Exploring the Universe with dark light scalars

Bugeon Jo[1,*], Hyeontae Kim,[1,†] Hyung Do Kim[1,‡] and Chang Sub Shin[2,§]

[1]*Department of Physics and Astronomy, Seoul National University, Seoul 08826, Korea*
[2]*Center for Theoretical Physics of the Universe, Institute for Basic Science (IBS), Daejeon 34051, Korea*

We study the cosmology of the dark sector consisting of (ultra)light scalars. Since the scalar mass is radiatively unstable, a special explanation is required to make the mass much smaller than the UV scale. There are two well-known mechanisms for the origin of scalar mass. The scalar can be identified as a pseudo-Goldstone boson, whose shift symmetry is explicitly broken by nonperturbative corrections, like the axion. Alternatively, it can be identified as a composite particle like the glueball, whose mass is limited by the confinement scale of the theory because no scalar degree of freedom exists at high scales. In both cases, the scalar can be naturally light, but interaction behavior is quite different. The lighter the axion (glueball), the weaker (stronger) its interaction. As the simplest nontrivial example, we consider the dark axion whose shift symmetry is anomalously broken by the hidden non-Abelian gauge symmetry. After the confinement of the gauge group, the dark axion and the dark glueball get masses and both form multicomponent dark matter. We carefully consider the effects of energy flow from the dark gluons to the dark axions and derive the full equations of motion for the background and the perturbed variables. The effect of the dark axion–dark gluon coupling on the evolution of the entropy and the isocurvature perturbations is also clarified. Finally, we discuss the gravothermal collapse of the glueball subcomponent dark matter after the halos form, in order to explore the potential to contribute to the formation of seeds for the supermassive black holes observed at high redshifts. With the simplified assumptions, the glueball subcomponent dark matter with the mass of 0.01–0.1 MeV and the axion main dark matter component with the decay constant $f_a = \mathcal{O}(10^{15}$–$10^{16})$ GeV and the mass of $\mathcal{O}(10^{-14}$–$10^{-18})$ eV can provide a hint on the origin of the supermassive black holes at high redshifts.

I. INTRODUCTION

The discovery of the Higgs boson completed the Standard Model, which explains most of the phenomena in the Universe including nuclei, atoms, and their interactions. However, various studies based on precise measurements of the cosmic microwave background (CMB), velocity distributions of stars and galaxies, and large-scale structures (LSS) show that only 4% of the Universe is understood by our knowledge of the Standard Model, and 96% must be filled with dark matter and dark energy, for which we do not know their origin [1].

Among dark matter candidates, particles classified as "weakly interacting massive particles" (WIMP) were considered as the best candidate. This is because their freeze-out relic abundance can naturally explain the present amount of dark matter, and they can be tested by different ongoing searches. Their mass scale can also be related to new physics that explains why the electroweak scale is stable against various quantum corrections.

However, there is no conclusive evidence of WIMP dark matter so far, and the alternative candidates got more attention in recent years. The masses of these candidates are not limited to a narrow range, and various ideas have been proposed, particularly focusing on the detection possibility [2]. For such a broad range of dark matter mass, there is no clear guiding principle to specify its natural range. Especially, it is extremely unnatural to consider a light scalar compared to the probing scales unless there is a special reason for it.

It is quite interesting to notice that an ultralight scalar dark matter is allowed by cosmological and astrophysical observations, because the scalar can act as an oscillating classical field whose averaged equation of state is the same as that of cold dark matter (CDM). Further interesting phenomena can arise from its field nature at small scales. The fuzzy dark matter [3] and the QCD axion [4–7] are

[*]whqnrjs@gmail.com
[†]htkim428@gmail.com
[‡]hdkim@phya.snu.ac.kr
[§]csshin@ibs.re.kr

good examples. Both candidates, the axion or axionlike particle, can be naturally light by their approximate shift symmetry, which is explicitly broken by the controllable nonperturbative corrections. The consequence of the global symmetry is that the lighter the axion becomes, the weaker its interaction is. Therefore, a large occupation number of the (ultra)light axion is allowed, and it can be described by the evolution of scalar condensate.

On the one hand, there is another natural way to obtain a light scalar dark matter. When the asymptotically free gauge group has confinement at low energy, the gauge fields are confined into scalar particles—glueballs—whose mass is limited by the confining scale. Unlike in the case of the axion, the lighter the glueballs, the stronger the scattering cross section among glueballs. The number-changing interactions are active, so the occupation number is always limited by its temperature during its cosmological evolution. In this case, the relic density is determined by the freeze-out mechanism. Because of this property, the light dark glueball becomes a good candidate for self-interacting dark matter (SIDM) [8,9], which is one of the ways to make the cored density profile around the centers of galaxies [10]. As a subcomponent hot dark matter, it can also play the role to suppress small-scale perturbations [11].

These two mechanisms to obtain a light scalar dark matter provide a completely different microscopic nature of dark matter and yield different predictions for small-scale evolution.

In this paper, we study the cosmology of light scalar dark matter, focusing on the origin of its mass and the consequences associated with it. We cover two mechanisms discussed above at the same time in a minimal setup: the dark sector, consisting of the axion, and the confining hidden gauge symmetry without light fermions. In this setup, the axion's shift symmetry is nonperturbatively broken by a Chern-Simons-type coupling between the axion and the dark gauge field—i.e., gluon. After the confinement of the gauge group, the axion and the glueball get masses and become a part of multicomponent dark matter. The idea that the mass of the ultralight scalar dark matter originates from a confining hidden gauge symmetry was studied in Refs. [12,13], but comprehensive study on its cosmological evolution is still lacking.

We derive a complete set of the equations of motion for the background and the perturbed variables of the dark axion and the dark gluon/glueball densities. Since the axion-gluon coupling provides energy transfer from the dark gluon to the dark axion, we clarify its effect on the thermodynamics of the dark gluon fluid and entropy evolution. We also quantify the transfer of the isocurvature perturbation through the same coupling.

The multicomponent dark matter which simultaneously contains feebly interacting and strongly interacting particles has an interesting cosmological consequence. When the glueball dark matter becomes a subcomponent, the stronger self-interaction between the glueballs is allowed and opens the possibility to form a black hole in the early Universe [14,15]. This may provide a possible answer to the question about the origin of observed supermassive black holes at high redshifts [16–19]. We discuss the parameter space to provide the solution and possible caveats.

The paper is organized as follows: In Sec. II, we establish basic formalism from the Lagrangian to the dynamical equations of the background and perturbation variables of the coupled axion-gluon fluid. In Sec. III, we focus on the background evolution of the glueball and the axion dark matter. The parametric dependence of the relic abundance of the glueball and the axion is also presented. Section IV is devoted to the evolution of the perturbed variables. For the initial conditions, we have three modes: adiabatic perturbation, isocurvature perturbation induced by the initial misalignment of the axion, and the temperature fluctuation of the dark gluon fluid. In Sec. V, we discuss the implication of the glueball subcomponent dark matter for the early formation of the supermassive black holes. Section VI presents conclusions.

II. AXION DARK MATTER AND CONFINING DARK SECTOR

A. General description of the model

Our starting Lagrangian of the dark sector is composed of the ultralight axion ϕ, whose field range is $2\pi f_a$, and the dark gauge symmetry with the confinement scale Λ ($\Lambda \ll f_a$). The coupling between the axion and the dark gauge bosons is given as

$$-\frac{\mathcal{L}_h}{\sqrt{-g}} = \frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{4}(G^a_{\mu\nu})^2 + \frac{g_h^2 \phi}{32\pi^2 f_a} G^a_{\mu\nu}\tilde{G}^{a\mu\nu}, \quad (2.1)$$

where $G^a_{\mu\nu}$ is the dark gluon field strength and g_h denotes the dark (hidden) gauge coupling. For illustration, $SU(N)$ is taken as our dark gauge group. Although we will not consider an extremely large value of N and only take $N^2 = \mathcal{O}(10)$ in concrete examples, we keep N-dependence explicitly in our discussion in order to organize the results using the large-N expansion.

Below the confinement scale, the dynamics of the gauge fields can be described by the composite bosons, the glueballs. The most relevant glueball for dark matter physics is the lightest glueball, φ_g. Considering the large-N limit (dominated by planar diagrams) and the 4π factor from the naive dimensional analysis (with the cutoff of the order of Λ), the effective Lagrangian of φ_g can be expanded in $(4\pi/N)(\varphi_g/m_g)$ as [20–24]

$$-\frac{\mathcal{L}_{heff}}{\sqrt{-g}} = \frac{1}{2}(\partial_\mu\phi)^2 + V(\phi) + \frac{1}{2}(\partial_\mu\varphi_g)^2 + \frac{1}{2}m_g^2\varphi_g^2$$

$$+ \frac{a_3}{3!}\left(\frac{4\pi}{N}\right)m_g\varphi_g^3 + \frac{a_4}{4!}\left(\frac{4\pi}{N}\right)^2\varphi_g^4$$

$$+ \frac{a_5}{5!}\left(\frac{4\pi}{N}\right)^3\frac{\varphi_g^5}{m_g} + \cdots, \tag{2.2}$$

where the lightest glueball's mass is denoted by $m_g = \mathcal{O}(\Lambda)$, and the coefficients a_i are expected to be $\mathcal{O}(1)$.

The axion also gets a scalar potential from the gluodynamics, which can be written as a power series in $(\phi/Nf_a)^2$ around its CP-conserving minimum [25,26]:

$$V(\phi) = N^2\Lambda^4\left(\frac{c_2}{2}\frac{\phi^2}{N^2f_a^2} + \frac{c_4}{4!}\frac{\phi^4}{N^4f_a^4} + \cdots\right)$$

$$= \frac{1}{2}m_a^2\phi^2 + \frac{c_4}{4!c_2}\frac{m_a^2}{N^2f_a^2}\phi^4 + \cdots. \tag{2.3}$$

In this expansion, the axion mass is given by

$$m_a^2 = \frac{1}{f_a^2}\int d^4x_E\left\langle\frac{g_h^2}{32\pi^2}G_E\tilde{G}_E(x_E)\frac{g_h^2}{32\pi^2}G_E\tilde{G}_E(0)\right\rangle_{\phi=0}$$

$$= (10^{-12}\sqrt{c_2}\text{eV})^2\left(\frac{\Lambda}{\text{MeV}}\right)^4\left(\frac{10^{15}\text{ GeV}}{f_a}\right)^2, \tag{2.4}$$

where the integral is evaluated for Euclidean continuation of Eq. (2.1). The $1/N^2$ factor for the quartic term of the axion potential leads to the suppression of anharmonic effects as long as the initial misalignment of the axion field is $\phi_i \lesssim f_a$.

The glueball is not the lightest particle in our dark sector, and the symmetry allows the decay of the glueball to two axions as $\varphi_g \to \phi\phi$. We can infer the glueball lifetime from the lattice calculation. The leading axion-glueball interaction can be obtained from the axion-dependent glueball mass term:

$$m_g = m_g(\phi=0)\left(1 + g_2\frac{\phi^2}{N^2f_a^2} + \mathcal{O}\left(\frac{\phi^4}{N^4f_a^4}\right)\right), \tag{2.5}$$

where $g_2 \simeq -0.5\,(-0.6)$ for $N = 3\,(4)$ [27,28]. Because the coefficient g_2 is not suppressed in the large-N limit [27], its value is expected to remain as $\mathcal{O}(1)$ for all $N \geq 3$. Through the interaction term $\phi^2\varphi_g^2$ from Eq. (2.5) and the cubic term φ_g^3 in Eq. (2.2), the one-loop diagram of the glueball provides the following effective Lagrangian, which is relevant for the glueball decay:

$$\left.\frac{\Delta\mathcal{L}_{eff}}{\sqrt{-g}}\right|_{\text{one-loop}} = c_*\frac{g_2a_3m_g^3}{4\pi N^3f_a^2}\phi^2\varphi_g. \tag{2.6}$$

Here c_* is the $\mathcal{O}(1)$ coefficient whose explicit value is not available at this moment. From Eq. (2.6), the lifetime of the glueball is estimated as

$$\tau_{\varphi_g} \sim 10^{18}\text{ Gyr}\left(\frac{N}{3}\right)^6\left(\frac{f_a}{10^{13}\text{ GeV}}\right)^4\left(\frac{\text{GeV}}{m_g}\right)^5. \tag{2.7}$$

In the parameter space we will focus on, the glueball is cosmologically stable, so that both the axion and glueball are dark matter of the Universe.

For cosmology, we consider the case that the dark sector and the visible sector are thermally decoupled at the beginning. In such a case, dark gluons/glueballs are thermalized by their own interactions at a temperature T_g that could be different from the SM photon temperature T_γ. Starting from the gluon fluid ($T_g \gg \Lambda$), as the Universe expands, T_g drops and crosses the dark critical temperature $T_{g,c} = \mathcal{O}(\Lambda)$, and the confining phase transition occurs. Below $T_{g,c}$, all gluons are confined into the glueballs, and the evolution is described by the massive glueball fluids.

The dark gluon temperature also affects the evolution of the dark axions. The leading term of the axion potential induced by the gluothermodynamics is

$$V(T_g, \phi) = \frac{1}{2}m_a^2(T_g)\phi^2. \tag{2.8}$$

The axion mass $m_a(T_g)$ is well described by the dilute instanton gas approximation in the deconfining phase,

$$m_a(T_g) \simeq m_a\left(\frac{T_{g,c}}{T_g}\right)^{\eta_a} \quad \text{for } T_g \gtrsim T_{g,c}, \tag{2.9}$$

with $\eta_a = 11N/6 - 2$. For $N = 3\,(4)$, $\eta_a = 3.5\,(5.3)$ [29]. After the confinement ($T_g \lesssim T_{g,c}$), the axion mass is saturated to its zero-temperature value, $m_a(T_g) \simeq m_a$ [30].

Actually, the temperature dependence of the axion potential implies the existence of the energy flow from the gluon fluid to the axions as the temperature decreases. Then, a natural question is whether or not the entropy of the dark gluon also evolves during the energy transfer. In order to make it clearer, let us discuss the gluothermodynamics in more detail. The free energy density of the gluon/glueball fluid can be evaluated from the gluon partition function for a given temperature T_g. If the topological θ term

$$\theta \equiv \frac{\phi}{f_a} \tag{2.10}$$

is vanishing, the free energy density f_g is only the function of the temperature as $f_g(T_g) = -p_g$, where p_g is the pressure of the gluon/glueball fluid. The energy (ρ_g) and entropy (s_g) densities are obtained by the thermodynamic relations, $s_g = -df_g/dT_g$ and $\rho_g = T_gs_g - p_g$. On the other hand, the situation is a little bit different for the

nonvanishing θ term. Because the gluon partition function also depends on θ, the free energy density (the negative of the pressure for the gluon fluid with the θ term) is evaluated as [27]

$$f_{g+\theta}(T_g) = -p_g + V(T_g, \phi), \qquad (2.11)$$

where $p_g \equiv -f_g(T_g)$. The second term of the rhs represents the contribution of the vacuum energy density generated by the nonperturbative gluothermodynamics. The energy density of the gluon fluid with the θ term also can be decomposed into the sum of the vacuum energy density and the pure gluonic contribution as $\rho_{g+\theta}(T_g) = \rho_g + V(T_g, \phi)$. Then, the thermodynamics relations provide

$$
\begin{aligned}
s_{g+\theta} &= -\frac{df_{g+\theta}(T_g)}{dT_g} = \frac{dp_g}{dT_g} - \frac{\partial V(T_g, \phi)}{\partial T_g} \\
&= \frac{\rho_{g+\theta}(T_g) - f_{g+\theta}(T_g)}{T_g} = \frac{\rho_g + p_g}{T_g} = s_g.
\end{aligned} \qquad (2.12)
$$

The relation $s_{g+\theta} = s_g$ implies that the entropy of the dark sector is mostly given by the gluonic excitations, not by the axionic excitations. This is the natural consequence, because the axion is homogeneously distributed in space, and its time evolution is negligible compared to the thermal process of the gluon plasma. On the other hand, the entropy and the energy density of the gluon fluid depend not only on its temperature but also on the axion field value as Eq. (2.12) and

$$\rho_g = T_g \frac{dp_g}{dT_g} - p_g - T_g \frac{\partial V(T_g, \phi)}{\partial T_g}. \qquad (2.13)$$

From the continuity equation of the dark sector, we can explicitly show that the entropy of the dark sector in a comoving volume $s_g a^3$ is conserved for whatever value of ϕ during adiabatic evolution.

Based on this observation, before discussing the explicit evolution of each component, we address general equations of motion for axion and gluon as the fluids including their homogenous and perturbation parts.

B. Dynamics of the axion-gluon/glueball fluids

The axion dark matter is described by the evolution of the classical field, $\phi(x)$. The dark gluon/glueball densities and their perturbations can be parametrized by its temperature evolution $T_g(x)$ and $\phi(x)$, as discussed in the previous section. In this context, $\{\phi(x), T_g(x)\}$ are good variables to derive full equations of motion of the dark sector, including their perturbations. Considering the fluid description, the evolution of energy densities and pressures are deduced from the evolution of ϕ and T_g with the help of the Einstein equations, gluothermodynamics, and the lattice calculation.

We introduce the conformal Newtonian gauge for the inhomogeneous part of the metric tensor:

$$ds^2 = a(\tau)^2(-(1 + 2\Psi)d\tau^2 + (1 + 2\Phi)d\vec{x}^2), \qquad (2.14)$$

where the conformal time τ and the conformal Hubble rate $\mathcal{H}$ are related with the proper time t and the Hubble expansion rate $H = \dot{a}/a$ as

$$\tau = \int \frac{dt}{a}, \qquad \mathcal{H} \equiv \frac{1}{a}\frac{da}{d\tau} = \frac{a'}{a} = aH. \qquad (2.15)$$

Here, we use the notation

$$' \equiv \frac{d}{d\tau}, \qquad \dot{} \equiv \frac{d}{dt} \qquad (2.16)$$

for the time derivatives. For most of the discussion in Secs. II and IV, we take the conformal time τ as the argument of the time-dependent variables, whereas the proper time t is mainly used in Sec. V when discussing the evolution after dark matter halos form.

Expanding ϕ and T_g near the homogeneous solutions

$$\phi(\tau, \vec{x}) = \phi(\tau) + \delta\phi(\tau, \vec{x}), \qquad (2.17a)$$

$$T_g(\tau, \vec{x}) = T_g(\tau) + \delta T_g(\tau, \vec{x}), \qquad (2.17b)$$

the equation of motion of the background axion field is given by

$$\phi'' + 2\mathcal{H}\phi' + a^2 \frac{\partial V(T_g, \phi)}{\partial \phi} = 0. \qquad (2.18)$$

The corresponding background energy density and pressure are

$$\rho_a = \frac{\phi'^2}{2a^2} + V(T_g, \phi), \qquad p_a = \frac{\phi'^2}{2a^2} - V(T_g, \phi). \qquad (2.19)$$

From Eqs. (2.18) and (2.19), the continuity equation for the background axion is obtained:

$$\rho_a' + 3\mathcal{H}(1 + w_a)\rho_a = \frac{\partial V}{\partial T_g} T_g'. \qquad (2.20)$$

$w_a \equiv p_a/\rho_a$ is the equation of state for the axion. Since the axion-gluon fluid is isolated from the visible sector, the total energy and pressure of dark sector should satisfy the continuity equation without source terms:

$$\rho_{a+g}' + 3\mathcal{H}(\rho_{a+g} + p_{a+g}) = 0. \qquad (2.21)$$

This leads to the evolution of the gluon/glueball fluid as

$$\rho'_g + 3\mathcal{H}(1 + w_g)\rho_g = -\frac{\partial V}{\partial T_g} T'_g, \qquad (2.22)$$

where $w_g = p_g/\rho_g$. Together with Eqs. (2.12) and (2.13), we can derive the conservation of the dark entropy as we claimed:

$$s'_g + 3\mathcal{H}s_g = 0. \qquad (2.23)$$

The evolution of the Fourier-transformed perturbed axion field variable $\delta\phi$ for a given wave number k is described by

$$\delta\phi'' + 2\mathcal{H}\delta\phi' + \left(k^2 + a^2\frac{\partial^2 V}{\partial\phi^2}\right)\delta\phi + \phi'(-\Psi' + 3\Phi')$$
$$+ 2a^2\frac{\partial V}{\partial\phi}\Psi + a^2\frac{\partial^2 V}{\partial\phi\partial T_g}\delta T_g = 0. \qquad (2.24)$$

From this, the evolution of the fluid perturbation variables $\delta\rho_a$, δp_a, v_a and π_a,

$$\delta\rho_a = \frac{\phi'\delta\phi' + \phi'^2\Phi}{a^2} + \frac{\partial V}{\partial\phi}\delta\phi + \frac{\partial V}{\partial T_g}\delta T_g, \qquad (2.25a)$$

$$\delta p_a = \delta\rho_a - 2\frac{\partial V}{\partial\phi}\delta\phi - 2\frac{\partial V}{\partial T_g}\delta T_g, \qquad (2.25b)$$

$$(\rho_a + p_a)v_a = a^{-2}k\phi'\delta\phi, \qquad p_a\pi_a = 0, \qquad (2.25c)$$

can also be calculated. We follow the definition and convention of the variables in Refs. [31–33].

On the one hand, from Eq. (2.13), the gluon/glueball fluid perturbations δp_g and $\delta\rho_g$ can be related with δT_g and $\delta\phi$ as

$$\delta\rho_g = \left(T_g^2\frac{d^2 p_g}{dT_g^2} - \frac{\partial^2 V}{\partial(\ln T_g)^2}\right)\frac{\delta T_g}{T_g} - \left(\frac{\partial^2 V}{\partial\phi\partial\ln T_g}\right)\delta\phi,$$

$$\delta p_g = \frac{dp_g}{dT_g}\delta T_g. \qquad (2.26)$$

Similarly, from Eqs. (2.24) and (2.25), we derive the equations of motion for the fluid perturbations:

$$\delta'_a = -ku_a - 3(1 + w_a)\Phi'$$
$$- \left(3\mathcal{H} + \frac{1}{2}\frac{\partial\ln V}{\partial\ln T_g}\frac{T'_g}{T_g}\right)\left(\frac{\delta p_a}{\delta\rho_a} - w_a\right)\delta_a$$
$$+ \frac{1}{2}(1 - w_a)\frac{d}{d\tau}\left(\frac{\partial\ln V}{\partial\ln T_g}\frac{\delta T_g}{T_g}\right), \qquad (2.27a)$$

$$u'_a = -\mathcal{H}(1 - 3w_a)u_a + k(1 + w_a)\Psi + k\frac{\delta p_a}{\delta\rho_a}\delta_a$$
$$- \frac{1}{2}(1 - w_a)\frac{\partial\ln V}{\partial\ln T_g}\left(\frac{T'_g}{T_g}u_a - k\frac{\delta T_g}{T_g}\right), \qquad (2.27b)$$

$$\delta'_g = -ku_g - 3(1 + w_g)\Phi' - 3\mathcal{H}\left(\frac{\delta p_g}{\delta\rho_g} - w_g\right)\delta_g$$
$$+ \frac{1}{2}\frac{\rho_a}{\rho_g}\left[\frac{\partial\ln V}{\partial\ln T_g}\frac{T'_g}{T_g}\left\{\left(\frac{\delta p_a}{\delta\rho_a} - 1\right)\delta_a + (1 - w_a)\delta_g\right\}\right.$$
$$\left. - (1 - w_a)\frac{d}{d\tau}\left(\frac{\partial\ln V}{\partial\ln T_g}\frac{\delta T_g}{T_g}\right)\right], \qquad (2.27c)$$

$$u'_g = -\mathcal{H}(1 - 3w_g)u_g + (1 + w_g)k\Psi + k\frac{\delta p_g}{\delta\rho_g}\delta_g$$
$$+ \frac{1}{2}(1 - w_a)\frac{\partial\ln V}{\partial\ln T_g}\left(\frac{T'_g}{T_g}u_g - k\frac{\delta T_g}{T_g}\right), \qquad (2.27d)$$

where $\delta_a = \delta\rho_a/\rho_a$, $u_a = (1 + w_a)v_a$, and we use the same definitions for δ_g, u_g.

Although Eq. (2.27) is not a closed form, it is straightforward to express $\{\delta p_a, \delta p_g, \delta T_g, \delta\phi\}$ in terms of $\{\delta_a, \delta_g, u_a\}$. One can also take the perturbed variables as $\{\delta\phi, \delta T_g, u_g\}$ using Eqs. (2.25) and (2.26) for solving Eq. (2.27). The nontrivial ingredients in our differential equations are the terms proportional to $\partial V/\partial T_g$, which originate from the nonperturbative interactions between the gluon and the axion. The effect of these terms becomes larger as T_g approaches $T_{g,c}$ and suddenly disappears after the confinement.

As the initial conditions for cosmological evolution, the amount of dark gluons can be parametrized by the ratio between the entropies of the dark sector and the visible sector, s_g/s_{SM}. Even though dealing with the entropy ratio between two sectors would make it easier to trace the evolution of the densities, in order to get a more intuitive picture about how cold the gluons are compared to the visible sector, we will use the ratio parameter between the temperatures. Taking a period around the phase transition, the photon temperature when T_g arrives at $T_{g,c}$ is denoted by $T_{\gamma,c}$. Then, we define the ratio parameter r as

$$r \equiv \left(\frac{g_{*S}(T_{\gamma,c})}{2(N^2 - 1)}\frac{s_g}{s_{\mathrm{SM}}}\right)^{1/3} \simeq \frac{T_{g,c}}{T_{\gamma,c}}, \qquad (2.28)$$

where s_g (s_{SM}) is the entropy density of the gluon fluid (the SM sector), and g_{*S} is the effective number of degrees of freedom in entropy for the SM sector.

So far, we have ignored the effect of dissipation for the axion's motion induced by the background dark gluon plasma. It is not crucial in our discussion, but let us clarify how small its effect is. Including the friction term (γ_{fr}) induced by gluon plasma, the equation of motion of the background axion is written as [34]

$$\ddot{\phi} + (3H + \gamma_{\mathrm{fr}})\dot{\phi} + m_a^2(T_g)\phi = 0. \qquad (2.29)$$

In the deconfining phase, $\gamma_{\mathrm{fr}} = \Gamma_{\mathrm{sph}}(T_g)/2f_a^2 T_g$, where the sphaleron rate is estimated by [35]

$$\Gamma_{\text{sph}}(T_g) = \int d^4x \left\langle \frac{g_h^2}{32\pi^2} G\tilde{G}(x) \frac{g_h^2}{32\pi^2} G\tilde{G}(0) \right\rangle_{T_g}$$

$$= \mathcal{O}(0.1\text{--}1) \left(\frac{g_h^2 N}{4\pi}\right)^5 \left(\frac{N^2-1}{N}\right) T_g^4 \qquad (2.30)$$

for $g_h^2/4\pi \lesssim 0.1$. Note that this is not a Euclidean correlator like Eq. (2.4), but evaluated in real spacetime. Around the critical temperature $T_g \sim T_{g,c}$, the gauge coupling can be as large as $g_h^2 N/4\pi = \mathcal{O}(1)$. In this regime, the sphaleron rate is expected as $\Gamma_{\text{sph}}(T_{g,c}) \sim T_{g,c}^4$ from the argument of dimensional analysis and the calculation using the AdS/CFT correspondence [36]. After the confining phase transition, no reliable calculation has been done so far. A crude estimation based on the dimensional analysis is that the dissipation rate is at most proportional to the entropy density (or number density) of the glueballs as $\gamma_{\text{fr}} \sim s_g/f_a^2$. Because the time dependence of the Hubble rate and the dissipation rate are given as $H \propto a^{-2}$ ($a^{-3/2}$) in the radiation-dominated era (matter-dominated era) and $\gamma_{\text{fr}} \propto a^{-3}$, the gluon-induced friction term is important only when the temperature of the visible sector becomes greater than

$$T_\gamma > \mathcal{O}(1) \left(\frac{10^9 \text{ GeV}}{Nr^3}\right) \left(\frac{10^{14} \text{ GeV}}{f_a}\right)^2 \left(\frac{g_h^2 N}{4\pi}\right)^{-5}. \qquad (2.31)$$

Comparing this with the temperature when the axion starts to oscillate in our scenario ($T_\gamma < \text{TeV}$), it is always irrelevant.

III. EVOLUTION HISTORY

A. Evolution of the background gluon and glueballs

In Eqs. (2.20)–(2.27), we establish the continuity equations for the background and perturbative variables based on the equations of motion of the axion field and gluothermodynamics. In this section, some features of the background evolution of gluon and glueball fluids are discussed in more detail.

The evolution of ρ_g and p_g is particularly easy for $T_g \gg T_{g,c}$ and $T_g \ll T_{g,c}$. In the limit of $T_g \gg T_{g,c}$, gluons are a relativistic fluid, and the contribution of the axion to the evolution of ρ_g, p_g is negligible, because from Eq. (2.13) and $T_{g,c} \sim \Lambda$,

$$\frac{T_g}{\rho_g} \frac{\partial V(T_g, \phi)}{\partial T_g} \sim \frac{\phi^2}{Nf_a^2} \left(\frac{\Lambda}{T_g}\right)^{11N/3} \ll 1. \qquad (3.1)$$

Therefore, the usual scaling relations hold as

$$p_g = \frac{\pi^2}{45}(N^2-1)T_g^4,$$

$$\rho_g \approx \frac{\pi^2}{15}(N^2-1)T_g^4 = 3p_g. \qquad (3.2)$$

As the gluon temperature approaches $T_{g,c}$, Eq. (3.2) does not hold anymore, because the strong interactions among gluons become significant [37] and the axion contribution in Eq. (2.13) is also increasing. The true evolution can only be figured out by the lattice calculation. We adopt the lattice data for $\theta = 0$ to evaluate $p_g(T_g)$ [38–40] and deduce the densities of the gluons for the nonzero θ using Eqs. (2.12) and (2.13).

At $T_g = T_{g,c}$, the dark gluons are combined into dark glueballs, whose masses are multiples of the confining scale Λ [41–44]. Here, Λ should be defined with a certain regularization scheme. Taking the $\overline{MS}$ scheme [45] shows the relation $\Lambda/\sqrt{\sigma} = 0.5 + 0.34/N^2$ for $N \geq 3$, where σ is the string tension. The relation between the critical temperature and the string tension is evaluated as $T_{g,c}/\sqrt{\sigma} = 0.59 + 0.46/N^2$ based on the study for $2 \leq N \leq 8$ [46]. As a result, we get

$$\frac{T_{g,c}}{\Lambda} \simeq 1.2 + \frac{0.1}{N^2}. \qquad (3.3)$$

The phase transition is first-order if $N > 2$. In order to understand how long the phase transition happens, we can compare the energy density in the deconfined phase with that in the confined phase at $T_{g,c}$. The former is the energy density of the gluon plasma of $\mathcal{O}(0.1N^2 T_{g,c}^4)$, while the latter is the sum of the glueball tower of $\mathcal{O}(0.1 T_{g,c}^4)$. Thus, as N increases, larger latent heat is released and the transition period becomes longer.

Let us shortly discuss how we evaluate the energy density of the glueballs. Since the lightest glueball mass is calculated as

$$\frac{m_g}{T_{g,c}} \simeq 5.7 - \frac{1.2}{N^2}, \qquad (3.4)$$

where we use the relation $m_g/\sqrt{\sigma} = 3.64$ for $N = 3$ [42], and $m_g/\sqrt{\sigma} = 3.37 + 1.93/N^2$ based on the results for $2 \leq N \leq 5$ [47], all glueballs are nonrelativistic at $T_{g,c}$. Using the spectral density $\hat{\rho}(m)$, the energy density of glueballs at $T_g \leq T_{g,c}$ can be written as

$$\rho_g(T_g) = \int_0^\infty dm \hat{\rho}(m) m \left(\frac{mT_g}{2\pi}\right)^{3/2} e^{-m/T_g}. \qquad (3.5)$$

For $\theta = 0$ (the effect of the axion-induced θ term will be discussed later), the glueballs are well described by the eigenstates of the spin (J), the parity (P), and the charge conjugation (C): J^{PC}. The lightest glueball corresponds to

0^{++} [48]. $\hat{\rho}(m)$ contains information about the tower of the glueballs. It turns out that the spectral density can be successfully approximated by the sum of the discrete low-lying resonances with a mass $m_{J^{PC}}$ ($<m_{\rm th}$) and the continuum spectrum of the Hagedorn tower [49–51],

$$\hat{\rho}(m) \simeq \sum_{m<m_{\rm th}} (2J+1)\delta(m-m_{J^{PC}})$$
$$+ \frac{n_N}{m}\left(\frac{2\pi T_H}{3m}\right)^3 e^{m/T_H}\Theta(m-m_{\rm th}), \quad (3.6)$$

where $n_2 = 1$, $n_{N\geq 3} = 2$. In the large-N limit, the Hagedorn temperature T_H is related with $T_{g,c}$ as [50–52]

$$\frac{T_H}{T_{g,c}} \simeq 1.16 - \frac{0.9}{N^2}. \quad (3.7)$$

We use the closed Nambu-Goto string model for glueballs to provide the relation between T_H and the string tension, $T_H = \sqrt{3\sigma/2\pi}$ [50]. The threshold mass $m_{\rm th}$ is not a physical quantity. It is shown that taking $m_{\rm th} = 2m_g \simeq 10T_H$ is good enough to reproduce the lattice results [51]. Actually, from the closed string description of the glueball spectrum, the continuum approximation works well for $T_H \ll m$.

Note that even if N increases, the number of glueball degrees of freedom does not increase. Therefore, the contribution from the Hagedorn glueballs to ρ_g near the confinement is insensitive to N and becomes $\mathcal{O}(0.01–0.1)T_{g,c}^4$. The low-lying glueball contributions with $m_{J^{PC}} < 2m_g$ are $\mathcal{O}(10–50)\%$ of it. For this estimation, the glueball spectrum of Refs. [42,43] is used in the case of $N = 3$. For the larger value of N, we adopt the spectrum $0^{++(*)}$, 2^{++} of Refs. [53,54]. In fact, the glueball spectrum calculated using the lattice has a large uncertainty except for a few low-lying modes in the case of $N > 3$. However, we can reasonably assume that there is no significant change in the spectrum even for higher values of N [55].

Now, we turn our attention to the effect of the axion on the glueball spectrum. First of all, the nonzero axion field value can change the glueball mass as $\delta m/m = \mathcal{O}(\phi^2/N^2 f_a^2)$ from Eq. (2.5). Its contribution is smaller than 10% for $N \geq 3$, so it is ignorable. Secondly, for the nonzero axion value, the parity is no longer a good quantum number. This leads to mixing between the glueballs with different P eigenvalues. For instance, Ref. [56] shows that the mixing between 0^{++} and 0^{-+} is not suppressed even in the large-N limit as $\varphi_g = 0^{++} \rightarrow \varphi_g = 0^{++} + \mathcal{O}(\theta)0^{-+}$. It is also noticed that this mixing only shifts the lightest state, while the heavier state $\varphi_{-+} = 0^{-+}$ remains intact. Through the cubic interaction of 0^{++} as in Eq. (2.2), the interaction between φ_g and φ_{-+},

$$\Delta\mathcal{L} \sim \frac{4\pi\phi}{Nf_a}m_g\varphi_g^2\varphi_{-+}, \quad (3.8)$$

is induced, and it may allow the decay of φ_{-+} to two φ_g's when the axion has the nonzero expectation value. However, $m_{\varphi_{-+}} \simeq 1.5m_g < 2m_g$ for $N = 3$ [42,43], and this inequality is expected to hold for $N > 3$ from the argument of the large-N expansion. Therefore, the axion-induced decay to the lightest glueballs seems to be kinematically forbidden for $N \geq 3$. On the one hand, the existence of the mixing between 0^{-+} and 0^{++} in the presence of a θ term can also imply the direct decay of φ_{-+} to the lightest glueball and the axion from an interaction like $(m_g^2/f_a)\phi 0^{-+}0^{++}$. If this term is not canceled in the mass eigenbasis, the corresponding lifetime of φ_{-+} is estimated as $\tau_{\varphi_{-+}\rightarrow\varphi_g\phi} \sim 20\ {\rm Myr}(f_a/10^{15}\ {\rm GeV})^2({\rm MeV}/m_g)^3$, which can be shorter than the age of the Universe.

Although there are large uncertainties in estimating the effect of the axion on the dynamics of heavier glueballs, we expect that (i) its effect on the glueball masses is small, (ii) the heavier glueballs become less stable, and (iii) they interact more with other glueballs. This means that the approach using the spectral density Eq. (3.5) works well near the critical temperature, $T_g \sim T_{g,c}$, even if the axion degree of freedom is included.

The lower limit of the actual transition period is given by the period obtained assuming the quasiequilibrium transition. This is the case when the pressures of the deconfining/confining phases are equal and the latent heat is released adiabatically as the Universe expands. In this situation, the temperatures of the confining and deconfining phases are the same and maintain at around $T_{g,c}$ during the transition, and the entropy is conserved. The duration of the phase transition is estimated by the conservation of dark entropy:

$$a_{cf} = a_{ci}\left(\frac{s_{\rm gluon}(T_{g,c})}{s_{\rm glueball}(T_{g,c})}\right)^{1/3} \simeq a_{ci}N^{2/3}, \quad (3.9)$$

where a_{cf} (a_{ci}) is the scale factor when the phase transition ends (starts), while $s_{\rm gluon}(T_{g,c})$ [$s_{\rm glueball}(T_{g,c})$] denotes the entropy density of the dark gluon (glueball) at $T_g = T_{g,c}$. The N dependence of the duration is obtained from $s_{\rm gluon}(T_{g,c}) \simeq 0.1N^2 T_{g,c}^3 \simeq N^2 s_{\rm glueball}(T_{g,c})$.

The phase transition becomes stronger in first order as N increases. Thus, the additional entropy is generated during the transition and makes the glueballs hotter than the previous estimation. However, this effect is negligible unless $N \gtrsim 4\pi$, because the nucleation temperature is just around $T_{g,c}$, and the strong interactions of the gluon and glueball fluids provide a large friction coefficient for the bubble wall propagation. For $N = 3$, $a_{cf} \simeq 2a_{ci}$ is obtained numerically, which is well matched with our parametric estimation $N^{2/3}$. The assumption of dark

entropy conservation will be kept in the following discussion.

After the phase transition, the evolution of the glueball temperature is calculated using the conservation of the dark entropy. The Hagedorn spectrum can slow down the decreasing rate of T_g, but numerically its contribution becomes gradually suppressed and negligible in comparison with that of the lightest glueball when $a \gtrsim 10 a_{cf}$ ($T \lesssim 0.55 T_{g,c}$). The detailed freeze-out process of the low-lying stable glueballs other than the lightest one is provided in Ref. [57] for $N = 3$ with $\theta = 0$. Their abundance is also quite suppressed after the glueball temperature becomes smaller than $0.5 T_{g,c}$. As we already discussed, the mass spectrum of the heavier glueballs for $N > 3$ is not much affected by the evolution of the axion. Furthermore, the role of the axion is generically to make the heavier glueballs unstable, so that the corresponding abundance could be further suppressed compared to the case with $\theta = 0$. Therefore, from now on, we will just focus on the evolution of the lightest glueball.

The relevant scattering processes to maintain the thermal equilibrium of φ_g are two-to-two ($2 \rightarrow 2$) and three-to-two ($3 \rightarrow 2$) scatterings, whose rates are estimated as

$$
\sigma_{2 \rightarrow 2} v \sim \frac{v_f (4\pi/N)^4}{32\pi m_g^2}, \qquad \sigma_{3 \rightarrow 2} v^2 \sim \frac{(4\pi/N)^6}{(4\pi)^3 m_g^5}, \quad (3.10)
$$

where v_f is the relative velocity of the final particles from the scattering.

As the Universe expands, the $3 \rightarrow 2$ process freezes out when most of the $2 \rightarrow 2$ processes are still active. This is because the interaction rate of the $3 \rightarrow 2$ process is proportional to the square of the number density of the glueballs, while that of the $2 \rightarrow 2$ processes is linearly proportional to the number density of the glueballs. Before the freeze-out of $3 \rightarrow 2$ interactions, the glueball density and pressure are the function of its temperature as

$$
\begin{aligned}
\rho_g(T_g) &= m_g \left(\frac{m_g T_g}{2\pi} \right)^{3/2} e^{-\frac{m_g}{T_g}} \left(1 + \frac{27}{8} \frac{T_g}{m_g} + \mathcal{O}\left(\frac{T_g^2}{m_g^2} \right) \right), \\
p_g(T_g) &= T_g \left(\frac{m_g T_g}{2\pi} \right)^{3/2} e^{-\frac{m_g}{T_g}} \left(1 + \frac{15}{8} \frac{T_g}{m_g} + \mathcal{O}\left(\frac{T_g^2}{m_g^2} \right) \right).
\end{aligned}
$$
$$(3.11)$$

A distinguishing property of the chemical equilibrium maintained by the number-changing self-interaction is that its temperature drops much slower than the photon temperature due to the entropy-damping effect. From entropy conservation $s_g \propto 1/a^3$, the temperature scales as $T_g \sim 1/\ln a$ [8]. As a consequence, the energy density drops faster than that of a cold dark matter,

$$
\rho_g \simeq T_g s_g \propto \frac{1}{a^3 \ln a}, \quad (3.12)
$$

since the $3 \rightarrow 2$ self-interaction converts the mass energy to kinetic energy. This behavior ends when the process freezes out at $T_g = T_{g,fo}$ with

$$
\rho_g^2(T_{g,fo}) \simeq \frac{(3 m_g T_{g,fo})(H|_{T_g = T_{g,fo}})}{\langle \sigma_{3 \rightarrow 2} v^2 \rangle}. \quad (3.13)
$$

After that, the glueballs still maintain kinetic equilibrium by the $2 \rightarrow 2$ interactions, but they act as free-streaming particles for their background evolution. Using Eq. (3.13) and the dark entropy conservation, the freeze-out temperature is evaluated as

$$
\frac{m_g}{T_{g,fo}} + \frac{5}{4} \ln \frac{m_g}{T_{g,fo}} + \frac{3}{4} \ln \frac{m_g}{\text{MeV}} \simeq 28.2 + \frac{3}{2} \ln \frac{r}{0.01} - \frac{7}{2} \ln \frac{N}{3}
$$
$$(3.14)$$

when it happens during the radiation-dominated era. As a specific example, for $N = 3$, $r = 0.01$, and $m_g = 1$ MeV, we get

$$
T_{g,fo} \simeq 0.04 m_g \simeq 0.2 T_{g,c}. \quad (3.15)
$$

The relation $T_{g,fo} = \mathcal{O}(0.2) T_{g,c}$ is not much sensitive to the values of r and m_g that we are interested in.

Meanwhile, we can argue that the impact of the axion dynamics on the evolution of the lightest glueballs is negligible. The main reason is that after the confining phase transition, the axion potential becomes independent of the glueball temperature—i.e., $\partial_{T_g} V(T_g, \phi) = 0$. According to Eqs. (2.12) and (2.13), the gluothermodynamic quantities (p_g, s_g, ρ_g) are decoupled from the axion dynamics as long as the relaxation time for the number-changing process $\Gamma_{3 \rightarrow 2}^{-1} \sim m_g^{-1}(a/a_{cf})^6$ is shorter than the axion oscillation period $m_a^{-1} \sim m_g^{-1}(f_a/m_g)$. If the axion starts oscillating before the glueball freeze-out, it is also plausible that the glueball energy density is modulated by the axion-induced $\theta(t)$ term after $\Gamma_{3 \rightarrow 2}^{-1} \gtrsim m_a^{-1}$. However, its maximum contribution, as in Eq. (2.5) for $\phi \sim f_a$, is less than 10% from the beginning, and the oscillating amplitude is redshifted such as $\delta m_g/m_g \sim \theta(t)^2/N^2 \propto \cos(m_a t)/a^3$. As its oscillation period becomes shorter than the relaxation time of the glueball's chemical process, its effect is also averaged out to be zero. Hence, the impact of the axion is negligible over time and does not change our main results.

During the evolution of the glueballs, the photon temperature also evolves. When the dark glueballs freeze out, the photon temperature becomes

$$T_{\gamma,fo} \simeq 3 \text{ keV} \left(\frac{N}{3}\right)^{1/2} \left(\frac{0.01}{r}\right)^{3/2} \left(\frac{m_g}{\text{MeV}}\right)^{5/4}. \quad (3.16)$$

One can also easily evaluate the case that the freeze-out of the dark glueball happens after the matter-radiation equality for $m_g < \text{keV}$.

So far, we have specified all the history of the gluons and glueballs in order to identify the time dependence of the glueball temperature $T_g(\tau)$, which is relevant to the evolution of the perturbative variables [Eq. (2.27)]. However, for the final relic density of the glueballs, it can be evaluated in a much simpler way from the conservation of the entropy of dark sector as

$$\begin{aligned}
\Omega_g h^2 &\simeq 0.014 \left(\frac{N^2-1}{10}\right) \left(\frac{r}{0.01}\right)^3 \left(\frac{T_{g,fo}}{10 \text{ keV}}\right) \left(\frac{3.94}{g_{*S}(T_{\gamma,c})}\right) \\
&\simeq 0.014 \left(\frac{N^2-1}{10}\right) \left(\frac{r}{0.01}\right)^4 \left(\frac{T_{\gamma,c}}{5 \text{ MeV}}\right) \left(\frac{3.94}{g_{*S}(T_{\gamma,c})}\right) \\
&\simeq 0.12 \left(\frac{N^2-1}{10}\right) \left(\frac{r}{0.003}\right)^3 \left(\frac{m_g}{100 \text{ MeV}}\right) \left(\frac{3.94}{g_{*S}(T_{\gamma,c})}\right).
\end{aligned} \quad (3.17)$$

B. Evolution of the background axion

The dilute instanton gas approximation works well for the axion potential before the confining transition occurs. However, it is no longer valid to describe the axion potential in the confining phase. The lattice studies can provide a part of the information for the axion potential— i.e., the coefficient of each term in the perturbative expansion [Eq. (2.3)] as [58]

$$c_2 \simeq 0.3 + \frac{1}{N^2}, \qquad c_4 \simeq -2.7c_2. \quad (3.18)$$

For the evolution of the axion field, further information is necessary. An interesting feature of the axion potential induced by the confinement without light fermions is that it is not a single branch, but is composed of multiple (N) branches, where for each branch the period of the scalar potential is $2\pi N f_a$ [59]. The general expression of the scalar potential for a kth branch is

$$V_k = N^2 \Lambda^4 h\left(\frac{\phi}{Nf_a} + \frac{2\pi(k-1)}{N}\right), \quad (3.19)$$

where $k = 1, \ldots, N$, and $h(\psi)$ is the 2π-periodic function. The full shape of the axion potential is not available. However, the analytic form of $h(\psi)$ in a certain range of the axion was studied in the large-N limit using the holographic description of the pure $SU(N)$ gauge theory [60,61]. The shape of the potential highly relies on the size of the 't Hooft coupling $\lambda_h = g_h^2 N$ at the KK scale. By comparing the axion potential in the dual gravity theory and

that of the lattice calculation given by Eq. (3.18), we find that $\lambda_h = 10\text{–}20$ gives a reasonable matching.

At high temperatures for the gluons, the instanton approximation for the axion potential is valid, and there is a single branch. During the phase transition, branches will emerge, and the axion can be located in a different branch in a different patch of the Universe. If each branch provides a stable axion trajectory, we have to consider the effect of them seriously.

Following the approach of the holographic description [60], we can estimate the tunneling rate per volume between the kth and $(k-1)$th branches as

$$\Gamma_{\text{tunneling}} \sim \Lambda^4 e^{-S_{(k\to k-1)}}, \quad (3.20)$$

where the Euclidean action is

$$S_{(k\to k-1)} = \mathcal{O}(10^{-11}) N \frac{(N/k)^3}{(1 + \frac{\mathcal{O}(1)k^2}{N^2})^2}. \quad (3.21)$$

This can be significantly large only when $N \gtrsim 10^3$. Therefore, in our consideration with $N^2 = \mathcal{O}(10)$, all branches with higher energy densities are quite unstable, and the transition to the lowest energy state will occur almost immediately. As a result, the effective potential of the axion is well described by

$$V(\phi) = \min_k V_k(\phi), \quad (3.22)$$

and one can think of the evolution of the axion within the range $2\pi f_a$.

Without worrying about the effect of other branches, Eq. (2.18) gives

$$\phi'' + 2\mathcal{H}\phi' + a^2 m_a^2(T_g)\phi = 0 \quad (3.23)$$

for $\phi \lesssim f_a$. If the second term of the lhs is much larger than the third term, the axion field is approximately constant because of the large Hubble friction. This is the slow-roll limit. In the opposite case, the axion field oscillates with the oscillation frequency $m_a(T_g)$. Such evolution can be well approximated by the simple transition at $a = a_{\text{osc}}$, where a_{osc} is the scale factor to give $3\mathcal{H} = am_a(T_g)$ [$3H = m_a(T_g)$]. For each epoch,

$$\begin{aligned}
\phi(\tau) &\simeq \phi_i \equiv f_a \theta_i (a < a_{\text{osc}}) \\
&\simeq \mathcal{A}(\tau) \cos\left(\int^\tau d\tilde{\tau} a(\tilde{\tau}) m_a(T_g(\tilde{\tau}))\right). \quad (a > a_{\text{osc}})
\end{aligned} \quad (3.24)$$

Here, θ_i is the initial misalignment angle of the axion field, and $\mathcal{A}(\tau)$ is a slowly varying function with $\mathcal{A}'/\mathcal{A} \ll am_a(T_g)$. The axion acts like dark energy during

$a < a_{\rm osc}$, while for $a > a_{\rm osc}$, the axion plays the role of cold dark matter because $\langle w_a \rangle \simeq 0$ by averaging out the fast oscillation. The initial axion value ϕ_i is not deterministic. Since both $\theta_i \ll 1$, and $|\theta_i - \pi| \ll 1$ need some tuning or special model building, here we take

$$\theta_i = \mathcal{O}(1). \tag{3.25}$$

Since the axion's mass depends on the history of the dark gluons [Eq. (2.9)], there are two characteristic scales which determine the evolution history of the axion: $a_{\rm osc}$ (onset of the axion oscillation) and a_{ci} (onset of the confining phase transition). As the scale factor approaches a_{ci}, the contribution of the gluons to the axion's potential becomes substantial, and the axion mass is saturated. The evolution of the axion mass is smooth compared to the gluon thermodynamic relaxation timescale—i.e., $\dot{m}_a/m_a \sim NH \ll T_g$, unless N is very large.

Using Eq. (2.9) and the definition of $a_{\rm osc}$, we find that the quantity

$$R(r, f_a) \equiv \left(\frac{r}{0.01}\right)^2 \left(\frac{6 \times 10^{13} \text{ GeV}}{f_a}\right) \tag{3.26}$$

determines whether or not the axion starts to oscillate before the confining transition. If $R(r, f_a) > 1$, the axion starts to oscillate in the deconfining phase. The corresponding photon temperature is

$$T_{\gamma,\rm osc} \simeq R(r, f_a)^{\frac{1}{2+\eta_a}} T_{\gamma,c}. \tag{3.27}$$

Otherwise $[R(r, f_a) < 1]$, the axion oscillation occurs after the phase transition. It happens when the photon temperature becomes

$$T_{\gamma,\rm osc} \simeq R(r, f_a)^{\frac{1}{2}} T_{\gamma,c}. \tag{3.28}$$

The initial energy density of the axion at $T_\gamma = T_{\gamma,\rm osc}$ is approximated as

$$\rho_a \simeq \frac{1}{2} m_a^2(T_{g,\rm osc}) f_a^2 \theta_i^2. \tag{3.29}$$

After that, the axion field oscillates with the time-dependent frequency. We notice that for the combination

$$N_a = \frac{a^3 \rho_a}{m_a(T_g)}, \tag{3.30}$$

Eq. (2.20) gives

$$N_a' + \left(3\mathcal{H} + \frac{m_a'(T_g(\tau))}{m_a(T_g(\tau))}\right) w_a N_a = 0. \tag{3.31}$$

For $N\mathcal{H} \ll a m_a(T_g)$, the mass of the axion changes much more slowly compared to the oscillation timescale. Such a

fast oscillation implies a vanishing averaged equation of state $\langle w_a \rangle = 0$ during the cosmological evolution. Therefore, N_a is nearly conserved and $\rho_a/m_a(T_g) \propto 1/a^3$.

In summary, if $R(r, f_a) > 1$, the axion starts to oscillate before the confining phase transition ($T_{\gamma,\rm osc} > T_{\gamma,c}$), and the present relic density of the axion dark matter becomes

$$\Omega_a h^2 \simeq 0.8 \times 10^{-3} \theta_i^2 \left(\frac{r}{0.01}\right)^4 \left(\frac{T_{\gamma,c}}{5 \text{ MeV}}\right) R(r, f_a)^{-\frac{3+\eta_a}{2+\eta_a}}. \tag{3.32}$$

If $R(r, f_a) < 1$, the axion oscillates after the confining phase transition ($T_{\gamma,\rm osc} < T_{\gamma,c}$). The corresponding axion dark matter density is estimated as

$$\begin{aligned} \Omega_a h^2 &\simeq 0.8 \times 10^{-3} \theta_i^2 \left(\frac{r}{0.01}\right)^4 \left(\frac{T_{\gamma,c}}{5 \text{ MeV}}\right) R(r, f_a)^{-\frac{3}{2}} \\ &\simeq 0.05 \theta_i^2 \left(\frac{r}{0.01}\right) \left(\frac{T_{\gamma,c}}{5 \text{ MeV}}\right) \left(\frac{f_a}{10^{15} \text{ GeV}}\right)^{3/2} \\ &\simeq 0.15 \left(\frac{m_a}{10^{-22} \text{ eV}}\right)^{1/2} \left(\frac{f_a}{10^{17} \text{ GeV}}\right)^2. \end{aligned} \tag{3.33}$$

As shown in Eqs. (3.32) and (3.17), the glueballs dominate the dark matter density if the axion oscillates earlier than the confining phase transition. The reason is simply that the initial axion energy density is bounded by the confining scale $\Lambda^4 \sim T_{g,c}^4$. On the one hand, when r^2/f_a is small enough, so that the axion starts to oscillate after the phase transition, the axion becomes a dominant component of dark matter.

The left and right panels of Fig. 1 show the evolution of energy densities for axion- and glueball-dominated dark matter scenarios, respectively. Figure 2 shows the parametric dependence of the dark matter which gives the correct relic density. The axes are represented by $T_{\gamma,c}$, the photon temperature when the confining phase transition of the dark gauge sector starts, and m_a, the zero-temperature axion mass, defined in Eq. (2.4) with Eq. (3.18), respectively. As we discussed, the dark matter today is dominated by the axion in the region $T_{\gamma,\rm osc} < T_{\gamma,c}$, and by the glueball in the opposite region $T_{\gamma,c} < T_{\gamma,\rm osc}$.

The initial amount of the dark gluon plasma is limited by the constraint on the effective extrarelativistic degrees of freedom $\Delta N_{\rm eff}$ as [62]

$$\begin{aligned} \Delta N_{\rm eff} &= \left(\frac{\rho_g}{\rho_{\nu_e}}\right)_{\rm BBN} = \frac{2(N^2 - 1)r^4}{(7/4)(4/11)^{4/3}} \\ &= 0.07 \left(\frac{N^2 - 1}{10}\right) \left(\frac{r}{0.2}\right)^4 \lesssim 0.3. \end{aligned} \tag{3.34}$$

There are various astrophysical observations to constrain the mass of the glueball and axion dark matter. We shortly summarize the relevant bounds. When the
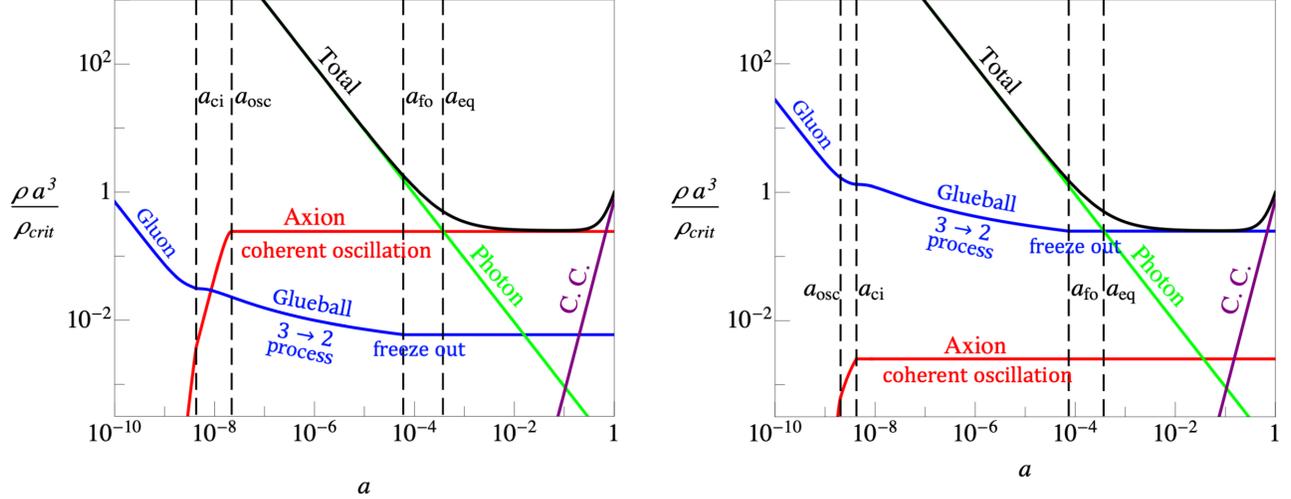
FIG. 1. The examples of dark matter density evolutions. Here we take $N = 3$. The left panel corresponds to the axion-dominated scenario, while the right panel is for the glueball-dominated case. The parameters in the right panel are taken for illustration only, because the corresponding values are ruled out by the self-interacting and warm dark matter constraints.

glueball dominates dark matter, its self-interaction gives observable effects if the scattering rate is large enough to reach the isothermal profile around the center of the halo. The self-interaction can also be detectable from the merger of dark matter halos, because the glueballs will be slowed

FIG. 2. Parametric dependence of the relic abundances of the glueball and the axion for $\Omega_{\mathrm{DM}} h^2 = 0.11$. Here we take $N = 3$. m_a is the zero-temperature axion mass, and $T_{\gamma,c} \simeq T_{g,c}/r$ is the photon temperature when the confining phase transition of the dark sector starts. The parameters $\{r$ (real lines), f_a (dashed lines)$\}$ are evaluated to give $\Omega_{\mathrm{DM}} h^2 = 0.11$ for each point in the plot. In the region above the line $R(r, f_a) = 1$, the oscillation of the axion starts earlier than the confinement phase transition and the glueball dominates the dark matter abundance with the mass $m_g \simeq 6r T_{\gamma,c}$. Below $R(r, f_a) = 1$, the axion starts to oscillate after the transition and becomes the dominant component of dark matter. Here, we did not impose the constraints from the current bound, which are discussed in text.

down during the collision if their scattering cross section is large enough. This leads to the offset between the dark matter and the collisionless components like stars. From these observations, the cross section of the glueball like self-interacting dark matter is bounded as (see Ref. [63] and references therein)

$$\frac{\sigma_{2\to2}}{m_g} \simeq \left(\frac{4}{N}\right)^4 \frac{1}{m_g^3} \simeq \left(\frac{4}{N}\right)^4 \left(\frac{60 \text{ MeV}}{m_g}\right)^3 \text{ cm}^2/\text{g}$$
$$\lesssim \mathcal{O}(0.5\text{--}5) \text{ cm}^2/\text{g}. \tag{3.35}$$

In terms of the glueball mass, it should be greater than $\mathcal{O}(50)$ MeV if it is the dominant component of dark matter. The phenomenology of heavier glueball dark matter was studied in Ref. [64].

If the axion is the dominant component of dark matter, there is a lower bound on the axion mass due to its fuzziness. The de Broglie wavelength reaches the astrophysical scale ($\sim$kpc) if m_a is around 10^{-22} eV, and it suppresses the structure formation. The ultralight axion can act like wave dark matter, with axions that are bound to or interact with each other by gravity inside the halo, which leads to the formation of solitonic cores and macroscopic quasiparticles moving around the center. These structures can have a great influence on the motion of stars. All these considerations give a strong constraint on the axion mass in the range $m_a \lesssim 10^{-22}$–10^{-20} eV (see Refs. [65,66] and references therein). There is another constraint on the mass of the axion from the observation of highly spinning black holes. This is because, if the axion mass is close to the inverse of the size of the spinning black hole, a superradiance phenomenon occurs, and parts of black hole's mass and spin are removed by the superradiant axion cloud. Current observations of the spinning supermassive black

holes with masses of 10^6–$10^7\,M_\odot$ provide interesting constraints for the axion mass range 10^{-20}–10^{-16} eV (see Ref. [67] and references therein). Since the efficiency of the black hole superradiance depends on the axion self-interaction and the surrounding environment, the constraint is rather model dependent.

When the dark matter is mostly composed of the axions ($\Omega_a h^2 \simeq 0.11$), the fraction of the dark glueball subcomponent dark matter becomes

$$f_g \equiv \frac{\Omega_g}{\Omega_{\rm DM}} \simeq 0.28 \left(\frac{N^2-1}{10\theta_i^2}\right)\left(\frac{r}{0.01}\right)^3 \left(\frac{10^{15}\,\text{GeV}}{f_a}\right)^{3/2}$$

$$\simeq 0.02 \left(\frac{N^2-1}{10}\right)\left(\frac{r}{0.01}\right)^3 \left(\frac{m_g}{0.05\,\text{MeV}}\right). \quad (3.36)$$

One can think that there is no strong constraint on the self-interactions of the glueball dark matter if $f_g \lesssim 0.1$. However, as discussed in Sec. V, the evolution of the glueball dark matter after structures form may alter the cosmological history of the Universe from $z = 7$–15.

IV. PERTURBATIONS

We now study the evolution of the cosmological perturbations for the axion and glueball dark matter. Generically, both have nontrivial features compared to the CDM. For example, the late-time transition of the axion from dark energy to dark matter modifies the early ISW effect [68]. The perturbation at scales smaller than the effective de Broglie wavelength of the axion is suppressed by its wave nature [3,69,70]. For glueballs, the number-changing self-interaction also disturbs the growth of the density perturbation at scales which enter the horizon well before the freeze-out [11,71].

On the one hand, in our setup, the dark sector is decoupled from the visible sector and the origin of their abundance can be totally different from that of the SM particles. Let us provide a simple example.

As the origin of Eq. (2.1), the axion field can be the phase of a complex scalar field X. The corresponding matter Lagrangian at high scales is

$$-\frac{\mathcal{L}_M}{\sqrt{-g}} = |\partial_\mu X|^2 + \bar{Q} i\gamma^\mu D_\mu Q + \frac{\lambda}{4}\left(|X|^2 - \frac{f_a^2}{2}\right)^2$$

$$+ yX\bar{Q}P_L Q + \text{H.c.} \quad (4.1)$$

Here Q is the vector-like fermion charged under the dark gauge group. The anomalous global symmetry

$$U(1)_{\rm PQ}: X \to e^{-2i\alpha}X, \qquad Q \to e^{-i\gamma_5\alpha}Q \quad (4.2)$$

is spontaneously broken by the nonzero vacuum expectation value of X, $f_a/\sqrt{2}$. Around the potential minimum, X can be decomposed as

$$X(x) = \frac{f_a + s(x)}{\sqrt{2}} e^{-i\phi(x)/f_a}. \quad (4.3)$$

The axion ϕ is identified as the Goldstone boson, so it is massless at perturbative level. The radial scalar s gets a mass as $m_s = \sqrt{\lambda/2}f_a$, and the dark fermion mass is given by $M_Q = yf_a/\sqrt{2}$. Assuming the hierarchy $m_s \ll M_Q$, integrating out the heavy fermion yields the following effective Lagrangian:

$$-\frac{\mathcal{L}_{M\rm eff}}{\sqrt{-g}} \simeq \frac{1}{2}(\partial_\mu s)^2 + \frac{1}{2}m_s^2 s^2 + \frac{1}{4}\left(1 + \frac{g_h^2 s}{8\pi^2 f_a}\right)(G_{\mu\nu}^a)^2$$

$$+ \frac{1}{2}\left(1 + \frac{2s}{f_a}\right)(\partial_\mu \phi)^2 + \frac{g_h^2 \phi}{32\pi^2 f_a}G_{\mu\nu}^a \tilde{G}^{a\mu\nu}. \quad (4.4)$$

The interaction between s (ϕ) and dark gluons is coming from the one-loop diagram mediated by Q. Finally, we obtain Eq. (2.1) at scales well below the mass of the radial scalar m_s, which is much larger than the confining scale of the dark gauge symmetry.

If the inflation Hubble rate H_I is given as $m_s \lesssim H_I \ll M_Q$, $U(1)_{\rm PQ}$ is not restored during inflation. In the case that the dominant scalar density is coming from the misalignment mechanism, the oscillation of the radial field happens after inflation around the potential minimum when $H \sim m_s$. These scalars will eventually decay to axions with the decay rate $\Gamma_s \sim m_s^3/8\pi f_a^2$ and gluons with the branching fraction $\text{Br}(s \to gg) \sim (Ng_h^2/8\pi^2)^2$. The gluons are quickly thermalized and form a thermal bath with temperature T_g, while the produced axions are just redshifted. These relativistic axions form dark radiation, whose abundance is negligible when r is small enough.

Although this is just one of the production mechanisms of the dark sector, it gives a good motivation to study the isocurvature perturbation of dark matter from the initial fluctuation of dark gluon temperature $\delta T_{g,i}$. In this example, in addition to the adiabatic perturbation, there are two sources of the isocurvature perturbation. One is the fluctuation of the axion field during inflation $\delta\phi_i$, and the other is the fluctuation of the gluon temperature $\delta T_{g,i}$ induced by the initial perturbation of the decaying scalar s. Because of the condition $m_s \lesssim H_I$, the superhorizon modes of both the radial scalar and the axion fields get independent fluctuations as $\delta s_i(k) \simeq H_I/2\pi$ and $\delta\phi_i(k) \simeq H_I/2\pi$ during inflation. Therefore, $\delta s_i/s_i \sim (\delta\rho_s/\rho_s)_{\rm iso} \sim (\delta T_{g,i}/T_{g,i})_{\rm iso}$.

Because the dark axions and the dark gluons are coupled with each other by the term $\partial V/\partial T_g$, as in Eq. (2.27), both perturbations could be important for the final isocurvature perturbation of dark matter. Based on the evolution of the background dark axion and dark gluon/glueball, we solve the equations for the density perturbations focusing on the effect of isocurvature perturbation transfer and obtain the approximated solutions for the superhorizon modes ($k = 0$)

in order to understand the parametric dependence more clearly.

A. Adiabatic perturbation

For the evolution of the multicomponent fields or fluids, the perturbations can be decomposed into the curvature (adiabatic) perturbation and the isocurvature (entropy) perturbations. The adiabatic perturbation is the modes perturbed along the direction of the background evolution, so that

$$
S_{XY} = \mathcal{H} \left(\frac{\delta \rho_X}{\rho'_X} - \frac{\delta \rho_Y}{\rho'_Y} \right) = 0 \tag{4.5}
$$

for any different species X and Y [72–74]. S_{XY} is the relative entropy perturbation, whose name can be easily understood from thermodynamics. For an isolated species which satisfies the continuity equation $\rho'_X + 3\mathcal{H}(\rho_X + p_X) = 0$, the perturbation of the entropy density s_X is given as $\delta s_X/s_X = \mathcal{H}\delta\rho_X/\rho'_X$, hence $S_{XY} = 3\delta \ln(s_X/s_Y)$.

The adiabatic mode can be described by the evolution of the comoving curvature perturbation [75],

$$
\mathcal{R} = \Phi - \frac{\mathcal{H}(\Phi' - \mathcal{H}\Psi)}{\mathcal{H}' - \mathcal{H}^2}. \tag{4.6}
$$

The corresponding initial condition for the adiabatic mode is derived as

$$
\Psi_i = -\Phi_i = -\frac{2}{3}\mathcal{R}_i, \qquad \delta_{\gamma,i} = \frac{4}{3}\mathcal{R}_i, \tag{4.7a}
$$

$$
\delta_{a,i} = -\frac{2\eta_a}{3}\mathcal{R}_i, \qquad \delta_{g,i} = \frac{4}{3}\mathcal{R}_i, \tag{4.7b}
$$

where δ_γ is for the photon fluid, and the index i indicates the time at which the initial perturbation is defined. Here we use the axion potential and mass in Eq. (2.9).

From Eqs. (2.20) and (2.22), we derive the solutions for the superhorizon modes in the radiation-dominated era:

$$
\Psi = -\Phi = -\frac{2}{3}\mathcal{R}_i, \qquad \delta_\gamma = \frac{4}{3}\mathcal{R}_i, \tag{4.8a}
$$

$$
\delta_a = \left[(1 + w_a) - (1 - w_a)\frac{m'_a(T_g(\tau))}{3\mathcal{H}m_a(T_g(\tau))} \right]\mathcal{R}_i, \tag{4.8b}
$$

$$
\delta_g = \left[(1 + w_g) + \frac{(1 - w_a)\rho_a}{\rho_g}\frac{m'_a(T_g(\tau))}{3\mathcal{H}m_a(T_g(\tau))} \right]\mathcal{R}_i, \tag{4.8c}
$$

where $m'_a(T_g(\tau)) \equiv dm_a(T_g(\tau))/d\tau$. As the scale factor becomes larger than a_{ci}, the terms proportional to m'_a are rapidly vanishing. One can show that from the continuity equation for the coupled gluon-axion fluid,

$$
\Delta(\rho_a\delta_a + \rho_g\delta_g)|_{a=a_{ci}^\pm} = 0. \tag{4.9}
$$

The detailed evolution of $m_a(T_g)$ around $a = a_{ci}$ does not lead to a different final result.

Equation (4.8) states that the adiabatic perturbation shares the same form as $\delta_X = (1 + w_X)\mathcal{R}_i$ after the confining phase transition of the dark sector, and no history dependence happens, because $S_{XY} = 0$ holds under the time evolution for the superhorizon modes. This is the characteristic feature of the adiabatic perturbation.

B. Isocurvature perturbation

The isocurvature perturbation is a mode perturbed along a direction orthogonal to the direction of background evolution. Taking S_X as

$$
S_X \equiv \mathcal{H} \left(\frac{\delta \rho_X}{\rho'_X} - \frac{\delta \rho_{\text{tot}}}{\rho'_{\text{tot}}} \right), \tag{4.10}
$$

we can trace the evolution of the individual component of the isocurvature perturbation. At the linear perturbation level, the curvature perturbation cannot generate the isocurvature perturbations, and it is also conserved on the superhorizon scales [73]. Thus, for the evolution of isocurvature perturbations, we can safely take $\mathcal{R}_i = 0$, so that $\Phi_i = 0$, $\Psi_i = 0$, and $\delta\rho_{\text{tot},i} = 0$ as the initial conditions at high temperatures, and we can solve the perturbation equations for X with the initial nonzero $\delta_{X,i}$.

The actual evolution of the density perturbation can be numerically calculated based on Eq. (2.27) and compared with the CMB and matter power spectrum. The form of the perturbation becomes particularly simple if both the axions and glueballs become CDM-like well before the matter-radiation equality. In this case, the constraints on the isocurvature perturbation can be easily provided by comparing the analytic formula in the superhorizon limit with the criteria of the Planck 2018 [1]. For this reason, let us focus on such a case.

The isocurvature perturbation of dark matter is expressed as

$$
(\hat{\delta}_{\text{DM}})_{\text{iso}} = \frac{\Omega_a}{\Omega_{\text{DM}}}(\hat{\delta}_a)_{\text{iso}} + \frac{\Omega_g}{\Omega_{\text{DM}}}(\hat{\delta}_g)_{\text{iso}}, \tag{4.11}
$$

where the "hat" denotes the Gaussian random variables satisfying $\langle \hat{\delta}_{a,i}\hat{\delta}_{g,i} \rangle = 0$, and $(\hat{\delta}_{a,g})_{\text{iso}}$ are related with $\hat{\delta}_{a,g,i}$ as

$$
\begin{bmatrix} (\hat{\delta}_a)_{\text{iso}} \\ (\hat{\delta}_g)_{\text{iso}} \end{bmatrix} = \begin{bmatrix} \mathcal{T}_{aa} & \mathcal{T}_{ag} \\ \mathcal{T}_{ga} & \mathcal{T}_{gg} \end{bmatrix} \begin{bmatrix} \hat{\delta}_{a,i} \\ \hat{\delta}_{g,i} \end{bmatrix}. \tag{4.12}
$$

For the superhorizon modes, the diagonal components of the transfer matrix are $\mathcal{T}_{aa} \simeq \mathcal{T}_{gg} \simeq 1$. The off-diagonal elements are calculated as follows.

1. Induced by the initial displacement of the axion field

The evolution of the isocurvature perturbation induced by an initial density perturbation of the axion $\delta_{a,i} = 2\delta\phi_i/\phi_i$ can be described by the input value of $\delta_{a,i}$ with the condition $\mathcal{R}_i = 0$ and the associated solutions from Eq. (2.27c). For $T_g \gg T_{g,c}$, $\partial V/\partial T_g \approx 0$, and

$$\Psi_i = -\Phi_i = 0, \qquad \delta_{\gamma,i} = 0, \qquad \delta_{g,i} = 0. \qquad (4.13)$$

In the axion-dominated dark matter scenario ($\Omega_a \gg \Omega_g$), the dominant contribution is trivial: $(\hat{\delta}_{\mathrm{DM}})_{\mathrm{iso}} \simeq \hat{\delta}_{a,i}$. In the opposite case ($\Omega_a \ll \Omega_g$), the relevant equation for $(\hat{\delta}_g)_{\mathrm{iso}}$ induced by the axion is

$$\delta_g' + 3\mathcal{H}(c_g^2 - w_g)\delta_g$$
$$\simeq \frac{(w_a - 1)\rho_a}{\rho_g}\frac{m_a'(T_g(\tau))}{m_a(T_g(\tau))}\left(1 + 3\mathcal{H}c_g^2\frac{T_g}{T_g'}\right)\delta_{a,i}, \qquad (4.14)$$

where $c_g^2 \equiv (dp_g/d\ln T_g)/(\partial\rho_g/\partial\ln T_g)$. Note that the combination of $1 + 3\mathcal{H}c_g^2 T_g/T_g'$ is vanishing in the limit of $\rho_a \ll \rho_g$ because

$$\frac{T_g'}{T_g} \simeq \left(\frac{\partial\rho_g}{\partial\ln T_g}\right)^{-1}\frac{d\rho_g}{d\tau} \simeq -3\mathcal{H}c_g^2. \qquad (4.15)$$

This implies that the transfer matrix element $\mathcal{T}_{ga}$ is of $\mathcal{O}(\Omega_a^2/\Omega_g^2)$. It is the consequence of the entropy conservation of the dark sector. So, the dominant contribution to the isocurvature perturbation of dark matter is just that from the axion dark matter.

In summary,

$$(\hat{\delta}_{\mathrm{DM}})_{\mathrm{iso}} \simeq \hat{\delta}_{a,i} \quad \text{for} \quad \Omega_g \ll \Omega_a,$$
$$(\hat{\delta}_{\mathrm{DM}})_{\mathrm{iso}} \simeq \frac{\Omega_a}{\Omega_{\mathrm{DM}}}\hat{\delta}_{a,i} \quad \text{for} \quad \Omega_a \ll \Omega_g. \qquad (4.16)$$

2. Induced by the initial fluctuation of the gluon temperature

For the initial fluctuation of the gluon temperature $\delta_{g,i} \approx 4\delta T_{g,i}/T_{g,i}$, the conditions $\mathcal{R}_i = 0$ and $\delta\phi_i = 0$ for the perturbative variables give the following initial values:

$$\Psi_i = -\Phi_i = 0, \qquad (4.17a)$$

$$\delta_{\gamma,i} = -(N^2 - 1)r^4\delta_{g,i}, \qquad (4.17b)$$

$$\delta_{a,i} = -\frac{\eta_a}{2}\delta_{g,i}. \qquad (4.17c)$$

In the glueball-dominated dark matter case, the contribution of the axion is suppressed by its energy density, so $(\hat{\delta}_{\mathrm{DM}})_{\mathrm{iso}} \simeq \hat{\delta}_{g,i}$. In the opposite case ($\Omega_g \ll \Omega_a$), the effect of the gluon temperature fluctuation to the axion perturbation can be captured by Eq. (2.24). There are three stages of the axion evolution: (I) a slow-rolling period [$H \gg m_a(T_g)$], (II) a confining phase transition to give the saturation of the axion mass $m_a(T_g) = m_a$, and (III) an axion oscillating period ($H \ll m_a$). For (I), the axion mass term is negligible and

$$\delta\phi'' + 2\mathcal{H}\delta\phi' + a^2\phi\left(\frac{3c_g^2}{4}\frac{dm_a^2(T_g)}{d\ln T_g}\right)\delta_{g,i} \simeq 0. \qquad (4.18)$$

The solution becomes

$$\frac{\delta\phi}{\phi} \simeq \frac{\eta_a}{2(2\eta_a + 4)(2\eta_a + 5)}\left(\frac{m_a^2(T_g)}{H^2}\right)\delta_{g,i}. \qquad (4.19)$$

After the confining phase transition, the perturbation of the axion in the periods (II) and (III) obeys the equation of motion without a δm_a term:

$$\delta\phi'' + 2\mathcal{H}\delta\phi' + a^2 m_a^2\delta\phi = 0. \qquad (4.20)$$

The general solution to Eq. (4.20) can be written as the sum of the Bessel functions,

$$\delta\phi = \left(\frac{4H}{m_a}\right)^{\frac{1}{4}}\sum_{\lambda=\pm}\delta\phi_\lambda J_{\lambda/4}\left(\frac{1}{2}\frac{m_a^2}{H^2}\right), \qquad (4.21)$$

with the constant coefficients $\delta\phi_\pm$. Matching Eq. (4.19) with Eq. (4.21) at $a = a_{ci}$ determines $\delta\phi_\pm$ and gives the solution during the periods (II) and (III). In the period (III), δ_a is given by

$$\delta_a \simeq \frac{\eta_a}{2\eta_a + 4}\left(\frac{m_a}{H_{ci}}\right)^2\delta_{g,i}, \qquad (4.22)$$

where H_{ci} is the Hubble rate at $a = a_{ci}$. Note that the transfer matrix element $\mathcal{T}_{ag}$ is suppressed by the factor of m_a^2/H_{ci}^2 whenever the axion dominates the dark matter, but no further suppression happens. Therefore,

$$(\hat{\delta}_{\mathrm{DM}})_{\mathrm{iso}} \simeq \hat{\delta}_{g,i} \quad \text{for } \Omega_a \ll \Omega_g,$$
$$(\hat{\delta}_{\mathrm{DM}})_{\mathrm{iso}} \simeq \left(\frac{\eta_a}{2\eta_a + 4}\left(\frac{m_a}{H_{ci}}\right)^2 + \frac{\Omega_g}{\Omega_{\mathrm{DM}}}\right)\hat{\delta}_{g,i} \quad \text{for } \Omega_g \ll \Omega_a.$$
$$(4.23)$$

C. Bound on the isocurvature perturbation

Since $\delta\phi_i$ and $\delta T_{g,i}$ are independent random fluctuations, the power spectrum can be decomposed as

$$P(k, z) = P_{\mathcal{R}\mathcal{R}}(k, z) + \sum_{X=a,g} P_{\mathcal{I}\mathcal{I},X}(k, z), \qquad (4.24)$$

where $X = a$, g stand for the isocurvature perturbations induced by $\delta_{a,i}$ and $\delta_{g,i}$, respectively. For the decomposition

of the power spectrum as $P(k,z) = (2\pi^2/k^3)\mathcal{P}(k)T(k,z)$, we can match the primordial spectrum $\mathcal{P}(k)$ with the values we obtained for the superhorizon modes in the previous section. From the observations, the adiabatic mode is nearly scale independent. For the isocurvature perturbations, we can also naturally assume they are nearly scale invariant if they originate from the fluctuations of the scalar fields during inflation as discussed in the beginning of Sec. IV. Then,

$$\mathcal{P}_{\mathcal{RR}} = A_s \left(\frac{k}{k_*}\right)^{n_s - 1}, \qquad \mathcal{P}_{\mathcal{II},X} = A_X \left(\frac{k}{k_*}\right)^{n_X - 1}, \quad (4.25)$$

where k_* is the pivot scale of the wave number, and the parameters $\{A_X, n_X\}$ are constants.

Observation of the CMB presents the upper bound on the isocurvature perturbation [1,76–78]. The constraint is expressed by the bound on the isocurvature fraction β_{iso}, which is defined by

$$\beta_{\text{iso}}(k) \equiv \frac{\mathcal{P}_{\mathcal{II}}(k)}{\mathcal{P}_{\mathcal{RR}}(k) + \mathcal{P}_{\mathcal{II}}(k)}, \qquad (4.26)$$

where $\mathcal{P}_{\mathcal{RR}}$ and $\mathcal{P}_{\mathcal{II}}$ are the power spectra defined in Eq. (4.25), and related with the density perturbations as $\mathcal{P}_{\mathcal{RR}} = k^3 \langle \mathcal{R}_i^2 \rangle / 2\pi^2$, $\mathcal{P}_{\mathcal{II}} = \sum_X k^3 \langle (\hat{\delta}_{\text{DM}})_{\text{iso},X}^2 \rangle / 2\pi^2$. We focus on the large scales in order to constrain the primordial perturbations from the CMB data. The constraint on β_{iso} for the pivot scale is given by [1]

$$\beta_{\text{iso}}(k_* = 0.002 \text{ Mpc}^{-1}) < 0.035. \qquad (4.27)$$

This can be compared with the value calculated in our scenario.

If the axion dominates dark matter—i.e., $\Omega_g \ll \Omega_a \simeq \Omega_{\text{DM}}$ [$R(r, f_a) < 1$ for Eq. (3.26)]—the fraction β_{iso} can be expressed as

$$\beta_{\text{iso}} \simeq \frac{\delta_{a,i}^2}{\mathcal{R}_i^2} + \left(\frac{\eta_a}{2\eta_a + 4}\left(\frac{m_a}{H_{ci}}\right)^2 + \frac{\Omega_g}{\Omega_{\text{DM}}}\right)^2 \frac{\delta_{g,i}^2}{\mathcal{R}_i^2}, \qquad (4.28)$$

where

$$\frac{m_a}{H_{ci}} \simeq 3R(r, f_a), \qquad \frac{\Omega_g}{\Omega_{\text{DM}}} \simeq (N^2 - 1)R(r, f_a)^{\frac{3}{2}}. \qquad (4.29)$$

Because $R(r, f_a) < 1$, the $\Omega_g/\Omega_{\text{DM}}$ part is always dominant in the second term of the rhs of Eq. (4.28).

In the opposite case, if the glueball is the main dark matter component—i.e., $\Omega_a \ll \Omega_g \simeq \Omega_{\text{DM}}$ [$R(r, f_a) > 1$]—we have

$$\beta_{\text{iso}} \simeq \frac{\delta_{g,i}^2}{\mathcal{R}_i^2} + \left(\frac{\Omega_a}{\Omega_{\text{DM}}}\right)^2 \frac{\delta_{a,i}^2}{\mathcal{R}_i^2} \qquad (4.30)$$

and

$$\frac{\Omega_a}{\Omega_{\text{DM}}} \simeq \frac{1}{(N^2 - 1)R(r, f_a)^{\frac{3 + \eta_a}{2 + \eta_a}}}. \qquad (4.31)$$

Although the coupling between the axion and the gluon is the key for the amount of axion dark matter, the contributions of the same coupling to the isocurvature perturbation are always subdominant. The perturbation is just close to the sum of the independent elements, which allows a large isocurvature perturbation of the subcomponent dark matter.

For the initial isocurvature perturbation induced by the axion misalignment, we can naturally take the form of $\delta_{a,i} \simeq H_I/2\pi f_a$, where H_I is the inflation Hubble rate. In the axion-dominated case, then, Eq. (4.27) provides the constraint as

$$\left(\frac{N^2 - 1}{8}\right)^2 \left(\frac{r}{0.004}\right)^6 \left(\frac{10^{15} \text{ GeV}}{f_a}\right)^3 \left(\frac{\delta_{g,i}}{10^{-3}}\right)^2$$
$$+ \left(\frac{H_I}{5 \times 10^{10} \text{ GeV}}\right)^2 \left(\frac{10^{15} \text{ GeV}}{f_a}\right)^2 \lesssim 1. \qquad (4.32)$$

In the glueball-dominated case (taking $N = 3$),

$$\left(\frac{0.001}{r}\right)^{4.73} \left(\frac{f_a}{10^{10} \text{ GeV}}\right)^{2.36} \left(\frac{H_I/2\pi f_a}{0.008}\right)^2$$
$$+ \left(\frac{\delta_{g,i}}{4 \times 10^{-6}}\right)^2 \lesssim 1. \qquad (4.33)$$

Considering both cases, we find the upper bound on the inflation Hubble rate as $H_I \lesssim 10^{11}$ GeV within the assumption of $H_I < f_a$. In the case of $H_I > f_a$, the axion cosmology is more UV dependent. We have to consider the restoration of $U(1)_{\text{PQ}}$ symmetry during inflation, the thermalization of the axions with the dark gluons, and the formation of dark axion strings. If the axions are thermalized with the gluons, the isocurvature perturbation will be mostly given by $\delta_{g,i}$. At the same time, the axion cosmic strings can leave the large density fluctuation at small scales. We relegate the study on the related cosmology to future work.

The effect of subcomponent isocurvature perturbation is not clear yet. Since the glueballs are strongly self-interacting particles, it may provide nontrivial effects when the glueball is the subcomponent dark matter with a large isocurvature perturbation.

In the following section, we study the somewhat different aspect of the subcomponent glueball dark matter in the late-time Universe.

V. SUBCOMPONENT GLUEBALL DM: FORMATION OF SUPERMASSIVE BLACK HOLES

The subcomponent self-interacting dark matter can play a certain role in the formation of supermassive black holes (SMBHs). If the self-interaction is strong enough, the gravothermal collapse of the subcomponent dark matter can occur at the center of the dark matter halo, leading to black hole formation at high redshifts $z \gtrsim 7$ [14]. From the quasar observations, we have the list of SMBHs $(z_{\rm obs}, M_{\rm BH})$ as J1342 + 0928 $(7.54, 7.8 \times 10^8 \, M_\odot)$, J1120 + 0641 $(7.09, 2.0 \times 10^9 \, M_\odot)$, J2348−3054 $(6.89, 2.1 \times 10^9 \, M_\odot)$, and also J0100 + 2802 $(6.3, 1.2 \times 10^{10} \, M_\odot)$ [16–19]. The idea is that the formation of these SMBHs can be explained by the evolution of the subcomponent dark matter.

In the standard mechanism on the formation and growth of black holes, SMBHs can exponentially increase their mass by the accretion of baryonic material. However, because the radiation pressure slows down the absorption of baryons, the rate is limited. The maximal growth rate is captured by the Salpeter time based on the Eddington limit [79,80],

$$t_{\rm Sal} = \frac{\epsilon \sigma_T}{4\pi G m_p} = \left(\frac{\epsilon}{0.1}\right) 45 \text{ Myr}, \qquad (5.1)$$

where m_p is the proton mass, σ_T is the Thomson scattering cross section, G is the gravitational constant, and ϵ is the efficiency factor which depends on the environment of the black hole. If the seed black hole is generated at t_i with a mass $M_{\rm seed}$, the black hole mass is bounded as

$$M_{\rm BH}(t) \lesssim M_{\rm seed} e^{\frac{t-t_i}{t_{\rm Sal}}}. \qquad (5.2)$$

If the seed black hole is formed at $z = 15$, the maximal black hole mass becomes $(2\text{–}6) \times 10^4 \, M_{\rm seed}$ at $z = 7$. If the seed is formed at $z = 30$, its mass becomes $(6\text{–}10) \times 10^5 \, M_{\rm seed}$. Therefore, in order to explain the SMBHs with masses of $\mathcal{O}(10^9 \, M_\odot)$ at $z \sim 7$, a seed mass should be greater than $(10^4\text{–}10^5) \, M_\odot$. This is quite challenging in the standard theory of black hole formation.

On the other hand, by solving the gravothermal fluid equations [14] and performing N-body simulation [15] with the assumption that the host halo is isolated, it is shown that such a heavy seed black hole could be generated from the gravothermal collapse of the subcomponent dark matter. Given the NFW density profile

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2} \qquad (5.3)$$

for the dominant DM component, the seed black hole is formed with the mass

$$M_{\rm seed} \simeq \beta_1 f_g M_h, \qquad (5.4)$$

when the age of the Universe becomes

$$t(z_{\rm col}) = t(z_i) + \Delta t_{\rm col}. \qquad (5.5)$$

Here, M_h is the mass of the host halo, and z corresponds to the redshift. $t(z_i)$ is the time when the virialized dark matter halo is isolated, as we assume. $\Delta t_{\rm col}$ is the duration of the gravothermal collapse of the subcomponent dark matter for given initial conditions. β_1 and $\Delta t_{\rm col}$ are both calculated numerically. The fraction factor $\beta_1 \simeq 0.025/(\ln(1 + c) - c/(1 + c))$ in Ref. [14], where c is the concentration of the NFW profile $[M_h = 4\pi \rho_s r_s^3(\ln(1 + c) - c/(1 + c))]$, and $\beta_1 \simeq 0.006$ in Ref. [15]. By comparing the dark matter halo density profiles of the two papers, we find that both results are well matched. The formation period $\Delta t_{\rm col}$ is estimated as the form

$$\Delta t_{\rm col} \simeq \beta_2 f_g^{-p} t_{\rm rel}, \qquad (5.6)$$

where $\beta_2 \simeq 456(480)$, $p = 0(2)$ in Ref. [14] ([15]), and the apparent relaxation time of the subcomponent dark matter at $t = t(z_i)$ is defined as

$$t_{\rm rel} \equiv \frac{m_g}{f_g \sigma_g \rho_s v_s}$$
$$= 0.28 \text{ Myr} \left(\frac{10 \text{ cm}^2/\text{g}}{f_g \sigma_g/m_g}\right) \left(\frac{10^9 \, M_\odot/\text{kpc}^3}{\rho_s}\right)^{3/2} \left(\frac{3 \text{ kpc}}{r_s}\right). \qquad (5.7)$$

σ_g is the elastic scattering cross section between two subcomponent dark matters (dark glueballs in our case), and v_s is the virialized velocity at $r = r_s$. Then the seed black hole can form after the period

$$\Delta t_{\rm col} \simeq 130 \text{ Myr} \left(\frac{10 \text{ cm}^2/\text{g}}{f_g^{p+1} \sigma_g/m_g}\right)$$
$$\times \left(\frac{10^9 \, M_\odot/\text{kpc}^3}{\rho_s}\right)^{3/2} \left(\frac{3 \text{ kpc}}{r_s}\right). \qquad (5.8)$$

Note that $\Delta t_{\rm col}$ can be shorter than the age of the Universe for a given z, $t(z) \simeq 550 \text{ Myr}(\frac{10}{1+z})^{3/2}$. Therefore, for the isolated halo with a mass $M_h = 10^{12} \, M_\odot$, $f_g^{p+1} \sigma_g/m_g \gtrsim (1\text{–}10) \text{ cm}^2/\text{g}$, and $f_g \lesssim 0.001\text{–}0.01$ can explain the SMBH around $z = 7$. We illustrate the formation of the seed black hole and its growth history in Fig. 3 for the halo mass $M_h = 10^{12} \, M_\odot$.

In our scenario, the dark glueball dark matter provides such a strongly interacting subcomponent dark matter. Since β_2 and p are directly estimated in N-body simulation, we take the result of Ref. [15] $(\beta_2 = 480, p = 2)$ as the benchmark value. Then, the relevant combination of the model parameters is $f_g^3 \sigma_g/m_g$, which is estimated as
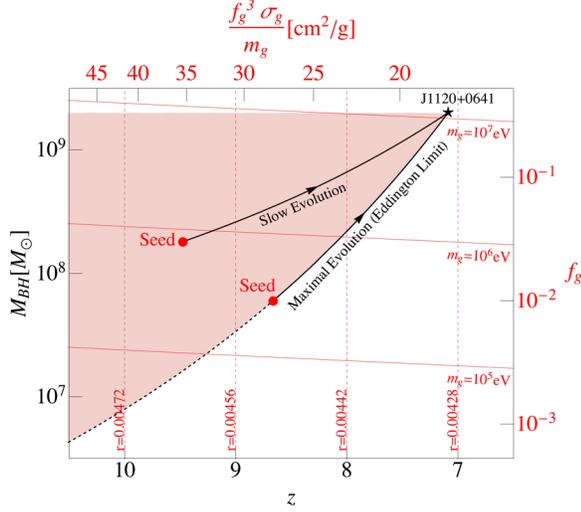
FIG. 3. Illustration of the black hole growth history for the observed high-z black hole J1120 + 0641 with the assumption of the isolated host halo ($M_h = 10^{12} \ M_\odot$) as in Refs. [14,15]. All information in red illustrates parameter space for a seed black hole (red dot). The seed black hole can be on the Eddington curve or on the shaded area in which the observations are explained by slower growth of the seed black hole. The time of collapse ($z_{\rm col}$) and the mass of the seed black hole $M_{\rm seed}$ are determined by model parameters $\{f_g, \sigma_g/m_g\}$ or $\{m_g, r\}$ for a given value of N. We take $N = 3$.

$$\frac{f_g^3 \sigma_g}{m_g} = \left(\frac{3}{N}\right)^4 \left(\frac{f_g}{m_g}\right)^3$$

$$\simeq 40 \ {\rm cm}^2/{\rm g} \left(\frac{3}{N}\right)^4 \left(\frac{N^2 - 1}{10}\right)^3 \left(\frac{r}{0.005}\right)^9. \quad (5.9)$$

For the final expression, Eq. (3.36) is used. Because it is very sensitive to r, the ratio parameter is nearly predicted from the explanation of the SMBH at high z. The corresponding allowed range of the glueball mass is also provided as $m_g = \mathcal{O}(0.05)$ MeV for $f_g = \mathcal{O}(0.01)$. As to the parameters of the dominant component of dark matter, the axion, its decay constant is $f_a = \mathcal{O}(10^{15}$–$10^{16}$ GeV), and the axion mass becomes $m_a = \mathcal{O}(10^{-14}$–$10^{-18})$ eV. This is safe from the current fuzzy dark matter constraints. Interestingly, this axion mass of 10^{-18} eV is also related to a supermassive black hole with the mass of $M_{\rm BH} \sim 10^7 \ M_\odot$ through superradiance, as we discussed before. The axions can be efficiently generated from the spinning black hole by superradiant amplification. During the amplification, the axion also takes away a sizable amount of the black hole's angular momentum, which gives a contradiction to the observation [81]. However, if the self-interaction among the axions is sizable, they will collapse before the axion cloud is saturated [82], and the loss of angular momentum is limited. For $m_a \sim 10^{-18}$ eV, the GUT-scale decay constant provides a sizable axion self-interaction to trigger bosenovae. In our case, the situation is more subtle because of

some nontrivial features of the axion potential. The perturbative quartic coupling of the axion could be suppressed by increasing N, but there is a kink structure of the potential at around $\phi \sim \pi f_a$. More detailed study is necessary to provide the correct constraints on f_a from superradiance.

Several simplifications are used in the previous discussion. Let us discuss possible caveats and an alternative history of the seed black hole formation. The host halo mass is taken as $10^{12} \ M_\odot$. This is because the halo mass is expected to be greater than $\mathcal{O}(10^3)$ times the mass of its SMBH [83,84]. In N-body simulations [85–87], the comoving number density of the cold dark matter halos with $M_h \geq 10^{12} \ M_\odot$ is evaluated as $(10^{-5}$–$10^{-6}) \ ({\rm Mpc})^{-3}$ at $z = 7$. Thus, the halo is also heavy enough to coincide with the fact that observations of SMBHs around $z = 7$ are rare.

However, since we consider the formation of the seed black hole at higher redshifts ($z > 7$), the existence of such an (isolated) heavy halo is questionable. If we extrapolate the halo mass function obtained by the N-body simulation [87], the comoving number density of the halos with $M_h \geq 10^{12} \ M_\odot$ becomes $(10^{-8}$–$10^{-9}) ({\rm Mpc})^{-3}$ at $z = 10$, and $10^{-15} \ ({\rm Mpc})^{-3}$ at $z = 15$. In this context, the issue of the formation of heavy seed black holes is just transferred to the problem of supermassive halo formation at high redshifts.

On the one hand, based on N-body simulations, we can define $M_h(z)$ at a given z in such a way that the comoving number density of the halos with their masses greater than $M_h(z)$ is given by $10^{-6}({\rm Mpc})^{-3}$. Then, $M_h(z)$ is evaluated as $10^{12} \ M_\odot$ at $z = 7$, $10^{11} \ M_\odot$ at $z = 10$, and $10^{10} \ M_\odot$ at $z = 15$. It is more natural to think of the possibility that when the seed black hole is formed at $z > 7$, the mass of the host halo is smaller than $10^{12} \ M_\odot$, although it is still one of the heaviest halos at z_i. These heaviest halos get bigger and bigger by mergers with nearby smaller halos or by accretion of the gases. The actual merger history is quite complex, but the heaviest halo is likely to remain the heaviest. In this sense, we consider $M_h(z)$ as the evolution of the host halo mass, and estimate the growth rate $\Gamma_h(z)$ as

$$\Gamma_h(z) \equiv \frac{1}{M_h(z)} \frac{dM_h(z)}{dt} \simeq \frac{4}{t(z)}. \quad (5.10)$$

The last equality holds numerically for $7 \lesssim z \lesssim 15$. The black hole growth rate by the accretion of baryons is much greater than the halo growth rate. However, the halo mass is still hierarchically larger than the black hole mass during the evolution.

Another important feature is that in terms of the halo mass, the relaxation time defined by Eq. (5.7) depends on z, c, and M_h as

$$t_{\rm rel} \propto \frac{(\ln(1+c) - \frac{c}{1+c})^{\frac{3}{2}}}{(1+z)^{\frac{7}{2}} c^{\frac{7}{2}} M_h^{\frac{1}{3}}}. \tag{5.11}$$

The concentration parameter c also depends on the halo mass and the redshift. The recent N-body simulation [88] calculates the concentration parameter $c(M_h, z)$ as the function of M_h and z in a wide range of M_h and z. With reasonable extrapolation, we find $c(10^{10}\,M_\odot, 13.8) \simeq c(10^{11}\,M_\odot, 9.5) \simeq c(10^{12}\,M_\odot, 7) = 4$–$5$. Thus, $c(M_h(z), z)$ does not have significant z dependence. Including all these considerations, Fig. 4 shows the apparent gravothermal collapse period $\Delta t_{\rm col}$ as a function of M_h and z in the units of $(f_g^3 \sigma_g / m_g)^{-1}$. The formation of the seed black hole is more efficient for heavier halos at a given z. In order to see whether or not the early formation of the seed is preferred (z dependence), we have to compare $\Delta t_{\rm col}$ with the Hubble time. Numerically, we find that the z dependence of $\Delta t_{\rm col}$ for $M_h = M_h(z)$ approximately scales as $1/(1+z)^{1.5}$ in the range $z = 7$–15 like the Hubble time. Therefore, if the seed black hole can form, the formation happens at an earlier time with a smaller mass.

Even if $\Delta t_{\rm col}$ is shorter than the age of the Universe at z_i, the isolated halo assumption may not be valid if the period of the gravothermal collapse is longer than the halo growth timescale $1/\Gamma_h$. The general expectation is that the merger process will hinder the gravothermal collapse. We consider the conservative criterion for the formation of the seed black hole as
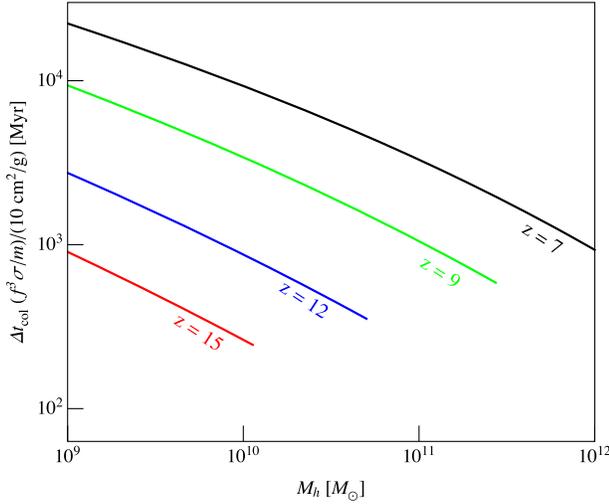
FIG. 4. The expected duration of the gravothermal collapse of the subcomponent dark matter $\Delta t_{\rm col}$ in the units of $(f_g^3 \sigma_g / m_g)^{-1}$, defined as in Eqs. (5.6) and (5.7) with $\beta_2 = 480$, $p = 2$. It is plotted for the different halo masses and redshifts based on the NFW profile of the dominant dark matter component with the fitted concentration parameter $c(M_h, z)$ [88]. The end point of each line corresponds to the halo mass $M_h = M_h(z)$. The actual collapse time of the subcomponent dark matter will depend on the halo growth history.

$$\Gamma_h(z) \Delta t_{\rm col}(z) \lesssim 4 \Delta t_{\rm col}(z) / t(z) \lesssim 1. \tag{5.12}$$

This condition means that the seed black hole can only form when the collapse process is faster than the growth rate of the halo mass. We take $z_i = 15$ as the initial redshift for the virialized heaviest host halo. Then, Eq. (5.12) is satisfied if

$$\frac{f_g^3 \sigma_g}{m_g} \gtrsim 40 \text{ cm}^2/\text{g}. \tag{5.13}$$

After the seed black hole is formed around $z = 15$, its mass is exponentially growing, and it becomes $M_{\rm BH} = 10^9\,M_\odot$ at $z = 7$ if the fraction of the glueball dark matter is given as $f_g = 2 \times 10^{-4}$. This is the case of the fastest growth, so the lower bound of f_g to explain current observations of the SMBHs is given by

$$f_g \gtrsim 2 \times 10^{-4}. \tag{5.14}$$

So far, we have ignored the effect of the number-changing interactions of the dark glueballs during the gravothermal collapse. If the number-changing process becomes efficient as the density increases, the sizable pressure of the glueballs may disturb the collapse. To simplify our discussion, in terms of the temperature of the glueball dark matter (T_g) inside the dark matter halo ($r < r_s$), there are two totally different sources to increase T_g. One is the gravothermal collapse. Because the gravitationally bound system has a negative specific heat, as heat flows outward, the glueballs become more and more concentrated in a smaller volume with a larger virial velocity. This results in temperature increasing, and leads to the collapse as the heat outflow accelerates. On the other hand, the $3 \to 2$ scatterings directly produce the large kinetic energies of the daughter glueballs as $E_{\rm kin} \simeq m_g/2$, respectively. These energies will be redistributed among glueballs within the relaxation time, so that the overall glueball temperature will increase compared to the virial temperature, and inhibit the collapse.

In order to figure out the condition for the gravothermal collapse to start, we require the criterion that the rate of glueball temperature increase be small enough to satisfy

$$\frac{\Delta t_{\rm col}}{T_g} \left(\frac{dT_g}{dt} \right)_{3 \to 2} \ll 1. \tag{5.15}$$

The temperature increase rate by the $3 \to 2$ scatterings is estimated for the given glueball density ρ_g and the velocity v_g:

$$\frac{1}{T_g} \left(\frac{dT_g}{dt} \right)_{3 \to 2} = \xi_{\rm eff} \langle \sigma_{3 \to 2} v_g^2 \rangle n_g^2 \frac{m_g}{T_g} \simeq \frac{\langle \sigma_{3 \to 2} v_g^2 \rangle \rho_g^2}{m_g^2 \langle v_g^2 \rangle}, \tag{5.16}$$

where $T_g = m_g \langle v_g^2 \rangle$, and ξ_{eff} is the efficiency factor of the energy redistribution. ξ_{eff} could be suppressed if the mean-free path of the glueball is much larger than the size of the core. In our case, most of the glueballs are trapped by the elastic scattering, so $\xi_{\text{eff}} \simeq 1$. Before the gravothermal collapse accelerates, the glueball density and the velocity are not much changed. For $\rho_g = f_g \rho_s$, $v_g = v_s$,

$$\frac{\Delta t_{\text{col}}}{T_g} \left(\frac{dT_g}{dt} \right)_{3 \to 2} \simeq 0.06 \left(\frac{10^{-3}}{f_g} \right) \left(\frac{3}{N} \right)^2 \left(\frac{10^{-3}}{v_s} \right)^2$$
$$\times \left(\frac{\text{keV}}{m_g} \right)^4 \left(\frac{\rho_s}{10^{12} \, M_\odot / \text{kpc}^3} \right). \quad (5.17)$$

Therefore, we expect that the gravothermal collapse for the SMBH would not be triggered if $m_g \lesssim \text{keV}$.

If m_g is much larger than $\mathcal{O}(\text{keV})$, the $3 \to 2$ interaction is not effective before the gravothermal collapse happens. The gravothermal collapse begins to accelerate after Δt_{col}. During the collapse, the diffusion of the dark matter mass is inefficient, and the glueballs concentrate their mass of $\mathcal{O}(M_{\text{seed}})$ around the center by increasing the core density and its temperature [89,90]. Then, the number-changing interaction becomes gradually important. It is not clear how it affects the last stage of the gravothermal collapse (the formation of the seed black hole). This is because the temperature increase rate caused by gravothermal collapse is not known yet for such a high mass density of the core. We leave it for future work.

There is also the lower bound on the glueball mass from the cosmological evolution. If the glueball is light enough, it becomes a warm or hot dark matter, so that its speed around $z = 7$–15 is greater than the escape velocity of the halo. This means that the subcomponent dark matter is not clustered, and it cannot provide a good initial condition. After the dark glueball freeze-out, its velocity scales as $1/a$. The corresponding redshifted glueball velocity at a given z is

$$v_g(z) \simeq 10^{-3} \left(\frac{1+z}{16} \right) \left(\frac{r}{0.001} \right)^{\frac{3}{2}} \left(\frac{100 \text{ eV}}{m_g} \right)^{\frac{5}{4}} \quad (5.18)$$

if the freeze-out happens before the epoch of matter-radiation equality, and

$$v_g(z) \simeq 10^{-3} \left(\frac{1+z}{16} \right) \left(\frac{r}{0.001} \right)^{\frac{4}{3}} \left(\frac{100 \text{ eV}}{m_g} \right)^{\frac{10}{9}} \quad (5.19)$$

if the freeze-out happens in the dark-matter-dominated era. In order to explain the SMBH formation, this value should be hierarchically smaller than the virial velocity $v_s \sim 10^{-3}$ during the period $z = 7$–15. In our scenario, r is nearly

fixed as 0.005; see Eq. (5.9). This implies the lower bound on m_g as 100 eV.

VI. CONCLUSIONS

We have studied the cosmological evolution of dark light scalars, whose masses and interactions originate from the approximate global symmetry and the nonperturbative dynamics of the hidden gauge symmetry. One is the feebly interacting dark axion, and the other is the strongly interacting dark glueball. Both can be dark matter if they are light enough. The equations of motion are derived and evaluated to identify the dark matter abundance and the perturbation evolution induced by the coupling between the axion and the dark gluon. We also explore the possibility that the subcomponent glueball dark matter contributes to the formation of the supermassive black hole at redshift $z \sim 7$.

Although we have dealt with the problems as closely as possible, there are still many questions that have not been covered by this paper. What would be the observable consequences of the first-order confining phase transition? In our discussion, we ignore gravitational wave productions during the confining phase transition, because it is just weakly first order unless N is very large. However, if the phase transition happens around the recombination era, it may leave a footprint on the CMB. What is the exact form of the axion scalar potential and the effect of self-interactions? The scalar potential of the axion is not a simple cosine form, and a multibranch structure may provide the nontrivial effects if the axion is produced around the spinning supermassive black hole by superradiant amplification. What is the correct period of the gravothermal collapse when the fraction of the subcomponent dark matter is small enough? So far, there is no intensive study on the gravothermal collapse of the subcomponent dark matter for such a small fraction. The empirical form of the collapse time scale Δt_{col} should be confirmed for $f_g \ll 10\%$ and higher scattering cross sections. What is the effect of the number-changing interactions of the glueball dark matter for the final stage of the black hole formation? During the gravothermal collapse, one may think of the possibility that the defining phase transition occurs, because of the large density of the glueball dark matter inside the core. It would be very interesting to study the implication of such a microscopic nature of the dark matter for the final formation of the black hole.

[1] Y. Akrami *et al.* (Planck Collaboration), Astron. Astrophys. **641,** A10 (2020).

[2] M. Battaglieri *et al.*, arXiv:1707.04591.

[3] W. Hu, R. Barkana, and A. Gruzinov, Phys. Rev. Lett. **85,** 1158 (2000).

[4] M. Dine, W. Fischler, and M. Srednicki, Phys. Lett. **104B,** 199 (1981).

[5] A. Zhitnitsky, Sov. J. Nucl. Phys. **31,** 260 (1980).

[6] J. E. Kim, Phys. Rev. Lett. **43,** 103 (1979).

[7] M. A. Shifman, A. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B166,** 493 (1980).

[8] E. D. Carlson, M. E. Machacek, and L. J. Hall, Astrophys. J. **398,** 43 (1992).

[9] K. K. Boddy, J. L. Feng, M. Kaplinghat, and T. M. P. Tait, Phys. Rev. D **89,** 115017 (2014).

[10] D. N. Spergel and P. J. Steinhardt, Phys. Rev. Lett. **84,** 3760 (2000).

[11] M. A. Buen-Abad, R. Emami, and M. Schmaltz, Phys. Rev. D **98,** 083517 (2018).

[12] H. Davoudiasl and C. W. Murphy, Phys. Rev. Lett. **118,** 141801 (2017).

[13] J. Halverson, B. D. Nelson, F. Ruehle, and G. Salinas, Phys. Rev. D **98,** 043502 (2018).

[14] J. Pollack, D. N. Spergel, and P. J. Steinhardt, Astrophys. J. **804,** 131 (2015).

[15] J. Choquette, J. M. Cline, and J. M. Cornell, J. Cosmol. Astropart. Phys. 07 (2019) 036.

[16] D. J. Mortlock, S. J. Warren, B. P. Venemans, M. Patel, P. C. Hewett, R. G. McMahon, C. Simpson, T. Theuns, E. A. Gonzles-Solares, A. Adamson *et al.*, Nature (London) **474,** 616 (2011).

[17] G. De Rosa *et al.*, Astrophys. J. **790,** 145 (2014).

[18] X.-B. Wu, F. Wang, X. Fan, W. Yi, W. Zuo, F. Bian, L. Jiang, I. McGreer, R. Wang, J. Yang, Q. Yang, D. Thompson, and Y. Beletsky, Nature (London) **518,** 512 (2015).

[19] E. Banados *et al.*, Nature (London) **553,** 473 (2018).

[20] G. 't Hooft, Nucl. Phys. **B72,** 461 (1974).

[21] E. Witten, Nucl. Phys. **B160,** 57 (1979).

[22] A. Manohar and H. Georgi, Nucl. Phys. **B234,** 189 (1984).

[23] S. Coleman, *Aspects of Symmetry: Selected Erice Lectures* (Cambridge University Press, Cambridge, England, 1985).

[24] A. G. Cohen, D. B. Kaplan, and A. E. Nelson, Phys. Lett. B **412,** 301 (1997).

[25] E. Witten, Ann. Phys. (N.Y.) **128,** 363 (1980).

[26] E. Witten, Phys. Rev. Lett. **81,** 2862 (1998).

[27] L. Del Debbio, G. M. Manca, H. Panagopoulos, A. Skouroupathis, and E. Vicari, J. High Energy Phys. 06 (2006) 005.

[28] E. Vicari and H. Panagopoulos, Phys. Rep. **470,** 93 (2009).

[29] D. J. Gross, R. D. Pisarski, and L. G. Yaffe, Rev. Mod. Phys. **53,** 43 (1981).

[30] S. Borsanyi *et al.*, Nature (London) **539,** 69 (2016).

[31] N. Bartolo, P. S. Corasaniti, A. R. Liddle, and M. Malquarti, Phys. Rev. D **70,** 043532 (2004).

[32] W. Hu, Astrophys. J. **506,** 485 (1998).

[33] C.-P. Ma and E. Bertschinger, Astrophys. J. **455,** 7 (1995).

[34] L. D. McLerran, E. Mottola, and M. E. Shaposhnikov, Phys. Rev. D **43,** 2027 (1991).

[35] G. D. Moore and M. Tassler, J. High Energy Phys. 02 (2011) 105.

[36] D. T. Son and A. O. Starinets, J. High Energy Phys. 09 (2002) 042.

[37] G. Boyd, J. Engels, F. Karsch, E. Laermann, C. Legeland, M. Lutgemeier, and B. Petersson, Nucl. Phys. **B469,** 419 (1996).

[38] Y. Deng, Nucl. Phys. B, Proc. Suppl. **9,** 334 (1989).

[39] S. Datta and S. Gupta, Phys. Rev. D **82,** 114505 (2010).

[40] S. Borsanyi, G. Endrodi, Z. Fodor, S. Katz, and K. Szabo, J. High Energy Phys. 07 (**2012**) 056.

[41] M. J. Teper, in *Confinement, Duality, and Nonperturbative Aspects of QCD* (1997), pp. 43–74, https://doi.org/10.1007/b114635.

[42] M. J. Teper, arXiv:hep-th/9812187.

[43] C. J. Morningstar and M. J. Peardon, Phys. Rev. D **60,** 034509 (1999).

[44] A. Athenodorou and M. Teper, J. High Energy Phys. 11 (2020) 172.

[45] B. Lucini and G. Moraitis, Phys. Lett. B **668,** 226 (2008).

[46] B. Lucini, A. Rago, and E. Rinaldi, Phys. Lett. B **712,** 279 (2012).

[47] B. Lucini and M. Teper, J. High Energy Phys. 06 (2001) 050.

[48] G. B. West, Phys. Rev. Lett. **77,** 2622 (1996).

[49] R. Hagedorn and J. Rafelski, Phys. Lett. **97B,** 136 (1980).

[50] M. Caselle, L. Castagnini, A. Feo, F. Gliozzi, and M. Panero, J. High Energy Phys. 06 (2011) 142.

[51] M. Caselle, A. Nada, and M. Panero, J. High Energy Phys. 07 (2015) 143; 11 (2017) 016(E).

[52] H. B. Meyer, Thesis, University of Oxford, 2004.

[53] B. Lucini, M. Teper, and U. Wenger, J. High Energy Phys. 06 (2004) 012.

[54] E. Bennett, J. Holligan, D. K. Hong, J.-W. Lee, C. J. D. Lin, B. Lucini, M. Piai, and D. Vadacchino, Phys. Rev. D **102,** 011501 (2020).

[55] B. Lucini, A. Rago, and E. Rinaldi, J. High Energy Phys. 08 (2010) 119.

[56] G. Gabadadze and A. Iglesias, Phys. Lett. B **609,** 167 (2005).

[57] L. Forestell, D. E. Morrissey, and K. Sigurdson, Phys. Rev. D **95,** 015032 (2017).

[58] C. Bonati, M. D'Elia, P. Rossi, and E. Vicari, Proc. Sci., LATTICE2016 (**2017**) 348 [arXiv:1702.01049].

[59] D. Gaiotto, A. Kapustin, Z. Komargodski, and N. Seiberg, J. High Energy Phys. 05 (**2017**) 091.

[60] S. Dubovsky, A. Lawrence, and M. M. Roberts, J. High Energy Phys. 02 (2012) 053.

[61] F. Bigazzi, A. L. Cotrone, and R. Sisca, J. High Energy Phys. 08 (2015) 090.

[62] N. Aghanim *et al.* (Planck Collaboration), Astron. Astrophys. **641,** A6 (2020).

[63] S. Tulin and H.-B. Yu, Phys. Rep. **730,** 1 (2018).

[64] B. S. Acharya, M. Fairbairn, and E. Hardy, J. High Energy Phys. 07 (2017) 100.

[65] D. Grin, M. A. Amin, V. Gluscevic, R. Hlzek, D. J. Marsh, V. Poulin, C. Prescod-Weinstein, and T. L. Smith, arXiv:1904.09003.

[66] E. G. Ferreira, arXiv:2005.03254.

[67] R. Brito, V. Cardoso, and P. Pani, *Superradiance: Energy Extraction, Black-Hole Bombs and Implications for*

Astrophysics and Particle Physics (Springer, New York, 2015), Vol. 906.

[68] R. Hlozek, D. J. E. Marsh, and D. Grin, Mon. Not. R. Astron. Soc. **476,** 3063 (2018).

[69] J.-c. Hwang and H. Noh, Phys. Lett. B **680,** 1 (2009).

[70] V. Iršič, M. Viel, M. G. Haehnelt, J. S. Bolton, and G. D. Becker, Phys. Rev. Lett. **119,** 031302 (2017).

[71] A. Soni and Y. Zhang, Phys. Rev. D **93,** 115025 (2016).

[72] A. R. Liddle, D. H. Lyth, K. A. Malik, and D. Wands, Phys. Rev. D **61,** 103509 (2000).

[73] D. Wands, K. A. Malik, D. H. Lyth, and A. R. Liddle, Phys. Rev. D **62,** 043527 (2000).

[74] S. Weinberg, Phys. Rev. D **67,** 123504 (2003).

[75] V. F. Mukhanov, H. Feldman, and R. H. Brandenberger, Phys. Rep. **215,** 203 (1992).

[76] P. Crotty, J. Garcia-Bellido, J. Lesgourgues, and A. Riazuelo, Phys. Rev. Lett. **91,** 171301 (2003).

[77] H. Li, J. Liu, J.-Q. Xia, and Y.-F. Cai, Phys. Rev. D **83,** 123517 (2011).

[78] J. Väliviita and V. Muhonen, Phys. Rev. Lett. **91,** 131302 (2003).

[79] E. Salpeter, Astrophys. J. **140,** 796 (1964).

[80] M. Volonteri, Astron. Astrophys. Rev. **18,** 279 (2010).

[81] A. Arvanitaki, M. Baryakhtar, and X. Huang, Phys. Rev. D **91,** 084011 (2015).

[82] H. Yoshino and H. Kodama, Classical Quantum Gravity **32,** 214001 (2015).

[83] J. Kormendy and D. Richstone, Annu. Rev. Astron. Astrophys. **33,** 581 (1995).

[84] J. Magorrian *et al.*, Astron. J. **115,** 2285 (1998).

[85] K. Heitmann, Z. Lukic, S. Habib, and P. M. Ricker, Astrophys. J. Lett. **642,** L85 (2006).

[86] Z. Lukic, K. Heitmann, S. Habib, S. Bashinsky, and P. M. Ricker, Astrophys. J. **671,** 1160 (2007).

[87] S. Tacchella, S. Bose, C. Conroy, D. J. Eisenstein, and B. D. Johnson, Astrophys. J. **868,** 92 (2018).

[88] T. Ishiyama *et al.*, arXiv:2007.14720.

[89] S. Balberg, S. L. Shapiro, and S. Inagaki, Astrophys. J. **568,** 475 (2002).

[90] S. Balberg and S. L. Shapiro, Phys. Rev. Lett. **88,** 101301 (2002).