PHYSICAL REVIEW D **96,** 075033 (2017)

BBN for the LHC: Constraints on lifetimes of the Higgs portal scalars

Anthony Fradette and Maxim Pospelov
Department of Physics and Astronomy, University of Victoria, Victoria, British Columbia V8P 5C2, Canada
Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2J 2W9, Canada
(Received 23 June 2017; published 23 October 2017)

LHC experiments can provide a remarkable sensitivity to exotic metastable massive particles, decaying with significant displacement from the interaction point. The best sensitivity is achieved with models where the production and decay occur due to different coupling constants, and the lifetime of exotic particles determines the probability of decay within a detector. The lifetimes of such particles can be independently limited from standard cosmology, in particular, the big bang nucleosynthesis (BBN). In this paper, we analyze the constraints on the simplest scalar model coupled through the Higgs portal, where the production occurs via $h \to SS$, and the decay is induced by the small mixing angle of the Higgs field h and scalar S. We find that throughout most of the parameter space, $2m_\mu < m_S < m_h/2$, the lifetime of an exotic particle has to be less than 0.1 s, while below $2m_\mu$ it could grow to about a second. The strong constraints on lifetimes are induced by the nucleonic and mesonic decays of scalars that tend to raise the n/p ratio. Strong constraints on lifetimes of the minimal singlet extensions of the Higgs potential are welcome news for the MATHUSLA proposal that seeks to detect displaced decays of exotic particles produced in the LHC collisions. We also point out how more complicated exotic sectors could evade the BBN lifetime constraints.

DOI: 10.1103/PhysRevD.96.075033

I. INTRODUCTION

The absence of readily discoverable new physics (NP) at the LHC has presented the physics community with a formidable puzzle. While the arguments for NP "not too far" from the weak scale still loom large, there is a distinct desire to explore wider (and wilder) theoretical options away from a simply realized weak-scale supersymmetry, or extra space dimensions. One possible strategy to look for new physics is to abandon theoretical preconceptions and to start looking for nonstandard signatures that the NP could present.

Large classes of models offer promising avenues for a nonstandard signal in the production of new exotic particles (possibly of electroweak-scale mass) with subsequent decay away from the interaction point (see, e.g., [1–4]). While both ATLAS and CMS have performed corresponding studies in a variety of contexts and for different ranges of displacement [5–7], it has recently been pointed out that a dedicated and relatively inexpensive detector [8] could extend the physics reach into cases where the decay lengths are on the order of $O(100\ \mathrm{m})$ and beyond.

When both the production and the decay of an exotic state S occur through one and the same coupling constant, the chances of detecting such NP at the LHC experiments are not great. Indeed, a large displacement implies a very small value for the coupling, which in turn leads to very inefficient production rates. Therefore, an ideal case for the collider studies would be when the production and decay occur through different coupling constants, and $\lambda_{\mathrm{production}} \gg \lambda_{\mathrm{decay}}$. For the pair-produced exotics, such a hierarchy can be made "natural" as the $\lambda_{\mathrm{decay}} \to 0$ limit could lead to an enhanced symmetry.

If the main signal to search for is an *appearance* of abnormal energy deposition or exotic vertex some distance from the interaction point, it is then very important to know how small λ_{decay} is allowed to be. In more practical terms, one would like to know if there is an external to the LHC physics constraint on the lifetimes τ_S of such exotic particles. An obvious source for such a constraint can be early cosmology. The big bang nucleosynthesis (BBN) and its overall agreement with observations [9] (apart from the unclear status of ^{7}Li) can provide a limit on the lifetimes of such particles. To derive such limits, one would have to make a fairly natural assumption that the Universe was indeed as hot as $T \sim m_S \sim$ electroweak scale at some point in its history. Subsequent thermal evolution to the BBN temperatures involves self-depletion via $SS \to$ SM due to $\lambda_{\mathrm{production}}$, in an expected weakly interacting massive particle (WIMP)-type annihilation process, and late-time decay of $S \to$ SM where depending on lifetimes and the properties of the decay products the BBN outcome may get affected. These mechanisms are well understood in the BBN literature (see, e.g., [10,11] for reviews). We will require that the late decay of S provides a small and acceptable perturbation to the standard BBN (SBBN) outcome, which in turn will limit τ_S.

In this paper, we analyze a fairly minimal model, where a new singlet scalar has predominantly a quadratic coupling to the Higgs boson that regulates both its production at

colliders and the intermediate cosmological abundance at $T_{\rm BBN} \ll T \ll m_S$. Given that the model is very predictive, it allows one to place robust bounds on lifetimes of such particles with a minimum amount of model dependence. We find that for most of the analyzed parameter space with $m_S < m_h/2$, the intermediate abundance of such particles is large enough to affect the neutron-proton freeze out ratios at relevant temperatures. This allows us to set fairly robust bounds on lifetimes of such particles, which come out to be remarkably strong, and shorter than 0.1 s. In what follows we describe the model and the cosmological history of S (Sec. II), derive the impact on the BBN (Sec. III), present our results (Sec. IV), and provide related discussion (Sec. V).

II. THE MINIMAL HIGGS PORTAL MODEL

We consider the simplest extension of the Standard Model (SM) by a singlet scalar field S. A new singlet scalar S can have two interaction terms with the SM at the renormalizable level, in addition to trilinear and quartic self-interactions. In this scenario, the Lagrangian of the singlet sector (including the SM) generically takes the form

$$\mathcal{L}_{H/S} = \mu^2 H^\dagger H - \lambda_H (H^\dagger H)^2 - V(S) - A S H^\dagger H - \lambda_S S^2 H^\dagger H + \text{kin terms}. \quad (1)$$

The Higgs expectation value $v = 246$ GeV is assumed to correspond to a global minimum. The self-interaction potential $V(S) = \lambda_4 S^4 + \lambda_3 S^3 + \frac{m_{S0}^2}{2} S^2$ can be redefined in such a way that the linear term is absent. It is important that the A, $\lambda_3 \to 0$, and $\langle S \rangle = 0$ limit would correspond to the case of stable S particles. To simplify the discussion without sacrificing much generality, we take $\lambda_{3,4} \to 0$ and assume $Av \ll m_{S0}^2$, $\lambda_S v^2$.

The physical mass of S receives a contribution from the electroweak symmetry breaking, $m_S = \sqrt{m_{S0}^2 + \lambda_S v^2}$. At linear order in A, the mixing angle θ between physical excitations S and h is

$$\theta = \frac{Av}{m_h^2 - m_S^2}\left(1 - \frac{\lambda_S v^2}{m_S^2}\right). \quad (2)$$

The λ_S term arises because the S field develops a small A-controlled vacuum expectation value. The mixing parameter θ leads, via the A coupling constant, to the decay of S particles, which can readily be derived from

$$\mathcal{L}_{\rm decay} = S \times \theta \sum_{\rm SM} O_h, \quad (3)$$

where O_h is the set of the standard Higgs interaction terms, with the Higgs field removed: e.g., $O_h = (m_f/v)\bar{f}f$ for an elementary SM fermion f.

This Yukawa-type coupling to the SM has been tested in rare meson decays [12–16] and in proton fixed-target experiments [17]. The model is mostly ruled out for large mixing angles $\theta \gtrsim 10^{-4}$–10^{-2} over the $m_S \sim$ MeV–5 GeV mass range. The proposed experiment SHiP could potentially improve current sensitivity down to $\theta \sim 10^{-6}$ for $m_S \sim$ few GeV [17], and there are additional islands of sensitivity at even lower mixing angles from BBN and the cosmic microwave background [18].

In the limit of $\theta \to 0$, S is stable and could be the dark matter [19–21]. Various limits arise from searches in direct and indirect detection if the particle is stable (see Refs. [22,23] for recent reviews), but λ_S is generically bounded from the constraints on invisible Higgs decay, independently of the direct detection limits. The Standard Model Higgs has a well-predicted decay rate into SM particles of $\Gamma_{\rm SM} = 4.07$ MeV. So far, the properties of 125 GeV resonance are remarkably consistent with the SM Higgs, and therefore there is little doubt that its width is close to $\Gamma_{\rm SM}$. The invisible branching ratio of Higgs decay to the SS final state is

$$\Gamma_{h\to SS} = \frac{\lambda_S^2 v^2}{8\pi m_h}\sqrt{1 - \frac{4m_S^2}{m_h^2}}, \quad (4)$$

$$\mathrm{Br}(h \to SS) = \frac{\Gamma_S}{\Gamma_S + \Gamma_{\rm SM}} \simeq 10^{-2}\left(\frac{\lambda_S}{0.0015}\right)^2, \quad (5)$$

where in the last line we assumed $\mathrm{Br}(h \to SS) \ll 1$ and $m_S \ll m_h$. The experimental upper bound on the invisible branching ratio of a SM Higgs is 0.19 (at 2σ) [24], which translates into an upper bound on λ_S,

$$\lambda_S \lesssim \frac{0.007}{\left(1 - \frac{4m_S^2}{m_h^2}\right)^{1/4}}. \quad (6)$$

If S is to be stable, such small couplings would lead to an excessive abundance of S, which invalidates the Z_2 symmetric case and forces us to include the decay term. From now on, we will consider $\theta \neq 0$, or in other words the case of unstable S particles. Since our analysis is motivated by the LHC physics, we will use $\mathrm{Br}(h \to SS)$ as an input parameter and substitute λ_S everywhere employing (4) and (5).

A. Decay products

Since S interacts with the SM in the same fashion as the Higgs with an additional θ mixing factor (3), its decay properties are similar to those of a light Higgs boson. For the derivations of the actual limits on the lifetime of S, we need to know its mesonic and nucleonic decay branching ratios.

The decay channels of a light Higgs have been considered in the early years of the Weinberg-Salam electroweak model [25], with additional refinements as SM particles and hadronic resonances were being discovered and final-state

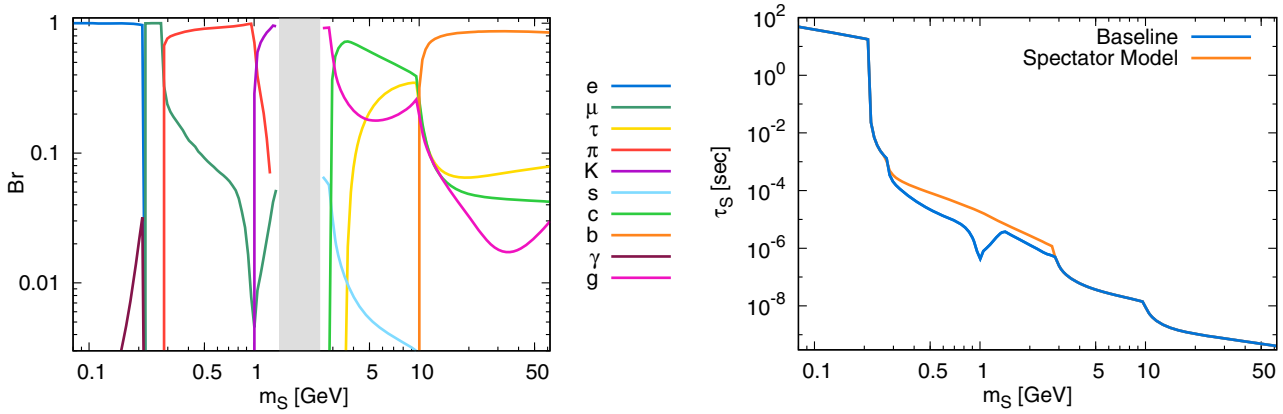

FIG. 1. *Left*: Branching ratios of the scalar S in our baseline decay model. See text for details. *Right*: Scalar S lifetime of our baseline model and the spectator model for the mixing angle $\theta = 10^{-6}$.

interactions were becoming better understood [26–28]. Hadronic decays in the mass range $2m_\pi < m_S \lesssim 4$ GeV are still poorly understood, with models varying by as much as a few orders of magnitude near the di-kaon threshold [16].

The leptonic decay channels are straightforward, with the decay rate given by

$$\Gamma_{S\to l\bar{l}} = \frac{\theta^2}{8\pi}\frac{m_l^2}{v^2} m_S \left(1 - \frac{4m_l^2}{m_S^2}\right)^{3/2}. \tag{7}$$

If the decaying product is a pair of heavy quarks, there are $\mathcal{O}(1)$ corrections coming from the 1-loop QCD vertex correction [22], which yields the following correction factor [29] to the fermionic expressions (7):

$$f_q = 3\left[1 + \frac{4\alpha_s}{3\pi}\left(\frac{9}{4} + \frac{3}{2}\ln\frac{m_q^2}{m_S^2}\right)\right], \tag{8}$$

and the factor of 3 comes from the number of color charges. For better accuracy, we use the higher order perturbative results from the HDECAY code [30] for $m_S > 2.5$ GeV.

Metastable mesons, such as $\pi^\pm$ and $K^\pm$, $\bar{K}^0$, K^0 are "important" decay products, as they can participate in the charge-exchange reactions with nucleons and shift the $n-p$ balance, hence affecting the whole nucleosynthetic chain. In the mass range where the perturbative QCD calculations are no longer valid, we base our baseline calculations on Ref. [31]. The scalar-pion interaction can be extracted from the low-energy expansion of the trace of the QCD energy-momentum tensor (see, e.g., [32,33]) by integrating out the three heavy quarks and using chiral perturbation theory on the remainder, yielding the effective Lagrangian [31]

$$\mathcal{L}_{S\pi\pi} = \frac{4}{9}\frac{\theta}{v} S\left(\frac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0 + \partial_\mu\pi^+\partial^\mu\pi^-\right) - \frac{5}{3}\frac{\theta m_\pi^2}{v} S\left(\frac{1}{2}\pi^0\pi^0 + \pi^+\pi^-\right), \tag{9}$$

where we have inserted the SM numerical values for the number of heavy quarks and the first coefficient of the QCD beta function. This leads to decay width to charged pions,

$$\Gamma_{S\to\pi^+\pi^-} = 2\Gamma_{S\to\pi^0\pi^0} = \frac{\theta^2}{16\pi}\frac{m_S^3}{v^2}\left(\frac{2}{9} + \frac{11}{9}\frac{m_\pi^2}{m_S^2}\right)^2\sqrt{1 - \frac{4m_\pi^2}{m_S^2}}. \tag{10}$$

This result, however, is not applicable far above the pion threshold, as final-state resonances would drastically affect this prediction. Instead, we use the pion and kaon decay width described in Ref. [34], where the authors matched the next-to-leading order corrections of the low-energy theorems to the dispersion results from the $\pi\pi$ phase-shift analysis above 600 MeV from the CERN-Munich group [35]. The photon decay channel is added with the prescription detailed in Ref. [36]. Finally, there is a gap for 1.4 GeV $< m_S <$ 2.5 GeV where no analytical treatment is entirely trustworthy, as this includes new resonances strongly coupled to $\eta\eta$ and other potential hadronic channels. We simply follow Ref. [31] and interpolate between the two regimes, under the assumption that there is no order of magnitude deviation in this mass range. The branching ratios and the lifetime for $\theta = 10^{-6}$ are displayed in Fig. 1.

As an alternative decay spectrum model, we also display the perturbative spectator approach [17,37,38], where the relative decay widths above the kaon threshold are given by

$$\Gamma_{\mu^+\mu^-} : \Gamma_{KK} : \Gamma_{\eta\eta} = m_\mu^2\beta_\mu^3 : 3\frac{9}{13}m_s^2\beta_K^3 : 3\frac{4}{13}m_s^2\beta_\eta^3, \tag{11}$$

with $\beta_i = \sqrt{1 - 4m_i^2/m_S^2}\Theta(m_S - 2m_i)$, Θ being the step-function, and we adopt the running of s quark mass following Ref. [36]. The pion contribution is kept as in Eq. (10), and then we use the HDECAY output at the c-quark threshold and above to match our baseline model.

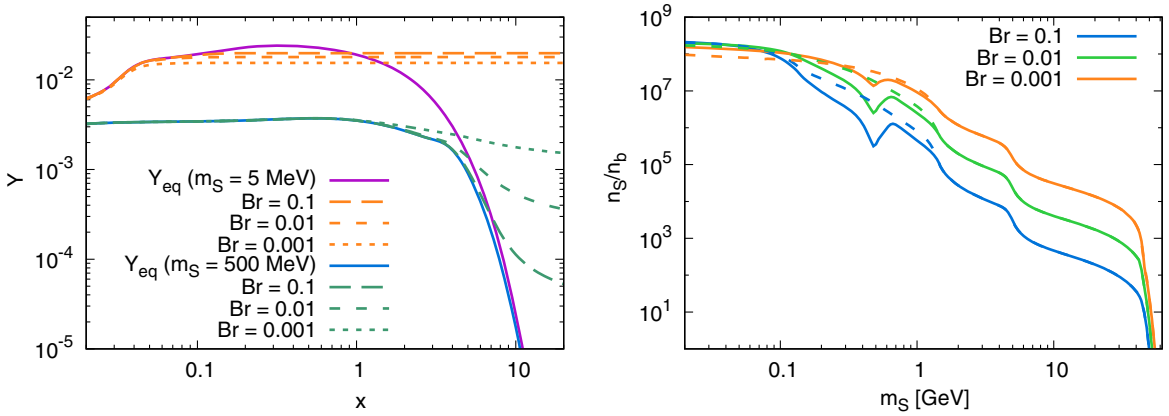

FIG. 2. *Left*: Temperature evolution ($x = m/T$) of the Y_S intermediate abundance for $m_S = 5$ MeV and 500 MeV for the three benchmark Higgs branching ratios. *Right*: Metastable abundance of S prior to its decay normalized over the baryon density. Values shown for $\mathrm{Br}(h \to SS) = 10^{-1}$, 10^{-2}, and 10^{-3}. The dashed lines correspond to the perturbative spectator model.

For m_S of several GeV and heavier, decays with final state nucleon-antinucleon pairs are possible. Even though the branching to such states are generally lower than 10%, the effect on BBN can be quite significant, and therefore these are by far the most important channels for $\tau_S \gtrsim 1$ s. On top of direct and for the most part subdominant contributions from $S \to \bar{n}n, \ldots$, we need to take into account the (anti)nucleon states that emerge from the hadronization of the quark decay products and heavy B-meson fragmentations.

B. Cosmological metastable abundance

After the temperature drops below m_S, the interaction of SS pairs with the SM shifts toward the annihilation, resulting in an intermediate (metastable) population of S bosons. In the mass range that we consider, the S annihilation is dominated by the s-channel reactions $SS \to h^* \to XX$, where on the receiving end are the pairs of the SM states XX created by a Higgs-mediation process. The annihilation cross section σv generically takes the form

$$\sigma v(s) = \frac{8\lambda_S^2 v^2}{(s - m_h^2)^2 + m_h^2 \Gamma^2_{\mathrm{SM}+S}} \frac{\Gamma_{\mathrm{SM}}^{m_h \to \sqrt{s}}}{\sqrt{s}},$$
$$\langle \sigma v \rangle = \frac{\int_{4m_S^2}^{\infty} ds \sigma v(s) s \sqrt{s - 4m_S^2} K_1(\frac{\sqrt{s}}{T})}{16 T m_S^4 K_2^2(\frac{m_S}{T})}. \quad (12)$$

This formula recast the rate in terms of a Higgs width $\Gamma_{\mathrm{SM}}^{m_h \to \sqrt{s}}$ with a fictitious mass of $\sqrt{s}$. This form encompasses both perturbative and nonperturbative channels in the h^* decay rate (with the substitution $m_h^* \to \sqrt{s}$), which we have described above. In the standard WIMP freeze-out paradigm, a dark matter (DM) particle freezes out at $T_{\mathrm{f.o.}} \sim m_{\mathrm{DM}}/20$, $\langle \sigma v \rangle$ is simply the nonrelativistic limit $\sigma v(\sqrt{s} = 2m_{\mathrm{DM}})$, and the relic density can be conveniently approximated as $\Omega_{\mathrm{DM}} h^2 \sim 0.11 \times 1\ \mathrm{pb}/\langle \sigma v \rangle$. This result emerges as a solution to the Boltzmann equation[1] [39]

$$\frac{dY}{dx} = \frac{s\langle \sigma v \rangle}{Hx} \left[1 + \frac{1}{3} \frac{d(\ln h_{\mathrm{eff}})}{d(\ln T)} \right] (Y_{\mathrm{eq}}^2 - Y^2), \quad (13)$$

when the freeze-out occurs in the exponentially falling region of the equilibrium density $Y_{\mathrm{eq}}(T)$. For a much smaller annihilation cross section, $\langle \sigma v \rangle \ll 1$ pb, Y departs from the equilibrium value earlier, possibly near the relativistic plateau $Y_{\mathrm{eq}} = n_{\mathrm{eq}}/s \to 45\zeta(3)/2\pi^4 h_{\mathrm{eff}}(T)$ for $x \ll 1$. Since the nonrelativistic annihilation cross section in the minimal Higgs portal model ranges from 10^{-3} to 10^{-14} pb for $m_S \sim 1$ MeV–60 GeV and $\mathrm{Br}(h \to SS) \sim 0.1 - 0.001$, we numerically integrate Eq. (13) to determine the metastable S abundance. The results are shown in Fig. 2, normalized to the baryon number density for a more intuitive interpretation of its impact on BBN in the following section.

For $m_S \simeq m_h/2$, the σv cross section evaluated at $s = 4m_S^2$ is a poor approximation, as it fails to capture the strong energy dependence of the cross section near the resonance at $\sqrt{s} = m_h/2$ [40]. The sharp drop in the abundance above $m_S \sim 45$ GeV is due to the resonant contribution to the thermally averaged cross section, leading to a delayed freeze-out and drastic decrease in metastable S abundance. Our numerical results agree with the semianalytic treatment of Ref. [22]. For very light m_S, one can see that the freeze-out abundances are large, and the relative spread between different input values of $\mathrm{Br}(h \to SS)$ gets smaller, as the annihilation cross section becomes very small and the

[1]We use the standard variable definitions, where $Y = n_S/s$ is the S abundance normalized on the entropy density s, $x = m/T$ is the dimensionless inverse temperature, H is the Hubble rate, h_{eff} is the number of entropic relativistic degrees of freedom, and Y_{eq} is the normalized thermal equilibrium S number density.

freeze-out happens in the semirelativistic regime $x_{\mathrm{f.o.}} \sim \mathcal{O}(1)$ with an asymptote at the Y_{eq} relativistic plateau for small m_S. The only difference at the lightest masses is from $Y_{\mathrm{eq}}^{\mathrm{rel}} \propto 1/h_{\mathrm{eff}}(T)$. Since h_{eff} is a monotonic function of temperature, weaker annihilation cross sections freeze out earlier, at a higher temperature, thus yielding smaller abundances (as seen in the $m_S = 5$ MeV curves in Fig. 2). This is in contrast with the standard freeze-out in the nonrelativistic regime, with final abundances inversely proportional to the cross section. We note in passing that the strong-interaction-related uncertainty "propagates" outside the $m_S \sim 2m_\pi - 2m_c$ window. For example, because of the relativistic freeze-out, for m_S smaller than $2m_\pi$ the hadronic channels may turn out to be important.

Corrections to the metastable abundance from the Z_2-breaking mixing angle are negligible. Parametrically, we will be interested in values of θ that have $\Gamma_S^{\mathrm{dec}} \sim (0.1\text{–}1\ \mathrm{s})^{-1}$. Modifications to the relic abundance would arise from decay or inverse decay processes with approximately

$$\left.\frac{\delta n_S}{n_S}\right|_\theta \sim \frac{\Gamma_S^{\mathrm{dec,inv-dec}}}{H(t_{fo})} \ll 1, \tag{14}$$

which justifies neglecting θ at the time of metastable S freeze-out.

III. BIG BANG NUCLEOSYNTHESIS

The formation of light nuclei is one of the earliest probes of NP in cosmology along with far less certain constraints imposed by the inflationary framework. BBN is well understood within SM physics, and its outcome agrees with observational data for ^{4}He and D. ^{7}Li has an outstanding factor of ∼2–3 discrepancy between theory and observations [9], with the caveat that the observed abundances may have been affected by stellar evolution. Nevertheless, the overall success over a wide range of abundances can be used to constrain various types of NP [11].

The initial BBN stage is the neutron-proton ratio n/p freeze-out. Maintained in equilibrium by electroweak interactions at high temperatures, the neutron abundance follows $n/p \sim e^{-Q/T}$, where $Q = m_n - m_p - m_e \simeq 1.293$ MeV, until the epoch when the weak processes decouple around temperatures of 0.7 MeV. The outcome, $n/p \simeq 1/6$, is quasistable, decreasing to $n/p \simeq 1/7$ at the end of the "deuterium bottleneck." The latter terminology is used to indicate a much delayed onset of nuclear reactions controlled by a relatively shallow $n - p$ binding energy. Once the Universe runs out of photons that can efficiently dissociate deuterium, the bulk of the nucleosynthetic reactions occurs at $t_{\mathrm{deut}} \sim 200$ s. ^{4}He has a large binding energy per nucleon, and the reactions leading to it are less Coulomb-suppressed than for heavier elements. Consequently, most neutrons end up in the final ^{4}He abundance (expressed in mass fraction from the total baryon mass) $Y_p \simeq 2(n/p)/(1 + n/p) \simeq 0.25$.

Traces of neutrons and incomplete nuclear burning of $A = 2$, 3 nuclei light nuclei result in the leftover abundances of ^{3}He and D. Beyond the ^{4}He atomic number, the deepest bound nucleus is ^{12}C, but its formation is completely suppressed since it would need to be produced by a triple ^{4}He collision. The $2 \to 2$ reactions $p + {}^4\mathrm{He}$ and $^4\mathrm{He} + {}^4\mathrm{He}$ are also ineffective at producing heavier nuclei as the $A = 5$ and $A = 8$ elements are all unstable. The only remaining possibilities are $^4\mathrm{He} + {}^3\mathrm{He} \to {}^7\mathrm{Be} + \gamma$ followed by an electron capture to yield $^7\mathrm{Li/H} \sim \mathcal{O}(10^{-10})$ and ^{6}Li formed at the $^6\mathrm{Li/H} \sim \mathcal{O}(10^{-14})$ via ^{4}He-D fusion. For the problem at hand—the determination of the upper limit on the S lifetime—few of these details matter. This is because of relatively large metastable abundances affecting the earliest stages of nucleosynthesis, primarily via the n/p ratio.

A. Neutron enrichment

Ample abundances of S particles ($n_S \sim 10^2\text{–}10^9 \times n_b$) flood the Universe with final state mesons and nucleons that in turn could spoil the final light nuclei abundances. For example, at temperatures $T \sim 0.5$ MeV, the protons are ∼6 times more abundant than neutrons, but this ratio can easily be changed due to meson-induced charge exchange reactions. At these temperatures, the probability of $p \to n$ conversion from charged pions is

$$P_{n\to p} \simeq n_p \langle \sigma v \rangle_{pn} c\tau_{\pi^+} \simeq 2 \times \frac{10^{21}}{\mathrm{cm}^3} \times 1.5\ \mathrm{mb} \times 2.6 \times 10^{-8}\ \mathrm{s} \times c \simeq 2.5 \times 10^{-3}. \tag{15}$$

It is then clear that injection of $O(10^3)$ mesons per nucleon at these temperatures can drastically increase the n/p freeze-out abundance. Similarly, direct baryonic injection of $n\bar{n}$ and $p\bar{p}$ will have a comparable effect on the n/p ratio. On the other hand, if S decays happen before the n/p freeze-out, the additional $p \to n$ conversions would not be as efficient, being washed out by the ongoing weak interaction conversions.

The limit of the exclusion region in the Y_S/τ_S parameter space ($Y_S \equiv n_S/n_b$ from now on) is determined by solving the Boltzmann equation with the injection of charge exchange inducing particles. Given that the abundances of S particles are large, the main constraints can be derived from the n/p freeze-out ratio. To that effect, we would not need a complete BBN framework, but only a subset of the whole code that deals with $n \leftrightarrow p$ conversions. We follow the semianalytic treatment by Mukhanov [41] that approximates $n \leftrightarrow p$ weak conversion rates by a few integrals over thermal distributions and assumes a "steplike" disappearance of charged leptons below $T = m_e$,

$$\Gamma_{n\nu_e\to pe^-} = \frac{1+3g_a^2}{2\pi^3} G_F^2 Q^5 J(1;\infty),$$
$$\Gamma_{pe^-\to n\nu_e} = e^{-Q/T}\Gamma_{n\nu_e\to pe^-}, \tag{16}$$

$$\Gamma_{ne^+\to p\bar{\nu}_e} = \frac{1+3g_a^2}{2\pi^3} G_F^2 Q^5 J\left(-\infty;-\frac{m_e}{Q}\right),$$
$$\Gamma_{p\bar{\nu}_e\to ne^+} = e^{-Q/T}\Gamma_{ne^+\to p\bar{\nu}_e}, \tag{17}$$

$$J(a,b) \equiv \int_a^b \sqrt{1-\frac{(m_e/Q)^2}{q^2}} \frac{q^2(q-1)^2 dq}{(1+e^{\frac{Q}{T_\nu}(q-1)})(1+e^{-\frac{Q}{T}q})}, \tag{18}$$

where $g_a \simeq 1.27$ is the standard nucleon axial-vector coupling, $Q = m_n - m_p - m_e \simeq 1.293$ MeV, and G_F is the Fermi constant. The reverse reaction rates are found by detailed balance. We evaluate J numerically and solve for the electron-neutrino temperature T_ν by entropy conservation, assuming a ν_e decoupling temperature of 2 MeV, which reproduces the correct entropy degrees of freedom at lower temperature [42]. It is then straightforward to solve numerically the differential equation for $X_n = n_n/n_b$,

$$\frac{dX_n}{dT} = \frac{\Gamma_{n\nu_e\to pe^-} + \Gamma_{ne^+\to p\bar{\nu}_e}}{TH(T)}(X_n - (1-X_n)e^{-Q/T}) + \frac{\Gamma_n X_n}{TH(T)}, \tag{19}$$

where the last term represents the neutron decay with $\Gamma_n^{-1} = 880$ s. This equation is approximately valid until the rapid switch-on of the nuclear reaction rates at the end of the deuterium bottleneck. Within this approximation, one can determine the final temperature where the equation is valid by starting with $X_n = 1/2$ at early times and solving for the deuterium bottleneck temperature by imposing $Y_p = 2X_n(T_{\rm deut}) = 0.25$. This results in $T_{\rm deut} \simeq 0.068$ MeV or $t_{\rm deut} \simeq 276$ s. We take this approximation as our baseline SBBN model, which is then modified by the inclusion of extra sources and sinks for n, p, and new $n \leftrightarrow p$ reactions. To constrain the parameter space of a species decaying into charged mesons or baryons, we proceed by solving the Boltzmann equation that includes new interactions. We will require that Y_p does not deviate from SBBN by more than 4%,

$$\Delta Y_p \equiv |Y_p - Y_p^{\rm SBBN}| < 0.01, \tag{20}$$

which is a rather generous allowance for the errors, considering the tight observational constraints on primordial helium abundance [9]. Consequently, it will result in conservative limits of τ_S.

1. Meson-mediated mechanism

Only long-lived mesons have an opportunity to interact with the baryon bath and induce proton-neutron conversions. As such, only $\pi^\pm$, $K^\pm$, and K_L have lifetimes in excess of $\tau \sim 10^{-8}$ s and can induce $p \leftrightarrow n$ via strong interactions. For temperatures relevant for the n/p freeze-out, the density of charged leptons is very high, and mesons are efficiently stopped by the primordial plasma. We assume that they are efficiently thermalized, and we take the relevant pion-induced reactions at threshold [43,44] ($c = 1$),

$$\pi^- + p \to n + \gamma: \qquad (\sigma v)^{\pi^-}_{pn(\gamma)} \simeq 0.57 \text{ mb}, \quad Q = 138.3 \text{ MeV}, \tag{21}$$

$$\pi^- + p \to n + \pi^0: \qquad (\sigma v)^{\pi^-}_{pn(\pi^0)} \simeq 0.88 \text{ mb}, \quad Q = 3.3 \text{ MeV}, \tag{22}$$

$$\pi^+ + n \to p + \gamma: \qquad (\sigma v)^{\pi^+}_{np(\gamma)} \simeq 0.44 \text{ mb}, \quad Q = 140.9 \text{ MeV}, \tag{23}$$

$$\pi^+ + n \to p + \pi^0: \qquad (\sigma v)^{\pi^+}_{np(\pi^0)} \simeq 1.26 \text{ mb}, \quad Q = 5.9 \text{ MeV}. \tag{24}$$

The reverse reactions are irrelevant due to the short lifetime of π^0's and the need for nonthermal γ's of ~140 MeV energy. The π^- reactions are to be added to the right-hand side of the Boltzmann Eq. (19) via the additional term

$$\left.\frac{dX_n}{dT}\right|_{\pi-} = \frac{-1}{TH(T)} n^{\rm inj}_{\pi^-}(\langle\sigma v\rangle^{\pi^-}_{pn(\pi^0)} + \langle\sigma v\rangle^{\pi^-}_{pn(\gamma)})(1-X_n), \tag{25}$$

and similarly for the π^+ reactions. The ambient population of injected pions from a S decay with $\mathrm{Br}(S \to \pi^+\pi^-) = \xi_{\pi^\pm}$ is $n^{\rm inj}_{\pi^\pm} \simeq \xi_{\pi^\pm}\Gamma_S\tau_{\pi^\pm}Y_S n_b(T)e^{-t\Gamma_S}$, $t \simeq 2.42\,\mathrm{s}(\mathrm{MeV}/T)^2/\sqrt{g_\star}$, and the thermal cross sections are taken at their threshold value $\langle\sigma v\rangle^{\pi^+}_{np} = (\sigma v)^{\pi^+}_{np}$. Reactions with pairs of charged particles in the initial states, such as $\pi^- p$, will be somewhat enhanced due to the Coulomb attraction, which provides a small but non-negligible correction. We account for it following Ref. [44].

The implementation of the charged kaon reactions is similar to the pion case, but the dominant reactions are rather different. The direct charge exchange between neutral and charged kaons is

$$\bar{K}^0 + n \to K^- + p: \qquad (\sigma v)^{\bar{K}^0}_{pn(K^-)} \simeq 10 \text{ mb}, \quad Q = 5.3 \text{ MeV}, \tag{26}$$

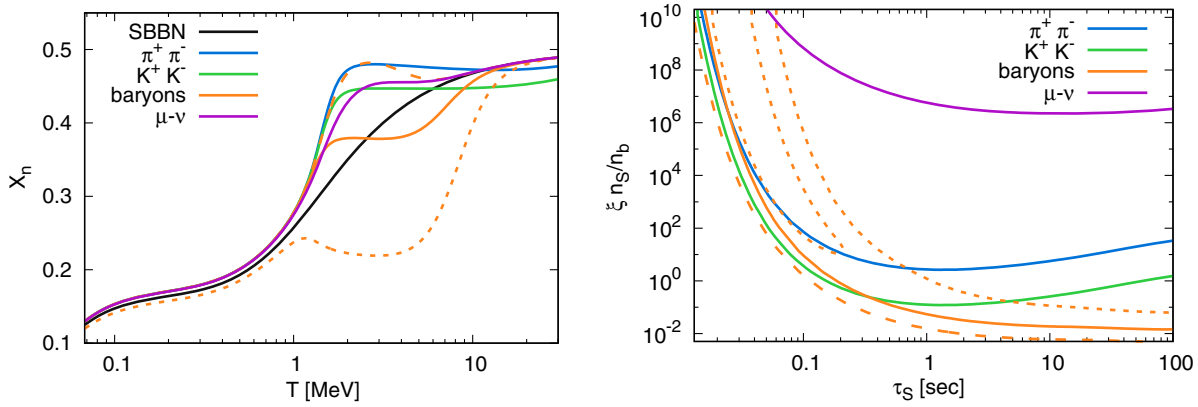

FIG. 3. *Left*: X_n evolution for the SBBN and the injection of pions, kaons, baryons, and muons (neutrinos) as described in the text (see Appendix A 1 for the muon case) for lifetimes of 0.05 s with the initial Y_S abundance tuned to yield $\Delta Y_p = 0.01$ (maximum allowed shift of Y_p). The baryonic injection is taken at $\kappa = 0.5$ (full line); the lines for $\kappa = 1$ (dashed) and $\kappa = 0.2$ (dotted) are also displayed. *Right*: Limit of injected pairs for each channel as a function of the S lifetime. The upper-right dotted line for $\kappa = 0.2$ is at $Y_p = 0.26$, and the upper-left dotted island yields $Y_p = 0.24$.

with a similar cross section for a charge-conjugated reaction, $K^0 p \to K^+ n$. For neutral kaons, the effects induced by K_L are the most important, and we use $\sigma(K_L n \to K^- p) \simeq \frac{1}{2}\sigma(\bar{K}^0 n \to K^- p)$ and (26) to find $(\sigma v)^{\bar{K}^0}_{pn(K^-)} \simeq 4.5$ mb. Additionally, efficient reactions can also proceed via s-quark being incorporated inside an unstable hyperon and a pion

$$K^- + p \to \Sigma^{\pm}\pi^{\mp}, \qquad \Sigma^0\pi^0, \qquad \Lambda\pi^0, \tag{27}$$

$$K^- + n \to \Sigma^{-}\pi^{0}, \qquad \Sigma^0\pi^-, \qquad \Lambda\pi^-, \tag{28}$$

with the hyperon subsequently decaying into $p/n + X$. The inclusive threshold cross sections found by weighting each hyperon-producing cross sections by their branching ratios to p/n are [44]

$$K^- + p \to n + X\text{:} \qquad (\sigma v)^{K^-}_{pn} \simeq 32 \text{ mb}, \tag{29}$$

$$K^- + n \to p + X\text{:} \qquad (\sigma v)^{K^-}_{np} \simeq 13 \text{ mb}, \tag{30}$$

$$K_L + p \to n + X\text{:} \qquad (\sigma v)^{K^-}_{pn} \simeq 6.5 \text{ mb}, \tag{31}$$

$$K_L + n \to p + X\text{:} \qquad (\sigma v)^{K^-}_{np} \simeq 16 \text{ mb}. \tag{32}$$

Notice the absence of corresponding hyperon reactions initiated by K^+ due to the presence of an anti-s quark.

Representative examples of $X_n(T)$ evolutions and the sensitivity to $\xi_{\pi^\pm} Y_S/\tau_S$ parameter space are shown in Fig. 3. The left panel displays significant modifications to the evolution of neutron abundance at $\tau_S = 0.05$ s with adjustable initial abundance, yielding $\Delta Y_p = 0.01$. The departure from $X_n = 0.5$ at high temperatures is clearly visible. [In fact, for short τ_S, the kaon injection channel at early times leads to a shift of the equilibrium value of X_n to $(\sigma v)^{K^-+K_L}_{pn}/((\sigma v)^{K^-+K_L}_{pn} + (\sigma v)^{K^-+K_L}_{np}) \simeq 0.45$.] As the temperature lowers, the Coulomb-enhanced reaction becomes stronger. For meson injection, these reactions enhance the $p \to n$ conversion, keeping X_n away from the SBBN value. The right panel gives a boundary of the exclusion regions for different injection modes. In addition to the already described channels, charged kaons also give rise to a population of secondary charged pions that should also be included in the analysis of $p \leftrightarrow n$ transitions. Since the constraints are already stronger than for the charged pion case, we neglect this effect, which leads to more conservative bounds.

2. Direct baryonic injection mechanism

If S is heavy enough, the end products after hadronization of the primary decay products (e.g., b or c quarks) may contain baryons. Since S has no baryon number, one should expect an equal number of baryons and antibaryons in the final states. Therefore, one should expect the injection of $n\bar{n}$, $p\bar{p}$, $\bar{n}p$, and $p\bar{n}$ pairs, as well as (in principle) baryonic states with higher multiplicities. The hadronization process and decay of heavy quarks produce many more light mesons than baryons, and a complete analysis must include a Monte Carlo study of the hadronization process (see Ref. [45] for benchmarks of heavy unstable particles decaying into 2 hadronic +1 leptonic jets in the early BBN epoch). Assuming that the heavy quarks inside baryons decay due to the "main" weak decay sequence, $b \to c \to s \to u$, one should also expect a somewhat large number of the final states with a proton or antiproton over neutron or antineutron. We will tune the branching models of S to available particle data on proton production, and take $N_n = \kappa N_p$ and $N_{\bar{n}} = \kappa N_{\bar{p}}$. Furthermore, because of a more frequent appearance of the up-quark at the end of the decay chain, we would take $\kappa \simeq 0.5$ on average.

As in the case of mesons, the thermalization of baryonic decay products is quick (see, e.g., [46]). As a baryonic pair is created in the decay, the baryon is added to the existing population of n or p. The antibaryon will, however, annihilate with either p or n and dissipate into lighter mesons. If it annihilates with its own antiparticle, there is no net change in n/p, but an annihilation with the other species induces a net $n-p$ change. The probability $P_{k\bar{l}}^{i\to j}$ of a net charge exchange $i\to j$ from a $k\bar{l}$ injection is simply given by the weighted annihilation rates

$$P_{p\bar{p}}^{n\to p}=\frac{X_n\langle\sigma v\rangle_{n\bar{p}}}{X_n\langle\sigma v\rangle_{n\bar{p}}+(1-X_n)\langle\sigma v\rangle_{p\bar{p}}},$$
$$P_{p\bar{n}}^{n\to p}=\frac{(1-X_n)\langle\sigma v\rangle_{p\bar{n}}}{X_n\langle\sigma v\rangle_{n\bar{n}}+(1-X_n)\langle\sigma v\rangle_{p\bar{n}}} \quad (33)$$

and similarly for the $n\bar{p}$ and $n\bar{n}$ injections. The baryonic annihilation rates are given by [43]

$$\langle\sigma v\rangle_{n\bar{n}}=\langle\sigma v\rangle_{p\bar{p}}/C=37\ \text{mb},$$
$$\langle\sigma v\rangle_{n\bar{p}}=\langle\sigma v\rangle_{p\bar{n}}=28\ \text{mb}, \quad (34)$$

where the $p\bar{p}$ has the low-v Coulomb correction $C(v)$. The implementation of these processes in the Boltzmann equation then require additional terms,

$$\left.\frac{dX_n}{dT}\right|_{pn}=\frac{-\xi_p\Gamma_S e^{-t\Gamma_S}}{TH(T)}\times(-P_{p\bar{p}}^{n\to p}-\kappa P_{p\bar{n}}^{n\to p}+\kappa P_{p\bar{n}}^{n\to p}+\kappa^2 P_{n\bar{n}}^{n\to p}). \quad (35)$$

As before, the outcome is displayed in Fig. 3. Again, for short S lifetimes and large Y_S, the large numbers of injected particles completely dictate the early X_n value. The constraint on Y_S goes up more sharply in the short S lifetime limit. There is a significant dependence on κ for $\tau_S\gtrsim 0.1$ s, which is washed out by the SM electroweak interactions at earlier times. If we take the extreme limit $\kappa\to 0$, no neutrons are injected and the $p\bar{p}$ pair can only further decrease the n/p ratio, thus constrained by the lower Y_p limit 0.24. On the other hand, a symmetric injection $\kappa=1$ enhances the n/p ratio as the antibaryon mostly annihilates on protons, more abundant than neutrons by a factor of ~6–7 after the standard n/p freeze-out. For $\kappa\gtrsim 0$, the final Y_p can be either increased or decreased, depending on whether the S particles decay away before or after the displaced X_n equilibrium crosses the SBBN n/p freeze-out curve. As shown for $\kappa=0.2$ in Fig. 3, there is a $Y_p=0.24$ exclusion island at low lifetimes and larger lifetimes are constrained by $Y_p=0.26$. We use $\kappa=0.5$ as our baryonic injection constraint benchmark.

Similar results were found in Ref. [47] in the context of MeV-scale reheating with hadronic energy injection. The importance of hadron injection on Y_p completely dominates the second order effects of incomplete neutrino thermalization after late reheating. Even though we have checked that S does not reheat the SM bath in our model, $\rho_S\ll\rho_{\rm SM}$, the importance of hadronic injection reinforces the validity of our treatment even in the absence of the full treatment of Boltzmann equation for neutrinos.

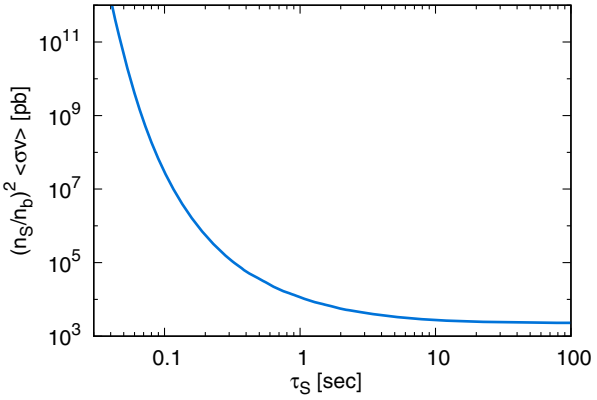

FIG. 4. Constraints on $Y_S^2\langle\sigma v\rangle_{\pi^+\pi^-}$ from SS annihilations into charged pions from the BBN ^{4}He abundance at $Y_p=0.26$.

3. Meson injection from residual annihilations

In addition to its decay products, S can also inject particles in the cosmic medium via SS annihilations to charged pions. The injected pions interact with the cosmic medium in the same way as from the S decays described above. The Boltzmann equation takes the addition term (25), with the injected pion density now given by

$$n_{\pi^\pm}^{\rm ann}=\tau_{\pi^\pm}n_S^2(T)\langle\sigma v\rangle_{\pi^+\pi^-}=\tau_{\pi^\pm}Y_S^2 n_b^2(T)e^{-2t\Gamma_S}\langle\sigma v\rangle_{\pi^+\pi^-}, \quad (36)$$

where $\langle\sigma v\rangle_{\pi^+\pi^-}$ is the nonrelativistic annihilation cross section $\sigma v(2m_S)$ as per Eq. (12), rescaled by the pionic branching ratio at $\sqrt{s}=2m_S$. The $n_{\pi^\pm}^{\rm ann}\propto n_S^2\propto T^6$ dependence implies a much stronger impact at high energies, enforcing the displaced initial condition $X_n^i\simeq 0.47$. As S decays away, its impact on $X_n(T)$ is even more rapidly exponentially suppressed, and its constraints are less stringent than decays at very short lifetimes. The bounds from annihilation are given in the $Y_S^2\langle\sigma v\rangle_{\pi^+\pi^-}-\tau_S$ parameter space and displayed in Fig. 4.

B. Energy density requirements

The resultant BBN abundances depend on the nuclear reaction rates and how efficient they are as the Universe expands. One by one, the reaction rates drop out of equilibrium, as the Universe expands and cools. If the Hubble rate is increased due to a large energy density locked in a dark sector, the active reaction time would shorten, potentially spoiling the SBBN results. For our study, the most important effect is the change of the Hubble

rate during the n/p freeze-out, which again affects Y_p. However, we can also use as a constraint a well-measured quantity in cosmology, the total energy density carried by neutrinos.

The neutrinos decouple from thermal processes at $T \sim 2$ MeV. If the decaying particle is heavy and does not decay into neutrinos, it will reheat electron-photon fluid with respect to the neutrinos, decrease T_ν/T_γ, and equivalently lower $N_{\rm eff}$. The Planck Collaboration measured $N_{\rm eff} = 3.04 \pm 0.33$ at 2σ, including their cosmic microwave background (CMB) results and external cosmological data [48], which imposes $N_{\rm eff} > 2.71$ as a lower bound. In the extreme case of a full reheating of the SM bath in the 1–10 MeV temperature range, numerical solutions show a decrease in $N_{\rm eff}$ due to inefficient thermalization of neutrinos [49], which slows down the Hubble rate resulting in a decreased Y_p [47]. In our model, we verified *a posteriori* that S never dominates the cosmic energy density $\rho_S < \rho_{\rm SM}$ prior to its decay. We can then treat the energy injection as a perturbation around the standard case, without the need for a full numerical machinery of the neutrino thermalization.

The energy densities and Hubble rate form a closed system of differential equations,

$$\dot\rho_S + 3H\rho_S = -\Gamma_S\rho_S, \qquad \dot\rho_{\rm rad} + 4H\rho_{\rm rad} = \Gamma_S\rho_S,$$
$$H^2 = \frac{8\pi G}{3}(\rho_{\rm rad} + \rho_S), \tag{37}$$

where we have assumed a nonrelativistic S and omitted the variation in relativistic degrees of freedom. Assuming steplike decoupling and changes in relativistic degrees of freedom, the T evolution separates into three regions. For $T > T_\nu^{\rm decoup}$, neutrinos are in equilibrium with the electromagnetic bath and ρ_S is injected equally in $e^\pm$'s, ν's, and γ's. For $T_\nu^{\rm decoup} > T > T_{m_e}$, the neutrinos are simply redshifted while the electron-photon bath is heated by the S decays. For $T_{m_e} > T$, electrons become nonrelativistic and transfer their entropy to photons, additionally heating the photon bath compared to the neutrino bath.

If S does not dominate the energy density of the Universe before its decay, we can write $\rho_S = \delta_S\rho_{\rm rad}^{\rm SM}$, $\rho_{\rm rad} = \rho_{\rm rad}^{\rm SM}(1+\delta_{\rm rad})$ and expand (37) around the δ perturbations to solve the system analytically. At linear order, we find the solutions

$$\rho_S(t) = \frac{c_S}{t^{3/2}}e^{-\Gamma_S t}, \qquad \rho_{\rm rad}(t) = \frac{c_{\rm rad}^i}{t^2}[1+F(t)],$$
$$F(t) = \frac{c_S}{c_{\rm rad}^i\sqrt{\Gamma_S}}\frac{1}{\Gamma_S t}\left[\Gamma_{3/2}(\sqrt{\Gamma_S t}) - \Gamma_{5/2}(\sqrt{\Gamma_S t}) + \frac{\sqrt{\pi}}{4}\right], \tag{38}$$

where $\Gamma_{3/2}$, $\Gamma_{5/2}$ are incomplete Gamma functions and the integration constants c_S, $c_{\rm rad}$ are set to have $\rho_S = m_S n_S$ and $\rho_{\rm rad} = \rho_{\rm rad}^{\rm SM}$ at some early time $\Gamma_S t \ll 1$. After the neutrinos decouple, the injected energy is distributed to the photon-electron bath and its energy density departs for the neutrino bath,

$$\rho_\gamma^{\rm mid}(t) = \tilde g_{\gamma+e}\frac{c_{\rm rad}^i}{t^2}[1+F(t)] + \tilde g_\nu\frac{c_{\rm rad}^i}{t^2}[G(t) - G(t_\nu^{\rm decoup})],$$
$$\rho_\nu^{\rm mid}(t) = \tilde g_\nu\frac{c_{\rm rad}^i}{t^2}[1+F(t) - G(t) + G(t_\nu^{\rm decoup})], \tag{39}$$

where $\tilde g_i \equiv g_i/(g_{\gamma+e}+g_\nu)$ is the fraction of relativistic degrees of freedom of each bath, $t_\nu^{\rm decoup}$ the neutrino decoupling time, and

$$G(t) = \frac{c_S}{2c_{\rm rad}^i}\sqrt{\frac{\pi}{\Gamma_S}}{\rm erf}(\sqrt{\Gamma_S t}) - \frac{c_S}{c_{\rm rad}^i}\sqrt{t}e^{-\Gamma_S t}. \tag{40}$$

Finally, after the electrons become nonrelativistic, they effectively transfer their entropy to the photon bath. Assuming an instantaneous transition, entropy continuity implies an increase of energy density by a factor of $\delta = (g_e+g_\gamma)^{1/3}/g_\gamma^{1/3} = (11/4)^{1/3}$. Matching boundary conditions, the energy densities at late times are

$$\rho_\gamma^{\rm late}(t) = \tilde g_{\gamma+e}\delta\frac{c_{\rm rad}^i}{t^2}\left[1+\alpha F(t) + c\frac{t_e}{t}\right] + \tilde g_\nu\alpha\frac{c_{\rm rad}^i}{t^2}[G(t) - G(t_e) + \delta(G(t_e) - G(t_\nu^{\rm decoup}))], \tag{41}$$

$$\rho_\nu^{\rm late}(t) = \tilde g_\nu\frac{c_{\rm rad}^i}{t^2}\left[1+\alpha\{F(t) - G(t) + G(t_e) + \delta(G(t_\nu^{\rm decoup}) - G(t_e))\} + c\frac{t_e}{t}\right], \tag{42}$$

with $\alpha = 1/(\delta\tilde g_{\gamma+e} + \tilde g_\nu)$ and c a boundary condition that is irrelevant in the $t \to \infty$ limit.

The temperature-time dependence is found via $\rho_{\rm rad}(t) = \pi^2 g_\star T^4/30$. Since the neutrino interaction rate scales as $\Gamma_{\nu_e} \sim T^5$, we find the neutrino decoupling time in the modified cosmology by equating $(T_\nu^{\rm decoupl})^5/H(T_\nu^{\rm decoupl}) = (T_\nu^0)^5/H_0(T_\nu^0)$, with H the perturbed Hubble rate and T_ν^0 the neutrino decoupling temperature in the SM. In the Maxwell-Boltzmann approximation, $T_\nu^0 = 2$ MeV, but thermal refinements in the interaction rates and phase space tend to yield a lower value $T_\nu^0 = 1.4$ MeV [50].

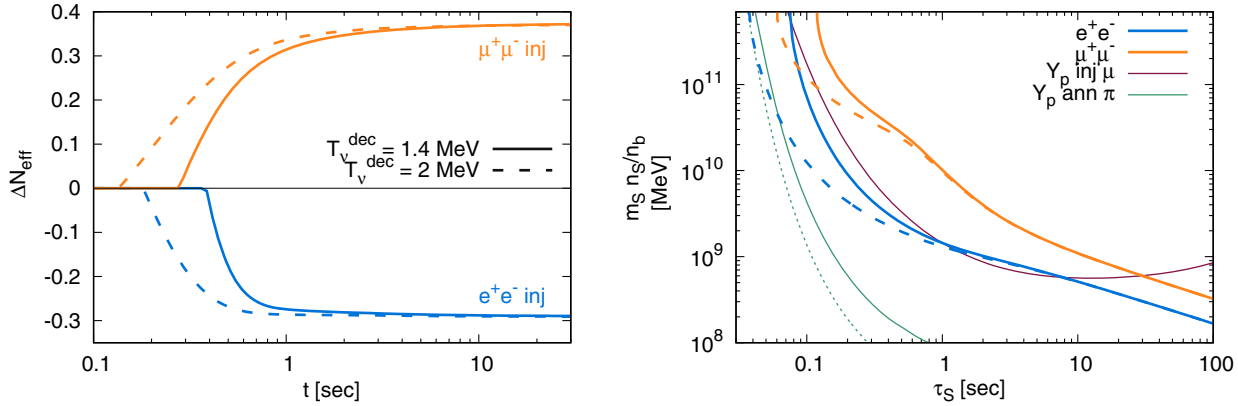

FIG. 5. *Left*: Departure from the SM $N_{\rm eff}$ as the Universe cools down for electron injections (blue lines) and muon injections (orange lines; see Appendix A 2 for details). The extrema of the neutrino decoupling temperature ranges are shown in full lines and dashed lines as labeled in the figure. *Right*: Bound of maximal stored energy decaying into electrons or muons as a function of particle lifetimes. The full lines and dashed lines represent the neutrino decoupling temperatures as on the left. We also show for comparison some benchmark bounds in this parameter space from the BBN Y_p results. The thin olive curves are the neutron enrichment constraint from annihilation into pions for $m_S = 140$ MeV (solid lines) and $m_S = 275$ MeV (dotted lines). The thin purple line is the Y_p constraint for a $m_S = 250$ MeV particle decaying into muons.

Then, at $\Gamma_S t \gg 1$, we can evaluate T_ν/T_γ and find

$$N_{\rm eff} = 3\left(\frac{T_\nu}{T_\gamma}\right)^4\left(\frac{11}{4}\right)^{4/3} \simeq 3\times\frac{\delta\tilde{g}_{\gamma+e}+\tilde{g}_\nu-\frac{c_S}{2c^i_{\rm rad}}\sqrt{\frac{\pi}{\Gamma_S}}+(1-\delta)G(t_e)+\delta G(t_\nu^{\rm decoup})}{\delta\tilde{g}_{\gamma+e}+\tilde{g}_\nu+\frac{\tilde{g}_\nu}{\delta\tilde{g}_{\gamma+e}}\left(\frac{c_S}{2c^i_{\rm rad}}\sqrt{\frac{\pi}{\Gamma_S}}-(1-\delta)G(t_e)-\delta G(T_\nu^{\rm decoup})\right)} \tag{43}$$

to constrain energy injection into electrons. We display in Fig. 5 the departure from $N_{\rm eff} = 3$ as a function of time for $\tau_S = 0.1$ s and the two neutrino decoupling temperature benchmarks. The limits are also shown in units of stored energy density $m_S n_S/n_b$, where m_S is in MeV. If the S decay happens after the neutrino decoupling, all energy is deposited in the photon bath and the result is independent of our choice of T_ν^0. If decays happen earlier, the photon and neutrinos are potentially still coupled and the energy emitted in S decays only influences $N_{\rm eff}$ after decoupling. As such, the constraints has a $t \propto (1/T_\nu^0)^2$ dependence. We adopt the conservative side,[2] $T_\nu^0 = 1.4$ MeV, as our bounds on the $m_S - \tau_S$ parameter space. Notice the constraints for SS annihilations to pions are much stronger and will be dominant when the pionic annihilation channel is open, i.e., for $m_\pi < m_S < 2m_\pi$.

[2]The neutrino decoupling temperature is not a well-defined quantity as the weak cross section with electrons scales as E^2 resulting in energetic neutrinos remaining coupled slightly longer, in addition to the ν_μ and ν_τ decoupling slightly earlier [50]. Our choice of $T_\nu^0 = 1.4$ MeV is on the lowest end of decoupling temperatures, underestimating the energy fraction going into the electromagnetic bath as higher generations of neutrinos and lower energy electron-neutrinos have already decoupled. A full Boltzmann evolution of the neutrino spectrum will yield stronger constraints.

C. Late-time energy injection

In the example of the S particles coupled through the Higgs portal, the most stringent constraints on lifetime come from the considerations of n/p freeze-out. In other models, with additional channels of annihilation that can suppress metastable abundances, the constraints on lifetime would not be as stringent and would mostly come from the considerations of late energy injection. For completeness, we also discuss these constraints here. Modification of BBN by unstable particles with lifetimes in excess of 200 s has been considered in detail, both through hadronic [43,51], electromagnetic [52], or combined [44,46,53] energy cascades.

Hadronic injection after $t \gtrsim 200$ s is most efficient at modifying the final yields of the less abundant light nuclei D, ^{3}He, ^{6}Li, and ^{7}Li. After most of ^{4}He has been synthesized, the BBN enters the regime ($T \sim 50$ keV) when neutrons are rare, $O(10^{-5})$ or so, yet their abundances are critical in determining the final abundance of deuterium. At that stage, any additional neutrons brought into the system through external processes such as heavy particle decays lead to the increase of the deuterium abundance. (Incidentally, it also leads to the suppression of ^{7}Be and consequently of ^{7}Li [43].) The increase of D production can be exacerbated by the hadro-dissociation of ^{4}He in the process of slowing down of injected hadrons. Additional

production of ^{3}He through spallation can also affect the ^{3}He/D ratio [54]. Secondary and tertiary processes may also generate ^{6}Li and ^{9}Be [55,56]. Detailed studies of the ensuing constraints [53] show strong sensitivity to hadronic (mostly nucleonic) decays of metastable particles with lifetimes in the hundreds of seconds and longer, and initial abundances comparable to or even smaller than that of baryons. In recent years, these constraints have only gotten stronger, primarily due to steady observational progress in determination of primordial D/H [57].

If for some reason, hadrons and specifically nucleons are absent from the decay chains, the abundances of light elements can be modified by the late injection of electromagnetic energy. At early times this mechanism is inefficient, as radiation quanta with energy in excess of nuclear binding are quickly energetically degraded by ambient plasma. The photodissociation therefore sets in at late times leading to a suppression D ($t \gtrsim 10^4$ s) and additional production of ^{3}He for $t \gtrsim 10^6$ s. Since typically 45% of hadronic energy injection is dissipated electromagnetically in the hadronization cascade [53], the late-time energy injection constraints on a heavy particle are dominated by the electromagnetic reactions in the BBN network.

IV. RESULTS

We are now in a position to perform a scan in parameter space of the minimal Higgs model, constrained by the consistency with BBN. In Fig. 6, we display the parameter space, in both the lifetime and an effective decay length $L_{\rm dec} = c\tau_S\beta_S(E_S/m_S)$. We assume an average E_S of 200 GeV, from a Higgs typically boosted at 400 GeV at the LHC. The resulting constraints, along with the assumptions considered in each mass range are described below.

(i) *Region A* $2m_e < m_S < 2m_\mu$: The constraint comes from the decrease in $N_{\rm eff}$ with the entropy dump in the SM bath after neutrino decoupling. We take the neutrino decoupling temperature to be $T_\nu^0 = 1.4\,$MeV as a conservative limit.

(ii) *Region B* $m_\pi < m_S < 2m_\pi$: This region is dominated by the SS annihilation to $\pi^+\pi^-$. We also derived the same constraint as region A from $N_{\rm eff}$ up to $m_S = 2m_\mu$, in addition to the raised $N_{\rm eff}$ from decays into muons in the $2m_\mu < m_S < 2m_\pi$ and the Y_p constraints from S decaying into muons (see Appendix for muon injection considerations). They all yield weaker bounds than the $SS \to \pi^+\pi^-$ constraints, of $\tau_S > 0.3$ s or longer.

(iii) *Region C* $2m_\pi < m_S < 2m_K$: The abundance Y_S weighted by the pion branching ratio constrains the region via direct charged pion decays. We assume 2/3 go into charged pions and 1/3 is radiated away in π^0.

(iv) *Region D* $2m_K < m_S < 1.4$ GeV: The abundance Y_S weighted by the kaon branching ratio constrains the region via direct charged kaon decays. We assume 1/2 go into charged kaons and 1/2 into $K^0\bar{K}^0$. Only half of the neutral kaons survive as K_L, creating similar in numbers metastable populations of K_L, K^+, and K^-.

(v) *Region E* 1.4 GeV $< m_S < 2m_D$: By strangeness conservation, we assume that all s quarks yield a kaon, half charged and half neutral. Since we do not have model-independent branching ratios of S in this mass regime, we vary the description according to the assumptions in each decay model. For the baseline model, we assume that 100% decays to the kaons and apply our kaon injection constraints. For the perturbative spectator model, the kaon branching ratio is given by (11), with non-negligible contributions from decays to pions, muons, and eta mesons, resulting in weaker bounds until the c-quark

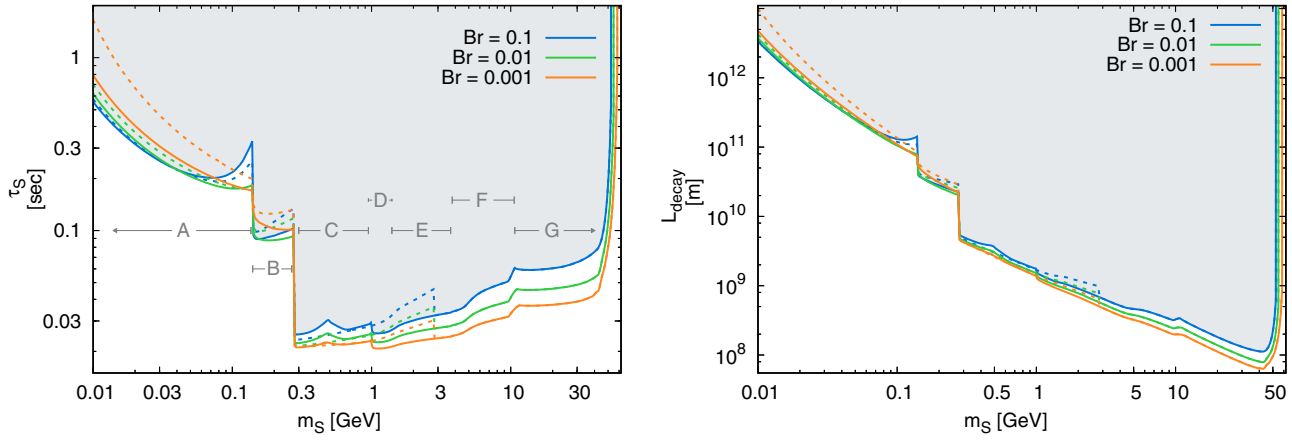

FIG. 6. *Left*: Lifetime constraint as a function of the S mass for three $h \to SS$ branching ratios. The lettered regions represent different assumptions or physics and are described in the text. The dotted lines correspond to the perturbative spectator model. *Right*: Same as left, except transposed in the decay length of S, assuming it is boosted to $E_S = 200$ GeV.

threshold. At $m_S = 2m_c$ the hadronic modeling dependence largely goes away.

(vi) *Region F* $2m_D < m_S < 2m_b$: We utilize the branching fractions of $c\bar{c}$ from e^+e^- at $\sqrt{s} = 10.5$ GeV into D mesons from Ref. [58] and weight each channel by its inclusive $K^\pm$ branching ratios to find a hadronization yield of 0.63 K^+K^- pair per S decay into c quarks. Rescaled by $\mathrm{Br}(S \to c\bar{c})$, the same constraints from kaon injection apply. Above the $2m_{\Lambda_c}$ threshold, a $c\bar{c}$ typically forms a c baryon with a 0.06 probability [58], which then hadronizes to p or n. We find this constraint weaker than the kaons injection and use the K^+K^- result across this entire range.

(vii) *Region G* $m_S > 2m_b$: The main decay channel here are pairs of $b\bar{b}$ quarks. The charged pion, charged kaons, and proton multiplicities in the $b\bar{b}$ decay of a Z boson are measured to be 18.44 ± 0.63, 2.63 ± 0.14, and 1.00 ± 0.08, respectively, by the ALEPH Collaboration [59]. We assume the ratio holds in the hadronization of lower center-of-mass decays into $b\bar{b}$ and scale by the mean charge multiplicity fit [60]

$$N_{\mathrm{ch}}(s) = -0.577 + 0.394\ln(s/s_0) + 0.213\ln^2(s/s_0) + 0.005(s/s_0)^{0.55}, \quad (44)$$

where $s_0 = 1\ \mathrm{GeV}^2$. This fit agrees well in both e^+e^- and $p\bar{p}$ collisions between $\sqrt{s} \sim 2$ GeV and 2 TeV. This gives us an estimate for the baryon injection of the $b\bar{b}$ branching fraction of S. We further assume a 50% smaller injection of $n(\bar{n})$ to utilize our baryon injection constraints. The accompanying pions and kaons also independently yield comparable constraints, not shown in the figure.

V. DISCUSSION

We have considered, in some detail, constraints on the lifetimes of the scalar particles, coupled to the Higgs portal via a minimal set of couplings. To stay relevant for the LHC, we have concentrated on the $m_S < m_h/2$ case, that allows pair production of S states in the decay of Higgs bosons. The same coupling is responsible for the cosmological depletion of S particles, leading to their metastable abundance in the early universe.

We find that throughout almost the whole mass range considered in this work, $2m_\mu < m_S < m_h/2$, the constraints on the lifetime of S particles are stronger than 0.1 s. Moreover, the results have a relatively mild dependence on the $\mathrm{Br}(h \to SS)$. The reason for that is as follows: the experimental limits on $\mathrm{Br}(h \to SS)$ are already strong enough to limit the annihilation rate of SS pairs to the SM states to be much less than one picobarn, and consequently the metastable abundance of S particles per nucleon is quite high, $Y_p \gg 1$. This leads to a massive injection of nucleons and mesons at early times, which raises the n/p ratio, and creates larger yields of ^{4}He compared to SBBN. Contributions of very light S particles to the Hubble rate during the n/p freeze-out also raises Y_p. The limits on τ_S are robust, and have rather mild dependence on the uncertainties in our treatment. This is because the initial large metastable Y_S abundance needs to be depleted prior to the n/p freeze-out time $t_{n/p}$, leading to the requirement $\tau_S \ll t_{n/p}$. Consequently $O(1)$ variations in the yields of mesons and nucleons in the final states can be compensated by small variations in τ_S, parametrically on the order $\log^{-1}(t_{n/p}/\tau_S)$, to produce the same influence on BBN. For the same reasons, our limits are also very insensitive to the exact observational constraint on ΔY_p, and we take a rather conservative limit of 0.01 (allowing ± 0.01 deviations from the observed/calculated mean). From the point of the LHC physics, the most promising is a scenario with a mass m_S not far below $m_h/2$. In that case, the effective decay length has to be on the order of or smaller than $\sim 10^8$ m (see Fig. 6), providing a 10^{-6} *minimum* probability for a decay within a 100 m length purposely built detector. Given that the high-luminosity LHC would produce copious numbers of the Higgs bosons, there is a chance to cover the entire lifetime range for masses within the 10–50 GeV range.

In our paper we have neglected the effect of S self-interaction. It turns out that the obtained results are also robust to a self-interaction induced reduction in number densities. Retaining the approximate Z_2-symmetry considered, a quartic $\lambda_4 S^4$ interaction could be considered, which would lead to, e.g., $4S \to 2S$ processes. Thus, large self-interaction can maintain chemical equilibrium within the dark sector through efficient number changing processes and further deplete the S number density [61,62] (see Ref. [63] for a realization of strongly interacting Z_2 Higgs portal dark matter). After thermal decoupling from the SM, entropy conservation in the S bath with chemical equilibrium imposes the scaling $n_S \propto 1/a^3 \ln a$ [61], which implies an additional depletion of $Y_S \sim Y_S^{\mathrm{fo,SM}} \ln(a/a_{\mathrm{fo,SM}})$, where "fo,SM" denotes the values without the self-interaction. The S metastable abundance will therefore be at most lowered by a factor of a few (certainly less than an order of magnitude), which would have a very limited impact on the τ_S limit.

It is easy to see that the above considerations can be generalized to other models of the Higgs-portal-coupled particles. For example, consider a fermion χ, coupled to the Higgs via $H^\dagger H(\bar{\chi}\chi)$ or $H^\dagger H(\bar{\chi}i\gamma_5\chi)$ dimension-five operators, and having a small decay term such as neutrino portal $LH\chi$. The main analysis of our work can be recast for that model, especially in the part that connects Higgs decays with a metastable abundance of χ. Evidently, for

$\mathrm{Br}(h \to \chi\bar{\chi}) \sim \mathrm{Br}(h \to SS)$ input, one will end up with $Y_\chi \sim Y_S$. The only change will be in the yields of mesons and baryons in the decays of χ compared to S. However, it is well known that already for m_χ above 250 MeV, the yields of pions and kaons are substantial [17], giving confidence that for the most part the same constraints we have derived for τ_S will translate to similar limits on τ_χ.

The analysis performed in this paper can easily be generalized to other models of metastable particles, with different types of interactions, via Z, Z', etc. In the limit when Z' is outside of the LHC reach, one could have a set of effective operators connecting χ with the SM fields, such as $\frac{1}{\Lambda^2}\bar{\chi}\gamma_\mu\chi\bar{q}\gamma_\mu q$, where Λ is some energy scale. The χ pair-production cross section in this case will scale as $\sigma_{q\bar{q}\to\chi\bar{\chi}} \propto E_q^2\Lambda^{-4}$, where E_q is a typical (anti)quark energy, while the cosmological annihilation cross section has $\sigma_{\chi\bar{\chi}\to q\bar{q}}v \propto m_\chi^2\Lambda^{-4}$ scaling. Therefore, the LHC-relevant cross section can be enhanced relative to the annihilation rate by a parametrically large ratio, E_q^2/m_χ^2 if m_χ is parametrically smaller than the TeV scale. Therefore, one can easily have a range of parameters with a relatively large $\chi\bar{\chi}$ pair-production cross section, while having very small annihilation rates, rendering $n_\chi/n_b \gg 1$, and resulting again in strong BBN constraints on lifetimes, $\tau_\chi < 0.1$ s. Therefore, we conclude that some simple Z' mediated models of metastable particles can also be strongly restricted by cosmology, making them a perfect candidate for the searches of metastable particles at the LHC.

It is also instructive to consider models where constraints on the lifetime of metastable particles are *much weaker*. Clearly, one needs an effective new mechanism for the self-annihilation in the early universe, as the Higgs channel is too inefficient. Staying within the Higgs portal models, consider the following potential with two real scalars:

$$V(H,S_1,S_2) = H^\dagger H(\lambda_1 S_1^2 + \lambda_2 S_2^2 + A_1 S_1 + A_2 S_2) + \lambda_{12}S_1^2S_2^2 + V(S_1) + V(S_2) + V(H^\dagger H), \tag{45}$$

with the following hierarchy of couplings:

$$\lambda_1 \gg \lambda_2; \quad A_1 \ll A_2; \quad \lambda_{12} \sim O(1); \quad m_{S_1} > m_{S_2}. \tag{46}$$

These choices will lead to a long-lived S_1, somewhat shorter-lived S_2, a predominant decay of the Higgs boson to pairs of S_1, and cosmological depletion of S_1 via $S_1S_1 \to S_2S_2$ annihilation before BBN (even if S1 has a lifetime $\tau_S \gg 1$ s) with potentially a large cross section due to a sizable λ_{12} coupling. Most importantly, in this model the Higgs decay to pairs of S_1 does not result in a prediction of Y_{S1} abundance, which can be quite small even for small values of $\mathrm{Br}(H \to S_1S_1)$. If $Y_{S1} \ll 1$, there would not be enough decay mesons and nucleons to affect early n/p freeze-out, and constraints on τ_{S1} will be coming only from the considerations of late decays with hadronic or electromagnetic energy injection. Instead of $\tau_S < 0.1$ s, one expects to have sensitivity to $\tau_{S1} \sim 10^3$ s, or even worse, beyond 10^4 s, if decays of S_1 are mostly leptonic. This example is not unique, and there are other models where constraints on lifetimes and decay lengths are relatively lax, provided that there are extra channels that ensure efficient cosmological annihilation of metastable particles.

ACKNOWLEDGMENTS

We thank D. Curtin, M. McCullough, P. Meade, M. Papucci, and J. Shelton for soliciting this study, as well as D. Curtin and B. Shuve for very helpful discussions. The work of M. P. is supported in part by NSERC, Canada, and research at the Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MEDT.

APPENDIX: MUON INJECTIONS IN EARLY BBN

Although muon injection related constraints are relatively weak and do not affect our main result, Fig. 6, in this appendix we provide additional details concerning the effects of muon injections on Y_p and N_{eff} to verify that they are subdominant to the residual annihilation to pions in the $m_\pi < m_S < 2m_\pi$ mass range.

1. Neutron enrichment

Muon injection physics differs from the previous scenarios of meson and baryon injection considered in Sec. III A. The direct charge exchange is through the weak force, as opposed to the strong force in the other cases, and is completely negligible over the lifetime of the muon. Instead, the reactions can happen via the energetic neutrinos emitted by the muon decays. The general case for $\nu + X$ injection in BBN is treated in Ref. [64], and the specific case of muon injection after $t \sim 100$ s has been covered in Ref. [44], to which we refer the reader for details. Assuming stopped muons, the authors solved for the injected neutrino energy spectrum, including redshifting and averaged over flavor oscillations, to be integrated in the $n - p$ conversion rate. At earlier times, we know background neutrinos are coupled to $e^\pm$ down to $T \simeq 2$ MeV, and energetic injected neutrinos must accordingly deplete their energy efficiently as well. Summing over the possible interactions with the background neutrinos and $e^\pm$ [65], the collision rate of an injected electron neutrino with the bath is given by

$$\Gamma^{\nu_e}_{\rm coll}(E_\nu, T) = \frac{7\pi}{135} G_{\rm F}^2 E_\nu [(5 + g_{\rm L}^2 + g_{\rm R}^2) T_\nu^4 + 4(g_{\rm L}^2 + g_{\rm R}^2)\eta(T) T_\gamma^4], \simeq \left(\frac{E_\nu}{32\ \text{MeV}}\right)\left[\frac{5.7}{\text{sec}}\left(\frac{T_\nu}{1\ \text{MeV}}\right)^4 + \frac{1.3}{\text{sec}}\eta(T_\gamma)\left(\frac{T_\gamma}{1\ \text{MeV}}\right)^4\right], \tag{A1}$$

$g_{\rm L} = 1/2 + \sin^2\theta_{\rm w}$, $g_{\rm R} = \sin^2\theta_{\rm w}$, while $\eta(T) = 1$ for $T \gtrsim m_e$ and exponentially falls to 0 at lower temperatures. We follow the implementation of Ref. [44] and correct for the removal of energetic neutrinos coupled to the electromagnetic plasma by adding an effective collision lifetime in the neutrino energy distribution (normalized on n_b),

$$f_e(T, E_\nu) = \Gamma_S Y_S \int_T^\infty \frac{dT_1 e^{-t_1\Gamma_S}}{H(T_1)T_1} \times F_e\left(E_\nu, \frac{E_0 T}{T_1}\right) e^{-\int_T^{T_1} dT_2 \frac{\Gamma_{\rm coll}(E_\nu \frac{T_2}{T}, T_2)}{H(T_2)T_2}}, \tag{A2}$$

where F_e is the distribution at injection time T_1, averaged over flavor oscillations. The charge-exchange rate to be inserted in the Boltzmann Eq. (19) is

$$\Gamma^\nu_{pn} = n_b(T) \int_0^{E_0} \sigma^{\bar\nu}_{pn} f_e(T, E_\nu) dE_\nu \tag{A3}$$

and similarly for the reverse np direction. The resulting constraints are shown in Fig. 3. Our results lean on the conservative side on a few assumptions. For simplicity, we assumed one collision for the neutrino thermalization, instead of following energy degradation over a shower of multiple interactions. Moreover, we took the collision time of the electron-neutrino, even though there are muon-neutrino states in the oscillations. Since $\Gamma^{\nu_e}_{\rm coll} > \Gamma^{\nu_\mu}_{\rm coll}$, we overestimate the actual collision time and the overall conversion rate should be slightly larger.

2. Energy injection partitioned between photon and neutrino baths (e.g., muon injection)

The case for energy injection from muons with respect to $N_{\rm eff}$ is somewhat interesting as its decay products, neutrinos and electrons, clearly deposit their energy in two different baths, once everything is decoupled. Both T_γ and T_ν will rise, but since the two neutrinos carry more energy than the electron for a muon decay, we expect a rise in $N_{\rm eff}$. More precisely, we solve a similar set of equations as (37), except the photon bath absorbs a ξ proportion of the S decay energy and the neutrino bath gets the remaining $(1-\xi)$ portion. Before neutrino decoupling, the radiation bath evolves as in Eq. (38). Each decay product carries on average the energy [44]

$$\langle E_e\rangle = 37.0\ \text{MeV}, \qquad \langle E_{\nu_e}\rangle = 31.7\ \text{MeV}, \qquad \langle E_{\nu_\mu}\rangle = 37.0\ \text{MeV}. \tag{A4}$$

After neutrino decoupling, the energetic neutrinos can still collide with the ambient electrons until $\Gamma^{\nu_e}_{{\rm coll}-e} < H$, where $\Gamma^{\nu_e}_{{\rm coll}-e}$ is the collision rate with electrons only, the T_γ-dependent term in Eq. (A1). Then, the energy distributed to the photon bath separates into two regimes,

$$\xi_1 = \frac{\langle E_e\rangle + \frac{\Gamma^{\nu_e}_{{\rm coll}-e}}{\Gamma^{\nu_e}_{\rm coll}}\langle E_{\nu_e}\rangle + \frac{\Gamma^{\nu_\mu}_{{\rm coll}-e}}{\Gamma^{\nu_\mu}_{\rm coll}}\langle E_{\nu_\mu}\rangle}{m_\mu} \simeq 0.47, \qquad \xi_2 = \frac{\langle E_e\rangle}{m_\mu} = 0.35, \tag{A5}$$

where the muon-neutrino collision term is given by

$$\Gamma^{\nu_\mu}_{\rm coll}(E_\nu, T) = \frac{7\pi}{135} G_{\rm F}^2 E_\nu [(5 + (g_{\rm L}-1)^2 + g_{\rm R}^2) T_\nu^4 + 4((g_{\rm L}-1)^2 + g_{\rm R}^2)\eta(T) T_\gamma^4]. \tag{A6}$$

Following the same procedure as in Sec. III B, we find

$$N_{\rm eff} = 3 \times \frac{\delta\tilde g_{\gamma+e+\tilde g_\nu} + \frac{c_S(\delta\tilde g_{\gamma+e}(1-\xi_2) - \xi_2\tilde g_\nu)}{2c^2_{rad}\tilde g_\nu}\sqrt{\frac{\pi}{\Gamma_S}} + \frac{C}{\tilde g_\nu}}{\delta\tilde g_{\gamma+e} + \tilde g_\nu - \frac{c_S(\delta\tilde g_{\gamma+e}(1-\xi_2) - \xi_2\tilde g_\nu)}{2c^2_{rad}\delta\tilde g_{\gamma+e}}\sqrt{\frac{\pi}{\Gamma_S}} - \frac{C}{\delta\tilde g_{\gamma+e}}}, \tag{A7}$$

$$C = \delta(\xi_2 - \xi_1) G(t_{\rm coll}) + \delta(\xi_1 - \tilde g_{\gamma+e}) G(t_\nu^{\rm decoup}) + \xi_2(1-\delta) G(t_e), \tag{A8}$$

with $t_{\rm coll}$ found by solving $\Gamma^{\nu_e}_{{\rm coll}-e} = H$. The physics is constrained by $N_{\rm eff} < 3.37$. The time dependence of the departure from $N_{\rm eff} = 3$ is shown for $\tau_S = 0.2$ s and the two choices of T^0_ν in Fig. 5. The corresponding constraints on the maximal stored energy for a given lifetime are shown on the right. For comparison with the muon-induced Y_p bound, we display the curve for $m_S = 250$ MeV from neutron enrichment in purple. Independently from the choice of T^0_ν, the Y_p bounds from μ injection are more constraining for $\tau_S \gtrsim 0.2$ s than its $N_{\rm eff}$ impact, while the annihilation to $\pi^+\pi^-$ provides the dominant constraint in the entire $2m_\mu < m_S < 2m_\pi$ range.

[1] P. W. Graham, D. E. Kaplan, S. Rajendran, and P. Saraswat, J. High Energy Phys. 07 (2012) 149.
[2] N. Craig, A. Katz, M. Strassler, and R. Sundrum, J. High Energy Phys. 07 (2015) 105.
[3] E. Izaguirre and B. Shuve, Phys. Rev. D **91**, 093010 (2015).
[4] B. Batell, M. Pospelov, and B. Shuve, J. High Energy Phys. 08 (2016) 052.
[5] V. Khachatryan *et al.* (CMS Collaboration), Phys. Rev. D **91**, 052012 (2015).
[6] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **92**, 072004 (2015).
[7] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **92**, 012010 (2015).
[8] J. P. Chou, D. Curtin, and H. J. Lubatti, Phys. Lett. B **767**, 29 (2017).
[9] R. H. Cyburt, B. D. Fields, K. A. Olive, and T. H. Yeh, Rev. Mod. Phys. **88**, 015004 (2016).
[10] K. Jedamzik and M. Pospelov, New J. Phys. **11**, 105028 (2009).
[11] M. Pospelov and J. Pradler, Annu. Rev. Nucl. Part. Sci. **60**, 539 (2010).
[12] D. O'Connell, M. J. Ramsey-Musolf, and M. B. Wise, Phys. Rev. D **75**, 037701 (2007).
[13] M. Pospelov, A. Ritz, and M. B. Voloshin, Phys. Lett. B **662**, 53 (2008).
[14] B. Batell, M. Pospelov, and A. Ritz, Phys. Rev. D **83**, 054005 (2011).
[15] K. Schmidt-Hoberg, F. Staub, and M. W. Winkler, Phys. Lett. B **727**, 506 (2013).
[16] J. D. Clarke, R. Foot, and R. R. Volkas, J. High Energy Phys. 02 (2014) 123.
[17] S. Alekhin *et al.*, Rep. Prog. Phys. **79**, 124201 (2016).
[18] J. Berger, K. Jedamzik, and D. G. E. Walker, J. Cosmol. Astropart. Phys. 11 (2016) 032.
[19] V. Silveira and A. Zee, Phys. Lett. **161B**, 136 (1985).
[20] J. McDonald, Phys. Rev. D **50**, 3637 (1994).
[21] C. P. Burgess, M. Pospelov, and T. t. Veldhuis, Nucl. Phys. **B619**, 709 (2001).
[22] J. M. Cline, K. Kainulainen, P. Scott, and C. Weniger, Phys. Rev. D **88**, 055025 (2013); **92**, 039906(E) (2015).
[23] P. Athron *et al.* (GAMBIT Collaboration), Eur. Phys. J. C **77**, 568 (2017).
[24] G. Belanger, B. Dumont, U. Ellwanger, J. F. Gunion, and S. Kraml, Phys. Rev. D **88**, 075008 (2013).
[25] J. R. Ellis, M. K. Gaillard, and D. V. Nanopoulos, Nucl. Phys. **B106**, 292 (1976).
[26] J. R. Ellis, M. K. Gaillard, D. V. Nanopoulos, and C. T. Sachrajda, Phys. Lett. **83B**, 339 (1979).
[27] S. Raby and G. B. West, Phys. Rev. D **38**, 3488 (1988).
[28] T. N. Truong and R. S. Willey, Phys. Rev. D **40**, 3635 (1989).
[29] M. Drees and K. i. Hikasa, Phys. Lett. B **240**, 455 (1990); **262**, 497(E) (1991).
[30] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. **108**, 56 (1998).
[31] F. Bezrukov and D. Gorbunov, J. High Energy Phys. 05 (2010) 010.
[32] M. B. Voloshin, Yad. Fiz. **44**, 738 (1986) [Sov. J. Nucl. Phys. **44**, 478 (1986)].
[33] H. Leutwyler and M. A. Shifman, Nucl. Phys. **B343**, 369 (1990).
[34] J. F. Donoghue, J. Gasser, and H. Leutwyler, Nucl. Phys. **B343**, 341 (1990).
[35] B. Hyams *et al.*, Nucl. Phys. **B64**, 134 (1973).
[36] M. Spira, Fortschr. Phys. **46**, 203 (1998).
[37] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, Front. Phys. **80**, 1 (2000).
[38] D. McKeen, Phys. Rev. D **79**, 015007 (2009).
[39] G. Steigman, B. Dasgupta, and J. F. Beacom, Phys. Rev. D **86**, 023506 (2012).
[40] K. Griest and D. Seckel, Phys. Rev. D **43**, 3191 (1991).
[41] V. F. Mukhanov, Int. J. Theor. Phys. **43**, 669 (2004).
[42] M. Srednicki, R. Watkins, and K. A. Olive, Nucl. Phys. **B310**, 693 (1988).
[43] M. H. Reno and D. Seckel, Phys. Rev. D **37**, 3441 (1988).
[44] M. Pospelov and J. Pradler, Phys. Rev. D **82**, 103514 (2010).
[45] K. Kohri, Phys. Rev. D **64**, 043515 (2001).
[46] M. Kawasaki, K. Kohri, and T. Moroi, Phys. Rev. D **71**, 083502 (2005).
[47] M. Kawasaki, K. Kohri, and N. Sugiyama, Phys. Rev. D **62**, 023506 (2000).
[48] P. A. R. Ade *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A13 (2016).
[49] M. Kawasaki, K. Kohri, and N. Sugiyama, Phys. Rev. Lett. **82**, 4168 (1999).
[50] A. D. Dolgov, Phys. Rep. **370**, 333 (2002).
[51] M. Kawasaki, K. Kohri, and T. Moroi, Phys. Lett. B **625**, 7 (2005).
[52] R. H. Cyburt, J. R. Ellis, B. D. Fields, and K. A. Olive, Phys. Rev. D **67**, 103521 (2003).
[53] K. Jedamzik, Phys. Rev. D **74**, 103509 (2006).
[54] J. R. Ellis, K. A. Olive, and E. Vangioni, Phys. Lett. B **619**, 30 (2005).
[55] S. Dimopoulos, R. Esmailzadeh, L. J. Hall, and G. D. Starkman, Astrophys. J. **330**, 545 (1988).
[56] M. Pospelov and J. Pradler, Phys. Rev. Lett. **106**, 121305 (2011).
[57] R. Cooke, M. Pettini, R. A. Jorgenson, M. T. Murphy, and C. C. Steidel, Astrophys. J. **781**, 31 (2014).
[58] M. Lisovyi, A. Verbytskyi, and O. Zenaiev, Eur. Phys. J. C **76**, 397 (2016).
[59] R. Barate *et al.*, Eur. Phys. J. C **5**, 205 (1998).
[60] E. K. G. Sarkisyan, A. N. Mishra, R. Sahoo, and A. S. Sakharov, Phys. Rev. D **93**, 054046 (2016); **93**, 079904 (E) (2016).
[61] E. D. Carlson, M. E. Machacek, and L. J. Hall, Astrophys. J. **398**, 43 (1992).
[62] Y. Hochberg, E. Kuflik, T. Volansky, and J. G. Wacker, Phys. Rev. Lett. **113**, 171301 (2014).
[63] N. Bernal and X. Chu, J. Cosmol. Astropart. Phys. 01 (2016) 006.
[64] T. Kanzaki, M. Kawasaki, K. Kohri, and T. Moroi, Phys. Rev. D **76**, 105017 (2007).
[65] A. D. Dolgov, S. H. Hansen, and D. V. Semikoz, Nucl. Phys. **B503**, 426 (1997).