UNIVERSITY OF BERGEN

Department of Physics and Technology
Master Thesis in Particle Physics

Effects of dark sector interactions on active-sterile neutrino mixing

Author:
Tarje Solberg Hillersøy

June, 2021

Abstract

This thesis presents methods and theory for exploring new physics by using neutrinos as the portal. A thorough introduction on the state of neutrino physics in, and beyond the Standard Model framework, is presented. Anomalies arise from short baseline neutrino oscillation experiments, which may hint towards a fourth neutrino with a mass at the eV-scale. This is explained by the addition of a dark, Abelian symmetry group. The new symmetry is broken at the MeV-scale. To obtain sizeable active-sterile mixing, a hierarchy between Yukawa couplings is required. Majorana masses are included for right-handed neutrinos, which by the see-saw mechanism, provide a spectrum of light neutrino masses. The dark sector neutrinos are allowed to couple to the active neutrinos by Yukawa couplings which form a non-diagonal Yukawa matrix.

The interplay between neutrino masses, Yukawa couplings, mixing angles and vacuum expectation values, is studied to provide solutions for the short baseline anomalies while remaining within experimental limits. Simple parameter scans are employed to further investigate the structure of the model. A parameter region compatible with global fits of reactor neutrino data is presented. When applied to bounds arising from resonant leptogenesis, the model confines right-handed neutrino masses in the $1\text{TeV} - 100\text{TeV}$ range.

Acknowledgements

I would like to thank my supervisor Jörn Kersten for his support and guidance. On the other side of the globe, Jörn has taken his fair share of late evening Zoom-meetings, which have been of utmost importance. I would like to thank my parents, Siri and Steinar, for their inspiration and support throughout my educational years. A big thank you to my friends, the old ones and the new ones met along the road, for making these past years exciting. Thanks to Erlend and Helge for discussions about the weird and wonderful world of physics. Special thanks goes out to fellow co-inhabitants of the third floor; Bendik, Birger and Vegard for helpful discussion on programming and for your, although incorrect, interesting views on particle physics.

Lastly I would like to thank Lene, the one who makes it all worth it.

Contents

1 Introduction

The Standard Model is the framework for particle physics, it describes the interactions between the elementary particles and how they come to be. Members of the Standard Model include particles such as the familiar electron, the quantum of light, the photon, the long-awaited Higgs boson and many more. Three peculiar members of the Standard Model are the elusive neutrinos, ghostly particles which seldom interact with the others. Neutrinos come in three "flavours" (ν_e, ν_μ and ν_τ), and they are known to oscillate between flavours as they propagate through spacetime.

Remarkably, despite their elusive nature, neutrinos can lead the way towards discovering new intriguing physics. Although massive, neutrinos have a near vanishing mass in comparison to the other fermions and the origin of neutrino masses is currently an open question. The discovery of neutrino masses imply the existence of right-handed neutrinos; even more elusive particles which can only interact gravitationally and through Yukawa couplings with the already elusive neutrinos. Although the gravitational footprint of right-handed neutrinos may be too faint to discover, their existence brings about a window into the dark, unexplored realm of physics beyond the Standard Model.

Neutrino oscillation data can provide the mass squared difference between the three neutrino masses, which along with cosmological data confines the neutrino mass scale to the sub-eV-scale. This aforementioned mass squared difference between neutrino masses is the driving force of neutrino oscillations. Studies of neutrino oscillations at short baselines suggest, however, that the Standard Model picture of oscillations is incomplete. Oscillation experiments at short baselines deviate from the predicted values and point towards a mass splitting at the eV-scale responsible for the observed deviations. As the three neutrinos of the Standard Model are too light to provide the mass necessary for the observed mass splitting, a fourth neutrino at the eV-scale may be responsible. Not only is this neutrino much heavier than the known neutrinos, but it must also be a gauge singlet of the Standard Model as the decay of the Z boson is only compatible with three active neutrinos. The short baseline anomalies may point towards a whole new class of particles, sterile neutrinos.

A natural explanation of the lightness of neutrino masses is the celebrated see-saw mechanism [1–5]. The see-saw mechanism generates light neutrinos masses at the cost of generally very heavy right-handed neutrino masses. Although unconstrained, a popular scale to put right-handed neutrino masses is at the Grand Unified Theory scale, $\Lambda \sim 10^{16}$GeV, however, this scale may be much lower. The short baseline anomaly may be reconciled if one, or more, right-handed neutrinos obtain their mass at the eV-scale, which inevitably leads to very small Yukawa couplings. Another possibility is, that in their mixing, neutrinos probe an unexplored dark sector through active-sterile neutrino oscillations. This has striking implications, an active neutrino may oscillate into a sterile state, which is unable to communicate with the Standard Model, effectively rendering the once active neutrino a missing particle. This process may also occur in reverse; an active neutrino may blossom into existence from the point of view of observers made of Standard Model particles, which corresponds to a sterile neutrino oscillating into an active neutrino.

From a model-building perspective, the addition of sterile neutrinos as a springboard into additional dark interactions and dark sectors is a hot topic. The reader is assumed to be familiar with the existence of the non-luminous matter which accounts for the majority of the matter-energy in the universe, so-called dark matter. Barring a vast

array of exclusion plots, regions of parameter space where dark matter *cannot* reside, the particle content of dark matter is still as open as it was upon its discovery. If dark matter can interact with itself, then there will inevitably be additional dark particles that are responsible for the mediation of some dark charge. Until dark matter is resolved, the topic of modelling dark sectors will remain a hot topic as no scientist can turn away the urge to discover something new.

This thesis will examine the effects of adding a dark Abelian gauge group to the Standard Model symmetry. The group is endowed with a complex scalar that yields particle masses through the Brout-Englert-Higgs (BEH) mechanism [6–8]. The model includes a dark neutrino at the eV-scale and a dark, massive vector boson. The goal of this thesis is to provide the structure needed to incorporate the short baseline anomaly while providing a model which is compatible with existing limits on the properties of the dark particle content. To achieve this goal, the relevant tools are extracted from known results from quantum field theory and the Standard Model. Bounds from leptogenesis are imposed to restrict the parameter space of right-handed neutrino masses. Adapting these tools to the model is the main discussion of this thesis. In addition, simple parameter scans are conducted to further explore the parameter space the model provides.

The thesis is structured as follows; first, a basic introduction to the theories underlying quantum field theory is given. Following this, a thorough discussion is conducted on the Standard Model and the mathematical structure needed to explain it. After this, a discussion on the state of neutrino physics, structured to provide a natural springboard into the realm of sterile neutrinos, is provided. A review on physics beyond the Standard Model and how it can be probed, with emphasis on sterile neutrinos, is presented. The rest of the thesis provides a delve into model specifics; the masses, the mixing, the Yukawa structure and more. An effective $1 + 1$ mixing scenario is studied, which reveals the qualitative structure needed to yield three mixing regions; small, moderate and large. Afterwards, the complete $3 + 1$ mixing scenario is analyzed and the structure of the Yukawa matrices is inferred. The model is then subjected to bounds from leptogenesis, which provides a parameter space suitable for explaining both leptogenesis and the short baseline anomaly.

1.1 Classical mechanics

Newtonian mechanics has from its inception served as a formidable tool for calculating more or less everything humans can perceive. As first formulated by Newton in 1687, the laws of motion were given by the all-time classic equation

$$\vec{F} = \frac{d\vec{p}}{dt} = m\vec{a}. \tag{1.1.1}$$

The next hundred years gave rise to a new formulation of classical mechanics. The combined effort of mathematicians and physicists such as d'Alembert, Maupertuis, Euler and Lagrange gave rise to what is today called Lagrangian mechanics. The key principle of Lagrangian mechanics is invariance of the *action* S, from which the equations of motion (EOF) follow. In classical mechanics, the action is defined as the integral from a prior time t_1, to some later time t_2, given by the functional

$$S = \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t)\, dt. \tag{1.1.2}$$

As usual, overhead dots are time derivatives and q is a vector of coordinates from configuration space. The integrand is the revered Lagrangian, which in the simplest form is given by

$$L = T - V, \tag{1.1.3}$$

where T is the kinetic energy and V is the potential energy of the system under consideration. The principle of least action, also known as the principle of stationary action, states that any physical path is one such that the action is stationary. Hence, any physical path can be obtained as solutions to

$$\delta S = 0, \tag{1.1.4}$$

where δ is the functional derivative. The resulting equations are called the Euler-Lagrange equations and serve as an alternative, yet equivalent method of solving problems in classical mechanics.

The third formalism of classical mechanics came in 1833 with the advent of Hamiltonian mechanics. Although similar to the Lagrangian approach, the two differ both in derivation and in how the equations are solved. Lagrangian mechanics yields a second-order differential equation for each generalized coordinate, while Hamiltonian mechanics yields two equations, but to sweeten the deal, Hamilton's equations are first-order differential equations. There is much more theory to these formalism's than covered here, however as the two are related by the Legendre transformation and the Lagrangian approach is best suited for quantum field theory (QFT), only the Lagrangian formalism will be further discussed. The interested reader is referred to a textbook on classical mechanics, e.g. [9].

After some time, inconsistencies of the Newtonian theory were discovered. Essentially all these problems are related to extreme cases, e.g. the very fast, the very small or even disjoint subjects such as the notion of simultaneity. To resolve these issues, three new theories were needed, special relativity (SR), general relativity (GR) and quantum mechanics (QM). In the following sections, a brief introduction/refresher on both special relativity and quantum mechanics is given.

1.2 Special relativity and relativistic notation

In standard theoretical physics convention, *natural units* are used throughout the thesis, $\hbar = c = k = 1$.

The Einstein summation convention is also used throughout. This convention states that repeated covariant (written in subscript) and contravariant indices (written in superscript) are to be summed over

$$\sum_{\mu=0}^{3} A_\mu B^\mu \equiv A_\mu B^\mu. \tag{1.2.1}$$

In standard convention, greek letters such as α, β, μ, ν are understood to range from 0 to 3, while roman letters such as i, j, k, are taken from 1 to 3. Unless specified otherwise, the distinction between greek and roman letters is the one specified here.

Regardless of definition, the summation convention is often severely violated, especially in papers with convoluted notation. As an example, many objects have various add-ons to them, being names, labels or some overhead symbol (or in extreme settings, a combination of all). In these cases, indices are placed where ever there is room, while summation is still implied. As a less convoluted example, quantities such as $A_i B_i$ or even $A_i B_i C_i$ will appear while summation is still implied. This may often be quite confusing, so the thesis is written to respect the summation convention given by (1.2.1).

The thesis is written in the framework of special relativity, which is a four-dimensional, flat, pseudo-Riemannian manifold defined by the metric signature in the (+,-,-,-) particle physics convention

$$g_{\mu\nu} = \mathrm{diag}\left(1, -1, -1, -1\right). \tag{1.2.2}$$

In general relativity, the nomenclature is to assign $g_{\mu\nu}$[1] as the metric for any arbitrary region of spacetime (i.e. not necessarily flat) while $\eta_{\mu\nu}$ is taken as the flat spacetime metric, sometimes called the *Minkowski* metric. In this thesis, effects arising from general relativity are not considered, and hence the metric defined in (1.2.2) is used.

Elements of Minkowski space are four-dimensional vectors x, which are creatively called *four-vectors* in the standard literature. Dual to the vectors, *forms*[2] are defined as linear maps from the vector space to the underlying field, $\mathbb{F}$. Here, *field* is taken in the mathematical sense of the word, which for practical purposes is either the set of real numbers $\mathbb{R}$, or the set of complex numbers $\mathbb{C}$. Starting from a vector, the corresponding form is obtained by *index-lowering*, which is the physicist's name for the isomorphism of obtaining elements of the form given the corresponding vector and metric. The components of the form (also called the covariant components of the original vector) are given by

$$x_\mu = g_{\mu\nu} x^\nu. \tag{1.2.3}$$

[1]Strictly speaking, these are the components of the metric. However, in most papers, the distinction between the metric (g) and elements of the metric ($g_{\mu\nu}$) is assumed to be understood from context and they are often used interchangeably about each other.

[2]Forms are elements of the dual space, which is the space dual to the original vector space. Forms are called many different names, including, but not limited to, co-vectors, one-forms, dual-vectors and bras, the latter arising from Dirac's bra-ket notation.

The two most relevant four-vectors are the spacetime coordinate x and the four-momentum P. Their covariant components are given in terms of their contravariant components by

$$x_\mu = g_{\mu\nu}x^\nu = (x^0, -x^1, -x^2, -x^3) = (t, -\vec{x}) \tag{1.2.4}$$
$$P_\mu = g_{\mu\nu}P^\nu = (E, -p^1, -p^2, -p^3) = (E, -\vec{p}). \tag{1.2.5}$$

The standard vector notation of an overhead arrow is reserved for spatial three-vectors like the three-position $\vec{x}$.

Vectors and forms belong to a much broader class of objects called *tensors*. One may define a tensor in many ways, highly dependent on where one learns the topic. A mathematician will give the rigorous definition of a (p, q) tensor T, as the multi-linear map from p copies of the dual space V^* and q copies of the vector space V

$$T : (V^*)^p \times (V)^q \to \mathbb{F}. \tag{1.2.6}$$

Physicists usually refrain from this definition, instead opting to define a tensor by how it transforms under a change of coordinates. On a differentiable manifold [3], e.g. Minkowski space, components of tensors transform p times covariantly and q times contravariantly. As a concrete example, under a change of coordinates, a $(p, q) = (2, 2)$ tensor transforms as

$$T^{\mu_1'\mu_2'}_{\nu_1'\nu_2'} = \frac{\partial x^{\mu_1'}}{\partial x^{\mu_1}}\frac{\partial x^{\mu_2'}}{\partial x^{\mu_2}}\frac{\partial x^{\nu_1}}{\partial x^{\nu_1'}}\frac{\partial x^{\nu_2}}{\partial x^{\nu_2'}}T^{\mu_1\mu_2}_{\nu_1\nu_2}. \tag{1.2.7}$$

In the physics department, the definition of a tensor is usually taken as the one given in (1.2.7). Although rather tedious to compute by hand, as each component of the transformed tensor is a sum of n^4 terms (in relativistic physics $n = 4$), it is no match for a computer, unless of course, n is extremely large. When defined from transformation properties, it is clear why tensors are invaluable; tensors preserve their form under change of coordinates. The coordinate transformation imposed in equation (1.2.7) is a general one. That is, the transformed coordinates can be any differentiable functions of the previous coordinates. For special relativity, this is too general. However, if the transformed coordinates are chosen as linear functions of the previous ones and the underlying manifold is Minkowskian, then equation (1.2.7) represents a Lorentz transformation of a rank $(2, 2)$ tensor.

A consequence of how tensors transform is that their *contraction(s)* are invariant under any change of coordinates. A contraction is an operation on tensors, which acts as the map from a (p, q) tensor to a $(p - 1, q - 1)$ tensor by summing over one covariant and one contravariant index. As an example, consider the spacetime separation dx^μ. The contraction of dx^μ with itself [4] is the Lorentz invariant spacetime interval

$$dx^\mu dx_\mu = g_{\mu\nu}dx^\mu dx^\nu = dt^2 - d\vec{x}^2 \equiv ds^2. \tag{1.2.8}$$

The use of the spacetime separation, as opposed to the spacetime coordinate itself, is due to the former being observable and the latter not. Experiments can only measure

[3] A differentiable manifold is essentially a continuous space which locally resembles $\mathbb{R}^n$.

[4] Rigorously, the contraction of dx^μ is not by itself, but rather the covariant object dx_μ. However, as they are related by the metric, this distinction is to be understood implicitly.

objects relative to some background; a difference. To prove invariance of the contraction, consider again the spacetime coordinate x^μ. Under a change of coordinates from unprimed coordinates to primed ones, the contraction transforms as

$$x^\mu x_\mu = \frac{\partial x^\mu}{\partial x^{\mu'}} x^{\mu'} \frac{\partial x^{\nu'}}{\partial x^\mu} x_{\nu'} = \delta^{\nu'}_{\mu'} x^{\mu'} x_{\nu'} = x^{\mu'} x_{\mu'} \tag{1.2.9}$$

where the Kronecker delta is defined as

$$\delta^{\nu'}_{\mu'} = \begin{cases} 1, & \mu' = \nu' \\ 0, & \mu' \neq \nu'. \end{cases} \tag{1.2.10}$$

A quick, but index heavy calculation, shows that equation (1.2.9) generalizes to contractions of general tensors. For a more thorough introduction to these concepts see an introductory book on tensor calculus or a standard textbook on general relativity, e.g. [10–13].

The Lorentz invariance of tensor contractions is fundamental in physical theories and especially for the Lagrangian formulation of QFT. On a differentiable manifold, derivatives are introduced to track the change in neighbouring tensors. Derivatives can be thought of as linear maps from a (p, q) tensor to a $(p + 1, q)$ tensor

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} : T \in (p, q) \to \partial_\mu T \in (p + 1, q). \tag{1.2.11}$$

As derivatives track changes, they are associated with the movement, hence expressions involving derivatives are called *kinetic terms* in QFT jargon. Examples of kinetic terms in the Standard Model (SM) Lagrangian are vast as each field corresponding to a particle has one such term. A concrete example is the electromagnetic (EM) field strength tensor $F_{\mu\nu}$[5], defined as the anti-symmetric two-form

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{1.2.12}$$

where A_μ are the covariant components of the four-potential, defined in terms of the scalar and vector potential as

$$A_\mu = g_{\mu\nu} A^\nu = (\varphi, -\vec{A}). \tag{1.2.13}$$

In fact, all bosonic fields with spin one, so called vector fields, share the same structure of the kinetic term.

1.3 Quantum mechanics

Having discussed special relativity, the next foundation this thesis is built upon is the theory of quantum mechanics. From the point of view of quantum mechanics, particles are not represented by points in spacetime, but rather as quantum states described by vectors in Hilbert space; an infinite-dimensional vector space of square-integrable functions. Fundamental to the theory of QM is the notion of *observables*, which are quantities one can measure in an experiment. In Hilbert space, observables are the eigenvalues of Hermitian operators acting on states. In contrast to both relativity and classical mechanics, QM is a probabilistic theory, i.e. one cannot predict the outcome of an experiment with absolute

[5] Also called the EM tensor or the Faraday tensor.

certainty. Instead one rather talks about the probability that an experiment will yield a certain outcome, which are the eigenvalues of the operator the experiment is sensitive to. Another peculiarity of QM is that the theory is based upon complex numbers, as the theory does not work when restricted to real numbers. Conservation of probability is imposed by

$$\langle\psi|\psi\rangle = 1, \tag{1.3.1}$$

which says that the square of the norm of any quantum state should be unity. If a discrete basis is used, then the above equation can be written as an infinite sum, while if the basis is continuous, then the sum is replaced by an integral. For example, bound states yield quantized spectra of allowed energies when acted upon by the Hamiltonian, but the location of a particle necessarily yields a continuous spectrum. A problem with QM is that while position is an observable, time is not, rather acting as an evolution parameter. This whole discussion has sparked countless papers and books debating the role of time in quantum mechanics. Although a very interesting topic, this is not the time to venture into that realm.

A quantum state is denoted $|\psi\rangle$ and the time evolution of the state is given by the Schrödinger equation

$$i\partial_t|\psi(t)\rangle = \hat{H}|\psi(t)\rangle, \tag{1.3.2}$$

where $\hat{H}$ is the Hamiltonian; an Hermitian operator corresponding to total energy. The solution to the Schrödinger is formally given by

$$|\psi(t)\rangle = e^{-i\hat{H}t}|\psi(0)\rangle = U(t)|\psi(0)\rangle, \tag{1.3.3}$$

which can be interpreted vaguely as a rotation in Hilbert space, albeit the whole exponentiating an operator limits the interpolation of that view. The operator $U(t)$ is called the propagator as it propagates the state in time. For this reason, one can define a basis in Hilbert space which rotates at the same rate as the exponent in the above equation. In this basis, the states are stationary, while the operators receive time-dependence in what is called the Heisenberg picture of quantum mechanics. The case when operators carry time dependence while operators do not is called the Schrödinger picture of quantum mechanics.

Returning to the solution of the Schrödinger equation, as the Hamiltonian is Hermitian $\hat{H}^\dagger = \hat{H}$, the propagator is unitary, which implies conservation of probability. Combining the quantum mechanical formalism with special relativity results in quantum field theory; the foundation of this thesis.

2 The Standard Model of particle physics

2.1 Lagrangian field theory

A preliminary for understanding QFT is the Lagrangian field theory approach. As discussed previously in Section 1.1, the key to a relativistic field theory is the principle of least action. However, to account for relativity, some new insight is needed. First of all, in the non-relativistic approach, the action is given as an integral over time, which is problematic as different observers experience different flows of time. Secondly, the generalized coordinates only depend on time, which again is ambiguous in a relativistic

framework. To resolve the first issue, the Lagrangian is replaced with a new object, called the Lagrangian density, denoted $\mathcal{L}$. These are related by

$$L = \int d^3\vec{x}\, \mathcal{L}. \tag{2.1.1}$$

To resolve the second issue, the notion of a *field* is defined. In physics, a field is a tensor defined at each point in spacetime. This is a very general definition. To get a field-theoretic Lagrangian density, the generalized coordinates are replaced by fields and their derivatives. In particular all derivatives, not just the time derivatives. Let φ be a field and Ω be some arbitrary spacetime region with boundary $\partial\Omega$. This boundary is usually taken at spatial and temporal infinity where all fields vanish. The principle of least action states that

$$\delta S[\varphi] = \delta \int_\Omega d^4x\, \mathcal{L}\left(\phi, \partial_\mu \phi\right) = 0, \tag{2.1.2}$$

where d^4x is the spacetime differential element. Proceeding with the integral leads to

$$\begin{aligned} 0 &= \int_\Omega d^4x\, \left(\frac{\partial \mathcal{L}}{\partial \varphi}\delta\varphi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}\delta(\partial_\mu \varphi) \right) \\ &= \int_\Omega d^4x\, \left(\frac{\partial \mathcal{L}}{\partial \varphi}\delta\varphi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}\delta\varphi \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \right) \delta\varphi \right). \end{aligned} \tag{2.1.3}$$

Applying the divergence theorem to the middle term yields

$$\int_\Omega d^4x\, \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}\delta\varphi \right) = \int_{\partial\Omega} d^4x\, \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}\delta\varphi, \tag{2.1.4}$$

which vanishes as the variation at the boundary is zero. Hence, the principle of least action demands that

$$0 = \int_\Omega d^4x\, \left(\frac{\partial \mathcal{L}}{\partial \varphi}\delta\varphi - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \right) \delta\varphi \right), \tag{2.1.5}$$

which due to the fundamental lemma of calculus of variations, yields the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} = 0. \tag{2.1.6}$$

Given a Lagrangian density, which from now on will be referred to simply as the Lagrangian, the equations of motions are obtained by plugging said Lagrangian into the Euler-Lagrange equations. Obtaining the equations of motion is rather easy, the tricky part, however, is to choose the right Lagrangian for the problem at hand. There are a few symmetries a physically meaningful Lagrangian has to respect, e.g. Poincaré invariance. However, before discussing this subject further, some new theory has to be introduced.

2.2 Noether's theorem and conserved currents

Many modern physical theories rely heavily on symmetry and invariance, for which the most powerful tool is Noether's Theorem. Proven by Emmy Noether in 1915, her theorem states that [14]: Every differentiable symmetry of the action has a corresponding conserved current and conserved charge.

There is, however, a subtle criterion for the theorem to hold; the fields have to satisfy the Euler-Lagrange equations. At first, this seems obvious as physical fields are by definition solutions to the Euler-Lagrange equations, a direct consequence of the principle of least action. The crux of this argument is in the word *physical*. Particle interactions are mediated through other particles, which need not be physical particles. These are particles created and annihilated as a means of obtaining a final state from an initial state. Particles that do not satisfy the equations of motion are called virtual particles.

Returning to Noether's theorem, a symmetry of the action is a variation on the field φ which leaves the action invariant. A variation in the field is generally expressed as $\varphi \to \varphi'$. As a standard example, consider a Lagrangian with a $U(1)$ invariance. This means that the Lagrangian is invariant under the field variation $\varphi \to e^{i\alpha}\varphi$ for some real number α. The variation of the action with respect to α is

$$0 = \frac{\delta S[e^{i\alpha}\varphi]}{\delta\alpha} = \int_\Omega d^4x \, \frac{\delta\mathcal{L}[e^{i\alpha}\varphi, \partial_\mu e^{i\alpha}\varphi]}{\delta\alpha}. \tag{2.2.1}$$

In a similar fashion to the derivation of the Euler-Lagrange equations, one can show that the above equation leads to (the explicit dependence of the Lagrangian is suppressed for cleaner notation)

$$0 = \frac{\delta\mathcal{L}}{\delta\alpha} = \left(\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right)\frac{\delta\varphi}{\delta\alpha} + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\frac{\delta\varphi}{\delta\alpha}\right). \tag{2.2.2}$$

If φ is a physical field, i.e. it satisfies the equations of motion, then the terms in the first parenthesis vanish and the equation reduces to the continuity equation

$$\partial_\mu J^\mu = 0, \tag{2.2.3}$$

for the conserved current[6]

$$J^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\frac{\delta\varphi}{\delta\alpha}. \tag{2.2.4}$$

From a conserved current, a conserved charge follows. The physical meaning of the conserved charge is dependent on the particular symmetry imposed on the action. The conserved charge is given as the integral over the zeroth component of the conserved current taken over all space

$$Q = \int d^3\vec{x}\, J^0. \tag{2.2.5}$$

To prove that Q is conserved, simply differentiate the above equation with respect to time and use the continuity equation

$$\partial_t Q = \int d^3\vec{x}\, \partial_t J^0 = \int d^3\vec{x}\, \vec{\nabla}\cdot\vec{J} = 0. \tag{2.2.6}$$

The last integral is a total derivative, hence using the divergence theorem one can move the integral to the boundary at spatial infinity where all fields vanish.

The previous example illustrates how one would go about obtaining Noether currents for a Lagrangian that does not change under the symmetry. This is not always the case. The Lagrangian itself may change under the symmetry as long as the variation of the

[6]Conserved currents are also called Noether currents

action remains zero. For example, under infinitesimal spacetime translations, a Lorentz scalar like the Lagrangian transforms as

$$\mathcal{L}(x+\varepsilon) = \mathcal{L}(x) + \varepsilon^\nu \partial_\nu \mathcal{L}(x) + \mathcal{O}(\varepsilon^\nu)^2, \tag{2.2.7}$$

which implies that the variation of the Lagrangian is given as

$$\frac{\delta \mathcal{L}}{\delta \varepsilon^\nu} = \partial_\nu \mathcal{L}. \tag{2.2.8}$$

The above equation is a total derivative, meaning that the integral can be relocated to the boundary where the variation is assumed to vanish, leaving the action unchanged. Therefore, a more correct statement is that the symmetry is that of the action, not the Lagrangian. From the Euler-Lagrange equations (2.1.6) of the Lagrangian itself $\mathcal{L}[\varphi, \partial_\mu \varphi]$, one can show that spacetime translations lead to the conserved tensor current [15]

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial^\nu \varphi - g^{\mu\nu}\mathcal{L}, \quad \partial_\mu T^{\mu\nu} = 0. \tag{2.2.9}$$

This object is the famous energy-momentum tensor, where for each ν, there is a conserved charge. The conserved charges of the energy-momentum tensor is the energy and momentum of the system.

So far the examples used to illustrate Noether's theorem were variations of the field with respect to scalars. If the field has several components, say φ is some vector, then a natural question to ask is: what happens if the variation is a matrix acting on φ? To answer this, it is convenient to take a brief detour into the mathematical structure of *groups*.

2.3 Elements of group theory

A group is a set G, together with a binary operation $\circ$. The elements of the group satisfy the four group axioms:

1. Associativity: For each $a, b, c \in G$: $a \circ (b \circ c) = (a \circ b) \circ c$
2. Identity element: $\exists\, e \in G$ such that $\forall\, a \in G$: $a \circ e = e \circ a = a$
3. Inverse element: For each $a \in G\ \exists\, b \in G$ such that $a \circ b = b \circ a = e$.
4. Closure: $\forall\, a, c \in G$ the following holds: $a \circ c \in G$

If in addition, the group is commutative, then the group is said to be Abelian. In physics, the relevant groups are usually matrix groups, which are in general not commutative. Hence a theory based on matrix groups is non-Abelian[7]. There are several important matrix groups in physics, but for this thesis, there are two, in particular, that stand out: the orthogonal group (also called the rotation group) $O(N)$, and the special unitary group $SU(N)$. Both of these are matrix groups that preserve the norm of the vector acted upon. Generally, a $N \times N$ matrix has N^2 degrees of freedom, i.e. none of the matrix elements are constrained. However, to get something interesting, some kind of symmetry has to be imposed on the group. As an example of a symmetry, elements of the rotation group are defined to satisfy

[7]Non-Abelian theories are also called Yang-Mills theories in the physics literature.

$$A^T A = I \quad \forall A \in O(N), \tag{2.3.1}$$

where I is the identity matrix in N dimensions. This is an equation between two symmetric matrices which have $N(N+1)/2$ degrees of freedom. This is due to $S_{ij} = S_{ji}$ for any symmetric matrix S. Hence members of $O(N)$ can be written in terms of $N^2 - N(N+1)/2 = N(N-1)/2$ independent numbers. This number is called the dimension of the group. The determinant of orthogonal matrices is found by taking the determinant of equation (2.3.1), yielding

$$\det(A^T A) = \det(A^T)\det(A) = (\det(A))^2 = \det(I) = 1, \rightarrow \det(A) = \pm 1. \tag{2.3.2}$$

Orthogonal matrices with positive determinant constitute a group of their own, namely the special orthogonal group, defined as

$$SO(N) = \{A \in O(N) : \det(A) = +1\}. \tag{2.3.3}$$

The dimensions of $SO(N)$ and $O(N)$ are the same, as the requirement of positive determinant is not an independent equation, but rather a choice of sign. Orthogonal matrices with negative determinants are combinations of rotations and reflections.

The special unitary group $SU(N)$ can be viewed as the complex analogue of the special orthogonal group $SO(N)$. Quantum mechanics says that the norm of a state, represented by a vector in Hilbert space, is related to the probability of observing said state. Conservation of probability, a general mathematical statement, is related to the conservation of the norm of quantum states. Furthermore, as all states live in Hilbert space, they can be expressed in terms of one another with linear transformations. In the language of Dirac's bra-ket notation, quantum states evolve as $|\psi\rangle \rightarrow U|\psi\rangle$, where $|\psi\rangle$ is the wave function and U is a matrix. Equating the norm leads to

$$\langle\psi|\psi\rangle \rightarrow \langle\psi|U^\dagger U|\psi\rangle = 1, \rightarrow U^\dagger U = I, \tag{2.3.4}$$

which expresses the very fundamental fact: Conservation of probability implies unitarity. A general $N \times N$ complex valued matrix has N^2 complex degrees of freedom which are $2N^2$ real degrees of freedom. One can show in a similar fashion to that of orthogonal matrices, that the dimension of $SU(N)$ is $N^2 - 1$. To summarize, the dimensions of the relevant groups are

$$\dim(SO(N)) = \frac{N(N-1)}{2}. \qquad \dim(SU(N)) = N^2 - 1. \tag{2.3.5}$$

Invertible matrix groups, which $SO(N)$ and $SU(N)$ happens to be, are called Lie groups. Lie groups can be viewed as differentiable manifolds on which each element of the group can be written in terms of the group generators [16]. The number of generators is equal to the dimension of the group. The generators are a subset of the group and act, loosely speaking, as a direction on the manifold, serving a similar role to the usual basis vectors from linear algebra. Let T^α be the generators of the group G, then any member $g \in G$ connected to the origin can be written as

$$g = exp\left(i\theta_\alpha T^\alpha\right), \quad \theta_\alpha \in \mathbb{R}. \tag{2.3.6}$$

The generators form a Lie algebra, defined through the Lie bracket. For matrix Lie groups, the Lie bracket is a commutator

$$[T^\alpha, T^\beta] = if^{\alpha\beta\gamma}T^\gamma, \tag{2.3.7}$$

where $f^{\alpha\beta\gamma}$ are the structure constants. The form of the structure constants varies depending on the group. As an example, $SU(N)$ and its corresponding algebra $\mathfrak{su}(N)$ is totally antisymmetric; $f^{\alpha\beta\gamma} \propto \varepsilon^{\alpha\beta\gamma}$. In fact, the Standard Model is based upon the Lagrangian being invariant under the combined *gauge* group

$$G_{SM} = SU(3) \otimes SU(2) \otimes U(1). \tag{2.3.8}$$

Another important group is the Lorentz group. To understand the structure of the Lorentz group, consider a general Lorentz transformation of a vector

$$x = \Lambda x. \tag{2.3.9}$$

As discussed in Section 1.2, the contraction of a vector with its covariant counterpart is invariant. The relativistic inner product under the above Lorentz transformation is

$$x^T g x \to x^T \Lambda^T g \Lambda x, \tag{2.3.10}$$

where $g = (g_{\mu\nu})$ is the matrix representation of the spacetime metric. By definition, the above transformation must leave the inner product invariant, meaning that Lorentz transformations are required to satisfy

$$\Lambda^T g \Lambda = g. \tag{2.3.11}$$

The above equation shows the Lorentz matrices are orthogonal matrices with respect to the Minkowski metric, as opposed to the Euclidean metric (which is the identity matrix). The Lorentz group is denoted as $O(1,3)$ (or $O(3,1)$ in the (-,+,+,+) convention) and is defined as

$$O(1,3) = \{\Lambda \in O(1,3) : \Lambda^T g \Lambda = g\}. \tag{2.3.12}$$

The discussion on the degrees of freedom for orthogonal matrices applies to the Lorentz group as well. The reasoning is simple; the sign of the diagonal entries of the metric does not add nor remove any constraining equations. Hence, the dimension of the Lorentz group is $4(4-1)/2 = 6$. The full Lorentz group $O(1,3)$ includes matrices, when applied to four-vectors, yield effects such as e.g. time reversion, T. Discrete symmetries such as T, and space inversion, P, are particular members of the Lorentz group which cannot be connected to the origin. The subgroup of Lorentz transformations which exclude both time-reversal and space inversion is called the proper, orthochronous Lorentz group and is defined as

$$SO^+(1,3) \equiv SO(1,3) = \{\Lambda \in O(1,3) : \det(\Lambda) = 1\}. \tag{2.3.13}$$

The last group which will be mentioned in this section is the Poincaré group; the combined group of spacetime translations, rotations and boosts. The latter two make up the Lorentz group. Spacetime translations are parametrized by four numbers, meaning that the Poincaré group is a ten-dimensional non-Abelian group. Invariance under the Poincaré group is a fundamental requirement for any physical theory including relativity.

Before applying this to the Standard Model, there are two more things that needs to be addressed:

1. The field content of the Standard Model.

2. The transition from global symmetries to gauged ones.

2.4 The field content of the Standard Model

Recall that a field was defined as some tensor that depends on spacetime. The field content of the universe can be broadly categorized into two types depending on the intrinsic spin of the field.

2.4.1 Bosonic fields

Fields that have an integer value of the spin quantum number, referred to as just spin, are called *bosons*. Bosons are particles following Bose-Einstein statistics and are the mediators of the three fundamental forces described by the Standard Model; the electromagnetic, weak and strong interaction. Bosonic fields with spin 0 are called scalar fields, which indicated by the name, means they transform as scalars under Lorentz transformations. Moving incrementally upwards, the next type of bosonic fields are the ones with spin 1, called vector fields as they transform as vectors under Lorentz transformations. Depending on the mass of the vector fields, they have either two or three degrees of freedom, the massless ones having the least. An example of a vector field is the familiar photon field, which from electromagnetic theory is known to have two degrees of freedom. These are the oscillating electric and magnetic fields perpendicular to the direction of motion. As far as the SM goes, scalar and vectors fields are all one needs to describe the bosonic field content of the observed fundamental particles.

Before continuing to the fermion sector, an honourable mention is given to the graviton; the quantum of gravity. If it exists, the graviton is a massless spin 2 tensor field. Gravitons require spin 2 as they would be produced from the stress-energy tensor $T_{\mu\nu}$, which generates gravity through the famous Einstein field equations (EFE)

$$R_{\mu\nu} - \frac{R}{2} g_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{2.4.1}$$

The EFE are a set of six coupled, partial, non-linear differential equations for the components of the general (not Minkowskian) metric $g_{\mu\nu}$. Technicalities aside, the EFE relates energy to curvature. The solutions to the EFE (which are infamously hard to come by) are tensor fields, of which the gravitons must also be. At the time of this thesis, there is no experimentally verified theory which both accounts for gravity and quantum mechanics.

2.4.2 Fermionic fields

Fermionic fields are classified as fields with half-integer spin. All the fundamental fermionic fields have spin 1/2 and follow Fermi-Dirac statistics. Hence, this discussion will only focus on spin half fields, called *spinors*. As with every other field, spinors are required to be representations of the Lorentz group. The algebra of the Lorentz group can be decomposed as [16]

$$\mathfrak{so}(1,3) = \mathfrak{su}(2) \oplus \mathfrak{su}(2). \tag{2.4.2}$$

The algebra $\mathfrak{su}(2)$ is the antisymmetric algebra of the Pauli matrices, which acts on two-dimensional, complex vector space. The elements of these vector spaces are called Weyl spinors and they are further distinguished based on their particular representation. A $(\frac{1}{2}, 0)$ Weyl spinor is called left-chiral ψ_L^W, while a $(0, \frac{1}{2})$ Weyl spinor is called right-chiral, ψ_R^W. Under infinitesimal Lorentz transformations, the chiral Weyl spinors transform as [15]

$$\delta\psi_L^W = \frac{1}{2}\left(i\theta^j - \beta^j\right)\sigma_j\psi_L^W \tag{2.4.3}$$

$$\delta\psi_R^W = \frac{1}{2}\left(i\theta^j + \beta^j\right)\sigma_j\psi_R^W, \tag{2.4.4}$$

where θ^j are real rotation angles, β^j are boost angles and σ_j are the Pauli matrices. Oddly enough, nature cares a great deal about the chirality of spinors, going so far as to neglect right chiral spinors in weak interactions altogether. From two Weyl spinors, one can construct a Dirac spinor as the direct sum

$$\psi = \psi_L^W \oplus \psi_R^W = \begin{pmatrix}\psi_L^W \\ \psi_R^W\end{pmatrix}, \tag{2.4.5}$$

which is a four-component object. The Dirac spinors, spinors for short, are the fundamental building blocks of fermionic matter. All free fermion fields are described in terms of spinors through the Dirac equation

$$\left(i\gamma^\mu\partial_\mu - mI\right)\psi = 0. \tag{2.4.6}$$

For each component ψ_α, the Dirac equation acts as a relativistic wave equation. The γ^μ in the Dirac equation are called the gamma matrices and they satisfy the Clifford algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \tag{2.4.7}$$

where the curly brackets refer to the anti-commutator

$$\{A, B\} \equiv AB + BA. \tag{2.4.8}$$

In addition to the four gamma matrices, it is convenient to define a fifth gamma matrix as the matrix product of all the other gamma matrices

$$\gamma^5 \equiv \gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{2.4.9}$$

2.5 Gauge symmetries

As stated in previous sections, the Standard Model is based on the gauge group G defined in (2.3.8). The *gauge* in gauge group means that the symmetry is local. A local symmetry can be thought of as a variation in a field, confined to some region of spacetime. In mathematical terms, a field transform as

$$\varphi \to X(x)\varphi, \tag{2.5.1}$$

where $X(x)$ is a differentiable object of spacetime, i.e. a field.

2.5.1 The electromagnetic interaction U(1)

Consider the free Dirac Lagrangian, which is trivially invariant under global $U(1)$ transformations, given as

$$\mathcal{L} = \overline{\psi}\left(i\not{\partial} - m\right)\psi. \tag{2.5.2}$$

In the above equation some new notation is introduced, the "barred" spinor $\overline{\psi}$ is defined as

$$\overline{\psi} \equiv \psi^\dagger \gamma^0, \tag{2.5.3}$$

while the "slashed" partial derivative is defined as the contraction

$$\not{\partial} \equiv \gamma^\mu \partial_\mu. \tag{2.5.4}$$

If instead, a local $U(1)$ symmetry is imposed, the spinors transform as

$$\psi \to e^{i\alpha(x)}\psi, \tag{2.5.5}$$

which in turn transforms the free Dirac Lagrangian to

$$\mathcal{L} \to \overline{\psi}e^{-i\alpha(x)}\left(i\not{\partial} - m\right)e^{i\alpha(x)}\psi = \mathcal{L} - \overline{\psi}\gamma^\mu\psi\partial_\mu\alpha(x). \tag{2.5.6}$$

The last term in the above equation is not generally zero. Thus, the free Dirac Lagrangian is not invariant under a gauged $U(1)$ symmetry. The culprit is the derivative operator, as it also acts on the field variation to create an extra term. To compensate for this extra term, a new form of the derivative is introduced; the *covariant* derivative. From general relativity, the covariant derivative is defined as a way to account for movement along trajectories on a curved space with a defined metric. The gravitational covariant derivative is defined in terms of the metric through the Christoffel symbols which are combinations of metric derivatives.

Although $U(1)$ constitutes a Lie group, which as discussed in Section 2.3 can be considered a differentiable manifold, it is not defined with respect to a metric. Regardless of approach, the gauge covariant derivative for $U(1)$ is given as

$$D_\mu \equiv \partial_\mu - igA_\mu, \tag{2.5.7}$$

for some real number g called a coupling constant, and a vector field A_μ. This vector field will soon be identified as the photon field. A standard result from classical electrodynamics is that the equations of motion (Maxwell's equations) are invariant under the following gauge transformation of the four potential

$$A'_\mu \to A'_\mu + \partial_\mu\alpha'(x), \tag{2.5.8}$$

where $\alpha'(x)$ is a smooth function with spacetime dependence. If one makes the following identification:

$$A' = A \tag{2.5.9}$$

$$\alpha' = \frac{\alpha}{g}, \tag{2.5.10}$$

then the problematic derivative term, which sparked this whole conversation, transforms under the combined gauge transformation of ψ and A_μ as

$$D_\mu\psi = (\partial_\mu - igA_\mu)\,\psi \to (\partial_\mu - igA_\mu - i\alpha(x))\,e^{i\alpha(x)}\psi = e^{i\alpha(x)}D_\mu\psi. \tag{2.5.11}$$

This is remarkable, by demanding gauge invariance of the free Dirac Lagrangian, the spinors acquired an interaction with the photon field! Bosons introduced by demanding gauge invariance are often called gauge bosons. Adding a kinetic term for the photon and relabeling the coupling constant g to the familiar electric charge, the Lagrangian describing quantum electrodynamics (QED) is given as

$$\mathcal{L}_{QED} = \overline{\psi}\left(i\not{D} - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.5.12}$$

The QED Lagrangian can be split into two parts: free and interacting. The free part is just the sum of the free Dirac Lagrangian and the photon kinetic term, while the interacting Lagrangian governs dynamics and is given as

$$\mathcal{L}^{int}_{QED} = e\overline{\psi}A_\mu\gamma^\mu\psi. \tag{2.5.13}$$

Invoking Noether's theorem on the $U(1)$ symmetry yields the conserved current

$$J^\mu_{QED} = \frac{\partial\mathcal{L}_{QED}}{\partial(\partial_\mu\psi)}\frac{\delta\psi}{\delta\alpha} = -e\overline{\psi}\gamma^\mu\psi, \tag{2.5.14}$$

which has conserved charge

$$Q = \int d^3\vec{x}\, J^0_{QED} = -e\int d^3\vec{x}\,\psi^\dagger\psi. \tag{2.5.15}$$

Given that the current description of electromagnetism is correct, then electric charge must be conserved. This appears to be the case as there are currently no observations indicating that the conservation of electric charge is violated. The strongest bound on conservation of electric charge is from non-observations of the process

$$e \to \nu_e + \gamma, \tag{2.5.16}$$

which has a lifetime of $\tau > 6.6 \times 10^{28}$ years [17].

2.5.2 The strong interaction SU(3)

After discussing the electromagnetic interaction and its inception from demanding gauge invariance under $U(1)$, the next natural step is to consider the strong interaction. The strong interaction governs how nuclei are held together, amongst other things. The particles charged under the strong force are the quarks; fermionic fields which are additionally charged under electromagnetism. Experiments can deduce that quarks come in three colors, however colour is not a direct observable. Hence, when constructing a theory for the strong interaction, this has to be respected. As quarks are fermionic fields with three degrees of freedom, i.e. the three colour charges, they are described by the three-component spinor

$$\psi = \begin{pmatrix}\psi_r & \psi_g & \psi_b\end{pmatrix}^T. \tag{2.5.17}$$

The subscript refers to the colour charges red, green and blue, respectively. The free quark Lagrangian has the same form as the free QED Lagrangian, the only difference being the dimension of the spinors. As ψ is a three-component object, transformations acting on it are represented by 3×3 matrices. The matrix group describing strong interactions $SU(3)$. The special unitary group $SU(N)$ depends on N^2-1 generators, which for $N=3$ means that $SU(3)$ is generated by eight matrices. These generators are called the Gell-Mann matrices. They are denoted by λ^j and they satisfy the commutation relation (2.3.7). An explicit form for these can be found in any textbook on QFT. The theory follows largely from QED, so in the spirit of gauge invariance, the logical next step is to gauge $SU(3)$. In doing so, coloured spinors transform as

$$\psi \to \psi e^{iT^j\alpha'_j(x)} \equiv \psi e^{ig_s\lambda^j\alpha_j(x)/2}, \tag{2.5.18}$$

where the eight real functions $\alpha'_j(x)$ are rescaled in terms of the strong coupling constant g_s. The one-half in the second exponent is a conventional construct relating the generators T_j to the Gell-Mann matrices λ_j. Under the gauge transformation (2.5.18), the free Quark Lagrangian written in terms of the partial derivative ∂_μ will not stay invariant. The reasoning is the same as for QED; the derivative will hit the exponent of the transformed field resulting in new terms which are generally not zero. In a similar fashion to QED, the $SU(3)$ covariant derivative is introduced as

$$D_\mu \equiv \partial_\mu - \frac{i}{2}g_s\lambda^j A_{j\mu}, \tag{2.5.19}$$

for eight real vector fields $A_{j\mu}$, labeled by $j=1,2,\ldots,8$. The covariant derivative couples spinors to vector fields through the interacting Lagrangian

$$\mathcal{L}^{int}_{QCD} = \frac{1}{2}g_s\overline{\psi}\gamma^\mu A_{j\mu}\lambda^j\psi, \tag{2.5.20}$$

which is a sum of $4\times8=32$ terms. The fact that $SU(3)$ is non-Abelian, as opposed to QED which is Abelian, makes creating a kinetic term for the vectors fields a bit more tedious. The vectors fields of QCD are called gluon fields. The usual field strength tensor contraction $F_{\mu\nu i}F^{\mu\nu i}$ is not gauge invariant when applied to the gluon fields. To compensate, an additional term must be added to the field strength tensor, now defined as [18]

$$G_{i\mu\nu} \equiv F_{i\mu\nu} + g_s f_{ijk}A_{j\mu}A_{k\mu}, \tag{2.5.21}$$

where f_{ijk} are the structure constants of $\mathfrak{su}(3)$ defined in (2.3.7). The non-Abelian nature of $SU(3)$ also affects how the gluon fields transform under infinitesimal gauge transformations

$$A_{i\mu} \to A_{i\mu} - \partial_\mu\alpha_i - g_s f_{ijk}\alpha_j A_{k\mu}. \tag{2.5.22}$$

By adding the quark and gluon terms, the QCD Lagrangian is given by

$$\mathcal{L}_{QCD} = \overline{\psi}\left(i\not{D} - m\right)\psi - \frac{1}{4}G_{i\mu\nu}G^{i\mu\nu}. \tag{2.5.23}$$

From Noether's theorem, the $SU(3)$ symmetry imposed on the QCD Lagrangian leads to eight conserved currents. Up to a constant multiple, the eight conserved currents in the quark sector are given as

$$J^{i\mu}_{QCD} = \overline{\psi}\gamma^{\mu}\lambda^{i}\psi, \tag{2.5.24}$$

which in turn, yields the conserved charges

$$Q^{i}_{QCD} = \int d^{3}\vec{x}\, \psi^{\dagger}\lambda^{i}\psi. \tag{2.5.25}$$

Quark colour is a conserved charge, but the Gell-Mann matrices are not all diagonal. This has important consequences for the nature of the quark sector. In non-diagonal interactions, quark colour is exchanged between quarks through interactions with gluons, which inevitably implies that the gluons themselves are charged under QCD; gluons have colour charge. The fact that gluons have colour means that they interact in-between themselves, leading to cubic and quartic gluon-gluon interactions. Any theory based on $SU(N)$ with $N \geq 2$ leads to such boson-boson interactions. It is at this point one starts to appreciate the simplicity of QED. There are no boson-boson interactions in QED as the photon is electrically neutral and hence does not couple to itself.

2.5.3 The weak interaction $SU(2)$

A long-standing principle in physics is that of parity: the universe should look the same if all spatial directions are reversed, i.e. physics in the mirror universe should be identical to the physics in the universe where this thesis is written. Until the 1950's parity was taken more or less like an axiom of physics, how could such a seemingly obvious statement be wrong? New things are discovered when inconsistencies with existing models are found, when fundamental ideas and assumptions need to be revisited and challenged. The theory of weak interactions is one such example. In 1956, Wu and collaborators found that parity is indeed violated in nuclear decay by studying the interaction [19]

$$^{60}_{27}\mathrm{Co} \rightarrow {}^{60}_{28}\mathrm{Ni} + e^{-} + \overline{\nu_e} + 2\gamma. \tag{2.5.26}$$

The experiment was set up such that the end state photons should be emitted in an isotropic way if parity was conserved. This was not the case, the photons had a preferred direction in space, a mismatch between left and right. On that day parity fell, which shook the world of physics to the core. Related to the study of parity two important quantities are introduced, *helicity* and *chirality*. Helicity is defined as the projection of the spin operator along the direction of motion given by

$$\frac{\vec{S}\cdot\vec{p}}{|s\vec{p}|}\psi = \pm\psi, \tag{2.5.27}$$

where the spin operator is given by $\vec{S} = \vec{\sigma}/2$ with eigenvalues $s = \pm 1/2$. In macroscopic terms, one can visualize helicity as the rotation of an object relative to the direction of motion. Chirality on the other hand, as briefly touched upon in the section on fermionic fields, is defined as a particular representation of the Lorentz group. The chiral projection operators are defined as

$$P^{L/R} \equiv \frac{1}{2}(1 \mp \gamma_5), \tag{2.5.28}$$

which act on Dirac spinors to project out the chiral Weyl spinors

$$P^L\psi \equiv \psi^L = \begin{pmatrix} \psi_L^W \\ 0 \end{pmatrix}, \qquad P^R\psi \equiv \psi^R = \begin{pmatrix} 0 \\ \psi_R^W \end{pmatrix}. \tag{2.5.29}$$

It can be shown that in the massless case, the helicity operator and the chiral projection operator coincide. For massive particles, they are equal up to a correction of order m/E. In the case of neutrinos, which is the main topic of the thesis, the approximation will hold to an excellent degree for neutrinos energetic enough to be of experimental interest. The theory of weak interactions is introduced as a *chiral* theory, in which only left-handed fields can interact. In gauge theory language, this implies that there should exist different gauge transformations for the different chiral fields. In standard terminology, the left-handed fermions are grouped into two-component fields, called doublets, with the respective fermion from the same generation. The lepton doublet is introduced as

$$\Psi_l^L = \begin{pmatrix} \psi_{\nu_l}^L \\ \psi_l^L \end{pmatrix}, \tag{2.5.30}$$

where l labels the generations, $l = \{e, \mu, \tau\}$. The quark doublets are introduced in a similar fashion. The left-handed doublets are the fundamental building blocks of the weak interaction. The observed parity violation is imposed by the following gauge transformations

$$\Psi_l^L \to e^{ig\sigma^j\alpha_j(x)/2}\Psi_l^L \tag{2.5.31}$$
$$\psi_l^R \to \psi_l^R \tag{2.5.32}$$
$$\psi_{\nu_l}^R \to \psi_{\nu_l}^R. \tag{2.5.33}$$

The lepton doublet gauge transformation will inevitably introduce three gauge bosons, identified as $W^\pm$ and Z. All gauge bosons are inherently massless when created by gauge invariance, however, this is not what experiments show. Furthermore, when imposing asymmetric gauge transformations for the different chiral fields, the Dirac mass term is no longer gauge invariant. To showcase why this is, recall that the mass term appearing in the Dirac Lagrangian is the bilinear combination

$$\mathcal{L} = m\overline{\psi}\psi = m\left(\overline{\psi}^L + \overline{\psi}^R\right)\left(\psi^L + \psi^R\right) = m\left(\overline{\psi}^R\psi^L + \overline{\psi}^L\psi^R\right), \tag{2.5.34}$$

where the L-L and R-R terms vanish due to

$$\overline{\psi}^L\psi^L = P^L\overline{\psi}P^L\psi = \overline{\psi}P^RP^L\psi = 0, \tag{2.5.35}$$

and similar for R-R. Under the chiral $SU(2)$ gauge transformations, the $R-L$ structure of the Dirac mass terms will not be invariant due to the different transformations of the chiral fields. This is opposed to QED and QCD, which do not differentiate between chiral fields and hence these theories are non-chiral. These issues need to be circumvented to obtain a gauge-invariant Lagrangian accounting for massive particles. The way forth is two-fold, first electroweak (EW) theory is introduced and then finally the EW symmetry is broken to yield particles masses through the BEH mechanism.

2.5.4 Electroweak theory $SU(2)_L \otimes U(1)_Y$

In the late '60s following the work of Glashow, Salam and Weinberg, the theory of electroweak interactions was created. The theory is based on the combined gauge symmetry $SU(2)_L \otimes U(1)_Y$ as a high-energy unification of the weak and electromagnetic interaction into a combined symmetry group. Electroweak unification occurred in the hot, early universe at temperatures $T \sim 200\text{GeV}$. The combined EW gauge transformations acting on chiral leptons are given by

$$\Psi_l^L \to \exp(ig\sigma^j\alpha_j(x)/2 + ig'Y\beta(x)/2)\Psi_l^L, \tag{2.5.36}$$
$$\psi_{l/\nu_l}^R \to \exp(ig'Y\beta(x)/2)\psi_{l/\nu_l}^R, \tag{2.5.37}$$

where ψ_{l/ν_l} is shorthand notation for the right chiral leptons l^R and ν_l^R. Here g is the coupling to $SU(2)_L$, g' the coupling to $U(1)_Y$ and Y is the hypercharge; a real number. Furthermore, in standard gauge theory language, $\alpha_j(x)$ and $\beta(x)$ are arbitrary differentiable functions of spacetime. The Lagrangian is still that of Dirac, albeit written such to reflect the chiral structure of the weak interaction. The kinetic terms are

$$\mathcal{L} = \overline{\Psi}_l^L i\partial\!\!\!/ \Psi_l^L + \overline{\psi}_{l/\nu_l}^R i\partial\!\!\!/ \psi_{l/\nu_l}^R. \tag{2.5.38}$$

Imposing the EW gauge transformations leads to a variation $\delta\mathcal{L}$ in the Lagrangian given as

$$\delta\mathcal{L} = -\frac{1}{2}\overline{\Psi}_l^L\gamma^\mu\left(g\sigma^j\partial_\mu\alpha_j(x) + g'Y\partial_\mu\beta(x)\right)\Psi_l^L - \frac{1}{2}\overline{\psi}_{l/\nu_l}^R\gamma^\mu g'Y\partial_\mu\beta(x)\psi_{l/\nu_l}^R. \tag{2.5.39}$$

As usual in gauge theories, which the reader should have some familiarity with by now, the next step is to introduce the covariant derivatives, defined by how they act on the different fields by

$$D_\mu\Psi_l^L = \left(\partial_\mu - ig\sigma^j W_{j\mu}/2 - ig'YB_\mu/2\right)\Psi_l^L, \tag{2.5.40}$$
$$D_\mu\psi_{l/\nu_l}^R = \left(\partial_\mu - ig'YB_\mu/2\right)\psi_{l/\nu_l}^R, \tag{2.5.41}$$

for four real vector fields $W_{j\mu}$ and B_μ. The conserved weak isospin currents for $SU(2)_L$ are found from Noether's theorem, which up to a constant are given as

$$J_{\text{WIS}}^{j\mu} = \frac{1}{2}\overline{\Psi}\gamma^\mu\sigma^j\Psi, \tag{2.5.42}$$

with corresponding conserved charges, called weak isospin, given by

$$I_W^j = \int d^3\vec{x}\, J_{\text{WIS}}^{j0} = \frac{1}{2}\int d^3\vec{x}\, \Psi_l^{L\dagger}\sigma^j\Psi_l^L. \tag{2.5.43}$$

These currents are however not the ones that belong to the weak bosons, rather the physical currents are defined as linear combinations of the weak isospin currents given as

$$J_{CC}^\mu = 2\left(J_{\text{WIS}}^{1\mu} - iJ_{\text{WIS}}^{2\mu}\right) = \overline{\psi}_l^L\gamma^\mu(1-\gamma_5)\psi_{\nu_l}^L, \tag{2.5.44}$$
$$J_{CC}^{\mu\dagger} = 2\left(J_{\text{WIS}}^{1\mu} + iJ_{\text{WIS}}^{2\mu}\right) = \overline{\psi}_{\nu_l}^L\gamma^\mu(1-\gamma_5)\psi_l^L, \tag{2.5.45}$$
$$J_{NC}^\mu = \frac{1}{2}\overline{\Psi}_l^L\gamma^\mu\sigma^3\Psi_l^L = -\frac{1}{2}\left(\overline{\psi}_{\nu_l}^L\gamma^\mu\psi_{\nu_l}^L - \overline{\psi}_l^L\gamma^\mu\psi_l^L\right). \tag{2.5.46}$$

The first two currents couple charged leptons to their respective electrically neutral neutrinos, hence the currents themselves must be electrically charged. Thus, these currents are called *charged-currents* and the exchange boson is $W^{\pm}$. The third current, however, cannot have any electric charge as it either couples two neutrinos or two charged leptons, promoting the name *neutral-current*. The exchange boson here is Z. Observe that the neutral current interaction includes the electromagnetic interaction up to a constant multiple of the electric charge.

In the electroweak gauge, the hypercharge is introduced as the charge of $U(1)_Y$ as is given by

$$Y \equiv Q/e - I_W^3. \tag{2.5.47}$$

Hypercharge is conserved as both electric charge and the third component of weak isospin are conserved. From this definition, the hypercharges of the various lepton fields are given as

$$Y\Psi_l^L = -\frac{1}{2}\Psi_l^L, \quad Y\psi_l^R = -1\psi_l^R, \quad Y\psi_{\nu_l}^R = 0. \tag{2.5.48}$$

A striking consequence of how hypercharge is defined is that the right chiral neutrino fields are uncharged under every gauge group discussed so far. For this exact reason, the right-handed neutrinos are not included in the original Standard Model, as there was no a priori reason to include them. A direct consequence of the exclusion of right-handed neutrinos is that neutrinos in the SM are massless. This is due to the lack of a right-handed field to construct the right-left Dirac mass. Experiments reveal that neutrinos do have mass, which is the first direct observational evidence for physics beyond the SM. A much more thorough review of neutrino physics can be found in Section 3 and will play a key role in this thesis.

As a last point of discussion, it is worth mentioning that the EW gauge fields are given as linear combinations of the generators of $SU(2)_L$, much akin to how the physical currents were defined. The familiar $W^{\pm}$ are given in terms of the first two $SU(2)_L$ gauge fields as the linear combinations

$$W_\mu^{\pm} = \frac{1}{\sqrt{2}}\left(W_{1\mu} \pm iW_{2\mu}\right), \tag{2.5.49}$$

while the neutral gauge fields can be viewed as a rotation in the two dimensional $W_{3\mu} - B_\mu$ space through the Weinberg angle θ_W

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & \sin\theta_w \\ -\sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} B_\mu \\ W_{3\mu} \end{pmatrix}. \tag{2.5.50}$$

A more comprehensive introduction to electroweak theory can be found in most textbooks on QFT.

2.5.5 Electroweak symmetry breaking $SU(2)_L \otimes U(1)_Y \to U(1)_{em}$

The final ingredient which is needed to create the Standard Model is the complex Higgs doublet H. The Higgs doublet has hypercharge $Y = 1/2$, meaning that the EW covariant derivative acting on H is given by

$$D_\mu H = \partial_\mu H - igW_{j\mu}\sigma^j H - \frac{1}{2}ig'B_\mu H, \tag{2.5.51}$$

where $W_{j\mu}$ and B_μ are the $SU(2)_L \otimes U(1)_Y$ gauge bosons. The most general *renormalizable* Lagrangian one can construct using only the Higgs doublet and its covariant derivative is

$$\mathcal{L}_H = (D_\mu H)^\dagger (D^\mu H) + m^2 (H^\dagger H) - \lambda (H^\dagger H)^2, \tag{2.5.52}$$

where m^2 and λ are real numbers. The "wrong" sign for the mass term allows for some interesting physics, namely a non-zero vacuum state.

A non-zero vacuum state is rather odd, it means that the universe is in a state of lower energy for a non-zero value of the field, as opposed to the standard ground state of no field excitation. The ground state is found by taking the derivative of the Higgs potential and setting that zero

$$\frac{\partial V(H)}{\partial (H^\dagger H)} = m^2 - 2\lambda H^\dagger H = 0. \tag{2.5.53}$$

The solution is straightforward to compute and is given as

$$H^\dagger H = |H_1|^2 + |H_2|^2 = \frac{m}{\sqrt{2\lambda}}, \tag{2.5.54}$$

which in unitary gauge, can be chosen such that H_1 is zero and H_2 is real. In choosing unitary gauge, the vacuum is spontaneously broken from a symmetric state, i.e. when the minimum value of the potential is for $H = 0$, into a non-zero ground state. The chosen ground state is only one of a continuum of ground states, which can be rotated into one another by an arbitrary phase. This arbitrariness allows for this phase to be set to zero as it is merely a matter of convenience, allowing for easier computations. Expanding about the ground state with a real scalar field h[8] leads to the spontaneously broken theory in which the Higgs doublet is written as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}, \quad v \equiv \frac{m}{\sqrt{\lambda}}, \tag{2.5.55}$$

where v is the vacuum expectation value (vev)

$$v = 246\text{GeV} \tag{2.5.56}$$

True to its name, the vev is the expectation value of the Higgs doublet in the vacuum state, leading to the alternative notation $v \equiv \langle H \rangle$. Only scalar fields may acquire a vev as fields transforming non-trivially under Lorentz transformations cannot be chosen to have a constant vev everywhere. In choosing unitary gauge, the degrees of freedom (d.o.f) of the Higgs doublet reduced from four to one, but these cannot simply vanish. To see where these seemingly lost degrees of freedom went, one can return to the Higgs Lagrangian and write out the term with the covariant derivatives. Written out, the relevant terms are

$$(D_\mu H)^\dagger (D^\mu H) = \frac{g^2 v^2}{8} \left(W_{1\mu} W_1^\mu + W_{2\mu} W_2^\mu + (\frac{g'}{g} B_\mu - W_{3\mu})(\frac{g'}{g} B^\mu - W_3^\mu) \right), \tag{2.5.57}$$

[8]Little h is what physicists usually mean when talking about the Higgs field. Of course, as with everything else, this may be subject to change depending on the text at hand.

which are three mass terms. From equation (2.5.50), the last term in the above equation is the mass term for the Z boson. Writing out the masses for W_i^μ in terms of the weak gauge fields $W^\pm$, the resulting Lagrangian obtains new mass terms, one for each of the weak gauge fields, but importantly, none for the photon. By turning massive, vector fields obtain a new d.o.f; longitudinal polarization. To summarize, when the Higgs doublet is spontaneously broken, its degrees of freedom transfer to the gauge bosons, which in turn makes them massive. This is the BEH mechanism, which yielded their creators the 2013 Nobel Prize following the discovery of the Higgs boson at CERN. Note that the BEH mechanism does not have to include $SU(2)_L \otimes U(1)_Y$, any Yang-Mills theory can be broken in a similar fashion. With that said, the discussion on gauge symmetries is finally complete.

2.6 Fermion masses

At this point, electroweak symmetry breaking yields masses for the weak gauge bosons through the BEH mechanism, what about the fermions? In the unbroken phase, one can construct gauge-invariant trilinear terms with chiral spinors and the Higgs doublet as e.g.

$$\mathcal{L}_Y = -Y_l^{ij}\overline{\Psi}_{li}^L H \psi_{lj}^R + h.c, \tag{2.6.1}$$

where Y_l^{ij} is a matrix of generally complex numbers. When H acquires a vev, the Lagrangian will produce Dirac mass terms for the charged leptons. By convention, the broken Higgs doublet is non-zero in its lower component so the above expression does not yield neutrino masses. This is usually not an issue as right chiral neutrinos have not been observed, but as neutrinos do have mass, a general theory needs to account for this. Although interesting in their own right, quarks are not important for this thesis and henceforth only leptons will be considered. It is nevertheless worth mentioning that the content of this section is similarly applicable to the quark sector.

Essentially, to get neutrino masses from the Lagrangian in (2.6.1) a modified Higgs term must be used to move the vev into the upper component, while the overall trilinear term is left gauge invariant. The term

$$\widetilde{H} \equiv i\sigma^2 H^*, \tag{2.6.2}$$

which can be shown to transform under $SU(2)_L$ such that the overall trilinear is invariant, will suffice. In unitary gauge, after symmetry breaking, it is straightforward to show that

$$\widetilde{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} v+h \\ 0 \end{pmatrix}, \tag{2.6.3}$$

which is suitable to construct Dirac mass terms for neutrinos. The Lagrangian encapsulating lepton masses is given by

$$\mathcal{L}_Y = -Y_l^{ij}\overline{\Psi}_{li}^L H \psi_{lj}^R - Y_\nu^{ij}\overline{\Psi}_{li}^L \widetilde{H} \psi_{\nu_l j}^R + h.c, \tag{2.6.4}$$

where Y_ν^{ij} is another generally complex matrix. These matrices are called Yukawa-matrices and their matrix elements are called Yukawa couplings. Historically, Yukawa interactions were introduced to study the strong force via the exchange of mesons, from which Hideki Yukawa was awarded the 1949 Nobel Prize after their discovery. The interactions take form as trilinear combinations of two spinor fields and one (pseudo)-scalar which, are the same form as the terms in the above Lagrangian.

The Standard Model was created when neutrinos were assumed massless. If neutrino masses are SM phenomena or not is subject to debate. Some will claim that the SM is the theory created when neutrinos were assumed massless. The addition of neutrino masses requires the existence of right-chiral neutrinos, which are gauge-singlets. As neutrino masses exist and right-chiral neutrinos can be used to create them, a modern view of the SM is to include neutrino masses, as the required ingredients are already available. This is the view of the SM which will be used and referred to throughout the rest of the thesis. The addition of right-chiral neutrinos does however open up the possibility of Majorana masses, which will be discussed in Section 3.5.

Particles masses are real, positive numbers, yet the mass terms introduced in the Lagrangian (2.6.4) depend on linear combinations of complex numbers. One might ask why one cannot just demand real, diagonal Yukawa matrices and skip this whole discussion entirely. This is an important question and will be answered in a later section when neutrino physics is discussed and the answer involves neutrino mixing and the nature of right-handed neutrinos.

2.7 The Standard Model

With the discussion on fermion masses complete, it is worthwhile to take a step back and review the whole Standard Model. The SM is the foundation of how the known fundamental particles interact with one another and its building blocks are fermions and bosons. The SM describes how the universe at small scales can be viewed as the Yang-Mills theory $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, which is spontaneously broken to $SU(3)_C \otimes U(1)_{\text{em}}$ at low-energy. The theory is governed by a Lagrangian, which is the sum of all possible terms which are

1. Lorentz Invariant
2. Gauge Invariant
3. Renormalizable

the former two being discussed in detail already. The last criteria; renormalizability, requires more discussion than this thesis can provide to understand and appreciate. For this thesis, a renormalizable term has energy dimension, dimension for short, less than or equal to four. All terms in the SM have dimension four, which is evident from dimensional analysis of the action

$$[S] = [\int_\Omega d^4x \, \mathcal{L}] = E^{-4}[\mathcal{L}], \tag{2.7.1}$$

which reveals that the dimension of the Lagrangian needs to be four. A direct consequence of $[\mathcal{L}] = E^4$ is that fermion fields have fractional energy dimension $[\psi] = E^{\frac{3}{2}}$, while bosonic fields have unit energy dimension $[\phi] = E$.

From the three guiding principles on constructing the Lagrangian, the SM more or less follows[9]. As one might expect, there are a lot of terms and writing all of them down explicitly is not recommended. Especially all the boson-boson interactions take up quite a lot of ink. Luckily, in applications, this is never needed, as one is rarely concerned with

[9] Writing out the Standard Model Lagrangian is left as an exercise for the reader

more than a few terms. As a leading example, when considering neutrinos, QCD effects are irrelevant and vice versa.

Putting this whole section together, the Standard Model Lagrangian can be written as some variation of

$$\mathcal{L} = \sum_{\text{Fermions}} i\overline{\Psi}\not{D}\Psi - \frac{1}{4}F_{\mu\nu j}F^{\mu\nu j} - (Y^{ij}\Psi_i^L H\psi_j^R + h.c) + (D_\mu H)^\dagger(D^\mu H) - V(H), \quad (2.7.2)$$

where $F_{\mu\nu j}$ is the collection of all bosonic field strength tensors, $V(H)$ is the Higgs potential and the sum is taken over all 12 (six quarks, three charged leptons and three neutrinos) fermionic fields. Do not take this equation too seriously, as it is meant to showcase how one can combine every interaction in the SM into a single line. This equation can be put on a coffee mug or printed on a t-shirt, however, it is not of much practical use. However, with that being said, the groundwork is laid and the way forth is to apply the developed theory to investigate the universe.

2.8 Problems with the Standard Model

As with every physical theory so far, also the SM has its flaws. It is certainly not wrong per se, as some of the predictions it makes are among the best tested and verified in all of science. In particular, the magnetic moment of the electron is the best-measured quantity in all of science, accurate to 0.28 parts per trillion [20], which is truly outstanding. However, the measured magnetic moment of the muon introduces a strong tension between the theoretically calculated value and the observed value. There are several problems with the SM, some are rather subtle and others very clear. An incomplete list of some of these problems are:

1. **Energy content of the universe**: Observations of the universe at large scales reveal that only about 4% of the matter in the universe is baryonic i.e. made of Standard Model particles. The remaining 96% can be categorized into roughly 26% dark matter and about 70% dark energy. The SM has no explanation for this.

2. **Gravity**: Simply put, the SM does not include gravity.

3. **Dirac vs Majorana neutrinos**: Neutrinos may acquire both a Dirac mass term and a Majorana mass term. The question of the origin of neutrino masses is still unknown.

4. **Free parameters**: The SM has 19 free parameters which cannot be deduced from theory; they have to be measured. Many physicists would like a theory with fewer free parameters like string theory, which has one; the string length. However, string theory can give rise to $\sim 10^{723}$ Standard Models [21]. This is the landscape problem.

5. **Muon g-2**: Although the revered 5σ signal has not been confirmed yet, there is a 4.2σ tension between the experimental and theoretical value of the magnetic moment of the muon [22].

This should, however, not be taken as a defeat, but rather an opportunity to discover new physics beyond the SM.

3 Neutrino physics

3.1 The discovery of neutrino oscillations

In the late 1960s, the first hint towards neutrino masses came with the discovery of solar neutrinos and the solar neutrino problem (SNP) from the Homestake experiment which monitored the solar flux for 24 years. The solar flux is the collection of all particles emitted by the sun, of which neutrinos constitute a respectable portion. The Homestake experiment was a neutrino capture experiment where a medium rich in chlorine awaits the incoming solar neutrino flux. The nuclear reaction in this case is

$$^{37}\mathrm{Cl} + \nu_e \to {}^{37}\mathrm{Ar} + e^-. \tag{3.1.1}$$

The sun provides the solar system with energy through constant thermonuclear fusion in its core, which radiates outwards and heats the solar system. In the core of the sun, the most prominent method of converting mass into energy is the proton-proton (pp) chain, which converts hydrogen into heavier elements and radiation through nuclear fusion. The whole process releases 26.73MeV worth of energy, which neutrinos obtain a few percent of depending on the step of the chain [23]. The inevitable first step is the fusion reaction

$$p + p \to {}^2\mathrm{H} + e^+ + \nu_e, \tag{3.1.2}$$

where the neutrino escapes the reaction with an energy $\mathcal{O}(\mathrm{MeV})$. Hence, the sun is an extremely powerful source of electron neutrinos, which propagate approximately freely until detected at Earth with a neutrino flux at Earth of $\Phi_\nu \sim 6 \times 10^{10} cm^{-2} s^{-1}$. The solar neutrino problem was the observed deficit of ν_e in the observed solar neutrino flux in contrast to theoretical predictions.

Over the next decades with experiments such as (Super)-Kamiokande, GALLEX, SNO and more, the SNP was resolved with the advent of neutrino oscillations. More specifically, a significant fraction of the ν_e produced in the sun are converted to ν_μ and ν_τ during propagation from the core of the sun to the awaiting detector at Earth. Due to this conversion, there is necessarily a deficit in the number of events one would expect from the nuclear reaction (3.1.1) without the possibility of neutrino oscillations. Note that mixing and oscillatory phenomena are not unique to the neutrino sector, quarks can mix between flavours, through the non-diagonal Cabibbo–Kobayashi–Maskawa (CKM) matrix. While neutral mesons, bound states of quark-antiquark pairs, can oscillate into their respective antiparticles.

3.2 Neutrino mixing

To see how adding right-chiral neutrinos to the SM results in neutrino oscillations and neutrino mixing, recall from equation (2.6.4) that the lepton Yukawa Lagrangian is given by

$$\mathcal{L}_Y = -Y_l^{ij} \overline{\Psi}^L_{li} H \psi^l_{Rj} - Y_\nu^{ij} \overline{\Psi}^L_{li} \widetilde{H} \psi^{\nu_l}_{RJ} + h.c, \tag{3.2.1}$$

which after the Higgs doublet acquires a vev, yields the leptonic mass terms

$$\mathcal{L}_M = -\frac{v}{\sqrt{2}} \bar{l}_L Y_l l_R - \frac{v}{\sqrt{2}} \overline{\nu}_L Y_\nu \nu_R + h.c. \tag{3.2.2}$$

The previous notation (Ψ_{li}^L) etc. is a bit tedious, so to avoid a notational nightmare, the three-component vectors $l_{L/R}$ and $\nu_{L/R}$ are introduced as

$$l_{L/R} \equiv (e_{L/R}\ \mu_{L/R}\ \tau_{L/R})^T, \quad \nu_{L/R} \equiv (\nu_{e_{L/R}}\ \nu_{\mu_{L/R}}\ \nu_{\tau_{L/R}})^T. \tag{3.2.3}$$

Physical particles have real masses, which means that the above Lagrangian is defined in some other basis. The basis in which the SM is usually written, which is the basis implicitly used in constructing the SM, is called the *flavour basis*. The SM is formulated in terms of interactions, particles are created and destroyed in interactions with other particles. The basis where these interactions are diagonal is called the flavour basis. As an example, neutrinos are produced as eigenstates of the weak interactions, so-called weak eigenstates. That is, neutrinos are produced in a definite flavour, not as a linear combination of flavours. The easiest option is if the flavour basis coincides with the basis in which the mass matrix is diagonal, the so-called *mass basis*. Physics is of course basis independent, but some bases are much easier to work with than others. For example, it is rather convoluted and unintuitive to describe a propagating particle in a non-diagonal mass basis. On the other hand, describing interactions in a non-diagonal flavour basis is similarly ambiguous.

Out of all the SM fermions, only the charged leptons have a shared mass and flavour basis; charged leptons do not mix. That is, a muon will not appear in a pure electron beam. This is however not clear from the formulation of Yukawa couplings. The charged lepton Yukawa interactions share the same form as neutrino Yukawa interactions; a trilinear coupling of left and right-handed fields contracted with the Higgs doublet times some complex matrix of Yukawa couplings.

To understand why charged leptons do not oscillate, quoting [24] "Flavour oscillations of charged leptons are not possible, because the flavour of a charged lepton is defined by its mass, which is the only property that distinguishes between differently charged leptons." For charged leptons, this means that flavour is measured through mass, $Y_l^{ij} = diag(y_l^1, y_l^2, y_l^3)$ for real, positive numbers y_l^i. The shared mass and flavour basis for the charged leptons means that their mixing matrices, without any loss of generality, can be taken as the identity matrix.

The mass and flavour bases are two bases of the same vector space. Thus, starting from one basis, the other one is obtained by a basis transformation. Consider the following unitary basis transformations

$$\nu_L \to U_L \nu_L, \quad \nu_R \to U_R \nu_R, \quad U_{L/R} \in U(3), \tag{3.2.4}$$

which rotates the flavour basis into the mass basis

$$\mathcal{L}_M \to -\frac{v}{\sqrt{2}} \bar{l}_L Y_l l_R - \frac{v}{2} \overline{\nu}_L U_L^\dagger Y_\nu U_R \nu_R + h.c, \tag{3.2.5}$$

where

$$U_L^\dagger Y_\nu U_R = diag(y_\nu^1, y_\nu^2, y_\nu^3), \quad y_\nu^i \geq 0. \tag{3.2.6}$$

Mathematically, this amounts to diagonalizing a complex matrix by a bi-unitary transformation, for a proof that this can be done see [25]. In rotating the neutrino basis, the terms in the Lagrangian will generally change. Neutrinos enter interactions only through the charged and neutral current interactions (2.5.44)-(2.5.46). Under the neutrino basis change (3.2.4), the neutral current (NC) transforms as

$$J^{\mu}_{\text{NC}} \to -\frac{1}{2}\left(\overline{\nu}_L U_L^{\dagger}\gamma^{\mu}U_L\nu_L - \bar{l}_L\gamma^{\mu}l_L\right) = J^{\mu}_{\text{NC}}. \tag{3.2.7}$$

As U_L is unitary, the NC interaction does not change under neutrino basis transformations, and to be general, it does not change under any unitary lepton basis transformation. This is due to the NC connecting charged leptons to charged leptons and neutrinos to neutrinos. The only physical imprint neutrino basis transformations can have is on the charged current (CC) interaction. The CC interaction is the contraction of the charged current (2.5.44) with the electroweak gauge boson W, given as

$$\mathcal{L}_{CC} = -\frac{g}{2\sqrt{2}}J^{\mu\dagger}_{CC}W_{\mu} + h.c. = -\frac{g}{2\sqrt{2}}\overline{\nu}_L\gamma^{\mu}W_{\mu}l_L + h.c, \tag{3.2.8}$$

which under the neutrino basis change, transforms as

$$\mathcal{L}_{CC} \to -\frac{g}{2\sqrt{2}}\overline{\nu}_L U_L^{\dagger}\gamma^{\mu}W_{\mu}l_L + h.c. \tag{3.2.9}$$

This expression is not generally diagonal as U_L is a general unitary matrix. The fact that the charged current is sensitive to a basis change reveals that neutrinos mix, and that effects of neutrino mixing can be attributed to a single unitary mixing matrix $U_L^{\dagger} \equiv U_{PMNS}$ called the Pontecorvo-Maki–Nakagawa–Sakata (PMNS) matrix. The mixing matrix is conventionally defined by acting on the mass basis to yield the flavour basis through

$$\begin{pmatrix}\nu_e\\ \nu_{\mu}\\ \nu_{\tau}\end{pmatrix} = \begin{pmatrix}U_{e1} & U_{e2} & U_{e3}\\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3}\\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3}\end{pmatrix}\begin{pmatrix}\nu_1\\ \nu_2\\ \nu_3\end{pmatrix}. \tag{3.2.10}$$

The elements of the PMNS matrix should be interpreted as measures of how much a flavour neutrino couples to a given mass eigenstate, e.g. $|U_{e3}|$ is a measure of how much the electron neutrino couples to the third mass eigenstate. Measurements of the PMNS matrix elements are powerful tests of the SM, as deviations from unitarity are direct violations and hence acts as a probe to physics beyond the SM, so-called BSM physics. As the PMNS matrix is a member of $U(3)$ it has nine degrees of freedom. However, many of these degrees of freedom are unphysical. It is also worth mentioning that the parameterization of the PMNS matrix is identical to the CKM matrix up to Majorana phases. These claims will be justified in Section 3.7

3.3 Neutrino oscillations

The fact that neutrinos mix implies that neutrinos may oscillate into different flavours as they propagate through spacetime. It is of interest to deduce a formula that quantifies the probability that a given flavour neutrino will oscillate into some other flavour neutrino. There are several approaches to obtaining the transition probability, one might treat it quantum-mechanically, either through a plane wave or a wave packet treatment. It is also possible to treat this in QFT, although significantly more involved. These approaches will generally lead to different transition probabilities, e.g. the QM approach is less complete than the QFT approach. In general, these approaches lead to different results. However, accounting for the sensitivity and energy threshold of detectors, they converge to the same observed transition probability [24]. For this reason, the simple quantum mechanical, plane wave treatment is presented.

3.3.1 N-dimensional oscillations

Although the SM picture concerns only three *active* neutrinos, which mix through three mass eigenstates, one can generalize the discussion to include oscillations between N flavour neutrinos mixing with N massive states. An active neutrino is defined as an eigenstate of the weak interaction, while flavour neutrino is enlarged to include both the three active neutrinos as well as possible *sterile* states. Sterile states are used as an umbrella for neutrinos which are non-trivial eigenstates of the weak interaction and does not interact directly with the SM. Sterile states are eigenstates of the weak interaction with vanishing coupling, i.e. sterile states are not produced in the weak interaction. Consider now the case of N massive neutrinos, all of which can be treated as ultra-relativistic and which are related to the flavour states by a unitary matrix U. The eigenstates of the Hamiltonian acting on the massive states are given by

$$\hat{H}|\nu_i\rangle = E_i|\nu_i\rangle, \tag{3.3.1}$$

where the energy of each massive state is approximated as

$$E_i = \sqrt{\vec{p}^2 + m_i^2} \approx |\vec{p}| + \frac{m_i^2}{2|\vec{p}|} \approx E + \frac{m_i^2}{2E}. \tag{3.3.2}$$

In general, the group velocities of each massive state will be different leading to different momenta. However, in a more complete approach [26], the dependence of the different momenta cancels out in the final transition probability. The massive states evolve according to the Schrödinger equation (1.3.2), and they are freely propagating. Therefore, they can be described by plane waves

$$|\nu_i(t)\rangle = e^{-iE_i t}|\nu_i\rangle. \tag{3.3.3}$$

Active neutrinos are produced in weak interactions as pure flavour states, say $|\nu_\alpha\rangle$ for $\alpha \in \{e, \mu, \tau\}$, at some initial time $t = 0$. In terms of the massive states, the flavour neutrino evolves in time as the superposition

$$|\nu_\alpha(t)\rangle = \sum_{i=1}^{N} U_{\alpha i}\, e^{-iE_i t}|\nu_i\rangle. \tag{3.3.4}$$

After some time $t > 0$, the neutrino is no longer in a pure flavour state and the transition amplitude of observing the neutrino in some other flavour state $|\nu_\beta\rangle$ where $\beta \in \{e, \mu, \tau, s_1, \ldots, s_{N-3}\}$, is the projection of that state onto the original state, given as the inner product

$$A(\nu_\alpha \to \nu_\beta; t) = \langle\nu_\beta|\nu_\alpha(t)\rangle = \sum_{i,j=1}^{N} U^*_{\beta i} U_{\alpha j}\, e^{-iE_j t}\langle\nu_i|\nu_j\rangle. \tag{3.3.5}$$

Using an orthonormal basis for the massive states, their inner product is the Kronecker delta, which yields the transition amplitude as the single sum

$$A(\nu_\alpha \to \nu_\beta; t) = \sum_{i=1}^{N} U^*_{\beta i} U_{\alpha i}\, e^{-iE_i t}. \tag{3.3.6}$$

The transition probability is given as the square modulus of the transition amplitude

$$P(\nu_\alpha \to \nu_\beta; t) = \sum_{i,j=1}^{N} U^*_{\alpha j} U_{\beta j} U^*_{\beta i} U_{\alpha i}\, e^{-i(E_i - E_j)t}. \tag{3.3.7}$$

Rewriting this expression in terms of the different neutrino masses and using the ultra-relativistic approximation $t = L$, as the time since a neutrino was created is not something which is measured, one obtains the master formula for neutrino oscillation probabilities

$$P(\nu_\alpha \to \nu_\beta; t) = \sum_{i,j=1}^{N} U^*_{\alpha j} U_{\beta j} U^*_{\beta i} U_{\alpha i} \exp\left(-i\frac{\Delta m^2_{ij} L}{2E}\right). \tag{3.3.8}$$

The mass squared difference Δm^2_{ij} is defined as

$$\Delta m^2_{ij} \equiv m^2_i - m^2_j. \tag{3.3.9}$$

Although not obvious from the expression for the oscillation probability, neutrino oscillations in vacuum are not sensitive to the sign of the exponent. That is, a neutrino oscillation experiment performed using the vacuum approximation, cannot determine the sign of the mass squared difference.

The elements of the mixing matrix are constants, meaning that the dynamics of neutrino oscillations can be attributed entirely to the term in the exponent. This is rather useful as the neutrino energy E is determined from the nuclear reaction which produced it, and the baseline L is easily measurable given a neutrino produced in say, the sun or some nuclear reactor. Knowing these variables, one can deduce the mass squared differences between the massive neutrino states. The main neutrino oscillation experiments and their characteristic parameters are listed in Table 3.1.

Experiment	$L(m)$	$E(MeV)$	$\Delta m^2(eV^2)$
Reactor SBL	10^2	1	10^{-2}
Reactor LBL	10^3	1	10^{-3}
Accelerator SBL	10^3	10^3	1
Accelerator LBL	10^6	10^3	10^{-3}
Atmospheric	10^7	10^3	10^{-4}
Solar	10^{11}	1	10^{-11}

Table 3.1: Order of magnitude estimations for observable mass squared splittings probed in different experiments. These values are not representative for all experiments, but rather serve to illustrate the generic values of neutrino oscillation parameters. The values are taken from [27].

A detectable[10] neutrino is necessarily a weak eigenstate, which means that both the produced and the detected neutrino are active. Remarkably, corrections due to oscillations into sterile states are still present in the expression for the transition probability as the sum is taken over the columns corresponding to the initial and final state neutrinos. Hence, deviations from the SM picture of neutrino oscillations, i.e. non-unitarity of the PMNS matrix, requires BSM physics.

[10]Detectable in the sense that an experiment can deduce that a neutrino interacted by measuring the end product of the interaction.

3.3.2 Recovering Standard Model oscillations

The previous section introduced the transition probability in the case of N ultra-relativistic, oscillating, massive neutrinos. To recover the SM picture of neutrino oscillations one can either put $N = 3$ or consider some peculiar forms of the mixing matrix. The simplest, non-trivial case is when the top left block is identified as the PMNS matrix, the remaining diagonal elements are ones, and the rest of the entries are zero. Too see this explicitly, consider oscillations where both the initial and final state neutrinos are active, say $|\nu_\alpha\rangle$ oscillates into $|\nu_\gamma\rangle$ for $(\alpha, \gamma) \in \{e, \mu, \tau\}$. The structure imposed on the mixing matrix results in the transition probability

$$P(\nu_\alpha \to \nu_\gamma; t) = \sum_{i,j=1}^{3} U^*_{\alpha j} U_{\gamma j} U^*_{\gamma i} U_{\alpha i} \exp\left(-i \frac{\Delta m^2_{ij} L}{2E}\right), \tag{3.3.10}$$

which is precisely the SM picture of neutrino oscillations as the sum is restricted to $(i, j) \in (1, 2, 3)$. This expression relied on the assumption that the except for the PMNS block, the rest of the mixing matrix were ones on the diagonal and zeros elsewhere. This assumption can in fact be relaxed and still provide the SM result. Consider the block structure of the mixing matrix given as

$$U = \begin{pmatrix} U_{\text{PMNS}} & 0_{3\times(N-3)} \\ 0_{(N-3)\times 3} & U_{\text{Dark}} \end{pmatrix}, \tag{3.3.11}$$

where $0_{A\times B}$ is the zero matrix with A rows and B columns and U_{Dark} is a unitary $(N - 3) \times (N - 3)$ matrix. The terminology, dark, is used to emphasize that regardless of what U_{Dark} is, the physics that produced it is insensitive to neutrino oscillation experiments. The transition probability of active neutrinos is blind to U_{Dark}, hence the name. Matrices on the form (3.3.11) are called *decoupled* as the dynamics of the different sectors are decoupled from each other.

3.4 Experimental values of neutrino observables

The previous sections illustrated how massive neutrinos can mix through the PMNS matrix and the general formula for neutrino oscillations was introduced. The measurable neutrino parameters introduced are the magnitudes of the PMNS matrix elements and the neutrino mass splittings, not the actual masses themselves. Oscillation data are fitted using two mass splittings, the *solar* mass splitting, defined as

$$\Delta m^2_{\text{sol}} \equiv \Delta m^2_{21} = m^2_2 - m^2_1, \tag{3.4.1}$$

which as the name suggests, is measured from solar neutrinos. The sign of the solar mass splitting is known due to the effects of neutrino oscillations in matter, which constitute an effective potential due to interactions with charged leptons. This is known as the Mikheyev–Smirnov–Wolfenstein (MSW) effect [28, 29]. When solar neutrinos propagate through the sun they are subject to matter effects, enabling the determination of the sign of the solar mass splitting through the MSW effect. The second mass splitting is the *atmospheric* mass splitting, defined as

$$\Delta m^2_{\text{atm}} \equiv |\Delta m^2_{31}| = |m^2_3 - m^2_1|, \tag{3.4.2}$$

where the sign is currently unknown. The source of atmospheric neutrinos is cosmic rays impacting the atmosphere. When these cosmic rays interact with particles in the atmosphere, they may produce charged pions. These charged pions almost exclusively decay to muons, which further decays into two neutrinos and an electron through the charged current interaction. A consequence of the undetermined sign of the atmospheric mass splitting is that it leaves the mass hierarchy of neutrinos undetermined. There are two possibilities; neutrinos may follow a *normal* hierarchy (NO)

$$\text{NO} : m_1 < m_2 < m_3, \tag{3.4.3}$$

or they may follow an *inverted* hierarchy (IO)

$$\text{IO} : m_3 < m_1 < m_2. \tag{3.4.4}$$

An illustration of the two hierarchies is given in Figure 3.1 which also reveals the scale of the mass splittings.

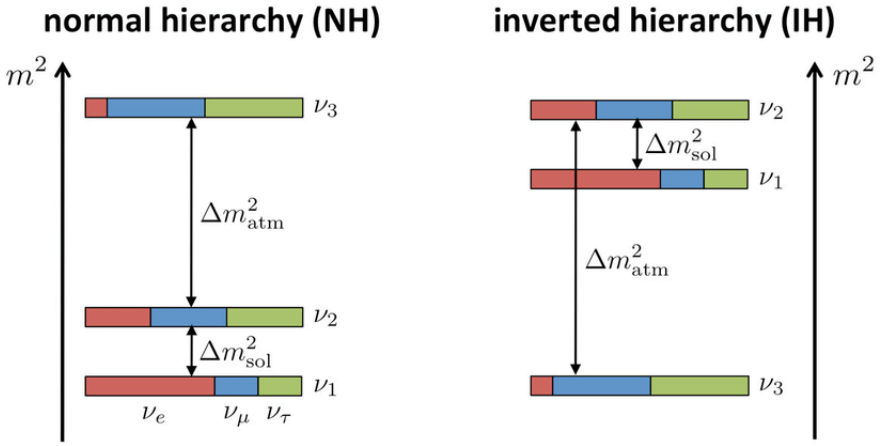

Figure 3.1: Illustration of the two mass hierarchies taken from [30]. The figures plot the relative difference between the mass splittings versus increasing mass squared for both hierarchies. The colours on horizontal bars represents the probability of observing of a given flavour in the corresponding mass eigenstate.

NO(3σ)	IO(3σ)
$\Delta m^2_{21} = 6.82 - 8.04 \times 10^{-5} eV^2$	$\Delta m^2_{21} = 6.82 - 8.04 \times 10^{-5} eV^2$
$\Delta m^2_{31} = 2.431 - 2.598 \times 10^{-3} eV^2$	$\Delta m^2_{31} = (-2.583) - (-2.412) \times 10^{-3} eV^2$

Table 3.2: Table showcasing the 3σ ranges of neutrino mass splittings from a 2020 global analysis [31]. The value of the atmospheric mass splitting is positive in NO while negative in IO, which is due to the undermined nature of neutrino masses.

The 3σ ranges of the neutrino mass splittings from a 2020 global analysis are given in Table 3.2. Remark that even though two neutrinos are massive, the lightest neutrino mass may still be vanishing, i.e. zero. This case only requires the addition of two right-handed neutrinos to provide the necessary neutrino masses. However, as this scenario is far from general it will not be pursued. Orthogonal to the constraints from neutrino

oscillations, cosmological data from the Planck collaboration [32] constrains the sum of active neutrino masses to

$$\sum_i m_i < 0.12\ eV \quad (95\% C.L). \tag{3.4.5}$$

For fixed values of the solar and atmospheric mass splittings, the sum of neutrino masses depends only on the value of one of the mass eigenstates. Conventionally, the independent mass eigenstate is taken as the lightest one. In NO the lightest massive state is m_1, yielding the sum of neutrino masses

$$\sum_i m_i(m_1) = m_1 + \sqrt{\Delta m^2_{21} + m_1^2} + \sqrt{\Delta m^2_{31} + m_1^2}, \tag{3.4.6}$$

while in IO, where m_3 is the lightest mass eigenstate, the formula is slightly modified

$$\sum_i m_i(m_3) = m_3 + \sqrt{m_3^2 - \Delta m^2_{31}} + \sqrt{m_3^2 - \Delta m^2_{31} + \Delta m^2_{21}}. \tag{3.4.7}$$

A plot of the sum of active neutrino masses versus the lightest mass eigenstate in both hierarchies is illustrated in Figure 3.2. From the figure, it is evident that NO allows for larger values of the lightest neutrino mass (m_1) within the cosmological bound on the sum of neutrino masses. The width between lines of equal colour is also narrower in NO, which is a direct consequence of the magnitude difference between the scale and uncertainty of atmospheric and solar neutrino mass splittings. Moreover, there are clear bounds on the mass of the lightest mass eigenstate, $m_1 < 3 \times 10^{-2}\ eV$ in NO and $m_3 < 2 \times 10^{-2}\ eV$ in IO.

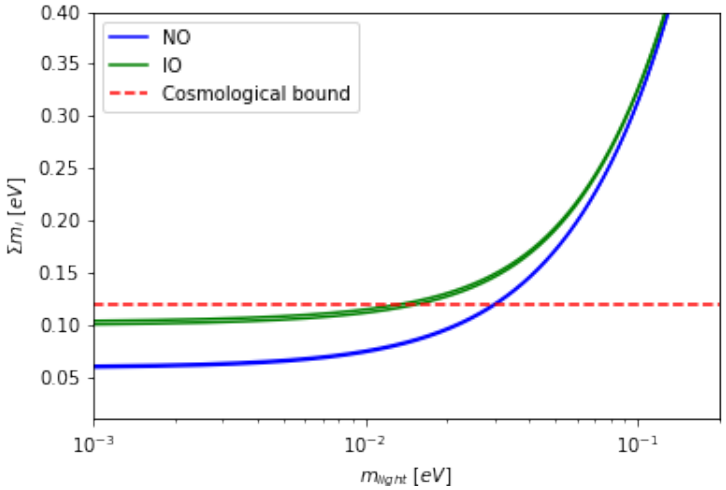

Figure 3.2: Plot of the sum of active neutrino masses versus the lightest mass eigenstate in both hierarchies. The nearly vanishing width of the graphs represent the 2σ uncertainties from measurements of the neutrino mass splittings. The horizontal, dashed, red line is the 2σ bound on neutrino masses from cosmological observations.

On the other side of neutrino oscillations, the magnitudes of PMNS matrix elements are also measured. The values of mixing matrix elements are obtained from the same 2020 global fit [31], which is under the critical assumption that the PMNS matrix is unitary.

The latter refers to the fact that the matrix elements are not independent; changing one affects the others due to unitarity $U^\dagger U = I$. The 3σ regions of the PMNS matrix are given as

$$|U_{\text{PMNS}}|_{3\sigma} = \begin{pmatrix} 0.801 \to 0.845 & 0.513 \to 0.579 & 0.143 \to 0.156 \\ 0.233 \to 0.507 & 0.461 \to 0.694 & 0.631 \to 0.778 \\ 0.261 \to 0.526 & 0.471 \to 0.701 & 0.611 \to 0.761 \end{pmatrix}. \tag{3.4.8}$$

Evident from the PMNS matrix, neutrino mixing cannot be modelled as small perturbations from the unit matrix. In particular, some off-diagonal elements are comparable or even larger than some of the diagonal elements, further showcasing that neutrinos mix quite significantly. For reference, this is in contrast to the quark sector, in which mixing is not too prevalent. The largest quark mixing, which is between first and second-generation quarks, is of order $|V_{us}| \approx |V_{cd}| = \mathcal{O}(0.2)$ [33].

3.5 Majorana masses

After spontaneous symmetry breaking, all fermions of the Standard Model share the same Dirac mass term; a bilinear term connecting left and right chiral spinors. There is, however, another way a chiral fermion may acquire a mass term, which is by obtaining a Majorana mass. Recall that the infinitesimal Lorentz boosts of right chiral Weyl spinors (2.4.4) are given by

$$\delta\psi_R^W = \frac{1}{2}\left(i\theta^j + \beta^j\right)\sigma_j\psi_R^W, \tag{3.5.1}$$

for rotation and boost angles θ^j and β^j, respectively. The Pauli matrices σ_j are given as

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.5.2}$$

and they satisfy the anti-commutator

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}. \tag{3.5.3}$$

The Majorana mass term is given as the bilinear (up to constants)

$$\mathcal{L}_{\text{Majorana}} = -\psi^{W^T}\sigma_2\psi^W, \tag{3.5.4}$$

which is a mass term given in terms of a single Weyl spinor. This is in contrast to the Dirac mass term which requires two Weyl spinors. To show that the Majorana mass term is Lorentz invariant, consider an infinitesimal Lorentz boost of the bilinear Majorana mass term

$$\delta(\psi_R^{W^T}\sigma_2\psi_R^W) = \frac{1}{2}\left(i\theta^j + \beta^j\right)\left(\psi_R^{W^T}\sigma_j^T\sigma_2 + \psi_R^{W^T}\sigma_2\sigma_j\psi_R^W\right). \tag{3.5.5}$$

Due to the anti-commutator (3.5.3) and the Pauli matrix relation $\sigma_j^T\sigma_2 = -\sigma_2\sigma_j$, the Majorana term can be written as

$$\delta(\psi_R^{W^T}\sigma_2\psi_R^W) = \frac{1}{2}\left(i\theta^j + \beta^j\right)\left(-\psi_R^{W^T}\sigma_2\sigma_j\psi_R^W + \psi_R^{W^T}\sigma_2\sigma_j\psi_R^W\right) = 0. \tag{3.5.6}$$

The Majorana mass term is Lorentz invariant! There is, however, a caveat in the construction of the Majorana mass term; particles described by a Majorana mass term cannot be charged under any gauge group. This means that particles described by a Majorana mass term, so-called Majorana particles, must be their own antiparticle. To showcase that Majorana particles must be uncharged under all gauge groups, consider a generic gauge transformation

$$\varphi \to X(x)\varphi = \exp\left(-iT^k\alpha_k(x)\right)\varphi, \quad X^\dagger(x) = X^{-1}(x), \tag{3.5.7}$$

where T^k are a set of generators and $\alpha_k(x)$ are differentiable functions of spacetime. For a field with a Majorana mass, the above gauge transformation transforms the Majorana mass term

$$-\mathcal{L}_{\text{Majorana}} = \varphi^T\sigma_2\varphi \to \varphi^T X^T(x)\sigma X(x)\varphi = -\mathcal{L}_{\text{Majorana}}X(2x), \tag{3.5.8}$$

which is clearly not invariant. Thus, Majorana particles must be uncharged in order to respect gauge invariance. In standard nomenclature, particles which do not transform under gauge transformation are called *singlets*. Out of all known particles, only right-chiral neutrinos can have Majorana masses, as all other particles are charged under a subgroup of the SM symmetry group. A consequence of their nature is that Majorana particles are not constrained to any mass scale, as the other charged particles are. In particular, the SM mass scale is the vacuum expectation value of the Higgs doublet $v = 246\text{GeV}$, as all particle masses in the SM are directly proportional to it.

To avoid cluttered notation, the superscript on Weyl spinors is dropped and the distinction between Weyl and Dirac spinors should be clear from context. The Lagrangian for a single Weyl spinor with a Majorana mass term, conveniently taken as the right-chiral Weyl spinor, is given by [15]

$$\mathcal{L}_{\text{Weyl + Maj}} = i\psi_R^\dagger\sigma^\mu\partial_\mu\psi_R + i\frac{m}{2}\left(\psi_R^\dagger\sigma^2\psi_R^* - \psi_R^T\sigma^2\psi_R\right), \tag{3.5.9}$$

where $\sigma^\mu(\overline{\sigma}^\mu)$ are introduced as

$$\sigma^\mu \equiv (1, \sigma^i), \quad \overline{\sigma}^\mu \equiv (1, -\sigma^i). \tag{3.5.10}$$

The new $\overline{\sigma}^\mu$ will be useful later when introducing the Weyl representation of the gamma matrices. Fermions described by this Lagrangian are called Majorana fermions. It is useful to embed the Weyl spinor into a Dirac spinor, which for Majorana fermions is given by

$$\psi = \begin{pmatrix} i\sigma^2\psi_R^* \\ \psi_R \end{pmatrix}. \tag{3.5.11}$$

The first entry of ψ is a left-handed field written in terms of a right-handed one. Using the newly constructed Dirac spinor, it is possible to write the Majorana mass term concisely as

$$\frac{m}{2}\overline{\psi}\psi = \frac{m}{2}\left(\psi_R^\dagger i\sigma^2\psi_R^* - \psi_R^T i\sigma^2\psi_R + \psi_R^\dagger\psi_R + \psi_R^T\psi_R^*\right), \tag{3.5.12}$$

where the last two terms cancel due to anti-commutativity of fermions. In the Weyl representation, the gamma matrices are given in block form as

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \overline{\sigma}^\mu & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}. \tag{3.5.13}$$

Majorana fermions described in terms of Dirac spinors satisfy

$$i\gamma^2\psi^* = \psi, \tag{3.5.14}$$

which is called the Majorana condition. Closely related to the Majorana condition is charge conjugation, in fact the operation

$$C : \psi \to i\gamma^2\psi^* \equiv \psi^c, \quad C : \overline{\psi} \to i\psi^T\gamma^2\gamma^0 \equiv \overline{\psi^c}, \tag{3.5.15}$$

is defined as the charge conjugation operator (in the Weyl basis). The effects of charge conjugation are best showcased when acting on the Dirac equation in an electromagnetic field, where charge conjugation flips the sign of the electric charge. Charge conjugation yields equations of motion for antiparticles and as Majorana particles do not transform under charge conjugation they must be gauge singlets. In summary: Majorana masses only require a Weyl spinor to construct a mass term, which makes them simpler to envision than Dirac masses, which require two independent Weyl spinors to construct a mass term. However, simplicity comes at a great cost in gauge theories; Majorana particles must be singlets under any gauge group. The effect of right-chiral neutrinos obtaining Majorana mass terms will be discussed in the next section.

3.6 The duality of neutrino masses

The previous section illustrated that neutral particles such as right-chiral neutrinos may obtain two separate mass terms. To explore the effects of the Majorana mass term, it is convenient to introduce some new notation. The vector of right-chiral states is introduced as

$$N = \begin{pmatrix} N_1 & \dots & N_N \end{pmatrix}^T \equiv \begin{pmatrix} \psi_{R_1} & \dots & \psi_{R_1} \end{pmatrix}^T, \tag{3.6.1}$$

which is a N-dimensional vector in flavour space. Although the number of active neutrinos is constrained by the invisible decay of the Z boson, the number of right-chiral neutrinos is unconstrained. Hence, one generally includes N (the number) such states. The general neutrino Yukawa Lagrangian is given as

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{i=1}^{N}\sum_{j=1}^{3}\left(Y_\nu^{ij}\overline{N}_i\widetilde{H}^\dagger L_j\right) - \frac{1}{2}\sum_{i,k=1}^{N}\left(M_R^{ik}\overline{N^c}_i N_k\right) + h.c, \tag{3.6.2}$$

where M_R^{ik} are the components of the Majorana mass matrix. After spontaneous symmetry breaking the Lagrangian acquires the following mass terms

$$\mathcal{L}_{\text{Mass}} = -\frac{v}{\sqrt{2}}\overline{N}Y\nu_L - \frac{1}{2}\overline{N^c}M_R N + h.c, \tag{3.6.3}$$

where $\nu_L \equiv \nu$ is the defined as the vector of left-chiral states

$$\nu \equiv \begin{pmatrix} \nu_e & \nu_\mu & \nu_\tau \end{pmatrix}^T. \tag{3.6.4}$$

Proceeding in standard convention, the Dirac mass matrix is defined as

$$M_D \equiv \frac{v}{\sqrt{2}} Y_\nu, \tag{3.6.5}$$

which reflects how Dirac mass term appears in the SM; as bilinear terms between left and right chiral fields. To obtain a single matrix term in the mass Lagrangian, which can then be diagonalized to yield physical masses, the vector Ψ is defined as the collection of all left-handed neutrinos

$$\Psi \equiv \begin{pmatrix} \nu & N^c \end{pmatrix}^T. \tag{3.6.6}$$

In addition, the enlarged mass matrix $\widetilde{\mathcal{M}}$ is introduced as

$$\widetilde{\mathcal{M}} \equiv \begin{pmatrix} 0 & M_D^T \\ M_D & M_R \end{pmatrix}, \tag{3.6.7}$$

which enables the mass Lagrangian to be expressed compactly as

$$\mathcal{L}_{\text{Mass}} = -\frac{1}{2} \overline{\Psi^c} \widetilde{\mathcal{M}} \Psi + h.c. \tag{3.6.8}$$

Remarkably, in the general case of both Majorana and Dirac masses, the combined mass term is a Majorana mass. The nature of neutrino masses is an open question in physics and there are observable differences between Dirac and Majorana neutrinos. A peculiar effect of Majorana neutrinos is that lepton number is no longer conserved as the global $U(1)$ symmetry corresponding to lepton number

$$\nu_l \to e^{i\theta} \nu_l, \quad N \to e^{i\theta} N, \tag{3.6.9}$$

is violated by the Majorana mass term by two units. This means that Majorana masses can produce interactions where the overall number of leptons is changed by two. Thus, by searching for lepton number violating interactions one can deduce if neutrinos are Majorana or Dirac. In particular, the search for neutrino-less double beta decays ($2\beta0\nu$) is of great interest as a positive signal would resolve this discussion. However $2\beta0\nu$ is second order in the weak interaction, which in turn yields a very low cross-section, see e.g. [34,35] for reviews on the status and importance of $2\beta0\nu$. The leading order Feynman diagram for $2\beta0\nu$ is illustrated in 3.3.

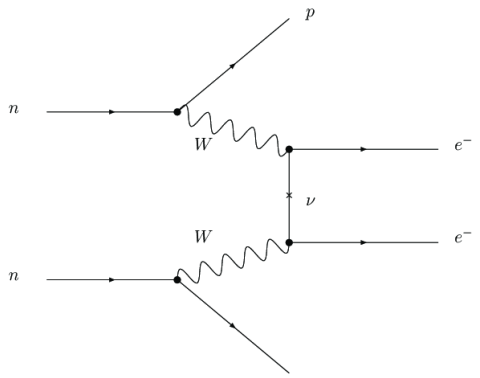

Figure 3.3: Feynman diagram showcasing neutrino-less double beta decay. The illustration is taken from [35].

3.6.1 Dirac neutrinos

If neutrinos are Dirac particles, then there is only one term in the mass Lagrangian. As illustrated in Section 3.4, cosmological data confine the sum of neutrino masses to $\sum_i m_i < 0.12\text{eV}$. On the other hand, atmospheric data suggest that at least one of the active neutrinos has a mass that is bounded from below by

$$m_{\text{heavy}} \geq \sqrt{|\Delta m^2_{atm}|} \simeq 0.05\text{eV}. \tag{3.6.10}$$

As Dirac masses are given as products of Yukawa couplings with the Higgs vev, the scale of neutrino Yukawa couplings is confined to

$$y_\nu \lesssim 10^{-12}. \tag{3.6.11}$$

This is a very small number. To understand which Yukawa couplings are considered small, it is instructive to consider the perturbative limit. A perturbative theory is one in which higher-order terms, originating from loops in QFT, can be treated as corrections. To obtain a perturbative theory, neutral couplings like Yukawa couplings are required to satisfy the perturbative bound

$$\frac{|y|}{4\pi} < 1. \tag{3.6.12}$$

The idea that a free parameter of a theory takes on such an abysmal value compared to the perturbative range may be experienced as quite unsettling. Created in a Dirac way, the neutrino Yukawa couplings are at least six orders of magnitude lower than the electron Yukawa coupling, which is the lightest Standard Model fermion. This unnatural small value of a free parameter has led physicists to alternative methods of neutrino mass generation. On the other hand, the electron Yukawa coupling $y_e = \mathcal{O}(10^{-6})$ is also a very small number in comparison to 4π. These issues are so-called smallness problems. The working definition of a smallness problem is for this thesis defined as a free parameter whose value is low in comparison to the range it is confined to.

3.6.2 Majorana neutrinos and the see-saw mechanism

Dirac neutrinos suffer a severe smallness problem to explain neutrino masses. If instead, neutrinos are Majorana particles, then the Lagrangian is given by (3.6.8). The see-saw mechanism can provide a natural explanation for the smallness of neutrino masses on the condition that all eigenvalues of the Dirac mass matrix are much smaller than all eigenvalues of the Majorana mass matrix. The see-saw story goes as follows. Consider the block-diagonalization of the enlarged mass matrix

$$\mathcal{B}^T \widetilde{\mathcal{M}} \mathcal{B} = \mathcal{M} = \begin{pmatrix} \mathcal{M}_{\text{light}} & 0 \\ 0 & \mathcal{M}_{\text{heavy}} \end{pmatrix} \tag{3.6.13}$$

which separates the light, active neutrinos from the heavy singlets. The see-saw mechanism decouples the light sector from the heavy sector, which enables one to work in an effective theory where the singlets are mere heavy degrees of freedom with no dynamics. The diagonalization matrix $\mathcal{B}$ is given in block form as

$$\mathcal{B} = \begin{pmatrix} I & M_D^\dagger M_R^{-1\dagger} \\ -M_R^{-1} M_D & I \end{pmatrix}, \tag{3.6.14}$$

which is approximately unitary

$$\mathcal{B}^\dagger \mathcal{B} = I + \mathcal{O}(M_D M_R^{-1})^2. \tag{3.6.15}$$

The fact that the block diagonalization matrix $\mathcal{B}$ is not unitary means that the working theory is an effective theory. In particular, the non-unitarity should be explainable by the physics which UV-completes the theory. Block-diagonalizing yields the light neutrino mass matrix

$$\mathcal{M}_{\text{light}} = -M_D^T M_R^{-1} M_D, \tag{3.6.16}$$

where neutrino masses are suppressed by the singlet mass matrix. The see-saw can mechanism avoids unnaturally small Yukawa couplings under the assumption that the singlet mass scale is much larger than the EW scale. When doing numerical work, the negative sign in the light mass matrix can be ignored. The correct sign can then be reconciled either by field definitions, or an overall phase of the Dirac mass matrix ($M_D \to iM_D$). In the see-saw picture, neutrino masses are small as an algebraic byproduct of physics at vastly different scales. The heavy mass matrix, on the other hand, is nearly unchanged from the singlet mass matrix

$$\mathcal{M}_{\text{heavy}} = M_R + \mathcal{O}(M_D^2 M_R^{-1}) \simeq M_R. \tag{3.6.17}$$

The see-saw mechanism is widely used in high-energy theories, especially grand-unified-theory (GUT) models, as by choosing Yukawa couplings $y = \mathcal{O}(1)$ and using $m_\nu = 0.1\text{eV}$, leads to singlet masses

$$M_R = \frac{v^2}{m_\nu} \sim 10^{14}\text{GeV}, \tag{3.6.18}$$

which is just below the GUT-scale, $\Lambda \sim 10^{16} GeV$. A peculiar phenomenon in QFT is that the couplings "constants" are not constant, they obtain an energy dependence. The GUT-scale is the scale at which the running couplings constants converge to the same strength. At this point it is possible to describe the three quantized forces by a single Lie group, the most common being $SO(10)$. Although the GUT-scale mass is the most common place to assign the singlet masses, all singlet masses are technically natural as a symmetry (lepton number) is restored in the limit of vanishing singlet masses. This definition of naturalness is credited to 't Hooft [36]. The see-saw mechanism has an astoundingly large range as

$$(M_D M_R^{-1})_{ij} \ll 1, \tag{3.6.19}$$

is satisfied automatically for singlet eigenvalues larger than about 10TeV. The see-saw limit is defined as the limit where (3.6.19) holds and the dynamics of light neutrinos can be decoupled from the heavy singlet masses. In this limit, the effective Lagrangian describing neutrino masses is given as

$$\mathcal{L}_{\text{Mass}} = -\frac{1}{2}\overline{\nu^c}\mathcal{M}_{\text{light}}\nu + h.c. \tag{3.6.20}$$

This Lagrangian describes three active, Majorana neutrinos in the flavour basis. To obtain physical masses, a basis transformation from the flavour basis to the mass basis is required. Consider the following unitary basis transformation which rotates the flavour basis

$$\overline{\nu^c} \to \overline{\nu^c} U_L^T, \quad \nu \to U_L \nu. \tag{3.6.21}$$

In turn, the Lagrangian transforms as

$$\mathcal{L}_{\text{Mass}} \to -\frac{1}{2}\overline{\nu^c} U_L^T \mathcal{M}_{\text{light}} U_L \nu + h.c, \tag{3.6.22}$$

for the diagonalized light mass matrix

$$U_L^T \mathcal{M}_{\text{light}} U_L = \mathcal{M}_{\text{diag}} = \text{diag}(m_1, m_2, m_3), \quad m_i \geq 0. \tag{3.6.23}$$

Before proceeding, it is worth remarking that singlet masses can be smaller than 10^4GeV if either the Yukawa couplings are small, or there are strong cancellations imposed on the Yukawa couplings. Although the latter requires fine-tuning, the possibility of it should not be completely discarded. These "accidentally" small neutrino masses obtained from cancellations of Yukawa couplings can translate to singlet masses within the reach of detectors such as large hadron collider (LHC). If singlet masses are lighter than the Higgs boson, then the decay channel $\Gamma(H \to N\nu_\alpha)$ is possible and potentially detectable given large enough active-singlet mixing [37, 38].

Before ending this section, it is worth noting that neutrino masses can be generated within the SM without invoking the need for right-chiral neutrinos if non-renormalizable terms are included. The most prevalent is the dimension five term credited to Weinberg [39], which considering one generation for simplicity, is given by

$$\mathcal{L}_{\text{Weinberg}} = -\frac{g^*}{\Lambda^*}(\overline{L}\widetilde{H})(\widetilde{H}L^c)^\dagger + h.c. \tag{3.6.24}$$

Here g^* is a dimensionless coupling constant and Λ^* is some energy cutoff-scale. After symmetry breaking, the Weinberg Lagrangian yields neutrino masses

$$m_\nu = \frac{g^* v^2}{2\Lambda^*}, \tag{3.6.25}$$

which is identical to the see-saw relation for one generation of neutrinos. In the picture, neutrino masses are obtained by tuning the value g^*/Λ^* to fit observations. However, to obtain a renormalizable theory the see-saw mechanism will be pursued instead.

To summarize this section: Majorana masses can provide an elegant solution to the smallness problems introduced by Dirac neutrinos. However, this comes at a cost, Majorana masses violate lepton number, whereas Dirac masses do not. For further reference, the case where neutrinos obtain their masses through Dirac bilinears is defined as an SM phenomenon for the following reason: The ingredients required to construct neutrino masses are already there, it only requires the addition of gauge-singlets to do so.

3.7 Parametrization of mixing matrices

The previous sections illustrate how fermion mixing and oscillations are described in terms of unitary matrices acting on either the flavour basis or the mass basis to produce the other one. Often, it is convenient to have an explicit parametrization of these matrices. The most illustrative being the "consecutive rotations" parametrization. If the mixing matrices are real, then the three-dimensional bases can be rotated into each other by the product of three real rotations performed in succession. In general, if R_{ij} is a real

rotation in the $i-j$ plane of a N-dimensional, Euclidean vector space, then any combined rotation can be written as the product

$$\mathcal{R} = \prod_{i<j} R_{ij}(\theta_{ij}), \quad \mathcal{R}^T = \mathcal{R}^{-1}. \tag{3.7.1}$$

The rotation matrices $R_{ij}(\theta_{ij})$ are dependent on single parameter; the angle θ_{ij} in the $i-j$ plane. The order in which rotations are taken is irrelevant, but the parametrization will look different depending on the order. This fact is useful, as experiments are seldom sensitive to all elements of the mixing matrix. Hence, a parametrization that simplifies the row(s) and/or column(s) of interest is practical. As an example of the explicit form of rotation matrices, consider a four-dimensional space. A rotation in the $2-3$ plane of that space is given by

$$R_{23} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c_{23} & -s_{23} & 0 \\ 0 & s_{23} & c_{23} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.7.2}$$

where in standard convention, the shorthand notations for the trigonometric functions are defined as

$$c_{ij} \equiv \cos\theta_{ij}, \quad s_{ij} \equiv \sin\theta_{ij}. \tag{3.7.3}$$

Unfortunately, the interesting mixing matrices are complex and unitary, instead of real and orthogonal. Nevertheless, unitary matrices share a lot of the same structure that orthogonal matrices do as unitary matrices can be viewed as the complex analogue of ordinary, real rotations. In particular, any member of $U(N)$ can be written in terms of

$$\frac{N(N-1)}{2} \quad \text{mixing angles and} \quad \frac{N(N+1)}{2} \quad \text{phases}, \tag{3.7.4}$$

for a total on N^2 degrees of freedom. The number of mixing angles coincides with the number of independent rotations one can perform in a N-dimensional space, hence they represent the usual rotation part. The phases on the other hand, are constructs which do not share the same illustrative description as the mixing angles. A general unitary matrix can be constructed in a similar fashion to the general rotation matrix introduced in equation (3.7.1), i.e. as the product of consecutive unitary matrices confined to the complex $i-j$ planes. Denote unitary matrices acting on the complex $i-j$ plane as $W_{ij}(\theta_{ij}, \eta_{ij})$ for angles θ_{ij} and phases η_{ij}, whose components are given as [24, 40]

$$\left(W_{ij}(\theta_{ij}, \eta_{ij})\right)_{rs} = \delta_{rs} + (c_{ij} - 1)\left(\delta_{ri}\delta_{si} + \delta_{rj}\delta_{sj}\right) + s_{ij}\left(e^{i\eta_{ij}}\delta_{ri}\delta_{sj} - e^{-i\eta_{ij}}\delta_{rj}\delta_{si}\right). \tag{3.7.5}$$

To illustrate the structure of unitary matrices given by the above definition, consider a three-dimensional complex space. The unitary matrix $W_{13}(\theta_{13}, \eta_{13})$ in this space is given by

$$W_{13}(\theta_{13}, \eta_{13}) = \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\eta_{13}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\eta_{13}} & 0 & c_{13} \end{pmatrix}. \tag{3.7.6}$$

This matrix is the source of the CP phase in the CKM and PMNS matrices. Conventionally, the phase η_{13} is relabeled to δ or δ_{13}. Informally, the parametrization (3.7.5) yields rotation matrices where the phase is coupled to the terms with the sine of the mixing angle. A direct consequence of this is that for small mixing angles, the CP phases will challenging to observe. In particular, for possible sterile states, the mixing angle is required to be small. Hence, the CP phases corresponding to sterile states will be very elusive.

A general member of $U(N)$ is parametrized as the product of these matrices along with a diagonal matrix of phases, given as

$$\mathcal{W} = D(\varphi) \prod_{i<j} W_{ij}(\theta_{ij}, \eta_{ij}), \tag{3.7.7}$$

where

$$D(\varphi) = \text{diag}\left(e^{i\varphi_1}, \ldots, e^{i\varphi_N}\right). \tag{3.7.8}$$

When the phase of a unitary matrix W_{ij} is zero, the matrix reduces to a real rotation matrix

$$W_{ij}(\theta_{ij}, \eta_{ij} = 0) = R_{ij}(\theta_{ij}). \tag{3.7.9}$$

The unitary mixing matrix appearing in the CC interaction can be parametrized by (3.7.7). However, the physical mixing matrix is not dependent on all phases appearing in the general parametrization, some phases are unphysical, here is why. Consider neutrinos as Dirac particles, i.e. the Lagrangian with no lepton number violating Majorana mass term. The Lagrangian exhibits a global $U(1)^6$ symmetry in the lepton sector, which can be used to perform the following field redefinitions

$$l_\alpha \to e^{i\theta_\alpha} l_\alpha, \quad \nu_\beta \to e^{i\varphi_\beta} \nu_\beta, \quad (\alpha, \beta) \in (e, \mu, \tau), \tag{3.7.10}$$

which leaves the Lagrangian unchanged. These phases are arbitrary. In the mass basis, the CC interaction (3.2.9) transformed by the inclusion of a unitary matrix U_L, which is the PMNS matrix. In general, before arriving at the PMNS matrix, the matrix appearing in the CC interaction is a general unitary matrix with nine d.o.f. Call this matrix $\mathcal{W}$. Under the leptonic field redefinitions, the CC Lagrangian changes to

$$\mathcal{L}_{\text{CC}} \supset -\frac{g}{2\sqrt{2}} W_\mu \sum_{\alpha,\beta} \overline{\nu_\beta} e^{-i\varphi_\beta} \gamma^\mu \mathcal{W}_{\alpha\beta} e^{i\theta_\alpha} l_\alpha + h.c. \tag{3.7.11}$$

By factoring an arbitrary phase difference outside the sum, the Lagrangian can be written as

$$\mathcal{L}_{\text{CC}} \supset -\frac{g}{2\sqrt{2}} e^{-i(\varphi_1 - \theta_1)} \sum_{\alpha,\beta} \overline{\nu_\beta} e^{-i(\varphi_\beta - \varphi_1)} \gamma^\mu \mathcal{W}_{\alpha\beta} e^{i(\theta_\alpha - \theta_1)} l_\alpha + h.c. \tag{3.7.12}$$

There are a total of five arbitrary phase differences in this Lagrangian. These can be used to remove five phases from the unitary matrix $\mathcal{W}$, resulting in the familiar PMNS matrix. An equivalent approach is used in the quark sector to arrive at the CKM matrix. This is why the CKM and PMNS matrices have only one physical phase. The lepton field redefinitions (3.7.10) are possible whenever the only mass term in the Lagrangian is a Dirac term. This is a given for quarks and charged leptons. Neutrinos on the other hand

may acquire a Majorana mass term, which violates the global neutrino $U(1)$ symmetry. The phases due to Majorana masses can be factored to the right of the Dirac-type mixing matrix. Let U_D be the Dirac mixing matrix; an umbrella for the matrix containing mixing angles and the Dirac phase(s), and U_M is the diagonal matrix of Majorana phases, then the full mixing matrix can be written as

$$U = U_D U_M. \tag{3.7.13}$$

In addition, one can show that one of the Majorana phases can be redefined away [40]. In the $N = 3$ case, for which both the CKM and the PMNS matrix reside, their standard parametrization is given as $U = R_{23}W_{13}R_{12}$ times a possible matrix of Majorana phases for neutrinos. Written out explicitly yields

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\phi_1} & 0 \\ 0 & 0 & e^{i\phi_2} \end{pmatrix}. \tag{3.7.14}$$

One can show that under the combined Charge-Parity (CP) symmetry, the Lagrangian changes unless $U^* = U$, that is U must be real to avoid CP-violation. Hence, any CP violation in the weak sector is equivalent to non-zero values of the physical phases. For Dirac particles, there is only one such phase. For Majorana particles, there are three; one Dirac phase and two Majorana phases.

A consequence of the fact that Majorana phases can be factored to the right of the mixing matrix has important implications for neutrino oscillations. Recall that the neutrino transition probability (3.3.8) is dependent on the quartic product

$$A_{\alpha\beta} = \sum_{i,j=1}^{3} U^*_{\alpha j} U_{\beta j} U^*_{\beta i} U_{\alpha i}, \tag{3.7.15}$$

and that elements of the PMNS matrix are products of Dirac terms and Majorana terms, given as

$$U_{\alpha j} = \sum_{k=1}^{3} (U_D)_{\alpha k} (U_M)_{kj} = (U_D)_{\alpha j} (U_M)_{jj}. \tag{3.7.16}$$

The sum reduced to a single term as U_M is diagonal. When PMNS matrix elements are separated into Dirac and Majorana terms like in the above equation, then the quartic product can be expressed as

$$A_{\alpha\beta} = \sum_{i,j=1}^{3} (U_D)^*_{\alpha j} (U_M)^*_{jj} (U_D)_{\beta j} (U_M)_{jj} (U_D)^*_{\beta i} (U_M)^*_{ii} (U_D)_{\alpha i} (U_M)_{ii}. \tag{3.7.17}$$

From this equation it is clear that all Majorana phases cancel out; neutrino oscillations are insensitive to whether neutrinos are Dirac or Majorana particles! Hence, oscillation experiments cannot provide any hints towards the nature of how neutrinos masses originate.

This outlined approach has been suited to the standard case of three active neutrinos mixing with three singlets. With the inclusion of sterile neutrinos, non-trivial eigenstates

of the weak interaction, the situation is more complicated. In the presence of N_S sterile states, the number of Dirac phases is [24]

$$N_D = 1 + 2N_S, \tag{3.7.18}$$

while the number of Majorana phases is

$$N_M = 2 + N_S. \tag{3.7.19}$$

Sterile neutrinos will be discussed in later sections.

4 Physics beyond the Standard Model

The discussion on the Standard Model ended with a foreshadowing of physics which is not explainable by the SM, an outlook for phenomena that require additional, new physics to explain. This section is devoted to the intriguing physics which must lay beyond the Standard Model and how such physics can be probed.

4.1 Dark sectors

The journey so far reveals that the particles of the universe are to our best understanding described by the Standard Model of particle physics; a gauge theory under which all particles are charged. But is that all there is to it? The best way to start this section may be to ask the simple question: When we point our telescopes to the sky, what do we see?

The simplified answer is: Not what we expected. If general relativity and the current understanding of gravity is right, then the universe is filled with non-luminous matter, i.e. matter which does not interact electromagnetically. This provides the descriptive name, dark matter (DM). dark matter does not interact electromagnetically, leaves only the neutrinos as SM candidates to explain dark matter. Without venturing too far into cosmology, neutrinos are generally disfavoured as they simply have too little mass to account for DM. The general framework for particle cosmology is the ΛCDM model where CDM stands for cold dark matter. In cosmological terminology, neutrinos are considered hot, which although hot dark matter models exist, the majority of research into DM assumes more massive particles. If dark matter can be described in terms of particles, then how do they fit into the picture and how can they be probed? Correctly answering either one of these questions is a guaranteed Nobel Prize.

From the point of view of the SM, all particles are charged under the SM symmetry group, but there is no reason why every particle in existence should be charged under the SM gauge group

$$G_{SM} \equiv SU(3)_C \otimes SU(2)_L \otimes U(1)_Y. \tag{4.1.1}$$

An example of particles that are not charged under this group are the right-chiral neutrinos, which poses the question: can there be more? In the gauge theory description, the answer is simply, yes. All SM particles arise from singlet contractions of fields which can be fit into a Lagrangian. A singlet contraction is defined as a term in the Lagrangian which is both Lorentz and gauge invariant, of which there are a lot. Imposing a new symmetry on the Lagrangian leads to the possibility of new particles and interactions,

which given the right one, can account for DM. The SM group is then extended to a larger symmetry group

$$G = G_{SM} \otimes G_{\text{Dark}}, \tag{4.1.2}$$

where G_{Dark} is some symmetry group whose particle content does not interact electromagnetically. Physics that is unaffected by electromagnetism are bestowed their own sector, the dark sector. The new gauge group may be a single group or the product of several groups. If the gauge-theoretic picture of particles is correct, then the particle content of the universe should be contained in G. The obvious problem, however, is that there are no clues for what G_{Dark} is. In particular, it would be extremely obnoxious if, for example, the dark sector is described in terms of the symmetry group

$$G_{\text{Dark}} = SU(100), \tag{4.1.3}$$

as the $100^2 - 1 = 9999$ generators would imply 9999 gauge bosons and some rather long sums. Although entirely possible that the dark sector is described by a high-dimensional group, any sensible person would start by exploring the simple solutions, the familiar $U(1)$'s and the $SU(2)$'s. The former sentence may be thought of as a group-theoretic version of Occam's Razor: "Entities should not be multiplied without necessity". I.e. the simplest solution is often the best.

For a dark sector to be of interest, it must somehow yield predictions that are observable through the physics of the SM. If not, then the only footprint it leaves on the universe is gravitational, which cannot be used to probe quantum properties of the particles inhabiting the dark sectors. In the following discussion, DM is used as a collective measure for all particles belonging to dark sectors. In a particle physics setting, interactions between dark sectors and the SM can be visualized through various Feynman diagrams. Depending on the particle content of the dark sector, several of the following pseudo-Feynman graphs illustrated in Figures 4.1, 4.2, 4.3 may be used to detect DM. In all Feynman graphs, time is taken to move from left to right.

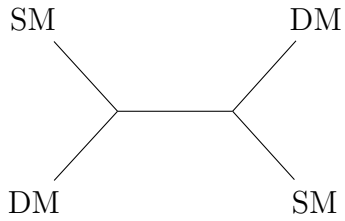

Figure 4.1: Scattering

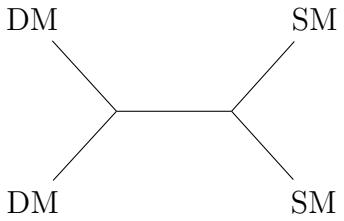

Figure 4.2: Indirect

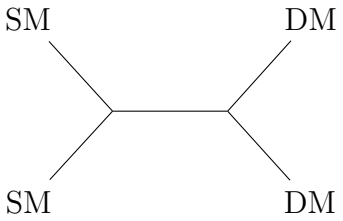

Figure 4.3: Collider method

Figure 4.1 illustrates the most straightforward method of discovering dark physics, it represents scattering of a DM particle with a SM particle; direct observation. Apart from the apparent simplicity of direct searches, they suffer two main drawbacks. First, the lack of direct observations of DM restricts the coupling to the SM to be small; DM needs to be weakly interacting. Secondly, it requires a region of the universe that is dense in DM, something the solar system is not. Hence, direct searches require cosmological data, which are famously plagued with large uncertainties. This is due to the large number of interactions that take place from a particles inception in some far-away galaxy until it interacts in a detector at Earth. However, the precision is only getting better.

Figure 4.2 illustrates indirect searches. If DM is charged under some symmetry group it can interact with itself and hence annihilate. If the particle created in a DM annihilation can further decay into SM particles, then these may be detected and studied to reveal the invariant mass of the decaying particle. If the invariant mass is incompatible with SM particles, a discovery is made. However, this method suffers from the same drawbacks as direct searches.

Figure 4.3 illustrates how DM can be probed in detectors. From a precision standpoint, the collider method is greatly preferred for the simple reason that the environment can be controlled. Colliders provide vast amounts of data from experiments, in which the key to understanding DM may already lay. Of course, reality is more complicated than summarized in the previous sentences, and all experimental data require proper analysis subject to a model, which yields statistical confidence of varying degree bases upon the data provided. This is why a large subset of papers on particle physics is titled as a permutation of "Search for X at experiment Y at energies Z". Here, X is a signal, a collective measure of either a particle or a prediction from theory. There is, however, a fundamental drawback to collider searches; the centre-of-mass energy. At the time of writing this thesis, the largest centre-of-mass energy achieved is 13TeV at the LHC at CERN. The problem is apparent, if the masses of DM particles are larger than the energy scale of colliders, it cannot be probed as pictured in the last Feynman diagram Figure 4.3. The non-observation of DM may indicate that the DM mass scale is beyond present collider energies, in this case, the only way to probe DM is through cosmological events, or larger accelerators. Limited to Earth, however, the latter may not be realistic.

These arguments are general and apply to all generic DM models, as they rely on model-independent arguments. Dark sectors do not, however, have to be high-energy phenomena. As will be clear from the next section, there is a low-energy phenomenon in the neutrino sector which may hint towards dark sectors.

4.2 Short baseline neutrino anomalies

Recall that the neutrino flavour transition amplitude (3.3.6) is oscillating with frequencies

$$\Phi_{ij} = \frac{\Delta m_{ij}^2 L}{2E}. \tag{4.2.1}$$

For oscillations to occur, the oscillation frequency must be comparable to unity or larger. For smaller values, oscillations are essentially frozen in place as they have not had enough time to develop. Thus, the baselines for which solar and atmospheric oscillations are relevant are of order

$$L_{\text{sol}} \gtrsim \frac{E}{\Delta m_{\text{sol}}^2} \approx 50km\frac{E}{MeV}, \quad L_{\text{atm}} \gtrsim \frac{E}{\Delta m_{\text{atm}}^2} \approx 1km\frac{E}{MeV}. \tag{4.2.2}$$

For shorter distances than these, any deviations in neutrino beams cannot be solved by solar and/or atmospheric neutrino oscillations. A convenient location to place a detector to measure properties of neutrinos is, well, right beside a source. Limited to Earth, these sources are mainly nuclear power plants and particle colliders, as converting mass to energy in a power plant releases a large flux of neutrinos through beta decay of heavy nuclei. For colliders, a smart choice to place a neutrino detector is besides a potent muon beam. Neutrinos produced in power plants are called reactor neutrinos and they typically lay either a couple of meters away from the core up to some kilometres.

4.2.1 The liquid scintillator appearance anomaly

The liquid neutrino detector (LSND) at Los Alamos National Laboratory was an experiment measuring the anti-neutrino oscillations $\overline{\nu}_\mu \to \overline{\nu}_e$ and they detected an excess in the number of electron anti-neutrinos in the muon anti-neutrino beam [41] at 3.8σ. Today this is known as the LSND anti-neutrino anomaly, LSND anomaly for short. To follow up on the LSND anomaly, the MiniBooNE short-baseline (SBL) [32, 42] experiment measured neutrino fluxes at comparable baselines. They also found an excess in anti-electron neutrinos with respect to the standard three flavour predictions at 4.8σ. Seeing that these experiments are in strict tension with ordinary oscillation predictions, the anomalies can be remarkably well solved with the addition of a new, independent mass splitting

$$\Delta m^2_{\text{SBL}} \sim 1\text{eV}^2, \tag{4.2.3}$$

which is much larger than both the solar and the atmospheric mass splittings. Hence, the SBL mass splitting requires a fourth mass eigenstate at the eV-scale! There is a subtlety to this new eigenstate, it cannot be an active state due to constraints from the invisible decay of the Z boson, which is only compatible with three active neutrinos [43]. Therefore, the fourth eigenstate implies that the fourth flavour neutrino is sterile, i.e. a SM singlet. Models which consider an additional sterile neutrino are collectively called $3+1$ active-sterile neutrino models, or simply $3+1$ models. Calling the fourth neutrino sterile is slightly wrong, as it does interact with the other neutrinos through oscillations. Thus, some papers may refer to the fourth neutrino as mostly-sterile or some variation thereof. However, this thesis will refer to the fourth flavour neutrino as sterile when no particular model of the fourth neutrino is considered.

4.2.2 The reactor disappearance anomalies

Neutrino experiments rely on counting the number of neutrinos that hit the detector per time unit and comparing that number to the expected number. This expected number is called the neutrino flux and is a construct of many quantities; baseline, energy, the efficiency of the detector, cross-sections, decay rates and more. The latter two are dependent on the specific interactions between the heavy nuclei used in the reactor, which is usually isotopes of uranium and plutonium, which requires knowledge of the nuclear matrix element. This is notoriously difficult and leads to moderate theoretical uncertainties. The reactor antineutrino anomaly [44] was unearthed in 2011 following the updated Huber-Mueller flux calculation [45, 46] of the $\overline{\nu}_e$ flux from nuclear reactors. Before the Huber-Mueller prediction, the measured flux was within that of earlier theoretical predictions. However, the Huber-Mueller calculation showed that the predicted flux should be 3% larger than previously assumed. Hence, from the point of the Huber-Mueller calculation, there is a deficit to the number density of $\overline{\nu}_e$ emitted from reactors, which cannot be attributed to active-active neutrino oscillations. The observed deficit can, however, be explained by $\overline{\nu}_e$ oscillating into a sterile state while propagating from the nuclear reactor to the detector. Tangential to the reactor antineutrino anomaly, the Gallium anomaly [47] is the observed deficit in the flux of ν_e measured from radioactive Gallium sources. These phenomena may be attributed to the disappearance of ν_e due to oscillations into sterile states. For an overview of the Gallium anomaly in terms of $3+1$ active-sterile oscillations see [48], in which they conclude that the deficit is compatible with an eV-scale sterile neutrino.

4.2.3 The SBL transition probability

The LSND, MiniBooNE and the reactor anomalies are collectively referred to as short baseline experiments. The first two constitute SBL appearance experiments, while the reactor anomalies govern SBL disappearance experiments. This distinction will be made clear shortly.

Having discussed the SBL anomalies, it is of interest to derive an expression for the transition probability subject to the SBL condition

$$\Phi_{4j} \sim 1, \quad \Phi_{ij} \ll 1, \quad (i,j) \in (1,2,3). \tag{4.2.4}$$

This states that for short baselines, the oscillatory phases corresponding to active-active neutrino oscillations have not yet developed. Due to the hierarchy of oscillation frequencies, it is possible to approximate all active-sterile frequencies as equal, that is

$$\Phi_{41} \approx \Phi_{42} \approx \Phi_{43} \equiv \Phi, \tag{4.2.5}$$

which enables the transition probability (3.3.6) to be expressed as

$$P_{\text{SBL}}\left(\nu_\alpha \to \nu_\beta\right) = \left| \sum_{k\leq 3} U^*_{\alpha k} U_{\beta k} e^{-i\Phi_{k1}} + U^*_{\alpha 4} U_{\beta 4} e^{-i\Phi} \right|^2 . \tag{4.2.6}$$

Following the assumption that the active-sterile oscillation frequencies are frozen at the relevant baselines and using unitary of the mixing matrix $\sum_k U^*_{\alpha k} U_{\beta k} = \delta_{\alpha\beta}$, the effective SBL transition probability reads

$$P_{\text{SBL}}\left(\nu_\alpha \to \nu_\beta\right) \simeq \left| \delta_{\alpha\beta} - U^*_{\alpha 4} U_{\beta 4} + U^*_{\alpha 4} U_{\beta 4} e^{-i\Phi} \right|^2 . \tag{4.2.7}$$

Expanding the square modulus, the SBL transition probability yields

$$\begin{aligned} P_{\text{SBL}}\left(\nu_\alpha \to \nu_\beta\right) = \delta_{\alpha\beta} - \delta_{\alpha\beta} U_{\alpha 4} U^*_{\beta 4} \left(1 - e^{i\Phi}\right) - \delta_{\alpha\beta} U^*_{\alpha 4} U_{\beta 4} \left(1 - e^{-i\Phi}\right) \\ + |U_{\alpha 4}|^2 |U_{\beta 4}|^2 \left(e^{i\Phi} - 1\right)\left(e^{-i\Phi} - 1\right) . \end{aligned} \tag{4.2.8}$$

Neutrino oscillation experiments are classified into two classes:

1. **Appearance** experiments: These are experiments where the goal is to observe a different flavoured neutrino than originally produced, e.g. the appearance of electron neutrinos in a muon neutrino beam. Appearance experiments require $\nu_\alpha \to \nu_\beta : \alpha \neq \beta$.

2. **Disappearance** experiments: This type of experiment searches for deficits in the number density of a neutrino flux as a function of propagation distance. It is a measure of the likelihood that a given flavour neutrino will oscillate into the same flavour, hence disappearance experiments require $\nu_\alpha \to \nu_\alpha$.

In the former case, observing a different flavour neutrino means that all terms proportional to the Kronecker delta vanish. Using elementary properties of complex numbers and some trigonometric relations, the SBL appearance transition amplitude can be written as

$$P^{\text{App}}_{\text{SBL}}\left(\nu_\alpha \to \nu_\beta\right) = 4|U_{\alpha 4}|^2|U_{\beta 4}|^2 \sin^2\left(\frac{\Phi}{2}\right) = 4|U_{\alpha 4}|^2|U_{\beta 4}|^2 \sin^2\left(\frac{\Delta m^2_{41} L}{4E}\right). \tag{4.2.9}$$

For disappearance experiments, it is similarly straightforward to show that the transition amplitude is given by

$$P^{\text{Dis}}_{\text{SBL}}\left(\nu_\alpha \to \nu_\alpha\right) = 1 - 4|U_{\alpha 4}|^2\left(1 - |U_{\alpha 4}|^2\right)\sin^2\left(\frac{\Delta m^2_{41} L}{4E}\right). \tag{4.2.10}$$

It is noteworthy that the constant term in both of these expressions relies solely on the elements of the last column of the mixing matrix. That is, the amplitude of SBL oscillations are fully determined by how much the flavour states couple to the fourth mass eigenstate. Furthermore, only the square modulus of mixing matrix elements enter the transition probabilities. Thus, SBL transition probabilities are insensitive to any possible CP violation in the neutrino sector, both Dirac and Majorana phases. For completeness, it is worth remarking that the coefficients in front of the oscillatory terms are often interpreted as effective mixing angles, conventionally defined as

$$\sin^2 2\theta_{\alpha\beta} \equiv 4|U_{\alpha 4}|^2|U_{\beta 4}|^2, \quad \sin^2 2\theta_{\alpha\alpha} \equiv 4|U_{\alpha 4}|^2\left(1 - |U_{\alpha 4}|^2\right). \tag{4.2.11}$$

The SBL transition amplitudes were derived under the assumption that only one mass splitting is relevant while the others can be approximated as static. Following the same steps as outlined in this section, one can apply these arguments to the three flavour scenario. In this case, there is a hierarchical mass splitting

$$|\Delta m_{21}| \ll |\Delta m_{31}| \approx |\Delta m_{32}|, \tag{4.2.12}$$

which for oscillation frequencies

$$\Phi_{31} \approx \Phi_{32} \sim 1, \quad \Phi_{21} \ll 1, \tag{4.2.13}$$

yield the atmospheric baseline transition probabilities

$$P^{\text{App}}_{\text{Atm}}\left(\nu_\alpha \to \nu_\beta\right) = 4|U_{\alpha 3}|^2|U_{\beta 3}|^2 \sin^2\left(\frac{\Delta m^2_{31} L}{4E}\right). \tag{4.2.14}$$

Similar modifications yield the atmospheric disappearance transition probability. Again, the transition probabilities are only dependent on the last column (now the 3×3 matrix PMNS matrix), which are measures of the flavour eigenstates couplings to the third mass eigenstate.

4.2.4 Appearance-disappearance tension

The eV-scale sterile neutrino needed to account for the short baseline anomalies necessarily imply that other experiments should be seeing the effects of the sterile neutrino as well. In particular, the active-sterile mixing angles needed to account for the appearance experiments are in strict tension with non-observations of disappearance experiments [49]. With more data, the overlapping regions between appearance and disappearance experiments have shrunk and the trend points towards alternative explanations for the SBL anomalies. However, the current situation is evolving and will remain unclear until more

experiments are performed and more data is collected. In particular, until future experiments have been conducted with similar parameters as the LSND and MiniBooNE, the final word is left unspoken.

If the appearance signals cannot be reconstructed by future experiments, the 3+1 model will be excluded as a solution to the SBL anomaly, but not as a theory of its own. In that case, one could always say that the active-sterile mixing angle is small, and therefore it has not been discovered yet. Either way, the only way forth is to continue searching and continuously testing the model. Even though the 3+1 model may not suffice for SBL anomalies, the theory can certainly be adapted to other problems. In particular, the methods devised in the rest of the thesis apply to dark sectors where the neutrino portal is open. This will be discussed in Section 5.3.

4.3 Mixing in general 3+1 neutrino models

The recently introduced SBL anomaly has sparked great interest in neutrino physics and sterile neutrinos, which naturally has led to several different models on sterile neutrinos. Before delving into a specific model, it is important to cover the accessible parameter space which a general 3+1 model provides. That is, to put model-independent constraints on 3+1 mixing. First of all, a fourth neutrino necessarily implies that the 3×3 PMNS matrix is no longer unitary, as the mixing matrix is now a 4×4 matrix, which generally implies active-sterile neutrino mixing. In the 3+1 model, flavour eigenstates are now related to the massive states by

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \\ \nu_s \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} & U_{e4} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} & U_{\mu4} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} & U_{\tau4} \\ U_{s1} & U_{s1} & U_{s3} & U_{s4} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{pmatrix}, \tag{4.3.1}$$

where ν_s is the fourth flavour neutrino, the subscript refers to it being sterile. The elements of the last column of the mixing matrix are measures of how much the flavour neutrinos couple to the fourth eigenstate. General considerations of the success of three flavour neutrino oscillations yields the generic bounds

$$|U_{\alpha4}| \lesssim 0.1, \quad |U_{s4}| \approx 1, \quad \alpha \in (e, \mu, \tau). \tag{4.3.2}$$

From unitarity of U it also follows that the sterile neutrino is mostly composed of the fourth mass eigenstate. Theoretically, there is, however, no reason to restrain neutrino mixing to only one sterile state. The arguments above apply identically when n_s sterile states are included. However, by adding n_s sterile states, the number of elements of the mixing matrix scales as n_s^2. Therefore, recalling Occam's Razor, the $3 + 1$ case is the main focus of both this thesis and most papers on active-sterile neutrino mixing. Furthermore, although the SBL anomaly hints towards eV-scale sterile neutrinos, the mixing matrix is oblivious to the masses of sterile neutrinos, it is just a mathematical relation. Hence, adding a new sterile neutrino to the mix does not fix a certain mass scale of the new massive state. The mass scale is rather an auxiliary constrain to the model at hand. Having said that, the SBL mass scale is of the largest phenomenological interest. The umbrella of models which consider either SBL mass scales or somewhat larger ones are called light sterile neutrino models. As a general rule in physics, the definition of an adjective such as e.g. light or high, are vague at best and is better used

as a relative measure within given models[11]. For a comprehensive review on the recent status of eV-scale neutrinos in the $3+1$ model, the reader is referred to [50].

5 Sterile neutrinos from a dark sector

The SBL anomaly is a powerful probe towards BSM physics and possible dark sectors. If the 3+1 neutrino hypothesis is correct, a mechanism for the creation of the sterile neutrino is required. The simplest solution is that at least one of the right-chiral neutrinos has a mass at the eV-scale, which is still sufficiently heavy enough to generate light, active neutrinos through the see-saw mechanism [51]. Although this is the most straightforward approach, nominating a right-chiral neutrino to serve as the eV-scale sterile neutrino requires a large hierarchy within both the Dirac mass matrix and the singlet mass matrix. Take the Dirac mass matrix which is generated by the BEH mechanism as an example, to respect constraints on active-sterile mixing within the aforementioned model, elements of the Dirac mass matrix span over ten orders of magnitude [51]. The large hierarchy between neutrino Yukawa couplings and the rest of the SM Yukawa couplings is the reason why Majorana masses are considered in the first place, so even though $3+1$ mixing can be explained this way, it does not answer the question of why the neutrino Yukawa couplings are so small.

Another method of constructing eV-scale sterile neutrinos is to conjecture that their origin is attributed to a dark sector. If the dark sector is capable of generating particle masses and does not directly couple to the SM, then the fermions of the dark sector may be responsible for the SBL anomaly through neutrino mixing and hence act as a portal to new physics!

5.1 The particle content of a minimal, Abelian, dark sector

Consider an Abelian dark sector $U(1)_X$, where X is the charge of the new group. The group is endowed with a fermion ν_s and a mediator X, which is a vector field. To yield particle masses, the group contains a complex scalar φ, inhabiting a Higgs-like potential which is spontaneously broken below a critical temperature. Lastly, to obtain a non-anomalous theory, a secondary fermion ν_s' with opposite $U(1)_X$ charge of ν_s is added. The terminology of dark neutrinos and dark fermions is reserved for the new fermions. The former will mainly be used. These newly introduced particles are defined to transform as singlets under the SM symmetry group, while SM particles are taken to transform as singlets under $U(1)_X$. The right-chiral neutrinos are taken as gauge-singlets under both groups. The enlarged symmetry group of interest is given by

$$G \equiv G_{SM} \otimes U(1)_X. \tag{5.1.1}$$

Particles that do not transform under G are defined as singlets. Let the charge of $U(1)_X$ in units of X be denoted g_x, the charges of all relevant particles are given in Table 5.1

The remaining unspecified field content in Table 5.1 is the following; L^i are the lepton doublets, e_{R^i} are the right-chiral charged leptons, N^i are the right-chiral neutrinos and

[11] In general relativity, the gravitational field of the Sun is weak, which of course in the grand picture, it is. But to e.g. someone learning the subject for the first time, it appears as a somewhat obscure statement.

	L^i	e_{R^i}	N^i	ν_s	ν'_s	φ	H
$SU(2)_L$	**2**	**1**	**1**	**1**	**1**	**1**	**2**
$U(1)_Y$	$-1/2$	-1	0	0	0	0	$1/2$
$U(1)_X$	0	0	0	1	-1	-1	0

Table 5.1: Table of charges for the selected particles under the relevant symmetry groups. All particles listed in this table transform as singlets under $SU(3)_C$.

H is the Higgs doublet. The assumption that the new particles are massive is realized by considering the scalar potential

$$V(\varphi) = -\mu_\varphi^2 \varphi^\dagger \varphi + \lambda_\varphi (\varphi^\dagger \varphi)^2, \quad (\mu_\varphi^2, \lambda_\varphi) > 0, \tag{5.1.2}$$

which has a non-zero vacuum expectation value

$$v_\varphi \equiv \langle \varphi \rangle = \sqrt{\frac{\mu_\varphi^2}{\lambda_\varphi}}. \tag{5.1.3}$$

The relevant Lagrangian can be split into three parts; the kinetic terms, the interaction terms and the mass terms. The latter one is defined as the terms which, after symmetry breaking, includes the mass terms. The kinetic Lagrangian is given by

$$\mathcal{L} \supset i\overline{\nu}_s \partial\!\!\!/ \nu_s + i\overline{\nu}'_s \partial\!\!\!/ \nu'_s + i\overline{N} \partial\!\!\!/ N - \frac{1}{4} X_{\mu\nu} X^{\mu\nu} + (\partial_\mu \varphi)^\dagger (\partial^\mu \varphi), \tag{5.1.4}$$

where $X_{\mu\nu}$ is the bosonic field strength tensor. Qualitatively, these kinetic terms describe the motion of free fermions, a vector boson and a scalar boson.

5.2 The bosonic portals

The interaction terms include both the dark scalar potential and the Higgs potential along with the rest of the relevant interactions, which are given by the interaction Lagrangian

$$-\mathcal{L} \supset V(\varphi) + V(H) + g_x \overline{\nu}_s X_\mu \gamma^\mu \nu_s - g_x \overline{\nu}'_s X_\mu \gamma^\mu \nu'_s + \lambda_{\varphi H} (\varphi^\dagger \varphi)(H^\dagger H) + \frac{\varepsilon}{2} X_{\mu\nu} B^{\mu\nu}. \tag{5.2.1}$$

The last two terms in this Lagrangian couple $U(1)_X$ to the SM and are called *portals*. Noteworthy, scalar mixing inevitably mixes the Higgs doublet with the dark scalar which opens up the Higgs portal, a gateway to dark physics enabled by the Higgs doublet. To respect the observed decay channels of the Higgs, the portal must be somehow suppressed. As discussed in [52], low mixing can be achieved if there is a large hierarchy between the induced vevs. In particular, if the Higgs vev is taken to be much larger than the dark vev

$$\langle H \rangle \gg \langle \varphi \rangle, \tag{5.2.2}$$

then [52] predicts that constraints from LHC data should be well within the model capabilities. The last term in the interaction Lagrangian is called kinetic mixing, which allows the hypercharge boson B_μ to mix with the dark scalar X_μ. The hypercharge boson is related to the photon and Z through the Weinberg angle. Thus, kinetic mixing acts as a portal to the dark sector, proportional to the dimensionless coupling ε. A large value of ε would open the decay channel $Z \to XX$, which is constrained by the invisible decay of Z,

on the condition that it is kinematically accessible. Therefore, the kinetic mixing needs to be suppressed, similar to the Higgs portal mentioned previously. A subtle consequence of kinetic mixing is that particles charged under $U(1)_X$ with zero electric charge ($Q_e = 0$), will acquire a minuscule effective electric charge [53]

$$Q_{\text{EM}}^{\text{Eff}} = -\varepsilon \frac{g_x}{e} Q_x. \tag{5.2.3}$$

Both the dark scalar and the dark neutrinos will obtain such an effective electric charge due to kinetic mixing. If ε is a large number ($\varepsilon \sim 1$), then the dark sector would not be as dark as required from non-observations of additional dark sectors. Therefore, the kinetic mixing coefficient ε is required to be small, hence the terminology minuscule. From cosmological observations of effective minuscule charges, one can deduce an upper bound on kinetic mixing from a broken $U(1)$, given as [53]

$$\varepsilon \lesssim 10^{-14} \frac{e}{g_x}, \quad \text{for} \quad m_X \lesssim 10\text{keV}. \tag{5.2.4}$$

Such a small number for a dimensionless coupling is unnatural and hints towards additional new physics to describe the apparent smallness, see e.g. [54] for a discussion on how to suppress kinetic mixing. The bound on kinetic mixing and more precisely, the bound on the mediator mass coincides nicely with the hierarchy of vacuum expectation values $\langle\varphi\rangle \ll \langle H\rangle$. To showcase this, consider the mass Lagrangian

$$-\mathcal{L} \supset g_x^2 \varphi^\dagger \varphi X_\mu^\dagger X^\mu + Y_\nu \overline{N} \widetilde{H}^\dagger L_L + Y_s \overline{N} \varphi \nu_s + Y_s' \overline{N} \varphi^\dagger \nu_s' + \frac{1}{2} M_R \overline{N} N^c + h.c, \tag{5.2.5}$$

where Y_ν is the matrix of active Yukawa couplings, Y_s and Y_s' are vectors of dark Yukawa couplings, and M_R is the Majorana mass matrix. After both EW and $U(1)_X$ symmetry breaking, the following mass terms appear in the Lagrangian

$$-\mathcal{L} \supset \frac{1}{2} g_x^2 \langle\varphi\rangle^2 X_\mu^\dagger X^\mu + \frac{1}{\sqrt{2}} \langle H\rangle Y_\nu \overline{N} \nu_L + \frac{1}{\sqrt{2}} \langle\varphi\rangle Y_s \overline{N} \nu_s + \frac{1}{\sqrt{2}} \langle\varphi\rangle Y_s' \overline{N} \nu_s' + \frac{1}{2} M_R \overline{N} N^c + h, . \tag{5.2.6}$$

from which the dark vector boson mass is given by

$$m_X = g_x \langle\varphi\rangle. \tag{5.2.7}$$

If the dark sector is broken at the MeV-scale and $g_x \lesssim \mathcal{O}(10^{-2})$, then both the bosonic portals are suppressed, leaving the sector as dark as required by observations.

5.3 The neutrino portal

There is a third way to peek into the physics of the dark sector; the neutrino portal. This portal is enabled by the non-diagonal matrix structure of the enlarged mass term $\widetilde{\mathcal{M}}$, constructed by combining all neutrino mass terms into a single matrix. The Yukawa Lagrangian consists of three Dirac terms ($Y_\nu\langle H\rangle$, $Y_s\langle\varphi\rangle$ and $Y_s'\langle\varphi\rangle$) and one Majorana term M_R, the latter can be shown to be symmetric, while the Dirac terms do not share this feature. In general, the number of singlets is unconstrained, but for numerical simplicity, the number of singlets is chosen equal to the number of light neutrinos. In doing so the

mass matrices are square. In analogue to the discussion on Majorana masses, the mass terms can be combined into a single 10×10 matrix $\widetilde{\mathcal{M}}$, acting on the vector

$$\Psi^T \equiv \begin{pmatrix} \nu_e & \nu_\mu & \nu_\tau & \nu_s & \nu_s' & N_1^c & N_2^c & N_3^c & N_4^c & N_5^c \end{pmatrix}, \tag{5.3.1}$$

to yield the combined Majorana mass Lagrangian

$$\mathcal{L}_{\text{Mass}} = -\frac{1}{2}\overline{\Psi^c}\widetilde{\mathcal{M}}\Psi + h.c. \tag{5.3.2}$$

The enlarged mass matrix is given in terms of 5×5 blocks as

$$\widetilde{\mathcal{M}} = \begin{pmatrix} 0 & M_D^T \\ M_D & M_R \end{pmatrix}, \tag{5.3.3}$$

where the Dirac mass matrix is extended to include both the SM term as well as the dark sector Dirac terms, explicitly

$$M_D = \frac{1}{\sqrt{2}} \begin{pmatrix} y_{e1}\langle H\rangle & y_{\mu 1}\langle H\rangle & y_{\tau 1}\langle H\rangle & y_{s1}\langle\varphi\rangle & y_{s'1}\langle\varphi\rangle \\ y_{e2}\langle H\rangle & y_{\mu 2}\langle H\rangle & y_{\tau 2}\langle H\rangle & y_{s2}\langle\varphi\rangle & y_{s'2}\langle\varphi\rangle \\ y_{e3}\langle H\rangle & y_{\mu 3}\langle H\rangle & y_{\tau 3}\langle H\rangle & y_{s3}\langle\varphi\rangle & y_{s'3}\langle\varphi\rangle \\ y_{e4}\langle H\rangle & y_{\mu 4}\langle H\rangle & y_{\tau 4}\langle H\rangle & y_{s4}\langle\varphi\rangle & y_{s'4}\langle\varphi\rangle \\ y_{e5}\langle H\rangle & y_{\mu 5}\langle H\rangle & y_{\tau 5}\langle H\rangle & y_{s5}\langle\varphi\rangle & y_{s'5}\langle\varphi\rangle \end{pmatrix}. \tag{5.3.4}$$

The Dirac matrix is constructed such that each column is a measure of each light neutrinos $(\nu_\alpha, \nu_s, \nu_s')$ coupling strength to each singlet. Conversely, the rows of the Dirac matrix are measures of how much each singlet couples to the light neutrinos. To constrain the Higgs portal and kinetic mixing to be compatible with observations, a hierarchy on the vevs had to be imposed. As the Higgs vev is much larger than the dark scalar vev, the Dirac matrix will generally take on a hierarchical form, where elements of the last two columns are much smaller than elements of the other columns, given that the Yukawa couplings are not too different. The combined mass matrix can be block diagonalized completely analogous to the three-neutrino case, by the see-saw mechanism. The goal is to deduce the structure of neutrino mixing and masses, which amounts to solving for the eigenvalues of a complex 10×10 matrix, where all entries are unknown. At first glance, this is a total of 200 free parameters. However, this number is much less due to symmetry arguments, but still rather large.

5.4 Parameter filtering

To restrict the number of parameters, one of the dark neutrinos is assumed to be decoupled from the remaining light spectra. This can be motivated by the following arguments; the vast majority of active-sterile oscillation studies are limited to a $3+1$ mixing scenario as the addition of several sterile neutrinos $(3+N_s)$ have the same qualitative effects [55]. See e.g. [56] for a comparison between a 3+1 and a 3+2 global fit. It also boils down to what the desired outcome is: For example, nothing is preventing the newly added $U(1)$ to include N pairs of dark fermions. If all of these are massive, then then the Dirac matrix is $(3+N_s+N_N)^2$ dimensional, where N_s is the number of dark fermions and N_N is the number of singlets. However, this is too general to be of use. Invoking Occam's razor, the number of dark neutrinos is taken to be a single pair, where half of this particle duo, say ν_s', is assumed to have negligible mixing with the active neutrinos. There are two ways

to make this happen. The mass of ν_s' can be significantly larger than the eV-scale, which would make it unable to account for the SBL mass splitting. If that is the case, it may be integrated out. The other possibility is that it simply does not mix much with the rest of the light neutrinos. The extreme case of this latter statement occurs whenever ν_s' is decoupled from the other light neutrinos, equivalent to the following Dirac mass matrix structure

$$M_D = \frac{1}{\sqrt{2}} \begin{pmatrix} y_{e1}\langle H\rangle & y_{\mu 1}\langle H\rangle & y_{\tau 1}\langle H\rangle & y_{s1}\langle\varphi\rangle & 0 \\ y_{e2}\langle H\rangle & y_{\mu 2}\langle H\rangle & y_{\tau 2}\langle H\rangle & y_{s2}\langle\varphi\rangle & 0 \\ y_{e3}\langle H\rangle & y_{\mu 3}\langle H\rangle & y_{\tau 3}\langle H\rangle & y_{s3}\langle\varphi\rangle & 0 \\ y_{e4}\langle H\rangle & y_{\mu 4}\langle H\rangle & y_{\tau 4}\langle H\rangle & y_{s4}\langle\varphi\rangle & 0 \\ 0 & 0 & 0 & 0 & y_{s'5}\langle\varphi\rangle \end{pmatrix}. \tag{5.4.1}$$

This matrix structure describes four light neutrinos coupling and mixing with four singlets. In addition, there is a light neutrino, decoupled from the rest of the light neutrinos, which only couples to one singlet. The working scenario for this thesis will assume that the fifth light neutrino can be ignored.

With the second dark neutrino out of the picture, the combined mass matrix is now an 8×8 complex matrix. The low-energy observables constitute six mixing angles, four masses, three Dirac phases and three Majorana phases for a total of 16 observables. Therefore, there is bound to be a large degree of redundancy if all 128 parameters are to be applied, whether in a parameter scan or some other procedure. Fortunately, this number can be severely reduced. From the block structure of the symmetric mass matrix $\widetilde{\mathcal{M}}$, the two upper blocks add no new information, which cuts the number of parameters in half. Furthermore, the symmetric, singlet Majorana matrix M_R can be chosen real and diagonal without any loss of generality. To prove this, note that any symmetric, complex matrix can be diagonalized by a unitary matrix W

$$W^T \widetilde{M_R} W = M_R = \mathrm{diag}(M_1, \ldots, M_N), \tag{5.4.2}$$

where M_i are real and positive. In field theory language, such a diagonalization amounts to a basis rotation of N

$$N \to WN, \quad \overline{N} \to \overline{N} W^\dagger, \quad \overline{N^c} \to \overline{N^c} W^T, \tag{5.4.3}$$

where the latter two follow directly from the first. Basis transformations are only physical if the Lagrangian changes under the transformation. The only term in the Lagrangian sensitive to choosing a real, diagonal singlet basis, is the singlet kinetic term, which transforms as

$$\mathcal{L} \supset i\overline{N}\partial\!\!\!/ N \to i\overline{N} W^\dagger \partial\!\!\!/ W N = i\overline{N}\partial\!\!\!/ N. \tag{5.4.4}$$

As the Lagrangian is unchanged under the singlet basis transformation, one can therefore consider a real, diagonal singlet matrix without any loss of generality. Note that by performing the above basis rotation, there is no more freedom left in the singlet basis. Recounting free parameters in $\widetilde{\mathcal{M}}$, there are 32 from the Dirac matrix and four from the singlet Majorana matrix, for a total of 36, which is more than enough. In complete generality, this is the minimum number of parameters needed to describe the neutrino mixing at hand. However, the anomaly which sparked the discussion on sterile neutrinos, leading to introducing dark neutrinos, were the SBL anomalies. Recall that the SBL transition amplitude only depends on the magnitude of mixing matrix elements, e.g.

$|U_{e4}|^2$. As SBL experiments cannot probe possible CP-violating phases, it is justifiable to consider real, orthogonal mixing matrices, which implies that Yukawa couplings can be taken as real numbers. To illustrate this point, recall that the light neutrino mass matrix is diagonalized by the mixing matrix

$$U^T \mathcal{M}_{\text{light}} U = -\frac{1}{2} U^T M_D^T M_R^{-1} M_D U = \text{diag}(m_1, m_2, m_3, m_4), \quad m_i \geq 0, \tag{5.4.5}$$

where, by assumption, U and M_R are real. Pathologically, there might be some extreme level of cancellation which removes all possible complex numbers arising from the Dirac matrix when summed through in the matrix product. However, this seems very unlikely and would require significant fine-tuning. Therefore, the choice of a real mixing matrix coincides with restricting the spectrum of Yukawa couplings to real numbers. These assumptions will be used in upcoming sections discussing the impact of dark neutrinos on neutrino mixing and neutrino masses.

6 1+1 Active-sterile neutrino mixing

6.1 Explicit expressions for 1+1 mixing

The simplest case showcasing qualitative effects of neutrino mixing is the $1 + 1$ mixing scenario, interpreted as mixing one active neutrino with the dark neutrino. There is still the ambiguity of how many singlets to include, as the number of singlets is unconstrained. Therefore, it is useful to define the terminology $2\nu nN$ mixing as the mixing configuration including one active neutrino, one dark neutrino and n singlets. The first step is to consider $2\nu 1N$, which is the simplest configuration in which the see-saw mechanism can be applied. In the basis (ν_α, ν_s, N_1), the mass matrix is given by

$$\widetilde{\mathcal{M}} = \begin{pmatrix} 0 & 0 & y_\alpha \langle H \rangle \\ 0 & 0 & y_s \langle \varphi \rangle \\ y_\alpha \langle H \rangle & y_s \langle \varphi \rangle & M_1 \end{pmatrix}, \tag{6.1.1}$$

which up to trivial field redefinitions to account for the possible wrong signs, is given by the see-saw formula yields the light neutrino mass matrix

$$\mathcal{M}_{\text{light}} = -M_D^T M_R^{-1} M_D = \frac{1}{2M_1} \begin{pmatrix} y_\alpha^2 \langle H \rangle^2 & y_\alpha y_s \langle \varphi \rangle \langle H \rangle \\ y_\alpha y_s \langle \varphi \rangle \langle H \rangle & y_s^2 \langle \varphi \rangle^2 \end{pmatrix}. \tag{6.1.2}$$

The eigenvalues of this matrix are straightforward to compute using the characteristic equation, and since the square of the off-diagonal elements equals the product of the diagonal elements, the eigenvalues are given as

$$\lambda_1 = m_1 = 0, \quad \lambda_2 = m_2 = \text{Tr}(\mathcal{M}_{\text{light}}). \tag{6.1.3}$$

The fact that one of the masses is vanishing means that $2\nu 1N$ cannot be used as an effective model of neutrino mixing unless applied to the special case in which the lightest active neutrino is massless. Proceeding, the next natural step is to consider $2\nu 2N$, which has the following mass matrix (in the $(\nu_\alpha, \nu_s, N_1, N_2)$ basis)

$$\widetilde{\mathcal{M}} = \begin{pmatrix} 0 & 0 & y_{\alpha 1}\langle H\rangle & y_{\alpha 2}\langle H\rangle \\ 0 & 0 & y_{s1}\langle\varphi\rangle & y_{s2}\langle\varphi\rangle \\ y_{\alpha 1}\langle H\rangle & y_{s1}\langle\varphi\rangle & M_1 & 0 \\ y_{\alpha 2}\langle H\rangle & y_{s2}\langle\varphi\rangle & 0 & M_2 \end{pmatrix}, \tag{6.1.4}$$

which by the see-saw formula yields the light neutrino mass matrix

$$\mathcal{M}_{\text{light}} = \frac{-1}{2M_1M_2}\begin{pmatrix} y_{\alpha 1}^2\langle H\rangle^2 M_2 + y_{\alpha 2}^2\langle H\rangle^2 M_1 & y_{\alpha 1}y_{s1}\langle\varphi\rangle\langle H\rangle M_2 + y_{\alpha 2}y_{s2}\langle\varphi\rangle\langle H\rangle M_1 \\ y_{\alpha 1}y_{s1}\langle\varphi\rangle\langle H\rangle M_2 + y_{\alpha 2}y_{s2}\langle\varphi\rangle\langle H\rangle M_1 & y_{s1}^2\langle\varphi\rangle^2 M_2 + y_{s2}\langle\varphi\rangle^2 M_1 \end{pmatrix}. \tag{6.1.5}$$

This matrix is real and symmetric. Therefore, it can be diagonalized by an orthogonal matrix $U \in O(2)$, parametrized in terms of a single mixing angle θ, by the following matrix

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{6.1.6}$$

In cases where the light mixing matrix is a simple, symmetric, two-dimensional matrix, one can obtain an explicit expression for the mixing angle. In particular, by writing out the matrix product $U^T\mathcal{M}_{\text{light}}U$ and demanding zero off-diagonal elements, the mixing angle can be expressed as (the superscript on the light mass matrix refers to components of the matrix)

$$\tan 2\theta = \frac{2\mathcal{M}^{12}_{\text{light}}}{\mathcal{M}^{22}_{\text{light}} - \mathcal{M}^{11}_{\text{light}}} = \frac{2y_{\alpha 1}y_{s1}\langle\varphi\rangle\langle H\rangle M_2 + 2y_{\alpha 2}y_{s2}\langle\varphi\rangle\langle H\rangle M_1}{y_{s1}^2\langle\varphi\rangle^2 M_2 + y_{s2}^2\langle\varphi\rangle^2 M_1 - y_{\alpha 1}^2\langle H\rangle^2 M_2 + y_{\alpha 2}^2\langle H\rangle^2 M_1}. \tag{6.1.7}$$

The mass eigenstates are the diagonal entries of $U^T\mathcal{M}_{\text{light}}U$, given in terms of the mixing angle as the linear combinations as

$$\begin{aligned} m_1 &= \cos^2\theta\mathcal{M}^{11}_{\text{light}} - 2\sin\theta\cos\theta\mathcal{M}^{12}_{\text{light}} + \sin^2\theta\mathcal{M}^{22}_{\text{light}} \\ m_2 &= \sin^2\theta\mathcal{M}^{11}_{\text{light}} + 2\sin\theta\cos\theta\mathcal{M}^{12}_{\text{light}} + \cos^2\theta\mathcal{M}^{22}_{\text{light}} \end{aligned} \tag{6.1.8}$$

6.2 Small active-sterile mixing

From the vev hierarchy, it is not hard to see that for Yukawa couplings of the same order ($y_{\alpha i} \sim y_{si}$), that the term which is quadratic in the Higgs vev is dominant. In fact, under this assumption, the mixing angle reduces to

$$\tan 2\theta \approx 2\theta \approx -\frac{2\langle\varphi\rangle}{\langle H\rangle}, \quad \langle\varphi\rangle \ll \langle H\rangle, \tag{6.2.1}$$

which is just the ratio of the vevs. This is the "vanilla" case of active-sterile neutrino oscillation and applications of this to other models involving a broken $U(1)$ is discussed briefly Section 6.5 Noteworthy, the singlet mass dependence cancels under the assumption that Yukawa couplings are similar. Thus, this result will hold for any configuration of the singlet mass scales when Yukawa couplings are similar. As the dark sector is broken at a much lower scale than the EW sector, the mixing is very small. The active-sterile neutrino transition probability is quadratic in the mixing angle, which means that the

effects are unobservable. For numerical purposes, consider the recurring example where the dark sector is broken at the MeV-scale ($\langle\varphi\rangle = 1\text{MeV}$), then the transition probability is proportional to

$$P(\nu_\alpha \to \nu_s) \propto \left(\frac{\langle\varphi\rangle}{\langle H\rangle}\right)^2 < 10^{-10}. \tag{6.2.2}$$

Although the numerical value obtained here is model specific, the ratio of the EW vev and the vev of a broken dark sector is a fundamental quantity. The vev ratio defines the ratio between the SM and the dark sector mass scales, as all particle masses are directly proportional to these vevs. Interestingly enough, by first assuming that there is some singlet mass scale that is much larger than both the EW scale and the dark scale, and then assuming there is a hierarchy between these, a double see-saw mechanism is realized [57]. To see this more explicitly, performing a first-order expansion of the mixing matrix in the small angle θ yields

$$U = \begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix} + \mathcal{O}(\theta^2) \simeq \begin{pmatrix} 1 & -\frac{\langle\varphi\rangle}{\langle H\rangle} \\ \frac{\langle\varphi\rangle}{\langle H\rangle} & 1 \end{pmatrix} + \mathcal{O}\left(\frac{\langle\varphi\rangle}{\langle H\rangle}\right)^2, \tag{6.2.3}$$

which up to an irrelevant sign on the off-diagonal, is the see-saw matrix $\mathcal{B}$ from (3.6.14)

$$\mathcal{B} = \begin{pmatrix} 1 & M_D^T M_R^{-1} \\ -M_D^{-1} M_D & 1 \end{pmatrix}, \quad M_D \propto \langle\varphi\rangle\langle H\rangle, \quad M_R \propto \langle H\rangle^2. \tag{6.2.4}$$

If for some reason the Yukawa couplings to the dark sector are much larger than the Yukawa couplings to the SM, then the relevant ratio

$$R = \frac{y_s\langle\varphi\rangle}{y_\alpha\langle H\rangle}, \tag{6.2.5}$$

might not be small, invalidating the above approach. A schematic visualization of the double see-saw mechanism is given in Figure 6.1.

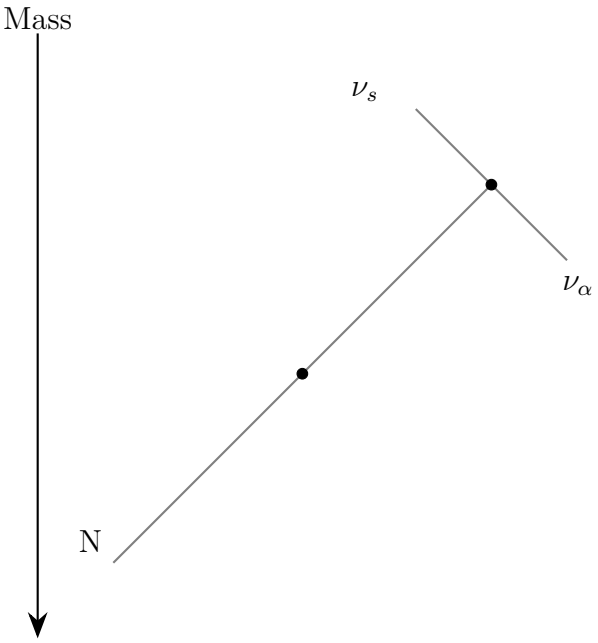

Figure 6.1: Illustration of the double see-saw mechanism. Note that mass increases downwards. N is meant to denote the singlet mass scale which by the standard see-saw mechanism provides small masses to the light neutrinos. At low-energy where the singlets are integrated out, a secondary see-saw mechanism is applied as the SM scale is much larger than the dark scale. If the dark scale is broken at a much higher energy than the SM scale, then the active and sterile neutrino would switch places on the upper see-saw. The latter is often used if sterile neutrinos are modeled as a DM candidate.

6.3 Large active-sterile mixing

Having seen that Yukawa couplings of the same order cannot accommodate large mixing due to the large hierarchy between the vevs, it is of interest to probe the Yukawa structure needed for large active-sterile neutrino mixing. Large mixing occurs whenever the denominator in the expression for the mixing angle (6.1.7) is small, which is whenever the vev-Yukawa ratio $R = \mathcal{O}(1)$. To see this, consider maximal mixing, i.e. when the denominator of (6.1.7) vanishes

$$\mathcal{M}^{11}_{\text{light}} = \mathcal{M}^{22}_{\text{light}}. \tag{6.3.1}$$

In the case of maximal mixing, the masses (6.1.8) reduce to

$$m_{2,1} = \mathcal{M}^{11}_{\text{light}} \pm \mathcal{M}^{12}_{\text{light}}. \tag{6.3.2}$$

If the off-diagonal elements are comparable to the diagonal elements, then this approach will yield m_1 close to zero and m_2 similar to two times each matrix element, comparable to the case of $(2\nu, 1N)$ mixing. Due to the structure of the light mass matrix, maximal mixing can only occur if

$$y_s \sim \frac{\langle H \rangle}{\langle \varphi \rangle} y_\alpha. \tag{6.3.3}$$

The finer structure between mass and mixing is determined by the interplay between the Yukawa couplings.

To further probe the structure of active-sterile mixing, a Monte-Carlo Markow-Chain (MCMC) parameter scan is performed. The MCMC method of choice is the Metropolis-Hastings algorithm (MHA), which is explained in the Appendix. The scan is performed in the four-dimensional Yukawa space $\mathcal{Y} = \text{span}\{y_{\alpha 1}, y_{\alpha 2}, y_{s1}, y_{s2}\}$ for fixed values of the singlet masses. That is, the singlet mass scale is taken as a scaling factor to obtain the desired values of light neutrino masses. For simplicity, the largest active neutrino mass is chosen, as it constrained by both the atmospheric mass splitting and the sum of neutrino masses. In the following discussion, let m_α be the mass of the active neutrino. Rewriting the expression for the sum of neutrino masses in terms of the lightest mass (3.4.6) in NO as a function of the heaviest mass, yields an upper bound on the heaviest mass eigenstate at $m_\alpha \leq 0.0584\text{eV}$. The lower bound on the heaviest mass eigenstate is obtained through the atmospheric mass splitting (3.4.2), $m_\alpha \geq \sqrt{\Delta m^2_{\text{atm}}} = 0.050\text{eV}$.

Furthermore, as large mixing is of interest, the probability distribution function (PDF) imposed on the mixing angle is chosen as a monotonic function of the absolute value of the mixing angle. Lastly, although the generic SBL transition amplitude requires a new mass splitting of one eV^2 [50], global fits of active-sterile mixing prefer slightly larger values. The best-fit value of the SBL mass splitting from two 2019 global analyses is $\Delta m^2_{s\alpha} = 1.3\text{eV}^2$ [49, 58]. For this reason, a PDF skewed to the right of the generic eV^2-scale mass splitting is considered.

In slight notational violation, the different PDF's are distinguished based on their argument. Furthermore, as the Metropolis-Hastings algorithm concerns only the ratio of PDF's, normalization constants are irrelevant as they cancel out when taking the ratio. Therefore, the following PDF's will not include any normalization factors. The PDF's used for the parameter scan are given as

$$P(m_\alpha) = \begin{cases} \epsilon, & \text{for} \quad m_\alpha < \sqrt{\Delta m^2_{atm}}, \\ 1, & \text{for} \quad \sqrt{\Delta m^2_{\text{atm}}} \leq m_\alpha \leq 0.058eV, , \\ \epsilon, & \text{for} \quad m_\alpha > 0.058eV, \end{cases} \tag{6.3.4}$$

where $\epsilon \neq 0$ is a small number taken positive to prevent dividing by zero when taking the ratio of PDF's. At first, the edges of the PDF for the active mass was taken as decaying Gaussian distributions. However, when normalizing these to be compatible with the 2σ uncertainties from global fits of the atmospheric mass splitting and the sum of neutrino masses, the standard deviation of the Gaussians were of order $\sigma^2 \sim 4 \times 10^6$. This resulted is an extremely narrow distribution, which accounting for the step size of the random walk, is therefore approximated by the sharp cut-off imposed in (6.3.4). As convergence of the active mass was not an issue, a more elaborate PDF was not chosen. Proceeding, the PDF for the SBL mass squared difference is given as a skewed Gaussian distribution

$$P(\Delta m^2_{s\alpha}) = G(\Delta m^2_{s\alpha}; \mu, \sigma)\Phi(\Delta m^2_{s\alpha}; \alpha), \quad \alpha, \mu, \sigma \in \mathbb{R}, \tag{6.3.5}$$

where μ is the mean, σ is the standard deviation and α is the skewness of the distribution. Further, $\Phi(\Delta m^2_{s\alpha}; \alpha)$ is the cumulative distribution function (CDF) of the Gaussian, skewed to either the side of the mean by a non-zero value of the skewness parameter α. A more detailed discussion on skewed Gaussian distributions is given in the Appendix. Lastly, the PDF for the mixing angle is taken as a linear function in the mixing angle

$$P(\theta) = |\theta|. \tag{6.3.6}$$

The parameter scan is performed in the subspace defined by the parameter values defined in the following Tables.

Parameter	Range
$y_{\alpha 1}$	$10^{-6} - 10^{-8}$
$y_{\alpha 2}$	$10^{-6} - 10^{-8}$
y_{s1}	$10^{0} - 10^{-2}$
y_{s2}	$10^{0} - 10^{-2}$

Table 6.1: Ranges of scanned parameters

Parameter	Value
α	4
σ	1
μ	1
ϵ	10^{-10}

Table 6.2: Values of PDF parameters

Parameter	Value
$\langle H \rangle$	$246GeV$
$\langle \varphi \rangle$	$10^{-3}GeV$
M_1	$10^{3}GeV$
M_2	$10^{3}GeV$

Table 6.3: Values of physical parameters

The two observables in SBL experiments are the active-sterile mixing angle and the mass squared difference. For each successful step of the random walk, the value of the observables are appended and their posterior distributions are illustrated in Figures 6.2 and 6.3.

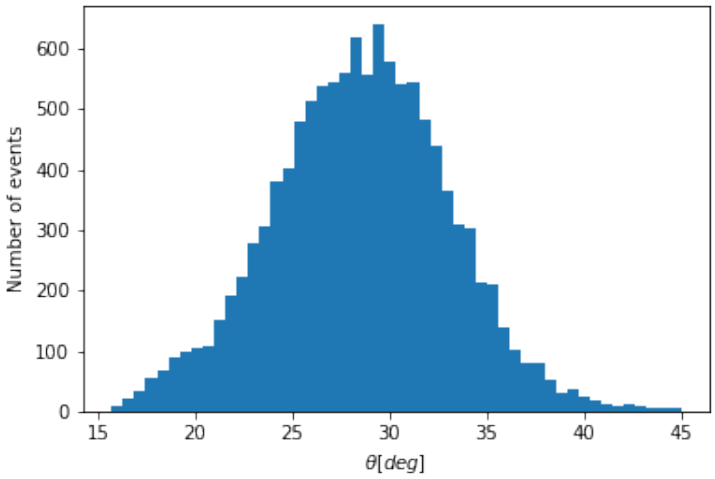

Figure 6.2: Histogram of the active-sterile mixing angle θ from the parameter scan.

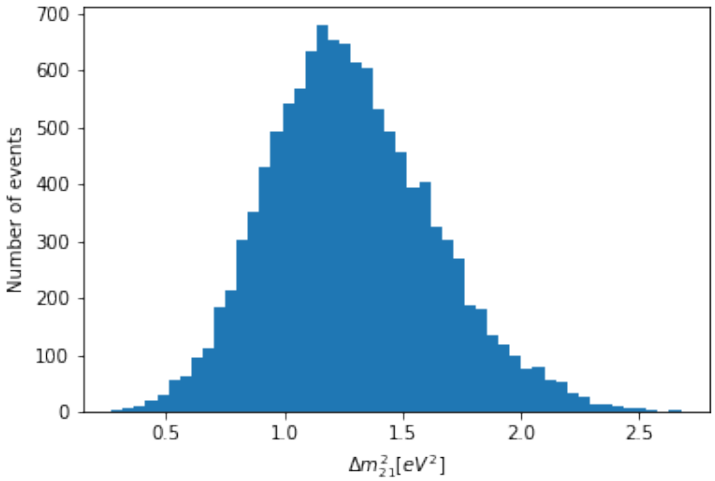

Figure 6.3: Histogram of the SBL mass squared difference $\Delta m^2_{21} = \Delta m^2_{s\alpha}$ from the parameter scan.

From Figure 6.2 it is evident that the posterior distribution of the mixing angle is a Gaussian about a mean value of about 30°. Notably, the posterior distribution does not resemble the proposed, linear prior. The posterior distributions are not required to be equal to the proposed, prior distributions. The prior distributions should rather be thought of as input, which the MCMC algorithm tries to replicate, if possible. Further, the PDF imposed on the mixing angle is not particularly strict, that is, for a slight increase in the mixing angle, the acceptance ratio only slightly increases. If the same step which increased the mixing moves the SBL difference[12] away from the mean, then as the PDF for the SBL difference is a Gaussian, the acceptance ratio will greatly decrease. It can be fruitful to think of the strict PDF as the dominant contribution to determining if a step is taken, and once the strict PDF has settled into a high-density region, then the less strict PDF's are allowed to dictate where the walk will continue.

Large mixing is due to the chosen regions of the Yukawa parameters, which include parts of parameter space where mixing can be maximal. Although a small subset of the sampled points corresponds to maximal mixing, the scan does not prefer maximal mixing. This is due to maximal mixing leaving the lightest mass eigenstate zero. Figure 6.3 illustrates that the posterior distribution of the SBL mass squared difference follows approximately the same skewed Gaussian distribution imposed on the prior distribution (6.3.5). However, the peak of the posterior distribution is shifted somewhat to the right, which indicates that the allowed parameter space prefers a slightly larger value of the SBL mass splitting. The remaining scan specific parameters can be found in Table 6.4.

Steps proposed (n)	50000
Steps accepted	11827
Efficiency	23.7%
Proposal distribution width	0.05
Burn-in period (b)	430

Table 6.4: Scan specific parameters

[12]The SBL mass squared difference will sometimes be abbreviated by the SBL difference.

The efficiency of any random walk is inversely proportional to the width of the proposal distribution and the latter should be seen as parameter tuned to reach a particular efficiency. After reaching the target distribution, the theoretically desired efficiency for an N-dimensional Gaussian converges to about 23% [59]. The probability distributions used are not N-dimensional Gaussian, but seeing that there are no other such bounds on the efficiency on random walks, a 23% efficiency was deemed reasonable. On the note of convergence, it brings up the last parameter listed in the Table 6.4. A burn-in period is the removal of the first b elements of the posterior distribution. A burn-in is required as the random walk is initialized at a random point in the parameter space, which may not reflect the desired model features. The random walk will eventually reach the target, but it will take some time before doing so. The sampled points corresponding to the low-density regions the walk has to traverse to arrive at the target does not encode any important information. Therefore, these points are removed from the posterior distribution to obtain better statistics and to remove any unwanted parameter values. There are no concrete rules for how long the burn-in process is, only general considerations. This is unless there is a way to figure out when the target distribution is reached. Luckily, for this particular scan, there is a way to obtain the burn-in length, b. Although not an SBL observable, the solution to burn-in length can be obtained from the active neutrino mass. Trace-plots of the active mass versus the number of successful iterations is illustrated in Figures 6.4 and 6.5.

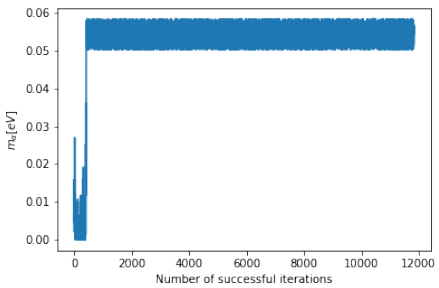

Figure 6.4: Trace-plot of the active neutrino mass versus number of successful iterations

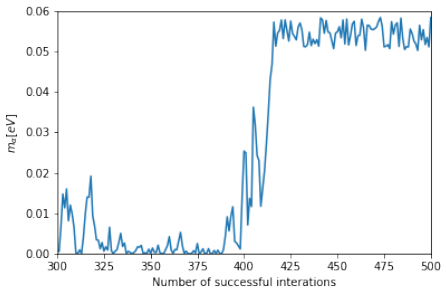

Figure 6.5: Snippet of the same trace-plot as Figure 6.4 in the interesting region.

Figures 6.4 and 6.5 illustrate the importance of burn-in periods when dealing with MCMC algorithms. The first 430 accepted points do not correspond to local maxima of the distribution. Hence, the burn-in period is taken as $b = 430$. The mixing angle enters the transition probability (4.2.8) as $\sin^2 2\theta$. Hence, the final illustration from this parameter scan is the scatter plot of this mixing term versus the SBL mass squared difference, which is illustrated in Figure 6.6.

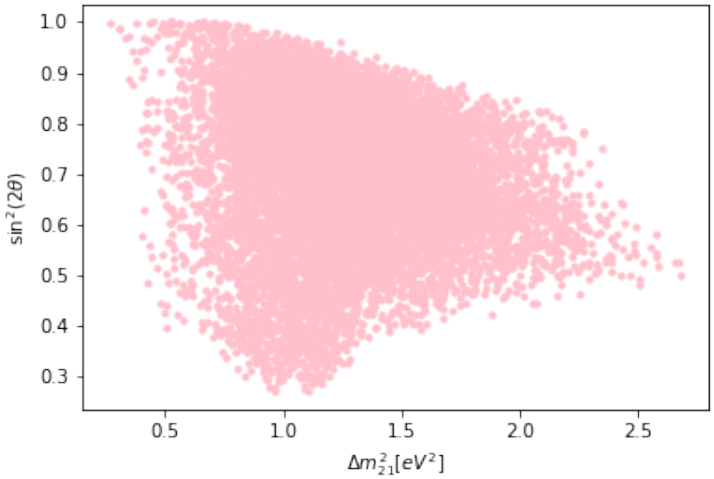

Figure 6.6: Scatter plot of the $\sin^2 2\theta - \Delta m^2_{21}$ plane from the parameter scan.

From Figure 6.6 one can conclude that there is no particular trend between mixing and the SBL mass squared difference in the selected part of the Yukawa parameter space. There seems, however, to be a lower bound on the mixing angle, $\sin^2 2\theta \gtrsim 0.3$. Large mixing is obtained as the parameter space was set up in the vicinity of points where the vev-Yukawa ratio is of order unity. From the get-go, large mixing was obvious due to the parameter space boundaries, but that the right masses were also obtainable was not obvious. This parameter scan illustrates how the model can incorporate potential large, observable mixing while confining the SBL mass squared difference to the eV^2-scale and the active neutrino within experimental bounds. That is, regions were found where all observables were within their desired ranges.

6.4 Moderate active-sterile mixing

The former parameter scan illustrated how one could go about locating regions of parameter space with large mixing, which is the most interesting scenario. However, regardless of how interesting a physical theory is, the most important quality a theory can inhabit is experimental support. To create a bridge between the unobservable, tiny mixing introduced by choosing all Yukawa couplings similar, and the experimentally rejected case of large, almost maximal mixing, a region somewhere in between these two extremes is of interest to locate. There are several methods to obtain such a region, but the simplest in respect to the recently introduced parameter scan is to alter the PDF for the mixing angle. The previous parameter scan was all about searching for regions of large mixing, which was attributed to a monotonic PDF in the mixing angle. If instead, the PDF is chosen as say a Gaussian, then the posterior distribution should be different than the one from the previous scan. The PDF's for the mass squared difference and the active neutrino mass are given by (6.3.5) and (6.3.4), respectively. Consider the following PDF for the mixing angle, given in degrees for easier visualization,

$$P(\theta) = G(\theta; \mu = 10, \sigma = 2). \tag{6.4.1}$$

In terms of the active-sterile mixing matrix element, this corresponds to $|U_{\alpha 2}| = 0.17$, which is comparable to the lowest active-active mixing element $|U_{e3}| = 0.15$. The posterior

distributions subject to burn-in are illustrated in Figures 6.7 and 6.8

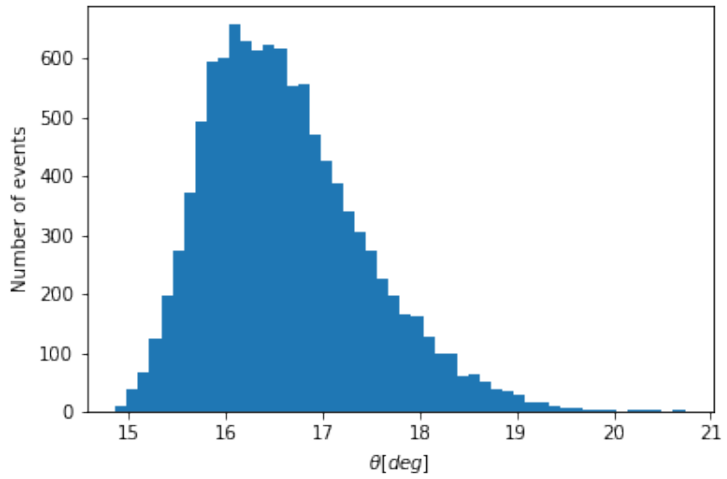

Figure 6.7: Histogram of the active-sterile mixing angle from the parameter scan opting to moderate the mixing angle.

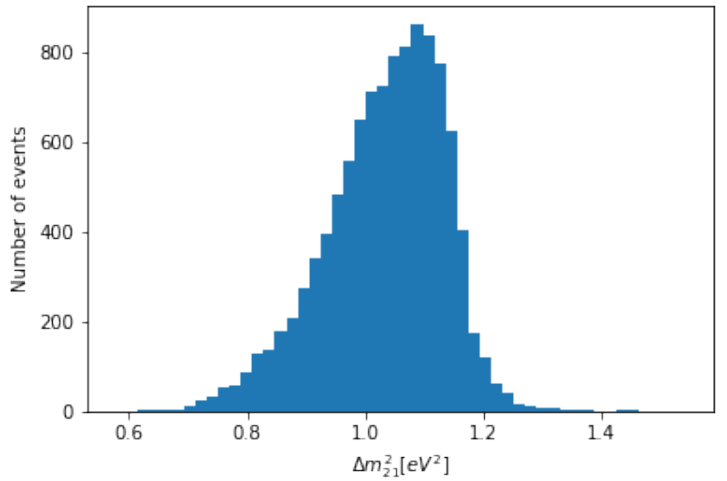

Figure 6.8: Histogram of the SBL mass squared difference from the parameter scan opting to moderate the mixing angle.

From Figure 6.7 it is clear that the parameter space is too narrow to include mixing of 10°. Imposing the Gaussian (6.4.1) has served a qualitative purpose of illustrating how (slightly over) moderate mixing can be achieved when the parameter space $\mathcal{Y}$ is unchanged. From Figure 6.8, the SBL difference is also shifted slightly to the left with respect to the prior distribution, although still in approximate agreement. This is likely due to the boundaries imposed on the parameter space as it was implicitly designed for large mixing and the singlet masses were chosen based on back-of-the-envelope calculations for eV-scale sterile neutrinos under the assumption of maximal mixing. A scatter plot of the mixing versus the SBL difference is illustrated in Figure 6.9.

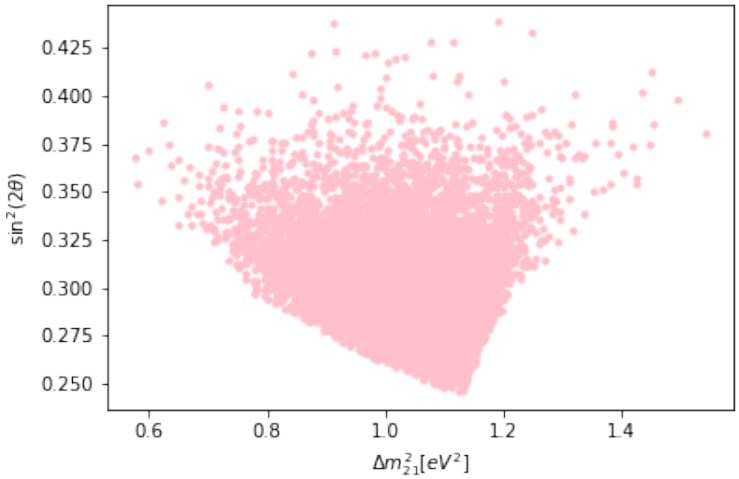

Figure 6.9: Scatter plot of the $\sin^2 2\theta - \Delta m^2_{21}$ plane from the parameter scan opting for a moderate mixing.

From Figure 6.9, a new feature is that the mixing is necessarily suppressed to smaller values due to the PDF imposed on it. Apart from the shift, there is no particular trend to this scatter plot. There is, however, a confirmation on the lower bound on the mixing angle obtained from the parameter scan for large mixing. That is, $\sin^2 2\theta \gtrsim 0.25$ which is approximately 15 degrees.

Although a simplification of full $3 + 1$ mixing, this section has illustrated several qualitative features of active-sterile neutrino mixing, which will hold also in the context of full mixing. The most important is the hierarchical structure of the Yukawa couplings.

6.5 Applications to other models

The elusive dark matter has long troubled the physics community and the particle content of the dark sector from which dark matter possibly originates is a mystery. Theorists have been at work devising models and experimentalists have been searching, however, neither has been successful. In trying to grasp dark matter or to study dark sectors as a whole, a vast number of models have been constructed. In particular, if the model at hand has quantum charges, such that the neutrino portal is open, then the general expressions outlined in this section should be applicable. A common feature of high-energy models is to assume Yukawa couplings of order unity such that the vev-Yukawa ratio R (6.2.5) reduces to the ratio of vevs. To put this into context, the detection of the unexplained 3.5keV X-ray line from the EM spectrum of several galaxy clusters [60, 61] is compatible with the decay of a sterile neutrino with mass $m_s = 7$keV. Following this observation, there have been several papers discussing the origin of this mysterious signal as a possible DM candidate. An interesting approach to explain the 3.5keV line is to assume there is a broken $U(1)$ symmetry from which these the keV-scale sterile neutrinos originate. Albeit from a supersymmetric standpoint, one can show that if the new symmetry is broken at the PeV scale, then the see-saw mechanism can both yield light active neutrinos and a keV-GeV sterile neutrino [62, 63]. Besides the technicalities of supersymmetry (SUSY), the relevant part of the super-potential, which is responsible for the see-saw mechanism,

includes two hierarchical terms which together form a non-diagonal mass matrix[13]. This matrix is then diagonalized equivalently as discussed at the start of this section. For additional examples showcasing how the neutrino sector can be used to probe hidden sectors through the (inverse) see-saw mechanism, see e.g. [64–66].

7 3+1 Active-sterile neutrino mixing

In the full picture of active-sterile oscillations, three active neutrinos mix with one dark neutrino and an unconstrained number of singlets, taken to be four for numerical purposes. In similar notation to $1+1$ mixing, the mixing configuration is $4\nu4N$. The Dirac mass matrix in the $\nu = (\nu_e, \nu_\mu, \nu_\tau, \nu_s)$ basis is given by (5.3.4) and the mixing matrix is given by (4.3.1).

7.1 The Casas-Ibarra parametrization

The simplicity of $1+1$ mixing does not carry over to the $3+1$ case. In particular, the matrix relations between observables and Yukawa couplings is somewhat tedious. A simple method of relating the unobservable Yukawa couplings to the observables of the theory is with the use of the Casas-Ibarra parametrization. Recall that the see-saw mechanism yields the following mass term in the Lagrangian

$$\mathcal{L} \supset -\frac{1}{2}\overline{\nu^c}\mathcal{M}_{\text{light}}\nu + h.c, \tag{7.1.1}$$

where as usual, the light mass matrix is given by $\mathcal{M}_{\text{light}} = -M_D^T M_R^{-1} M_D$. The mixing matrix U is by definition the matrix which diagonalizes the light mass matrix into a diagonal matrix of physical masses

$$U^T \mathcal{M}_{\text{light}} U = D_m = \text{diag}(m_1, m_2, m_3, m_4), \quad m_i \geq 0. \tag{7.1.2}$$

Up to field redefinition's to account for possible wrong signs of the light neutrino masses, the matrix structure which yields light neutrino masses is

$$D_m = U^T M_D D_M^{-1/2} D_M^{-1/2} M_D U, \quad D_M = M_R. \tag{7.1.3}$$

As the singlet mass matrix is diagonal, it can be decomposed as $D_M = D_M^{1/2} D_M^{1/2}$, where the square root is applied to each element of the diagonal. The above expression can be turned into an expression for the identity matrix by right-left multiplying by $D_m^{-1/2}$. This leads to

$$I = \left(D_M^{-1/2} M_D U D_m^{-1/2}\right)^T \left(D_M^{-1/2} M_D U D_m^{-1/2}\right) = R^T R, \tag{7.1.4}$$

for some generally complex matrix R. Rewriting in terms of the Dirac mass matrix, the Casas-Ibarra parametrization [67] is realized

$$M_D = D_M^{1/2} R D_m^{1/2} U^\dagger. \tag{7.1.5}$$

[13] A quick and rather effective method of deducing if a paper is written in a supersymmetric framework is to scout for excessive use of calligraphic symbols.

In their original paper [67], the parametrization is given in terms of the Yukawa matrix as only the three active neutrinos are considered. This means that the common vev can be divided upon. In the dark neutrino model, there are two vevs, hence dividing away the Higgs vev is of little use. This is, however, just a technicality and does not affect physics in any way. The Casas-Ibarra parametrization is useful as it relates the unobservable Yukawa couplings to the observables of the theory (light masses and mixing angles) up to singlet masses and the newly introduced matrix R. The price to pay for using this parametrization is that R is unconstrained apart from $R^T R = I$. In component form, this relation reads $\sum_k R_{ki} R_{kj} = \delta_{ij}$, which for $i = j$ translates to the inequality

$$|\delta_{ii}| = 1 = |\sum_k R_{ki} R_{ki}| \leq \sum_k |R_{ki} R_{ki}| = \sum_k \left|R_{ki}^2\right|, \quad R_{ij} \in \mathbb{C}. \tag{7.1.6}$$

For complex numbers, this means unconstrained matrix elements, as the square of a complex number can be negative, which may lead to an arbitrary degree of cancellation. If R is constrained to real numbers, then the above inequality would reduce to an equality. For this reason, and to be consistent with the previous assumptions of real Yukawa couplings, the matrix R is chosen to be real. In doing so, R is promoted to an orthogonal matrix, $R \in O(4)$.

Although R is some unknown matrix, it is also orthogonal. Therefore, it should not alter the magnitude of the matrices on which it acts, barring the possibility of cancellations. A particularly simple orthogonal matrix is the identity matrix, and choosing $R = I$ is equivalent to the simultaneous diagonalization of the Dirac and the singlet mass matrices. To see this, recall that the Dirac term in the Lagrangian, where the freedom in rotating the singlet basis is used up by choosing a real and diagonal singlet mass matrix, is

$$\mathcal{L} \supset -\overline{N} M_D \nu + h.c. \tag{7.1.7}$$

Written out with Casas-Ibarra parametrization for $R = I$, this reads

$$\mathcal{L} \supset -\overline{N} D_M^{1/2} D_m^{1/2} U^\dagger \nu + h.c. \tag{7.1.8}$$

By rotating the light neutrino basis with the mixing matrix ($\nu \to U\nu$), the diagonal Lagrangian is realized

$$\mathcal{L} \supset -\overline{N} D_M^{1/2} D_m^{1/2} U^\dagger U \nu + h.c = -\overline{N} D_M^{1/2} D_m^{1/2} \nu + h.c. \tag{7.1.9}$$

Hence, choosing $R = I$ simultaneously diagonalizes the Dirac and singlet mass matrices. In general, the choice of using a diagonal singlet basis and a non-diagonal Dirac basis, as opposed to the contrary, is a matter of choice and leads to the same physical predictions. This latter statement is true as physics is basis-independent.

7.2 Bounds on singlet masses

Choosing $R = I$ and using a diagonal singlet basis is not general, but it can be used to figure out different scales of the interesting matrices. In particular, from the perturbative bound on neutral couplings (3.6.12), the Casas-Ibarra parametrization provides an approximate upper bound on singlet masses through ($R = I$)

$$(M_D)_{ij} = \sqrt{M_i m_i} U_{ji}^*. \tag{7.2.1}$$

There is no sum over indices as both D_M and D_m are diagonal. A dark vev at the MeV-scale, severely limits how large singlet masses can be. The last column of the Dirac matrix and in particular the last element of that column is the crux to the bound on singlet masses. This is due to the sterile-sterile mixing matrix element being of order one, while the active-sterile mixing elements are much lower. To illustrate the approximate bound on the singlet mass scale, writing out $|(M_D)_{44}|$ yields

$$|(M_D)_{44}| = |y_{s4}\langle\varphi\rangle| = |\sqrt{M_4 m_4}U_{44}^*| \to M_4 = \frac{|y_{s4}|^2\langle\varphi\rangle^2}{m_4|U_{44}|^2}. \tag{7.2.2}$$

Using generic values; $m_1 = 1\text{eV}$, $|U_{44}| = 1$, $\langle\varphi\rangle = 1\text{MeV}$ and $|y_{s4}| \leq 4\pi$ yields the bound

$$M_4 \lesssim 16\pi^2 \times 10^3\text{GeV} = 1.58 \times 10^5\text{GeV}. \tag{7.2.3}$$

Although this bound is generally violated for $R \neq I$, the singlet scale will not change too dramatically. To show this, consider a general, orthogonal matrix R. In this case, the elements of the Dirac mass matrix are given as the sum

$$(M_D)_{ij} = \sum_l (D_M^{1/2})_{ii} R_{il} (D_m^{1/2})_{ll} U_{jl}^*. \tag{7.2.4}$$

In the following discussion and the rest of the thesis, the mixing matrix is taken to be real because SBL experiments cannot probe any CP structure related to complex numbers. As in the case of $R = I$, the bound on singlet masses arise from the elements of the last column on the Dirac matrix (as opposed to the last entry of the last column for $R = I$)

$$(M_D)_{k4} = y_{sk}\langle\varphi\rangle = \sqrt{M_k}\sum_l R_{kl}\sqrt{m_l}U_{4l}, \quad k \in (1,2,3,4). \tag{7.2.5}$$

The last term of the sum will dominate unless R_{k4} is a small number. To get an estimate of how small, the last term will no longer dominate once it is proportional to the other terms in the sum, which occurs whenever

$$R_{k4} \simeq \sqrt{\frac{m_{k'}}{m_4}}\frac{U_{k'4}}{U_{44}}, \quad k' \in (1,2,3). \tag{7.2.6}$$

For generic values of neutrino masses and constraints on active-sterile mixing elements, this above equation is realized for $R_{k4} \simeq 0.03$. This can be realized by handpicking for three out of four R_{k4}, but at least one of these has to be comparable to unity, as R_{k4} are elements of a column of an orthogonal matrix. Hence, it is hence subject to the orthogonality condition $\sum_k R_{k4}^2 = 1$. Therefore, at least one of the singlet masses is constrained by the bound obtained in (7.2.3).

To obtain knowledge about some unknown set of parameters, some statistical procedure is often employed, say R is sampled from a set of orthogonal matrices. The expectation value of random numbers a_i, subject to the constraint $\sum_{i=1}^4 a_i^2 = 1$, is $a_i = \pm 0.5$, which is what one would expect from the columns of an orthogonal matrix. Therefore, the bound given in (7.2.3) should apply to all singlets. A consequence of the bound on singlet masses is that active Yukawa couplings $y_{\alpha i}$ are confined to scales much below the dark Yukawa couplings y_{si}. This complements the results obtained in $1+1$ mixing.

7.3 Parametrization of 3+1 neutrino mixing

The full active-sterile mixing matrix is a four by four unitary matrix U, given in terms of six mixing angles, three Dirac phases and three Majorana phases. As with the three-dimensional PMNS/CKM matrices, it is useful to parametrize the four-dimensional mixing matrix also in terms of consecutive rotations in the (complex) four-dimensional flavour space. As discussed in Section 3.7, it is irrelevant which order the rotations are taken in, but a sequence of rotations that leaves ν_e mixing simple is usually preferred. This is particularly useful when it comes to SBL anomalies as these are dependent only on the $\nu_e - \nu_s$ and $\nu_\mu - \nu_s$ mixing angles, and their respective antiparticle counterparts. The latter mixing is only relevant when considering appearance experiments such as the LSND experiment, while reactor SBL anomalies are only sensitive to the former. As SBL experiments are insensitive to CP violation, the results obtained in this section will apply equally well to particles and antiparticles alike. The row corresponding to ν_τ mixing is usually left as the most convoluted as there are no good ν_τ sources, leaving scarce information on the elusive $\nu_\tau - \nu_s$ mixing angle. Further, it is also conventional to let the new rotations act on the PMNS matrix, which up to Majorana phases is given by $U_{\text{PMNS}} = R_{23}W_{13}R_{12}$. The mixing matrix parametrization used is the following unitary matrix

$$U = W_{34}R_{24}W_{14}R_{23}W_{13}R_{12}D(\varphi), \tag{7.3.1}$$

where $D(\varphi)$ is a diagonal matrix of three Majorana phases. As discussed, SBL experiments cannot probe and CP structure, so only the real, rotation part of the above mixing matrix is relevant. Therefore, the mixing parametrization reduces to the orthogonal matrix

$$U = R_{34}R_{24}R_{14}R_{23}R_{13}R_{12}, \quad U^T = U^{-1}, \tag{7.3.2}$$

determined completely by the six mixing angles. The (somewhat) explicit form is given as

$$U = \begin{pmatrix} c_{12}c_{13}c_{14} & c_{13}c_{14}s_{12} & c_{14}s_{13} & s_{14} \\ \cdots & \cdots & \cdots & c_{14}s_{24} \\ \cdots & \cdots & \cdots & c_{14}c_{24}s_{34} \\ \cdots & \cdots & \cdots & c_{14}c_{24}c_{34} \end{pmatrix}, \tag{7.3.3}$$

where the dots refer to long and convoluted expression of trigonometric functions. Using the parametrization given in (7.3.1), yields simple expressions for the ν_e mixing with the mass eigenstates. It also provides a simple form of the last column, which is the coupling of the flavour neutrinos to the fourth eigenstate. In particular, the mixing between ν_e and ν_s is given by a single term, $U_{e4} = \sin\theta_{14}$.

In the $1+1$ mixing scenario, both the masses and the mixing angle could be derived in terms of simple functions, as the mixing matrix was composed of a single rotation in a two-dimensional, flavour space. This does not generalize to larger mixing scenarios. Although, $R_{12}^T\mathcal{M}_{\text{light}}R_{12}$ does set $(\mathcal{M}_{\text{light}})_{12}$ and $(\mathcal{M}_{\text{light}})_{21}$ zero for a particular value of $\tan 2\theta_{12}$, this zero will become nonzero when the rest of the rotation matrices act on it. This hinders any attempts at an analytical approach for obtaining mixing angles in terms of light mass matrix elements. That is, mixing angles and light mass matrix elements are treated as independent quantities. A caveat to this, is that parameter scans are generally ineffective as there are no dependencies to exploit in say an MCMC approach.

7.4 3+1 global reactor fits

Having introduced both the Casas-Ibarra parametrization of the Dirac mass matrix, as well as the standard $3+1$ parametrization of the mixing matrix, it is time to verify the $3+1$ model in light of global fits. As eV-scale sterile neutrinos are characterised by SBL experiments, it is natural to consider fits from said experiments. In particular, the $3+1$ global fit of reactor disappearance data [68] from 2019 is used. In their analysis [68], they used data from the Karlsruhe Tritium Neutrino (KATRIN) experiment, which is the leading experiment on direct observation of neutrino masses from a sensitivity standpoint. KATRIN is sensitive to effective neutrino masses ν_β down to the eV-scale [69], which is the range where eV-scale sterile neutrino enter. The KATRIN experiment relies on measurements of the spectrum of electron energies from tritium decay

$$^3\mathrm{H} \rightarrow ^3\mathrm{He}^+ + e^- + \overline{\nu}_e, \tag{7.4.1}$$

in which the neutrino observable is the effective ν_e mass, given in the 3+1 scenario as

$$m_\beta^2 = \sum_{k=1}^{3} \frac{|U_{ek}|^2}{1-|U_{e4}|^2} m_k^2. \tag{7.4.2}$$

Note that to return to the three neutrino picture, simply set $U_{e4} = 0$. Additionally, as the masses m_k are not known, the above expression has to be rewritten in terms of the neutrino mass splittings whose 2020 values are listed in Table 3.2. In their analysis [68], they used the 2019 best-fit values of the mass splittings, but the difference between the 2019 and 2020 values are marginal. In addition, they show that while accounting for the sensitivity of KATRIN, the effective neutrino mass reduces to $m_\beta^2 \simeq m_1^2$. Therefore, any marginal update on the BF values of neutrino mass splittings is irrelevant. The contours corresponding to overlapping regions from various reactor experiments subject to KATRIN data are illustrated in Figure 7.1. Reactor anomalies are disappearance experiments $P(\nu_e \rightarrow \nu_e)$, and the notation used in [68] is the same as introduced in (4.2.11) on effective mixing angles. Using the parametrization of the mixing matrix introduced in (7.3.3) the relation between the effective mixing angle and the angle θ_{14} is trivial, $\sin^2 2\theta_{ee} = \sin^2 2\theta_{14}$.

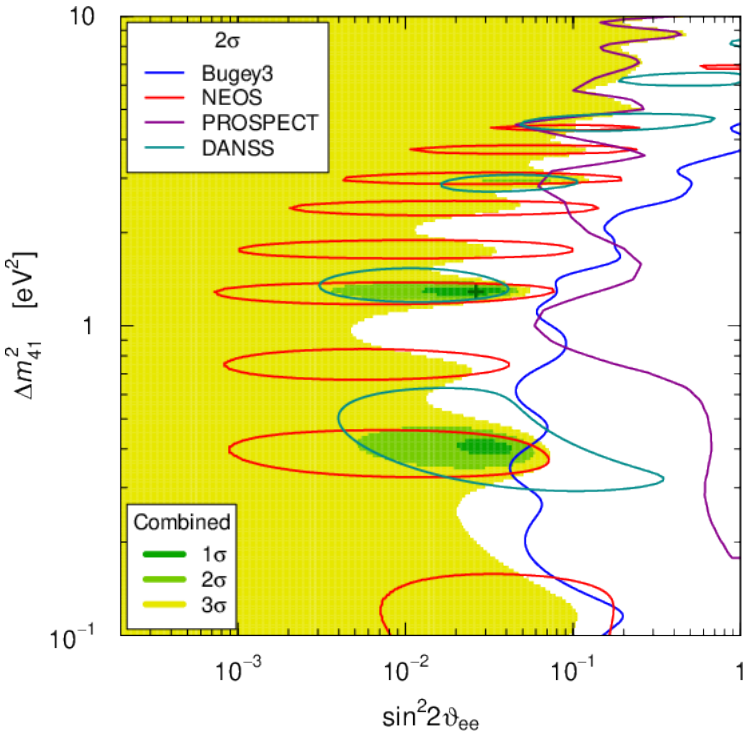

Figure 7.1: Contours of 2σ regions in the $\sin^2 2\theta_{ee} - \Delta m^2_{41}$ space from a collection of reactor spectral measurements, taken from [68]. The best-fit point is indicated by the black cross.

From Figure 7.1 there are three regions which are both to the left of the Bugey3 and PROSPECT exclusion curves and within the overlapping contours of NEOS and DANSS. The best-fit point is indicated with a black cross and corresponds to the pair of best-fit values $\Delta m^2_{41} = 1.3\text{eV}^2$ and $\sin^2 2\theta_{ee} = 0.026$. The other two regions lay approximately on the same vertical line, i.e. they have the same mixing, but for mass splittings, $\Delta m^2_{41} \approx 0.4\text{eV}^2$ and $\Delta m^2_{41} \approx 3\text{eV}^2$ for the lower and upper region, respectively.

Seeing that the orthogonal matrix R from the Casas-Ibarra is unknown, at first, to compare Dirac matrix elements corresponding to each of these overlapping regions, R is taken as the identity matrix. As R is orthogonal, this particular choice should not change the overall scale. Furthermore, as reactor SBL experiments are insensitive to $\theta_{\mu 4}$ and $\theta_{\tau 4}$, these are taken to zero. It will become clear later that regardless of what these mixing angles are, the overall scale does not change. As usual, the singlet basis in unknown up to the approximate bound of 10^5GeV, but to get a numeric value, the basis $M_1 = 10^3\text{GeV}, M_2 = M_3 = M_4 = 10^4\text{GeV}$ is chosen. Choosing M_1 a magnitude lower than the other singlet masses is arbitrary, but serves as an important technicality for leptogenesis bounds, which will be discussed in Section 8.3. Lastly, the independent mass eigenvalue m_1 is taken to be maximal $m_1 = 3 \times 10^{-2}\text{eV}$. This value is obtained from the cosmological bound on the sum of neutrino masses in NO. Again, this is a particular value of a particular hierarchy, but varying the independent active mass eigenstate and changing hierarchies does not affect the scale, as these are marginal. The Dirac mass matrix corresponding to the best-fit values from [68] is

$$\frac{M_D}{\langle\varphi\rangle} = \begin{pmatrix} 0.14 & -0.08 & 0.06 & -0.01 \\ 0.30 & 0.27 & -0.38 & -0.02 \\ 0.11 & 0.57 & 0.49 & -0.009 \\ 0.27 & 0 & 0 & 3.37 \end{pmatrix}, \tag{7.4.3}$$

while for the other two regions, the results are very similar and hence not shown. The matrix is divided by the dark vev such that the dark Yukawa couplings can be read straight from the last column. As expected, when $R = I$, the issue is the last entry of the Dirac matrix, is closing in on the perturbative bound. The two zeros in the last row are constructs of choosing $\theta_{\mu 4}$ and $\theta_{\tau 4}$ zero. Furthermore, the first three entries of the fourth column are small, as the fourth flavour neutrino couples only scarcely to the first three mass eigenstates. This is reflected by the PMNS matrix being approximately unitary, leaving only room for small $|U_{sk}|, k \in (1, 2, 3)$. The numerical values of active Yukawa couplings are found by multiplying the first three columns of the Dirac matrix with the vev ratio $\langle\varphi\rangle/\langle H\rangle$. This is quantitatively the same result as obtained from $1 + 1$ mixing; due to the large difference in vevs, the Yukawa structure has to compensate by taking on a hierarchical structure. Furthermore, in the case of $R = I$, which as shown earlier is equivalent to the simultaneous diagonalization of both the singlet and the Dirac mass matrices, the Yukawa couplings satisfy the following hierarchy

$$|y_{\alpha i}| \ll |y_{sj}| \ll |y_{s4}|, \quad i \in (1, 2, 3, 4), \quad j \in (1, 2, 3) \tag{7.4.4}$$

In the more general case of an orthogonal matrix R, the dominant term responsible for the large Dirac matrix element $(M_D)_{44}$, will mix into the last column of the Dirac matrix. Thus, turning all dark sector Yukawa couplings similar. It is instructive to consider an example of the effects of a randomly generated orthogonal matrix R. For a numerical example, consider R drawn from a set of orthogonal matrices

$$R = \begin{pmatrix} 0.52 & -0.56 & -0.51 & 0.38 \\ 0.16 & 0.72 & -0.12 & 0.66 \\ 0.81 & 0.29 & 0.16 & -0.48 \\ 0.19 & -0.29 & 0.83 & 0.43 \end{pmatrix}, \quad \frac{M_D}{\langle\varphi\rangle} = \begin{pmatrix} 0.04 & -0.13 & 0.07 & 0.52 \\ 0.50 & 0.36 & -0.04 & 2.73 \\ 0.31 & -0.23 & -0.08 & -2.00 \\ 0.23 & 0.53 & 0.73 & 1.67 \end{pmatrix}, \tag{7.4.5}$$

which makes it clear that introducing an orthogonal matrix washes away the dark Yukawa hierarchy defined in (7.4.4). By including a non-diagonal R, the zeros of (7.4.3) become nonzero, further smoothing out the structure of the Dirac matrix. Although the elements of the Dirac matrix are rather uniform, elements of the last column will generally be slightly larger than the other columns due to them obtaining a term $(M_D)_{k4} \propto \sqrt{m_4} U_{44}$, which will always dominate. Although some of the active-active mixing angles are large, e.g. $|U_{e1}| \approx 0.8$, they appear alongside active neutrino masses as shown in (7.2.4), which are much smaller than m_4. Therefore, the general structure of the Dirac mass matrix is somewhat uniform, while promoting slightly larger values for the last column.

7.5 Structure of the mass matrices

As mentioned in the previous section, changing the mixing angles and light neutrino masses does not alter the structure of the Dirac matrix much. The only parameters which can significantly change the Dirac matrix are the singlet masses, but they change

the overall scale and cannot be tuned to only increase the active Yukawa couplings. The last resort to obtain larger active Yukawa couplings is to attempt to fine-tune the values of the last column of the Dirac matrix, such that the singlet masses can be taken large enough to provide large active Yukawa couplings. However as discussed in Section 7.2, this is not possible as it requires that all elements of an orthonormal column are small ($\ll 1$), something orthogonal matrices cannot do. Therefore, the structure of Yukawa couplings and the Dirac matrix elements are required to satisfy

$$y_{\alpha k} \lesssim \frac{\langle\varphi\rangle}{\langle H\rangle} y_{sk}, \quad (M_D)_{ij} \lesssim (M_D)_{k4}, \quad j \in (1,2,3), \quad i,k \in (1,2,3,4), \tag{7.5.1}$$

regardless of chosen mass squared difference hierarchy. Furthermore, as mentioned earlier, the dark Yukawa couplings have to overcompensate past the vev ratio to yield the desired light neutrino masses. An observation related to this result is that there is some dominant term arising from the dark sector, which washes away any structure, a kind of inverse fine-tuning, as it dominates regardless of how tuned the whole sector is.

This section has shed light on the structure of the Dirac mass matrix by using bounds on singlet masses, experimental bounds on mixing angles and active neutrino masses, and finally using the SBL mass squared difference to yield the sterile neutrino mass. The last piece of study is the symmetric, light mass matrix $\mathcal{M}_{\text{light}}$. Given in terms of the mixing matrix and the diagonal matrix of light neutrino masses $\mathcal{M}_{\text{diag}}$, the light mass matrix can be expressed as

$$\mathcal{M}_{\text{light}} = U\mathcal{M}_{\text{diag}}U^T, \quad (\mathcal{M}_{\text{light}})_{ij} = \sum_k U_{ik} m_k U_{jk}, \tag{7.5.2}$$

where m_k are the mass eigenstates. Note that each element of the light mass matrix acquires a term which is proportional to the fourth mass eigenstate. In particular, the active part of the light mass matrix, which is the upper 3×3 block, will each obtain a term $(\mathcal{M}_{\text{light}})_{ij} \supset U_{i4} m_4 U_{j4}$. Although this is quadratically suppressed in terms of active-sterile mixing elements, it is also directly proportional to the heaviest mass eigenvalue.

To double down on the earlier statement, that there is a term arising from the dark sector which dominates the structure of the light mass matrix, and to prove that there is not any fine-tuning in the light sector, a scan of fine-tuning is performed. An added benefit of checking for fine-tuning is that the bound on singlet masses from leptogenesis, which will be given in Section 8.3 is valid. Informally, a fine-tuned parameter is very small relative to its constituents. Say an observable B is the sum of some terms $B = \sum_i B_i$, then a working definition of fine-tuning is defined as

$$FT(B) \equiv \frac{\max(|B_i|)}{|B|}. \tag{7.5.3}$$

More elaborate approaches to fine-tuning may be defined in terms of how changing model parameters change the mass of the Z boson using Bayesian interference, but this is beyond the scope of this thesis. The interested reader is referred to [70]. Regardless of the approach, a large value of the fine-tuning parameter is generally considered undesired.

To obtain graphs, some model parameter has to be chosen as the evolution parameter. In this case, a natural choice is m_1 for the following reason: Given the value of the lightest mass eigenstate, the others follow. In particular, given a value of m_1 assuming normal ordering, then m_2 and m_3 follow from the solar and atmospheric mass splittings,

respectively. In addition, m_4 follows from the SBL mass splitting. Hence, m_1 is chosen as the evolution parameter of fine-tuning for components of the light matrix. The solar and atmospheric mass splittings are varied in their 3σ ranges given in Table 3.2.

This approach, however, means that the mixing angles have to be chosen. A reasonable choice is to simply use best-fit values of the PMNS angles, while for θ_{14}, the best-fit value from the global fit is chosen. This leaves the two remaining active-sterile angles. Two cases are chosen, the first in which these angles are zero, and in the second case they are taken to their maximum values, corresponding to $|U_{\mu 4}| = |U_{\tau 4}| = 0.1$. As the light mass matrix $\mathcal{M}_{\text{light}}$ is symmetric, it has ten independent entries. To avoid messy plots, only the three worst offenders of the fine-tuning measure are included, i.e. the largest values of FT. In the following, let $FT(M_{ij}) \equiv FT_{ij}$.

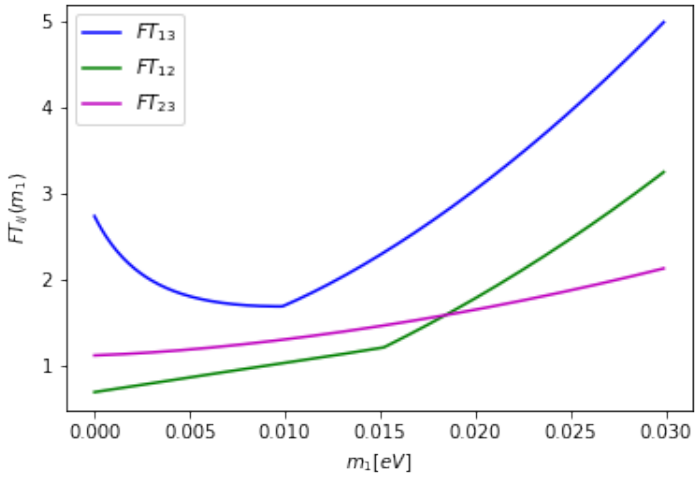

Figure 7.2: The largest fine-tuning versus m_1 for $U_{\mu 4} = U_{\tau 4} = 0$.

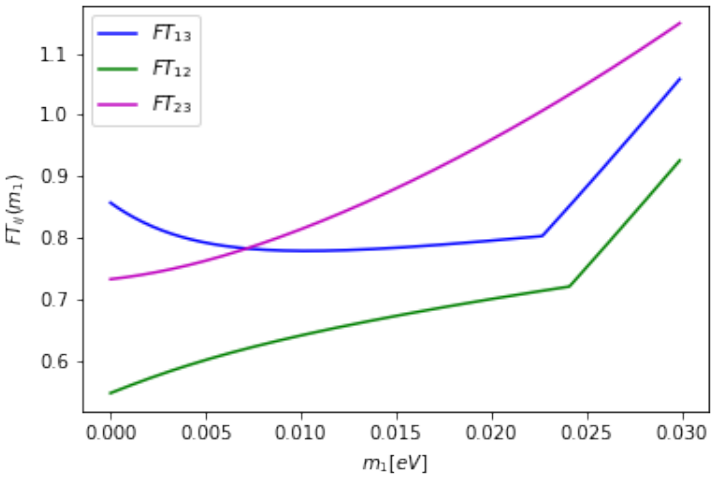

Figure 7.3: The largest fine-tuning versus m_1 for $|U_{\mu 4}| = |U_{\tau 4}| = 0.1$.

From the above figures, it is evident that there is no fine-tuning in the light sector. There is a difference in the overall scale when active-sterile mixing is turned on as illus-

trated in these figures. This is expected, as the active-sterile mixing angles couple to the large fourth mass eigenstate. The fact that there is no tuning in the light mass matrix means that light neutrino masses are not accidentally small due to cancellations from a higher scale.

This section has shown that the mass matrices needed to explain the SBL anomaly are rather bland, i.e. they do not have a very detailed structure apart from the last column on the Dirac matrix obtaining slightly larger values. The Yukawa matrix is necessarily hierarchical to account for the hierarchical vev structure, but when combined the hierarchies cancel out. Furthermore, there is no tuning in the sector, which means that everything acts as one would expect.

8 The role of right-handed neutrinos

8.1 A biased universe

The discovery of neutrino oscillations and neutrino masses is without a doubt the clearest hint towards new physics, as it is a direct violation of the Standard Model. Neutrino masses and their smallness can be elegantly explained by the see-saw mechanism, which has been the assumption throughout this thesis. A large part of the discussion has been on the mixing between active and dark neutrinos, where the singlets have merely acted as bystanders, a scale factor with no dynamics. Generally, singlets are much too heavy to explain the apparent SBL anomalies, but their large mass may hold the key to a more fundamental question; Why does the universe prefer matter over antimatter? From cosmological observations, it is clear that the universe we live in is composed of matter. Stars, galaxies, cluster of galaxies, all made of matter, not antimatter. Why this is, is a peculiar question, to which there is no answer, yet. A conventional measure of the matter-antimatter asymmetry in the universe is the asymmetry parameter [71] defined as

$$\eta_B \equiv \frac{n_B - n_{\overline{B}}}{n_\gamma} = (6.12 \pm 0.04) \times 10^{-10}. \tag{8.1.1}$$

One might ask, what if the universe simply started with more matter than antimatter i.e. the observed baryon asymmetry is merely a consequence of an initial condition. The answer to this is, no [72, 73]! Another seemingly plausible solution to baryon asymmetry is to assume that we are in a patch of the universe made of matter, and that outside our patch, there are patches composed of antimatter, where someone may be writing this very anti-thesis. However, if this was the case, then at the boundary between regions of matter and antimatter, a significant amount of radiation would be released when matter annihilates with antimatter. Cosmological observations exclude this possibility.

Models which set their eyes on describing the baryon asymmetry are collectively called models of *baryogenesis*. Leptogenesis models [74] are subsets of baryogenesis models, where the asymmetry is generated by the addition of lepton number violation, right-chiral neutrinos, which can decay into leptons and antileptons.

A physical theory is deemed good based on the number of predictions it can make. Therefore, if the addition of right-chiral neutrinos can solve both the origin of neutrino masses and explain why the observed universe is matter-dominated, then that model should be pursued over models which can explain only one of the mentioned problems.

The goal is thus to figure out how the introduction of a dark sector interplay with leptogenesis through the singlets. Now it is time for the dark sector to play a supporting role, to spectate how the dynamics of singlets in the hot, early universe can yield the observed universe we see today.

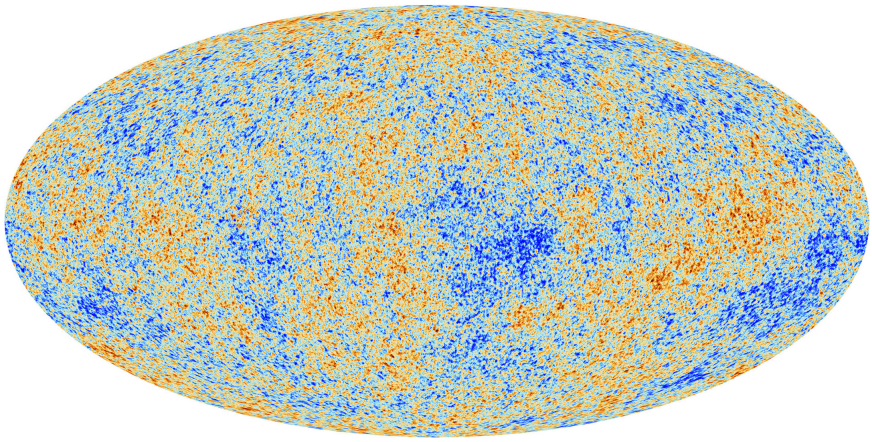

Figure 8.1: The Cosmic Microwave Background (CMB); the first light created in the universe after it became transparent 3.8×10^5 years ago. The temperature fluctuations of the CMB, according to the theory of inflation, are macroscopic reflections of quantum fluctuations that occurred when the universe was less than 10^{-32} seconds old. This image is taken from the Planck Collaboration [75]

8.2 The Sakharov conditions

In his seminal paper [76], Sakharov proposed three conditions a theory would need to exhibit to explain the observed baryon asymmetry of the universe:

1. Baryon number violation.
2. C and CP violation.
3. A departure from thermal equilibrium.

These are known as the Sakharov conditions. Baryon number violation is obvious, as, without it, there would be no baryon asymmetry. C and CP violation is needed such that the sum of all interactions produce more matter than antimatter, which is not washed away by opposite interactions. The last condition of a departure from thermal equilibrium is imposed to prevent CPT from washing away any matter-antimatter imbalance.

All of these conditions are met by the Standard Model. Baryon number is violated by the chiral anomaly, C and CP are violated by e.g. the phase in the CKM matrix. Finally, a departure from thermal equilibrium can be realized by EW symmetry breaking while the universe cools. EW symmetry breaking is a phase transition where matter-antimatter symmetry can be generated. Lengthy calculations show, however, that the SM does not violate baryon number enough, and the EW phase transition is not strong enough to generate the observed matter-antimatter asymmetry.

8.3 The singlet masses required

As stated earlier, leptogenesis is a model of baryogenesis where the asymmetry originates from the lepton sector. To obtain a baryon asymmetry starting from a lepton asymmetry, there must be some process that converts leptons to baryons. It is known within the Standard Model that both lepton and baryon number are violated by the chiral anomaly. Let B be baryon number and L be lepton number, then the corresponding non-conserved currents related to $U(1)_B$ and $U(1)_L$ are given as [77]

$$\partial_\mu J^\mu_B = \partial_\mu J^\mu_L = \frac{n_g g^2}{32\pi^2}\varepsilon^{\mu\nu\alpha\beta}W^a_{\mu\nu}W^{a\alpha\beta}, \tag{8.3.1}$$

where $n_g = 3$ is the number of generations and $\varepsilon^{\mu\nu\alpha\beta}$ is the Levi-Civita tensor. Although both B and L are violated by the anomaly, their difference $B - L$ vanishes. The conservation of $B - L$ is broken by the Majorana mass term, which violated L by two units, but preserves B. A possibility is that $B - L$ is a symmetry of its own, and that it is broken at the GUT-scale and that right-chiral neutrinos obtain their mass from this broken symmetry. The addition of a gauged $U(1)_{B-L}$ is considered in e.g. [78], while the detectability the particles arising from this group can be found in [79].

The lepton number violation created by right-chiral neutrinos can be translated into a baryon number violation in terms of a $B - L$ conserving, but $B + L$ violating *sphaleron* processes. A sphaleron process is the conversion of three baryons into antileptons or three antibaryons into leptons, however, this requires temperatures of order $T_{\text{Sph}} \sim m_W/\alpha_W$ to become relevant. For sphaleron processes to produce the observed baryon asymmetry of the universe, there must have been a lepton asymmetry at temperatures $T \sim T_{\text{Sph}}$. This can occur if the decay channels of right-chiral neutrinos are CP-violating following the Sakharov conditions. Outside thermal equilibrium, a lepton asymmetry may be formed, directly proportional to the CP parameter

$$\varepsilon_i = \frac{\Gamma(N_i \to lH) - \Gamma(N_i \to \bar{l}H^\dagger)}{\Gamma(N_i \to lH) + \Gamma(N_i \to \bar{l}H^\dagger)}, \tag{8.3.2}$$

which is the difference of decays into particles versus antiparticles, normalized to the total decay width. For a hierarchical spectrum of singlet masses, say $M_1 \ll M_2, \dots M_N$, one can using standard values of the CP parameter, show that the lightest singlet mass required for leptogenesis is $M_1 \gtrsim 3 \times 10^9$GeV. This is under the assumption that there is no fine-tuning in the light sector [80]. The condition of no fine-tuning in the light sector was shown in Section 7.5, hence the bound can be applied. Following the upper bound on singlet masses derived in (7.2.3), which was related to the perturbative upper limit on Yukawa couplings, it is clear that the lightest singlet mass is not sufficiently heavy enough.

The former, standard approach is based on the assumption that singlet masses are hierarchical. If instead, at least two of the singlet masses are nearly degenerate $(M_1 + M_2)/|M_1 - M_2| \ll 1$, then the regime is called resonant leptogenesis [81, 82], and the singlet masses required for leptogenesis drops significantly to $M_1 \simeq M_2 = \mathcal{O}(\text{TeV})$ [82]. The price to pay for TeV-scale singlet masses is the inevitable tuning of the singlet sector. Coincidentally, singlet masses of order 1TeV is the scale needed to produce dark Yukawa couplings $y_{si} = \mathcal{O}(1)$, see e.g. (7.2.4). From the discussion of the bound on singlet mass scale in Section 7.2, only an upper bound is achieved. Following the discussion in that section, one might consider much lower singlet masses, but the price to pay for lowering

the singlet masses is a more severe smallness problem in the Yukawa sector. Although this is possible, it is certainly not an attractive feature of the model. Therefore, it is refreshing that resonant leptogenesis can provide a lower bound on singlet masses, hence confining the singlet mass scale to

$$1\text{TeV} \lesssim M_i \lesssim 10^2\text{TeV}. \tag{8.3.3}$$

This bound justifies the chosen, numerical values used in the parameter scans in Sections 6.3 and 6.4, and in reproducing the global fit using in Section 7.4. Furthermore, this bound necessarily confines the dark Yukawa couplings to large values $y_{si} \gtrsim 1$, while the active Yukawa couplings are suppressed with respect to the vev ratio $y_{\alpha i} \sim y_{si}\langle\varphi\rangle/\langle H\rangle$.

8.4 The 3+1 global fit revisited

With the addition of leptogenesis to the theory, singlet masses are confined to a narrow range. This fixes the last free parameters of the theory, barring the Yukawa couplings and the orthogonal matrix R, see e.g. (7.1.4). In light of the new constraint on the singlet scale, the Yukawa parameter space is now sufficiently small to be scanned. However, as the Yukawa couplings are not observables, no PDF's are chosen for them. Hence, the scan is blind. The scan is performed using the Casas-Ibarra parametrization given in (7.1.5), where singlet masses are confined to the range (8.3.3). In particular, the two lightest singlet masses, say M_1 and M_2 are taken degenerate. The scan is performed in normal ordering where m_1 is the lightest neutrino mass and is drawn from the uniform distribution

$$m_1 = \mathcal{U}(0, 3 \times 10^{-2})\text{eV}. \tag{8.4.1}$$

The remaining active neutrino masses are expressed from m_1 with the use of the observed mass splittings within the 3σ listed in Table 3.2. The fourth mass eigenstate and the active-sterile angle θ_{e4} are taken as the best-fit values of the global fit [68]. Furthermore, the PMNS angles are taken within their respective 3σ regions (3.4.8), while the remaining two active-sterile angles are taken to respect the generic bounds $|U_{\mu 4}|, |U_{\tau 4}| \leq 0.1$. Lastly, R is drawn from a set of random orthogonal matrices.

The Yukawa couplings form a 16-dimensional space, a bit large to project onto a two-dimensional paper/screen. In addition, the Yukawa couplings may be of both signs. Therefore, the averages of the absolute values of Yukawa couplings are introduced

$$\langle y_\alpha \rangle = \frac{1}{12}\sum_{i=1}^{4}\sum_{j=1}^{3}|Y_{ij}|, \quad \langle y_s \rangle = \frac{1}{4}\sum_{k=1}^{4}|Y_{k4}|. \tag{8.4.2}$$

In addition, the ratio of the vevs is defined as

$$\beta \equiv \frac{\langle H\rangle}{\langle\varphi\rangle} = 2.46 \times 10^5. \tag{8.4.3}$$

The scan is performed by considering 10^4 draws from the distributions in the outlined regions discussed above. The results from two example scans for different singlet mass configurations are illustrated in Figures 8.2 and 8.3.

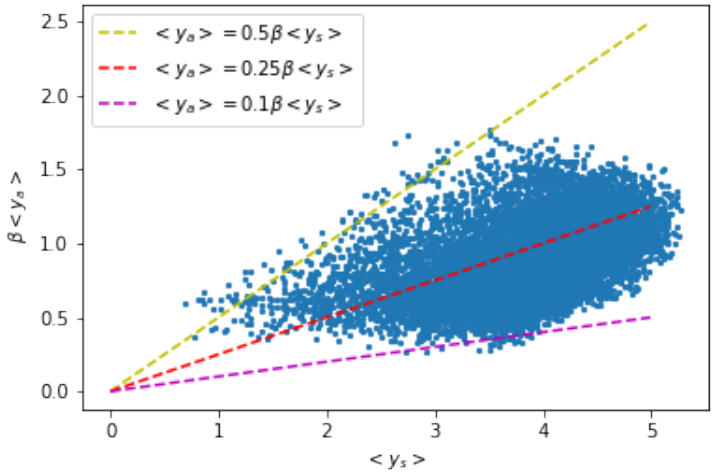

Figure 8.2: The parameter space of 3+1 global mixing illustrated in the $\langle y_s \rangle - \langle y_\alpha \rangle$ plane. The dashed lines represent lines of symmetry related to how the distribution varies relative to the vev ratio β. This scan is performed with $M_3 = M_4 = 100\text{TeV}$ and $M_1 = M_2 = \mathcal{U}(1, 100)\text{TeV}$

From this figure, it is evident that the parameter space of $3 + 1$ mixing, subject to the global fit is once again hierarchical. From the symmetry lines, one can observe that the parameter space is approximately symmetric about the red line, which means that the dark Yukawa couplings overcompensate by a factor of four relative to the vev ratio β. This is due to the fourth mass eigenstate outweighing the other mass eigenstates by a good margin. This parameter scan included the maximal values obtained from the bound on singlet masses (8.3.3) and is in good agreement with the perturbative upper bound on Yukawa couplings (3.6.12). As Yukawa couplings are monotonic in singlet masses, Figure 8.2 provides the largest Yukawa couplings allowed by $3 + 1$ active-mixing.

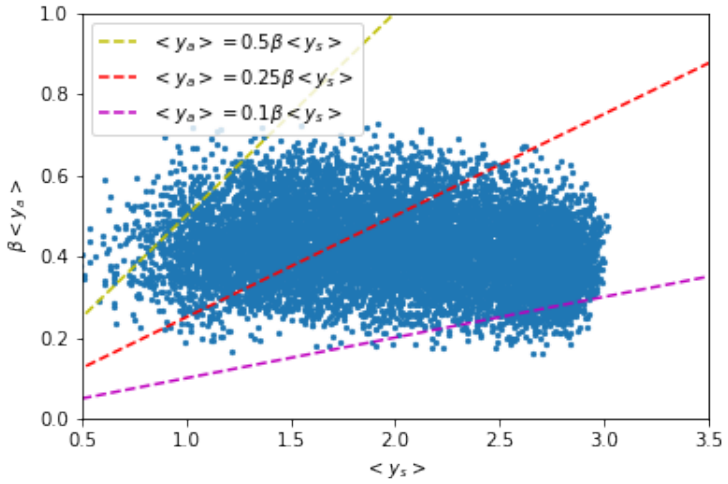

Figure 8.3: The parameter space of 3+1 global mixing illustrated in the $\langle y_s \rangle - \langle y_\alpha \rangle$ plane. The dashed lines represent lines of symmetry related to how the distribution varies relative to the vev ratio β. This scan is performed with $M_3 = 10\text{TeV}$, $M_4 = 100\text{TeV}$ and $M_1 = M_2 = \mathcal{U}(1, 10)\text{TeV}$

The above figure is taken for $M_3 = 10\text{TeV}$ as opposed to Figure 8.2 which was taken for $M_3 = 100\text{GeV}$. The figures are qualitatively similar, the only difference is the overall lower Yukawa scale. As discussed, this is due to Yukawa couplings being monotonic in the singlet masses. Hence, for a lower singlet scale, the Yukawa couplings (and their absolute averages) are lower. These plots were obtained in assuming a normal hierarchy between the active neutrino mass eigenvalues. One can do the same thing using an inverted hierarchy. This was done, but the distributions were very similar to those obtained using a normal hierarchy, and are hence not plotted. The reasoning for their similarity is that changing mass hierarchy only marginally changes the light mass eigenvalues. That is, these plots are hardly sensitive to variations of active masses, which are at the 10^{-2}eV-scale. Therefore, it is safe to conclude that for the recreation of the global fit, the Yukawa structure is hierarchical, where the dark Yukawa couplings overcompensate relative to the vev ratio β by a factor of about four. Due to the properties of the orthogonal matrix R in the Casas-Ibarra parametrization (7.1.5), there is no amount of fine-tuning which can change the hierarchical structure of Yukawa coupling with the addition of a fourth mass eigenstate at the eV-scale.

9 Conclusions and outlook

9.1 Conclusions

In this thesis, the effects of dark sector interactions on active-sterile neutrino mixing were discussed and analyzed. To perform the analysis, a thorough introduction into the properties of neutrinos, in and beyond, the standard picture was given. Majorana neutrinos and their peculiar properties were also discussed. Dark neutrinos are introduced by the addition of a minimal dark Abelian gauge, broken at the MeV-scale. These dark neutrinos were allowed to mix with the active neutrinos through non-diagonal Yukawa interactions with right-handed neutrinos, resulting in active-sterile neutrino mixing and

oscillations. The particle content of the added dark sector was discussed, and their masses and couplings were constrained based on observations. The mass scale of the dark sector was deduced based on the suppression of bosonic portals. This suppression is necessary, otherwise, the section would not be "dark".

The neutrino portal and the parameters it contains were studied. For a minimal dark sector, the low-energy observables (light masses and mixing angles) were formulated in terms of the Dirac mass matrix and the singlet mass matrix through the see-saw mechanism. These matrices had a large degree of redundancy which was removed based on the physics of SBL neutrino oscillation observables.

The model was first applied to a simplified version of neutrino oscillations, namely 1+1 active-sterile mixing. Three parameter regions corresponding to increasing active-sterile mixing were found using simple parameter scans. Large active-sterile mixing was found to require a hierarchical spectrum of Yukawa couplings. The addition of a fourth mass eigenstate at the eV-scale required at least one dark Yukawa coupling to be much larger than the active neutrino Yukawa couplings. This was argued to be due to the MeV-scale at which the dark sector was broken.

The model was then applied to 3+1 neutrino mixing. The Casas-Ibarra parametrization was used to obtain expressions for the Yukawa structure. This led to an upper bound on the singlet scale. The model was then used to replicate results from a global fit. The resulting Yukawa structure was found to be equivalent to the 1+1 mixing case. The reasoning was argued to be similar to that of 1+1 mixing. Due to the properties of the matrices occurring in the Casas-Ibarra parametrization, fine-tuning arguments cannot be used to flatten the hierarchical Yukawa structure. The general structure of the mass matrices was also discussed in light of the global fit.

The singlet masses were constrained from leptogenesis bounds. In particular, the model cannot accommodate singlet masses to account for standard thermal leptogenesis. The resonant regime had to be invoked. In light of resonant leptogenesis, the singlet masses were confined to the $1\text{TeV} - 100\text{TeV}$ scale. Lastly, the Yukawa parameter space, subject to these bounds, were scanned. The results indicate that the dark sector Yukawa couplings have to overcompensate the vev-ratio $\langle H\rangle\langle\varphi\rangle^{-1}$ by a factor of about 4. The overcompensation relative to the vev-ratio was expected due to the fourth mass eigenstate being much larger than the three others. In conclusion, the model can provide explanations for the SBL anomalies and resonant leptogenesis, while staying within experimental limits. The condition is that Yukawa couplings are hierarchical.

By studying active-sterile neutrino mixing, the tools developed can be applied more generally to neutrino oscillations with arbitrarily broken sectors. As an optimistic, yet possible example, if the particle content of dark matter allows for the construction of non-diagonal mass terms for dark fermions, then they might be probed in neutrino oscillation experiments. This could enlighten some of the mysteries surrounding the elusive dark matter, thus possibly provide a clue towards the largest mystery in all of spacetime.

9.2 Outlook

The main topic of this thesis is active-sterile neutrino mixing, enabled by the addition of a dark sector to the Standard Model. As discussed in Section 5.2, two additional portals can be used to connect the dark sector with the Standard Model, kinetic and scalar mixing. For future work, it would be interesting to analyze these portals and the physics within them. This would provide a greater understanding of the model, as the bosonic portals

could impose new constraints, thus leading to a more refined set of model parameters. Furthermore, studying the bosonic portals may provide a better understanding of how an arbitrary dark sector can be probed. This is important knowledge for further studies, of say, dark matter.

For future work, it could also be of interest to perform more sophisticated scans of the parameter space this model provides, either Bayesian methods or machine learning. Lastly, it would be interesting to consider the cosmological effect of the model and how the model fits into the current description of the expanding universe.

A Parameter scans

When constructing a particular model of, well anything, more often than not, there will necessarily be some unknown constant quantities, which the model relies on. These unknown constants are known as parameters. The space of all parameters from a given model is known as the parameter space of the model, which may range everywhere from a one-dimensional space to an infinite-dimensional space. Let y_i be a model parameter, which is a component of the vector $y \in \mathcal{Y}$, where $\mathcal{Y}$ is the parameter space. If all parameters are independent, then $\mathcal{Y}$ is a flat vector space, an assumption that is used throughout. For a model to be physical, it must be able to provide observable predictions. That is, there should some measurable quantity, which can be expressed as a function of the model parameters. Denote any observable quantity by $\mathcal{O}$. The fact that a physical model is required to make at least one physical prediction can be expressed by $\mathcal{O} = \mathcal{O}(y)$.

If the parameter space of a given model is small, say one or two-dimensional and the parameters are constrained in some fashion, e.g. $y_1 \in (a, b)$, then the brute force method of simply dividing the space into chunks of equal spacing, and computing the observables at each lattice point can be viable. This method suffers from two main drawbacks. First, if the parameter space is large, then the number of computations needed to scan the lattice grows exponentially: Consider an n-dimensional parameter space which is divided into p equally spaced lattice points along each axis. At each lattice point, q number of calculations is done to extract the relevant observables and store them. The total number of calculations needed to scan the lattice is then $N = qp^n$. For a fine scan of a large parameter space, this method is computationally infeasible. The second issue is that this method is ignorant of any structure hidden between lattice points. For example, a sharp, local maxima may occur in the middle of two lattice points. If the maxima are sharp enough, the underlying distribution will appear as the background at the evaluated lattice points.

To circumvent these issues, Monte-Carlo methods are used. Monte-Carlo methods/algorithms are fundamentally stochastic, which means that in choosing to use Monte-Carlo algorithms, determinism is lost. Regardless, the stochastic approach is invaluable when dealing with large spaces where deterministic approaches become inefficient. In this thesis, the MCMC algorithm of choice is the Metropolis-Hastings algorithm (MHA). The goal of the MHA is to find optimal regions of parameter space, which are defined by the desired qualities of the model. For example, consider a model of neutrino mixing where the masses are constrained, but the mixing angle is not. If large mixing is desirable, then the optimal regions(s) of parameter space are those such that the mixing is large, while the masses satisfy experimental constraints. It is then of interest to figure out which combinations of model parameters yields these optimal regions. To do this, probability distribution functions are constructed for the model observables $\mathcal{O}$ in terms of model parameters y. A standard choice is to model observables in terms of the symmetric, two-parameter Gaussian distribution

$$G(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{A.0.1}$$

where $\mu \in \mathbb{R}$ is the mean and $\sigma > 0$ is the standard deviation. In standard statistical notation, $P(x; a, b, \ldots)$ should be read as: $P(x)$ given $a, b, \ldots$. The Gaussian distribution is widely used as measurements of physical quantities often follow a Gaussian distribution. The cumulative Gaussian distribution is defined as

$$\Phi(x;\mu,\sigma) = \int_{-\infty}^{x} G(t;\mu,\sigma)dt. \tag{A.0.2}$$

The product of these distributions yields the skewed-Gaussian distribution

$$F(x;\mu,\sigma,\alpha) = 2G(x;\mu,\sigma)\Phi(\alpha x;\mu,\sigma). \tag{A.0.3}$$

A positive(negative) value of α skews $F(x;\mu,\sigma,\alpha)$ to the right(left). In the case where the skewness parameter α is zero, the skewed Gaussian distribution reduces to the ordinary Gaussian distribution. Figure A.1 illustrates three examples of skewed Gaussian distributions.

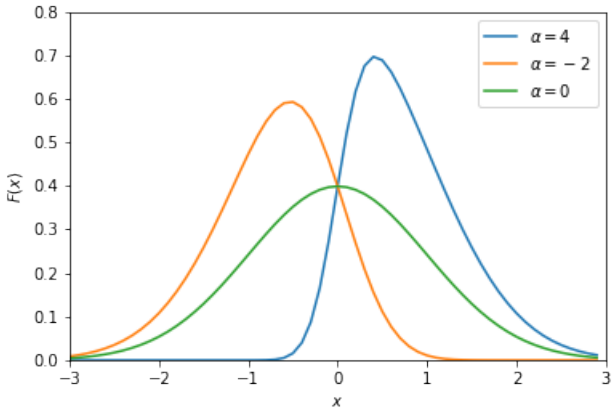

Figure A.1: Graphs of the example skewed Gaussian distributions. All graphs have a standard deviation of one and a mean at zero.

Another PDF that is commonly used is the uniform distribution $\mathcal{U}(a,b), b > a$, which returns a point between a and b with equal probability. The uniform distribution is used whenever there is no reason to choose a particular point in a range over another. As an example of applicability, the uniform distribution can be used when a variable is confined within some range, but no more information is given.

The MHA can be understood as the following sequence of steps:

1. Pick an initial point $y \in \mathcal{Y}$, either randomly or by hand and set $n = 0$.
2. Propose a new point y', obtained from the previous point through some PDF $Q(y';y_t)$, where the proposal PDF is usually taken to be a Gaussian.
3. Evaluate the acceptance ratio $R = P(y')/P(y_t)$ and generate a random number $u \in \mathcal{U}(0,1)$. If $R \geq u$, take the step by setting $y' = y_{n+1}$. Else, reject the step by setting $y_{n+1} = y_n$.
4. Repeat n times.

As each step is only dependent on the previous point, this sequence will form a Markow chain. The fact that the MHA constitutes a Markow chain means that the formal properties of Markow chains can be applied to the MHA. These include e.g. ergodicity, which means that every point reached by the chain will, at some later point, be reached

again. This is equivalent to stating that the system has a finite recurrence time. Another important property of Markow chains is coverage. This means that every point which can be reached from a series of steps with non-zero probability will be reached given enough time(steps). These two fundamental properties are a consequence of the acceptance criteria $R \geq u$. If on the contrary, the acceptance criterium was $R \geq 1$, then the walk would remain inert once either the global maxima is reached, or more dangerously, if a local maxima is reached. Such local maxima are called *islands of high density*, which refers to the fact that once the walk has landed on one of these islands, it is stranded. A region of high density is defined as a region in which a respectable portion of the PDF is located, it does not necessarily have to be the bulk of the PDF. To show the dangers of local maxima, consider the twin Gaussian distribution

$$K(x;\mu,\sigma) = K_1 G(x;-\mu,\sigma) + K_2 G(x;\mu,\sigma), \quad K_1 + K_2 = 1. \tag{A.0.4}$$

An example plot of this distribution is illustrated in Figure A.2.

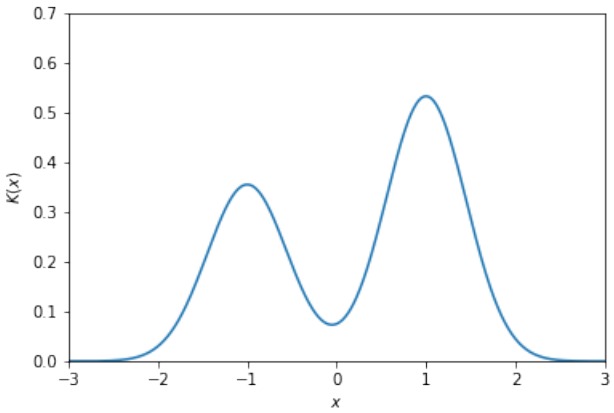

Figure A.2: The twin Gaussian distribution for $K_1 = 0.4$ and $K_2 = 0.6$.

Figure A.2 illustrates why the acceptance criterium is taken such to allow possible steps away from a local maxima. Additionally, there is also the issue with a small step size. As an example, if the random walk has reached the local maxima corresponding to the left peak of the twin Gaussian distribution in Figure A.2, and the step size is too low, then although by the properties of Markow chains, it is guaranteed to reach the global maxima (the right peak), it may take a very long time (longer than the programs runs). If the proposal distribution is Gaussian, then the step size is identified with the standard deviation. If the chosen distribution is of a high dimension, then distinguishing a local maxima from a global one can be difficult, if for some reason, after some repeated scans, the random walk ends at the local maxima each time. For optimization purposes, this is not ideal. However, if the local maxima satisfy the desired model features, then the local maxima can be deemed a sufficiently good region. This latter sentence can be identified with the question: Can the model do X?. X is interpreted as some condition(s). If the local maxima can satisfy X, then the model is valid. Which of these two approaches is chosen is based on the problem at hand.

In the Metropolis-Hastings algorithm, each accepted point is appended to some list along with the value of the observables at that point. Generally, not all of these sampled points correspond to regions of high density. If the walk is initialized in a region of low

density, e.g. far to the right of the right peak in twin Gaussian in Figure A.2, then the first sampled points will be unrepresentative of the target distribution. Therefore, in any MCMC algorithm, the first b sampled points are removed from consideration in a process known as *burn-in*. This number depends on how large the average step size is, but there is no consensus on how long the burn-in period should be. This is a disadvantage of MCMC algorithms, as to obtain good statistics, a large number of sampled points needs to be discarded.

On the note of step size, once the target is reached, the acceptance ratio should be 50% for a one-dimensional Gaussian distribution and decaying down to about 23% for a N-dimensional Gaussian distribution [59]. This proof does not generalize to other distributions, however, there is a general idea of how the acceptance should be for an arbitrary distribution, which is related to step size. At the target distribution, if the acceptance is large (almost one), then the step size is too small as it only reaches points that have approximately the same probability. This reflects the fact that every continuous space is locally flat. On the contrary, if the acceptance is too low (almost zero) then the step size is too large, as a large step away from, say a global maxima, will most likely result in a region of much lower density, i.e. the step is rarely taken.

The last technicality regarding parameter scans is the sampling near the edges of the space. This occurs only if there are some boundaries to the space under consideration. Consider a model parameter that is confined to $x \in (0, 1)$, where the proposal distribution is a Gaussian. As Gaussian distributions have compact support on the whole real line, there is a possibility that the proposed step will be outside the allowed region. To compensate for the possibility of such run-away parameters, one option is to simply reflect the runaway point back onto the allowed range. However, this is only a good approximation if the width of the proposal distribution is small, i.e. the proposal distribution is narrow. The reflection method is only analytic if the proposal distribution is a Dirac delta, which is rather boring, to say the least, or the sampled point is exactly on the boundary. If the sampled point is close to the boundary and has a large width, then the area outside the boundary is not representative of the area within the boundary and the reflection method is a bad approximation. To illustrate this point, three Gaussian proposal distributions are plotted near the right boundary of the model parameter.

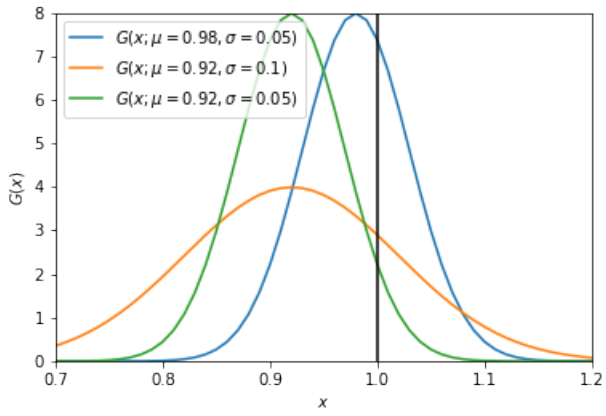

Figure A.3: Gaussian distributions near the boundary of a one-dimensional parameter space. The boundary is illustrated by the vertical, black line.

From Figure A.3 it is clear that for the narrow distributions, reflection at the edge

is a valid approximation. The rightmost, narrow distribution (blue) is almost symmetric about the boundary, which means that any reflected point is representative of the bulk of the distribution. For the leftmost narrow distribution (green), due to the sharp peak, there is only a small tail of the distribution which is outside the boundary. On the contrary, for the wide distribution (orange), there is a considerable bulk outside the boundary, while the peak is relatively far from the boundary. Hence, narrow proposal distributions are required whenever there is a boundary to parameter space, which is dealt with by the reflection method. A more sophisticated approach to the boundary issue would be to renormalize the distribution relative to the boundaries.

References

[1] P. Minkowski, $\mu \to e\gamma$ *at a Rate of One Out of* 10^9 *Muon Decays?* Phys. Lett. B **67** (1977) 421–428.

[2] M. Gell-Mann, P. Ramond, and R. Slansky, *Complex Spinors and Unified Theories*. Conf. Proc. C **790927** (1979) 315–321, `arXiv:1306.4669 [hep-th]`.

[3] T. Yanagida, *Horizontal gauge symmetry and masses of neutrinos*. Conf. Proc. C **7902131** (1979) 95–99.

[4] S. L. Glashow, *The Future of Elementary Particle Physics*. NATO Sci. Ser. B **61** (1980) 687.

[5] R. N. Mohapatra and G. Senjanovic, *Neutrino Mass and Spontaneous Parity Nonconservation*. Phys. Rev. Lett. **44** (1980) 912.

[6] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*. Phys. Rev. Lett. **13** (1964) 321–323.

[7] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*. Phys. Rev. Lett. **13** (1964) 508–509.

[8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*. Phys. Rev. Lett. **13** (1964) 585–587.

[9] H. Goldstein, *Classical mechanics*. Pearson, Essex, England, 2014.

[10] S. Carroll, *Spacetime and geometry : an introduction to general relativity*. Cambridge University Press, Cambridge, United Kingdom New York, NY, 2019.

[11] O. Gron, *Einstein's general theory of relativity : with modern applications in cosmology*. Springer, New York, 2007.

[12] B. Schutz, *A first course in general relativity*. Cambridge University Press, Cambridge New York, 2009.

[13] C. Misner, *Gravitation*. Princeton University Press, Princeton, N.J, 2017.

[14] E. Noether, *Invariante Variationsprobleme*. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1918** (1918) 235–257. `http://eudml.org/doc/59024`.

[15] M. Schwartz, *Quantum field theory and the standard model*. Cambridge University Press, Cambridge, United Kingdom New York, 2014.

[16] B. Hall, *Lie groups, Lie algebras, and representations : an elementary introduction*. Springer, New York, 2003.

[17] **Borexino** Collaboration, M. Agostini *et al.*, *A test of electric charge conservation with Borexino*. Phys. Rev. Lett. **115** (2015) 231802, `arXiv:1509.01223 [hep-ex]`.

[18] F. Mandl, *Quantum field theory*. Wiley, Hoboken, N.J, 2010.

[19] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in β Decay*. Phys. Rev. **105** (1957) 1413–1414.

[20] D. Hanneke, S. Fogwell, and G. Gabrielse, *New Measurement of the Electron Magnetic Moment and the Fine Structure Constant*. Phys. Rev. Lett. **100** (2008) 120801, arXiv:0801.1134 [physics.atom-ph].

[21] A. Constantin, Y.-H. He, and A. Lukas, *Counting String Theory Standard Models*. Phys. Lett. B **792** (2019) 258–262, arXiv:1810.00444 [hep-th].

[22] **Muon g-2** Collaboration, B. Abi *et al.*, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*. Phys. Rev. Lett. **126** (2021) no. 14, 141801, arXiv:2104.03281 [hep-ex].

[23] G. D. O. Gann, *Everything Under the Sun: A Review of Solar Neutrinos*, 2015. arXiv:1504.02154 [nucl-ex].

[24] C. Giunti, *Fundamentals of Neutrino Physics and Astrophysics*. Oxford University Press, UK, Oxford, 2007.

[25] S. M. Bilenky and S. T. Petcov, *Massive Neutrinos and Neutrino Oscillations*. Rev. Mod. Phys. **59** (1987) 671. [Erratum: Rev.Mod.Phys. 61, 169 (1989), Erratum: Rev.Mod.Phys. 60, 575–575 (1988)].

[26] C. Giunti and C. W. Kim, *Quantum mechanics of neutrino oscillations*. Found. Phys. Lett. **14** (2001) no. 3, 213–229, arXiv:hep-ph/0011074.

[27] S. M. Bilenky, C. Giunti, and W. Grimus, *Phenomenology of neutrino oscillations*. Prog. Part. Nucl. Phys. **43** (1999) 1–86, arXiv:hep-ph/9812360.

[28] L. Wolfenstein, *Neutrino Oscillations in Matter*. Phys. Rev. D **17** (1978) 2369–2374.

[29] S. P. Mikheyev and A. Y. Smirnov, *Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos*. Sov. J. Nucl. Phys. **42** (1985) 913–917.

[30] S. Vagnozzi, *Weigh them all! - Cosmological searches for the neutrino mass scale and mass ordering*. PhD thesis, 06, 2019.

[31] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou, *The fate of hints: updated global analysis of three-flavor neutrino oscillations*. JHEP **09** (2020) 178, arXiv:2007.14792 [hep-ph].

[32] **Planck** Collaboration, N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*. Astron. Astrophys. **641** (2020) A6, arXiv:1807.06209 [astro-ph.CO].

[33] **Particle Data Group** Collaboration, P. A. Zyla *et al.*, *Review of Particle Physics*. PTEP **2020** (2020) no. 8, 083C01.

[34] S. M. Bilenky and C. Giunti, *Neutrinoless Double-Beta Decay: a Probe of Physics Beyond the Standard Model*. Int. J. Mod. Phys. A **30** (2015) no. 04n05, 1530001, arXiv:1411.4791 [hep-ph].

[35] **NEXT** Collaboration, F. Granena *et al.*, *NEXT, a HPGXe TPC for neutrinoless double beta decay searches*. arXiv:0907.4054 [hep-ex].

[36] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*. NATO Sci. Ser. B **59** (1980) 135–157.

[37] A. Pilaftsis, *Radiatively induced neutrino masses and large Higgs neutrino couplings in the standard model with Majorana fields*. Z. Phys. C **55** (1992) 275–282, arXiv:hep-ph/9901206.

[38] A. de Gouvea, *GeV seesaw, accidentally small neutrino masses, and Higgs decays to neutrinos*. arXiv:0706.1732 [hep-ph].

[39] S. Weinberg, *Baryon and Lepton Nonconserving Processes*. Phys. Rev. Lett. **43** (1979) 1566–1570.

[40] J. Schechter and J. W. F. Valle, *Neutrino Masses in SU(2) x U(1) Theories*. Phys. Rev. D **22** (1980) 2227.

[41] **LSND** Collaboration, A. Aguilar-Arevalo *et al.*, *Evidence for neutrino oscillations from the observation of $\bar{\nu}_e$ appearance in a $\bar{\nu}_\mu$ beam*. Phys. Rev. D **64** (2001) 112007, arXiv:hep-ex/0104049.

[42] **MiniBooNE** Collaboration, A. A. Aguilar-Arevalo *et al.*, *Event Excess in the MiniBooNE Search for $\bar{\nu}_\mu \to \bar{\nu}_e$ Oscillations*. Phys. Rev. Lett. **105** (2010) 181801, arXiv:1007.1150 [hep-ex].

[43] **ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group** Collaboration, S. Schael *et al.*, *Precision electroweak measurements on the Z resonance*. Phys. Rept. **427** (2006) 257–454, arXiv:hep-ex/0509008.

[44] G. Mention, M. Fechner, T. Lasserre, T. A. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau, *The Reactor Antineutrino Anomaly*. Phys. Rev. D **83** (2011) 073006, arXiv:1101.2755 [hep-ex].

[45] P. Huber, *On the determination of anti-neutrino spectra from nuclear reactors*. Phys. Rev. C **84** (2011) 024617, arXiv:1106.0687 [hep-ph]. [Erratum: Phys.Rev.C 85, 029901 (2012)].

[46] T. A. Mueller *et al.*, *Improved Predictions of Reactor Antineutrino Spectra*. Phys. Rev. C **83** (2011) 054615, arXiv:1101.2663 [hep-ex].

[47] J. N. Abdurashitov *et al.*, *Measurement of the response of a Ga solar neutrino experiment to neutrinos from an Ar-37 source*. Phys. Rev. C **73** (2006) 045805, arXiv:nucl-ex/0512041.

[48] C. Giunti, M. Laveder, Y. F. Li, Q. Y. Liu, and H. W. Long, *Update of Short-Baseline Electron Neutrino and Antineutrino Disappearance*. Phys. Rev. D **86** (2012) 113014, arXiv:1210.5715 [hep-ph].

[49] M. Dentler, A. Hernández-Cabezudo, J. Kopp, P. A. N. Machado, M. Maltoni, I. Martinez-Soler, and T. Schwetz, *Updated Global Analysis of Neutrino Oscillations in the Presence of eV-Scale Sterile Neutrinos*. JHEP **08** (2018) 010, arXiv:1803.10661 [hep-ph].

[50] S. Böser, C. Buck, C. Giunti, J. Lesgourgues, L. Ludhova, S. Mertens, A. Schukraft, and M. Wurm, *Status of Light Sterile Neutrino Searches*. Prog. Part. Nucl. Phys. **111** (2020) 103736, arXiv:1906.01739 [hep-ex].

[51] G. C. Branco, J. T. Penedo, P. M. F. Pereira, M. N. Rebelo, and J. I. Silva-Marcos, *Type-I Seesaw with eV-Scale Neutrinos*. JHEP **07** (2020) 164, arXiv:1912.05875 [hep-ph].

[52] E. Bertuzzo, S. Jana, P. A. N. Machado, and R. Zukanovich Funchal, *Neutrino Masses and Mixings Dynamically Generated by a Light Dark Sector*. Phys. Lett. B **791** (2019) 210–214, arXiv:1808.02500 [hep-ph].

[53] M. Ahlers, J. Jaeckel, J. Redondo, and A. Ringwald, *Probing Hidden Sector Photons through the Higgs Window*. Phys. Rev. D **78** (2008) 075005, arXiv:0807.4143 [hep-ph].

[54] T. Gherghetta, J. Kersten, K. Olive, and M. Pospelov, *Evaluating the price of tiny kinetic mixing*. Phys. Rev. D **100** (2019) no. 9, 095001, arXiv:1909.00696 [hep-ph].

[55] C. Giunti and E. M. Zavanin, *Appearance–disappearance relation in* $3 + N_s$ *short-baseline neutrino oscillations*. Mod. Phys. Lett. A **31** (2015) no. 01, 1650003, arXiv:1508.03172 [hep-ph].

[56] C. Giunti and M. Laveder, *3+1 and 3+2 Sterile Neutrino Fits*. Phys. Rev. D **84** (2011) 073008, arXiv:1107.1452 [hep-ph].

[57] W. Grimus and L. Lavoura, *Double seesaw mechanism and lepton mixing*. JHEP **03** (2014) 004, arXiv:1309.3186 [hep-ph].

[58] S. Gariazzo, C. Giunti, M. Laveder, and Y. F. Li, *Model-independent* $\bar{\nu}_e$ *short-baseline oscillations from reactor spectral ratios*. Phys. Lett. B **782** (2018) 13–21, arXiv:1801.06467 [hep-ph].

[59] A. Gelman, W. R. Gilks, and G. O. Roberts, *Weak convergence and optimal scaling of random walk Metropolis algorithms*. The Annals of Applied Probability **7** (1997) no. 1, 110 – 120. https://doi.org/10.1214/aoap/1034625254.

[60] E. Bulbul, M. Markevitch, A. Foster, R. K. Smith, M. Loewenstein, and S. W. Randall, *Detection of An Unidentified Emission Line in the Stacked X-ray spectrum of Galaxy Clusters*. Astrophys. J. **789** (2014) 13, arXiv:1402.2301 [astro-ph.CO].

[61] A. Boyarsky, O. Ruchayskiy, D. Iakubovskyi, and J. Franse, *Unidentified Line in X-Ray Spectra of the Andromeda Galaxy and Perseus Galaxy Cluster*. Phys. Rev. Lett. **113** (2014) 251301, arXiv:1402.4119 [astro-ph.CO].

[62] S. B. Roland, B. Shakya, and J. D. Wells, *Neutrino Masses and Sterile Neutrino Dark Matter from the PeV Scale*. Phys. Rev. D **92** (2015) no. 11, 113009, arXiv:1412.4791 [hep-ph].

[63] S. B. Roland, B. Shakya, and J. D. Wells, *PeV neutrinos and a 3.5 keV x-ray line from a PeV-scale supersymmetric neutrino sector*. Phys. Rev. D **92** (2015) no. 9, 095018, arXiv:1506.08195 [hep-ph].

[64] B. Shakya and J. D. Wells, *Exotic Sterile Neutrinos and Pseudo-Goldstone Phenomenology*. JHEP **02** (2019) 174, arXiv:1801.02640 [hep-ph].

[65] J. M. Berryman, A. de Gouvêa, K. J. Kelly, and Y. Zhang, *Dark Matter and Neutrino Mass from the Smallest Non-Abelian Chiral Dark Sector*. Phys. Rev. D **96** (2017) no. 7, 075010, arXiv:1706.02722 [hep-ph].

[66] A. Das, S. Goswami, V. K. N., and T. K. Poddar, *Freeze-in sterile neutrino dark matter in a class of $U(1)'$ models with inverse seesaw*. arXiv:2104.13986 [hep-ph].

[67] J. A. Casas and A. Ibarra, *Oscillating neutrinos and $\mu \to e, \gamma$*. Nucl. Phys. B **618** (2001) 171–204, arXiv:hep-ph/0103065.

[68] C. Giunti, Y. F. Li, and Y. Y. Zhang, *KATRIN bound on 3+1 active-sterile neutrino mixing and the reactor antineutrino anomaly*. JHEP **05** (2020) 061, arXiv:1912.12956 [hep-ph].

[69] **KATRIN** Collaboration, M. Aker *et al.*, *Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN*. Phys. Rev. Lett. **123** (2019) no. 22, 221802, arXiv:1909.06048 [hep-ex].

[70] S. Fichet, *Quantified naturalness from Bayesian statistics*. Phys. Rev. D **86** (2012) 125029, arXiv:1204.4940 [hep-ph].

[71] **Planck** Collaboration, N. Aghanim *et al.*, *Planck 2018 results. I. Overview and the cosmological legacy of Planck*. Astron. Astrophys. **641** (2020) A1, arXiv:1807.06205 [astro-ph.CO].

[72] P. Di Bari, *An introduction to leptogenesis and neutrino properties*. Contemp. Phys. **53** (2012) no. 4, 315–338, arXiv:1206.3168 [hep-ph].

[73] D. Bödeker and W. Buchmüller, *Baryogenesis from the weak scale to the grand unification scale*. arXiv:2009.07294 [hep-ph].

[74] M. Fukugita and T. Yanagida, *Baryogenesis Without Grand Unification*. Phys. Lett. B **174** (1986) 45–47.

[75] *Planck and the cosmic microwave background*, http://www.esa.int/Science_Exploration/Space_Science/Planck/Planck_and_the_cosmic_microwave_background, note = Accessed: 2021-05-25.

[76] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*. Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32–35.

[77] G. 't Hooft, *Symmetry Breaking Through Bell-Jackiw Anomalies*. Phys. Rev. Lett. **37** (1976) 8–11.

[78] A. E. Nelson and J. Walsh, *Short Baseline Neutrino Oscillations and a New Light Gauge Boson*. Phys. Rev. D **77** (2008) 033001, arXiv:0711.1363 [hep-ph].

[79] R. Harnik, J. Kopp, and P. A. N. Machado, *Exploring nu Signals in Dark Matter Detectors*. JCAP **07** (2012) 026, arXiv:1202.6073 [hep-ph].

[80] S. Davidson and A. Ibarra, *A Lower bound on the right-handed neutrino mass from leptogenesis*. Phys. Lett. B **535** (2002) 25–32, arXiv:hep-ph/0202239.

[81] A. Pilaftsis and T. E. J. Underwood, *Resonant leptogenesis*. Nucl. Phys. B **692** (2004) 303–345, arXiv:hep-ph/0309342.

[82] A. Pilaftsis, *CP violation and baryogenesis due to heavy Majorana neutrinos*. Phys. Rev. D **56** (1997) 5431–5451, arXiv:hep-ph/9707235.