

Self-heating of strongly interacting massive particles

Ayuki Kamada,^{1,*} Hee Jung Kim,^{2,†} and Hyungjin Kim^{3,‡}

¹*Center for Theoretical Physics of the Universe, Institute for Basic Science (IBS), Daejeon 34126, Korea*

²*Department of Physics, KAIST, Daejeon 34141, Korea*

³*Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 7610001, Israel*



(Received 24 May 2018; published 13 July 2018)

It was recently pointed out that semiannihilating dark matter (DM) may experience a novel temperature evolution dubbed as *self-heating*. Exothermic semiannihilation converts the DM mass to the kinetic energy. This yields a unique DM temperature evolution, $T_\chi \propto 1/a$, in contrast to $T_\chi \propto 1/a^2$ for free-streaming nonrelativistic particles. Self-heating continues as long as self-scattering sufficiently redistributes the energy of DM particles. In this paper, we study the evolution of cosmological perturbations in self-heating DM. We find that sub-GeV self-heating DM leaves a cutoff on the subgalactic scale of the matter power spectrum when the self-scattering cross section is $\sigma_{\text{self}}/m_\chi \sim \mathcal{O}(1) \text{ cm}^2/\text{g}$. Then we present a particle physics realization of the self-heating DM scenario. The model is based on recently proposed strongly interacting massive particles with pionlike particles in a QCD-like sector. Pionlike particles semiannihilate into an axionlike particle, which is thermalized with dark radiation. The dark radiation temperature is smaller than the standard model temperature, evading the constraint from the effective number of neutrino degrees of freedom. It is easily realized when the dark sector is populated from the standard model sector through a small coupling.

DOI: [10.1103/PhysRevD.98.023509](https://doi.org/10.1103/PhysRevD.98.023509)

I. INTRODUCTION

Accumulated observational data begin to test our understanding of how dark matter (DM) is distributed over the Universe and how the DM distribution evolves in time [1]. In regard to the structure formation of the Universe, DM is often described as a perfect fluid with zero temperature, which is referred to as cold dark matter (CDM). The CDM paradigm is successful in reproducing the large-scale structure of the Universe observed through cosmic microwave background (CMB) anisotropies [2] and galaxy clustering [3].

Contrary to its success on large scales, CDM predictions appear to be incompatible with the observations on smaller scales [4]. Although baryonic physics may play an important role [5–7], these small-scale issues may be hinting to alternative DM models. One example of the CDM failure is the *missing satellite problem*, indicating that CDM overpredicts the number of dwarf-size subhalos in a Milky Way–size halo when compared to that of the observed

satellite galaxies [8,9]. Warm dark matter (WDM) is an interesting possibility in this respect. Gravitational clustering of WDM particles is interrupted on a subgalactic scale because of a sizable thermal velocity of $v/c \sim 10^{-3} - 10^{-4}$ at the matter-radiation equality. This suppresses dwarf-size halo formation [10,11]. A sterile neutrino with a keV mass is a good benchmark model of WDM, where its phenomenology is described by the mass and the mixing angle with an active neutrino in the simplest setup [12].

Another example is the *core-cusp problem*: Some dwarf-size halos have a cuspy profile as predicted by CDM, while others have a cored profile [13–15]. Self-interacting dark matter (SIDM) is an interesting solution. The self-scattering cross section of $\sigma_{\text{self}}/m_\chi \sim 1 \text{ cm}^2/\text{g}$ with the DM mass m_χ leads to isothermalization of DM particles, whose distribution is characterized by a kiloparsec core [16,17]. SIDM reproduces the observed diversity of the rotation velocity among similar-size halos by adjusting its distribution sensitively to the baryon distribution [18,19].

A new possibility, which is called self-heating DM, has been proposed recently [20]. A characteristic feature of this scenario is that the strength of self-scattering is related to the thermal velocity of DM and, hence, potentially solves several small-scale issues simultaneously. The original proposal was based on semiannihilating DM [21–23], $\chi\chi \rightarrow \chi\phi$ with a light particle ϕ in the thermal bath. A key observation is that semiannihilation converts the mass of DM into the kinetic energy, which leads to the novel DM temperature evolution, $T_\chi \propto 1/a$ after the freeze-out,

*akamada@ibs.re.kr

†hyzer333@kaist.ac.kr

‡hyungjin.kim@weizmann.ac.il

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). 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.

instead of $T_\chi \propto 1/a^2$ for free-streaming nonrelativistic particles. Self-heating continues as long as self-scattering occurs rapidly. DM self-scattering is an essential ingredient of the scenario, because it redistributes the large kinetic energy of the boosted DM particles through the semiannihilation over the whole DM particles. Stronger self-scattering elongates the duration of self-heating, resulting in a larger thermal motion of DM particles.

In this paper, we investigate the impact of DM self-heating on the matter distribution of the Universe and propose a viable particle physics realization of self-heating DM. To study the structure formation, we derive the evolution equation of the cosmological perturbations in the self-heating DM scenario. We show that semiannihilation not only changes the DM temperature evolution but also affects the entropic perturbation of DM. We follow the evolution of the primordial perturbations numerically and show that self-heating DM can leave a subgalactic-scale cutoff in the linear matter power spectrum. We extend the recently proposed strongly interacting massive particle (SIMP) model [24,25] to realize self-heating DM. Pionlike particles in a QCD-like sector semiannihilate into an axionlike particle (ALP), which is thermalized with dark radiation. We identify a parameter region that is compatible with observational constraints.

This paper is organized as follows. In Sec. II, we describe the thermal history of the self-heating DM from its freeze-out to structure formation of the Universe. In Sec. III, we discuss an extension of the SIMP model with a light ALP and dark radiation. We conclude in Sec. IV. In Appendix A, we provide a detailed derivation of the coevolution equations of the DM number density and temperature. In Appendix B, we present the evolution equations of cosmological perturbations in the self-heating DM scenario.

II. THERMAL HISTORY OF SELF-HEATING DM FROM THE FREEZE-OUT TO STRUCTURE FORMATION

In this section, we discuss the freeze-out of the DM number density, its novel temperature evolution afterwards, and the evolution of cosmological perturbations in the self-heating DM scenario. For our purpose, we consider a scalar DM χ and a light mediator ϕ . We mainly consider two types of interaction: One is self-scattering $\chi\chi \leftrightarrow \chi\chi$, and the other one is semiannihilation $\chi\chi \leftrightarrow \chi\phi$, which are the minimal ingredients to realize self-heating of DM. The presence of efficient self-scattering enforces the DM distribution to be $f_\chi = (n_\chi/n_\chi^{\text{eq}}) \exp[-E_\chi/T_\chi]$, where $n_\chi^{\text{eq}} = (m_\chi^2 T_\chi / 2\pi^2) K_2(m_\chi/T_\chi)$ and K_2 is the second-order modified Bessel function of the second kind. We also assume that a light mediator remains in thermal equilibrium by contacting either to the standard model (SM) sector or to the dark sector and, hence, $f_\phi = \exp[-E_\phi/T_\phi]$, where

$T_\phi = T_{\text{SM}}$ or $T_\phi = T_{\text{DR}}$, respectively. Its number density is thus given by $n_\phi^{\text{eq}} = (m_\phi^2 T_\phi / 2\pi^2) K_2(m_\phi/T_\phi)$. In general, annihilation $\chi\chi \leftrightarrow \phi\phi$ and also elastic scattering $\chi\phi \leftrightarrow \chi\phi$ may exist.

A. Homogeneous and isotropic evolution

The coevolution equations of n_χ and T_χ are given, respectively, as

$$\dot{n}_\chi + 3Hn_\chi = -n_\chi \langle \sigma_{\text{semi}} v_{\text{rel}} \rangle_{T_\chi T_\chi} \times [n_\chi - \mathcal{J}(T_\chi, T_\phi) n_\chi^{\text{eq}}(T_\chi)], \quad (1)$$

$$\begin{aligned} \dot{T}_\chi + 3HT_\chi \left(\frac{T_\chi}{\sigma_E} \right)^2 &= - \left(\frac{T_\chi}{\sigma_E} \right)^2 \frac{n_\phi^{\text{eq}}(T_\chi)}{n_\chi^{\text{eq}}(T_\chi)} \\ &\times \langle \Delta E \sigma_{\text{inv}} v_{\text{rel}} \rangle_{T_\chi, T_\phi = T_\chi} \\ &\times [n_\chi - n_\chi^{\text{eq}}(T_\chi) \mathcal{K}(T_\chi, T_\phi)] \\ &+ 2\gamma_{\chi\phi \rightarrow \chi\phi} (T_\phi - T_\chi), \end{aligned} \quad (2)$$

where σ_{semi} (σ_{inv}) is the cross section for the $\chi\chi \rightarrow \chi\phi$ ($\chi\phi \rightarrow \chi\chi$) process. See Appendix A for the derivation. We have added the elastic scattering term for completeness. The expression of momentum exchange rate $\gamma_{\chi\phi \rightarrow \chi\phi}$ can be found in Ref. [26] and thus is not repeated here. We remark that self-scattering does not contribute to these equations, because it conserves the number and energy of DM particles. The relic abundance of χ coincides with the observed one when

$$\langle \sigma_{\text{semi}} v_{\text{rel}} \rangle |_{T_\chi = T_{\chi, \text{fo}}} = (\sigma v_{\text{rel}})_{\text{can}} (T_\phi / T_{\text{SM}})_{\text{fo}}, \quad (3)$$

where $(\sigma v_{\text{rel}})_{\text{can}} \simeq (3 \times 10^{-26} \text{ cm}^3/\text{s})$ is a canonical cross section reproducing the observed relic density of thermal DM [27]. The canonical cross section is rescaled by the factor of $(T_\phi / T_{\text{SM}})_{\text{fo}}$, because $T_{\text{SM, fo}} \simeq (m_\chi / 20) (T_{\text{SM}} / T_\phi)_{\text{fo}}$.

We present numerical results in Fig. 1, which show the evolution of DM yield $Y_\chi = n_\chi/s$ and the temperature ratio T_χ/T_ϕ . Here $s = (2\pi^2/45) g_{*s, \text{SM}} T^3$ is the entropy density, and $g_{*s, \text{SM}}$ is the effective number of relativistic degrees of freedom (d.o.f.) in the SM sector. In this figure, we assume that $\gamma_{\chi\phi \rightarrow \chi\phi}/H \ll 1$ and $T_\phi = T_{\text{SM}}$. Like usual thermal DM with $T_\chi = T_\phi$, the relic abundance is determined around $T_{\phi, \text{fo}} \simeq m_\chi/20$ (see Appendix A for a small difference). On the other hand, the DM temperature evolution shows a unique behavior. Especially, after the freeze-out, the DM temperature scales as $T_\chi \propto 1/a$ (see Appendix A for a thermodynamic derivation) despite the fact that DM particles are nonrelativistic and that no elastic scattering equilibrates T_χ and T_ϕ . Self-heating of DM occurs because a small portion of DM still undergoes the semiannihilation after the freeze-out and gain the kinetic energy of the order

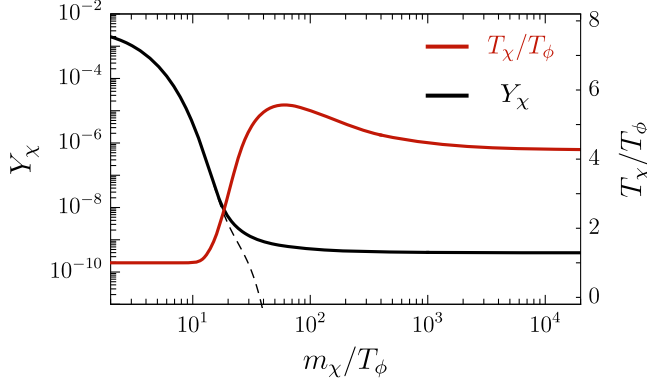


FIG. 1. Thermal history of self-heating DM. The plot shows the evolution of DM yield (black curve) and also the evolution of the DM temperature (red curve). The dashed line corresponds to the equilibrium value. For this plot, we assume no elastic scatterings; thus, the chemical and kinetic decoupling take place at the same time as semiannihilation decouples from the plasma. One can see that the DM temperature scales as T_ϕ scales even though DM is nonrelativistic and kinetically decoupled from the thermal bath.

of its mass, which is much larger than the DM temperature. We find that the ratio between the two temperatures approaches¹

$$r_{\chi\phi} \equiv \frac{T_\chi}{T_\phi} = (\gamma - 1) \frac{2m_\chi}{3T_{\phi,\text{fo}}}, \quad (4)$$

where $\gamma = (5/4)[1 - m_\phi^2/(5m_\chi^2)]$ is the Lorentz boost factor of the final state DM particle.

Although we ignored elastic scatterings above, we emphasize that the self-heating is a generic feature of exothermic semiannihilation even in the presence of elastic scattering. If $(\gamma_{\chi\phi \rightarrow \chi\phi}/H)|_{T_\chi=T_{\chi,\text{fo}}} \gg 1$, the elastic scattering is able to maintain the kinetic equilibrium of DM after the freeze-out, resulting in $T_\chi = T_\phi$. In this case, Eq. (1) reproduce an usual discussion found in Refs. [21–23], since $\mathcal{J}(T_\chi = T_\phi, T_\phi) = 1$. Eventually, elastic scattering decouples. After this kinetic decoupling, DM begins to self-heat, and the DM temperature continues to scale as $T_\chi \propto 1/a$. See Fig. 2 for schematic picture of the DM temperature evolution.

We stress that DM self-scattering is necessary for DM self-heating. If self-scattering is absent, semiannihilation merely produces a small portion of boosted DM particles, which act as a hot component of DM. Thus, the unique temperature evolution continues only until $\Gamma_{\text{self}} = \langle \sigma_{\text{self}} v_{\text{rel}} \rangle n_\chi \simeq H$. The larger the self-scattering cross section is, the longer the self-heating lasts and thus results in a larger thermal motion of DM. The self-scattering decouples at

¹The temperature ratio is constant up to changes in the number of relativistic d.o.f.

$$T_{\text{SM,self}} \simeq 1 \text{ eV} r_{\chi\phi}^{-n} \left(\frac{1 \text{ cm}^2/\text{g}}{\sigma_{\text{self}}/m_\chi} \right)^{2n} \left(\frac{m_\chi}{1 \text{ GeV}} \right)^n \left(\frac{T_{\text{SM}}}{T_\phi} \right)_{\text{self}}^n, \quad (5)$$

where $n = 1/3$ for $T_{\text{SM,self}} > T_{\text{SM,eq}}$, while $n = 1/4$ when $T_{\text{SM,self}} < T_{\text{SM,eq}}$. Here, $T_{\text{SM,eq}} \simeq 0.8 \text{ eV}$ is the SM temperature at the matter-radiation equality. When deriving Eq. (5), we have used $\langle v_{\text{rel}} \rangle = (4/\sqrt{\pi}) \sqrt{T_{\text{SM}}/m_\chi} r_{\chi\phi}^{1/2} \times (T_\phi/T_{\text{SM}})^{1/2}$ and solve $\Gamma_{\text{self}} = H$ for T_{SM} . Remember that $\sigma_{\text{self}}/m_\chi \sim 1 \text{ cm}^2/\text{g}$ forms a kiloparsec core in a subgalactic halo [28].² A resultant large thermal motion of sub-GeV SIDM leaves a subgalactic-scale cutoff in the linear matter power spectrum like keV WDM. We will discuss the structure formation of self-heating DM in the next section.

Before closing this section, we emphasize that the thermal history of ϕ plays a significant role both in DM searches and in the evolution of the Universe. If ϕ is massless, it changes the expansion rate of the Universe, as it contributes to the total energy density. Its impact is described by the change in the effective number of neutrino d.o.f. ΔN_{eff} and is constrained by big bang nucleosynthesis (BBN) and CMB. The latest constraint from the Planck Collaboration is $\Delta N_{\text{eff}} = 3.15 \pm 0.24$ [2]. For massive ϕ , it may overclose the Universe unless it decays or they annihilate into light particles. If ϕ decays into visible particles, the late time production of ϕ through semiannihilation is severely constrained by the Galactic and extra-Galactic gamma-ray searches as well as the CMB measurement. These constraints require $m_\chi \gtrsim 10 \text{ GeV}$ [30,31], which results in a shorter duration of self-heating [see Eq. (5)].³ The other possibility is that ϕ decays into dark radiation. In this case, constraints from DM indirect searches cannot be applied, although ΔN_{eff} still constrains the scenario. It will be discussed in Sec. III.

B. Perturbed evolution

The novel evolution of the DM temperature affects the evolution of cosmological perturbations. We derive evolution equations of cosmological perturbations in the self-heating scenario in Appendix B. In the case of self-scattering DM, relevant variables are density contrast δ_χ , velocity divergence θ_χ , and entropy perturbation π_χ . The linearized equations for the cosmological perturbation for self-heating DM are given as

²It is claimed that semiannihilation cooperates with self-interaction to flatten the inner density profile of a subgalactic halo [29]. The impact of the self-heating is more significant in smaller halos, and it may alleviate a required strength of self-scattering for solving the core-cusp problem. We do not take this effect into account for simplicity in this paper.

³Thermal relic DM with a sub-GeV mass is still viable if DM particles annihilate into heavier particles [32]. In the case of semiannihilation, this can be realized if ϕ is heavier than the DM mass [33]. In this case, we expect no self-heating of DM, because semiannihilation is no longer exothermic.

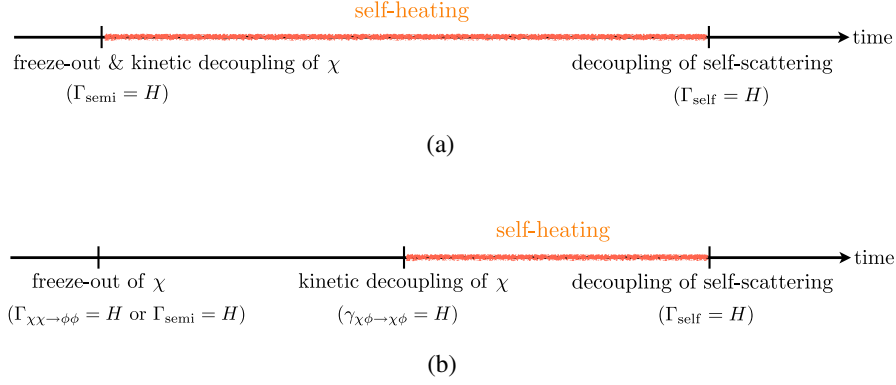


FIG. 2. The self-heating takes place between the kinetic decoupling and the freeze-out of DM self-interaction. (a) $(\gamma_{\chi\phi\rightarrow\chi\phi}/H)|_{T_\chi=T_{\chi,\text{fo}}} \ll 1$. The chemical and kinetic decoupling take place simultaneously. The self-heating of DM begins from the freeze-out of semiannihilation to the freeze-out of self-interaction. (b) $(\gamma_{\chi\phi\rightarrow\chi\phi}/H)|_{T_\chi=T_{\chi,\text{fo}}} \gg 1$. The kinetic equilibrium could be maintained after the chemical freeze-out. In this case, the self-heating begins after the decoupling of elastic scattering and continues until $\Gamma_{\text{self}} = H$.

$$\delta'_\chi = -\theta_\chi - 3\Phi', \quad (6)$$

$$\theta'_\chi = -\mathcal{H}\theta_\chi + k^2\Psi + k^2(c_{s\chi}^2\delta_\chi + \pi_\chi), \quad (7)$$

$$\begin{aligned} \pi'_\chi = & -2\mathcal{H}\pi_\chi + \left(c_{s\chi}^2 - \frac{5}{3}\omega_\chi\right) \left[-\mathcal{H}\left(1 - \frac{\mathcal{H}'}{\mathcal{H}^2}\right)\delta_\chi + \theta_\chi \right. \\ & \left. + 3(\Phi' - \mathcal{H}\Psi)\right], \quad (8) \end{aligned}$$

where the prime is a derivative with respect to the conformal time and $\mathcal{H} = a'/a$. Here we take the conformal Newtonian gauge:

$$ds^2 = a^2(\tau)[-(1 + 2\Psi)d\tau^2 + (1 + 2\Phi)\delta_{ij}dx^i dx^j]. \quad (9)$$

We cross-check the result in the synchronous gauge in Appendix B. Semiannihilation affects δ_χ mainly through the relatively large $c_{s\chi}^2 \simeq (4/3)T_\chi/m_\chi$. One can obtain the

matter power spectrum by solving the above equations with equations of the equation of state ω_χ and sound speed squared $c_{s\chi}^2$:

$$\omega'_\chi = -2\mathcal{H}\omega_\chi + \frac{2}{3}(\gamma - 1)a\Gamma_{\text{semi}}(1 - e^{-\Gamma_{\text{self}}/H}), \quad (10)$$

$$c_{s\chi}^2 = \frac{5}{3}\omega_\chi - \frac{2}{9}(\gamma - 1)\frac{a\Gamma_{\text{semi}}}{\mathcal{H}}(1 - e^{-\Gamma_{\text{self}}/H}), \quad (11)$$

with $\Gamma_{\text{semi}} = n_\chi\langle\sigma_{\text{semi}}v_{\text{rel}}\rangle \propto 1/a^3$. As the interaction rate for self-scattering becomes smaller than the Hubble expansion rate, the DM temperature behaves as that of free-streaming nonrelativistic particles, i.e., $T_\chi \propto 1/a^2$.

By modifying the publicly available Boltzmann solver CLASS [34], we obtain the present linear matter power spectrum as shown in Fig. 3. The matter power spectrum exhibits a sharp cutoff around $k = \mathcal{O}(100) \text{ Mpc}^{-1}$. This

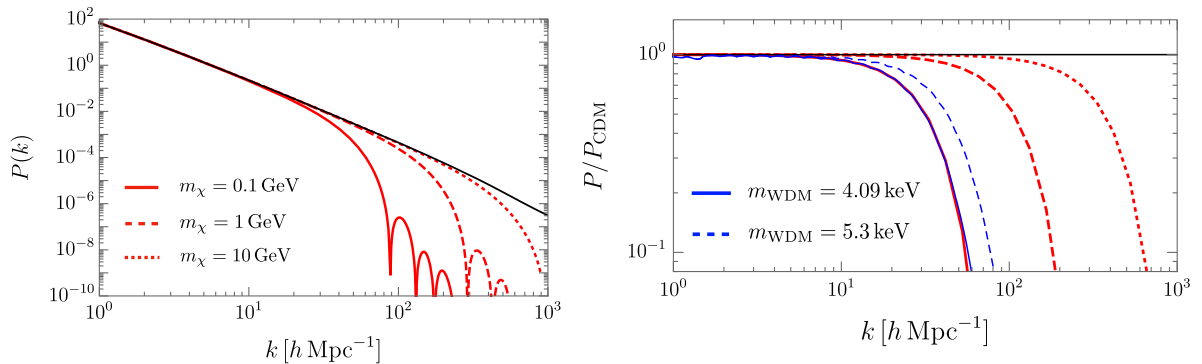


FIG. 3. (Left) Linear matter power spectrum for self-heating DM (red lines) and CDM (black line) at the present Universe. (Right) Linear matter power spectrum normalized with respect to that of CDM. We also show the power spectrum in the thermal WDM models (blue lines) with $m_{\text{WDM}} = 4.09 \text{ keV}$ and $m_{\text{WDM}} = 5.3 \text{ keV}$, where the latter is the latest lower bound on the mass of WDM [36]. We choose $\sigma_{\text{self}}/m_\chi = 1 \text{ cm}^2/\text{g}$ such that self-scattering decouples at $T_{\text{SM,self}} = 0.5, 0.8, \text{ and } 1.6 \text{ eV}$ for $m_\chi = 0.1, 1, \text{ and } 10 \text{ GeV}$, respectively. We also assume $T_\phi = T_{\text{SM}}$.

scale turns out to well match the Jeans instability scale at the matter-radiation equality [20,35]:

$$k_J \simeq 180 \text{ Mpc}^{-1} r_{\chi\phi}^{-1/2} \max\left(1, \sqrt{a_{\text{eq}}/a_{\text{self}}}\right) \times \left(\frac{m_\chi}{1 \text{ GeV}}\right)^{1/2} \left(\frac{T_\gamma}{T_\phi}\right)^{1/2}, \quad (12)$$

where the temperature ratio $r_{\chi\phi} = T_\chi/T_\phi$ should be evaluated at $a = \min(a_{\text{self}}, a_{\text{eq}})$.

We compare the resultant matter power spectra in self-heating DM to those in the thermal WDM model (see, e.g., Ref. [37] for details). The Lyman- α forest data constrain the WDM mass as $m_{\text{WDM}} > 5.3 \text{ keV}$ [36]. In the case of WDM, the Jeans instability scale appears to be $k_{J,\text{WDM}} \simeq 180 \text{ Mpc}^{-1} (m_{\text{WDM}}/5.3 \text{ keV})^{4/3}$, while a drop of power in the matter power spectrum takes place around $k_{J,\text{WDM}}/4$, where an additional order one factor can be attributed to the free-streaming of WDM during the radiation-dominated Universe [38]. By equating

$k_J \simeq k_{J,\text{WDM}}/4$, we find the following correspondence between WDM and self-heating DM:

$$\frac{m_{\text{WDM}}}{5.3 \text{ keV}} \simeq \left(\frac{r_{\chi\phi}}{2.4}\right)^{-3/8} \left(\frac{m_\chi}{0.1 \text{ GeV}}\right)^{3/8} \times \max\left(1, \sqrt{a_{\text{eq}}/a_{\text{self}}}\right)^{3/4} \left(\frac{T_\gamma}{T_\phi}\right)^{3/8}. \quad (13)$$

For the massless mediator ϕ sharing the temperature with SM particles, the current limit on the thermal WDM mass, $m_{\text{WDM}} \geq 5.3 \text{ keV}$, translates into the mass bound of self-heating DM as $m_\chi \geq 0.1 \text{ GeV}$ if the self-interaction decouples after the matter-radiation equality.

III. EXTENDED SIMP MODEL WITH AN ALP AND DARK RADIATION

In this section, we propose a realization of self-heating DM. We extend the SIMP model [39] with an ALP [33] by introducing dark radiation. The Lagrangian density is given by

$$\begin{aligned} \mathcal{L} = & |\partial_\mu \Phi|^2 - V(|\Phi|^2) + N^\dagger i\bar{\sigma}^\mu D_\mu N + \bar{N}^\dagger i\bar{\sigma}^\mu D_\mu \bar{N} - (m_N \bar{N} N + \text{H.c.}) + Q^\dagger i\bar{\sigma}^\mu D_\mu Q + \bar{Q}^\dagger i\bar{\sigma}^\mu D_\mu \bar{Q} + L^\dagger i\bar{\sigma}^\mu D_\mu L \\ & + \bar{L}^\dagger i\bar{\sigma}^\mu D_\mu \bar{L} - \Phi(y_Q Q \bar{Q} + y_L L \bar{L} + \text{H.c.}) - \frac{1}{4} H^a{}_{\mu\nu} H^{a\mu\nu} - \frac{1}{4} X^a{}_{\mu\nu} X^{a\mu\nu} + \theta_H \frac{g_H^2}{32\pi^2} H^a{}_{\mu\nu} \tilde{H}^{a\mu\nu} + \theta_X \frac{g_X^2}{32\pi^2} X^a{}_{\mu\nu} \tilde{X}^{a\mu\nu}, \end{aligned} \quad (14)$$

where $H^a{}_{\mu\nu}$ ($\tilde{H}^a{}_{\mu\nu}$) is the (dual) field strength of the confining $\text{SU}(N_H)$ gauge field and $X^a{}_{\mu\nu}$ ($\tilde{X}^a{}_{\mu\nu}$) is the (dual) field strength of the quasiperturbative $\text{SU}(N_X)$ gauge field. Gauge charges of the matter contents are summarized in Table I. N and \bar{N} are N_f -flavored ($N_f \geq 3$).

As the PQ symmetry is spontaneously broken, an ALP arises. The mass of Q and L originates from vacuum expectation value $\langle \Phi \rangle$ and is assumed to be heavier than the confinement scale of the $\text{SU}(N_H)$ gauge group. After integrating out Q and L , we find that the ALP couples to both $H^a{}_{\mu\nu} \tilde{H}^{a\mu\nu}$ and $X^a{}_{\mu\nu} \tilde{X}^{a\mu\nu}$. Once $\text{SU}(N_H)$ confines, $\text{SU}(N_f)_L \times \text{SU}(N_f)_R$ global symmetry breaks down to the

diagonal subgroup $\text{SU}(N_f)_V$, and pionlike particles χ emerge. For low-energy phenomenology, we obtain the following effective Lagrangian:

$$\mathcal{L}_{\text{eff}} = \frac{1}{2} (\partial_\mu \phi)^2 - \frac{1}{4} X^a{}_{\mu\nu} X^{a\mu\nu} + \frac{g_X^2}{32\pi^2} \frac{\phi}{f} X^a{}_{\mu\nu} \tilde{X}^{a\mu\nu} + \mathcal{L}_{\text{chiral}}. \quad (15)$$

The expression of $\mathcal{L}_{\text{chiral}}$ can be found in Ref. [33] and thus is not repeated here. We emphasize that $\text{SU}(N_f)_V$ is the exact symmetry of the theory; hence, no decay operator of DM pions is allowed. The ALP ϕ and dark radiation X form a dark plasma with the temperature $T_\phi = T_{\text{DR}}$. In the following, we take $N_f = 4$, $N_H = 3$, and $N_X = 2$ as a benchmark for the analysis.

Low-energy phenomenology is described by three parameters: DM pion mass m_χ , DM pion decay constant f_χ , and ALP decay constant f . The self-scattering cross section for solving the core-cusp problem and the semi-annihilation cross section for the observed DM relic density [see Eq. (3)] are achieved as [33]

$$\frac{\sigma_{\text{self}}}{m_\chi} = 1 \text{ cm}^2/\text{g} \left(\frac{m_\chi}{50 \text{ MeV}}\right) \left(\frac{40 \text{ MeV}}{f_\chi}\right)^4, \quad (16)$$

TABLE I. Gauge charges of matter contents.

	$\text{SU}(N_H)$	$\text{SU}(N_X)$	$\text{U}(1)_{\text{PQ}}$
Q	\square	1	-1/2
\bar{Q}	$\bar{\square}$	1	-1/2
L	1	\square	-1/2
\bar{L}	1	$\bar{\square}$	-1/2
N	\square	1	0
\bar{N}	$\bar{\square}$	1	0
Φ	1	1	1

$$\langle \sigma_{\text{semi}} v_{\text{rel}} \rangle = (\sigma v_{\text{rel}})_{\text{can}} \left(\frac{m_\chi}{50 \text{ MeV}} \right)^2 \left(\frac{40 \text{ MeV}}{f_\chi} \right)^2 \times \left(\frac{300 \text{ GeV}}{f} \right)^2. \quad (17)$$

Note that we only consider CP -conserving interactions of DM pions and ALPs, because the CP -violating vacuum angle in the hidden confining sector dynamically vanishes due to an ALP. Although we consider semiannihilation as a dominant process for the freeze-out of the DM number density, there exist two additional number-changing processes: the 3-to-2 process of $\chi\chi\chi \rightarrow \chi\chi$ and annihilation $\chi\chi \rightarrow \phi\phi$. Since the annihilation cross section is suppressed by $(f_\chi/f)^2$ compared to that of semiannihilation, it is subdominant. Meanwhile, the 3-to-2 process is subdominant for

$$m_\chi < 230 \text{ MeV} \left(\frac{T_{\text{DR}}}{T_{\text{SM}}}_{\text{fo}} \right)^{8/9}, \quad (18)$$

where we have chosen f_χ and f such that $\sigma_{\text{self}}/m_\chi \simeq 1 \text{ cm}^2/\text{g}$ and $\langle \sigma_{\text{semi}} v_{\text{rel}} \rangle = (\sigma v_{\text{rel}})_{\text{can}} (T_{\text{DR}}/T_{\text{SM}})_{\text{fo}}$. Note that the ALP obtains its mass from explicit breaking of chiral symmetry of the QCD-like sector, and its mass is given as [33]

$$m_\phi = \frac{m_\chi f_\chi}{2\sqrt{2}N_f f} \simeq 1 \text{ keV} \left(\frac{m_\chi}{50 \text{ MeV}} \right) \left(\frac{f_\chi}{40 \text{ MeV}} \right) \left(\frac{300 \text{ GeV}}{f} \right). \quad (19)$$

The elastic scattering between the ALP and DM also exists in this model. For the self-heating to occur, the elastic scattering should decouple before the self-scattering decouples. This can be trivially achieved in our model, because the strength of elastic scattering is suppressed by $(f_\chi/f)^2$ compared to that of semiannihilation and by $(f_\chi/f)^4$ compared to that of self-interaction. As a consequence, the momentum exchange rate due to elastic scattering is

$$\frac{\gamma_{\chi\phi \rightarrow \chi\phi}}{H} \simeq 5 \times 10^{-4} r_{\chi\phi}^{-1} \left(\frac{300 \text{ GeV}}{f} \right)^4 \left(\frac{m_\chi}{50 \text{ MeV}} \right) \times \left(\frac{10.75}{g_{*,\text{SM}}(T_{\text{SM}})} \right)^{1/2} \left(\frac{T_{\text{DR}}}{T_{\text{SM}}} \right)^4 \left(\frac{T_{\text{SM}}}{1 \text{ MeV}} \right)^2. \quad (20)$$

Thus, elastic scattering is inefficient during and after the freeze-out of DM, and the self-heating begins roughly after the freeze-out of DM.

We assumed that the ALP and dark gauge boson form a thermal bath with T_{DR} during the freeze-out of DM. The $XX \rightarrow XXX$ rate [40,41] is sufficiently large for $\alpha_X > \mathcal{O}(10^{-10})$. The $\phi X \rightarrow XX$ keeps ϕ in the thermal bath of X until $T_{\text{SM}} = 1 \text{ MeV}$ as [42,43]

$$\frac{\Gamma_{\phi X \rightarrow XX}}{H} \simeq 4 \times 10^3 \left(\frac{\alpha_X}{10^{-2}} \right)^3 \left(\frac{300 \text{ GeV}}{f} \right)^2 \left(\frac{10.75}{g_{*,\text{SM}}(T_{\text{SM}})} \right)^{1/2} \times \left(\frac{T_{\text{DR}}}{T_{\text{SM}}} \right)^3 \left(\frac{T_{\text{SM}}}{1 \text{ MeV}} \right). \quad (21)$$

An ALP decays when it is still semirelativistic with the thermally averaged decay rate of

$$\frac{\langle \Gamma_{\phi \rightarrow XX} \rangle}{H} \simeq 0.4 \left(\frac{\alpha_X}{10^{-2}} \right)^2 \left(\frac{m_\phi}{1 \text{ keV}} \right)^4 \left(\frac{300 \text{ GeV}}{f} \right)^2 \times \left(\frac{10.75}{g_{*,\text{SM}}(T_{\text{SM}})} \right)^{1/2} \left(\frac{T_{\text{DR}}}{T_{\text{SM}}} \right)^{-1} \left(\frac{1 \text{ keV}}{T_{\text{SM}}} \right)^3. \quad (22)$$

Non-Abelian dark radiation does not confine until the present Universe for $\alpha_X(T_{\text{DR}} = T_{\text{DR,fo}}) \lesssim 0.03(2/N_X)$, since the confinement scale is given by $\Lambda = \mu_0 \exp[-6\pi/(11N_X\alpha_{X,0})]$, where $\alpha_{X,0} = \alpha_X(T_{\text{DR}} = \mu_0)$. Neither the ALP nor dark radiation overcloses the Universe.

The model Lagrangian (14) does not contain any interaction that equilibrates the SM and dark sector. Indeed, if two sectors are in thermal equilibrium with each other in the early Universe, ΔN_{eff} tends to exceed unity, which is strongly disfavored by BBN and CMB. The inflaton can decay both into the SM sector and into the dark sector. In this case, $(T_{\text{DM}}/T_{\text{SM}})$ is determined by the branching ratio. Even if the inflaton predominantly decays into the SM sector, the dark sector can be populated through a feeble interaction to the SM sector while not being completely thermalized with SM particles. In addition to the interactions in Eq. (14), we may consider a Higgs portal coupling:

$$\mathcal{L}_{H\Phi} = \lambda_{H\Phi} |\Phi|^2 |H|^2. \quad (23)$$

The continuous production of the dark sector increases the temperature ratio between the dark sector and the SM plasma until the electroweak phase transition. The temperature ratio between the two sectors at the electroweak phase transition is estimated as

$$\left(\frac{T_{\text{DR}}}{T_{\text{SM}}} \right)_{\text{ew}} = \left(\frac{g_{*,\text{SM}} \rho_{\text{DR}}}{g_{*,\text{DR}} \rho_{\text{SM}}} \right)^{1/4} \simeq 0.5 \left(\frac{\lambda_{H\Phi}}{2.2 \times 10^{-6}} \right)^{1/2} \left(\frac{106.75}{g_{*,\text{SM,ew}}} \right)^{1/8} \left(\frac{83.5}{g_{*,\text{DR,ew}}} \right)^{1/4}, \quad (24)$$

where $g_{*,\text{DR,ew}} = 83.5$ takes into account all the d.o.f. of particles in Eq. (14). The contribution of the dark sector to ΔN_{eff} at the neutrino decoupling is given by

$$\begin{aligned}\Delta N_{\text{eff}} &= \frac{8}{7} \times \frac{1}{2} \times g_{*,s,\text{DR}} \times \left(\frac{T_{\text{DR}}}{T_{\text{SM}}} \right)_{\nu\text{-dec}}^4 \\ &\simeq 9.8 g_{*,s,\text{DR}}^{-1/3} \left(\frac{g_{*,s,\text{DR,ew}}}{83.5} \right)^{4/3} \left(\frac{T_{\text{DR}}}{T_{\text{SM}}} \right)_{\text{ew}}^4.\end{aligned}\quad (25)$$

After the freeze-out of DM but before the decay of the ALP, we find $g_{*,s,\text{DR}} = 7$. Thus, $(T_{\text{DR}}/T_{\text{SM}})|_{\text{ew}} \simeq 0.5$ is consistent with $\Delta N_{\text{eff}} < 3.4$ [2].

IV. CONCLUSION

Self-heating of semiannihilating DM can suppress subgalactic-scale structure formation when it lasts until the matter-radiation equality. The reduced number of dwarf-size halos can reconcile the possible tension between the CDM paradigm and the observation. The self-heating of sub-GeV DM is maintained until the matter-radiation equality for $\sigma_{\text{self}}/m_\chi \simeq \mathcal{O}(0.1 - 1)$ cm²/g. We have followed the evolution of cosmological perturbations and demonstrated that self-heating sub-GeV DM indeed leaves a cutoff on the subgalactic scale of the linear matter power spectrum.

It is interesting that self-heating DM interrelates a subgalactic cutoff in the linear matter power spectrum and a kiloparsec core of the DM distribution in halos through the thermalization of DM particles. We can take full advantage of astrophysical and cosmological searches of WDM and SIDM to probe self-heating DM. For example, we could tighten the range of the self-interaction strength by analyzing line-of-sight velocity dispersions of dwarf spheroidal galaxies [44] and rotation curves of low-surface brightness galaxies [45] with a larger number of samples. The satellite number counts restrict the cutoff in the linear matter power spectrum [46,47]. The matter distribution smoother than the CDM prediction will be tested by the perturbations on strongly lensed systems [48–51]. The top-down structure formation in contrast to the bottom-up one in the CDM paradigm is tested by the further discoveries of high- z galaxies [52–54] and by multiple probes of the reionization

epoch such as the 21 cm brightness temperature [55–57] and its fluctuations due to minihalos [58].

We have proposed an extension of the SIMP model with an ALP and dark radiation as a particle physics realization of sub-GeV self-heating DM. DM pions semiannihilate into an ALP, which decays into dark radiation. Dark radiation and an ALP forms thermal equilibrium. We have shown that self-heating can be realized for a certain range of model parameters. When the dark sector is populated from the SM sector through a Higgs portal, we can produce the dark sector particles while being consistent with the constraints from BBN and CMB measurements.

We have focused on a velocity-independent self-scattering cross section. On the other hand, the self-scattering cross section diminishing with an increasing velocity may be favored by the constraints from galaxy cluster ellipticities and bullet clusters. One way to realize the velocity-dependent self-scattering cross section is to introduce a light mediator coupling to two DM particles. Extending our discussion to such a case will be intriguing.

ACKNOWLEDGMENTS

We thank Toyokazu Sekiguchi for useful discussions in the early stage of this work. We also thank Andrew Spray for carefully reading our manuscript. This work was supported by IBS under Project Code No. IBS-R018-D1.

APPENDIX A: EVOLUTION EQUATIONS OF THE SELF-HEATING DM NUMBER DENSITY AND TEMPERATURE

The evolution of the DM phase-space distribution is governed by the Boltzmann equation

$$E_\chi \left[\frac{\partial f_\chi}{\partial t} - H p_\chi \frac{\partial f_\chi}{\partial p_\chi} \right] = C[f_\chi], \quad (\text{A1})$$

with

$$\begin{aligned}C_{\text{semi}}[f_\chi] &= \int d\Pi_2 d\Pi_3 d\Pi_4 (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) |\mathcal{M}_{\text{semi}}|^2 [f_\chi(p_3) f_\phi(p_4) - f_\chi(p_1) f_\chi(p_2)] \\ &\quad + \frac{1}{4} \int d\Pi_2 d\Pi_3 d\Pi_4 (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) |\mathcal{M}_{\text{semi}}|^2 [f_\chi(p_3) f_\chi(p_4) - f_\chi(p_1) f_\phi(p_2)],\end{aligned}\quad (\text{A2})$$

$$C_{\text{self}}[f_\chi] = \frac{1}{2} \int d\Pi_2 d\Pi_3 d\Pi_4 (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) |\mathcal{M}_{\text{self}}|^2 [f_\chi(p_3) f_\chi(p_4) - f_\chi(p_1) f_\chi(p_2)]. \quad (\text{A3})$$

$|\mathcal{M}_{\text{semi}}|^2$ and $d\Pi = d^3\vec{p}/(2\pi)^3/2E$ denote the invariant amplitude squared and the invariant phase space measure, respectively. Symmetry factors are multiplied in each integral in order not to overcount the phase space of identical particles.

We derive the evolution equations of the DM number density n_χ and temperature T_χ from the number and

energy conservation equations. We remark that self-scattering does not contribute to these equations, because it conserves the number and energy of DM particles. Integrating the Boltzmann equation with $1/E_\chi$ over the phase space of DM, we find the evolution equation of n_χ as

$$\begin{aligned}\dot{n}_\chi + 3Hn_\chi &= \frac{1}{2} \int \prod_{i=1}^4 d\Pi_i (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) |\mathcal{M}_{\text{semi}}|^2 [f_\chi(p_3)f_\phi(p_4) - f_\chi(p_1)f_\chi(p_2)], \\ &= -\langle \sigma_{\text{semi}} v_{\text{rel}} \rangle_{T_\chi T_\chi} [n_\chi^2 - n_\chi n_\chi^{\text{eq}}(T_\chi) \mathcal{J}(T_\chi, T_\phi)],\end{aligned}\quad (\text{A4})$$

where $\mathcal{J}(T_\chi, T_\phi)$ is defined as

$$\mathcal{J}(T_\chi, T_\phi) = \frac{n_\phi^{\text{eq}}(T_\phi) \langle \sigma_{\text{inv}} v_{\text{rel}} \rangle_{T_\chi T_\phi}}{n_\phi^{\text{eq}}(T_\chi) \langle \sigma_{\text{inv}} v_{\text{rel}} \rangle_{T_\chi T_\phi = T_\chi}}. \quad (\text{A5})$$

In this derivation, we have used $(n_\chi^{\text{eq}})^2 \langle \sigma_{\text{semi}} v_{\text{rel}} \rangle_{T_\chi T_\chi} = n_\chi^{\text{eq}} n_\phi^{\text{eq}} \langle \sigma_{\text{inv}} v_{\text{rel}} \rangle_{T_\chi T_\chi}$. The subscript in $\langle \sigma_{\text{semi}} v_{\text{rel}} \rangle_{T_\chi T_\chi}$ denotes that this thermal average is defined as a thermal average with respect to the Boltzmann distribution with temperature T_χ . Because of the explicit dependence on T_χ , the equation for the number density is not closed by itself. To correctly describe the evolution of the system, one must track the DM temperature evolution as well.

To obtain the evolution equation of the DM temperature, we integrate the Boltzmann equation without any weight. We find the evolution equation of the energy density:

$$\begin{aligned}\dot{\rho}_\chi + 3H(\rho_\chi + P_\chi) &= \frac{1}{2} \int \prod_{i=1}^4 d\Pi_i (E_1 + E_2 - E_3)(2\pi)^4 \\ &\quad \times \delta^{(4)}(p_1 + p_2 - p_3 - p_4) |\mathcal{M}_{\text{semi}}|^2 \\ &\quad \times [f_\chi(p_3)f_\phi(p_4) - f_\chi(p_1)f_\chi(p_2)],\end{aligned}\quad (\text{A6})$$

with pressure P_χ . Since $\rho_\chi = \langle E_\chi \rangle n_\chi$ and $P_\chi = \langle p_\chi^2 / (3E_\chi) \rangle n_\chi = T_\chi n_\chi$, the above equation reads as

$$\begin{aligned}\dot{\langle E_\chi \rangle} + 3HT_\chi &= -\frac{n_\phi^{\text{eq}}(T_\phi)}{n_\chi^{\text{eq}}(T_\chi)} \langle \Delta E \sigma_{\text{inv}} v_{\text{rel}} \rangle_{T_\chi T_\chi} \\ &\quad \times [n_\chi - n_\chi^{\text{eq}}(T_\chi) \mathcal{K}(T_\chi, T_\phi)],\end{aligned}\quad (\text{A7})$$

where $\Delta E = E_\phi - \langle E_\chi \rangle_{T_\chi}$ and the function $\mathcal{K}(T_\chi, T_\phi)$ is defined as

$$\mathcal{K}(T_\chi, T_\phi) = \frac{n_\phi^{\text{eq}}(T_\phi) \langle \Delta E \sigma_{\text{inv}} v_{\text{rel}} \rangle_{T_\chi T_\phi}}{n_\phi^{\text{eq}}(T_\chi) \langle \Delta E \sigma_{\text{inv}} v_{\text{rel}} \rangle_{T_\chi T_\phi = T_\chi}}. \quad (\text{A8})$$

Using $\dot{\langle E_\chi \rangle} = (\dot{T}_\chi / T_\chi^2) (\langle E_\chi^2 \rangle - \langle E_\chi \rangle^2) = (\dot{T}_\chi / T_\chi^2) \sigma_E^2$, one obtains

$$\begin{aligned}\frac{\dot{T}_\chi}{T_\chi} + 3H \frac{T_\chi}{\sigma_E^2} &= -\frac{1}{\sigma_E^2} \frac{n_\phi^{\text{eq}}(T_\phi)}{n_\chi^{\text{eq}}(T_\chi)} \langle \Delta E \sigma_{\text{inv}} v_{\text{rel}} \rangle_{T_\chi T_\chi} \\ &\quad \times [n_\chi - n_\chi^{\text{eq}}(T_\chi) \mathcal{K}(T_\chi, T_\phi)].\end{aligned}\quad (\text{A9})$$

Figure 1 presents our numerical result of the coevolution of n_χ and T_χ . One can see that the freeze-out of DM yield Y_χ proceeds in a similar way to a usual discussion found in Refs. [21–23]. The DM yield is estimated as

$$Y_{\chi, \infty} \simeq \frac{H}{s \langle \sigma_{\text{semi}} v_{\text{rel}} \rangle} \Big|_{T_\phi = T_{\phi, \text{fo}}}, \quad (\text{A10})$$

while determining the freeze-out temperature is ambiguous in our case. In the case of $T_\chi = T_\phi$, the freeze-out temperature is usually determined by $\Delta(x_{\text{fo}}) = Y_\chi - Y_\chi^{\text{eq}} = c Y_\chi^{\text{eq}}(x_{\text{fo}})$ with c being a numerical constant of order unity, where $\Delta = (d \ln Y_\chi^{\text{eq}} / dx) / (-s \langle \sigma_{\text{semi}} v_{\text{rel}} \rangle / H)$. In the self-heating scenario, however, the DM freeze-out is delayed, because the DM temperature shortly increases relative to T_ϕ and enhances the backward semiannihilation process. This can be seen from Fig. 4, where we present numerically computed $\Delta = Y_\chi - \mathcal{J} Y_\chi^{\text{eq}}$ as a function of m_χ / T_ϕ . This delay results only in an $\mathcal{O}(10)\%$ change of the final DM relic abundance.

It may be nontrivial why semiannihilation still affects the evolution of T_χ even after the kinetic decoupling. To illustrate the temperature evolution more clearly, it is useful to consider the thermodynamics of the SM bath and DM bath. Here we assume $T_\phi = T_{\text{SM}}$. From the first law of thermodynamics, we find

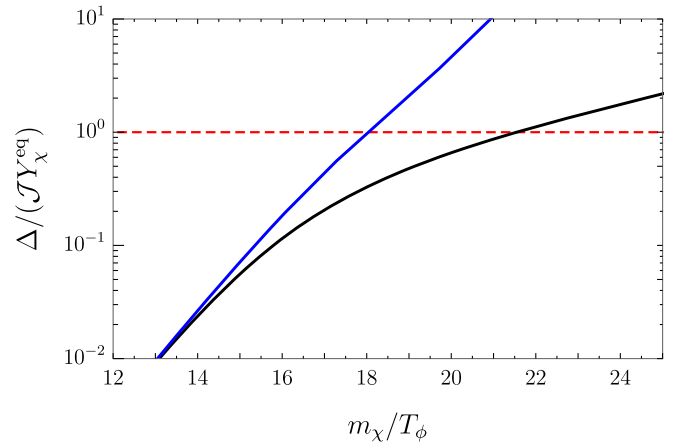


FIG. 4. $\Delta / (\mathcal{J} Y_\chi^{\text{eq}})$ in the self-heating DM scenario (black line). Here we assume $T_\phi = T_{\text{SM}}$ for simplicity. Compared to the $T_\chi = T_{\text{SM}}$ case (blue line), the freeze-out is delayed.

$$d(\rho_{\text{SM}}V) = dQ_{\text{SM}} - p_{\text{SM}}dV, \quad (\text{A11})$$

$$d(\rho_\chi V) = dQ_\chi - p_\chi dV + m_\chi dN_\chi. \quad (\text{A12})$$

Since we are interested in the dynamics after the freeze-out, we consider only forward semiannihilation $\chi\chi \rightarrow \chi\phi$. For a single forward semiannihilation process,

$$dQ_{\text{SM}} = -(2 - \gamma)m_\chi dN_\chi, \quad (\text{A13})$$

$$dQ_\chi = -(\gamma - 1)m_\chi dN_\chi, \quad (\text{A14})$$

where $\gamma = (5/4)[1 - m_\phi^2/(5m_\chi^2)]$ is a Lorentz boost factor. We see that the total energy of the whole system is conserved: $dQ_{\text{SM}} + dQ_\chi + m_\chi dN_\chi = 0$. To investigate how the DM temperature evolves, we note that $\rho_\chi = (m_\chi + 3T_\chi/2)n_\chi$ and $p_\chi = T_\chi n_\chi$ in the nonrelativistic limit of DM. In addition, a small fraction of DM undergoes semiannihilation even after the freeze-out, $d \ln N_\chi \simeq -(\Gamma_{\text{semi}}/H)d \ln a$. We find

$$d \ln T_\chi \simeq \left[-2 + (\gamma - 1) \frac{2m_\chi \Gamma_{\text{semi}}}{3T_\chi H} \right] d \ln a. \quad (\text{A15})$$

The first term in the square brackets represents the adiabatic cooling due to the expansion of the Universe, while the second term is due to the energy injection through semiannihilation. If $T_\chi \propto 1/a^2$ as for free-streaming nonrelativistic DM particles, then the second term increases with the expansion of the Universe as $\propto (m_\chi/T_\chi)(\Gamma_{\text{semi}}/H) \propto a$ and thus heats the DM particles. Hence, the DM temperature is determined by the balance between the adiabatic cooling and semiannihilation heating as $T_\chi \propto 1/a$.

APPENDIX B: EVOLUTION EQUATIONS OF SELF-HEATING DM COSMOLOGICAL PERTURBATIONS

We derive the evolution equations of cosmological perturbations. Taking the conformal Newtonian gauge [see Eq. (9)], we get the linearized Boltzmann equation given by

$$\begin{aligned} \delta f' + i \left(\frac{\vec{k} \cdot \vec{q}}{\epsilon} \right) \delta f - \left[\Phi' + i(\vec{k} \cdot \vec{q}) \frac{\epsilon}{q} \Psi \right] \frac{\partial \bar{f}}{\partial \ln q} \\ = a \left(\frac{1}{E} C^{(1)} + \frac{1}{E} C^{(0)} \Psi \right). \end{aligned} \quad (\text{B1})$$

We expand δf as a function of $\mu = \hat{k} \cdot \hat{q}$ in terms of the Legendre polynomial:

$$\delta f(\tau, \vec{k}, q, \hat{q}) = \sum_{\ell=0}^{\infty} (-i)^\ell (2\ell + 1) F_\ell(\tau, k, q) P_\ell(\mu). \quad (\text{B2})$$

Multiplying the Legendre polynomial by the linearized Boltzmann equation and integrating it with respect to μ , we find the Boltzmann hierarchy as

$$\begin{aligned} F'_0 = -\frac{kq}{\epsilon} F_1 + \Phi' \frac{\partial \bar{f}}{\partial \ln q} \\ + a \int_{-1}^1 \frac{d\mu}{2} \left(\frac{1}{E} C^{(1)} + \frac{1}{E} C^{(0)} \Psi \right), \end{aligned} \quad (\text{B3})$$

$$\begin{aligned} F'_1 = -\frac{kq}{3\epsilon} (2F_2 - F_0) - \frac{k\epsilon}{3q} \Psi \frac{\partial \bar{f}}{\partial \ln q} \\ + ia \int_{-1}^1 \frac{d\mu}{2} P_1(\mu) \left(\frac{1}{E} C^{(1)} + \frac{1}{E} C^{(0)} \Psi \right), \end{aligned} \quad (\text{B4})$$

$$\begin{aligned} F'_2 = -\frac{kq}{5\epsilon} (3F_3 - 2F_1) \\ - a \int_{-1}^1 \frac{d\mu}{2} P_2(\mu) \left(\frac{1}{E} C^{(1)} + \frac{1}{E} C^{(0)} \Psi \right), \end{aligned} \quad (\text{B5})$$

$$\begin{aligned} F'_\ell (\ell \geq 3) = -\frac{kq}{(2\ell + 1)\epsilon} [(\ell + 1)F_{\ell+1} - \ell F_{\ell-1}] \\ + (-i)^{-\ell} a \int_{-1}^1 \frac{d\mu}{2} P_\ell(\mu) \left(\frac{1}{E} C^{(1)} + \frac{1}{E} C^{(0)} \Psi \right). \end{aligned} \quad (\text{B6})$$

In terms of the DM fluid variables [59], one obtains

$$\begin{aligned} \delta'_\chi = -3 \left(c_{s\chi}^2 - \omega_\chi + \frac{a}{3\mathcal{H}\bar{\rho}_\chi} \int \frac{d^3 \vec{p}}{(2\pi)^3} E \frac{1}{E} C^{(0)} \right) \mathcal{H} \delta_\chi \\ - 3\mathcal{H}\pi_\chi - (1 + \omega_\chi)(\theta_\chi + 3\Phi') \\ + \frac{a}{\bar{\rho}_\chi} \int \frac{d^3 \vec{p}}{(2\pi)^3} E \left(\frac{1}{E} C^{(1)} + \frac{1}{E} C^{(0)} \Psi \right), \end{aligned} \quad (\text{B7})$$

$$\begin{aligned} \theta'_\chi = -(1 - 3\omega_\chi) \mathcal{H} \theta_\chi - \frac{\omega'_\chi}{1 + \omega_\chi} \theta_\chi - k^2 \sigma_\chi + k^2 \Psi \\ + \frac{1}{1 + \omega_\chi} k^2 (c_{s\chi}^2 \delta_\chi + \pi_\chi) \\ + \frac{iak}{\bar{\rho}_\chi (1 + \omega_\chi)} \int \frac{d^3 \vec{p}}{(2\pi)^3} p P_1(\mu) \left(\frac{1}{E} C^{(1)} + \frac{1}{E} C^{(0)} \Psi \right), \end{aligned} \quad (\text{B8})$$

$$\begin{aligned} \pi'_\chi = & - \left(2 - 3c_s^2 - 3\omega_\chi - \alpha_0\omega_\chi + \frac{a}{3\mathcal{H}\bar{\rho}_\chi} \int \frac{d^3\vec{p}}{(2\pi)^3} E \frac{1}{E} C^{(0)} \right) \mathcal{H}\pi_\chi - \left(\frac{c_{s\chi}^{2\prime}}{\mathcal{H}c_{s\chi}^2} + 2 - 3c_{s\chi}^2 - \alpha_0\omega_\chi \right) c_{s\chi}^2 \mathcal{H}\delta_\chi \\ & + \left(c_{s\chi}^2 - \frac{\alpha_2\omega_\chi}{3} \right) (1 + \omega_\chi)\theta + 3 \left[c_{s\chi}^2 - \frac{5\omega_\chi}{3} + \omega_\chi \left(c_{s\chi}^2 + \frac{\alpha_3\omega_\chi}{3} \right) \right] \Phi' + \frac{a}{\bar{\rho}_\chi} \int \frac{d^3\vec{p}}{(2\pi)^3} \left(\frac{p^2}{3E} - c_s^2 E \right) \left(\frac{1}{E} C^{(1)} + \frac{1}{E} C^{(0)} \Psi \right), \end{aligned} \quad (\text{B9})$$

$$\sigma'_\chi = -(2 - 3\omega - \alpha_1\omega)\mathcal{H}\sigma_\chi + \frac{4}{15}\alpha_2\omega_\chi\theta_\chi - \frac{2k}{5\bar{\rho}_\chi(1 + \omega_\chi)} \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{p^3}{E^2} F_3 - \frac{2a}{3\bar{\rho}_\chi(1 + \omega_\chi)} \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{p^2}{E} P_2(\mu) \left(\frac{1}{E} C^{(1)} + \frac{1}{E} C^{(0)} \Psi \right). \quad (\text{B10})$$

We have decomposed the pressure perturbation into the isentropic and entropic parts as $\delta P_\chi = \bar{\rho}_\chi(c_{s\chi}^2\delta_\chi + \pi_\chi)$, where the adiabatic sound speed squared is given by

$$c_{s\chi}^2 = \frac{\bar{P}'_\chi}{\bar{\rho}'_\chi} = \omega_\chi + \bar{\rho}_\chi \frac{\omega'_\chi}{\bar{\rho}'_\chi}. \quad (\text{B11})$$

The dimensionless constants are defined as

$$\alpha_0 = \left(\int \frac{d^3\vec{p}}{(2\pi)^3} \frac{p^2}{E} \frac{p^2}{E^2} F_0 \right) / \left(\omega \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{p^2}{E} F_0 \right), \quad (\text{B12})$$

$$\alpha_1 = \left(\int \frac{d^3\vec{p}}{(2\pi)^3} \frac{p^2}{E} \frac{p^2}{E^2} F_2 \right) / \left(\omega \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{p^2}{E} F_2 \right), \quad (\text{B13})$$

$$\alpha_2 = \left(\int \frac{d^3\vec{p}}{(2\pi)^3} p \frac{p^2}{E^2} F_1 \right) / \left(\omega \int \frac{d^3\vec{p}}{(2\pi)^3} p F_1 \right). \quad (\text{B14})$$

We consider cosmological perturbations entering the horizon well after the freeze-out of the DM number density. $\omega_\chi \simeq (T_\chi/m_\chi)$, $c_{s\chi}^2 \simeq (T_\chi/m_\chi)[1 - d \ln T_\chi / (3d \ln a)]$, and F_ℓ ($\ell \geq 2$) are suppressed by the low DM velocity. Furthermore, F_ℓ ($\ell \geq 2$) are erased by self-scattering of DM. In this limit, we find Eqs. (6)–(8). We have substituted $\alpha_2 = 5$, which can be derived as follows.

In general, $F_\ell(\tau, k, q)$ can be expanded by a complete system of functions of q . We can expand $F_\ell(\tau, k, q) = \sum_n 1/(2\pi a^2 m_\chi T_\chi)^{2/3} y^{\ell/2} L_n^{\ell+1/2}(y) e^{-y} \mathcal{F}_{n\ell}$ with $y = q^2/(2\pi a^2 m_\chi T_\chi)$ and L_n^α being the associated Legendre function, as in Refs. [26,60]. An advantage of this complete system is that $n = 0$ gives the dominant contribution in the nonrelativistic limit. Since $L_0^\alpha = 1$, one can find $\alpha_2 = 5$. One may also calculate other α 's in a similar manner, while they do not appear in Eqs. (6)–(8) and thus are not given here. One solves Eqs. (6)–(8) simultaneously with

$$\omega'_\chi = -2\mathcal{H}\omega_\chi + \frac{2}{3}(\gamma - 1)a\Gamma_{\text{semi}}, \quad (\text{B15})$$

$$c_{s\chi}^2 = \frac{5}{3}\omega_\chi - \frac{2}{9}(\gamma - 1)\frac{a\Gamma_{\text{semi}}}{\mathcal{H}}, \quad (\text{B16})$$

which can be directly obtained from Eq. (2). These relations are valid as long as $\Gamma_{\text{self}}/H \gtrsim 1$.

By taking the synchronous gauge,

$$ds^2 = a^2(\tau)[-d\tau^2 + (\delta_{ij} + h_{ij})dx^i dx^j], \quad (\text{B17})$$

one can derive the evolution equations given as

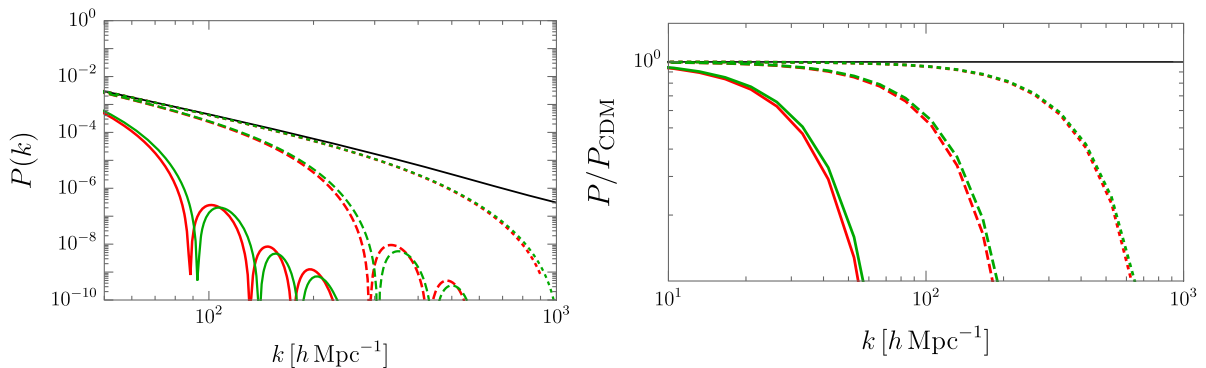


FIG. 5. (Left) Linear matter power spectrum when the decoupling of self-scattering is described by $(1 - e^{-\Gamma_{\text{self}}/H})$ (red lines) and when it is described by $e^{-a/a_{\text{self}}}$ (green lines). (Right) Power spectrum in self-heating DM scenario relative to CDM.

$$\delta'_\chi = -\theta_\chi - \frac{h'}{2}, \quad (\text{B18})$$

$$\theta'_\chi = -\mathcal{H}\theta_\chi + k^2(c_{s\chi}^2\delta_\chi + \pi_\chi), \quad (\text{B19})$$

$$\pi'_\chi = -2\mathcal{H}\pi_\chi + \left(c_{s\chi}^2 - \frac{5}{3}\omega_\chi\right) \left[-\mathcal{H}\left(1 - \frac{\mathcal{H}'}{\mathcal{H}^2}\right)\delta_\chi + \theta_\chi + \frac{h'}{2}\right], \quad (\text{B20})$$

where h denotes the trace of h_{ij} . They are equivalent to those in the conformal Newtonian gauge under the gauge transformation [see Eq. (27) of Ref. [59]]. To check the consistency, one needs to note that $c_s^2\mathcal{H}k^2\alpha = c_s^2\mathcal{H}(h' + 6\eta')/2$ is negligible when compared to θ'_χ for nonrelativistic DM.

After decoupling of self-scattering [see Eq. (5)], energetic DM particles through semiannihilation freely stream, while the majority of DM follows the Maxwell-Boltzmann distribution with the temperature $T_\chi \propto 1/a^2$. After that,

Eqs. (6)–(8) still describe the evolution of the majority of DM, although Eqs. (B15) and (B16) are no longer valid. For the decoupling of self-scattering, we take a phenomenological approach, i.e., multiplying $(1 - e^{-\Gamma_{\text{self}}/H})$ to the equations of ω_χ and $c_{s\chi}^2$ [see Eqs. (10) and (11)]. We may introduce an alternative cutoff, for instance, like

$$\omega'_\chi = -2\mathcal{H}\omega_\chi + \frac{2}{3}(\gamma - 1)a\Gamma_{\text{semi}} \exp\left[-\left(\frac{a}{a_{\text{self}}}\right)\right], \quad (\text{B21})$$

$$c_{s\chi}^2 = \frac{5}{3}\omega_\chi - \frac{2}{9}(\gamma - 1)\frac{a\Gamma_{\text{semi}}}{\mathcal{H}} \exp\left[-\left(\frac{a}{a_{\text{self}}}\right)\right], \quad (\text{B22})$$

where a_{self} is the scale factor when $\Gamma_{\text{self}} = H$. In this case, the decoupling takes place more rapidly compared to the case presented in the main text, and the cutoff in the matter power spectrum appears for larger k . See Fig. 5 for the numerical difference in the two descriptions.

-
- [1] M. R. Buckley and A. H. G. Peter, Gravitational probes of dark matter physics, [arXiv:1712.06615](https://arxiv.org/abs/1712.06615).
- [2] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2015 results. XIII. Cosmological parameters, *Astron. Astrophys.* **594**, A13 (2016).
- [3] S. Alam *et al.* (BOSS Collaboration), The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological analysis of the DR12 galaxy sample, *Mon. Not. R. Astron. Soc.* **470**, 2617 (2017).
- [4] J. S. Bullock and M. Boylan-Kolchin, Small-scale challenges to the Λ CDM paradigm, *Annu. Rev. Astron. Astrophys.* **55**, 343 (2017).
- [5] T. Sawala *et al.*, The APOSTLE simulations: Solutions to the Local Group's cosmic puzzles, *Mon. Not. R. Astron. Soc.* **457**, 1931 (2016).
- [6] A. A. Dutton, A. V. Macciò, J. Frings, L. Wang, G. S. Stinson, C. Penzo, and X. Kang, NIHAO V: Too big does not fail reconciling the conflict between Λ CDM predictions and the circular velocities of nearby field galaxies, *Mon. Not. R. Astron. Soc.* **457**, L74 (2016).
- [7] A. R. Wetzel, P. F. Hopkins, J.-h. Kim, C.-A. Faucher-Giguère, D. Kereš, and E. Quataert, Reconciling dwarf galaxies with Λ CDM cosmology: Simulating a realistic population of satellites around a Milky Way-mass galaxy, *Astrophys. J.* **827**, L23 (2016).
- [8] B. Moore, S. Ghigna, F. Governato, G. Lake, T. R. Quinn, J. Stadel, and P. Tozzi, Dark matter substructure within galactic halos, *Astrophys. J.* **524**, L19 (1999).
- [9] A. V. Kravtsov, Dark matter substructure and dwarf galactic satellites, *Adv. Astron.* **2010**, 281913 (2010).
- [10] P. Bode, J. P. Ostriker, and N. Turok, Halo formation in warm dark matter models, *Astrophys. J.* **556**, 93 (2001).
- [11] M. R. Lovell, V. Eke, C. S. Frenk, L. Gao, A. Jenkins, T. Theuns, J. Wang, S. D. M. White, A. Boyarsky, and O. Ruchayskiy, The haloes of bright satellite galaxies in a warm dark matter universe, *Mon. Not. R. Astron. Soc.* **420**, 2318 (2012).
- [12] R. Adhikari *et al.*, A White Paper on keV sterile neutrino dark matter, *J. Cosmol. Astropart. Phys.* **01** (2017) 025.
- [13] B. Moore, T. Quinn, F. Governato, J. Stadel, and G. Lake, Cold collapse and the core catastrophe, *Mon. Not. R. Astron. Soc.* **310**, 1147 (1999).
- [14] W. J. G. de Blok, The core-cusp problem, *Adv. Astron.* **2010**, 789293 (2010).
- [15] K. A. Oman *et al.*, The unexpected diversity of dwarf galaxy rotation curves, *Mon. Not. R. Astron. Soc.* **452**, 3650 (2015).
- [16] D. N. Spergel and P. J. Steinhardt, Observational Evidence for Self-Interacting Cold Dark Matter, *Phys. Rev. Lett.* **84**, 3760 (2000).
- [17] M. Rocha, A. H. G. Peter, J. S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, J. Oñorbe, and L. A. Moustakas, Cosmological simulations with self-interacting dark matter - I. Constant density cores and substructure, *Mon. Not. R. Astron. Soc.* **430**, 81 (2013).
- [18] M. Kaplinghat, S. Tulin, and H.-B. Yu, Dark Matter Halos as Particle Colliders: Unified Solution to Small-Scale Structure Puzzles from Dwarfs to Clusters, *Phys. Rev. Lett.* **116**, 041302 (2016).

- [19] A. Kamada, M. Kaplinghat, A. B. Pace, and H.-B. Yu, Self-Interacting Dark Matter Can Explain Diverse Galactic Rotation Curves, *Phys. Rev. Lett.* **119**, 111102 (2017).
- [20] A. Kamada, H. J. Kim, H. Kim, and T. Sekiguchi, Self-Heating Dark Matter via Semiannihilation, *Phys. Rev. Lett.* **120**, 131802 (2018).
- [21] T. Hambye, Hidden vector dark matter, *J. High Energy Phys.* **01** (2009) 028.
- [22] F. D’Eramo and J. Thaler, Semi-annihilation of dark matter, *J. High Energy Phys.* **06** (2010) 109.
- [23] G. Bélanger, K. Kannike, A. Pukhov, and M. Raidal, Impact of semi-annihilations on dark matter phenomenology. An example of Z_N symmetric scalar dark matter, *J. Cosmol. Astropart. Phys.* **04** (2012) 010.
- [24] E. D. Carlson, M. E. Machacek, and L. J. Hall, Self-interacting dark matter, *Astrophys. J.* **398**, 43 (1992).
- [25] Y. Hochberg, E. Kuflik, T. Volansky, and J. G. Wacker, Mechanism for Thermal Relic Dark Matter of Strongly Interacting Massive Particles, *Phys. Rev. Lett.* **113**, 171301 (2014).
- [26] T. Binder, L. Covi, A. Kamada, H. Murayama, T. Takahashi, and N. Yoshida, Matter power spectrum in hidden neutrino interacting dark matter models: A closer look at the collision term, *J. Cosmol. Astropart. Phys.* **11** (2016) 043.
- [27] P. Gondolo and G. Gelmini, Cosmic abundances of stable particles: Improved analysis, *Nucl. Phys.* **B360**, 145 (1991).
- [28] S. Tulin and H.-B. Yu, Dark matter self-interactions and small scale structure, *Phys. Rep.* **730**, 1 (2018).
- [29] X. Chu and C. Garcia-Cely, Core formation from self-heating dark matter, [arXiv:1803.09762](https://arxiv.org/abs/1803.09762).
- [30] G. Arcadi, M. Dutra, P. Ghosh, M. Lindner, Y. Mambrini, M. Pierre, S. Profumo, and F. S. Queiroz, The waning of the WIMP? A review of models, searches, and constraints, *Eur. Phys. J. C* **78**, 203 (2018).
- [31] L. Roszkowski, E. M. Sessolo, and S. Trojanowski, WIMP dark matter candidates and searches—Current status and future prospects, *Rep. Prog. Phys.* **81**, 066201 (2018).
- [32] R. T. D’Agnolo and J. T. Ruderman, Light Dark Matter from Forbidden Channels, *Phys. Rev. Lett.* **115**, 061301 (2015).
- [33] A. Kamada, H. Kim, and T. Sekiguchi, Axionlike particle assisted strongly interacting massive particle, *Phys. Rev. D* **96**, 016007 (2017).
- [34] D. Blas, J. Lesgourgues, and T. Tram, The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes, *J. Cosmol. Astropart. Phys.* **07** (2011) 034.
- [35] A. Kamada, N. Yoshida, K. Kohri, and T. Takahashi, Structure of dark matter halos in warm dark matter models and in models with long-lived charged massive particles, *J. Cosmol. Astropart. Phys.* **03** (2013) 008.
- [36] V. Iršič *et al.*, New constraints on the free-streaming of warm dark matter from intermediate and small scale Lyman- α forest data, *Phys. Rev. D* **96**, 023522 (2017).
- [37] K. J. Bae, A. Kamada, S. P. Liew, and K. Yanagi, Light axinos from freeze-in: Production processes, phase space distributions, and Ly- α forest constraints, *J. Cosmol. Astropart. Phys.* **01** (2018) 054.
- [38] A. Kamada, K. T. Inoue, K. Kohri, and T. Takahashi, Constraints on long-lived electrically charged massive particles from anomalous strong lens systems, *J. Cosmol. Astropart. Phys.* **11** (2017) 008.
- [39] Y. Hochberg, E. Kuflik, H. Murayama, T. Volansky, and J. G. Wacker, Model for Thermal Relic Dark Matter of Strongly Interacting Massive Particles, *Phys. Rev. Lett.* **115**, 021301 (2015).
- [40] T. S. Biró, E. van Doorn, B. Müller, M. H. Thoma, and X.-N. Wang, Parton equilibration in relativistic heavy ion collisions, *Phys. Rev. C* **48**, 1275 (1993).
- [41] Z. Xu and C. Greiner, Shear Viscosity in a Gluon Gas, *Phys. Rev. Lett.* **100**, 172301 (2008).
- [42] E. Massó, F. Rota, and G. Zsembinszki, Axion thermalization in the early universe, *Phys. Rev. D* **66**, 023004 (2002).
- [43] P. Graf and F. D. Steffen, Thermal axion production in the primordial quark-gluon plasma, *Phys. Rev. D* **83**, 075011 (2011).
- [44] M. Valli and H.-B. Yu, Dark matter self-interactions from the internal dynamics of dwarf spheroidals, [arXiv:1711.03502](https://arxiv.org/abs/1711.03502).
- [45] K. Bondarenko, A. Boyarsky, T. Bringmann, and A. Sokolenko, Constraining self-interacting dark matter with scaling laws of observed halo surface densities, *J. Cosmol. Astropart. Phys.* **04** (2018) 049.
- [46] E. Polisensky and M. Ricotti, Constraints on the dark matter particle mass from the number of Milky Way satellites, *Phys. Rev. D* **83**, 043506 (2011).
- [47] S. Y. Kim, A. H. G. Peter, and J. R. Hargis, There is no missing satellites problem, [arXiv:1711.06267](https://arxiv.org/abs/1711.06267).
- [48] K. T. Inoue, R. Takahashi, T. Takahashi, and T. Ishiyama, Constraints on warm dark matter from weak lensing in anomalous quadruple lenses, *Mon. Not. R. Astron. Soc.* **448**, 2704 (2015).
- [49] A. Kamada, K. T. Inoue, and T. Takahashi, Constraints on mixed dark matter from anomalous strong lens systems, *Phys. Rev. D* **94**, 023522 (2016).
- [50] S. Birrer, A. Amara, and A. Refregier, Lensing substructure quantification in RXJ1131-1231: A 2 keV lower bound on dark matter thermal relic mass, *J. Cosmol. Astropart. Phys.* **05** (2017) 037.
- [51] D. Gilman, S. Birrer, T. Treu, and C. R. Keeton, Probing the nature of dark matter by forward modeling flux ratios in strong gravitational lenses, [arXiv:1712.04945](https://arxiv.org/abs/1712.04945).
- [52] R. Barkana, Z. Haiman, and J. P. Ostriker, Constraints on warm dark matter from cosmological reionization, *Astrophys. J.* **558**, 482 (2001).
- [53] F. Pacucci, A. Mesinger, and Z. Haiman, Focusing on warm dark matter with lensed high-redshift galaxies, *Mon. Not. R. Astron. Soc.* **435**, L53 (2013).
- [54] M. R. Lovell, J. Zavala, M. Vogelsberger, X. Shen, F.-Y. Cyr-Racine, C. Pfrommer, K. Sigurdson, M. Boylan-Kolchin, and A. Pillepich, ETHOS—An effective theory of structure formation: Predictions for the high-redshift Universe—Abundance of galaxies and reionization, *Mon. Not. R. Astron. Soc.* **477**, 2886 (2018).
- [55] M. Sitwell, A. Mesinger, Y.-Z. Ma, and K. Sigurdson, The imprint of warm dark matter on the cosmological 21-cm signal, *Mon. Not. R. Astron. Soc.* **438**, 2664 (2014).
- [56] M. Safarzadeh, E. Scannapieco, and A. Babul, A limit on the warm dark matter particle mass from the redshifted

- 21 cm absorption line, *Astrophys. J. Lett.* **859**, L18 (2018).
- [57] A. Schneider, Constraining non-cold dark matter models with the global 21-cm signal, [arXiv:1805.00021](https://arxiv.org/abs/1805.00021).
- [58] T. Sekiguchi and H. Tashiro, Constraining warm dark matter with 21 cm line fluctuations due to minihalos, *J. Cosmol. Astropart. Phys.* 08 (2014) 007.
- [59] C.-P. Ma and E. Bertschinger, Cosmological perturbation theory in the synchronous and conformal Newtonian gauges, *Astrophys. J.* **455**, 7 (1995).
- [60] E. Bertschinger, Effects of cold dark matter decoupling and pair annihilation on cosmological perturbations, *Phys. Rev. D* **74**, 063509 (2006).