University of Silesia
Faculty of Science and Technology
August Chełkowski Institute of Physics

Wojciech Flieger

Constraints on neutrino mixing from matrix theory

PhD Thesis

PhD Supervisor:
prof. dr hab. Janusz Gluza
co-Supervisor:
dr Bartosz Dziewit

UNIVERSITY OF SILESIA
IN KATOWICE

Katowice 2021

słowa kluczowe: neutrino, mechanizm mieszania, teoria macierzy, geometria wypukła
key words: neutrino, mixing mechanism, matrix theory, convex geometry

Oświadczenie autora pracy

Ja, niżej podpisany: Wojciech Flieger, autor pracy doktorskiej pt.
'*Constraints on neutrino mixing from matrix theory*'.

Oświadczam, że ww. praca doktorska:

- została przygotowana przeze mnie samodzielnie,
- nie narusza praw autorskich w rozumieniu ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (tekst jednolity Dz. U. z 2006 r. Nr 90, poz. 631, z późn. zm.) oraz dóbr osobistych chronionych prawem cywilnym,
- nie zawiera danych i informacji, które uzyskałem w sposób niedozwolony, nie była podstawą nadania stopnia doktora nauk, dyplomu wyższej uczelni lub tytułu zawodowego ani mnie, ani innej osobie.

Oświadczam również, że treść pracy doktorskiej zapisanej na przekazanym przeze mnie jednocześnie nośniku elektronicznym jest identyczna z treścią zawartą w wydrukowanej wersji pracy.

Jestem świadomy odpowiedzialności karnej za złożenie fałszywego oświadczenia.

....

Miejscowość, data Podpis autora pracy

Acknowledgments

This work was supported by the Polish National Science Center (NCN) under grants 2013/11/B/ST2/04023 and 2017/25/B/ST2/01987, and by the COST (European Cooperation in Science and Technology) Action CA16201 PARTICLEFACE.

Abstract

One of the long-standing puzzles in neutrino physics concerns the number of neutrino flavours in nature. So far the existence of three types of active neutrinos has been established. However, it is crucial to ascertain if more neutrino flavour states exist. Such neutrinos are dubbed sterile as their weak interaction with ordinary matter is below available detection limits. Nonetheless, sterile neutrinos can mix with active neutrinos leaving visible imprints in the form of a deviation from the unitarity of the Standard Model neutrino mixing matrix. Thus, studies of non-unitarity of the mixing matrix are crucial in understanding neutrino physics.

We develop a novel approach of studying neutrino mixing matrices based on matrix theory. It has been built on quantities known as singular values and the notion of contractions. Based on that we define a region of physically admissible mixing matrices as a convex hull over experimentally determined three-dimensional unitary mixing matrices. We study the geometrical properties of this physical region by measuring its volume expressed by the Haar measure of the singular value decomposition and exploring its internal structure corresponding to a different number of sterile neutrinos.

We show how to identify unitarity-breaking cases based on singular values and construct their unitary extensions yielding a complete theory of minimal dimensionality larger than three through the theory of unitary matrix dilations. Using that we find stringent constraints on active-sterile neutrino mixings in models with three active and one sterile neutrino states.

Streszczenie

Jeden z kluczowych problemów współczesnej fizyki cząstek elementarnych dotyczy liczby zapachów neutrin występujących w naturze. Do tej pory udało się ustalić, że istnieją trzy rodzaje neutrin aktywnych. Istotnym problemem jest ustalenie, czy istnieją inne dodatkowe stany neutrinowe. Neutrina takie nazywamy sterylnymi ze względu na fakt, że ich oddziaływanie słabe ze znaną materią jest jak do tej pory poniżej eksperymentalnego progu detekcji. Niemniej jednak neutrina sterylne mogą się mieszać z neutrinami aktywnymi pozostawiając tym samym ślady swojego istnienia na poziomie Modelu Standardowego w postaci nieunitarności macierzy mieszania neutrin. Z tego powodu badanie nieunitarności macierzy mieszania jest tak istotne dla pełnego zrozumienia fizyki neutrin.

W rozprawie przedstawiamy nową metodę analizy macierzy mieszania neutrin opartą na teorii macierzy. Fundament naszego podejścia do badania macierzy mieszania neutrin stanowią pojęcia wartości osobliwych oraz kontrakcji. Dzięki tym pojęciom zdefiniowaliśmy obszar fizycznie dopuszczalnych macierzy mieszania jako powłokę wypukłą rozpiętą na trójwymiarowych unitarnych macierzach mieszania wyznaczonych na podstawie danych eksperymentalnych. W rozprawie badamy geometryczne własności tego obszaru wyznaczając jego objętość wyrażoną poprzez miarę Haara rozkładu na wartości osobliwe oraz studiując jego strukturę wewnętrzną zależną od minimalnej liczby dodatkowych sterylnych neutrin.

Stosując teorię unitarnej dylatacji pokazujemy jak wartości osobliwe pozwalają zidentyfikować nieunitarne macierze mieszania oraz jak tworzyć ich rozszerzenia do pełnej macierzy unitarnej wymiaru większego niż trzy, opisującej kompletną teorię zawierającą neutrina sterylne. Na tej podstawie wyznaczamy nowe ograniczenia w modelach gdzie aktywne neutrina mieszają się z jednym dodatkowym neutrinem sterylnym.

List of papers

This thesis is based in part on the following papers and conference contributions:

1. K. Bielas, W. Flieger, J. Gluza and M. Gluza, 'Neutrino mixing, interval matrices and singular values', Phys. Rev. D**98** (2018) no.5, 053001, doi:10.1103/PhysRevD.98.053001 .

2. W. Flieger, J. Gluza and K. Porwit, 'New limits on neutrino non-unitary mixings based on prescribed singular values', JHEP **03** (2020), 169, doi:10.1007/JHEP03(2020)169.

3. W. Flieger and J. Gluza, 'General neutrino mass spectrum and mixing properties in seesaw mechanisms', Chin. Phys. C**45** (2021) no.2, 023106, doi:10.1088/1674-1137/abcd2f.

4. W. Flieger and K. Grzanka, PoS ICHEP2020 (2021), 129, doi:10.22323/1.390.0129.

5. W. Flieger, F. Pindel and K. Porwit, PoS CORFU2018 (2019), 050, doi:10.22323/1.347.0050.

6. W. Flieger, K. Porwit and J. Gluza, Acta Phys. Polon. B **50** (2019), 1729-1736, doi:10.5506/APhysPolB.50.1729.

7. K. Bielas and W. Flieger, Acta Phys. Polon. B **48** (2017), 2213, doi:10.5506/APhysPolB.48.2213.

Contents

1 Introduction

Since its earliest history, neutrino physics has stimulated discoveries that break the paradigms of particle physics. Even today, when we know much more about neutrinos, there are still many unsolved problems. Solutions to them could change the physics we know today drastically. Among them is the question concerning the number of neutrino types in nature. At the moment three types of neutrinos are known, each one corresponding to the different type[1] of the charged lepton (electron, muon, tau). That there are three types of neutrinos has been known since the LEP era. The central value for the effective number of light neutrinos N_ν was determined by analyzing around 17 million Z-boson decays, yielding $N_\nu = 2.9840 \pm 0.0082$ [1,2]. It is worth mentioning that the recent reevaluation of the data [3,4], including higher-order QED corrections to the Bhabha process, constrains further the value of N_ν, which is now $N_\nu = 2.9963 \pm 0.0074$. Moreover, from oscillation experiments it is known that neutrinos are massive particles [5, 6]. Massive neutrinos can be either self-conjugate Majorana particles or of Dirac type. In the second case left-handed neutrino fields are accompanying by right-handed neutrino states. These additional right-handed states do not couple directly with the Standard Model (SM) W and Z bosons. Such non-weakly interacting neutrinos are known as "sterile" and are in the limelight of modern particle physics research. However, they may influence the Standard Model physics, as they can mix with active Standard Model left-handed states. As a result, new neutrino states modify the mixing matrix, so that it is no longer unitary, and the mixing between extended flavour and mass states is described by a matrix of dimension larger than three. This extended matrix should, in general, itself be unitary, meaning completeness of the active-sterile mixing is restored.

The concept of sterile neutrinos is very appealing. They can explain important physical phenomena such as small masses of active neutrinos, e.g. by the seesaw mechanism [7–10]. Interestingly, some recent experiments and experimental signals suggest that the fourth type of neutrino may exist. Namely, measurements of the $\bar{\nu}_e$ flux at small distances from nuclear reactors give 6% fewer events than expected [11]. Such a deficit referenced as a "Reactor Antineutrino Anomaly" can be explained as active-sterile antineutrino oscillations at very short baselines [12]. The recent results of Daya Bay [13], NEOS [14], RENO [15] and Double Chooz [16] experiments confirm this reactor antineutrino anomaly. Moreover, the NEUTRINO-4 reactor experiment claims to detect an electron antineutrino to sterile neutrino oscillation at the 3σ significance

[1] Various lepton types are also often named flavours or species.

level [17–19]. However, there are some controversies concerning this result and a more refined analysis of the data shows that claimed results are rather doubtful [20–22]. Gallium solar neutrinos experiments also observe fewer events than predicted. A deficit in data is reported by GALEX and SAGE collaborations [23]. Statistical significance for such "Gallium Anomalies" in terms of neutrino oscillations was recently estimated at the level of 2.3σ [24] which decreases the previous estimation of 3σ level [25].

Clues for sterile neutrino oscillations exist also in the long-baseline experiments. The first reported abnormalities were reported by the LSND collaboration in 1995. More efficient analysis from 1996 showed an excess of electron antineutrinos events from the muon antineutrino beam [26]. Recent results from the MiniBooNE experiment [27] show a massive excess of electron (anti)neutrino events from the muon (anti)neutrino beam. There also exist some hints towards two additional sterile neutrinos with eV scale masses [28–31]. However, they contradict the latest muon neutrino disappearance results from `MINOS/MINOS+` and `IceCube` [32, 33]. So, the situation is not clear concerning scales and the number of additional neutrino states in general, and further scrutinized studies are needed, both on experimental and theoretical sides. Hence, studies of the violation of unitarity of the SM mixing matrix are crucial for finding hints for new neutrino states.

Theoretical approach to the deviation from unitarity of the mixing matrix is based on η and α parametrizations [34–39]. The use of these parametrizations is mostly phenomenological, focused mainly on the estimation of this deviation without touching the nature of the mixing mechanism. However, there are great possibilities for formal studies of neutrino mixing. On the one hand, the structure of the neutrino mixing matrix can be studied in terms of group theory using additional symmetries (A_4, S_3, etc., [40]). On the other hand, matrix theory opens new insight to the neutrino mixing studies. In this work, we will develop the latter approach which allows for a unified treatment of both the Standard Model and beyond the Standard Model (BSM) scenarios by focusing on the 3×3 mixing matrices available from oscillation experiments. The crucial step in the development of this method is the recognition of the significance of the singular values in studies of mixing matrices [41]. They underlie the notion of contractions which impose a strong restriction on the mixing matrices allowing to sift the physically meaningful mixing matrices from those non-physical. Following this idea, all physically admissible mixing matrices form a geometric region whose structure reflects important physical properties. It appears that the minimal number of additional sterile neutrinos is not arbitrary, but is encoded in singular values. This property is inscribed in the internal structure of the geometric region of physically meaningful mixing matrices.

The study of complete unitary matrices of some BSM scenarios agreeing with experimental results is possible by a procedure called a unitary dilation. Besides its formal insight, the matrix theory approach to mixing studies enables also to make quantitative predictions of the physical effects of sterile neutrinos.

The dissertation has the following structure. In the next chapter, we present details of the theory of massive neutrinos important for further studies. Chapter three contains the basic notation of matrix theory presented from the neutrino mixing perspective. In chapter four, we define and study geometrical properties of the region of physically admissible mixing matrices. Chapter five is devoted to the study of phenomenological predictions of effects of sterile neutrinos by singular values. In the final chapter formal aspects of the connection between masses and mixing are presented, focusing on the seesaw family. The work ends with a summary and discussion of the possible directions of further studies.

2 Theory of neutrino masses and mixing

2.1 Basic framework

The history of neutrinos has been inseparably connected to the β-decay and development of weak interactions. The β-decay possesses the continuous energy spectrum of the decaying particle which cannot be reconciled with a discrete energy spectrum of directly detectable particles seen in a two-body decay, ${}^{A}_{Z}X \to {}^{A}_{Z+1}X + e$, unless the energy conservation law is broken. On the other hand, if an additional neutral particle takes part in this process, ${}^{A}_{Z}X \to {}^{A}_{Z+1}X + e + \nu$, the continuous spectrum of the electron can be obtained. Such a postulate was made by Wolfgang Pauli and announced in 1930 during the nuclear conference in Tübingen. In his original statement, Pauli names the new particle ν a neutron and assumes that its spin is 1/2, and mass must be about the same as the electron mass and certainly not larger than one percent of the proton mass. After the discovery of the neutron, as it is known today, in 1932 [42], Fermi changed the name of Pauli's particle to the neutrino. Further development of the neutrino theory was driven by the necessity of the description of experimental results of the β-decay. The first step in this direction was undertaken by Fermi in 1934 [43] who constructed a suitable Hamiltonian assuming that the neutron decays to the proton and electron-neutrino pair, $n \to p + e + \nu$, and the interaction is similar to the electromagnetic one

$$\mathcal{H}^{\beta}(x) = G_F \bar{p}(x)\gamma_\mu n(x)\bar{e}(x)\gamma^\mu \nu(x) + H.c., \tag{2.1}$$

where G_F is the interaction constant, $p(x)$, $n(x)$, $e(x)$ and $\nu(x)$ are proton, neutron, electron and neutrino fields, respectively. Further experimental discoveries, such as the discovery of the parity violation in the β-decay [44], were followed by the development of the theory [45–47] in which the two-component neutrinos are massless particles. In this framework, fermion fields are the sum of the left-handed and right-handed chiral components

$$\nu(x) = \nu_L(x) + \nu_R(x), \tag{2.2}$$

where

$$\nu_{L,R}(x) = \frac{I \mp \gamma_5}{2}\nu(x) \equiv P_{L,R}\nu(x), \tag{2.3}$$

with I being the identity matrix and $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$ (gammas are defined along with the Dirac equation). The $P_{L,R}$ are called chirality projectors and satisfies the following

properties

$$\begin{aligned} P_L + P_R &= \frac{I - \gamma_5 + I + \gamma_5}{2} = I, \\ P_L P_R &= P_R P_L = 0. \end{aligned} \tag{2.4}$$

In the two-component theory, it is assumed that the neutrino fields exist only in one chiral state ν_L or ν_R. Let us express the neutrino mass Lagrangian in terms of chiral fields

$$\begin{aligned} \mathcal{L} &= -m\bar{\nu}(x)\nu(x) = -m(\bar{\nu}_L(x) + \bar{\nu}_R(x))(\nu_L(x) + \nu_R(x)) \\ &= -m(\bar{\nu}_L(x)\nu_L(x) + \bar{\nu}_L(x)\nu_R(x) + \bar{\nu}_R(x)\nu_L(x) + \bar{\nu}_R(x)\nu_R(x)), \end{aligned} \tag{2.5}$$

where $\bar{\nu} = \nu^\dagger\gamma^0$. By using the fact that $\bar{\nu}_R = \bar{\nu}P_L$ and $\bar{\nu}_L = \bar{\nu}P_R$ and applying the second identity of (2.4), we end up with

$$\mathcal{L} = -m(\bar{\nu}_L(x)\nu_R(x) + \bar{\nu}_R(x)\nu_L(x)), \tag{2.6}$$

which is a standard form for a fermion mass term. However, in the two-component neutrino theory, one of the fields does not exist and as a result, neutrinos must be massless particles. Experimental confirmation of the two-component neutrino theory was given in 1958 [48], showing that neutrinos are left-handed particles. If we assume that the Standard Model is based on the two-component neutrino theory, where only left-handed neutrino fields ν_L are present in the SM Lagrangian, then neutrinos are massless particles. However, there is experimental evidence that neutrinos are massive particles [5,6]. That being said, massive neutrinos can be easily accommodated within the SM by introducing right-handed neutrino fields ν'_R which are singlets of the $SU(2)_L$ gauge group. Then, the Yukawa Lagrangian reads

$$\mathcal{L}^\nu_Y = -\frac{\sqrt{2}}{v}\sum_{l',l}\bar{\psi}_{l'L}M'_{l'l}\nu'_{lR}\tilde{\phi} + H.c. \tag{2.7}$$

where $\psi_L = (\nu'_L, l'_L)^T$ for $l = e, \mu, \tau$, M' is a complex matrix and $\tilde{\phi}$ is a conjugated Higgs doublet with the vacuum expectation value v. After the spontaneous symmetry breaking we get

$$\mathcal{L}^\nu_Y = -\bar{\nu}'_L M'\nu'_R\left(1 + \frac{H}{v}\right) + H.c. \tag{2.8}$$

with

$$\begin{aligned} \nu'_L &= (\nu'_{eL}, \nu'_{\mu L}, \nu'_{\tau L})^T, \\ \nu'_R &= (\nu'_{eR}, \nu'_{\mu R}, \nu'_{\tau R})^T. \end{aligned} \tag{2.9}$$

The left-handed fields ν'_L can be transformed to the flavour basis by a transformation

$$\nu_L^f = U_L^\dagger \nu'_L = (\nu_{eL}, \nu_{\mu L}, \nu_{\tau L})^T, \tag{2.10}$$

where U_L is a unitary matrix used to diagonalize a charged leptons mass matrix. This gives

$$\mathcal{L}_Y^\nu = -\bar{\nu}_L^f M_D \nu'_R \left(1 + \frac{H}{v}\right) + H.c. \tag{2.11}$$

where $M_D = U_L^\dagger M'$. The first part represents the neutrino mass term

$$\mathcal{L}^D = -\bar{\nu}_L^f M_D \nu'_R + H.c. \tag{2.12}$$

The complex mass matrix M_D can be diagonalized by the singular value decomposition $M_D = UmV^\dagger$, where U and V are unitary matrices and $m = diag(m_1, m_2, m_3)$. After the diagonalization we obtain

$$\mathcal{L}^D = -\bar{\nu}_L m \nu_R + H.c. \tag{2.13}$$

where

$$\begin{aligned} \nu_L &= U^\dagger \nu_L^f = U^\dagger U_L^\dagger \nu'_L, \\ \nu_R &= V^\dagger \nu'_R. \end{aligned} \tag{2.14}$$

The left- and right-handed fields can be combined into neutrino fields with a definite mass

$$\mathcal{L}^D(x) = -\bar{\nu}(x) m \nu(x), \tag{2.15}$$

where $\nu = \nu_L + \nu_R$ and $\nu = (\nu_1, \nu_2, \nu_3)^T$ and ν_i are neutrino fields with definite mass m_i for $i = 1, 2, 3$. This mass term is called the Dirac mass term. The relation (2.14) is known as the neutrino mixing and the matrix $U_{\texttt{PMNS}} = U_L U$ is called the PMNS (Pontecorvo-Maki-Nagawa-Sakata) mixing matrix [49, 50]. This relation describes the fact that neutrino fields with definite flavour differ from the fields with definite mass, but are related as the linear combinations of each other. Experimental data shows that for charged leptons the flavour fields are physically measurable fields, i.e fields with definite mass. Thus, for charged leptons, flavour and massive fields coincide within experimental accuracy. This implies that the charged lepton mass matrix is already in a diagonal form, hence the unitary U_L matrix is the identity matrix and the PMNS mixing matrix is equal to the neutrino mixing matrix U.

However, as neutrinos are neutral particles, there is another possibility to construct a mass term involving only one chiral type of the field, namely the Majorana field. For

neutral particles the charge conjugation of left-handed fields $(\nu_{lL})^{\mathcal{C}} = \mathcal{C}\bar{\nu}_{lL}^{T}$, where $\mathcal{C}$ is a charge conjugate unitary operator satisfying $\mathcal{C}\gamma_{\mu}^{T}\mathcal{C}^{-1} = -\gamma_{\mu}$ and $\mathcal{C}^{T} = -\mathcal{C}$, behave as right-handed fields and, conversely, the charge conjugation of right-handed fields $(\nu_{lR})^{\mathcal{C}}$ are left-handed fields. To construct a mass term we need a combination of right- and left-handed fields, thus in the case of Majorana particles, we can build a mass term only from one type of neutrino fields. It is called a Majorana mass term and for left-handed neutrino fields its matrix form can be written as

$$\mathcal{L}_{L}^{M} = -\frac{1}{2}\bar{\nu}_{L}M_{L}\nu_{L}^{\mathcal{C}} + H.c. \tag{2.16}$$

where $\nu_{L} = (\nu_{eL}, \nu_{\mu L}, \nu_{\tau L})^{T}$ and M_{L} is a complex symmetric matrix. Such a matrix can be diagonalized by the congruence transformation

$$M_{L} = UmU^{T}, \tag{2.17}$$

where U is a unitary matrix and $m = diag(m_1, m_2, m_3)$. After the diagonalization we get

$$\mathcal{L}_{L}^{M} = -\frac{1}{2}\bar{\nu}^{M}m\nu^{M}, \tag{2.18}$$

where we defined

$$\nu^{M} = U^{\dagger}\nu_{L} + (U^{\dagger}\nu_{L})^{\mathcal{C}} = (\nu_1, \nu_2, \nu_3)^{T} = \nu_{L}^{M} + \nu_{R}^{M}. \tag{2.19}$$

The ν_i are Majorana neutrino fields with mass m_i for $i = 1, 2, 3$. The Majorana fields are invariant under the charge conjugation

$$(\nu^{M})^{\mathcal{C}} = \nu^{M} \text{ or for individual fields } \nu_{i}^{\mathcal{C}} = \nu_{i}. \tag{2.20}$$

This is the so-called Majorana condition and physically it means that Majorana particles are their own antiparticles. The mixing between flavour and massive fields is given by

$$\nu_{L} = U\nu_{L}^{M}. \tag{2.21}$$

A similar mass term can be constructed for the right-handed neutrino fields

$$\mathcal{L}_{R}^{M} = -\frac{1}{2}\overline{(\nu_{R})^{\mathcal{C}}}M_{R}(\nu_{R}) + H.c. \tag{2.22}$$

Finally, the most general neutrino mass term consists of all three types of mass terms, i.e. Dirac, left- and right-handed Majorana mass terms

$$\mathcal{L}_{m} = -\frac{1}{2}\bar{\nu}_{L}M_{L}(\nu_{L})^{\mathcal{C}} - \bar{\nu}_{L}M_{D}\nu_{R} - \frac{1}{2}\overline{(\nu_{R})^{\mathcal{C}}}M_{R}(\nu_{R}) + H.c. \tag{2.23}$$

where M_L is a complex 3×3 matrix, M_D is a complex $3 \times N_R$ matrix and M_R is a complex $N_R \times N_R$ matrix. This term can be written in the following compact form

$$\mathcal{L}_m = -\frac{1}{2}\overline{(n_L)^{\mathcal{C}}} M n_L + H.c. \tag{2.24}$$

where $n_L = (\nu_L, (\nu_R)^{\mathcal{C}})^T$, and

$$M = \begin{pmatrix} M_L & M_D \\ M_D^T & M_R \end{pmatrix}. \tag{2.25}$$

The M matrix is a $(3+N_R)\times(3+N_R)$ complex symmetric matrix. As in the case of the Majorana mass matrix, the M can be diagonalized by the congruence transformation

$$M = UmU^T, \tag{2.26}$$

where U is $(3+N_R) \times (3+N_R)$ unitary matrix and $m = diag(m_1, \ldots, m_{3+N_R})$.

2.2 Modelling the neutrino sector

The introduction of neutrino masses requires abandoning the two-component neutrino theory and accepting either the existence of right-handed neutrino fields or the Majorana nature of neutrinos. The theory of massive neutrinos is in the limelight of particle physics research and there is a variety of models trying to explain the origin of neutrino masses. We will focus on the models based on the Dirac-Majorana mass term (2.24) as an interesting, general framework. Two main classes of possible extensions of the neutrino sector dictated by the renormalizability and gauge symmetries of the SM will be discussed. Apart from Dirac or Majorana neutrino types, there are also pseudo-Dirac (or quasi-Dirac) [51], schizophrenic [52], or vanilla [53] neutrinos, to name some of them. Popular seesaw mechanisms give a possibility for a dynamical explanation of why the known active neutrino states are so light. They appear to be of Majorana type (recently, a dynamical explanation for Dirac light neutrinos was proposed [54]). By including more types of new fields we can approach neutrino masses by inverse or linear seesaw models [55–58].

2.2.1 Minimal extension of the Standard Model

To stay within the SM framework, i.e. keeping the Lagrangian invariant under the $SU(2)_L \times U(1)$ gauge group with the SM fields, we must abandon the left-handed Majorana mass term in (2.23) since it does not leave the Lagrangian invariant under the

SM gauge symmetry. This imposes that $M_L = 0$. On the other hand, as right-handed neutrino fields are singlets of the SM gauge symmetry, the right-handed Majorana mass term is allowed. This setup we call the minimal extension of the SM. Thus, the most general neutrino mass term compatible with the SM is given by

$$\mathcal{L}_m = -\bar{\nu}_L M_D \nu_R - \frac{1}{2}\overline{(\nu_R)^{\mathcal{C}}} M_R (\nu_R) + H.c. \tag{2.27}$$

This Lagrangian can be written in a compact form as

$$\mathcal{L}_m = -\frac{1}{2}\overline{(n_L)^{\mathcal{C}}} M n_L + H.c. \tag{2.28}$$

where

$$\begin{aligned} n_L &= (\nu_L, (\nu_R)^{\mathcal{C}})^T, \\ M &= \begin{pmatrix} 0 & M_D \\ M_D^T & M_R \end{pmatrix}. \end{aligned} \tag{2.29}$$

The complex symmetric mass matrix M can be diagonalized by the congruence transformation $U^T M U = m$, where $m = diag(m_1, m_2, m_3, ..., m_{3+N_R})$. The unitary matrix U is responsible for the transition between flavour and massive neutrino bases, i.e. the mixing mechanism for neutrinos

$$n_L = U\tilde{n}_L \quad \text{with} \quad \tilde{n}_L = (\tilde{n}_{1L}, \tilde{n}_{1L}, ..., \tilde{n}_{3+N_R L}). \tag{2.30}$$

The fields $\tilde{n}_{iL}$ are neutrino fields with definite mass m_i for $i = 1, 2, \ldots, 3 + N_r$. The current experimental data [59] ensures that at least two standard neutrinos must be massive

$$\begin{aligned} \Delta m_{21}^2 &= (7.53 \pm 0.18) \times 10^{-5} eV^2, \\ \Delta m_{32}^2 &= (2.453 \pm 0.034) \times 10^{-3} eV^2. \end{aligned} \tag{2.31}$$

This data restricts the minimal number of additional sterile neutrinos allowed in the minimal extension of the SM. If we add only one right-handed neutrino the mass matrix (2.29) takes the form

$$M = \begin{pmatrix} 0 & 0 & 0 & m_{14} \\ 0 & 0 & 0 & m_{24} \\ 0 & 0 & 0 & m_{34} \\ m_{14} & m_{24} & m_{34} & m_{44} \end{pmatrix}. \tag{2.32}$$

Such a matrix has two eigenvalues equal to zero. It is easily seen by solving the eigenvalue problem for the M matrix, $Mx = \lambda x$, which gives

$$\lambda^4 - (m_{14}^2\lambda^2 + m_{24}^2\lambda^2 + m_{34}^2\lambda^2 + m_{44}\lambda^3) = -\lambda^2(m_{14}^2 + m_{24}^2 + m_{34}^2 + m_{44}\lambda - \lambda^2) = 0, \tag{2.33}$$

which is true only if $\lambda^2 = 0$ or $m_{14}^2+m_{24}^2+m_{34}^2+m_{44}\lambda-\lambda^2 = 0$. From the first condition, we see that two eigenvalues are equal to zero which contradicts experimental results (2.31). The same is true in the general setting, where masses are taken to be singular values. Thus, the minimally extended SM defined here excludes the possibility that only one additional right-handed neutrino exists and the allowed minimal number of sterile neutrinos is two. The fact that the M_L submatrix is equal to zero in the minimal extension restricts also the structure of the mixing matrix.

2.2.2 Beyond the minimal SM extension

In order to allow the most general structure of the neutrino mass matrix (2.25), i.e. $M_L \neq 0$, it is necessary to go beyond the SM framework. The simplest way to do this is by introducing a weak isospin triplet $\Delta = (\Delta^{++}, \Delta^{+}, \Delta^{0})$ into the model. The effects of interaction between new fields such as the isospin triplet and SM fields can be incorporated into a model by an effective Lagrangian of dimension 5 consisting only of SM fields [60]

$$\mathcal{L}_5 = -\frac{1}{\Lambda}\sum_{l'l} y_{l'l}(\psi_{l'L}^T\sigma_2\phi)\mathcal{C}^\dagger(\phi^T\sigma_2\psi_{lL}) + H.c. \tag{2.34}$$

where $\psi_{lL} = (\nu_{lL}, l_L)^T$ and $\phi = (\phi^+, \phi^0)^T$ are lepton and Higgs doublets, respectively. The $\mathcal{L}_5$ operator is invariant under the SM symmetries, however, it is also non-renormalizable. It can be treated as an effective low-energy Lagrangian of the high-energy physics, which can be generated by integrating out heavy fields. At the tree level, this can be done in only three ways, one of them involves heavy-scalar triplet Δ. After the symmetry breaking $\mathcal{L}_5$ gives

$$\mathcal{L}^M = -\frac{1}{2}\sum_{l'l} \overline{(\nu_{l'L})^\mathcal{C}} M_{l'l}\nu_{lL} + H.c. \tag{2.35}$$

where $M_{l'l} = \frac{y_{l'l}v^2}{\Lambda}$, which is exactly the Majorana mass term for neutrinos. Thus, by adding such a term to the Dirac mass term and right-handed Majorana mass term, we recover the complete mass matrix (2.25). In that way, we can consider scenarios with only one sterile right-handed neutrino (three active and one sterile neutrinos) which are not allowed in the minimal extension of the SM. The 3+1 scenario is promoted by some oscillation experiments [61]. As mentioned the $\mathcal{L}_5$ operator can be realized in three different ways by involving only one type of new heavy fields. The one way has been discussed above and it involves heavy scalar triplet Δ. All these realizations provide the small mass spectrum of the known neutrinos. The realization involving the Δ is known as the seesaw type II mechanism [10, 62]. The other two are known

as seesaw type I [7–10] and III [63]. Seesaw type I is generated by the introduction of heavy-sterile neutrino fields ν_R, whereas the seesaw type III is generated by the heavy fermion triplet $\Sigma = (\Sigma^+, \Sigma^0, \Sigma^-)$. A hybrid mechanism is also possible [64].

2.3 Neutrino mixing matrix

Neutrino flavour fields are (linear) combinations of the massive fields

$$\nu_l^f = \sum_{i=1} U_{li} \nu_i^m \tag{2.36}$$

This property of neutrino fields is called the neutrino mixing mechanism. The mixing of neutrinos occurs regardless if they are Dirac or Majorana particles. As the massive and flavour fields form two orthogonal bases in the state space, the transition from one base to another can be done by the unitary matrix. This restricts coefficients of the linear combination, the sum of squares of their absolute values must equal one

$$\nu_l^f = \sum_{i=1} U_{li} \nu_i^m \text{ with } \sum_{i=1} |U_{li}|^2 = 1. \tag{2.37}$$

This unitary matrix is a major object of study in the theory of massive neutrinos. In the Standard Model, it is known as the PMNS mixing matrix.

It is useful to study the neutrino mixing matrix via a specific parametrization. The general $n \times n$ complex matrix has n^2 complex parameters or equivalently $2n^2$ real parameters. The unitarity condition $UU^\dagger = I$ imposes n^2 additional constraints on the elements. It can be seen from the $UU^\dagger$ which is a Hermitian matrix and has n independent diagonal elements and $n^2 - n$ independent off-diagonal elements which together give n^2 independent elements or conditions imposed on the unitary matrix. Thus, the $n \times n$ unitary matrix has $2n^2 - n^2 = n^2$ independent real parameters. An alternative way to see this is by writing a unitary matrix as the matrix exponent of the Hermitian matrix $U = e^{iH}$, where the H matrix is Hermitian and thus has n^2 independent real parameters which implies that U also has n^2 independent real parameters. These parameters can be split into two categories: rotation angles and complex phases. The number of angles corresponds to the number of parameters of the orthogonal matrix which has $\frac{n(n-1)}{2}$ independent real parameters. The remaining parameters correspond to phases. Thus, the n^2 independent real parameters of the unitary matrix split into

$$\begin{aligned} &\text{angles:} && \frac{n(n-1)}{2}, \\ &\text{phases:} && \frac{n(n+1)}{2}. \end{aligned} \tag{2.38}$$

However, not all phases are physical observables. The charged leptons and neutrino fields can be redefined as

$$\nu_i \to e^{i\alpha_i}\nu_i \text{ and } l \to e^{i\beta_l} l. \tag{2.39}$$

The α_i and β_l phases can be chosen in such a way that they eliminate $2n-1$ phases from the mixing matrix leaving the Lagrangian invariant. This reduces the number of phases of the mixing matrix. The number of remaining free parameters is $(n-1)^2$ which divides into

$$\begin{aligned} &\text{angles:} \quad \frac{n(n-1)}{2}, \\ &\text{phases:} \quad \frac{(n-1)(n-2)}{2}. \end{aligned} \tag{2.40}$$

These are the numbers under consideration when neutrinos are of the Dirac type. However, we know already that neutrinos can also be particles of the Majorana type. Then the Majorana condition (2.20) fixes phases of the neutrino fields, which no longer can be chosen to eliminate phases in the mixing matrix. On the other hand, the phases of charged leptons are still arbitrary and can be chosen in such a way as to eliminate phases from the mixing matrix. Thus, from all $\frac{n(n+1)}{2}$ phases of the unitary matrix, n phases can be eliminated. Finally, for the Majorana neutrinos, the number of free parameters of the mixing matrix is as follows

$$\begin{aligned} &\text{angles:} \quad \frac{n(n-1)}{2}, \\ &\text{phases:} \quad \frac{n(n-1)}{2}. \end{aligned} \tag{2.41}$$

Knowing the number of parameters necessary to describe the mixing matrix, we can find its explicit form by invoking a particular parametrization. In the SM, the mixing matrix is a 3×3 matrix and thus for the Dirac case we have three mixing angles and one complex phase. The standard way of parametrizing the PMNS mixing matrix is as the product of three rotation matrices with additional complex phase in one of them, i.e. in terms of Euler angles θ_{12}, θ_{13}, θ_{23} and complex phase δ

$$\begin{aligned} U_{\text{PMNS}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &\equiv \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & u_{\tau 3} \end{pmatrix}. \end{aligned} \tag{2.42}$$

In the case of Majorana neutrinos we must include additional phases, which is done typically by multiplying the PMNS mixing matrix from the right-hand side by the diagonal matrix of phases P^M. For the 3×3 mixing matrix, we must add two more complex phases. The Majorana neutrino mixing matrix is then given by

$$U^M_{\texttt{PMNS}} = U_{\texttt{PMNS}} P^M, \tag{2.43}$$

where $P^M = diag(e^{i\gamma_1}, e^{i\gamma_2}, 1)$.

2.4 CP transformation in the neutrino sector

The decays of neutral kaons show that CP symmetry is violated by weak interactions in the quark sector of elementary particles. In principle, strong interactions could violate CP symmetry, but the size of the CP violation found there is insufficient to explain the matter- antimatter imbalance. However, leptogenesis with hypothetical heavy Majorana neutrinos and their matter-antimatter asymmetric decays can produce desired effects. Thus, it is very important to study possible CP effects in the neutrino sector. As we will see, CP transformations affect possible parametrization of the neutrino mixing and mass matrices.

2.4.1 C, P and CP transformations

The C charge conjugation transformation is responsible for the particle to anti-particle transition. The spinor field $\psi(x)$ under the C transformation changes as follows

$$\psi(x) \xrightarrow{\mathcal{C}} \psi^{\mathcal{C}}(x) = \zeta_{\mathcal{C}} \mathcal{C} \overline{\psi}^T(x) = \zeta_{\mathcal{C}} \mathcal{C} \gamma_0^T \psi^*(x) = -\zeta_{\mathcal{C}} \gamma_0 \mathcal{C} \psi^*(x), \tag{2.44}$$

where $\mathcal{C}$ is a charge conjugation matrix satisfying

$$\begin{aligned} &\mathcal{C}\gamma_\mu^T \mathcal{C}^{-1} = -\gamma_\mu, \\ &\mathcal{C}^\dagger = \mathcal{C}^{-1}, \\ &\mathcal{C}^T = -\mathcal{C}. \end{aligned} \tag{2.45}$$

The coefficient $\zeta_{\mathcal{C}}$, which is a phase, is restricted by the fact that two consecutive charge conjugation transformations must leave the field unchanged

$$\psi(x) \xrightarrow{\mathcal{C}} \zeta_{\mathcal{C}} \mathcal{C} \overline{\psi}^T(x) \xrightarrow{\mathcal{C}} \zeta_{\mathcal{C}} \mathcal{C} (-\zeta_{\mathcal{C}}^* \psi^T \mathcal{C}^\dagger(x))^T = -|\zeta_{\mathcal{C}}|^2 \mathcal{C}\mathcal{C}^* \psi(x) = |\zeta_{\mathcal{C}}|^2 \psi(x). \tag{2.46}$$

Thus, $\zeta_{\mathcal{C}}$ must satisfy $|\zeta_{\mathcal{C}}|^2 = 1$.

The P transformation or parity transformation is responsible for the spatial reflection

$$x^\mu = (x^0, x) \xrightarrow{\mathcal{P}} x'^\mu = (x^0, -x) = x_\mu. \tag{2.47}$$

The spinor field $\psi(x)$ under the P transformation is transformed as

$$\psi(x) \xrightarrow{\mathcal{P}} \psi^{\mathcal{P}}(x') = \zeta_{\mathcal{P}} \gamma^0 \psi(x'). \tag{2.48}$$

The phase $\zeta_{\mathcal{P}}$ is constrained by the fact that two parity transformations must reproduce the initial state of the field

$$\psi(x) \xrightarrow{\mathcal{P}} \zeta_{\mathcal{P}} \gamma^0 \psi(x') \xrightarrow{\mathcal{P}} \zeta_{\mathcal{P}}^2 \gamma^0 \gamma^0 \psi(x) = \zeta_{\mathcal{P}}^2 \psi(x), \tag{2.49}$$

which gives $\zeta^2 = 1$. However, as the sign of fermion field changes with the rotation by 2π, the values of ζ^2 are ± 1 which gives $\zeta_{\mathcal{P}} = \pm 1, \pm i$.

The combination of C and P transformations is called the CP transformation. The fermion field under the CP transformation changes as follows

$$\psi(x) \xrightarrow{\mathcal{CP}} \zeta_{\mathcal{C}} \zeta_{\mathcal{P}} \gamma^0 \mathcal{C} \overline{\psi}^T(x'). \tag{2.50}$$

We will call the product of coefficients $\zeta_{\mathcal{C}}$ and $\zeta_{\mathcal{P}}$ the CP phase and denote it by $\zeta_{\mathcal{CP}}$, i.e. $\zeta_{\mathcal{CP}} = \zeta_{\mathcal{C}} \zeta_{\mathcal{P}}$. From the already established restrictions on $\zeta_{\mathcal{C}}$ and $\zeta_{\mathcal{P}}$, the CP phase must satisfy

$$|\zeta_{\mathcal{CP}}|^2 = 1. \tag{2.51}$$

To check if it is true, we use the fact that two successive CP transformations must reproduce the initial field

$$\psi(x) \xrightarrow{\mathcal{CP}} \zeta_{\mathcal{CP}} \gamma^0 \mathcal{C} \overline{\psi}^T(x') \xrightarrow{\mathcal{CP}} -|\zeta_{\mathcal{CP}}|^2 \psi(x). \tag{2.52}$$

Since the sign of the field has no physical meaning, we get $|\zeta_{\mathcal{CP}}|^2 = 1$.

2.4.2 CP conservation

Let us discuss conditions for the CP invariance of the neutrino sector by investigating the charged-current (CC) Lagrangian

$$\mathcal{L}_L^{CC} = -\frac{g}{\sqrt{2}} \left(\sum_{\alpha=e,\mu,\tau} \bar{\nu}_{\alpha L} \gamma^\mu l_{\alpha L} W_\mu + \sum_{\alpha=e,\mu,\tau} \bar{l}_{\alpha L} \gamma^\mu \nu_{\alpha L} W_\mu^\dagger \right). \tag{2.53}$$

It is necessary to discuss the CP invariance condition for both Dirac and Majorana type of neutrinos as the Majorana condition (2.20) imposes additional restrictions. Let us first investigate the Dirac case.

The CP transformation (2.50) is applied to the fermion fields by the unitary operator $U_{\mathcal{CP}}$

$$U_{\mathcal{CP}}\psi(x)U_{\mathcal{CP}}^{-1} = \zeta_{\mathcal{CP}}\gamma^0\mathcal{C}\overline{\psi}^T(x'). \tag{2.54}$$

The conjugate field transforms then as

$$U_{\mathcal{CP}}\overline{\psi}(x)U_{\mathcal{CP}}^{-1} = -\zeta_{\mathcal{CP}}^*\psi^T\mathcal{C}^\dagger\gamma^0(x'). \tag{2.55}$$

As the fields are the sum of left- and right-handed fields, these components transform in the same way

$$\begin{aligned} U_{\mathcal{CP}}\psi_{L,R}(x)U_{\mathcal{CP}}^{-1} &= \zeta_{\mathcal{CP}}\gamma^0\mathcal{C}\overline{\psi_{L,R}}^T(x'), \\ U_{\mathcal{CP}}\overline{\psi_{L,R}}(x)U_{\mathcal{CP}}^{-1} &= -\zeta_{\mathcal{CP}}^*\psi_{L,R}^T\mathcal{C}^\dagger\gamma^0(x'). \end{aligned} \tag{2.56}$$

Thus, we require the $\mathcal{L}_L^{CC}$ to be invariant under the CP transformation

$$U_{\mathcal{CP}}\mathcal{L}_L^{CC}(x)U_{\mathcal{CP}}^{-1} = \mathcal{L}_L^{CC}(x'). \tag{2.57}$$

The charged-current Lagrangian can be expressed in terms of massive neutrino fields

$$\mathcal{L}_L^{CC} = -\frac{g}{\sqrt{2}}\left(\sum_{\alpha=e,\mu,\tau}\sum_i U_{\alpha i}^*\bar{\nu}_{iL}\gamma^\mu l_{\alpha L}W_\mu + \sum_{\alpha=e,\mu,\tau}\sum_i U_{\alpha i}\bar{l}_{\alpha L}\gamma^\mu\nu_{iL}W_\mu^\dagger\right). \tag{2.58}$$

Under the CP transformation, the W boson transforms as

$$U_{\mathcal{CP}}W_\mu(x)U_{\mathcal{CP}}^{-1} = e^{i\xi_W}W^{\mu\dagger}(x'), \tag{2.59}$$

which gives the following CC current Lagrangian transformation

$$\begin{aligned} &U_{\mathcal{CP}}\mathcal{L}_L^{CC}U_{\mathcal{CP}}^{-1} = \\ &= \frac{g}{\sqrt{2}}\sum_{\alpha,i}U_{\alpha i}^*\eta_i^*\nu_{iL}^T\mathcal{C}^\dagger\gamma^0\gamma^\mu\zeta_\alpha\gamma^0\mathcal{C}\bar{l}_{\alpha L}^T e^{i\xi_W}W^{\mu\dagger} + \frac{g}{\sqrt{2}}\sum_{\alpha,i}U_{\alpha i}\zeta_\alpha^* l_{\alpha L}^T\mathcal{C}^\dagger\gamma^0\gamma^\mu\eta_i\gamma^0\mathcal{C}\bar{\nu}_{iL}^T e^{-i\xi_W}W^\mu \\ &= -\frac{g}{\sqrt{2}}\sum_{\alpha,i}U_{\alpha i}^*\eta_i^*\zeta_\alpha e^{i\xi_W}\nu_{iL}^T\gamma^{\mu *}\bar{l}_{\alpha L}^T W^{\mu\dagger} - \frac{g}{\sqrt{2}}\sum_{\alpha,i}U_{\alpha i}\zeta_\alpha^*\eta_i e^{-i\xi_W}l_{\alpha L}^T\gamma^{\mu *}\bar{\nu}_{iL}^T W^\mu = \\ &= \frac{g}{\sqrt{2}}\sum_{\alpha,i}U_{\alpha i}^*\eta_i^*\zeta_\alpha e^{i\xi_W}\bar{l}_{\alpha L}\gamma_\mu\nu_{iL}W^{\mu\dagger} + \frac{g}{\sqrt{2}}\sum_{\alpha,i}U_{\alpha i}\zeta_\alpha^*\eta_i e^{-i\xi_W}\bar{\nu}_{iL}\gamma_\mu l_{\alpha L}W^\mu, \end{aligned} \tag{2.60}$$

where the change of the sign in the second line is due to the properties of the $\mathcal{C}$ matrix (2.45), and in the last row owing to the anti-commutation of fermion fields. As the phase $e^{i\xi_W}$ of the W boson is not a physical observable and is arbitrary, we can choose it to be equal to one, which is achieved by taking $\xi_W = \pi$, this gives

$$U_{\mathcal{CP}}\mathcal{L}_L^{CC}U_{\mathcal{CP}}^{-1} = -\frac{g}{\sqrt{2}}\sum_{\alpha,i}U_{\alpha i}^*\eta_i^*\zeta_\alpha\bar{l}_{\alpha L}\gamma^\mu\nu_{iL}W_\mu^\dagger - \frac{g}{\sqrt{2}}\sum_{\alpha,i}U_{\alpha i}\zeta_\alpha^*\eta_i\bar{\nu}_{iL}\gamma^\mu l_{\alpha L}W_\mu. \tag{2.61}$$

Comparing this with the initial Lagrangian, we see that to maintain CP invariance the mixing matrix must satisfy

$$U^*_{\alpha i} = \eta_i \zeta^*_\alpha U_{\alpha i} \Leftrightarrow U^*_{\alpha i} = e^{-i\zeta_\alpha} U_{\alpha i} e^{i\eta_i}. \tag{2.62}$$

As the phases of lepton fields are not observable, we can set up $e^{-i\zeta_\alpha} = e^{i\eta_i} = 1$, this gives

$$U^*_{\alpha i} = U_{\alpha i}. \tag{2.63}$$

From this condition follows that the CC Lagrangian is invariant under CP transformation if the mixing matrix is a real orthogonal matrix

$$U^*_{\alpha i} = U_{\alpha i} = O_{\alpha i}. \tag{2.64}$$

For the Majorana neutrinos situation is slightly different. The neutrino field with definite mass transforms under the CP transformation as

$$U_{\mathcal{CP}} \nu(x) U^{-1}_{\mathcal{CP}} = \zeta_{\mathcal{CP}} \gamma^0 \mathcal{C} \bar{\nu}^T(x') \tag{2.65}$$

Then, the charge conjugated neutrino field $\nu(x)^{\mathcal{C}} = \mathcal{C}\bar{\nu}^T(x)$ transforms as

$$U_{\mathcal{CP}} \nu^{\mathcal{C}}(x) U^{-1}_{\mathcal{CP}} = -\zeta^*_{\mathcal{CP}} \gamma^0 \nu(x'). \tag{2.66}$$

If $\nu(x)$ is a Majorana field, i.e. $\nu(x) = \nu^{\mathcal{C}}(x)$ from (2.65) and (2.66) we have

$$\begin{aligned} U_{\mathcal{CP}} \nu(x) U^{-1}_{\mathcal{CP}} &= \zeta_{\mathcal{CP}} \gamma^0 \nu^{\mathcal{C}}(x') = \zeta_{\mathcal{CP}} \gamma^0 \nu(x'), \\ U_{\mathcal{CP}} \nu^{\mathcal{C}}(x) U^{-1}_{\mathcal{CP}} &= U_{\mathcal{CP}} \nu(x) U^{-1}_{\mathcal{CP}} = -\zeta^*_{\mathcal{CP}} \gamma^0 \nu(x'). \end{aligned} \tag{2.67}$$

By comparing these two relations we get

$$\zeta^*_{\mathcal{CP}} = -\zeta_{\mathcal{CP}}, \tag{2.68}$$

which means that $\zeta_{\mathcal{CP}}$ in the Majorna case is purely imaginary and combining this with (2.51) we establish that

$$\zeta_{\mathcal{CP}} = \pm i. \tag{2.69}$$

Thus, in comparison with the Dirac case (2.51), the CP phase for Majorana neutrinos is not arbitrary.

Let us now study the CP invariance of the $\mathcal{L}^{CC}$ to obtain the necessary conditions for the invariance under the CP transformation in the case of Majorana neutrinos. The left-handed massive Majorana field transforms as

$$U_{\mathcal{CP}} \nu_{kL}(x) U^{-1}_{\mathcal{CP}} = \rho_k i \gamma^0 \mathcal{C} \bar{\nu}^T_{kL}(x'), \tag{2.70}$$

where $\rho_k = \pm 1$. Thus the CC interaction Lagrangian

$$\mathcal{L}_L^{CC} = -\frac{g}{\sqrt{2}} \left(\sum_{\alpha=e,\mu,\tau} \sum_i U_{\alpha i}^* \bar{\nu}_{iL} \gamma^\mu l_{\alpha L} W_\mu + \sum_{\alpha=e,\mu,\tau} \sum_i U_{\alpha i} \bar{l}_{\alpha L} \gamma^\mu \nu_{iL} W_\mu^\dagger \right). \tag{2.71}$$

is transformed into

$$\begin{aligned} U_{\mathcal{CP}} \mathcal{L}_L^{CC} U_{\mathcal{CP}}^{-1} = & -\frac{g}{\sqrt{2}} \sum_{\alpha,i} U_{\alpha i}^* \rho_i i \nu_{iL}^T \mathcal{C}^\dagger \gamma^0 \gamma^\mu \zeta_\alpha \gamma^0 \mathcal{C} \bar{l}_{\alpha L}^T e^{i\xi_W} W^{\mu\dagger} + \\ & \frac{g}{\sqrt{2}} \sum_{\alpha,i} U_{\alpha i} \zeta_\alpha^* l_{\alpha L}^T \mathcal{C}^\dagger \gamma^0 \gamma^\mu \rho_i i \gamma^0 \mathcal{C} \bar{\nu}_{iL}^T e^{-i\xi_W} W^\mu \end{aligned} \tag{2.72}$$

After similar steps as for the Dirac case, we finally obtain

$$U_{\mathcal{CP}} \mathcal{L}_L^{CC} U_{\mathcal{CP}}^{-1} = -\frac{g}{\sqrt{2}} \left(\sum_{\alpha,i} U_{\alpha i} \rho_i i \zeta_\alpha^* \bar{\nu}_{iL} \gamma^\mu l_{\alpha L} W_\mu - \sum_{\alpha,i} U_{\alpha i}^* \rho_i i \zeta_\alpha \bar{l}_{\alpha L} \gamma^\mu \nu_{iL} W_\mu^\dagger \right). \tag{2.73}$$

Comparing this with the initial Lagrangian, we get the following condition imposed on the neutrino mixing matrix providing CP conservation

$$i U_{\alpha i} \rho_i \zeta_\alpha^* = U_{\alpha i}^*. \tag{2.74}$$

The charged lepton CP phase ζ_α can be chosen arbitrary, and if we chose it to be equal $-i$, we get

$$U_{\alpha i} \rho_i = U_{\alpha i}^*. \tag{2.75}$$

With this choice of the leptonic CP phase, the mixing matrix can be either real for $\rho_i = 1$ *or purely imaginary for* $\rho_i = -1$.

So far we have discussed the CP invariance in the neutrino sector by considering the charged-current interaction. However, as the mixing matrix diagonalize the mass matrix, it is important to check whether the same conditions follow from the mass terms and what kind of restriction is imposed by the demand of the CP invariance on the neutrino mass matrix. The Dirac mass term

$$\mathcal{L}^D = -\sum_{l',l} \bar{\nu}_{l'L} M_{l'l} \nu_{lR} + H.c. \tag{2.76}$$

is invariant under the CP transformation if

$$U_{\mathcal{CP}} \mathcal{L}^D(x) U_{\mathcal{CP}}^{-1} = \mathcal{L}^D(x'). \tag{2.77}$$

After the CP transformation $\mathcal{L}^D$ looks as follows

$$\begin{aligned} U_{\mathcal{CP}} \mathcal{L}^D(x) U_{\mathcal{CP}}^{-1} &= -\sum_{l',l} (-\zeta_{l'}^*) \nu_{l'L}^T \mathcal{C}^\dagger \gamma^0 M_{l'l} \eta_l \gamma^0 \mathcal{C} \bar{\nu}_{lR}^T + H.c. \\ &= \sum_{l',l} \zeta_{l'}^* \eta_l \nu_{l'L}^T \mathcal{C}^\dagger \gamma^0 \gamma^0 \mathcal{C} M_{l'l} \bar{\nu}_{lR}^T + H.c. \\ &= \sum_{l',l} \zeta_{l'}^* \eta_l \nu_{l'L}^T M_{l'l} \bar{\nu}_{lR}^T + H.c. = -\sum_{l',l} \zeta_{l'}^* \eta_l \bar{\nu}_{lR} M_{ll'} \nu_{l'L} + H.c. \end{aligned} \tag{2.78}$$

By comparing it with the initial mass term it is invariant under the CP transformation if

$$\zeta_{l'}^{*}\eta_{l}M_{ll'} = M_{ll'}^{*}. \tag{2.79}$$

The phases of the leptons are not physical observables, thus we can always set $\zeta_{l'}^{*} = \eta_{l} = 1$, which gives

$$M_{ll'} = M_{ll'}^{*}. \tag{2.80}$$

In other words, the CP symmetry of the Dirac mass term imposes that the mass matrix must be a real matrix. Any real matrix can be diagonalized by the singular value decomposition with orthogonal matrices

$$M = O'mO^{T}, \tag{2.81}$$

which means that if CP symmetry holds, then the mixing matrix is a real orthogonal matrix. What follows, we came to the same conclusion as by investigating the CP invariance of the CC Lagrangian.

Let us now consider Majorana neutrinos and suppose that the left-handed flavour neutrino fields transform under CP as given in (2.54) with the CP phase equals to i, i.e.

$$U_{\mathcal{CP}}\nu_{\alpha L}(x)U_{\mathcal{CP}}^{-1} = i\gamma^{0}\mathcal{C}\bar{\nu}_{\alpha L}^{T}(x'). \tag{2.82}$$

Together with the choice of not physically observable phases of charged leptons and W boson (equal to $-i$ and 1, respectively) the charged-current Lagrangian, in the flavour basis, is invariant under the CP transformation. The Majorna mass term

$$\mathcal{L}^{M} = -\frac{1}{2}\bar{\nu}_{L}M_{L}\nu_{L}^{\mathcal{C}} - \frac{1}{2}\overline{(\nu_{L}^{\mathcal{C}})}M_{L}^{\dagger}\nu_{L}, \tag{2.83}$$

transforms under the CP transformation as follows

$$\begin{aligned} U_{\mathcal{CP}}\mathcal{L}^{M}U_{\mathcal{CP}}^{-1} &= -\frac{1}{2}i\nu_{L}\mathcal{C}^{\dagger}\gamma^{0}M_{L}i\gamma^{0}\nu_{L} + H.c. = \frac{1}{2}\nu_{L}\mathcal{C}^{\dagger}\gamma^{0}\gamma^{0}M_{L}\nu_{L} + H.c. \\ &= -\frac{1}{2}\overline{(\nu_{L}^{\mathcal{C}})}M_{L}\nu_{L} + H.c. \end{aligned} \tag{2.84}$$

Comparing this result with the initial Lagrangian, we see that Majorana mass term is CP invariant if

$$M_{L} = M_{L}^{\dagger}. \tag{2.85}$$

Moreover, M_{L} is a symmetric matrix which implies that

$$M_{L} = M_{L}^{*}. \tag{2.86}$$

Thus, the Majorana mass matrix is a real symmetric matrix. This type of matrices can be diagonalized by the orthogonal similarity transformation

$$M_L = OmO^T. \tag{2.87}$$

However, eigenvalues of the real symmetric matrix can be either positive or negative

$$m_k = |m_k|\rho_k, \tag{2.88}$$

where $\rho_k = \pm 1$. The sign factor ρ_k can be incorporated into the diagonalizing matrix resulting in positive masses and giving the same condition as established from inspection of the CP invariance of the CC Lagrangian (2.75).

When sterile neutrinos are taken into account, the mixing and mass matrices have a dimension larger than three and hence the number of complex phases increases accordingly (2.41). As the effects of CP can be visible in physical processes [65, 66] it is crucial to provide a method of studying these extended complex mass and mixing matrices.

2.4.3 Rephasing invariants

Let $U \in \mathbb{M}_{3\times 3}$ be unitary matrix with elements $U_{\alpha j}$ $\alpha, j = 1, \ldots, 3$. Let us seek for the invariants of the following phase transformation

$$U_{\alpha j} \to e^{-\phi_\alpha} U_{\alpha j} e^{\phi_j}. \tag{2.89}$$

It is immediate that invariants of this transformation must contain the same amount of elements with indices i, j as their complex conjugate. Thus, the simplest of such invariants is the square of the absolute value of the element $U_{\alpha j}$

$$|U_{\alpha j}|^2 = U_{ij}U_{ij}^* \to e^{-\phi_i} U_{ij} e^{\phi_j} e^{\phi_i} U_{ij}^* e^{-\phi_j} = U_{ij}U_{ij}^*. \tag{2.90}$$

The next invariant consists of four elements and will be called a square invariant

$$^{\alpha j}\square_{\beta k} \equiv U_{\alpha j} U_{\beta k} U_{\alpha k}^* U_{\beta j}^*. \tag{2.91}$$

The box ($\square$) notation has been established in [67]. Each more complicated invariant can be constructed from these two basic types of invariants. The square invariant behaves in the following way under the index interchange

$$^{\alpha j}\square_{\beta k} = ^{\beta k}\square_{\alpha j} = ^{\alpha k}\square_{\beta j}^* = ^{\beta j}\square_{\alpha k}^*. \tag{2.92}$$

Rephasing invariants are very useful in the study of mixing matrices since they are parametrization independent and reveal important properties without invoking any specific parametrization. Especially they provide information about the CP invariance. To see this let us look at what we can take out from the unitarity relation $UU^\dagger = I$

$$\sum_j U_{\alpha j} U^*_{\beta j} = \delta_{\alpha\beta}. \tag{2.93}$$

Let us multiply this relation by $U^*_{\alpha k} U_{\beta k}$ for a given k

$$\sum_j U_{\alpha j} U^*_{\beta j} U^*_{\alpha k} U_{\beta k} = |U_{\alpha k}|^2 |U_{\beta k}|^2 + \sum_{j\neq k} U_{\alpha j} U^*_{\beta j} U^*_{\alpha k} U_{\beta k} = |U_{\alpha k}|^2 \delta_{\alpha\beta} \tag{2.94}$$

The imaginary part of this formula gives

$$\sum_{j\neq k} Im(U_{\alpha j} U^*_{\beta j} U^*_{\alpha k} U_{\beta k}) = 0. \tag{2.95}$$

The second unitarity condition gives

$$\sum_{\alpha\neq\beta} Im(U_{\alpha j} U^*_{\beta j} U^*_{\alpha k} U_{\beta k}) = 0. \tag{2.96}$$

These two relations impose that the imaginary parts of square rephasing invariants are equal up to sign. Thus, all of them can be expressed by the so-called Jarlskog invariant [68–70]

$$J \equiv Im(U_{12} U_{23} U^*_{13} U^*_{22}). \tag{2.97}$$

The sign of the other square invariants is expressed by the Jarlskog invariant as

$$Im(U_{\alpha j} U_{\beta k} U^*_{\alpha k} U^*_{\beta j}) = J \sum_\gamma \epsilon_{\alpha\beta\gamma} \sum_l \epsilon_{jkl}. \tag{2.98}$$

2.5 Neutrino oscillations

The discovery of neutrino oscillations is one of the major achievements of particle physics. In the core of neutrino oscillation lies the mixing mechanism. To relate the discussed earlier theory of CP transformations and mass and mixing neutrino matrices with neutrino oscillation formalism and experimental data, we will sketch now a simplified plane wave description of the oscillation phenomenon which is sufficient for studies undertaken in this thesis. For more refined analysis, see e.g. [71–74].

Let the neutrino of definite flavour be created via the charged-current interaction process

$$\mathcal{L}^{CC}_L = -\frac{g}{2\sqrt{2}} \left(j^\mu_L W_\mu + j^{\mu\dagger}_L W^\dagger_\mu \right), \tag{2.99}$$

with the charged-current

$$j_L^\mu = 2 \sum_{\alpha=e,\mu,\tau} \bar{\nu}_{\alpha L} \gamma^\mu l_{\alpha L} = 2 \sum_{\alpha=e,\mu,\tau} \sum_i U_{\alpha i}^* \bar{\nu}_{iL} \gamma^\mu l_{\alpha L}. \tag{2.100}$$

The field operator $\bar{\nu}_{iL}$ creates the neutrino with mass m_i whereas $l_{\alpha L}$ creates either lepton l^- or anti-lepton l^+. Thus in the CC process neutrino is either created for the lepton or in a pair with the anti-lepton. In the process, the neutrino of definite flavour is created and is defined as

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i}^* |\nu_i\rangle \quad \text{for } \alpha = e, \mu, \tau. \tag{2.101}$$

The massive neutrino states $|\nu_\alpha\rangle$ are eigenstates of the Hamiltonian $\mathcal{H}$

$$\mathcal{H}|\nu_i\rangle = E_i |\nu_i\rangle, \tag{2.102}$$

with eigenvalues $E_k = \sqrt{p^2 + m_k^2}$. The time evolution of the massive state is governed by the Schrödinger equation

$$i \frac{d}{dt} |\nu_i(t)\rangle = \mathcal{H} |\nu_i(t)\rangle \tag{2.103}$$

which gives the plane wave solution $|\nu_i(t)\rangle = e^{-iE_i t} |\nu_i\rangle$. From (2.101) we get the time evolution of the flavour state

$$|\nu_\alpha(t)\rangle = \sum_i U_{\alpha i}^* e^{-iE_i t} |\nu_i\rangle \tag{2.104}$$

with $|\nu_\alpha(t=0)\rangle = |\nu_\alpha\rangle$. By using the unitarity of the mixing matrix U we can express massive states as the linear combination of the flavour states

$$|\nu_i\rangle = \sum_\alpha U_{\alpha i} |\nu_\alpha\rangle \tag{2.105}$$

which can be inputted into the time evolution of the flavour state

$$|\nu_\alpha(t)\rangle = \sum_{\beta=e,\mu,\tau} \sum_i U_{\alpha i}^* e^{-iE_i t} U_{\beta i} |\nu_\beta\rangle. \tag{2.106}$$

Thus, in the time evolution, the pure flavour state (t=0) becomes the combination of all the flavour states. Then the transition amplitude for the evolution of a given pure flavour state to another state is given by

$$\begin{aligned} \mathcal{A}_{\nu_\alpha \to \nu_\beta}(t) = \langle \nu_\beta | \nu_\alpha(t) \rangle = \sum_{\rho=e,\mu,\tau} \sum_i U_{\alpha i}^* e^{-iE_i t} U_{\rho i} \langle \nu_\beta | \nu_\rho \rangle = \\ = \sum_{\rho=e,\mu,\tau} \sum_i U_{\alpha i}^* e^{-iE_i t} U_{\rho i} \delta_{\beta\rho} = \sum_i U_{\alpha i}^* U_{\beta i} e^{-iE_i t}. \end{aligned} \tag{2.107}$$

This gives the probability of transition from the flavour state α to β as

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta}(L,E)(t) = |\mathcal{A}_{\nu_\alpha \to \nu_\beta}(t)|^2 = \sum_{i,j} U^*_{\alpha i} U_{\beta i} e^{-iE_i t} U_{\alpha j} U^*_{\beta j} e^{iE_j t} = \\ = \sum_{i,j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} e^{-i(E_i - E_j)t} \end{aligned} \tag{2.108}$$

For the ultra-relativistic neutrinos when the momentum is much larger than mass, energy eigenvalues can be approximated as

$$E_k \simeq E + \frac{m_k^2}{E}, \tag{2.109}$$

where $E = |p|$. Moreover, the time t equals almost the distance travelled by the neutrinos $t \simeq L$ which gives

$$P_{\nu_\alpha \to \nu_\beta}(L,E) = \sum_{i,j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} exp\left(-i\frac{\Delta m^2_{ij} L}{2E}\right), \tag{2.110}$$

where $\Delta m^2_{ij} \equiv m^2_i - m^2_j$. The probability of transition from one neutrino flavour to another becomes a function of the travelled distance and the neutrino energy. The probability in (2.110) can be rewritten as

$$P_{\nu_\alpha \to \nu_\beta}(L,E) = \sum_i |U_{\alpha i}|^2 |U_{\beta i}|^2 + 2Re \sum_{i>j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} exp\left(-i\frac{\Delta m^2_{ij} L}{2E}\right). \tag{2.111}$$

Further, we can use the following relation

$$\sum_i |U_{\alpha i}|^2 |U_{\beta i}|^2 = \delta_{\alpha\beta} - 2Re \sum_{i>j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} \tag{2.112}$$

which gives

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta}(L,E) = \delta_{\alpha\beta} - 2Re \sum_{i>j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} \left(1 - exp\left(-i\frac{\Delta m^2_{ij} L}{2E}\right)\right) = \\ \delta_{\alpha\beta} - 2Re \sum_{i>j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} \left(1 - \cos\left(\frac{\Delta m^2_{ij} L}{2E}\right) - i \sin\left(\frac{\Delta m^2_{ij} L}{2E}\right)\right) = \\ \delta_{\alpha\beta} - 2Re \sum_{i>j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} \left(1 - \cos\left(2\frac{\Delta m^2_{ij} L}{4E}\right) - i \sin\left(\frac{\Delta m^2_{ij} L}{2E}\right)\right) = \\ \delta_{\alpha\beta} - 2Re \sum_{i>j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} \left(2\sin^2\left(\frac{\Delta m^2_{ij} L}{4E}\right) - i \sin\left(\frac{\Delta m^2_{ij} L}{2E}\right)\right). \end{aligned} \tag{2.113}$$

This finally can be written as

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta}(L,E) = \delta_{\alpha\beta} - 4 \sum_{i>j} Re(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}) \sin^2\left(\frac{\Delta m^2_{ij} L}{4E}\right) \\ + 2 \sum_{i>j} Im(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}) \sin\left(\frac{\Delta m^2_{ij} L}{2E}\right). \end{aligned} \tag{2.114}$$

The transition probability can also be written in terms of square rephasing invariants ${}^{\alpha j}\square_{\beta i}$

$$P_{\nu_\alpha \to \nu_\beta}(L,E) = \delta_{\alpha\beta} - 4\sum_{i>j} Re({}^{\alpha j}\square_{\beta i}) \sin^2\left(\frac{\Delta m^2_{ij} L}{4E}\right) + 2\sum_{i>j} Im({}^{\alpha j}\square_{\beta i}) \sin\left(\frac{\Delta m^2_{ij} L}{2E}\right). \tag{2.115}$$

The antineutrino flavour state is defined as

$$|\bar{\nu}_\alpha\rangle = \sum_i U_{\alpha i} |\bar{\nu}_i\rangle \tag{2.116}$$

which gives the following transition probability for antineutrinos

$$P_{\bar{\nu}_\alpha \to \bar{\nu}_\beta}(L,E) = \sum_{i,j} U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j} exp\left(-i\frac{\Delta m^2_{ij} L}{2E}\right). \tag{2.117}$$

It differs from the neutrino probability by the complex conjugation of the square rephasing invariant and thus, after the similar steps as for neutrinos, it results in

$$P_{\nu_\alpha \to \nu_\beta}(L,E) = \delta_{\alpha\beta} - 4\sum_{i>j} Re(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}) \sin^2\left(\frac{\Delta m^2_{ij} L}{4E}\right) - 2\sum_{i>j} Im(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}) \sin\left(\frac{\Delta m^2_{ij} L}{2E}\right). \tag{2.118}$$

or in terms of the rephasing invariants

$$P_{\nu_\alpha \to \nu_\beta}(L,E) = \delta_{\alpha\beta} - 4\sum_{i>j} Re({}^{\alpha j}\square_{\beta i}) \sin^2\left(\frac{\Delta m^2_{ij} L}{4E}\right) - 2\sum_{i>j} Im({}^{\alpha j}\square_{\beta i}) \sin\left(\frac{\Delta m^2_{ij} L}{2E}\right). \tag{2.119}$$

Thus, we see that the CP violation is visible only through the imaginary term.

The oscillation experiments provide the major information about the structure of the neutrino mixing matrix. The current data gives the following limits for the mixing parameters [59, 75]

$$\begin{aligned} &\theta_{12} \in [31.27^\circ, 35.86^\circ], \quad &&\theta_{23} \in [40.1^\circ, 51.7^\circ], \\ &\theta_{13} \in [8.20^\circ, 8.93^\circ], &&\delta \in [120^\circ, 369^\circ]. \end{aligned} \tag{2.120}$$

By inputting these ranges into (2.42) we get allowed ranges for the mixing matrix elements [75] (at the 3σ confidence level)

$$|U|_{3\sigma} = \begin{pmatrix} [0.801, 0.845] & [0.513, 0.579] & [0.143, 0.155] \\ [0.243, 0.500] & [0.471, 0.689] & [0.637, 0.776] \\ [0.271, 0.525] & [0.477, 0.694] & [0.613, 0.756] \end{pmatrix}. \tag{2.121}$$

The exact values of the allowed ranges in the CP invariant case presented as the interval matrix are

$$U_{int} = \begin{pmatrix} [0.801, 0.845] & [0.513, 0.579] & [0.143, 0.155] \\ [-0.529, -0.417] & [0.431, 0.606] & [0.637, 0.776] \\ [0.233, 0.388] & [-0.721, -0.586] & [0.613, 0.756] \end{pmatrix}, \tag{2.122}$$

whereas when the non-zero CP phase δ is included, the elements of the U_{int} are within the following ranges

$$\begin{aligned} U_{e1} &\in [0.801, 0.845]\,, \\ U_{e2} &\in [0.513, 0.579]\,, \\ U_{e3} &\in [-0.155 - 0.155i, 0.155 + 0.134i]\,, \\ U_{\mu 1} &\in [-0.528 - 0.0901i, -0.218 + 0.104i]\,, \\ U_{\mu 2} &\in [0.432 - 0.0616i, 0.707 + 0.0711i]\,, \\ U_{\mu 3} &\in [0.637, 0.776]\,, \\ U_{\tau 1} &\in [0.233 - 0.0878i, 0.538 + 0.101i]\,, \\ U_{\tau 2} &\in [-0.721 - 0.060i, -0.453 + 0.0693i]\,, \\ U_{\tau 3} &\in [0.613, 0.756]\,. \end{aligned} \tag{2.123}$$

2.6 Non-unitarity of the PMNS mixing matrix

If sterile neutrinos exist, they can mix with the SM neutrinos. In such case the PMNS mixing matrix is no longer unitary, but is a part of a larger unitary matrix where the mixing between known neutrinos is preserved

$$\begin{pmatrix} \nu^f \\ \tilde{\nu}^f \end{pmatrix} = \begin{pmatrix} U_{\texttt{PMNS}} & U_{lh} \\ U_{hl} & U_{hh} \end{pmatrix} \begin{pmatrix} \nu^m \\ \tilde{\nu}^m \end{pmatrix}. \tag{2.124}$$

ν^f and ν^m are SM neutrinos in the flavour and massive basis respectively, and $\tilde{\nu}^f$ and $\tilde{\nu}^m$ are sterile neutrinos in both bases. So far, there is no restriction to the number of

sterile neutrinos which will be denoted by N_R. Thus, if we exclude the case where SM and sterile neutrino sectors are decoupled

$$\begin{pmatrix} \nu^f \\ \tilde{\nu}^f \end{pmatrix} = \begin{pmatrix} U_{\texttt{PMNS}} & 0 \\ 0 & U_{hh} \end{pmatrix} \begin{pmatrix} \nu^m \\ \tilde{\nu}^m \end{pmatrix}, \tag{2.125}$$

signals of sterile neutrinos should be visible as a deviation from unitarity of the $U_{\texttt{PMNS}}$ mixing matrix

$$\nu^f_\alpha = \underbrace{(U_{\texttt{PMNS}})_{\alpha i}\nu^m_i}_{\text{SM part}} + \underbrace{(U_{lh})_{\alpha j}\tilde{\nu}^m_j}_{\text{BSM part}}. \tag{2.126}$$

Hence, the study of non-unitarity of the PMNS mixing matrix is of unprecedented importance for particle physics. Presently there are two commonly used ways of describing the deviation from unitarity of the neutrino mixing matrix. These are known as the η- and α-parametrization [34–39]. In these approaches the mixing matrix is expressed by a product of a unitary matrix with the second matrix of some special type. The η-parametrization decomposes the mixing matrix into the product of the Hermitian and unitary matrix

$$U_{\texttt{PMNS}} = HU = (I - \eta)U, \tag{2.127}$$

where H is a Hermitian matrix and U is a unitary matrix. The current data shows that such a deviation from unitarity must be very small, thus it is common to represent the H part as $I - \eta$ where η is a Hermitian matrix representing a small deviation from the unitarity. On the other hand, the α-parametrization decomposes the mixing matrix into the product of the lower triangular and unitary matrix

$$U_{\texttt{PMNS}} = TU' = (I - \alpha)U', \tag{2.128}$$

where T is a lower triangular matrix and U' is a unitary matrix. Similarly to the η-parametrization, it is common to write the T term as $I - \alpha$, with α being lower triangular and parametrizing small deviation from unitarity of the PMNS mixing matrix. These two parametrizations can be related to each other as they decompose the same matrix [39]. Let us calculate the Hermitian product of the PMNS mixing matrix in (2.127)

$$U_{\texttt{PMNS}}U^\dagger_{\texttt{PMNS}} = (I - \eta)UU^\dagger(I - \eta)^\dagger = (I - \eta)(I - \eta) = I - 2\eta + \eta^2. \tag{2.129}$$

On the other hand, calculating the same product for (2.128), we get

$$U_{\texttt{PMNS}}U^\dagger_{\texttt{PMNS}} = (I - \alpha)U'(U')^\dagger(I - \alpha)^\dagger = (I - \alpha)(I - \alpha^\dagger) = I - \alpha - \alpha^\dagger + \alpha\alpha^\dagger. \tag{2.130}$$

By comparing (2.129) with (2.130) and neglecting the square terms in η and α, we get

$$2\eta \approx \alpha + \alpha^{\dagger}. \tag{2.131}$$

Let us write this in the explicit form

$$2\begin{pmatrix} \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{12}^* & \eta_{22} & \eta_{23} \\ \eta_{13}^* & \eta_{23}^* & \eta_{33} \end{pmatrix} \approx \begin{pmatrix} \alpha_{11} & \alpha_{21}^* & \alpha_{31}^* \\ \alpha_{21} & \alpha_{22} & \alpha_{32}^* \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}. \tag{2.132}$$

This gives us

$$\begin{aligned} &\alpha_{11} = \eta_{11}, \quad \alpha_{22} = \eta_{22}, \quad \alpha_{33} = \eta_{33}, \\ &\alpha_{21} = 2\eta_{12}^*, \quad \alpha_{31} = 2\eta_{13}^*, \quad \alpha_{32} = 2\eta_{23}^*. \end{aligned} \tag{2.133}$$

Thus, η and α matrices are related as

$$\begin{pmatrix} \alpha_{11} & 0 & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \approx \begin{pmatrix} \eta_{11} & 0 & 0 \\ 2\eta_{12}^* & \eta_{22} & 0 \\ 2\eta_{13}^* & 2\eta_{23}^* & \eta_{33} \end{pmatrix}. \tag{2.134}$$

The current bounds for the η matrix are as follows [39]

$$|\eta_{\alpha\beta}| \leq \begin{pmatrix} 1.3 \cdot 10^{-3} & 1.2 \cdot 10^{-5} & 1.4 \cdot 10^{-3} \\ 1.2 \cdot 10^{-5} & 2.2 \cdot 10^{-4} & 6.0 \cdot 10^{-4} \\ 1.4 \cdot 10^{-3} & 6.0 \cdot 10^{-4} & 2.8 \cdot 10^{-3} \end{pmatrix}. \tag{2.135}$$

From (2.134) it is clear that that knowing limits for one parametrization, we can calculate limits for the second one.

3 Basic notions of matrix theory in mixing mechanism

The prominent feature of the neutrino mixing matrix (2.14) is its unitarity and it has important consequences. The neutrino states with definite masses are orthonormal, i.e.

$$\langle \nu_i | \nu_k \rangle = \delta_{ik}. \tag{3.1}$$

The unitarity of the neutrino mixing matrix ensures that flavour neutrino states are also orthonormal

$$\begin{aligned} \langle \nu_{l'} | \nu_l \rangle &= \langle \sum_{i=1}^{3} U_{il'}^{*} \nu_i | \sum_{k=1}^{3} U_{lk} \nu_k \rangle = \sum_{i=1}^{3} \sum_{k=1}^{3} U_{il'}^{*} U_{lk} \langle \nu_i | \nu_k \rangle = \sum_{i=1}^{3} \sum_{k=1}^{3} U_{il'}^{*} U_{lk} \delta_{ik} \\ &= \sum_{i=1}^{3} U_{li} U_{il'}^{*} = \delta_{ll'}. \end{aligned} \tag{3.2}$$

The unitary mixing matrix is also necessary to diagonalize the neutrino mass matrix in order to obtain neutrinos with definite masses. These are mostly theoretical requirements of the unitarity of the mixing matrix. The last consequence of the unitarity has important physical significance, namely, it ensures that the probability of oscillation of a given flavour state to any other is equal to one

$$\sum_{\beta} P_{\nu_\alpha \to \nu_\beta} = 1. \tag{3.3}$$

Similarly, the probability of transition to a given neutrino flavour state from any neutrino state is also conserved

$$\sum_{\alpha} P_{\nu_\alpha \to \nu_\beta} = 1. \tag{3.4}$$

In other words, the unitarity of the mixing matrix provides that no neutrino states disappear during the propagation or are created out of nowhere.

If sterile neutrinos exist, exactly the same requirements must be satisfied and the complete, extended mixing matrix (2.124) must also be unitary. Thus, the PMNS mixing matrix is either unitary or is a part of a larger unitary matrix. In the latter case, if active and sterile neutrinos do not decouple, it is no longer unitary. These are two strong restrictions imposed on the 3×3 PMNS mixing matrix. Thus, the natural question arises, whether it is possible to characterize matrices by these two requirements.

3.1 Singular values and operator matrix norm

To answer this question we first need to introduce some notions coming from matrix theory. Let A be an arbitrary matrix of size $n \times m$, the space of all $n \times m$ matrices will be denoted $\mathbb{M}_{n\times m}$, thus $A \in \mathbb{M}_{n\times m}$.

Definition 1. *Singular values of a matrix A are the positive square roots of the eigenvalues of the matrix $A^\dagger A$, i.e.*

$$\sigma_i(A) = \sqrt{\lambda_i(A^\dagger A)} \quad i = 1, 2, ..., m. \tag{3.5}$$

The set of singular values of a matrix A will be denoted by $S(A)$. The matrix $A^\dagger A$ is a Hermitian matrix which ensures that its eigenvalues are real numbers and moreover they are also non-negative. The matrix of this type is called non-negative definite. From the definition it is clear that singular values are also non-negative quantities. Every matrix can be decomposed into the diagonal form according to its singular values [76, 77].

Theorem 3.1. *(Singular value decomposition)*
Let $A \in \mathbb{M}_{m\times n}$ be given and let $q = \min\{m, n\}$. Then there is a matrix $\Sigma = (\sigma_{ij}) \in \mathbb{M}_{m\times n}$ with $\sigma_{ij} = 0$ for all $i \neq j$ and $\sigma_{11} \geq \sigma_{22} \geq ... \geq \sigma_{qq}$, and there are two unitary matrices $U \in \mathbb{M}_{m\times m}$ and $V \in \mathbb{M}_{n\times n}$, which will be called left- and right-hand singular matrices respectively, such that $A = U\Sigma V^\dagger$.

From the singular value decomposition $A = U\Sigma V^\dagger$ it can be seen that the nonzero eigenvalues of $AA^\dagger$ are equal to the nonzero eigenvalues of $A^\dagger A$

$$\begin{aligned} A^\dagger A &= V\Sigma^\dagger U^\dagger U\Sigma V^\dagger = V\Sigma^T\Sigma V^\dagger \\ AA^\dagger &= U\Sigma V^\dagger V\Sigma^\dagger U^\dagger = U\Sigma\Sigma^T U^\dagger. \end{aligned} \tag{3.6}$$

These are eigenvalue decompositions of $A^\dagger A$ and $AA^\dagger$ respectively with nonzero eigenvalues $\sigma_1^2, \ldots, \sigma_q^2$. In other words, we can also use a matrix $AA^\dagger$ in the definition of singular values. Moreover, if we take the Hermitian conjugation of the singular value decomposition $A^\dagger = (U\Sigma V^\dagger)^\dagger = V\Sigma U^\dagger$, we see that $S(A) = S(A^\dagger)$. As the singular values are real numbers, they can be ordered. In this work, we use the convention that singular values are ordered in non-increasing order, i.e. if $\sigma_1, \ldots, \sigma_q$ are nonzero singular values of a matrix A, then

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q. \tag{3.7}$$

The problem of finding singular values can be transformed into a problem of finding eigenvalues of Hermitian matrices $A^\dagger A$ or $AA^\dagger$. However, there is another way to obtain results about singular values by considering eigenvalues of Hermitian matrices [76, 77].

Theorem 3.2. *Let $A \in M_{m\times n}$, let $q = min\{m, n\}$, let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q$ be ordered singular values of A. Then the ordered eigenvalues of the Hermitian matrix*

$$\mathcal{A} = \begin{pmatrix} 0 & A \\ A^\dagger & 0 \end{pmatrix} \tag{3.8}$$

are

$$-\sigma_1 \leq -\sigma_2 \leq \cdots \leq \sigma_q \leq \underbrace{0 = \cdots = 0}_{|m-n|} \leq \sigma_q \leq \cdots \leq \sigma_2 \leq \sigma_1. \tag{3.9}$$

Another important notion associated with matrices is that of a norm. As the matrices form vector space, we can use a standard vector norm to measure the size of objects. However, for matrices we can define another function that suits matrices better. This function is called a matrix norm.

Definition 2. *A matrix norm is a function $\| \cdot \|$ from the set of all matrices $\mathbb{M}_{n\times m}$ into $\mathbb{R}$ that satisfies the following properties*

$$\begin{aligned} &\|A\| \geq 0 \text{ and } \|A\| = 0 \Leftrightarrow A = 0, \\ &\|\alpha A\| = |\alpha|\|A\|, \\ &\|A + B\| \leq \|A\| + \|B\|, \\ &\|AB\| \leq \|A\|\|B\|. \end{aligned} \tag{3.10}$$

In other words, the matrix norm is a vector norm with the additional condition of submultiplicativity.

There exists an important class of matrix norms consisting of matrix norms which do not change by the unitary multiplication.

Definition 3. *A matrix norm $\| \cdot \|$ is called unitarily invariant if for every unitary matrices U, V and a given matrix A it satisfies*

$$\|UAV\| = \|A\|. \tag{3.11}$$

Another important class of matrix norms, called the induced matrix norms, contains matrix norms that are obtained from the vector norms in the following way

$$\|A\|_\star = \max_{\|x\|_\star = 1} \|Ax\|_\star, \tag{3.12}$$

where $\|\cdot\|_\star$ stands for the corresponding vector norm. In our case, of particular interest is the matrix norm induced from the Euclidean 2-norm $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{(x,x)} = \sqrt{x^\dagger x}$ for $x = (x_1, \ldots, x_n)^T$. From the Rayleigh quotient $\lambda_{\max}(A) = \max_{\|x\|_2=1} x^\dagger A x$ [77], we have

$$\begin{aligned} \|A\|_2^2 &= \max_{\|x\|_2=1} \|Ax\|_2^2 = \max_{\|x\|_2=1} (Ax)^\dagger Ax = \max_{\|x\|_2=1} x^\dagger A^\dagger Ax \\ &= \lambda_{\max}(A^\dagger A) = \sigma_1^2(A). \end{aligned} \tag{3.13}$$

Thus, the matrix norm $\|\cdot\|_2$ can be defined as the largest singular value of a given matrix. This matrix norm is called an operator norm or spectral norm and will be denoted throughout the text as $\|\cdot\|$. Thus,

Definition 4. *A spectral norm of a matrix $A \in \mathbb{M}_{n\times m}$ is the matrix norm defined as*

$$\|A\| = \max_{\|x\|_2=1} \|Ax\|_2 = \sigma_1(A). \tag{3.14}$$

Moreover, the spectral norm is also a unitary invariant norm (3.11).

Let us define a special class of matrices that have operator norm less than or equal to one, i.e. matrices for which the largest singular value is less than or equal to one.

Definition 5. *A matrix $A \in \mathbb{M}_{n\times m}$ is called a contraction if*

$$\|A\| \leq 1, \tag{3.15}$$

and a strict contraction if $\|A\| < 1$.

3.2 Mixing matrices as contractions

As we have already seen, the crucial property of the mixing matrix is its unitarity. Let us investigate singular values of the unitary matrix $U \in \mathbb{M}_{n\times n}$ which is determined by the unitarity condition

$$U^\dagger U = I, \tag{3.16}$$

where I is the $n \times n$ identity matrix, i.e. $I = diag(1, \ldots, 1)$. From the definition of singular values (3.5), it immediately follows that all singular values of the unitary matrix are equal to one, i.e. $S(U) = \{1, \ldots, 1\}$. This is the case for the PMNS mixing matrix in the Standard Model. However, in the presence of sterile neutrinos, the PMNS mixing matrix is no longer unitary but is a part of a larger unitary matrix. Moreover, to preserve mixing between SM neutrinos untouched, the PMNS mixing matrix must

occupy a top diagonal block of a complete unitary matrix. Thus, it is important to know the mutual relation between the spectral norm of a matrix and its submatrices. For principal submatrices such as the PMNS mixing matrix, this relation is presented in the following proposition.

Proposition 3.1. *Let $A \in \mathbb{M}_{r\times r}$ be a principal submatrix of $B \in \mathbb{M}_{n\times n}$, $r \leq n$. Then*

$$\|A\| \leq \|B\|. \tag{3.17}$$

Proof. Let $x \in \mathbb{C}^r$ with $\|x\|_2 = 1$. Then there is a vector $y \in \mathbb{C}^n$ which is a unit embedding of x in $\mathbb{C}^n$ with $\|y\|_2 = 1$ such that

$$\|Ax\|_2 \leq \|By\|_2. \tag{3.18}$$

Then,

$$\max_{\|x\|_2=1} \|Ax\|_2 \leq \max_{\|y\|_2=1} \|By\|_2. \tag{3.19}$$

□

From the neutrino mixing point of view, it is crucial to know singular values of submatrices of a unitary matrix.

Corollary 1. *Every principal submatrix A of a unitary matrix U is a contraction*

$$\|A\| \leq 1. \tag{3.20}$$

Thus, every 3×3 PMNS mixing matrix must be a contraction, either unitary, if there are only three neutrinos, or a non-unitary contraction, if additional sterile neutrinos exist. This allows us to make the following definition

Definition 6. *A mixing matrix is called physically admissible if it is a contraction.*

As the result, the notion of singular values and contractions allows us to describe completely neutrino mixing matrices including both the SM and BSM scenarios.

3.3 Unitary dilation

We have established that the physically interesting mixing matrices are contractions. These are matrices that are either unitary or can be a part of a larger unitary matrix. If we are interested in the BSM scenarios with additional sterile neutrinos we need to

study the structure of the complete unitary mixing matrix. However, at our disposal we have only data concerning the SM 3-dimensional mixing matrix. Thus, we need to find a way to extend these 3×3 contractions to a higher dimensional unitary matrix and maintain the mixing between the SM neutrinos. In mathematics, such a procedure is called a unitary dilation and it takes lower-dimensional contraction and puts it as the principal submatrix of a larger unitary matrix

$$U_{\texttt{PMNS}} \xrightarrow{\text{dilation}} \begin{pmatrix} U_{\texttt{PMNS}} & U_{lh} \\ U_{hl} & U_{hh} \end{pmatrix} \equiv U \to UU^\dagger = I. \tag{3.21}$$

The inverse operation is called the compression of the unitary matrix to the lower dimensional contraction. Practically the dilation procedure can be applied through the so-called cosine-sine decomposition (CS decomposition) [78].

Theorem 3.3. *Let the unitary matrix $U \in \mathbb{M}_{(n+m)\times(n+m)}$ be partitioned as*

$$U = \begin{array}{c} \begin{matrix} n & m \end{matrix} \\ \begin{pmatrix} U_{PMNS} & U_{lh} \\ U_{hl} & U_{hh} \end{pmatrix} \end{array} \begin{array}{c} n \\ m \end{array}, \tag{3.22}$$

If $m \geq n$, then there are unitary matrices $W_1, Q_1 \in M_{n\times n}$ and unitary matrices $W_2, Q_2 \in \mathbb{M}_{m\times m}$ such that

$$\begin{pmatrix} U_{PMNS} & U_{lh} \\ U_{hl} & U_{hh} \end{pmatrix} = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix} \left(\begin{array}{c|cc} C & -S & 0 \\ \hline S & C & 0 \\ 0 & 0 & I_{m-n} \end{array}\right) \begin{pmatrix} Q_1^\dagger & 0 \\ 0 & Q_2^\dagger \end{pmatrix}, \tag{3.23}$$

where $C \geq 0$ and $S \geq 0$ are diagonal matrices satisfying $C^2 + S^2 = I_n$.

There exists another form of the CS decomposition which is more important from the neutrino physics perspective. Let $U_{\texttt{PMNS}}$ have the singular value decomposition $U_{\texttt{PMNS}} = W_1 diag(I_r, C) Q_1^\dagger$, where I_r denotes r singular values equal to one, and C contains singular values that are strictly less than one. The structure of the CS decomposition reveals the intriguing fact, namely the minimal dimension of the unitary dilation of a given contraction is not arbitrary, but is encoded in the number of singular values strictly less than one.

Corollary 2. *The parametrization of the unitary dilation of the smallest size is given by*

$$\begin{pmatrix} U_{PMNS} & U_{lh} \\ U_{hl} & U_{hh} \end{pmatrix} = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix} \left(\begin{array}{cc|c} I_r & 0 & 0 \\ 0 & C & -S \\ \hline 0 & S & C \end{array}\right) \begin{pmatrix} Q_1^\dagger & 0 \\ 0 & Q_2^\dagger \end{pmatrix}, \tag{3.24}$$

where $r = n-m$ *is the number of singular values equal to 1 and* $C = diag(\cos\theta_1, ..., \cos\theta_m)$ *with* $|\cos\theta_i| < 1$ *for* $i = 1, ..., m$.

This is crucial from the physical point of view, since it tells that the minimal number of sterile neutrinos is not arbitrary, but depends on the singular values of the PMNS mixing matrix.

3.4 Remarks on the parametrization of the non-unitarity of the mixing matrix

In section 2.6 we have presented the two most commonly used parametrizations of non-unitarity of the neutrino mixing matrix and a connection between them. Let us discuss now more formal details. The α- and η-parametrization represent the deviation from unitarity as a multiplicative perturbation of a unitary matrix, which requires decomposition of a matrix into a product of matrices from which exactly one is a unitary matrix. There exist three different ways to decompose the matrix into a product of matrices of which one is unitary, namely [77, 79]

1. Polar decomposition,

2. QR decomposition,

3. Mostow decomposition.

The first two are already used in neutrino physics in the context of parametrization of non-unitarity effects in the neutrino mixing matrix [34–39, 80]. These are the polar decomposition and a modified version of the QR decomposition. Thus, let us take a closer look at these two parametrizations. The polar decomposition factorizes a given square matrix A into the following product

$$A = PU \tag{3.25}$$

where matrix P is a positive semidefinite Hermitian matrix and U is a unitary matrix. The polar factor P is uniquely determined and is given by $\sqrt{AA^\dagger}$ while the unitary part is uniquely determined if the initial matrix is non-singular.

In the physical application, the polar factor is further decomposed in the following way

$$P = I - \eta \tag{3.26}$$

where a Hermitian matrix η describes the deviation from unitarity of the neutrino mixing matrix and this is the idea behind the η-parametrization. As we already know, physical mixing matrices must be contractions, i.e. matrices with spectral norm less than or equal to one or equivalently with the largest singular value less than or equal to one. Let us notice that in general the polar decomposition does not provide this property. To see this let us look at a simple example, where we take matrix η in simple diagonal form [2]

$$\eta = \begin{pmatrix} \epsilon & 0 \\ 0 & -\epsilon \end{pmatrix} \tag{3.27}$$

where $0 < \epsilon \leq 1$.

Observe that this results in a positive semidefinite matrix $P = I - \eta$ which is necessary for a polar factor. However, such P is not a contraction since one singular value will be always larger than one, independently of how small ϵ is.

Recently, the polar factor in a form of (3.26) was identified with a matrix $I - \frac{\Theta\Theta^\dagger}{2}$ which arises in the context of the complete unitary mixing matrix [81] (for similar construction see also [82, 83]). Hence, in scenarios where the polar factor is considered as a principal submatrix of a complete unitary mixing matrix, it must be a contraction (Proposition 3.1). Thus, to ensure that the polar factor $I - \eta$ is a contraction, a necessary condition must be imposed on the η matrix, namely it must be a positive semidefinite matrix. On the other hand, using the fact that the operator norm is unitarily invariant, it can be shown that for sufficiently small entries of the matrix η also the inverse is true, i.e if the η matrix is positive semidefinite then $P = I - \eta$ must be a contraction [41].

The second of currently used factorizations in neutrino physics is the QR decomposition. It factorizes a given matrix into a product of a unitary matrix Q and an upper triangular matrix R. For the purpose of neutrino physics, a modified version of the QR factorization is used, namely the LQ decomposition, where L corresponds to a lower triangular matrix and Q is a unitary matrix. Moreover, in the context of the neutrino mixing this lower triangular matrix is further split into the following form

$$L = I - \alpha \tag{3.28}$$

where the matrix α is lower triangular and describes a deviation from unitarity of the $U_{\texttt{PMNS}}$.

[2] Since by the unitary invariance (3.11) of the spectral norm the unitary part is irrelevant here, we can focus only on the polar factor given by (3.26).

Both QR and polar decompositions can be related to each other by either applying the polar decomposition to the lower triangular matrix L or the LQ decomposition to the polar factor P.

Finally, let us look briefly at the last factorization, i.e. Mostow decomposition. It decomposes any non-singular complex matrix A in the following way

$$A = Ue^{iK}e^{S} \tag{3.29}$$

where U is a unitary matrix, K is a real skew-symmetric matrix and S corresponds to a real symmetric matrix.

At the end, let us emphasize the relation between the approach based on singular values and the polar decomposition. In our analysis, we use singular values as an indicator of whether a given matrix is a contraction. However, it is known that eigenvalues of the polar factor, which follows from its definition, are equal to singular values of the initial matrix. Thus, from that perspective, a polar decomposition can be treated as an intermediate step between the matrix under consideration and its singular value decomposition. Nevertheless, from the numerical analysis point of view, singular value decomposition algorithms are more natural since they arise from eigenvalues decomposition of matrices $AA^{\dagger}$ and $A^{\dagger}A$. Thus, in most cases, in order to obtain an algorithm for a polar decomposition, we have to adjust algorithms for the singular value decomposition.

3.5 Non-standard parametrizations, norms and contractions

For the SM $U_{\texttt{PMNS}}$ matrix, it holds that sum of probabilities of neutrino oscillations equals 1, i.e.

$$\sum_{\beta} P_{\alpha\beta} = 1, \quad e.g. \quad P_{ee} + P_{e\mu} + P_{e\tau} = 1. \tag{3.30}$$

However, for a non-unitary matrix U the analogous relation is not fulfilled. Let us see it in a simple case of two flavours (the same can be done for dimension 3 modified $U_{\texttt{PMNS}}$ matrix), when U is defined as

$$U = \begin{pmatrix} \cos\Theta_1 & \sin\Theta_1 \\ -\sin\Theta_2 & \cos\Theta_2 \end{pmatrix}, \tag{3.31}$$

where $\Theta_2 = \Theta_1 + \epsilon$.

In this case we get, $\Delta_{ij} \propto (m_i^2 - m_j^2)\frac{L}{E}$

$$\begin{aligned} \sum_{\alpha=e,\mu} P_{e\alpha} &= P_{ee} + P_{e\mu} \\ &= 1 + 4\epsilon \sin^2 \Delta_{21} \sin\Theta_1 \cos\Theta_1 \cos 2\Theta_1 + \mathcal{O}(\epsilon^2), \end{aligned} \tag{3.32}$$

$$\begin{aligned} \sum_{\alpha=e,\mu} P_{\mu\alpha} &= P_{\mu e} + P_{\mu\mu} \\ &= 1 - 4\epsilon \sin^2 \Delta_{21} \sin\Theta_1 \cos\Theta_1 \cos 2\Theta_1 + \mathcal{O}(\epsilon^2). \end{aligned} \tag{3.33}$$

We can see that the sum can be either larger or smaller than 1. This example was given in [84], however, no clue at that time was how to interpret possible results where the sum of probabilities does not equal 1. Here we show that matrix (3.31) is not the right way to parametrize BSM effects since it is not representing physically meaningful mixing matrices. This will be shown by considering the spectral norm of the U matrix (3.31).

First, we calculate UU^T and U^TU, and use the following abbreviation $s(c)_i \equiv \sin(\cos)\Theta_i$ for $i = 1, 2$

$$UU^T = \begin{pmatrix} 1 & s_1c_2 - s_2c_1 \\ s_1c_2 - s_2c_1 & 1 \end{pmatrix}, \tag{3.34}$$

$$U^TU = \begin{pmatrix} c_1^2 + s_2^2 & c_1s_1 - s_2c_2 \\ c_1s_1 - s_2c_2 & s_1^2 + c_2^2 \end{pmatrix}. \tag{3.35}$$

The spectral norm satisfies the following relation $\|U^TU\| = \|UU^T\| = \|U\|^2$, thus we can focus on either of these products. Let us write the first product in (3.35) as the sum

$$\begin{aligned} UU^T &= \begin{pmatrix} 1 & s_1c_2 - s_2c_1 \\ s_1c_2 - s_2c_1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & s_1c_2 - s_2c_1 \\ s_1c_2 - s_2c_1 & 0 \end{pmatrix}. \end{aligned} \tag{3.36}$$

This can be simplified into

$$UU^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & s_3 \\ s_3 & 0 \end{pmatrix} \equiv I + B \tag{3.37}$$

where $s_3 \equiv \sin\Theta_3 = \sin(\Theta_1 - \Theta_2)$. The B is symmetric matrix which eigenvalues are equal $\pm s_3$. Let V be a unitary matrix such that $V^T B V = D = \mathrm{diag}(s_3, -s_3)$. Since operator norm is unitarily invariant (3.11), we can write

$$\begin{aligned} \|UU^T\| &= \|I + B\| = \|V^T(I + B)V\| = \|I + V^T B V\| \\ &= \|I + D\|. \end{aligned} \tag{3.38}$$

The $I + D$ equals

$$\begin{pmatrix} 1 + s_3 & 0 \\ 0 & 1 - s_3 \end{pmatrix}, \tag{3.39}$$

hence its spectral norm, i.e. the largest singular value, is equal

$$\begin{aligned} 1 + s_3 \quad &\text{if} \quad s_3 \geq 0, \\ 1 - s_3 \quad &\text{if} \quad s_3 < 0. \end{aligned} \tag{3.40}$$

So we can see that by adding B to the identity matrix we cannot decrease the operator norm

$$1 = \|I\| \leq \|I + B\| = \|UU^T\| = \|U\|^2 = 1 + |s_3|. \tag{3.41}$$

Thus

$$\|U\| \geq 1. \tag{3.42}$$

It is already clear that physically meaningful theory should include only mixing matrices for which contraction relation (3.15) is fulfilled, which is obviously not satisfied by the discussed matrix U (3.31). The result (3.42) implies that not all parametrizations which violate unitarity are a proper choice, and a simple mixing matrix (3.31) is superfluous from the physical point of view. It fulfills $\|U\| = 1$ for $\epsilon = 0$, but then trivially a unitary matrix is recovered.

3.6 Numerical precision of singular values determination

In the next sections and chapters we will determine numerically singular values. Thus it is important to estimate errors connected with matrix evaluations. Singular values are continuous functions of matrix elements [77], and hence a small change in elements does not change drastically singular values. Quantitatively this can be described with the use of the Weyl inequality for singular values [85,86] which describes the behaviour of the singular values of a given matrix U under the additive perturbation gathered in the matrix E

$$|\sigma_i(U + E) - \sigma_i(V)| \leq \|E\|, \tag{3.43}$$

where $\|\cdot\|$ is a spectral norm. In our analysis the elements of the perturbation matrix E are taken to be equal to the finest precision of the elements of the matrix under the consideration. For example, if the best precision of elements of the PMNS matrix is 10^{-3}, all the elements of the matrix E are equal to 10^{-3}. This gives the following estimation of the error for singular values

$$|\sigma_i(U+E) - \sigma_i(V)| \leq 3 \times 10^{-3}. \tag{3.44}$$

The experimental data with which we will work on (Chapter 5) has the matrix elements given with 10^{-5} accuracy. We keep such accuracy, which is correct for considered interval mixing matrix elements. This precision of matrix elements implies that the singular values will be given with the 3×10^{-5} accuracy. In other words, the threshold put on the largest singular value of a given matrix to be acknowledged as a contraction is 1.00003.

On that basis, we can establish the allowed lower bounds for singular values for physically admissible mixing matrices by considering the interval matrix U_{int} (2.122) or (2.123). The results are as follows

$$\begin{aligned} &\text{CP invariant scenario: } \{\sigma_1 = 0.95954, \sigma_2 = 0.88186, \sigma_3 = 0.84189\}, \\ &\text{General scenario: } \{\sigma_1 = 0.95592, \sigma_2 = 0.84112, \sigma_3 = 0.70275\}. \end{aligned} \tag{3.45}$$

As the deviation below unity of any of singular values is a signal of the BSM scenario, these results clearly show that within current experimental limits there is a lot of space for additional neutrinos. Moreover, the largest singular value tells us whether the mixing matrix is physically meaningful. Thus, if we abandon the restriction of contraction, i.e. $\sigma_1 \leq 1.00003$, then the study of the upper limit for the σ_1 of the U_{int} (2.122)-(2.123) will measure the amount of non-physical mixing matrices within the present experimental limits

$$\begin{aligned} &\text{CP invariant scenario: } \{\sigma_1 = 1.1369, \sigma_2 = 1.09429, \sigma_3 = 1.02917\}, \\ &\text{General scenario: } \{\sigma_1 = 1.28378, \sigma_2 = 1.17208, \sigma_3 = 1.02917\}. \end{aligned} \tag{3.46}$$

We can see that the largest possible σ_1 within the U_{int} is much above the limit of contractions. On top of that, its deviation from unity is much larger than for the lower bound for physically admissible matrices (3.45). Moreover, all other singular values also exceed the unity from above in (3.46), which suggests that within current experimental limits there is a lot of physically meaningless mixing matrices. In the next chapter we will construct a general space of physically meaningful mixing matrices.

4 Region of physically admissible mixing matrices

We will show that a matrix constructed as a finite convex combination of unitary matrices is a contraction. Let U_i, $i = 1, \ldots, n$, be a unitary matrix, and let $A = \sum_{i=1}^{n} \alpha_i U_i$ with $\alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$, then

$$\|A\| = \|\sum_{i=1}^{n} \alpha_i U_i\| \leq \sum_{i=1}^{n} \alpha_i \|U_i\| = \sum_{i=1}^{n} \alpha_i = 1 \Rightarrow \|A\| \leq 1. \tag{4.1}$$

The converse is also true [87], i.e.

Theorem 4.1. *A matrix A is a contraction if and only if A is a finite convex combination of unitary matrices.*

This characterization of contractions has physical consequences. It allows gathering physically meaningful mixing matrices into the geometric region.

Definition 7. *The region of all physically admissible mixing matrices, denoted Ω, is the set of all finite convex combinations of 3×3 unitary matrices with parameters restricted by experiments*

$$\begin{aligned} \Omega :=& \text{conv}(U_{PMNS}) = \{\sum_{i=1}^{m} \alpha_i U_i \mid U_i \in U(3), \alpha_1, ..., \alpha_m \geq 0, \sum_{i=1}^{m} \alpha_i = 1, \\ & \theta_{12}, \theta_{13}, \theta_{23} \text{ and } \delta \text{ given by experimental values}\}. \end{aligned} \tag{4.2}$$

There is another equivalent definition of the Ω region, which reflects its geometric nature, namely as the convex hull spanned on the unitary PMNS matrices.

The Corollary 2 in section 3.3 restricts the minimal dimension of the unitary extension of the contractions. This allows us to divide the Ω region into four disjoint subsets according to the minimal dimension of the unitary dilation

$$\Omega_1 : \quad \text{3+1 scenario: } \Sigma = \{\sigma_1 = 1.0, \sigma_2 = 1.0, \sigma_3 < 1.0\}, \tag{4.3}$$

$$\Omega_2 : \quad \text{3+2 scenario: } \Sigma = \{\sigma_1 = 1.0, \sigma_2 < 1.0, \sigma_3 < 1.0\}\ , \tag{4.4}$$

$$\Omega_3 : \quad \text{3+3 scenario: } \Sigma = \{\sigma_1 < 1.0, \sigma_2 < 1.0, \sigma_3 < 1.0\}, \tag{4.5}$$

$$\Omega_4 : \quad \text{PMNS scenario: } \Sigma = \{\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1\}. \tag{4.6}$$

This division allows to analyze individually scenarios with a different number of sterile neutrinos. Thus, the study of geometric features of this region is useful for a better understanding of neutrino physics, especially regarding the number of additional sterile neutrinos and the structure of the complete mixing matrix.

It is important to notice that matrices from the Ω_1 subset can be extended to unitary matrices of arbitrary dimension, starting from the dimension four. The same is true for contractions from the subset Ω_2 which can produce any unitary matrices of dimension five or higher. It may look that there is an overlapping between matrices from different subsets of the Ω region and some of them may be redundant. This is however not true, as unitary matrices produced by the contraction from each subset are unique. It is so because contractions must end up in the 3×3 top diagonal block of a complete unitary matrix and as the subsets are disjoint, we cannot reproduce the same unitary matrices using contractions from different subsets. Thus, instead of overlapping, we should treat dilations of a given dimension of contractions from different subsets as complementary to each other.

4.1 Geometry of the region of physically admissible mixing matrices

The Ω region is a subset of the unit ball of the spectral norm

$$\mathcal{B}(n) = \{A \in \mathbb{C}^{n\times n} : \|A\| \leq 1\}. \tag{4.7}$$

This fact allows us to give another characterization of the Ω region as the intersection of the $\mathcal{B}(3)$ with the interval matrix U_{int} (2.123)-(2.123), i.e.

$$\Omega = \mathcal{B}(3) \cap U_{int}. \tag{4.8}$$

The geometry of $\mathcal{B}(n)$ is strictly connected to the geometry of symmetric gauge functions [76].

Definition 8. *A function $\Phi : \mathbb{R}^n \to \mathbb{R}$ is a symmetric gauge function if it satisfies the following conditions*

1. *Φ is a vector norm,*
2. *For any permutation matrix P we have $\Phi(Px) = \Phi(x)$,*
3. *$\Phi(|x|) = \Phi(x)$.*

Von Neumann proved that symmetric gauge functions and unitarily invariant norms (3.11) are connected to each other [88], namely

Theorem 4.2. *$\|\cdot\|$ is a unitary invariant norm if and only if there exists a symmetric gauge function Φ such that $\|A\| = \Phi(S(A))$ for all $A \in \mathbb{C}^{n\times n}$, where $S(A)$ is the set of singular values of A.*

The spectral norm is a unitarily invariant norm and its corresponding symmetric gauge function is an infinite norm, i.e.

$$\Phi_\infty(x) = max\{|x_1|, |x_2|, \ldots, |x_n|\}. \tag{4.9}$$

The unit ball of the infinite norm is a hypercube

$$\mathcal{B}_\infty(n) = \{x \in \mathbb{R}^n : \Phi_\infty(x) \leq 1\} = [-1, 1]^n . \tag{4.10}$$

The Von Neuman's relation between unitary invariant norms and symmetric gauge functions is also reflected in the geometry of the corresponding unit balls. The characterization of the extreme points and facial structure of unit balls of unitarily invariant norms by the corresponding structure of unit balls of symmetric gauge functions have been studied in [89–93]. Faces and extreme points are defined as follows [94, 95]

Definition 9. *Let $C \in \mathbb{R}^n$ be a convex set. A convex set $F \subseteq C$ is called a face of C if for every $x \in F$ and every $y, z \in C$ such that $x \in (y, z)$, we have $y, z \in F$.*

Definition 10. *The zero dimensional faces of a complex set C are called extreme points of C. Thus a point $x \in C$ is an extreme point of C if and only if there is no way to express x as a convex combination $(1 - \lambda)y + \lambda z$ such that $y, z \in C$ and $0 < \lambda < 1$, except by taking $x = y = z$.*

It appears that the extreme points of the $\mathcal{B}(n)$ are exactly unitary matrices. This result has also been obtained in a more general setting by Stoer [96]. The facial structure of the $\mathcal{B}(n)$ is given in the following theorem [90, 93]

Theorem 4.3. *$\mathcal{F}$ is a face of $\mathcal{B}(n)$ if and only if there exist $0 \leq r \leq n$ and unitary matrices U and V such that*

$$\mathcal{F} = \{U \begin{pmatrix} I_r & 0 \\ 0 & A \end{pmatrix} V : A \in \mathcal{B}(n - r)\}. \tag{4.11}$$

As the Ω region is a subset of $\mathcal{B}(n)$, its geometric structure is inherited from $\mathcal{B}(n)$. Thus, the facial structure of the Ω region is the same as for $\mathcal{B}(n)$ with restriction of parameters of unitary matrices U and V to experimental results (2.120) and with established ranges of singular values (3.45). However, if we look at the definition of the faces of $\mathcal{B}(n)$ (4.11), we will observe that they do not correspond completely to physically interesting subsets (4.3)-(4.5) of the Ω. Namely, higher-dimensional faces contain lower-dimensional faces, e.g. for $r = 1$ the face contains not only matrices with

two singular values strictly less than one, but also unitary matrices and contractions with only one singular value strictly less than one. In other words faces of $\mathcal{B}(n)$ comprise matrices from different subsets of Ω. To restrict faces to subsets containing only matrices with the specific number of singular values strictly less than one, we can use the notion of the relative interior [94].

Definition 11. *The relative interior of a convex set $C \subset \mathbb{R}^n$, which is denoted by $ri(C)$, is defined as the interior which results when C is regarded as a subset of its affine hull aff(C).*

In this definition by the affine hull of the set C we understand the set of all finite affine combinations of elements of C [97], i.e. $\text{aff}(C) = \{\sum_{i=1}^{k} \alpha_i x_i : x_i \in C, \sum_{i=1}^{k} \alpha_i = 1\}$. In that way, the subsets of $\mathcal{B}(n)$ corresponding to different minimal unitary extensions are the relative interiors of $\mathcal{F}$, i.e. subsets of faces for which singular values of the A submatrix are strictly smaller than one.

Definition 12. *The subsets $\Omega_1, \ldots, \Omega_4$ of the Ω region are relative interiors of the faces $\mathcal{F}$ of $\mathcal{B}(3)$ for $r = 2, 1, 0, 3$, respectively, with parameters of unitary matrices U and V restricted by experimental data and with allowed ranges of singular values.*

There is another way to characterize subsets of the Ω region, namely in terms of Ky-Fan k-norms.

Definition 13. *For a given matrix $A \in \mathbb{C}^{n\times n}$ the Ky-Fan k-norm is defined as the sum of k largest singular values*

$$\|A\|_k = \sum_{i=1}^{k} \sigma_i(A), \text{ for k=1,...,n.} \tag{4.12}$$

In particular for matrices in $\mathbb{C}^{3\times 3}$ the three possible Ky-Fan norms are

$$\begin{aligned} \|A\|_1 &= \sigma_1(A) \text{ (spectral norm)}, \\ \|A\|_2 &= \sigma_1(A) + \sigma_2(A), \\ \|A\|_3 &= \sigma_1(A) + \sigma_2(A) + \sigma_3(A) \text{ (nuclear norm)}. \end{aligned} \tag{4.13}$$

Let us define for $k = 1, \ldots, 3$ the following sets

$$\begin{aligned} S_k(r) &= \{A \in \mathbb{C}^{n\times n} : \|A\|_k = r\}, \\ A_k(r_1, r_2) &= \{A \in \mathbb{C}^{n\times n} : r_1 \leq \|A\|_k < r_2\}, \end{aligned} \tag{4.14}$$

i.e. we defined the sphere of radius r and the annulus with radii r_1 and r_2 for Ky-Fan norms centered at the origin. Then, the subsets of the Ω can be defined as

$$\begin{aligned}
\Omega_1 &= S_1(1) \cap S_2(2) \cap A_3(2 + \sigma_{3min}, 3), \\
\Omega_2 &= S_1(1) \cap A_2(1 + \sigma_{2min}, 2) \cap A_3(1 + \sigma_{2min} + \sigma_{3min}, 3), \\
\Omega_3 &= A_1(\sigma_{1min}, 1) \cap A_2(\sigma_{1min} + \sigma_{2min}, 2) \cap A_3(\sigma_{1min} + \sigma_{2min} + \sigma_{3min}, 3), \\
\Omega_4 &= S_1(1) \cap S_2(2) \cap S_3(3),
\end{aligned} \tag{4.15}$$

where σ_{imin} for $i = 1, 2, 3$ are lower limits of singular values allowed by current experimental data (3.45).

4.1.1 Physically admissible mixing matrices as a convex combination of PMNS matrices

The Ω region is defined as the convex hull of unitary PMNS mixing matrices or equivalently as the set of a finite convex combination of PMNS mixing matrices. In this way, we claim that every physically admissible mixing matrix can be represented as a fine convex combination of unitary PMNS matrices. This agrees with the Krein-Milman theorem which states [97, 98]

Theorem 4.4. *Let $C \subset \mathbb{R}^n$ be a nonempty compact convex set, and let $ext(C)$ be the set of extreme points of C, then*

$$C = \overline{conv}(ext(C)), \tag{4.16}$$

As we have discussed, the extreme points of the unit ball of the spectral norm are unitary matrices and as the Ω is a subset of the $\mathcal{B}(n)$, the above theorem justifies our definition. However, this theorem does not put any restriction on the number of extreme points necessary to construct any point of a convex set as a convex combination of its extreme points. The upper bound for this number has been given by Carathéodry [97, 99]

Theorem 4.5. *If $K \subset \mathbb{R}^n$, then each point of $conv(K)$ is a convex combination of at most $n + 1$ points of K.*

The natural question arises: What is the Carathéodry number for the $\mathcal{B}(n)$ and Ω which are embedded in $\mathbb{C}^{n^2} \simeq \mathbb{R}^{2n^2}$? In the physically interesting case where $n = 3$ according to the Carathéodry theorem we would need 19 unitary matrices. However, we will prove that for $\mathcal{B}(3)$ this number can be significantly reduced.

Proposition 4.1. *The Carathéodory's number for the conv*$(\mathcal{U}(3))$=$\mathcal{B}(3)$ *is* 4.

Proof. Let $B_\infty = \{x \in \mathbb{R}^3 : \|x\|_\infty \leq 1\}$ be a unit ball of the infinite norm in $\mathbb{R}^3$, i.e. the cube $[-1,1]^3$. The extreme points of the B_∞ are vertices of the cube, i.e. vectors $v_j = (\pm 1, \pm 1, \pm 1)^T$ for $j = 1, \ldots, 8$. Let $\psi : B_\infty \to \mathbb{M}_{3\times 3}$ be a mapping which sends a vector from the unit ball B_∞ into the diagonal matrix. Then, the ψ sends the extreme points of the B_∞ into the diagonal unitary matrices $U_j = diag(\pm 1, \pm 1, \pm 1)$ for $j = 1, \ldots, 8$. The Carathéodory's number for the cube is 4. Thus, every point in B_∞ can be written as the convex combination of at most 4 extreme points v_j. In particular every point of the positive octant can be written in this way. This means that every diagonal matrix $D \in \mathbb{M}_{3\times 3}$ with diagonal elements in $[0,1]$ can be written as convex combination of at most 4 diagonal unitary matrices U_j, i.e. $D = \sum_{i=1}^4 \alpha_i U_i$, with $\alpha_i \geq 0$ and $\sum_{i=1}^4 \alpha_i = 1$. Now, let A be a contraction with a singular value decomposition $A = WDV^\dagger$, where W and V are unitary matrices. This gives

$$A = WDV^\dagger = \sum_{i=1}^{4} \alpha_i W U_i V^\dagger. \tag{4.17}$$

As the *conv*$(U(3))$=$\mathcal{B}(3)$ is the set of all 3×3 contractions, this completes the proof. □

As an immediate consequence of this proposition and the construction used in the proof, matrices from the Ω_2 subset, i.e. with two singular values strictly less than one, can be constructed as the convex combination of 3 unitary matrices. Whereas, matrices from the Ω_1 subset, i.e. with only one singular value strictly less than one, can be constructed as the convex combination of two unitary matrices.

Following the idea of Stoer [96], we will show how to construct contractions with two and one singular values strictly less than one as a convex combination of three and two unitary matrices, respectively. Let us take the following diagonal matrix

$$D_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix}, \tag{4.18}$$

where $a, b < 1$. It can be written as the following sum

$$D_1 = \frac{1-a}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} + \frac{a-b}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} + \frac{1+b}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.19}$$

Now let us take another diagonal matrix. This time with only one diagonal element strictly less than one

$$D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a \end{pmatrix}, \tag{4.20}$$

where $a < 1$. The D_2 matrix can be written as

$$D_2 = \frac{1-a}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} + \frac{1+a}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.21}$$

Multiplying D_1 and D_2 matrices from left- and right-hand side by unitary matrices, we end up with a singular value decomposition of a given matrix with singular values gatherd in D_1 and D_2, respectively. As a result contractions with two and one singular values strictly smaller than one can be written as convex combinations of unitary matrices with singular values encoded in coefficients of the combination.

4.2 Volume

Lie groups, besides being of course groups, are also manifolds [100], i.e. geometric structures. Thus, we can associate with them geometrical properties such as the surface area, also called the volume. Two very important groups in physics fall into this category, namely an orthogonal group, and its complex counterpart: a unitary group. These groups are also very important in neutrino physics as the mixing matrix is either an orthogonal matrix or, if the CP phase is non-zero, a unitary matrix.

The set of all orthogonal matrices of dimension $n \times n$, i.e. $\mathcal{O}(n) = \{O \in \mathbb{R}^{n\times n} : OO^T = I\}$, is an example of a Stiefel manifold [101]. As the orthogonal matrices have $\frac{n(n-1)}{2}$ independent parameters, the Stiefel manifold of the orthogonal group is a $\frac{n(n-1)}{2}$ dimensional manifold embedded in n^2 space. We can associate to it a volume which is expressed as the Haar measure over the orthogonal group [101–105]

$$vol(\mathcal{O}(n)) = \int_{\mathcal{O}(n)} [O^T dO]^\wedge, \tag{4.22}$$

where $[O^T dO]^\wedge$ denotes the wedge product of the matrix $O^T dO$ and dO is the matrix of the differentials of the orthogonal matrix O. This volume can be expressed in the following compact form [105, 106]

$$vol(\mathcal{O}(n)) = \frac{2^n \pi^{\frac{n^2}{2}}}{\Gamma_n(\frac{n}{2})} = \frac{2^n \pi^{\frac{n(n+1)}{4}}}{\prod_{k=1}^{n} \Gamma(\frac{k}{2})}, \tag{4.23}$$

where $\Gamma_n(x) = \pi^{\frac{n(n-1)}{4}}\Gamma(x)\Gamma\left(x-\frac{1}{2}\right)\ldots\Gamma\left(x-\frac{n-1}{2}\right)$. In the case interesting from the neutrino physics perspective, i.e. for $n=3$, this gives

$$vol(\mathcal{O}(3)) = 16\pi^2. \tag{4.24}$$

However, as the determinant of the PMNS mixing matrix is equal to 1, it belongs to even a smaller subset, namely the special orthogonal group $\mathcal{SO}(3)$. The special orthogonal group is a subgroup of $\mathcal{O}(3)$ and its volume is half of the volume of the orthogonal group, i.e.

$$vol(\mathcal{SO}(3)) = 8\pi^2. \tag{4.25}$$

Moreover, the PMNS matrix does not cover the entire range of parameters and hence we must start from $O^T dO$ in order to calculate the volume of this submanifold. Taking the standard PMNS parametrization (2.42) we get

$$O^T dO = \begin{pmatrix} 0 & d\theta_{12} + s_{13}d\theta_{23} & c_{12}d\theta_{13} - c_{13}s_{12}d\theta_{23} \\ -d\theta_{12} - s_{13}d\theta_{23} & 0 & c_{13}c_{12}d\theta_{23} + s_{12}d\theta_{13} \\ -c_{12}d\theta_{13} + c_{13}s_{12}d\theta_{23} & -c_{13}c_{12}d\theta_{23} - s_{12}d\theta_{13} & 0 \end{pmatrix} \tag{4.26}$$

The wedge product of the independent elements of this matrix is equal to

$$[O^T dO]^\wedge = \cos(\theta_{13}) d\theta_{12} d\theta_{13} d\theta_{23}. \tag{4.27}$$

Thus, the volume of PMNS matrices is given by

$$vol(PMNS) = \int_{\theta_{12_L}}^{\theta_{12_U}} \int_{\theta_{13_L}}^{\theta_{13_U}} \int_{\theta_{23_L}}^{\theta_{23_U}} = \cos(\theta_{13}) d\theta_{12} d\theta_{13} d\theta_{23}, \tag{4.28}$$

which with the current experimental limits on θ_{12}, θ_{13} and θ_{23} (2.120) gives

$$vol(PMNS) = 2.2667 \times 10^{-4}. \tag{4.29}$$

As we can see the PMNS matrices contribute only in a small portion, 0.00029 percent, to the entire $\mathcal{SO}(3)$, and if compared to the orthogonal group the PMNS matrices constitute only 0.00014 percent of the $\mathcal{O}(3)$.

The unitary group $\mathcal{U}(n)$, i.e the group of all unitary matrices $\mathcal{U}(n) = \{U \in \mathbb{C}^{n\times n} : UU^T = I\}$, is another example of Stiefel manifold. Similarly, as for the orthogonal group, the volume of the unitary group is given by the Haar measure over the group

$$vol(\mathcal{U}(n)) = \int_{U(n)} [U^\dagger dU]^\wedge, \tag{4.30}$$

where $[U^\dagger dU]^\wedge$ denotes the exterior product of the matrix $U^\dagger dU$ and dU is the matrix of the differentials of the unitary matrix U. The volume of the unitary group can be expressed in a compact form as [102–106]

$$vol(\mathcal{U}(n)) = \frac{2^n \pi^{n^2}}{\tilde{\Gamma}_n(n)} = \frac{2^n \pi^{\frac{n(n+1)}{2}}}{1!2!\dots(n-1)!}, \tag{4.31}$$

where $\tilde{\Gamma}_n(x) = \pi^{\frac{n(n-1)}{2}} \Gamma(x)\Gamma(x-1)\dots\Gamma(x-n+1)$. Thus, the volume of the 3×3 unitary matrices equals

$$vol(\mathcal{U}(3)) = 4\pi^6. \tag{4.32}$$

Moreover, the determinant of the PMNS matrix is equal to one, which means it belongs to the special unitary group $\mathcal{SU}(n)$. The volume of the $\mathcal{SU}(n)$ written in the compact form is [102–104]

$$vol(\mathcal{SU}(n)) = \frac{2^{\frac{(n-1)}{2}} \pi^{\frac{(n-1)(n+2)}{2}}}{1!2!\dots(n-1)!}. \tag{4.33}$$

For the physically interesting dimension, i.e. $n = 3$ this volume is equal to

$$vol(\mathcal{SU}(3)) = \sqrt{3}\pi^5. \tag{4.34}$$

The PMNS mixing matrix with non-zero CP phase, however, does not have the full set of parameters, as some of them can be incorporated into the redefinition of the fermion fields, and moreover ranges of these parameters are restricted by experiments. Hence, in order to calculate its volume, it is necessary to start from a specific parametrization of the mixing matrix (2.42). It can be done in the same way as for its real counterpart, i.e. by calculating the wedge product of the matrix $U^\dagger dU$. However, for the complex matrices, it is much more complicated. Alternatively, it can be calculated by determining the Jacobian matrix of the PMNS matrix in the parametrization (2.42)

$$J = \left(\frac{\partial u_{ij}}{\partial y_k} \right), \quad i, j = 1, \dots, n \tag{4.35}$$

and the y_k are parameters (for the PMNS matrix $k = 1, \dots, 4$). Then, the volume element is multiplied by the Jacobian $|J| = \sqrt{det(\frac{1}{2} J^\dagger J)}$.

The volume of complex PMNS matrices can be calculated in one more way, namely by using the Cartan-Killing metric [107–109]

$$ds^2 = (V, V)dt^2, \tag{4.36}$$

where $(A, B) = \frac{1}{2} Tr(A^\dagger B)$ is the inner product induced by the Frobenius norm and $V = U^\dagger dU$. The V is anti-Hermitian and thus $(V, V) = \frac{1}{2} Tr(V^\dagger V) = -\frac{1}{2} Tr(V^2)$.

The Hermitian product of the Jacobian matrix for the PMNS matrix is given by

$$\frac{1}{2}J^{\dagger}J = \begin{pmatrix} 1 & 0 & \sin(\theta_{13})\cos(\delta) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\theta_{13})\cos(\delta) & 0 & 1 & 0 \\ 0 & 0 & 0 & \sin^2(\theta_{13}) \end{pmatrix}. \tag{4.37}$$

Let us look also at the expression for the Cartan-Killing metric

$$ds^2 = d\theta_{23}^2 + d\theta_{13}^2 + d\theta_{12}^2 + 2\sin(\theta_{13})\cos(\delta) + \sin^2(\theta_{13})d\delta^2 \tag{4.38}$$

which as expected gives the same matrix as (4.37). Finally, the Jacobian for the PMNS matrices is equal to

$$|J| = \sqrt{\sin^2(\theta_{13}) - \cos^2(\delta)\sin^4(\theta_{13})}. \tag{4.39}$$

Thus, the volume of the complex PMNS matrices is given by

$$\begin{aligned} vol(PMNS) &= \int_{PMNS} |J| dV \\ &= \int \sqrt{\sin^2(\theta_{13}) - \cos^2(\delta)\sin^4(\theta_{13})} d\theta_{23} d\theta_{13} d\theta_{12} d\delta. \end{aligned} \tag{4.40}$$

Taking into account current experimental limits for mixing parameters (2.120) the numerical value for the volume of the complex PMNS mixing matrices is

$$vol(PMNS) = 1.4777 \times 10^{-4}. \tag{4.41}$$

So far we have established the volume of the neutrino mixing matrices only for the SM scenario. However, if we want to consider also BSM scenarios, it is required to consider the entire Ω region and not only its extreme points. In order to do this, we will use the fact that the region of all physically admissible mixing matrices is a subset of the unit ball in the spectral norm $\mathcal{B}(n) = \{A \in \mathbb{C}^{n\times n} : \|A\| \leq 1\}$ and for the CP conserving case it is restricted to the real matrices $\tilde{\mathcal{B}}(n) = \{A \in \mathbb{R}^{n\times n} : \|A\| \leq 1\}$. Volumes of the $\mathcal{B}(n)$ and $\tilde{\mathcal{B}}(n)$ can be calculated from the singular value decomposition. The differential of the singular value decomposition of a given matrix $A = U\Sigma V^{\dagger}$ is equal to

$$dA = dU\Sigma V^{\dagger} + Ud\Sigma V^{\dagger} + U\Sigma dV^{\dagger}. \tag{4.42}$$

By multiplying this from the left-hand side by $U^{\dagger}$ and from the right-hand side by V we get

$$dX \equiv U^{\dagger}dAV = U^{\dagger}dU\Sigma + d\Sigma + \Sigma dV^{\dagger}V, \tag{4.43}$$

which can be rewritten using $dV^\dagger V = -V^\dagger dV$ in the following form

$$dX = U^\dagger dAV = U^\dagger dU\Sigma + d\Sigma - \Sigma V^\dagger dV. \tag{4.44}$$

The Haar measure is left- and right-invariant, thus $[dA]^\wedge = [U^\dagger dAV]^\wedge = [dX]^\wedge$. The entrywise analysis of the dX in the CP invariant scenario gives

$$[dX]^\wedge = \prod_{i<j} |\sigma_j^2 - \sigma_i^2| \wedge_{i=1}^n d\sigma_i [O^T dO]^\wedge [Q^T dQ]^\wedge. \tag{4.45}$$

Hence, the volume of the unit ball of the spectral norm in a real case is given by

$$vol(\tilde{\mathcal{B}}(n)) = \frac{1}{2^n n!} vol(\mathcal{O}(n))^2 \int_0^1 \prod_{i<j} |\sigma_j^2 - \sigma_i^2| \prod_{k=1}^n d\sigma_k. \tag{4.46}$$

The inclusion of the factor $\frac{1}{2^n}$ assures the uniqueness of the singular value decomposition. For the physically interesting dimension $n = 3$, we have

$$vol(\tilde{\mathcal{B}}(3)) = \frac{8\pi^4}{45}. \tag{4.47}$$

Similar entrywise analysis of the (4.44) provides the volume element for the singular value decomposition of complex matrices

$$[dX]^\wedge = \prod_{i=1}^n \sigma_i \prod_{i<j} |\sigma_j^2 - \sigma_i^2|^2 \wedge_{i=1}^n d\sigma_i [U^\dagger dU]^\wedge [V^\dagger dV]^\wedge \tag{4.48}$$

Thus, the volume of the unit ball of the spectral norm is given by

$$vol(\mathcal{B}(n)) = \frac{1}{(2\pi)^n n!} vol(\mathcal{U}(n))^2 \int_0^1 \prod_{k=1}^n \sigma_k \prod_{i<j} |\sigma_j^2 - \sigma_i^2|^2 \prod_{k=1}^n d\sigma_k. \tag{4.49}$$

The factor $\frac{1}{(2\pi)^n}$ ensures the uniqueness of the singular value decomposition. In the case interesting from the neutrino physics perspective, i.e. $n = 3$, this gives

$$vol(\mathcal{B}(3)) = \frac{\pi^9}{8640}. \tag{4.50}$$

We can use the formulas for the volumes of $\mathcal{B}(3)$ and $\tilde{\mathcal{B}}(3)$ as the basis in the calculation of the volume of the Ω region. As the Ω region is defined as the convex hull of the PMNS matrices, to find its volume we need to replace in the formulas (4.46) and (4.49) $vol(\mathcal{U}(n))$ and $vol(\mathcal{O}(n))$ by $vol(PMNS)$ in the general and CP conserving case, respectively. Moreover, it is also necessary to restrict ranges of singular values for

those allowed by current experimental data (3.45). As the result, for the CP invariant scenario, we have the following formula

$$vol(\tilde{\Omega}) = \frac{1}{2^n n!} vol(PMNS)^2 \int_{\sigma_{min}}^{1} \prod_{i<j} |\sigma_j^2 - \sigma_i^2| \prod_{k=1}^{n} d\sigma_k. \tag{4.51}$$

Taking into account current experimental bounds (2.120) and allowed ranges for singular values (3.45), the numerical value is equal

$$vol(\tilde{\Omega}) = 6.45 \times 10^{-16}. \tag{4.52}$$

Thus, the Ω region in the CP invariant case constitutes only 1.84×10^{-18} percent of the unit ball (4.46).

For the general case including the CP phase, the formula for the volume of the Ω region is given by

$$vol(\Omega) = \frac{1}{(2\pi)^n n!} vol(PMNS)^2 \int_{\sigma_{min}}^{1} \prod_{k=1}^{n} \sigma_k \prod_{i<j} |\sigma_j^2 - \sigma_i^2|^2 \prod_{k=1}^{n} d\sigma_k, \tag{4.53}$$

and its numerical value is

$$vol(\Omega) = 1.12 \times 10^{-18}. \tag{4.54}$$

In the complex case, the contribution of the Ω region is even smaller than in the CP invariant scenario and it constitutes only 4.34×10^{-25} percent of the respective unit ball in the spectral norm (4.49). It may look like that $vol(\tilde{\Omega})$ is larger than the $vol(\Omega)$, however, we must keep in mind that $\tilde{\Omega}$ and Ω are structures of different dimensions, and thus cannot be compared directly.

We have established earlier the characterization of the Ω region as the intersection of the $\mathcal{B}(3)$ and U_{int} (4.8). The U_{int} can be treated as a hyperrectangle in $\mathbb{R}^9$ or $\mathbb{C}^9 \simeq \mathbb{R}^{18}$ respectively for the CP invariant case and the general case. As such, they also are geometric structures with associated volume. This volume is simply the product of the length of its sides, i.e. given intervals. Thus for the CP conserving case, it gives

$$vol(U_{int}) = 2.84 \times 10^{-10}. \tag{4.55}$$

Whereas when the CP phase is taken into account it is equal to

$$vol(U_{int}) = 2.27 \times 10^{-11}. \tag{4.56}$$

In [41] statistical analysis was performed concerning the amount of physically admissible mixing matrices contained within the interval matrix U_{int}. The analysis establishes

that for the CP invariant scenario only about 4% of matrices within the interval matrix are contractions. Comparison of volumes gives a similar qualitative result, namely contractions make a small part of the U_{int}. The exact calculation reveals that the volume of the Ω region constitutes 0.00023 percent of U_{int} in the CP conserving case and 8.12×10^{-6} percent for the general complex scenario.

5 Phenomenological implications

So far we have discussed formal aspects of the region of physically admissible mixing matrices. Now, we will focus on the application of the matrix theory methods in order to make phenomenological predictions, focusing in particular on the estimation of the active-sterile neutrino mixing.

5.1 Numerical analysis based on the α-parametrization

Most of the numerical estimations will be based on the α-parametrization (2.128).

$$U_{\texttt{PMNS}} = TU' = (I - \alpha)U'. \tag{5.1}$$

Since the singular values are unitarily invariant [110], i.e. the unitary matrix U' does not change their values, we can focus solely on the triangular matrix T which simplifies the calculations. Moreover, it is important that the left-hand singular matrix (see Theorem 3.1 in section 3.1) is the same for $U_{\texttt{PMNS}}$ and T. This can be seen if we consider the singular value decomposition of the T matrix $T = U\Sigma V^\dagger$. Then,

$$U_{\texttt{PMNS}} = TU' = U\Sigma V^\dagger U' = U\Sigma W^\dagger. \tag{5.2}$$

The next reason to work with the α-parametrization lies in the fact that it is possible to construct lower triangular matrices with a prescribed set of singular values [111]. The process is known as the inverse singular value problem. Thus, right from the beginning we control the number of singular values that are strictly less than one, i.e. we can specify the minimal number of additional neutrinos. Moreover, constructed lower triangular matrices have prescribed diagonal elements which in the case of triangular matrices equal to eigenvalues. As a result, we are able to construct the following matrix

$$T = \begin{pmatrix} \mathbf{t_{11}} & 0 & 0 \\ t_{21} & \mathbf{t_{22}} & 0 \\ t_{31} & t_{32} & \mathbf{t_{33}} \end{pmatrix} \tag{5.3}$$

with $\mathbf{S(T)} = \{\sigma_\mathbf{1}(\mathbf{T}), \sigma_\mathbf{2}(\mathbf{T}), \sigma_\mathbf{3}(\mathbf{T})\}$ (the highlighted quantities are prescribed).

The elements t_{ij} are restricted by the experimental data. The current bounds are presented in Tab. 1. These limits are given for three different massive scenarios: $m >$ EW (EW stands for the electroweak scale), $\Delta m^2 \gtrsim 100$ eV^2 and $\Delta m^2 \sim 0.1 - 1$ eV^2. The numerical estimation will also be performed for these three different massive scenarios.

Entry	(I): $m >$ EW	(II): $\Delta m^2 \gtrsim 100$ eV^2	(III): $\Delta m^2 \sim 0.1 - 1$ eV^2
$T_{11} = 1 - \alpha_{11}$	$0.99870 \div 1$	$0.976 \div 1$	$0.990 \div 1$
$T_{22} = 1 - \alpha_{22}$	$0.99978 \div 1$	$0.978 \div 1$	$0.986 \div 1$
$T_{33} = 1 - \alpha_{33}$	$0.99720 \div 1$	$0.900 \div 1$	$0.900 \div 1$
$T_{21} = \lvert\alpha_{21}\rvert$	$0.0 \div 0.00068$	$0.0 \div 0.025$	$0.0 \div 0.017$
$T_{31} = \lvert\alpha_{31}\rvert$	$0.0 \div 0.00270$	$0.0 \div 0.069$	$0,0 \div 0.045$
$T_{32} = \lvert\alpha_{32}\rvert$	$0.0 \div 0.00120$	$0.0 \div 0.012$	$0.0 \div 0.053$

Table 1: Limits on the elements of the $T = I - \alpha$ matrix for different non-standard neutrino mass scenarios (I)-(III). Limits for the α matrix elements α_{ij} are taken from [39] (with 95% CL), which are obtained from the global fits [112] to the experimental data [113–117]. Moreover, these data also include formalism connected with non-standard interactions [39].

5.2 Numerical distinguishability and continuity of singular values

It is important to distinguish numerically, with the chosen precision, matrices which exhibit two distinct sets of singular values. This is crucial when we consider subsets of the Ω region (4.3)-(4.5). To see under which circumstances it is possible, let us consider a simple two dimensional model of a lower triangular matrix.

We start from the following diagonal matrix

$$\begin{pmatrix} 0.99 & 0 \\ 0 & 0.99 \end{pmatrix}. \tag{5.4}$$

This matrix has obviously two singular values strictly less than one, equal to 0.99. Now we transform this matrix to the lower triangular form by adding to the position $(2,1)$ a parameter ϵ:

$$\begin{pmatrix} 0.99 & 0 \\ \epsilon & 0.99 \end{pmatrix}. \tag{5.5}$$

We will increase ϵ from 0 to 0.1 with different precision (Tab. 2) to see which precision guarantees such distinction. The behaviour of singular values for the matrix (5.5) is shown in Fig. 1.

Singular values will be fixed with the $3 \cdot 10^{-5}$ accuracy, as discussed in section 3.6. Thus, we are interested in the precision of ϵ which allows us distinguishing singular

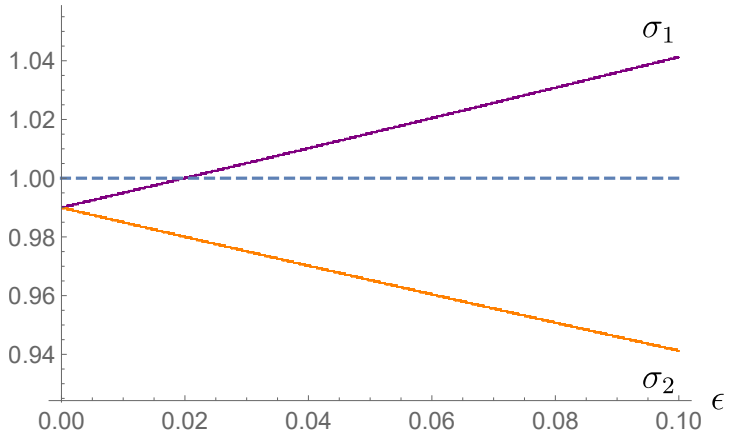

Figure 1: The behaviour of singular values σ_1 and σ_2 as a function of the ϵ parameter which changes from 0 to 0.1. The continuity in changes of singular values with ϵ is evident. We are looking for the precision of the parameter ϵ for which σ_1 has within the precision of calculations a common value with the horizontal dashed line at 1, which represents the physical limit, see Table 2.

values with this precision. In particular, we will determine when the first singular value can reach unambiguously the value one, i.e. $1 \pm 3 \cdot 10^{-5}$. The results are gathered in Tab. 2. We can see that the exact determination of the singular value is possible for the precision of ϵ at the level of 10^{-4}. It cannot be achieved with the precision of the matrix element ϵ at the level of 10^{-3}. Further increase of the ϵ precision results in a larger number of solutions so, as long as we stay within allowed intervals, by adjusting a proper density of sifting we are able to distinguish two singular values with imposed accuracy.

Initial singular value (σ_1)	0.99
Size of the ϵ step	Number of matrices
0.001	0
0.0001	1
0.00001	11

Table 2: A number of matrices that can reach the value of the first singular value (σ_1) equal to $1 \pm 3 \cdot 10^{-5}$ starting from 0.99. When the precision of ϵ increases, it is more probably to find non-zero solutions.

5.3 Numerical separation between the subsets of the Ω region

According to the division of the Ω region into the subsets corresponding to the minimal number of additional sterile neutrinos (4.3)-(4.5), it is natural to ask whether scenarios with different number of sterile neutrinos can be distinguished on the level of experimental limits. To answer this we have created lower triangular matrices with prescribed singular values which correspond to the subsets (4.3)-(4.5). After the multiple repetition of the procedure we have established upper and lower limits for elements of the T matrix for each subset. The results show that on the current experimental level of accuracy it is not possible to distinguish between scenarios with two and three additional neutrinos. However, for the scenario with only one additional neutrino, small deviations from experimental bounds given in Tab. 1 for some matrix elements have been observed. These results are presented in Tab. 3 where elements with the largest deviation from the experimental limits are highlighted in red.

	(I): $m >$ EW	(II): $\Delta m^2 \gtrsim 100$ eV^2	(III): $\Delta m^2 \sim 0.1 - 1$ eV^2
$(1,1)$	$0.99885 \div 0.99999$	$0.97641 \div 0.99996$	$0.99020 \div 0.99999$
Exp:	$0.99870 \div 1$	$0.976 \div 1$	$0.990 \div 1$
$(2,2)$	$0.99980 \div 0.99999$	**0.99331** $\div 0.99999$	$0.98646 \div 0.99999$
Exp:	$0.99978 \div 1$	**0.978** $\div 1$	$0.986 \div 1$
$(3,3)$	$0.99721 \div 0.99996$	$0.90040 \div 0.99985$	$0.90015 \div 0.99958$
Exp:	$0.99720 \div 1$	$0.900 \div 1$	$0.900 \div 1$
$(2,1)$	$0.00001 \div 0.00062$	$0.00031 \div$ **0.02214**	$0.00014 \div 0.01615$
Exp:	$0.0 \div 0.00068$	$0.0 \div$ **0.025**	$0.0 \div 0.017$
$(3,1)$	$0.00002 \div 0.00266$	$0.00048 \div 0.06892$	$0.00012 \div 0.04500$
Exp:	$0.0 \div 0.00270$	$0.0 \div 0.069$	$0.0 \div 0.045$
$(3,2)$	$0.00008 \div 0.00113$	$0.00052 - 0.01196$	$0.00024 \div 0.05281$
Exp:	$0.0 \div 0.00120$	$0.0 \div 0.012$	$0.0 \div 0.053$

Table 3: Lower and upper bounds for elements of the matrix T in the 3+1 scenario. For scenarios 3+2 and 3+3 we do not observe any changes of experimental limits given in Tab. 1.

5.4 Active-sterile mixing

In scenarios with sterile neutrinos we have at our reach only the 3×3 submatrix of the complete unitary mixing matrix, corresponding to the SM neutrinos mixing. Thus, it is important to estimate from the currently available experimental bounds possible limits on the active-sterile mixing, i.e. limits on the elements of the submatrix U_{lh} of (2.124). Such estimation can be provided on the matrix theory basis using the CS decomposition (discussed in section 3.3). Let us first focus on the 3+1 scenario, i.e. the scenario with one additional neutrino, where the complete unitary matrix is 4×4. In this case only matrices from the Ω_1 subset can be extended to a full 4×4 unitary matrix, i.e. contractions with only one singular value strictly less than one. This complete unitary matrix can be decomposed via CS decomposition as

$$\begin{pmatrix} U_{\texttt{PMNS}} & U_{lh} \\ U_{hl} & U_{hh} \end{pmatrix} = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & -s \\ \hline 0 & 0 & s & c \end{array}\right) \begin{pmatrix} Q_1^\dagger & 0 \\ 0 & Q_2^\dagger \end{pmatrix}. \tag{5.6}$$

We are interested in the estimation of the U_{lh}, which is given by

$$U_{lh} = W_1 S_{12} Q_2^\dagger, \tag{5.7}$$

where $W_1 \in \mathbb{C}^{3\times 3}$ is unitary, $S_{12} = (0, 0, -s)^T$ and $Q_2 = e^{i\theta}, \theta \in (0, 2\pi]$. The structure of the S_{12} matrix implies that only the third column of the W_1 matrix is important in our estimation

$$U_{lh} = -(w_{e3}, w_{\mu 3}, w_{\tau 3})^T s e^{-i\theta}. \tag{5.8}$$

Focusing on the absolute value of the active-sterile mixing we get

$$|(U_{lh})_{\alpha 3}| = |w_{\alpha 3} s e^{-i\theta}| = |w_{\alpha 3}| \cdot |s|, \ \alpha = e, \mu\, \tau. \tag{5.9}$$

The absolute value of the s can be expressed by the third non-unit singular value

$$|s| = |\sqrt{1 - c^2}| = |\sqrt{1 - \sigma_3^2}|. \tag{5.10}$$

Finally, the complete formula for the active-sterile mixing in the 3+1 scenario as a function of the third singular value is given by

$$|(U_{lh})_{\alpha 3}| = |w_{\alpha 3}| \cdot |\sqrt{1 - \sigma_3^2}|, \ \alpha = e, \mu\, \tau. \tag{5.11}$$

The similar steps provide the formula for the active-sterile mixing in the 3+2 scenario as functions of the σ_2 and σ_3. The CS decomposition of the 5×5 unitary matrix gives

$$\begin{pmatrix} U_{\texttt{PMNS}} & U_{lh} \\ U_{hl} & U_{hh} \end{pmatrix} = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix} \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & c_2 & 0 & -s_2 & 0 \\ 0 & 0 & c_3 & 0 & -s_3 \\ \hline 0 & s_2 & 0 & c_2 & 0 \\ 0 & 0 & s_3 & 0 & c_3 \end{array}\right) \begin{pmatrix} Q_1^\dagger & 0 \\ 0 & Q_2^\dagger \end{pmatrix}. \tag{5.12}$$

From which the submatrix corresponding to the active-sterile mixing reads

$$U_{lh} = W_1 S_{12} Q_2^\dagger, \tag{5.13}$$

where $W_1 \in \mathbb{C}^{3\times3}$, $Q \in \mathbb{C}^{2\times2}$ are unitary matrices and

$$S_{12} = \begin{pmatrix} 0 & 0 \\ -s_2 & 0 \\ 0 & -s_3 \end{pmatrix}. \tag{5.14}$$

The matrix form of the U_{lh} is as follows

$$U_{lh} = -\begin{pmatrix} w_{12}s_2q_{11} + w_{13}s_3q_{21} & w_{12}s_2q_{12} + w_{13}s_3q_{22} \\ w_{22}s_2q_{11} + w_{23}s_3q_{21} & w_{22}s_2q_{12} + w_{23}s_3q_{22} \\ w_{32}s_2q_{11} + w_{33}s_3q_{21} & w_{32}s_2q_{12} + w_{33}s_3q_{22} \end{pmatrix}. \tag{5.15}$$

As in the 3+1 scenario, we are interested in the estimation of the absolute values of the active-sterile mixing which is given by

$$|(U_{lh})_{\alpha j}| = |w_{\alpha2}s_1q_{1j} + w_{\alpha3}s_2q_{2j}| = |w_{\alpha2}q_{1j}\sqrt{1-\sigma_2^2} + w_{\alpha3}q_{2j}\sqrt{1-\sigma_3^2}|. \tag{5.16}$$

The Q_2 matrix is arbitrary which makes calculation much more difficult. However, if we are only interested in the largest value for the active-sterile mixing for each flavour, we can use triangle inequality and take $|q_{ij}| = 1$ to simplify the calculation

$$|(U_{lh})_{\alpha j}| \leq |w_{\alpha2}| \cdot |\sqrt{1-\sigma_2^2}| + |w_{\alpha3}| \cdot |\sqrt{1-\sigma_3^2}|. \tag{5.17}$$

Finally, we can derive the formula for the active-sterile mixing for the 3+3 scenario as a function of all three singular values. As before to accomplish this we use the CS

decomposition

$$\begin{pmatrix} U_{\text{PMNS}} & U_{lh} \\ U_{hl} & U_{hh} \end{pmatrix} = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix} \left(\begin{array}{ccc|ccc} c_1 & 0 & 0 & -s_1 & 0 & 0 \\ 0 & c_2 & 0 & 0 & -s_2 & 0 \\ 0 & 0 & c_3 & 0 & 0 & -s_3 \\ \hline s_1 & 0 & 0 & c_1 & 0 & 0 \\ 0 & s_2 & 0 & 0 & c_2 & 0 \\ 0 & 0 & s_3 & 0 & 0 & c_3 \end{array}\right) \begin{pmatrix} Q_1^\dagger & 0 \\ 0 & Q_2^\dagger \end{pmatrix}. \tag{5.18}$$

The submatrix corresponding to the active-sterile mixing is given by

$$U_{lh} = W_1 S_{12} Q_2^\dagger, \tag{5.19}$$

where $W_1 \in \mathbb{C}^{3\times3}$, $Q_2 \in \mathbb{C}^{3\times3}$ are unitary matrices and

$$S_{12} = \begin{pmatrix} -s_1 & 0 & 0 \\ 0 & -s_2 & 0 \\ 0 & 0 & -s_3 \end{pmatrix}. \tag{5.20}$$

After multiplying these three matrices, the matrix form of the active-sterile submatrix takes the following form

$$U_{lh} = \begin{pmatrix} -w_{11}s_1q_{11} - w_{12}s_2q_{21} - w_{13}s_3q_{31} & -w_{11}s_1q_{12} - w_{12}s_2q_{22} - w_{13}s_3q_{32} & -w_{11}s_1q_{13} - w_{12}s_2q_{23} - w_{13}s_3q_{33} \\ -w_{21}s_1q_{11} - w_{22}s_2q_{21} - w_{23}s_3q_{31} & -w_{21}s_1q_{12} - w_{22}s_2q_{22} - w_{23}s_3q_{32} & -w_{21}s_1q_{13} - w_{22}s_2q_{23} - w_{23}s_3q_{33} \\ -w_{31}s_1q_{11} - w_{32}s_2q_{21} - w_{33}s_3q_{31} & -w_{31}s_1q_{12} - w_{32}s_2q_{22} - w_{33}s_3q_{32} & -w_{31}s_1q_{13} - w_{32}s_2q_{23} - w_{33}s_3q_{33} \end{pmatrix}. \tag{5.21}$$

We are interested in the estimation of the absolute value of the active-sterile mixing which is given by

$$\begin{aligned} |(U_{lh})_{\alpha j}| =& |w_{\alpha1}s_1q_{1j} + w_{\alpha2}s_2q_{2j} + w_{\alpha3}s_3q_{3j}| = \\ & |w_{\alpha1}q_{1j}\sqrt{1-\sigma_1^2} + w_{\alpha2}q_{2j}\sqrt{1-\sigma_2^2} + w_{\alpha3}q_{3j}\sqrt{1-\sigma_3^2}|. \end{aligned} \tag{5.22}$$

The Q_2 matrix is an arbitrary unitary matrix which makes the calculation more difficult. However, if we are only interested in the estimation of the largest value for each flavour, we can use triangle inequality and take $|q_{ij}| = 1$

$$|(U_{lh})_{\alpha j}| \leq |w_{\alpha1}| \cdot |\sqrt{1-\sigma_1^2}| + |w_{\alpha2}| \cdot |\sqrt{1-\sigma_2^2}| + |w_{\alpha3}| \cdot |\sqrt{1-\sigma_3^2}|. \tag{5.23}$$

In this way we have established the analytical formulas (5.11),(5.16) and (5.22) for the active-sterile mixing as a function of singular values for scenarios with a different

minimal number of additional neutrinos. The next natural step is to perform numerical analysis to obtain bounds for the active-sterile mixing. As it is clearly seen from the formulas, the difficulty of the calculation grows with the number of additional sterile neutrinos. The necessary ingredients, i.e singular values and singular vectors w_{ij}, come from the singular value decomposition of the lower triangular matrix T. As we have shown in (5.2) the left singular matrix W_1 is the same for the T matrix and for the $U_{\texttt{PMNS}}$ matrix, and hence we can perform analysis on a much simpler T matrix. What follows, we present the numerical results for the scenario with one additional neutrino:

- `(I)`: $m >$ EW.

$$\begin{aligned} &|U_{e4}| \in [0, 0.021]\,, \quad |U_{\mu 4}| \in [0.00013, 0.021]\,, \quad |U_{\tau 4}| \in [0.0115, 0.075]\,. \\ &|U_{e4}| \leq 0.041 \;\; [118], \quad |U_{\mu 4}| \leq 0.030 \;\; [118], \qquad\qquad |U_{\tau 4}| \leq 0.087 \;\; [118]. \end{aligned} \tag{5.24}$$

- `(II)`: $\Delta m^2 \gtrsim 100$ eV^2.

$$|U_{e4}| \in [0, 0.082]\,, \quad |U_{\mu 4}| \in [0.00052, 0.099]\,, \quad |U_{\tau 4}| \in [0.0365, 0.44]\,. \tag{5.25}$$

- `(III)`: $\Delta m^2 \sim 0.1 - 1$ eV^2.

$$\begin{aligned} &|U_{e4}| \in [0, 0.130]\,, \qquad\qquad |U_{\mu 4}| \in [0.00052, 0.167]\,, \quad |U_{\tau 4}| \in [0.0365, 0.436]\,. \\ &|U_{e4}| \in [0.114, 0.167] \;\; [119], \quad |U_{\mu 4}| \in [0.0911, 0.148] \;\; [119], \quad |U_{\tau 4}| \leq 0.361 \;\; [120]. \end{aligned} \tag{5.26}$$

As we can see, for the seesaw like neutrinos, i.e. very heavy sterile neutrinos (scenario `I`), results obtained with the use of singular values give stringent upper bounds for the active-sterile mixing when comparing to results obtained by others. The analysis in [118] assumes that each new lepton mixes with only one SM family. What is interesting, in some cases ($|U_{\mu 4}|, |U_{\tau 4}|$) we find a preference for non-zero mixings. For the intermediate massive scenario (`II`), to our knowledge, these are the first established limits for the active-sterile mixing. In the case of light sterile neutrinos (`III`), the results obtained with our method are comparable with other results available in literature.

5.5 Norms as non-unitarity quantifiers

If any singular value of a studied matrix differs from unity, then the matrix is surely non-unitary. However, not all the time we need such precise analysis. To check unitarity

condition, we can calculate the Hermitian product of a matrix, and check whether it satisfies unitarity conditions $UU^\dagger = I$ and $U^\dagger U = I$. In particular, we want to calculate the difference between the identity matrix and the Hermitian product of the matrix

$$\begin{aligned} I - UU^\dagger, \\ I - U^\dagger U. \end{aligned} \tag{5.27}$$

This tells us how much a given matrix differs from unitarity. However, if we would like to measure how this deviation differs from matrix to matrix, it is not clear how to do that using the above formulas. Thus, it is preferable to find a matrix function that measures the deviation from unitarity as one number.

Definition 14. *Let $A \in \mathbb{C}^{n\times n}$ be a given matrix and let $\| \cdot \|$ be any matrix norm, then the function*

$$\|I - AA^\dagger\| \tag{5.28}$$

measures how far A is from the unitary matrix.

So far we have used only spectral norm, however, there is an abundance of matrix norms, which can be used to measure the deviation from unitarity. Many of them are simple functions of matrix elements and thus are easy to calculate. We will discuss some of them, each measuring non-unitarity on different matrix level and highlighting different physical properties. The norms we are interested in are

1. Spectral norm : $\|A\|_2 = max_{\|x\|_2=1}\|Ax\|_2 = \sigma_1(A)$. (5.29)
2. Frobenius norm : $\|A\|_F = \sqrt{Tr(A^\dagger A)} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}|a_{ij}|^2} = \sqrt{\sum_{i=1}^{min\{n,m\}}\sigma_i(A)^2}$. (5.30)
3. Maximum absolute column sum norm (MACN): $\|A\|_1 = \max_j \sum_{i=1}^{m}|a_{ij}|$. (5.31)
4. Maximum absolute row sum norm (MARN): $\|A\|_\infty = \max_i \sum_{j=1}^{m}|a_{ij}|$. (5.32)

We will also consider the maximal absolute value of the elements $|A|_{max} = \max_{ij}|a_{ij}|$ as it is usually used to quantify the deviation from unitarity.

Let us construct a matrix from the Ω region and calculate its deviation from unitarity using the above norms. The matrix under consideration will be constructed as a convex combination of four PMNS matrices

$$U = \alpha_1 U_1 + \alpha_2 U_2 + \alpha_3 U_3 + \alpha_4 U_4. \tag{5.33}$$

As an example, we have chosen the following values for the mixing parameters for U_i matrices

$$\begin{aligned} U_1 &: \theta_{12} = 31.27^\circ, \theta_{13} = 8.20^\circ, \theta_{23} = 40.1^\circ, \delta = 120^\circ, \\ U_2 &: \theta_{12} = 35.86^\circ, \theta_{13} = 8.93^\circ, \theta_{23} = 51.7^\circ, \delta = 369^\circ, \\ U_3 &: \theta_{12} = 33.21^\circ, \theta_{13} = 8.54^\circ, \theta_{23} = 44.4^\circ, \delta = 238^\circ, \\ U_4 &: \theta_{12} = 32.67^\circ, \theta_{13} = 8.78^\circ, \theta_{23} = 48.5^\circ, \delta = 176^\circ, \end{aligned} \tag{5.34}$$

and the following values for α_i coefficients

$$\alpha_1 = \frac{12}{100}, \alpha_2 = \frac{48}{100}, \alpha_3 = \frac{27}{100}, \alpha_4 = \frac{13}{100}. \tag{5.35}$$

The resulting matrix is as follows

$$U = \begin{pmatrix} 0.817367 & 0.555024 & 0.0239921 - 0.00614i \\ -0.391253 + 0.00345799i & 0.538647 + 0.00215621i & 0.731669 \\ 0.409187 + 0.00400813i & -0.619299 + 0.00269373i & 0.66064 \end{pmatrix}, \tag{5.36}$$

with the following set of singular values $S(U) = \{0.99905, 0.99165, 0.98046\}$, and as expected it is a contraction. The evaluation of the deviation from unitarity gives

$$\begin{aligned} \|I - UU^\dagger\|_2 &= 0.0386998, \\ \|I - UU^\dagger\|_F &= 0.0421683, \\ \|I - UU^\dagger\|_1 &= 0.0433827, \\ \|I - UU^\dagger\|_\infty &= 0.0433827, \\ |I - UU^\dagger|_{\max} &= 0.0232465. \end{aligned} \tag{5.37}$$

We can see that the maximal absolute value of elements gives the smallest result. It is useful if we are concerned about the largest deviation amongst the elements. However, it does not provide information about the deviation of other elements beside that it is smaller. Other considered norms treat the deviation from uniatarity more globally, taking into account entire column or all matrix elements. From these norms the spectral norm gives the smallest result and this is expected from the compatibility of matrix norms [77]. The MACN and MARN norms gave the same result, as the calculation is performed on a symmetric matrix. The spectral and Frobenius norm can be viewed as global indicators of the deviation from unitarity as they involve all matrix elements. On the other hand, the MACN and MARN norms measure this deviation on the level of rows, i.e. on the flavour level, even if are disturbed by the addition of cross product

of elements, which is relatively small for matrices close to unitarity. To measure the deviation on the column level, i.e. on the massive level, we must calculate MACN or MARN for $I - U^\dagger U$ instead of $I - UU^\dagger$.

It is worth mentioning that all information about deviation from unitarity important from the physical point of view is encoded in the diagonal of the matrices $I - U^\dagger U$ and $I - UU^\dagger$. The elements on the diagonal contain information about the deviation for a given flavour or massive state, depending on the choice of $U^\dagger U$ or $UU^\dagger$. Whereas the trace gives a square of the Frobenius norm, which measures the deviation from unitarity globally.

5.6 Quark sector

On the occasion, as the discussed matrix theory formalism is universal, let us consider the quark sector. In the quark sector, similarly to the lepton sector, the mass matrices, in general, are not diagonal. Thus, in order to obtain quarks with definite masses, these matrices must be diagonalized. For quarks in general the mass matrix is a 3×3 complex matrix and as such it must be diagonalized by the singular value decomposition. As the result there exists a mixing between weak quark states and mass quark states

$$\begin{aligned} q_L^U &= V_L^{U\dagger} q_L'^U = (u_L, c_L, t_L)^T, \quad q_R^U = V_R^{U\dagger} q_R'^U = (u_R, c_R, t_R)^T, \\ q_L^D &= V_L^{D\dagger} q_L'^D = (d_L, s_L, b_L)^T, \quad q_R^D = V_R^{D\dagger} q_R'^D = (d_R, s_R, b_R)^T, \end{aligned} \tag{5.38}$$

where $q_{L,R}^{U,D}$ are arrays of quark fields with definite masses and $q_{L,R}'^{U,D}$ are arrays of quark fields with definite flavours. However, the only physical effects of the mixing between weak and mass eigenstates are visible through the charged-current interactions which contain only left-handed chiral fields

$$j_{Q,W}^\mu = 2\overline{q_L'^U}\gamma^\mu q_L'^D. \tag{5.39}$$

By expressing the weak eigenstate quark fields in the mass eigenstate we get

$$j_{Q,W}^\mu = 2\overline{q_L^U} V_L^{U\dagger}\gamma^\mu V_L^D q_L^D = 2\overline{q_L^U}\gamma^\mu V_L^{U\dagger} V_L^D q_L^D. \tag{5.40}$$

Thus, physically meaningful effects of quark mixing are encoded in the product $V_L^{U\dagger} V_L^D$. In the quark sector, this matrix is called Cabibbo-Kobayashi-Masakawa (CKM) mixing matrix [121, 122], i.e.

$$V_{CKM} = V_L^{U\dagger} V_L^D = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{5.41}$$

As the product of two unitary matrices, the CKM mixing matrix is also unitary. It shares many similarities with the PMNS mixing matrix of the neutrino sector. Alike for Dirac neutrinos, $2n-1$ complex phases can be eliminated from the CKM matrix by the redefinition of the quarks fields. As the result, we have exactly same the number of parameters as for the PMNS mixing matrix in the case of Dirac neutrinos

$$\begin{aligned} &\text{angles:} \quad \frac{n(n-1)}{2}, \\ &\text{phases:} \quad \frac{(n-1)(n-2)}{2}, \end{aligned} \tag{5.42}$$

which for a 3×3 matrix gives 3 angles and 1 phase. This means that it can be parametrized in the exactly same way as the PMNS matrix

$$V_{CKM} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.43}$$

The main difference between the CKM mixing matrix and the PMNS matrix is the value of mixing parameters and the resulting exact structure of the mixing matrix. The current bounds for the mixing parameters in the quark sector are [59]

$$\begin{aligned} &\sin\theta_{12} = 0.22650 \pm 0.00048, \quad &&\sin\theta_{23} = 0.04053^{+0.00083}_{-0.00061}, \\ &\sin\theta_{13} = 0.00361^{+0.00011}_{-0.00009}, &&\delta = 1.196^{+0.045}_{-0.043}. \end{aligned} \tag{5.44}$$

These parameters exhibit the following hierarchy $s_{13} \ll s_{23} \ll s_{12} \ll 1$ which translates into the allowed ranges of the mixing elements

$$|V_{CKM}| = \begin{pmatrix} [0.9739, 0.97412] & [0.22602, 0.22698] & [0.00352, 0.00372] \\ [0.22588, 0.22684] & [0.97309, 0.97331] & [0.03992, 0.04136] \\ [0.00838, 0.00877] & [0.03918, 0.0406] & [0.999137, 0.999196] \end{pmatrix}, \tag{5.45}$$

or written in an explicit interval form

$$\begin{aligned} &V_{ud} \in [0.9739, 0.97412]\,, \\ &V_{us} \in [0.22602, 0.22698]\,, \\ &V_{ub} \in [0.00114 - 0.00352i, 0.00151 - 0.00322i]\,, \\ &V_{cd} \in [-0.22686 - 0.00014i, -0.22587 - 0.00013i]\,, \\ &V_{cs} \in [0.97305 - 0.000033i, 0.97334 - 0.000029i]\,, \\ &V_{cb} \in [0.03992, 0.04136]\,, \\ &V_{td} \in [0.00755 - 0.00343i, 0.00827 - 0.00313i]\,, \\ &V_{ts} \in [-0.04059 - 0.00079i, -0.03914 - 0.00073i]\,, \\ &V_{tb} \in [0.999137, 0.999196]\,. \end{aligned} \tag{5.46}$$

The CKM matrix has almost a diagonal structure in comparison to the neutrinos sector where the elements of the mixing matrix are much more uniformly distributed. This almost diagonal structure of the quark mixing matrix allows to represent it in the so-called Wolfenstein parametrization [123, 124]

$$V_{CKM} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{5.47}$$

The Wolfenstein parameters are defined as follows

$$s_{12} = \lambda, \quad s_{23} = A\lambda^2, \quad s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta). \tag{5.48}$$

The present limits for the Wolfenstein parameters are [59]

$$\begin{aligned} &\lambda = 0.22650 \pm 0.00048, \quad A = 0.790^{+0.017}_{-0.012}, \\ &\bar{\rho} = 0.141^{+0.016}_{-0.017}, \qquad \bar{\eta} = 0.357 \pm 0.011, \end{aligned} \tag{5.49}$$

where $\bar{\rho} = \rho(1 - \frac{\lambda^2}{2} + \dots)$ and $\bar{\eta} = \eta(1 - \frac{\lambda^2}{2} + \dots)$.

Regardless of the difference in the exact structure between neutrino and quark mixing matrices, the methodology developed for the neutrino sector translates one-to-one to the quark sector. As a deviation from the unity of any singular value is a signal of non-unitarity of the matrix, it is even more desirable to perform singular value analysis of CKM mixing matrices, as we have experimental access from different processes to all the matrix elements with much higher precision than for neutrinos. The calculated maximal and minimal singular values allowed for the interval quark mixing matrix (5.46) are:

$$\begin{aligned} \sigma_1 &\in [0.99997, 1.00101], \\ \sigma_2 &\in [0.99965, 1.00037], \\ \sigma_3 &\in [0.99890, 1.00002]. \end{aligned} \tag{5.50}$$

The obtained ranges for allowed singular values for the CKM mixing matrices are much narrower than those for neutrinos (3.46). It means that the CKM mixing matrix is much closer to the unitary matrix than the PMNS mixing matrix. Nevertheless, there is still a possibility that the CKM mixing matrix is not unitary, however, the physically interesting deviation from unitarity, i.e. $\sigma_1 \leq 1.00003$, is visible only by investigating the second and the third singular value.

It might seem that with the considered precision three additional quarks are forbidden. However, referring to the discussion at the beginning of Chapter 4, contraction

with minimal unitary dilation of dimension smaller than 6 can also be extended to a 6×6 unitary matrix. Nevertheless, physically meaningful mixing matrices for quarks are restricted to the unit sphere in the spectral norm $\mathcal{S}(3) = \{A \in \mathbb{C}^{3\times 3} : \|A\| = 1\}$. In other words, they lie on the boundary of the unit ball (4.7) and do not go inside.

6 Connection between masses and mixing

The neutrino masses and mixing are inseparably connected. The mixing matrix diagonalizes the mass matrix (2.26). To study them simultaneously it is necessary to solve the eigenvalue problem. This becomes an issue when dimensions grow, which is the case when sterile neutrinos are involved. When, the CP phase is non-zero, which is currently strongly supported by experimental results, the problem is even more difficult, as we have to find singular values of the mass matrix. Thus, often we study these two matrices separately guided by experimental insight. One way to study these two matrices together as a system is in terms of angles between spaces spanned by the eigenvectors. This is a very broad and interesting subject, to our knowledge, so far not discussed in the literature in the context of neutrino physics.

Before we move to the main theorem which allows such description, let us state the auxiliary theorem which highlights geometric aspects of the relation between subspaces [76, 125].

Theorem 6.1. *Let X_1, Y_1 be $n \times l$ matrices with orthonormal columns. Then there exist $l \times l$ unitary matrices U_1 and V_1, and an $n \times n$ unitary matrix Q, such that if $2l \leq n$, then*

$$QX_1U_1 = \begin{pmatrix} I \\ 0 \\ 0 \end{pmatrix}, \tag{6.1}$$

$$QY_1V_1 = \begin{pmatrix} C \\ S \\ 0 \end{pmatrix}, \tag{6.2}$$

where C, S are diagonal matrices with diagonal entries $0 \leq c_1 \leq ... \leq c_l \leq 1$ and $1 \geq s_1 \geq ... \geq s_l \geq 0$, respectively, and $C^2 + S^2 = I$.

The geometric interpretation of this theorem is as follows. Let $\mathcal{E}$ and $\mathcal{F}$ be l-dimensional subspaces of $\mathbb{C}^n$. Then, we can transform $\mathbb{C}^n$ by the unitary transformation Q such that the columns of matrices

$$\begin{pmatrix} I \\ 0 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} C \\ S \\ 0 \end{pmatrix} \tag{6.3}$$

form orthogonal bases for $Q\mathcal{E}$ and $Q\mathcal{F}$, respectively. Space spanned by the columns for which $c_i = 0$, is the orthogonal complement of $Q\mathcal{E}$ in $Q\mathcal{F}$. The relation $s_i^2 + c_i^2 = 1$

suggests that s_i and c_i can be treated as the sines and cosines between these subspaces. Moreover, these two spaces are equal to each other if and only if $S = 0$. Thus, the matrix S measures how close subspaces $\mathcal{E}$ and $\mathcal{F}$ are to each other. This allows us to make the following definition

Definition 15. *Let $\mathcal{E}$ and $\mathcal{F}$ be $l-$dimensional subspaces of $\mathbb{C}^n$. The angel operator between $\mathcal{E}$ and $\mathcal{F}$ is defined as follows*

$$\Theta(\mathcal{E}, \mathcal{F}) = \arcsin S. \tag{6.4}$$

It is a diagonal matrix whose diagonal elements are called the canonical (principal) angles between subspaces $\mathcal{E}$ and $\mathcal{F}$.

Moreover, using this operator, we can define the distance or gap between two subspaces as

Definition 16. *Let $\mathcal{E}$ and $\mathcal{F}$ let be $l-$dimensional subspaces of $\mathbb{C}^n$. Let E and F be orthogonal projection onto $\mathcal{E}$ and $\mathcal{F}$ respectively. The distance between subspaces $\mathcal{E}$ and $\mathcal{F}$ is defined to be*

$$\|E - F\| = \|E^{\perp}F\| = \|\sin\Theta\| \tag{6.5}$$

The perturbation behaviour of individual eigenvectors is much more complicated than the behaviour of eigenvalues. However, the spaces spanned by eigenvectors behave quite nicely. This behaviour is described by the Davis-Kahan theorem [126].

Theorem 6.2. *Let A and B be Hermitian operators, and let S_1 be an interval $[a, b]$ and S_2 be the complement of $(a - \delta, b + \delta)$ in $\mathbb{R}$. Let $E = P_A(S_1), F^{\perp} = P_B(S_2)$ be orthogonal projections onto subspaces spanned by eigenvectors of A and B corresponding to eigenvalues from S_1 and S_2 respectively. Then for every unitarily invariant norm,*

$$\|EF^{\perp}\| \leq \frac{1}{\delta}\|E(A - B)F^{\perp}\| \leq \frac{1}{\delta}\|A - B\|, \tag{6.6}$$

where

$$\delta = dist(\sigma(A), \sigma(B)) = min\{|\lambda - \mu| : \lambda \in \sigma(A), \mu \in \sigma(B)\}, \tag{6.7}$$

and $\sigma(A)$, $\sigma(B)$ are spectra of matrices A and B, respectively.

6.1 Seesaw family

The most controllable behaviour of eigenspaces is ensured for the Hermitian matrices. However, every problem for singular values can be transformed into an equivalent

problem for eigenvalues of Hermitian matrices (Theorem 3.2). Thus, the most important scenarios from the neutrino physics perspective can be covered. At our disposal, we have the Davis-Kahan theorem which describes the behaviour of eigenspaces when the matrices experience perturbation. Such a situation can be identified in the mass matrices of the seesaw family.

Definition 17. *The neutrino mass matrix is called the seesaw matrix if it has the following form*

$$M = \begin{pmatrix} 0 & M_D \\ M_D^T & M_R \end{pmatrix} \tag{6.8}$$

with elements of M_D *and* M_R *satisfying* $|m_D| \ll |m_R|$.

All the matrices with a zero top diagonal block and appropriate scale of elements of off-diagonal blocks can be reorganized into the seesaw mass matrix. For instance, this is the case for the inverse seesaw and the linear seesaw [127]. Further, the seesaw mass matrix can be written in a perturbative way

$$M = \begin{pmatrix} 0 & M_D \\ M_D^T & M_R \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & M_R \end{pmatrix} + \begin{pmatrix} 0 & M_D \\ M_D^T & 0 \end{pmatrix} \equiv \mathcal{M}_R + \mathcal{M}_D, \tag{6.9}$$

where the $\mathcal{M}_D$ matrix is treated as the perturbation of the $\mathcal{M}_R$ matrix. This representation of the M matrix allows for the immediate application of the Davis-Kahan theorem (Theorem 6.2).

6.2 Separation between eigenspaces in the seesaw scenario

We are interested in how masses and mixings are connected to each other in the seesaw scenario[3]. To pose the problem, we will study the behaviour of the eigenspace of the matrix $\mathcal{M}_R$ under the perturbation $\mathcal{M}_D$ (6.9). Thus, we are interested in the estimation of the difference between eigenspaces spanned by eigenvectors of $\mathcal{M}_R$ and M (6.9). As an example, we will study the CP invariant case that allows for the direct application of the Davis-Kahan theorem (Theorem 6.2). The CP-violating case can also be treated by this approach by rephrasing the problem for singular values to the eigenproblem for a Hermitian matrix (Theorem 3.2).

As a starting point let us consider the eigenproblem for the matrix $\mathcal{M}_R$. For a block-diagonal matrix eigenvalues correspond to the eigenvalues of its diagonal blocks.

[3]Recently, an interesting relation has been found between eigenvectors and eigenvalues for neutrino oscillations in [128, 129].

In this case, one of these blocks is a zero matrix. Thus, this block has a threefold eigenvalue 0 and the corresponding eigenvectors are

$$(1,0,0,0,0,0,...,0)^T, (0,1,0,0,0,0,...,0)^T,$$
$$(0,0,1,0,0,0,...,0)^T. \tag{6.10}$$

They span a standard 3-dimensional Euclidean space embedded in a $(3+N_R)$-dimensional space. The rest of the eigenvalues of $\mathcal{M}_R$ corresponds to those of the M_R submatrix. The Davis-Kahan theorem allows us to estimate the sine of the angle between subspaces, denoted as $\sin\Theta$ in (6.5), spanned by the eigenvectors. Since the eigenspace spanned by the zero eigenvalues of $\mathcal{M}_R$ has a very simple structure, we will focus on the estimation of the angle between spaces corresponding to light neutrinos. Information about the other pair of subspaces follows immediately from the orthogonality of the mixing matrix. Let us denote the eigenspaces spanned by the eigenvectors corresponding to small eigenvalues by V_L and $V_L^{'}$, respectively for $\mathcal{M}_R$ and M. Then in the seesaw scenario we have

$$\|\sin\Theta(V_L, V_L^{'})\| \leq \frac{1}{\delta}\|M - \mathcal{M}_R\| = \frac{1}{\delta}\|\mathcal{M}_D\|, \tag{6.11}$$

where δ is the distance between the largest of the light masses and the smallest of the heavy masses. The relation is illustrated in Fig. 2.

The above inequality says that $\sin\Theta(V_L, V_L^{'})$ can be estimated using a gap between spectra and the size of the perturbation. It is clear that if the subspaces V_L and $V_L^{'}$ are close to each other then the sine between them will tend to zero. Therefore, from (6.11) we can draw the following general conclusions:

- If the separation between light and heavy neutrinos is pronounced like in the seesaw case, then the subspace spanned by light neutrinos is almost parallel to the 3-dimensional Euclidean space. However, when these two spectra approach each other not much information can be retrieved from the Davis-Kahan theorem (Theorem 6.2).

- Even if the δ is not that large, these two subspaces still can be almost parallel when $\mathcal{M}_D$ is very small.

We have considered the perturbation behaviour of eigenspaces of the neutrino mass matrix in the CP conserving seesaw scenario. The general complex seesaw mass matrix analysis is much more difficult and requires studying singular values. This can

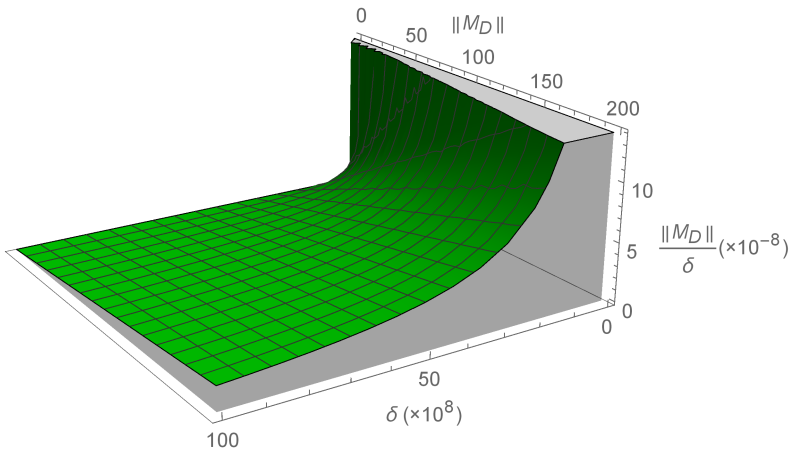

Figure 2: The behaviour of the $\sin\Theta$, controlled by the Davis-Kahan theorem, in the seesaw scenario. The region below the graph represents allowed values. In this case, $\sin\Theta$ is bounded by a function depending on the norm of $\mathcal{M}_D$ and the gap between the spectrum. As heavy neutrinos become lighter and lighter, the blowout of the bound is observed.

be transformed to the eigenproblem for Hermitian matrices (Theorem 3.2) but the dimensionality of the matrix grows. On top of that, it would be interesting to find a CS decomposition of the seesaw mixing matrix, as it encodes the principal angles between subspaces.

7 Summary and outlook

The concepts which have their origin in matrix theory can enrich neutrino studies significantly. The major role in the development of neutrino mixing analysis by matrix theory methods has been played by singular values. First of all, they allow defining physically admissible mixing matrices as contractions, i.e. matrices with the largest singular value less than or equal to one. The recognition of contractions as physically meaningful mixing matrices was crucial for further development, since only those matrices can be extended to a larger unitary matrix which appears when sterile neutrinos are under consideration. We have found that all physically admissible mixing matrices form a geometric region defined as the convex hull spanned by the 3-dimensional unitary PMNS mixing matrices (Chapter 4). The structure of this region has important physical consequences, namely, it can be divided into four disjoint subsets corresponding to the different minimal number of additional neutrinos. *This means that the minimal number of additional neutrinos is not arbitrary and is encoded in the number of singular values strictly less than one.* In this way we can consider independently scenarios with different number of sterile neutrinos, resulting in separate constraints on active-sterile mixing. Motivated by this observation, we have undertaken the geometric analysis of this region. Based on the fact that the region of physically admissible mixing matrices is a subset of the unit ball of the spectral norm, we found out that the subsets corresponding to the different minimal number of additional neutrinos are the relative interiors of the faces of the unit ball with parameters restricted by experiments. We have calculated the volume of this region expressed by the Haar measure of the singular value decomposition, providing information about its size. We have also made a step towards an efficient way of constructing physically interesting matrices as every contraction can be written as a convex combination of unitary matrices. The upper limit for the number of unitary matrices necessary to construct any contraction is known as the Carathéodry number. We have proven that for 3×3 contractions the Carathéodry number is equal to four. Moreover, for matrices corresponding to one and two additional neutrinos, this number can be reduced to two and three, respectively.

The second part of the dissertation (Chapter 5) focused on the application of matrix theory methods to phenomenological studies. The emphasis has been put on the estimation of the active-sterile mixing. The approach we have undertaken is based on the so-called unitary dilation procedure, which puts a given contraction on the top diagonal block of a larger unitary matrix. The practical application of this procedure has been done by using the cosine-sine decomposition. We were able to derive analytic

formulas for active-sterile mixing as a function of singular values for each scenario with a different minimal number of additional neutrinos. On that basis, numerical estimation of the active-sterile mixing in the scenario with one sterile neutrino has been performed resulting in new stringent upper bounds, and in some cases lower bounds. On top of that, we have discussed the application of matrix norms as the indicators of non-unitarity of the mixing matrix. A simple example has been presented showing that different norms measure deviations from unitarity on different matrix levels. Finally, as the matrix theory methods translate one-to-one into the quark sector, we made a brief comparison of both neutrino and quark sectors based on singular values.

In the last chapter, we studied the connection between neutrino masses and mixing in terms angles between subspaces spanned by the eigenvectors of the mass matrix. As an example, we analysed the seesaw family mass matrices. We used the Davis-Kahan theorem to quantify angles between subspaces in terms of the Dirac mass matrix and the gap between active and sterile neutrino masses.

Our findings provide new concepts in neutrino mixing analysis, allowing for a consistent treatment of neutrino phenomena, including extra neutrino species, and to go towards refined neutrino mass and mixing model constructions. The material presented in this dissertation is by no means exhaustive. However, it gives a solid basis for further studies. The first step should be the numerical estimation of the active-sterile mixing in scenarios with two and three sterile neutrinos. On the theoretical side, it would be important to study carefully the sequence of unitary dilations of increasingly higher dimensions. Such analysis can shed a light on the upper limit of additional neutrinos. On top of that, the connection between masses and mixing deserves in-depth studies.

Matrix theory is a broad subject, and more interesting connections between physics and mathematics are waiting for discovery. Another particularly interesting direction from the neutrino physics perspective is the application of the inverse eigenvalue problem in studies of structures of the mass and mixing matrices.

References

[1] S. Schael, et al., Precision electroweak measurements on the Z resonance, Phys. Rept. 427 (2006) 257–454. arXiv:hep-ex/0509008, doi:10.1016/j.physrep.2005.12.006.

[2] V. A. Novikov, L. B. Okun, A. N. Rozanov, M. I. Vysotsky, Theory of Z boson decays, Rept. Prog. Phys. 62 (1999) 1275–1332. arXiv:hep-ph/9906465, doi:10.1088/0034-4885/62/9/201.

[3] G. Voutsinas, E. Perez, M. Dam, P. Janot, Beam-beam effects on the luminosity measurement at LEP and the number of light neutrino species, Phys. Lett. B800 (2020) 135068. arXiv:1908.01704, doi:10.1016/j.physletb.2019.135068.

[4] P. Janot, S. Jadach, Improved Bhabha cross section at LEP and the number of light neutrino species, Phys. Lett. B803 (2020) 135319. arXiv:1912.02067, doi:10.1016/j.physletb.2020.135319.

[5] Y. Fukuda, et al., Evidence for oscillation of atmospheric neutrinos, Phys. Rev. Lett. 81 (1998) 1562–1567. arXiv:hep-ex/9807003, doi:10.1103/PhysRevLett.81.1562.

[6] Q. R. Ahmad, et al., Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory, Phys. Rev. Lett. 89 (2002) 011301. arXiv:nucl-ex/0204008, doi:10.1103/PhysRevLett.89.011301.

[7] P. Minkowski, $\mu \to e\gamma$ at a Rate of One Out of 10^9 Muon Decays?, Phys. Lett. B67 (1977) 421–428. doi:10.1016/0370-2693(77)90435-X.

[8] T. Yanagida, Horizontal Symmetry and Masses of Neutrinos, Prog. Theor. Phys. 64 (1980) 1103. doi:10.1143/PTP.64.1103.

[9] M. Gell-Mann, P. Ramond, R. Slansky, Complex Spinors and Unified Theories, Conf. Proc. C790927 (1979) 315–321. arXiv:1306.4669.

[10] R. N. Mohapatra, G. Senjanovic, Neutrino Masses and Mixings in Gauge Models with Spontaneous Parity Violation, Phys. Rev. D23 (1981) 165. doi:10.1103/PhysRevD.23.165.

[11] T. A. Mueller, et al., Improved Predictions of Reactor Antineutrino Spectra, Phys. Rev. C83 (2011) 054615. arXiv:1101.2663, doi:10.1103/PhysRevC.83.054615.

[12] G. Mention, et. al., Reactor antineutrino anomaly, Phys. Rev. D 83 (2011) 073006. doi:10.1103/PhysRevD.83.073006.

[13] F. P. An, et al., Measurement of the Reactor Antineutrino Flux and Spectrum at Daya Bay, Phys. Rev. Lett. 116 (6) (2016) 061801, [Erratum: Phys.Rev.Lett. 118, 099902 (2017)]. arXiv:1508.04233, doi:10.1103/PhysRevLett.116.061801.

[14] Y. Ko, et al., Sterile Neutrino Search at the NEOS Experiment, Phys. Rev. Lett. 118 (12) (2017) 121802. arXiv:1610.05134, doi:10.1103/PhysRevLett.118.121802.

[15] G. Bak, et al., Measurement of Reactor Antineutrino Oscillation Amplitude and Frequency at RENO, Phys. Rev. Lett. 121 (20) (2018) 201801. arXiv:1806.00248, doi:10.1103/PhysRevLett.121.201801.

[16] H. de Kerret, et al., First Double Chooz θ_{13} Measurement via Total Neutron Capture Detection, Nature Phys. 16 (5) (2020) 558–564. arXiv:1901.09445, doi:10.1038/s41567-020-0831-y.

[17] A. P. Serebrov, et. al, First observation of the oscillation effect in the neutrino-4 experiment on the search for the sterile neutrino, JETP Letters 109 (4) (2019) 213–221. doi:10.1134/S0021364019040040.

[18] A. P. Serebrov, R. M. Samoilov, Analysis of the Results of the Neutrino-4 Experiment on the Search for the Sterile Neutrino and Comparison with Results of Other Experiments, JETP Lett. 112 (4) (2020) 199–212. arXiv:2003.03199, doi:10.31857/S1234567820160016.

[19] C. P. Serebrov, et al., Experiment Neutrino-4 search for sterile neutrino and results of measurements. arXiv:2005.05301.

[20] M. Andriamirado, et al., Note on arXiv:2005.05301, 'Preparation of the Neutrino-4 experiment on search for sterile neutrino and the obtained results of measurements'. arXiv:2006.13147.

[21] M. V. Danilov, N. A. Skrobova, Comment on "Analysis of the Results of the Neutrino-4 Experiment on the Search for the Sterile Neutrino and Comparison with Results of Other Experiments" (JETP Letters 112, 199 (2020)), JETP Lett. 112 (7) (2020) 452–454. doi:10.1134/S0021364020190066.

[22] C. Giunti, Y. F. Li, C. A. Ternes, Y. Y. Zhang, Neutrino-4 anomaly: oscillations or fluctuations?, Phys. Lett. B 816 (2021) 136214. arXiv:2101.06785, doi:10.1016/j.physletb.2021.136214.

[23] J. N. Abdurashitov, et al., Measurement of the response of a Ga solar neutrino experiment to neutrinos from a ^{37}Ar source, Phys. Rev. C73 (2006) 045805. arXiv:nucl-ex/0512041, doi:10.1103/PhysRevC.73.045805.

[24] J. Kostensalo, J. Suhonen, C. Giunti, P. C. Srivastava, The gallium anomaly revisited, Phys. Lett. B 795 (2019) 542–547. arXiv:1906.10980, doi:10.1016/j.physletb.2019.06.057.

[25] C. Giunti, M. Laveder, Statistical Significance of the Gallium Anomaly, Phys. Rev. C83 (2011) 065504. arXiv:1006.3244, doi:10.1103/PhysRevC.83.065504.

[26] C. Athanassopoulos, et al., Evidence for $\overline{\nu}_\mu \to \overline{\nu}_e$ oscillations from the LSND experiment at LAMPF, Phys. Rev. Lett. 77 (1996) 3082–3085. arXiv:nucl-ex/9605003, doi:10.1103/PhysRevLett.77.3082.

[27] A. A. Aguilar-Arevalo, et al., Significant Excess of Electronlike Events in the MiniBooNE Short-Baseline Neutrino Experiment, Phys. Rev. Lett. 121 (22) (2018) 221801. arXiv:1805.12028, doi:10.1103/PhysRevLett.121.221801.

[28] O. L. G. Peres, A. Yu. Smirnov, (3+1) spectrum of neutrino masses: A Chance for LSND?, Nucl. Phys. B599 (2001) 3. arXiv:hep-ph/0011054, doi:10.1016/S0550-3213(01)00012-8.

[29] M. Sorel, J. M. Conrad, M. Shaevitz, A Combined analysis of short baseline neutrino experiments in the (3+1) and (3+2) sterile neutrino oscillation hypotheses, Phys. Rev. D70 (2004) 073004. arXiv:hep-ph/0305255, doi:10.1103/PhysRevD.70.073004.

[30] J. Kopp, M. Maltoni, T. Schwetz, Are there sterile neutrinos at the eV scale?, Phys. Rev. Lett. 107 (2011) 091801. arXiv:1103.4570, doi:10.1103/PhysRevLett.107.091801.

[31] C. Giunti, M. Laveder, 3+1 and 3+2 Sterile Neutrino Fits, Phys. Rev. D84 (2011) 073008. arXiv:1107.1452, doi:10.1103/PhysRevD.84.073008.

[32] P. Adamson, et al., Search for sterile neutrinos in MINOS and MINOS+ using a two-detector fit, Phys. Rev. Lett. 122 (9) (2019) 091803. arXiv:1710.06488, doi:10.1103/PhysRevLett.122.091803.

[33] M. G. Aartsen, et al., Search for sterile neutrino mixing using three years of IceCube DeepCore data, Phys. Rev. D95 (11) (2017) 112002. arXiv:1702.05160, doi:10.1103/PhysRevD.95.112002.

[34] S. Antusch, C. Biggio, E. Fernandez-Martinez, M. B. Gavela, J. Lopez-Pavon, Unitarity of the Leptonic Mixing Matrix, JHEP 10 (2006) 084. arXiv:hep-ph/0607020, doi:10.1088/1126-6708/2006/10/084.

[35] E. Fernandez-Martinez, M. B. Gavela, J. Lopez-Pavon, O. Yasuda, CP-violation from non-unitary leptonic mixing, Phys. Lett. B649 (2007) 427–435. arXiv:hep-ph/0703098, doi:10.1016/j.physletb.2007.03.069.

[36] Z.-z. Xing, Correlation between the Charged Current Interactions of Light and Heavy Majorana Neutrinos, Phys. Lett. B660 (2008) 515–521. arXiv:0709.2220, doi:10.1016/j.physletb.2008.01.038.

[37] Z.-z. Xing, A full parametrization of the 6×6 flavor mixing matrix in the presence of three light or heavy sterile neutrinos, Phys. Rev. D85 (2012) 013008. arXiv:1110.0083, doi:10.1103/PhysRevD.85.013008.

[38] F. J. Escrihuela, D. V. Forero, O. G. Miranda, M. Tortola, J. W. F. Valle, On the description of nonunitary neutrino mixing, Phys. Rev. D92 (5) (2015) 053009, [Erratum: Phys. Rev.D93,no.11,119905(2016)]. arXiv:1503.08879, doi:10.1103/PhysRevD.93.119905,10.1103/PhysRevD.92.053009.

[39] M. Blennow, P. Coloma, E. Fernandez-Martinez, J. Hernandez-Garcia, J. Lopez-Pavon, Non-Unitarity, sterile neutrinos, and Non-Standard neutrino Interactions, JHEP 04 (2017) 153. arXiv:1609.08637, doi:10.1007/JHEP04(2017)153.

[40] S. F. King, C. Luhn, Neutrino Mass and Mixing with Discrete Symmetry, Rept. Prog. Phys. 76 (2013) 056201. arXiv:1301.1340, doi:10.1088/0034-4885/76/5/056201.

[41] K. Bielas, W. Flieger, J. Gluza, M. Gluza, Neutrino mixing, interval matrices and singular values, Phys. Rev. D98 (5) (2018) 053001. arXiv:1708.09196, doi:10.1103/PhysRevD.98.053001.

[42] J. Chadwick, Possible Existence of a Neutron, Nature 129 (1932) 312. doi:10.1038/129312a0.

[43] E. Fermi, An attempt of a theory of beta radiation. 1., Z. Phys. 88 (1934) 161–177. doi:10.1007/BF01351864.

[44] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, R. P. Hudson, Experimental test of parity conservation in beta decay, Phys. Rev. 105 (1957) 1413–1415. doi:10.1103/PhysRev.105.1413.

[45] L. D. Landau, On the conservation laws for weak interactions, Nucl. Phys. 3 (1957) 127–131. doi:10.1016/0029-5582(57)90061-5.

[46] T. D. Lee, C.-N. Yang, Parity Nonconservation and a Two Component Theory of the Neutrino, Phys. Rev. 105 (1957) 1671–1675. doi:10.1103/PhysRev.105.1671.

[47] A. Salam, On parity conservation and neutrino mass, Nuovo Cim. 5 (1957) 299–301. doi:10.1007/BF02812841.

[48] M. Goldhaber, L. Grodzins, A. W. Sunyar, Helicity of Neutrinos, Phys. Rev. 109 (1958) 1015–1017. doi:10.1103/PhysRev.109.1015.

[49] B. Pontecorvo, Inverse beta processes and nonconservation of lepton charge, Sov. Phys. JETP 7 (1958) 172–173, [Zh. Eksp. Teor. Fiz.34,247(1957)].

[50] Z. Maki, M. Nakagawa, S. Sakata, Remarks on the unified model of elementary particles, Prog. Theor. Phys. 28 (1962) 870–880, [34(1962)]. doi:10.1143/PTP.28.870.

[51] E. Akhmedov, Majorana neutrinos and other Majorana particles:Theory and experiment, 2014. arXiv:1412.3320.

[52] R. Allahverdi, B. Dutta, R. N. Mohapatra, Schizophrenic Neutrinos and ν-less Double Beta Decay, Phys. Lett. B695 (2011) 181–184. arXiv:1008.1232, doi:10.1016/j.physletb.2010.11.006.

[53] P. S. B. Dev, R. N. Mohapatra, Probing TeV Left-Right Seesaw at Energy and Intensity Frontiers: a Snowmass White Paper, in: Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013, 2013. arXiv:1308.2151.

[54] J. W. F. Valle, C. A. Vaquera-Araujo, Dynamical seesaw mechanism for Dirac neutrinos, Phys. Lett. B755 (2016) 363–366. arXiv:1601.05237, doi:10.1016/j.physletb.2016.02.031.

[55] R. N. Mohapatra, Mechanism for Understanding Small Neutrino Mass in Superstring Theories, Phys. Rev. Lett. 56 (1986) 561–563. doi:10.1103/PhysRevLett.56.561.

[56] R. N. Mohapatra, J. W. F. Valle, Neutrino Mass and Baryon Number Nonconservation in Superstring Models, Phys. Rev. D34 (1986) 1642. doi:10.1103/PhysRevD.34.1642.

[57] E. K. Akhmedov, M. Lindner, E. Schnapka, J. W. F. Valle, Dynamical left-right symmetry breaking, Phys. Rev. D53 (1996) 2752–2780. arXiv:hep-ph/9509255, doi:10.1103/PhysRevD.53.2752.

[58] S. M. Barr, A Different seesaw formula for neutrino masses, Phys. Rev. Lett. 92 (2004) 101601. arXiv:hep-ph/0309152, doi:10.1103/PhysRevLett.92.101601.

[59] P. Zyla, et al., Review of Particle Physics, PTEP 2020 (8) (2020) 083C01. doi:10.1093/ptep/ptaa104.

[60] S. Weinberg, Baryon and Lepton Nonconserving Processes, Phys. Rev. Lett. 43 (1979) 1566–1570. doi:10.1103/PhysRevLett.43.1566.

[61] C. Giunti, T. Lasserre, eV-scale Sterile Neutrinos, Ann. Rev. Nucl. Part. Sci. 69 (2019) 163–190. arXiv:1901.08330, doi:10.1146/annurev-nucl-101918-023755.

[62] M. Magg, C. Wetterich, Neutrino Mass Problem and Gauge Hierarchy, Phys. Lett. B94 (1980) 61–64. doi:10.1016/0370-2693(80)90825-4.

[63] R. Foot, H. Lew, X. G. He, G. C. Joshi, Seesaw Neutrino Masses Induced by a Triplet of Leptons, Z. Phys. C44 (1989) 441. doi:10.1007/BF01415558.

[64] E. T. Franco, Type I+III seesaw mechanism and CP violation for leptogenesis, Phys. Rev. D92 (11) (2015) 113010. arXiv:1510.06240, doi:10.1103/PhysRevD.92.113010.

[65] J. Gluza, M. Zralek, CP violation in the heavy neutrinos production process $e^+e^- \to N(1)N(2)$, Phys. Rev. D 51 (1995) 4707–4711. arXiv:hep-ph/9409224, doi:10.1103/PhysRevD.51.4707.

[66] J. Gluza, M. Zralek, Heavy neutrinos production and decay in future e^+e^- colliders, Phys. Rev. D 55 (1997) 7030–7037. arXiv:hep-ph/9612227, doi:10.1103/PhysRevD.55.7030.

[67] D. J. Wagner, T. J. Weiler, Boxing with neutrino oscillations, Phys. Rev. D 59 (1999) 113007. arXiv:hep-ph/9801327, doi:10.1103/PhysRevD.59.113007.

[68] C. Jarlskog, Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Violation, Phys. Rev. Lett. 55 (1985) 1039. doi:10.1103/PhysRevLett.55.1039.

[69] D.-d. Wu, The Rephasing Invariants and CP, Phys. Rev. D 33 (1986) 860. doi:10.1103/PhysRevD.33.860.

[70] I. Dunietz, O. W. Greenberg, D.-d. Wu, A Priori Definition of Maximal CP Violation, Phys. Rev. Lett. 55 (1985) 2935. doi:10.1103/PhysRevLett.55.2935.

[71] C. Giunti, Neutrino wave packets in quantum field theory, JHEP 11 (2002) 017. arXiv:hep-ph/0205014, doi:10.1088/1126-6708/2002/11/017.

[72] M. Fukugita, T. Yanagida, Physics of neutrinos and applications to astrophysics, Springer, , 2003.

[73] M. Ochman, R. Szafron, M. Zralek, Neutrino production states in oscillation phenomena: Are they pure or mixed?, J. Phys. G 35 (2008) 065003. arXiv:0707.4089, doi:10.1088/0954-3899/35/6/065003.

[74] D. V. Naumov, V. A. Naumov, Quantum Field Theory of Neutrino Oscillations, Phys. Part. Nucl. 51 (1) (2020) 1–106. doi:10.1134/S1063779620010050.

[75] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, A. Zhou, The fate of hints: updated global analysis of three-flavor neutrino oscillations, JHEP 09 (2020) 178. arXiv:2007.14792, doi:10.1007/JHEP09(2020)178.

[76] G. Stewart, J. guang Sun, Matrix Perturbation Theory, Computer science and scientific computing, Academic Press, 1990.

[77] R. A. Horn, C. R. Johnson, Matrix Analysis, 2nd Edition, Cambridge University Press, 2012. doi:10.1017/9781139020411.

[78] A. Allen, D. Arceo, Corporate Author: Space and Naval Warfare Systems Center San Diego CA, Matrix Dilations via Cosine-Sine Decomposition, Defense Technical Information Center. https://apps.dtic.mil/dtic/tr/fulltext/u2/a446226.pdf.

[79] P. Grover, P. Mishra, Perturbation bounds for mostow's decomposition and the bipolar decomposition, Linear Algebra and its Applications 529 (2017) 12–29. doi:https://doi.org/10.1016/j.laa.2017.03.027.

[80] F. J. Escrihuela, D. V. Forero, O. G. Miranda, M. Tórtola, J. W. F. Valle, Probing CP violation with non-unitary mixing in long-baseline neutrino oscillation experiments: DUNE as a case study, New J. Phys. 19 (9) (2017) 093005. arXiv:1612.07377, doi:10.1088/1367-2630/aa79ec.

[81] S. Antusch, M. Blennow, E. Fernandez-Martinez, J. Lopez-Pavon, Probing non-unitary mixing and CP-violation at a Neutrino Factory, Phys. Rev. D80 (2009) 033002. arXiv:0903.3986, doi:10.1103/PhysRevD.80.033002.

[82] M. Blennow, E. Fernandez-Martinez, Parametrization of Seesaw Models and Light Sterile Neutrinos, Phys. Lett. B704 (2011) 223–229. arXiv:1107.3992, doi:10.1016/j.physletb.2011.09.028.

[83] N. R. Agostinho, G. C. Branco, P. M. F. Pereira, M. N. Rebelo, J. I. Silva-Marcos, Can One have Significant Deviations from Leptonic 3×3 Unitarity in the Framework of Type I Seesaw Mechanism? arXiv:1711.06229.

[84] M. Czakon, J. Gluza, M. Zrałek, Nonunitary neutrino mixing matrix and CP violating neutrino oscillations, Acta Phys. Polon. B32 (2001) 3735–3744. arXiv:hep-ph/0109245.

[85] H. Weyl, Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung), Mathematische Annalen 71 (4) (1912) 441–479. doi:10.1007/BF01456804.

[86] C. R. Rao, M. B. Rao, Matrix Algebra and Its Applications to Statistics and Econometrics, World Scientific, 1998. doi:10.1142/3599.

[87] F. Zhang, Matrix Theory: Basic Results and Techniques, 2nd Edition, Springer, 2011. doi:10.1007/978-1-4614-1099-7.

[88] J. von Neumann, Some Matrix-Inequalities and Metrization of Matrix-Space, Tomsk. Univ. Rev. Vol. 1 (1937) 286–300, In: A. H. Taub, Ed., John von Neumann Collected Works, Vol. IV, Pergamon, Oxford, 1962, pp. 205-218.

[89] K. Ziętak, On the characterization of the extremal points of the unit sphere of matrices, Linear Algebra and its Applications 106 (1988) 57 – 75. doi:10.1016/0024-3795(88)90023-7.

[90] W. So, Facial structures of schatten p-norms, Linear and Multilinear Algebra 27 (3) (1990) 207–212. doi:10.1080/03081089008818012.

[91] E. M. de Sá, Faces and traces of the unit ball of a symmetric gauge function, Linear Algebra and its Applications 197-198 (1994) 349 – 395. doi:10.1016/0024-3795(94)90496-0.

[92] E. M. de Sá, Exposed faces and duality for symmetric and unitarily invariant norms, Linear Algebra and its Applications 197-198 (1994) 429 – 450. doi:10.1016/0024-3795(94)90499-5.

[93] E. M. de Sá, Faces of the unit ball of a unitarily invariant norm, Linear Algebra and its Applications 197-198 (1994) 451–493. doi:10.1016/0024-3795(94)90500-2.

[94] R. T. Rockafellar, Convex Analysis, Princeton University Press, 1970.

[95] J.-B. Hiriart-Urruty, C. Lemaréchal, Fundamentals of Convex Analysis, 2001, pp. 121–162. doi:10.1007/978-3-642-56468-0_4.

[96] J. Stoer, On the characterization of least upper bound norms in matrix space, Numer. Math. 6 (1) (1964) 302–314. doi:10.1007/BF01386078.

[97] I. E. Leonard, J. E. Lewis, Geometry of Convex Sets, Wiley, 2015.

[98] M. Krein, D. Milman, On extreme points of regular convex sets, Studia Mathematica 9 (1) (1940) 133–138.

[99] C. Carathéodory, Über den variabilitätsbereich der fourier'schen konstanten von positiven harmonischen funktionen, Rendiconti del Circolo Matematico di Palermo (1884-1940) 32 193–217.

[100] L. Tu, An Introduction to Manifolds, Universitext, Springer New York, 2010.

[101] R. J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley-Interscience, 2005.

[102] M. S. Marinov, Invariant volumes of compact groups, Journal of Physics A: Mathematical and General 13 (11) (1980) 3357–3366. doi:10.1088/0305-4470/13/11/009.

[103] M. S. Marinov, Correction to 'invariant volumes of compact groups', Journal of Physics A: Mathematical and General 14 (2) (1981) 543–544. doi:10.1088/0305-4470/14/2/030.

[104] L. J. Boya, E. Sudarshan, T. Tilma, Volumes of compact manifolds, Reports on Mathematical Physics 52 (3) (2003) 401–422. doi:10.1016/s0034-4877(03)80038-1.

[105] L. Zhang, Volumes of orthogonal groups and unitary groups (2017). arXiv:1509.00537.

[106] F. W. Ponting, H. S. A. Potter, The volume of orthogonal and unitary space, The Quarterly Journal of Mathematics os-20 (1) (1949) 146–154. doi:10.1093/qmath/os-20.1.146.

[107] S. Kobayashi, K. Nomizu, Foundations of Differential Geometry Vol. I, Wiley-Interscience, 1996.

[108] R. L. Mkrtchyan, A. P. Veselov, A universal formula for the volume of compact lie groups (2013). arXiv:1304.3031.

[109] K. Zyczkowski, H.-J. Sommers, Hilbert–schmidt volume of the set of mixed quantum states, Journal of Physics A: Mathematical and General 36 (39) (2003) 10115–10130. doi:10.1088/0305-4470/36/39/310.

[110] R. A. Horn, C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991. doi:10.1017/CBO9780511840371.

[111] C. Kwong-Li, R. Mathias, Construction of Matrices with Prescribed Singular Values and Eigenvalues, BIT Numerical Mathematics 41 (2001) 115. doi:10.1023/A:1021969818438.

[112] E. Fernandez-Martinez, J. Hernandez-Garcia, J. Lopez-Pavon, Global constraints on heavy neutrino mixing, JHEP 08 (2016) 033. arXiv:1605.08774, doi:10.1007/JHEP08(2016)033.

[113] Y. Declais, et al., Search for neutrino oscillations at 15-meters, 40-meters, and 95-meters from a nuclear power reactor at Bugey, Nucl. Phys. B434 (1995) 503–534. doi:10.1016/0550-3213(94)00513-E.

[114] K. Abe, et al., Limits on sterile neutrino mixing using atmospheric neutrinos in Super-Kamiokande, Phys. Rev. D91 (2015) 052019. arXiv:1410.2008, doi:10.1103/PhysRevD.91.052019.

[115] P. Adamson, et al., Search for Sterile Neutrinos Mixing with Muon Neutrinos in MINOS, Phys. Rev. Lett. 117 (15) (2016) 151803. arXiv:1607.01176, doi:10.1103/PhysRevLett.117.151803.

[116] P. Astier, et al., Final NOMAD results on $\nu_\mu \to \nu_\tau$ and $\nu_e \to \nu_\tau$ oscillations including a new search for ν_τ appearance using hadronic τ decays, Nucl. Phys. B611 (2001) 3–39. arXiv:hep-ex/0106102, doi:10.1016/S0550-3213(01)00339-X.

[117] P. Astier, et al., Search for $\nu_\mu \to \nu_e$ oscillations in the NOMAD experiment, Phys. Lett. B570 (2003) 19–31. arXiv:hep-ex/0306037, doi:10.1016/j.physletb.2003.07.029.

[118] J. de Blas, Electroweak limits on physics beyond the Standard Model, EPJ Web Conf. 60 (2013) 19008. arXiv:1307.6173, doi:10.1051/epjconf/20136019008.

[119] S. Gariazzo, C. Giunti, M. Laveder, Y. F. Li, Updated Global 3+1 Analysis of Short-BaseLine Neutrino Oscillations, JHEP 06 (2017) 135. arXiv:1703.00860, doi:10.1007/JHEP06(2017)135.

[120] M. Dentler, A. Hernández-Cabezudo, J. Kopp, P. A. N. Machado, M. Maltoni, I. Martinez-Soler, T. Schwetz, Updated Global Analysis of Neutrino Oscillations in the Presence of eV-Scale Sterile Neutrinos, JHEP 08 (2018) 010. arXiv:1803.10661, doi:10.1007/JHEP08(2018)010.

[121] N. Cabibbo, Unitary Symmetry and Leptonic Decays, Phys. Rev. Lett. 10 (1963) 531–533. doi:10.1103/PhysRevLett.10.531.

[122] M. Kobayashi, T. Maskawa, CP Violation in the Renormalizable Theory of Weak Interaction, Prog. Theor. Phys. 49 (1973) 652–657. doi:10.1143/PTP.49.652.

[123] L. Wolfenstein, Parametrization of the Kobayashi-Maskawa Matrix, Phys. Rev. Lett. 51 (1983) 1945. doi:10.1103/PhysRevLett.51.1945.

[124] A. J. Buras, M. E. Lautenbacher, G. Ostermaier, Waiting for the top quark mass, K+ → pi+ neutrino anti-neutrino, B(s)0 - anti-B(s)0 mixing and CP asymmetries in B decays, Phys. Rev. D 50 (1994) 3433–3446. arXiv:hep-ph/9403384, doi:10.1103/PhysRevD.50.3433.

[125] R. Bhatia, Matrix Analysis, Graduate Texts in Mathematics, Springer New York, 1996.

[126] C. Davis, W. Kahan, The Rotation of Eigenvectors by a Perturbation. III, Siam Journal on Numerical Analysis 7 (1970) 1–46. doi:10.1137/0707001.

[127] W. Flieger, J. Gluza, General neutrino mass spectrum and mixing properties in seesaw mechanisms, Chin. Phys. C 45 (2) (2021) 023106. arXiv:2004.00354, doi:10.1088/1674-1137/abcd2f.

[128] P. B. Denton, S. J. Parke, X. Zhang, Eigenvalues: the Rosetta Stone for Neutrino Oscillations in Matter. arXiv:1907.02534.

[129] P. B. Denton, S. J. Parke, T. Tao, X. Zhang, Eigenvectors from Eigenvalues: a survey of a basic identity in linear algebra. arXiv:1908.03795.