Ultra-light scalar saving the 3+1 neutrino scheme from the cosmological bounds

Yasaman Farzan

Institute for Research in Fundamental Sciences (IPM),
P.O. Box 19395-5531, Tehran, Iran
The Abdus Salam ICTP, Strada Costiera 11, 34151, Trieste, Italy

The LSND and MiniBooNE results as well as the reactor and Gallium anomalies seem to indicate the presence of a sterile neutrino with a mass of $\sim 1$ eV mixed with active neutrinos. Such sterile neutrino can be produced in the early universe before the neutrino decoupling, leading to a contribution to the effective number of neutrinos ($N_{\text{eff}}$) as well as to a contribution to the sum of neutrino masses which are in tension with cosmological observations. We propose a scenario to relax this tension by a Yukawa coupling of the sterile neutrinos to ultra-light scalar particles which contribute to the dark matter in the background. The coupling induces an effective mass for $\nu_s$ which prevents its production in the early universe. We discuss the implications for the upcoming KATRIN experiment and future relic neutrino search experiments such as PTOLEMY. We also briefly comment on certain non-renormalizable forms of interaction between $\nu_s$ and the scalar and their consequences for the $\nu_s$ production in the early universe.
I. INTRODUCTION

The 3 neutrino mass and mixing scheme has been established as the standard paradigm to explain the results from various solar, atmospheric, long baseline and reactor neutrino experiments. However, there are a few hints that may point out the existence of a fourth sterile neutrino ($\nu_s$) with a mass of $\sim 1$ eV mixed with active neutrinos. This is the essence of the so-called 3+1 neutrino scheme which has been invoked to explain the LSND [1] and MiniBooNE [2] as well as the Gallium and reactor neutrino anomalies [3, 4, 5]. To explain the LSND and MiniBooNE anomalies, $\nu_\mu$ should partially convert en-route into $\nu_s$ which implies that the sterile neutrino has to be mixed with $\nu_e$ and $\nu_\mu$, simultaneously. From ICECUBE and MINOS+, strong bounds are derived on the $\nu_s$ mixing with $\nu_\mu$ shedding doubt on the 3+1 oscillation solution to the LSND and MiniBooNE anomalies [3, 4, 5]. However, the reactor [4, 5, 6] and Gallium [3] anomalies (the observation that at short baselines $P(\nu_e \rightarrow \nu_e), P(\nu_\mu \rightarrow \bar{\nu}_e) < 1$) can be explained even if $\nu_s$ mixes only with $\nu_e$ so this solution is not ruled out by the ICECUBE or MINOS+ results which are based on the $\nu_\mu \rightarrow \nu_\mu$ observation. A recent analysis shows that the solutions to the Gallium and reactor anomalies are compatible with each other within the 3+1 neutrino mixing scheme with $|U_{e4}|^2 \sim 0.01 - 0.02$ [10]. What makes this possibility even more exciting is that a sterile neutrino mixed with $\nu_e$ will lead to observable kinks in the spectrum of beta decay [11].

On the other hand, the mixing of $\nu_s$ with $\nu_e$ implies that in the early universe before neutrinos decouple from the plasma, the neutrino oscillation brings the sterile neutrinos to thermal equilibrium with the active ones [13]. This means the effective relativistic degrees of freedom will increase by 1 unit ($N_{\text{eff}} = 4$) which is disfavored by the CMB data [14] as well as by the Big Bang Nucleosynthesis (BBN). Moreover, the production of $\nu_s$ with a mass of 1 eV in the early universe will violate the upper bound on the sum of neutrino masses which are derived by combining the CMB and BAO results [14].

To avoid the bounds from cosmology, various models for self-interaction of $\nu_s$ has been proposed [15–19]. (See, however, [20, 22].) The essence of all these scenarios is that the self-interaction of $\nu_s$ will induce an effective mass for $\nu_s$ at $T > \text{MeV}$ which will suppress the effective mixing and will therefore decrease the $\nu_s$ to $\nu_\alpha$ oscillation probability, $P(\nu_s \rightarrow \nu_\alpha)$. Notice that within these scenarios, the generated effective mass itself is given by the $\nu_s$ density. That is in order for the mechanism to be efficient, a nonzero $\nu_s$ density is required in the first place. This way the bound on $N_{\text{eff}}$ can be satisfied but the bound on the sum of masses cannot be avoided. A recent thorough study shows that by adding the BAO data, the self-interaction scenario of $\nu_s$ characterized by an effective four-Fermion interaction will be still ruled out by the BAO data [23]. However, if the scenario involves light states coupled to $\nu_4$ which open up the possibility of the removal of $\nu_4$ by annihilation [24] or decay before the onset of structure formation (before the matter radiation equality) the BAO+CMB bound can be avoided, too. An alternative remedy is the late phase transition scenario proposed in [22]. A non-renormalizable coupling of $\nu_s$ to background scalar is suggested in [20] and its consequences for the $\nu_s$ abundance is discussed.

Ref. [27] proposes a $U(1)$ gauge model with a gauge boson of mass 10 eV coupled to $\nu_s$ as well as to asymmetric dark matter. The coupling creates an effective mass for $\nu_4$ proportional to dark matter density which is sizable even for vanishing $\nu_s$ density. The effective mixing at $T > \text{MeV}$ will be then suppressed, preventing the $\nu_s$ production. Moreover, the new gauge interaction opens up the possibility of relatively fast decay of the $\nu_4$ components of the active neutrinos before the onset of structure formation. As a result, both the bound on $N_{\text{eff}}$ from CMB and BBN and the bound on the sum of neutrino masses from BAO and CMB can be satisfied.

It is well-known that if dark matter (or a component of it) is of bosonic type, it can be as light as $\sim 10^{-21}$ eV. Despite its small mass, the ultra-light dark matter is considered to be cold because its production is non-thermal. These particles can be non-relativistic even at high temperatures. Recently such dark matter has gained popularity in the literature as it has been advocated as a solution to the small scale structure tensions within the WIMP scenario [28]. As long as their de Broglie wavelength is larger than their average distance with each other, they can be described

II. THE MODEL

It is well-known that if dark matter (or a component of it) is of bosonic type, it can be as light as $\sim 10^{-21}$ eV. Despite its small mass, the ultra-light dark matter is considered to be cold because its production is non-thermal. These particles can be non-relativistic even at high temperatures. Recently such dark matter has gained popularity in the literature as it has been advocated as a solution to the small scale structure tensions within the WIMP scenario [28]. As long as their de Broglie wavelength is larger than their average distance with each other, they can be described
by a classical field. In particular, a real ultra-light scalar dark matter can be described as

$$\phi = \frac{\sqrt{2\rho_0}}{m_\phi} \cos(m_\phi t - \vec{p}_\phi \cdot \vec{x})$$  \hspace{1cm} (1)$$

where $|\vec{p}_\phi| \ll m_\phi$. For $t > 1/m_\phi$, $\rho_0$ like all non-relativistic relics scales as $T^3$. For $t \ll 1/m_\phi$, it can be shown that $\rho_0$ (the 00 element of the energy-momentum tensor, $T_{\mu\nu}$) is equal to minus the pressure, $-\rho_0$ ($T_{ii}$). Thus, the relation $T_{\mu\nu}^{\text{int}} = 0$ or equivalently $\rho_0 + 3H(\rho_0 + p_0) = 0$ implies that for $t \ll 1/m_\phi$, $\rho_0$ and therefore the amplitude of $\phi$ remains constant.

Refs. \[29–31\] assume a Yukawa coupling between complex ultra-light scalar DM and leptons and studies its impact on the flavor ratios of cosmic neutrinos detected by ICECUBE. Here, we assume a Yukawa coupling of the following form between $\phi$ and $\nu_s$

$$\lambda_\phi \nu_s^T c \nu_s + \text{H.c.}$$  \hspace{1cm} (2)$$

where $c$ is an asymmetric $2 \times 2$ matrix with components equal to $\pm 1$. Notice that this coupling is renormalizable and invariant under the SM gauge group. As long as $c$ is lighter than the lightest neutrino mass eigenstate, $\phi$ remains stable and therefore a suitable dark matter candidate. Notice that we could write the interaction of type $\phi \bar{\nu}_\tau \nu_s$ with similar results but to avoid adding new degrees of freedom, we stick to this Majorana form which does not require right-handed component for $\nu_s$.

The coupling in Eq. (2) induces an effective mass for $\nu_s$ given by

$$m_{\text{eff}} = \lambda \sqrt{2\rho_0/m_\phi} \cos(m_\phi t).$$  \hspace{1cm} (3)$$

Taking $m_\phi < 5 \times 10^{-17} \text{eV} = \frac{1}{13 \text{sec}}$, for up to after neutrino decoupling (to be precise until $T \sim 0.22 \text{MeV}(m_\phi/(5 \times 10^{-17} \text{eV}))^{1/2}$, $m_{\text{eff}}$ remains almost constant and equal to $m_{\text{eff}} = \lambda \sqrt{2\rho_0/m_\phi}$ where $\rho_0$ is the value of $\rho_0$ at $t = 1/m_\phi$. Taking for example $\rho_0^{\text{int}} = \rho_{\text{DM}}(0.22 \text{MeV} \sqrt{m_\phi/(5 \times 10^{-17} \text{eV})})^3$ (where the 0 superscript denotes the values today), we find $m_{\text{eff}} = 2.3 \times 10^{24} \text{eV} \sqrt{5 \times 10^{-17} \text{eV}/m_\phi}$.

Notice that the format of the effective mass that $\nu_s$ receives is of the Lorentz invariant Majorana type which should be summed with the $\nu_s$ mass in vacuum ($m_{\nu_s}$) to obtain dispersion relation i.e., $E_{\nu_s}^2 - |\vec{p}_{\nu_s}|^2 = (m_{\text{eff}} + m_{\nu_s})^2$. Using the superradiance argument, a vector dark matter with a mass of $6 \times 10^{-20} - 2 \times 10^{-17} \text{eV}$ is constrained \[33\] but these bounds do not apply for the scalar dark matter. The superradiance bound from M87* rules out only the scalar dark matter of mass of $10^{-21} \text{eV}$ and lower \[34\].

Remember that in the case of the propagation of the active neutrinos in medium, the Lorentz violating effective mass of active neutrinos in medium (e.g., $2\sqrt{2} G_F n_a \nu^i \nu_e$) is added to $m_{\nu_s}^2/(2E_{\nu})$ to obtain the Hamiltonian governing the neutrino flavor evolution. Here, the Lorentz conserving $m_{\text{eff}}$ should be added to $m_{\nu_s}$ rather than to $m_{\nu_s}^2/E_{\nu}$. In the presence of $m_{\text{eff}} \gg m_{\nu_s}$, we can write the effective active sterile mixing angle as

$$\sin 2\theta_m|_T = \sin 2\theta_m \frac{m_{\nu_s}}{m_{\text{eff}}} = \left\{ \begin{array}{ll}
\sin 2\theta_m & \text{at } t \ll m_\phi^{-1} \\
\sin 2\theta_m \left( \frac{0.22 \text{MeV} \sqrt{m_\phi/5 \times 10^{-17} \text{eV}}}{T} \right)^{3/2} & \text{at } t \gg m_\phi^{-1} \end{array} \right.$$  \hspace{1cm} (4)$$

where $\theta$ is the mixing angle in vacuum. At early universe when $T > 1 \text{MeV}$, the active neutrinos undergo scattering off the neighboring neutrinos and electrons. Each electroweak scattering will convert them to coherent active states without any $\nu_s$ component. To compute the oscillation probabilities, the evolution of full density matrix has to be computed \[13\] which is beyond the scope of the present paper. However, for $\sin^2 2\theta_m \ll 1$, a simplified estimate can be made as follows \[33\]: The rate of $\nu_a$ to $\nu_s$ conversion, $\Gamma_{\nu_a \rightarrow \nu_s}$, can be estimated as

$$\Gamma_{\nu_a \rightarrow \nu_s} = \frac{\sin^2 2\theta_m}{4\tau_\nu}$$

where $\tau_\nu^{-1}$ is the interaction rate of neutrinos $\tau_\nu^{-1} \sim G_F^2 T^5$. Thus, the contribution to $N_{\text{eff}}$ can be evaluated as

$$\delta N_{\text{eff}} = \int_{T_{\min}}^{T_{\max}} \Gamma_{\nu_a \rightarrow \nu_s} dt = \frac{\sin^2 2\theta_m}{4} \int_{T_{\min}}^{T_{\max}} \frac{1}{\tau_\nu} dt,$$
where $T_{\text{min}}$ is the neutrino decoupling temperature and $T_{\text{max}}$ is the maximum temperature for which $(\Delta m^2/T) t \gtrsim 1$. Notice that we use the fact that for $m_\phi \lesssim 5 \times 10^{-17}$ eV up until $T_{\text{min}}$, $m_{\nu f f}$ and therefore $\sin^2 2\theta^\text{int}_{m}$ remains constant. Within the canonical 3+1 scheme (in the limit $\sin 2\theta = \sin 2\theta_0$, Ref. [12] shows that for $\sin^2 2\theta \sim 4 \times 10^{-4}$ and $\Delta m^2 \sim 1$ eV$^2$, the contribution to $N_{\nu f f}$ is reduced to 0.1. Scaling these results, we conclude that taking $\sin^2 2\theta^\text{int}_{m} = 4 \times 10^{-5}$, the contribution will be less than $O(0.01)$ and therefore negligible. For $\Delta m^2 \sim 3$ eV$^2$ and $|U_{e4}|^2 \sim 2 \times 10^{-2}$ (a typical solution to the Gallium and reactor neutrino anomalies [10] which is consistent with the most recent DANSS and STEREO bounds [26]), $\sin^2 2\theta^\text{int}_{m} = 4 \times 10^{-5}$ can be achieved with $m_{\nu f f} > 40$ eV which for $\phi^\text{int}$ corresponding to $\rho_{DM}$ implies $\lambda > 2 \times 10^{-23}$. That is taking $\lambda \lesssim 2 \times 10^{-23}(m_{\nu f f}/1$ eV), the bound on $N_{\nu f f}$ can be safely relaxed but below $T \sim 0.01$ MeV (well above the matter radiation equality era) as well as in the Milky Way, $m_{\nu f f}$ can be neglected because $\rho_\phi$ and therefore the amplitude of $\phi$ will be suppressed.

Let us now discuss how the bounds from BAO and CMB on the sum of neutrino masses can be avoided. As we discussed, by choosing $\lambda > 10^{-23}$, the density of the $\nu_s$ particles produced at $T \sim$ MeV can be reduced to an arbitrarily small value. The contribution of them to the sum of the neutrino masses can be estimated as $\delta N_{\nu f f} m_{\nu s}$. Thus, as long as $\delta N_{\nu f f} \lesssim 0.01$, the contribution is well below the bound on the sum of neutrino masses, $\sum_{\nu} m_{\nu}$ [14].

For $t \gtrsim 1/m_{\phi}$, the $\phi$ field will start oscillating so $m_{\nu f f}$ can be even negative. This will have two dramatic consequences: (1) At certain epochs, $\nu_s$ can become lighter than even $m_{\phi}$, opening the possibility of decay of $\phi$ to $\nu_s$; (2) $\nu_s$ can become degenerate with active neutrino, paving the way for non-adiabatic conversion of active neutrinos to $\nu_s$ despite the fact that $m_{\nu f f} = m_\phi m_{\nu f f} \tan(m_\phi t) \ll m_{\nu s}$. Let us discuss the consequences of each case.

Even when $\nu_s$ becomes lighter than $\phi$, the perturbative lifetime of $\phi$ (i.e., $4\pi/\lambda m_\phi$) will be greater than the age of universe, however; as shown in [38], the $\phi$ field can convert into $\nu_s$ and $\bar{\nu}_s$ pairs through a mechanism known as parametric resonance production. During the epoch of our interest, the radiation dominates so $\rho_\phi \ll \rho_\nu$. Thus, even if $\phi$ completely decays into $\nu_s$, the effects of the produced $\nu_s$ on cosmological observation will be negligible. If $\phi$ completely decays, another particle should play the role of dark matter.

The non-adiabatic conversion of $\nu_1$, $\nu_2$, and $\nu_3$ to $\nu_4$ can lead to a tension with the bounds from BAO and CMB on the sum of neutrino masses. However, such tension can be avoided by opening up decay modes for $\nu_4$, which can result in decay $\nu_1 \rightarrow \bar{\nu}_1 \phi$ with a rate of $\lambda^2 \Delta m^2_4 |U_{e4}|^2 |U_{\nu_1}|^2/(4\pi E_\nu)$ where we have neglected the $\phi$ mass. Notice that the lifetime of $\nu_1$ relative to that of $\nu_4$ will be longer by a factor of $(\Delta m^2_3 |U_{e3}|^2)/(|\Delta m^2_4| |U_{e4}|^2)$. This means if we choose $\lambda'$ in a range that $\nu_4$ decays during $T = \text{few eV-10 eV}$, the rest of $\nu_1$ will be free streaming at the matter radiation equality era [39] but they can decay after recombination era. With a single nonzero $U_{e4}$ (a $\{\ell, \mu, \tau\}$), the unitarity of the neutrino $4 \times 4$ mixing matrix implies that all elements $U_{e1}$, $U_{e2}$ and $U_{e3}$ should be nonzero so the coupling $\lambda' \phi^2 u_{\nu_3}^T c_{\nu_3}$ leads to eventual decay of all neutrinos to the lightest mass eigenstate; i.e., $j = 1$ for normal ordering and $j = 3$ for inverted ordering. This also means that the decay of lighter $\nu_i$ does not takes place for cosmic neutrinos [39]. The required lifetime of $\nu_i$ can be achieved with $\lambda' = 3.5 \times 10^{-12}$. With such small $\lambda'$, the lightest neutrino mass eigenstates, $\nu_1$, $\nu_2$, and $\nu_3$ as well as the produced $\phi$ will be free streaming during 0.1 eV < $T$ < MeV. Taking $\phi$ massless or with a mass much smaller than that of $\nu_i$, its contribution to the $\sum_{\nu} m_{\nu}$ measurement from CMB and BAO will be negligible. The total energy in the form $l \nu_1$, $\nu_2$, and $\nu_3$ and $\phi$ will be equal to that of three neutrinos within the standard scheme leading to the same signatures as the standard 3$\nu$ scheme. Thus, the bounds from CMB and BAO can be safely satisfied. With $\lambda' = 3.5 \times 10^{-12}$, the lifetime of $\nu_4$ will be too long to be relevant for terrestrial, solar and even galactic supernova neutrinos.

If instead of the renormalizable Yukawa coupling in Eq. (2), we had taken a non-renormalizable coupling of form $\phi^2 u_{\nu_3}^T c_{\nu_3}$ as $[28]$ or of form $i(\phi^* \partial_\mu \phi - \partial_\mu \phi^*) \bar{\nu} \gamma^\mu \nu$ as $[32]$, $m_{\nu f f}$ could not have become negative so the two consequences of $m_{\nu} + m_{\nu f f} \rightarrow 0$ enumerated above would not have applied. The active background neutrinos at the start of the $\phi$ oscillation are mainly composed of $\nu_1$, $\nu_2$ and $\nu_3$, with a small contribution given by $\sin^2 \theta_{m}$ of $\bar{\nu}_4$ where $\sim$ emphasizes that these are the energy eigenstates inside the (dark) matter medium. The neutrino propagation for the aforementioned non-renormalizable coupling will remain adiabatic so after the amplitude of $\phi$ diminishes due to the expansion, the background will mainly consist of the vacuum mass eigenstates $\nu_1$, $\nu_2$ and $\nu_3$ with a small contribution from $\nu_4$ given by $\sin^2 \theta_{m}/(\sin \theta_{m}) \sim 10^{-5}$. As a result, the contribution from $\nu_4$ to $\sum_{\nu} m_{\nu}$ will be negligible so satisfying the bounds from CMB and BAO on $\sum_{\nu} m_{\nu}$ will not require a $\nu_4$ decay mechanism.

---

1. To be precise, the degeneracy between mass eigenstates will be broken by $\Delta E = V_\nu$ where $V_\nu$ is the effective potential for active neutrinos which at $T < m_{\phi}$ is composed of a contribution from symmetric neutrino background $2G\nu^2 T^3$ plus a contribution from the asymmetric electron background $\sqrt{2}G F_6 (m_{\nu} - n_{e})$ [25]. However, at these temperatures both of these quantities are very small, satisfying the non-adiabaticity condition: $4\phi_{m}(\Delta E)_{\text{resonance}} = \lambda \sqrt{2 \phi_{m}^2 / (\sin \theta \cos \theta_{m})} \gg 1$.

2. However, see [41] which argues that the condition of free streaming might be relaxed.
Let us now discuss the stability of the $\phi$ mass in the presence of the $\lambda$ coupling. This coupling is similar to the top Yukawa coupling in the SM and will similarly induce a quadratically divergent mass for $\phi$. Like the standard model, we assume that there is a yet unknown mechanism (e.g., SUSY) which cancels this divergent contribution. Still to have a “natural model”, we should check whether the finite part of the contribution, $\lambda m_{\nu_s}/(4\pi)$ is smaller than $m_\phi$. Taking $\lambda \sim 10^{-23}$, we see that this condition is readily satisfied. At $T \sim 10 - 20$ MeV when $\nu_a \rightarrow \nu_s$ may start, a “thermal” mass of $m_{\nu_s} \sim n_{\nu_s}/\sqrt{2}$ is induced in which $n_{\nu_s}$ is the number density of the produced $\nu_s$. Remembering that $\sqrt{n_{\nu_s}} \ll T \sim 10 - 20$ MeV, we find the contribution is much smaller than $(\lambda/10^{-23})10^{-17}$ eV which is smaller than our benchmark value for $m_\phi$. Similar consideration holds valid for $T < 0.1$ MeV where $\nu_s$ can be resonantly populated. As a result the thermal stability is guaranteed. For the larger values of $\lambda$, the stability can be jeopardized and a more careful study is required.

### III. PROSPECTS FOR KATRIN AND PTOLEMY

The KArlsruhe TRItium Neutrino (KATRIN) experiment is designed to measure the neutrino mass by studying the endpoint of the spectrum of the emitted electron in the beta decay of Tritium. The experiment, which will soon release its first data, can be sensitive to the neutrino mass (or to be more precise to $m_{\nu_e} = m_1|U_{e1}|^2 + m_2|U_{e2}|^2 + m_3|U_{e3}|^2$ $^{11}$) down to 0.2 eV $^{12}$. On the other hand, in the framework of the $\Lambda$CDM and the standard model of particles (including neutrino mass) combining the CMB and BAO $^{14}$ implies that neutrino mass should be smaller than this threshold and KATRIN cannot therefore discern the shift of the endpoint of the spectrum. However as shown in $^{13, 14}$, there are ways to relax the bounds from cosmology on the sum of the neutrino masses opening up the hope for KATRIN to resolve a sizable shift of the endpoint and to measure $m_{\nu_e}$.

If $\nu_e$ has a $\nu_4$ component with a mass of $\sim 1$ eV, it will show up as a kink $^{11, 12, 43, 46}$ in the spectrum of the emitted electron at $E_e = Q - m_{\nu_e}$ where $Q$ is the mass difference between the mother and daughter nuclei. The height of the kink will be characterized by $|U_{e4}|^2$. Within the 3+1 solution to the LSND and MiniBooNE anomalies or the 3+1 solution to the reactor and Gallium anomalies, the size of the kink can be large enough to be resolved $^{12}$. Let us discuss the implications of KATRIN observations combined with other observations within our scenario.

If future studies establish a deficit in the reactor $\bar{\nu}_e$ flux compatible with the 3+1 scheme with $\Delta m^2 \sim 1$ eV$^2$ and $|U_{e4}|^2 \sim 0.01$ and on the other hand if KATRIN observes a kink with the corresponding position and amplitude, this will be a strong hint in favor of the 3+1 scheme. There is a similar concept to detect relic neutrinos by the $\nu_e$ capture on Tritium. The PTOLEMY experiment is proposed to search for relic neutrinos invoking this concept $^{47}$. Similarly to KATRIN, it can also study the beta decay spectrum. Within the 3$\nu$ mass scheme, we expect a peak further away from the endpoint at $E_e = Q + m_{\nu_e}$ due to the $\nu_e$ capture on Tritium (To be more precise, we expect three peaks at $Q + m_{\nu_1}$, $Q + m_{\nu_2}$ and $Q + m_{\nu_3}$ which overlap with each other, looking like a single peak. Since within the 3$\nu$ scheme, we expect $F_{\nu_e} : F_{\nu_2} : F_{\nu_1} = 1 : 1 : 1$, the heights of these three overlapping peaks are given by $|U_{e1}|^2$, $|U_{e2}|^2$ and $|U_{e3}|^2$.) Within the 3+1 scheme in addition to this peak, there will be another peak at $E_e = Q + m_{\nu_4}$ but with a height suppressed by $|U_{e4}|^2$. As we discussed in the previous section, within our scenario neutrinos will eventually decay into the lightest mass eigenstate; i.e., $\nu_1$ for the normal mass ordering and $\nu_3$ for the inverted mass ordering. Thus, we expect a single peak at PTOLEMY-like setups at $Q + m_{\nu_4}$ ($Q + m_{\nu_5}$) with a height enhanced (suppressed) by a factor of $3|U_{e1}|^2$ (a factor of $3|U_{e3}|^2$) relative to the peak for the 3$\nu$ scheme for normal (inverted) mass ordering scheme. As a result, the presence of the kink at KATRIN results (or at PTOLEMY itself) but the absence of a second peak in the $\nu_e$ capture experiments might be taken as an indication for the $\nu_4$ decay with a lifetime shorter than the age of the universe. We should however notice that in a scenario with a non-renormalizable coupling between $\nu$ and $\phi$, as discussed in the previous section, the contribution of $\nu_4$ to the background will be also negligible so similarly to our scenario, there will be no second peak. In this case however the height of the peak will be the same as that in the standard 3$\nu$ scheme rather than being enhanced (suppressed) by $3|U_{e1}|^2$ (or by $3|U_{e3}|^2$).

### IV. SUMMARY

We have proposed a scenario to make the 3+1 solution to the short baseline neutrino anomalies compatible with the cosmological observations. The scenario is based on a small Yukawa coupling between the sterile neutrino and ultra-light background scalar with a mass of $m_\phi < 5 \times 10^{-17}$ eV. This coupling will induce an effective mass for $\nu_s$ in the early universe when active neutrinos are still in thermal equilibrium with plasma, suppressing the effective active sterile mixing and therefore the $\nu_s$ production. This way the bound on $N_{\nu eff}$ is satisfied. Below $T \sim 0.01$ MeV (and for sure at present) the effective mass induced by the coupling to the dark matter is negligible.
After the neutrino decoupling when the $\phi$ field starts oscillating, the effective mass induced for $\nu_s$ ($m_{\nu_s}^{eff}$) can become negative, canceling the vacuum mass, $m_{\nu_s}$. During the instants of total cancellation, active neutrinos can be resonantly converted to $\nu_s$, causing a tension with the total neutrino mass bounds from BAO and CMB. A remedy is to open up the possibility of $\nu_4$ decay before matter radiation equality. This rather fast decay can be achieved by coupling $\nu_s$ to another singlet scalar which is lighter than $\nu_4$. We have discussed the interpretation of possible results from future observations of KATRIN and PTOLEMY within the framework of the present scenario and compare it with the predictions of certain alternative frameworks.

Throughout this letter, our main focus was on the 3 + 1 solution to the short baseline anomalies but these results can be applied to even the 3 + 1 solution to the ANITA events \cite{48} which relies on nonzero $|U_{\tau4}|$ instead of nonzero $|U_{e4}|$ or $|U_{\mu4}|$.

ACKNOWLEDGMENTS

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 674896 and No 690575. YF has received partial financial support from Saramadan under contract No. ISEF/M/98223. YF would like also to thank the ICTP staff and the INFN node of the INVISIBLES network in Padova. The author is also grateful to M M Sheikh-Jabbari for useful discussions and thanks P. Denton and J. Cline for fruitful remarks.

\[1\] A. Aguilar-Arevalo et al. [LSND Collaboration], Phys. Rev. D 64 (2001) 112007 doi:10.1103/PhysRevD.64.112007 [hep-ex/0104049].
\[2\] A. A. Aguilar-Arevalo et al. [MiniBooNE Collaboration], Phys. Rev. Lett. 121 (2018) no.22, 221801 doi:10.1103/PhysRevLett.121.221801 [arXiv:1805.12028 [hep-ex]].
\[3\] C. Giunti, M. Laveder, Y. F. Li, Q. Y. Liu and H. W. Long, Phys. Rev. D 86 (2012) 113014 doi:10.1103/PhysRevD.86.113014 [arXiv:1210.5715 [hep-ph]]; C. Giunti and M. Laveder, Phys. Rev. C 83 (2011) 065504 doi:10.1103/PhysRevC.83.065504 [arXiv:1006.3244 [hep-ph]].
\[4\] T. A. Mueller et al., Phys. Rev. C 83 (2011) 054615 doi:10.1103/PhysRevC.83.054615 [arXiv:1101.2663 [hep-ex]].
\[5\] P. Huber, Phys. Rev. C 84 (2011) 024617 Erratum: [Phys. Rev. C 85 (2012) 029901] doi:10.1103/PhysRevC.84.024617 [arXiv:1106.0687 [hep-ph]].
\[6\] P. Adamson et al. [MINOS+ Collaboration], Phys. Rev. Lett. 122 (2019) no.9, 091803 doi:10.1103/PhysRevLett.122.091803 [arXiv:1710.06488 [hep-ex]].
\[7\] B. J. P. Jones [IceCube Collaboration], EPJ Web Conf. 207 (2019) 04005 doi:10.1051/epjconf/201920704005 [arXiv:1902.06185 [hep-ex]].
\[8\] G. Mention, M. Fechner, T. Lasserre, T. A. Mueller, D. Lhuillier, M. Cribier and A. Letourneau, Phys. Rev. D 83 (2011) 073006 doi:10.1103/PhysRevD.83.073006 [arXiv:1101.2755 [hep-ex]].
\[9\] S. Gariazzo, C. Giunti, M. Laveder, Y. F. Li and E. M. Zavain, J. Phys. G 43 (2016) 033001 doi:10.1088/0954-3899/43/3/033001 [arXiv:1507.08204 [hep-ph]].
\[10\] J. Kostensalo, J. Suhonen, C. Giunti and P. C. Srivastava, arXiv:1906.10980 [nucl-th].
\[11\] Y. Farzan, O. L. G. Peres and A. Y. Smirnov, Nucl. Phys. B 612 (2001) 59 doi:10.1016/S0550-3213(01)00361-3 [hep-ph/0105105]; S. M. Bilenky, S. Pascoli and S. T. Petcov, Phys. Rev. D 64 (2001) 113003 doi:10.1103/PhysRevD.64.113003 [hep-ph/0104218].
\[12\] A. Esmaili and O. L. G. Peres, Phys. Rev. D 85 (2012) 117301 doi:10.1103/PhysRevD.85.117301 [arXiv:1203.2632 [hep-ph]].
\[13\] S. Gariazzo, P. F. de Salas and S. Pastor, arXiv:1005.1290 [astro-ph.CO].
\[14\] N. Aghanim et al. [Planck Collaboration], arXiv:1807.06209 [astro-ph.CO].
\[15\] S. Hannestad, R. S. Hansen and T. Tram, Phys. Rev. Lett. 112 (2014) no.3, 031802 doi:10.1103/PhysRevLett.112.031802 [arXiv:1310.5926 [astro-ph.CO]].
\[16\] X. Chu, B. Dasgupta, M. Dentler, J. Kopp and N. Saviano, JCAP 1811 (2018) no.11, 049 doi:10.1088/1475-7516/2018/11/049 [arXiv:1806.06079 [hep-ph]].
\[17\] B. Dasgupta and J. Kopp, Phys. Rev. Lett. 112 (2014) no.3, 031803 doi:10.1103/PhysRevLett.112.031803 [arXiv:1310.6337 [hep-ph]].
\[18\] X. Chu, B. Dasgupta and J. Kopp, JCAP 1510 (2015) no.10, 011 doi:10.1088/1475-7516/2015/10/011 [arXiv:1505.02795 [hep-ph]].
\[19\] A. Paul, A. Ghoshal, A. Chatterjee and S. Pal, arXiv:1808.09706 [astro-ph.CO].
\[20\] A. Mirizzi, G. Mangano, O. Pisanti and N. Saviano, Phys. Rev. D 91 (2015) no.2, 025019 doi:10.1103/PhysRevD.91.025019 [arXiv:1410.1385 [hep-ph]].
[21] J. F. Cherry, A. Friedland and I. M. Shoemaker, arXiv:1605.06506 [hep-ph].
[22] N. Saviano, O. Pisanti, G. Mangano and A. Mirizzi, Phys. Rev. D 90 (2014) no.11, 113009 doi:10.1103/PhysRevD.90.113009 [arXiv:1405.1680 [astro-ph.CO]].
[23] N. Song, M. C. Gonzalez-Garcia and J. Salvado, JCAP 1810 (2018) no.10, 055 doi:10.1088/1475-7516/2018/10/055 [arXiv:1805.08218 [astro-ph.CO]].
[24] M. Archidiacono, S. Hannestad, R. S. Hansen and T. Tram, Phys. Rev. D 91 (2015) no.6, 065021 doi:10.1103/PhysRevD.91.065021 [arXiv:1405.5915 [astro-ph.CO]].
[25] L. Vecchi, Phys. Rev. D 94 (2016) no.11, 113015 doi:10.1103/PhysRevD.94.113015 [arXiv:1607.04161 [hep-ph]].
[26] Y. Zhao, Phys. Rev. D 95 (2017) no.11, 115002 doi:10.1103/PhysRevD.95.115002 [arXiv:1701.02735 [hep-ph]].
[27] P. B. Denton, Y. Farzan and I. M. Shoemaker, Phys. Rev. D 99 (2019) no.3, 035003 doi:10.1103/PhysRevD.99.035003 [arXiv:1811.01310 [hep-ph]].
[28] N. Song, M. C. Gonzalez-Garcia and J. Salvado, JCAP 1810 (2018) no.10, 055 doi:10.1088/1475-7516/2018/10/055 [arXiv:1409.1680 [astro-ph.CO]].
[29] N. Saviano, O. Pisanti, G. Mangano and A. Mirizzi, Phys. Rev. D 90 (2014) no.11, 113009 doi:10.1103/PhysRevD.90.113009 [arXiv:1405.1680 [astro-ph.CO]].
[30] M. Archidiacono, S. Hannestad, R. S. Hansen and T. Tram, Phys. Rev. D 91 (2015) no.6, 065021 doi:10.1103/PhysRevD.91.065021 [arXiv:1405.5915 [astro-ph.CO]].
[31] L. Vecchi, Phys. Rev. D 94 (2016) no.11, 113015 doi:10.1103/PhysRevD.94.113015 [arXiv:1607.04161 [hep-ph]].
[32] Y. Zhao, Phys. Rev. D 95 (2017) no.11, 115002 doi:10.1103/PhysRevD.95.115002 [arXiv:1701.02735 [hep-ph]].
[33] P. B. Denton, Y. Farzan and I. M. Shoemaker, Phys. Rev. D 99 (2019) no.3, 035003 doi:10.1103/PhysRevD.99.035003 [arXiv:1811.01310 [hep-ph]].
[34] M. Archidiacono, S. Hannestad, R. S. Hansen and T. Tram, Phys. Rev. D 91 (2015) no.6, 065021 doi:10.1103/PhysRevD.91.065021 [arXiv:1405.5915 [astro-ph.CO]].
[35] L. Vecchi, Phys. Rev. D 94 (2016) no.11, 113015 doi:10.1103/PhysRevD.94.113015 [arXiv:1607.04161 [hep-ph]].
[36] Y. Zhao, Phys. Rev. D 95 (2017) no.11, 115002 doi:10.1103/PhysRevD.95.115002 [arXiv:1701.02735 [hep-ph]].
[37] P. B. Denton, Y. Farzan and I. M. Shoemaker, Phys. Rev. D 99 (2019) no.3, 035003 doi:10.1103/PhysRevD.99.035003 [arXiv:1811.01310 [hep-ph]].
[38] M. Archidiacono, S. Hannestad, R. S. Hansen and T. Tram, Phys. Rev. D 91 (2015) no.6, 065021 doi:10.1103/PhysRevD.91.065021 [arXiv:1405.5915 [astro-ph.CO]].
[39] L. Vecchi, Phys. Rev. D 94 (2016) no.11, 113015 doi:10.1103/PhysRevD.94.113015 [arXiv:1607.04161 [hep-ph]].
[40] Y. Zhao, Phys. Rev. D 95 (2017) no.11, 115002 doi:10.1103/PhysRevD.95.115002 [arXiv:1701.02735 [hep-ph]].
[41] P. B. Denton, Y. Farzan and I. M. Shoemaker, Phys. Rev. D 99 (2019) no.3, 035003 doi:10.1103/PhysRevD.99.035003 [arXiv:1811.01310 [hep-ph]].
[42] M. Archidiacono, S. Hannestad, R. S. Hansen and T. Tram, Phys. Rev. D 91 (2015) no.6, 065021 doi:10.1103/PhysRevD.91.065021 [arXiv:1405.5915 [astro-ph.CO]].
[43] L. Vecchi, Phys. Rev. D 94 (2016) no.11, 113015 doi:10.1103/PhysRevD.94.113015 [arXiv:1607.04161 [hep-ph]].
[44] Y. Zhao, Phys. Rev. D 95 (2017) no.11, 115002 doi:10.1103/PhysRevD.95.115002 [arXiv:1701.02735 [hep-ph]].
[45] P. B. Denton, Y. Farzan and I. M. Shoemaker, Phys. Rev. D 99 (2019) no.3, 035003 doi:10.1103/PhysRevD.99.035003 [arXiv:1811.01310 [hep-ph]].