Anatomy of diluted dark matter in the minimal left-right symmetric model

Miha Nemevšek[1,2,*] and Yue Zhang[3,†]

[1]*Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia*
[2]*Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia*
[3]*Department of Physics, Carleton University, Ottawa, Ontario K1S 5B6, Canada*

(Received 16 December 2023; accepted 4 March 2024; published 26 March 2024)

Temporary matter domination and late entropy dilution, injected by a "long-lived" particle in the early Universe, serves as a standard mechanism for yielding the correct dark matter relic density. We recently pointed out the cosmological significance of diluting particle's partial decay into dark matter. When repopulated in such a way, dark matter carries higher momentum than its thermal counterpart, resulting in a suppression of the linear matter power spectrum that is constrained by the large scale structure observations. In this work, we study the impact of such constraints on the minimal left-right symmetric model that accounts for the origin of neutrino mass. We map out a systematic anatomy of possible dilution scenarios with viable parameter spaces, allowed by cosmology and various astrophysical and terrestrial constraints. We show that to accommodate the observed dark matter relic abundance the spontaneous left-right symmetry breaking scale must be above PeV and cosmology will continue to provide the most sensitive probes of it. In case the dilutor is one of the heavier right-handed neutrinos, it can be much lighter and lie near the electroweak scale.

DOI: 10.1103/PhysRevD.109.056021

I. INTRODUCTION

The field of particle physics and cosmology is facing at least three unresolved issues, driven by experiments: the nature of dark matter, the origin of neutrino mass, and the origin of the matter-anti-matter asymmetry in the Universe. They are likely linked to new fundamental laws of nature. Conceptually, it would be very appealing to have these puzzles solved within a single unified framework.

A heavy neutrino is one of the oldest, simplest and most obvious of dark matter candidates. It was first introduced as a Standard Model (SM) gauge singlet with a small mixing with the active neutrinos, produced via active-sterile neutrino oscillations in the early Universe. With improved astrophysical observations, both the Dodelson-Widrow [1] and the Shi-Fuller [2] mechanisms are already excluded [3,4]. To save such oscillation mechanisms, one must resort to novel neutrino self-interactions [5–8]. A common assumption here is a vanishing dark matter population at very early times, which can easily be affected by high-scale new physics. Right-handed neutrinos are often mandatory for gauge anomaly cancellation in many extensions of the SM [9–11]. Assuming the Universe was once sufficiently hot, new gauge interactions can then bring them into thermal equilibrium with the SM. A right-handed neutrino can be made cosmologically stable and comprise 100% of dark matter in the Universe. While it appears nearly sterile at low energies, the origin of dark matter (its abundance and momentum distribution) is governed by details of the high-scale theory.

The minimal left-right symmetric model (LRSM), based on $SU(3)_c \times SU(2)_L \times SU(2)_R \times U(1)_{B-L}$ [12–14], was originally proposed as a theory for nonzero neutrino mass [11]. In the model, parity is implemented as a left-right Z_2 symmetry, acting between left- and right-handed fermions. Since $SU(2)_R$ is gauged, three generations of right-handed neutrinos need to be present to cancel the anomalies. Parity and new gauge symmetries are spontaneously broken above the electroweak scale and light neutrino masses originate from both type-I and -II seesaw [11].

Remarkably enough, the Dirac couplings also get predicted [15], which makes the LRSM a particularly complete and predictive theory for neutrino masses. The breaking of lepton number manifests itself in a number of processes, ranging from direct production at high energy colliders [16,17], neutrinoless double beta decay [18,19], lepton flavour violation [20,21], and cosmology [22,23]. Moreover, the model can explain the origin of cosmic baryon asymmetry via leptogenesis, as long as the LR scale is sufficiently high [24,25].

*miha.nemevsek@ijs.si
†yzhang@physics.carleton.ca

The lightest right-handed neutrino (called N_1 hereafter) as a viable dark matter candidate in the minimal LRSM was first considered by Bezrukov, Hettmansperger, and Lindner [26]. It features a dark matter candidate with a light mass below ~MeV, in order to be cosmologically stable. On the other hand, its mass is bounded from below by several keV scale from phase space packing in dwarf galaxies, because N_1 is a fermion and subject to Pauli blocking [27,28]. Slightly stronger mass lower bound applies if N_1 had thermal contact with the SM plasma in the early Universe (warm dark matter) [3,29–34]. The right-handed current gauge interaction, mediated by the W_R, decouples in the early Universe and always leaves N_1 freezing out ultrarelativistically. This overproduces the dark matter relic abundance, unless there is a "long-lived" matter component that temporarily dominates the energy density of the Universe before decaying dominantly into the SM [35]. The late entropy release causes a relative dilution of the final dark matter abundance and brings it down to the observed value. In [26], it was suggested that the role of the diluting particle can be taken on within the LRSM by one of the heavier right-handed neutrinos. The corresponding W_R boson mass scale is typically constrained to be rather high to facilitate the relativistic freeze-out and longevity of the diluting particle. The possibility of having a lighter W_R boson was investigated in [36], which resorts to a decaying phase space suppression to keep the lifetime of the diluting particle sufficiently long.

Recently, the dark matter dilution mechanism was revisited and a new, model-independent constraint has been discovered [37]. This new opportunity for testing dark matter dynamics lies in the partial decay of the dilutor into dark matter. Such decay modes exist quite generically, either at tree or loop level, and the branching ratio is sometimes fixed by the internal structure of a UV complete model. With dark matter repopulated this way, the relic density obtains a secondary component on top of the primary one that comes from the usual freeze-out. Most importantly, this component is predicted to be much more energetic than the original thermal one.

Under a reasonable assumption that both the dark matter and the dilutor freeze-out relativistically,[1] the secondary dark matter particles stay relativistic until the temperature of the Universe cools down to around eV scale. This temperature is nearly independent of parameters including the dark matter mass and the dilutor's mass and lifetime. As a result, dark matter free-streaming strongly impacts the matter power spectrum and the formation of large scale structures. Using the existing data from the Sloan Digital Sky Survey (SDSS), an upper bound on the branching ratio for dilutor into dark matter is set at about $\lesssim 1\%$. Because the primordial perturbations remain linear on large scales, a robust cosmological constraint can be applied on the fundamental theories for the origin of dark matter.

In the context of LRSM, the decay of a heavy right-handed neutrino to the lighter one could occur via the exchange of the W_R gauge boson, similar to weak decays in the SM. If this is the dominant mode, then the branching ratio is predicted by the number of lighter fermions and universality of the $SU(2)_R$ gauge interactions and comes out to be larger than 10%, which turns out to be forbidden by the large scale structure observations [37]. To mitigate this exclusion, the right-handed neutrino dilutor must have other significant decay channels to reduce the dark matter repopulation. Such an important constraint has been ignored in previous analysis. We are therefore strongly motivated to revisit the viability of right-handed neutrino dark matter in the LRSM and, as we shall see, they strongly affect the allowed parameter space where an appropriate dark matter relic density can be obtained. In performing a systematic and thorough analysis, we first focus on the usual dilutor in the form of another right-handed neutrino, and then identify a new candidate from the scalar sector of the LRSM that can also play the role of dilution.

Before our journey begins, we would like to stress that the entropy dilution explored here in the context of LRSM is a generic new physics scenario for fixing the dark matter relic density, or for suppressing the amount of extra radiation ($\Delta N_{\rm eff}$) in the early Universe. It has been employed in a broad range of dark matter models including the gravitino, twin-Higgs models, and various dark sector incarnations [38–55]. The same large scale structure constraint would also affect the viability of dark matter in these models and must be taken into account in future studies.

This article is organized as follows. In Sec. II, we give a lightning review of the minimal LRSM and highlight several aspects of the model that are important for the dark matter study in this work. In Sec. III, we discuss the entropy dilution mechanism that features temporary matter domination by a "long-lived" diluting particle. We go beyond the earlier work [37] and provide a detailed derivation of the Boltzmann equation and distribution function of secondary dark matter component from dilutor's decay. The discussion in this section is model independent and easily applicable to other models, besides the LRSM, that resort to a similar dilution mechanism. In Sec. IV, we present the anatomy of right-handed neutrino dark matter in the minimal LRSM by exhausting all the possible dilution scenarios that we have envisioned. This includes the heavier right-handed neutrino or the Higgs boson counterpart of left-right symmetry breaking (the right-handed triplet) playing the role of dilutor. For each of the cases, we establish the viable parameter space for DM, compatible with the large scale structure limits from SDSS, along with other constraints that include the correct relic

[1]We will comment on what happens to the constraint from dilution when this assumption is relaxed. In the LRSM, relativistic freeze-out is always valid, regardless of which particle plays the role of the dilutor.

density, generation of neutrino mass, big-bang nucleosynthesis and x-ray line searches. We also comment on supernova cooling and existing laboratory constraints on the LR scale. We conclude and provide an outlook of opportunities in Sec. V.

II. DARK MATTER IN THE MINIMAL LEFT-RIGHT SYMMETRIC MODEL

We start with a brief overview of the structure of the minimal LRSM and highlight several ingredients that are important for understanding the cosmology in the later sections. We refer to [12–14] for the original works and in-depth reviews of the model.

A. The minimal left-right symmetric model

The LRSM is based on the gauge group $G_{\rm LR} = SU(3)_c \times SU(2)_L \times SU(2)_R \times U(1)_{B-L}$, with a discrete Z_2 symmetry interchanging the left and right $SU(2)$ sectors. At low energies, the Z_2 symmetry manifests itself as the parity symmetry of QCD and QED. Quarks and leptons come in parity symmetric representations

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} = \left(3, 2, 1, \frac{1}{3}\right),$$
$$Q_R = \begin{pmatrix} u_R \\ d_R \end{pmatrix} = \left(3, 1, 2, \frac{1}{3}\right),$$
$$L_L = \begin{pmatrix} \nu \\ \ell_L \end{pmatrix} = (1, 2, 1, -1),$$
$$L_R = \begin{pmatrix} N \\ \ell_R \end{pmatrix} = (1, 2, 1, -1), \tag{2.1}$$

where N stands for the right-handed neutrino. The scalar potential of the minimal model is also parity symmetric. It consists of three complex fields: a bidoublet $\Phi = (1, 2, 2, 0)$ and two triplets $\Delta_L = (1, 3, 1, 2)$ and $\Delta_R = (1, 1, 3, 2)$ under $G_{\rm LR}$, with the following field assignments:

$$\Phi = \begin{pmatrix} \phi_1^0 & \phi_2^+ \\ \phi_1^- & \phi_2^0 \end{pmatrix}, \quad \Delta_{L,R} = \begin{pmatrix} \delta^+/\sqrt{2} & \delta^{++} \\ \delta^0 & -\delta^+/\sqrt{2} \end{pmatrix}_{L,R}. \tag{2.2}$$

Under parity, $\Phi \to \Phi^\dagger$, $\Delta_L \leftrightarrow \Delta_R$. Starting from the LR and parity symmetric potential, it was shown that parity is broken spontaneously [13,14], with either with two doublets [14] or two triplets [11,18]. The complete form of the potential was discussed in [56] and studied in some depth in subsequent years [57–61] with more recent works focusing on phenomenological signals [62–64]. A strong lower bound from perturbativity of the potential was worked out in [65], with constraints from vacuum the vacuum structure [66] and opportunities for gravitational waves [67].

The above quantum numbers allow for the following Yukawa terms that couple the fermions to scalars

$$\mathcal{L}_{\rm Yuk} = \bar{Q}_L (Y_q \Phi + \tilde{Y}_q \tilde{\Phi}) Q_R + \bar{L}_L (Y_l \Phi + \tilde{Y}_l \tilde{\Phi}) L_R + Y_{\Delta_L} L_L^T i\sigma_2 \Delta_L L_L + Y_{\Delta_R} L_R^T i\sigma_2 \Delta_R L_R + \text{H.c.}, \tag{2.3}$$

where $\tilde{\Phi} = i\sigma_2 \Phi^* i\sigma_2$ and the family indices are suppressed. The σ_2 matrices operate within the two $SU(2)_{L,R}$ group spaces and ensure gauge invariance. The first two terms are of the Dirac type and give the usual mass terms that connects the left and right chiral fields, as in the SM. An important component of the LRSM are the Dirac mass term for neutrinos

$$M_D = \frac{v}{\sqrt{2}} (\cos\beta Y_l + \sin\beta e^{-i\alpha} \tilde{Y}_l), \tag{2.4}$$

and the Majorana-type Yukawa couplings $Y_{\Delta_{L,R}}$ in the second line, that generate lepton number violating masses for neutrinos.

The spontaneous symmetry breaking occurs in two steps. First, the $SU(2)_R \times U(1)_{B-L}$ symmetry is broken down to $U(1)_Y$ for hypercharge by the vacuum expectation value (VEV) of the right-handed scalar triplet $\langle \Delta_R^0 \rangle = v_R/\sqrt{2}$, which lies well above the electroweak scale. It generates masses for the new gauge bosons $W_R^\pm$ and Z', with a mass relation $M_{Z'} = \sqrt{3} M_{W_R}$. Through the Y_{Δ_R} Yukawa coupling term in Eq. (2.3), the v_R condensate also gives a Majorana mass to the right-handed neutrinos N.

The second stage of spontaneous symmetry breaking is triggered by the VEV of the bidoublet scalar, $\langle \phi_1^0 \rangle = v\cos\beta/\sqrt{2}$ and $\langle \phi_2^0 \rangle = v\sin\beta e^{i\alpha}/\sqrt{2}$, where $v = 246$ GeV. Without loss of generality, we choose $\beta \in (0, \pi/2)$. In view of $SU(2)_L$, the bidoublet effectively behaves as two Higgs doublets, thus their VEVs can give masses to the regular $W^\pm$ and Z gauge bosons that mediate the weak interactions. Through the Yukawa coupling in Eq. (2.3), the electroweak VEVs also provide Dirac mass matrices for all the fermions, including the one between left-handed neutrinos ν and right-handed ones N.

B. Essential model ingredients for the early Universe

1. Right-handed currents

The most important ingredient of the LRSM, relevant for dark matter cosmology, are the new gauge interactions, mediated by the $W_R^\pm$ and Z' gauge bosons. The right-handed charged-current interactions for fermions take on the form

$$\mathcal{L}_{\rm gauge} = \frac{g}{\sqrt{2}} W_R^\mu (\bar{N} \gamma_\mu V_{\rm PMNS}^{R\dagger} \ell_R + \bar{u}_R \gamma_\mu V_{\rm CKM}^R d_R) + \text{H.c.}, \tag{2.5}$$

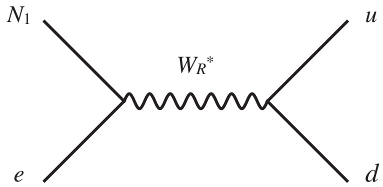


FIG. 1. Feynman diagram for right-handed charged-current interaction of dark matter N_1 in LRSM.

where all the fermions fields are now in their mass eigenstates. We have introduced the right-handed Cabibbo–Kobayashi–Maskawa (CKM) and Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrices that transform the fermions from the weak flavor into the mass basis. Here we neglected the mixing between the W and W_R gauge bosons, which is constrained to be small (see Sec. II B 4) and its effects in dark matter pehenomenology will be accounted for in Sec. IV A 4 and the Appendix.

In the minimal LRSM with parity symmetry, the right-handed CKM matrix has been solved both numerically [68–70] and analytically [71,72]. Its off-diagonal elements are suppressed by similar powers of the Wolfenstein parameter as the regular CKM matrix, but all the matrix elements carry extra phase factors with implications for the strong CP violation [70,73–75]. For the purposes of this work, the exact form of the quark charged currents is not really important and we shall approximate $V^R_{\rm CKM}$ with a unit matrix. The situation is quite different for the leptons. There, the form of right-handed PMNS matrix is much less constrained and allowed to be wildly different from the regular PMNS matrix. The only special case is when the light neutrino mass contribution is dominated by the type-II seesaw contribution. In this case, parity requires that $V^R_{\rm PMNS}$ be identical to its left-handed counterpart.

2. Lightest right-handed neutrino as dark matter

The minimal LRSM does not preserve any exact Z_2 symmetry that would stabilize DM, even LR parity gets broken spontaneously and its quality is not crucial for dark matter stability. As a result, none of the new particles beyond the SM are absolutely stable. As argued in [36], the only candidate for dark matter in the model is the lightest right-handed neutrino, N_1.[2] Without fine-tuning the flavor structure of $V^R_{\rm PMNS}$, the Feynman diagram in Fig. 1 shows the dominant interaction for N_1 to couple with the charged leptons and quarks.

If heavy enough, then N_1 could decay into an electron and a charged pion, and it would not be cosmologically stable, unless the W_R boson is ultraheavy, near the grand unified theories (GUT) scale.[3] In this work, we are interested in finding lower (and upper) bounds of W_R, and we will focus on the N_1 dark matter mass below $\sim$100 MeV, where there is a wide portion of parameter space for N_1 to be sufficiently stable. To be more specific, the upper bound on N_1 mass can be written as

$$\begin{aligned} m_{N_1} &< \max\left(\left(\frac{96\pi^3}{\tau G_F^2}\right)^{1/5}\left(\frac{M_{W_R}}{M_W}\right)^{4/5}, m_\pi\right) \\ &\simeq m_\pi \times \max\left(\left(\frac{M_{W_R}}{10^{10}\ \text{GeV}}\right)^{4/5}, 1\right), \end{aligned} \tag{2.6}$$

where we require dark matter decay rate $\tau^{-1} \lesssim 10^{-50}$ GeV, which corresponds to a typical bound from cosmic ray positrons and x-ray searches. As for laboratory constraints, the mass of the W_R boson is chiefly bounded by searches at the LHC. For right-handed (RH) neutrinos below a few 10 of GeV, the signal from $W_R \to \ell N_1$ looks like a very energetic charged lepton and missing energy, because N_1 escapes detection. Recasting the generic $W' \to \ell \not{E}$ searches [80], the current bounds set the LR scale to be $M_{W_R} \gtrsim 5$ TeV, depending on the flavor of the charged lepton. This channel is particularly clean and may probe the scales up to 37 TeV at a future 100 TeV collider [81].

With such super-weak interactions, mediated by the Feynman diagram in Fig. 1, the RH neutrinos thermalize in the early Universe and decouple at temperatures

$$T_d \sim 1\ \text{MeV}\left(\frac{M_{W_R}}{M_W}\right)^{4/3} \gtrsim 300\ \text{MeV}, \tag{2.7}$$

where in the second step we apply the existing lower bound on M_{W_R} from the LHC. Comparing with Eq. (2.6), we find that

$$T_d > m_{N_1} \tag{2.8}$$

always holds. We will assume that the reheating temperature of the early Universe was sufficiently high, such that all the RH neutrinos (and other particles in the LRSM) were once kept in thermal equilibrium by gauge interactions. The above comparison implies that N_1 must decouple when it was still ultra relativistic, similar to the decoupling of SM neutrinos. As will be sharpened in Sec. III A, this leads to a severe dark matter overproduction problem and requires a nonstandard cosmology after the freeze out.

3. Neutrino mass contributions

Neutrino masses in the minimal LRSM come from two sources. First, the Dirac neutrino mass term together with

[2]The real part of Δ_R^0 in (2.2) may also be cosmologically stable if it is made to be much lighter than M_{W_R}. This brings in issues with vacuum stability, similar to the case of the SM [76,77], and may require fine-tuning of couplings in the LRSM [65,78,79], therefore we do not pursue this option any further.

[3]One may consider the PMNS coupling of N_1 to τ only, which would prevent the tree-level decay to $e\pi$ and allow for slightly heavier N_1.

the Majorana mass for N, generated by spontaneous symmetry breaking, can give mass to the active neutrinos through the type-I seesaw mechanism. In addition, the model features another contribution through the type-II seesaw. It comes from the VEV of the left-handed scalar triplet $\langle \Delta_L^0 \rangle = v_L/\sqrt{2}$ and the λ Yukawa coupling term in Eq. (2.3). The v_L condensate originates from terms in the scalar potential of the form, e.g., $\mathrm{Tr}(\Delta_L \Phi \Delta_R \Phi^\dagger)$, which is a tadpole term for Δ_L after Φ and Δ_R have obtained their condensates.

The full neutrino mass matrix in the model is then given by

$$\begin{aligned} M_\nu &= -M_D^T M_N^{-1} M_D + M_L, \\ M_D &= \frac{v}{\sqrt{2}}(\cos\beta Y_l + \sin\beta e^{-i\alpha} \tilde{Y}_l), \end{aligned} \tag{2.9}$$

$$M_N = \frac{v_R}{\sqrt{2}} Y_{\Delta_R}, \qquad M_L = \frac{v_L}{\sqrt{2}} Y_{\Delta_L}. \tag{2.10}$$

Accommodating the neutrino masses and mixings, needed to explain neutrino masses and fit the neutrino oscillations, is one of the primary motivations for considering the LRSM as a plausible beyond the Standard Model theory. We will use it as an important guiding principle when exploring cosmological aspects of the model.

Within the LRSM, the Dirac Yukawa couplings can be solved for in terms of the Majorana mass matrix in case of $\mathcal{C}$ [15] and $\mathcal{P}$ [82–84]. Thus one can predict the M_D mass matrix from m_N and V_R and calculate the heavy-light neutrino mixing. Note that the charged lepton mass matrix M_l comes from a different linear combination of Y_l and $\tilde{Y}_l$ and is therefore independent of M_D. This renders the mass scale of right-handed neutrinos free parameters in the LRSM. In this work we do not concern ourselves much with fine-tunings of Y_l and $\tilde{Y}_l$ but take a phenomenological approach by scanning the entire parameter space and figuring out which parts can accommodate the dark matter relic abundance.

4. $W - W_R$ gauge boson mixing

Because the scalar bidoublet Φ transforms under both $SU(2)_L$ and $SU(2)_R$, its VEVs allow for a mixing between the W and W_R gauge bosons. The corresponding mass terms take on the form

$$\mathcal{L}_W = -(A^-_{L\mu} \quad A^-_{R\mu}) \frac{g^2}{2} \begin{pmatrix} \frac{1}{2}(v^2 + 2v_L^2) & -v^2 \sin\beta\cos\beta e^{-i\alpha} \\ -v^2 \sin\beta\cos\beta e^{i\alpha} & \frac{1}{2}(v^2 + 2v_R^2) \end{pmatrix} \begin{pmatrix} A_L^{\mu+} \\ A_R^{\mu+} \end{pmatrix}, \tag{2.11}$$

where $A^\pm_{L,R}$ are the gauge bosons in the flavor basis. After diagonalization, the mass eigenstates are the SM-like W boson, which is a linear superposition of mostly A_L and a small admixture of A_R that mediates RH currents. Vice versa, the W_R mass eigenstate is mostly right handed. Phenomenologically, the relevant scales in the LRSM need to hierarchical, such that $v_R \gg v \gg v_L$ and the $W - W_R$ mixing is approximated by

$$\xi_{\mathrm{LR}} \simeq \sin\beta\cos\beta e^{i\alpha}\left(\frac{v}{v_R}\right)^2 \simeq \sin(2\beta) e^{i\alpha}\left(\frac{M_W}{M_{W_R}}\right)^2, \tag{2.12}$$

where in the last step we used $M_W^2 \simeq g^2 v^2/4$ and $M_{W_R}^2 \simeq g^2 v_R^2/2$. As we will see, ξ_{LR} is one of the key LRSM parameters for resolving the dark matter repopulation issue.

Clearly ξ_{LR} is suppressed due to the small mass ratio $(M_W/M_{W_R})^2$. The magnitude of ξ_{LR} also depends on $\tan\beta$, i.e., the ratio of the two VEVs from the bidoublet for which there exist an upper bound from the perturbativity of Yukawa couplings. To make the point, we consider the quark mass generation in the minimal LRSM here and explain the logic by approximating with the third family only, which has the largest Yukawa couplings. For a single generation, the top and bottom quark masses are given by

$$\begin{aligned} m_t &= \frac{v}{\sqrt{2}}(Y_q \cos\beta + \tilde{Y}_q \sin\beta e^{-i\alpha}), \\ m_b &= \frac{v}{\sqrt{2}}(Y_q \sin\beta e^{i\alpha} + \tilde{Y}_q \cos\beta). \end{aligned} \tag{2.13}$$

These can be inverted and solved for the Yukawa couplings

$$\begin{aligned} Y_q e^{i\alpha} &= \frac{\sqrt{2}}{v\cos 2\beta}(m_t e^{i\alpha}\cos\beta - m_b \sin\beta), \\ \tilde{Y}_q &= \frac{\sqrt{2}}{v\cos 2\beta}(-m_t e^{i\alpha}\sin\beta + m_b\cos\beta). \end{aligned} \tag{2.14}$$

By requiring $|Y_q|$ and $|\tilde{Y}_q|$ to take on perturbative values ($\lesssim 3$) for $\alpha \in [0, 2\pi)$, we get the following allowed range for $\tan\beta$

$$\tan\beta < -1.24 \cup -0.77 < \tan\beta < 0.77 \cup \tan\beta > 1.24. \tag{2.15}$$

In other words, the large difference between the top and bottom quark masses forbids the angle β to be close to $\pm\pi/4$, where $\cos 2\beta$ approaches to zero blowing up Y_q

and $\tilde{Y}_q$. See [70] for more details and a full numerical study with three generations.

5. The Majorana Higgs

As mentioned earlier, the gauge symmetry breaking from LRSM to the SM is triggered by the VEV of the $SU(2)_R$ scalar triplet Δ_R. Its components $\delta_R^\pm$, and the imaginary part of δ_R^0, become the longitudinal components of the $W_R^\pm$ and Z' bosons, respectively. The real part of δ_R^0 is the "Higgs boson" for this step of symmetry breaking, a massive propagating particle that reveals the nature of spontaneous breaking. We denote the properly normalized physical state as Δ, where

$$\Delta \equiv \sqrt{2}\delta_R^0 - v_R. \tag{2.16}$$

Because v_R violates lepton number and serves as the source of Majorana neutrino mass, Δ is referred to as the Majorana Higgs [79]. The doubly charged component Δ_R^{++} is left over as another physical state.

As will be discussed Sec. IV B, Δ can also play a crucial role of dilution for addressing the dark matter relic density. Here, we list its interactions that are relevant for understanding its role in the early Universe. The Δ is the excitation above the VEV v_R, which is mostly responsible for the mass generation for the right-handed neutrinos N and $W_R^\pm$, Z' gauge bosons. The corresponding couplings can be derived by shifting $v_R \to v_R(1+\Delta/v_R)$,

$$\mathcal{L} = -\frac{m_N}{v_R}\bar{N}N\Delta + 2\frac{M_{W_R}^2}{v_R}W_{R\mu}^+ W_R^{-\mu}\Delta + \frac{M_{Z'}^2}{v_R}Z'_\mu Z'^\mu\Delta, \tag{2.17}$$

where we keep the interaction terms linear in Δ, which are useful for calculating its decay rates. The $\Delta - N$ coupling is diagonal in the mass basis of N.

Another important parameter that controls the decay rates of Δ is its mixing with the Higgs boson, $\theta_{\Delta h}$. In the presence of such mixing, Δ can decay into all the SM particles that the Higgs boson couples to. In particular, when the mass of Δ is much above the electroweak scale, it mainly decays into W^+W^-, ZZ, and hh, with a ratio of 2:1:1, as dictated by the equivalence principle. The scalar potential terms that couple Δ to Higgs are $\alpha_1 \mathrm{Tr}(\Phi^\dagger\Phi)\mathrm{Tr}(\Delta_R^\dagger\Delta_R) + [\alpha_2 \mathrm{Tr}(\Phi^\dagger\tilde{\Phi})\mathrm{Tr}(\Delta_R^\dagger\Delta_R) + \mathrm{H.c.}] + \alpha_3 \mathrm{Tr}(\Phi^\dagger\Phi\Delta_R^\dagger\Delta_R)$ [59]. After the right-handed triplet Δ_R develops the VEV, but before the electroweak symmetry breaking, these terms allow Δ to decay into a pair of SM Higgs bosons, as well as the wouldbe Goldstone bosons that eventually become the longitudinal components of the W and Z bosons.

III. GENERAL DILUTION MECHANISM VS LARGE SCALE STRUCTURE

We review the dark matter dilution mechanism under the sudden decay approximation [35], which is a useful tool for exploring the late decay of long-lived particles in the early Universe. This approximation allows us to analytically derive the important parametrical dependence in relevant quantities, such as the final dark matter relic density Ω_X, and the reheating temperature T_{RH} immediately after the decay of the dilutor. To keep the discussion here as general as possible, we call here the dark matter particle X and the cosmologically "long-lived" particle for entropy dilution Y. We will assign their identities within the LRSM ($X \to N_1$, $Y \to N_2, \Delta$) in the next section, when we discuss the concrete dark matter dilution scenarios.

A. Relativistic freeze out and overproduction problem

Consider the dark matter X, which freezes out from the SM thermal plasma relativistically. The yield $Y_X = n_X/s$ is then defined as the ratio of number density n_X to the total entropy density s of the SM plasma, and is given by

$$Y_X = \frac{135\zeta(3)}{4\pi^4 g_*(T_{\mathrm{fo}})}, \tag{3.1}$$

where we assumed that X is a Majorana fermion with two degrees of freedom. The T_{fo} is the photon temperature when X freezes out, and $g_*(T_{\mathrm{fo}})$ counts the corresponding number of relativistic degrees of freedom in the Universe in the plasma. Because most of our discussion will be restricted to temperatures above the MeV scale for successful big bang nucleosynthesis (BBN), we will not distinguish $g_*(T)$ and $g_{*S}(T)$ hereafter. If nothing else happened after the freeze out, then Y_X would be a conserved quantity, and the dark matter relic density today would be

$$\Omega_X^0 = \frac{m_X Y_X s_0}{\rho_0} \simeq 2.6\left(\frac{m_X}{1\ \mathrm{keV}}\right)\left(\frac{100}{g_*(T_{\mathrm{fo}})}\right), \tag{3.2}$$

where $s_0 = 2891.2\ \mathrm{cm}^{-3}$ is the entropy density in the Universe today, and $\rho_0 = 1.05\times10^{-5}h^2\ \mathrm{GeV/cm^3}$ represents today's critical density with $h = 0.67$ [85]. In contrast, the *Planck* experiment observes that the value of $\Omega_{\mathrm{dark\ matter}}$ is 0.26 [86]. Because m_X is constrained to be heavier than several keV due to various warm dark matter constraints, the above result creates the dark matter overproduction problem.

B. Entropy dilution mechanism with a long-lived particle

To address the issue of overproduction, we introduce a dilutor particle Y. For simplicity, we assume it is also a Majorana fermion that freezes out relativistically and has a similar yield as the dark matter before decaying away, mostly into the SM particles. To achieve sufficient dilution, Y must dominate the total energy density of the Universe (as matter) before it decays away and dumps most of its energy (or entropy) into the SM sector.

In the sudden decay approximation, we have

$$\tau_Y^{-1} = H_{\text{before}} = H_{\text{after}}, \tag{3.3}$$

where τ_Y is the lifetime of Y, and $H_{\text{before, after}}$ are the Hubble parameters ($H \equiv \sqrt{8\pi G_N \rho/3}$) immediately before and after the decay, respectively.

Before the decay of Y, the Universe is dominated by the energy density of nonrelativistic massive Y particles,

$$\rho = \rho_Y = Y_Y s_{\text{before}} m_Y, \tag{3.4}$$

where s_{before} is the total entropy density of relativistic species before the decay. Here we assume that Y experiences a similar relativistic freeze out as dark matter and Y_Y is the same as Y_X given in Eq. (3.1). Note that the Universe is already matter dominated right before Y decays. However, this does not prevent us from defining Y as the ratio of n_Y to the relativistic entropy density s, and $Y_Y = n_Y/s$ remains conserved in the time window between the freeze-out and Y's decay. With these inputs, the first of Eq. (3.3) leads to

$$s_{\text{before}} = \frac{\pi^3 g_*(T_{\text{fo}})}{90\zeta(3)} \frac{M_{\text{pl}}^2}{m_Y \tau_Y^2}. \tag{3.5}$$

Assuming that the decay of Y takes no time, the energy density of Y immediately before its decay is equal to the radiation energy density immediately after. The latter is related to the corresponding "reheating" temperature T_{RH} of the SM plasma,

$$\rho = \rho_R = \frac{\pi^2}{30} g_*(T_{\text{RH}}) T_{\text{RH}}^4. \tag{3.6}$$

The second equation of Eq. (3.3), $\tau_Y^{-1} = H_{\text{after}}$, leads to

$$T_{\text{RH}} \simeq 0.6 g_*(T_{\text{RH}})^{-1/4} \sqrt{\frac{M_{\text{pl}}}{\tau_Y}} \simeq \frac{0.93\,\text{MeV}}{g_*(T_{\text{RH}})^{1/4}} \sqrt{\frac{1\ \text{sec}}{\tau_Y}}, \tag{3.7}$$

where $M_{\text{pl}} = \sqrt{1/G_N} = 1.2 \times 10^{19}$ GeV is the Planck constant. The entropy density of the SM plasma immediately after Y decay can then be calculated in terms of T_{RH},

$$s_{\text{after}} = \frac{2\pi^2}{45} g_*(T_{\text{RH}}) T_{\text{RH}}^3. \tag{3.8}$$

With Eqs. (3.5) and (3.8) we can derive the dilution factor $\mathcal{S}$,

$$\mathcal{S} \equiv \frac{s_{\text{after}}}{s_{\text{before}}} \simeq \frac{0.7 g_*(T_{\text{RH}})^{1/4}}{g_*(T_{\text{fo}})} \frac{m_Y \sqrt{\tau_Y}}{\sqrt{M_{\text{pl}}}}. \tag{3.9}$$

The diluted relic density of X today is given by

$$\Omega_X = \frac{\Omega_X^0}{\mathcal{S}} \simeq \frac{0.72 g_*(T_{\text{RH}})^{1/4}}{g_*(T_{\text{fo}})} \frac{m_Y \sqrt{\tau_Y}}{\sqrt{M_{\text{pl}}}}$$
$$\simeq 0.26 \left(\frac{m_X}{1\ \text{keV}}\right) \left(\frac{2.2\ \text{GeV}}{m_Y}\right) \sqrt{\frac{1\ \text{sec}}{\tau_Y}}. \tag{3.10}$$

This is the standard dark matter dilution mechanism that has been employed in various contexts for addressing the dark matter relic density.

C. Dilutor to dark matter decay

We recently pointed out [37] new opportunities to test the dark matter dilution mechanism. We showed that the repopulation of dark matter in Y decays leaves an imprint on structure formation and gets constrained by the existing data (or gives a characteristic signal upcoming data). While in [37] we worked in a largely model independent way, in a concrete UV realization of such mechanisms, the model predicts not only the relic density, but also the spectrum—the phase space distribution of dark matter with a primary and secondary component. To account for all these possibilities, we consider the following decay channels of Y, where it can decay into SM particles as well as dark matter X,

$$Y \to SM, \qquad Y \to nX(+mSM), \tag{3.11}$$

where $n, m \in \mathbb{Z}^+$ count the multiplicity of X and SM particles, respectively, in the final states. The first decay channel is desired for dumping entropy into the visible sector and dilutes the primordial thermal population of X. If this were the only final state of Y decay, then the resulting X would remain a purely thermal distribution with a temperature T_X, which would be relatively lower than the counterpart in the absence of dilution.

The other decay mode, whose branching ratio is assumed to be Br_X, produces a secondary nonthermal population of X that also contributes to the final dark matter relic abundance. The bracket in (3.11) also includes the possibility that this second decay channel is completely dark, without any "SM" in the final state. In the absence of extended dark sectors, the branching ratio of the first channel is simply $1 - \text{Br}_X$. With a nonzero Br_X, the final dark matter relic density becomes

$$\Omega_X \simeq 0.26(1 + n\text{Br}_X) \left(\frac{m_X}{1\ \text{keV}}\right) \left(\frac{2.2\ \text{GeV}}{m_Y}\right) \sqrt{\frac{1\ \text{sec}}{\tau_Y}}. \tag{3.12}$$

Requiring Ω_X to agree with the *Planck* measured value fixes τ_Y in terms of m_X and m_Y, we can rewrite Eq. (3.7) as

$$T_{\rm RH} \simeq \frac{0.4 \text{ MeV}}{g_*(T_{\rm RH})^{1/4}} \frac{m_Y}{10^6 m_X}. \tag{3.13}$$

Because g_* is always larger than 1, and for successful big bang nucleosynthesis to work, for which $T_{\rm RH} \gtrsim 1$ MeV is needed, the dilutor Y must be heavier than dark matter X by a factor of at least 10^6.

Before deriving and solving the equation for dark matter phase space distribution, we first give a qualitative discussion and introduce an important temperature relevant for large scale structure of the Universe. Immediately after the $Y \to nX + mSM$ decay, each secondary X particle roughly carries the energy of $m_Y/(m+n)$. Under the sudden decay approximation, the corresponding temperature of the Universe is given by $T_{\rm RH}$ in Eq. (3.7). The velocity of X particles will then redshift with the expansion of the Universe. After a while the X particles start to turn nonrelativistic when the energy drops to around their mass. This requires the scale factor of the Universe to grow by a factor of

$$\frac{a_{\rm NR}}{a_{\rm RH}} \simeq \frac{m_Y}{(m+n)m_X}. \tag{3.14}$$

The corresponding temperature $T_{\rm NR}$ can be found with entropy conservation in the SM sector

$$g_{*S}(T_{\rm NR}) T_{\rm NR}^3 a_{\rm NR}^3 = g_*(T_{\rm RH}) T_{\rm RH}^3 a_{\rm RH}^3, \tag{3.15}$$

which leads to

$$T_{\rm NR} = T_{\rm RH} \left(\frac{g_*(T_{\rm RH})}{g_*(T_{\rm NR})} \right)^{1/3} \frac{a_{\rm RH}}{a_{\rm NR}} \simeq 0.25 \text{ eV} n g_*(T_{\rm RH})^{\frac{1}{12}}. \tag{3.16}$$

In the second step we used Eq. (3.13) and the late-time value for $g_{*S}(T_{\rm NR}) = 3.91$, valid for $T_{\rm NR}$ well below the electron mass.

Note that $T_{\rm NR}$ defines the time when both the primordial X particles and the secondary ones from Y decay have become matterlike. Dialing the clock back to temperatures above $T_{\rm NR}$, the dark matter fluid is made out of the nonrelativistic primordial and the relativistic secondary component. The energy density of the latter is more important at temperatures above $T_{\rm NR}/\mathrm{Br}_X$. In this regime, the overall X fluid is relativistic and features a large pressure, which can interrupt the regular logarithmic growth of matter density perturbations in X. This suppresses the matter power spectrum $P(k)$ for wavelengths of the perturbation smaller than the Hubble radius at temperature equal to $T_{\rm NR}/\mathrm{Br}_X$. The resulting $P(k)$ may then potentially disagree with the large scale structure (LSS) measurements, unless $\mathrm{Br}_X \ll 1$.

D. Phase space distribution of dark matter

Let us go beyond the sudden decay approximation and derive the equations governing the dark matter phase space distribution. Because the temperature $T_{\rm NR}$, when the secondary dark matter particles from dilutor decay turn nonrelativistic, is found to be rather low, they act as a hot dark matter component for a period of time that overlaps with the observational data. Consequently, they may suppress the large and small scale structures, which in turn allows one to derive a powerful constraint on the dilutor → dark matter branching ratio using cosmological data from the SDSS [37].

Without loss of generality, we write the dilutor to dark matter decay channel as

$$Y \to X + 2 + 3 + \cdots + N = nX + m\mathrm{SM}, \tag{3.17}$$

where particles $2, 3, \ldots, N = m + n$ represent either the SM or additional X particles in the final state. As before, we assume that nX particles are produced in this decay. The phase space distribution for the produced X is governed by the Liouville's equation that relates the phase space distribution functions of dark matter f_X and the dilutor f_Y

$$\left(\frac{\partial}{\partial t} - H \frac{|\vec{p}_X|^2}{E_X} \frac{\partial}{\partial E_X} \right) f_X(E_X, t) = \int \frac{\mathrm{d}^3 \vec{p}_Y}{(2\pi)^3} \frac{1}{2E_Y} f_Y(E_Y, t) A, \tag{3.18}$$

$$A = \frac{1}{2E_X} \prod_{i=1}^{N} \int \frac{\mathrm{d}^3 \vec{p}_i}{(2\pi)^3} \frac{1}{2E_i} (2\pi)^4 \delta^4 \left(p_Y - p_X - \sum_i p_i \right) |\mathcal{M}|^2. \tag{3.19}$$

Hereafter we work in the ultrarelativistic X limit (from Y decay) and approximate $|\vec{p}_X| \simeq E_X$. $\mathcal{M}$ is the decay matrix element related to the decay in (3.17) and A can be rewritten in terms of the partial decay rate of $Y \to X + \cdots + N$ in its rest frame as

$$\Gamma_{Y\to X} = \frac{1}{2m_Y} \int \frac{\mathrm{d}^3 \vec{p}_X}{(2\pi)^3} A = \frac{1}{4\pi^2 m_Y} \int \mathrm{d}E_X E_X^2 A, \tag{3.20}$$

where it is assumed that the X particles from Y decay are ultrarelativistic throughout most of the phase space. Here it is useful to introduce a dimensionless spectral function $g(\omega)$, which satisfies

$$n \frac{\mathrm{d}\Gamma_{Y\to X}}{\mathrm{d}\omega} = \Gamma_{Y\to X} g(\omega), \quad \int \mathrm{d}\omega g(\omega) = n, \quad \omega = \frac{E_X}{m_Y}. \tag{3.21}$$

It allows us to establish a relation between A and g and simplify the Liouville's equation to

$$\left(\frac{\partial}{\partial t} - HE_X \frac{\partial}{\partial E_X}\right) f_X(E_X, t) = \frac{4\pi^2}{E_X^2} \Gamma_{Y\to X} g\left(\frac{E_X}{m_Y}\right) \int \frac{\mathrm{d}^3\vec{p}_Y}{(2\pi)^3} \frac{1}{2E_Y} f_Y(E_Y, t). \tag{3.22}$$

Next, we assume that Y had already turned nonrelativistic when the decay occurs, i.e., $E_Y \simeq m_Y$. This is a necessary condition for the dilution mechanism to work, because Y is assumed to dominate the energy content in the Universe as a matter component. It allows us to complete the p_Y integral in (3.22) and express it with the number density, $n_Y(t) = \rho_Y(t)/m_Y$. ρ_Y is the energy density of nonrelativistic Y particles.

For the left-hand side of Eq. (3.22), we change the variables of f_X to $x \equiv E_X/T_X$ and t, where T_X is the temperature of primordial X warm dark matter defined above. The appealing reason for such a change is that as long as the X particles remain ultrarelativistic, the ratio E_X/T_X stays invariant in the expanding Universe. Using the identity,

$$\left(\frac{\partial}{\partial t} - HE_X \frac{\partial}{\partial E_X}\right) f_X(E_X, t) = \frac{\partial}{\partial t} f_X(x, t), \tag{3.23}$$

we finally obtain the phase space equation for secondary dark matter from dilutor decay,

$$\frac{T_X^3}{2\pi^2} x^2 \frac{\partial}{\partial t} f_X(x, t) = n_Y(t) \Gamma_{Y\to X} \frac{T_X}{m_Y} g\left(\frac{T_X}{m_Y} x\right). \tag{3.24}$$

This equation is to be solved along with the following set of energy density Boltzmann equations for X, Y, and the SM particles

$$\dot{\rho}_Y + 3H\rho_Y = -\Gamma_Y \rho_Y, \tag{3.25}$$

$$\dot{\rho}_X + 4H\rho_X = y\mathrm{Br}_X \Gamma_Y \rho_Y, \tag{3.26}$$

$$\dot{\rho}_{\mathrm{SM}} + \left(4H - \frac{\dot{g}_*}{3g_*}\right)\rho_{\mathrm{SM}} = (1 - y\mathrm{Br}_X)\Gamma_Y \rho_Y, \tag{3.27}$$

where ρ_{SM} is the energy density carried by relativistic visible particles and $H^2 = 8\pi G_N(\rho_Y + \rho_X + \rho_{\mathrm{SM}})/3$ is the Hubble parameter. This set of equations applies for nonrelativistic Y, while the X population remains ultrarelativistic.

Using T_X to keep track of time, the phase space function f_X at late times can be solved

$$f_X(x) = \frac{1}{e^x + 1} + \frac{2\pi^2}{x^2} \frac{\Gamma_Y \mathrm{Br}_X}{m_Y^2} \int_{T_{\mathrm{fin}}}^{T_{\mathrm{ini}}} \frac{\mathrm{d}T_X}{T_X} \frac{\rho_Y}{T_X^2 H} g\left(\frac{T_X}{m_Y} x\right). \tag{3.28}$$

The first term of (3.28) is the primordial Fermi-Dirac distribution of X and the second is the nonthermal repopulation of X. The T_X integral should cover the entire temperature range relevant for the production of the secondary component of dark matter. It goes from an arbitrary high initial temperature T_{ini}, which in practice we take $T_{\mathrm{ini}} = m_Y/10$ in order for Y to already be nonrelativistic, as assumed below (3.22). The final result is insensitive to the exact choice of T_{ini}, because the $\rho_Y/(T_X^2 H)$ factor in the integrand is suppressed at higher T_X, as long as $T_X \ll m_Y$. Moreover, the g function cannot lift this suppression, because T_X is bounded from above for a given fixed x.

On the lower limit of integration, we need to go to sufficiently low temperatures, such that all of the Y is depleted by decays into SM and X. Because of the exponential suppression in ρ_Y, the exact T_{fin} is also not relevant. In practice we integrate down to temperatures corresponding to $t = 10\tau_Y$, which sufficiently covers the entire period of nonrelativistic Y decay.

The resulting distributions f_X are shown in Fig. 2, which are plotted at late times after the dilution has completed. The two options for two and three body decays correspond to two scenarios that are relevant for the minimal LRSM under consideration in this work.

(1) The case when X is the lightest right-handed neutrino N_1 and Y is a heavier N_2, which undergoes a three-body decay into N_1 plus two charged leptons, mediated by the W_R gauge boson.
(2) Y is a long-lived scalar Δ_R with a partial decay width into two N_1.

The corresponding g functions and n, y integrals are summarized in Table I, where the masses of final-state charged leptons were neglected.

It is clear from Fig. 2 that the primary component of dark matter dominates at small x, while the secondary

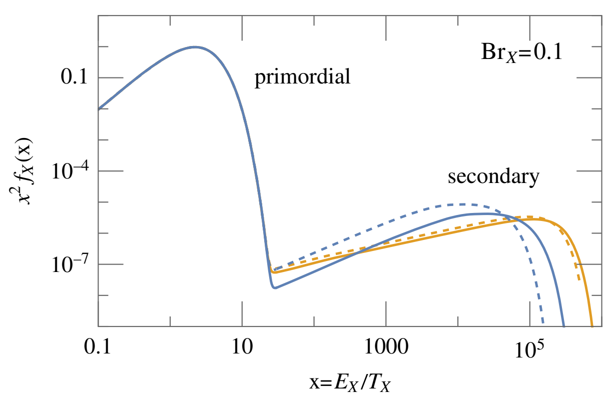


FIG. 2. Phase space distribution of ultrarelativistic dark matter species N_1. We hold $m_{N_1} = 10$ keV for all the curves. The blue (orange) curve corresponds to the three-body decay of N_2 (two-body decay of Δ) doing the job of dilution, as listed in Table I. The solid (dashed) curves correspond to $m_{N_{2},\Delta} = 100$ GeV (1 PeV), respectively. We set $\mathrm{Br}_X = 0.1$ for all the cases.

TABLE I. The energy fraction distribution $g(\omega)$, taken away by dark matter X in the rest frame of the decaying dilutor Y, and its integrals $n = \int g, y = \int \omega g$. In the context of LRSM, the first row corresponds to N_2 as dilutor which can undergo a three-body decay into a N_1 plus two charged leptons ($n = 1$) and θ is the Heaviside unit step function. The second corresponds to the Majorana scalar boson Δ dilution scenario, where each $\Delta \to N_1 N_1$ produces two dark matter states ($n = 2$) and δ is the Dirac delta function.

Dark matter X	Dilutor Y	$Y \to X$ decay	n	$g(\omega)$	y
N_1	N_2	$N_2 \to N_1 + SM$	1	$16\omega^2(3-4\omega)\theta(\frac{1}{2}-\omega)$	$\frac{7}{20}$
N_1	Δ	$\Delta \to N_1 N_1$	2	$2\delta(\omega - 1/2)$	1

component features a significantly smaller occupancy, but carries more energy and thereby affects structure formation. On the plotted curves we kept fixed m_X and Br_X and took two different values of m_Y to demonstrate that the shape of the secondary component is roughly independent of the mass of the dilutor. This follows from Eq. (3.12), where the relic density requires the scaling $m_Y \sim 1/\sqrt{\tau_Y}$ and the reheating temperature is set by the Hubble time $T_{\mathrm{RH}} \sim \sqrt{H} \sim 1/\sqrt{\tau_Y}$. Immediately after $Y \to X$ decay, $E_X \lesssim m_Y$ and $T_X \sim T_{\mathrm{RH}}$. As a result, the kinematic end point $x_{\max} \sim m_Y/T_{\mathrm{RH}}$ is roughly held constant for fixed m_X, irrespective of the values of m_Y or τ_Y.

In the case where the dilutor is a long-lived scalar Δ (see Sec. II B 5), the kinematics is so simple that we can further derive an explicit closed form for f_X. Since it is a two body decay, g is a Dirac-δ function and we can complete the T_X integral in Eq. (3.28) to obtain

$$x^2 f_X(x) = \frac{x^2}{e^x+1} + \frac{8\pi^2 \mathrm{Br}_X \Gamma_Y}{m_Y^2}\left(\frac{\rho_Y}{T_X^2 H}\right)_{T_{X\star}}, \tag{3.29}$$

$$\simeq \frac{x^2}{e^x+1} + 0.16\ \mathrm{sec}\,\mathrm{Br}_X\left(\frac{10\,\mathrm{keV}}{m_X}\right)^2\left(\frac{\rho_Y}{T_X^2 H}\right)_{T_{X\star}}, \tag{3.30}$$

where we used the relic equation (3.12), set $\Omega_X = 0.26$ and approximated with small Br_X in the second step. With such a simple expression in (3.30) we can understand the behavior of f_X in Fig. 2 that emerges from solving Eq. (3.28). First of all, g is a δ function in temperature, which essentially selects a particular moment in $T_{X\star} = m_Y/(2x)$ for a fixed x and thus completely removes any dependence on the boundary conditions $T_{\mathrm{ini,fin}}$.

Furthermore, we can derive the explicit dependence on x for various moments in the expansion of the Universe during the dilutors' decay. For this, we only have to examine the x dependence of the factor $(\rho_Y/T_X^2 H)_{T_{X\star}}$ in Eq. (3.30); the rescaling with Ω_X and m_X is trivial. Note that for the purpose of this discussion, Y is always nonrelativistic and $\rho_Y \propto m_Y T_X^3$. In the early stages of radiation domination we have $H \propto T_X^4$ and therefore $(\rho_Y/T_X^2 H)_{T_{X\star}} \propto m_Y/T_{X\star} \sim x$. Once Y starts to dominate, the Hubble parameter goes as $H \sim \sqrt{\rho_Y}$ and the relevant term goes as $(\rho_Y/T_X^2 H)_{T_{X\star}} \propto \sqrt{m_Y/T_{X\star}} \sim \sqrt{x}$. In both cases, m_Y cancels out, that is why the orange curve in Fig. 2 remains almost identical for difference choices of m_Y. Finally, large x corresponds to low $T_{X\star}$ when all the Y has decayed away and the secondary part of f_X is exponentially suppressed. These x dependencies explain the shape of the two-body orange curves in the $x > 10$ region.

For the three-body decay, the spectral function g is not as sharply peaked as the Dirac δ, but has a maximum at $\omega = 1/2$ and similar considerations go through, with transitions between different x dependencies becoming less sharp. The bottom line is that, once we fix Br_X and m_X, the f_X does not dependent on the mass of the dilutor Y, the behavior of f_X is roughly the same for different decay topologies and it extends to large x, beyond the usual primary component. In the following section we will examine how this behavior translates onto the physical matter power spectrum $P(k)$.

Let us emphasize that we assume the X particles remain collisionless after the Y decay throughout this work. This means that the imprint of the dark matter model (fundamental physics) on the phase space distribution is preserved until later times and can directly affect cosmological observations. We do not consider the possibility of having strong DM self-interactions. They could rethermalize the dark sector and soften the above phase space distribution. At the same time, they would facilitate excessive dark matter production, which is adverse to the dilution mechanism considered here.

E. Imprint on the matter power spectrum

Let us turn to a quantitative numerical analysis in the parameter space of m_X versus m_Y. For each point we first set the dilutor lifetime τ_Y using Eq. (3.13). Next, we determine the phase space distribution f_X with Eq. (3.28) and evolve the density perturbations using the linear Boltzmann solver code CLASS [87–89] to obtain the corresponding matter power spectrum $P(k)$. We scan over 200 points in the mass range $m_X \in (1\ \mathrm{keV}, 1\ \mathrm{MeV})$ and $m_Y \in (1\ \mathrm{GeV}, 10^{16}\ \mathrm{GeV})$ for both decay channels

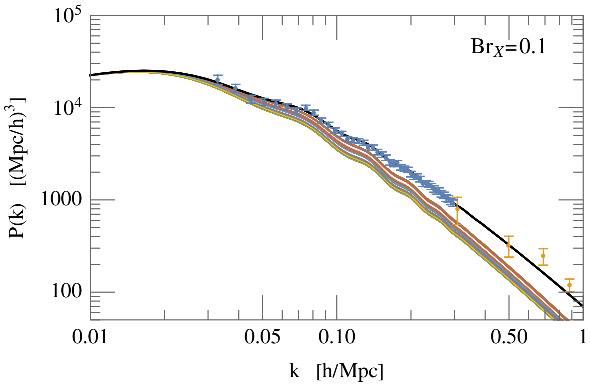


FIG. 3. Primordial matter power spectrum in standard ΛCDM (black solid curve) and a set of diluted dark matter models listed in Table I (colorful curves). Like in Fig. 2, we set $\mathrm{Br}_X = 0.1$. Data points from SDSS DR7 LRG and Lyman-α observations are shown in blue and orange, respectively.

considered in Table I. The results are shown by the colored curves in Fig. 3, where we set $\mathrm{Br}_X = 0.1$. The black solid curve is the fiducial ΛCDM (Λ cold dark matter) model.

The experimental data points come from the SDSS DR7 on luminous red galaxies [90] (blue) and the Lyman-α forest [91] (orange) measurements. All the curves in scenarios with secondary X share a common feature with significant deviations from data in the $k \gtrsim 0.03\ h/\mathrm{Mpc}$ region. These occur at a much lower k compared to other dark matter production mechanisms such as thermal freeze-in [92,93]. This is mainly due to the large hierarchy between the dilutor and the dark matter mass, required by Eq. (3.13). Based on a simple $\Delta\chi^2$ fit to the data, we find that the LSS data from SDSS sets a much stronger constraint on these scenarios than Lyman-α, making this probe particularly robust. The conflict with data increases with Br_X, which translates into an upper bound, shown in Fig. 4 for the two models in Table I. The bound does not depend much on the precise shape of the phase space distribution (because it is integrated over) and the message is similar for both cases: the branching ratio of the dilutor decaying into dark matter is constrained to be

$$\mathrm{Br}_X \lesssim 1\%, \qquad @95\%\ \mathrm{CL}. \tag{3.31}$$

This bound is nearly independent of m_Y, simply because the secondary component of the phase space distribution f_X in (3.28) is mostly independent of m_Y, as explained in the paragraph below (3.28). The Br_X limit gets slightly relaxed for larger m_Y, because holding the dark matter relic density fixed in Eq. (3.13) requires the lifetime τ_Y to be shorter, leading to a higher reheating temperature after the decay of dilutor. The corresponding temperature for the secondary dark matter component to become nonrelativistic also increases, which is a $\sqrt{12 g_*}$ effect, see Eq. (3.16). Eventually, this shifts the deviation of $P(k)$ to a slightly higher k, where the data is less precise and thus becomes less constraining.

The constraint derived here comes predominantly from the LSS data, which relies only on the evolution of matter density perturbations in the linear regime. LSS thus provides a robust test of these models and we expect similar constraints to apply broadly for other dilutor → dark matter decay topologies.

Our result can be generalized to initial abundances for Y and X beyond the relativistic freeze-out. A subthermal initial population of Y needs to be heavier and/or longer lived in order to provide the same amount of entropy

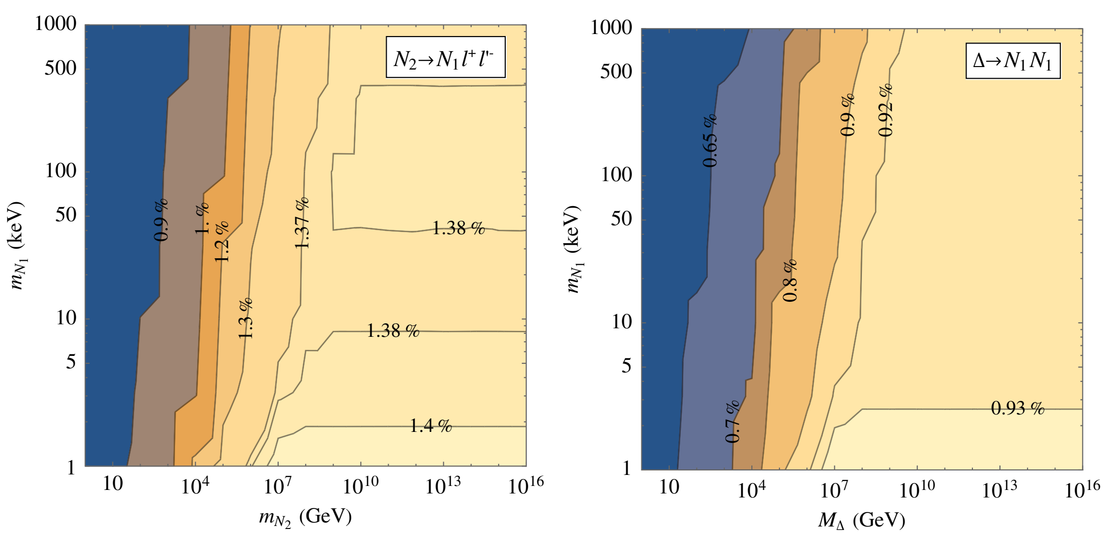


FIG. 4. Upper bound on Br_X, the branching ratio of dilutor Y decaying into X from the fit to LSS data (SDSS DR7 LRG), for the two dilution scenarios considered in Table I in the context of LRSM. For each point in the $m_X - m_Y$ parameter space, the Y lifetime is determined by requiring X to comprise all of the dark matter in the Universe and the upper bound on Br_X can be read from the value of the contour passing the point.

injection. The secondary X particles from Y decay become more energetic and take even longer to become matterlike. This impacts the primordial matter power spectrum down to even lower k and leads to a more stringent constraint on Br_X than Eq. (3.31). On the contrary, starting with a smaller overpopulation of X, the constraint on Br_X will be weaker.

IV. ANATOMY OF DILUTION SCENARIOS IN LRSM

Remarkably, the minimal LRSM contains all the ingredients for the dark matter dilution mechanism, described in the previous section, to occur. Throughout this work, we consider the lightest right-handed neutrino N_1 to be the dark matter, i.e.,

$$X = N_1. \tag{4.1}$$

Following the discussion in Sec. II B 2, with a mass below 100 MeV, it always freezes out relativistically and is generically overproduced. The role of the diluting particle Y can be played by either a heavier right-handed neutrino or the Majorana Higgs boson Δ; the latter option is explored here for the first time. In this section, we carefully examine the viability of each scenario and point out the corresponding opportunities for experimental tests.

A. Heavy neutrino dilutor

In this subsection, we first consider one of the two heavier right-handed neutrinos N_2 to play the role of the dilutor Y. We begin with the implications of LSS constraint obtained from the previous section, which results in a nontrivial no-go theorem. We then discuss the options for bypassing such constraints and describe the viable scenarios.

1. Implication of the SDSS constraint

Let us label the dilutor of Sec. IV A as $Y = N_2$. The most obvious decay channels of N_2 are those mediated by the heavy W_R boson in the LRSM. Corresponding Feynman diagrams for the decays are depicted in Fig. 5 below. Similarly to weak decays of the τ lepton in the SM, there are semileptonic and pure-leptonic decay channels.

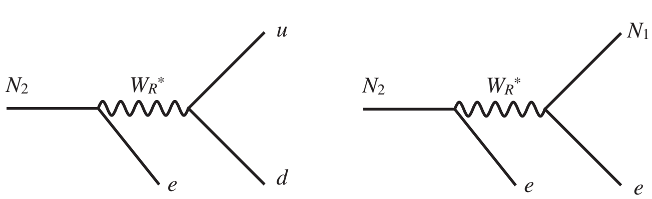


FIG. 5. Feynman diagram for dilutor N_2 decay via right-handed charged-current interaction mediated by W_R. The flavor or final state charged leptons and quarks is dictated by the matrix elements of V_{PMNS}^R and V_{CKM}^R.

All the decay products in the first channel are SM particles, which makes it a perfect decay mode for the DM dilution mechanism to work. In contrast, the second channel has a dark matter N_1 in the final state. The corresponding decay branching ratio is tightly constrained by large scale structure observations, to be smaller than 1%, as pointed out in Sec. III E.

In the minimal LRSM, the ratios among the above W_R mediated decay are fixed by the structure (particle content and gauge interactions) of the model. If these were the only N_2 decay channels, then the decay branching into the final state containing N_1 would be 10% for mass of N_2 well above the electroweak scale. This branching ratio only gets higher for lighter N_2. The large scale structure constraint from SDSS thus firmly excludes such simplest dilution scenario. What may save the day are other possible decay channels of N_2. These may arise either due to neutrino or gauge boson mixing in the LRSM, as will be discussed next.

2. A no-go theorem

If minimality is one's first priority, then it would be most desirable to open up the extra decay channel(s) for the right-handed neutrino dilutor and also allow it to participate in the type-I seesaw mechanism for generating the active neutrino masses. Naïvely, both could occur through a mixing of N_2 with the active neutrinos. However, there is a no-go theorem against such a possibility.

If the dilutor N_2 participates in the seesaw mechanism, then the mixing between N_2 and light neutrinos is bounded from below

$$\theta_{N_2\nu} \gtrsim \sqrt{\frac{m_\nu}{m_{N_2}}}. \tag{4.2}$$

The value of $\theta_{N_2\nu}$ could be much larger than $\sqrt{m_\nu^\odot/m_{N_2}}$ due to additional degrees of freedom involved in type-I seesaw mechanism [94], but it cannot be made smaller if N_2 is responsible for neutrino mass generation. For this contribution to the neutrino mass to be significant and at least explain the mass difference for solar neutrino oscillation, we require $m_\nu \gtrsim m_\nu^\odot \simeq \sqrt{8\times 10^{-5}\ \mathrm{eV}^2}$.

We first consider the case where the N_2 mass lies below the weak scale. The partial decay width of N_2 via this mixing and an off-shell W boson is then

$$\tau_{N_2}^{-1} \gtrsim \frac{\mathfrak{m} G_F^2 m_{N_2}^5 \theta_{N_2\nu}^2}{96\pi^3} \gtrsim 10^{-23}\ \mathrm{GeV}\left(\frac{m_{N_2}}{1\ \mathrm{GeV}}\right)^4, \tag{4.3}$$

where $\mathfrak{m}$ is the final state multiplicity factor within the range $\sim$1–10. In the first step, we neglect contributions to the decay via the Z boson or possible interference effects. This approximation would only affect our estimate by an order of 1 factor, but keep our conclusion intact. The first inequality also accounts for other possible (subdominant) N_2 decay channels (e.g., via W_R) and the second inequality follows from Eq. (4.2). Plugging Eq. (4.3) into Eq. (3.12), we find a lower bound on the dark matter relic density

$$\Omega_{N_1} \gtrsim 2.9\left(\frac{m_{N_1}}{6.5\ \mathrm{keV}}\right)\left(\frac{m_{N_2}}{1\ \mathrm{GeV}}\right). \tag{4.4}$$

The reference mass for N_1 is the lower bound on warm dark matter found by the DES collaboration [3], which is consistent with other constraints using the Lyman-α forest [29,30], strong gravitational lensing observations [31,32], and recent combined analysis [33,34]. The observed dark matter relic abundance then sets an upper bound on the mass of N_2,

$$m_{N_2} \lesssim 90\ \mathrm{MeV}. \tag{4.5}$$

Applying this bound again back in Eq. (3.12), we obtain a lower bound on the lifetime of N_2,

$$\tau_{N_2} \gtrsim 160\ \ \mathrm{sec}. \tag{4.6}$$

Because the Universe was matter-dominated before the N_2 decayed away, the above lower bound on its lifetime is strongly excluded by the big bang nucleosynthesis, which would require $\tau_{N_2} \lesssim 1$ sec.

On the other hand, if N_2 is heavier than the weak scale, the decay induced by the mixing parameter Eq. (4.2) would be into an on-shell W boson at a much higher rate,

$$\begin{aligned}\tau_{N_2}^{-1} &\gtrsim \frac{G_F m_{N_2}^3 \theta_{N_2\nu}^2}{4\sqrt{2}\pi}\left(1-\frac{M_W^2}{m_{N_2}^2}\right)\left(1+\frac{M_W^2}{m_{N_2}^2}-\frac{2M_W^4}{m_{N_2}^4}\right)\\ &\gtrsim 10^{-12}\ \mathrm{GeV}\left(\frac{m_{N_2}}{100\ \mathrm{GeV}}\right)^2. \end{aligned}\tag{4.7}$$

The factor $G_F m_{N_2}^2$ properly accounts for a longitudinal enhancement in the limit when $m_{N_2} \gg M_W$. Plugging this into Eq. (3.12) leads to a lower bound on the dark matter relic density

$$\Omega_{N_1} \gtrsim 1.7\times 10^5\left(\frac{m_{N_1}}{6.5\ \mathrm{keV}}\right). \tag{4.8}$$

All the N_2 mass dependency cancels out completely and dark matter is considerably overproduced.

This completes the proof of the no-go theorem. It is derived by combining the constraints on the dark matter relic density and the dilutor lifetime. Although we have used some $\sim$ in the above reasoning, the results in Eqs. (4.6) and (4.8) are in sharp contradiction with the existing constraints, making it convincing that there is no room to avoid the theorem. It is also worth pointing out that the theorem not only applies to the LRSM focused in this work, but also to other gauge extensions, such as the $U(1)_{B-L}$ model [26,95,96].

3. Type-II seesaw dominance and dilutor decay via $N-\nu$ mixing

The no-go theorem presented above implies that if one of the heavier right-handed neutrinos (e.g., N_2) plays the role of dilutor, its mixing with light active neutrinos must be much smaller than Eq. (4.2). In other words, the contribution to neutrino mass from N_2 via the type-I seesaw must be well below what is needed for explaining the observed neutrino oscillation phenomena. At the same time, the dark matter candidate N_1 also cannot fully participate in the seesaw, because of the x-ray constraints. The only remaining right-handed neutrino N_3 is free from constraints and does contribute to neutrino masses in the type-I seesaw, but is unable to explain the two mass square differences needed for neutrino oscillations. As a consequence, light neutrino masses must be accounted for by additional sources.[4]

A way out within the minimal LRSM is by considering another source of mass for the light neutrinos, which comes from the vacuum condensate of the left-handed scalar triplet Δ_L, through the type-II seesaw mechanism. By relieving the dilutor N_2 from the role of neutrino mass generation, we can treat its mixing with light neutrinos $\theta_{N_2\nu}$ as a free parameter, which can be arbitrarily small. This enables a viable window in the model parameter space for the dark matter dilution mechanism to work.

In Fig. 6, the blue and orange curves show where in the $\theta_{N_2\nu}$ versus m_{N_2} plane the dark matter N_1 can obtain the correct relic density after the N_2 dilution, for two choices of N_1 mass. The value 6.5 keV is the lowest allowed warm dark matter mass by the DES result. The purple region is excluded by BBN because the lifetime of N_2 is longer than a second. Clearly, viable values of $\theta_{N_2\nu}$ must be very tiny $\lesssim 10^{-9}$ to satisfy both constraints. In contrast, the region above the green line indicates the required values of $\theta_{N_2\nu}$ if N_2 participates in the type-I seesaw mechanism, which offers a way to visualize and quantify the above no-go theorem. We also find a lower bound on the dilutor N_2 mass of around 20 GeV.

The dark matter relic curves in Fig. 6 are valid under the assumption that the W_R mediated decay modes of the dilutor N_2 (see Fig. 5) are subdominant to those induced by the $N_2-\nu$ mixing. This condition is mostly easily satisfied if N_2 is heavier than the W boson but still close to

[4]This argument also implies that the minimal $U(1)_{\mathrm{B-L}}$ model where the new gauge symmetry is broken by a Standard Model singlet scalar is unable to account for both the dark matter dilution mechanism and neutrino masses. Additional degrees of freedom (e.g., the counterpart of Δ_L, see below) are needed.

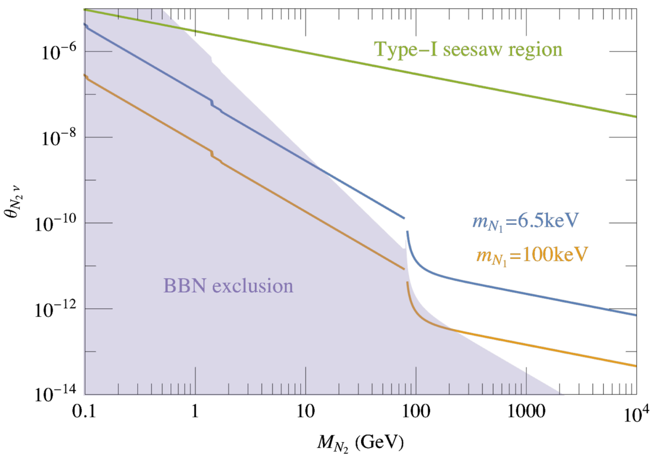


FIG. 6. Parameter space for correct dark matter relic density, where N_2 serves as the dilutor and decays via a mixing with the active neutrino. We obtain $\Omega_{N_1} = 0.26$ along the blue and orange curves for $m_{N_1} = 6.5$ and 100 keV, respectively. The purple shaded region is excluded by BBN, because the lifetime of N_2 is longer than a second. Along with the relic curves, it sets an upper bound on the mixing $\theta_{N_2\nu}$. This upper bound is stronger for heavier N_1. In contrast, the region above the green line in the upper-right corner of the figure shows the mixing angle needed for N_2 to participate in the type-I seesaw mechanism and explain the neutrino mass difference for solar neutrino oscillation. This is clearly incompatible with the dilution mechanism and verifies the no-go theorem presented in Sec. IVA 2.

the weak scale. This leads to a lower bound on the mass scale of W_R boson,

$$M_{W_R} \gtrsim \frac{10 \text{ GeV}}{\sqrt{\theta_{N_2\nu}}} \gtrsim 10^6 \text{ GeV} = 1 \text{ PeV}, \tag{4.9}$$

where in the second step we read from Fig. 6 that for $m_{N_2} > M_W$, the highest value of $\theta_{N_2\nu}$ is around 10^{-10}, which is also orders of magnitude below the seesaw line, required for neutrino mass generation.

4. Dilutor decay via $W - W_R$ mixing and x-ray limits

The previous subsections explored the possibility of the dilutor N_2 decaying dominantly through its mixing with the active light neutrinos. Here, we discuss another possible N_2 decaying channel inherent to the minimal LRSM, via the gauge boson mixing $\xi_{\rm LR}$, mentioned in Sec. II B 4 and shown in the Feynman diagram on Fig. 7 below. Because the absolute value of $\xi_{\rm LR}$ is bounded from above by $M_W^2/M_{W_R}^2$ [see Eq. (2.12)], if N_2 is lighter than the W boson and decays via off-shell W, the corresponding partial decay rate will have the same parametric dependence as those in Fig. 5. It cannot provide sufficient suppression to the branching ratio of $N_2 \to N_1$ decay and the resulting dark matter production/dilution mechanism still suffers from the strong constraint from large scale structure.

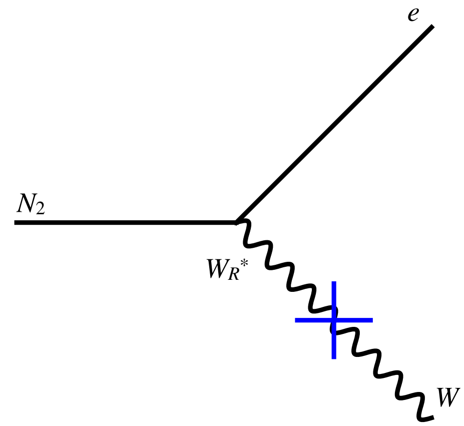


FIG. 7. Feynman diagram for N_2 decay via $W - W_R$ mixing in LRSM. Blue cross indicates an insertion of $\xi_{\rm LR}$ mixing.

This observation forces the viable parameter space to the window where $M_{W_R} > m_{N_2} > M_W$.

In this case, the available decay rates for N_2 are

$$\begin{aligned}
\Gamma_{N_2\to N_1\ell^+\ell'^-} &= \frac{G_F^2 m_{N_2}^5}{96\pi^3}\left(\frac{M_W}{M_{W_R}}\right)^4, \\
\Gamma_{N_2\to \ell q\bar{q}'} &= \frac{\mathfrak{m} G_F^2 m_{N_2}^5}{96\pi^3}\left(\frac{M_W}{M_{W_R}}\right)^4, \\
\Gamma_{N_2\to \ell W} &= \frac{g^2|\xi_{\rm LR}|^2 m_{N_2}}{32\pi}\left(\frac{m_{N_2}}{M_W}\right)^2\left(1-\frac{M_W^2}{m_{N_2}^2}\right) \\
&\quad \times\left(1+\frac{M_W^2}{m_{N_2}^2}-\frac{2M_W^4}{m_{N_2}^4}\right),
\end{aligned} \tag{4.10}$$

where the first two decays occur via diagrams in Fig. 5 and the last one via Fig. 7. Again, $\mathfrak{m}$ is the final state multiplicity factor, which equals 12 (9) for $m_{N_2} > (<) m_t + m_b$. In the absence of $\theta_{N_2\nu}$, the three rates in (4.10) sum up to the total decay rate of N_2.

Like before, we impose three requirements on the dilution scenario here:

(1) Dark matter N_1 obtains the correct relic density.
(2) The decay branching ratio of dilutor N_2 to dark matter is smaller than 1%.
(3) Dilutor decays faster than 1 second.

Our main results are then summarized in Fig. 8.

The left-panel shows the parameter space of m_{N_2} versus $\tan\beta$. Outside the darkest green shaded region, the dilutor to dark matter decay branching ratio exceeds 1% and the parameter space is excluded by the LSS data. From Eq. (4.10), it is useful to note that the branching ratio and the LSS constraint is independent of other parameters of the model such as M_{W_R} or m_{N_1}. In contrast, the total decay rate of dilutor N_2 does depends on M_{W_R} and so is the allowed parameter space that is consistent with the BBN constraint. The two blue curves show the lower limit on

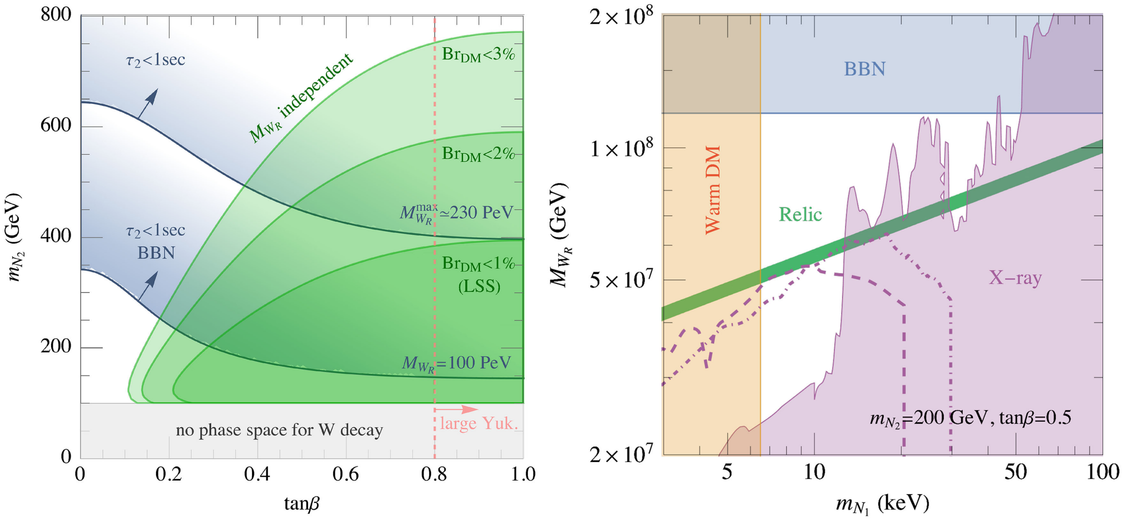


FIG. 8. Left: viable parameter space for the dark matter dilution to work in the m_{N_2} versus $\tan\beta$ plane, where the diluting particle N_2 dominantly decays via the mixing between W and W_R gauge bosons. The region outside the darkest green is excluded by the LSS constraint in Eq. (3.31). The region to the right of the vertical pined dashed line is excluded by the theoretical constraint on the range of $\tan\beta$, Eq. (2.15). The two dark blue curves corresponds to lower bound on m_{N_2} to pass the BBN constraint, for two choices of W_R mass, 10^8 and 2.25×10^8 GeV, respectively. Right: further constraints in the M_{W_R} versus m_{N_1} plane, with other parameters fixed, $m_{N_2}=200$ GeV and $\tan\beta=0.5$. The blue shaded region again indicates the upper bound on M_{W_R} from the BBN constraint. The purple region is excluded by the existing x-ray line searches for dark matter decay $N_1\to\nu\gamma$ via the loop processes shown in Fig. 11. The dashed and dot-dashed purple curves corresponds to the reach of future x-ray experiments ATHENA and XRISM, respectively. The orange shaded region is the warm dark matter exclusion limit set by the DES experiment. The green band is where dark matter obtains the correct relic density after the entropy dilution.

m_{N_2} for two choices of $M_{W_R}=10^8$ GeV (lower) and 2.25×10^8 GeV (upper), respectively. Clearly, the latter case is marginal where the available parameter space for dilution mechanism closes. From this, we derive a upper bound on $M_{W_R}<2.25\times 10^8$ GeV. In the same plot, the vertical pink dashed line corresponds to the theoretical upper bound on $\tan\beta$ derived in Eq. (2.15). Viable parameter space for dark matter relic only occurs in the darkest green region with proper arrangement of other parameters.

Next, we address the x-ray line search limits on dark matter N_1 decay, as shown in the right panel of Fig. 8. In the minimal LRSM, the dark matter candidate is not absolutely stable and there are in fact two contributions to the radiative decay of $N_1\to\nu\gamma$. One is via the $N_1-\nu$ mixing and applies also to the regular sterile neutrino dark matter. The other is through the $W-W_R$ mixing and both occur at the one loop level with coherent amplitudes. In the presence of a nonzero $\theta_{N_1\nu}$, the radiative decay rate of dark matter is a well-known result [97,98],

$$\Gamma_{N_1\to\nu\gamma}=\frac{9\alpha}{256\pi^4}G_F^2 m_{N_1}^5\sin^2\theta_{N_1\nu}. \tag{4.11}$$

In the presence of a nonzero ξ_{LR}, there are new Feynman diagrams for the radiative decay of dark matter N_1. We derive the leading-order decay rate

$$\Gamma_{N_1\to\nu\gamma}=\frac{\alpha\xi_{\mathrm{LR}}^2}{8\pi^4}G_F^2 m_{N_1}^3\sum_\ell |(V_{\mathrm{PMNS}}^R)_{\ell 1}|^2 m_\ell^2. \tag{4.12}$$

See the Appendix for a detailed derivation of this rate. In this case, because of the W_R and $\tan\beta$ dependencies in ξ_{LR}, we find a closer interplay between the x-ray search bounds and the requirements on the dilution mechanism found in the previous subsection.

The implications from x-ray constraints are shown in Fig. 8 (right). We fix $m_{N_2}=200$ GeV and $\tan\beta=0.5$ which is an allowed point in Fig. 8 (left), and show the other constraint in the dark matter mass m_{N_1} versus M_{W_R} parameter space. As discussed earlier, for the dilutor N_2 to decay before BBN, there is an upper bound on M_{W_R} for given m_{N_2}. This excludes the blue shaded region. The purple shaded region is then excluded by the existing x-ray line searches for dark matter decay [99–101], which sets a lower bound on M_{W_R} and upper bound on m_{N_1}. Interestingly, the remaining window for viable dark matter in this scenario can be tested by the upcoming x-ray experiments ATHENA [102] and XRISM [103,104], as shown by the dashed and dot-dashed purple curves. Here, we assume generic flavor mixing matrix where all the elements are $\mathcal{O}(1)$ in magnitude and approximate $\sum_\ell |(V_{\mathrm{PMNS}}^R)_{\ell 1}|^2 m_\ell^2\sim m_\tau^2$ in Eq. (4.12). The x-ray limit could be weakened if the right-handed leptonic mixing matrix element $(V_{\mathrm{PMNS}}^R)_{\tau 1}$ is suppressed or if destructive interference between amplitudes is strong enough. As explained in Sec. III D [see also Eq. (3.28)], for sufficiently small $\mathrm{Br}_{Y\to X}$, the phase space distribution of dark matter N_1 follows exactly the thermal distribution, exactly like a warm dark matter is defined [105,106]. The orange shaded region corresponds to a lower bound of 6.5 keV on warm dark matter mass.

5. Lower bound on the W_R mass scale

So far, we have discussed several options of having the light right-handed neutrino N_1 to comprise all the dark matter in the Universe, through the dilution mechanism where the dilutor is a heavier right-handed neutrino N_2. In all cases, we find that the mass scale of the W_R gauge boson must be rather high. In the case where N_2 dominantly decays via its mixing with light neutrinos, the lower bound is around PeV scale, as found in Eq. (4.9). In the case where N_2 dominantly decays via $W - W_R$ gauge boson mixing, lower bound on M_{W_R} is higher (tens of PeV), due to the relic density explanation and constraints on the radiative decay of dark matter N_1, as shown in Fig. 8 (right). In both cases when calculating the decay rate of dilutor N_2, we have made the assumption that the masses of its decay products are much smaller than m_{N_2}. The lower bounds on M_{W_R} are derived based on this assumption, which are generic and does require special arrangement of the parameters of the model.

Here, we wish to scrutinize if the mass scale of W_R is allowed to be even lighter at all if some amount of tuning of parameters is arranged. While this might be less appealing, the main motivation behind is the prospect of other experimental probes (such as high-energy colliders) of the LR symmetry scale. Such a possibility was first explored in [36], which resorts to a compress spectrum with the mass of dilutor N_2 being close to the sum of charged pion and a charged lepton masses. This leads to a phase space suppression in the dilutor decay rate ($N_2 \to \pi + \ell$) and enables sufficient longevity while keeping W_R mass near the TeV scale. Moreover, the flavor structure of the right-handed lepton mixing matrix V^R_{PMNS} must also be tuned, such that N_1 primarily couples to the τ lepton in the right-handed current interaction, thus kinematically forbidding the $N_2 \to N_1$ decay (via off-shell W_R) that is constrained by large scale structure as discussed in Sec. IVA 1. In this scenario, light neutrino masses can be explained via a mixed type-I (where N_3 mainly contributes) and type-II seesaw mechanism.

To explain the dark matter relic density using the dilution mechanism [see Eq. (3.12)], it seems challenging to have a dilutor N_2 mass well below 2.2 GeV, because BBN forbids the lifetime of N_2 to be longer than a second, while at the same time the Tremaine-Gunn bound forbids dark matter mass to be well below a keV. To get around this difficulty, [36] noticed a special mass window around $M_{W_R} \sim 5$ TeV could work, where the above flavor structure allows dark matter N_1 to freeze out slightly before the QCD phase transition whereas the dilutor N_2 freezes out slightly after. It leads to an enhancement factor in the dilution factor $\mathcal{S}$ in Eq. (3.9), given by the ratio of g_* at temperatures above and below Λ_{QCD}, and in turn a suppression in the final dark matter relic density. Thanks to this effect, [36] found viable solutions for dark matter mass around 0.5 keV. However, after the recent substantial progress in constraining the warm dark matter mass, e.g., $m_{N_1} > 6.5$ keV, found by the DES collaboration from ultrafaint dwarfs [3], such a low mass W_R window has been firmly closed. This leads us to conclude that with a right-handed neutrino dilutor, the up-to-date lower bound on M_{W_R} for consistent dark matter cosmology in LRSM is pushed to above the PeV scale, given by Eq. (4.9).

B. Majorana Higgs dilutor

In this subsection, we explore the other dilutor candidate in LRSM, the Majorana Higgs Δ, introduced in Sec. II B 5. The role of dark matter is still played by N_1. To our knowledge, such a possibility has not been considered in any previous dark matter analysis of the model.

Through the spontaneous gauge symmetry breaking $SU(2)_R \times U(1)_{B-L} \to U(1)_Y$, the couplings of Δ are tied to mass generation of the gauge bosons $W_R^\pm, Z'$ and the right-handed neutrinos N_i. To be long lived and qualify as the diluting particle, the mass of Δ must be well below those of W_R and Z'. We will work in the parameter space where Δ is also much lighter than two of the right-handed neutrinos N_2, N_3. As a result, the decays of Δ into these on-shell final state particles are forbidden. Its possible decay channels are shown by the Feynman diagrams in Fig. 9.

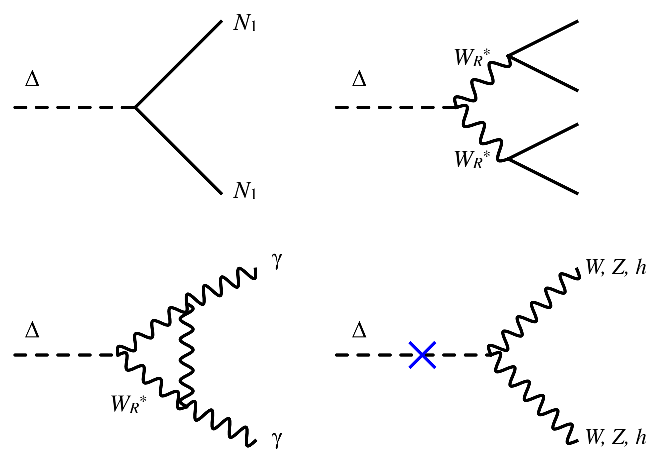


FIG. 9. Feynman diagrams for the decay rate of the dilutor Δ in the minimal LRSM. The coupling of Δ to right-handed neutrino and $W_R^\pm$ are proportional to their masses, whereas the blue cross in the last diagram represents Δ-Higgs boson mixing, generated by the scalar potential. In the second diagram, the fermion final states that connect to virtual W_R include quark pairs and N_1 plus a charged lepton.

The decay rate of Δ into two dark matter particles N_1 is suppressed by the small mass m_{N_1}. The decays via off-shell W_R into light fermions or two photons are suppressed by the small ratio of M_Δ/M_{W_R}. The decay via off-shell Z' is always subdominant, because Z' is heavier than W_R by a factor of $\sqrt{3}$ and a relatively smaller decay branching ratio into N_1 [107]. Finally, Δ could decay via a mixing with the SM Higgs boson. For M_Δ well above the electroweak scale, the dominant decays via the Higgs mixing are into W^+W^-, ZZ, and hh final states.

With the mass hierarchy $M_W \ll M_\Delta \ll M_{W_R}$, the partial decay rates of Δ are

$$\Gamma_{\Delta\to N_1N_1} = \frac{G_F M_W^2 m_{N_1}^2 M_\Delta}{4\sqrt{2}\pi M_{W_R}^2}, \quad \Gamma_{\Delta\to W_R^* W_R^*} = \frac{5G_F^3 M_W^6 M_\Delta^7}{576\sqrt{2}\pi^3 M_{W_R}^6}, \tag{4.13}$$

$$\Gamma_{\Delta\to\gamma\gamma} = \frac{49\alpha^2 G_F M_W^2 M_\Delta^3}{128\sqrt{2}\pi^3 M_{W_R}^2}, \quad \Gamma_{\Delta\to h^*} \simeq \frac{\theta_{\Delta h}^2 G_F M_\Delta^3}{4\sqrt{2}\pi}. \tag{4.14}$$

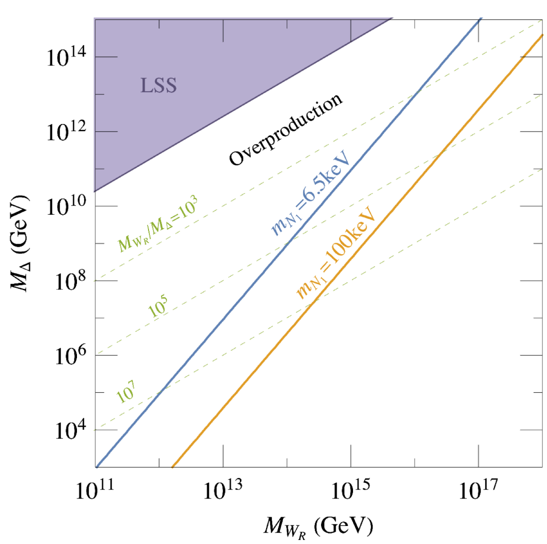
For simplicity, we work in the limit where all final state particle masses are negligible. The $\Delta \to W_R^* W_R^*$ decay occurs through two off-shell $W_R^\pm$ bosons and has four right-handed fermions in the final states. For this partial rate, we apply the four-body decay formula Eq. (2.35) of [108] and work in the heavy W_R limit. The kinematically allowed fermion final states are quark pairs and N_1 plus a charged lepton.

Among the above four decay channels of the dilutor, the first two can produce energetic dark matter N_1 in the final state and get constrained by the large scale structure observations. The third channel, where Δ decays into a pair of photons, can bypass the LSS constraint. Indeed, we find that the ratios

$$\frac{\Gamma_{\Delta\to N_1N_1}}{\Gamma_{\Delta\to\gamma\gamma}} \simeq 1.2\times 10^{-7}\left(\frac{m_{N_1}}{M_\Delta}\right)^2, \quad \frac{\Gamma_{\Delta\to W_R^*W_R^*}}{\Gamma_{\Delta\to\gamma\gamma}} \simeq 2.3\times\left(\frac{M_\Delta}{M_{W_R}}\right)^4, \tag{4.15}$$

can both be made much smaller than 1% if $m_{N_1} \ll M_\Delta \ll M_{W_R}$. This mass-scale hierarchy is consistent with the above mass spectrum assumptions. It allows the LRSM to evade the LSS constraint, even in the absence of Δ-Higgs boson mixing.

In the left panel of Fig. 10, we first work in the $\theta_{\Delta h} = 0$ limit and the blue and orange curves show the M_Δ versus M_{W_R} parameter space, where dark matter N_1 obtains the correct relic abundance through the Δ-dilution mechanism, for two values of $m_{N_1} = 6.5$ and 100 keV, respectively. We apply Eq. (3.10) by identifying $Y = \Delta$. Dark matter is overproduced in regions to the left of the curves. The purple region is excluded by LSS observations, because the branching ratio for $\Delta \to W_R^* W_R^* \to$ light fermions decay is too high [see Eq. (4.15)]. We find that the mass scale of W_R must be very high, above $\sim 10^{11}$ GeV, but the mass scale of Δ can be much lower. However, the price of having a lighter Δ is to increase the mass hierarchy between Δ and W_R, as indicated by the green dashed lines. Similar to the argument in footnote 2 against a very light Δ to be the dark matter, we do not consider Δ to be lighter than W_R by much more than a loop factor. Taking into account of this theoretical constraint, we end up finding that both Δ and



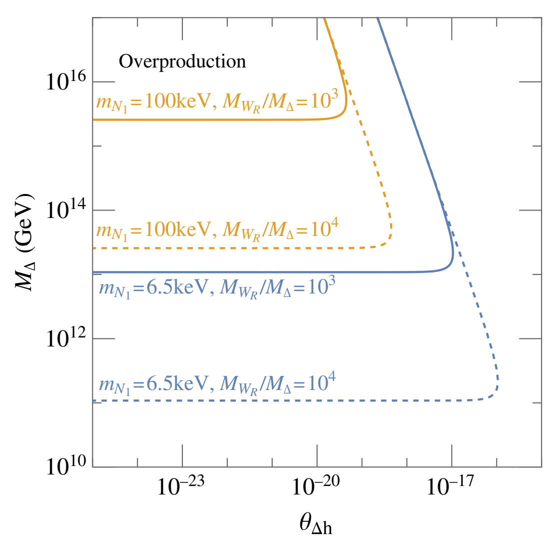


FIG. 10. Blue and orange curves show the parameter space where dark matter obtains the correct relic abundance via the dilution mechanism, where the Majorana Higgs Δ plays the role of dilutor. In the left panel, we set $\theta_{\Delta h} = 0$. The purple region is excluded by the LSS observations. The green dashed curves indicate different M_{W_R}/M_Δ mass ratios. In the right panel, we turn on $\theta_{\Delta h}$ but fix the M_{W_R}/M_Δ mass ratio for each relic curve.

W_R masses are pushed to rather high values, close to the GUT scale. We have also checked that the lifetime of Δ is much shorter than 1 second along the entire curves, thus safely evading the BBN constraint.

In the right panel of Fig. 10, we project the relic curves to the $\theta_{\Delta h}$ versus M_Δ parameter space for two realistic choices of the mass ratio M_{W_R}/M_Δ. On each relic curve, the horizontal part has no $\theta_{\Delta h}$ dependence because the total decay rate is dominated by $\Delta \to \gamma\gamma$. The decay via Higgs mixing takes over in the region with larger $\theta_{\Delta h}$ and heavier Δ. On each relic curve, there is also an upper bound on $\theta_{\Delta h}$ otherwise Δ would decay too fast.

To justify the use of Eq. (3.10), we must verify that Δ decouples from the rest of the plasma while it was still ultrarelativistic. First of all, at temperatures around the LR symmetry breaking scale, Δ is in thermal equilibrium with heavy particles that receive their mass from v_Δ, which are $N_{2,3}$ and W_R, Z'. When the temperature of the Universe falls around the M_Δ, the $N_{2,3}$ and W_R, Z' particles already decouple from the thermal plasma because they are much heavier. The remaining processes to consider are similar to those in Fig. 9. Among them, the process $\Delta \leftrightarrow \gamma\gamma$ is suppressed by the heavy W_R mass and remains decoupled until the temperature of the Universe cools down to (set by $\Gamma_{\Delta\leftrightarrow\gamma\gamma} = H$)

$$T \sim 10^9 \text{ GeV}\left(\frac{M_\Delta}{1013 \text{ GeV}}\right)^{1/2}\left(\frac{10^3}{M_{W_R}/M_\Delta}\right). \tag{4.16}$$

This temperature is well below M_Δ, thus the inverse decay will be Boltzmann suppressed and never reach equilibrium. As discussed in Eq. (4.15), the other processes $\Delta \leftrightarrow N_1 N_1$ and $\Delta \leftrightarrow 4q$ have rates much smaller than $\Delta \leftrightarrow \gamma\gamma$ and cannot keep Δ thermalized either. Therefore, the decoupling of Δ must occur at a temperature between M_{W_R} and M_Δ. The particle mass spectrum considered for the Δ-dilution mechanism is indeed compatible with the assumption that Δ freezes out relativistically.

V. CONCLUSION AND OUTLOOK

In this work, we explore the entropy dilution mechanism for dark matter relic density in the minimal LRSM that also addresses the origin of neutrino mass. In this model, the lightest right-handed neutrino (N_1) is the sole dark matter candidate and its mass must be below the QCD scale in order to stay cosmologically stable. We first emphasize that N_1 always decouples relativistically from the right-handed current interactions and an entropy release afterwards must happen for producing the observed dark matter relic abundance in the Universe. This requires the presence of a "long-lived" diluting particle which comes to dominate the energy content of the Universe as a matter component, before decaying away mainly into SM particles. One of the heavier right-handed neutrinos (N_2) or the Higgs boson from spontaneous $SU(2)_R \times U(1)_{B-L}$ gauge symmetry breaking (Δ) can play the role of the diluting particle.

Our original contribution here is a new opportunity of such a mechanism in cosmology. The matter power spectrum for the large scale structure of the Universe is sensitive to the diluting particle's partial decay model into dark matter. When produced this way, dark matter can remain relativistic until the onset of recombination and suppress the primordial matter density perturbations. Through a detailed analysis, we derive an upper bound on such a decay branching ratio to be less than $\sim 1\%$, using the existing SDSS data. Such a large scale structure constraint is generic and can be applied to various dark matter models that require an entropy dilution mechanism. In the context of LRSM, the decay of N_2 into N_1 can happen via right-handed charged-current interaction (mediated by the $W_R^\pm$ gauge boson) and the decay of Δ into N_1 is tied to dark matter mass generation. Therefore, the large scale structure constraint plays a crucial role in determining the viable parameter space for the dark matter relic density. We carry out an anatomy of possible dark matter dilution scenarios in the left-right symmetric model:

(1) In the scenario of N_2 dilution, we point out that the decays of N_2 cannot be dominated by the right-handed currents, otherwise the $N_2 \to N_1$ branching ratio is too high ($\gtrsim 10\%$). Thus, additional decay modes must be present due to a N_2-light-neutrino mixing or $W - W_R$ gauge boson mixing. In both cases, we find that the mass scale of the W_R boson must be rather high, above the PeV scale. On the other hand, the dilutor N_2 can have a mass as low as the weak scale. We also derive the monochromatic x-ray constraint on dark matter N_1 from $W - W_R$ mixing, which further narrows down the viable mass range of N_1 to a mass window of 6.5–30 keV.

(2) The possibility of Δ dilution is original to this work. This scenario requires Δ to be lighter than the right-handed gauge bosons and neutrinos (except for N_1). We find the $\Delta \to \gamma\gamma$ mode to be the most useful for suppressing the $\Delta \to N_1 N_1$ decay and passing the strong large scale structure constraint. The corresponding mass scales of Δ and W_R need to be very high, close to the GUT scale. Because of this, the $W - W_R$ mixing contribution to N_1 dark matter radiative decay is negligible. The $N_1 \to \nu\gamma$ decay will only proceed via its mixing with light active neutrinos, as is the case of a regular sterile neutrino.

Based on the above results, we point out the following opportunities of diluted dark matter in the light of the upcoming experimental efforts:

(1) The primordial dark matter power spectrum will be more precisely measured by the upcoming large scale galaxy surveys, including Euclid and Rubin LSST [109,110]. Future high-redshift surveys such as MegaMapper and PUMA have the promise to

extend the precision measurement up to wave number $k \sim 0.9\ h/\mathrm{Mpc}$ [111]. A discovery of suppressed matter power spectrum, starting from $k \gtrsim 0.03\ h/\mathrm{Mpc}$, will serve as a smoking-gun evidence for dark matter entropy dilution mechanism in the early Universe with a nonzero dilutor to dark matter decay branching ratio.

(2) Related to the thermal dark matter population that gets diluted, future experimental facilities exploring the small scale structure of the Universe will be instrumental as well. Observations of low mass dark matter halos and the lensing of cosmic microwave background may allow the discovery and measurements of the dark matter mass, if it lies not far above the current lower bound (~6.5 keV) [112,113].

(3) Future experiments including ATHENA and XRISM [102–104] will search for monochromatic x-ray emission from the radiative decay of N_1 dark matter in the Milky Way and nearby galaxies. A positive measurement will be useful as another input to discriminate between the various dark matter dilution scenarios and neutrino mass generation mechanisms in the LRSM, and map out the favored parameter space.

On the theory side, a follow-up exercise will be to calculate the evolution of primordial matter density perturbations by taking into account the nonlinear terms in the Boltzmann equations, which start to become non-negligible (bringing in corrections of percent level or higher) for wave numbers $k \gtrsim 0.2\ h/\mathrm{Mpc}$ and observations made at redshift $z \approx 0$ [89,114,115]. Because the hot dark matter component from dilutor decay acts to suppress the perturbations, we expect the nonlinear effect to be smaller than the case of regular cold/warm dark matter. Nonetheless, a more careful analysis is warranted and can serve as a useful tool for discovering this piece of new physics in the upcoming precision cosmology era.

We wrap up with a further comment on the validity of leptogenesis in LRSM given the new constraints found in this work. We have considered two options of entropy dilution, using either N_2 or Δ. In both cases, we have left the mass and couplings of N_3 arbitrary, so it may also be degenerate with N_2 to fulfill the resonant leptogenesis condition [116]. In the case of Δ dilution, both $N_{2,3}$ have arbitrary masses and they can be degenerate or heavy and take on the role of lepton number asymmetry generators.

ACKNOWLEDGMENTS

We thank Weiyi Deng, Diego Redigolo, Filippo Sala, and Katelin Schutz for useful discussions and correspondence. M. N. is supported by the Slovenian Research Agency under the research core funding Grant No. P1-0035 and in part by the research Grants No. J1-3013, No. N1-0253, and No. J1-4389. Y. Z. is supported by a Subatomic Physics Discovery Grant (individual) from the Natural Sciences and Engineering Research Council of Canada, and by the Canada First Research Excellence Fund through the Arthur B. McDonald Canadian Astroparticle Physics Research Institute.

APPENDIX: RADIATIVE N_1 DECAY VIA $W - W_R$ MIXING

The radiative decay of dark matter to a monochromatic photon $N_1 \to \nu\gamma$ gives a very stringent constraint on its couplings from the x-ray spectra measurements, shown in Fig. 8. Here we provide some details on the calculation of the rate in Eq. (4.12), which has new sources within the LRSM. There are two possible contributions, one from the Dirac mixing of N_1 with ν, which is well known and the same as for the sterile neutrinos. In the presence of gauge boson mixing ξ_{LR}, another amplitude is present, coming from the SM-like W having a coupling to the right-handed charged current. In the small ξ_{LR} limit we have

$$\mathcal{L}_{\mathrm{CC}} \simeq \frac{g}{\sqrt{2}}[\bar{\nu}_\ell \gamma^\mu \mathbb{P}_L \ell + \xi_{\mathrm{LR}} (V^{R\dagger}_{\mathrm{PMNS}})_{i\ell} \bar{N}_i \gamma^\mu \mathbb{P}_R \ell] W^+_\mu + \mathrm{H.c.}, \tag{A1}$$

where $V^{R\dagger}_{\mathrm{PMNS}}$ is the right-handed PMNS matrix introduced in Eq. (2.5). With this coupling turned on, there are two new diagrams contributing to radiative N_1 decay, and their topologies are shown in Fig. 11.

The $N_1 \to \nu\gamma$ decay always occurs via the dimension-five effective operator

$$\mathcal{L}_{\mathrm{eff}} = C\bar{\nu}\sigma_{\mu\nu}\mathbb{P}_R N_1 F^{\mu\nu} + \mathrm{H.c.}, \tag{A2}$$

where C is the Wilson coefficient, to be determined next, and $\sigma_{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. The chiral projection operator in front of the N_1 field must be $\mathbb{P}_R = (1+\gamma_5)/2$. The corresponding decay amplitude for $N_1(p_1) \to \nu(p_2)\gamma(q)$ is

$$i\mathcal{M} = -iC\bar{u}_\nu(p_2, s_2)(\not{q}\not{\epsilon}^* - \not{\epsilon}^*\not{q})\mathbb{P}_R u_N(p_1, s_1), \tag{A3}$$

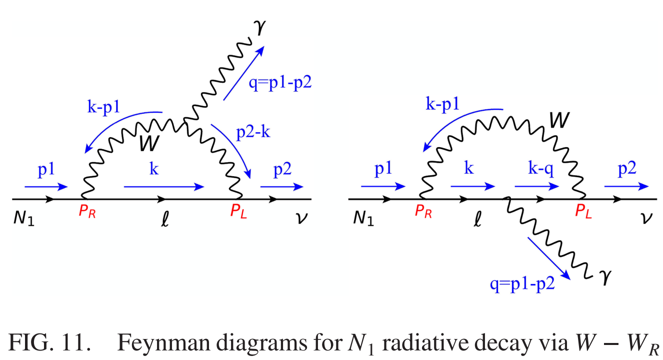


FIG. 11. Feynman diagrams for N_1 radiative decay via $W - W_R$ mixing in the LRSM. The $W - N_1 - \ell$ vertex labeled by $\mathbb{P}_R$ is induced by the $W - W_R$ mixing and the corresponding Feynman rule is obtained from the second term in Eq. (A1).

where ε^*_μ is the photon polarization vector and $q\cdot\varepsilon^*(q)=0$ for an on-shell transverse photon. The partial decay rate of $N_1\to\nu\gamma$ is

$$\Gamma_{N_1\to\nu\gamma}=\frac{1}{4\pi}|C|^2 m_{N_1}^3. \tag{A4}$$

For a Majorana N_1, it can also decay into $\bar{\nu}\gamma$ with the same partial rate.

With the momentum assignments shown in Fig. 11, the first diagram has an amplitude

$$\begin{aligned} i\mathcal{M}_1=&-\frac{eg^2}{2}\xi_{\rm LR}\sum_\ell (V^{R\dagger}_{\rm PMNS})_{1\ell}m_\ell\int\frac{{\rm d}^4k}{(2\pi)^4}\frac{1}{(k^2-m_\ell^2)[(k-p_1)^2-M_W^2][(p_2-k)^2-M_W^2]}\\ &\times\{(2k-p_1-p_2)\cdot\varepsilon^*\bar{u}_\nu(p_2,s_2)\gamma^\mu\gamma_\mu\mathbb{P}_R u_N(p_1,s_1)+\bar{u}_\nu(p_2,s_2)\not{\varepsilon}^*(\not{p}_2-\not{k}-\not{q})\mathbb{P}_R u_N(p_1,s_1)\\ &+\bar{u}_\nu(p_2,s_2)(\not{q}-\not{k}+\not{p}_1)\not{\varepsilon}^*\mathbb{P}_R u_N(p_1,s_1)\}, \end{aligned}\tag{A5}$$

where we have dropped terms that are suppressed by additional powers of $1/M_W$. For the γ matrices between the fermion spinors, we are interested in the structure $\not{q}\not{\varepsilon}^*-\not{\varepsilon}^*\not{q}$, as in Eq. (A3). This immediately implies that the first term in $\{\}$ does not contribute. In addition, because the external fermions already have the correct chirality, we can drop the chirality-flipping terms (upon equation of motion) such as $\not{p}_1$ acting on $u_N(p_1,s_1)$ and $\not{p}_2$ acting on $\bar{u}_\nu(p_2,s_2)$. This allows us to reduce the $\{\}$ bracket in Eq. (A5) into

$$\{\}\to\bar{u}(p_2,s_2)(\not{\varepsilon}^*(-\not{k}-2\not{q})+(2\not{q}-\not{k})\not{\varepsilon}^*)\mathbb{P}_R u(p_1,s_1). \tag{A6}$$

After completing the k integral, the remaining relevant term is

$$i\mathcal{M}_1=i\frac{eg^2}{16\pi^2}\xi_{\rm LR}\sum_\ell (V^{R\dagger}_{\rm PMNS})_{1\ell}\frac{m_\ell}{M_W^2}\bar{u}(p_2,s_2)(\not{q}\not{\varepsilon}^*-\not{\varepsilon}^*\not{q})\mathbb{P}_R u(p_1,s_1). \tag{A7}$$

The second diagram of Fig. 11 has an amplitude

$$\begin{aligned} i\mathcal{M}_2=&-\frac{eg^2}{2}\xi_{\rm LR}\sum_\ell (V^{R\dagger}_{\rm PMNS})_{1\ell}m_\ell\int\frac{{\rm d}^4k}{(2\pi)^4}\frac{1}{(k^2-m_\ell^2)[(k-q)^2-m_\ell^2][(k-p_1)^2-M_W^2]}\\ &\times\{\bar{u}_\nu(p_2,s_2)\gamma^\mu\not{\varepsilon}^*\not{k}\gamma_\mu\mathbb{P}_R u_N(p_1,s_1)+\bar{u}_\nu(p_2,s_2)\gamma^\mu(\not{k}-\not{q})\not{\varepsilon}^*\gamma_\mu\mathbb{P}_R u_N(p_1,s_1)\}. \end{aligned}\tag{A8}$$

Using the identity $\gamma^\mu\gamma^\beta\gamma^\rho\gamma_\mu=4g^{\beta\rho}$, all the γ matrices between $\bar{u}_\nu(p_2,s_2)$ and $u_N(p_1,s_1)$ are gone. Thus, we conclude that this diagram dose not contribute to the $N_1\to\nu\gamma$ decay.

Comparing Eqs. (A3) and (A7), we get

$$C=-\frac{eG_F\xi_{\rm LR}\sum_\ell (V^{R\dagger}_{\rm PMNS})_{1\ell}m_\ell}{2\sqrt{2}\pi^2}. \tag{A9}$$

The corresponding N_1 radiative decay rate is

$$\Gamma_{N_1\to\nu\gamma}=\frac{\alpha\xi_{\rm LR}^2}{8\pi^4}G_F^2 m_{N_1}^3\sum_\ell |(V^R_{\rm PMNS})_{\ell 1}|^2 m_\ell^2. \tag{A10}$$

This is how we get Eq. (4.12) in the main text.

[1] S. Dodelson and L. M. Widrow, Phys. Rev. Lett. **72**, 17 (1994).
[2] X.-D. Shi and G. M. Fuller, Phys. Rev. Lett. **82**, 2832 (1999).
[3] E. O. Nadler *et al.* (DES Collaboration), Phys. Rev. Lett. **126**, 091101 (2021).
[4] K. N. Abazajian, Phys. Rep. **711–712**, 1 (2017).
[5] A. De Gouvêa, M. Sen, W. Tangarife, and Y. Zhang, Phys. Rev. Lett. **124**, 081802 (2020).
[6] K. J. Kelly, M. Sen, W. Tangarife, and Y. Zhang, Phys. Rev. D **101**, 115031 (2020).

[7] K. J. Kelly, M. Sen, and Y. Zhang, Phys. Rev. Lett. **127**, 041101 (2021).
[8] R. An, V. Gluscevic, E. O. Nadler, and Y. Zhang, Astrophys. J. Lett. **954**, L18 (2023).
[9] J. C. Pati and A. Salam, Phys. Rev. D **10**, 275 (1974); **11**, 703(E) (1975).
[10] H. Fritzsch and P. Minkowski, Ann. Phys. (N.Y.) **93**, 193 (1975).
[11] R. N. Mohapatra and G. Senjanović, Phys. Rev. Lett. **44**, 912 (1980).
[12] R. N. Mohapatra and J. C. Pati, Phys. Rev. D **11**, 2558 (1975).
[13] G. Senjanović and R. N. Mohapatra, Phys. Rev. D **12**, 1502 (1975).
[14] G. Senjanović, Nucl. Phys. **B153**, 334 (1979).
[15] M. Nemevšek, G. Senjanović, and V. Tello, Phys. Rev. Lett. **110**, 151802 (2013).
[16] W.-Y. Keung and G. Senjanović, Phys. Rev. Lett. **50**, 1427 (1983).
[17] M. Nemevsek, F. Nesti, G. Senjanovic, and Y. Zhang, Phys. Rev. D **83**, 115014 (2011).
[18] R. N. Mohapatra and G. Senjanović, Phys. Rev. D **23**, 165 (1981).
[19] V. Tello, M. Nemevšek, F. Nesti, G. Senjanović, and F. Vissani, Phys. Rev. Lett. **106**, 151801 (2011).
[20] V. Cirigliano, A. Kurylov, M. J. Ramsey-Musolf, and P. Vogel, Phys. Rev. D **70**, 075007 (2004).
[21] M. Nemevsek, F. Nesti, G. Senjanovic, and V. Tello, arXiv:1112.3061.
[22] R. N. Mohapatra and X. Zhang, Phys. Rev. D **46**, 5331 (1992).
[23] A. S. Joshipura, E. A. Paschos, and W. Rodejohann, Nucl. Phys. **B611**, 227 (2001).
[24] J.-M. Frere, T. Hambye, and G. Vertongen, J. High Energy Phys. 01 (2008) 051.
[25] P. S. Bhupal Dev, C.-H. Lee, and R. N. Mohapatra, Phys. Rev. D **90**, 095012 (2014).
[26] F. Bezrukov, H. Hettmansperger, and M. Lindner, Phys. Rev. D **81**, 085032 (2010).
[27] S. Tremaine and J. Gunn, Phys. Rev. Lett. **42**, 407 (1979).
[28] A. Boyarsky, O. Ruchayskiy, and D. Iakubovskyi, J. Cosmol. Astropart. Phys. 03 (2008) 005.
[29] M. Viel, G. D. Becker, J. S. Bolton, and M. G. Haehnelt, Phys. Rev. D **88**, 043502 (2013).
[30] V. Iršič *et al.*, Phys. Rev. D **96**, 023522 (2017).
[31] D. Gilman, S. Birrer, A. Nierenberg, T. Treu, X. Du, and A. Benson, Mon. Not. R. Astron. Soc. **491**, 6077 (2020).
[32] J.-W. Hsueh, W. Enzi, S. Vegetti, M. Auger, C. D. Fassnacht, G. Despali, L. V. E. Koopmans, and J. P. McKean, Mon. Not. R. Astron. Soc. **492**, 3047 (2020).
[33] W. Enzi *et al.*, Mon. Not. R. Astron. Soc. **506**, 5848 (2021).
[34] E. O. Nadler, S. Birrer, D. Gilman, R. H. Wechsler, X. Du, A. Benson, A. M. Nierenberg, and T. Treu, Astrophys. J. **917**, 7 (2021).
[35] R. J. Scherrer and M. S. Turner, Phys. Rev. D **31**, 681 (1985).
[36] M. Nemevšek, G. Senjanović, and Y. Zhang, J. Cosmol. Astropart. Phys. 07 (2012) 006.
[37] M. Nemevšek and Y. Zhang, Phys. Rev. Lett. **130**, 121002 (2023).
[38] T. Moroi and L. Randall, Nucl. Phys. **B570**, 455 (2000).
[39] E. A. Baltz and H. Murayama, J. High Energy Phys. 05 (2001) 067.
[40] T. Asaka, M. Shaposhnikov, and A. Kusenko, Phys. Lett. B **638**, 401 (2006).
[41] J. Hasenkamp and J. Kersten, Phys. Rev. D **82**, 115029 (2010).
[42] G. Arcadi and P. Ullio, Phys. Rev. D **84**, 043520 (2011).
[43] Y. Zhang, J. Cosmol. Astropart. Phys. 05 (2015) 008.
[44] A. V. Patwardhan, G. M. Fuller, C. T. Kishimoto, and A. Kusenko, Phys. Rev. D **92**, 103509 (2015).
[45] Z. Chacko, N. Craig, P. J. Fox, and R. Harnik, J. High Energy Phys. 07 (2016) 023.
[46] M. Cirelli, P. Panci, K. Petraki, F. Sala, and M. Taoso, J. Cosmol. Astropart. Phys. 05 (2016) 036.
[47] A. Soni, H. Xiao, and Y. Zhang, Phys. Rev. D **96**, 083514 (2017).
[48] R. Contino, A. Mitridate, A. Podo, and M. Redi, J. High Energy Phys. 02 (2019) 187.
[49] J. A. Evans, A. Ghalsasi, S. Gori, M. Tammaro, and J. Zupan, J. High Energy Phys. 02 (2019) 151.
[50] C. Cosme, M. Dutra, T. Ma, Y. Wu, and L. Yang, J. High Energy Phys. 03 (2020) 026.
[51] J. A. Dror, D. Dunsky, L. J. Hall, and K. Harigaya, J. High Energy Phys. 07 (2020) 168.
[52] P. Chanda and J. Unwin, J. Cosmol. Astropart. Phys. 06 (2021) 032.
[53] P. Asadi, T. R. Slatyer, and J. Smirnov, Phys. Rev. D **106**, 015012 (2022).
[54] S. Baumholzer and P. Schwaller, J. Cosmol. Astropart. Phys. 06 (2021) 013.
[55] K. Bleau, J. Bramante, and C. Cappiello, J. Cosmol. Astropart. Phys. 01 (2024) 021.
[56] J. Basecq, Scalar sector in $SU(2)_L \times SU(2)_R \times U(1)_{(B-L)}$ models, Other thesis, Carnegie Mellon University, 1986.
[57] J. F. Gunion, J. Grifols, A. Mendez, B. Kayser, and F. I. Olness, Phys. Rev. D **40**, 1546 (1989).
[58] K. Kiers, J. Kolb, J. Lee, A. Soni, and G.-H. Wu, Phys. Rev. D **66**, 095002 (2002).
[59] N. G. Deshpande, J. F. Gunion, B. Kayser, and F. I. Olness, Phys. Rev. D **44**, 837 (1991).
[60] P. Duka, J. Gluza, and M. Zralek, Ann. Phys. (N.Y.) **280**, 336 (2000).
[61] O. Khasanov and G. Perez, Phys. Rev. D **65**, 053007 (2002).
[62] W. Dekens and D. Boer, Nucl. Phys. **B889**, 727 (2014).
[63] G. Bambhaniya, J. Chakrabortty, J. Gluza, T. Jelinski, and R. Szafron, Phys. Rev. D **92**, 015016 (2015).
[64] P. S. B. Dev, R. N. Mohapatra, and Y. Zhang, J. High Energy Phys. 05 (2016) 174.
[65] A. Maiezza, M. Nemevšek, and F. Nesti, Phys. Rev. D **94**, 035008 (2016).
[66] P. S. Bhupal Dev, R. N. Mohapatra, W. Rodejohann, and X.-J. Xu, J. High Energy Phys. 02 (2018) 154.
[67] V. Brdar, L. Graf, A. J. Helmboldt, and X.-J. Xu, J. Cosmol. Astropart. Phys. 12 (2019) 027.
[68] Y. Zhang, H. An, X. Ji, and R. N. Mohapatra, Nucl. Phys. **B802**, 247 (2008).

[69] Y. Zhang, H. An, X. Ji, and R. N. Mohapatra, Phys. Rev. D **76**, 091301 (2007).
[70] A. Maiezza, M. Nemevšek, F. Nesti, and G. Senjanović, Phys. Rev. D **82**, 055022 (2010).
[71] G. Senjanović and V. Tello, Phys. Rev. Lett. **114**, 071801 (2015).
[72] G. Senjanović and V. Tello, Phys. Rev. D **94**, 095023 (2016).
[73] A. Maiezza and M. Nemevšek, Phys. Rev. D **90**, 095002 (2014).
[74] S. Bertolini, A. Maiezza, and F. Nesti, Phys. Rev. D **101**, 035036 (2020).
[75] S. Bertolini, L. Di Luzio, and F. Nesti, Phys. Rev. Lett. **126**, 081801 (2021).
[76] S. Weinberg, Phys. Rev. Lett. **36**, 294 (1976).
[77] A. D. Linde, JETP Lett. **23**, 64 (1976).
[78] J. Basecq and D. Wyler, Phys. Rev. D **39**, 870 (1989).
[79] M. Nemevšek, F. Nesti, and J. C. Vasquez, J. High Energy Phys. 04 (2017) 114.
[80] M. Nemevšek, F. Nesti, and G. Popara, Phys. Rev. D **97**, 115018 (2018).
[81] M. Nemevšek and F. Nesti, Phys. Rev. D **108**, 015030 (2023).
[82] G. Senjanović and V. Tello, Phys. Rev. Lett. **119**, 201803 (2017).
[83] G. Senjanović and V. Tello, Phys. Rev. D **100**, 115031 (2019).
[84] J. Kiers, K. Kiers, A. Szynkman, and T. Tarutina, Phys. Rev. D **107**, 075001 (2023).
[85] P. A. Zyla *et al.* (Particle Data Group), Prog. Theor. Exp. Phys. **2020**, 083C01 (2020).
[86] N. Aghanim *et al.* (Planck Collaboration), Astron. Astrophys. **641**, A6 (2020).
[87] J. Lesgourgues, arXiv:1104.2932.
[88] D. Blas, J. Lesgourgues, and T. Tram, J. Cosmol. Astropart. Phys. 07 (2011) 034.
[89] J. Lesgourgues and T. Tram, J. Cosmol. Astropart. Phys. 09 (2011) 032.
[90] S. Roychowdhury, J. N. Chengalur, A. Begum, and I. D. Karachentsev, Mon. Not. R. Astron. Soc. **404**, L60 (2010).
[91] K.-G. Lee *et al.*, Astron. J. **145**, 69 (2013).
[92] F. D'Eramo and A. Lenoci, J. Cosmol. Astropart. Phys. 10 (2020) 045.
[93] Q. Decant, J. Heisig, D. C. Hooper, and L. Lopez-Honorez, J. Cosmol. Astropart. Phys. 03 (2021) 041.
[94] J. A. Casas and A. Ibarra, Nucl. Phys. **B618**, 171 (2001).
[95] A. Davidson, Phys. Rev. D **20**, 776 (1979).
[96] R. N. Mohapatra and R. E. Marshak, Phys. Rev. Lett. **44**, 1316 (1980); **44**, 1643(E) (1980).
[97] R. Shrock, Phys. Rev. D **9**, 743 (1974).
[98] P. B. Pal and L. Wolfenstein, Phys. Rev. D **25**, 766 (1982).
[99] A. Boyarsky, D. Malyshev, A. Neronov, and O. Ruchayskiy, Mon. Not. R. Astron. Soc. **387**, 1345 (2008).
[100] A. Merle and V. Niro, Phys. Rev. D **88**, 113004 (2013).
[101] K. C. Y. Ng, B. M. Roach, K. Perez, J. F. Beacom, S. Horiuchi, R. Krivonos, and D. R. Wik, Phys. Rev. D **99**, 083005 (2019).
[102] A. Neronov and D. Malyshev, Phys. Rev. D **93**, 063518 (2016).
[103] XRISM Science Team, arXiv:2003.04962.
[104] C. Dessert, O. Ning, N. L. Rodd, and B. R. Safdi, arXiv: 2305.17160.
[105] S. Colombi, S. Dodelson, and L. M. Widrow, Astrophys. J. **458**, 1 (1996).
[106] M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese, and A. Riotto, Phys. Rev. D **71**, 063534 (2005).
[107] R. N. Mohapatra, Adv. Ser. Dir. High Energy Phys. **3**, 384 (1989).
[108] A. Djouadi, Phys. Rep. **457**, 1 (2008).
[109] A. Ealet, Cosmology with the EUCLID satellite: Mapping the large structures of the universe probing dark energy (2017), https://indico.cern.ch/event/527550/.
[110] S. Ferraro, N. Sailer, A. Slosar, and M. White, arXiv: 2203.07506.
[111] N. Sailer, E. Castorina, S. Ferraro, and M. White, J. Cosmol. Astropart. Phys. 12 (2021) 049.
[112] S. Chakrabarti *et al.*, in *Snowmass 2021* (2022), arXiv: 2203.06200.
[113] A. Drlica-Wagner *et al.*, arXiv:2209.08215.
[114] V. Junk and E. Komatsu, Phys. Rev. D **85**, 123524 (2012).
[115] J. J. M. Carrasco, S. Foreman, D. Green, and L. Senatore, J. Cosmol. Astropart. Phys. 07 (2013) 057.
[116] A. Pilaftsis and T. E. J. Underwood, Nucl. Phys. **B692**, 303 (2004).