Cosmological beam dump: Constraints on dark scalars mixed with the Higgs boson

Anthony Fradette,[1,2] Maxim Pospelov,[1,2] Josef Pradler,[3] and Adam Ritz[1]

[1]*Department of Physics and Astronomy, University of Victoria, Victoria, British Columbia V8P 5C2, Canada*
[2]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2J 2W9, Canada*
[3]*Institute of High Energy Physics, Austrian Academy of Sciences, Nikolsdorfergasse 18, 1050 Vienna, Austria*

Precision cosmology provides a sensitive probe of extremely weakly coupled states due to thermal freeze-in production, with subsequent decays impacting physics during well-tested cosmological epochs. We explore the cosmological implications of the freeze-in production of a new scalar S via the super-renormalizable Higgs portal. If the mass of S is at or below the electroweak scale, peak freeze-in production occurs during the electroweak epoch. We improve the calculation of the freeze-in abundance by including all relevant QCD and electroweak production channels. The resulting abundance and subsequent decay of S is constrained by a combination of x-ray data, cosmic microwave background anisotropies and spectral distortions, $N_{\rm eff}$, and the consistency of big bang nucleosynthesis with observations. These probes constrain technically natural couplings for such scalars from $m_S \sim 10$ keV all the way to $m_S \sim 100$ GeV. The ensuing constraints are similar in spirit to typical beam dump limits, but extend to much smaller couplings, down to mixing angles as small as $\theta_{Sh} \sim 10^{-16}$, and to masses all the way to the electroweak scale.

I. INTRODUCTION

A pragmatic approach to searching for new physics is to focus on generic interactions that have the potential to be detected experimentally with current or upcoming technology. Classifying the interactions of new neutral states with the Standard Model (SM) according to the dimensionality of the couplings, there are only three "portal" operators that are unsuppressed by a new energy scale. The so-called scalar, vector, and neutrino portals could provide the leading connection to a hidden or dark sector, motivated by considerations of neutrino mass and dark matter, but possibly comprising a rich structure of yet-unseen particles and forces [1].

The three portals have recently been under intense experimental scrutiny (see, e.g., Refs. [2,3]), with a forth-coming program to increase sensitivity into unexplored regions of the parameter space. While collider and beam-dump experiments provide sensitivity to relatively large portal couplings, astrophysical phenomena and cosmology can provide complementary reach to much weaker couplings. Constraints generically arise as follows: thermal production of new states in the very early universe can occur with a sub-Hubble rate (a process often called "freeze-in"), which necessarily leads to a small but non-negligible abundance of such particles in the thermal bath. If the lifetime of these particles is large, they may survive to later epochs and decay during or after big bang nucleosynthesis (BBN) altering the light element yield. Longer lifetimes may lead to decays during or after the formation of the cosmic microwave background (CMB), potentially altering the detailed features observed in precision CMB experiments.

The origin of these cosmological constraints is reminiscent of the detection strategy behind a generic particle beam-dump experiment. Typically, very energetic particles in the beam initiate the production of exotic states in the target, which then propagate through a rock or dirt filter, and decay/scatter in a relatively background-free environment inside the detector, thus generating a signal of anomalous energy deposition. In the cosmological setting, the analogue of the initial beam on target is the stage of the very early hot universe, and the analogue of propagation through a dirt filter is the long stage of subsequent expansion and cooling, as the Universe evolves into a well-understood stage associated with BBN or the CMB, which is then a direct analogue of the calorimeter-type detector that measures abnormal energy deposition.

Therefore, it is appropriate to name this method of studying rare long lived particles the *cosmological beam dump*.

Cosmological constraints of this kind were first applied to the heavy neutral lepton (HNL) portal and were considered in a number of publications [4–6], resulting in stringent constraints on sterile neutrino degrees of freedom (d.o.f.) N. Cosmological constraints on the ultraweak regime of the dark photon parameter space (a new particle A'_μ that has an $\epsilon F'_{\mu\nu} F_{\mu\nu}$ coupling to photons) were explored in Ref. [7] (see also Ref. [8]). For the very small coupling constants relevant for cosmological probes, dark photons never thermalize and the abundance is determined by freeze-in production via inverse decay reactions. Subsequent energy injection from dark photon decays can alter the path of BBN and the CMB, and agreement with precision observations excludes certain disconnected regions in the parameter space. Unlike the case of HNL, thermal production of dark photons may not exhaust all channels, as bosonic states can also be copiously produced during inflation. This extra production channel has, however, a wide range of possible outcomes depending on the Hubble scale during inflation. In that sense, the limits presented in [7] are conservative.

A similar study can be carried out for the scalar portal. Unlike the cases of HNL and dark photons where the leading portal operator is of dimension four, the neutral scalar S can have a dimension three coupling to the Higgs bilinear, $SH^\dagger H$. This represents the only superrenormalizable portal that exists between the SM and any potential dark sector. Moreover, the radiative corrections to the scalar mass created by such an interaction can naturally be under control [9], ensuring technical naturalness of a small m_S. There has been significant attention paid to this interaction over the past few years, due, for example, to the idea of cosmological self-tuning of the Higgs mass through scanning via the small $SH^\dagger H$ interaction [10].

With the $SH^\dagger H$ interaction alone, S is guaranteed to be produced in the early universe, and its subsequent decay may occur during cosmological epochs that are sensitive to energy injection, as was pointed out in Ref. [11] in an application to BBN. It is therefore interesting in our search for new physics to investigate the phenomenology of this interaction in the superweak regime, in which the abundance is determined by the freeze-in mechanism. The electroweak era can be identified as the main contributor to the freeze-in abundance of S, at temperatures where $t(\bar{t})$, $W^\pm$, Z, h are thermally excited, due to the preferential coupling of the SM Higgs to heavy particles. This was first recognized in [12] in the context of the quadratic $S^2 H^\dagger H$ interaction for $m_S > 1$ MeV, but the same conclusion holds for the $SH^\dagger H$ interaction. This is markedly different from N and A' freeze-in production [13,14], where the peak occurs at noticeably different temperatures T_{max} (for kinematically accessible particles):

$$T_{max}(A') \propto m_V,$$
$$T_{max}(N) \propto O(100 \text{ MeV}),$$
$$T_{max}(S) \propto M_W. \tag{1}$$

The goal of the present work is to determine the cosmological constraints on S, due to its Higgs portal coupling. This requires a computation of the cosmological S abundance $Y_S \equiv n_S/\tilde{s}$ due to freeze-in production, where n_S is the S number density and $\tilde{s}$ the entropy density. Existing estimates of S freeze-in have considered QCD production via top quarks, leading to $Y_S \sim 1.6 \times 10^{12}\theta^2$ [8], where θ is the mixing angle between S and the SM Higgs, and also a lower bound $Y_S \gtrsim 2.9 \times 10^9\theta^2$ on the abundance from Primakoff and Compton processes at low temperatures $T < 20$ GeV, in the context of relaxion-Higgs mixing [15]. (Further in-depth considerations of the freeze-in production due to quadratic and linear couplings of S were performed very recently in [16] and in a more generic setting in [17].) Our analysis of the S abundance and decays indicates that QCD and electroweak processes are both significant, and the conclusions are summarized below:

(i) *Freeze-in yield*: The tree-level freeze-in production of S is computed for all electroweak and QCD channels, with a T-dependent electroweak vacuum expectation value (VEV) $v(T)$ used as the first approximation of the relevant thermal effects and to provide an estimate of the precision of the calculation. Solving the Boltzmann equation numerically and incorporating a full set of QCD cross sections result in a reduction of the total QCD yield relative to the channel analyzed in [8]. We also assess the accuracy of the Maxwell-Boltzmann approximation in the production calculation, leading to the following result for the abundance from QCD and electroweak processes for the $m_S \ll M_W$ case:

$$Y_S \sim 2.8 - 5.2 \times 10^{11}\theta^2, \tag{2}$$

This estimate becomes more uncertain for masses $m_S \sim 100$ GeV, i.e., for masses close in value to the temperature/energy scale of the electroweak phase transition.

(ii) *Decay rate*: We find that there is a sizable uncertainty in the constraints for m_S in the QCD range, due to the poorly known S decay rate to pions and kaons. We show two decay profiles in this case. Additionally, we improve the calculation of the $S \to \gamma\gamma$ decay rate, which is important for $m_S < 1$ MeV, by incorporating the light quark contribution via mesonic loops, thus decreasing (or increasing) the decay rate by a factor of 4 over the rate used when u, d, s are assumed to be massive (or are neglected).

(iii) *Early decays*: Reference [8] performed a thorough analysis of the BBN constraints, but noted that their analysis of early decays for $m_S < 2m_\pi$ did not consider energy density considerations due to the large stored energy in the S bath. We include a treatment of early decays, transitioning from the freeze-in abundance to the thermalized freeze-out relic, and consider the impact on the relative neutrino and electromagnetic energy baths.

In what follows, we first review the model and describe its features in Sec. II. The freeze-in abundance calculation is described in Sec. IV, including details of several subtleties. We provide a complete scan of the S parameter space at small mixing angles in Sec. V, with details of the cosmological constraints updated in this work. The results are summarized in Fig. 1, which shows that precision cosmology provides an efficient probe of the parameter space of the model many orders of magnitude inside the region where it is "technically natural" (i.e., not plagued by the issue of fine-tuning). Finally, we conclude the paper with a general discussion of the robustness of these results and final conclusions in Sec VI. Several technical results are relegated to appendixes.

II. THE SUPERRENORMALIZABLE HIGGS PORTAL MODEL

We consider a subset of the minimal Higgs portal model, the superrenormalizable Higgs portal. The scalar part of the SM Lagrangian involving the Higgs doublet H is

augmented by a mass term for the singlet S and a dimension three interaction:

$$\mathcal{L}_{H/S} \supset \mu^2 H^\dagger H - \lambda_H (H^\dagger H)^2 - \frac{1}{2} m_S^2 S^2 - ASH^\dagger H. \quad (3)$$

The A term induces a small mixing angle θ between the physical excitations S and h. At linear order in A, the mixing angle is given by

$$\theta = \frac{Av}{m_h^2 - m_S^2}, \quad (4)$$

and leads to Yukawa interactions between S and SM particles, equivalent to the SM Higgs boson interactions rescaled by the suppression factor θ. In the unitary gauge for the broken electroweak phase, after diagonalizing to find the physical states h and S, we have the scalar potential

$$V_{H/S} = \frac{m_h^2}{2} h^2 + \frac{m_S^2}{2} S^2 + \lambda v h^3 + \frac{\lambda}{4} h^4 \quad (5)$$

$$+ \left(\frac{A}{2} - 3\theta\lambda v \right) h^2 S - \theta\lambda h^3 S, \quad (6)$$

which exhibits the hhS and $hhhS$ contact interactions.

The S sector could include additional self-interactions, e.g., $\lambda_3 S^3$. Such self-interactions would combine with the A term and contribute to the S freeze-in production via Higgs boson decays $h \to SS$. Large self-interactions can also
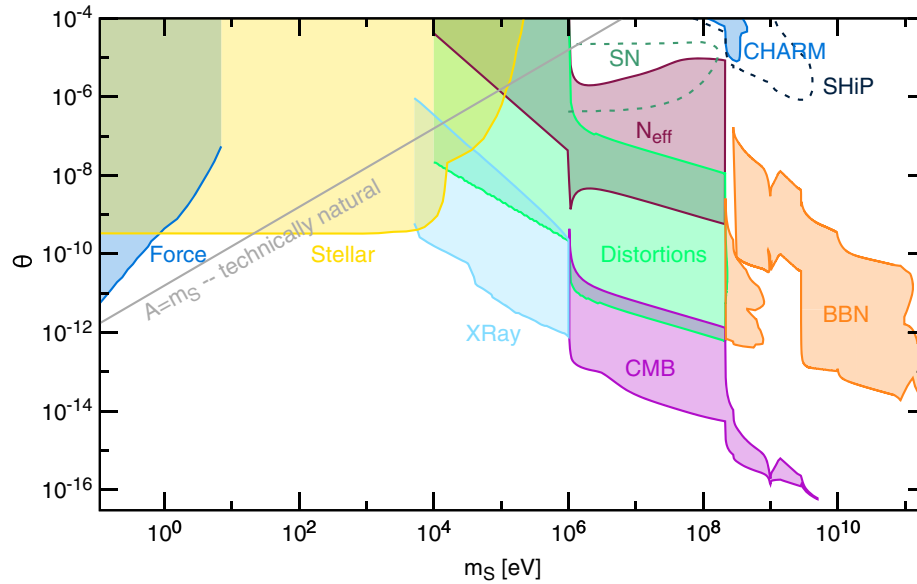

FIG. 1. An overview of the excluded parameter space for the superrenormalizable Higgs portal scalar, including the updated constraints from this work due to the diffuse x-ray background (XRay), CMB anisotropies, spectral distortions, N_{eff}, and BBN. Constraints from new short-range forces (Force) [18–21] and stellar cooling (Stellar) [22] from other authors are also shown. We also display the projected SHiP sensitivity [2] and an estimate of supernova (SN) constraints [23]. Note that the primary uncertainties impacting the limits are for the S decay rate for $m_S \sim$ GeV (see Fig. 2), and the cosmological S abundance for $m_S \gtrsim m_W$. See the text for further details.

influence the S metastable abundance after freeze-in by maintaining thermalization of the dark sector prior to subsequent decays [24,25]. We will neglect this type of interaction and focus on the pure freeze-in regime of Lagrangian (3).

A. S decay rate

The S decay rate has well-known theoretical uncertainties associated with mesonic decay channels in the mass range $2m_\pi < m_S < 4$ GeV [26]. We follow Ref. [27] and use two different decay models to demonstrate the magnitude of the theoretical uncertainty in the final S freeze-in parameter space. The baseline decay model matches low-energy theorems near the pion threshold to a $\pi\pi$ phase-shift analysis above 600 MeV by the CERN-Munich group [28] up to $m_S \lesssim 1.4$ GeV, and interpolates to $m_S \lesssim 2.5$ GeV where analytical results are expected to be valid [29]. For comparison, the spectator model uses perturbative results up to the c-quark threshold. In this case, the decay rate into pions is given by low-energy theorems and the kaon or η meson contributions are estimated by rescaling the muon branching ratio appropriately [2,30,31]. The S lifetime in this mass range for the two decay models is shown in Fig. 2. (See also the recent work [32–34].)

Below the electron threshold, a Higgs-like particle decays to two photons through a loop of heavy particles. The leading order decay rate is found by summing over the massive charged particles entering the loop [35],

$$\Gamma(S \to \gamma\gamma) = \frac{\theta^2 \alpha^2 m_S^3}{256\pi^3 v^2}|C|^2, \tag{7}$$

where C is a loop function, given explicitly in Appendix A. In this prescription, the light quark d.o.f. are incorporated through their explicit breaking of chiral symmetry, and the associated mass of pions, kaons, and eta mesons [36], i.e., through virtual loops of pions and kaons [37]. Adding the contributions from all SM particles, for $m_S \ll 2m_e$ we find
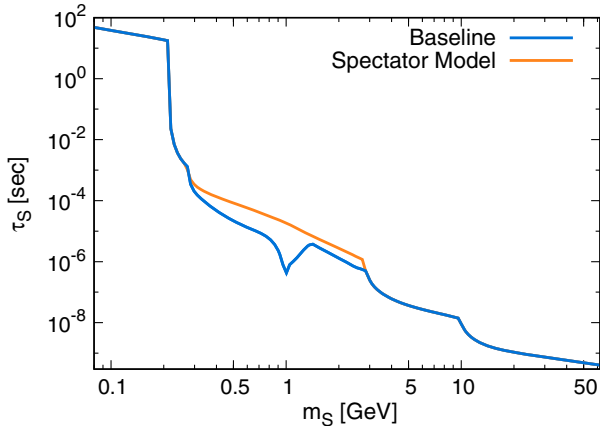

FIG. 2. The S lifetime as a function of m_S for $\theta = 10^{-6}$ (reproduced from Ref. [27]).

$$C = \begin{cases} 11/3 \simeq 3.67 & \text{for } 0+6 \text{ quarks} \\ 989/522 \simeq 1.89 & \text{for } 2+4 \text{ quarks} \\ 50/27 \simeq 1.85 & \text{for } 3+3 \text{ quarks} \\ 1 = 1 & \text{for } 0+3 \text{ quarks} \end{cases}, \tag{8}$$

where different scenarios of $(a$ light$) + (b$ heavy$)$ quarks are shown. For the case of 2 (3) light flavors the pion (pion and kaon) loops are taken into account, while for 0 light flavors they are neglected. The true physical value should be close to the $3+3$ or $2+4$ scenarios. Since the difference in decay rate between the two cases, $\mathcal{O}(4\%)$, is negligible for the analysis of new physics, we simply choose $C = 50/27$.

III. SCALAR MIXING IN THE COSMOLOGICAL THERMAL BATH

In vacuum, the relevant SM masses are generated via the Higgs mechanism and are proportional to the electroweak VEV v. In the cosmological thermal bath, and in particular near the electroweak symmetry restoration temperature, long-range interactions are screened by the plasma. Particles effectively develop a thermal mass as a representation of this screening. The mass of a particle at a given temperature T can generally be written as [38]

$$m^2(T) = m_0^2(v(T)) + m_T^2(T), \tag{9}$$

where m_0 is the zero-temperature mass that depends on the VEV and m_T is the thermal mass. Note that the VEV depends on T, so that m_0 also has a temperature dependence. A simple analytic formulation of the high-T Higgs thermal mass parameter in the effective potential is given by [39]

$$m_{h,T}^2(T) = c_h T^2, \tag{10}$$

where

$$c_h = \frac{1}{16}(8\lambda_H + 4y_t^2 + 3g_2^2 + g_1^2). \tag{11}$$

Inserting the additional term $\frac{c_h}{2}T^2 h^2$ into the Lagrangian (3) to generate the equivalent thermal mass, we can solve for $v(T)$,

$$v(T) = \sqrt{v_0^2 - \frac{c_h T^2}{\lambda_h}}, \tag{12}$$

which predicts an electroweak symmetry restoration at the critical temperature $T_c \simeq 140$ GeV.[1] The Higgs thermal

---

[1]This value is $\mathcal{O}(10\%)$ smaller than the full SM value of $T_c^{\text{SM}} \simeq 160$ GeV from lattice simulations [40,41].

mass (10) applies to the SM eigenstate prior to mass diagonalization.[2] After this diagonalization of $h - S$ mixing, we obtain a temperature-dependent mixing angle,

$$\theta(T) = \frac{Av(T)}{m_h^2(T) - m_S^2}, \qquad (13)$$

which incorporates the leading dependence on temperature for small m_S. However, this expression also signals the presence of a thermal resonance when $m_S \sim m_h(T)$, which can arise on scanning T for $m_h(T)_{\min} \lesssim m_S \lesssim m_{h,0}$. Lattice results indicate the Higgs thermal mass drops to $m_h(T)_{\min} \sim 15$ GeV at the electroweak crossover [41].

The apparent divergence in (13) at $m_S = m_h(T)$ is resolved by thermal broadening, which amounts to replacing the factor of $1/(m_h^2 - m_S^2)$ with a Breit-Wigner propagator for the intermediate metastable Higgs in the rest frame of the thermal bath. This is conveniently derived by considering the thermal rate Γ_S at which S approaches equilibrium, given by $\Gamma_S = -\mathrm{Im}\Pi_S/E$ where E is the particle energy. Γ_S in turn is related to the S production rate Γ_{prod} by a Boltzmann factor, $\Gamma_S = \Gamma_{\mathrm{prod}} - \Gamma_{\mathrm{dest}} = (e^{E/T} - 1)\Gamma_{\mathrm{prod}}$. The S self-energy takes the form

$$\Pi_S(k) = Av(T) \times \frac{1}{k^2 - m_0^2(v) - \Pi_h(k)} \times Av(T), \quad (14)$$

where Π_h is the Higgs self-energy. Computing the imaginary part in the on-shell limit, and with $\mathrm{Re}\Pi_h = m_{h,T}^2$, leads to

$$\Gamma_S \equiv \theta_{\mathrm{eff}}^2(T)\Gamma_h$$
$$= A^2 v(T)^2 \frac{\Gamma_h}{(m_S^2 - m_h^2(T))^2 + (E\Gamma_h)^2}, \qquad (15)$$

allowing us to read off the thermally broadened mixing angle,

$$\theta_{\mathrm{eff}}^2(T) = \frac{A^2 v(T)^2}{(m_S^2 - m_h(T)^2)^2 + (E\Gamma_h)^2}, \qquad (16)$$

where Γ_h is the Higgs width, or more generically, damping rate. The zero temperature width $\Gamma_{h,0} = 4.07$ MeV gives a reasonable approximation for the decay rate, since for the parameter regime of interest here, these decays occur late in the cosmological evolution when the temperature is low. Notice, however, that at temperatures around the electroweak scale, the damping rate (set by interactions with top quarks and weak gauge bosons) is expected to scale as $\Gamma_h(T) \propto T$, and $\Gamma_h(T \sim m_W) \gg \Gamma_{h,0}$. The effect of thermal broadening at high temperatures can be relevant for the epoch of freeze-in

production. A density plot of $\theta_{\mathrm{eff}}(T)$ from (16) is shown in Fig. 3. As $T \to T_c$, simulations suggest that $m_h(T)$ drops rapidly near T_c to 15–20 GeV [41], potentially allowing a resonance for any $m_S > 15$ GeV, an effect that is not well captured by our $v(T)$-scaling model. The potential importance of the thermal resonance will be considered in more detail in the next section.

We conclude this section by noting an apparent discontinuity in the behavior of the mixing angle at zero temperature and at temperatures close to the phase transition, assuming for simplicity that m_S is a small parameter. From (4), it follows that $\theta \propto A/(\lambda_H v)$, while at finite temperature $\theta_{\mathrm{eff}}(T)$, from (16), scales quite differently as $\theta_{\mathrm{eff}}(T) \propto Av(T)/(\mathrm{couplings} \times T^2)$. As $v(T)$ approaches zero, these two formulas have completely different behavior. Taken at face value, this suggests that vertices with Feynman rules proportional to $v(T)$ will not contribute at all in the electroweak symmetric phase. However, this is only true at tree level, and the surviving diagrams are generated at higher order in perturbation theory and *do not vanish* in the limit of $v(T) \to 0$. These higher order corrections are discussed further in Appendix B, and example diagrams in the higher order expansion are shown in Fig. 14. Therefore, *very near* the phase transition, the thermally corrected mixing angle (16) will not provide an
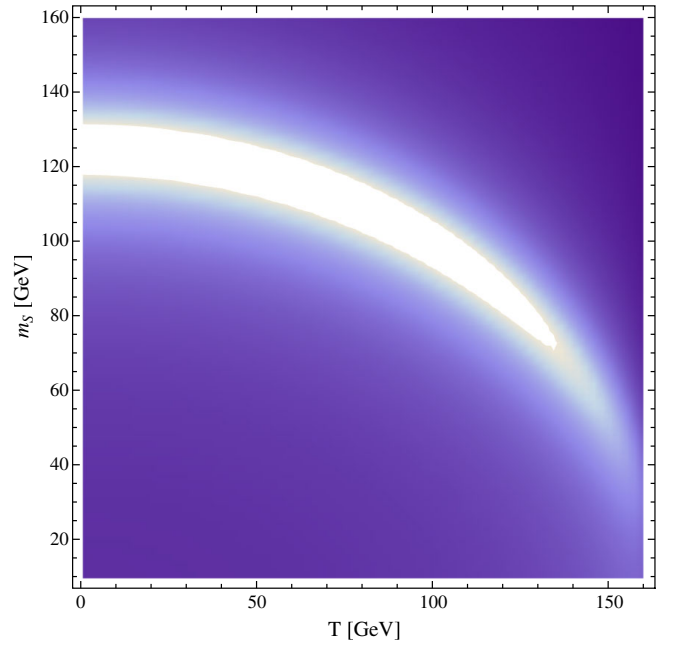

FIG. 3. A density plot of the thermal mixing angle $\theta_{\mathrm{eff}}(T)$ from (16), showing the location of the thermal resonance for $T < 140$ GeV. The peak of the resonance defines the resonance temperature T_{res} as a function of m_S. The behavior of $m_h(T)$ follows (9) with the naive $v(T)$ model, but with an additional T-dependent contribution added to ensure that $m_h(T)$ tracks down to the minimum value of $m_h(T)_{\min} \sim 15$ GeV near the crossover transition, as suggested by lattice simulations [41]. A finite Higgs damping rate of $0.05T$ was also added for illustration.

---

[2] As we are considering *very small* mixing, the effect of S on the SM thermal masses is negligible. Similarly, the S thermal mass will be $m_S^2(T) \sim \theta^2 T^2$ and thus can be neglected for this study.

adequate description of thermal effects, and a complete treatment of thermal loops would be necessary.

IV. COSMOLOGICAL PRODUCTION VIA FREEZE-IN

The cosmological production rate of a new species S, due to $2 \to 2$ interactions, is given by the Boltzmann equation

$$\tilde{s} \dot{Y} = \int \sum \prod_{i=1}^{4} \left(\frac{d^3 p_i}{2E_i (2\pi)^3} \right) \Lambda(f_1, f_2, f_3, f_4) |\mathcal{M}|^2 \\ \times (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4), \tag{17}$$

where $\tilde{s}$ is the entropy density, $Y \equiv n_S/\tilde{s}$, while $\Lambda = f_1 f_2 (1 \pm f_3)(1 \pm f_4)$ represents the thermal distribution of each species and $|\mathcal{M}|^2$ is the spin-summed squared amplitude. In the Maxwell-Boltzmann (MB) approximation for the freeze-in mechanism $\Lambda \to f_1^{\text{MB}} f_2^{\text{MB}} = e^{-\frac{(E_1+E_2)}{T}}$. The sum goes over various multiplicity factors, such as spin and color. For four different species Eq. (17) takes the form [42]

$$\tilde{s} \dot{Y}_{12 \to 3S} = \frac{g_1 g_2}{8\pi^4} T \int_{s_{\min}}^{\infty} ds \, p_{12}^2 \sqrt{s} \sigma_{12 \to 3S} K_1 \left(\frac{\sqrt{s}}{T} \right), \tag{18}$$

where $s_{\min} = \text{Max}[(m_1 + m_2)^2, (m_3 + m_4)^2]$ and

$$p_{12}^2 = \frac{s}{4} \left(1 - \frac{(m_1 - m_2)^2}{s} \right) \left(1 - \frac{(m_1 + m_2)^2}{s} \right), \tag{19}$$

while $\sigma = \sigma_{12 \to 3S}$ is the standard cross section averaged over initial state d.o.f., while $g_{1(2)}$ are the spin and color multiplicity factors for initial particles.

The total S yield is found by summing all possible $12 \to 3S$ interactions where 1, 2, and 3 are SM particles. Production channels of the form $12 \to SS$ are suppressed

by an extra factor of θ and are neglected. Since S preferentially interacts with massive particles, we anticipate a large number of possible production channels around the electroweak scale. We classify the different channels by their asymptotic behavior in the electroweak unbroken phase. According to the Goldstone boson equivalence theorem [43,44], in the $v^2/s \to 0$ limit the behavior must be determined by the corresponding Goldstone bosons interactions. Expanding the Higgs doublet in the form

$$H = \begin{pmatrix} \phi^+ \\ (h + i\phi^0)/\sqrt{2} \end{pmatrix}, \tag{20}$$

we find that the only interactions producing S in the symmetric phase will be $2 \to 2$ scattering channels, $t_R Q_L \to HS, VH \to HS$ [V is a SU(2)/U(1) gauge boson], and $t_R H \to Q_L S$, shown in Fig. 4. We can therefore categorize the S-producing interactions as follows:

(i) *QCD production*, which includes all diagrams with gluons and top quarks such as $tg \to tS$.

(ii) *Yukawa annihilation*, which includes the four reactions contributing to $t_R Q_L \to HS$.

(iii) *Compton-like scattering*, which includes reactions with a quark scattering off a boson in the form of $t_R H \to Q_L S$.

(iv) *Gauge boson scattering*, which includes the reactions purely with electroweak bosons and Higgses.

We segment the production calculation into two regimes, first for $T < T_c$ with $v/(v - v(T)) > 1$, and then for $T > T_c$, where the vacuum expectation value is negligible and the dimensionful SM couplings proportional to v vanish. T close to T_c can be treated by continuity. In all instances, we compute the cross sections at tree level, with a few phenomenological improvements justified below.

In the broken phase, we incorporate the first thermal corrections by explicitly varying the electroweak VEV as in (12), treating all SM masses as VEV-dependent variables
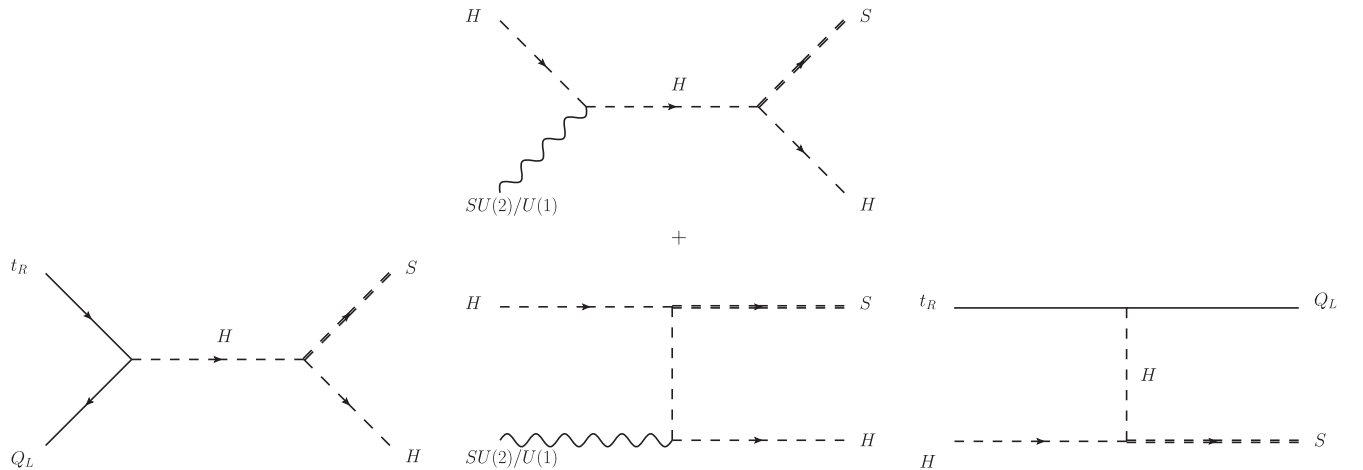

FIG. 4. S-producing interactions in the electroweak symmetric phase. *Left:* Yukawa annihilations. *Center:* Gauge boson scatterings. *Right:* Compton-like scatterings.

$$m_{\mathrm{SM}}(T) = m_{\mathrm{SM}}^0 \times \frac{v(T)}{v_0}, \qquad (21)$$

and dropping the T^2 term in Eq. (9). As we will see in Sec. IV A, we expect this approximation to hold for $v(T) \gtrsim gT$, i.e., until the temperature is high enough that the thermal masses become dominated by plasma contributions. For $m_h(T)$, the T^2 term is retained for consistency in the definition of $v(T)$ and to make sure $\theta(T)$ does not have an unphysical divergence for small m_S.

In the symmetric phase, we retain the quark masses in the cross sections and promote them to thermal masses acquired from the QCD plasma [45]

$$m_q^2(T) = \frac{g_s^2 C_F}{8} T^2 = \frac{g_s^2}{6} T^2, \qquad (22)$$

which affects the kinematic phase space available for interactions. The Higgs doublet components all obtain the Higgs thermal mass (10). We neglect the gauge boson transverse mass. From a finite-temperature point of view, the magnetic thermal mass of a non-Abelian SU(N) gauge boson vanishes at one loop [46]. A nonvanishing value is generated at higher order as a nonperturbative quantity $m_T^2 \sim (g/3\pi)gT$ [47,48], which is subleading compared to the other masses.

In the intermediate regime where a full finite-temperature calculation is needed $\frac{v(T)}{g} \lesssim T \le T_c$, we extrapolate from the two limiting regimes to obtain an uncertainty band for the model. In either case, we obtain results for the relic density that are consistent to within a factor of 2, which is acceptable for the problem at hand.

Retaining the top quark Yukawa coupling y_t, the electroweak couplings g, g', and the Higgs self-coupling λ_H as the only nonzero coupling constants, the yields from each nonvanishing production channel in the $m_S \ll m_h$ limit are compiled in Table I. In total, on including the $v(T)$ model, we obtain the following result for the abundance from QCD and electroweak processes,

$$Y_S|_{\mathrm{MB\,approx}} \sim 3.1 - 3.8 \times 10^{11}\theta^2, \qquad (23)$$

for m_S well below the thermal resonance region. The quoted uncertainty band corresponds to whether we cut off the production at $T \le v/g_s \simeq 121$ GeV or we push the extrapolation to $T \le T_c \simeq 140$ GeV. Later in this section, we will also make an estimate of the precision of the Maxwell-Boltzmann approximation, which will enlarge the precision band somewhat [see Eq. (43)]. Note that the total yield from QCD reactions is found to be $Y_S^{(\mathrm{QCD})} \simeq (6.3 - 8.2) \times 10^{10}\theta^2$, a factor of approximately 20 smaller than the value quoted by Berger *et al.* [8].[3] As listed in

---

TABLE I. S freeze-in yields for small m_S. For each production channel, the yield is given in units of $10^{10}\theta^2$, separated into the near-vacuum contribution $Y_i^{v \gg 0}$ and the additional yield $Y_i^{v \gtrsim 0}$ if extrapolated to the phase transition. The yield from each production category in the symmetric phase is shown as $Y_{\mathrm{cat}}^{\mathrm{sym}}$ with the range displaying the total yield for T above T_c or extrapolated down to $T \gtrsim gv(T)$.

Production channel i	$Y_i^{v \gg 0}$	$Y_i^{v \gtrsim 0}$	Y_i^{sym}	$Y_i^{\mathrm{tot}}[10^{10}\theta^2]$
$t\bar{t} \to gS$	2.11	0.93		
$tg \to tS$ (×2)	4.17	0.90	0	6.29–8.11
$t\bar{t} \to hS$	0.41	0.08		
$t\bar{t} \to ZS$	0.44	0.11	0.03–0.05	1.72–2.01
$t\bar{b} \to W^+S$ (×2)	0.82	0.11		
$th \to tS$ (×2)	0.38	0.13		
$tZ \to tS$ (×2)	1.46	0.77		
$tW \to bS$ (×2)	3.66	1.43	0.14–0.21	14.40–17.77
$bW \to tS$ (×2)	8.70	1.11		
$Zh \to ZS$	0.26	0.10		
$ZZ \to hS$	0.33	0.17		
$WW \to hS$	0.57	0.25		
$WW \to ZS$	3.47	0.89	0.01–0.02	8.68–10.93
$Wh \to WS$ (×2)	0.46	0.16		
$WZ \to WS$ (×2)	3.57	0.69		
$hh \to hS$	0.01	<0.01	0	
Total	30.81	7.84	0.19–0.28	31.1–38.8

Table I, we also find many other channels with electroweak gauge bosons that contribute at or above the level of these QCD-induced reactions. We use FEYNCALC [49,50] to compute the relevant cross sections, and they are listed for completeness in Appendix C, primarily in the limit of large Mandelstam s. The respective emissivity as a function of temperature is shown in Figs. 5 and 6.
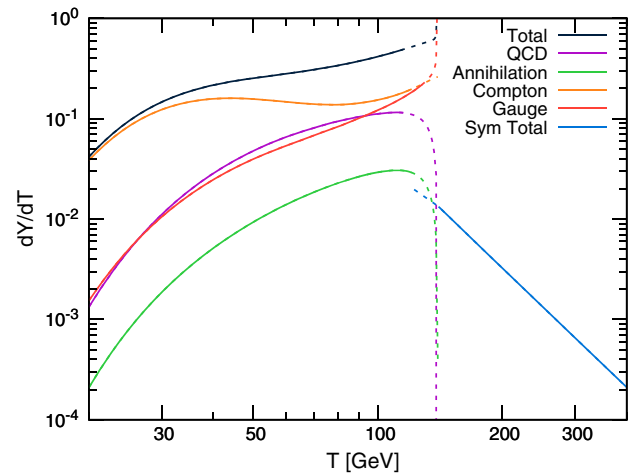

FIG. 5. Total S freeze-in emissivity and the contribution from each production channel category as a function of temperature for $\theta = 10^{-5}$.
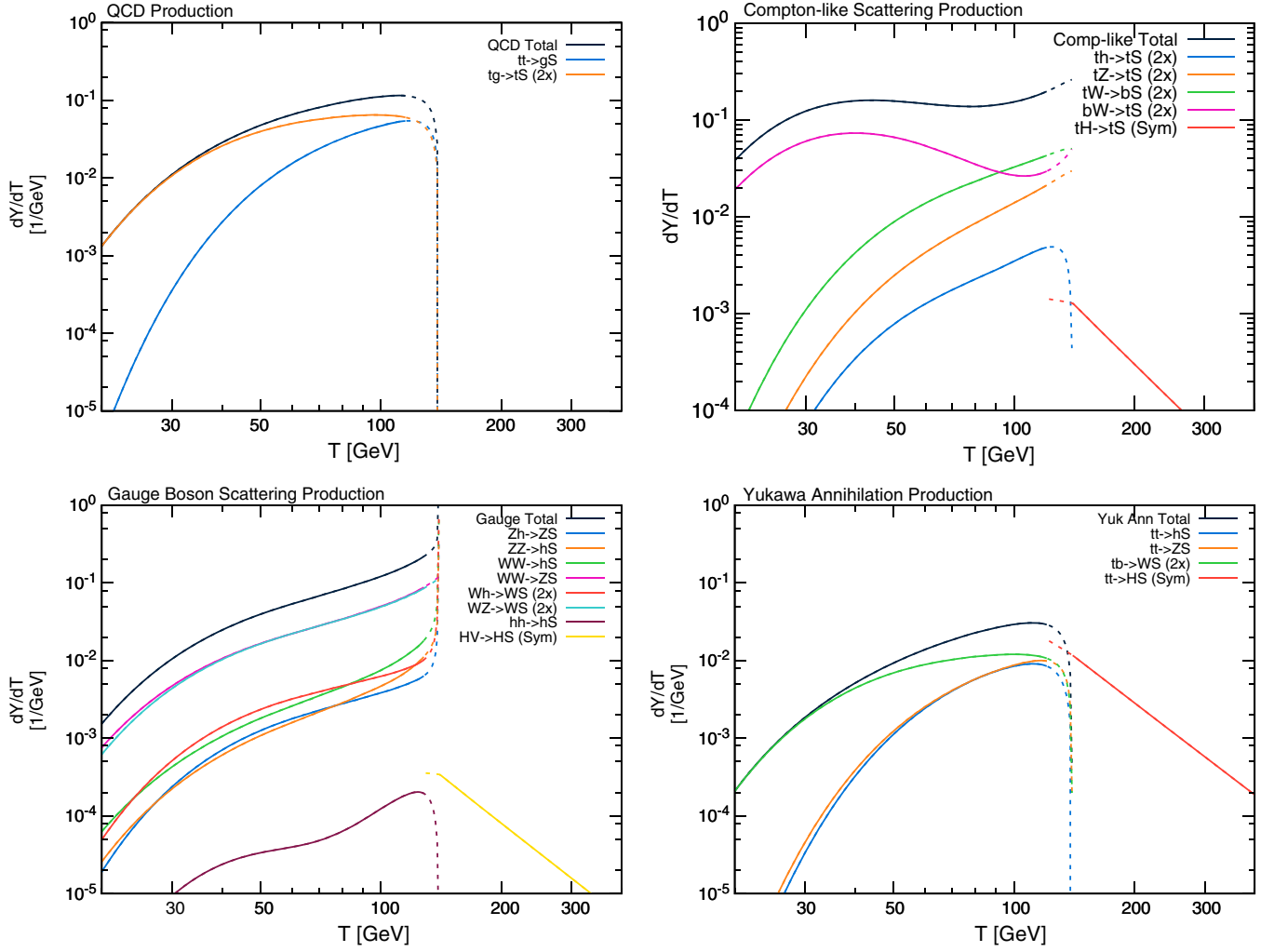
---

[3]We also point out a disagreement with one of the QCD cross sections quoted by [8], with correct expressions given in Appendix C.

FIG. 6.   The S abundance yield from each production channel category with $\theta = 10^{-5}$. *Top left:* QCD production. *Top right:* Compton-like scattering. *Bottom left:* Gauge boson scattering. *Bottom right:* Yukawa annihilation.

Thus far, we have discussed the parameter range where m_S is parametrically smaller than the weak scale. In this regime, with production occurring mostly at weak scale temperatures, the prediction for Y_S is approximately independent of m_S. However, this will change once m_S is increased to a scale comparable to the thermal masses of particles in the SM bath. In Fig. 7, we plot the resulting value of $Y_S(m_S)$ that follows from generalizing the $2 \rightarrow 2$ production mechanisms discussed above to finite m_S. We note that once m_S reaches a few tens of GeV, the $2 \rightarrow 2$ production channels may no longer be dominant, and other processes such as resonant oscillations ($1 \rightarrow 1$ production), and inverse decays ($2 \rightarrow 1$ production), may also contribute significantly to Y_S. We discuss such contributions in separate subsections below.

A. Infrared divergences

When calculating various $2 \rightarrow 2$ production processes using the simplified $v(T)$ approach, we encounter
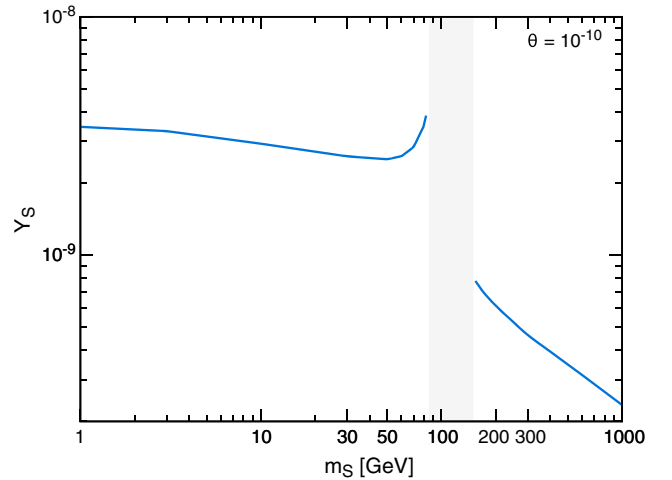

FIG. 7.   The total S abundance yield from nonresonant $2 \rightarrow 2$ production channels as a function of m_S (with $\theta = 10^{-10}$). The mass region where S is resonant with the Higgs boson is excluded, and we note that for m_S above a few tens of GeV, other production channels may also be significant. (See the text for details.)

additional complications due to the infrared sensitivity of the production cross sections. There are two types of infrared divergences in the interactions that require special care, both present in the channel with the largest yield contribution, $bW \rightarrow tS$, schematically shown in Fig. 8.

At lower temperatures, where the vacuum cross sections are clearly applicable, the emission of a soft S is enhanced by the near-on-shell t and W mediators. Removing the S emission, the inverse decay process $bW \rightarrow t$ is kinematically allowed. Since we are considering an arbitrarily light scalar, $m_S \ll m_W, m_t$, the $2 \rightarrow 2$ reactions creating the S have a kinematic cutoff $s \geq (m_t + m_S)^2$, which approaches the propagator singularity $1/(s - m_t^2)$ as $m_S \rightarrow 0$. This type of divergence is regulated by the finite width of the propagator. For this channel, we promote the denominator of the t quark and W boson propagators to their Breit-Wigner equivalent

$$\frac{1}{p^2 - m_{t/W}^2} \rightarrow \frac{1}{p^2 - m_{t/W}(T)^2 - i\Gamma_{t/W}(T)E_{t/W}}, \quad (24)$$

where $\Gamma_i(T) = \Gamma_i^0 \times \frac{v(T)}{v_0}$ is consistent with our $v(T)$ model throughout the calculation and the SM values are $\Gamma_t^0 = 1.4$ GeV and $\Gamma_W^0 = 2.1$ GeV [51]. The resonances are further broadened by thermal effects. Multiple schemes for calculating cross sections with unstable particles have been proposed beyond the simple substitution (24). The basic Breit-Wigner is technically incompatible with gauge invariance and Ward identities [52–54], a problem that can lead to dramatic inconsistencies in the small-angle scattering away from the resonance [52]. Explicitly comparing the
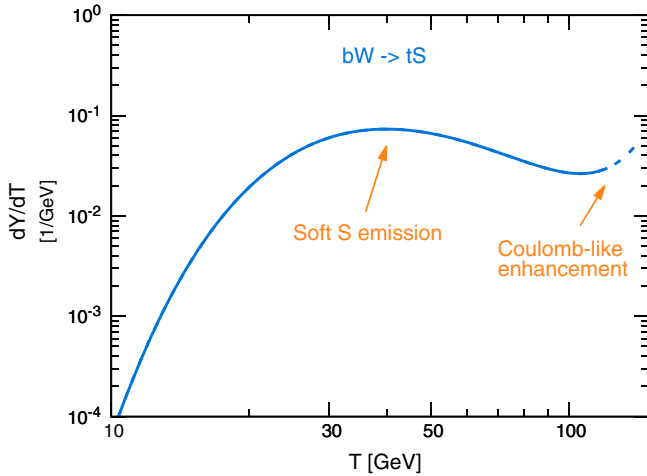

FIG. 8. The emissivity of the production channel $bW \rightarrow tS$ showing the two types of IR divergences present in the calculation. The soft S emission is physical and unique to this production channel. The Coulomb-like enhancement is present in all reactions with a t- or u-channel spin-1 mediator, is unphysical, and signifies the breakdown of our calculations. (See the text for details.)

cross section with and without the substitution (24) away from the resonance, we find

$$\frac{\sigma_{bW \rightarrow tS}^{BW}}{\sigma_{bW \rightarrow tS}^0} \xrightarrow{s \gg m_t^2, m_W^2} \frac{m_W}{\Gamma_W} \arctan \frac{\Gamma_W}{m_W} = 0.9998, \quad (25)$$

which justifies the use of Breit-Wigner propagators in this reaction near the singular point.

The second class of infrared divergences appear near T_c, where $v \rightarrow 0$. The exchange of a massless spin-1 particle in the $t(u)$ channel generates a well-known collinear divergence in forward (backward) scattering [55]. Therefore, with the approximation $m_W^2 = \frac{1}{4}g^2v^2(T)$, when the mass of m_W approaches zero, all the t-channel exchange diagrams are necessarily enhanced. Once again, thermal effects will come to the rescue and stabilize these divergences.

In the context of S freeze-in production, the total cross sections with the t- or u-channel gauge boson propagators do not fall off as $1/s$ in the high energy limit (see Appendix C for expressions). For the example shown in Fig. 8, we have

$$\sigma_{bW^+ \rightarrow tS} \rightarrow \frac{\theta^2 m_W^2}{12\pi v^4} + \mathcal{O}(1/s) \rightarrow \frac{\theta^2 g^2}{48\pi v^2} + \mathcal{O}(1/s).$$

At low temperatures, this asymptotic behavior does not matter as large values of s are exponentially suppressed by the energy available in the initial particle distributions, which clearly diverges as $v(T) \rightarrow 0$. Conceptually, this IR divergence should be regularized in the same fashion as the scalar QED example. In particular, $g^2v^2(T)/4$ will receive an additional $m_{W,T}^2(T) \propto g^2T^2$ temperature-dependent correction. Thus the Coulomb-like enhancement near $T \rightarrow T_c$ obtained from a simple extrapolation of the vacuum cross sections, with VEV-dependent masses [as in Eq. (21)], signifies the breakdown of our calculation as thermal effects are not incorporated. The formal strategy to deal with collinear IR divergences in thermal field theory has been laid out by Braaten and Yuan [56] in the weak coupling limit $g \ll 1$. We will not perform this full calculation, but simply use the limit $gT \lesssim v(T)$ as the boundary of validity for our tree-level cross sections.

B. Resonant S production

As seen in Fig. 3, the mixing angle has a physical resonance when $m_h(T) \simeq m_S$. Near resonance, the $h \rightarrow S$ oscillation may become efficient and contribute to the overall yield $Y_S(m_S)$. Below, we are going to show that the contribution of the resonance is not important for the production of very light S particles, while it can contribute significantly, starting at m_S in the range of a few tens of GeV. In practice, there is a significant uncertainty in the behavior of $m_h(T)$, and the lowest value $m_h(T)$ can acquire as a function of temperature. Recent lattice simulations suggest that near T_c the thermal Higgs mass $m_h(T)$ drops

rapidly to 15–20 GeV, in a manner reminiscent of a second-order phase transition. In principle, this allows the resonance to arise for any m_S above this scale. Physically, the resonance arises when the virtual Higgs that rotates into S is allowed to go on-shell, and the corresponding mixing angle develops a Breit-Wigner form (16) associated with the Higgs thermal width.

Although our primary interest is in lower values of m_S, where the resonance is not present, it is interesting to consider the enhancement associated with the thermal resonance. Since $\Gamma_h \ll m_h$ even after accounting for thermal broadening, we can use the narrow-width approximation (NWA) to estimate the S yield from the resonance. Taking the NWA in (16), we obtain

$$\theta_{\text{eff}}^2 \to \theta_{\text{NWA}}^2 = \frac{A^2 v^2(T_{\text{res}})\pi}{2m_S^2 E\Gamma_h(T_{\text{res}})|m_h'(T_{\text{res}})|}\delta(T - T_{\text{res}}), \quad (26)$$

where $m_h'(T_{\text{res}})$ is the temperature derivative of $m_h(T)$ evaluated at resonance. Substituting into (18), we find the simplified integral

$$Y_{S,\text{res}} \simeq \int \frac{dn_h(E, T_{\text{res}})}{s(T_{\text{res}})H(T_{\text{res}})T_{\text{res}}} \frac{\pi A^2 v^2(T_{\text{res}})}{m_S E|m_h'(T_{\text{res}})|}, \quad (27)$$

where $dn_h(T, E)$ is the Bose-Einstein distribution for the Higgs boson. Notice that the overall damping rate Γ_h for the Higgs boson drops out of this formula, and its main uncertainty is encapsulated in the value for T_{res} and $m_h'(T_{\text{res}})$. Evaluating the remaining integral (and using simplified Maxwell-Boltzmann statistics in the process), we arrive at an analytic estimate for the $h \to S$ oscillation-induced abundance,

$$Y_{S,\text{res}} \simeq \frac{\pi\theta_{\text{vac}}^2}{2(2\pi)^3} \frac{(m_{h,\text{vac}}^2 - m_S^2)^2 v(T)^2 K_1(m_S/T)}{v^2 H(T)s(T)|m_h'(T)|}, \quad (28)$$

where all thermal quantities are to be evaluated at $T = T_{\text{res}}, m_h(T_{\text{res}}) = m_S$.

The simplicity of (28) is deceiving. Depending on the assumed behavior of $m_h(T)$, the results can vary substantially. It is possible, however, to conclude that if one takes the most extreme behavior, $m_h(T) = (v(T)/v) \times m_{h,\text{vac}}$, which should conservatively overestimate the resonant contribution, the result is still quite small for small m_S. In particular, we find

$$Y_{S,\text{res}}(m_S < 2m_b) \lesssim 10^{10}\theta^2. \quad (29)$$

On the other hand, for $m_h \simeq 100$ GeV, our results indicate that $Y_{S,\text{res}}$ can reach $\sim 4 \times 10^{11}\theta^2$ and become comparable to or even larger than the nonresonant contributions. Interestingly, for m_S as high as 100 GeV, the uncertainty in the resonant contribution becomes *smaller*, due to the

fact that the resonance occurs at temperatures significantly lower than the crossover temperature.)

Another interesting observation is that for $m_S \simeq m_h$ the actual cosmological constraints are weaker than for $m_S \neq m_h$. In that limit, Eq. (28) is not applicable, and one has to retain the proper thermal damping rate for the production calculation. The point is that all constraints are very asensitive to the lifetime of S, and the effect of close-to-resonance mixing on the decay rate is very pronounced, leading to significant shortening of the lifetime and relaxation of the bounds despite enhanced production.

To ensure that our constraints are conservative, we focus on m_S below the weak scale and do not include the resonant contribution to Y_S.

C. Production via inverse decays at large m_S

We have concentrated on $2 \to 2$ and resonant production modes of the S scalar, which are dominant for m_S below the weak scale. However, at much larger values of m_S, there is also an inverse decay, or $2 \to 1$ type production channel, that we comment on briefly in this subsection.

The treatment is simplest when the mass of S is asymptotically larger than the weak scale. The decay is then predominantly to longitudinal WW, ZZ and to hh pairs, or equivalently into four pairs of real scalars. The total width is

$$\Gamma_S = \frac{A^2}{8\pi m_S} = \left.\frac{\theta^2 m_S^3}{8\pi v^2}\right|_{m_S \gg m_h}. \quad (30)$$

Production is governed by the very same width, and appropriately modifying previous results for dark photons [7], we obtain the corresponding estimate for the yield,

$$Y_S = \frac{3}{4\pi} \times \frac{\Gamma_S m_S^3}{(Hs)_{T=m_S}}. \quad (31)$$

Parametrically, this result scales as $\theta^2 M_{\text{Pl}} m_S/v^2$ or $M_{\text{Pl}} A^2/m_S^3$, where M_{Pl} is the Planck mass, while numerically we find

$$Y_{S,\text{ID}}(m_S = 1 \text{ TeV}) \simeq 2.5 \times 10^{12}\theta^2. \quad (32)$$

Notice the slightly larger overall numerical coefficient, which results from less phase space suppression for the inverse decay process compared to $2 \to 2$ processes.

Jumping forward to consider potential cosmological sensitivity in this high m_S regime, we note that the best chance of constraining the model is provided by BBN (as the lifetime is too short for other probes). Normalizing the decay width to the most sensitive lifetime window, $\Gamma_S = 1/(1000 \text{ s})$, we have

$$Y_{S,\mathrm{ID}} = 1.3 \times 10^{-18} \times \frac{\Gamma_S}{10^{-3}\,\mathrm{Hz}} \times \left(\frac{1\,\mathrm{TeV}}{m_S}\right)^2$$

$$\Rightarrow m_S Y_S = 1.3 \times 10^{-15}\,\mathrm{GeV}|_{m_S = 1\,\mathrm{TeV}}. \tag{33}$$

In spite of the large abundance, this falls about 1 order of magnitude short of the current best sensitivity [57]. In practice, the sensitivity to singlet scalars is enhanced below their decay threshold to weak bosons, $m_S < 2m_W$, where the decay width is set by the Yukawa coupling of b quarks. Correspondingly, the same lifetime is achieved through a parametrically larger value of θ, which also translates into a larger abundance, and as a consequence, tighter BBN constraints.

For $m_S \sim 100$ GeV, the inverse decay processes are subject to significant uncertainties. In particular, it is not entirely clear what the true kinematic threshold is for $WW, ZZ, hh \to S$ production via the $2 \to 1$ mechanism. For example, the thermal mass of longitudinal W's is expected to be of order $m_{W,L} \sim (0.5 - 0.6) \times T$ at temperatures around the electroweak scale. Thus, for $m_S \sim 100$ GeV, it is difficult to determine for how long the $WW \to S$ process is kinematically accessible, which renders predictions for the inverse decay processes very uncertain, and sensitive to the details of thermal physics near the electroweak crossover.

Bearing these uncertainties in mind, and given our focus on the low m_S range, to be conservative we will retain only the $2 \to 2$ production channels in analyzing the constraints below. Note that an in-depth analysis would be required to achieve a higher precision calculation of the resonant and inverse decay production channels of $O(100$ GeV$)$ singlet scalar bosons.

D. Thermalization of the S sector with the SM

The freeze-in abundance of S applies for mixing angles sufficiently small that the production rate remains below the Hubble rate,

$$\Gamma_{\mathrm{prod}} = n_i \langle \sigma v \rangle_{12 \to 3S} \lesssim H(T). \tag{34}$$

Summing all the production channels, we find that the freeze-in relic abundance obtained with Eq. (18) is valid for

$$\theta \lesssim \theta_{\mathrm{therm}} \sim 10^{-6}. \tag{35}$$

Larger mixing angles ensure complete thermalization with the SM bath before S decouples and the relic abundance is simply given by the standard freeze-out paradigm. In this case, Y_S is maintained at its relativistic equilibrium value

$$Y_{\mathrm{eq}} = \frac{45\zeta(3)}{2\pi^2 g_\star(T)} \simeq \frac{0.28}{g_\star(T)}, \tag{36}$$

until m_S becomes nonrelativistic, $T \lesssim m_S$, or the coannihilation rate becomes inefficient and S decouples with its freeze-out abundance. Since S interacts dominantly with heavy particles, S remains relativistic while the coannihilating partners become nonrelativistic and the annihilation efficiency is exponentially lowered by the phase-space suppression of the other particles. Thus, S freezes out according to Eq. (36) and the abundance depends only on the number of relativistic d.o.f. $g_\star$ at the decoupling temperature. Above the QCD confinement scale $T_{\mathrm{QCD}} \sim 200$ MeV, $g_\star$ varies by at most a factor of 2, in the range of $g_\star \sim 205/4$–$427/4$. A conservative estimate for the thermalized S relic abundance is therefore

$$Y_S^{\mathrm{f-o}} \simeq \frac{1}{400}. \tag{37}$$

E. Validity of the Maxwell-Boltzmann approximation

In the classic case of weakly interacting massive particle freeze-out, the decoupling temperature of the species is typically in the nonrelativistic regime $T_{\mathrm{decoupl}} \sim m/20$. The statistical ensemble of particle energies is well described by the MB distribution, which allows for an analytical simplification of the phase-space integrals in the Boltzmann equation. In the freeze-in scenario considered here, this simplification is not necessarily justified, and we must verify its validity. We derive in Appendix D the analytical three-dimensional expression to be numerically integrated for the S abundance including the correct statistical distribution for all particles.

Instead of proceeding with the full treatment, we can verify the MB approximation with a simpler integration. Keeping the exact statistical distributions in the Boltzmann equation, we obtain [42]

$$s\dot{Y} = \frac{1}{32\pi^4} \int ds\, p_{ij} \sqrt{s}\,\sigma \int_{\sqrt{s}}^\infty dE_+ \int_{E_-^-}^{E_-^+} dE_- f_1 f_2 (1 \pm f_3) \tag{38}$$

with

$$E_-^\pm = \frac{|m_1^2 - m_2^2|E_+}{s} \pm 2p_{ij}\sqrt{\frac{E_+^2 - s}{s}}, \tag{39}$$

where the initial energies were rewritten in terms of $E_+ = E_1 + E_2$, $E_- = |E_1 - E_2|$, f_i is the Fermi-Dirac (FD) or Bose-Einstein (BE) distribution of species i, and the $+$ $(-)$ is chosen for bosons (fermions) in the last term. The MB approximation (18) arises as an analytic solution in the MB limit $f_{1,2} = e^{-E_{1,2}/T}$ and $(1 \pm f_3) \to 1$. We should stress that Eq. (39) is not mathematically correct as E_3 should have been included in the cross section phase-space integration over the end products. This integration is in general nontrivial and includes an additional angular dependence with s (see Appendix D). Nonetheless, we can

use Eq. (39) as an estimate to the full result. To obtain the first correction beyond the MB approximation, we can expand

$$f_i = \frac{1}{e^{E_i/T} \pm 1} \simeq e^{-E_i/T}(1 \mp e^{-E_i/T} + \cdots). \quad (40)$$

It is important to notice that the first order correction for the initial particles is equal to the MB limit of the $(1 \pm f)$ term. At first order in $e^{-E_i/T}$, we have

$$f_1 f_2 (1 \pm f_3) \simeq e^{-E_+/T}(1 + \kappa_1 e^{-(E_+ + E_-)/2T}$$
$$+ \kappa_2 e^{-(E_+ - E_-)/2T} + \kappa_3 e^{-E_3/T}), \quad (41)$$

where $\kappa_i = \pm 1$, with $+$ $(-)$ for bosons (fermions). As expected, the bosonic distribution enhances the overall yield, while the fermionic distribution decreases it. In principle, E_3 is a function of $\sqrt{s}$ and the angular kinematics for the final state particles. As mentioned, the $1 \pm f_3$ factor should be included in the annihilation cross section, modifying σ. However, we know by conservation of energy that $E_1 + E_2 = E_3 + E_4$, with the following bounds on E_3

$$m_3 \leq E_3 \leq E_+. \quad (42)$$

To estimate the range of possible yield values from the first correction to the full quantum distribution, we can integrate Eq. (39) with Eq. (41) for each of the E_3 extremum values. The potential spread in total emissivity is shown in Fig. 9 with an estimated range,

$$2.8 \times 10^{11} \theta^2 < Y_S < 5.2 \times 10^{11} \theta^2. \quad (43)$$

The total error for the MB approximation is thus expected to be within a factor of 2. The first order correction band in
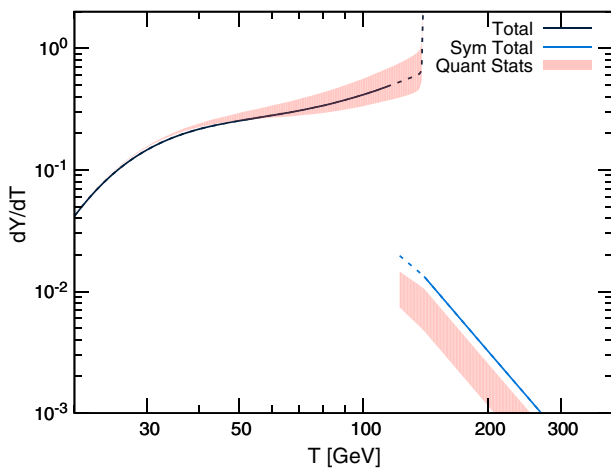

FIG. 9. The total S emissivity as a function of temperature, including the estimated error range from adopting the MB approximation compared to the correct emissivity with quantum distributions of particles 1, 2, and 3.

the symmetric phase lies completely below the MB value, because the top quark thermal masses in the QCD plasma dominate and suppress the available Fermi-Dirac distribution phase space.

V. COSMOLOGICAL CONSTRAINTS

Having determined the freeze-in abundance for small mixing angles, we can now place the minimal set of bounds on the S parameter space with $\theta \lesssim \theta_{\text{therm}} \sim 10^{-6}$. We update and improve the cosmological constraints partially presented in both Ref. [8] and Ref. [15], and the final results for the low-θ parameter space are shown in Fig. 1. The cosmological constraints that depend on Y_S are discussed in the following subsections. Below $m_S \lesssim 5$ keV, the strongest constraint on S comes from stellar energy loss [22], and even lighter scalars in the sub-eV range are constrained by fifth force experiments [18–21] (as discussed in [9]). We also show the projected sensitivity from the SHiP beam-dump experiment [2] and an order-of-magnitude estimate of supernova energy loss [23] (which should be modified to account for in-medium effects [22]). Above the pion threshold, we find strong sensitivity to the S decay model. The colored exclusion regions presented in Fig. 1 utilize the baseline decay model, which has a decay width that is larger than or equal to that in the spectator model. The baseline model therefore provides more conservative results due to the reduced abundance for a fixed lifetime.

A. Diffuse x-ray background

Many present-day satellites observe the galactic and extragalactic photon spectrum in various wavelength bands and provide upper bounds on the luminosity of x-ray or Gamma-ray emission. These bounds apply, for example, to photons from decaying or annihilating dark matter. Below $m_S = 2m_e$ and at s small mixing angle, the lifetime of S is longer than the age of the Universe, and therefore the model is constrained by these observations. In particular, Ref. [58] derived the lifetime constraint on scalar dark matter particles decaying into two photons in the 4 keV $< m_S <$ 10 GeV mass range assuming $\tau_S \gg \tau_{\text{universe}}$. We directly rescale their constraint from the HEAO-1 [59] and INTEGRAL [60] satellites to obtain an exclusion band for 4 keV $< m_S < 1$ MeV with 10^{16} s $\lesssim \tau_S \lesssim 10^{22}$ s displayed as *x-Ray* in Fig. 1.

B. CMB anisotropies

Precision measurements of the temperature and polarization anisotropies in the CMB by the WMAP [61] and Planck [62] satellites provide strong constraints on energy injection that can ionize cosmic neutral hydrogen after recombination [63–68]. The raised ionization fraction at lower redshifts allows for delayed photon interactions, which modifies the visibility function that weighs the probability of last scattering for a given CMB photon at

a specific time. This effectively damps the high-l tail of the temperature anisotropy (TT) power spectrum and increases the low-l E-mode polarization [63,64].

At redshift z_{dep}, the efficiency of energy deposition in the cosmic plasma by an energetic electron-positron pair or photons injected at an earlier redshift $z_{inj} > z_{dep}$ has been determined in Ref. [69] and updated in Ref. [70]. This update provides the energy fractions that go into ionization, excitations, heating, and emission of low-energy photons. Given the process-dependent and z-dependent ionization efficiency, comparing the modified power spectra to the CMB data is computationally intensive. In practice, principal component analysis of modified recombination histories shows that a decaying particle is well described by a constant deposition efficiency taken at $z_{dep} = 300$ [70,71]. We can then simply utilize the derived constraints for decaying particles in Ref. [7] and translate to the current model with

$$\zeta = f_{eff} \frac{m_S Y_S s_0}{m_p n_{b,0}}, \qquad (44)$$

where $f_{eff} = f(z = 300)$ is the ratio of energy absorbed leading to ionization over energy emitted at $z_{dep} = 300$. In the mass range where $S \to \mu^+\mu^-$ is the main decay channel, we solve for f_{eff} by integrating over the electron from muon decay, which decreases the ionization efficiency by a factor of 3 due to neutrinos radiating away energy. We repeat the procedure for the decay chains $S \to \pi^+\pi^- \to \mu^+\mu^-\nu\nu$, $S \to \pi^0\pi^0 \to \gamma\gamma\gamma\gamma$ and find that it is well approximated by evaluating the decay products at their average energy from the decay. We evaluate the efficiency of kaons by weighting the main branching ratios and the decay products by their average energy, percolating down to their final $e^\pm - \gamma - \nu$ spectra. Above the dicharm threshold, the light quarks, charm quark, and gluon all have similar deposition efficiencies that lie somewhere between those of electrons and muons [68]. In general, for $m_S \lesssim 10$ GeV, the efficiency tends to approach the muonic case [72]. We adopt the same ionization efficiency as muons for conservative results. The overall f_{eff} for S with $\tau_S = 10^{14}$ s is shown in Fig. 10 for both the baseline and spectator decay models.

We will not extrapolate the CMB constraints down to lifetimes $\tau_S < 10^{13}$ s because f_{eff} is not numerically stable for decays before recombination [70] and the on-the-spot approximation at $z_{dep} = 300$ fails to represent the correct physics for short lifetimes [71]. The excluded band is shown in orange within Fig. 11 for the baseline model, with the would-be exclusion region for the spectator decay model delimited by a thin gray line.

C. CMB spectral distortions

While energy injection after recombination may be observed in the CMB as variations in the anisotropies,
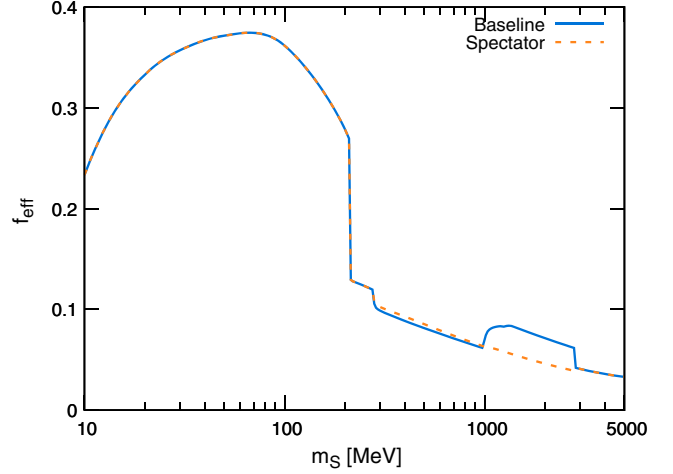

FIG. 10. Effective fraction of energy deposited leading to ionization of the cosmic plasma at $z = 300$ for $\Gamma_S = 10^{14}$ s, for the baseline and spectator decay models.

earlier energy injection can induce spectral distortion of the blackbody distribution (see Ref. [74] for a recent review) and can be used as probes of decaying particles [75,76].

Cosmological thermalization is very efficient at arbitrarily early times, and the photon plasma becomes susceptible to incomplete reequilibration of its spectrum for energy injected at $z \lesssim z_\mu \simeq 2 \times 10^6$. The CMB photons are still efficient at redistributing their energy across the energy spectrum, but double-Compton scattering and bremsstrahlung interactions that adjust the number of photons become inefficient. The bath thus develops a nonzero chemical potential in its high-energy tail resulting in the μ distortion.
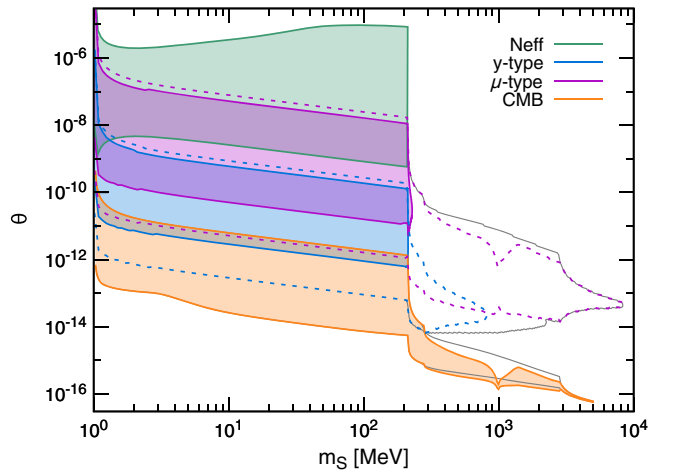

FIG. 11. Detailed cosmological constraints on S in the MeV–GeV mass range, excluding BBN (see Fig. 13). The solid lines and shaded areas represent the parameters excluded in the baseline decay model. Dashed lines refer to the projected sensitivity to spectral distortions of a PIXIE-like detector [73] and the thin gray line exhibits the would-be PIXIE sensitivity in the spectator S decay model.

At lower redshifts, $z \lesssim z_{\mu y} \simeq 5 \times 10^4$, Compton scattering between electrons and photons fails to maintain both species at a common temperature. The photon bath inherits a reduced temperature at low energies while high frequencies receive a relative gain in temperature, a phenomenon called the Compton y distortion [74].

Distortion due to arbitrary energy injection can be approximated by [77]

$$y = \frac{1}{4} \int_{z_{rec}}^{z_{\mu y}} \frac{d(Q/\rho_\gamma)}{dz'} dz', \quad (45)$$

$$\mu = 1.401 \int_{z_{\mu y}}^{\infty} e^{-\left(\frac{z'}{z_\mu}\right)^{5/2}} \frac{d(Q/\rho_\gamma)}{dz'} dz', \quad (46)$$

where $z_{rec} = 1000$ and the normalized injected electromagnetic energy is

$$\frac{d(Q/\rho_\gamma)}{dz'} = \frac{1}{\rho_\gamma} \frac{dE}{dtdV} Br_{em} \frac{1}{H(1+z')}. \quad (47)$$

In this expression, $\frac{dE}{dtdV}$ is the total energy injected and Br_{em} is the branching ratio to electromagnetic end products. In a radiation-dominated universe, the y distortion can be evaluated analytically,

$$y \simeq \frac{\sqrt{\pi}}{8} \frac{Y_S m_S s_0}{\rho_{\gamma 0} \sqrt{\Gamma_S t_0}} Br_{em} \mathcal{I}(\Gamma), \quad (48)$$

with the current entropy density $s_0 = 2891$ cm^{-3}, the current photon energy density $\rho_{\gamma 0} = 0.26$ eV cm^{-3}, a time normalization of $t_0 = 2.4 \times 10^{19}$ s and where the integral $\mathcal{I}(\Gamma)$ is defined as

$$\mathcal{I}(\Gamma) = \frac{2}{\sqrt{\pi}} \int_{\frac{\Gamma t_0}{z_{\mu y}^2}}^{\frac{\Gamma t_0}{z_{rec}^2}} e^{-\xi} \sqrt{\xi} d\xi$$

$$\rightarrow \begin{cases} 1, & \text{if } 10^{-13} \lesssim \Gamma \times s \lesssim 10^{-10}, \\ 0, & \text{if } \Gamma \ll 10^{-13} \text{ s}^{-1} \text{ or } \Gamma \gg 10^{-10} \text{ s}^{-1}. \end{cases} \quad (49)$$

The measured bounds from COBE/FIRAS [78] and the projected sensitivity from a primordial inflation explorer (PIXIE)-like detector [73] are

$$\text{COBE/FIRAS: } |y| \leq 1.5 \times 10^{-5}, \quad |\mu| \leq 9 \times 10^{-5}, \quad (50)$$

$$\text{PIXIE: } |y| \leq 2 \times 10^{-9}, \quad |\mu| \leq 1 \times 10^{-8}. \quad (51)$$

We approximate the electromagnetic branching ratio by weighting the average energy carried by end products from initial decays at rest with their respective branching ratios from S. Since the averaged energy carried away by the electron in a muon decay is $\langle E_e \rangle / m_\mu = 0.35$ [11], we have $Br_{em}^{S \to \mu^+ \mu^-} = 0.35$. The electromagnetic fractions for heavier

decay products can be found from the corresponding fractions of their lighter decay products. We find the following electromagnetic energy injection ratios:

$$E_{em}^{\mu^\pm} \simeq 0.35 \times E_{\mu^\pm}, \qquad E_{em}^{\pi^\pm} \simeq 0.27 \times E_{\pi^\pm}, \quad (52)$$

$$E_{em}^{\pi^0} = 1.00 \times E_{\pi^0}, \qquad E_{em}^{K^\pm} \simeq 0.29 \times E_{K^\pm}, \quad (53)$$

$$E_{em}^{K_S^0} \simeq 0.49 \times E_{K_S^0}, \qquad E_{em}^{K_L^0} \simeq 0.48 \times E_{K_L^0}, \quad (54)$$

where we neglected the kaon decay channels that contribute less than 10% of the kaon decay width. The total electromagnetic branching ratios for the S decay channels are

$$Br_{em}^{S \to \mu^+ \mu^-} = 0.35, \qquad Br_{em}^{S \to \pi\pi} = 0.51, \quad (55)$$

$$Br_{em}^{S \to KK} = 0.39, \qquad Br_{em}^{S \to q\bar{q}, gg} = 0.45, \quad (56)$$

where $2/3$ $(1/3)$ of pions are charged (neutral), $1/2$ $(1/4 + 1/4)$ of kaons are charged (neutral short + long), and we assume the electromagnetic yield of high energy quarks and gluons of 0.45 is maintained for c quarks. The total Br_{em} as a function of m_S is shown for the baseline and the spectator decay models in Fig. 12. The excluded regions from COBE/FIRAS are shown in Fig. 11, with a robust conservative overlap between all probes in the 1 MeV $< m_S < 2m_\mu$ mass range. A PIXIE-like detector would not change the constraints in the $m_S < 2m_\mu$ mass range, but has the potential to increase the sensitivity range to $m_S \lesssim 8$ GeV, with a sensitivity band that somewhat depends on the S decay model.
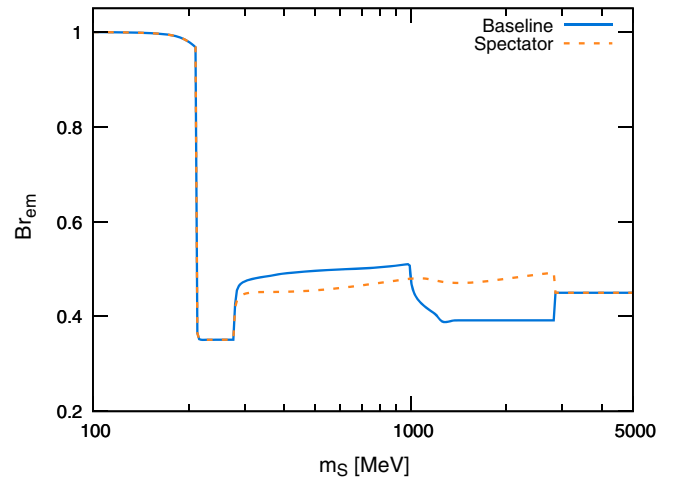

FIG. 12. Fraction of S rest energy decaying into electromagnetic energy as a function of its mass for the baseline and spectator decay models.

D. Relativistic d.o.f. ($N_{\rm eff}$)

The total relativistic energy density in the Universe at decoupling is well constrained by the CMB. The temperature of the photon bath determines its contribution to the total radiation energy density. Any additional component is parametrized in $N_{\rm eff}$, the effective number of neutrinos with a temperature $T_\nu = (4/11)^{1/3} T_\gamma$. The Planck Collaboration measurement of $N_{\rm eff} = 3.04 \pm 0.33$ [62] at 2σ is in agreement with the SM predicted value of 3.046 [79].

The constraints on early injection $\tau_S < 1$ s of $S \to \{e^+ e^-, \gamma\gamma\}$ or $S \to \mu^+ \mu^-$ have been derived in Ref. [27] including neutrino decoupling effects, which we apply using the freeze-out abundance from Eq. (37) in the mass range $10~{\rm MeV} \lesssim m_S < 2m_\mu$. For lower masses, the S lifetime at the border of the exclusion band is longer than $\tau_S = 100$ s, the maximal range derived in Ref. [27]. For longer lifetimes, we simply compare the energy density of the S sector with the SM energy densities in the neutrino and EM baths prior to the energy release (at $t = x\tau_S$) and assume an instantaneous decay. $N_{\rm eff}$ can then be estimated by comparing the relative energy density in the neutrino and EM baths

$$N_{\rm eff} = \frac{8}{7}\left(\frac{11}{4}\right)^{4/3} \frac{\rho_\nu^0 + \rho_\nu^S}{\rho_\gamma^0 + \rho_\gamma^S}, \qquad (57)$$

where $\rho_\gamma^S = Br_{\rm em}\rho_S(t = x\tau_S)$, $\rho_\nu^S = (1 - Br_{\rm em})\rho_S(t = x\tau_S)$ and the electromagnetic/neutrino energy partition of the end products $Br_{\rm em}$ is taken from Fig. 12. We choose $x = 1/10$ to match the constraints for $\tau_S < 100$ s. This choice of $x = 1/10$ is conservative. The nonrelativistic S energy density decreases less with time than the relativistic SM energy density, and larger values of x would yield stronger constraints. Using the 2σ range for $N_{\rm eff} = 3.04 \pm 0.33$, we find the constraints labeled $N_{\rm eff}$ in Figs. 1 and 11.

E. Big bang nucleosynthesis

The synthesis of light nuclei during BBN is well understood, with the final abundance of stable light elements in good agreement with predicted values (see, for example, [80–82] for recent overviews and discussions of the discrepancy with ^{7}Li). The concordance between predicted and observed abundances of ^{4}He, ^{3}He, and D can be used to constrain electromagnetic and hadronic energy injection (see Refs. [57,83,84] for reviews).

The impact of decaying particles in the BBN era depends on the ability of the decay products to efficiently interact with light nuclei, which varies as the BBN reaction network evolves and the universe cools down. Early meson decays were thoroughly discussed in Ref. [27], effectively increasing the n/p ratio before they freeze out, thus raising the ^{4}He yield above its observational limit. After most neutrons have converted to ^{4}He, the negatively charged mesons, π^- and K^-, can dissociate the copious ^{4}He, producing lighter

^{3}He, T, D, n, and p that are fed back into the reaction network [11]. This mechanism was suggested to decrease the ^{7}Li prediction by reducing the amount of ^{7}Be and resolving the discrepancy with observations [11], but incidentally also raising the D/H ratio above 3×10^{-5} from inefficient D burning, which is now excluded by observations [85,86]. Beyond $\tau_S \gtrsim 10^4$ s, the mesonic interaction rate with ambient nuclei is suppressed below the decay rate by the dilution due to expansion. The mesons instead have enough survival time to decay away in a shower of electromagnetic energy. Photodissociation of nuclei becomes efficient when photons have cascaded down below the $e^+ e^-$ thermal pair creation energy $E_{\rm th} = m_e^2/22T$. These γ rays can photodissociate D with a binding energy of $E_{\rm D}^{\rm bind} = 2.2$ MeV at $t \sim 5 \times 10^4$ s and similarly for ^{4}He with $E_{^4{\rm He}}^{\rm bind} = 19.8$ MeV at $t \sim 4 \times 10^6$ s [87].

We implement mesonic decays of $S \to \pi\pi$ and $S \to KK$ (charged and neutral) by weighting the freeze-in abundance by their respective ratios. We can then apply the constraint on early decays from Ref. [27], constraining the S parameter space by an overproduction of ^{4}He shown in blue in Fig. 13. Moreover, we use the $D/H < 3 \times 10^{-5}$ limit from Ref. [11], utilizing their stopped pion and kaon analysis for conservative results, displayed as the orange region on the S parameter space. Longer lifetimes with electromagnetic shower constraints are shown in green. We weight the S abundance by the electromagnetic branching ratio from Fig. 12 and compare the value with the electromagnetic injection limit from Ref. [57]. The upper protruding band is from a decreased D/H ratio while the lower green region is from an increased ^{3}He/D ratio. We have assumed 100%
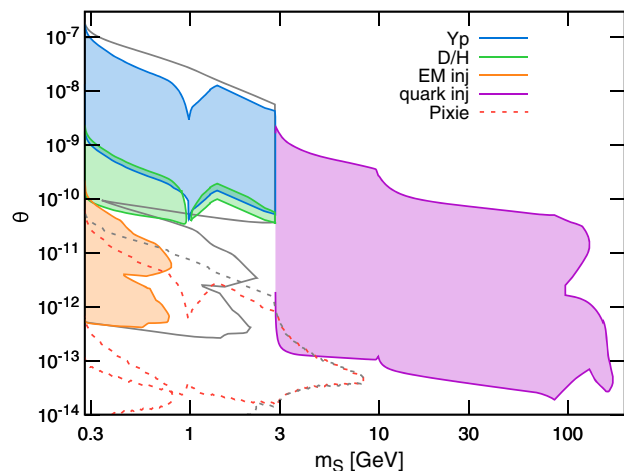

FIG. 13. BBN constraints above the di-pion threshold. Solid lines and shaded areas are ruled out if S follows the baseline decay model. Thin gray lines show the region for the spectator decay model instead. For comparison, dashed lines represent the future spectral distortion sensitivity from a PIXIE-like detector [73].

decays to kaons in the uncertain region, from $m_S \sim$ 1.4 GeV to the dicharm threshold. Note that the BBN constraint on S from mesonic decays has large uncertainties due to the unknown decay spectrum. This is illustrated by the rather different exclusion region if the spectator decay model is assumed instead, as shown by the thin gray line in Fig. 13.

Above the dicharm threshold, the S decay products are well understood and the model allows for more accurate predictions. BBN is sensitive to the total energy injected via quark pairs, with a negligible dependence on the quark flavor [57]. The sensitivity is mostly dominated by the number of hadrons produced in the hadronic shower, with a $m_S^{0.3}$ dependence on the total energy input (assuming a nonrelativistic S). Constraints from quark injection therefore increase with a smaller m_S [57]. We use the $b\bar{b}$ limits from a 30 GeV initial particle from Ref. [57] for $c\bar{c}$ and $b\bar{b}$ injections, both rescaled by their respective branching ratios. The resulting exclusion is shown in red in Fig. 13, with the upper θ range constrained by ^{4}He overabundance and the lower θ values excluded by a D overproduction.

Electromagnetic injection below the pion threshold will also affect the BBN network, through decays to muons and electrons with a nontrivial dependence on the S mass [8]. However, the S decay rate is much weaker and requires θ values close to S thermalization in order to decay during the active BBN epoch. In similar regions of parameter space, there are also significant constraints from the limits on $N_{\rm eff}$, derived in Sec. V D, along with the spectral distortion results from Sec. V C. These already exclude the parameter space with $\tau_S \sim 0.1\text{--}10^{13}$ s for $2m_e < m_S < 2m_\mu$. We therefore do not need to perform a detailed analysis of the BBN constraints below the di-pion threshold and, by the same token, the potential solution to the ^{7}Li problem tagged in Ref. [8] is also ruled out.

VI. DISCUSSION AND CONCLUSIONS

Interacting light scalar bosons are usually associated with a familiar technical naturalness problem, $m_{\rm scalar} \sim$ (couplings) $\times \Lambda_{\rm UV}$, resulting in the dilemma of justifying fine-tuning. The superrenormalizable portal offers an exception: the mass of the scalar can be natural at both tree and loop levels, as long as m_S is larger than the trilinear coupling A, or in terms of the mixing angle, $m_S > \theta \times (100\ {\rm GeV})$. Only a very few particle physics experiments (notably flavor changing decays $b \to sS$, $s \to dS$ [88]) and the fifth force experiments [9] can probe the parameter space with natural values of couplings. In contrast, cosmology within the ΛCDM framework provides a sensitive probe of this model, with couplings reaching down to $\theta \sim 10^{-16}$, which in terms of the coupling to electrons translates to $g_{Se} = \theta \times (m_e/v) \propto 10^{-21}$. (Incidentally, gravitons have approximately the same coupling strength to nonrelativistic electrons.) In this paper, we have attempted to take full advantage of the available cosmological sensitivity, by improving the calculations of metastable S abundance produced through the unique superrenormalizable portal the SM has to offer: $ASH^\dagger H$.

We have revisited the calculation of freeze-in production of the Higgs portal scalar and shown that electroweak gauge bosons and the $t - b$ quarks have non-negligible contributions to the total yield. We find the largest yield from top quark coalescence with a soft S emission, regulated by the finite top quark width. Even for $m_S \ll m_h$, the favored interactions with heavy particles push the bulk of the production just below the electroweak phase transition, in the $10\ {\rm GeV} \le T \le T_c$ temperature range. Improvements in the precision of the calculation require the full finite-temperature quantum field theory machinery that models the phase transition and includes plasma screening effects. This incremental effort is substantial and not necessarily justified for new physics searches, especially because we estimate the current level of precision to be a factor of ~ 2.

Control of the abundance calculation has allowed us to provide a full parameter space scan of the low mixing angle constraints, from sub-eV masses to 100 GeV. The cosmological constraints that depend on the primordial S abundance appear for $m_S > 10$ keV. The Higgs portal scalar has well-defined decay products below the pion threshold. It allows for rigorous constraints in the $10\ {\rm keV} < m_S < 2m_\mu$ range that fully cover 6 to 10 orders of magnitude in the mixing angle. The low θ boundary is probed by the late decays seen in the diffuse x-ray background and the CMB. In both cases, the lifetimes exceed the sensitive injection time window. This means that the number of decaying particles scales as $dN/dt \propto Y_S \Gamma_S \propto \theta^4$. We note at this point that the dependence on (coupling)^4 is a familiar one: it appears in any beam-dump experiment searching for a decay within a detector in the limit where the decay length is much larger than the size of the experiment. It also appears in the "cosmological beam-dump" constraints considered here, for the case of extremely long-lived particles. The change in S abundance by a factor of 2 would change the θ sensitivity by a factor of $2^{1/4}$, which would not be visible on the scale of the final log-log plot. For this reason, improvements in the calculation of Y_S would provide a minimal gain in precision and our approximate framework with near-vacuum cross sections is justified, at least for $m_S < 2m_\mu$. Similarly, the upper boundary for large θ is set by the freeze-out relic abundance, rather than the freeze-in abundance. Early cosmological energy injection has a logarithmic dependence on the S primordial metastable abundance [27]. This log dependence renders the exclusion region robust to variations of a few in $Y_S^{\rm f-o}$, confirming the accuracy of our conservative estimate, Eq. (37), without the need for a relativistic freeze-out computation.

Above the pion threshold, in the $2m_\pi < m_S \lesssim 2m_c$ mass range, we have derived the CMB anisotropy, spectral distortion (plus the forecast for a PIXIE-like detector), and BBN constraints for two different decay models, with significant differences in the exclusion bands. Especially near the kaon resonance, $m_S \sim 1$ GeV, the sensitivity of each probe can vary by an order of magnitude in θ. Given this poorly defined feature, we stress that improvements in the determination of the S mesonic decay width are needed before finite-temperature abundance calculations are warranted to improve the accuracy of cosmological probes.

For higher masses, where S predominantly decays to quarks, the large θ sensitivity boundary diverges sharply for short lifetimes in the $m_S Y_S$ vs τ_S plane [57]. Thus the θ constraint is insensitive to minor changes in Y_S. For $\tau_S > 10^4$ s, however, the required energy stored in the dark sector, $m_S Y_S$, is almost flat as a function of increasing lifetimes. Changes in the freeze-in yield by a factor of 2 can thus be compensated by a factor of $2^{1/4}$ in the mixing angle. We therefore conclude that cosmological constraints on S are robust to variations of Y_S by a factor of a few, except where S decays primarily to mesons, where the determination of the decay width results in a larger uncertainty.

The derived constraints are relatively conservative with respect to the structure of the dark sector provided that a quartic interaction $\lambda_S S^2 H^\dagger H$ does not thermalize the dark sector. The only additional requirement for the existence of these constraints is that S decays visibly and not into some additional stable dark state. In this case, the freeze-in production mechanism provides the minimal metastable abundance S can have for a given lifetime. Any additional interactions with other states will increase its population, until it thermalizes with the SM. Large self-interactions can potentially dilute Y_S before it decays, which would reduce the S abundance by a factor of $\ln(a_{\tau_S}/a_{f-i})$ [24]. Given that the freeze-in relic abundance is set at $T_{f-i} \sim 10$ GeV, this is a negligible factor of a few for early decays probed by BBN or N_{eff}, but can potentially be more than an order of magnitude in Y_S for lifetimes relevant to the CMB and x-ray measurements. Large self-interactions could therefore reduce the low θ sensitivity by a factor of a few.

APPENDIX A: SCALAR DECAYS BELOW THE ELECTRON THRESHOLD

As summarized in Sec. II A, below the electron mass threshold a Higgs-like particle decays to two photons at leading order through a loop of heavy particles [35],

$$\Gamma(S \to \gamma\gamma) = \frac{\theta^2 \alpha^2 m_S^3}{256\pi^3 v^2}|C|^2, \tag{A1}$$

where

$$C = \sum_f N_c Q_f^2 A_f(\tau_f) + A_W(\tau_W), \tag{A2}$$

in terms of loop functions A_f and A_W which are functions of $\tau_X = m_S^2/4m_X^2$. Note that the W contribution has the opposite interference sign, and the functions are given by

$$A_f(\tau) = \frac{2}{\tau^2}[\tau + (\tau - 1)f(\tau)] \xrightarrow[\tau \to 0]{} 4/3, \tag{A3}$$

$$A_W(\tau) = \frac{-1}{\tau^2}[2\tau^2 + 3\tau + 3(2\tau - 1)f(\tau)] \xrightarrow[\tau \to 0]{} -7, \tag{A4}$$

with

$$f(\tau) = \begin{cases} \arcsin^2 \sqrt{\tau} & \tau \le 1, \\ -\frac{1}{4}\left[\ln\frac{1+\sqrt{1-\tau^{-1}}}{1-\sqrt{1-\tau^{-1}}} - i\pi\right]^2 & \tau > 1. \end{cases} \tag{A5}$$

In this prescription, the inclusion of light quarks u, d, and s is rather ambiguous. Their masses are not physical rest masses since $m_{u,d,s} \ll \Lambda_{QCD} \simeq 350$ MeV. Instead, they are determined as chiral symmetry breaking variables in QCD which generate the nonzero mass of the Goldstone bosons of the approximate symmetry, the pions, kaons, and eta mesons [36]. A more appropriate interpretation of their contribution to $S \to \gamma\gamma$ is through virtual loops of pions and kaons [37]. Adding the contributions from all SM particles, for $m_S \ll 2m_e$ we find

$$C = \begin{cases} 11/3 \simeq 3.67 & \text{for } 0 + 6 \text{ quarks,} \\ 989/522 \simeq 1.89 & \text{for } 2 + 4 \text{ quarks,} \\ 50/27 \simeq 1.85 & \text{for } 3 + 3 \text{ quarks,} \\ 1 = 1 & \text{for } 0 + 3 \text{ quarks,} \end{cases} \tag{A6}$$

where different scenarios of $(a \text{ light}) + (b \text{ heavy})$ quarks are shown. For the case of 2(3) light flavors the pion (pion and kaon) loops are taken into account, while for 0 light flavors they are neglected. The true physical value should be close to the $3 + 3$ or $2 + 4$ scenarios.

APPENDIX B: HIGHER ORDER CORRECTIONS TO THE THERMAL MIXING ANGLE

As discussed in Sec. III, for small m_S the mixing angle θ scales quite differently at zero temperature and at temperatures close to the phase transition. From Eq. (4), we find $\theta \propto A/(\lambda_H v)$, while at finite temperature $\theta_{\text{eff}}(T)$, given by Eq. (16), scales quite differently as $\theta_{\text{eff}}(T) \propto Av(T)/(\text{couplings} \times T^2)$, particularly as $v(T) \to 0$ at the phase transition. The resolution is that there are additional diagrams which contribute to these vertices at higher order in perturbation theory and survive in the limit of $v(T) \to 0$. Example diagrams in the higher order expansion are shown in Fig. 14.

The tree-level coupling between Z and S is proportional to $g^2 Av^2(T)/m_h^2(T)$ and vanishes in the limit $v(T) \to 0$. However, it is clear that the thermal average of the full Higgs including VEV and fluctuations, $\langle (v+h)^2(T) \rangle$, does not vanish. This follows from the first thermal correction, with one thermal loop of Goldstone and Higgs bosons. Taking the soft S limit so that the Matsubara sum of the Goldstone boson loop is tractable, we find that at finite T the first vertex correction is

$$\Gamma_{ZZS}^{\mu\nu} \sim \frac{Ag^2}{4} g^{\mu\nu} \int \frac{d^3 p}{(2\pi)^3} \frac{\beta\omega e^{\beta\omega} + e^{\beta\omega} - 1}{\omega^3(1 - e^{\beta\omega})^2}, \quad \text{(B1)}$$

where $\beta = 1/T$ and ω is the energy of the boson in the loop. For a massless boson, this integral is infrared divergent $\propto \int \frac{d\omega}{\omega^2}$, a well-known feature of finite-temperature corrections [45]. However, accounting for screening with the effective thermal mass $m_{h,T}^2 = c_h T^2$, the regulated corrections scale as follows: $\Gamma_{ZZSm}^{\mu\nu} \propto (Ag^2/\sqrt{c_h})g^{\mu\nu}$. Notice that despite the associated loop suppression, this thermal loop correction is finite in the limit $v(T) \to 0$. Therefore, near the phase transition, a more complete treatment of thermal loops would be necessary.

APPENDIX C: S PRODUCTION CROSS SECTIONS

In this appendix, we exhibit the production cross sections used in the calculation of the S freeze-in abundance, in the limit of small m_S. As a couple of primary examples of QCD

production, the cross sections for a gluon scattering with a top quark to produce a S (in the $m_S \to 0$ limit) are

$$\sigma_{t\bar{t} \to gS} = \frac{\alpha_s \theta^2 y_t^2}{9s} \left[\left(1 - \frac{4m_t^2}{s} \right) \ln \left(\frac{1 + \sqrt{1 - \frac{4m_t^2}{s}}}{1 - \sqrt{1 - \frac{4m_t^2}{s}}} \right) \right.$$
$$\left. + \frac{8m_t^2}{s\sqrt{1 - \frac{4m_t^2}{s}}} \right], \quad \text{(C1)}$$

$$\sigma_{tg \to tS} = \frac{\alpha_s \theta^2 y_t^2}{96s} \left[\frac{2s(s + 3m_t^2)^2}{(s - m_t^2)^3} \ln \left(\frac{s}{m_t^2} \right) \right.$$
$$\left. - \frac{1}{(1 - \frac{s}{m_t^2})^3} \left(3 + 22\frac{m_t^2}{s} - 20\frac{m_t^4}{s^2} - 6\frac{m_t^6}{s^3} + \frac{m_t^8}{s^4} \right) \right], \quad \text{(C2)}$$

where $y_t = \sqrt{2}m_t/v$ and α_s is the strong coupling constant, evaluated at the cosmic temperature $\alpha_s = \alpha_s(T)$.

The cross sections in other channels are quite lengthy, so we collect here just the large s limits to simplify the presentation. Note, however, that full expressions are used in the numerical computation. For the Yukawa-type annihilation, we have

$$\sigma_{t\bar{t} \to hS} \to \frac{y_t^4 \theta^2}{192\pi s} \left(\ln \frac{s}{m_t^2} - 2 \right), \quad \text{(C3)}$$

$$\sigma_{t\bar{t} \to ZS} \to \frac{\theta^2}{576\pi v^4 s} \left[6m_t^2(2m_t^2 + (1 + c_v^2)m_Z^2) \ln \frac{s}{m_t^2} \right.$$
$$\left. + m_Z^2((1 + c_v^2)m_Z^2 - 24m_t^2) \right], \quad \text{(C4)}$$

$$\sigma_{t\bar{b} \to W^+S} \to \frac{\theta^2}{288\pi v^4 s} \left[3m_t^2(m_t^2 + 2m_W^2) \ln \frac{s}{m_t^2} + 2m_W^2 \right.$$
$$\left. - 12m_t^2 m_W^2 - 3m_t^2 \right], \quad \text{(C5)}$$

where $c_v = I_3 - 2Q \sin^2 \theta_W$, in terms of the eigenvalues of charge Q and isospin I_3 for the relevant fermion. The leading forms for the Compton-like scattering cross sections are
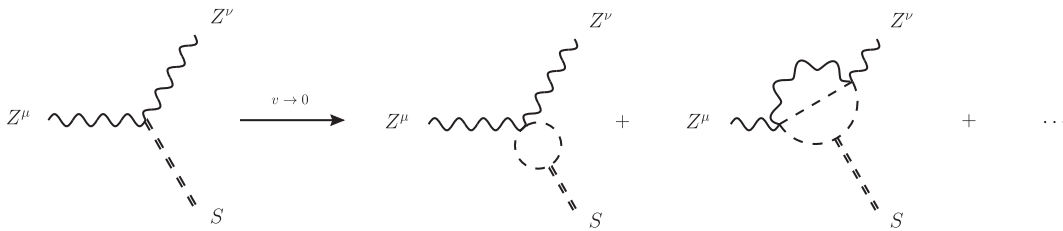

FIG. 14. Survival of the ZZS vertex at higher order in the symmetric phase.

$$\sigma_{th \to tS} \to \frac{\theta^2 y_t^4}{128\pi s}\left(2\ln\frac{s}{m_t^2}+5\right), \tag{C6}$$

$$\sigma_{tZ \to tS} \to \frac{\theta^2 m_Z^2}{48\pi v^4}(1+c_v^2)+\mathcal{O}\left(\frac{1}{s}\right), \tag{C7}$$

$$\sigma_{tW^- \to bS} \to \frac{\theta^2 m_W^2}{12\pi v^4}+\mathcal{O}\left(\frac{1}{s}\right), \tag{C8}$$

$$\sigma_{bW^+ \to tS} \to \frac{\theta^2 m_W^2}{12\pi v^4}+\mathcal{O}\left(\frac{1}{s}\right), \tag{C9}$$

and the bosonic scattering cross sections are

$$\sigma_{Zh \to ZS} \to \frac{\theta^2 m_Z^2}{12\pi v^4}, \tag{C10}$$

$$\sigma_{ZZ \to hS} \to \frac{\theta^2 m_Z^2}{36\pi v^4}, \tag{C11}$$

$$\sigma_{W^+W^- \to hS} \to \frac{\theta^2 m_W^2}{18\pi v^4}, \tag{C12}$$

$$\sigma_{W^+W^- \to ZS} \to \frac{\theta^2 m_W^2(8m_W^2+m_Z^2)}{18\pi m_Z^2 v^4}, \tag{C13}$$

$$\sigma_{W^\pm h \to W^\pm S} \to \frac{\theta^2 m_W^2}{36\pi v^4}, \tag{C14}$$

$$\sigma_{W^\pm Z \to W^\pm S} \to \frac{\theta^2(20m_W^4-3m_W^2 m_Z^2+m_Z^4)}{36\pi m_Z^2 v^4}, \tag{C15}$$

$$\sigma_{hh \to hS} \to \frac{9\theta^2 \lambda^2}{8\pi s}. \tag{C16}$$

APPENDIX D: NUMERICAL INTEGRATION WITH QUANTUM STATISTICS

We use an integration strategy based on Ref. [89] in which the authors reduced the collision integral in the context of neutrino decoupling from nine dimensions to two dimensions retaining the quantum distributions of particles. The heavy mediator limit was assumed in this reference, which is not appropriate here, but we present a strategy to reduce the number of integrals requiring numerical treatment.

We wish to integrate

$$s\dot{Y} = \int \prod_{i=1}^{4}\left(\frac{d^3 p_i}{2E_i(2\pi)^3}\right)\Lambda(f_1,f_2,f_3,f_4)$$
$$\times |\mathcal{M}|^2(2\pi)^4\delta^4(p_1+p_2-p_3-p_4), \tag{D1}$$

where Λ represents the thermal distribution of each species and $|\mathcal{M}|^2$ is the spin-summed squared amplitude. Working

in a reference frame where species 1 travels in the $\hat{x}$ direction, we define the four vectors as

$$p_1 = (E_1, \mathrm{p}_1, 0, 0), \tag{D2}$$

$$p_2 = (E_2, \mathrm{p}_2\cos\alpha, \mathrm{p}_2\sin\alpha\sin\beta, \mathrm{p}_2\sin\alpha\cos\beta), \tag{D3}$$

$$p_2 = (E_3, \mathrm{p}_3\cos\theta, 0, \mathrm{p}_3\sin\theta), \tag{D4}$$

$$p_4 = p_1 + p_2 - p_3, \tag{D5}$$

where $\mathrm{p}_i = |\vec{p}_i|$; the angle between $\vec{p}_1$ and $\vec{p}_2$ is α and between $\vec{p}_1$ and $\vec{p}_3$ is θ. Both $\vec{p}_2$ and $\vec{p}_3$ have an azimuthal angle with $\vec{p}_1$, but there is an overall azimuthal symmetry and only the difference between the two azimuthal angles matters, denoted by μ. We have used the azimuthal symmetry to fix the $\vec{p}_3$ azimuthal angle to 0. Then we have $d^3\mathrm{p}_1 d^3\mathrm{p}_2 d^3\mathrm{p}_3 = \mathrm{p}_1 E_1 dE_1 d\Omega_1 \mathrm{p}_2 E_2 dE_2 d(\cos\alpha)d\beta \mathrm{p}_3 E_3 dE_3 d(\cos\theta)d\mu$, and our overall integral reduces to

$$s\dot{Y} = \frac{2(2\pi)^2}{8(2\pi)^8}\int \prod_{i=1}^{3}(\mathrm{p}_i dE_i)\frac{d^3 p_4}{2E_4}d(\cos\alpha)d\beta d(\cos\theta)$$
$$\times \Lambda(f_1,f_2,f_3,f_4)|\mathcal{M}|^2\delta^4(p_1+p_2-p_3-p_4), \tag{D6}$$

on performing the trivial integrals over Ω_1 and μ. We recall that the three-dimensional integral $d^3\mathrm{p}_4$ comes from $\frac{d^3\mathrm{p}_4}{2E_4} = d^4 p_4\delta(p_4^2-m_4^2)\Theta(p_4^0)$, and we can use the four-dimensional δ function to perform the $d^4 p_4$ integral,

$$s\dot{Y} = \frac{1}{4(2\pi)^6}\int \prod_{i=1}^{3}(\mathrm{p}_i dE_i)d(\cos\alpha)d\beta d(\cos\theta)$$
$$\times \Lambda(f_1,f_2,f_3,f_4)|\mathcal{M}|^2\delta(p_4^2-m_4^2)\Theta(p_4^0), \tag{D7}$$

which fixes $p_4^2 = p_1^2+p_2^2+p_3^2+2(p_1\cdot p_2-p_1\cdot p_3-p_2\cdot p_3)$ from now on. The dot products can be evaluated via our angle definitions ($p_i \cdot p_j \equiv p_{ij}$)

$$p_{12} = E_1 E_2 - \mathrm{p}_1 \mathrm{p}_2\cos\alpha, \tag{D8}$$

$$p_{23} = E_2 E_3 - \mathrm{p}_2 \mathrm{p}_3(\cos\alpha\cos\theta + \sin\alpha\sin\theta\cos\beta),$$

$$p_{13} = E_1 E_3 - \mathrm{p}_1 \mathrm{p}_3\cos\theta, \tag{D9}$$

$$p_{24} = m_2^3 + p_{12} - p_{23}, \tag{D10}$$

$$p_{14} = m_1^2 + p_{12} - p_{13}, \tag{D11}$$

$$p_{34} = -m_3^3 + p_{13} + p_{23}. \tag{D12}$$

The argument of the last delta function can be expressed as a function of β,

$$f(\beta) = p_4^2 - m_4^2$$
$$= \omega + 2(p_2 p_3 \cos\alpha \cos\theta + p_2 p_3 \sin\alpha \sin\theta \cos\beta$$
$$- p_1 p_2 \cos\alpha), \tag{D13}$$

where

$$\omega = Q + 2(E_1 E_2 - E_1 E_3 - E_2 E_3 + p_1 p_3 \cos\theta)$$

with $Q = m_1^2 + m_2^2 + m_3^2 - m_4^2$. The β integral can be evaluated using $f'(\beta) = -2p_2 p_3 \sin\alpha \sin\theta \sin\beta$, forcing $\beta \to \beta_0$, where

$$\cos\beta_0 = -\frac{\omega + 2(p_2 p_3 \cos\alpha \cos\theta - p_1 p_2 \cos\alpha)}{2p_2 p_3 \sin\alpha \sin\theta} \tag{D14}$$

is found by solving $f(\beta_0) = 0$. There are actually two β_0 solutions given by $\sin\beta_0 = \pm\sqrt{1 - \cos^2\beta_0}$. Since everything is symmetric in β (all dot products are $\cos\beta$ dependent and the $\partial f/\partial\beta$ factor that appears in the dominator is an absolute value), we can simply use the positive root and multiply by 2. Hence, we get

$$s\dot{Y} = \frac{1}{2(2\pi)^6} \int \prod_{i=1}^{3} (p_i dE_i) d(\cos\alpha) d(\cos\theta) \Lambda(f_1, f_2, f_3, f_4)$$
$$\times \frac{|\mathcal{M}|^2 \Theta(p_4^0) \Theta(4p_2^2 p_3^2 \sin^2\alpha \sin^2\theta \sin^2\beta_0)}{2p_2 p_3 \sin\alpha \sin\theta \sin\beta_0}. \tag{D15}$$

The extra step function arises via an obligation to maintain β_0 in the physical phase space

$$\cos^2\beta_0 \le 0 \leftrightarrow (2p_2 p_3 \sin\alpha \sin\theta \sin\beta_0)^2 \ge 0$$
$$\leftrightarrow \left|\frac{\partial f}{\partial\beta_0}\right|^2 \ge 0. \tag{D16}$$

We can now focus on the angular integrations.

$$s\dot{Y} = \frac{1}{2(2\pi)^6} \int dE_1 dE_2 dE_3 p_1 p_2 p_3 \Lambda(f_1, f_2, f_3, f_4) \times \mathcal{I}$$

$$\mathcal{I} = \int d(\cos\theta) d(\cos\alpha) \frac{|\mathcal{M}|^2 \Theta(p_4^0) \Theta(|\frac{\partial f}{\partial\beta_0}|^2)}{|\frac{\partial f}{\partial\beta_0}|}. \tag{D17}$$

Expanding f' as

$$\left|\frac{\partial f}{\partial\beta_0}\right|^2 = a\cos^2\alpha + b\cos\alpha + c \tag{D18}$$

$$a = -4p_2^2(p_1^2 + p_3^2 - 2p_1 p_3 \cos\theta) \tag{D19}$$

$$b = 4p_2(p_1 - p_3 \cos\theta)\omega \tag{D20}$$

$$c = 4p_2^2 p_3^2 \sin^2\theta - \omega^2, \tag{D21}$$

the step function ensures that the denominator is real,

$$\mathcal{I} = \int d(\cos\theta) \int_{\cos\alpha_-}^{\cos\alpha_+} d(\cos\alpha)$$
$$\times \frac{|\mathcal{M}|^2 \Theta(p_4^0) \Theta(a\cos^2\alpha + b\cos\alpha + c)}{\sqrt{a\cos^2\alpha + b\cos\alpha + c}}. \tag{D22}$$

Since $|\mathcal{M}|^2$ only consists of simple functions of $\cos\alpha$, the $\cos\alpha$ integration can be performed straightforwardly for each process. Since $a \le 0$, the integration bounds are set by the real-valued criterion, between which the quadratic function is positive. Given the roots $\cos\alpha_\pm = (-b \mp \sqrt{b^2 - 4ac})/(2a)$, notice that we always have $-1 \le \cos\alpha_-$ and $\cos\alpha_+ \le 1$. The θ integral can be performed in a similar fashion. Requiring $\cos\alpha_\pm$ to be real implies the condition $b^2 - 4ac \ge 0$, and

$$\cos\theta_\pm = -\frac{Q + 2p_2^2 + 2\gamma \mp 2p_2\sqrt{Q + p_1^2 + p_2^2 + p_3^2 + 2\gamma}}{2p_1 p_3},$$

where we use the shorthand $\gamma = E_1 E_2 - E_1 E_3 - E_2 E_3$ and we have

$$\mathcal{I} = \int_{\max(-1,\cos\theta_-)}^{\min(1,\cos\theta_+)} d(\cos\theta) \int_{\cos\alpha_-}^{\cos\alpha_+} d(\cos\alpha)$$
$$\times \frac{|\mathcal{M}|^2 \Theta(p_4^0) \Theta(a\cos^2\alpha + b\cos\alpha + c)}{\sqrt{a\cos^2\alpha + b\cos\alpha + c}}. \tag{D23}$$

These two integrals can be carried out analytically for each reaction. The final integral to be performed numerically for the emissivity is three dimensional, and a step function guarantees that the phase space is physical and $\cos\alpha_\pm$ is real valued,

$$s\dot{Y} = \frac{1}{2(2\pi)^6} \int dE_1 dE_2 dE_3 p_1 p_2 p_3 \Lambda(f_1, f_2, f_3, f_4)$$
$$\times \mathcal{I} \times \Theta(Q + p_1^2 + p_2^2 + p_3^2 + 2\gamma), \tag{D24}$$

$$\mathcal{I} = \int_{\max(-1,\cos\theta_-)}^{\min(1,\cos\theta_+)} d(\cos\theta)$$
$$\times \int_{\cos\alpha_-}^{\cos\alpha_+} d(\cos\alpha) \frac{|\mathcal{M}|^2}{\sqrt{a\cos^2\alpha + b\cos\alpha + c}}. \tag{D25}$$

[1] R. Essig *et al.*, arXiv:1311.0029.
[2] S. Alekhin *et al.*, Rep. Prog. Phys. **79**, 124201 (2016).
[3] J. Alexander *et al.*, arXiv:1608.08632.
[4] J. A. Adams, S. Sarkar, and D. W. Sciama, Mon. Not. R. Astron. Soc. **301**, 210 (1998).
[5] O. Ruchayskiy and A. Ivashko, J. Cosmol. Astropart. Phys. 10 (2012) 014.
[6] A. C. Vincent, E. F. Martinez, P. Hernndez, M. Lattanzi, and O. Mena, J. Cosmol. Astropart. Phys. 04 (2015) 006.
[7] A. Fradette, M. Pospelov, J. Pradler, and A. Ritz, Phys. Rev. D **90**, 035022 (2014).
[8] J. Berger, K. Jedamzik, and D. G. E. Walker, J. Cosmol. Astropart. Phys. 11 (2016) 032.
[9] F. Piazza and M. Pospelov, Phys. Rev. D **82**, 043533 (2010).
[10] P. W. Graham, D. E. Kaplan, and S. Rajendran, Phys. Rev. Lett. **115**, 221801 (2015).
[11] M. Pospelov and J. Pradler, Phys. Rev. D **82**, 103514 (2010).
[12] J. McDonald, Phys. Rev. Lett. **88**, 091304 (2002).
[13] S. Dodelson and L. M. Widrow, Phys. Rev. Lett. **72**, 17 (1994).
[14] J. Redondo and M. Postma, J. Cosmol. Astropart. Phys. 02 (2009) 005.
[15] T. Flacke, C. Frugiuele, E. Fuchs, R. S. Gupta, and G. Perez, J. High Energy Phys. 06 (2017) 050.
[16] S. Heeba, F. Kahlhoefer, and P. Stcker, J. Cosmol. Astropart. Phys. 11 (2018) 048.
[17] S. Banerjee, D. Barducci, G. Blanger, B. Fuks, A. Goudelis, and B. Zaldivar, J. High Energy Phys. 07 (2017) 080.
[18] D. J. Kapner, T. S. Cook, E. G. Adelberger, J. H. Gundlach, B. R. Heckel, C. D. Hoyle, and H. E. Swanson, Phys. Rev. Lett. **98**, 021101 (2007).
[19] R. S. Decca, D. Lopez, E. Fischbach, G. L. Klimchitskaya, D. E. Krause, and V. M. Mostepanenko, Eur. Phys. J. C **51**, 963 (2007).
[20] A. A. Geraci, S. J. Smullin, D. M. Weld, J. Chiaverini, and A. Kapitulnik, Phys. Rev. D **78**, 022002 (2008).
[21] A. O. Sushkov, W. J. Kim, D. A. R. Dalvit, and S. K. Lamoreaux, Phys. Rev. Lett. **107**, 171101 (2011).
[22] E. Hardy and R. Lasenby, J. High Energy Phys. 02 (2017) 033.
[23] G. Krnjaic, Phys. Rev. D **94**, 073009 (2016).
[24] E. D. Carlson, M. E. Machacek, and L. J. Hall, Astrophys. J. **398**, 43 (1992).
[25] Y. Hochberg, E. Kuflik, T. Volansky, and J. G. Wacker, Phys. Rev. Lett. **113**, 171301 (2014).
[26] J. D. Clarke, R. Foot, and R. R. Volkas, J. High Energy Phys. 02 (2014) 123.
[27] A. Fradette and M. Pospelov, Phys. Rev. D **96**, 075033 (2017).
[28] B. Hyams *et al.*, Nucl. Phys. **B64**, 134 (1973).
[29] F. Bezrukov and D. Gorbunov, J. High Energy Phys. 05 (2010) 010.
[30] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, Front. Phys. **80**, 1 (2000).
[31] D. McKeen, Phys. Rev. D **79**, 015007 (2009).
[32] A. Monin, A. Boyarsky, and O. Ruchayskiy, Phys. Rev. D **99**, 015019 (2019).
[33] M. W. Winkler, Phys. Rev. D **99**, 015018 (2019).
[34] F. Bezrukov, D. Gorbunov, and I. Timiryasov, arXiv: 1812.08088.
[35] A. Djouadi, Phys. Rep. **457**, 1 (2008).
[36] A. Pich, Rep. Prog. Phys. **58**, 563 (1995).
[37] H. Leutwyler and M. A. Shifman, Phys. Lett. B **221**, 384 (1989).
[38] M. E. Carrington, Phys. Rev. D **45**, 2933 (1992).
[39] K. Enqvist, S. Nurmi, T. Tenkanen, and K. Tuominen, J. Cosmol. Astropart. Phys. 08 (**2014**) 035.
[40] M. D'Onofrio, K. Rummukainen, and A. Tranberg, Phys. Rev. Lett. **113**, 141602 (2014).
[41] M. D'Onofrio and K. Rummukainen, Phys. Rev. D **93**, 025003 (2016).
[42] J. Edsjo and P. Gondolo, Phys. Rev. D **56**, 1879 (1997).
[43] J. M. Cornwall, D. N. Levin, and G. Tiktopoulos, Phys. Rev. D **10**, 1145 (1974); **11**, 972(E) (1975).
[44] B. W. Lee, C. Quigg, and H. B. Thacker, Phys. Rev. D **16**, 1519 (1977).
[45] J. I. Kapusta and C. Gale, *Finite-temperature field theory: Principles and applications*, 2nd ed., (Cambridge University Press, Cambridge, England, 2009).
[46] D. J. Gross, R. D. Pisarski, and L. G. Yaffe, Rev. Mod. Phys. **53**, 43 (1981).
[47] J. R. Espinosa, M. Quiros, and F. Zwirner, Phys. Lett. B **314**, 206 (1993).
[48] O. K. Kalashnikov, Phys. Lett. B **279**, 367 (1992).
[49] R. Mertig, M. Bohm, and A. Denner, Comput. Phys. Commun. **64**, 345 (1991).
[50] V. Shtabovenko, R. Mertig, and F. Orellana, Comput. Phys. Commun. **207**, 432 (2016).
[51] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. C **40**, 100001 (2016).
[52] E. N. Argyres, W. Beenakker, G. Jan van Oldenborgh, A. Denner, S. Dittmaier, J. Hoogland, R. Kleiss, C. G. Papadopoulos, and G. Passarino, Phys. Lett. B **358**, 339 (1995).
[53] N. Kauer and D. Zeppenfeld, Phys. Rev. D **65**, 014021 (2001).
[54] C. Schwinn, arXiv:hep-ph/0307057.
[55] V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum electrodynamics: Volume 4 (course of theoretical physics)*, 2nd ed., (Butterworth-Heinemann, 1982).
[56] E. Braaten and T. C. Yuan, Phys. Rev. Lett. **66**, 2183 (1991).
[57] M. Kawasaki, K. Kohri, T. Moroi, and Y. Takaesu, Phys. Rev. D **97**, 023502 (2018).
[58] R. Essig, E. Kuflik, S. D. McDermott, T. Volansky, and K. M. Zurek, J. High Energy Phys. 11 (2013) 193.
[59] D. E. Gruber, J. L. Matteson, L. E. Peterson, and G. V. Jung, Astrophys. J. **520**, 124 (1999).
[60] L. Bouchet, E. Jourdain, J. P. Roques, A. Strong, R. Diehl, F. Lebrun, and R. Terrier, Astrophys. J. **679**, 1315 (2008).
[61] G. Hinshaw *et al.* (WMAP Collaboration), Astrophys. J. Suppl. Ser. **208**, 19 (2013).
[62] P. A. R. Ade *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A13 (2016).
[63] X. L. Chen and M. Kamionkowski, Phys. Rev. D **70**, 043502 (2004).
[64] E. Pierpaoli, Phys. Rev. Lett. **92**, 031301 (2004).
[65] N. Padmanabhan and D. P. Finkbeiner, Phys. Rev. D **72**, 023508 (2005).

[66] L. Zhang, X. Chen, M. Kamionkowski, Z. g. Si, and Z. Zheng, Phys. Rev. D **76,** 061301 (2007).

[67] D. P. Finkbeiner, S. Galli, T. Lin, and T. R. Slatyer, Phys. Rev. D **85,** 043522 (2012).

[68] T. R. Slatyer and C. L. Wu, Phys. Rev. D **95,** 023010 (2017).

[69] T. R. Slatyer, Phys. Rev. D **87,** 123513 (2013).

[70] T. R. Slatyer, Phys. Rev. D **93,** 023521 (2016).

[71] V. Poulin, J. Lesgourgues, and P. D. Serpico, J. Cosmol. Astropart. Phys. 03 (2017) 043.

[72] J. M. Cline and P. Scott, J. Cosmol. Astropart. Phys. 03 (2013) 044; 05 (2013) E01.

[73] A. Kogut et al., J. Cosmol. Astropart. Phys. 07 (2011) 025.

[74] J. Chluba and R. A. Sunyaev, Mon. Not. R. Astron. Soc. **419,** 1294 (2012).

[75] J. Chluba, Mon. Not. R. Astron. Soc. **436,** 2232 (2013).

[76] J. Chluba and D. Jeong, Mon. Not. R. Astron. Soc. **438,** 2065 (2014).

[77] J. Chluba, Mon. Not. R. Astron. Soc. **460,** 227 (2016).

[78] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, Astrophys. J. **473,** 576 (1996).

[79] G. Mangano, G. Miele, S. Pastor, T. Pinto, O. Pisanti, and P. D. Serpico, Nucl. Phys. **B729,** 221 (2005).

[80] R. H. Cyburt, B. D. Fields, K. A. Olive, and T. H. Yeh, Rev. Mod. Phys. **88,** 015004 (2016).

[81] A. Coc and E. Vangioni, Int. J. Mod. Phys. E **26,** 1741002 (2017).

[82] C. Pitrou, A. Coc, J. P. Uzan, and E. Vangioni, Phys. Rep. **754,** 1 (2018).

[83] K. Jedamzik and M. Pospelov, New J. Phys. **11,** 105028 (2009).

[84] M. Pospelov and J. Pradler, Annu. Rev. Nucl. Part. Sci. **60,** 539 (2010).

[85] R. Cooke, M. Pettini, R. A. Jorgenson, M. T. Murphy, and C. C. Steidel, Astrophys. J. **781,** 31 (2014).

[86] R. J. Cooke, M. Pettini, K. M. Nollett, and R. Jorgenson, Astrophys. J. **830,** 148 (2016).

[87] K. Jedamzik, Phys. Rev. D **74,** 103509 (2006).

[88] B. Batell, M. Pospelov, and A. Ritz, Phys. Rev. D **83,** 054005 (2011).

[89] S. Hannestad and J. Madsen, Phys. Rev. D **52,** 1764 (1995).