure of ALP DM over a range of scales. As a result the DM distribution is clumpy, and the minihalo spectrum reflects fundamental properties of the ALPs and cosmology (the ALP mass and reheating temperature), as well as the tidal encounter history. Larger minihalos are assembled from smaller ones, but tidal disruption of the larger halos can "free" some of the component subhalos. Because the EMD-enhanced scales collapse early, all of DM is expected to be in minihalos by the time of galaxy formation. This observation has important implications for direct and indirect searches for ALP DM. First, the clumpiness of DM typically decreases the encounter rate of DM objects with laboratories on Earth, reducing the probability of a signal. On the other hand, if we are lucky and such an encounter occurs, the signal is much stronger, since the minihalo density is much larger than the average local DM density. The DM substructure also opens up a range of other probes, e.g., through pulsar timing or through gravitational lensing, as we discuss below.

A. Minicluster survival

Hierarchical structure formation assembled miniclusters into larger and larger objects. Both the early-time assembly and late-time encounters with other minihalos and dense baryonic objects (such as stars in the galactic disk) can disrupt miniclusters. The precise nature of the substructure in our vicinity must be modeled numerically, but we can get a sense of which objects survive using simple estimates, following Refs. [53,97]. Evolution of DM substructure has been extensively studied in the context of CDM and axionlike particles. Reviews of these subjects include Refs. [54,98].

1. Disruption by other miniclusters

Whether the EMD-enhanced minihalos at a given redshift themselves have substructure depends on the precise form of the power spectrum, as well as the age of the clumps. If the power spectrum features an isolated peak (which is the case if $k_{\rm osc} \sim k_{\rm RH}$), then numerical simulations of Ref. [99] indicate that the minihalos lack structure as one would naively expect; these objects then evolve in isolation until they are assembled into galaxies. In the other limit $k_{\rm RH} \ll k_{\rm osc}$, a wide range of scales is enhanced by EMD, and clumps are formed from smaller clumps. Such an initial power spectrum was also studied in Refs. [99,100], and their results suggest that minihalos retain their substructure for at least a factor of ~10 in redshift after formation. In both cases, the analytical arguments of Ref. [101] imply that tidal heating and stripping through minihalo-minihalo encounters is a subdominant effect to stellar and galactic disk encounters. We therefore neglect this effect in what follows, but note that a definitive confirmation of this approximation would require an extremely high-resolution simulation over ~10 Gyr timescales.

Minihalo-minihalo encounters at early times are, however, important for determining the internal density profile $\rho(r)$ of clumps. References [99,100] found that $\rho(r)$ depends on whether the minihalos evolved in isolation (corresponding to the power-spectrum spike mentioned above) or continually accrete other clumps (corresponding to a scale-invariant enhancement). The latter case is analogous to standard hierarchical structure formation in CDM (albeit at much smaller scales) and results in a Navarro-Frenk-White (NFW) density profile:

$$\rho(r) = \frac{4\rho_s}{(r/r_s)(1+r/r_s)^2}, \tag{49}$$

where r_s is the scale radius and $\rho_s = \rho(r_s)$ is the scale density. The scale quantities, including M_s (the mass interior to r_s), can be related to R_* and M_* used previously, and to virial quantities at any redshift. For reference, we catalog these relationships in Appendix E. The spiked initial power spectrum leads to a steeper inner slope, $\rho(r) \propto r^{-3/2}$ for $r < r_s$. The EMD effective power spectrum in Fig. 6 is likely to result in minihalos with density profiles that interpolate between these two limits. In what follows, we take the minihalos to have the NFW profile in Eq. (49). This assumption is conservative for the gravitational probes we consider in Sec. V C, since a shallower inner slope leads to more diffuse structures, thus weakening observational prospects.

2. Disruption in stellar encounters

If a minicluster survives structure formation, it still might be disrupted within the galaxy. Compared to CDM minihalos, the EMD-enhanced substructures form earlier, resulting in denser, more compact objects, which have a higher probability of surviving tidal stripping. We estimate the disruption probability following Refs. [53,97]. An encounter with a star transfers energy to the subcomponents of the minicluster, effectively heating them and decreasing their binding energy. Collisions with an impact parameter $b < b_c$ transfer enough energy to completely unbind the minicluster, where b_c is the critical impact parameter [53,97],

$$b_c^2 \sim \frac{G m_s r_{\rm vir}}{v_{\rm rel} v_{\rm vir}}. \tag{50}$$

Here m_s is the typical stellar mass, $v_{\rm rel}$ is the relative velocity of the star and minicluster, and $r_{\rm vir}$ and $v_{\rm vir} = \sqrt{GM_{\rm vir}/r_{\rm vir}}$ are the virial radius and velocity of the minicluster; these quantities can be related to the NFW scale parameters as described in Appendix E. Collisions with $b > b_c$ transfer less energy but are more likely. Integrating over the entire range of impact parameters yields the approximate probability of disruption in a single traversal of the disk of $p = 2\pi b_c^2 S$, where S is the orbit-averaged column mass density of stars along the minicluster orbit. A typical minicluster has experienced $n_{\rm cross} \sim 100$ crossings in the age of the Milky Way, leading to a total disruption probability of [53,97]

$$p = \frac{2\pi n_{\rm cross} G r_{\rm vir} S}{v_{\rm rel} v_{\rm vir}} \tag{51}$$

$$\approx \left(\frac{n_{\rm cross}}{100}\right)\left(\frac{250}{1+z_c}\right)^{3/2}. \tag{52}$$

Here it is assumed that $v_{\rm rel} = 10^{-3}$, the disc has a constant stellar density, and the distribution of minicluster orbits is isotropic, leading to $S \approx 140\ M_\odot/{\rm pc}^2$ [53]. Note that since this is an estimate for the disruption probability of an individual clump, it only depends on its density [through $1 + z_c$ and Eq. (46)], and it is not sensitive to other model details such as m_a and $T_{\rm RH}$ which determine the abundances of clumps of various sizes. Reference [102] has improved on these approximations by carefully modeling the disc and considering interactions with halo and bulge stars. However, the total disruption probability remains numerically the same. Therefore we see that miniclusters that have formed after $z \sim 250$ are expected to have been disrupted. Conversely, less than 2% of halos that formed around MRE have been disrupted. These numbers should be considered as guidelines rather than hard boundaries for the destruction or survival of substructure. We note that high-resolution simulations of subhalos indicate that even if tidal encounters transfer energy far in excess of the subhalo binding energy, the subhalo is never completely disrupted, even for CDM subhalos [101,103]. This is because the energy injected is not efficiently redistributed among minihalo particles and therefore it is not directly correlated with minihalo survival. Moreover, EMD-enhanced clumps are more compact and have an even larger chance of withstanding such encounters. Thus, the above estimate is most likely conservative and a more realistic calculation of the DM substructure distribution today requires a simulation.

Encounters with baryonic objects can also alter the density profiles of minihalos, recently studied with N-body simulation in Ref. [104]. These events have a dramatic effect on the minihalo density profiles at $r > r_s$, efficiently stripping away the outer minihalos, and resulting in a much steeper falloff for $r > r_s$. The inner core $r < r_s$ remains NFW-like ($\rho \propto 1/r$), with scale radius and density modified by $\mathcal{O}(1)$ in the encounter. The gravitational probes we consider in Sec. V C are mainly sensitive to the "core" mass M_s, so this modification of the density profile does not qualitatively affect our results in the following sections, and we continue to use the NFW profile for simplicity.

B. Direct detection

Dark matter substructure has important implications for direct searches for ALPs. Direct detection experiments in terrestrial laboratories will detect ALPs only when Earth encounters a clump or its remnant. If the clumps remain intact, the typical time between such encounters, $\tau_{\rm enc}$, is

$$\tau_{\rm enc} = \frac{1}{n\sigma v_{\rm rel}}, \tag{53}$$

where, assuming all of DM is in minihalos of similar mass and size, $n = \rho_{\rm dm}/M_*$ is the local clump number density, $\sigma \sim \pi R_*^2$ is their geometric cross section and $v_{\rm rel}$ velocity relative to Earth. Since $\sigma \propto M^{(5+n_s)/3}$, the time between encounters is smaller for heavier miniclusters. This is a

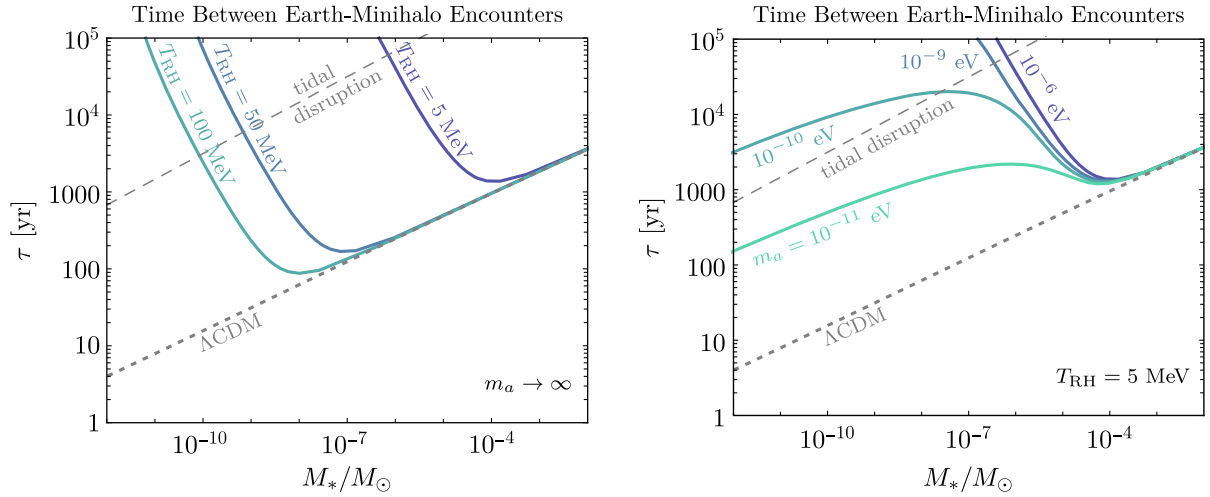

FIG. 12. Time between Earth-minihalo encounters assuming all of DM is inside minihalos of a single mass that survive tidal disruption until today. In the left panel we fix $m_a \gtrsim 10^{-6}$ eV so that the small-scale cutoff due to the ALP effective pressure is irrelevant; in the right panel we take $T_{\rm RH} = 5$ MeV and vary the ALP mass. Early matter domination produces minihalos at high redshift, leading to dense and therefore compact minihalos. The resulting reduced geometric cross section increases the time between encounters for $M < M_{\rm RH}$, despite the increasing number density for smaller M. Smaller ALP masses suppress growth of small scales, leading to the formation of more diffuse objects with larger cross sections and encounter rates. Gray dotted lines correspond to encounter rates for ΛCDM minihalos. Above the thin gray dashed line, tidal disruption of minihalos due to stellar encounters is expected to be unimportant.

consequence of the fact that the heavier miniclusters are less dense (due to their later formation), leading to a cross section that grows faster than the constant density expectation $\propto M^{2/3}$; this effect compensates for the decreasing number density of heavier miniclusters at fixed local DM density $\rho_{\rm dm}$. As we showed previously, in cosmologies with EMD, the smallest bound objects are assembled into larger miniclusters. Further evolution may disrupt these, so their final mass and size distribution depends on the merger history of the Milky Way halo and interactions with the disk. With this caveat in mind, we show the time between Earth-minihalo encounters in Fig. 12. The left panel corresponds to the CDM case with various values of $T_{\rm RH}$, where ALP effective pressure effects are not important. Note that regime is reached already for $m_a \gtrsim 10^{-6}$ eV for scales of interest. In the right panel we fix $T_{\rm RH} = 5$ MeV and show the effect of decreasing m_a, which suppresses the growth of structure at small scales, delaying collapse and leading to clumps that are more diffuse at a given mass. Their larger size leads to larger cross sections and therefore higher encounter rates compared to the $m_a \to \infty$ case.

A minihalo-Earth encounter implies a higher density of DM in the laboratory than expected from the local volume average and a different velocity dispersion compared to the unclustered scenario. The latter fact means that signal frequency linewidth $\delta f/f$ in resonant detectors is orders of magnitude smaller. The naive expectation from the typical kinetic energy of ALPs is $\delta f/f \approx v_s v_{\rm rel}$, where $v_s^2 \sim GM_s/r_s$ is the scale velocity and $v_{\rm rel} \sim 10^{-3}$ is the relative Earth-minihalo velocity. For example, if the effective pressure effects are not important, Eqs. (44) and (46) imply $v_s \sim 10^{-8}(5\ {\rm MeV}/T_{\rm RH})$ [we approximated $n_s \approx 1$ and dropped factors of $g_*(T_{\rm RH})$], which is much smaller than the galactic virial velocity of $\sim 10^{-3}$. This suggests that the signal can be orders of magnitude narrower than in unclustered models. However, as the minihalo enters the solar system, it will experience tidal forces from the Sun, which impart different velocities to different parts of the minihalo, resulting in an additional drift in the signal frequency as the minihalo crosses the laboratory [62]. The measurement time can be limited to ensure that the drift can be ignored or the drift can be incorporated as part of the signal template. In the former case, Ref. [62] estimates that the minimum signal width is $\delta f/f \sim 10^{-11}(\mu{\rm eV}/m_a)^{1/2}$, which can be comparable or larger than the intrinsic minihalo dispersion (depending on m_a and $T_{\rm RH}$). As emphasized in that work, the narrow width of the signal in models with ALP substructure suggests that a broadband search strategy is beneficial in efficiently probing the ALP parameter space.

The previous discussion is rather optimistic, since Fig. 12 indicates that in the absence of disruption the expected encounter rates in EMD cosmologies with low $T_{\rm RH}$ are small. However, tidal disruption through interactions with the disk or other clumps can lead to a population of DM streams, which occupy a more significant fraction of the local volume at the price of reducing the density enhancement compared to an intact clump [53,102,105,106]. In Fig. 12, the region of parameter space below the thin dashed line is where tidal disruption due to interactions with stars may be important; this line

scales as $M^{1/3}$ and corresponds to the time between encounters, Eq. (53), evaluated at fixed disruption probability as obtained from Eq. (52). Thus in some parts of parameter space minihalos are expected to undergo disruption in the galactic gravitational field and form tidal streams. We have estimated the stream-crossing rate following Refs. [53,102], finding that it can approach once per year at the price of losing the density enhancement of the progenitor minihalo. However, the stream may still retain a lower velocity dispersion which can have an impact on direct detection as discussed above. We leave a detailed study of stream formation and properties, including the effects of a realistic minihalo distribution to future work.

C. Lensing and pulsar timing

The DM clumps formed after EMD are much denser than minihalos in ΛCDM, but are far from compact. The size of these miniclusters is much larger than the typical Einstein radii of stellar and supernova gravitational lens systems [58]. As a result, these objects do not lead to signals in microlensing searches such as Refs. [107–109]. However, new techniques to search for more diffuse objects have recently been proposed in Refs. [56–58]. We estimate the sensitivity of photometric microlensing [57] and pulsar timing [58] to axion minihalos produced from EMD below (the astrometry proposal of Ref. [56] is sensitive to heavier subhalos that cannot arise from EMD).

Photometric monitoring of individual stars behind galaxy cluster lenses can be used to observe the imprints of substructure on microlensing light curves [57]. The stars of interest are those which are very near to a cluster lens caustic; these stars are therefore highly magnified. If such a star then undergoes a microlensing event due to compact objects inside the cluster (stars or black holes), their brightness becomes variable on the timescale of the observation and the individual star can be studied [110]. The lensed stars can experience magnifications of up to $\mu \sim 10^{3-4}$ under these circumstances. The proximity to the caustic and the resulting large amplification means that the magnification matrix is nearly singular—its determinant is tuned close to zero at the level of μ^{-1}. This tuning means that the brightness of the observed star is sensitive to surface density fluctuations in the cluster lens at that level. Thus, the presence of DM clumps within the cluster can then lead to $\mathcal{O}(1)$ brightness fluctuations if they produce surface density fluctuations of $\mathcal{O}(\mu^{-1})$ on the relevant time and length scales [57]. This phenomenon has recently been observed for a small number of stars [111,112] using the Hubble Space Telescope. Those initial measurements were not sensitive to the presence of miniclusters, which would require dedicated monitoring over a period of days or weeks either with HST or with future telescopes such as James Webb Space Telescope.

To estimate how sensitive this technique would be to the presence of ALP minicluster we follow the simplified analysis of Ref. [62], based on the proposal of Ref. [57]. We consider the miniclusters to have a standard NFW density profile [Eq. (49)], with a distribution described by a fractional halo mass function $df/d\ln M$. Assuming there is a large number of DM minihalos in the cluster along the line of sight, the resulting random surface density fluctuations can be described by the lensing convergence power spectrum [57], given by

$$P_\kappa(q) = \frac{\Sigma_{\rm cl}}{\Sigma_{\rm cr}^2}\int dM \frac{df}{d\ln M}\left(\frac{|\tilde{\rho}(q,M)|}{M}\right)^2, \tag{54}$$

where $\Sigma_{\rm cl}$ is the cluster surface density, $\Sigma_{\rm cr}$ is the critical surface density, and $\tilde{\rho}$ is the Fourier transform of the NFW density profile, and $q = 2\pi/r$ is the inverse length scale of the fluctuations. The critical surface density can be expressed in terms of an effective distance $D_{\rm eff}$ by $\Sigma_{\rm cr} = 1/(4\pi G D_{\rm eff})$. Following Ref. [62], we make the simplifying assumption that some fraction f of the DM is contained entirely within clumps with mass $M = M_s$, so that

$$\frac{df}{d\ln M} = fM\delta(M - M_s). \tag{55}$$

Note that we do not use the Press-Schechter estimate of $df/d\ln M$ derived in Sec. IV, since it does not account for minihalo disruption or sub-substructure.

We obtain the sensitivity of photometric microlensing observations by comparing the dimensionless power spectrum $\Delta_\kappa(q)$ to the (expected) amplitude of observable surface density fluctuations $\mathcal{O}(10^{-3-4})$ for realistic lenses. The power spectrum is

$$\Delta_\kappa(q) = \left[\frac{q^2 P_\kappa(q)}{2\pi}\right]^{1/2} = \frac{1}{\ln(2/\sqrt{e})}\frac{[\Sigma_{\rm cl} f M_s]^{1/2} q r_s g(q r_s)}{\Sigma_{\rm cr} r_s} \tag{56}$$

where r_s is the NFW scale radius, M_s is the mass within the scale radius, and

$$g(x) = \frac{1}{2}\sin(x)(\pi - 2Six) - \cos(x)Cix \tag{57}$$

comes from the Fourier transform of the halo density profile. The quantity $q r_s g(q r_s)$ is maximized at $q r_s = 0.77$ with a value of 0.35. We take $D_{\rm eff} = 1$ Gpc and $\Sigma_{\rm cl} = 0.8\Sigma_{\rm cr}$; these numbers roughly correspond to the observed highly magnified star LS1 [111]. We define the sensitivity of photometric lensing by requiring $\Delta_\kappa(q)$ in Eq. (56) to be larger than 10^{-3}, while a number of other conditions are simultaneously satisfied. First, the length scale of the fluctuations $r = 2\pi/q$ must be larger than the minimum length scale ∼10 AU probed in the lens plane; this sets a lower bound in M_s for the sensitivity, since smaller

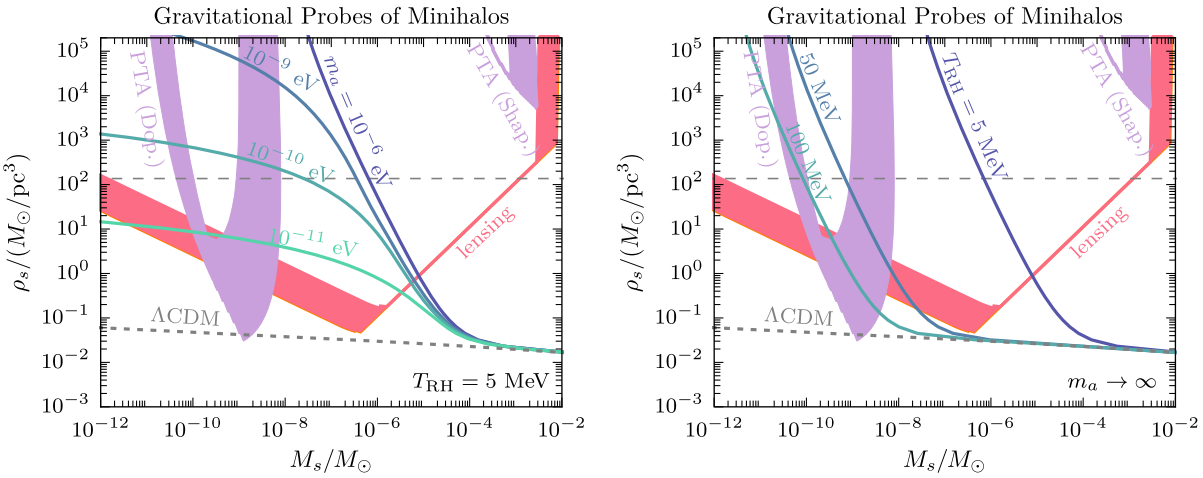

FIG. 13. The reach of future pulsar timing array (PTA) Doppler and Shapiro dynamic measurements (purple) and photometric microlensing (red) in the M_s–ρ_s plane. In each panel, the (upper) gray dashed line corresponds to a collapse redshift of 250: the region of the plane above this has $z_c > 250$ with minihalos that are likely to survive tidal disruption in stellar encounters. The (lower) gray dotted line shows the prediction from the standard ΛCDM scenario. The left panel shows EMD predictions for $m_a = 10^{-6}$, 10^{-9}, 10^{-10} and 10^{-11} eV and fixed $T_{\rm RH} = 5$ MeV. The right panel shows EMD predictions for $T_{\rm RH} = 10$, 50 and 100 MeV for fixed $m_a \to \infty$ (this limit is already reached for $m_a \gtrsim 10^{-6}$ eV). The thickness of the PTA and lensing projections corresponds to varying the DM fraction in minihalos of mass M_s between 0.3 and 1. The actual fraction of ALPs in minihalos can span a wide range and depends on model parameters (see Fig. 10) and tidal disruption history.

minihalos would give density fluctuations that are too rapid to be detected. Second, the characteristic size of the clump must be smaller than the largest length scale of the microlensing event, $r_s < d$ where $d \sim 10^3$ AU; minihalos that exceed this size do not give rise to star magnification fluctuations on the timescale of the lensing event. Note that this condition does not depend on the fraction of DM in minihalos. Finally, to ensure that there are many clumps along the line of sight we require $f\pi(d/2)^2\Sigma_{\rm cl}/M_s > 10$ such that the fluctuations can be described by a power spectrum [62]. This sets the sharp cutoff of $M_s \sim 10^{-2}\ M_\odot$. To obtain the sensitivity of the future lensing search in the M_s–ρ_s plane we evaluate Eq. (56) at the value of $q = 2\pi/r$ that maximizes the sensitivity, subject to the constraints above. We show the projected limits as filled red regions in Fig. 13, where the thickness of the bands arises from varying the DM fraction in clumps of mass M_s between $f = 0.3$ and 1 in the M_s–ρ_s plane. DM substructures with scale densities above these bands will be testable with photometric lensing. Figure 13 also shows the region which can be constrained by future pulsar timing array searches, which we now discuss.

Reference [58] has argued that even comparatively diffuse minihalos can be probed with pulsar-timing array (PTA) measurements. They consider two types of signal. The first is a Doppler-shift in the pulsar frequency as a DM clump passes near the star or Earth. The second is a Shapiro time delay if a minihalo crosses our line of sight to the pulsar. Near-term facilities (and in particular the Square Kilometre Array) will be sensitive to these signals. Following [58], we assume that 73 currently known pulsars will continue to be monitored for the next 30 years, and that SKA will discover 200 more which it will measure for 20 years with 50 ns timing accuracy. We show the sensitivity of a search for such anomalous frequency shifts in Fig. 13 in purple. The projections shown here correspond to 30 years of observations with the current pulsar dataset, and 20 years of observations with the Square Kilometre Array (SKA) (i.e., assuming that SKA starts in 10 years from the time of writing). The main sensitivity occurs at $M_s \sim 10^{-9}\ M_\odot$ for the Doppler dynamic signal, with the small spike feature at $M_s \sim 10^{-3}\ M_\odot$ corresponding to the Shapiro dynamic signal. Again, the width of the band corresponds to scanning over the clump fraction from 0.3 to 1. These projections were obtained by rescaling the projected sensitivity to primordial black holes by assuming the NFW density profile in Eq. (49) as described in Ref. [58]. The necessary conversions between the virial quantities used in Ref. [58] and M_s and ρ_s are discussed in Appendix E.

The left panel of Fig. 13 shows the target parameter space corresponding to $m_a = 10^{-6}$, 10^{-9}, 10^{-10} and 10^{-11} eV for $T_{\rm RH} = 5$ MeV. In order to relate R_* and M_* obtained in Sec. IV to M_s, r_s and ρ_s, we assumed a particular concentration parameter at formation, $c_* = R_*/r_s$ (R_* is just the virial radius at collapse; see Appendix E). Based on the compilation of simulated Earth-mass minihalos in ΛCDM from Ref. [113], we take $c_* \approx 2$.[7] Larger concentration

[7]Reference [57] instead took $c_* = 4$, which is tuned to simulations of galaxy-sized DM halos.

parameters at formation imply more compact halos (larger ρ_s at fixed M_s) that are easier to detect with lensing or PTA. Each point on the lines in Fig. 13 corresponds to a specific value of the collapse epoch z_c for the clumps of that mass, since z_c relates the scale mass to the scale radius (and hence scale density). The horizontal gray-dashed line at $\rho_s \sim 10^2\ M_\odot/\mathrm{pc}^3$ corresponds to a collapse redshift of 250; according to Eq. (52), minihalos that form at this redshift or earlier are likely to survive tidal disruption in the galaxy. We therefore expect EMD clumps which collapse before $z_c = 250$ and where the ALP mass is larger than around 10^{-10} eV to be detectable by photometric microlensing, and possibly by Doppler dynamic PTA searches. This mass range was indicated in the ALP mass-coupling parameter space of Fig. 1 as a green arrow. Finally, the gray dotted line shows the results for the standard ΛCDM scenario. Since collapse happens relatively late in that case, none of the CDM miniclusters would be dense enough to turn up in these searches (even without considering the fact that they are quite easily disrupted in hierarchical structure formation). The right panel of Fig. 13 shows the same projections as the left panel but with the model curves corresponding to different reheating temperatures $T_{\mathrm{RH}} = 5$, 50 and 100 MeV with the ALP mass fixed to the CDM limit of $m_a \to \infty$. We see that lensing and PTA will be sensitive to a wide range of cosmologies with reheat temperatures as high as $T_{\mathrm{RH}} \sim 100$ MeV (and possibly higher if our estimates of the minihalo survival probability are too conservative).

VI. CONCLUSION

Early matter domination is a natural feature of many UV completions of the Standard Model, including supersymmetric theories and various hidden sector models. If the cosmological history included a period of EMD, both the relic abundance of dark matter and the growth of its density perturbations are modified relative to ΛCDM. Nonthermally produced dark matter candidates, including axions produced through misalignment, are particularly sensitive to the expansion history of the Universe. In the axion case, EMD yields a relic abundance that is independent of the axion mass and favors high f_a, including a QCD axion window of roughly $m_a \sim 10^{-(8-9)}$ eV and $f_a \sim 10^{15-16}$ GeV.

In this work we studied the evolution of ALP density perturbations and the resulting DM substructure in cosmologies with early matter domination. During EMD density perturbations grow linearly with the scale factor, enhancing the density contrast on scales smaller than the horizon size at reheating and larger than the Jeans scale set by the effective ALP pressure. This enhancement of substructure results in early formation of ALP minihalos, and subsequently their hierarchical assembly into larger and larger objects. The high redshift of formation results in DM structures that have a typical density much larger than DM halos in the standard cosmology. Given the constraints on late reheating, the largest objects that can benefit from EMD-enhanced growth have a mass of $\mathcal{O}(1-100)M_\oplus$. Since all of the DM ends up in minihalos, the direct detection rates are suppressed by the time between Earth-minihalo encounters. These times are longer than the timescale of typical experiments, making this search strategy impractical if all minihalos remain intact. However, it is not clear whether all minihalos survive until today. Tidal disruption processes include clump-clump and clump-star encounters. If a significant fraction of clumps are disrupted, the encounter rates with the resulting streams may be significantly larger than for isolated minihalos. Simulations are required to reliably estimate the disruption probability over the lifetime of the galaxy and the resulting ALP volume-filling fraction. It will also be important to study higher-order terms in the potential, and the formation and impact of ALP solitons inside miniclusters.

If the minihalo survival probability is high, we find that proposed astrophysical detection techniques offer strong sensitivity. Pulsar timing measurements are sensitive to the Doppler shift induced by a minihalo passing close a pulsar. The enhanced small-scale structure can also have an observable imprint on the microlensing lightcurves of highly magnified stars. These observations are sensitive to a wide range of the relic-density target regions for different reheating temperatures and natural misalignment angles. While both techniques require long-term observations on decadal timescales, it is important to note that the DM power spectrum at small scales is nearly unconstrained. Early matter domination provides an illustrative example of the interesting physics that can be imprinted on these scales. There are also other compelling possibilities leading to similar physics, including different modifications of the expansion history (e.g., a period of kination [114]), a running spectral index and gravitational particle production. Astrophysical searches for the small-scale structure of matter can thus provide a crucial window in the prenucleosynthesis Universe and offer hints about the origin of DM.

ACKNOWLEDGMENTS

We thank Adrienne Erickcek, Albert Stebbins, Gabriela Barenboim, Dan Hooper, Nickolay Gnedin, Tom LeCompte, Jessie Shelton and Yang Bai for useful conversations, Jonathan Kozaczuk for collaboration in early stages of this work, and Felix Draper for incisive criticism. We are especially grateful to Jeff Dror, Harikrishnan Ramani and Ken Van Tilburg for helpful discussions on pulsar timing and lensing. This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. M. J. D. is supported by the Australian

Research Council. The work of P. D. is supported by NSF Grant No. PHY-1719642. We thank the Galileo Galilei Institute for Theoretical Physics for the hospitality and the INFN for partial support during the completion of this work.

APPENDIX A: BOLTZMANN EQUATIONS IN DETAIL

In this section we express the linear Boltzmann equations (17) in convenient dimensionless variables defined in Eq. (9) and using the scale factor a as a time variable. This form of the Boltzmann system is easily implemented and solved numerically. The resulting perturbation equations are [43]

$$a^2E\delta_\phi' + \theta_\phi + 3a^2E\Phi' = a\Gamma_\phi\Phi, \tag{A1a}$$

$$a^2E\theta_\phi' + aE\theta_\phi + k^2\Phi = 0, \tag{A1b}$$

$$a^2E\delta_r' + \frac{4}{3}\theta_r + 4a^2E\Phi' = \frac{\rho_\phi}{\rho_r}a\Gamma_\phi[\delta_\phi - \delta_r - \Phi], \tag{A1c}$$

$$a^2E\theta_r' + k^2\Phi - k^2\frac{\delta_r}{4} = \frac{\rho_\phi}{\rho_r}a\Gamma_\phi\left[\frac{3}{4}\theta_\phi - \theta_r\right], \tag{A1d}$$

$$a^2E\delta_a' + \theta_a + 3a^2E\Phi' = -3c_{\rm nad}^2aE\delta_a - 9c_{\rm nad}^2a^2E^2\theta_a/k^2, \tag{A1e}$$

$$a^2E\theta_a' + aE\theta_a + k^2\Phi = +3c_{\rm nad}^2aE\theta_a + k^2c_{\rm nad}^2\delta_a, \tag{A1f}$$

$$k^2\Phi + 3aE^2[a^2\Phi' + a\Phi] = \frac{3}{2}a^2[\rho_\phi\delta_\phi + \rho_r\delta_r + \rho_a\delta_a], \tag{A1g}$$

where the prime denotes differentiation with respect to a, E is the dimensionless Hubble parameter

$$E^2 = \rho_\phi + \rho_r + \rho_a, \tag{A2}$$

and

$$c_{\rm nad}^2 = \frac{k^2}{k^2 + 4m_a^2a^2}. \tag{A3}$$

The energy densities are normalized as in Eq. (9), while Γ_ϕ, m_a and k are in units of H_1, the Hubble rate at an arbitrary initial time.

1. Initial conditions

To derive initial conditions for the perturbation equations in Eq. (A1) we follow [43], with some changes. First, we assume that all relevant modes are initially subhorizon, such that $k \ll 1$. In this limit we can approximately solve the perturbation equations while expanding in k. Superhorizon modes do not evolve, i.e., $\delta_i' = 0$ and $\theta_i \sim \mathcal{O}(k^2)$. This implies that the right-hand side of Eqs. (A1c) vanishes, leading to the following constraint at leading order in k:

$$\delta_\phi - \delta_r - \Phi = 0. \tag{A4}$$

Furthermore, Eq. (A1g) can be solved for δ_ϕ using the early-time background solutions in Eq. (13), giving

$$\delta_\phi = 2\Phi. \tag{A5}$$

The two above equations in turn imply that

$$\delta_r = \Phi \tag{A6}$$

to leading order in k.

For the ALP density perturbation, we assume that low-scale inflation or nontrivial inflationary dynamics prevents the generation of a large isocurvature mode. The adiabatic mode is zero before the ALP starts to oscillate [9]. After oscillations begin, the correct initial condition for the superhorizon density perturbations becomes adiabatic,

$$\delta_a = \delta_\phi. \tag{A7}$$

We show that this also follows from approximate solutions to the perturbed field equations in Appendix B.

Given the initial assumptions above, the right-hand side of Eq. (A1e) also vanishes. This allows us to relate δ_a to θ_a via $\delta_a = -3aE\theta_a/k^2$, where we consider $c_{\rm nad}^2 \sim \mathcal{O}(k^0)$. Substituting this into the right-hand side of Eq. (A1f), we observe that it vanishes, leaving

$$a^2E\theta_a' + aE\theta_a + k^2\Phi = 0. \tag{A8}$$

During early matter domination we have $E(a) \simeq a^{-3/2}$. This is solved by

$$\theta_a = -\frac{2}{3}k^2\sqrt{a}\Phi. \tag{A9}$$

We obtain an identical solution

$$\theta_\phi = -\frac{2}{3}k^2\sqrt{a}\Phi \tag{A10}$$

from Eq. (A1b) with the same logic. Using this in Eq. (A1d) along with $\delta_r \sim \Phi$ we finally obtain

$$\theta_r = -\frac{2}{3}k^2\sqrt{a}\Phi. \tag{A11}$$

The set of Eqs. (A5), (A6), (A7), (A9), (A10), (A11) form our set of initial conditions.

APPENDIX B: ALP INITIAL CONDITIONS FROM FIELD EQUATIONS

In this section we derive the initial conditions for the ALP density perturbation δ_a from the perturbed field equations. We will denote the ALP field by $\varphi = \varphi_0 + \varphi_1$ to avoid confusion with the scale factor a, and work in the dimensionless variables defined below Eq. (7). The modes of interest enter the horizon during EMD when $\mathcal{H} = 1/\sqrt{a}$ and the gravitational potentials are constant with $\Phi + \Psi = 0$. In this regime, the background and perturbed field equations, Eqs. (3) and (16), simplify to

$$\varphi_0'' + \frac{5}{2a}\varphi_0' + m_a^2 a\varphi_0 = 0, \tag{B1}$$

$$\varphi_1'' + \frac{5}{2a}\varphi_1' + (k^2/a + m_a^2 a)\varphi_1 = +2m_a^2 a\varphi_0\Phi, \tag{B2}$$

where primes denote derivatives with respect to the scale factor. We take the initial conditions for a completely smooth field initially at rest with a misalignment value φ_i:

$$\varphi_0(0) = \varphi_i, \qquad \varphi_0'(0) = 0, \tag{B3}$$

$$\varphi_1(0) = 0, \qquad \varphi_1'(0) = 0. \tag{B4}$$

The background equation is easily solved to give

$$\varphi_0 = \frac{3\varphi_i}{2m_a a^{3/2}} \sin\left(\frac{2}{3} m_a a^{3/2}\right). \tag{B5}$$

Next, we consider the superhorizon evolution of modes that enter the horizon after oscillations have begun, $k \ll k_{\rm osc} \sim m_a^{1/3}$. This allows us to drop the k^2 term in Eq. (B2) in comparison to either the Hubble damping term $\propto \varphi'$ or the mass term. We then find the $\mathcal{O}(k^0)$ solution

$$\varphi_1 \approx -\Phi\varphi_i \cos\left(\frac{2}{3} m_a a^{3/2}\right) + \frac{3\Phi\varphi_i}{2m_a a^{3/2}} \sin\left(\frac{2}{3} m_a a^{3/2}\right). \tag{B6}$$

We can now construct the energy density and its perturbation from [115]

$$\rho_a = \frac{1}{2a}(\varphi_0')^2 + \frac{1}{2} m_a^2 \varphi_0^2, \tag{B7}$$

$$\delta\rho_a = \frac{1}{a}(\varphi_0'\varphi_1' + (\varphi_0')^2\Phi) + m_a^2\varphi_0\varphi_1. \tag{B8}$$

Averaging over oscillations, we find that for superhorizon scales with $k \ll k_{\rm osc}$

$$\rho_a \approx \frac{9\varphi_i^2}{16a^3}, \tag{B9}$$

$$\delta\rho_a \approx \frac{9\varphi_i^2\Phi}{8a^3}, \tag{B10}$$

and therefore

$$\delta_a \approx 2\Phi \tag{B11}$$

at leading order in k. Thus, we see that even though the ALP field starts as completely homogeneous with $\delta_a = 0$, superhorizon evolution in the gravitational potential ensures that it locks onto the matter adiabatic mode after oscillations have begun.

APPENDIX C: FITTING FUNCTIONS FOR ALP DENSITY CONTRAST EVOLUTION

In this section we define the various semianalytical fitting functions that we use to evaluate the fluctuation variance for a wide range of parameters, without having to solve for the evolution of all modes numerically from horizon entry to now. These fitting functions are nearly identical to those presented in Ref. [43] despite slightly different initial conditions, owing to the fact the ALP DM cannot be produced in the decays of the EMD field ϕ. The reason for this similarity is that the late-time evolution of modes is insensitive to this initial condition. This is already evident in the approximate solution in Eq. (19), since the growing term quickly overtakes the initial value once the mode is inside the horizon. We confirmed this by comparing the various fitting functions below to the numerical solutions discussed in Sec. III C. The ALP fluid and CDM evolve differently at small scales, which we implement as a simple cutoff as we discuss below.

The differences between ΛCDM and EMD in the evolution of the DM density contrast are neatly encapsulated by the ratio

$$R(k) = \frac{\delta_a}{\delta_c} = \frac{A(k)\ln\left[\left(\frac{4}{e^3}\right)^{f_2/f_1}\frac{B(k)a_{\rm eq}}{a_{\rm hor}(k)}\right]}{9.11\ln\left[\left(\frac{4}{e^3}\right)^{f_2/f_1}\frac{0.594\sqrt{2}k}{k_{\rm eq}}\right]}, \tag{C1}$$

where δ_c refers to the evolution of the CDM density contrast in ΛCDM and [43,81]

$$f_1 = 1 - 0.568 f_b + 0.094 f_b^2, \tag{C2}$$

$$f_2 = 1 - 1.156 f_b + 0.149 f_b^2 - 0.074 f_b^3, \tag{C3}$$

with $f_b = \Omega_b/\Omega_m$ and [43]

$$\frac{a_{\rm hor}(k)}{a_{\rm eq}} \approx \frac{k_{\rm eq}}{\sqrt{2}k}\left[1 + \left(\frac{k}{k_{\rm RH}}\right)^{4.235}\right]^{-1/4.235}. \tag{C4}$$

This scaling relation was obtained by fitting this approximate form to numerical solutions of $k = \mathcal{H}(a_{\rm hor})$; we confirmed the results of Ref. [43]. The values of $A(k)$, $B(k)$ and in various limits are given by

$$A = \begin{cases} 0 & k > k_{\rm osc} \\ \frac{3}{5}\left(\frac{k}{k_{\rm RH}}\right)^2 & k_{\rm osc} > k > k_{\rm RH} \\ 9.11 & k < k_{\rm RH} \end{cases},$$
$$B = \begin{cases} e\left(\frac{k_{\rm RH}}{k}\right)^2 & k > k_{\rm RH} \\ 0.594 & k < k_{\rm RH} \end{cases}. \quad \text{(C5)}$$

The functions A and B interpolate between the linear EMD growth and logarithmic RD evolution. Reference [43] provides numerical functions that smoothly connect the two limits above. Note that we model the small-scale suppression of power due to the effective ALP mass by a hard cutoff at $k = k_{\rm osc}$. This is an approximation as the actual falloff is much smoother; see the right panel of Fig. 5. However, this approximation allows for fast exploration of the clump parameter space without having to solve the full Boltzmann system for each $(m_a, T_{\rm RH})$.

The density contrasts in Eq. (C1) are evaluated at matter-radiation equality, so the evolution at later times is captured by a scale-dependent growth function $TD_{\Lambda\rm CDM}(a,k)$ defined via Eqs. (31) and (35). Since the density variance in Eq. (30) involves an integral over all scales within the horizon, we need $TD_{\Lambda\rm CDM}(a,k)$ for a wide range of k and a. We obtain $TD_{\Lambda\rm CDM}$ by stitching together solutions from the Boltzmann solver `CLASS` [85,116] at small k and the Eisenstein-Hu interpolating formula [117] at high k. This matching is performed at $k/k_{\rm eq} \sim 10^6$ (where both the `CLASS` calculation and the Eisenstein-Hu formula are accurate) and scale factor a_m and the result is then propagated forward or backward in time using solutions of the Meszaros equation [81]:

$$\delta_c'' + \frac{2+3y}{2y(1+y)}\delta_c' = \frac{3}{2y(1+y)}(1-f_b)\delta_c, \quad \text{(C6)}$$

where primes denote derivatives with respect to $y = a/a_{\rm eq}$ and $f_b = \Omega_b/\Omega_m$. We denote the DM density contrast with a subscript c to emphasize that we are now discussing evolution in ΛCDM—the EMD and ALP dynamics are encapsulated by $R(k)$ defined above. This equation is valid before and after equality, and well after horizon entry; the two solutions $U_{1,2}$ are expressed in terms of hypergeometric functions in Ref. [81] and they can be matched onto the standard radiation-domination solution, given in Eq. (24). As a result, the evolution of the density contrast can be factorized as

$$\delta_c(a,k) = \mathcal{D}(a)\delta_c(a_{\rm eq},k), \quad \text{(C7)}$$

where

$$\mathcal{D}(a) = U_1(a/a_{\rm eq}) + \frac{A_1}{A_2}U_2(a/a_{\rm eq}). \quad \text{(C8)}$$

The coefficients $A_{1,2}$ (obtained by matching in the RD regime such that $\mathcal{D} \to 1$ as $a/a_{\rm eq} \to 0$) and the functional form of $U_{1,2}$ are given in Ref. [81]. The function $\mathcal{D}$ captures the linear growth during standard matter domination since $U_1(a/a_{\rm eq}) \sim (a/a_{\rm eq})$ for $f_b = 0$ and $a/a_{\rm eq} \gg 1$. The full expression is accurate near the transition from logarithmic growth (captured by U_2) to linear evolution, and $f_b \neq 0$.[8] The scale-dependent growth function at an arbitrary redshift is then given by

$$TD_{\Lambda\rm CDM}(a,k) = TD_{\Lambda\rm CDM}(a_m,k)\left[\frac{\mathcal{D}(a)}{\mathcal{D}(a_m)}\right]. \quad \text{(C9)}$$

The scale factor a_m at which the numerical and semi-analytic expressions for $TD_{\Lambda\rm CDM}$ are matched is arbitrary, and can be chosen to minimize the error made in the simple extrapolation using $\mathcal{D}$. Similar to Ref. [43], we find that matching at $z_m = 50$ and using $\mathcal{D}(a)$ leads to fractional errors of $< 4\%$ for a wide range of redshifts and scales of interest. This procedure ensures that the amplitude of density fluctuations on large scales is correctly normalized. In particular, using Eq. (30), we reproduce the observed value of $\sigma_8 = \sigma(z=0, R = 8/h{\rm Mpc}) \approx 0.8$ [70].

APPENDIX D: ISOCURVATURE CONSTRAINTS

Planck [118] constrains ratio of the scalar-to-isocurvature amplitude to be $\alpha < 0.038$. For theories with a period of EMD we proceed following Ref. [68]. The definition of the isocurvature perturbation is

$$S_i = \frac{\delta(n_i/s)}{n_i/s} = \frac{\delta n_i}{n_i} - 3\frac{\delta T}{T}. \quad \text{(D1)}$$

We assume that only the axion has $S_a \neq 0$, with all other fields satisfying $S_i = 0$. The invariance of the local energy density under isocurvature perturbations can then be used to relate S_a to the temperature perturbation $\delta T/T$ as follows. The total energy is

$$\rho = \sum_i m_i n_i + m_a n_a + \rho_r \quad \text{(D2)}$$

and we require

$$\delta\rho = \sum_i m_i \delta n_i + m_a \delta n_a + 4\rho_r \frac{\delta T}{T} \simeq 0. \quad \text{(D3)}$$

The second equality in Eq. (D1) implies $\delta n_i = n_i 3\delta T/T$ for $i \neq a$. Substituting that in the equation above we obtain

$$\frac{\delta T}{T} \simeq -\frac{\rho_a}{3\sum_i \rho_i + 4\rho_r} S_a, \quad \text{(D4)}$$

[8] On scales larger than the baryonic Jeans length, the $f_b = 0$ solution is appropriate since the baryons are no longer pressure-supported and collapse like CDM. We interpolate between the $f_b = 0$ and $f_b \neq 0$ regimes using the prescription in Ref. [43].

which implicitly requires $\delta T/T \simeq 0$ for the ALP. For fluctuations on superhorizon scales which enter the horizon during standard matter domination (after $T_{\rm eq}$) the above implies

$$\frac{\delta T}{T} \simeq -\frac{1}{3}\frac{\Omega_a}{\Omega_m} S_a. \tag{D5}$$

We add an extra $-1/15$ onto the prefactor to take the Sachs-Wolfe effect into account [68], so the total temperature fluctuation is

$$\frac{\delta T}{T}_{\rm iso} \simeq -\frac{6}{15}\frac{\Omega_a}{\Omega_m} S_a. \tag{D6}$$

Now we find an expression for S_a in terms of θ_i, f and H_I. The fractional fluctuation in the axion density is related to fluctuations around the initial misalignment angle θ_i in the early Universe

$$\begin{aligned} S_a &\simeq \frac{\delta n_a}{n_a} \simeq \frac{\delta(\theta^2)}{\langle\theta^2\rangle} \simeq \frac{\langle(\theta\rangle + \delta\theta)^2 - \langle\theta^2\rangle}{\langle\theta^2\rangle} \\ &\simeq \frac{2\theta_i\delta\theta + (\delta\theta)^2 - \sigma_\theta^2}{\theta_i^2 + \sigma_\theta^2} \end{aligned} \tag{D7}$$

where the variance is $\sigma_\theta^2 = \langle(\theta - \langle\theta\rangle)^2\rangle = H_I^2/(2\pi f)^2$ and $\langle\theta\rangle = \theta_i$. We will also need

$$\langle S_a^2\rangle = 2\sigma_\theta^2 \frac{2\theta_i^2 + \sigma_\theta^2}{(\theta_i^2 + \sigma_\theta^2)^2}. \tag{D8}$$

The isocurvature component of the total power in CMB temperature fluctuations is

$$\alpha = \frac{\langle(\delta T/T)^2_{\rm iso}\rangle}{\langle(\delta T/T)^2_{\rm tot}\rangle}, \tag{D9}$$

where COBE measured $\langle(\delta T/T)^2_{\rm tot}\rangle^{1/2} = 1.1 \times 10^{-5}$ [119]. Putting the pieces together we have

$$\alpha = \left(\frac{6}{15}\right)^2 \frac{(\Omega_a/\Omega_m)^2}{\langle(\delta T/T)^2_{\rm tot}\rangle} 2\sigma_\theta^2 \frac{2\theta_i^2 + \sigma_\theta^2}{(\theta_i^2 + \sigma_\theta^2)^2} < 0.038. \tag{D10}$$

where $\Omega_m \simeq 0.13$, the expression for the axion relic density in EMD theories is Eq. (6) and the rest is defined above.

As an example, for benchmark values of $\theta_i = 1$, $f_a = 9 \times 10^{14}$ GeV and $T_{\rm RH} = 10$ MeV Eq. (D10) implies $H_I < 2 \times 10^{10}$ GeV. The requirements on the scale of inflation in EMD theories are less onerous than standard cosmology by a factor of $\mathcal{O}(10^{2-3})$. As we raise the reheating temperature the bound on the H_I decreases: for $T_{\rm RH} = 500$ MeV we have $H_I < 2 \times 10^9$ GeV. The isocurvature bounds can be evaded in low-scale theories of inflation; see, e.g., Refs. [120,121].

APPENDIX E: RELATIONSHIPS BETWEEN SIZE AND MASS DEFINITIONS

In this section we relate the various mass and size scales used to characterize minihalos. We will assume that minihalos have an NFW profile:

$$\rho(r) = \frac{4\rho_s}{(r/r_s)(1 + r/r_s)^2}, \tag{E1}$$

where ρ_s and r_s are scale density and scale radius. The mass within a certain radius r is then

$$M(r) = 16\pi\rho_s r_s^3 f(r/r_s), \tag{E2}$$

where

$$f(c) = \ln(c+1) - \frac{c}{c+1}. \tag{E3}$$

The scale mass M_s is the mass within the scale radius:

$$M_s = M(r_s) = 16\pi\rho_s r_s^3 f(1). \tag{E4}$$

Given any two of (r_s, ρ_s, M_s) we can solve for the other one using this relationship.

1. Virial mass, radius and concentration

The virial quantities are defined for a sphere centered on the gravitational potential minimum that encloses a region within which the average density is $\Delta = 200$ times the critical density at some redshift $\rho_c(z) = 3H^2(z)/(8\pi G)$ [113,122] (sometimes only the average matter density $\Omega_m\rho_c(z)$ is used [97]). We therefore have the following relationship between $M_{\rm vir}$ and $r_{\rm vir}$:

$$M_{\rm vir} = \frac{4\pi}{3} 200\rho_c r_{\rm vir}^3. \tag{E5}$$

We want to relate virial quantities to the NFW scale parameters defined above. First,

$$M_{\rm vir} = M(r_{\rm vir}) = M_s f(c_{200})/f(1), \tag{E6}$$

where we defined the concentration parameter

$$c_{200} = r_{\rm vir}/r_s. \tag{E7}$$

We can take $(c_{200}, M_{\rm vir})$ as defining the halo and solve for the NFW scale quantities using the relations above, leading to

$$M_s = \frac{f(1)}{f(c_{200})} M_{\rm vir}, \tag{E8}$$

$$r_s = r_{\rm vir}/c_{200} \tag{E9}$$

and

$$\rho_s = \frac{1}{12} 200 \rho_c \frac{c_{200}^3}{f(c_{200})}. \tag{E10}$$

Note that at large concentration parameters $c_{200} \gg 1$, $f(c_{200}) \sim \ln c_{200}/e$ and Eq. (E8) implies that

$$\frac{M_{\rm vir}}{M_s} \sim 43 + 5 \ln\left(\frac{c_{200}}{10^4}\right), \tag{E11}$$

so the virial mass and scale mass can be quite different. This occurs if the redshift at which the virial quantities are calculated is long after the initial collapse that created the gravitationally bound core; subsequent evolution resulted in accretion of matter on this core.

2. M_*, R_* and concentration at formation

Recall that R_* is defined as the radius within which the density is $178\rho_a(z_c)$, where $\rho_a(z_c)$ is the background density at collapse. This means that M_* and R_* are like the virial quantities, evaluated at collapse, i.e., $M_{\rm vir}(z_c) \approx M_*$ and $r_{\rm vir}(z_c) \approx R_*$. Let c_* be the concentration parameter at formation, i.e.,

$$c_* = R_*/r_s. \tag{E12}$$

This is an $\mathcal{O}(1)$ number. Simulations of Earth-mass ΛCDM halos imply that $c_* \approx 2$ [113] for $M_s \sim 10^{-6}\, M_\odot$, but in principle this is a cosmology and mass-dependent quantity. The concentration parameter allows us to find M_s and ρ_s from c_*, R_* and M_*:

$$M_s = \frac{f(1)}{f(c_*)} M_* \tag{E13}$$

The scale density is then obtained from Eq. (E4).

[1] L. F. Abbott and P. Sikivie, A cosmological bound on the invisible axion, Phys. Lett. **120B**, 133 (1983).
[2] M. Dine and W. Fischler, The not so harmless axion, Phys. Lett. **120B**, 137 (1983).
[3] J. Preskill, M. B. Wise, and F. Wilczek, Cosmology of the invisible axion, Phys. Lett. **120B**, 127 (1983).
[4] P. Arias, D. Cadamuro, M. Goodsell, J. Jaeckel, J. Redondo, and A. Ringwald, WISPy cold dark matter, J. Cosmol. Astropart. Phys. 06 (2012) 013.
[5] P. Svrcek and E. Witten, Axions in string theory, J. High Energy Phys. 06 (2006) 051.
[6] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, String axiverse, Phys. Rev. D **81**, 123530 (2010).
[7] M. Cicoli, M. Goodsell, and A. Ringwald, The type IIB string axiverse and its low-energy phenomenology, J. High Energy Phys. 10 (2012) 146.
[8] P. Sikivie, Axion cosmology, Lect. Notes Phys. **741**, 19 (2008).
[9] D. J. E. Marsh, Axion cosmology, Phys. Rep. **643**, 1 (2016).
[10] N. Blinov, M. J. Dolan, P. Draper, and J. Kozaczuk, Dark matter targets for axionlike particle searches, Phys. Rev. D **100**, 015049 (2019).
[11] S. Chaudhuri, P. W. Graham, K. Irwin, J. Mardon, S. Rajendran, and Y. Zhao, Radio for hidden-photon dark matter detection, Phys. Rev. D **92**, 075012 (2015).
[12] M. Silva-Feaver *et al.*, Design overview of DM radio pathfinder experiment, IEEE Trans. Appl. Supercond. **27**, 1400204 (2017).
[13] Y. Kahn, B. R. Safdi, and J. Thaler, Broadband and Resonant Approaches to Axion Dark Matter Detection, Phys. Rev. Lett. **117**, 141801 (2016).
[14] J. L. Ouellet *et al.*, First Results from ABRACADABRA-10 cm: A Search for Sub-μeV Axion Dark Matter, Phys. Rev. Lett. **122**, 121802 (2019).
[15] R. Henning *et al.* (ABRACADABRA Collaboration), ABRACADABRA, A search for low-mass axion dark matter, in *Proceedings, 13th Patras Workshop on Axions, WIMPs and WISPs, (PATRAS 2017): Thessaloniki, Greece, 2017* (Verlag Deutsches Elektronen-Synchrotron, Hamburg, Germany, 2018), pp. 28–31.
[16] P. W. Graham and S. Rajendran, New observables for direct detection of axion dark matter, Phys. Rev. D **88**, 035023 (2013).
[17] D. Budker, P. W. Graham, M. Ledbetter, S. Rajendran, and A. Sushkov, Proposal for a Cosmic Axion Spin Precession Experiment (CASPEr), Phys. Rev. X **4**, 021030 (2014).
[18] D. F. J. Kimball *et al.*, Overview of the cosmic axion spin precession experiment (CASPEr), arXiv:1711.08999.
[19] D. Malyshev, A. Neronov, D. Semikoz, A. Santangelo, and J. Jochum, Improved limit on axion-like particles from γ-ray data on Perseus cluster, arXiv:1805.04388.
[20] T. Banks, M. Dine, and M. Graesser, Supersymmetry, axions and cosmology, Phys. Rev. D **68**, 075011 (2003).
[21] J. Polchinski, *String Theory. Vol. 1: An Introduction to the Bosonic String*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2007).
[22] J. Polchinski, *String Theory. Vol. 2: Superstring Theory and Beyond*, Cambridge Monographs on Mathematical

Physics (Cambridge University Press, Cambridge, England, 2007).
[23] M. R. Douglas and S. Kachru, Flux compactification, Rev. Mod. Phys. **79**, 733 (2007).
[24] B. S. Acharya, G. Kane, and P. Kumar, Compactified string theories: Generic predictions for particle physics, Int. J. Mod. Phys. A **27**, 1230012 (2012).
[25] G. D. Coughlan, W. Fischler, E. W. Kolb, S. Raby, and G. G. Ross, Cosmological problems for the Polonyi potential, Phys. Lett. **131B**, 59 (1983).
[26] T. Banks, D. B. Kaplan, and A. E. Nelson, Cosmological implications of dynamical supersymmetry breaking, Phys. Rev. D **49**, 779 (1994).
[27] B. de Carlos, J. A. Casas, F. Quevedo, and E. Roulet, Model independent properties and cosmological implications of the dilaton and moduli sectors of 4-d strings, Phys. Lett. B **318**, 447 (1993).
[28] M. Kawasaki, K. Kohri, and N. Sugiyama, MeV scale reheating temperature and thermalization of neutrino background, Phys. Rev. D **62**, 023506 (2000).
[29] S. Hannestad, What is the lowest possible reheating temperature?, Phys. Rev. D **70**, 043506 (2004).
[30] P. F. de Salas, M. Lattanzi, G. Mangano, G. Miele, S. Pastor, and O. Pisanti, Bounds on very low reheating scenarios after Planck, Phys. Rev. D **92**, 123534 (2015).
[31] T. Hasegawa, N. Hiroshima, K. Kohri, R. S. L. Hansen, T. Tram, and S. Hannestad, MeV-scale reheating temperature and thermalization of oscillating neutrinos by radiative and hadronic decays of massive particles, J. Cosmol. Astropart. Phys. 12 (2019) 012.
[32] T. Moroi and L. Randall, Wino cold dark matter from anomaly mediated SUSY breaking, Nucl. Phys. **B570**, 455 (2000).
[33] B. S. Acharya, P. Kumar, K. Bobkov, G. Kane, J. Shao, and S. Watson, Non-thermal dark matter and the moduli problem in string frameworks, J. High Energy Phys. 06 (2008) 064.
[34] M. Bose, M. Dine, and P. Draper, Moduli or not, Phys. Rev. D **88**, 023533 (2013).
[35] J. Fan and M. Reece, In wino veritas? Indirect searches shed light on neutralino dark matter, J. High Energy Phys. 10 (2013) 124.
[36] N. Blinov, J. Kozaczuk, A. Menon, and D. E. Morrissey, Confronting the moduli-induced lightest-superpartner problem, Phys. Rev. D **91**, 035026 (2015).
[37] T. Banks and M. Dine, The cosmology of string theoretic axions, Nucl. Phys. **B505**, 445 (1997).
[38] L. Visinelli and P. Gondolo, Axion cold dark matter in nonstandard cosmologies, Phys. Rev. D **81**, 063508 (2010).
[39] A. E. Nelson and H. Xiao, Axion cosmology with early matter domination, Phys. Rev. D **98**, 063516 (2018).
[40] L. Visinelli and J. Redondo, Axion miniclusters in modified cosmological histories, Phys. Rev. D **101**, 023008 (2020).
[41] N. Ramberg and L. Visinelli, Probing the early Universe with axion physics and gravitational waves, Phys. Rev. D **99**, 123513 (2019).
[42] G. F. Giudice, E. W. Kolb, and A. Riotto, Largest temperature of the radiation era and its cosmological implications, Phys. Rev. D **64**, 023508 (2001).
[43] A. L. Erickcek and K. Sigurdson, Reheating effects in the matter power spectrum and implications for substructure, Phys. Rev. D **84**, 083503 (2011).
[44] G. Barenboim and J. Rasero, Structure formation during an early period of matter domination, J. High Energy Phys. 04 (2014) 138.
[45] J. Fan, O. Özsoy, and S. Watson, Nonthermal histories and implications for structure formation, Phys. Rev. D **90**, 043536 (2014).
[46] A. L. Erickcek, K. Sinha, and S. Watson, Bringing isolated dark matter out of isolation: Late-time reheating and indirect detection, Phys. Rev. D **94**, 063502 (2016).
[47] C. J. Hogan and M. J. Rees, Axion miniclusters, Phys. Lett. B **205**, 228 (1988).
[48] E. W. Kolb and I. I. Tkachev, Nonlinear axion dynamics and formation of cosmological pseudosolitons, Phys. Rev. D **49**, 5040 (1994).
[49] E. W. Kolb and I. I. Tkachev, Axion Miniclusters and Bose Stars, Phys. Rev. Lett. **71**, 3051 (1993).
[50] E. W. Kolb and I. I. Tkachev, Large amplitude isothermal fluctuations and high density dark matter clumps, Phys. Rev. D **50**, 769 (1994).
[51] J. Enander, A. Pargner, and T. Schwetz, Axion minicluster power spectrum and mass function, J. Cosmol. Astropart. Phys. 12 (2017) 038.
[52] M. Fairbairn, D. J. E. Marsh, J. Quevillon, and S. Rozier, Structure formation and microlensing with axion miniclusters, Phys. Rev. D **97**, 083502 (2018).
[53] P. Tinyakov, I. Tkachev, and K. Zioutas, Tidal streams from axion miniclusters and direct axion searches, J. Cosmol. Astropart. Phys. 01 (2016) 035.
[54] V. S. Berezinsky, V. I. Dokuchaev, and Yu. N. Eroshenko, Small-scale clumps of dark matter, Usp. Fiz. Nauk **184**, 3 (2014) [Phys. Usp. **57**, 1 (2014)].
[55] E. W. Kolb and I. I. Tkachev, Femtolensing and picolensing by axion miniclusters, Astrophys. J. **460**, L25 (1996).
[56] K. Van Tilburg, A.-M. Taki, and N. Weiner, Halometry from astrometry, J. Cosmol. Astropart. Phys. 07 (2018) 041.
[57] L. Dai and J. Miralda-Escudé, Gravitational lensing signatures of axion dark matter minihalos in highly magnified Stars, Aston. J. **159**, 49 (2020).
[58] J. A. Dror, H. Ramani, T. Trickle, and K. M. Zurek, Pulsar timing probes of primordial black holes and Subhalos, Phys. Rev. D **100**, 023003 (2019).
[59] M. Buschmann, J. W. Foster, and B. R. Safdi, Early-universe simulations of the cosmological axion, arXiv: 1906.00967.
[60] A. Vaquero, J. Redondo, and J. Stadler, Early seeds of axion miniclusters, J. Cosmol. Astropart. Phys. 04 (2019) 012.
[61] J. Ollé, O. Pujolàs, and F. Rompineve, Oscillons and dark matter, arXiv:1906.06352.
[62] A. Arvanitaki, S. Dimopoulos, M. Galanis, L. Lehner, J. O. Thompson, and K. Van Tilburg, The large-misalignment mechanism for the formation of compact axion structures: Signatures from the QCD axion to fuzzy dark matter, arXiv:1909.11665.

[63] E. W. Kolb and M. S. Turner, The early universe, Front. Phys. **69**, 1 (1990).
[64] M. Fairbairn, D. J. E. Marsh, and J. Quevillon, Searching for the QCD Axion with Gravitational Microlensing, Phys. Rev. Lett. **119**, 021101 (2017).
[65] M. Axenides, R. H. Brandenberger, and M. S. Turner, Development of axion perturbations in an axion dominated universe, Phys. Lett. **126B**, 178 (1983).
[66] D. Seckel and M. S. Turner, Isothermal density perturbations in an axion dominated inflationary universe, Phys. Rev. D **32**, 3178 (1985).
[67] G. Efstathiou and J. R. Bond, Isocurvature cold dark matter fluctuations, Mon. Not. R. Astron. Soc. **218**, 103 (1986).
[68] P. Fox, A. Pierce, and S. D. Thomas, Probing a QCD string axion with precision cosmological measurements, arXiv:hep-th/0409059.
[69] C. Pitrou, A. Coc, J.-P. Uzan, and E. Vangioni, Precision big bang nucleosynthesis with improved helium-4 predictions, Phys. Rep. **754**, 1 (2018).
[70] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, arXiv:1807.06209.
[71] N. Blinov, K. J. Kelly, G. Z. Krnjaic, and S. D. McDermott, Constraining the Self-Interacting Neutrino Interpretation of the Hubble Tension, Phys. Rev. Lett. **123**, 191102 (2019).
[72] V. B. Klaer and G. D. Moore, The dark-matter axion mass, J. Cosmol. Astropart. Phys. 11 (2017) 049.
[73] M. Gorghetto, E. Hardy, and G. Villadoro, Axions from strings: The attractive solution, J. High Energy Phys. 07 (2018) 151.
[74] M. Kawasaki, T. Sekiguchi, M. Yamaguchi, and J. Yokoyama, Long-term dynamics of cosmological axion strings, Prog. Theor. Exp. Phys. **2018**, 091E01 (2018).
[75] P. Draper, J. Kozaczuk, and J.-H. Yu, Theta in new QCD-like sectors, Phys. Rev. D **98**, 015028 (2018).
[76] D. J. H. Chung, E. W. Kolb, and A. Riotto, Production of massive particles during reheating, Phys. Rev. D **60**, 063504 (1999).
[77] W. Hu, R. Barkana, and A. Gruzinov, Cold and Fuzzy Dark Matter, Phys. Rev. Lett. **85**, 1158 (2000).
[78] J.-C. Hwang and H. Noh, Axion as a cold dark matter candidate, Phys. Lett. B **680**, 1 (2009).
[79] W. Hu, Covariant linear perturbation formalism, ICTP Lect. Notes Ser. **14**, 145 (2003).
[80] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2015 results. XIII. Cosmological parameters, Astron. Astrophys. **594**, A13 (2016).
[81] W. Hu and N. Sugiyama, Small scale cosmological perturbations: An Analytic approach, Astrophys. J. **471**, 542 (1996).
[82] R. J. Scherrer and M. S. Turner, Decaying particles do not "heat up" the Universe, Phys. Rev. D **31**, 681 (1985).
[83] R. Hlozek, D. Grin, D. J. E. Marsh, and P. G. Ferreira, A search for ultralight axions using precision cosmological data, Phys. Rev. D **91**, 103512 (2015).
[84] A. Lewis, A. Challinor, and A. Lasenby, Efficient computation of CMB anisotropies in closed FRW models, Astrophys. J. **538**, 473 (2000).
[85] D. Blas, J. Lesgourgues, and T. Tram, The cosmic linear anisotropy solving system (CLASS) II: Approximation schemes, J. Cosmol. Astropart. Phys. 07 (2011) 034.
[86] J. Diemand, B. Moore, and J. Stadel, Earth-mass dark-matter haloes as the first structures in the early universe, Nature (London) **433**, 389 (2005).
[87] T. Bringmann, Particle models and the small-scale structure of dark matter, New J. Phys. **11**, 105027 (2009).
[88] W. H. Press and P. Schechter, Formation of galaxies and clusters of galaxies by selfsimilar gravitational condensation, Astrophys. J. **187**, 425 (1974).
[89] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser, Excursion set mass functions for hierarchical Gaussian fluctuations, Astrophys. J. **379**, 440 (1991).
[90] H. Mo, F. C. van den Bosch, and S. White, *Galaxy Formation and Evolution* (Cambridge University Press, Cambridge, England, 2010).
[91] C. Blanco, M. S. Delos, A. L. Erickcek, and D. Hooper, Annihilation signatures of hidden sector dark matter within early-forming microhalos, Phys. Rev. D **100**, 103010 (2019).
[92] A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida, The mass function of dark matter halos, Mon. Not. R. Astron. Soc. **321**, 372 (2001).
[93] R. K. Sheth, H. J. Mo, and G. Tormen, Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes, Mon. Not. R. Astron. Soc. **323**, 1 (2001).
[94] R. K. Sheth and G. Tormen, An excursion set model of hierarchical clustering: Ellipsoidal collapse and the moving barrier, Mon. Not. R. Astron. Soc. **329**, 61 (2002).
[95] C. G. Lacey and S. Cole, Merger rates in hierarchical models of galaxy formation, Mon. Not. R. Astron. Soc. **262**, 627 (1993).
[96] A. R. Zentner, The excursion set theory of halo mass functions, halo clustering, and halo growth, Int. J. Mod. Phys. D **16**, 763 (2007).
[97] T. Goerdt, O. Y. Gnedin, B. Moore, J. Diemand, and J. Stadel, The survival and disruption of CDM micro-haloes: Implications for direct and indirect detection experiments, Mon. Not. R. Astron. Soc. **375**, 191 (2007).
[98] S. M. Koushiappas, The detection of sub-solar mass dark matter halos, New J. Phys. **11**, 105012 (2009).
[99] M. S. Delos, A. L. Erickcek, A. P. Bailey, and M. A. Alvarez, Density profiles of ultracompact minihalos: Implications for constraining the primordial power spectrum, Phys. Rev. D **98**, 063527 (2018).
[100] M. S. Delos, A. L. Erickcek, A. P. Bailey, and M. A. Alvarez, Are ultracompact minihalos really ultracompact?, Phys. Rev. D **97**, 041303 (2018).
[101] F. C. van den Bosch, G. Ogiya, O. Hahn, and A. Burkert, Disruption of dark matter substructure: Fact or fiction?, Mon. Not. R. Astron. Soc. **474**, 3043 (2018).
[102] V. I. Dokuchaev, Yu. N. Eroshenko, and I. I. Tkachev, Destruction of axion miniclusters in the Galaxy, Zh. Eksp. Teor. Fiz. **152**, 511 (2017) [J. Exp. Theor. Phys. **125**, 434 (2017)].
[103] F. C. van den Bosch and G. Ogiya, Dark matter substructure in numerical simulations: A tale of discreteness noise,

runaway instabilities, and artificial disruption, Mon. Not. R. Astron. Soc. **475**, 4066 (2018).
[104] M. S. Delos, Evolution of dark matter microhalos through stellar encounters, Phys. Rev. D **100**, 083529 (2019).
[105] A. Schneider, L. Krauss, and B. Moore, Impact of dark matter microhalos on signatures for direct and indirect detection, Phys. Rev. D **82**, 063525 (2010).
[106] C. A. J. O'Hare and A. M. Green, Axion astronomy with microwave cavity experiments, Phys. Rev. D **95**, 063017 (2017).
[107] P. Tisserand *et al.* (EROS-2 Collaboration), Limits on the macho content of the Galactic halo from the EROS-2 survey of the magellanic clouds, Astron. Astrophys. **469**, 387 (2007).
[108] K. Griest, A. M. Cieplak, and M. J. Lehner, Experimental limits on primordial black hole dark matter from the first 2 yr of Kepler data, Astrophys. J. **786**, 158 (2014).
[109] H. Niikura *et al.*, Microlensing constraints on primordial black holes with Subaru/HSC Andromeda observations, Nat. Astron. **3**, 524 (2019).
[110] J. Miralda-Escude, The magnification of stars crossing a caustic. I—Lenses with smooth potentials, Astrophys. J. **379**, 94 (1991).
[111] P. L. Kelly *et al.*, Extreme magnification of an individual star at redshift 1.5 by a galaxy-cluster lens, Nat. Astron. **2**, 334 (2018).
[112] W. Chen *et al.*, Searching for highly magnified stars at cosmological distances: Discovery of a redshift 0.94 blue supergiant in archival images of the Galaxy Cluster MACS J0416.1-2403, Astrophys. J. **881**, 8 (2019).
[113] M. A. Sánchez-Conde and F. Prada, The flattening of the concentrationmass relation towards low halo masses and its implications for the annihilation signal boost, Mon. Not. R. Astron. Soc. **442**, 2271 (2014).
[114] K. Redmond, A. Trezza, and A. L. Erickcek, Growth of dark matter perturbations during kination, Phys. Rev. D **98**, 063504 (2018).
[115] W. Hu, Structure formation with generalized dark matter, Astrophys. J. **506**, 485 (1998).
[116] J. Lesgourgues, The cosmic linear anisotropy solving system (CLASS) I: Overview, arXiv:1104.2932.
[117] D. J. Eisenstein and W. Hu, Baryonic features in the matter transfer function, Astrophys. J. **496**, 605 (1998).
[118] Y. Akrami *et al.* (Planck Collaboration), Planck 2018 results. X. Constraints on inflation, Astrophys. Space Sci. **364**, 69 (2019).
[119] C. L. Bennett, A. Banday, K. M. Gorski, G. Hinshaw, P. Jackson, P. Keegstra, A. Kogut, G. F. Smoot, D. T. Wilkinson, and E. L. Wright, Four year COBE DMR cosmic microwave background observations: Maps and basic results, Astrophys. J. **464**, L1 (1996).
[120] G. German, G. G. Ross, and S. Sarkar, Low scale inflation, Nucl. Phys. **B608**, 423 (2001).
[121] J. Martin, C. Ringeval, and V. Vennin, Encyclopaedia inflationaris, Phys. Dark Universe **5–6**, 75 (2014).
[122] A. D. Ludlow, J. F. Navarro, R. E. Angulo, M. Boylan-Kolchin, V. Springel, C. Frenk, and S. D. M. White, The massconcentrationredshift relation of cold dark matter haloes, Mon. Not. R. Astron. Soc. **441**, 378 (2014).