Cosmological limits on strong dark forces

Peter W. Graham,[1,2] Harikrishnan Ramani,[3] Olivier Simon,[4,5] and Erwin H. Tanin[1,*]

[1]*Leinweber Institute for Theoretical Physics, Department of Physics, Stanford University, Stanford, California 94305, USA*

[2]*Kavli Institute for Particle Astrophysics and Cosmology, Department of Physics, Stanford University, Stanford, California 94305, USA*

[3]*Department of Physics and Astronomy, University of Delaware and the Bartol Research Institute, Newark, Delaware 19716, USA*

[4]*Princeton Center for Theoretical Science, Princeton University, Princeton, New Jersey 08544, USA*

[5]*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*

(Received 13 May 2026; accepted 21 August 2026; published 23 September 2026)

We showcase cosmology's ability to constrain long-range forces between dark matter particles. Specifically, we consider a fermionic dark matter interacting via a Yukawa-coupled light scalar, focusing on regimes where the dark forces are stronger than gravitational and yet unconstrained. We show that the dark sector dynamics, both at the background and perturbation levels, is far richer than what can be captured with just the static interparticle Yukawa potential. The background dynamics includes an attractor that funnels a wide range of initial conditions onto an evolution unique to each parameter space. In a large swath of parameter space beyond existing limits, the dark sector deviates drastically from cold dark matter in observable epochs. We rule out this parameter space using existing constraints on dark-sector equation of state and small-scale cosmic perturbations, thus setting the strongest constraints yet on dark matter self-interactions at length scales shorter than 100 kpc. In addition, we briefly discuss repulsive dark forces and place cosmological limits that are stricter than in the attractive case.

DOI: 10.1103/pkkh-t861

I. INTRODUCTION

Decades of dark matter (DM) searches have placed increasingly stringent bounds on its possible interactions with the Standard Model (SM). We are gradually approaching the sobering possibility that any nongravitational DM-SM coupling, if it exists, is too feeble to be detected by current or near-future experiments. Even in this "nightmare" scenario, the dark sector may still exhibit rich internal dynamics. Such dynamics would manifest as deviations from the ΛCDM paradigm, most notably through imprints on small-scale structure or via deviations from the well measured DM equation of state. Examples of things learned about the dark sector from such probes include lower limits on the dark matter particle mass [1], the earliest epoch by which dark matter needed to be cold [2], dark matter lifetime [3], stringent limits on its primordial power spectrum [4], postinflationary Peccei-Quinn breaking scenarios [5], the possible presence of strong self-interactions at small scales [6], as well as tight constraints [7] at these scales to name a few.

A particularly intriguing nontriviality is that dark matter experiences a long-range "fifth force." These forces are typically mediated by an ultralight scalar field, motivated in scenarios such as models with extra dimensions [8] or as moduli arising naturally in string theory [9,10]. Fifth forces between SM particles have been extensively searched for, with experimental constraints requiring them to be weaker than gravity down to the $\mathcal{O}(10\ \mu\text{m})$ scale. In general, such forces can violate the equivalence principle, and in the extreme case, may be confined entirely within the dark sector rendering them invisible to conventional probes.

Astrophysical considerations have constrained the strength of long-range forces between DM particles (dark forces) with ranges longer than about 1 Mpc to be weaker than gravity [11]; see also [12–16]. However, in the sub-Mpc range, dark forces with strengths orders of magnitude stronger than gravity are thus far still allowed [11]. Such dark forces could make drastic changes to the early Universe, when it was much smaller and denser. On the other hand, consistency with a wealth of cosmological data necessitates that the DM behaves like cold dark matter (CDM) beginning at redshifts of $z \sim 10^5$, whereupon the

*Contact author: ehtanin@stanford.edu

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

Universe was about 100 pc in size. Thus, the early Universe could in principle probe dark forces as short as 100 pc in range at near-gravitational strengths.

There have been efforts to constrain long-range forces between dark matter particles through their cosmological consequences [17–25]. In early studies [17–21], the interaction between two DM particles with mass m_χ separated by a distance r is simply modeled as a static Yukawa potential

$$|V_{\rm Yuk}| = \frac{\alpha_{\chi\chi} G m_\chi^2}{r} e^{-r/\lambda_\phi}, \tag{1}$$

where λ_ϕ and $\alpha_{\chi\chi}$ parametrize, respectively, the range of the dark force and its strength relative to gravity. A particle-physics realization of such a force requires a light field serving as a mediator. The potential in Eq. (1) corresponds to the static solution of the mediator field, analogous to electrostatics. While this approximation may be appropriate in static systems such as those of small-scale tests of gravity, it may not adequately describe the mediator field in the early Universe, which typically has its own rich dynamics that is not completely dictated by the DM configuration at an instance [26,27].

More recently, Refs. [22–25] carefully considered the dynamical nature of the mediator field in placing limits on dark forces. Their findings ruled out dark forces with ranges $\lambda_\phi \gtrsim 100$ kpc and strengths relative to gravity $\alpha_{\chi\chi} \gtrsim 10^{-3}$–$10^{-2}$, thereby placing the most stringent constraint in this regime of the parameter space. Note, however, that these analyses are based on a perturbative expansion in the coupling $\alpha_{\chi\chi}$ and are therefore restricted, for forces in those ranges, to dark forces with strengths at most comparable to that of gravity, $\alpha_{\chi\chi} \lesssim 1$.

Our aim in this paper is to determine cosmological constraints on long-range dark forces with coupling strengths larger than gravitational, beyond the current reach of astrophysical observations and small-coupling analyses of cosmology. We will thus focus mainly on new forces with range $\lambda_\phi \lesssim 100$ kpc and $\alpha_{\chi\chi} \gtrsim 1$. While a detailed likelihood analysis is crucial to disentangle the effects of dark forces with small-coupling from that of other cosmological parameters, large-coupling dark forces may lead to drastic modifications to the standard cosmology well beyond what is allowed by the inherent degeneracies of cosmological parameters. We focus on the following striking effects of dark forces. First, the interplay between the DM and the mediator field generically leads to a dynamical dark sector equation of state that, at times, deviates far from that of CDM. Second, long-range dark forces could cause orders of magnitude of extra growth in observable cosmological perturbation modes. The latter is an effect known to occur when the quasistatic exchange interaction between dark matter particles, Eq. (1), dominates dark sector dynamics [27–29], but, as we will discuss, this condition is often not realized in the early Universe. Even so, in some cases, a catastrophic amount of growth can take place during the limited period of applicability of Eq. (1). We leverage these phenomena to place conservative cosmological limits on dark forces.

The paper is organized as follows. In Sec. II, we describe the Yukawa dark-force model and specify the assumptions that enter our analyses. In Sec. III, we discuss the impact of dark forces on background cosmology and how it could be probed. In Sec. IV, we consider the enhanced growth of cosmic perturbations in the presence of dark forces and how it could be probed. In Sec. V, we detail our main results. In Sec. VI, we briefly discuss repulsive dark forces. Finally, we conclude in Sec. VII.

II. ASSUMPTIONS

A. Model

The Yukawa parametrization of dark forces, Eq. (1), enables a model-independent treatment of static dark forces. However, certain considerations, such as the early-Universe dynamics of the mediator field, are inherently model dependent. For this work, we adopt one of the simplest models of long-range forces, namely a fermion χ of mass m_χ serving as the primary DM with a Yukawa coupling g to a light scalar mediator ϕ of mass m_ϕ, as described by the following Lagrangian [17]:

$$\mathcal{L}_{\rm dark} = \frac{1}{2}(\partial\phi)^2 + \bar{\chi} i\partial\!\!\!/ \chi - (m_\chi - g\phi)\bar{\chi}\chi - \frac{1}{2}m_\phi^2\phi^2.$$

In the main part of our analysis, the dark coupling g and the fermion bare mass m_χ enter only through the combination g/m_χ. As long as the mass m_χ is limited to certain (g, m_ϕ)-dependent range, its exact value is not important for our main analysis; see Appendix D. Thus, it is convenient to express g in terms of the strength of the dark long-range force relative to that of gravity

$$\alpha_{\chi\chi} \equiv \frac{g^2}{4\pi G m_\chi^2}, \tag{2}$$

as in Eq. (1), where in this model the force range is given by the inverse mass of the mediator $\lambda_\phi = m_\phi^{-1}$. Throughout the paper, we will assume that the dark sector is completely decoupled from the Standard Model sector.

Note that for a *scalar* force mediator, the two-body static exchange force is universally *attractive* between particles and antiparticles of the same kind. This is in contrast to static forces mediated by a vector, such as the electrostatic force in the SM, which are repulsive between particle-particle and antiparticle-antiparticle pairs, but attractive between particle-antiparticle pairs. We briefly consider repulsive dark forces in Sec. VI.

Throughout our main analysis, we neglect the scalar field's quartic coupling, requiring it be smaller than a

certain threshold (see Appendix D) such that it can be safely ignored. We note that both the bare scalar masses m_ϕ and self-interactions considered in this paper are fine-tuned in the sense that their values are smaller than the expected loop-contributions to them. See Appendix E for more details. However, there are theoretical realizations [30–32] that ameliorate such fine-tuning and make the scalar generically light.

B. Initial condition

We begin our analysis at some scale factor a_i deep in radiation domination (RD), e.g., at $a_i = 10^{-9}$, early enough that the exact value of a_i is unimportant. In the simplest case, we start the ϕ background field at the value $\phi_i = 0$ and populate χ completely asymmetrically, with only particles and no antiparticles. We also assume that the quanta of the ϕ and χ fields are sufficiently cold that their temperatures do not affect our analysis. We consider the effects of finite χ density and finite χ temperature on the background dynamics of ϕ in Appendix D. We find that avoiding the regimes where these effects are important amounts to mild restrictions on the fermion mass m_χ.

Nevertheless, thanks to both the existence of a dynamical attractor in the ϕ field's background evolution and the insensitivity of the backreaction of a nonrelativistic χ to its precise momentum distribution, a broad range of initial conditions other than the simplest one, as mentioned in the previous paragraph, converge to virtually the same dynamics at sufficiently late times, relevant for cosmological observations. We describe these further in Appendix A and D. In all the cases we consider, the number density of χ is covariantly conserved, $n_\chi \propto a^{-3}$. Thus, we can write its initial value as

$$n_{\chi,i} = \frac{f_\chi \rho_{\mathrm{m},0}}{m_\chi}\left(\frac{a_i}{a_0}\right)^{-3}. \tag{3}$$

Here, the subscripts $_i$ and $_0$ refer to the initial and present-day values, and $\rho_{\mathrm{m},0} \approx 12.5\ \mathrm{meV}^4$ is the fiducial cosmic matter density today [33]. For the most part, we will assume that $f_\chi = f_{\mathrm{DM}} = 0.85$, corresponding to the current DM mass density of $\rho_{\mathrm{DM},0} = f_{\mathrm{DM}}\rho_{\mathrm{m},0} = 10.6\ \mathrm{meV}^4$. An exception to this occurs in regimes that predict the ϕ field behaves like an ultralight CDM component and contributes significantly to the DM abundance at the present epoch, in which case the f_χ should be appropriately calibrated such that the total mass density of ϕ and χ (not just χ's) equals $\rho_{\mathrm{DM},0}$ at the present epoch; we will return to this subtlety in Sec. V A 2.

III. BACKGROUND COSMOLOGY

A. Effective potential

In the presence of a background ϕ, the χ particles receives a mass correction such that its effective mass becomes[1]

$$M_\chi(\phi) = |m_\chi - g\phi|. \tag{4}$$

As long as χ is nonrelativistic, we can approximate $(m_\chi - g\phi)\langle\bar{\chi}\chi\rangle \approx M_\chi(\phi) n_\chi$.[2] The effective potential of ϕ is then a sum of a quadratic piece and a linear piece whose slope changes sign at $\phi = m_\chi/g$

$$\begin{aligned} V_{\mathrm{eff}}(\phi) &= \frac{1}{2}m_\phi^2\phi^2 + (m_\chi - g\phi)\langle\bar{\chi}\chi\rangle \\ &\approx \frac{1}{2}m_\phi^2\phi^2 + m_\phi^2\phi_*|\phi_0 - \phi|. \end{aligned} \tag{5}$$

Here, we have defined the following finite-density minima:

$$\phi_0 \equiv \frac{m_\chi}{g} \qquad \text{(linear minimum)}, \tag{6}$$

$$\phi_* \equiv \frac{g n_\chi}{m_\phi^2} \qquad \text{(quadratic minimum)}. \tag{7}$$

In Appendix D, we consider a more general finite-density effective potential of ϕ and clarify the regime of validity of the above Eq. (5). We plot in Fig. 1 the $V_{\mathrm{eff}}(\phi)$ for representative values of ϕ_*/ϕ_0. It can be seen that there are two types of minima, ϕ_* and ϕ_0, which we dub the quadratic minimum and linear minimum, respectively. The linear potential piece $m_\phi^2\phi_*|\phi_0 - \phi|$ arises from the tendency of the $\phi - \chi$ system to minimize the nonrelativistic χ's effective mass $|m_\chi - g\phi|$ at zero, which occurs at the linear minimum $\phi = \phi_0$.[3] The quadratic minimum ϕ_* is relevant when $\phi < \phi_0$. It appears because adding the linear

[1]Note that the sign of a fermion mass term is unphysical because it can be flipped by a chiral rotation $\chi \to e^{i\gamma_5\pi/2}\chi$, which leaves the rest of the Lagrangian invariant. The two theories related by this rotation are identical, so no physical quantity can depend on the sign of $m_\chi - g\phi$, and only its magnitude $|m_\chi - g\phi|$ is meaningful. For instance, the dispersion relation of χ in a ϕ background reads $E = \sqrt{p^2 + (m_\chi - g\phi)^2}$, which is manifestly independent of the sign of $m_\chi - g\phi$.

[2]Since fermions and bosons behave similarly when they are nonrelativistic and non-overlapping in wave function, many of our conclusions will also apply, e.g., to a real scalar $\tilde{\chi}$ in place of our fermion DM χ, with an effective mass in the presence of ϕ background mimicking that of our fermion χ, namely $M_{\tilde{\chi}}(\phi) = m_{\tilde{\chi}} - g\phi$. This specific effective mass of the scalar $\tilde{\chi}$ can arise, for example, from the Lagrangian term $-\mathcal{L} \supset (m_{\tilde{\chi}} - g\phi)^2\tilde{\chi}^2/2$ [17].

[3]The approximation that $(m_\chi - g\phi)\langle\bar{\chi}\chi\rangle$ reduces to a linear potential $|m_\chi - g\phi|n_\chi$ breaks down in a small region around $\phi = \phi_0$ where χ is relativistic. In this region, $V_{\mathrm{eff}}(\phi)$ is to leading order quadratic instead of linear in ϕ. Nevertheless, this quadratic regime corresponds to a very small range of ϕ around $\phi = 0$ and does not affect the dynamics of ϕ appreciably. We discuss this further in Appendix D.

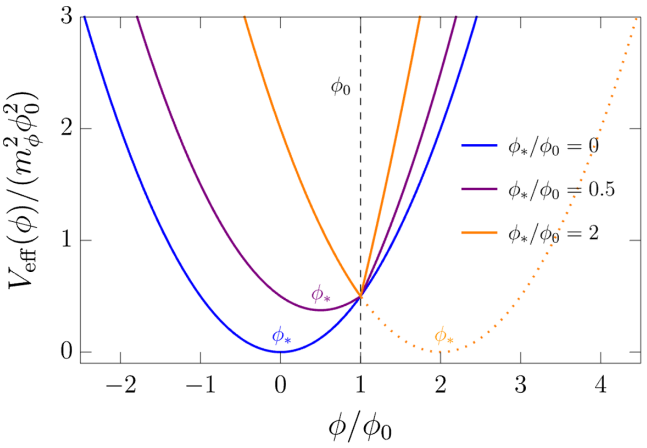


FIG. 1. Effective potential of ϕ; cf. Eq. (5). The labels "ϕ_0" and "ϕ_*" mark the locations of the linear finite-density minimum $\phi_0 = m_\chi/g$ and the quadratic finite-density minimum $\phi_* = gn_\chi/m_\phi^2$. In the $\phi_*/\phi_0 = 2$ case, the ϕ_* exists mathematically {the dotted line shows what $V_{\rm eff}$ would be if the $|\phi_0 - \phi|$ in Eq. (5) is replaced with $\phi_0 - \phi$ without the absolute value}, but in this case, ϕ_* cannot be reached because before ϕ can get there, the slope of the linear potential flips sign at $\phi = \phi_0$.

term $-m_\phi^2\phi_*\phi$ to the quadratic potential $m_\phi^2\phi^2/2$ amounts to a shift in the quadratic potential.[4] In the early Universe, the ratio $\phi_*/\phi_0 \propto n_\chi$ starts relatively high, likely greater than unity, in which case ϕ_0 is the only minimum of $V_{\rm eff}(\phi)$. Once ϕ_*/ϕ_0 goes below unity, the minimum of $V_{\rm eff}$ switches to ϕ_*, which moves progressively closer to the origin ($\phi = 0$) due to its $\phi_* \propto n_\chi \propto a^{-3}$ scaling. Notice also that the values of $V_{\rm eff}$ at these minima are nonzero and given by $V_{\rm eff}(\phi_0) = m_\phi^2\phi_0^2/2$ and $V_{\rm eff}(\phi_*) = m_\chi n_\chi - m_\phi^2\phi_*^2/2$ where $-m_\phi^2\phi_*^2/2 \propto a^{-6}$. As we will see, the constancy of $V_{\rm eff}(\phi_0)$ could lead to the dark sector as a whole, behaving as dark energy at the background level.

The quadratic piece of the potential dominates as long as $\phi \ll \phi_0$. Writing $\phi_0 = (4\pi G\alpha_{\chi\chi})^{-1/2}$ helps understand that a leading order perturbative expansion in $\alpha_{\chi\chi}$, such as the one performed in the previous precision cosmology studies of Refs. [22–25], is tantamount to effectively sending $\phi_0 \to \infty$, or equivalently, saying that the mass shift $g\phi$ is never comparable to the bare mass scale m_χ. While this is appropriate for the precision study of suitably small couplings, it cannot be assumed of regions of parameter space with large couplings.

B. Scalar evolution

In our setting, the static Yukawa parametrization in Eq. (1) corresponds to locking the cosmological background mediator field value ϕ at the quadratic minimum ϕ_*. However, as we will see, ϕ in general deviates far from ϕ_* and instead evolves nontrivially according to its equation of motion

$$\ddot{\phi} + 3H\dot{\phi} + m_\phi^2[\phi - \phi_*\mathrm{sign}(\phi_0 - \phi)] = 0. \tag{8}$$

We solve this equation numerically for several representative parameters λ_ϕ and $\alpha_{\chi\chi}$ and plot the resulting magnitude of ϕ as a function of scale factor a in Fig. 2. In solving Eq. (8), we assume that the Hubble rate H is given by the ΛCDM value. This is justified deep in RD, where the dark sector's energy density is subdominant, even if the background evolution of the dark sector deviates considerably from that in ΛCDM.

Rather insensitively to the initial condition, the ϕ field initially tracks an attractor solution $\phi_{\rm rise}$, whose form during RD is

$$\phi_{\rm rise} = \frac{3\alpha_{\chi\chi}f_\chi}{4}\phi_0\frac{a}{a_{\rm eq}}, \tag{9}$$

where $a_{\rm eq}^{-1} \approx 3400$ is the scale factor at matter-radiation equality. Physically, during this phase, where $\phi \ll \phi_0, \phi_*$, the ϕ field simply rolls down the initially dominant, linear potential piece $V_{\rm eff} \supset m_\phi^2\phi_*|\phi_0 - \phi|$. After a sufficient field excursion, this rolling becomes insensitive to the initial values of ϕ and $\dot{\phi}$, asymptoting to the dynamical attractor $\phi_{\rm rise}$; we clarify the basin of attraction toward $\phi_{\rm rise}$ in Appendix A. The existence of this attractor greatly reduces the sensitivity of our subsequent analyses to the initial conditions of ϕ. This universal attractor solution $\phi_{\rm rise}$ is seen in the rising behavior on the left sides of all three of the red, green, and blue curves in Fig. 2.

We show in Appendix B that as long as ϕ tracks the attractor solution $\phi_{\rm rise}$ at early times, big bang nucleosynthesis (BBN) does not place any constraint on $(\lambda_\phi, \alpha_{\chi\chi})$.

The tracking of the attractor $\phi_{\rm rise}$ continues until $\phi_{\rm rise}$ hits either the linear minimum ϕ_0 or the quadratic minimum ϕ_*. Depending on which of these comes first, ϕ subsequently oscillates around the linear minimum with a radiation-like equation of state (ϕ_0 phase) or around the quadratic minimum with a CDM like equation of state ($\phi_{\rm DM}$ phase). This binary possibility causes the evolution of ϕ to branch into two cases:

(1) *Small χ mass-change regime*: If $(3\alpha_{\chi\chi}f_\chi/4)(\lambda_\phi H_{\rm eq})^{1/2} \lesssim 1$, ϕ goes through the following phases, $\phi_{\rm rise} \to \phi_{\rm DM}$, and so χ's effective mass remains $M_\chi \approx m_\chi$ throughout. This is shown in the blue curve in Fig. 2. The increasing $\phi_{\rm rise} \propto a$ solution gets cut by the crossing with the decreasing $\phi_* \propto a^{-3}$. At that point, ϕ switches to oscillating around ϕ_* with its amplitude scaling as $\sim a^{-3/2}$ like that of an ultralight DM condensate [34–36] until the present epoch.

[4]When $\phi < \phi_0$, the effective potential can be written as $V_{\rm eff} = m_\phi^2(\phi - \phi_*)^2/2 - m_\phi^2\phi_*^2/2 + m_\chi n_\chi$.

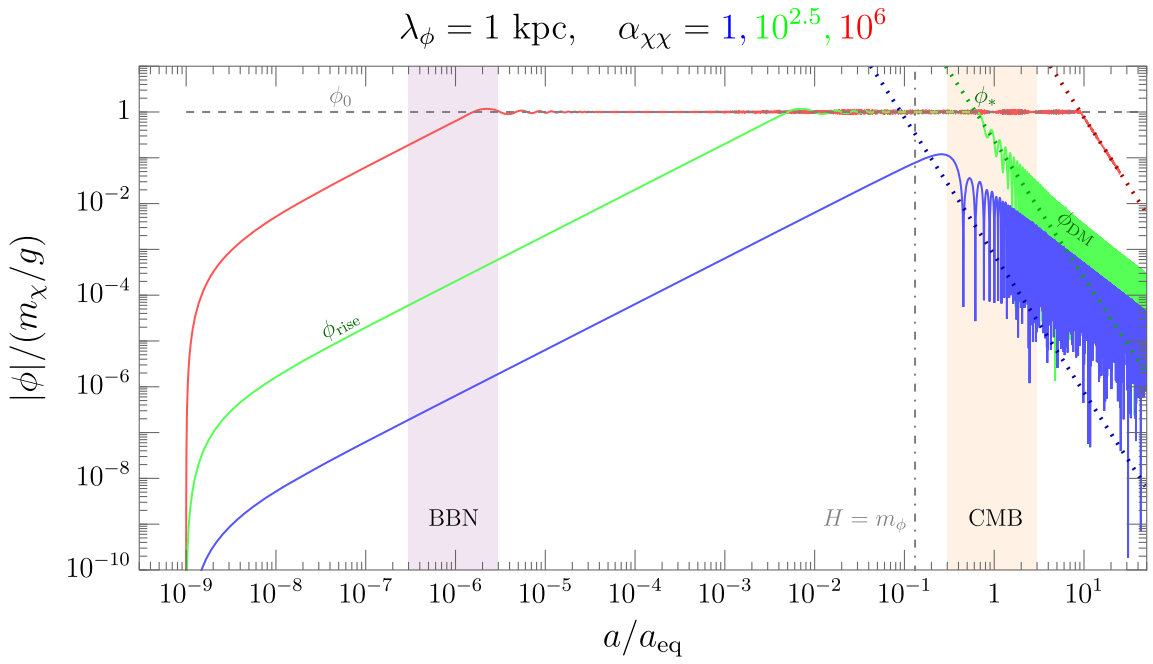


FIG. 2. Evolution of the background scalar field ϕ as a function of the scale factor a for $\lambda_\phi = 1$ kpc and $\alpha_{\chi\chi} = 1$ (solid blue), $\alpha_{\chi\chi} = 10^{2.5}$ (solid green), $\alpha_{\chi\chi} = 10^6$ (solid red). Here, we set $a_i = 10^{-9}$, $\phi(a_i) = 0$, and $\dot\phi(a_i) = 0$, and numerically evolve ϕ using the equation of motion Eq. (8). The gray dashed line is the linear finite-density minimum $\phi_0 \equiv m_\chi/g$ defined in Eq. (6). The dotted lines are the quadratic finite-density minima $\phi_* \equiv g n_\chi/m_\phi^2$, defined in Eq. (7), corresponding to the values of $\alpha_{\chi\chi}$ matching their colors. The dark green labels "$\phi_{\rm rise}$" and "$\phi_{\rm DM}$" indicate, respectively, where in the evolution of the green line the ϕ field tracks the attractor solution $\phi_{\rm rise} \equiv (3\alpha_{\chi\chi} f_\chi/4)\phi_0(a/a_{\rm eq})$, defined in Eq. (9), and oscillates as CDM, with an amplitude scaling as $\propto a^{-3/2}$. The colored bands labeled "BBN" and "CMB" indicate the very rough points where BBN and recombination take place.

(2) *Large χ mass-change regime*: If $(3\alpha_{\chi\chi} f_\chi/4)(\lambda_\phi H_{\rm eq})^{1/2} \gtrsim 1$, then ϕ goes through the following phases $\phi_{\rm rise} \to \phi_0 \to \phi_* \to \phi_{\rm DM}$, in which case M_χ sometimes deviates significantly from m_χ. This is shown in the red and green curves in Fig. 2. The increasing $\phi_{\rm rise} \propto a$ first crosses ϕ_0. Following that, ϕ oscillates around ϕ_0 with its maximum speed scaling as $\dot\phi \sim H\phi \propto a^{-2}$, thus behaving at background level as dark radiation, $(\dot\phi)^2/2 \sim H^2\phi^2 \propto a^{-4}$. Once the Hubble H goes below the bare mass m_ϕ, the energy density of the constant mass term $m_\phi^2\phi_0^2/2$ becomes dominant over the kinetic energy of ϕ, and the dark sector subsequently behaves as dark energy. This continues until ϕ_0 is crossed by $\phi_* \propto a^{-3}$. Then, ϕ turns to oscillating around ϕ_*. Because the oscillation of ϕ around ϕ_0 had been significantly Hubble damped between the $\phi_{\rm rise} \to \phi_0$ and $\phi_0 \to \phi_*$ points, the oscillation around ϕ_* is at first tiny compared to ϕ_*, and so ϕ appears to closely track ϕ_*. After a while, as the oscillation amplitude scales as $\propto a^{-3/2}$, which is slower than $\phi_* \propto a^{-3}$, this oscillation eventually dominates, and from then on, ϕ behaves as an ultralight DM condensate with an amplitude scaling as $\propto a^{-3/2}$.

The case boundary $(3\alpha_{\chi\chi} f_\chi/4)(\lambda_\phi H_{\rm eq})^{1/2} \sim 1$ applies only for $\lambda_\phi \lesssim H_{\rm eq}^{-1}$, in which case the $\phi_{\rm rise}$ does not extend beyond RD. For $\lambda_\phi \gg H_{\rm eq}^{-1}$, this boundary asymptotes to $\alpha_{\chi\chi} f_\chi \sim 0.1$.

Depending on the place in parameter space $(\lambda_\phi, \alpha_{\chi\chi})$, at matter-radiation equality, the ϕ field can be tracking $\phi_{\rm rise}$, tracking ϕ_0, oscillating around ϕ_0 (with dark-radiation like scaling), tracking ϕ_*, or oscillating around ϕ_* (with CDM-like scaling). These possibilities are mapped in Fig. 3. Given that the cosmic history is particularly well measured at $z \sim 10^3$, around matter-radiation equality, we can already deduce that parameter space points that yield $\phi \approx \phi_0$ then are almost certainly ruled out, as they predict a dark sector that behaves very differently from CDM. Less obviously, we will see in the next subsection that parameter space that yields $\phi \approx \phi_0$ prior to that, at $z \sim 10^3$–10^5, may also be ruled out when confronted with cosmological data, which have some sensitivity to this earlier epoch. Moreover, we have some probes of cosmic-perturbation modes that entered the horizon at even earlier times. As we will show in Sec. IV, these can rule out additional parameter space that predicts a dark sector deviating significantly from CDM at $z \gtrsim 10^5$.

We also show in Fig. 3 the existing limits from Bullet Cluster observation [11] and Ref. [22]'s analysis of Planck

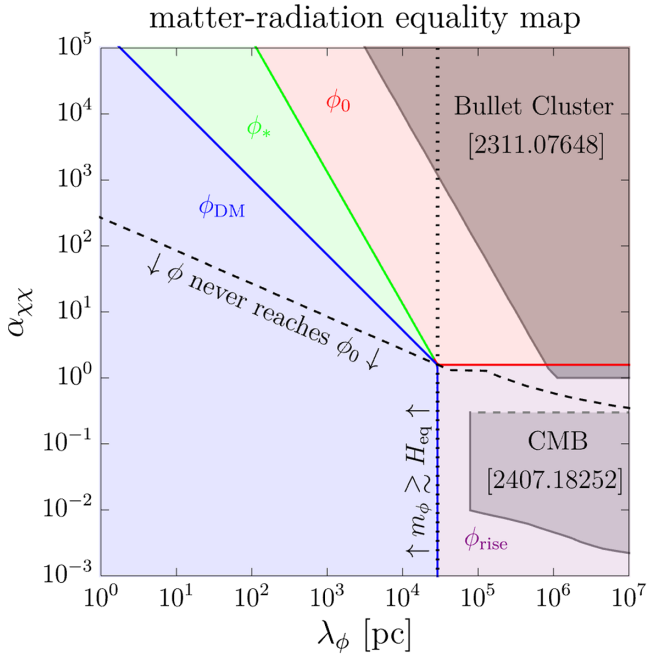


FIG. 3. Types of solutions of ϕ at matter-radiation equality. For example, in the red region, ϕ is at $\phi_0 \equiv m_\chi/g$, the linear finite-density minimum. In the green region, ϕ is at $\phi_* \equiv g n_\chi/m_\phi^2$, the quadratic finite-density minimum. In the purple region, it is at $\phi_{\rm rise} \equiv (3\alpha_{\chi\chi} f_\chi/4)\phi_0(a/a_{\rm eq})$, the attractor solution initially tracked by ϕ for a wide range of initial conditions, and in the blue region, it is at $\phi_{\rm DM}$, which is a CDM-like solution where ϕ oscillates with its bare mass with an amplitude that scales as $a^{-3/2}$. In the red region, the dark sector behaves as dark energy to the left of the dotted line and as dark radiation to the right of the dotted line. Also shown are the regions ruled out by Bullet Cluster observation [4] and small-coupling analysis of the CMB data [22]. The CMB limit has a ceiling because Ref. [22] employs a small-$\alpha_{\chi\chi}$ approximation in their analysis. We show a rough approximation to this ceiling as a dashed-gray upper boundary, corresponding to $\alpha_{\chi\chi} \sim 0.3$. Above (below) the black dashed line, ϕ does (does not) reach ϕ_0, which corresponds to the large (small) mass-change regime, i.e., the $\phi_{\rm rise} \to \phi_0 \to \phi_* \to \phi_{\rm DM}$ ($\phi_{\rm rise} \to \phi_{\rm DM}$) case.

CMB data [33]. Incidentally, the Bullet Cluster limit lies almost entirely in the ϕ_0 regime, whereas the small-$\alpha_{\chi\chi}$ CMB limit lies completely in the $\phi_{\rm rise}$ regime. In obtaining their limits, Ref. [22] employs a small-$\alpha_{\chi\chi}$ approximation in their analysis. Thus, this analysis breaks down at large enough $\alpha_{\chi\chi}$. For instance, a nonzero $\alpha_{\chi\chi}$ leads to a shift $g\phi$ in the effective mass of χ. The small-coupling treatment requires this mass shift stays small compared with the bare mass, $g\phi \ll m_\chi$, so that ϕ never reaches the linear minimum $\phi_0 = m_\chi/g$, where the dark sector deviates starkly from CDM. To roughly represent the $\alpha_{\chi\chi}$ cutoff above which their analysis is expected to be invalid, we draw a gray-dashed line at the top of the ruled-out region at $\alpha_{\chi\chi} \sim 0.3$. Note that this cutoff is not exact; its precise value is not to be taken literally.

Naively, as we increase the self-interaction of DM from zero, we would expect its behavior to deviate increasingly from that of CDM. However, it is possible that beyond the small-coupling regime, entirely new dynamics takes over and reverses this tendency, making the cosmology more viable. For instance, one could imagine in some models that when the DM self-interaction is increased beyond a certain threshold, the dark sector could clump into compact macroscopic objects whose large-scale behavior returns to resembling CDM. Therefore, small-coupling analyses that rule out couplings larger than a certain value should end their exclusion region at the largest couplings where their small-coupling approximation is valid. Moreover, existing data may be reinterpreted through the lens of different or less-minimal dark sector models, where the boundaries that delineate what is ruled out and what is not could change. It is therefore important to understand the actual physical effects that rule out a given parameter space.

C. Catastrophic dark sector equation of state

The background energy density $\rho_{\rm dark}$ and pressure $p_{\rm dark}$ of the ϕ and χ fields, collectively, are given by

$$\rho_{\rm dark} = |m_\chi - g\phi| n_\chi + \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}m_\phi^2\phi^2, \tag{10}$$

$$p_{\rm dark} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}m_\phi^2\phi^2, \tag{11}$$

where we have neglected the contribution to $p_{\rm dark}$ from χ, which we assume to be cold. On timescales longer than that of ϕ oscillation, the dark sector energy density scales as $\rho_{\rm dark} \propto a^{-3(\bar{w}_{\rm dark}+1)}$, with an average equation of state $\bar{w}_{\rm dark}$ given by

$$\bar{w}_{\rm dark} = \frac{\langle \dot{\phi}^2/2 - m_\phi^2\phi^2/2\rangle}{\langle |m_\chi - g\phi| n_\chi + \dot{\phi}^2/2 + m_\phi^2\phi^2/2\rangle} = \begin{cases} 0, & \phi_{\rm rise} \text{ or osc. around } \phi_* (|\phi| \ll \phi_0) \\ 1/3, & \text{osc. around } \phi_0 (|\phi - \phi_0| \gg \phi_0) \\ -1, & \text{tracking } \phi \approx \phi_0 (|\phi - \phi_0| \ll \phi_0), \end{cases} \tag{12}$$

which corresponds to $\rho_{\rm dark} \propto a^{-3}, a^{-4}, a^0$, respectively. Here, angled brackets $\langle \ldots \rangle = \int_0^{T_{\rm osc}} dt \ldots / T_{\rm osc}$ denote a time average over an oscillation period $T_{\rm osc}$, assumed $\ll H^{-1}$. For an oscillation around a polynomial potential $V(\phi) \propto \phi^n$, the time average of the kinetic energy of ϕ evaluates to $\langle \dot{\phi}^2/2 \rangle = n\langle V \rangle/2$ as per virial theorem. The evolution regimes where the dynamics of ϕ is mainly dictated by the linear potential $|m_\chi - g\phi| n_\chi$ and by the quadratic potential $m_\phi^2\phi^2/2$ correspond to $n = 1$ and $n = 2$, respectively.

According to Eq. (12), the $\bar{w}_{\rm dark}$ deviates significantly from zero when ϕ is either oscillating around or closely tracking ϕ_0. Specifically, our estimates (detailed in Appendix A) show that it is radiationlike ($\bar{w}_{\rm dark} \approx 1/3$) from $a = (3\alpha_{\chi\chi} f_\chi/4)^{-1} a_{\rm eq}$ to $H \sim m_\phi$ and dark-energy-like ($w_{\rm dark} \approx -1$) from $H \sim m_\phi$ to $a = (3\alpha_{\chi\chi} f_\chi/4)^{1/3} (\lambda_\phi H_{\rm eq})^{2/3} a_{\rm eq}$. Such a large deviation of $\bar{w}_{\rm dark}$ from the CDM-like $\bar{w}_{\rm dark} \approx 0$ only occurs in the large mass-change regime ($\phi_{\rm rise} \to \phi_0 \to \phi_* \to \phi_{\rm DM}$). In the small mass-change regime ($\phi_{\rm rise} \to \phi_{\rm DM}$), in which ϕ does not reach ϕ_0, the average $w_{\rm dark}$ remains CDM-like throughout. Thus, we expect cosmological observations to yield strong limits particularly in the large mass-change case.

We plot in Fig. 4 the evolution of the instantaneous $w_{\rm dark}$ as a function of the scale factor a for the parameters $\lambda_\phi = 10$ kpc and $\alpha_{\chi\chi} = 100$, representative of the large mass-change case ($\phi_{\rm rise} \to \phi_0 \to \phi_* \to \phi_{\rm DM}$). During most of $\phi_{\rm rise}$, the energy density of ϕ is negligible and $w_{\rm dark} \approx 0$. $w_{\rm dark}$ begins to rise when $\phi_{\rm rise}$ approaches ϕ_0. Once ϕ switches to oscillating under the linear potential around ϕ_0, $w_{\rm dark}$ becomes spiky, varying between 0 and 1, and averages to $1/3$. This is seen in Fig. 4 in the rapid oscillations of the gray line between $a \sim 10^{-9}$ and $a \sim 10^{-4}$. After the ϕ oscillation gets sufficiently Hubble damped, the value of ϕ essentially freezes at ϕ_0, and the bare mass term $m_\phi^2 \phi_0^2/2$ behaves as dark energy, with $\bar{w}_{\rm dark} \approx -1$. This is seen when the gray line is near -1 around $a \sim 10^{-3}$. At some point, ϕ_* crosses ϕ_0, whereupon ϕ begins oscillating around the moving ϕ_*. As that happens, $w_{\rm dark}$ departs from -1 and begins rising, while ϕ closely tracks ϕ_*. Once the amplitude of ϕ's oscillation around ϕ_* becomes large compared to ϕ_*, the ϕ field starts behaving as CDM, and $w_{\rm dark}$ then undergoes tiny oscillations around zero such that on average $\bar{w}_{\rm dark} \approx 0$.

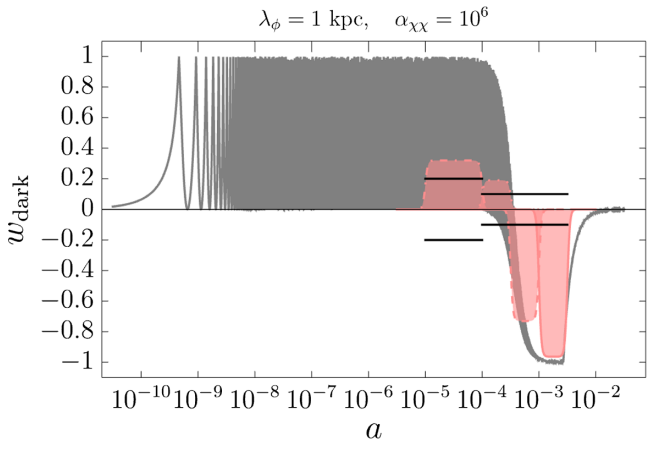


FIG. 4. Evolution of the instantaneous dark sector equation of state $w_{\rm dark} = p_{\rm dark}/\rho_{\rm dark}$ (solid gray) as a function of the scale factor a. The dark sector pressure $p_{\rm dark}$ and energy density $\rho_{\rm dark}$ are defined in Eqs. (10) and (11). The widths and amplitudes of the red bars represent the four scale-factor bins 1, 2, 3, 4 [see Eq. (13)] used in our analysis and the averaged equation of state $\bar{w}_{\rm dark}^{(i)}$ within them. The black lines mark the boundaries of the allowed ranges of $\bar{w}_{\rm dark}^{(i)}$; we consider parameter space with red bars protruding the black lines as ruled out.

D. Constraints on dark matter equation of state

References [37,38] have performed relatively model-independent analyses to place constraints on the evolution of the dark sector's equation of state $w_{\rm dark}(a)$ across the observable cosmic history. Instead of assuming a specific functional form of $w_{\rm dark}(a)$, as commonly done, they allow its values at different epochs to float freely in bins of scale factor.[5] The equations of state in the different bins then serve as extra parameters, in addition to the six standard ΛCDM parameters, with which to fit cosmological data, such as those of CMB anisotropies. Because the $w_{\rm dark}(a)$ assumed in their analyses are allowed to vary in a less-constrained and model-independent manner, their analyses lead to constraints that are not only conservative but also applicable to a wide range of scenarios, including ours.

Reference [37] divided the scale factor range from 10^{-5} to 1 into eight logarithmic scale-factor bins. We consider only the first four of these bins, corresponding to the scale factor ranges:

$$\underbrace{[10^{-5}, 10^{-4}]}_{\text{bin 1}}, \underbrace{[10^{-4}, 10^{-3.5}]}_{\text{bin 2}}, \underbrace{[10^{-3.5}, 10^{-3}]}_{\text{bin 3}}, \underbrace{[10^{-3}, 10^{-2.5}]}_{\text{bin 4}}. \tag{13}$$

These are the most constraining bins for the model because it behaves very closely to ΛCDM at scale factors later than these bins. We denote the average equation of state in bin $i = 1, 2, 3, 4$ as

$$\bar{w}_{\rm dark}^{(i)} \equiv \frac{\int_{{\rm bin}\, i} d\ln a\, w_{\rm dark}}{\int_{{\rm bin}\, i} d\ln a}. \tag{14}$$

The results of Ref. [37] show that consistency with CMB power spectrum (Planck 2015), CMB lensing likelihood (Planck 2015), BAO (6dF galaxy survey and SDDS-III BOSS), and the Hubble Space Telescope prior on H_0, restricts the $\bar{w}_i$s to certain ranges close to zero. In some bins close to matter-radiation equality, their limits are so stringent that they even ruled out ΛCDM at 99% confidence level. While our treatment is inspired by the analysis of Ref. [37], to keep our results robust, we instead consider the following very conservative approximations to their limits:

[5]See also Refs. [39,40], which are similar in spirit. These works constrain the Hubble rate at different epochs, instead of the dark sector equation of state. However, they only allow *positive* deviations in Hubble relative to the ΛCDM evolution. On the other hand, our scenario actually predicts local-in-time Hubble deficits relative to ΛCDM, and so their analyses unfortunately do not cover our scenario.

$$-0.2 < \bar{w}^{(1)}_{\rm dark} < +0.2, \qquad -0.1 < \bar{w}^{(2,3,4)}_{\rm dark} < +0.1. \quad (15)$$

These are shown as black solid lines in Fig. 4. We refer the reader to Fig. 13 in Appendix F for a comparison between the actual ranges of allowed $\bar{w}_i$ found in Ref. [37] and the very conservative ones we assume. Since the $w_{\rm dark}$ in our scenario can, at times, deviates far from zero (see Fig. 4), the resulting limits on dark forces do not depend sensitively on the precise values of the maximum and minimum allowed $\bar{w}^{(i)}_{\rm dark}$ assumed, as long as they are well below $1/3$ and well above -1, respectively. We discuss the dark-force constraints that result from imposing Eq. (15) in Sec. V. The results are shown in Fig. 7.

We note that any constraints on background cosmology based on probes of cosmic perturbations, e.g., the CMB power spectra, are unavoidably intertwined with details of the perturbation evolution, which we do not study in full in this work (we only consider a limited regime in Sec. IV where the perturbations behave relatively simply). Reference [37] takes a step further to make their analysis less sensitive to assumptions on cosmic perturbation evolution. They adopt the generalized DM parametrization [41], which in addition to $w_{\rm dark}$ includes also the dark sector's sound speed c_s and shear viscosity $c_{\rm vis}$ for each scale-factor bin, both allowed to float, though assumed wave number independent. After marginalizing over the c_s and $c_{\rm vis}$, Ref. [37] found constraints on $\bar{w}_{\rm dark}$ that are less stringent but comparable to that of Ref. [38], which assumes $c_s = c_{\rm vis} = 0$. This greatly reduces but does not completely remove the sensitivity of the results to unmodeled cosmic perturbations. In fact, in some cases of our interest, the evolution of the dark sector's sound speed is not only highly nontrivial (and so may not be well represented by coarse-grained values) but also wave number dependent. Nevertheless, achieving an exact cancellation between the effects of dark-sector perturbations and those of a modified background expansion would require an extraordinary degree of fine-tuning. We therefore regard such scenarios as highly implausible and consequently exclude regions of parameter space that produce extreme departures from the standard background evolution violating Eq. (15).

IV. COSMIC PERTURBATIONS

Here, we study the evolution of over-densities in DM in the presence of dark forces. Dark forces can have drastic consequences on cosmic perturbations even in regimes where the background evolution is consistent with ΛCDM. Hence, cosmic-perturbation considerations can place limits complementary to those obtained from background cosmology.

A. Meszaros-like perturbation equation

It has long been suggested that dark forces can induce exponential growth in DM perturbations [27–29]. Given the catastrophic outcome it may cause, we will focus on this effect in constraining dark forces through perturbation analyses. We note that the exponential growth is based on the assumption that the effect of coupling the DM χ to the mediator ϕ amounts to replacing the Newton's constant G with $G_{\rm eff}(k) = G[1 + \alpha_{\chi\chi}k^2/(a^2m_\phi^2 + k^2)]$ in cosmic perturbation equations, namely the Meszaros equation. This corresponds to considering only the particular solutions to the mediator's equation of motion Eq. (8) and its perturbation, which physically can be interpreted as the mediator field being sourced in a quasistatic manner by the DM particles, as in electrostatics. In Appendix C, we derive the Meszaros-like perturbation equation with $G \to G_{\rm eff}$ [Eq. (18)] from first principles, clarifying under which conditions this equation is valid. We find that the cosmic perturbation equation only reduces to such a form when ϕ is closely tracking its quadratic finite-density minimum $\phi_* = gn_\chi/m_\phi^2$ while its value is sufficiently small $\phi_* \ll \phi_0$.

We have seen in Sec. III that the mediator field ϕ has rich background dynamics where it generically deviates far from ϕ_*. While this suggests that the substitution $G \to G_{\rm eff}$ is rarely valid, we have sufficient understanding of ϕ's background evolution to pinpoint when $\phi \approx \phi_*$ is actually satisfied. As found in Sec. III, this occurs only in the large mass-change case ($\phi_{\rm rise} \to \phi_0 \to \phi_* \to \phi_{\rm DM}$), starting right after ϕ_0 is crossed by ϕ_*, at scale factor $a_{\phi_0=\phi_*}$, and ending when the ϕ background begins to redshift as CDM at scale factor $a_{\phi_*=\phi_{\rm DM}}$. For example, in Fig. 2, the tracking of ϕ_* occurs for the green curve at $a/a_{\rm eq} \sim 1$ and for the red curve at $a/a_{\rm eq} \gtrsim 10$. In terms of the model parameters, $a_{\phi_0=\phi_*}$ and $a_{\phi_*=\phi_{\rm DM}}$ are given by (see Appendix A)

$$a_{\phi_0=\phi_*} = \left(\frac{3\alpha_{\chi\chi}f_\chi}{4}\right)^{1/3}\left(\frac{\lambda_\phi}{H_{\rm eq}^{-1}}\right)^{2/3}a_{\rm eq}, \quad (16)$$

$$a_{\phi_*=\phi_{\rm DM}} = \left(\frac{3\alpha_{\chi\chi}f_\chi}{4}\right)^{7/9}\left(\frac{\lambda_\phi}{H_{\rm eq}^{-1}}\right)^{8/9}a_{\rm eq}. \quad (17)$$

In order to approximately capture the exponential growth of the DM density contrast $\delta_\chi \equiv \delta n_\chi/n_\chi$, we consider the following Meszaros-like equation [42]:

$$\ddot{\delta}_\chi + 2H\dot{\delta}_\chi = \frac{3H^2}{2}\Omega_\chi\delta_\chi\left[1 + \frac{\alpha_{\chi\chi}\Theta_*(t)}{1 + (\frac{a/k}{\lambda_\phi})^2}\right], \quad (18)$$

where we have defined

$$\Theta_*(t) = \Theta[a(t) - 2a_{\phi_0=\phi_*}]\Theta[a_{\phi_*=\phi_{\rm DM}} - a(t)], \quad (19)$$

and $\Omega_\chi = m_\chi n_\chi/\rho_{\rm crit}$; Θ without a subscript is the Heaviside step function, and we have defined Θ_* such that it is equal to unity when ϕ is closely tracking ϕ_* and zero otherwise. We assume that H is given by the ΛCDM value. We start the support at $a = 2a_{\phi_0=\phi_*}$ instead of at $a = a_{\phi_0=\phi_*}$ in order to ensure $\phi_* \ll \phi_0$ and that $w_{\rm dark}$ and its time derivative are negligible (which also implies $|\dot{M}_\chi/M_\chi| \ll H$). Enhanced growth of δ_χ occurs when the right-hand side of Eq. (18) is $\gtrsim H^2\delta_\chi$. If at some point during the $\phi \approx \phi_*$ tracking phase, the physical wave number enters the range of the dark force, $a/k \ll \lambda_\phi$, then it suffices to have $\alpha_{\chi\chi} \gtrsim 1$. In the opposite case, when $a/k \gg \lambda_\phi$, the $\alpha_{\chi\chi}$ needs to be large enough that $\Omega_\chi\alpha_{\chi\chi}(\lambda_\phi k/a)^2 \gtrsim 1$.

We plot in Fig. 5 the evolution of δ_χ in the presence of dark forces relative to that in ΛCDM. A dramatic, exponential increase in the growth of δ_χ is clearly visible. The figure also shows that to cause an exponential growth, the dark force need not be long range compared to the physical wavelength of the mode. Interestingly, δ_χ continues to grow even outside the region where $\Theta_* = 1$. This is due to a combination of the inertia of δ_χ, as indicated by the presence of $\ddot{\delta}_\chi$ term in Eq. (18), and the speed $\dot{\delta}_\chi$ acquired during the time window when $\Theta_* = 1$.

Note that the density contrast δ_χ grows considerably even in ΛCDM, in the absence of dark forces. In order to quantify the enhancement due to dark forces relative to the ΛCDM case (corresponding to $\alpha_{\chi\chi} = 0$), we define the following quantity:

$$\mathcal{G}_{\rm rec}(k) \equiv \left.\frac{\delta_\chi(k)}{\delta_\chi^{\Lambda\rm CDM}(k)}\right|_{a=a_{\rm rec}}, \tag{20}$$

which we will refer to as the *relative growth factor*. We have chosen to evaluate this ratio at the epoch of recombination, $a_{\rm rec} \approx (1100)^{-1}$, for practical reasons which will become clear later. Comoving modes with $k \gg k_{\rm eq} \approx 0.01$ Mpc^{-1} cross the horizon far before matter-radiation equality, at $a_k \approx (k_{\rm eq}/\sqrt{2}k)a_{\rm eq}$. For such modes and $a \gtrsim a_{\rm eq}$, we can approximate the DM density contrast in ΛCDM as [43]

$$\delta_\chi^{\Lambda\rm CDM}(k) \approx \frac{3}{5}\frac{k^2}{\Omega_m H_0^2}\Phi_p(k)T(k)D(a), \tag{21}$$

where the transfer function $T(k)$ and growth function $D(a)$ can be found in [43]. We assume Φ_p to be the usual adiabatic, scale-invariant, superhorizon perturbation, and take as fiducial values $\Omega_m = 0.27$ and $H_0 = 70$ km/s/Mpc. Figure 6 shows how the extra growth due to long-range forces distorts the DM density contrast spectrum as quantified by $\mathcal{G}_{\rm rec}(k)$. In k space, the dark-force induced enhancement starts at $k \sim \alpha_{\chi\chi}^{-1/3}(\lambda_\phi H_{\rm eq})^{-2/3}k_{\rm eq}$, which corresponds to $\Omega_\chi\alpha_{\chi\chi}(\lambda_\phi k/a)^2 \sim 1$ at $a_{\phi_0=\phi_*}$. The enhancement plateaus

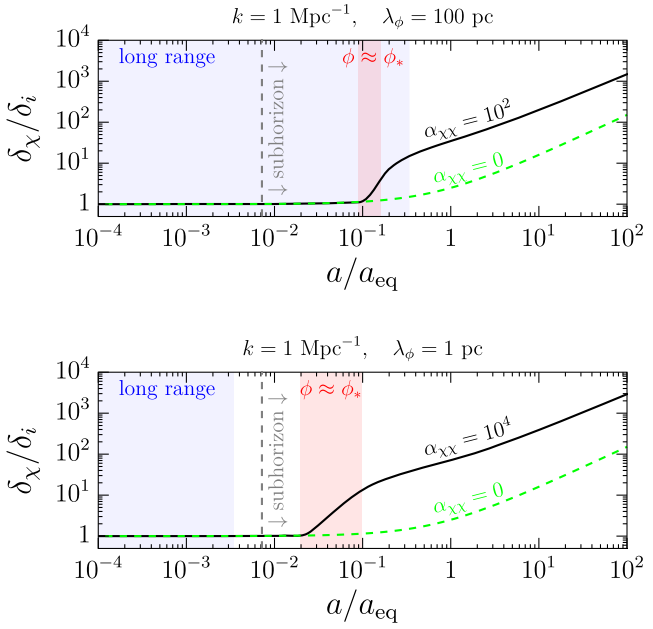


FIG. 5. The evolution of the DM density contrast $\delta_\chi = \delta n_\chi/n_\chi$ (solid black line, normalized to its initial, superhorizon value) as a function of the scale factor a (normalized to that at matter-radiation equality, $a_{\rm eq}$). Here, we follow the mode with comoving wave number $k = 1$ Mpc^{-1} and set the dark-force parameters as indicated in the plots. The gray dashed line marks the point where this mode becomes subhorizon. The black solid lines represent the evolution of δ_χ as dictated by Eq. (18), to be contrasted with the evolution in the absence of dark forces ($\alpha_{\chi\chi} = 0$) as given by Eq. (21) and shown in dashed green lines. The light red region shows the region where ϕ is expected to closely track ϕ_*, in which the Θ_* in Eqs. (18) and (19) is equal to unity. The light blue region indicates where the dark-force range is longer than the physical wavelength a/k of the mode. The *top* (*bottom*) plot corresponds to the case where the dark force is long-range $\lambda_\phi \gg a/k$ (short-range $\lambda_\phi \ll a/k$) when $\phi \approx \phi_*$. It can be seen that enhanced growth occurs in the light red region in either case. Furthermore, δ_χ appears to continue growing relative to the ΛCDM one even outside of the light red region. This is due to the speed $\dot{\delta}_\chi$ acquired in the light red region.

(becomes nearly k independent) when k exceeds $k \sim \alpha_{\chi\chi}^{7/9}(\lambda_\phi H_{\rm eq})^{-1/9}$, in which case the physical size of the mode stays smaller than λ_ϕ throughout the $\phi \approx \phi_*$ phase, and consequently the evolution of δ_χ becomes k independent. This happens when $a_{k\lambda} = k\lambda_\phi$ becomes larger than $a_{\phi_*=\phi_{\rm DM}}$.

A possible concern is that the aforementioned, approximate scale-invariant enhancement in δ_χ at high wave numbers is such a simple effect that it can easily be degenerate with other effects, such as a compensating change to the primordial perturbation amplitude. Nevertheless, requiring that ϕ begins tracking ϕ_* before recombination, $a_{\phi_0=\phi_*} \lesssim a_{\rm rec}$ implies that the wave number above which there is enhancement, $k \sim \alpha_{\chi\chi}^{-1/3}(\lambda_\phi H_{\rm eq})^{-2/3}k_{\rm eq}$, is higher than $k_{\rm eq}$. In other words, the turning on of the enhancement

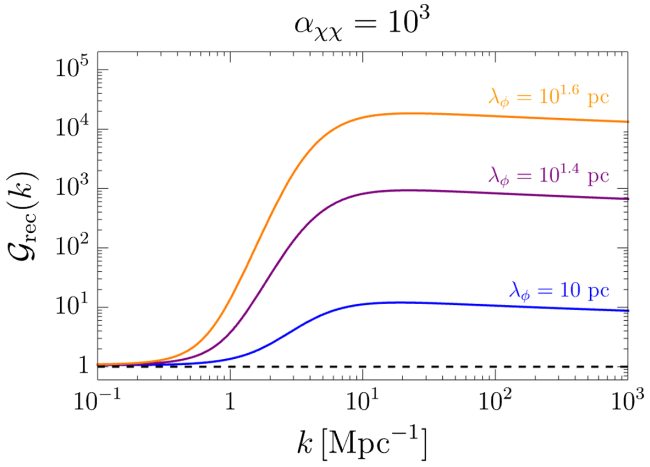


FIG. 6. Extra growth of DM density contrast $\mathcal{G}_{\rm rec}(k) = [\delta_\chi(k)/\delta_\chi^{\Lambda{\rm CDM}}]_{\rm rec}$ by the epoch of recombination as a function of comoving wave number k. Here, the δ_χ and $\delta_\chi^{\Lambda{\rm CDM}}$ are evolved according to Eq. (18) and Eq. (21), respectively, with the dark-force parameters as indicated in the plot.

occurs at some wave number $k \gtrsim k_{\rm eq}$, which generically lies within the range of wave numbers for which we have probes. Hence, the enhanced clustering due to the dark force can manifest as a nontrivial shape-altering effect on the observable power spectrum, and not just an increase in the overall amplitude of the power spectrum.

B. Scale-dependent growth enhancement of cosmic perturbations

In ΛCDM cosmology, modes that become subhorizon during RD typically grow by at most a factor of $\mathcal{O}(10)$ before the epoch of CMB decoupling. By contrast, we have seen that dark forces can cause orders of magnitude of extra growth in δ_χ, during (and also slightly after) the scale-factor window where background field ϕ is expected to closely track ϕ_*. We expect such a drastic modification to the evolution of δ_χ to have clearly observable implications in cosmological data.

The clustering of DM is imprinted in the linear matter power spectrum, which is particularly well probed for wave numbers $k \sim 10^{-3} - 10$ Mpc^{-1} by CMB anisotropy, galaxy surveys, weak lensing, and Lyman-α forest. At higher k's, matter perturbations have been relatively poorly probed. Recently, constraints on higher-k modes are strengthening [4,44–47]. In particular, observations of ultrafaint dwarf galaxies (UFDs) could be used to place constraints on the abundance of DM substructures expected from enhanced small-scale DM perturbations. Despite being relatively less stringent in terms of sensitivity to the amplitude of δ_χ (or equivalently the matter power spectrum), these higher-k constraints from UFDs are actually far better at probing the dark-force parameter space than the more-stringent, lower-k probes, since dark-force-induced growths tend to add orders of magnitude of power at higher wave numbers $k \gg k_{\rm eq}$.

Based on the constraints from Lyman-α and UFDs, as reported in [4], we consider the parameter space satisfying the following conditions (significantly weakened to be conservative) as ruled out:

$$\begin{aligned} \mathcal{G}_{\rm rec}(k = 2\ {\rm Mpc}^{-1}) &\gtrsim 10 \qquad (\text{Lyman-}\alpha), \\ \mathcal{G}_{\rm rec}(k = 300\ {\rm Mpc}^{-1}) &\gtrsim 10^2 \qquad (\text{UFD heating}), \end{aligned} \tag{22}$$

where $k = 2$ Mpc^{-1} and $k = 300$ Mpc^{-1} represent, respectively, the highest wave number k probed by Lyman-α and the k for which UFD is most sensitive. We will discuss the constraints that result from imposing these conditions in Sec. V. Note that we have deliberately chosen highly conservative $G_{\rm rec}(k)$ thresholds in Eq. (22) in order to make our results more robust. We have checked that the final constraints are insensitive to the exact threshold values used due to the rapid rate at which $G_{\rm rec}(k)$ increases as $\alpha_{\chi\chi}$ and λ_ϕ are varied near the approximate boundaries of the constrained regime.

Apart from Lyman-α and UFDs, FIRAS's null measurements of CMB spectral distortions should also place some constraints on extra DM perturbation growths. Spectral distortions can arise from Silk damping of small-scale baryon-photon fluid perturbations, which leads to inhomogeneous heating of the primordial plasma. Enhanced DM perturbations affect the predamping baryon-photon fluid perturbations through their contributions to the Newtonian potential Φ. The damping of a given baryon-photon fluid mode occurs primarily at the moment it becomes smaller than the (time-dependent) damping scale k_d, typically far before matter-radiation equality. Hence, we expect the impact of enhanced DM perturbations on Φ to suffer some suppression due to both the subdominance of the DM energy density and the high pressure of the baryon-photon fluid. Nevertheless, capturing these effects properly would require solving the full, coupled Boltzmann and Einstein equations, which we leave for future work.

V. RESULTS

A. Limits from background cosmology

1. Fixed χ abundance

The dark-force constraint that results from our cosmological-background analysis is shown on the right panel of Fig. 7 in beige color. Here, we set $f_\chi = f_{\rm DM} = 0.85$ and solve Eq. (8) numerically with the initial conditions $\phi = 0$ and $\dot\phi = 0$ at the scale factor $a = 10^{-9}$. We then use the solution to compute the average equation of state $\bar w_{\rm dark}^{(i)}$ in each bin, which we then confront with the allowed ranges of $\bar w_{\rm dark}$ we assumed in Eq. (15). Figure 7 shows that the constrained regime lies primarily in the large mass-change

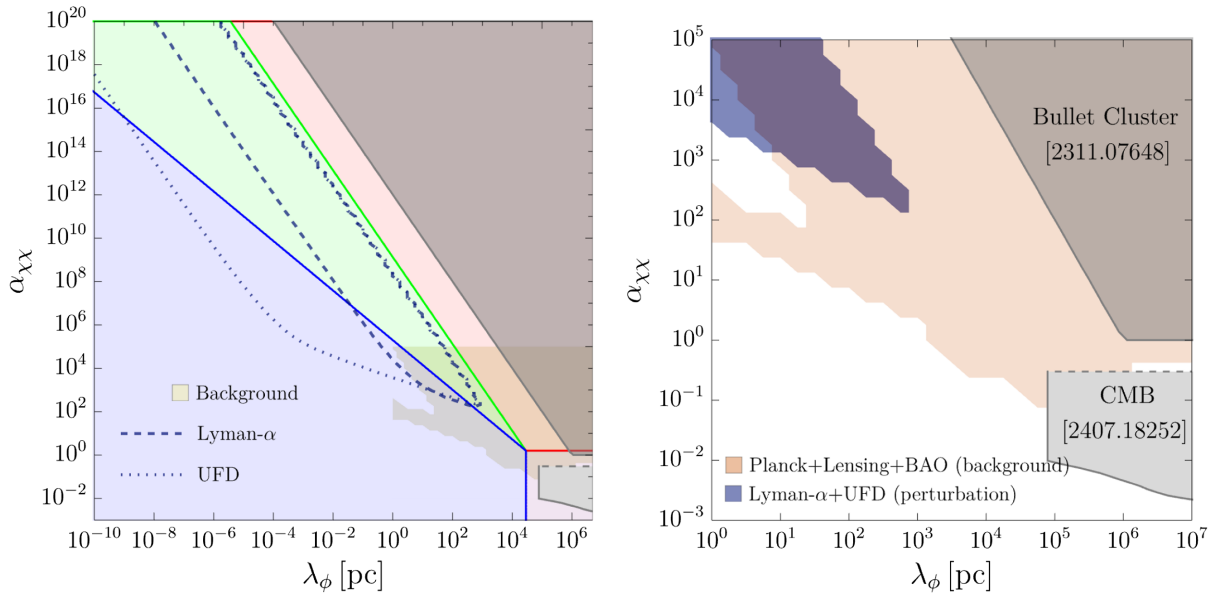


FIG. 7. Right: Combined limits on dark forces based on cosmological background analysis (beige) and perturbation analysis (dark blue). We also show existing limits from Bullet Cluster observation [11] and Ref. [22]'s analysis of CMB data. The CMB limit has a ceiling because Ref. [22] employs a small-$\alpha_{\chi\chi}$ approximation in their analysis. We indicate the rough upper boundary of the regime of validity of their result with dashed gray. We do not claim that the beige region continues to the left (small λ_ϕ direction) and upward (large $\alpha_{\chi\chi}$ direction) of the plot, although in principle it could. We stop our background analysis at $\lambda_\phi = 1$ pc because it becomes too computationally costly at $\lambda_\phi \ll 1$ pc. However, we expect the left edge of the beige region (apart from the part that sticks out, which we discuss in Sec. V A 2) to be covered by the dark blue region, which does extend to smaller λ_ϕ. Left: The same dark-force limit from perturbation analysis as in the right plot (where it was shown in dark blue), but extended to much smaller λ_ϕ and much larger $\alpha_{\chi\chi}$ to illustrate the extent of the limit. The individual limits from the two probes that make up the totality of perturbation limit in the right figure are shown in the left figure in dashed (Lyman-α) and dotted (UFD) lines. The light blue, light green, and light red colored regions correspond to that of the matter-radiation equality map of Fig. 3. For comparison, the same beige region in the right plot is reproduced in the left plot, where now its perceived color changes due to being overlaid by the light blue, green, and red regions.

part of the parameter space (see Fig. 3), as expected. For the most part, the constraint arises from the fact that when ϕ is oscillating around ϕ_0, the dark sector's equation of state can momentarily behave as dark radiation ($\bar{w}_{\rm dark} \approx 1/3$) and dark energy ($\bar{w}_{\rm dark} \approx -1$). These values of $\bar{w}_{\rm dark}$ are well outside of the allowed ranges of equation of state, Eq. (15), which apply for the scale-factor range $a^{-5} - a^{-2.5}$. Hence, parameter space that predicts ϕ oscillating around ϕ_0 during this scale-factor range is almost certainly ruled out. Thus, we expect parameter space that simultaneously satisfies the following three conditions to be ruled out: (1) in the large mass-change case ($\phi_{\rm rise} \to \phi_0 \to \phi_* \to \phi_{\rm DM}$), (2) $\phi_{\rm rise} \to \phi_0$ transition occurs at the latest before the end of bin 4, at a scale factor $a \lesssim 10^{-2.5}$, (3) $\phi_0 \to \phi_*$ occurs at the earliest after the start of bin 1, at a scale factor $a \gtrsim 10^{-5}$. These considerations predict that parameter space points that satisfy the following three inequalities are ruled out:

$$\text{Large mass-change}: \left(\frac{3\alpha_{\chi\chi} f_\chi}{4}\right)\left(\frac{\lambda_\phi}{H_{\rm eq}^{-1}}\right)^{\frac{1}{2}} \gtrsim 1, \tag{23}$$

$$a_{\phi_{\rm rise}\to\phi_0} \lesssim 10^{-2.5}: \ \frac{3\alpha_{\chi\chi} f_\chi}{4} \gtrsim \frac{a_{\rm eq}}{10^{-2.5}}, \tag{24}$$

$$a_{\phi_0\to\phi_*} \gtrsim 10^{-5}: \left(\frac{3\alpha_{\chi\chi} f_\chi}{4}\right)^{\frac{1}{3}}\left(\frac{\lambda_\phi}{H_{\rm eq}^{-1}}\right)^{\frac{2}{3}} \gtrsim \frac{10^{-5}}{a_{\rm eq}}. \tag{25}$$

This matches well with the numerically obtained limits shown in Fig. 7, except for the beige branch that sticks out. This branch is actually part of a strip corresponding to $(3\alpha_{\chi\chi} f_\chi/4)(\lambda_\phi H_{\rm eq})^{1/2} \sim 1$, which is ruled out by our analysis but is not captured by the approximate conditions of Eqs. (23)–(25). We discuss this strip in the next subsection.

2. Mixed-dark-matter strip and variation in χ abundance

As described in Sec. III, the oscillation of ϕ around ϕ_* eventually behaves like CDM. Therefore, ϕ contributes to the DM abundance at late times alongside χ. We can parametrize this contribution from ϕ the same way we did with χ in Eq. (3), namely as $f_\phi = \rho_{\phi,0}/\rho_{m,0}$. In nearly all cases, this contribution is negligible compared to

that of χ. However, close to the boundary between the regimes of small and large mass excursions, $(3\alpha_{\chi\chi}f_\chi/4)(\lambda_\phi H_{\rm eq})^{1/2} \sim 1$, the contribution from ϕ can be comparable to χ's, $f_\phi/f_\chi = \mathcal{O}(1)$. Hence, we call this approximate line the mixed-DM strip. In Fig. 7, a part of the beige region follows this line.

We reiterate that the limits shown in Fig. 7 were obtained by assuming the fermion χ makes up the entirety of the DM, which corresponds to setting $f_\chi = f_{\rm DM} = 0.85$. Since the mixed-DM strip predicts $f_\phi/f_\chi = \mathcal{O}(1)$, on it the total DM abundance today, as quantified by $f_\chi + f_\phi$, would exceed the fiducial value $f_{\rm DM} = 0.85$ by $\mathcal{O}(1)$ and would overclose the Universe. Furthermore, there are inherent, "geometric" degeneracies among cosmological parameters that allow a range of $f_{\rm DM}$ values different from 0.85 to be consistent with the CMB data [33,48]. Late-time cosmic expansion probes such as SN1a and BAO observations can break these degeneracies to some extent but still leave about 20% of $\Omega_{\rm DM,0}$ wiggle room. These uncertainties in $\Omega_{\rm DM,0}$ translate to uncertainties in the allowed value of $f_\chi + f_\phi$, and hence f_χ, since the value of f_ϕ is set once f_χ is specified.

The existence of the mixed-DM strip and the uncertainty in the allowed values of f_χ suggest that we need to reassess the validity of the beige limit in Fig. 7, which was produced with the assumption that $f_\chi = 0.85$, especially along the mixed-DM strip. We do so by checking the robustness of the limits against variations in f_χ. In general, we find that changing f_χ by $\mathcal{O}(1)$ in the range 0.7–1.2 amounts to mild (and not necessarily monotonic) shifts in the boundary of the region ruled out by each of the four scale-factor bins. For points deep inside the boundary of at least one of the bins, it is clear that there is no value of f_χ that yields $f_\chi + f_\phi$ consistent with SN1a and BAO that would save the points from being ruled out. Points along the edges of the ruled out region, including those in the mixed-DM strip, are close to the boundaries of two or more bins, which suggests that they are more likely to be salvageable. However, we checked multiple such points by varying $f_\chi/f_{\rm DM}$ in the range 0.7–1.2 and found that over this range of $f_\chi/f_{\rm DM}$, they remain ruled out by at least one of the bins, thus demonstrating the robustness of our background-cosmology limits.

B. Limits from cosmic perturbations

Our limits on the dark-force parameter space based on cosmic-perturbation analysis are shown in the right panel of Fig. 7 in dark blue. These limits gradually worsen with decreasing λ_ϕ, but can in principle extend to very small values of λ_ϕ, as shown in in dotted and dashed dark-blue lines in the left panel of Fig. 7. However, at some point, these limits are invalidated by model-dependent considerations, i.e., finite-density corrections to the mass and self-interactions of the ϕ field, as discussed in Appendix D.

At each parameter-space point, we solve the Meszaros-like Eq. (18) for δ_χ until the epoch of recombination. We then compute the relative growth factor defined in Eq. (20) and deem the parameter space point ruled out if it satisfies Eq. (22). The obtained limits are well approximated by the following conditions: (1) The $\phi \approx \phi_*$ period exists: $2a_{\phi_0=\phi_*} \lesssim \min[a_{\phi_*=\phi_{\rm DM}}, a_{\rm rec}]$, and (2) the dark force is effectively stronger than gravity when it is short-range compared to the physical wave number $k_{\rm probe}/a$ of the mode of interest ($k_{\rm probe}\lambda_\phi/a \gg 1$): $[(3H^2/2)\Omega_\chi\alpha_{\chi\chi}(k_{\rm probe}\lambda_\phi/a)^2]_{2a_{\phi_0=\phi_*}} \gg 1$. In terms of $\alpha_{\chi\chi}$ and λ_ϕ, this predicts that regimes satisfying the following are ruled out:

$$\left(\frac{\lambda_\phi}{H_{\rm eq}^{-1}}\right)^{-1/2} \lesssim \alpha_{\chi\chi} \lesssim \left(\frac{\lambda_\phi}{H_{\rm eq}^{-1}}\right)^{-2}, \tag{26}$$

$$\alpha_{\chi\chi} \gg \left(\frac{\lambda_\phi}{H_{\rm eq}^{-1}}\right)^{-2}\left(\frac{k_{\rm probe}}{k_{\rm eq}}\right)^{-3}, \tag{27}$$

where $k_{\rm eq} \approx 0.01\ {\rm Mpc}^{-1}$, and we have set $f_\chi = 0.85$. These limits provide a good approximation to the numerically obtained ones shown in Fig. 7. Note that the limits we obtained are not sensitive to the exact numerical values of the thresholds listed in (22), which we have chosen very conservatively. Accordingly, we have neglected various $\mathcal{O}(1)$ factors in writing Eqs. (26) and (27).

In our analysis, we do not consider the evolution of cosmic perturbations prior to the epoch when the scalar ϕ field is tracking its quadratic minimum, $\phi \approx \phi_*$. It is possible that during this earlier phase, some Fourier modes could grow to nonlinearity and collapse into bound clumps. If so, the typical interclump separation could exceed the range of the dark force, in which case dark-force effects between clumps would effectively switch off. Quantifying these effects requires a dedicated nonlinear treatment beyond the linear analysis used here. Nevertheless, it is unlikely that all of the dark sector ends up in compact objects. An $\mathcal{O}(1)$ fraction of the DM could still remain smooth and continue to participate in linear perturbation dynamics as captured by our analysis. Moreover, nonlinear clumping would imprint additional small-k power, which is readily constrained by the cosmic-perturbation probes discussed here.

VI. REPULSIVE DARK FORCES

The simplest way to make the DM interact repulsively is to give it a net charge under a dark photon. We shall see that the presence of a net charge density automatically leads to an energy component that scales as a stiff fluid $m_A^2(A^0)^2 \propto (J^0)^2 \propto a^{-6}$, since $J^0 \propto n_\chi \propto a^{-3}$. Even if this component makes up a tiny component of the Universe today, it would rapidly become increasingly important as we go back in time, potentially spoiling, e.g., BBN's

predictions. It can already be seen that cosmology will place strong constraints on such models. We show in this section that these simple considerations yield limits on repulsive dark forces that are even stronger than the limits we have obtained for the attractive Yukawa case.

The Lagrangian that we use to model a repulsive force acting on DM is

$$\mathcal{L} \supset -\frac{1}{4}V_{\mu\nu}V^{\mu\nu} + \frac{1}{2}m_V^2 V_\mu V^\mu - V_\mu J^\mu_{(\chi)}. \tag{28}$$

We assume that the DM χ is net charged under a dark $U(1)$ gauge boson V^μ whose mass is m_V and field strength tensor is $V^{\mu\nu}$. We treat χ as a nonrelativistic fluid whose dark-charge current is given by

$$J^\mu_{(\chi)} = \frac{q_\chi \rho_\chi}{m_\chi} u^\mu_{(\chi)}, \tag{29}$$

where ρ_χ, q_χ, and $u^\mu_{(\chi)}$ are, respectively, χ's mass density, χ particle's dark $U(1)$ charge and χ fluid's four-velocity. The Proca equation reads

$$\nabla_\nu V^{\nu\mu} + m_V^2 V^\mu = J^\mu_{(\chi)}. \tag{30}$$

We assume that the Universe is homogeneous and isotropic at the background level, and this implies $V^i = 0$ and $\nabla_\nu V^{\nu\mu} = 0$ (since spatial derivatives must be zero). The Proca equation then tells us that a background dark $U(1)$ charge density $J^0_{(\chi)}$ sources a background scalar potential

$$V^0 = J^0_{(\chi)}/m_V^2. \tag{31}$$

This highlights an important difference between the vector and scalar cases. In contrast to the heuristically similar "quasistatic" solution $\phi \approx \phi_* = g n_\chi / m_\phi^2$ for the scalar, which requires ignoring the effects of time derivatives on the dynamics, the relation $V^0 = J^0_{(\chi)}/m_V^2$ in the vector case applies in general for a homogeneous and isotropic universe. In other words, the scalar ϕ's "inertia" (i.e., second time-derivative term in the equation of motion) prevents it from perfectly tracking its (in general changing) potential minimum. In contrast, the V^0 component of a homogeneous vector field background does not have a kinetic term in its equation of motion, and it thereby "locks" onto the instantaneous value of the charge density $J^0_{(\chi)}$.

The scalar potential V^0 carries an energy density

$$\frac{1}{2}m_V^2 (V^0)^2 = \frac{1}{2}\left(\frac{J^0_{(\chi)}}{m_V}\right)^2 \propto a^{-6}. \tag{32}$$

Its scaling as a^{-6} follows from DM particle-number conservation, which ensures that the charge density scales as $J^0_{(\chi)} \propto a^{-3}$. Knowing this scaling, we compute the energy density of the background scalar potential V^0 at an arbitrary redshift z as

$$\frac{1}{2}m_V^2 (V^0)^2 \sim \frac{q^2 \rho_{\chi,0}^2 z^6}{m_\chi^2 m_V^2} \sim 3\times 10^5 \text{ MeV}^4 \, \alpha_{\chi\chi}^{\text{repulsive}} \left(\frac{m_V^{-1}}{\text{pc}}\right)^2 \left(\frac{z}{10^{10}}\right)^6, \tag{33}$$

where we have defined the coupling strength relative to gravity as

$$\alpha_{\chi\chi}^{\text{repulsive}} \equiv \frac{q^2}{4\pi G m_\chi^2}, \tag{34}$$

and set $\rho_{\chi,0} \approx 11 \text{ meV}^4$ [33]. Requiring $m_V^2 (V^0)^2/2$ at redshift $z_{\text{BBN}} \sim \text{MeV}/(0.1 \text{ meV}) \sim 10^{10}$ be much less than MeV^4 in order to avoid spoiling nucleosynthesis, we find a rough upper bound on $\alpha_{\chi\chi}^{\text{repulsive}}$

$$\alpha_{\chi\chi}^{\text{repulsive}} \lesssim 4\times 10^{-6}\left(\frac{m_V^{-1}}{\text{pc}}\right)^{-2}, \qquad \text{(BBN)} \tag{35}$$

which is far more stringent than the limits we found in the main text on attractive Yukawa forces. Furthermore, requiring $m_V^2 (V^0)^2/2 \lesssim \text{ eV}^4$ at $z \sim 10^3$ yields

$$\alpha_{\chi\chi}^{\text{repulsive}} \lesssim 3\times 10^{12}\left(\frac{m_V^{-1}}{\text{pc}}\right)^{-2}. \qquad \text{(CMB)} \tag{36}$$

If the above is not satisfied, then the Universe would be stiff-fluid dominated at what is supposed to be the epoch of matter-radiation equality. Such a cosmology is inconsistent with CMB data and is therefore ruled out.

The current best astrophysical limit on $\alpha_{\chi\chi}^{\text{repulsive}}$ comes from the existence of ultrafaint dwarf galaxies [11] and is approximately given by $\alpha_{\chi\chi}^{\text{repulsive}} \lesssim 1 + (m_V^{-1}/\text{kpc})^{-2}$. Dwarf galaxies would not be stable if the strength of repulsive dark forces exceeds this limit. We summarize the crude limits on repulsive dark forces discussed in this section in Fig. 8. We see that the bound from BBN is much stronger than the astrophysical bound and also stronger than the bounds on an attractive force. We should note that the BBN bound is somewhat model dependent. In particular, we would not have to satisfy the BBN bound if the dark matter does not exist at the time of BBN and is only created later in the Universe. We do not need to have DM at the time of BBN, so in principle, the BBN bound does not need to be satisfied; however, one would need to carefully build a DM model specifically to avoid it. Generic models for such DM would generally be subject to this BBN bound.

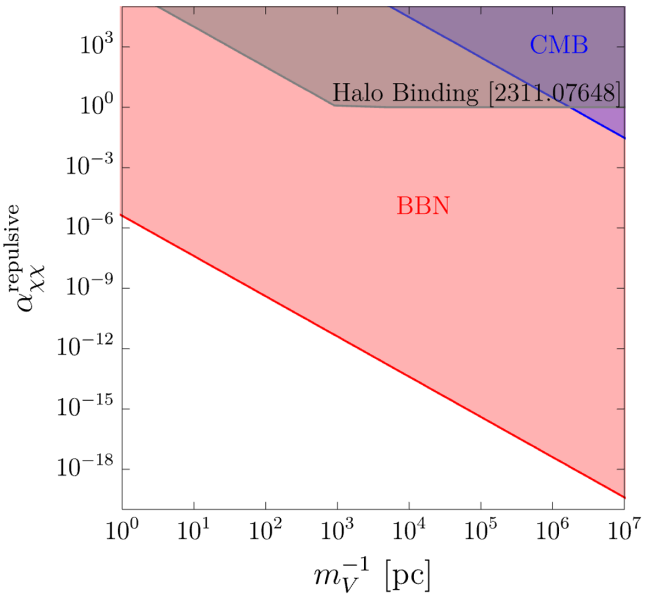


FIG. 8. Limits on repulsive dark forces. In the red (blue) region, the energy density of the Universe is dominated by the scalar potential's contribution $m_V^2(V^0)^2/2 \propto z^{-6}$ at redshifts $z \sim 10^{10}$ ($z \sim 10^3$), thus spoiling BBN (CMB). The region labeled "Halo Binding" is ruled out by the observation of ultrafaint dwarf galaxies because the repulsive forces would overcome gravity and destabilize the dwarf galaxies [11].

VII. DISCUSSION AND CONCLUSION

We have explored the cosmological implications of an attractive long-range force acting between DM particles that is stronger than gravity. We considered a model where the DM is fermionic and the dark force is mediated by a light scalar coupled to the fermionic DM. We found that such a dark sector can behave, at the background level, momentarily as dark radiation and dark energy, in stark deviation from cold dark matter, which is known to simultaneously fit a wide range of cosmological observables very well. At the perturbation level, dark forces can cause orders of magnitude of extra growth in small-scale perturbations, enough to be observable in existing probes of small-scale power spectra, including but not limited to Lyman-α forest and ultrafaint dwarf heating. Our analysis places some of the strongest constraints on the strength $\alpha_{\chi\chi}$ of dark forces with ranges $\lambda_\phi \lesssim 1$ Mpc, at least for the Yukawa model under consideration.

The Yukawa model is, of course, not the only possible dark-force model. Nevertheless, it illustrates a range of possible early-Universe dynamics one might find in a generic dark-force model, as well as the observable imprints it may leave. As another example, we briefly considered a model of repulsive dark forces, where the DM is net charged under a dark photon. By requiring consistency with BBN, we placed a limit on the strength of repulsive dark forces that is even more stringent than what we have obtained for the attractive case. It would be interesting to consider other dark-force models [49,50]. While we focus on inter-DM forces in this work, the mechanisms we studied here could be relevant for any ultralight scalars coupled to fermions, e.g., neutrinos [51–55].

The newly obtained limits on long-range inter-DM forces, combined with existing limits on, say, internucleon fifth forces, can be translated into new limits on DM-nucleon forces. In known models of long-range forces, the DM-DM coupling $\alpha_{\chi\chi}$, nucleon-nucleon coupling α_{nn}, and DM-nucleon coupling $\alpha_{\chi n}$ typically satisfy the geometric mean relation $\alpha_{\chi n} = \sqrt{\alpha_{\chi\chi}\alpha_{nn}}$. At $\lambda_\phi = 1$ AU, for instance, in the attractive case, our perturbation analysis constrains $\alpha_{\chi\chi} \lesssim 10^8$. This, in combination with the current limit on internucleon forces from MICROSCOPE, $\alpha_{nn} \lesssim 2 \times 10^{-12}$, implies that $\alpha_{\chi n} \lesssim 10^{-2}$ at AU scales. Since the current strongest limits on $\alpha_{\chi n}$ [56] are obtained with a similar method of taking the geometric mean of the limits on $\alpha_{\chi\chi}$ and on α_{nn}, our limits on $\alpha_{\chi\chi}$ also imply the strongest limits on $\alpha_{\chi n}$, for a wide range of λ_ϕ. This will have important implications, as DM-nucleon interactions far stronger than the limit we have obtained were often assumed in astrophysical contexts up until now.

We have seen that in the Yukawa model under consideration, the dark sector goes through phases with equation of state $w_{\text{dark}} \approx 1/3$ and $w_{\text{dark}} \approx -1$ when ϕ is oscillating around the linear-potential minimum ϕ_0. In certain parameter space, around $\alpha_{\chi\chi} \sim 1$ and $\lambda_\phi \sim H_{\text{eq}}^{-1} \approx 30$ kpc, the $w_{\text{dark}} \approx 1/3$ phase is nearly nonexistent, while the $w_{\text{dark}} \approx -1$ phase lasts a long time and occurs around matter-radiation equality. Since this behavior is similar to the early dark energy (EDE) scenario [57,58], it might be interesting to revisit this parameter space with a more detailed analysis. Unlike other EDE scenarios, which consider an extra dark component in an otherwise ΛCDM cosmology, we have a dark energy component that becomes the dark matter. Our scenario is thus qualitatively different from the standard EDE as it changes the evolution of the DM and does not merely add an extra energy component to ΛCDM.

At the perturbation level, the effects of long-range dark forces are not limited to enhanced matter power on small scales. They may also have important implications for the abundances and properties of DM halos. Dark forces may cause structures to collapse earlier than in the standard scenario, forming so-called ultracompact minihalos. Other considerations that we do not include here, such as NANOGrav's nondetection of scalar-induced GWs and microlensing limits on primordial black holes and other compact structures, may strengthen the limits on dark forces further [46,59]. These probes may access yet higher wave numbers and potentially extend the dark-force limits to even shorter λ_ϕ.

There are possible caveats to our results. We have assumed in our analysis that the primordial power spectrum

is scale invariant. While it is possible that the primordial power spectrum has a nontrivial running of scalar spectral index at higher wave numbers than probed by the CMB, it would require a tuning for the nonstandard primordial spectrum to exactly cancel out the scale-dependent growth enhancement due to dark forces. However, effects that lead to the dissipation of high-k perturbation power could serve as potential caveats, as they may wash away any prior enhancement caused by dark-forces. For instance, the χ particles may have non-negligible temperature by virtue of their production mechanism. In that case, they may cause enough diffusion damping to undo dark-force clustering effects on scales smaller than their free-streaming length. Apart from that, we reiterate that the scalar field ϕ's bare mass and bare quartic coupling considered in this work are generally smaller than their expected loop corrections. It would be interesting to address the model-building issues surrounding such fine-tuned scenarios and to consider the effects of larger quartic couplings in the future.

Earlier works prior to ours that study the cosmological implications of dark forces tend to focus on the small coupling $\alpha_{\chi\chi}$ regime, where the effects of the dark force amount to a small correction to the dominant gravitational force. In this regime, it is essential to perform a detailed likelihood analysis, as done in, e.g., [22], to disentangle the effects of dark forces from the degeneracies of cosmological parameters. Although these analyses are highly valuable, the practically accessible region of the $(\alpha_{\chi\chi}, \lambda_\phi)$ parameter space is limited to suitably small couplings, where the perturbative expansion in $\alpha_{\chi\chi}$ remains valid, and to $\lambda_\phi \gtrsim H_{\rm eq}^{-1}$, beyond which the computational cost of the analysis becomes prohibitive. Our analysis is complementary to such detailed numerical analyses as it covers the large coupling regime, $\alpha_{\chi\chi} \gtrsim 1$, and is far less computationally limited at smaller λ_ϕ values. In this regime of the parameter space, the cosmology is in such stark deviation from ΛCDM that a detailed numerical analysis would not be necessary to rule it out. Thus, our analysis offers a relatively economical way to cover this otherwise numerically and technically challenging regime.

ACKNOWLEDGMENTS

We thank Tom Abel, Cyril Creque-Sarbinowski, Sten Delos, Savas Dimopoulos, Dan Green, Junwu Huang, David E. Kaplan, Xuheng Luo, Joel Meyers, Oliver Philcox, Surjeet Rajendran, Tristan Smith, Ken Van Tilburg, and Zachary Weiner for useful discussions. This work was supported in part by NSF Grants No. PHY-2310429, No. PHY-2515007, Simons Investigator Award No. 824870, the Gordon and Betty Moore Foundation Grant No. GBMF7946, The University of Delaware Research Foundation and the John Templeton Foundation Award No. 63595.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: SCALAR EVOLUTION DETAILS

It is useful to define $x \equiv a/a_{\rm eq}$ and write the Klein-Gordon equation and the Hubble rate in terms of x

$$0 = \frac{d^2\phi}{dx^2} + \left(\frac{1}{H}\frac{dH}{dx} + \frac{4}{x}\right)\frac{d\phi}{dx} + \frac{m_\phi^2}{H^2x^2}[\phi - \phi_* \mathrm{sign}(\phi_0 - \phi)], \tag{A1}$$

$$H^2 = \frac{H_{\rm eq}^2}{2}\left(\frac{1-f_\chi}{x^3} + \frac{1}{x^4}\right) + \frac{H_{\rm eq}^2}{2}\left(\frac{f_\chi|1-\phi/\phi_0|}{x^3}\right) + \frac{\dot\phi^2}{2} + \frac{m_\phi^2\phi^2}{2}, \tag{A2}$$

where

$$f_\chi = \frac{\Omega_{\chi,0}}{\Omega_{\rm m,0}} \tag{A3}$$

is the fraction of the present-day matter density in χ. It will be set by the requirement that $\rho_{\rm DM,0}/\rho_{\rm m,0} \approx 0.85$. We start by assuming that the Hubble rate H is the same as in ΛCDM to the zeroth order. Later, we will check the validity of this assumption *a posteriori*. During the course of its evolution, the envelope of the background ϕ can track or oscillate around the following special solutions:

$$\phi_{\rm rise} = \frac{3\alpha_{\chi\chi}f_\chi}{4}\phi_0 x, \tag{A4}$$

$$\phi_0 = \sqrt{\frac{2}{\alpha_{\chi\chi}}}M_{\rm pl}, \tag{A5}$$

$$\phi_* = \frac{3\alpha_{\chi\chi}f_\chi}{4}\left(\frac{H_{\rm eq}}{m_\phi}\right)^2\frac{\phi_0}{x^3}, \tag{A6}$$

$$\phi_{\rm DM}^{\rm rms} = \sqrt{\frac{3\alpha_{\chi\chi}f_\phi}{4}}\left(\frac{H_{\rm eq}}{m_\phi}\right)\frac{\phi_0}{x^{3/2}}, \tag{A7}$$

where $M_{\rm pl} \equiv 1/\sqrt{8\pi G} = 2.4 \times 10^{18}$ GeV is the reduced Planck mass; $\phi_{\rm rise}$ is an attractor during RD when $\phi \ll \phi_*, \phi_0$; $\phi_0 \equiv m_\chi/g$ is the linear finite-density minimum; $\phi_* \equiv g n_\chi/m_\phi^2$ is the quadratic finite-density minimum; $\phi_{\rm DM}$ is the rms amplitude of a dark-matter-like solution normalized to the fraction of the total dark matter it comprises today, $f_\phi = \rho_{\phi,0}/\rho_{\rm DM,0}$.

There are two types of evolution histories that the envelope of ϕ can take:

(i) *Large mass-change case* ($\phi_{\rm rise} \to \phi_0 \to \phi_* \to \phi_{\rm DM}$):
If $3\alpha_{\chi\chi} f_\chi/4 \gtrsim (m_\phi/H_{\rm eq})^{1/2}$, then ϕ does reach ϕ_0 during its evolution, which goes schematically as follows: $\phi_{\rm rise} \to \phi_0 \to \phi_* \to \phi_{\rm DM}$. First ϕ approaches $\phi_{\rm rise} \propto x$ and proceeds to track it, and then ϕ tracks $\phi_0 =$ constant on average but oscillates with a bottom-potential velocity of $|\dot\phi_{\rm lin}| \approx H(x)\phi_0 \propto x^{-2}$, then ϕ tracks $\phi_* \propto x^{-3}$ adiabatically with $\dot\phi \approx \dot\phi_* \propto x^{-5}$ for a while, before redshifting like matter with a root-mean-square (rms) field-space velocity $\dot\phi^{\rm rms}_{\rm DM} \sim |\dot\phi_*|_{\phi_0=\phi_*}(x/x_{\phi_0=\phi_*})^{-3/2}$. In this case, the points where ϕ switches behavior are

$$x_{\phi_{\rm rise}=\phi_0} \approx \left(\frac{3\alpha_{\chi\chi} f_\chi}{4}\right)^{-1}, \tag{A8}$$

$$x_{\phi_0=\phi_*} \approx \left(\frac{3\alpha_{\chi\chi} f_\chi}{4}\right)^{1/3}\left(\frac{H_{\rm eq}}{m_\phi}\right)^{2/3}, \tag{A9}$$

$$x_{\phi_*=\phi^{\rm rms}_{\rm DM}} \approx \left(\frac{3\alpha_{\chi\chi} f_\chi}{4}\right)^{7/9}\left(\frac{H_{\rm eq}}{m_\phi}\right)^{8/9}. \tag{A10}$$

(ii) *Small mass-change case* ($\phi_{\rm rise} \to \phi_{\rm DM}$):
If $3\alpha_{\chi\chi} f_\chi/4 \lesssim (m_\phi/H_{\rm eq})^{1/2}$, then ϕ does not reach ϕ_0 during its evolution, which goes schematically as follows: $\phi_{\rm rise} \to \phi_{\rm DM}$. First ϕ approaches $\phi_{\rm rise} \propto x$ and tracks it, without ever reaching $\phi_0 =$ constant, and then it crosses ϕ_* at

$$x_{\phi_{\rm rise}=\phi_*} \approx \left(\frac{H_{\rm eq}}{m_\phi}\right)^{1/2}, \tag{A11}$$

and from then on, redshifts like dark matter with an rms field-space velocity $|\dot\phi^{\rm rms}_{\rm DM}| \approx |\dot\phi_*|_{\phi_{\rm rise}=\phi_*}(x/x_{\phi_{\rm rise}=\phi_*})^{-3/2}$.

Note that the large (small) mass-change case corresponds to $x_{\phi_{\rm rise}=\phi_0} \lesssim x_{\phi_{\rm rise}=\phi_*}$ ($x_{\phi_{\rm rise}=\phi_0} \gtrsim x_{\phi_{\rm rise}=\phi_*}$). Below, we derive the piecewise analytical formulas mentioned above.

1. $\phi_{\rm rise}$: initial attractor

Suppose we start deep in RD at $x = x_i \ll 1$ when the Hubble rate is $H(x_i) \approx H_{\rm eq}/\sqrt{2}x_i^2$, the mediator field is ϕ_i, and its derivative is $\dot\phi_i$. It turns out that for a wide range of initial conditions and parameter space, the solution $\phi_{\rm rise} = (3\alpha_{\chi\chi} f_\chi/4)\phi_0 x$ is a dynamical attractor for the initial evolution of ϕ during RD. When $\phi \ll \phi_*, \phi_0$, the equation of motion for ϕ reduces to

$$\frac{d}{dx}\left(x^2\frac{d\phi}{dx}\right) = \frac{3\alpha_{\chi\chi} f_\chi}{2}\phi_0 x, \tag{A12}$$

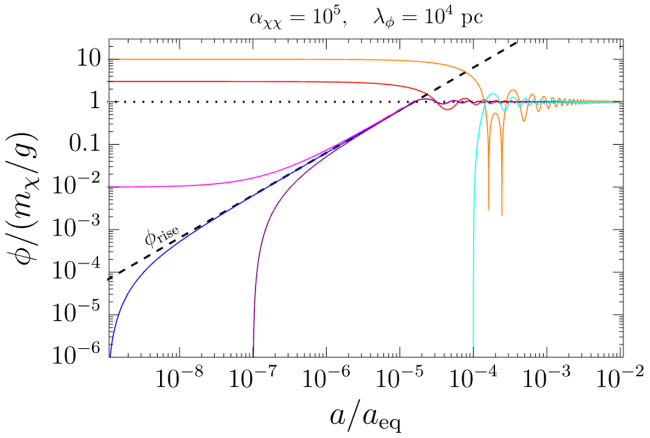


FIG. 9. Dynamical attractors of the background evolution of ϕ. The solid lines are the evolutions of ϕ obtained by solving Eq. (A1) with different initial conditions for ϕ (and with $\dot\phi = 0$ initially). The plot shows that the $\phi_{\rm rise}$ solution defined in Eq. (A4) (dashed line) and also the linear finite-density minimum $\phi_0 = m_\chi/g$ (dotted line) are attractors for a wide range of initial conditions. For the assumed values of λ_ϕ and $\alpha_{\chi\chi}$, the $\phi_* = gn_\chi/m_\phi^2$ is far greater than $\phi_0 = m_\chi/g$ and thus lies outside of the range of the plot.

whose general solution is

$$\phi = \phi_i + \frac{\dot\phi_i}{H(x_i)}\frac{x - x_i}{x} + \frac{3\alpha_{\chi\chi} f_\chi}{4}\phi_0\frac{(x - x_i)^2}{x}. \tag{A13}$$

Here, the first two terms are the initial field position and the field excursion due to the initial field velocity, while the last term captures the rolling of the field down the linear, finite-χ-density potential. This last term asymptotes to $\phi_{\rm rise}$ when $x \gg x_{\rm rise}$. Since in this limit $\phi_{\rm rise} \propto x$, it grows relative to the initial-condition terms, and will eventually becomes dominant, at

$$x \sim x_{\rm rise} \equiv \max\left[x_i, \frac{\phi_i + \dot\phi_i/H(x_i)}{(3\alpha_{\chi\chi} f_\chi/4)\phi_0}\right]. \tag{A14}$$

In other words, $\phi_{\rm rise}$ is an attractor. Figure 9 illustrates ϕ's tendency to approach $\phi_{\rm rise}$ in cases with $\dot\phi_i = 0$, for different values of ϕ_i. If $|\phi_i| \ll \phi_{\rm rise}$, then ϕ quickly increases toward $\phi_{\rm rise}$ in a few e folds and proceeds to track it. Next, if $\phi_{\rm rise} \ll |\phi_i| < \phi_0$, then ϕ remains frozen at ϕ_i until $\phi_{\rm rise}$ crosses it, whereupon ϕ switches to tracking $\phi_{\rm rise}$. Further, if $|\phi_i| > \phi_0$,[6] then ϕ would similarly stay frozen at ϕ_i until $\phi_{\rm rise}$ crosses it, however this time after the crossing ϕ switches to oscillating around ϕ_0 with an initial-condition-dependent amplitude instead of tracking $\phi_{\rm rise}$. Finally, any initial field velocity $\dot\phi_i$ will slow down due to

[6]When $|\phi_i| > \phi_0$, the ϕ field sees ϕ_0 but not ϕ_* as the instantaneous effective-potential minimum; see Fig. 1.

Hubble friction to approach the terminal slow rolling, which is captured by the third term of Eq. (A13). At $x \gg x_i$, accounting for a nonzero $\dot{\phi}_i$ and amounts to adding a constant field excursion $\approx\dot{\phi}_i/H(x_i)$ to the initial field value of ϕ.

Overall, ϕ would track $\phi_{\rm rise}$ at some point during its evolution if its initial value ϕ_i and initial velocity $\dot{\phi}_i$ are sufficiently small. If ϕ never tracked $\phi_{\rm rise}$, the late-time evolution of ϕ would be initial condition dependent and there would be more branching of cases. However, in those cases, the dark sector will carry more energy density than in the case where ϕ had tracked $\phi_{\rm rise}$, and this translates to larger oscillations around ϕ_0 and ϕ_* at late times. The latter suggests that deviating from the $\phi_{\rm rise}$-tracking case would lead to background evolutions that are more catastrophic and so easier to rule out. Thus, for the purpose of placing limits on dark forces, it is both generic and conservative to assume the attractor solution $\phi_{\rm rise}$ as an initial condition. We will also assume, for simplicity, that the initial conditions, ϕ_i and $\dot{\phi}_i$, are set (by an unspecified earlier dynamics) sufficiently early that the pre-$\phi_{\rm rise}$ evolution does not have observable consequences.

Once ϕ reaches $\phi_{\rm rise}$, it will continue tracking $\phi_{\rm rise}$ until whichever of the following comes first: radiation domination ending at $x = 1$ or $\phi_{\rm rise}$ crossing the instantaneous minimum of the effective potential of ϕ, namely either ϕ_0 or ϕ_*, whereupon ϕ switches into oscillating around ϕ_0 or ϕ_*, respectively. If $\lambda_\phi \gtrsim H_{\rm eq}^{-1}$, the $\phi_{\rm rise}$ solution continues into MD, where it asymptotes, at $x \gg 1$, to $\approx\alpha_{\chi\chi} f_\chi \phi_0 \ln x$. So, when $\lambda_\phi \gtrsim H_{\rm eq}^{-1}$, ϕ never reaches ϕ_0 if $\alpha_{\chi\chi} f_\chi \ln(10^3) \ll 1$.

2. $\phi_{\rm lin}$: oscillations around ϕ_0

The discussion in this subsection is relevant only for the large mass-change case, namely in the parameter space where $\phi_{\rm rise}$ crosses ϕ_0 before it crosses ϕ_*.

After the $\phi_{\rm rise} = \phi_0$ crossing point, in the absence of ϕ damping other than Hubble, ϕ oscillates around ϕ_0 with a bottom-potential velocity of

$$|\dot{\phi}_{\rm lin}| \approx \dot{\phi}_{\rm rise}|_{\phi_{\rm rise}=\phi_0} \approx \phi_0 H \propto x^{-2}. \tag{A15}$$

Note that the energy density associated to this oscillation scales like radiation, $|\dot{\phi}_{\rm lin}|^2/2 \propto x^{-4}$. This kinetic energy becomes negligible compared to the dark energy associated to the mass term $m_\phi^2\phi_0^2/2$ when the Hubble H goes below the mass m_ϕ. Given a velocity amplitude $\dot{\phi}_{\rm lin}$, the amplitude and period of $\phi - m_\chi/g$ oscillation under the linear potential can be found by simple kinematics

$$|\phi_{\rm lin} - \phi_0| = \frac{|\dot{\phi}_{\rm lin}|^2}{2gn_\chi} \approx \frac{\phi_0}{4}\left(\frac{x_{\phi_{\rm rise}=\phi_0}}{x}\right), \tag{A16}$$

$$T_{\rm lin} = 2 \times \frac{2|\dot{\phi}_{\rm lin}|}{gn_\chi} = 2\left(\frac{x_{\phi_{\rm rise}=\phi_0}}{x}\right)H^{-1} \propto x. \tag{A17}$$

Initially, when ϕ first crosses ϕ_0, a half oscillation takes a Hubble time. Subsequently, each oscillation takes a progressively lesser fraction of the Hubble time, but a longer time to complete.

One might worry that χ can become relativistic when its effective mass $M_\chi = |m_\chi - g\phi|$ vanishes near $\phi = \phi_0 \equiv m_\chi/g$. The relevant question is how much time ϕ spends in the neighborhood of ϕ_0. The χ particles are relativistic only where M_χ drops below their typical momentum p_χ, which is set by the Fermi momentum p_F in the degenerate case (or by the thermal momentum p_T in a thermal production scenario). Because we assume χ to be extremely cold, p_χ is tiny, so the relativistic condition $M_\chi \lesssim p_\chi$ is met only within very narrow field-space intervals $\Delta\phi \sim p_\chi/g$ around each $\phi = \phi_0$ crossing. Note that $|\phi_{\rm lin} - m_\chi/g|$ has the same scaling as that of the momenta of χ particles, so the typical velocity of χ, $v_\chi = p_\chi/|m_\chi - g\phi|$, does not change with scale factor. In the $T_\chi \approx 0$ case, χ is nonrelativistic *on average* if

$$\frac{p_F}{g|\phi_{\rm lin} - \phi_0|} \sim 2 \times 10^{-18}\alpha_{\chi\chi}^{1/2}\left(\frac{m_\chi}{\rm GeV}\right)^{-1/3} \lesssim 1; \tag{A18}$$

i.e., χ is nonrelativistic on average for a wide range of $\alpha_{\chi\chi}$ and m_χ. In another appendix section, we show that the χ particles can be relativistic briefly when $|m_\chi - g\phi| \lesssim p_F \sim n_\chi^{1/3}$, but ϕ always swings far away from this region, and outside this region, the χ particles are very nonrelativistic. In Eq. (A18), we find that the fraction of each oscillation spent with $M_\chi \lesssim p_\chi$ is extremely small in the parameter space we consider. Consequently χ is highly nonrelativistic during essentially all of its evolution. Furthermore, nonadiabatic particle production can occur whenever ϕ crosses m_χ/g, but this is almost always negligible, and m_χ dependent; see Appendix D 1 b.

3. $\phi_{\rm small}$: oscillations around ϕ_*

When the energy density in ϕ is sufficiently small, the minimum of $V_{\rm eff}$ is at ϕ_*. The transition to this quadratic finite-density minimum occurs approximately when ϕ crosses ϕ_* for the first time. Any relative deviation or velocity of ϕ from ϕ_* at that time, as a result of its previous dynamics, will contribute to later oscillations of ϕ around ϕ_*, with an amplitude that scales as $x^{-3/2}$. In general, the initial amplitude of this oscillation is set roughly by the relative speed $|\dot{\phi} - \dot{\phi}_*|$ around the moment when the oscillation around ϕ_* starts. In the large mass-change case

($\phi_{\rm rise} \to \phi_0 \to \phi_* \to \phi_{\rm DM}$), this occurs when ϕ_* crosses ϕ while it is oscillating around ϕ_0 with a bottom-potential velocity $|\dot{\phi}_{\rm lin}| \approx \phi_0 H$; in the small mass-change case ($\phi_{\rm rise} \to \phi_{\rm DM}$), this occurs when ϕ_* crosses $\phi \approx \phi_{\rm rise}$ that is rising at the rate $\dot{\phi}_{\rm rise} \approx \phi_{\rm rise} H$. Since $\dot{\phi}_* \approx 3H\phi_*$, we have $|\dot{\phi} - \dot{\phi}_*| \sim H\phi_*$ at the respective crossing points, in both cases. Knowing this, it can be shown that the envelope of ϕ is given by

$$\phi_{\rm small} = \phi_* + \phi_{\rm DM}^{\rm rms}(f_\phi), \tag{A19}$$

where

$$\frac{f_\phi}{f_\chi} \sim \begin{cases} \frac{9}{4}\left(\frac{3\alpha_{\chi\chi} f_\chi}{4}\right)^{-4/3}\left(\frac{H_{\rm eq}}{m_\phi}\right)^{-2/3}, & \text{large mass-change} \\ \frac{9}{4}\left(\frac{3\alpha_{\chi\chi} f_\chi}{4}\right)\left(\frac{H_{\rm eq}}{m_\phi}\right)^{1/2}. & \text{small mass-change} \end{cases} \tag{A20}$$

This is maximized at the boundary separating the large mass-change and small mass-change cases, namely $(3\alpha_{\chi\chi} f_\chi/4)(H_{\rm eq}/m_\phi)^{1/2} \sim 1$, where $f_\phi/f_\chi \sim 9/4$; see Fig. 10 for the time evolutions of energy densities in this regime. Along that boundary, the $\phi_{\rm DM}$ solution begins when $\phi \sim \phi_0$, which is the largest value ϕ can have, modulo oscillations around ϕ_0.

To fix f_χ, we require

$$f_\chi\left(1 + \frac{f_\phi}{f_\chi}\right) = \frac{\Omega_{\rm DM,0}}{\Omega_{\rm m,0}} \approx 0.85. \tag{A21}$$

If $x_{\phi_0=\phi_*} \ll x_{\phi_*=\phi_{\rm DM}^{\rm rms}}$, the oscillation of ϕ around ϕ_* is negligible at first, i.e., $|\phi - \phi_*| \ll \phi_*$. However, since $|\phi - \phi_*| \propto x^{-3/2}$ redshifts more slowly than $\phi_* \propto x^{-3}$, eventually at $x_{\phi_*=\phi_{\rm DM}^{\rm rms}}$, the oscillation around ϕ_* takes over as the envelope of ϕ.

APPENDIX B: BBN

The success of big bang nucleosynthesis is not spoiled if at $x_{\rm BBN} \approx 10^{-6}$, we have $\Delta N_{\rm eff} \lesssim 0.5$, which corresponds to $\rho_{\rm dark}/\rho_{\rm crit} \lesssim 1\%$ then. The largest $\rho_{\rm dark}$ achievable for initial conditions that lead to ϕ tracking the attractor $\phi_{\rm rise}$ corresponds to the case where $\dot{\phi} \sim \dot{\phi}_{\rm rise} \sim H\phi$ and $\phi \sim \phi_0$, which amounts to $\Omega_{\rm dark} = \rho_{\rm dark}/\rho_{\rm crit} \lesssim (1 + m_\phi^2/H^2)/3\alpha_{\chi\chi}$. This maximum density fraction is achieved at BBN if $x_{\phi_{\rm rise}=\phi_0} \sim x_{\rm BBN}$ and $x_{\phi_{\rm rise}=\phi_*} \gtrsim x_{\rm BBN}$, which translate to $(3\alpha_{\chi\chi} f_\chi/4)^{-1} \sim 10^{-6}$ and $(H_{\rm eq}/m_\phi)^{1/2} \gtrsim 10^{-6}$. Thus, for the fiducial $f_\chi = 0.85$, we find that the dark density fraction is at most $\rho_{\rm dark}/\rho_{\rm crit} \sim 3 \times 10^{-7}$, corresponding to $\alpha_{\chi\chi} \approx 2 \times 10^6$ and $\lambda_\phi \approx 3 \times 10^{-8}$ pc. Therefore, we conclude that BBN does not place any constraint on the scenarios we consider.

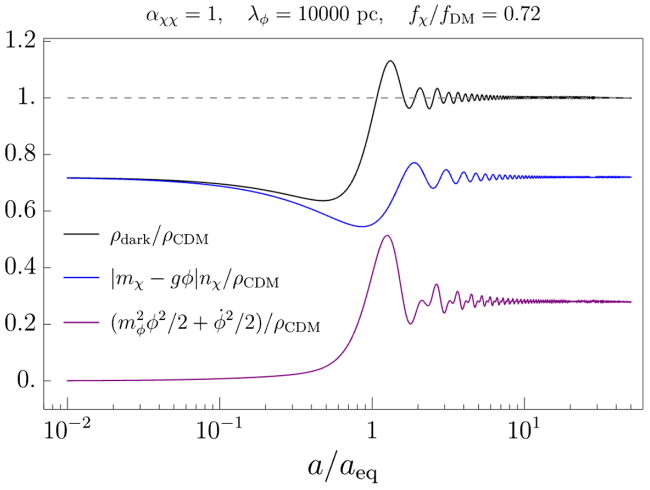


FIG. 10. Mixed DM scenario. Here, we plot as function of the scale factor χ's effective energy density $|m_\chi - g\phi|n_\chi$, ϕ's energy density $m_\phi^2\phi^2/2 + \dot{\phi}^2/2$, and their sum $\rho_{\rm dark}$, all normalized to the backward-in-time extrapolation of the CDM energy density $\rho_{\rm CDM}$. The plot shows that the initially negligible average energy density of ϕ rises considerably at some point and contributes significantly to $\rho_{\rm dark}$ at late times. Thus, the DM at late times is an $\mathcal{O}(1)$ mix of χ and ϕ, both behaving as CDM. In this plot, the parameters are chosen to hit the correct total DM abundance today.

APPENDIX C: DERIVATION OF MESZAROS-LIKE PERTURBATION EQUATION

Here, we use mostly positive metric signature, inline with most cosmic-perturbation literature. We work in the conformal Newtonian gauge [60] and write the metric as

$$ds^2 = a^2[-(1 + 2\Phi)d\tau^2 + (1 - 2\Phi)dx_i dx^i], \tag{C1}$$

where τ is the conformal time, Φ is the Newtonian potential, and we have assumed that anisotropic stress is absent. The perturbation parts of the Klein-Gordon equation for ϕ and energy-momentum conservation equations for the entire dark sector, $\nabla_\nu T^{\mu\nu}_{\rm dark} = 0$, can be written in spatial Fourier space as

$$\begin{aligned} &\delta\phi'' + 2\mathcal{H}\delta\phi' + (a^2 m_\phi^2 + k^2)\delta\phi \\ &= \frac{3\alpha_{\chi\chi}}{2}\mathcal{H}^2\Omega_\chi(\delta_\chi + 2\Phi)\,\mathrm{sign}(\phi_0 - \phi) \\ &\quad - 2a^2 m_\phi^2\Phi\phi + 4\Phi'\phi', \end{aligned} \tag{C2}$$

$$\delta'_{\rm dark} = -(1 + w_{\rm dark})(\theta_{\rm dark} - 3\Phi') - 3\mathcal{H}\left(\frac{\delta p_{\rm dark}}{\delta\rho_{\rm dark}} - w_{\rm dark}\right)\delta_{\rm dark}, \tag{C3}$$

$$\theta'_{\rm dark} = -(1-3w_{\rm dark})\mathcal{H}\theta_{\rm dark} - \frac{w'_{\rm dark}}{1+w_{\rm dark}}\theta_{\rm dark} + \frac{\delta p_{\rm dark}/\delta\rho_{\rm dark}}{1+w_{\rm dark}}k^2\delta_{\rm dark} + k^2\Phi, \tag{C4}$$

$$\delta_{\rm dark} = \frac{|m_\chi - g\phi|n_\chi\delta_\chi - a^{-2}\phi'^2\Phi + a^{-2}\phi'\delta\phi' + m_\phi^2\phi\delta\phi}{|m_\chi - g\phi|n_\chi + \frac{1}{2}a^{-2}\phi'^2 + \frac{1}{2}m_\phi^2\phi^2}, \tag{C5}$$

$$w_{\rm dark} = \frac{\frac{1}{2}a^{-2}\phi'^2 - \frac{1}{2}m_\phi^2\phi^2}{|m_\chi - g\phi|n_\chi + \frac{1}{2}a^{-2}\phi'^2 + \frac{1}{2}m_\phi^2\phi^2}, \tag{C6}$$

$$\delta\rho_{\rm dark} = |m_\chi - g\phi|n_\chi\delta_\chi - a^{-2}\phi'^2\Phi + a^{-2}\phi'\delta\phi' + m_\phi^2\phi\delta\phi, \tag{C7}$$

$$\delta p_{\rm dark} = -a^{-2}\phi'^2\Phi + a^{-2}\phi'\delta\phi' - m_\phi^2\phi\delta\phi, \tag{C8}$$

where a prime indicates a derivative with respect to the conformal time τ, $w_{\rm dark} \equiv p_{\rm dark}/\rho_{\rm dark}$, $\delta_{\rm dark} \equiv \delta\rho_{\rm dark}/\rho_{\rm dark}$, $\delta_\chi \equiv \delta n_\chi/n_\chi$, $(\ldots)' = d(\ldots)/d\tau$, and $\theta_{\rm dark} = ik_j v^j_{\rm dark}$ is the velocity divergence of the dark sector fluid.

Assuming $m_\chi \gg g\phi$ and ignoring ϕ', ϕ'', $\delta\phi'$, $\delta\phi''$, and Φ terms in the equations of motion of ϕ and $\delta\phi$, we find

$$\phi \approx \phi_* = \frac{gn_\chi}{m_\phi^2}, \tag{C9}$$

$$\delta\phi \approx \frac{a^2 g n_\chi \delta_\chi}{a^2 m_\phi^2 + k^2}, \tag{C10}$$

which, given the $m_\chi \gg g\phi$ assumption, imply $m_\chi n_\chi \gg m_\phi^2\phi^2$, $m_\chi n_\chi|\delta_\chi| \gg gn_\chi|\delta\phi|$, and

$$w_{\rm dark} \sim \frac{\phi_*}{\phi_0}, \tag{C11}$$

$$w'_{\rm dark} \sim H\frac{\phi_*}{\phi_0}, \tag{C12}$$

$$\frac{\delta p_{\rm dark}}{\delta\rho_{\rm dark}} \sim \frac{\phi_*}{\phi_0}, \tag{C13}$$

$$\delta_{\rm dark} \approx \delta_\chi. \tag{C14}$$

Thus, if $\phi_* \ll \phi_0$ (which is implied by the $m_\chi \gg g\phi$ assumption), then the dark sector behaves like matter to leading order: $w_{\rm dark} \ll 1$, $w'_{\rm dark} \ll \mathcal{H}$, and $\delta p_{\rm dark}/\delta\rho_{\rm dark} \ll 1$. The fluid perturbation equations then reduce to

$$\delta'_{\rm dark} \approx -\theta_{\rm dark} + 3\Phi', \tag{C15}$$

$$\theta'_{\rm dark} \approx -\mathcal{H}\theta_{\rm dark} + k^2\Phi + k^2\frac{\delta p_{\rm dark}}{\rho_{\rm dark}}. \tag{C16}$$

Even though $\delta p_{\rm dark}$ is small, we keep the last term of the second equation because it can be important when $k \gg \mathcal{H}$. Taking the conformal time derivative of $\delta'_{\rm dark}$ and substituting $\theta'_{\rm dark}$ equation into it, we find

$$\delta''_{\rm dark} + \mathcal{H}\delta'_{\rm dark} = 3(\Phi'' + \mathcal{H}\Phi') - k^2\Phi - k^2\frac{\delta p_{\rm dark}}{\rho_{\rm dark}}. \tag{C17}$$

The metric perturbation Φ is given by the Poisson equation

$$-k^2\Phi = 4\pi G(\rho_{\rm rad}\delta_{\rm rad} + \rho_B\delta_B + \rho_{\rm dark}\delta_{\rm dark}) \approx 4\pi G\rho_{\rm dark}\delta_{\rm dark}. \tag{C18}$$

In the last line, we neglect the radiation and baryon terms. While the baryon and photon fluids are tightly coupled, these contributions oscillate rapidly at subhorizon scales due to radiation pressure and so mostly average out. Then, using Eqs. (C13) and (C14), we simplify Eq. (C17) to the following Meszaros-like form:

$$\delta''_\chi + \mathcal{H}\delta'_\chi = 4\pi G a^2 m_\chi n_\chi \delta_\chi\left[\frac{\alpha_{\chi\chi}k^2}{a^2m_\phi^2 + k^2} + 1\right]. \tag{C19}$$

APPENDIX D: COMPLETE FINITE-DENSITY POTENTIAL OF ϕ

A finite density, the fermion χ affects the dynamics of ϕ through the term $g\langle\bar\chi\chi\rangle$ on the RHS of the Klein-Gordon equation. In the main text, we approximate this term as $g\langle\bar\chi\chi\rangle \approx gn_\chi{\rm sign}(m_\chi - g\phi)$. Here, we consider the exact expression of this term:

$$g\langle\bar\chi\chi\rangle = g\int\frac{d^3p}{(2\pi)^3}\frac{m_\chi - g\phi}{\sqrt{(m_\chi - g\phi)^2 + p^2}}f_\chi(p). \tag{D1}$$

We will clarify the regime of validity of the aforementioned approximation and briefly discuss possible consequences of going beyond this regime.

First, we derive Eq. (D1). We expand the χ field in creation and annihilation operators,

$$\chi = \int d\Pi\sum_s[a_s(\mathbf{p})u_s(\mathbf{p})e^{-ip.x} + b_s^\dagger(\mathbf{p})v_s(\mathbf{p})e^{+ip.x}], \tag{D2}$$

$$\bar\chi = \int d\Pi\sum_s[a_s^\dagger(\mathbf{p})\bar u_s(\mathbf{p})e^{+ip.x} + b_s(\mathbf{p})\bar v_s(\mathbf{p})e^{-ip.x}], \tag{D3}$$

with $d\Pi = d^3p/[(2\pi)^3\sqrt{2E_{\mathbf{p}}}]$ and $E_{\mathbf{p}} = \sqrt{(m_\chi - g\phi)^2 + |\mathbf{p}|^2}$. Multiplying the above, forming $\bar{\chi}\chi$, taking the expectation value, and then using the canonical anti-commutators $\{a_r(\mathbf{p}), a_s^\dagger(\mathbf{q})\} = (2\pi)^3\delta_{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q}) = \{b_r(\mathbf{p}), b_s^\dagger(\mathbf{q})\}$ (and zero for others) as well as the spinor normalizations $\bar{u}_r(\mathbf{p})u_s(\mathbf{p}) = 2(m_\chi - g\phi)\delta_{rs} = -\bar{v}_r(\mathbf{p})v_s(\mathbf{p})$ yields

$$\langle\bar{\chi}\chi\rangle_{\rm tot} = \int \frac{d^3p}{(2\pi)^3}\frac{m_\chi - g\phi}{E_{\mathbf{p}}} \times \sum_s [\langle a_s(\mathbf{p})a_s^\dagger(\mathbf{p})\rangle - \langle b_s^\dagger(\mathbf{p})b_s(\mathbf{p})\rangle]. \tag{D4}$$

Inserting $\langle a_r^\dagger(\mathbf{p})a_s(\mathbf{q})\rangle = (2\pi)^3\delta_{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q})f_\chi^{(s)}(p)$ and $\langle b_r(\mathbf{p})b_s^\dagger(\mathbf{q})\rangle = (2\pi)^3\delta_{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q})[1 - f_{\bar{\chi}}^{(s)}(p)]$, where the b relation uses one of the aforementioned anticommutation relations,

$$\langle\bar{\chi}\chi\rangle_{\rm tot} = \int \frac{d^3p}{(2\pi)^3}\frac{m_\chi - g\phi}{E_{\mathbf{p}}}\sum_s [f_\chi^{(s)}(p) + f_{\bar{\chi}}^{(s)}(p)) - 1]. \tag{D5}$$

The "-1" term is the (divergent) vacuum contribution. Subtracting it,

$$\langle\bar{\chi}\chi\rangle = \langle\bar{\chi}\chi\rangle_{\rm tot} - \langle\bar{\chi}\chi\rangle_{\rm vac} = \int \frac{d^3p}{(2\pi)^3}\frac{m_\chi - g\phi}{E_{\mathbf{p}}}[f_\chi(p) + f_{\bar{\chi}}(p)], \tag{D6}$$

where we have defined $f_{\chi,\bar{\chi}}(p) \equiv \sum_s f_{\chi,\bar{\chi}}^{(s)}(p)$. In the asymmetric case assumed in main text, $f_{\bar{\chi}} = 0$, recovering the expression in Eq. (D1).

1. Degenerate fermions

Assuming the fermion χ is degenerate and asymmetrically populated (with no antiparticles), such that $f_\chi(p) \approx \Theta(p_F - p)$, the above integral can be evaluated exactly, giving

$$g\langle\bar{\chi}\chi\rangle = gn_\chi {\rm sign}(m_\chi - g\phi)\left[\frac{M_\chi(\phi)\sqrt{M_\chi^2(\phi) + p_F^2}}{p_F^2} + \frac{M_\chi^3(\phi)}{p_F^3}\ln\left(\frac{-p_F + \sqrt{M_\chi^2(\phi) + p_F^2}}{M_\chi(\phi)}\right)\right], \tag{D7}$$

where $M_\chi(\phi) = |m_\chi - g\phi|$ is the effective mass of χ, and $p_F = (6\pi^2 n_\chi)^{1/3}$ is the Fermi momentum of χ, which is numerically given by

$$p_F = 3\ {\rm meV}\left(\frac{m_\chi}{\rm GeV}\right)^{-1/3} x^{-1}. \tag{D8}$$

Next, we discuss separately the nonrelativistic-fermion ($p_F \ll M_\chi(\phi)$ and relativistic-fermion ($p_F \gtrsim M_\chi(\phi)$) regimes.

The temperature of χ is not important as long as it is much colder than the Fermi energy $p_F^2/2M_\chi$.

a. Nonrelativistic regime

Expanding the exact expression of $g\langle\bar{\chi}\chi\rangle$ in both p_F and $g\phi$, regardless of the order expansion (p_F first or $g\phi$ first), we find that the leading-order corrections to the effective potential of ϕ, which can be written as $\delta V(\phi) = \delta m_\phi^2\phi^2/2 + \delta\mu\phi^3/3! + \delta\kappa\phi^4/4!$, are

$$|\delta m_\phi^2| \sim \frac{g^2 n_\chi p_F^2}{m_\chi^3}, \qquad |\delta\mu| \sim \frac{g^3 n_\chi p_F^2}{m_\chi^4}, \qquad |\delta\kappa| \sim \frac{g^4 n_\chi p_F^2}{m_\chi^5}. \tag{D9}$$

Typically, the $\propto gn_\chi$ linear-potential term used in the main text far dominates the effective potential of ϕ. In those cases, the finite-density mass correction is not important even if it were large compared to the bare mass (which is itself negligible compared to the linear potential). However, near the finite-density minimum, $\phi_* = gn_\chi/m_\phi^2$, this linear-potential term is canceled by the bare-mass term $m_\phi^2\phi \approx m_\phi^2\phi_* = gn_\chi$. Thus, the oscillation of ϕ around ϕ_* depends on the dominant mass term (term $\propto\phi$ in the equation of motion and $\propto\phi^2$ in the effective potential). During that evolution phase, the mass correction is unimportant if

$$\frac{|\delta m_\phi^2|}{m_\phi^2} \sim \frac{\phi_*}{\phi_0}\left(\frac{p_F}{m_\chi}\right)^2 \lesssim 1. \tag{D10}$$

The earliest time when the finite-density mass-squared correction δm_ϕ^2 might be important is when ϕ crosses ϕ_* for the first time, namely at $x = x_{\phi_0=\phi_*}$ in the large mass-change case and at $x = x_{\phi_{\rm rise}=\phi_*}$ in the small mass-change case, where $\phi_*/\phi_0 \sim 1$ and $\phi_*/\phi_0 \sim (3\alpha_{\chi\chi}/4)(H_{\rm eq}/m_\phi)^{1/2}$, respectively. Unless m_χ is extremely light (close to the Tremaine-Gunn lower bound of $m_\chi = 100$ eV), the parameter space where the δm_ϕ^2 is important lies in regimes with very small λ_ϕ, where our limits do not apply.

Anharmonic terms, in particular the cubic term $\delta\mu\phi^3/3$, whose coefficient $\delta\mu$ is lower order in g, may excite modes of ϕ perturbation around the ϕ background, $\delta\phi(x) \equiv \phi(x) - \bar{\phi}(t)$, through narrow parametric resonance. Assuming $\delta m_\phi^2 \ll m_\phi^2$, the equation of motion for the spatially Fourier-transformed mode $\delta\phi_k$ reads

$$\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \left[\frac{k^2}{a^2} + m_\phi^2 + \delta\mu\bar{\phi}(t)\right]\delta\phi_k = 0. \quad \text{(D11)}$$

Mapping this into the Mathieu equation, we find that the first resonance band exists and has the least stringent instability condition. That said, the first band only populates nonrelativistic modes, which would rejoin the background after a few e folds of redshifting. Thus, we focus on the second band, which populates relativistic modes with $k_{\rm res} \sim m_\phi$, $\delta k_{\rm res} \sim m_\phi\sqrt{q}$, $\Gamma^{(2)}_{\rm Floquet} \sim m_\phi q^2$, $\delta t^{(2)}_{\rm res} \sim \sqrt{q}H^{-1}$, and $q \sim \delta\mu\bar{\phi}/m_\phi^2$. The earliest time when $\bar{\phi}$ behaves sinusoidally is, again, at $x = x_{\phi_0=\phi_*}$ in the large mass-change case and at $x = x_{\phi_{\rm rise}=\phi_*}$ in the small mass-change case. At these points, the amplitude of the sinusoidal part of ϕ is $\bar{\phi} \sim \phi_* H/m_\phi$. Requiring $\Gamma^{(2)}_{\rm Floquet}\delta t^{(2)}_{\rm res} \lesssim 10$ then amounts to

$$\left(\frac{\phi_*}{\phi_0}\right)^5\left(\frac{H}{m_\phi}\right)^{3/2}\left(\frac{p_F}{m_\chi}\right)^5 \lesssim 10. \quad \text{(D12)}$$

Repeating the same arguments for quartic coupling, in which case $q \sim \delta\kappa\bar{\phi}^2/m_\phi^2$, we find

$$\left(\frac{\phi_*}{\phi_0}\right)^{6/5}\left(\frac{m_\phi}{H}\right)\left(\frac{p_F}{m_\chi}\right)^{4/5} \lesssim 10. \quad \text{(D13)}$$

The tree-level cubic coefficient μ and quartic coefficient κ are negligible if they are smaller than the values of the finite-density corrections, $\delta\mu$ and $\delta\kappa$, at the earliest point where ϕ oscillates sinusoidally, namely at $x = x_{\phi_0=\phi_*}$ in the large mass-change case and at $x = x_{\phi_{\rm rise}=\phi_*}$ in small mass-change case.

$$\frac{\mu}{m_\phi} \lesssim \left(\frac{m_\phi}{\phi_*}\right)\left(\frac{m_\phi}{H}\right)^{3/5} \quad \text{(D14)}$$

$$\kappa \lesssim \left(\frac{m_\phi}{\phi_*}\right)^2\left(\frac{m_\phi}{H}\right)^{8/5} \quad \text{(D15)}$$

We display in Fig. 11 parameter space regions where finite-density mass correction and parametric resonance are important for representative values of m_χ.

b. Relativistic regime

In the limit $M_\chi(\phi) \ll p_F$, which amounts to expanding around $\phi = \phi_0$, we have

$$g\langle\bar{\chi}\chi\rangle = \frac{g^2 p_F^2}{4\pi^2}(\phi_0 - \phi). \quad \text{(D16)}$$

Thus, the effective potential of ϕ close to ϕ_0 is quadratic instead of linear, with an effective mass squared of $m_\phi^2 + g^2 p_F^2/4\pi^2$. The ratio between the χ-induced mass and the bare mass of ϕ is

$$\sim \frac{g n_\chi^{1/3}}{m_\phi} = 3\times 10^4 \alpha_{\chi\chi}^{1/2}\left(\frac{m_\chi}{\rm GeV}\right)^{2/3}\left(\frac{\lambda_\phi}{1\ \rm kpc}\right)x^{-1}. \quad \text{(D17)}$$

So, it is likely very significant, although this is relevant only during the brief moments when $M_\chi(\phi) \ll p_F$.

Every time ϕ crosses ϕ_0, there can be efficient nonadiabatic production of relativistic χ-$\bar{\chi}$ pairs (this is because it is easy to have $|\dot{M}_\chi| \gtrsim M_\chi^2$ when $M_\chi \to 0$), with typical

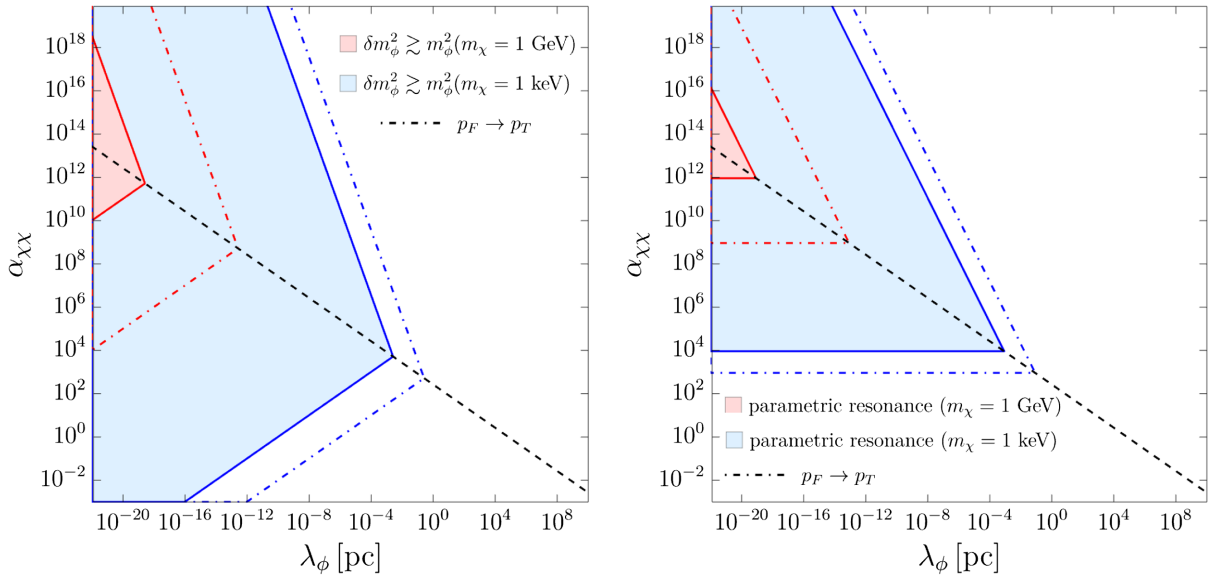


FIG. 11. Regimes where finite χ density and finite χ temperature are important. Left: regimes where the finite χ density induced mass is important when ϕ first oscillates around ϕ_*. Right: regimes where the finite-density induced cubic coupling causes ϕ to undergo parametric resonance when ϕ starts oscillating around ϕ_*. In both plots, the shared regions correspond to degenerate χ particles. The dot dashed lines are where the boundaries would be if χ has a finite-temperature such that p_T/p_F is given by Eq. (D24).

number density and momentum of

$$n_\chi^{\rm NA} \sim \frac{|\dot{M}_\chi|^{3/2}}{8\pi^3}, \qquad p_\chi^{\rm NA} \sim \frac{|\dot{M}_\chi|^{1/2}}{\pi^{1/2}}. \tag{D18}$$

Since χ is not a boson, the χ production rate does not grow exponentially. The fractional rate per Hubble time, $\dot{\rho}^{\rm NA}H^{-1}/(\dot{\phi}^2/2)$, at which energy is being drained from $\dot{\phi}^2/2$ to create $\chi-\bar{\chi}$ pairs, can thus be estimated as

$$\frac{1}{H}\frac{2p_\chi^{\rm NA}n_\chi^{\rm NA}/T_{\rm lin}}{\dot{\phi}^2/2} \sim \frac{\alpha_{\chi\chi}Gm_\chi^2}{\pi^{5/2}}\frac{x}{x_{\phi_{\rm rise}=\phi_0}} \tag{D19}$$

since $x_{\phi_{\rm rise}=\phi_0} \sim \alpha_{\chi\chi}^{-1}$. This is negligible unless m_χ is close to the Planck scale.

2. Nondegenerate fermions

The fermions under consideration would in general have a finite chemical potential but they need not be degenerate. If they were once in thermal equilibrium, their occupation number is given by the Fermi-Dirac distribution $f_\chi(p) = (e^{(\sqrt{M_\chi^2+p^2}-\mu)/T_\chi}+1)^{-1}$. Assuming that the fermion is nonrelativistic, $p^2/2M_\chi \lesssim T_\chi \ll M_\chi$, we can expand the Fermi-Dirac distribution in p/M_χ, giving $f_\chi(p) \approx (e^{(M_\chi-\mu)/T_\chi}e^{p^2/2M_\chi T_\chi}+1)^{-1}$. In this limit, the chemical potential μ is related to the total number density of χ as $n_\chi = (M_\chi T_\chi/2\pi)^{3/2}e^{-(M_\chi-\mu)/T_\chi}$. Thus, we find that $f_\chi \sim n_\chi/(p_T/4\pi)^{3/2}$ for $p \lesssim p_T \equiv \sqrt{2M_\chi T_\chi}$ and Boltzmann suppressed for $p \gg p_T$. For crude estimates, we can approximate this distribution function as a step function

$$f_\chi(p) = \frac{2\pi^2 n_\chi}{p_T^3}\Theta(p_T-p). \tag{D20}$$

Since the scalar background is homogeneous, by translational symmetry, we can deduce that it will keep momenta of the fermions χ unchanged. Thus, as $M_\chi(\phi)$ is changing, the effective temperature $T_\chi(\phi) = p_T^2/2M_\chi(\phi)$ would change to keep p_T constant. This yields

$$g\langle\bar{\chi}\chi\rangle = gn_\chi {\rm sign}(m_\chi - g\phi)\left[\frac{M_\chi(\phi)\sqrt{M_\chi^2(\phi)+p_T^2}}{p_T^2} + \frac{M_\chi^3(\phi)}{p_T^3(\phi)}\ln\left(\frac{-p_T+\sqrt{M_\chi^2(\phi)+p_T^2}}{M_\chi(\phi)}\right)\right], \tag{D21}$$

which is similar to the degenerate case. The only difference is that p_F is replaced with p_T. The condition for $\delta m_\phi^2 \lesssim m_\phi^2$ and $\Gamma^{(2)}_{\rm Floquet}\delta t^{(2)}_{\rm res} \lesssim 10$ thus becomes

$$\frac{\phi_*}{\phi_0}\left(\frac{p_T}{m_\chi}\right)^2 \lesssim 1, \tag{D22}$$

$$\left(\frac{\phi_*}{\phi_0}\right)^5\left(\frac{H}{m_\phi}\right)^{3/2}\left(\frac{p_T}{m_\chi}\right)^5 \lesssim 10. \tag{D23}$$

For instance, if the χ abundance was set by freeze-out, its typical momentum post freeze-out is given by $p_T \sim \sqrt{2M_\chi T_{\rm FO}}(T/T_{\rm FO})$, where usually $T_{\rm FO} \approx m_\chi/20$. Since, the Fermi momentum p_F scales the same way, in freeze-out scenarios, χ can only be degenerate at late times if it is degenerate at freeze-out. In this freeze-out case, the ratio p_T/p_F is approximately

$$\frac{p_T}{p_F} \sim 1\times 10^3\left(\frac{m_\chi}{\rm GeV}\right)^{1/3}, \tag{D24}$$

where we have assumed $M_\chi \approx m_\chi$ at freeze-out.

If χ was relativistic in the past, then $g\langle\bar{\chi}\chi\rangle$ evaluates to $gM_\chi(\phi)T_\chi^2/3$. If number-changing processes were efficient, ϕ could be subject to significant thermal damping [55,61,62].

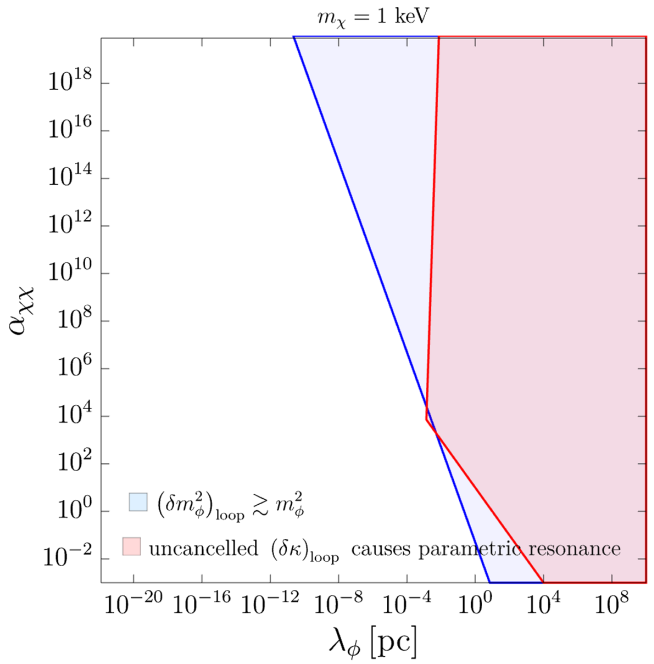


FIG. 12. Fine-tuned parameter space for $m_\chi = 1$ keV. In the blue region, the squared mass of the scalar mediator m_ϕ^2 is smaller than the expected correction to it from χ loop, $(\delta m_\phi^2)_{\rm loop}$. In the red region, the χ-loop-induced quartic coupling $(\delta\kappa)_{\rm loop}$ would cause parametric resonance, thus invalidating our analysis, unless this $(\delta\kappa)_{\rm loop}$ is somehow canceled by other contributions to the quartic coupling.

APPENDIX E: NATURALNESS

The bare mass of ϕ is greater than the loop correction from χ, $m_\phi^2 \gtrsim (\delta m_\phi^2)_{\rm loop} \sim g^2 m_\chi^2/16\pi^2$, if

$$\alpha_{\chi\chi} \lesssim \frac{32\pi^2 M_{\rm pl}^2}{m_\chi^4 \lambda_\phi^2} \sim 5\times 10^{-2}\left(\frac{\lambda_\phi}{1\ {\rm pc}}\right)^{-2}\left(\frac{m_\chi}{\rm keV}\right)^{-4}. \quad \text{(E1)}$$

In general, this is difficult to satisfy in the parameter space where our constraints apply. So, the scenarios we consider in this paper are somewhat fine-tuned. For m_χ values near the Tremaine-Gunn limit, $m_\chi \sim 100$ eV, part of the constrained parameter space can be natural in the above sense, although in the same regime, finite-density corrections to m_ϕ^2 are important, potentially removing the constraints in the first place. The fermion-loop contribution to the quartic coupling $\delta\kappa\phi^4/4$ is roughly $(\delta\kappa)_{\rm loop} \sim g^4/2\pi^2 \sim 4\times 10^{-100}\alpha_{\chi\chi}^2 \times (m_\chi/{\rm keV})^4$. This quartic coupling is negligible if it satisfies Eq. (D15)

$$4\times 10^{-100}\alpha_{\chi\chi}^2\left(\frac{m_\chi}{\rm keV}\right)^4 \lesssim \left(\frac{m_\phi}{\phi_*}\right)^2\left(\frac{m_\phi}{H}\right)^{8/5}. \quad \text{(E2)}$$

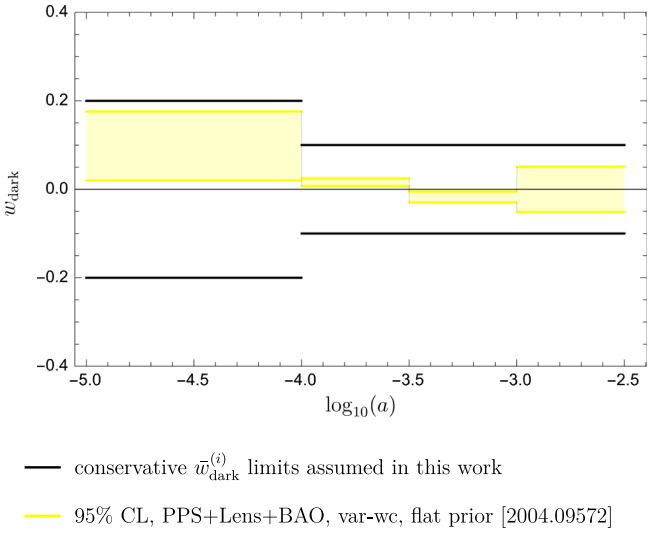


FIG. 13. Comparison between the conservative limits on $\bar{w}^{(i)}_{\rm dark}$ we assumed in our analysis, as given in Eq. (15), (black lines) and the allowed values of $\bar{w}^{(i)}_{\rm dark}$ from [37] (yellow).

We find that in all parameter space constrained by our main analysis, the mediator's mass, and quartic coupling are tuned; see Fig. 12.

APPENDIX F: EXTRA FIGURES

We compare the conservative limits on $\bar{w}^{(i)}_{\rm dark}$ assumed in our analysis [Eq. (15)] with the allowed values of $\bar{w}^{(i)}_{\rm dark}$ found by Ref. [37] in Fig. 13. We show the background limits (discussed in Sec. III D and Sec. V A) from each of the four bins in Fig. 14. In Fig. 15, we show contours of the fractional difference between the dark sector's energy density $\rho_{\rm dark}$ and the CDM value $\rho_{\rm CDM}$ on the $(\lambda_\phi, \alpha_{\chi\chi})$ parameter space, at scale factors $0.01a_{\rm eq}, 0.1a_{\rm eq}, a_{\rm eq}, 10a_{\rm eq}$. These contour plots give a rough indication for which parts of the parameter space have the dark sector deviating considerably from CDM at various epochs.

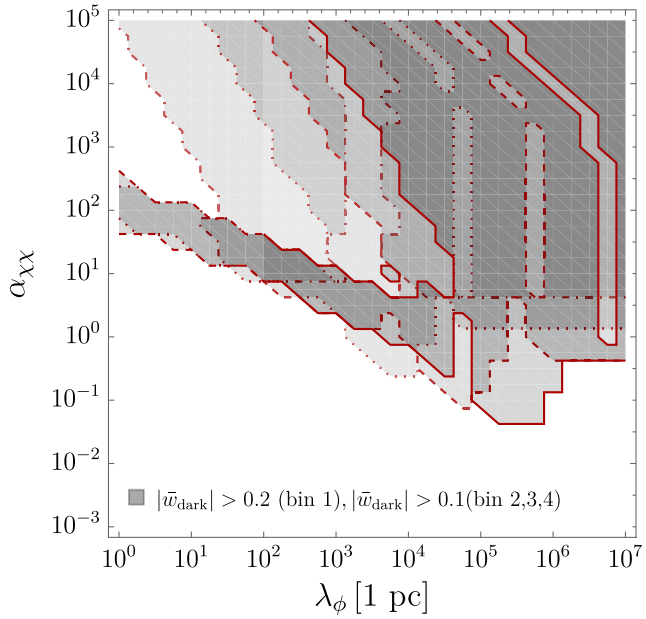


FIG. 14. Background limits from each scale-factor bin, Eq. (13). The limits from bin 1, 2, 3, and 4 are shown in solid, dashed, dotted, and dot dashed.

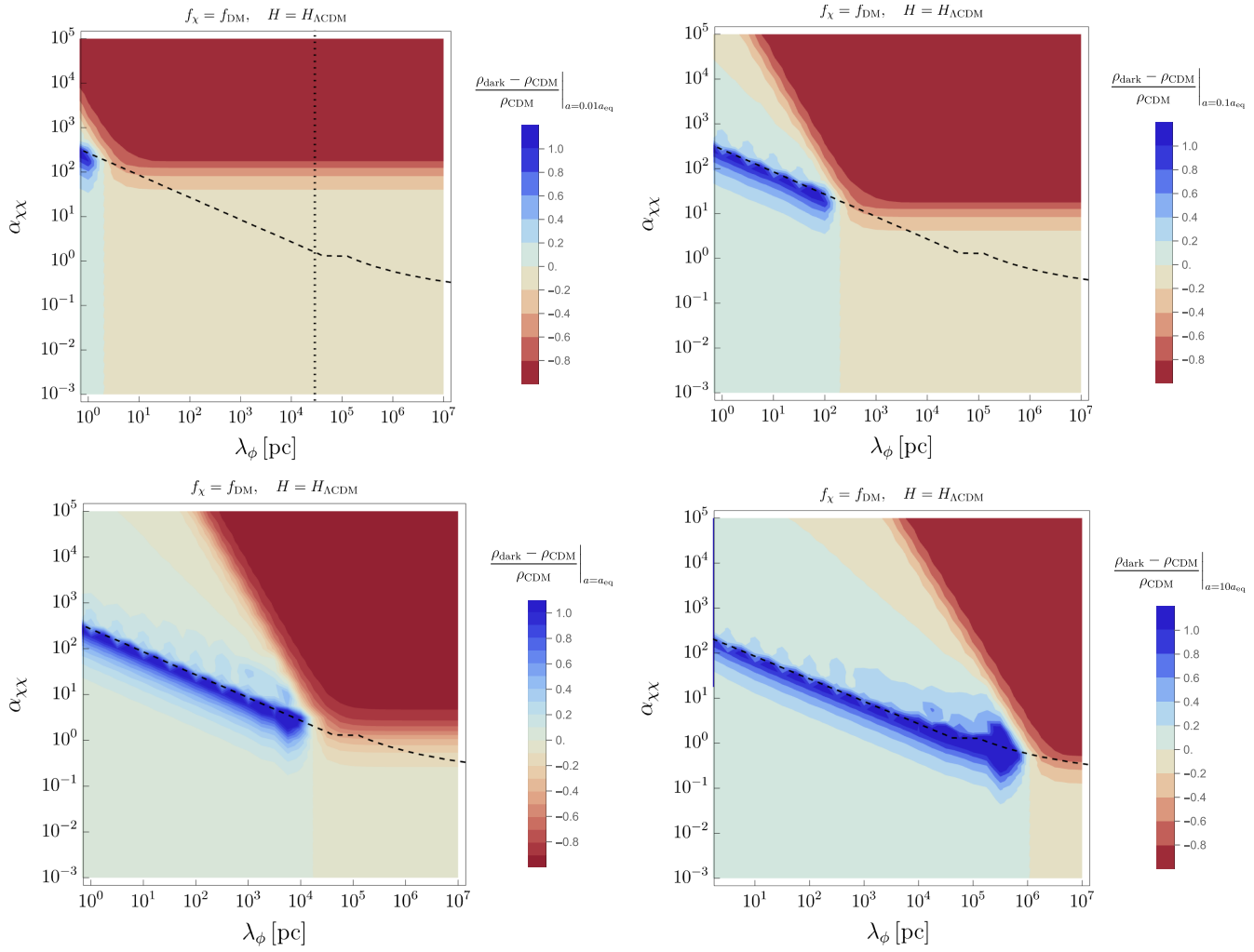


FIG. 15. Fractional of $\rho_{\rm dark}$ from $\rho_{\rm CDM}$. The blue region signifies the mixed DM $f_\phi/f_\chi = \mathcal{O}(1)$. The blobby part near the case $\phi_{\rm rise}\phi_0\phi_*\phi_{\rm DM}$-case $\phi_{\rm rise}\phi_{\rm DM}$ boundary is because the amplitude of $\phi_{\rm DM}$ is set by the relative field-space velocity $\dot{\phi}_{\rm lin-osc} - \dot{\phi}_*$ when $\phi_* \approx \phi_0$, and this relative velocity depends on the phase of $\phi_{\rm lin-osc}$ at that time. In the bottom right figure, $\lambda_\phi \lesssim 1$ pc was not simulated because that regime is computationally expensive.

[1] K. K. Rogers and H. V. Peiris, Phys. Rev. Lett. **126**, 071302 (2021).
[2] S. Das and E. O. Nadler, Phys. Rev. D **103**, 043517 (2021).
[3] V. Poulin, P. D. Serpico, and J. Lesgourgues, J. Cosmol. Astropart. Phys. 08 (2016) 036.
[4] P. W. Graham and H. Ramani, Phys. Rev. D **110**, 075012 (2024).
[5] J. H. Chang, P. J. Fox, and H. Xiao, J. Cosmol. Astropart. Phys. 08 023 (2024).
[6] S. Tulin and H.-B. Yu, Phys. Rep. **730**, 1 (2018).
[7] O. Slone, F. Jiang, M. Lisanti, and M. Kaplinghat, Phys. Rev. D **107**, 043014 (2023).
[8] R. Sundrum, arXiv:hep-th/0312212.
[9] T. Damour and A. M. Polyakov, Gen. Relativ. Gravit. **26**, 1171 (1994).
[10] T. Damour and A. M. Polyakov, Nucl. Phys. B **423**, 532 (1994).
[11] Z. Bogorad, P. W. Graham, and H. Ramani, J. Cosmol. Astropart. Phys. 03 (2025) 067.
[12] J. A. Keselman, A. Nusser, and P. J. E. Peebles, Phys. Rev. D **80**, 063517 (2009).
[13] M. Kesden and M. Kamionkowski, Phys. Rev. Lett. **97**, 131303 (2006).
[14] M. Kesden and M. Kamionkowski, Phys. Rev. D **74**, 083007 (2006).
[15] J. A. Frieman and B.-A. Gradwohl, Phys. Rev. Lett. **67**, 2926 (1991).

[16] B.-A. Gradwohl and J. A. Frieman, Astrophys. J. **398**, 407 (1992).
[17] G. R. Farrar and P. J. E. Peebles, Astrophys. J. **604**, 1 (2004).
[18] R. Bean, E. E. Flanagan, I. Laszlo, and M. Trodden, Phys. Rev. D **78**, 123514 (2008).
[19] R. Bean, E. E. Flanagan, and M. Trodden, Phys. Rev. D **78**, 023009 (2008).
[20] J. A. Keselman, A. Nusser, and P. J. E. Peebles, Phys. Rev. D **81**, 063521 (2010).
[21] W. A. Hellwing and R. Juszkiewicz, Phys. Rev. D **80**, 083522 (2009).
[22] S. Bottaro, E. Castorina, M. Costa, D. Redigolo, and E. Salvioni, Phys. Rev. D **112**, 023525 (2025).
[23] S. Bottaro, E. Castorina, M. Costa, D. Redigolo, and E. Salvioni, Phys. Rev. Lett. **132**, 201002 (2024).
[24] M. Archidiacono, E. Castorina, D. Redigolo, and E. Salvioni, J. Cosmol. Astropart. Phys. 10 (2022) 074.
[25] M. Costa, C. Creque-Sarbinowski, O. Simon, and Z. J. Weiner, J. Cosmol. Astropart. Phys. 06 (2026) 055.
[26] G. Domènech and M. Sasaki, J. Cosmol. Astropart. Phys. 06 (2021) 030.
[27] G. Domènech, D. Inman, A. Kusenko, and M. Sasaki, Phys. Rev. D **108**, 103543 (2023).
[28] N. Afshordi, M. Zaldarriaga, and K. Kohri, Phys. Rev. D **72**, 065024 (2005).
[29] S. Savastano, L. Amendola, J. Rubio, and C. Wetterich, Phys. Rev. D **100**, 083518 (2019).
[30] N. Blinov and A. Hook, J. High Energy Phys. 06 (2016) 176.
[31] A. Hook, Phys. Rev. Lett. **120**, 261802 (2018).
[32] D. Brzeminski, Z. Chacko, A. Dev, and A. Hook, Phys. Rev. D **104**, 075019 (2021).
[33] Planck, N. Aghanim *et al.*, Astron. Astrophys. **641**, A6 (2020); **652**, C4(E) (2021).
[34] W. Hu, R. Barkana, and A. Gruzinov, Phys. Rev. Lett. **85**, 1158 (2000).
[35] E. G. M. Ferreira, Astron. Astrophys. Rev. **29**, 7 (2021).
[36] L. Hui, Annu. Rev. Astron. Astrophys. **59**, 247 (2021).
[37] S. Ilić, M. Kopp, C. Skordis, and D. B. Thomas, Phys. Rev. D **104**, 043520 (2021).
[38] M. Kopp, C. Skordis, D. B. Thomas, and S. Ilić, Phys. Rev. Lett. **120**, 221102 (2018).
[39] J. Samsing, E. V. Linder, and T. L. Smith, Phys. Rev. D **86**, 123504 (2012).
[40] A. Hojjati, E. V. Linder, and J. Samsing, Phys. Rev. Lett. **111**, 041301 (2013).
[41] W. Hu, Astrophys. J. **506**, 485 (1998).
[42] P. Meszaros, Astron. Astrophys. **37**, 225 (1974).
[43] W. Hu and N. Sugiyama, Astrophys. J. **471**, 542 (1996).
[44] P. W. Graham and H. Ramani, Phys. Rev. D **110**, 075011 (2024).
[45] P. W. Graham, D. Green, and J. Meyers, Phys. Rev. D **113**, 043514 (2026).
[46] A. J. Iovino, G. Perna, A. Riotto, and H. Veermäe, J. Cosmol. Astropart. Phys. 10 (**2024**) 050.
[47] T. Bringmann, D. Croon, and S. Sevillano Muñoz, Phys. Rev. Lett. **137**, 061002 (2026).
[48] G. Efstathiou and J. R. Bond, Mon. Not. R. Astron. Soc. **304**, 75 (1999).
[49] A. Chakraborty, P. K. Chanda, K. L. Pandey, and S. Das, Astrophys. J. **932**, 119 (2022).
[50] A. Chakraborty, T. Ray, S. Das, A. Banerjee, and V. Ganesan, Astrophys. J. **998**, 83 (2026).
[51] R. Fardon, A. E. Nelson, and N. Weiner, J. Cosmol. Astropart. Phys. 10 (2004) 005.
[52] D. B. Kaplan, A. E. Nelson, and N. Weiner, Phys. Rev. Lett. **93**, 091801 (2004).
[53] J. Sakstein and M. Trodden, Phys. Rev. Lett. **124**, 161301 (2020).
[54] J. M. Berryman *et al.*, Phys. Dark Universe **42**, 101267 (2023).
[55] A. Banerjee, N. H. Nguyen, and E. H. Tanin, arXiv:2509.25308.
[56] Z. Bogorad, P. W. Graham, and H. Ramani, J. Cosmol. Astropart. Phys. 04 (2025) 006.
[57] V. Poulin, T. L. Smith, T. Karwal, and M. Kamionkowski, Phys. Rev. Lett. **122**, 221301 (2019).
[58] M. Kamionkowski and A. G. Riess, Annu. Rev. Nucl. Part. Sci. **73**, 153 (2023).
[59] K. Inomata and T. Nakama, Phys. Rev. D **99**, 043511 (2019).
[60] C.-P. Ma and E. Bertschinger, Astrophys. J. **455**, 7 (1995).
[61] E. H. Tanin and E. D. Stewart, J. Cosmol. Astropart. Phys. 11 (2017) 019.
[62] A. Banerjee, N. H. Nguyen, and E. H. Tanin, Phys. Rev. D **114**, 015032 (2026).