Title
Sterile neutrinos in cosmology

Permalink
https://escholarship.org/uc/item/0065t9js

Journal
PHYSICS REPORTS-REVIEW SECTION OF PHYSICS LETTERS, 711

ISSN
0370-1573

Author
Abazajian, KN

Publication Date
2017-11-10

DOI
10.1016/j.physrep.2017.10.003

License
CC BY 4.0

Peer reviewed
Sterile neutrinos in cosmology

Kevork N. Abazajian

Center for Cosmology, Department of Physics & Astronomy, University of California, Irvine, Irvine, California 92697, USA

Abstract
Sterile neutrinos are natural extensions to the standard model of particle physics in neutrino mass generation mechanisms. If they are relatively light, less than approximately 10 keV, they can alter cosmology significantly, from the early Universe to the matter and radiation energy density today. Here, we review the cosmological role such light sterile neutrinos can play from the early Universe, including production of keV-scale sterile neutrinos as dark matter candidates, and dynamics of light eV-scale sterile neutrinos during the weakly-coupled active neutrino era. We review proposed signatures of light sterile neutrinos in cosmic microwave background and large scale structure data. We also discuss keV-scale sterile neutrino dark matter decay signatures in X-ray observations, including recent candidate ∼3.5 keV X-ray line detections consistent with the decay of a ∼7 keV sterile neutrino dark matter particle.

Contents
1 Introduction 2
2 Neutrino Oscillation, Mass Generation and Sterile Neutrinos 3
  2.1 Neutrino Oscillation and Massive Neutrinos ............................................. 3
  2.2 Generating Neutrino Mass Beyond the Standard Model ................................. 5
     2.2.1 The Seesaw Mechanism and Related Models ......................................... 6
     2.2.2 Other Mass Generation Mechanisms ................................................... 7
3 Thermal History of the Universe and Sterile Neutrino Production 7
  3.1 Quantum Statistical Mechanics of Neutrinos .............................................. 8
     3.1.1 Partial to Full Thermalization in Symmetric Thermal Backgrounds .......... 11
     3.1.2 Creation of Lepton Number .............................................................. 12
4 Light eV-scale Sterile Neutrinos: Thermalization Conditions and Early Universe Constraints 13
5 Sterile Neutrino keV-scale Dark Matter 14
  5.1 Oscillation-based Production ................................................................. 14
  5.2 Low Reheating Temperature Universe and the Visible Sterile Neutrino .......... 17
  5.3 Non-oscillation Production: Particle Decays ............................................. 19
6 Sterile Neutrinos and Cosmological Structure Formation 20
  6.1 eV-Scale Sterile Neutrinos ................................................................. 20
  6.2 keV-Scale Neutrinos: Warm to Cold Dark Matter ...................................... 21
7 keV Sterile Neutrino Dark Matter Detectability in X-ray Observations 24
  7.1 Methods & Current Results ................................................................. 24
  7.2 Future Observations ................................................................. 28
8 Conclusions 28
1. Introduction

The conclusive determination of neutrinos as massive approximately two decades ago has opened a window into beyond the standard model of particle physics regarding efforts to determine their mass generation mechanism and the structure governing their mixing [1]. On nearly the same timescale, the standard model of cosmology has emerged with a robust framework, but requiring two new components that are little understood: dark matter and dark energy. The simplest and earliest mechanisms proposed to provide light neutrino masses invoke the presence of sterile neutrinos via the “seesaw” mechanism [2, 3, 4].

The Super-Kamiokande experiment has accomplished a precise measurement of the zenith-angle dependence of the atmospheric neutrino flux, leading to strong evidence of neutrino oscillations [5]. The mixing was consistent with a neutrino mass splitting of $|\delta m_{\text{atm}}| \approx 50$ meV between mass states associated largely with muon and tau flavor neutrinos. Evidence of the oscillations of solar neutrinos indicate the presence of an additional splitting of $\delta m_{\odot} \approx 9$ meV [6, 7]. The atmospheric and solar mass splittings are consistent with three flavor models that mirror the flavor structure of the charged leptons and quarks.

However, a persistence of anomalies in short baseline neutrino experiments, starting with the Liquid Scintillator Neutrino Detector (LSND) experiment, indicate a possibility of neutrino oscillations with at least one mass splitting at the $|\delta m_{\text{SBL}}| \sim 1$ eV scale [8]. An consistent signal of candidate flavor conversions of neutrinos at short baselines was seen in the MiniBooNE experiment [9]. A related signal exists in the discrepancy in reactor neutrino experiments that could be indicative of this higher-scale mass splitting [10, 11]. A third mass splitting would require, at a minimum, a fourth mass eigenstate in the neutrino sector, and the $Z_0$ width would require this fourth state to be sterile [1]. The experimental program and theoretical background of sterile neutrinos related to short baseline anomalies was reviewed in Refs. [12, 13].

There was a serendipitous start to both precision cosmology and precision neutrino oscillation experiments in the late twentieth century. This prompted a leading paper that the luminous red galaxy (LRG) sample of the Sloan Digital Sky Survey would be highly sensitive to the presence of massive neutrinos [14]. Though the focus of that paper was massive active neutrinos, the physics also applies to the cosmological presence of massive sterile neutrinos via partial or complete thermalization [15]. The presence of significant mixing between active and sterile neutrinos will also partially or completely thermalize the sterile neutrinos [16] and contribute to both the mass density as well as relativistic energy density at early times. A light sterile neutrino mixing with active neutrinos could also lead to lepton number generation [17] and numerous matter-affected neutrino mixing effects that distort neutrino spectra [18]. The effects of massive active and sterile neutrinos and extra relativistic energy density on both large scale structure (LSS) and the cosmic microwave background (CMB) were reviewed in Ref. [19]. We discuss the effects of massive eV-scale sterile neutrinos on cosmology in Section 4.

At a substantially higher mass scale, keV-scale mass sterile neutrinos with proper mixing with active neutrinos can be produced as dark matter via collisional processes in the early Universe in the Dodelson-Widrow mechanism [20]. Couplings to other particles can also produced in the early Universe at the abundance to be part or all of the cosmological dark matter [21]. Oscillation-based sterile neutrino dark matter production can also occur via an enhanced Mikheev-Smirnov-Wolfenstein (MSW) mechanism [22, 23] if there exists a primordial lepton number, through the Shi-Fuller mechanism [20]. Alternatively, the Dodelson-Widrow mechanism is often dubbed “non-resonant” sterile neutrino dark matter while the Shi-Fuller mechanism produces “resonant” sterile neutrino dark matter. Both mechanisms produce sterile neutrino dark matter that can behaves as “warm” dark matter, though resonant production could produce dark matter indistinguishable from absolutely cold dark matter (CDM) with current constraints. Non-oscillation production could also produce sufficiently cold relic sterile neutrino dark matter. Therefore, sterile neutrino dark matter can effectively be CDM.

The non-resonant Dodelson-Widrow mechanism collisionally produces the sterile neutrino dark matter in a non-equilibrium state, but with a momentum distribution that reflects the active neutrinos’. However, non-resonant sterile neutrino dark matter has a mapping between particle mass-scale different from thermally

---

1 Here, we describe mass splittings heuristically, as associated with their discovery method, but below the precise definition of a mass splitting is $\delta m_{\nu_\alpha}^2 = m_{\nu_\beta}^2 - m_{\nu_\alpha}^2$ between vacuum mass eigenstates $\nu_\alpha$ and $\nu_\beta$. 2
produced warm dark matter (WDM) [24] and therefore non-resonant sterile neutrino dark matter’s effect on structure formation is unique [25]. The resonant-production mechanism produces a momentum distribution that is on average lower in average momenta at a given temperature than non-resonant production due to the resonance occurring first through the lower momenta modes. The resonant sterile neutrino dark matter therefore is frequently called “cooler” than non-resonantly produced sterile neutrino dark matter. Neither Dodelson-Widrow nor Shi-Fuller mechanisms for dark matter produce a sterile neutrino in thermal equilibrium, therefore temperature-based concepts like warm and cold are inaccurate to apply. We discuss the connection between the thermal warm dark matter case and oscillation and non-oscillation production mechanisms in Section 6.2.

The value of the initial lepton asymmetry in the Universe, whether trivial or substantial but within constraints, connects the resonant-production mechanism with the non-resonant production mechanism. Resonant production relies on a substantial but observationally unconstrained primordial cosmological lepton asymmetry. The exploration of the full parameter space of sterile neutrino particle mass, mixing angle and lepton number was first performed in Abazajian, Fuller & Patel [26], as well as an exploration of constraints on sterile neutrino dark matter from primordial nucleosynthesis, the CMB, structure formation, and diffuse X-ray to gamma-ray backgrounds. Ref. [27] also discussed diffuse X-ray constraints on non-resonantly produced sterile neutrino dark matter. Abazajian, Fuller & Patel [26] were the first to point out the serendipitous connection between the allowable parameter space for sterile neutrino dark matter and the detectability of the radiative decay line of sterile neutrino dark matter with observations of dark-matter rich objects like galaxies and galaxy clusters by the then-recently launched Chandra and XMM-Newton X-ray Space Telescopes. The follow-up paper Ref. [28] was the first to study X-ray constraints from XMM-Newton observations of the Virgo Cluster, and also made forecasts for future observations, including those from proposed large-exposure and large-effective-area X-ray microcalorimeter spectrometer missions like Constellation-X. Ref. [28] also discussed that increased sensitivity could also be attained from large exposures via a stacking analysis of the spectra of a number of similar clusters.

For a considerable amount of time, only upper limits to decay rates were found from searches for the X-ray line. In 2014, Bulbul et al. [29] reported a high-significance, $4\sigma - 5\sigma$, detection in stacked observations of 73 clusters with the MOS and PN spectrometers aboard XMM-Newton, as well as a consistent signal from the Perseus cluster of galaxies observed with the Chandra telescope. Boyarsky et al. [30] found a consistent signal from the Andromeda galaxy as well as Perseus using data from the XMM-Newton satellite. There has been significant scrutiny of the results as well as follow-up observations that see commensurate signals in other astronomical observations. We review these signals and challenges in its interpretation in Section 7.

2. Neutrino Oscillation, Mass Generation and Sterile Neutrinos

2.1. Neutrino Oscillation and Massive Neutrinos

A quantum state created in some atomic, nuclear or particle interaction need not be created in an energy eigenstate that is stationary—that is, a state that does not evolve—in the vacuum. A “textbook” example of such a problem is the spin precession of a spin 1/2 system in an external magnetic field $\mathbf{B}$ (e.g., see Ref. [31]). A spin 1/2 particle that is, for example, created in a spin state aligned with the positive $x$-axis ($|S_x\rangle$), and enters a region with a magnetic field aligned along the positive $z$-axis, will evolve or “precess.” In the region with a magnetic field along the $z$-axis, the energy eigenstates are spin 1/2 particles with their spins aligned either in the positive or negative $z$-axis ($|S_z\rangle$). The state $|S_x\rangle$, $t = 0$ will have a probability at some later time, $t'$ to be either in a $|S_x\rangle$ or $|S_z\rangle$ state, since $|S_x\rangle$ is not a stationary state of the system. The probability of finding it in a $|S_x\rangle$ or $|S_z\rangle$ state has an oscillatory form.

The case of the massive neutrino, $\nu$, propagating in vacuum is very similar to spin precession. The neutrino interacts via the Weak Interaction and displays the three family ($e, \mu, \tau$) symmetry of the lepton sector of the standard model of particle physics. The simplifying assumption of the standard model is that the neutrino is massless, which automatically makes a neutrino of a particular flavor ($e, \mu, \tau$) an energy eigenstate. If neutrinos have mass, theory does not require that the mass states are identical with the flavor
states (|να⟩, where α = e, µ, τ). Most generally, Weak Interaction flavor and vacuum mass states are not identical. A flavor state |να⟩, is, most generally, a superposition of mass eigenstates |να⟩(α = 1, 2, 3...).

Since the time evolution operator \( U(t) = \exp(-iHt) \) in a field/force free vacuum is \( U(t) = \exp(-iE_\nu t) = \exp(-i\sqrt{m^2 + p^2}t) \approx \exp(-i\delta m^2t/2p) \), the mass states are energy eigenstates in vacuum (we have identified \( x = t \)).

If we make the often-applied approximation that there are only two neutrino flavors and mass states important to consider, the two flavors (α, β) are superpositions of the two mass states (i.e., “mixed”):

\[
|\nu_\alpha\rangle = \cos \theta |\nu_1\rangle + \sin \theta |\nu_2\rangle \\
|\nu_\beta\rangle = -\sin \theta |\nu_1\rangle + \cos \theta |\nu_2\rangle,
\]

and the analogy to the spin 1/2 precession system is apparent. A neutrino created in some Weak Interaction process as a flavor state (via flavor conservation), will evolve and the probability of finding it later in a state α will have an oscillatory form.

One of the most interesting features of neutrino physics is the coherent effect of the medium through which the neutrino is propagating. The effective mass and mixing angles can be very different for neutrinos propagating through dense environments. Such coherent effects in the sun’s interior can produce a coherent level crossing and transformation of neutrino flavor in the sun—the Mikheev-Smirnov-Wolfenstein (MSW) mechanism [22, 23]. Resonant transformations can also occur in the the dense environment of core-collapse supernovae, and can induce effects on neutrino heating behind the shock [32]. Resonant neutrino transformations can also occur in the early Universe, which we explore further below [33, 34, 35].

Considerable evidence suggests the existence of small neutrino masses and mixings, and that neutrino mass states that have different masses. The first indications of anomalies in the neutrino sector came from the measurement of the flux of electron neutrinos coming from the Sun by the Davis Experiment [36], which found that flux to be a third of that predicted flux from solar models [37, 38, 39]. Observations of the neutrino flux from the sun are found to be far below expected rates, and indicate neutrino flavor conversion via the MSW effect with relatively large mixing angles and at a mass splitting scale of |δm| ∼ 8 meV. The flavor conversion was confirmed by the Sudbury Neutrino Observatory which sees the full flux in flavor-independent neutral current interactions, while also seeing the deficit of electron neutrinos in charge-current interactions [7]. This flavor conversion was confirmed in the KamLAND laboratory reactor neutrino experiment, which sees the same flavor conversion but with reactor antineutrinos [40]. The Super-Kamiokande atmospheric neutrino experiment sees a zenith angle dependence of the atmospheric neutrino flux that is fit most simply with an model of neutrino masses for the µ and τ flavored neutrinos, νµ and ντ, split by a mass difference of |dνatm| ∼ 50 meV [5], and this has been verified and refined by long-baseline neutrino laboratory experiments [41, 42].

The splittings of the mass eigenstates are becoming more precisely measured in neutrino oscillation experiments [43], while determining the value of the absolute mass in the laboratory requires the accurate measurement of the endpoint of the spectrum of beta-decay neutrinos [44]. Neutrinoless double-beta decay experiments could simultaneously determine whether neutrinos are Majorana particles (i.e., their own antiparticles) as well as their absolute mass scale [45]. In addition, long baseline neutrino oscillation experiment will likely determine whether the neutrino eigenstates are “normal ordered” with \( m_1 < m_2 < m_3 \) or “inverted order” \( m_3 < m_1 < m_2 \). A measure of this ordering in concert with a measurement of the sum of neutrino masses can reveal the nature of the neutrino mass-generation mechanism and therefore open a window into high-energy scale physics [46].

The Los Alamos Liquid Scintillator Neutrino Detector (LSND) is a short-baseline beam-dump experiment that finds a flavor excess that may be consistent with neutrino oscillations at a eV-scale mass splittings, |δmSBL| ∼ 1 eV [47]. The MiniBooNE Experiment sees an anomaly in its antineutrino channel that may be consistent with the LSND signal [48]. In total, there are three mass-splitting scales of |δm| ≪ |δm atm| ≪ |δmSBL|, which indicates the need for at least four mass eigenstates to accommodate three disparate mass splittings. It is difficult to include all atmospheric, solar and short baseline oscillations, and constraints on short baseline oscillations, in a single neutrino mass and mixing model of four or more neutrino states [49].
2.2. Generating Neutrino Mass Beyond the Standard Model

In the canonical Standard Model (SM) of particle physics, neutrinos, combined with the charged leptons, form left-handed electroweak $SU(2)$ doublets, $L_{\alpha}$ with family $\alpha$. This is associated via CPT with a right-handed antiparticle state $L_{\alpha,R}^c$, which transform as \[ (\nu_e)^L \xleftarrow{\text{CPT}} (\nu_e)^R \] \[ (e^+)^L \xrightarrow{\text{CPT}} (\nu_e)^R. \] (1)

The $L$ and $R$ refer to left and right chiral projections, which in the case of zero mass also correspond to helicity states. The SM has that the neutrinos have no electric charge, no color, as well as the ansatz of masslessness with no inclusion of a “right-handed” $N_R$ that would be required to produce mass. $N_R$ would be $SU(2)$-singlets, with no weak interactions except those induced by mixing with the active neutrinos. Such “sterile neutrinos” can be added to the SM, and are predicted in most extensions. For sterile neutrinos, the $R$ state is the particle and the $L$ state is the anti-particle, \[ N_R \xleftarrow{\text{CPT}} N_L^c. \] (2)

Below, after this discussion of neutrino mass generation, we will refer to sterile neutrinos as $\nu_s$: \[ \nu_s \equiv N_R. \]

Neutrinos can have masses through the same mechanism that provides the charged leptons masses, via a Dirac mass term, where there are two distinct neutrinos $\nu_L$ and $N_R$: \[ -L_{\text{Dirac}} = m_D (\bar{\nu}_L N_R + \bar{N}_R \nu_L) = m_D \hat{\nu} \nu \] \[ \text{and the Dirac field is defined as } \nu \equiv \nu_L + N_R. \] (3)

For three or more families (flavors) of neutrinos, this mechanism is easily generalized, and the mass terms become matrices. The charged current interactions then involve a leptonic mixing matrix analogous to the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix, with the leptonic one dubbed the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [51, 52, 53]. This allows for oscillations between the light neutrinos.

The generation of mass requires a $SU(2)$ breaking generated by a Yukawa coupling \[ -L_{\text{Yukawa}} = h_{\nu} (\bar{\nu}_e \bar{e} \nu) L \left( \varphi^0 - \varphi \right) N_R + H.c. \] (4)

This is gives the Dirac mass for the neutrinos in the same way as for charged leptons with $m_D = h_{\nu} v / \sqrt{2}$. The vacuum expectation value (VEV) of the Higgs doublet is \[ v = \sqrt{2} (\varphi^0) = (\sqrt{2} G_F)^{-1/2} = 246 \text{ GeV}, \] and the Yukawa coupling is $h_{\nu}$. The question in this mass generation mechanism is why $h_{\nu}$ is so small, with $h_{\nu} \lesssim 10^{-11}$ for $m_{\nu_e} < 2$ eV [54, 55].

Neutrinos may also have masses via Majorana mass interaction, which involves the right-handed antineutrino, $\bar{\nu}_R^c$, instead of a separate Weyl neutrino. That is, it is a transition from an antineutrino to a neutrino, or equivalently as a creation or annihilation of two neutrinos. If present, this would lead to neutrinoless double beta decay. The Majorana mass term is \[ -L_{\text{Majorana}} = \frac{1}{2} m_T (\bar{\nu}_L \nu_R^c + \bar{\nu}_R^c \nu_L) = \frac{1}{2} m_T \nu \nu \] \[ \text{and the Majorana mass term among the sterile neutrinos alone,} \] \[ -L_{\text{Majorana, sterile}} = \frac{1}{2} M_N \left( N_L^c N_R + \bar{N}_R \bar{N}_L^c \right). \] (5)
2.2.1. The Seesaw Mechanism and Related Models

When $\nu_i$ is an active neutrino, isospin is violated by one unit by the Majorana term, and $m^T$ must be generated by an elementary Higgs triplet or by an effective operator from two Higgs doublets arranged to act as a Higgs triplet. The latter case is

$$\frac{\lambda_{ij}}{M_N} (L_i \varphi)^T (L_j \varphi), \quad i, j = e, \mu, \tau,$$

where $\varphi$ is the Higgs doublet, $\lambda_{ij}$ are dimensionless couplings, and $M_N$ is a cutoff mass scale above electroweak symmetry breaking. This generates the Majorana masses

$$m_{ij} = \frac{\lambda_{ij}(H)^2}{M_N} \sim \frac{m_D^2}{M_N}.$$  

(8)

For $M_N$ large, $m_{ij} \ll m_D$, the high-scale completion of this mechanism determines the type of seesaw mechanism \cite{46}. If $M_N$ is at the Planck scale, $M_N \sim M_{\text{Planck}} \sim 10^{19}$ GeV, and $\lambda_{ij} \sim 1$. The neutrino mass scales are smaller than the observed splittings, $m_{ij} \sim 10^{-5}$. This intriguingly indicates new physics below the Planck scale.

The candidate mass insertions of Eqs. (4),(6),(7) can be distilled into a generalized form of new terms to enter the SM to generate neutrino mass, dubbed the “new Standard Model” or “neutrino Standard Model” (νSM) \cite{56} or the “neutrino Minimal Standard Model” (νMSM) \cite{57},

$$\mathcal{L} \supset -h_{\alpha i} L_i N_i \varphi - \frac{1}{2} M_{ij} N_i N_j + H.c.,$$

(9)

where $h_{\alpha i}$ are the Yukawa couplings for the flavor states $\alpha = e, \mu, \tau$ and $M_{ij} = M_{ji}$ ($i, j = 1, 2, \ldots$) are the Majorana masses.

There is quite a bit of freedom as to the values of sterile neutrino masses $M_N$ and Yukawa couplings $h_{\alpha i}$ allowed \cite{58}. Experimental data limits currently places a wide range of values for these: $10^{-12} \lesssim h_{\alpha i} \lesssim 1$ while $M_{ij} = 0$ for all $i, j$, where Majorana neutrinos “fuse” to Dirac neutrinos, to $M_{ij} \sim 10^{16}$ GeV. There is a relatively small range at 0.001 eV $\lesssim M_{ij} \lesssim 1$ eV, which is disfavored since it induces large sterile neutrino mixing which is disfavored by oscillation data. The seesaw mechanism is simply the case where $M_i \gg h_{\alpha i} v$, and the three sterile neutrinos are split from the mass states associated with with active neutrinos.

Truly only two of the heavy sterile neutrinos are needed to produce the solar and atmospheric mass scales, and the third is free to play the role of either a largely sterile mass eigenstate that leads to the short baseline anomalies or could play the role of dark matter \cite{56, 57}. Being open about the sterile neutrino numbers reflecting the family structure of the standard model allows for three sterile neutrino states and therefore one of these to be a relative light state: a sterile neutrino associated with short baseline oscillations or a dark matter sterile neutrinos. It is often mistakenly taken that the short baseline anomalies indication of potential light eV-scale neutrinos are motivation for dark matter sterile neutrinos at the keV scale, but simplicity argues for only one of the two sterile neutrino mass states to be likely, either eV or keV in scale.

The mass scale of the active neutrinos is given by Eq. (9), $m_{\alpha} \sim h^2v^2/M_N$ while the active-sterile mixing angles scale as

$$\theta \sim \frac{hv}{M} \sim \sqrt{\frac{m_{\alpha}}{M}}.$$  

(10)

This allows for a predominantly sterile mass state to have a mass $m_{s} \equiv M \sim \text{keV}$ that has arbitrarily small (or large) mixings as long as the “active” mass scale $m_{\alpha}$ at the appropriate scale. Given that the lightest active neutrino mass eigenstate is not bounded from below, neither is the mixing angle. This is a natural mechanism for a sterile neutrino to have the properties like that of a dark matter sterile neutrino, or an eV-scale short baseline neutrino. However, this mechanism cannot provide both eV-scale and keV-scale sterile neutrinos within a three family structure. The origin of the splitting of the lower scale, eV or keV, sterile neutrinos from the higher scale sterile neutrinos ($\gg \text{keV}$) must be come from a correction to zero mass for the light sterile, or from a suppression relative to the higher scale. Several mechanisms regarding both methods were explored in Refs. \cite{59, 60}.
2.2.2. Other Mass Generation Mechanisms

There are numerous mechanisms for active sterile neutrino mass generation that are related to the original seesaw described above, or independent. As mentioned above, a triplet Majorana mass $m_T$ can be generated by a Higgs triplet VEV $v_T$, in what is dubbed a Type-II seesaw mechanism [61, 62, 63, 64]. In that case, $m_T = hv_T v$, where $h_T$ is the Yukawa coupling. The smallness of $m_T$ would be provided by a small $v_T$. A combination of type I terms and type II is called a “mixed seesaw.” Mohapatra & Smirnov [46] provide an excellent review of seesaw and related neutrino mass generation mechanisms.

One case where additional singlets contributing to the neutrino mass generation mechanism provided by a high-energy theory is dubbed an inverse seesaw [65]. In Merle [59], mass generation mechanisms for keV-scale (and for that matter eV-scale) masses were usefully categorized as those that arise from “bottom-up” models, where the natural sterile neutrino scale is zero, with has a model-specific correction to shift it to non-zero, or arise from “top-down” mechanisms, where the natural scale is high ($>\text{keV}$), but split downward to low mass scales by suppression mechanism. Examples of bottom-up mechanisms include two based on the flavor symmetries of $L_e - L_\mu - L_\tau$ [66, 67] and $Q_6$ [68]. In other bottom-up models, neutrino masses are only produced via higher order loops [69].

One top-down model uses an extra dimension compactified in a spherical orbifold geometry on $S_1/Z_2$, with the Yukawa couplings and right handed Majorana masses on a standard model brane, while the right handed neutrino wave functions exponentially localize on a hidden brane. This allows splitting of the scales of the seesaw, with two right handed neutrinos at high scale and one at a keV scale, in what is called a split seesaw mechanism [70]. This was extended to a flavor symmetry model in an $A_4$ extended split seesaw [71]. A version of the split seesaw has varied localizations of both the right handed Majorana masses on the hidden brane, and is dubbed the separate seesaw [72]. Another set of top-down models are based on the Froggat-Nielsen mechanism [73, 74, 75].

Another class of models include the extended seesaw mechanism, where a singlet fermion from a supersymmetric model can produce light sterile neutrinos [76, 77, 78, 79]. In Ref. [80] the axino plays the role as the singlet in $R$-parity violating supersymmetry models. Another class of models involve mirror sector [81, 82] or left-right symmetric models [83, 84] which in some cases can provide light eV-scale sterile neutrinos as well as the keV scale dark matter sterile neutrino [85]. Arguably just as natural as the seesaw mechanism is the possibility of Dirac neutrino masses, e.g. Ref. [86]. There are many further models of light sterile neutrino mass generation, which are well reviewed in Refs. [12, 59, 60].

3. Thermal History of the Universe and Sterile Neutrino Production

The cosmic microwave background (CMB) is now a very well-studied thermal black body with small anisotropies that carry information from the surface of last scattering in the early Universe. That information includes the mass-energy content of the Universe and primordial perturbation spectrum [87], as well as the background of cosmic neutrinos [19]. The CMB has a current photon temperature of $T_{\gamma,0} = 2.72548 \pm 0.00057$ K [88] and number density of $n_{\gamma,0} = 410.7(T_{\gamma,0}/2.72548 \text{K})^3$. The photon background had higher temperatures in the past, increasing inversely with the scale factor $T \propto a^{-1}$, during phases where the degrees of freedom coupled to the photons is constant. The photon number density correspondingly increased as $n_{\gamma} \propto a^{-3}$, which implies a hot, dense early phase. This led to the ionization of the primordial gas and Thompson scattering coupling of the photon-electron fluid at early times, prior to the CMB surface of last scattering.

At even earlier times, the photon background had sufficient energy to pair produce an electron-positron background $\gamma + \gamma \leftrightarrow e^+ + e^-$ at $T_e \gtrsim 0.1$ MeV, with a neutrino background being created at higher temperatures $T \gtrsim 1$ MeV: $e^+ + e^- \leftrightarrow \nu_\alpha + \bar{\nu}_\alpha$, and, above respective energy thresholds the entire standard model of particle physics catalog of particles is created in the thermal bath, along with any non-standard particles with sufficiently strong interactions [89].

In the neutrino-coupled era at $T \gtrsim 1$ MeV, the effects of neutrino scattering take place in an environment where neutrino oscillations would necessarily also need to be taken into account. In the standard flavor and $CPT$ symmetric neutrino background, there are not significant effects to due active neutrino oscillations,
but asymmetries can produce interesting collective effects [90, 91, 92]. These collective oscillations constrain
the total asymmetry of the Universe and are tied to the mixing angle to which solar neutrino oscillations
are largely sensitive, $\theta_{12}$ [92]. In the standard thermal history, active neutrinos are clearly created in their
flavor states but then freely propagate as mass eigenstates that are symmetric except for a the slight heating
that occurs from partial coupling of the neutrinos during electron-positron annihilations [93, 94].

3.1. Quantum Statistical Mechanics of Neutrinos

Most of the topics in this report reflect neutrino mass and mixing in the early Universe, where conditions
are very different from an ambient vacuum. To approach the evolution of neutrinos in the early Universe,
a quantum statistical formulation must be adopted. We follow here the descriptions of the statistical
formulation of a quantum system through the density matrix as described in Refs. [95, 31]. The density
operator, $\rho$, or its matrix is the most general form for the quantum description of a statistical system, and
contains all physical information of the system that we consider. This seems ambitious, but it is simply
the nature of the density matrix by definition. A general quantum system here is described by a wave
function $|\Psi(x, q)\rangle$, where $x$ are the coordinates or properties of the system that we are interested in and $q$
are the coordinates we are the rest of the coordinates of the system. The average expectation value of some
observable $\hat{f}$, $\langle f \rangle$ in wave-function integral form is

$$\langle f \rangle = \int \int \Psi^*(x, q) \hat{f} \Psi(x, q) \, dq \, dx.$$  \hfill (12)

We can reduce the general system by averaging over all other coordinates $q$, so that there remains only
a weight $w_x$ for the system to be characterized by $|\Psi(x)\rangle$. The ensemble average of an observable $\hat{f}$ is

$$\langle f \rangle = \sum_x \langle \Psi(x)|\hat{f}|\Psi(x)\rangle.$$  \hfill (13)

We can introduce the orthogonal complete bases $|a'\rangle$ and $|a''\rangle$ that may be of interest in our description
of the quantum system. The average (13) can be written, using completeness, as

$$\langle f \rangle = \sum_x w_x \sum_{a'} \sum_{a''} \langle \Psi(x)|a'\rangle \langle f|a''\rangle \langle a''|\Psi(x)\rangle$$

$$= \sum_{a'} \sum_{a''} \left( \sum_x w_x \langle a''|\Psi(x)\rangle \langle \Psi(x)|a'\rangle \right) \langle a'|f|a''\rangle.$$ \hfill (14)

In order to simplify the average over the coordinates of interest, we define the density operator as

$$\hat{\rho} = \sum_x w_x |\Psi(x)\rangle \langle \Psi(x)|,$$ \hfill (15)

which can be projected into a matrix in the basis of a set of states $|a'\rangle$ via

$$\langle a''|\hat{\rho}|a'\rangle = \sum_x w_x \langle a''|\Psi(x)\rangle \langle \Psi(x)|a'\rangle.$$ \hfill (16)

Therefore, the ensemble average Eq. (14) can be written simply as

$$\langle f \rangle = \sum_{a'} \sum_{a''} \langle a''|\hat{\rho}|a'\rangle \langle a'|f|a''\rangle$$

$$= \sum_{a''} \langle a''|\hat{\rho}f|a''\rangle$$

$$= \text{tr}(\hat{\rho}f).$$ \hfill (17)
Since this trace is independent of the basis representation, any convenient basis is used.

The diagonal elements of the density matrix describe the probability distribution of the system in a specific basis state \(|a'\rangle\):

\[
(f)\big|_{a'} = \langle a' | \hat{\rho} | a' \rangle.
\]  
Equation (18)

In this way, the average value of an observable in a state \(|a'\rangle\) can be found readily. Therefore, the analysis of the density matrix of a system is a powerful formulation. The diagonal elements, Eq. (18), serve as the statistical distribution functions for the system.

For the early Universe, spatial coordinates are unnecessary under the assumption of spatial homogeneity, which is the case in the standard picture. For neutrinos in this environment, we may project the density matrix into a mass eigenstate basis or a flavor basis. Now we have the density matrix as a tool for describing the statistical properties of a quantum mechanical system, but in order to find how neutrinos evolve in the early Universe, we need to be able to evaluate the time evolution of the relativistic quantum statistical ensemble.

The weak interactions of mixed active-sterile neutrinos have finite width to produce sterile states, and serve to couple sterile density operator amplitude to the thermal environment. The time evolution of the density operator can be described by the Heisenberg equation of motion:

\[
i\frac{\partial \hat{\rho}}{\partial t} = [\hat{\rho}, H],
\]  
Equation (19)

with an analogous equation for the antineutrino density matrix, \(\bar{\rho}\). Now, the problem lies in the choice of the Hamiltonian of the neutrino ensemble \(H\). The first application of the evolution of the system through a Heisenberg equation of the density matrix was done by Dolgov [96], who took the matrix elements of elastic (\(\nu e \leftrightarrow \nu e\)) and inelastic (\(\nu e \leftrightarrow e^+ e^-\)) to write a simple singlet neutrino production formulation.

Another way of looking at the time evolution of the density operator is through a scattering matrix approach such that the differential evolution of the density operator is given by [97]

\[
\hat{\rho}^{(f)} = \hat{S} \hat{\rho}^{(i)} \hat{S}^\dagger
\]  
Equation (20)

which can be used to derive the scattering matrix elements for population of singlet neutrinos through two-body elastic and inelastic processes, and the time rate of change of the density operator.

The integration over contributing scattering matrix elements in the right hand side of the density operator evolution equations or “quantum kinetic equations” (19) & (20) lead to the integrated and drastically less cumbersome “quantum rate equations.” The derivation of the rate equations from the kinetic equations has been described by McKellar & Thomson [97], Sigl & Raffelt [98], and Bell et al. [99].

One approach to solutions of the evolution of the sterile neutrino states approaches the solution of Eq. (19) directly. This is separately coupled to an additional relation governing the expansion of the Universe, the Friedmann equation, with approximations to the subsequent time-temperature evolution and scattering rates [100, 101, 102]. The most significant effects that enter the evolution of neutrinos in the early Universe are the self energies produced the asymmetry and thermal potentials. These self energies are from propagating active neutrinos’ interactions with the plasma [103]. There are three contributions to the neutrino self energy: (a) an imaginary part proportional to the net neutrino scattering rate or opacity, (b) a real part due to finite weak gauge boson masses \((V_{\text{th}})\), and (c) a real part proportional to asymmetries in weakly interacting particles \((V_{\text{L}})\). This exposition follows the treatment in Venumadhav et al. [104] that employs some of the definitions from Ref. [103]. We will specify to different regimes of applicability later.

Figure 1 shows the lowest-order contributions to the neutrinos’ self energy. Thick red lines are thermal propagators of weakly charged species in the background plasma. There are two contributions: bubbles and tadpoles, shown in Fig. 1a and 1b respectively. The background fermion is a lepton of the same flavor in the former, and any weakly charged species in the latter.

A massless active neutrino’s propagator is

\[
G^{-1}_{\nu_\alpha}(p_{\nu_\alpha}) = p_{\mu_{\nu_\alpha}} - b_{\nu_\alpha}(p_{\nu_\alpha}) \mathcal{U}(1 - \gamma_5)/2, \\
b_{\nu_\alpha}(p_{\nu_\alpha}) = b_{\nu_\alpha}^{(0)} + b_{\nu_\alpha}^{(1)} \omega_{\nu_\alpha}, \\
\omega_{\nu_\alpha} = -p_{\nu_\alpha} \cdot u,
\]  
Equations (21) & (22)
Figure 1: These are the lowest order contributions to a propagating active neutrino’s self energy. Red lines are thermal propagators. In (a), $f$ is any species with weak charge. In (b), $f = \nu_\alpha, \bar{\nu}^-$. 

\[ \Psi_{\nu_\alpha} \rightarrow W^\pm, Z_0 \]

(a) bubble diagram

\[ \Psi_{\nu_\alpha} \rightarrow Z_0 \]

(b) tadpole diagram

where $p_\nu$ and $u$ are the neutrino and plasma’s four-momenta, $y$ is shorthand for $\gamma^\mu v_\mu$, and $b_{\nu_\alpha}$ is the left handed neutrino’s self energy. The relation Eq. (22) divides the self energy into separate contributions that affect the particle and anti-particles in Eq. (21) differently.

The finite density potential in the early universe can be dominated by asymmetries in the lepton number. Therefore it is often referred to as the “lepton potential” $V^L$. It takes the form

\[ V^L = \frac{2\sqrt{2}\zeta(3)}{\pi^2} G_F T^3 \left( L^\alpha \pm \frac{\eta}{4} \right), \]

where we take “+” for $\alpha = e$ and “−” for $\alpha = \mu, \tau$. Here we define the lepton number $L^\alpha$ in terms of the lepton numbers in each active neutrino species as

\[ L^\alpha \equiv 2L_{\nu_\alpha} + \sum_{\beta \neq \alpha} L_{\nu_\beta}, \]

with the final sum over active neutrino flavors other than $\nu_\alpha$. Here $\eta \equiv n_b/n_\gamma$ is the baryon to photon ratio. An $\alpha$-type neutrino asymmetry is defined as

\[ L_{\nu_\alpha} \equiv \frac{n_{\nu_\alpha} - n_{\bar{\nu}_\alpha}}{n_\gamma}, \]

where the photon number density $n_\gamma = 2\zeta(3)T^3/\pi^2$.

There are a large number of scattering processes that contribute to the neutrino opacity (i.e. the imaginary part of the self-energy) at temperatures $10 \text{ MeV} \leq T \leq 10 \text{ GeV}$. Accurate neutrino opacities are needed since they control the production of of sterile neutrinos as well as the damping of the active neutrinos’ oscillations. Quantum damping, also described as the quantum Zeno effect, is the suppression of conversions due to the interaction length of neutrinos being much smaller than the oscillation length. That is, there is no appreciable conversion possible unless neutrinos are able to sufficiently propagate to have a non-trivial probability to convert to sterile neutrinos.

Early work on neutrino interactions in the early Universe proposed that neutrinos largely scattered off relativistic particles and therefore scaled their cross-sections with the center-of-mass (CM) energy [103, 97, 26]. In addition, these early calculations neglected the effects of particle statistics. Under these two simplifying assumptions, the opacity $\Gamma(E_{\nu_\alpha})$ for an input neutrino of energy $E_{\nu_\alpha}$ is of the form

\[ \Gamma(E_{\nu_\alpha}) = \lambda(T) G_F^2 T^4 E_{\nu_\alpha}, \]

where $G_F$ is the Fermi coupling constant, and $\lambda(T)$ is a constant that depends on the number and type of available relativistic species in the cosmic plasma. References [100, 101] subsequently developed a framework
to include particle masses, loop corrections, and particle statistics in the neutrino opacity calculation. These were also included in Venumadhav et al. [104], where previously-neglected contributions to the scattering rates such as two- and three-body fusion reactions were included, also. Ref. [104] was first in using chiral perturbation theory to compute the hadronic contribution to the scattering opacity below the quark-hadron transition, with significant qualitative and quantitative modifications to the form of Eq. (26).

3.1.1. Partial to Full Thermalization in Symmetric Thermal Backgrounds

Oscillations induced by the mass generation mechanism between light eV-scale to keV-scale sterile neutrinos allows for them to be populated by scattering processes in the cosmological background. The initial work on this mechanism used a quasi-classical treatment that included the effects of quantum damping and thermal self-energy. Langacker [16] showed that quantum damping suppresses active-sterile neutrino oscillations at the highest temperatures in the early Universe, and that active-sterile mass splittings of $\delta m^2 \gtrsim 10^{-7} \text{ eV}^2$ will fully thermalize at maximal mixing, while $\delta m_{\alpha s}^2 \sim 1 \text{ eV}^2$ will thermalize with mixings of $\sin^2 2\theta > 10^{-4}$. Subsequent work found that finite temperature effects further suppressed the mixing and conversion of active to sterile neutrinos, with similar regions of parameter space where sterile neutrinos would be thermalized, notably the parameter space associated with the large mixing angle solution to the solar neutrino problem as well as the short baseline anomalies [105, 106, 107, 108, 109, 110].

The essential discoveries of early work are in the matter and thermal affected contributions to the neutrino self-energy. They lead to an effective mixing between the active and sterile neutrinos that suppresses mixing at high temperatures:

$$\sin^2 2\theta_{\alpha s} = \frac{\Delta^2(p) \sin^2 2\theta}{\Delta^2(p) \sin^2 2\theta + \left[\Delta(p) \cos 2\theta - V_L - V_T(p)\right]^2}. \quad (27)$$

Combining this with quantum damping, the effective production rate of sterile neutrinos is

$$\Gamma(\nu_\alpha \rightarrow \nu_s; p, t) \approx \frac{\Gamma_\alpha}{2} \left( P_m(\nu_\alpha \rightarrow \nu_s; p, t) \right). \quad (28)$$

This probability $P_m$ depends on the amplitude of the matter mixing angle and the quantum damping rate

$$D(p) = \frac{\Gamma_\alpha(p)}{2} \quad (29)$$

for neutrinos and antineutrinos [111, 99],

$$\langle P_m(\nu_\alpha \rightarrow \nu_s; p, t) \rangle \approx \frac{1}{2} \frac{\Delta(p)^2 \sin^2 2\theta}{\Delta(p)^2 \sin^2 2\theta + D^2(p) + \left[\Delta(p) \cos 2\theta - V_L - V_T(p)\right]^2}, \quad (30)$$

$$\langle P_m(\bar{\nu}_\alpha \rightarrow \bar{\nu}_s; p, t) \rangle \approx \frac{1}{2} \frac{\Delta(p)^2 \sin^2 2\theta}{\Delta(p)^2 \sin^2 2\theta + D^2(p) + \left[\Delta(p) \cos 2\theta + V_L - V_T(p)\right]^2}. \quad (31)$$

There are a number of consequences of this collision-dominated conversion process:

1. With appreciable mixing and specific mass differences, sterile neutrinos will be fully thermalized by the oscillations via collision dominated production, when no other new physics is included. This makes sterile neutrinos that thermalize be subject to classic bounds from primordial nucleosynthesis and over-closure as well as new bounds on relativistic density and hot dark matter from the CMB and LSS. These effects are predominantly present for active-sterile mass splittings of $\delta m \gtrsim 10^{-7} \text{ eV}^2$ to $\delta m_{\alpha s} \sim 100 \text{ eV}^2$.

2. At keV-scales, there is a level of mixing at which the sterile neutrino is produced at the proper amount to be all or part of the dark matter, is mildly relativistic to nonrelativistic at matter-radiation equality and CMB decoupling, but can act as hot to “warm” dark matter since it has a significant free streaming scale. This is the Dodelson-Widrow mechanism. There are constraints on sufficiently large free streaming scales from observations of cosmological and even Milky Way Local Group structure. We discuss these constraints in Section 6.2 below.

11
3. Even during collision-domination, the system can have a resonances between the terms in brackets in the denominators of Eqs. (27), (30), (31). This allows for the resonant creation or destruction of lepton number with associated production or destruction of sterile neutrino density. The multitude of potential processes that can engender is described in the following section. The case of conversion of primordial lepton number to sterile neutrinos via this resonance is the Shi-Fuller mechanism.

3.1.2. Creation of Lepton Number

Since the sterile neutrino states are a bath of particles that are typically initially unpopulated and are placed in thermal contact with the active neutrinos via matter-affected oscillations, the system was found to be analogous to two thermal state condensed matter systems [112, 111]. In such cases a two-state neutrino system is projected into a polarization vector description

\[
\rho(p) = \frac{1}{2} P_0(p) \left[ 1 + P(p) \cdot \sigma \right],
\]

\[
\bar{\rho}(p) = \frac{1}{2} \bar{P}_0(p) \left[ 1 + \bar{P}(p) \cdot \sigma \right],
\]

where

\[
P(p) \equiv P_z(p) \hat{x} + P_x(p) \hat{y} + P_z(p) \hat{z},
\]

where \( P_0 \) and \( P \) are convenient representations of the magnitude of the elements of the density matrix \( \rho \) in terms of the Pauli matrices \( \sigma \). \( P \) is often called the polarization of the neutrino system, is dependent on time, and effectively represents the presence of asymmetry, such as the presence of active neutrinos over sterile neutrinos. Recall that diagonal elements of \( \rho \) and \( \bar{\rho} \) are the relative number density distributions or phase space distributions (PSDs) of \( \nu_\alpha \) or \( \bar{\nu}_\alpha \) and \( \nu_\nu \) or \( \bar{\nu}_\nu \).

The time evolution of the polarization vector \( P \) and magnitude \( P_0 \) is obtained from evolving the density matrix forward in time including the presence of thermal, matter and scattering effects [99]. The evolution of \( P(p), P_0(p), \bar{P}(p) \) and \( \bar{P}_0(p) \) are given by the following equations (e.g., [113]),

\[
\frac{dP}{dt} = V(p) \times P(p) - D(p)[P_z(p)\hat{x} + P_y(p)\hat{y}] + \frac{dP_0}{dt} \hat{z},
\]

\[
\frac{dP_0}{dt} \simeq \Gamma(p) \left[ f_{\nu}(p) - \frac{1}{2} (P_0(p) + P_z(p)) \right].
\]

The equations for the anti-particles are given by the substitutions \( P \rightarrow \bar{P}, \bar{P}_0 \rightarrow \bar{P}_0, V(p) \rightarrow V(p), f_{\bar{\nu}}(p) \rightarrow f_{\bar{\nu}}(p) \). \( \bar{V}(p) \) is obtained by replacing \( L^{(\alpha)} \) by \( -L^{(\alpha)} \). The total collision rates of the flavor neutrino and anti-neutrino are approximately equal, \( \Gamma(p) \simeq \bar{\Gamma}(p) \). These 8 equations form the quantum kinetic equations for active-sterile oscillation.

The damping coefficient is given by Eq. (29). The rotation vector \( V(p) \) has the following components

\[
V_x(p) = \frac{\delta m^2}{2p} \sin 2\theta_0
\]

\[
V_y(p) = 0
\]

\[
V_z(p) = V_0(p) + V_L(p),
\]

where

\[
V_0(p) = -\frac{\delta m^2}{2p} \cos 2\theta_0 + V_1
\]

\[
V_1(p) = -\frac{7\sqrt{2}}{2} \zeta(4) \frac{G_F}{\zeta(3) M_Z^2} n_{\nu_\alpha} n_T[ n_{\nu_\alpha} + n_{\bar{\nu}_\alpha} ]
\]

\[
V_L(p) = \sqrt{2} G_F n_\gamma L^{(\alpha)}.
\]
Here $n_{\gamma}$ is the photon equilibrium number density, $n_{\nu_{\alpha}}$ and $n_{\bar{\nu}_{\alpha}}$ are normalized to unity and $f_{eq}(p)$ is the Fermi-Dirac distribution with a chemical potential $\mu_{\alpha}$.

It was discovered that this active-sterile neutrino mixing system could generate a lepton number under the proper conditions, namely that the mass eigenstate more closely associated with the sterile neutrino is lighter than that of the predominantly active neutrino with which it mixes, $m_4 < m_1$ [17], and furthermore it was found that this process could be chaotic in nature with the potential indeterminacy in the lepton number [35].

There is one significant aspect of the fact that the chaos is introduced by the system: it leads to the sign of the lepton number not being predictable for certain parts of the active-sterile neutrino mixing parameter space. Furthermore, it has a fractal nature in parts of the active-sterile neutrino mixing parameter space, leading to the potential indeterminacy in the sign of any lepton number generated in the early Universe. In such a case, no matter the precision and accuracy of the parameters of any discovered light scale sterile neutrino, we would be unable to predict the sign of the lepton number generated, and it would be necessary to determine other ways, such as its effects on the weak rates setting the stage for primordial nucleosynthesis [114]. A view of the sign of the lepton number as a function of a sterile neutrino’s position in parameter space is shown in Fig. 2.

4. Light eV-scale Sterile Neutrinos: Thermalization Conditions and Early Universe Constraints

If a sterile neutrino is present, its effects on the early Universe may include lepton number generation, destruction, thermalization and potentially all of the above. The contingencies of what the sequence of what
would occur in the presence of arbitrary numbers of sterile neutrinos present with arbitrary parameters has not been completely explored. There have been a number of papers that have studied the evolution of a four neutrino system in the case of only one sterile neutrino with appreciable mixing, consistent with short baseline anomalies [115, 116, 117, 118].

A synthesis of the possibilities was in Abazajian et al. [18]. There, it was shown that the presence of a lepton number in the early Universe—in models with four or more sterile neutrinos—almost certainly produces nonthermal active and sterile neutrino spectra, regardless of the origin of the lepton asymmetry. It was also shown that a standard MSW resonance with smooth transformation until lepton number is depleted is impossible, but must be stochastic given the nature of the resonance evolution. Mixing among the active neutrinos almost certainly transfers some of the distortions and asymmetry to the electron neutrino sector [90, 92, 91], which will directly affect the weak rates an light element production in the early Universe.

A number of the possible four neutrino scenarios were studied in Ref. [119]. Much of the parameter space for active-sterile mixing is ruled out by light element abundances from primordial nucleosynthesis, plus hot dark matter ($\Omega_{\nu_s}$) constraints from the CMB plus LSS, but not all. The full quantum kinetic treatment and implications remain to be explored for 3 active neutrino—with all respective known mixings—and N sterile neutrinos and their implications for light elements, the CMB and LSS.

5. Sterile Neutrino keV-scale Dark Matter

5.1. Oscillation-based Production

In the case of production that primarily takes place via oscillations, the parameters governing production are closely tied to the detectability of sterile neutrino dark matter via X-ray observations and effects on structure formation. The primordial initial lepton number of the Universe, Eq. (25), is a single parameter that connects the resonant-production Shi-Fuller mechanism with the non-resonant production Dodelson-Widrow mechanism.

For the non-resonant Dodelson Widrow mechanism the production rate at high temperatures decreases with increasing temperature due to finite temperature and damping effects, so that the rate of production to the Hubble expansion rate is $\Gamma/H \propto T^3$, while at lower temperatures the collisions fade rapidly and $\Gamma/H \propto T^{-9}$. The peak rate of sterile neutrino production occurs at temperature $T \approx 100$ MeV.

For Dodelson-Widrow production the proper mixing required for a given sterile neutrino particle mass scale is best determined now by Ref. [120], which gives a fitting form for the proper level of production of dark matter at the proper density

$$T_{\text{max}} \approx 133 \text{ MeV} \left(\frac{m_s}{1\text{ keV}}\right)^{1/3},$$

where $m_s$ is the mass eigenstate that is predominantly sterile. Therefore, for keV-scale sterile neutrino dark matter, sterile neutrinos with sufficiently high masses—that avoid structure formation bounds—are produced above temperatures $T \gtrsim 100$ MeV.

For resonant production in universes with nontrivial primordial lepton numbers, the production temperature depends on the mass, mixing angle and the primordial lepton number, and has not been condensed.
into a three parameter fit. A sense of the production temperature scale is achieved by looking at where the resonance, which moves from low to high momenta, would be appreciably through the active neutrino momentum distribution and therefore have depleted the primordial lepton number and created the dark matter. The position of the resonance is given by [121]

$$\epsilon_{\text{res}} \equiv \frac{\delta m^2}{(8\sqrt{2} \zeta(3)/\pi^2) G_F T^4 L} \approx 3.65 \left( \frac{\delta m^2}{7 \text{keV}^2} \right) \left( \frac{10^{-3}}{L} \right) \left( \frac{170 \text{MeV}}{T} \right)^4,$$

where \( \epsilon_{\text{res}} \equiv p/T|_{\text{res}} \) is the position of the resonance.

The most accurate calculations for resonant production are given by Venumadhav et al. [104]. For lepton numbers and sterile neutrino particle masses consistent with constraints the production largely occurs at \( T \gtrsim 100 \) MeV, where the process is collisionally dominated, i.e. the interaction contribution dominates the vacuum oscillations. In this regime, the evolution of the density matrix separates out and yields a quasi-classical Boltzmann transport equation for the diagonal terms, which are the momentum space distributions of the active and sterile components [99, 122, 123]. The quasi-classical Boltzmann equation that governs the incomplete coming to equilibrium of the sterile neutrino momentum space distribution function is

$$\frac{\partial}{\partial t} f_{\nu_i}(p, t) - H p \frac{\partial}{\partial p} f_{\nu_i}(p, t) = \sum_{\nu_i + a + \cdots \rightarrow i + \cdots} \int \frac{d^3 p_a}{(2\pi)^3 2E_a} \cdots \frac{d^3 p_i}{(2\pi)^3 2E_i} \cdots (2\pi)^4 \delta^4(p + p_a + \cdots - p_i - \cdots) \times \frac{1}{2} \left[ (P_m(\nu_i \rightarrow \nu_a; p, t)) (1 - f_{\nu_i}) \sum |M|_{i+\cdots+a+\nu_a+\cdots}^2 f_i \cdot (1 \mp f_{\nu_i}) \cdots \right. \left. - (P_m(\nu_a \rightarrow \nu_i; p, t)) f_{\nu_a} (1 - f_{\nu_i}) \sum |M|_{\nu_a+a+\cdots+i+\cdots}^2 f_a \cdot (1 \mp f_i) \cdots \right].$$

There is an analogous equation for the antineutrinos. Here, the \( f(p) \) are momentum space distribution functions for particles with momentum \( p \) and energy \( E \), and \( H \) is the Hubble expansion rate at that time. The sum on the right hand side is over all reactions that consume or produce an active neutrino. The symbol \( \sum |M|^2 \) denotes the squared and spin-summed matrix element for the reaction, and the factors of \((1 \mp f)\) implement Pauli blocking/Bose enhancement, respectively. The factor of \(1/2\) accounts for the fact that only one active neutrino state in the two-state system interacts [112, 111, 124]. The \( P_m \) are matter-affected active-sterile oscillation probabilities, which depend on the vacuum mixing angle \( \theta \), and are modified by interactions with the medium as described above. In terms of these quantities, the oscillation probabilities are [122, 123]

$$\langle P_m(\nu_\alpha \leftrightarrow \nu_\beta; p, t) \rangle = (1/2) \Delta^2(p) \sin^2 2\theta \left\{ \Delta^2(p) \sin^2 2\theta + D^2(p) + [\Delta(p) \cos 2\theta - V^L - V^{\text{th}}(p)]^2 \right\}^{-1},$$

where \( D(p) \) is the quantum damping rate described above. Here, \( \Delta(p) \) is the momentum scaled vacuum oscillation rate, \( \Delta(p) \equiv \delta m^2_{\nu_\alpha, \nu_\beta}/2p \). The neutrino self energy is split into the lepton asymmetry potential \( V^L \), and the thermal potential \( V^{\text{th}} \) (the asymmetry contribution enters with the opposite sign in the version of Eq. (40) for antineutrinos). The net interaction rate for an active neutrino is

$$\Gamma_{\nu_\alpha}(p) = \sum_{\nu_\alpha + a + \cdots \rightarrow i + \cdots} \int \frac{d^3 p_a}{(2\pi)^3 2E_a} \cdots \frac{d^3 p_i}{(2\pi)^3 2E_i} \cdots (2\pi)^4 \delta^4(p + p_a + \cdots - p_i - \cdots) \times \sum |M|_{\nu_\alpha + a + \cdots \rightarrow i + \cdots}^2 f_a \cdot (1 \mp f_i) \cdots$$

---

The code for the numerical calculation of the production is available at https://github.com/ntveem/sterile-dm
The resulting Boltzmann equation for quantum-damped, collisionally-driven sterile neutrino production is [26]

$$\frac{\partial}{\partial t} f_{\nu_s}(p, t) - H p \frac{\partial}{\partial p} f_{\nu_s}(p, t) \approx \frac{\Gamma_{\nu_s}(p)}{2} \langle P_m(\nu_\alpha \leftrightarrow \nu_5; p, t) \rangle [f_{\nu_\alpha}(p, t) - f_{\nu_s}(p, t)],$$  \hspace{1cm} (42)

and there is a related equation for antineutrinos. There have been some questions as to the subtleties regarding the effects of quantum-damping in the case of resonance [125], but tests with the full density matrix formalism find that the quasi-classical treatment is appropriate [126].

An example of how resonant sterile neutrino dark matter production occurs is shown in Fig. 3, where one can see that there are two resonances, but one is dominant, proceeding from low to high momenta. An example of final momentum space distributions are shown in Fig. 4. These are for the the region of parameter space for the 7 keV resonantly-produced sterile neutrino decaying dark matter consistent with detections of an unidentified line at \( \approx 3.5 \) keV, shown as stars in Fig. 7. When compared to the Fermi-Dirac distribution of the active neutrinos, the sterile neutrino dark matter distribution is “cooler” (though explicitly non-thermal), with \( \langle p_{\text{sterile}} \rangle < \langle p_{\text{active}} \rangle \). This means that for a given particle mass, the free streaming scale is smaller, and specifically compared to Dodelson-Widrow active neutrinos, which have a momentum space distribution closer to that of the active neutrinos, yet still nonthermal [120]. This means that for a fixed particle mass, resonantly produced dark matter acts more like CDM than that particle mass Dodelson-Widrow mechanism produced sterile neutrino. We discuss implications for structure formation calculations in the following section.
5.2. Low Reheating Temperature Universe and the Visible Sterile Neutrino

One possibility in the case of oscillation-based production is that the Universe never heated up to the full peak-production temperatures of the given production mechanism. In this case, large mixing angles can produce the proper production levels because of the reduction of the level of a strong thermal bath. This possibility was proposed in Gelmini et al. [127]. Such low reheating scenarios can under-produce the cosmological relativistic energy density measured as the effective number of neutrinos, $N_{\text{eff}}$, but for reheating temperatures of $T_R = 5$ MeV, $N_{\text{eff}}$ can be within 10% of its canonical value, which remains consistent with current measures of $N_{\text{eff}} = 3.15 \pm 0.23$ (1σ) [128]. Future CMB experiments are forecast place stronger limits on the relativistic energy density, as $N_{\text{eff}}$, to within ~1% [129]. Though low reheating temperature universes are an unusual cosmology, they are possible, with baryogenesis occurring at low scale via an Affleck-Dine mechanism [130]. Perhaps most interestingly is that this plausible scenario marries quite disparate fields of cosmological dark matter production, cosmological large scale structure constraints on hot dark matter (HDM), galaxy formation limits on WDM and mixed WDM plus CDM, Type II supernova reheating models, as well as laboratory-based nuclear beta-decay searches and neutrinoless double beta decay searches.

For the case of a $T_R = 5$ MeV reheating temperature, the production of sterile neutrino dark matter proceeds via partial thermalization due to low temperatures and low mixing angles combined, allowing larger mixing angles that the standard thermal history to produce the requisite dark matter density. Here, the $\nu_s$ distribution function turns out to be

$$f_s(E, T) \approx 3.2 d_\alpha \left( \frac{T_R}{5 \text{MeV}} \right)^3 \sin^2 2\theta \left( \frac{E}{T} \right) f_\alpha(E, T),$$

$$\sin^2(2\theta) \times 10^{11}$$

Figure 4: Shown here are sterile neutrino momentum space distributions at close of production at $T = 10$ MeV for parameters consistent with the candidate signal at 3.5 keV. For reference we show the momentum space distribution for the active neutrinos, as the dotted line. The dark matter has a “colder” distribution with $\langle p_{\text{sterile}} \rangle < \langle p_{\text{active}} \rangle$. Parameters here are shown as stars in Fig. 7. This figure is from Ref. [104].
Figure 5: Shown here is the parameter space for a possible low reheating temperature universe with \( T_R = 5 \text{ MeV} \), for the case of \( \nu_s \leftrightarrow \nu_e \) mixing. The stacked galaxy cluster \([29]\) and M31 \([30]\) candidate signals are shown near 7 keV and \( \sin^2 2\theta \approx 10^{-7} \). The contours for fraction of dark matter production \( f_{DM} \) at this reheating are shown, with the lower one \( f_{DM} = 7 \times 10^{-4} \) required to match that needed for the candidate signals. The black solid regions are constrained by laboratory experiments, independent of any astrophysical or cosmological models: from neutrinoless double-beta decay searches in the hatched region labeled \( 0\nu\beta\beta \) \([131]\), or by nuclear beta decay kink searches in the solid black region labeled \( \beta\text{-decay} \) \([1]\). The black dot-dashed line is the forecast 1\( \sigma \) sensitivity of time-of-flight measures from the KATRIN \( \beta\text{-decay} \) experiment \([132]\). We also show constraints from large scale structure limits on the presence of hot dark matter (HDM) in light pink \([128]\), from mixed cold and warm dark matter models \([133]\), as well as local group dwarf galaxy count constraints on WDM (labeled CWDM) in red \([134]\). The blue region \( \Omega_s > \Omega_{DM} \) overproduces the dark matter. The constraint from M31 observations for the same fraction of dark matter as the signals is shown in purple, labeled H14 M31 \([134]\). The yellow region is where sterile neutrinos deplete energy in the core of a Type II supernova \([26,135,136]\), though portions of this region may also be responsible for supernova shock enhancement \([135]\) or the origination of pulsar kicks \([21]\).

where \( d_\alpha = 1.13 \) for \( \nu_\alpha = \nu_e \), and \( d_\alpha = 0.79 \) for \( \nu_\alpha = \nu_{\mu,\tau} \). The fraction of the sterile neutrino distribution produced is

\[
f \equiv \frac{n_{\nu_s}}{n_{\nu_\alpha}} \approx 10d_\alpha \sin^2 2\theta \left( \frac{T_R}{5 \text{ MeV}} \right)^3 .
\]  

(44)

In such a model, one must match the requisite decay rate with the fraction of dark matter that is sterile neutrino dark matter via the production calculation at \( T_R = 5 \text{ MeV} \). In such a case, the signals at 3.55 keV \([29,30]\) are at a relatively much higher mixing angle, \( \sin^2 2\theta \approx 10^{-7} \), as shown in Fig. 5, where they are within reach of laboratory experiments. To match the decay rate, the production makes sterile neutrino dark matter that is a small fraction of the total dark matter \( f_{DM} \equiv \Omega_s/\Omega_{DM} \approx 7 \times 10^{-4} \) for the case of the
signal from stacked clusters [29].

The parameter space of low-reheating temperature models of \( T_R = 5 \) MeV, for the case of \( \nu_s \leftrightarrow \nu_e \) mixing, is shown in Fig. 5. We also show regions that are constrained by laboratory experiments, independent of any astrophysical or cosmological models. Kink searches in nuclear beta-decay are direct kinematic searches [1]. Neutrinoless double-beta decay searches \( (0\nu\beta\beta) \) are sensitive to \( \langle m_\nu \rangle = m_s \sin^2 \theta e^{i\beta} \), where \( \beta \) is the Majorana \( CP \)-violating phase. The best \( 0\nu\beta\beta \) limits are currently \( \langle |m_\nu| \rangle < 61 - 165 \) meV, which correspond to \( m_s \sin^2 2\theta < 0.66 \) eV [131]. Kinematic reconstruction by time-of-flight in beta decay in the KATRIN \( \beta \)-decay experiment has sensitivity to this region [132]. Even partial thermalization of the sterile neutrino in cases where its free streaming scale is large gives constraints from the CMB and large scale structure, \( \sigma m_s + f m_s < 0.23 \) eV [128], where \( f \) is the fraction of the active neutrino density in sterile neutrinos (Eq. (44)). From affecting galaxy counts, there are constraints on specific ratios of mixed cold and warm dark matter models. Namely, mixed CWDM models constrain 5% WDM with equivalent \( m_s = 0.08 \) keV, 20% fraction WDM with equivalent \( m_s = 0.90 \) keV, 80% fraction WDM with equivalent \( m_s = 4.5 \) keV [133], plus local group dwarf galaxy count constraints on WDM at 100% fraction limits at \( m_s \approx 9 \) keV [134, 137]. Constraints from M31 observations apply, but move commensurately up for the same fraction of dark matter as the signals shown [134]. Sterile neutrinos deplete energy in the core of a Type II supernova [138, 26, 135, 136], though portions of this parameter space can also be responsible for supernova shock enhancement [135] or the origination of pulsar kicks in portions of that region [21].

The sterile neutrino dark matter’s large mixing angle in low-reheating scenarios is much more accessible to laboratory experiments like \( \beta \)-decay [139, 140, 132]. There are also efforts at testing these couplings via K-capture experiments [141, 142] including novel methods employing atom traps that search for the massive admixed neutrino in nuclear beta decay via complete kinematic reconstruction of the final state [141, 143].

5.3. Non-oscillation Production: Particle Decays

In addition to the non-resonant and resonant oscillation production models, a few other production mechanisms have been proposed via particle decays. For most cases, a generic scalar \( S \)-particle is introduced with an interaction Lagrangian

\[
\mathcal{L}_{\text{int}} = \frac{y}{2} (\nu_R) \nu_R S + h.c. \tag{45}
\]

\( S \)-particles are created in some process in the early Universe and could decay into sterile neutrinos at some later point. Clearly, the abundance is now independent of the active-sterile mixing angle, but in general, if this mechanism is responsible for the bulk of production, the mixing angle must be below that produced by non-resonant thermal production. In addition, depending on the nature of the mechanism, the scalar-decay sterile neutrino dark matter could be “warmer” or “colder” than in oscillation production. A thorough review of these production mechanisms is given in Ref. [60], and we provide a brief overview here.

Regardless of the mechanism, if the active-sterile mixing is large enough for Dodelson-Widrow or Shifuller production to be significant, that has to be included. In the various particle-decay mechanisms, the “parent” \( S \)-particle can itself be in or out of thermal equilibrium at the time the sterile neutrinos are produced. The decay into the sterile neutrino dark matter has been considered via several scalar particles as a single particle decay channel or via one part of multiple possible decay channels [144, 21, 145, 146, 147, 148]. In scalar decay cases, the production of sterile neutrinos is governed by a Boltzmann kinetic equation describing their distribution \( f(p, t) \) as a function of momentum \( p \) and time \( t \) [144]

\[
\frac{\partial f}{\partial t} - H p \frac{\partial f}{\partial p} = \frac{2m_S \Gamma}{p^2} \int_{p+m_s^2/4p}^{\infty} f(E) \, dE \tag{46}
\]

for a scalar particle of mass \( m_S \). \( \Gamma \) is the partial width for scalar decay into sterile neutrinos. If the decays occur when the number of degrees of freedom in the background plasma is constant, then this kinetic equation is analytically integrable to find the late time number density of sterile neutrinos,

\[
n_0 \approx \frac{3\Gamma M_p c(5)}{3.32 \pi m_S^3 \sqrt{g_*}} \, T^3, \tag{47}
\]
where $g_*$ is the number of statistical degrees of freedom during the production epoch, $M_{\text{Pl}}$ is the Planck mass, and $\zeta(5)$ is the zeta function. The average momentum of the sterile neutrinos is $\langle p \rangle = \pi^6/(378\zeta(5))T \approx 2.45T$, which is lower than thermal $\langle p \rangle \approx 3.15T$, but not overwhelmingly colder.

There have been explorations where the parent particle is not a scalar, but rather a vector [149] or fermion [150]. In these decay-production cases, or even the oscillation-production cases, further decays of particles, including related more massive sterile neutrinos, can further “cool” the sterile neutrino dark matter relative to the thermal bath [151, 152, 153, 154]. They often cannot be not arbitrarily cooled, and could still be constrained by measures of its effects on structure formation [155, 156].

6. Sterile Neutrinos and Cosmological Structure Formation

The effects of massive sterile neutrinos that are completely or partially thermalized are like that of the active neutrinos, except that the particle mass scales and therefore structure scales can vary more considerably. The primary physical effect for all mass scales is the damping of structure on scales below the free streaming scale of the particle in the early Universe before structure starts to grow appreciably, at matter-radiation equality,

$$\lambda_{\text{FS}} = \int_{0}^{\lambda_{\text{EQ}}} \frac{v(t)dt}{a(t)} \approx 1.2 \text{Mpc} \left(\frac{1 \text{keV}}{m_\nu}\right) \left(\frac{\langle p/T \rangle}{3.15}\right),$$

where $\langle p/T \rangle$ is the average momentum over temperature of the particle species. For a thermalized Fermi gas with small chemical potential, $\langle p/T \rangle = 7\pi^4/180\zeta(3) \approx 3.15$. In order to produce structure below galaxy cluster scales of $\sim 10$ Mpc, the sterile neutrino can either not be the entire dark matter for the eV-scale since it would be hot dark matter, or it must be at the keV-scale such that it is not fully thermalized in its production process. In order to avoid the Cowsik-McLelland/Gershtein-Zeldovich bound, the particle mass must be $M < 94 h^2$ eV so as to not over-close the Universe [89].

Because sterile neutrinos in many cases are nonthermal, including the case of dark matter, then the free-streaming scale change due to “cooler” $\langle p/T \rangle$ must be taken into account in Eq. (48). This is accurately performed by calculating the free streaming modification of the transfer function of the linear total matter power spectrum by employing numerical Boltzmann solvers such as CAMB [157]. CAMB now has the ability to include extra massive sterile neutrinos, and modified versions are employed to solve for the case of sterile neutrino warm dark matter with nonthermal momentum space distributions [120, 158, 104].

6.1. eV-Scale Sterile Neutrinos

For the case of eV-scale sterile neutrinos that are partially or completely thermalized, the constraints on their modification of the linear matter power spectrum as measured in large scale structure (LSS) are similar to the constraints on massive active neutrinos, except for the fact that a fourth neutrino species must be included in the calculation. That is, thermalizing a sterile neutrino clearly alters the total radiation energy density at early times, when they are relativistic, but then the massive neutrino modifies the matter density. Therefore, a partially or fully thermalized sterile neutrino is subject to constraints from big bang nucleosynthesis (BBN) [159] as well as LSS. Note that massive sterile neutrinos are sometimes thought to contribute to the relativistic energy density as quantified by the effective number of neutrinos, which is specifically defined as the relativistic energy density relative to a single massless active neutrino density,

$$N_{\text{eff}} \equiv \rho_\nu/n_\nu.$$ (49)

Therefore, when the sterile neutrino is nonrelativistic, as in the era of structure formation with photon temperatures $T_\gamma \lesssim 1$ eV, $N_{\text{eff}}$ is not an accurate descriptor. $N_{\text{eff}}$ is an accurate characterization of eV-scale neutrinos at the BBN epoch at $T_\gamma \approx 1$ MeV, where relativistic energy density modifies weak freeze-out, the neutron-to-proton ratio entering BBN, and therefore the light element abundances, which constrain the number of thermalized sterile neutrinos in the early Universe [160]. Note that keV-scale dark matter sterile neutrinos do not affect BBN since they are, by definition, produced at the abundance of dark matter, which is very small in relative density to radiation in the radiation-dominated BBN epoch.
The effects of massive sterile neutrinos on LSS observables are described in detail in the review by Abazajian & Kaplinghat [19]. The effect of neutrinos on large scale structure is reflected on their effects on the primordial power spectrum of perturbations in the early Universe. The evolution of the inhomogeneities are linear in early times. The solution of the evolution of inhomogeneities from inflation through the radiation dominated, matter dominated and vacuum dominated eras are delineated clearly in recent texts, e.g. Ref. [161]. In short, the gravitational potential, \( \Phi \), drives the linear growth of perturbations. The primordial potential \( \Phi_P \) is modified by the physics of the perturbations in the early epoch, described by the scale dependent transfer function. Growth of inhomogeneities further modifies the linear potential to give the scale which may be observed at late times, \( \Phi (k, a) \), where \( k \) is the wavenumber of the potential and \( a \) is the scale factor of the Universe:

\[
\Phi (k, a) = \Phi_P \times \{ \text{transfer function}(k) \} \times \{ \text{growth function}(k, a) \} .
\] (50)

Neutrinos affect these processes by not participating in gravitational collapse until they become nonrelativistic. Instead, neutrinos free-stream out of the gravitational potential wells that are populated by dark matter and baryons. Massive eV-scale sterile neutrinos produce the same effects as active neutrinos since they are both decoupled from baryons (via charged leptons) very early (\( T \sim 1 \text{ MeV} \)). All forms of massive neutrinos do participate as matter when they become nonrelativistic. This occurs at a scale when \( k_B T \nu (a) \sim m_\nu c^2 \), which corresponds to a redshift of \( z_{\nu} \sim 300 \) (\( m_\nu / 50 \text{ meV} \)). Primordial fluctuations in neutrinos are streamed away at scales below the horizon at \( z_{\nu} \). In comoving space, this corresponds to a wavenumber

\[
k_{\nu} \equiv a_{\nu} H (a_{\nu}) \approx 0.04 a^2 \sqrt{\Omega_m \Omega_\Lambda^{-1} + \Omega_\Lambda} \left( \frac{m_\nu}{50 \text{ meV}} \right) h/\text{Mpc}
\] (51)

where \( \Omega_m \equiv \rho_m / \rho_{\text{crit}} \) is the fraction of critical density in all matter. Larger than this scale, \( k_{\text{NR}} \), adiabatic perturbations in neutrinos, dark matter and baryons are coherent and are sufficiently described by a single perturbation \( \delta_m = \delta \rho_m / \rho_m \). However, even when the neutrinos are nonrelativistic, they have finite velocity dispersion that prevents clustering and that free-streaming scale is what is described in Eq. (48).

The detailed effects of massive eV-scale sterile neutrino free streaming below the scale defined by Eq. (48) is quantified accurately by Boltzmann solvers of perturbation evolution. The effects suppress structure below \( \lambda_{\text{FS}} \) in the linear power spectrum \( P(k) \) by an amount at small scales by

\[
\frac{\Delta P_k}{P_k} \approx -8 \frac{\Omega_\nu}{\Omega_M} = -8 \frac{\Sigma m_\nu}{94 \Omega_M h^2 \text{eV}} ,
\] (52)

where \( \Omega_\nu = \rho_\nu / \rho_{\text{crit},0} \) is the present critical density in neutrinos and \( \Sigma m_\nu \) is the total neutrino mass in active and sterile neutrinos. The amplitude of the matter power spectrum at small scales is typically parameterized by \( \sigma_8 \), the rms amplitude of fluctuations integrated over a scale of 8 \( h^{-1} \text{Mpc} \). There are a number of LSS data sets that indicate suppressed power at small scales, the so-called “\( \sigma_8 \) problem.” These include weak lensing, cluster abundance, and Lyman-\( \alpha \) forest measures of the small scale amplitude [162, 163, 164, 165]. These could indicate a non-trivial non-zero neutrino mass, possibly in the form of a sub-eV-scale active or sterile neutrino. However, for this to be in the form of a sterile neutrino, the extra relativistic energy density it would impose at early times is constrained by BBN [166], though partial thermalization could accommodate both [167]. The tension between \( H_0 \) inferred from the CMB and the \( H_0 \) measured directly in the local Universe could also indicate extra radiation density. However, evidence suggests that other new physics may be responsible such as features in the primordial power spectrum [168] or non-standard dark energy [169].

6.2. keV-Scale Neutrinos: Warm to Cold Dark Matter

For many simple models of warm dark matter particle production, there is a unique relationship between the particle mass and its effect on the free streaming scale, \( \lambda_{\text{FS}} \), Eq. (48). The topic of what particle-mass constraints arise from structure formation is often fraught with misunderstanding. Often constraints in the literature are provided on an ill-defined “warm dark matter” particle mass, but sometimes more carefully on “thermal warm dark matter” particle mass. The case of truly thermal warm dark matter is well defined,
as well as the standard Dodelson-Widrow production mechanism. These two cases were originally explored in Colombi, Dodelson & Widrow [25], which we follow here. Other cases typically have no clear mapping between particle mass and measurable structure formation effects \( \lambda_{FS} \), but are certainly quantifiable, which we also discuss below.

For the case of a purely thermal warm dark matter “x” particle, the models can be interpolated between hot dark matter (HDM) and CDM with the single function 

\[
 f_x(v) = \frac{\beta}{e^{p/\alpha T_x} + 1},
\]

where \( T_x \) is the photon temperature, \( v = p/(p^2 + m_x^2)^{1/2} \); \( m_x \) is the particle’s mass. Clearly, there are three parameters that define this distribution function, \( \alpha, \beta \) and \( m_x \). For the case of standard fermionic hot dark matter, \( \alpha = (4/11)^{1/3}, \beta = 1 \). The particle mass is then set to provide the requisite fraction of critical density \( \Omega_{HDM} \). Therefore, there is no freedom in the shape of the distribution function and, for pure HDM, a fixed distribution function for a given density. For the case of pure CDM, the velocity dispersion is negligible and the shape or form of the distribution function is not important. For the case of the parameterization in Eq. (53), CDM corresponds to the limiting case of \( \alpha = \) constant, \( \beta \to 0 \) and \( m_x \to \infty \) (equivalently, \( \alpha \to 0 \) and \( m_x \to \infty \)).

For what generic thermally-produced WDM like both sterile neutrinos and gravitinos, \( \alpha \) and \( \beta \) can be different from the HDM values. Fixing the density of particles to provide all or a fraction of the cosmological dark matter gives a constraint 

\[
 \Omega_x h^2 = \beta \left( \frac{\alpha^3}{4/11} \right) \left( \frac{m_x}{94 \text{ eV}} \right).
\]

We are left with a two free parameters, which are typically chosen to be \( m_x/\alpha \) (proportional to \( m_x/T_x \)) and \( \alpha \). The parameter \( m_x/\alpha \) (or \( m_x/T_x \)) defines the shape of the dark matter power spectrum and therefore the free streaming scale. A value of \( \alpha \) corresponds to early-decoupled dark matter—like gravitino dark matter—and another to the simplest production models of Dodelson-Widrow sterile neutrino dark matter. In the class of models where Eq. (53) describes the distribution function for sterile neutrino dark matter and early-decoupled thermal WDM, the particle masses are related as [170]

\[
 m_{DW,\text{ideal}} \approx 4.46 \text{ keV} \left( \frac{m_{\text{thermal}}}{1 \text{ keV}} \right) \left( \frac{0.12}{\Omega_x} \right).
\]

However, as discussed in Colombi, Dodelson & Widrow [25] and detailed in Abazajian [120], this relation is altered if the number of degrees of freedom in the plasma is changing during the production epoch, which is often a dramatic because peak production typically occurs near the quark-hadron transition in the early Universe [171], where the time-temperature evolution is significantly altered, and the number of weak-scatterers drastically changes. These effects combine to “cool” Dodelson-Widrow produced sterile neutrino dark matter [120], which alters the relation for Eq. (55) to be different by 20% for typical parameters, and alters the constraints for Dodelson-Widrow dark matter when these effects are included [172]. The full momentum space distributions even for the Dodelson-Widrow case must be included into cosmological Boltzmann solvers in order to arrive at accurate constraints, though Eq. (55) is approximately correct. Given that pure Dodelson-Widrow sterile neutrino dark matter is excluded by Local Group constraints [134], this is not a significant issue any longer, but pertains to accurate characterizations of WDM particle mass limits in general. For the case of resonantly-produced Shi-Fuller sterile neutrino dark matter the quasi-thermal momentum space distribution of Eq. (53) is invalid, with the momentum space distribution highly nonthermal and “cooler” than thermal (Fig. 4). Therefore, the full momentum space distribution must be employed in its effects on structure formation [104] and that then used for structure formation constraints [173, 137].

One of the most potentially powerful constraints on the matter power spectrum at small scales affected by WDM is the clustering of gas as measured along the line of sight to distant quasars in the Lyman-\( \alpha \) forest [170, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181]. The claimed level of the most recent constraints by Iršič
et al. [181] place an approximate limit on the ideal Dodelson-Widrow mass (Eq. (55)) when mapped from the thermal WDM limits of \(m_{\text{thermal}} \geq 5.3 \text{ keV}\) to be \(m_{\text{DW,ideal}} \geq 41 \text{ keV}\) (95% CL). When combined with X-ray limits, they strongly exclude Dodelson-Widrow sterile neutrino dark matter from being all of the dark matter. These are considerably weakened when sterile neutrinos are only partially the dark matter [182]. The Lyman-\(\alpha\) forest is a potentially very powerful tool to measure the small-scale matter power spectrum. It relies on mapping the clustering of neutral gas in one dimension to the 3-dimensional full matter power spectrum. For some time, this has been known to have potential systematic problems in entangling the thermal history of the intergalactic medium—via the thermal broadening and Jeans pressure-support of the gas—with the underlying matter power spectrum [183]. This has become more apparent with very high resolution hydrodynamic simulations like those in Kulkarni et al. [184], where pressure support in the gas was shown to greatly affect the flux power spectrum at high redshift, and the recovered nonlinear flux power spectrum at late time varied greatly from the linear theory methods typically used in cosmological analyses like that done to probe WDM. Kulkarni et al. also show the temperature density relation has a dispersion that is highly non-Gaussian and that temperature-density relation should be augmented with a third pressure smoothing scale parameter \(\lambda_F\). The Lyman-\(\alpha\) forest is argued further to be best used as a probe of the epoch of cosmological reionization [185].

A potentially strong probe of WDM vs. CDM is the formation of structure at high redshift probed by galaxy number counts [186] as well as reionization [187]. The limits from reionization from the optical depth to the cosmic microwave background as measured by Planck are at the level of \(m_{\text{thermal}} \gtrsim 1.3 \text{ keV}\) [186], while recent limits from the luminosity function of high redshift galaxies are at the level of \(m_{\text{thermal}} \geq 2.5 \text{ keV}\) at 2\(\sigma\) [156]. The sensitivity of the James Webb Space Telescope (JWST) to galaxy counts at even higher redshift, which are even more sensitive to WDM suppression of structure formation, will push to sensitivities to even higher thermal WDM particle masses [186]. JWST is planned to launch in October 2018.[4] Detailed observations of reionization, the Lyman-\(\alpha\) forest and high-redshift galaxy counts could differentiate between the variations of the shape of the suppression scales of different WDM production scenarios [188].

In studies of the formation of the small scale structure as probed in the Local Group of galaxies and the cores and central densities of galaxies, there remain too-low of a central density profile compared to that expected in CDM of dwarf galaxies that are satellites as well as in the field—i.e., not gravitationally bound to another galaxy [189]. This has been dubbed the too-big-to-fail problem, and it can be alleviated by WDM of the proper free streaming scale, at approximately the free streaming scale provided by \(m_{\text{thermal}} \approx 2 \text{ keV}\) [190, 133]. Very significantly, it was pointed out that this free streaming scale was matched by sterile neutrino dark matter in the region of parameter space consistent with the 3.5 keV candidate dark matter decay signal in the X-ray, discussed below [158, 191]. Resonantly-produced Shi-Fuller sterile neutrino dark matter in the 7 keV, \(\sin^2 2\theta \sim 10^{-10}\) region produce a range of cutoff scale consistent with 1.5 keV \(\lesssim m_{\text{thermal}} \lesssim 3.0 \text{ keV}\) [158, 104].

The Milky Way’s Local Group satellite galaxy counts can also provide limits on the free streaming scale since the free streaming suppresses dwarf galaxy formation [192, 193, 191, 173]. As dwarf galaxies are discovered by deep all-sky observations, the limits have increased to place tension with WDM suppression scales in the region consistent with resonantly-produced 7.1 keV sterile neutrino dark matter in the region of the 3.55 keV signal [137]. This tension may make more attractive the scenario where 10% to 20% of the dark matter is Dodelson-Widrow sterile neutrino dark matter, with the rest being some other form. A 10% to 20% fraction of Dodelson-Widrow sterile neutrinos can produce the 3.55 keV signal, with a mixing angle that is commensurately approximately five to then times larger, in order to continue to match the observed flux in the signal with a smaller sterile neutrino dark matter mass in the field of view [29]. This case, where there is a mixed cold plus warm dark matter, escapes constraints from galaxy counts and the Lyman-\(\alpha\) forest [182], and may still alleviate small-scale structure challenges [133].

---

7. keV Sterile Neutrino Dark Matter Detectability in X-ray Observations

7.1. Methods & Current Results

The fact that a light, neutral lepton, like a sterile neutrino, would have a radiative decay mode was first pointed out and calculated by Shrock [194] and independently by Pal & Wolfenstein [195]. For the Majorana neutrino case, the decay rate is

\[ \Gamma_\gamma(m_s, \sin^2 2\theta) \approx 1.36 \times 10^{-30} \text{s}^{-1}\left(\frac{\sin^2 2\theta}{10^{-7}}\right)\left(\frac{m_s}{1 \text{ keV}}\right)^5, \]  

(56)

where \( m_s \) is the mass eigenstate most closely associated with the sterile neutrino, and \( \theta \) is the mixing angle between the sterile and active neutrino. For the case of a Dirac sterile neutrino, the decay rate is reduced by a factor of two. The decay of a nonrelativistic sterile neutrino into two (nearly) massless particles produces a line at energy \( E_\gamma = m_s/2 \).

The radiative decay time is orders of magnitude greater than the age of the universe (which is necessary for a viable dark matter candidate) but the number of particles in the field of view of the Chandra or XMM-Newton observatories is approximately \( \sim 10^{78} \), which makes this decay for the candidate signal we will discuss below to be at the level of \( 10^{48} \) s\(^{-1}\) for a \( 10^{15} \) M\(_\odot\) cluster of galaxies halo.

The radiative decay of sterile neutrino dark matter was found to be constrained by the diffuse X-ray background by Drees & Wright [196], and this was reiterated in Ref. [27]. The first proposals of the estimated sensitivity of looking for the radiative decay toward massive dark matter halos like galaxy clusters and galaxies were presented in Abazajian, Fuller & Patel [26]. A detailed signal-to-background analysis of contemporary X-ray telescopes as well as forecasts for future telescopes was performed in Abazajian, Fuller & Tucker [28] (AFT). That work found that the sensitivity provided by microcalorimeter spectroscopy and much greater effective area of the at-the-time proposed Constellation-X telescope made it particularly sensitive to the radiative decay signal. In AFT, it was also proposed as a hedge that “An exposure equivalent to that of Constellation-X could be obtained by a stacking analysis of the spectra of a number of similar clusters.”

Since the initial proposal of the search of the radiative decay line in clusters of galaxies, field galaxies, and the cosmic X-ray background by AFT, many groups have conducted follow-up searches for the signature decay in the cosmic X-ray background [197], clusters of galaxies [198], individual dwarf galaxies [199, 200, 201, 202], the Andromeda galaxy [198, 134] and the Milky Way [203]. Among the best current constraints is that from an analysis of Chandra X-ray observations of the Andromeda galaxy by Horiiuchi et al. [134]. Constraints from stacked dwarf galaxy observations are comparable in strength [204]. There are also constraints from Perseus observations from Suzaku which extend to higher energies and masses [205]. At the highest energies, constraints exist from the Fermi Gamma-Ray Space Telescope Gamma-Ray Burst Monitor (GBM) [206] and INTEGRAL observations of the Milky Way halo [207]. Several of these constraints are shown in Fig. 6. For a considerable amount of time since these methods were proposed, no significant detections of unidentified candidate dark matter lines had been found, with only upper limits to the decay flux.

In early 2014, Bulbul et al. [29] used stacked cluster observations totaling over 6 Ms in exposure time and detected an unidentified line near 3.55 keV in energy at high significance, 4 to 5\( \sigma \), using both the PN and MOS CCDs aboard XMM-Newton. The signal was also detected in Chandra observations of the Perseus cluster, at 2.2\( \sigma \) in that work. As seen in Fig. 7, the signal immediately straddled the robust constraints from M31 [134], which used Chandra data and were available even at that time. The Bulbul et al. analysis was quite thorough in studying the possible atomic and instrumental sources of the line, and anticipated much of the followup work. They showed that potassium lines are far too low in emissivity and relative abundance to account for the line, and that limits on partner lines of Ar XVII and Cl XVII with stronger emissivity strongly constrain those lines near the 3.55 keV feature. Bulbul et al. also proposed that the planned ASTRO-H X-ray Space Telescope would be sensitive to the Doppler broadened energy of the dark matter line due to its velocity dispersion, which turned into one of the key analyses of the mission prior to...
Figure 6: The full parameter space for sterile neutrino dark matter, when it comprises all of the dark matter, is shown. Among the most stringent constraints at low energies and masses are constraints from X-ray observations M31 Horich et al. [134], as well as stacked dwarfs [204]. Also shown are constraints from the diffuse X-ray background [197], and individual clusters “Coma+Virgo” [208]. At higher masses and energies, we show the limits from Fermi GBM [206] and INTEGRAL [207]. The signals near 3.55 keV from M31 and stacked clusters are also shown [29, 30]. The vertical mass constraint only directly applies to the Dodelson-Widrow model being all of the dark matter, labeled “DW,” which is now excluded as all of the dark matter. The Dodelson-Widrow model could still produce sterile neutrinos as a fraction of the dark matter. We also show forecast sensitivity of the planned Athena X-ray Telescope [209].

Shortly after the Bulbul et al. result, Boyarsky et al. [30] reported the signal as being seen at 3σ in XMM-Newton observations of M31, which had more exposure time at that point than that available from Chandra, as well as at 2.3σ from XMM-Newton observations of the outskirts of the Perseus Cluster. They reported a combined statistical significance of 4.4σ. The M31 signal parameters are shown in Fig. 7.

Significant observational work followed. An early result was the detection of the signal in the Milky Way’s Galactic Center [210, 211]. Urban et al. [212] detected the signal toward the Perseus Cluster using Suzaku.

---

5This velocity broadening occurs for the plasma and dark matter virialized system because $m_s \approx 7\text{ keV} \ll m_{\text{Ar}} \sim m_{\text{K}} \sim 40\text{ GeV}$, leaving the dark matter particle to have a much higher velocity dispersion for a matched kinetic energy distribution. Note that ions such as Ar and K will have turbulent velocity broadening as well, at a level of 120 km/s relative to 100 km/s thermal broadening at $T = 4\text{ keV}$, though still very low relative to the dark matter.
Figure 7: X-ray line detections consistent with sterile neutrino dark matter are shown here. The dark colored regions are 1, 2 and 3 $\sigma$ from the MOS (blue) and PN (red) stacked clusters by Bulbul et al. [29], the Bulbul et al. core-removed Perseus cluster (green), and M31 (orange) from Boyarsky et al. [30]. Also shown are the 1 and 2 $\sigma$ regions of the detection in the Galactic Center (GC) [210] as well as the >2$\sigma$ line detections in 1. Abell 85; 2. Abell 2199; 3. Abell 496 (MOS); 4. Abell 496 (PN); 5. Abell 3266; 6. Abell S805; 7. Coma; 8. Abell 2319; 9. Perseus by Iakubovskyi et al. [215]. Numbers in the plot mark the centroid of the regions, with MOS detections in orange and PN in purple. We also show, in purple, the region consistent with the signal in Chandra Deep Field observations, with errors given by the flux uncertainty, i.e., not including dark matter profile uncertainties [216]. The lines show constraints at the 90% level from Chandra observations of M31 (14) [134], stacked dwarf galaxies (M14) [204], and Suzaku observations of Perseus (T15) [205]. Stars mark the models shown in Fig. 4.

data at the level of 7.4$\sigma$ (reported as a $C$-statistic). Urban et al. reported that the radial flux profile was inconsistent with that expected from a dark matter halo, but a subsequent analysis of the same data found that profile to be consistent with a halo and with dark matter decay [213]. The line had weak indications at $\sim$2$\sigma$ in stacked Suzaku cluster observations [214]. The 3.5 keV line was seen at >2$\sigma$ significance in eight new clusters by Iakubovskyi et al. [215], with a redshifting of the line energy consistent with cosmological origin, i.e., not instrumental.

In another follow-up, Anderson et al. [217] stacked approximately 80 galaxies from XMM-Newton and Chandra data, and claimed a high-significance exclusion of the 3.5 keV line as due to dark matter decay. The analysis relied on a model for the continuum of the stacked spectrum that had large positive and negative residuals, approximately the size of the 3.5 keV signal (c.f. Anderson et al.’s Fig. 4). Despite this overwhelming systematic uncertainty, this work used the statistical deviation from the continuum to place their quoted limit. Another follow-up analysis by Jeltema & Profumo (JP) [211] initially fit an unconstrained
potassium (K XVIII) line to XMM-Newton data, subtracted the fit from the data, and used the modified data to place a constraint on the presence of dark matter decay. This is prima facie a highly problematic methodology. JP also used line ratios to argue that the temperature of Perseus and stacked clusters could have a large contribution from cool phase gas, and therefore the emissivity of K XVIII is not well known. Those line ratios were shown to be inaccurate in Ref. [218]. JP also used an energy window in their M31 analysis that was too narrow to significantly detect the line toward M31 [219].

In related work, Ref. [220] attempted a characterization of the spatial origin of the 3.5 keV photons from the Milky Way Galactic Center and Perseus Cluster. The Galactic Center analysis in that work is compromised by the lack of incorporation of X-ray absorption, which is very high in the Galactic Center direction, and likely patchy and irregular, because of the irregular coverage by molecular clouds [221]. The observed variation in hydrogen column density gives a qualitative idea of the possible spatial variations of the brightness of any extended signal, like dark matter, toward the Galactic Center. Therefore, a proper dark matter template will not be symmetric. The column density of hydrogen on the sky is likely quadrupolar (cf. Fig. 2 in Ref. [220]) since molecular clouds tend to align with the Galactic plane. The spatial analysis toward the Perseus Cluster is also problematic in that the line in any spatial portion is of order 1% of the continuum flux (discussed in Appendix B of [213]). The spatial continuum templates that Ref. [220] used to produce a map of the 3.5 keV excess signal cannot be normalized to an accuracy anywhere near that required, 1% (given the broad range of continuum models used by the authors). Since the continuum templates are astrophysical, any residuals caused by their inaccurate normalization would also follow the spatial distribution of that astrophysical signal, regardless of the true distribution of the much weaker 3.5 keV signal [213].

The Hitomi Telescope did receive 275 ks of data toward the Perseus Cluster before its failure which was analyzed for the presence of the 3.5 keV line [222]. The data was unprecedented in its ability to do resolved X-ray spectrometry of the Fe lines [223]. Given that the protective closed gate valve was in place, the exposure was equivalent to 70 ks of normal operations, which was far short of what would be needed to be highly sensitive. Ref. [222] analyzed XMM-Newton MOS data in the same field of view as the Hitomi data, and found the MOS data to have a higher flux within that field. Hitomi excluded the central value of the new MOS detection by $3\sigma$. The prior detections were not appreciably constrained by the Hitomi data, as shown in Fig. 7.

The NuSTAR telescope was found to be sufficiently sensitive to 3.5 keV photons, with a wide field of view, $\sim 37 \text{ deg}^2$, from “zero bounce” photons allowed into the detector because the design of the telescope’s optical bench allows for these photons in without passing through the telescope’s optics [224, 225]. This was used by Neronov et al. [224] to place constraints in the high mass range of sterile neutrino decay parameters space with NuSTAR data toward the COSMOS and CDFS empty sky fields. A few unidentified lines that could be due to instrumental effects were also detected. A line at 3.51 $\pm$ 0.02 keV was detected at 11.1$\sigma$ in that work, which is consistent with flux expected from previous detections given the dark matter in the field of view. The response of NuSTAR is very poorly known at the lower energies near 3-4 keV, and it is thought that the line is likely instrumental since the line is seen in Earth occulted data [225]. Perez et al. [225] placed constraints from observations of the Galactic Center, which are shown in Fig. 6.

The Deep Field exposures of the Chandra telescope were studied to be potentially very sensitive to dark matter decays, and placed limits on the parameter space [226]. Recent work by Cappelluti et al. [216] used over three times the exposure, with $\sim 10 \text{ Ms of Chandra observations towards the COSMOS Legacy and CDFS survey fields}$. They see a line feature at 3$\sigma$ at 3.51 $\pm$ 0.02 keV with flux (mixing angle) consistent with prior detections. These parameters are shown in Fig. 7. Significantly, Cappelluti et al. show that the line feature cannot be due to instrumental effects given the background analyses by Ref. [227], and cannot be due to K or Ar lines near 3.5 keV due to the lack of the presence of partner lines of higher emissivity. Importantly, charge exchange lines that could also be present near 3.5 keV [228, 229] were constrained by partner lines by Cappelluti et al., and more analysis of the possibility of charge exchange emissivity contributions are warranted.
7.2. Future Observations

There are a number of proposed and planned missions that are dedicated to the search of X-ray lines from dark matter or as additional science. The XQC sounding rocket flight [230] is very sensitive to the presence of a line due to its large field of view and microcalorimeter-based high energy resolution [226]. The XQC and Micro-X sounding rocket X-ray experiments could be very sensitive to the presence of the dark matter line when exposed toward or near the Galactic Center, which is being proposed for summer of 2019 for a campaign in Australia, where the Galactic Center is visible from the southern sky [231].

In 2016, the Hitomi satellite was unfortunately lost. In March 2017, Paul Hertz, director of NASA’s astrophysics division, reported a formal start will happen in 2017 of the project known as the X-Ray Astronomy Recovery Mission (XARM), which will have a soft X-ray spectrometer like that of Hitomi, prepared by NASA and JAXA for launch in March 2021 [232]. As shown in Bulbul et al. [29], and discussed above, such a mission could see the telltale sign of the velocity broadening of a dark matter line. In addition, the high energy resolution velocity differential between dark matter and gas in the Milky Way Galaxy could be used to discriminate between gas and dark matter origins of the 3.5 keV line [233, 234]. The eROSITA is an all-sky mission built by the Max Planck Institute for Extraterrestrial Physics to follow up ROSAT, and is scheduled for launch in 2018 [235]. eROSITA will have sensitivity to sterile neutrino dark matter in cross-correlation analyses [236].

The future of X-ray astronomy after Chandra and XMM-Newton is taking two paths. The Athena X-ray Observatory is planned for launch in about 2028 as a European-Space-Agency-led mission that has goals of microcalorimeter spectroscopy with resolving power of $R \equiv \lambda / \delta \lambda \approx 1000$, similar to that of XARM, and a design for wide, medium-sensitivity surveys [237]. For these goals, area is built up at the expense of angular resolution and sensitivity. The sensitivity of Athena to sterile neutrino dark matter was forecast by Neronov & Malyshev [209] and is shown in Fig. 6.

In a time frame beyond 2030, the second path after Chandra is the Lynx Telescope, which is planned with 50 to 100 times the sensitivity over Chandra and Athena. Lynx has the goals of achieving $R \approx 1000$ spectroscopy on 1” scales, adding a true third dimension to high-angular resolution X-ray data, as well as $R \approx 5000$ spectroscopy for point sources. For Lynx, area is built up while preserving Chandra angular resolution of 0.5”, with 10 times the field of view. Since the technology and science goals are still being developed, a detailed study of Lynx sensitivity to dark matter has yet to be performed. It is certain, however, that if there is a dark matter line signature well within the sensitivity of Lynx like the 3.5 keV feature, the simultaneous high energy and angular resolution of a telescope like Lynx will open up an era of dark matter astronomy where the density and velocity space mapping of dark matter could take place.

8. Conclusions

Sterile neutrinos are the leading natural extension of the Standard Model of particle physics, accommodating neutrino mass. Their presence can have significant effects on the early Universe and structure formation, at the eV scale, and could even be the dark matter, at the keV scale. The light, eV-scale sterile neutrinos can thermalize in the early Universe, affecting primordial nucleosynthesis and structure formation.

Recent indication of tensions between large scale and local measures of the amplitude of fluctuations on small scales ($\sigma_8$) and the Hubble expansion rate ($H_0$) may be indicative of new physics related to new neutrino physics, including eV-scale sterile neutrinos. The light sterile neutrinos can also be responsible for chaotic lepton number generation or lepton number depletion. Detection of sterile neutrinos in the laboratory would have significant implications for our understanding of several aspects of cosmology.

Heavier, keV-scale sterile neutrinos can be produced to be the dark matter via a number of mechanisms, through couplings to other particles, or resonant and non-resonant oscillations. The evidence for the candidate line signal near 3.5 keV is significant and building, including detections in clusters of galaxies, field galaxies and dwarfs, but the signal also has significant challenges, primarily Chandra observations of the Andromeda Galaxy and XMM-Newton observations of the Draco Dwarf Galaxy. If sterile neutrinos

\footnote{https://wwwastro.msfc.nasa.gov/lynx}
are definitively detected astrophysically or in the laboratory as all—or as a portion—of the dark matter, a significant piece of our standard cosmological and fundamental physics paradigm would be found.

Acknowledgments

I would like to thank my fruitful collaborations and discussions with Prateek Agrawal, Mike Boylan-Kolchin, Ezra Bulbul, James Bullock, Mu-Chun Chen, J. J. Cherry, Francis-Yan Cyr-Racine, André de Gouvêa, Bhupal Dev, Pasquale Di Bari, Scott Dodelson, Marco Drewe, Enectali Figueroa-Feliciano, George M. Fuller, Chris Hirata, Shunsaku Horiuchi, Patrick Huber, Manoj Kaplinghat, Alex Kusenko, Ranjan Laha, Jon Link, Maxim Markevitch, Alexander Merle, Amol Patwardhan, Oleg Ruchayskiy, Aaron Tohuvavohu, Teja Venumadhav, and Alexey Vikhlinin. I would also like to thank the Mainz Institute for Theoretical Physics program on “The Energy Scales of the Universe” for hospitality, where many of these discussions took place and part of this review was written. This work was supported in part by NSF Grants PHY-1316792 and PHY-1620638.

References


[147] P. F. Smith, Proposed experiments to detect keV range sterile neutrinos using energy-momentum reconstruction of beta decay or K-capture eventsarXiv:1610.06876.


[157] N. Menci, A. Merle, M. Totzauer, A. Schneider, A. Grazian, M. Castellano, N. G. Sanchez, Fundamental physics with the...
galaxy clusters observed with XMM-Newton.


