CERN-TH-2024-XXX
DESY-24-XXX
HU-EP-24/XX

The flavor invariants of the νSM

Christophe Grojean,[a,b,c] **Jonathan Kley,**[a,b] **Damien Leflot,**[a,d] **Chang-Yuan Yao**[a,e]

[a] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*
[b] *Institut für Physik, Humboldt-Universität zu Berlin, 12489 Berlin, Germany*
[c] *Theoretical Theoretical Physics Department, CERN, 1211 Geneva 23, Switzerland*
[d] *École Normale Supérieure de Lyon, Lyon Cedex 7, France*
[e] *School of Physics, Nankai University, Tianjin 300071, China*

E-mail: christophe.grojean@desy.de, jonathan.kley@desy.de, damien.leflot@ens-lyon.fr, chang.yuan.yao@desy.de

Abstract: The νSM, the Standard Model (SM) extended with 3 generation of sterile neutrinos, is a supercalifragilisticexpialidocious explanation for the neutrino masses that cannot be explained by the SM. In this paper, we systematically study the algebraic structure of the flavor invariants constructed from the flavorful parameters of the νSM. We calculate the Hilbert series and plethystic logarithm of the theory. Together with a graph-based method for the construction of the invariants, this allows us to construct a generating and primary set of invariants for the νSM for the first time. The generating set contains 459 invariant, out of which 208 are CP-even and 251 are CP-odd. Furthermore, we discuss how the sources of CP violation, as well as the necessary and sufficient condition for CP violation can be captured with the use of these invariants. Along the way, we present useful algorithms, which allowed us to obtain the results in this paper.

Contents

1 Introduction

The smallness of the observed neutrino masses (needed to explain neutrino oscillation experiments [1, 2]) is yet another consequence/success of the chiral nature of the Standard Model (SM) gauge structure. In the absence, yet to be confirmed, of right-handed neutrino at *low* energy, these neutrino masses cannot emerge from marginal interactions among the SM degrees of freedom, like in the case for quarks and charged leptons, but require irrelevant interactions [3] possibly stemming from interactions with new physics degrees of freedom or new strong interactions. A particularly compelling scenario is the type-I seesaw mechanism with right-handed sterile neutrinos endowed with a *large* SM gauge-invariant Majorana mass and coupled via Yukawa interactions to the left-handed active neutrinos [4–8]. The value of the scale of this Majorana mass, which also measures the breaking of an accidental symmetry of the SM - the conservation of the total number of leptons, remains unknown and largely unconstrained.

In this paper, we will focus on the renormalizable theory of the SM extended with $n_N = 3$ flavors of right-handed sterile neutrinos N, often referred to as νSM or νMSM [9–13]. Depending on the mass we assign to these new particles our discussion will capture both a type-I seesaw scenario and a scenario with light right-handed sterile neutrinos. Just like in the quark sector of the SM, adding more parameters to the theory by introducing a mass term opens the possibility to violate CP through physical phases that cannot be removed from the theory using field redefinitions. These new sources of CP violation (CPV) could help generating the matter-antimatter asymmetry via a leptogenesis mechanism [14]. In the case of neutrinos, the option of having a Majorana mass allows for another peculiarity. Beyond the Dirac-like phase in the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix – in analogy to the CKM matrix in the quark sector of the SM – there can be additional Majorana-like phases which are unphysical in the case of a Dirac mass due to the conservation of lepton number. However, the parameters describing CPV in the neutrino sector have not been measured reliably to date and it is unclear if it is feasible to achieve sufficient experimental sensitivity to detect Majorana-like phases in the near future [15].

Therefore, it is an important task to classify all possible scenarios for the generation of neutrino masses from an IR perspective. In this paper, we want to tackle this task using the language of flavor invariants and in addition also classify new sources of CPV that are introduced through the neutrino masses with Jarlskog-like flavor invariants [16–18]. In this analysis we will make use of the fact that – as physical quantities – all observables should in principle be expressible in terms of finite number of flavor invariant objects. Such a set of invariants is referred to as the *generating* set or the *basic* set in the literature. This set enables us to express all observables as functions of them, which will allow us to learn about the structure of CP violation in the theory. Furthermore, it is straightforward to differentiate between Dirac and Majorana-like mass terms for the light sterile neutrinos in our formulation by taking the appropriate limits.

The generating set of flavor invariants is the maximal set capturing all physical effects of the theory. In terms of CPV effects, we will also try to address the questions of what is the minimal set of CP-odd invariants that are necessary to capture all sources of CPV,

and what is the minimal set that can capture the necessary and sufficient conditions of CP conservation (CPC). Taking all texture zeros and degenerate and vanishing masses into account, these sets are required to be valid in all special spectra in the theory.

In order to simplify the process of finding a generating set of flavor invariants, we will make use of tools from invariant theory that have been developed in recent years in the context of theoretical particle physics. Invariants have been used both in UV complete theories [19–33] and EFTs [34–37] to characterize the parameters of the theory with respect to CP and other symmetries of EFTs [38]. Some problems which are closely related to our analysis in this paper have been previously investigated in the literature. In particular, the νSM with only two generations for all fermions has been studied in Ref. [26] and the case of adding two generations of right-handed sterile neutrinos to the SM has been treated in Ref. [30]. There, the authors also use the flavor invariants to formulate the necessary and sufficient conditions for CPV in the model. As we will see throughout this paper, generalizing the discussion from two to three generations is not an easy task. The more complex flavor group structure leads to a significantly more complicated algebraic ring structure of the invariants in the set. We will find that the number of generating invariants in the theory explodes from 38 in Ref. [30] for two generations of sterile neutrinos to 459 for three generations of sterile neutrinos.

The paper is structured as follows. In Sec. 2, we will define more precisely the setup for neutrino masses that we consider in this paper and introduce all tools from invariant theory we will need to build a generating set of flavor invariants. In Sec. 3, we will present our results for 3 generations of right-handed neutrinos including the Hilbert series and the generating set. In Sec. 4, we distinguish between the minimal set of CP-odd invariants used to parameterize the sources of CPV and the set used to capture the CPC conditions. We emphasize that these sets have the property of being applicable to all special spectra of the theory that give rise to larger exact flavor symmetries. We show the results for the theory with 2 generations of sterile neutrinos. The case of 3 generations turns out to be too hard to solve, but we nonetheless discuss an approach towards finding such a set. Finally, in Sec. 5, we draw some conclusions. In App. A, we present useful parameterizations for the flavorful couplings that we use throughout this paper. In App. D, we list all invariants in the generating set along with their CP parities. In App. B, we show our results for the graded Hilbert series and the graded plethystic logarithm that we do not show in the main text due to their length. In App. C, we present the algorithms used in this paper. These algorithms cover several purposes: generating flavor invariants, reducing the invariants to a generating set, and finding the minimal set of CP-odd invariants to capture all sources of CPV. In App. F, we introduce the Hironaka decomposition that we use in some parts of the paper. In App. G, we introduce Hilbert's Nullstellensatz and present the relevant concepts that are useful for obtaining a minimal set of CP-odd invariants that capture the necessary and sufficient conditions for CPC. Finally, in App. H, we present the complete list of CPC conditions obtained through the ideal-related method.

	$SU(3)_L \times U(1)_L$	$SU(3)_e \times U(1)_e$	$SU(3)_N \times U(1)_N$
Y_e	$\mathbf{3}_{+1}$	$\bar{\mathbf{3}}_{-1}$	$\mathbf{1}_0$
Y_N	$\mathbf{3}_{+1}$	$\mathbf{1}_0$	$\bar{\mathbf{3}}_{-1}$
M_N	$\mathbf{1}_0$	$\mathbf{1}_0$	$(\bar{\mathbf{3}} \otimes_s \bar{\mathbf{3}})_{-2}$

Table 1: The flavor transformation properties of the relevant Yukawa matrices and Majorana mass matrix treated as spurions. The subscripts of the $SU(3)$ representations denote the charge under the $U(1)$ part of the flavor symmetry group. Furthermore, $\otimes_s$ denotes the symmetric tensor product of the simple representations.

2 Prerequisites

2.1 SM extended with right-handed neutrinos

Before we start with the analysis, we have to precisely define the theory we will work with. As already mentioned in the introduction, there are several ways to introduce a mass terms for neutrinos at low energies. In this paper, we will extend the SM particle spectrum by adding 3 generations of right-handed sterile neutrinos N, which in the literature usually goes under the name of νSM.[1] The most general renormalizable Lagrangian that can be built with these fields and the symmetries of the SM gauge group is

$$\begin{aligned}\mathcal{L}_{\nu\text{SM}} = & \sum_{\psi} \bar{\psi} i \not{D} \psi - \left[\frac{1}{2} \left(N C M_N N \right) + \bar{L} Y_N N \widetilde{H} + \bar{L} Y_e e H + \bar{Q} Y_u u \widetilde{H} + \bar{Q} Y_d d H + \text{H.c.} \right] \\ & - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} - \frac{1}{4} W^I_{\mu\nu} W^{I\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + (D_\mu H)^\dagger (D^\mu H) - \lambda \left(H^\dagger H - \frac{v^2}{2} \right)^2 , \end{aligned} \tag{2.1}$$

where ψ represents all fermion fields $\{Q, u, d, L, e, N\}$, H is the Higgs doublet with the vacuum expectation value v, $\widetilde{H}_i = \epsilon_{ij} H^*_j$, and $G^a_{\mu\nu}$, $W^I_{\mu\nu}$, $B_{\mu\nu}$ are the field strengths of the SM gauge fields. The right-handed singlet neutrino N is introduced to extend the SM. If lepton number is not a symmetry, a Majorana mass for the N fields is allowed, which corresponds to a symmetric matrix M_N.[2] The 3×3 $Y_{u,d,e,N}$ are the Yukawa couplings with the Higgs field. C is the charge-conjugation matrix, and D_μ is the gauge covariant derivative.

One can easily check that the fermion kinetic term in $\mathcal{L}_{\nu\text{SM}}$ is invariant under unitary $U(3)$ flavor transformations of the fermion fields. Assuming that this flavor symmetry is only softly broken by the Yukawa couplings and Majorana mass term, we promote all flavorful couplings to spurions under this symmetry, making the Lagrangian formally invariant. The corresponding transformation properties of the spurions under the non-Abelian part of the flavor group can be found in Tab. 1.

One immediate consequence of this assignment is that the presence of M_N breaks lepton number, as the transformation properties of M_N do not allow for rephasings of N.

[1] In principle 2 generations of sterile neutrinos are enough to generate the observed neutrino masses at low energies [39]. As mentioned before this case has been treated in Ref. [30].

[2] If there is no lepton number violation, the Majorana mass term is forbidden, and the neutrino will have a Dirac mass. Then, the flavor structure of the lepton sector will have exactly the same form as the quark sector in the SM.

Therefore, there will be additional Majorana type phases in the spectrum of the theory. These phases can be removed in the SM because all interactions enjoy a symmetry under respectively rephasing the lepton and quark fields. Furthermore, lepton family number, which is conserved in the SM, is broken due to the PMNS matrix which encodes the fact that in general Y_e and Y_N cannot be diagonalized simultaneously in analogy to the quark sector.

Later on in the paper it will prove useful to have an explicit parametrization for the flavorful matrices. We will use the following parametrization

$$Y_e = \text{diag}\,(y_e, y_\mu, y_\tau)\,, \quad Y_N = V \cdot \text{diag}\,(y_1, y_2, y_3) \cdot W^\dagger, \quad M_N = \text{diag}\,(m_1, m_2, m_3)\,, \tag{2.2}$$

where

$$V = U(\theta_{12}, \theta_{13}, \theta_{23}, \delta) \cdot \text{diag}\left(1, e^{i\phi_1}, e^{i\phi_2}\right), \; W = \text{diag}\left(1, e^{i\phi_1'}, e^{i\phi_2'}\right) \cdot U(\theta_{12}', \theta_{13}', \theta_{23}', \delta')\,, \tag{2.3}$$

and $U(\theta_{12}, \theta_{13}, \theta_{23}, \delta)$ has been defined in Eq. (A.9), which is a CKM-like matrix with a Dirac phase $\delta \in [0, 2\pi)$ and three mixing angles $\theta_{ij} \in [0, \pi/2]$. $\phi_{1,2} \in [0, 2\pi)$ and $\phi_{1,2}' \in [0, \pi)$ are additional Majorana phases. This parametrization correctly captures the 9 mass parameters, 6 mixing angles and 2 Dirac as well as 4 Majorana phases of the theory. This paper adopts a different convention, distinct from that which was previously discussed in Ref. [26]. As a result, we provide the details focus on the new convention in App. A. In addition, we introduce new parameterizations that can be more conveniently used for the study of algebraic structures of the polynomial rings.

2.2 The Hilbert series and plethystic logarithm

In this section we will briefly review useful tools from invariant theory, developed in Refs. [40–43] for operator bases and flavor invariants, that simplify the building process and characterization of the low-energy flavor invariants. We will mostly follow the notation introduced in Refs. [29, 42] here. Our goal in this paper is to find a minimal set of flavor invariants that allows us to parametrize all observables in the theory in terms of those invariants.

The central object of this paper will be flavor invariants $\mathcal{I}$ that are combinations of Lagrangian parameters of the theory invariant under the maximal possible flavor group of the renormalizable Lagrangian to be defined below. In a first step, we want to find a set of invariants that allows us to express all remaining invariants in the theory as a polynomial of the invariants in the set. As mentioned above, this set of invariants is called the set of *generating* or *basic* invariants and always has finite cardinality for reductive groups [44, 45].

The basic set is a set of invariants, in which no invariant can be expressed as a polynomial of all other invariants in the basic set. If any invariant can be expressed as a polynomial of other invariants, it will be termed as a linear relation in this paper. But the invariants in the basic set may still be *algebraically dependent*, i.e., there exist relations between them in the form of

$$P\,(\mathcal{I}_1, ..., \mathcal{I}_m) = 0\,, \tag{2.4}$$

which are called *syzygies* in the invariant literature.

Among the basic set of invariants there exists a set of invariants which are furthermore *algebraically independent*, they are the so-called *primary* invariants.[3] The fact that they are algebraically independent implies that there exists no syzygy only comprising of invariants from the primary set. Another interesting result is that the number of physical parameters, i.e., the minimal number of parameters that are left after all transformations allowed by the symmetry group of the theory are used, is equal to the number of invariants in the primary set [26, 44, 45].

A useful guide to construct those invariants is the so-called *Hilbert series* (HS)

$$\mathcal{H}(q) = \sum_{i=0}^{\infty} c_i q^i \tag{2.5}$$

which enumerates the number c_i of all possible invariants that can be built from the given set of building blocks labeled by the parameter q at a given order i.

It can be shown that the HS can always be written as a fraction of two polynomials [45]

$$\mathcal{H}(q) = \frac{\mathcal{N}(q)}{\mathcal{D}(q)}\,, \tag{2.6}$$

where the numerator is of palindromic form, i.e. $\mathcal{N}(q) = t^p \mathcal{N}(1/q)$ with p is the highest power of q in $\mathcal{N}(q)$ and all terms in $\mathcal{N}(q)$ come with a positive sign. The denominator is of the form $\mathcal{D}(q) = \prod_{i=1}^{m}(1-q^{d_i})$, where the total number of factors m counts the number of primary invariants corresponding to the physical parameters in the theory, while the exponents d_i in each factor give the power of the spurion in the invariant. If the numerator is trivial, i.e. $\mathcal{N}(q) = 1$ and the complete generating set is given by the set of primary invariants, the ring is called a *free ring*.

Of course, most theories contain more than one coupling and it can be convenient to count each coupling with its own spurion to simplify the identification of the invariants in the HS. For n independent couplings in the theory that are used to build invariants, one defines the *multi-graded* HS

$$\mathcal{H}(q_1,\ldots,q_n) = \sum_{i_1=0}^{\infty}\cdots\sum_{i_n=0}^{\infty} c_{i_1\ldots i_n} q_1^{i_1}\ldots q_n^{i_n}\,, \tag{2.7}$$

where the coefficient $c_{i_1\ldots i_n}$ now count the number of invariants containing the spurions $(q_1,\ldots,q_n)$ to the power $(i_1,\ldots,i_n)$. We call these powers the *degrees* of the invariant while we call the sum of the degrees the *order* of the invariant.

Note that the multi-graded HS is no longer guaranteed to come in the form of Eq. (2.6), which has a palindromic property in numerator with positive terms and the denominator counting the number of primary invariants. To still obtain this information, one can always

[3] Note that there is a difference between the math and physics invariant literature on what is called a primary invariant. In the physics literature a set of algebraically independent invariants is often called a primary set. In the math literature on the other hand, the primary set is defined by a unique decomposition, the so-called Hironaka decomposition (see App. F for more details), which automatically implies that the invariants are algebraically independent. Since we do not care about this unique decomposition, we will follow the definition that is commonly used in physics but will also make an effort to find a primary set which is consistent with the Hironaka decomposition.

take the single-graded limit of the HS, $\mathcal{H}(q_1,\ldots,q_n) = \mathcal{H}(q,\ldots,q)$, where all couplings in the theory are counted with the same spurion.

After describing the properties of the HS, we now introduce the mathematical methods for calculating it. One convenient way to calculate the HS for reductive Lie groups is the so-called Molien-Weyl formula, which for a single coupling transforming in the representation R of the group G is defined as

$$\mathcal{H}(q) = \int d\mu_G \ \exp\left(\sum_{k=1}^{\infty} \frac{q^k \chi_R\left(z_1^k,\ldots,z_d^k\right)}{k}\right) \equiv \int d\mu_G \mathrm{PE}\left[\chi_R(z_1,\ldots,z_d);q\right] , \tag{2.8}$$

where $d\mu_G$ is the Haar measure of the group, $\chi_R(z_1,\ldots,z_d)$ is the character of the representation R of the group G of rank d, and we have defined the *plethystic exponential* (PE) in the last step. There is a straightforward generalization of the Molien-Weyl formula for a multi-graded HS in a theory with several couplings transforming in different representations R_i

$$\mathcal{H}(q_1,\ldots,q_n) = \int d\mu_G \prod_{i=1}^{n} \mathrm{PE}\left[\chi_{R_i}(z_1,\ldots,z_d);q_i\right] . \tag{2.9}$$

To study some of the properties of the ungraded HS of a theory with several couplings, we can take the single-graded limit $q_i \to q$.

Another useful function is the so-called plethystic logarithm (PL) which is the inverse function of the PE that we just defined, i.e., $\mathrm{PE}^{-1}(f(x)) = \mathrm{PL}(f(x))$ and is defined as follows

$$\mathrm{PL}\left[f\left(x_1,\ldots,x_N\right)\right] = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \log\left[f\left(x_1^n,\ldots,x_N^n\right)\right] , \tag{2.10}$$

where $\mu(n)$ is the so-called Möbius function.[4] In most cases the PL proves extremely helpful because we can simply read off the number of basic invariants and syzygies from the coefficients of the spurions at a given order in the spurions.[5] In the case of a *complete intersection ring*,[6] where the PL is just a polynomial in the spurions, the positive terms can be identified with the basic invariants of the theory, and the negative terms correspond to the syzygies that exist among them [46, 47]. However, for *non-complete intersection rings*, the PL becomes a non-terminating series, and it has been noted in the literature [29, 46, 47]

[4] The Moebius function is defined as

$$\mu(n) = \begin{cases} 0 & \text{n has repeated prime factors} \\ 1 & n = 1 \\ (-1)^j & n \text{ is product of } j \text{ distinct prime numbers} \end{cases} . \tag{2.11}$$

[5] For instance, the PL of the quark sector of the SM can be easily calculated with the HS shown in Ref. [26], which is given by $\mathrm{PL}(t) = 2t^2 + 3t^4 + 4t^6 + t^8 + t^{12} - t^{24}$. From the positive terms, we can read off that there are two order 2 invariants, three order 4 invariants, four order 6 invariants, one order 8 invariant and one order 12 invariant, while the negative term shows there is a syzygy at order 24. Among these 11 basic invariants, 10 of them are algebraically independent, which map to the ten physical parameters of the quark sector, with an additional invariant capturing the sign of the CP phase.

[6] A ring is classified as a complete intersection if the difference between the number of generating invariants and the number of syzygies is equal to the Krull dimension (the Krull dimension being the maximal number of algebraically-independent invariants). Otherwise, it is categorized as a non-complete intersection [29, 45].

that the *leading positive* terms, i.e., all positive terms up to the first term with a negative sign in the PL, can be identified with the basic invariants, and the *leading negative*, i.e., the first negative terms that appear after the leading positive terms correspond to syzygies.[7]

We will see later for the νSM with 3 generations of sterile neutrinos, that this does not necessarily have to hold true and the interpretation of the positive and negative terms in the PL have to be slightly changed for more complicated invariant rings. This has also been pointed out in the literature [29]. As we will see in the next section, this is also the main difference between the SM quark sector and the νSM. While, the representations that the spurions of the SM quark sector transform in are still sufficiently simple to generate a complete intersection ring with a terminating PL, this is no longer true for the νSM. Here, the representation the Majorana mass M_N lives in complicates the ring structure significantly leading to a non-complete intersection ring for both two and three generations of sterile neutrinos.

3 Building an invariant basis for the νSM

3.1 Hilbert series of the νSM

Before we start building invariants, we will first compute the HS and PL to set our expectations for the *generating* and *primary* set of invariants. We use again the Molien–Weyl formula to calculate the Hilbert series for the νSM with the spurions Y_e, Y_N and M_N.[8] For that, we need the character of the fundamental and anti-fundamental representation and the Haar measure of $U(3)$ which are given by [42]

$$\begin{aligned}
\chi^{\mathbf{3}}_{U(3)} &= z_1 + z_2 + z_3\,, \\
\chi^{\bar{\mathbf{3}}}_{U(3)} &= z_1^{-1} + z_2^{-1} + z_3^{-1}\,, \\
d\mu_{U(3)} &= \frac{1}{6!}\left(\prod_{i=1}^{3}\frac{dz_i}{2\pi i z_i}\right)\left(-\frac{(z_2-z_1)^2\,(z_3-z_1)^2\,(z_3-z_2)^2}{z_1^2 z_2^2 z_3^2}\right).
\end{aligned} \tag{3.1}$$

From these, we can construct the characters for the representations of the flavorful Lagrangian parameters of the νSM following Tab. 1. For instance, the character for Y_N is given by

$$\chi_{Y_N} = \chi^{\mathbf{3}}_{U(3)_L}(z_1,z_2,z_3)\chi^{\bar{\mathbf{3}}}_{U(3)_N}(z_4,z_5,z_6) = (z_1+z_2+z_3)\left(z_4^{-1}+z_5^{-1}+z_6^{-1}\right). \tag{3.2}$$

The characters for all other spurions can be obtained in the same manner. Using the expression for the Molien–Weyl formula in Eq. (2.9) with the same grading for all spurions, we can calculate the ungraded HS. The calculation involves the integral over the six variables $z_1,\ldots,z_6$ over the contour $|z_i|=1$, which can be obtained by calculating the residues. The

[7] All other terms in the non-terminating PL of a non-complete intersection ring after the leading negative terms have – to our knowledge – no meaning for the construction of a generating set beyond the fact that they appear in a special form of the HS, the so-called Euler form [46]. In this form the HS can be written as $\mathcal{H}(q) = \prod_{n=1}^{\infty}(1-q^n)^{-b_n}$, where it can be shown that the b_n are exactly the coefficients in the PL $\mathrm{PL}[\mathcal{H}(q)] = \sum_{n=1}^{\infty} b_n q^n$.

[8] To get a dimensionless quantity we take M_N to be divided by the only other mass scale in the problem, the Higgs vev v. Only then, one can compare invariants with a different number of insertions of M_N.

same calculation has been presented before, please refer to Refs. [37, 47] for details. We find for the numerator of the ungraded HS

$$\begin{aligned}\mathcal{N}(t) \;=\; & 1 + t^4 + 5t^6 + 9t^8 + 22t^{10} + 61t^{12} + 126t^{14} + 273t^{16} + 552t^{18} + 1038t^{20} \\ & + 1880t^{22} + 3293t^{24} + 5441t^{26} + 8712t^{28} + 13417t^{30} + 19867t^{32} + 28414t^{34} + 39351t^{36} \\ & + 52604t^{38} + 68220t^{40} + 85783t^{42} + 104588t^{44} + 123852t^{46} + 142559t^{48} + 159328t^{50} \\ & + 173201t^{52} + 183138t^{54} + 188232t^{56} + 188232t^{58} + 183138t^{60} + 173201t^{62} + 159328t^{64} \\ & + 142559t^{66} + 123852t^{68} + 104588t^{70} + 85783t^{72} + 68220t^{74} + 52604t^{76} + 39351t^{78} \\ & + 28414t^{80} + 19867t^{82} + 13417t^{84} + 8712t^{86} + 5441t^{88} + 3293t^{90} + 1880t^{92} + 1038t^{94} \\ & + 552t^{96} + 273t^{98} + 126t^{100} + 61t^{102} + 22t^{104} + 9t^{106} + 5t^{108} + t^{110} + t^{114}, \end{aligned} \tag{3.3}$$

which has a palindromic form.[9] The denominator is

$$\mathcal{D}(t) \;=\; \left(1-t^2\right)^3 \left(1-t^4\right)^4 \left(1-t^6\right)^4 \left(1-t^8\right)^2 \left(1-t^{10}\right)^2 \left(1-t^{12}\right)^3 \left(1-t^{14}\right)^2 \left(1-t^{16}\right). \tag{3.4}$$

As expected, the powers of the factors in the denominator add up to 21, the number of physical parameters in the νSM which is also the cardinality of the primary set. [37, 47]. Our result of the ungraded HS is consistent with those found in Refs. [37, 47]. We have furthermore calculated the multi-graded HS with a different parameter $\{e, m, n\}$ counting the degrees of the couplings $\{Y_e, M_N, Y_N\}$, which we only show in App. B due to its length. To obtain the results in this paper, we have developed our own `Mathematica` code that can efficiently calculate the Hilbert series. The code will be published as a `Mathematica` package under the name `CHINCHILLA` [48].

Plugging the ungraded HS in Eq. (2.10) to calculate the PL, we find furthermore

$$\begin{aligned}\mathrm{PL}[\mathcal{H}(t)] =& 3t^2 + 5t^4 + 9t^6 + 10t^8 + 19t^{10} + 40t^{12} + 66t^{14} + 92t^{16} + 70t^{18} - 124t^{20} \\ & - 703t^{22} - 2039t^{24} - 4391t^{26} - 7472t^{28} - 8522t^{30} + 590t^{32} + O(t^{34}). \end{aligned} \tag{3.5}$$

We only show the PL up to order 32. For higher orders, both positive and negative terms will appear repeatedly in an infinite series, which implies the νSM has an algebraic structure that resembles a *non-complete intersection* ring.

The usual interpretation of the PL suggests that the leading positive terms indicate a total of 314 basic invariants. However, our analysis reveals that this number is incorrect. Discrepancies begin to arise at order 16 in the PL, where the count of basic invariants exceeds 92. This is due to the non-complete intersection nature of the ring, resulting in cancellations between the number of basic invariants and the number of syzygies. Concrete

[9] Note that in the numerator Eq. (3.3), a term $(1+t^2)$ can be factorized, and the same factor also appears in the denominator Eq. (3.4). If this factor were to be simplified, the Hilbert series would take a rational form with a numerator featuring some negative terms, in contradiction with the positivity requirement announced earlier However, one can always multiply a factor $(1+t^k)^m$ in both numerator and denominator if there exists a factor $(1-t^k)^n$ (where $m \le n$) in the denominator. This multiplication removes a factor of $(1-t^k)^m$ from the denominator while introducing a new factor of $(1-t^{2k})^m$. The total number of factors in the denominator does not change, and the numerator keeps its palindromic form with positive terms. This freedom indicates that there is ambiguity in determining the form of the Hilbert series if there is no further requirement of the Hilbert series.

examples will be provided in the following sections to clarify this matter. It is worth noting that a similar cancellation was observed in a low-energy neutrino model in Ref. [29], which also corresponds to a non-complete intersection ring. There is no rigorous proof for this cancellation, however we can try to understand it as follows. For a sufficiently simple invariant ring, the orders in the PL corresponding to the appearance of syzygies are well-separated from the orders corresponding to the appearance of the basic invariants. When the ring becomes more complicated, more invariants are needed to describe the full algebraic structure of the ring, hence there exist more basic invariants at higher orders. If syzygies containing the lower order basic invariants still appear at a similar order as in less complicated rings, there will be an overlap between the regions of positive and negative terms. This overlap will result in cancellations between the number of basic invariants and the number of syzygies. Therefore, one should be cautious when using the PL to count the number of basic invariants in a non-complete intersection ring. Observing a negative term in the PL does not necessarily imply the absence of basic invariants, but rather indicates the presence of more syzygies than basic invariants. It is preferable to define *the coefficient in the PL as the difference between the number of basic invariants and the number of syzygies*. Moreover, the coefficients in the ungraded PL at a specific order can be subject to cancellations from terms which have a different grading for the same total order in the multi-graded PL but cancel once the ungraded limit is taken. In this sense, we can not naively assume that the leading positive terms in Eq. (3.5) can capture all basic invariants as we expected for the theory with a complete intersection ring. Hence, the multi-graded PL (see App. B) is our main guide to check if we have found the correct number of basic invariants and syzygies at a given order in the spurions. It is also conjectured in Ref. [30] that the basic invariants are all captured by the terms prior to the pure negative order[10] in multi-graded PL. The pure negative terms occur at order 26 in Eq. (B.3) in our theory, so the basic set should already be found within order 24. However, to test the conjecture, we also construct invariants up to order 26 to see there is indeed no basic invariant can be found in this pure negative order.

3.2 Constructing the invariants

Although the PL may provide some clues regarding the number of basic invariants, their specific form remains unknown. While it is possible to construct invariants for some simple models manually, in the case of complex models such as the νSM, which involves hundreds of invariants in their generating sets, it becomes unfeasible to manually build them. Furthermore, as mentioned at the end of the last section, in a sufficiently complicated ring, the orders at which generating invariants and syzygies appear in the PL might overlap, hence leading to cancellations. To construct a generating set, one therefore cannot solely rely on the information provided by the PL. Instead, we will start by constructing all possible invariants up to a given order which is done by a method based on graphs introduced in this section. The objective is to eliminate redundant invariants with linear relations and create a generating set. This generating set will then be used to obtain a primary set.

[10] In the multi-graded PL, we sort the terms according to the order (total degrees) of $[emn]$ (c.f. Eq. (B.3)). In a given order, if all terms are negative, then it's called pure negative order.

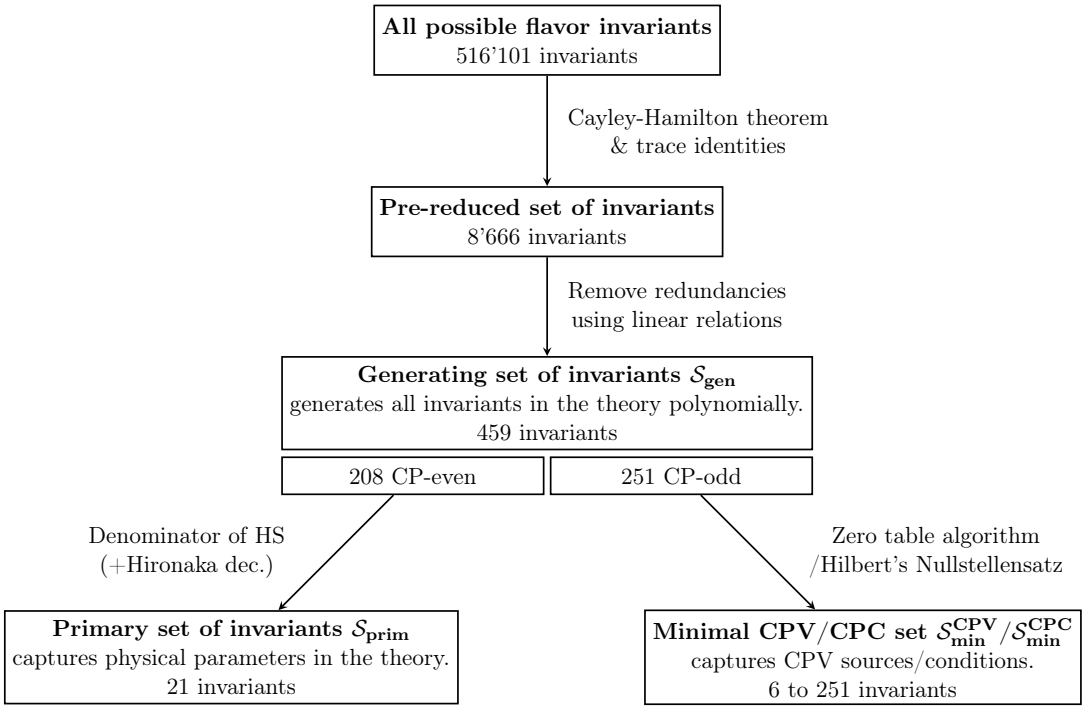
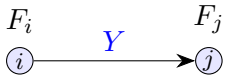


Figure 1: Flow graph of different invariant sets that appear in the analysis alongside the algorithms that are used to obtain one set from another. The numbers in the top two boxes correspond to all single trace invariants up to $\mathcal{O}([emn]^{26})$. Finally, we find 208 CP-even and 251 CP-odd invariants that make up the generating set of the ring defined by the Lagrangian parameters and their transformation properties. A complete list of these invariants can be found in App. D. The 21 algebraically independent invariants are selected from the CP-even generating set to form the primary set. These invariants accurately capture the 21 physical parameters of the theory. To capture CPV sources and determine the CPC conditions, the CP-odd invariants are selected from the CP-odd generating set. However, our program fails to find the minimal set due to the complexity of the theory. Detailed explanations on the Hironaka decomposition, the zero table algorithm and Hilbert's Nullstellensatz can be found in App. F, App. C.3 and App. G respectively.

Furthermore, we will analyze the CPV effects based on the CP-odd generating invariants of the theory. The entire process is summarized in Fig. 1, and we will provide a detailed explanation of each step in the following sections.

Flavor invariants from closed walks in the graph The construction of flavor invariants in our paper heavily relies on the flavor invariant graph, which is inspired by Ref. [49]. We show the graph of the νSM in Fig. 2, where the lepton sector and quark sector are presented in the top and bottom panels respectively. In a graph, the nodes represent the different fields and their conjugates, while the arrow lines are labeled with flavor matrices connecting the flavor indices of the various fields at the nodes. These graphs have the following two main advantages. First, they display the transformation rules for all flavor matrices. Taking the general building block of the graph as an example,



where the arrow labeled with a flavor matrix Y, starting from vertex i to vertex j, which

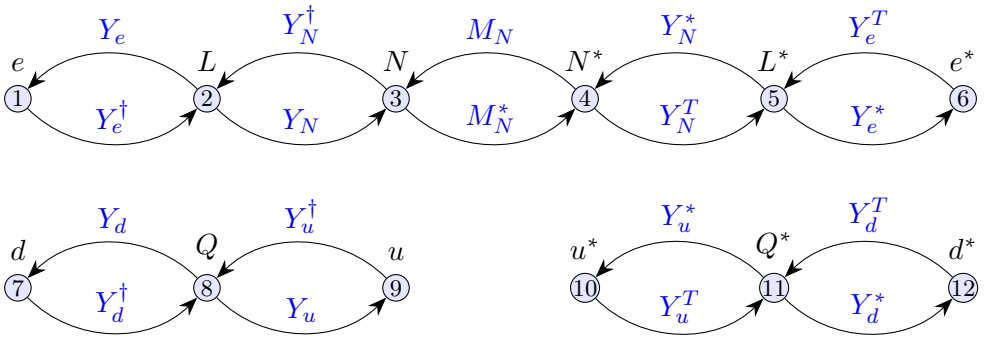
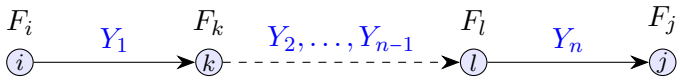


Figure 2: The flavor invariant graph that can be used to construct all possible single trace flavor invariants in the νSM. To any closed walk that follow the arrows, one can associate a single-trace invariant. Note that the graph for the SM would have a "holomorphic" structure, i.e., it has two separated branches involving separately only fields or only their conjugates. This changes in the νSM, where the transformation properties of the Majorana mass M_N connect the holomorphic and antiholomorphic branches. More details can be found in the main text.

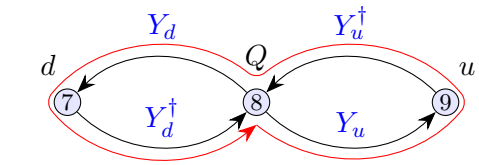
are labeled with the fields F_i and F_j respectively, indicating that the flavor matrix Y should transform as $Y \to U_{F_i} Y U_{F_j}^\dagger$ under the flavor group. For example, for the graph in Fig. 2, we can read off $Y_e \to U_L Y_e U_e^\dagger$ and $Y_e^* \to U_{L^*} Y_e^* U_{e^*}^\dagger = U_L^* Y_e^* U_e^T$, where we have used the fact $U_{f^*} = U_f^*$. Second, following the directional flow of the arrows, one can pass through vertices and edges, creating "paths". If the repetition of vertices and edges are allowed, these paths will be referred to as "walks" in mathematical terminology. To every walk, one can associate an object with specific transformation under the flavor transformations. For instance, in the graph below, starting from vertex i and following a sequence of arrows until reaching vertex j,



then the product of flavor matrices $X \equiv Y_1 \ldots Y_n$ will transform as $X \to U_{F_i} X U_{F_j}^\dagger$. If there is a walk, with $i = j$, which means the walk is closed and ends at the starting vertex, then we will have $X \to U_{F_i} X U_{F_i}^\dagger$. This implies $\mathrm{Tr}(X)$ is a single trace invariant. As a result, one can construct flavor invariants from closed walks in the graph. Conversely, one can also identify a closed walk with each single-trace invariant.

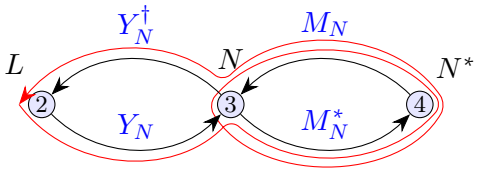
For instance, one can find the following correspondences between invariants and the closed walks in the quark and lepton sectors (the red paths represent closed walks on the

graphs denoted by the black arrows connecting the vertices)



$$\sim 8 \to 9 \to 8 \to 7 \to 8$$

$$\sim \mathrm{Tr}\Big(Y_u Y_u^\dagger Y_d Y_d^\dagger\Big)\,,$$



$$\sim 2 \to 3 \to 4 \to 3 \to 4 \to 3 \to 2 \tag{3.6}$$

$$\sim \mathrm{Tr}\Big(Y_N M_N^* M_N M_N^* M_N Y_N^\dagger\Big)\,.$$

Note that the walks are equivalently given as a chain of numbers corresponding to the vertices that are passed through. We only show here simple examples of walks in subsets of the νSM graph. Obviously, one can consider more complicated walks involving more vertices and obtain new flavor invariants accordingly. Due to the cyclicity of the trace, the invariant associated to a closed walk is independent of the starting vertex, e.g. $8 \to 9 \to 8 \to 7 \to 8 = 7 \to 8 \to 9 \to 8 \to 7$ for the first example above. In addition, since the last vertex in a closed walk is always the same as the first vertex, we can omit the last number in the chain. To further shorten the notation, we can also remove the arrow, resulting in an integer representation of the walk notation. Furthermore, due to the cyclicity of the walk, we can rotate the numbers in the chain to arrange them in lexicographically smallest order. By following this approach, all single trace invariants are uniquely represented as integers. For instance, the two invariants above should be represented by two integers 7898 and 234343 respectively.

The inclusion of the Majorana neutrino mass M_N creates an important difference between the lepton and the quark sectors. The quark sector has two separate parts in the graph: the closed walks involve either the fields or their complex conjugate at the vertices, but never mix both. The walks are holomorphic or anti-holomorphic. The set of invariants built from the anti-holomorphic closed walks is equivalent of the one built from the holomorphic walks (see App. C.1 for details). The introduction of the Majorana neutrino mass term M_N complicates the scenario in the lepton sector. The two conjugate parts are connected through this new flavor matrix, making the invariant structure much more complex, as we will see.

In the Dirac limit $M_N \to 0$, the quark and the lepton sectors obviously have the same flavor structure, and thus have the same forms of the flavor invariants. In Section C.1, we show how to construct the flavor invariants in the quark sector based on the graph approach. The generating flavor invariants are shown explicitly, which can be easily mapped to the flavor invariants of the lepton sector in the Dirac limit.

For nonzero and finite M_N, we can systematically enumerate the closed walks corresponding all single-trace flavor invariants up to arbitrarily high order. Based on the arguments of Ref. [30], we are particularly interested in invariants up to order 26 which is the order of the first pure negative term in the multi-graded PL of the νSM, and therefore it should be possible to obtain a set of basic invariants out of them. In App. C.1, we present the brute-force algorithm we used to construct all these single-trace invariants up to order 26. We obtain 516'001 of them. However, there is still a lot of redundancy among them, and we can immediately reduce the set of invariants using some simple relations.

- Transpose redundancy

  The flavor invariant associated with the walk $W_1 \equiv ij\ldots kl$ is always accompanied with another walk with primed vertices in reverse order $W_2 \equiv l'k'\ldots j'i'$, where $v' = 7-v$ for the graph in the lepton sector, featuring a left and right mirror symmetry. The invariants generated from these two walks are the same, and they are connected through the trace identity $\mathrm{Tr}\left(X^T\right) = \mathrm{Tr}\left(X\right)$. For instance, $\mathrm{Tr}\left(Y_e Y_e^\dagger\right) = \mathrm{Tr}\left(Y_e^* Y_e^T\right)$.

- Conjugate redundancy

  The walk W_1 defined above is also associated with $W_3 \equiv i'j'\ldots k'l'$. The invariants generated by these two walks are conjugate to each other. Thus, both $\mathrm{Tr}(X)$ and $\mathrm{Tr}(X^*)$ will be generated in our construction. As the CP properties of $\mathrm{Re}\,\mathrm{Tr}(X)$ and $\mathrm{Im}\,\mathrm{Tr}(X)$ are more transparent than those of $\mathrm{Tr}(X)$ and $\mathrm{Tr}(X^*)$, we will trade $\mathrm{Tr}(X)$ and $\mathrm{Tr}(X^*)$, which are generated by the graphs, with $\mathrm{Re}\,\mathrm{Tr}(X)$ and $\mathrm{Im}\,\mathrm{Tr}(X)$, whenever a complex invariant is found.

- Cayley–Hamilton theorem

  The Cayley–Hamilton theorem, along with its variations, enables us to eliminate invariants or establish relations among them. The trace identities induced by the Cayley–Hamilton theorem hold true for generic matrices, but further identities hold among the invariants since they end up involve products of matrices with special structure like $Y_e Y_e^\dagger$, or with special symmetry properties as for M_N.

The details of these redundancies are discussed in App. C.1. Making full use of the redundancies, we are left with a set of 8'666 invariants. This is still a too large set to form a generating set.

Construction of the generating set After pre-reducing the set of invariants with well-known trace relations, we will then search for linear relations of invariants in terms of the other invariants at the previous order. This will allow us to identify the generating set of invariants which, by definition, does not have any linear redundancy. To fully remove all *linear* dependence and obtain a *basic* set, we introduce a numerical algorithm outlined in App. C.2, which converts the problem of finding polynomial relations to a problem of solving finite system of linear equations. This algorithm has been used in different forms in Refs. [27, 29, 34, 38] before. We have improved the algorithm for this work to avoid

redundant syzygies[11] as we show in detail in App. C.2. By using this method, we are able to generate all possible polynomial relations among the invariants at each degree, which includes both linear relations and syzygies. Once a linear relation is found, the corresponding invariant will be removed from our set.

Running this algorithm up to order 26, the invariant list is reduced to a set without linear dependencies. Our final set includes 459 invariants which form a *generating* set of the flavor invariants in the νSM. Out of the 459 generating invariants 208 are CP-even and 251 are CP-odd. We want to stress again that all invariants in the theory can be captured by polynomials of these 459 flavor invariants. Hence, any observable in the theory is in principle expressible in terms of these invariants. In the following sections, we will further reduce the *generating* set to a *primary* set which captures all physical parameters in the theory. Additionally, we will reduce the CP-odd generating set to a minimal CPV set that captures all sources of CPV, as well as a minimal CPC set that captures all necessary and sufficient conditions for CPC in the νSM (c.f. Fig. 1). We provide the full generating set in App. D, which is split into CP-even set and CP-odd set. From now on, we will use $\mathcal{I}_i(\mathcal{J}_i)$ to represent the ith invariant in the CP-even(CP-odd) set. When referencing $\mathcal{I}_i(\mathcal{J}_i)$, one should bear in mind that these invariants are from our generating set.

Linear relations and syzygies To cross-check if the correct amount of generating invariants was found, we can use the information provided by the graded PL. In order to match the numbers in the PL given by the number of generating invariants minus the number of syzygies at each order, we also have to find all syzygies at a given order. All of this have been done with the process described above. Now, we will provide some examples for the polynomial relations we have found, which include both linear relations and syzygies. Our program scans and checks all terms in the PL from lowest order to highest order. Prior to order 12, our reduced invariant set accurately produces the terms in the PL. This means that no polynomial relation can be found, and all the invariants in our pre-reduced set are basic invariants. At degree e^6n^6 in spurions, two invariants are found in our pre-reduced set, they are given by $\mathcal{I} \equiv \mathrm{Re}\,\mathrm{Tr}\big(X_N^2 X_e^2 X_N X_e\big)$ and $\mathcal{J}_{10} \equiv \mathrm{Im}\,\mathrm{Tr}\big(X_N^2 X_e^2 X_N X_e\big)$. From the multi-graded PL in Eq. (B.3), a positive term $+e^6n^6$ is observed. According to the usual understanding of the PL, there will be only one basic invariant at this degree, which means one of the invariants $\mathcal{I}$ and $\mathcal{J}_{10}$ is redundant. This is confirmed by our algorithm, we find that $\mathcal{I}$ can be written as a polynomial of other lower degree CP-even invariants

$$\begin{aligned} 6\mathcal{I} =& \mathcal{I}_3^3\mathcal{I}_1^3 - \mathcal{I}_3\mathcal{I}_6\mathcal{I}_1^3 - 3\mathcal{I}_3^2\mathcal{I}_7\mathcal{I}_1^2 + 3\mathcal{I}_3\mathcal{I}_{12}\mathcal{I}_1^2 - \mathcal{I}_3^3\mathcal{I}_4\mathcal{I}_1 + \mathcal{I}_4\mathcal{I}_{11}\mathcal{I}_1 + \\ &+3\mathcal{I}_3^2\mathcal{I}_{13}\mathcal{I}_1 - 3\mathcal{I}_3\mathcal{I}_{18}\mathcal{I}_1 + \mathcal{I}_3\mathcal{I}_6\mathcal{I}_9 - \mathcal{I}_9\mathcal{I}_{11} + 3\mathcal{I}_{12}\mathcal{I}_{13} + 3\mathcal{I}_7\mathcal{I}_{18}\,. \end{aligned} \tag{3.7}$$

With this linear relation, the CP-even invariant $\mathcal{I}$ becomes redundant, while no such linear relation can be found for the CP-odd invariant $\mathcal{J}_{10}$; as a result it is collected in our basic set. The same linear relation was found in the discussion of the quark flavor invariants in

[11] Redundant syzygies are those which have previously appeared in the algorithm at a lower order in the spurions and are multiplied by another syzygy, which has previously appeared, or some invariant of the generating set, hence reappearing at a higher order. These kind of syzygies evidently do not carry any new information. This is also discussed in Ref. [27], where the term "old relation" is used.

Ref. [26]. The CP-odd invariant $\mathcal{J}_{10}$ in the lepton sector is analogous to the Jarlskog invariant $I^{(-)}_{6,6}$ in the quark sector as shown in Eq. (C.3), and is the only CP-odd generating invariant in the νSM that has no dependence on M_N. Note that $I^{(-)}_{6,6} = 2i\,\mathrm{Im}\,\mathrm{Tr}\big(X_u^2 X_d^2 X_u X_d\big)$, while in our notation the additional factor "$2i$" is omitted.

Running our algorithm up to order 14, we get exactly the same number of basic invariants as indicated by the positive terms in the multi-graded PL. However, we run into a mismatch at degree m^8n^8. There are 10 invariants from the pre-reduced set, we only find 9 linear relations. As a result, we are left with one invariant $\mathcal{J}_{76}$ that can not be written as any polynomial of other invariants, by definition, it should be identified as a *basic* invariant. But looking at the multi-graded PL in Eq. (B.3), one does not find a positive term at degree m^8n^8 and would naively believe that there are no basic invariants at that degree. However, as we mentioned in Sec. 3.1, there can be non-trivial cancellations between the number of basic invariants and syzygies. Indeed, we also find another relation at degree m^8n^8, which is given by

$$
\begin{aligned}
&3\mathcal{I}_5^2\mathcal{I}_3^4 - 3\mathcal{I}_2\mathcal{I}_{10}\mathcal{I}_3^4 - 8\mathcal{I}_2^3\mathcal{I}_8\mathcal{I}_3^3 + 12\mathcal{I}_2\mathcal{I}_5\mathcal{I}_8\mathcal{I}_3^3 + 8\mathcal{I}_8\mathcal{I}_{10}\mathcal{I}_3^3 + 12\mathcal{I}_2^2\mathcal{I}_{16}\mathcal{I}_3^3 - 24\mathcal{I}_5\mathcal{I}_{16}\mathcal{I}_3^3 + \\
&+ 30\mathcal{I}_2^2\mathcal{I}_8^2\mathcal{I}_3^2 - 18\mathcal{I}_5\mathcal{I}_8^2\mathcal{I}_3^2 + 36\mathcal{I}_{16}^2\mathcal{I}_3^2 - 6\mathcal{I}_5^2\mathcal{I}_6\mathcal{I}_3^2 + 6\mathcal{I}_2\mathcal{I}_6\mathcal{I}_{10}\mathcal{I}_3^2 + 8\mathcal{I}_2^3\mathcal{I}_{14}\mathcal{I}_3^2 - 12\mathcal{I}_2\mathcal{I}_5\mathcal{I}_{14}\mathcal{I}_3^2 + \\
&- 8\mathcal{I}_{10}\mathcal{I}_{14}\mathcal{I}_3^2 + 10\mathcal{I}_2^3\mathcal{I}_{15}\mathcal{I}_3^2 - 18\mathcal{I}_2\mathcal{I}_5\mathcal{I}_{15}\mathcal{I}_3^2 - 4\mathcal{I}_{10}\mathcal{I}_{15}\mathcal{I}_3^2 - 60\mathcal{I}_2\mathcal{I}_8\mathcal{I}_{16}\mathcal{I}_3^2 - 12\mathcal{I}_2^2\mathcal{I}_{20}\mathcal{I}_3^2 + \\
&+ 24\mathcal{I}_5\mathcal{I}_{20}\mathcal{I}_3^2 - 24\mathcal{I}_2^2\mathcal{I}_{21}\mathcal{I}_3^2 + 24\mathcal{I}_5\mathcal{I}_{21}\mathcal{I}_3^2 + 24\mathcal{I}_2\mathcal{I}_{28}\mathcal{I}_3^2 - 24\mathcal{I}_2\mathcal{I}_8^3\mathcal{I}_3 + 8\mathcal{I}_2^3\mathcal{I}_6\mathcal{I}_8\mathcal{I}_3 + \\
&- 12\mathcal{I}_2\mathcal{I}_5\mathcal{I}_6\mathcal{I}_8\mathcal{I}_3 - 8\mathcal{I}_6\mathcal{I}_8\mathcal{I}_{10}\mathcal{I}_3 - 24\mathcal{I}_2^2\mathcal{I}_8\mathcal{I}_{14}\mathcal{I}_3 + 24\mathcal{I}_5\mathcal{I}_8\mathcal{I}_{14}\mathcal{I}_3 - 48\mathcal{I}_2^2\mathcal{I}_8\mathcal{I}_{15}\mathcal{I}_3 + 24\mathcal{I}_5\mathcal{I}_8\mathcal{I}_{15}\mathcal{I}_3 + \\
&+ 48\mathcal{I}_8^2\mathcal{I}_{16}\mathcal{I}_3 - 12\mathcal{I}_2^2\mathcal{I}_6\mathcal{I}_{16}\mathcal{I}_3 + 24\mathcal{I}_5\mathcal{I}_6\mathcal{I}_{16}\mathcal{I}_3 + 24\mathcal{I}_2\mathcal{I}_{14}\mathcal{I}_{16}\mathcal{I}_3 + 48\mathcal{I}_2\mathcal{I}_{15}\mathcal{I}_{16}\mathcal{I}_3 - 16\mathcal{I}_2^3\mathcal{I}_{19}\mathcal{I}_3 + \\
&+ 24\mathcal{I}_2\mathcal{I}_5\mathcal{I}_{19}\mathcal{I}_3 + 16\mathcal{I}_{10}\mathcal{I}_{19}\mathcal{I}_3 + 48\mathcal{I}_2\mathcal{I}_8\mathcal{I}_{20}\mathcal{I}_3 - 48\mathcal{I}_{16}\mathcal{I}_{20}\mathcal{I}_3 + 48\mathcal{I}_2\mathcal{I}_8\mathcal{I}_{21}\mathcal{I}_3 - 48\mathcal{I}_{16}\mathcal{I}_{21}\mathcal{I}_3 + \\
&+ 48\mathcal{I}_2^2\mathcal{I}_{27}\mathcal{I}_3 - 48\mathcal{I}_5\mathcal{I}_{27}\mathcal{I}_3 - 48\mathcal{I}_8\mathcal{I}_{28}\mathcal{I}_3 - 48\mathcal{I}_2\mathcal{I}_{39}\mathcal{I}_3 + 3\mathcal{I}_5^2\mathcal{I}_6^2 - 6\mathcal{I}_2^2\mathcal{I}_6\mathcal{I}_8^2 + 6\mathcal{I}_5\mathcal{I}_6\mathcal{I}_8^2 + 6\mathcal{I}_2^2\mathcal{I}_{14}^2 + \\
&- 6\mathcal{I}_5\mathcal{I}_{14}^2 + 24\mathcal{I}_2^2\mathcal{I}_{15}^2 - 12\mathcal{I}_5\mathcal{I}_{15}^2 - 12\mathcal{I}_6\mathcal{I}_{16}^2 + 24\mathcal{I}_{20}^2 + 12\mathcal{I}_{21}^2 - 3\mathcal{I}_2\mathcal{I}_6^2\mathcal{I}_{10} + 12\mathcal{I}_2\mathcal{I}_8^2\mathcal{I}_{14} + \\
&+ 24\mathcal{I}_2\mathcal{I}_8^2\mathcal{I}_{15} - 10\mathcal{I}_2^3\mathcal{I}_6\mathcal{I}_{15} + 18\mathcal{I}_2\mathcal{I}_5\mathcal{I}_6\mathcal{I}_{15} + 4\mathcal{I}_6\mathcal{I}_{10}\mathcal{I}_{15} + 12\mathcal{I}_2\mathcal{I}_6\mathcal{I}_8\mathcal{I}_{16} - 24\mathcal{I}_8\mathcal{I}_{14}\mathcal{I}_{16} + \\
&- 24\mathcal{I}_8\mathcal{I}_{15}\mathcal{I}_{16} + 24\mathcal{I}_2^2\mathcal{I}_8\mathcal{I}_{19} - 24\mathcal{I}_5\mathcal{I}_8\mathcal{I}_{19} - 24\mathcal{I}_2\mathcal{I}_{16}\mathcal{I}_{19} - 24\mathcal{I}_8^2\mathcal{I}_{20} + 12\mathcal{I}_2^2\mathcal{I}_6\mathcal{I}_{20} - 24\mathcal{I}_5\mathcal{I}_6\mathcal{I}_{20} + \\
&- 48\mathcal{I}_2\mathcal{I}_{15}\mathcal{I}_{20} - 24\mathcal{I}_2\mathcal{I}_{14}\mathcal{I}_{21} + 6\mathcal{I}_2^3\mathcal{I}_{26} - 6\mathcal{I}_2\mathcal{I}_5\mathcal{I}_{26} - 12\mathcal{I}_{10}\mathcal{I}_{26} - 48\mathcal{I}_2\mathcal{I}_8\mathcal{I}_{27} + 48\mathcal{I}_{16}\mathcal{I}_{27} + \\
&+ 24\mathcal{I}_{14}\mathcal{I}_{28} - 24\mathcal{I}_2^2\mathcal{I}_{38} + 24\mathcal{I}_5\mathcal{I}_{38} + 48\mathcal{I}_8\mathcal{I}_{39} + 24\mathcal{I}_2\mathcal{I}_{55} + 12\mathcal{J}_1^2 = 0\,,
\end{aligned}
\tag{3.8}
$$

This polynomial relation indicates a redundancy of the CP-odd invariant squared $\mathcal{J}_1^2$ – not a linear redundancy of an invariant – and therefore by definition it is a syzygy. Hence, combining the two findings we have

$$
\mathrm{PL}\,(e,m,n) \supset +m^8n^8 - m^8n^8\,, \tag{3.9}
$$

explaining the non-trivial zero in the PL. At higher degrees in the spurions even stronger cancellations appear, that can even generate negative terms at orders where basic invariants may exist. Some examples are

$$
\mathrm{PL}\,(e,m,n) \supset (21-1)\,e^4m^4n^8 + (3-1)\,e^2m^4n^{12} + (3-6)\,e^2m^6n^{10} + (2-6)\,e^2m^8n^8\,, \tag{3.10}
$$

where we have used $(n_b - n_s)$ as a coefficient to indicate there are n_b basic invariants and n_s syzygies at the corresponding degree. Although there may be challenges when identifying

Order		2	4	6	8	10	12	14	16	18	20	22	24	26	Total
Basic	CP-even	3	5	9	8	12	17	25	33	41	34	17	4	0	208
	CP-odd	0	0	0	2	7	23	41	61	61	42	13	1	0	251
Syz.		0	0	0	0	0	0	0	2	32	200	733	2044	4391	7402
PL=Basic−Syz.		3	5	9	10	19	40	66	92	70	−124	−703	−2039	−4391	−6943

Table 2: The number of basic invariants and syzygies from order 2 to order 26, where the basic invariants are split into CP-even and CP-odd in the counting. The difference between the number of basic invariants and number of syzygies precisely aligns with the terms in the ungraded PL shown in Eq. (3.5). In the last column, we list the total number of CP-even, CP-odd basic invariants, syzygies, and their difference. In the complete intersection ring, the difference between the number of basic invariants and number of syzygies should be the Krull dimension, which is 21 in our theory. The negative number shown here featuring a non-complete intersection ring.

basic invariants and syzygies, as long as the terms in PL are correctly interpreted, we will be able to determine the correct number of the basic invariants and syzygies at each degree. Our algorithm accurately generates the terms in PL up to order 24 based on the updated description of the PL. However, at order 26, there are some mismatches due to the "redundant syzygies", which are products of syzygies that appeared at a lower order in the spurions. We provide solutions to address these mismatches. For more details, please refer to App. C.2. The order 26 is the first order that only has negative terms in the multi-graded PL. We also confirmed that there is no basic invariants at this order. Finally, our program successfully generate the correct number of basic invariants and syzygies at each degree of $[emn]$, and they can lead to the graded PL shown in Eq. (B.3). We have summarized the number of basic invariants and syzygies, as well as the CP-even and CP-odd counting at each degree in Tabs. (7,8). Under the new interpretation, the PL should be presented in a new form with cancellations. The coefficient of each term should be replaced with $(n_b - n_s)$ as shown and described around Eq. (3.10). According to the new graded PL, one should be able to read off the correct number of basic invariants and syzygies at each degree easily. Due to its length, we will not show it here. Furthermore, we summarize the counting information at each ungraded order in Tab. 2. With this table, the ungraded PL shown in Eq. (3.5) can also be revised to the new form

$$\begin{aligned}\mathrm{PL}[\mathcal{H}(t)] =&(3-0)\,t^2+(5-0)\,t^4+(9-0)\,t^6+(10-0)\,t^8+(19-0)\,t^{10}+(40-0)\,t^{12}\\&+(66-0)\,t^{14}+(94-2)\,t^{16}+(102-32)\,t^{18}+(76-200)\,t^{20}\\&+(30-733)\,t^{22}+(5-2044)\,t^{24}+(0-4391)\,t^{26}+O(t^{28})\,.\end{aligned}\tag{3.11}$$

One can find that up to order 26, the coefficients are exactly the same as those shown in Eq. (3.5).

At higher orders, the coefficients in the PL can no longer be connected to meaningful quantities like the number of generating invariants or the number of syzygies at a given degree [29, 46]. For instance, at order 28, we can find two positive terms

$$\mathrm{PL}\,(e,m,n) \supset +6m^{14}n^{14}+4m^{16}n^{12}\,.\tag{3.12}$$

However, all invariants constructed by brute force at these two degrees are redundant after applying the Cayley-Hamilton theorem. Therefore there is no basic invariant, and these

two positive terms must be misleading.[12] The scanning program should terminate at the first pure negative order [29], and we can be assured that we get a complete and minimal set of basic invariants. It is worth noting that identifying syzygies at each order is not necessary to obtain the basic set. The identification of linear relations is sufficient for this purpose. Counting correct number of syzygies adds much complexity to our analysis, and it is only necessary to verify that we have found the correct number of basic invariants as indicated by the PL.

3.3 A primary set for the νSM

To reduce the *generating* set of invariants to a *primary* set that captures all physical degrees of freedom of the theory, we can use the denominator of the HS as a guide. In addition, the Hironaka decomposition (c.f. App. F) documented in mathematical literature proves to be useful in the field of physics studies. The significance of the Hironaka decomposition lies in its ability to simplify the analysis of the invariants in the ring, particularly in scenarios where primary invariants are CP-even and secondary invariants are CP-odd. This enables a more focused investigation into the linear span of CP-odd secondary invariants, providing necessary and sufficient conditions for CP conservation [27]. Once a set of primary and secondary invariants is chosen, there will be a unique decomposition of any invariant in the theory in terms of the primary and secondary invariants in the form of Eq. (F.2).

If one is not interested in the Hironaka decomposition, the *primary* set is defined as a set of algebraically independent invariants. These invariants will capture all physical parameters in the theory. In order to reduce redundancies from the beginning, we choose the candidate sets with cardinalities equal to 21, which is the number of physical parameters in the case of the νSM. One way to check if a candidate set is algebraically independent is to calculate the Jacobian with respect to all parameters in a given parameterization of the Lagrangian, for which we will use the parametrization from Eq. (2.2). If the rank is equal to the number of physical parameters in the theory, a set of algebraically independent invariants is found. Following this procedure we find the following primary set of invariants

$$
\begin{aligned}
&\mathcal{I}_1 = \mathrm{Tr}\left(X_e\right),\ \mathcal{I}_2 = \mathrm{Tr}\left(X_M\right),\ \mathcal{I}_3 = \mathrm{Tr}\left(X_N\right),\ \mathcal{I}_5 = \mathrm{Tr}\left(X_M^2\right),\ \mathcal{I}_6 = \mathrm{Tr}\left(X_N^2\right),\\
&\mathcal{I}_7 = \mathrm{Tr}\left(X_e X_N\right),\ \mathcal{I}_8 = \mathrm{Tr}\left(Z_{MN}\right),\ \mathcal{I}_9 = \mathrm{Tr}\left(X_e^3\right),\ \mathcal{I}_{12} = \mathrm{Tr}\left(X_e X_N^2\right),\ \mathcal{I}_{13} = \mathrm{Tr}\left(X_e^2 X_N\right),\\
&\mathcal{I}_{15} = \mathrm{Tr}\left(X_N Z_{MN}\right),\ \mathcal{I}_{23} = \mathrm{Re}\,\mathrm{Tr}\left(X_e X_N Z_{MN}\right),\ \mathcal{I}_{25} = \mathrm{Tr}\left(X_e^2 Z_{MN}\right),\\
&\mathcal{I}_{34} = \mathrm{Tr}\left(X_e^2 Y_N M_N^* Y_N^T Y_N^* M_N Y_N^\dagger\right),\ \mathcal{I}_{35} = \mathrm{Tr}\left(X_e Y_N M_N^* Y_N^T X_e^* Y_N^* M_N Y_N^\dagger\right),\\
&\mathcal{I}_{47} = \mathrm{Tr}\left(X_e^2 Y_N M_N^* Y_N^T X_N^* Y_N^* M_N Y_N^\dagger\right),\ \mathcal{I}_{50} = \mathrm{Re}\,\mathrm{Tr}\left(X_e X_N Y_N M_N^* Y_N^T X_e^* Y_N^* M_N Y_N^\dagger\right),\\
&\mathcal{I}_{54} = \mathrm{Tr}\left(X_e^2 Y_N M_N^* Y_N^T X_e^* Y_N^* M_N Y_N^\dagger\right),\ \mathcal{I}_{65} = \mathrm{Re}\,\mathrm{Tr}\left(X_e^2 X_N^2 Y_N M_N^* Y_N^T Y_N^* M_N Y_N^\dagger\right),\\
&\mathcal{I}_{79} = \mathrm{Tr}\left(X_e^2 Y_N M_N^* Y_N^T X_e^{*2} Y_N^* M_N Y_N^\dagger\right),\ \mathcal{I}_{91} = \mathrm{Re}\,\mathrm{Tr}\left(X_e^2 X_N^2 Y_N M_N^* M_N M_N^* Y_N^T Y_N^* M_N Y_N^\dagger\right),
\end{aligned}
\tag{3.13}
$$

where we have defined $X_e = Y_e Y_e^\dagger, X_N = Y_N Y_N^\dagger, X_M = M_N M_N^*$ and $Z_{MN} = Y_N M_N^* M_N Y_N^\dagger$. We want to stress here that the set of algebraically independent invariants is not unique. In

[12] The only meaning we can give to the coefficients of these terms is due to the Euler form we mentioned in footnote 7.

particular, there are many sets that are compatible with the denominator of the ungraded HS and have a Jacobian rank of 21. Furthermore, the primary invariants are only chosen to be CP-even invariants in the generating set. However, it is possible to include CP-odd invariants, provided they are algebraically independent. For instance, some CP-odd invariants are included in the primary set of the seesaw effective field theory in Ref. [37].

In addition, we have chosen the orders of the invariants in our primary set to follow those of the denominator of the HS in Eq. (3.4). For instance, the first term $(1-t^2)^3$ of the denominator in Eq. (3.4) indicates that there should be 3 invariants of order 2. However, following our discussion in footnote 9, one can change the numbers in the denominator of the HS by multiplying the numerator and denominator of the HS with the same factor $(1+t^k)^m$. In this case, more algebraically independent subsets of the generating set with cardinality 21 can function as a primary set.

It is also worth noting that we chose our primary set in such a way that it is compatible with the Hironaka decomposition up to the maximal order $\mathcal{O}\left([emn]^{26}\right)$ we could verify. We have checked this explicitly by constructing all possible secondary invariants and verifying that the numerator of the HS counts these secondary invariants correctly.

We want to make one further comment about the primary set in the context of degenerate mass spectra here. No choice of a primary set can capture all physical parameters in all degenerate cases. In particular, in the Dirac limit by taking $M_N \to 0$, one should recover the invariants of the SM quark sector upon replacing $Y_N \to Y_u, Y_e \to Y_d$. Comparing our minimal set in Eq. (3.13) in the limit $M_N \to 0$ with the primary set of the quark sector in Eq. (C.3), one can easily verify that they are not the same after making the replacements $d \to e, u \to N$. On top of that our reduced set will not be a primary set after taking the limit as it will not have enough invariants to capture all physical degrees of freedom.

We propose the following to build a minimal set which captures all physical parameters in all degenerate cases. Instead of starting with the denominator of the ungraded HS, which we believe has the ambiguity of having unphysical parameters to be removed in the spurions, we start with the denominator of the *graded* HS. This will obviously lead to a set with algebraic dependence between invariants, hence not a primary set. However, this set – which we call *maximal primary set* – can be reduced to a primary set in every limit of taking spurions to zero.

For instance, for the set in Eq. (3.13) to retain the properties of a primary set in the degenerate case of setting $M_N \to 0$, the invariants

$$\{\mathcal{I}_4, \mathcal{I}_{11}, \mathcal{I}_{18}\} \tag{3.14}$$

have to be added to the primary set in Eq. (3.13). All of these invariants correspond to factors which appear in the denominator of the multi-graded HS in Eq. (B.1). Hence, to build a minimal set which captures all degrees of freedom in the theory in all special spectra, we believe one also has to add all the invariants corresponding to the remaining factors in the denominator of the multi-graded PL.

4 Studying CP violation with flavor invariants

4.1 Capturing the sources of CP violation

According to invariant theory, all invariants in the invariant ring can be expressed as polynomials in the generating invariants. On the other hand, physical observables are not necessarily polynomials of parameters in the theory, and therefore do not generally belong to the invariant ring. However, due to their flavor invariant nature, they must be some function of the generating invariants. In this sense, all observables can be parameterized by the generating invariants. This can in particular be helpful for observables which violate CP. For instance, in the SM there is a single flavor invariant object which captures all CPV, the Jarlskog invariant.[13] Hence, all CPV observables in the SM must be expressible by this single flavor invariant. For instance, the electron EDM d_e breaks both chiral symmetry and violates CP, therefore we immediately know that $d_e \propto m_e I^{(-)}_{6,6}$ with the Jarlskog invariant $I^{(-)}_{6,6}$ which we have defined in Eq. (C.3).

We want to achieve the same for the νSM and find a way to express all CPV observables in terms of a set of flavor invariant objects. For polynomial functions of the invariants we have already found this set, which is simply the generating set. If we however also allow for non-polynomial functions of flavor invariants, that observables in QFTs can be, we can further reduce this set. As we will see below, the primary set is however too small because it does not capture all CP-odd parameters in all possible degenerate cases of the mass spectrum and texture zeros etc. Hence, we want to find the smallest set which captures all sources of CP violation in the theory for all numerical values of the Lagrangian parameters, also including cases which simplify the spectrum and make some phases unphysical. Hence, it might be necessary to include more than one invariant for each phase in the theory. We call this set the minimal CPV set $\mathcal{S}^{\text{CPV}}_{\text{min}}$.

Starting with the 251 CP-odd invariants, there is no straightforward way to reduce them to a minimal set. One way is to find syzygies which allow to express one invariant as a rational function of other invariants, but this seems hopeless given the amount of CP-odd generating invariants in the νSM with 3 generations.

Therefore, we have come up with an algorithm that allows us to reduce the CP-odd invariants in the generating set to a smaller set. The details of the algorithm and a few examples can be found in App. C.3. However, even though the algorithm works in principle to reduce the CP-odd generating set to a minimal set, we did not manage to run it successfully for our generating set of the νSM. In our algorithm we consider all choices of parameters, which simplify the spectrum of the theory and enlarge its exact flavor symmetry group, to decide which invariant should be kept in the minimal set. Due to the amount of cases that have to be considered in the νSM with 3 generations, the size of the tables storing the information of which invariant is non-zero in each case becomes extremely large. This yields several TBs of data that have to be stored, which quickly becomes completely unmanage-

[13] The Jarlskog invariant does not only capture the phase in the CKM matrix. Furthermore, it captures all cases where simplifying the spectrum of the theory renders the CKM phase unphysical. This is for example the case when two quark masses of the same type are degenerate, which is captured by the Jarlskog invariant as follows $\prod_{i>j}(y^2_{u,i} - y^2_{u,j})(y^2_{d,i} - y^2_{d,j})$.

able even on a computing cluster with many cores and a lot of memory. Nonetheless, we think our algorithm can be useful for theories with less CP-odd generating invariants to reduce the CP-odd generating set to a smaller set which captures all sources of CPV. In App. C.3, we show a reduced set for the νSM with $n_f = n_N = 2$, $n_f = n_N = 3$[14] and also run our algorithm for the Two-Higgs-Doublet model (2HDM) considered in Ref. [27].

We want to stress that the algorithm can only reduce the CP-odd generating set to a smaller set which captures the sources of CPV. Without solving the equations that are obtained by setting the invariants to zero, to our knowledge, it cannot be established if the set is also sufficient to capture all sources of CPV. In particular, there can be cases where the invariant vanishes which does not correspond to a direction in parameter space with an enhanced flavor symmetry, but instead is just a single point which happens to be a solution of the invariant. These cases can only be checked for by explicitly considering all the roots of the invariants. We give an example for such a case in App. C.3.

The problem of finding a set of flavor invariants which captures the sources of CPV in the νSM has been discussed in the literature before. Assuming non-vanishing and non-degenerate neutrino masses, the authors of Ref. [22], translated to our notation, find the following set of flavor invariants

$$\begin{aligned}
\mathcal{J}_1 &= \mathrm{Im}\,\mathrm{Tr}\left(Y_N^\dagger Y_N M_N^* M_N M_N^* Y_N^T Y_N^* M_N\right),\\
\mathcal{J}_4 &= \mathrm{Im}\,\mathrm{Tr}\left(Y_N^\dagger Y_N M_N^* M_N M_N^* M_N M_N^* Y_N^T Y_N^* M_N\right),\\
\mathcal{J}_{15} &= \mathrm{Im}\,\mathrm{Tr}\left(Y_N^\dagger Y_N M_N^* M_N M_N^* M_N M_N^* Y_N^T Y_N^* M_N M_N^* M_N\right),\\
\mathcal{J}_{31} &= \mathrm{Im}\,\mathrm{Tr}\left(Y_N^\dagger Y_e Y_e^\dagger Y_N M_N^* M_N M_N^* Y_N^T Y_e^* Y_e^T Y_N^* M_N\right),\\
\mathcal{J}_{67} &= \mathrm{Im}\,\mathrm{Tr}\left(Y_N^\dagger Y_e Y_e^\dagger Y_N M_N^* M_N M_N^* M_N M_N^* Y_N^T Y_e^* Y_e^T Y_N^* M_N\right),\\
\mathcal{J}_{116} &= \mathrm{Im}\,\mathrm{Tr}\left(Y_N^\dagger Y_e Y_e^\dagger Y_N M_N^* M_N M_N^* M_N M_N^* Y_N^T Y_e^* Y_e^T Y_N^* M_N M_N^* M_N\right),
\end{aligned} \tag{4.1}$$

where $\mathcal{J}_i$ is the ith invariant in our CP-odd generating set in App. D. However, this set clearly does not pass our requirements for a minimal CPV set as this set, by construction, trivially vanishes for $M_N \to 0$ and hence does not capture the physical Dirac phase that is present in this limit.[15] One could add the equivalent of the Jarlskog invariant $\mathcal{J}_{10} = \mathrm{Im}\,\mathrm{Tr}\left(X_N^2 X_e^2 X_N X_e\right)$ to the set to make it work in this limit, which however still does not ensure that our set is a minimal CPV set according to our definition above. Making other well-motivated assumptions, one can build similar CPV sets which capture all sources of CPV given those assumptions are true. In most cases then only 6 invariants are required to capture the 6 physical phases of the theory.

We can also approach the problem from the other end by considering which invariants *must* be included for sure in order for our set to capture all sources of CPV for all possible values of the model parameters. Simplifying cases of the parameter spectrum that help identifying these invariants are those which drastically simplify (like setting

[14] Here, n_f is the number of generations of fermions for all fermions except the right-handed neutrinos N.

[15] In certain degenerate mass spectra of charged leptons, this set also fails to meet our criteria for a minimal CPV set.

$M_N \to 0$) and give rise to only few non-zero invariants. We have found 3 such cases where only one CP-odd invariant is non-zero assuring us that those invariants have to be included in the minimal CPV set in order to capture all sources of CPV. The first is the limit $M_N \to 0$ where the theory is reduced to a copy of the SM quark sector. Then the analogue of the Jarlskog invariant $\mathcal{J}_{10} = \mathrm{Im\,Tr}\left(X_N^2 X_e^2 X_N X_e\right)$ is the only non-vanishing CP-odd invariant and has to be included in the minimal CPV set. The other 2 cases are $\{M_N \to m_N \mathbb{1}, Y_e \to 0\}$ and $\{M_N \to m_N \mathbb{1}, Y_N \to y_N \mathbb{1}\}$ which force us to add $\mathcal{J}_{74} = \mathrm{Im\,Tr}\left(Y_N Y_N^\dagger Y_N M_N^* Y_N^T Y_N^* M_N Y_N^\dagger Y_N M_N^* Y_N^T Y_N^* Y_N^T Y_N^* M_N Y_N^\dagger\right)$ and $\mathcal{J}_{251} = \mathrm{Im\,Tr}\left(X_e^2 Y_N M_N^* Y_N^T X_e^{*2} Y_N^* M_N Y_N^\dagger X_e Y_N M_N^* Y_N^T X_e^* Y_N^* M_N Y_N^\dagger\right)$ to the minimal CPV set respectively. Here, as before, we define $X_{e,N} = Y_{e,N} Y_{e,N}^\dagger$.

We want to emphasize one point here. In the quark sector of the SM all parameters are measured to high precision, telling us that none of the cases which simplify the spectrum nor a zero CKM phase are realized in nature, making the Jarlskog invariant undoubtedly non-zero. The opposite is true for the neutrino sector of the SM where some parameters are not measured well enough today to even exclude that one of the neutrinos is massless. Therefore, having a set of flavor invariants which captures all those special cases is important to make general statements about the theory that hold true for all possible experimental results.

If on the other hand none of these scenarios simplifying the renormalizable Lagrangian is realized in nature, the set of Eq. (4.1) will be sufficient to parametrize all phases in the theory. More generally, once the precision of experimental data is good enough to exclude all of those cases, any algebraically independent set of CP-odd invariants will be sufficient to capture the physical phases in the theory. Then, choosing those invariants which are the last suppressed by small couplings is likely the best flavor invariant description of the physical phases in the theory.

4.2 Conditions for CP conservation

As mentioned in the previous subsection, the vanishing of all CP-odd invariants in the generating set $\mathcal{S}_{\mathrm{gen}}$ leads to the necessary and sufficient conditions for CP conservation. However, this does not mean that the CP-odd generating set is the minimal set characterizing the CPV property. The CP-odd generating set is required to generate any value of the CP-odd invariants in a given parameterization, while the CP-odd set that can determine if there exists CPV in the theory only captures their roots. In principle, the latter set should be a subset of the CP-odd generating set, and from now on, we will call it the minimal CPC set $\mathcal{S}_{\mathrm{min}}^{\mathrm{CPC}}$.

It is possible to find a minimal CPC set, such that

$$\mathcal{J}_{\mathrm{min}} = 0,\ \forall \mathcal{J}_{\mathrm{min}} \in \mathcal{S}_{\mathrm{min}}^{\mathrm{CPC}} \Longrightarrow \mathcal{J} = 0,\ \forall \mathcal{J} \in \mathcal{S}_{\mathrm{gen}}/\mathcal{S}_{\mathrm{min}}^{\mathrm{CPC}}. \tag{4.2}$$

The straightforward method one can try is to solve the common zeros of polynomials in a candidate minimal CPC set, then apply the solutions to the other CP-odd invariants to check whether they will vanish. However, this is not practical for complicated polynomials.

Without directly solving the polynomial equations, one can also use the syzygies to determine whether other invariants are automatically zero given that all invariants in a minimal set are set to zero. This approach was e.g. followed in Ref. [27], where the author found some syzygies that can determine the minimal CPC set in the 2HDM. In this section, we will show that there is a specific form of the syzygy that can help to determine whether an invariant will vanish.

The general problem concerning the common zeros of polynomials is closely connected to the Hilbert's Nullstellensatz [50, 51], a theorem that establishes a fundamental relationship between geometry and algebra. We have presented this theorem and relevant mathematical terms in App. G. In this section, we will employ the Hilbert's Nullstellensatz to reframe the problem of identifying the minimal CPC set. In the invariant ring, or more generally in the polynomial ring $R \coloneqq \mathbb{Q}[x_1, \ldots, x_n]$, where $x_{1,\ldots,n}$ are the parameters in the theory. The CP-odd generating invariants are polynomials in this ring. Let $\mathcal{J}_s, \mathcal{J}_1, \ldots, \mathcal{J}_m \in R$, Hilbert's Nullstellensatz says that if an invariant $\mathcal{J}_s$ vanishes on all the common zeros of the $\mathcal{J}_{1,\ldots,m}$, then there exist some integer t, such that $\mathcal{J}_s^t$ is a subset of the ideal $I \coloneqq (\mathcal{J}_1, \ldots, \mathcal{J}_m)$, i.e.,

$$\mathcal{J}_s^t = f_1\mathcal{J}_1 + \cdots + f_m\mathcal{J}_m\,, \quad f_i \in R\,, \tag{4.3}$$

where f_i are the ring elements, i.e. they are polynomials of the parameters. However, since $\mathcal{J}_s$ and $\mathcal{J}_{1,\ldots,m}$ are elements in the invariant ring, f_i should also be invariants. They can be parameterized by the generating invariants in a polynomial form

$$\mathcal{J}_s^t = \sum_{i=1}^{m} P_i(\mathcal{J}_l, \mathcal{I}_k)\mathcal{J}_i\,, \quad i \neq s\,. \tag{4.4}$$

Therefore, Hilbert's Nullstellensatz tells us that if a CP-odd invariant $\mathcal{J}_s$ is redundant in the presence of a given CPC set, there must exist a syzygy of some power of $\mathcal{J}_s$ that can be used to eliminate this invariant. This theorem is quite helpful when using the syzygy approach. For example, in Ref. [27], the syzygies are used to identify the CPC conditions in the 2HDM, one can find that the syzygies regarding the vanishing invariant $\mathcal{J}_s$ can be obtained at order $\mathcal{J}_s^2$, there is no need to discuss the spectrum of the syzygies as presented in their analysis. We discuss the CPC conditions of the 2HDM in the new framework of Hilbert's Nullstellensatz in App. E.2.

It is possible to come up with an elimination algorithm based on the Hilbert's Nullstellensatz to find the minimal CPC set. However, the problem of finding a syzygy like Eq. (4.4), or in more mathematical language, determining whether an ideal $\mathcal{J}_s^t$ is a subset of another ideal $I \coloneqq (\mathcal{J}_1, \ldots, \mathcal{J}_m)$ highly relies on the calculation of the Gröbner basis, which is computationally quite expensive in complicated polynomial rings. The undetermined power t also introduces a lot of complexity in this problem. Some software systems are devoted to studying the algebraic geometry and commutative algebra, such as `Macaulay2` [52] and `Singular` [53]. They have proved helpful when exploring algebraic structures of simple theory. In addition, there is the `Mathematica` function **PolynomialReduce**, that can also be used to solve these problems. However, it's also based on the expensive Gröbner basis

calculation. Despite the capabilities of these packages, they failed to generate results within a reasonable time frame for the three-generation case.

As outlined in App. C.2, the syzygy problem can also be converted to a finite system of linear equations, which can be addressed using standard linear algebra techniques. This algorithm relies on the information about upper bound on the power t and the total degree of parameters in P_i in Eq. (4.4). There is some known bound $t \leq d^n$, with d the maximal degree of $\mathcal{J}_{1,\ldots,m}$ with respect to the parameters, and n the number of the parameters [54]. In our case, the bound is an incredible large number d^{21} for some value of d, making it impossible to scan all possible values of t. The relevant aspect regarding the complexity of such a problem is often referred to as the effective Nullstellensatz, which is difficult to solve in general cases. However, if the objective is to reduce the CPV set to a smaller set, imposing a specific smaller limit on t is also a viable approach. In that case, one can not make sure if the obtained set is minimal. On the other hand, if the theory is simple enough, one should also be able to tackle this problem by this approach.

Although we were unable to construct the minimal CPV set for the theory with three generations of fermions, we did identify some example of syzygy that follows the form shown in Eq. (4.4). For instance,

$$\begin{aligned} 2\mathcal{J}_{14}^2 = {} & \mathcal{J}_1\left(2\mathcal{J}_{34}\mathcal{I}_2 - 2\mathcal{J}_{76} - \mathcal{J}_{11}\mathcal{I}_2^2 - \mathcal{J}_{11}\mathcal{I}_5\right) + \mathcal{J}_3\left(4\mathcal{J}_{36} - 4\mathcal{J}_{13}\mathcal{I}_2 + \mathcal{I}_2^2\mathcal{J}_3 + \mathcal{I}_5\mathcal{J}_3\right) \\ & + 2\mathcal{J}_4\left(\mathcal{J}_{11}\mathcal{I}_2 - \mathcal{J}_{34}\right) - 2\mathcal{J}_{11}\mathcal{J}_{15} + 2\mathcal{J}_{13}^2\,, \end{aligned} \tag{4.5}$$

one can find that if $\{\mathcal{J}_1, \mathcal{J}_3, \mathcal{J}_4, \mathcal{J}_{11}, \mathcal{J}_{13}\}$ is set to zero, then $\mathcal{J}_{14} = 0$ automatically.

In addition, some CP-odd invariants can be identified as candidates to the minimal CPC set. Specifically, if a unique CP-odd invariant $\mathcal{J}_s$ is non-zero under a specific spectrum, it would not be possible to establish Eq. (4.4). Consequently, $\mathcal{J}_s$ cannot be eliminated by any CPC set unless it is included in the set. In Sec. 4.1, several example candidates have been identified, such as $\mathcal{J}_{10}$, $\mathcal{J}_{74}$, and $\mathcal{J}_{251}$. These candidates must be included in the minimal CPC set.

In conclusion, we introduced Hilbert's Nullstellensatz as a reliable mathematical language to address the problem of the CPC condition. However, the problem is inherently complex and cannot be solved with our current efforts. It deserves further exploration as it requires larger computational resources to be successfully resolved. Additionally, it would be interesting to investigate this approach for some simple theories.

5 Conclusions

In this paper, we have analyzed the algebraic structure of the νSM, the SM extended with three generations of sterile neutrinos, by constructing its set of generating flavor invariants. In the quark sector, the theory exhibits a complete intersection ring, and the flavor invariants can be easily generated with the guidance of the Hilbert series. However, in the νSM, the theory corresponds to a non-complete intersection ring, introducing much complexities. The analysis of the Hilbert series in this context requires a more careful examination to extract correct information about the number of generating invariants and syzygies. We construct the flavor invariants by brute force based on a graph of flavor matrices, then

reduce it to a generating set by removing redundant invariants. The number of generating invariants and syzygies are carefully compared with the Hilbert series, a consistent agreement was observed at each order of flavor invariants. The generating set comprises of in total 459 invariants, out of which 208 are CP-even and 251 are CP-odd. We have furthermore reduced this set to a primary set of flavor invariants capturing the 21 physical parameters of the renormalizable Lagrangian. It's worth noting that the flavor structure of the Type-III seesaw is identical to that of the Type-I seesaw. Therefore, the flavor invariants identified in this paper can be directly applied to the Type-III seesaw model with some renaming of flavor matrices. In the spirit of the Jarlskog invariant, we have defined a minimal set of CPV invariants, that captures all sources of CPV in all possible cases of spectra for the theory. The reduction of the CP-odd generating set to a minimal CPV set turns out to be a hard task and we only managed to reduce the set of CP-odd flavor invariants to a more minimal set in the case of two generations of sterile neutrinos, while the reduction of the three generation case was too complicated to complete by our means. Along the way, we present useful algorithms which are essential to our analysis and hopefully prove to be useful for other analyses of flavor invariants in the future.

The huge number of flavor invariants that are needed to capture all sources of CPV in the νSM contests the idea of using flavor invariants to capture all sources of CPV in all degenerate cases of the spectrum of the theory, in theories with complicated flavor structures. An alternative approach was, for instance, put forward in Ref. [55].

An obvious next step for the flavor invariants is to connect them to phenomenological applications to study, for instance, CPV observables in a consistent way with the help of flavor invariants.

Our invariants in this paper are also useful to impose conservation of CP at the order of the renormalizable Lagrangian while studying the flavor invariants of the νSMEFT to characterize the violation of CP in the theory. This was the motivation to study the invariants of the renormalizable Lagrangian in the first place.

Acknowledgments

We thank Quentin Bonnefoy, Gabriele Dian, Emanuele Gendy, Gabriel Massoni Salla and Paula Pilatus for useful discussions. This work is supported by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy EXC 2121 "Quantum Universe" – 390833306, as well as by the grant 491245950. This project also has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie Staff Exchange grant agreement No 101086085 - ASYMMETRY. C.Y.Y. is supported in part by the Grants No. NSFC-11975130, No. NSFC-12035008, No. NSFC-12047533, the National Key Research and Development Program of China under Grant No. 2017YFA0402200, the China Postdoctoral Science Foundation under Grant No. 2018M641621 and the Helmholtz-OCPC International Postdoctoral Exchange Fellowship Program.

A Parameterization of flavor matrices

A.1 Standard parameterization

The parameterizations of the flavor matrices in the lepton and quark sectors have been discussed in detail in Ref. [26, 49]. However, we use a different convention in this paper, thus we will follow a similar discussion but in the new convention. The fermionic part of the Lagrangian is given by

$$\mathcal{L}_4 \supset \sum_{\Psi} \bar{\Psi} i \not{D} \Psi - \left[\frac{1}{2} (NCM_N N) + \bar{L} Y_N N \widetilde{H} + \bar{L} Y_e e H + \bar{Q} Y_u u \widetilde{H} + \bar{Q} Y_d d H + \text{H.c.} \right], \quad \text{(A.1)}$$

where the kinetic term sums all fermion fields $\Psi = \{Q, L, u, d, e, N\}$. The flavor transformations of the fermion fields are given by

$$\begin{aligned} &L \to U_L\, L\,, \quad e \to U_e\, e\,, \quad N \to U_N\, N\,, \\ &Q \to U_Q\, Q\,, \quad u \to U_u\, u\,, \quad d \to U_d\, d\,. \end{aligned} \quad \text{(A.2)}$$

The corresponding Yukawa matrices and the Majorana mass matrix transform as

$$\begin{aligned} &Y_e \to U_L Y_e U_e^\dagger\,, \quad Y_N \to U_L Y_N U_N^\dagger\,, \quad M_N \to U_N^* M_N U_N^\dagger\,, \\ &Y_u \to U_Q Y_u U_u^\dagger\,, \quad Y_d \to U_Q Y_d U_d^\dagger\,. \end{aligned} \quad \text{(A.3)}$$

The Yukawa matrices $Y_{N,e,u,d}$ are general complex matrices, while the Majorana mass matrix M_N is symmetric, which can be diagonalized as follows

$$\begin{aligned} &Y_e = V_e\, \widehat{Y}_e\, W_e^\dagger\,, \quad Y_N = V_N\, \widehat{Y}_N\, W_N^\dagger\,, \quad M_N = V_N'\, \widehat{M}_N\, V_N'^T\,, \\ &Y_u = V_u\, \widehat{Y}_u\, W_u^\dagger\,, \quad Y_d = V_d\, \widehat{Y}_d\, W_d^\dagger\,, \end{aligned} \quad \text{(A.4)}$$

where $\widehat{Y}_{N,e,u,d}$ and $\widehat{M}_N$ are diagonal matrices with real and non-negative entries. V_f, W_f with $f = N, e, u, d$ and V_N' are unitary matrices. We can choose specific flavor transformations in Eq. (A.3) to get a fixed mass basis. In this paper, we will work on the charged lepton diagonal basis in the lepton sector and up basis in the quark sector. This can be achieved by setting $U_L = V_e^\dagger, U_e = W_e^\dagger, U_N = V_N'^T, U_Q = V_u^\dagger, U_u = W_u^\dagger, U_d = W_d^\dagger$. The flavor matrices will be fixed to the following forms,

$$\begin{aligned} &Y_e = \widehat{Y}_e\,, \quad Y_N = V_L\, \widehat{Y}_N\, W^\dagger\,, \quad M_N = \widehat{M}_N\,, \\ &Y_u = \widehat{Y}_u\,, \quad Y_d = V_{\text{CKM}}\, \widehat{Y}_d\,, \end{aligned} \quad \text{(A.5)}$$

where $V_{\text{CKM}} = V_u^\dagger V_d$ is the CKM matrix, corresponding to the mismatch between the diagonalization matrices of the up and down sectors. $V_L = V_e^\dagger V_N$ is a similar matrix in the lepton sector, describing the mismatch between the diagonalization matrices of lepton Yukawa matrices. The existence of the Majorana mass matrix in the lepton sector introduces another mixing matrix $W = V_N'^T W_N$, which describes the mismatch between the diagonalization matrices of Y_N and M_N. If the Majorana mass term is forbidden in some theory, the neutrino will have a Dirac mass, then we can simply choose $U_N = W_N$, which will lead to $W = \mathbb{1}$.

The lepton sector will have the same structure as the quark sector, and V_L can be seen as the PMNS matrix phenomenologically.

From Eq. (A.4), if we assume that the flavor matrices have no degenerate or vanishing eigenvalues[16], the diagonalization matrices will have at least column phase redefinition degrees of freedom, $V_f \to V_f e^{i\widehat{\Phi}_f}, W_f \to W_f e^{i\widehat{\Phi}_f}, f = \{N, e, u, d\}$ and $V'_N \to V'_N \eta_N$. This can be expressed as the a rephasing invariance of the diagonal matrices.

$$\begin{aligned} &\widehat{Y}_f \to e^{i\widehat{\Phi}_f} \widehat{Y}_f e^{-i\widehat{\Phi}_f}, f = N, e, u, d, \\ &\widehat{M}_N \to \eta_N \widehat{M}_N \eta_N, \end{aligned} \tag{A.6}$$

where $\widehat{\Phi}_f \equiv \mathrm{diag}(\phi_{f_1}, \phi_{f_2}, \phi_{f_3}), f = N, e, u, d$ are diagonal complex phase matrices, and η_N is a diagonal matrix with ± 1 eigenvalues. Under these rephasing, the mixing matrices transform as

$$V_{\mathrm{CKM}} \to e^{-i\widehat{\Phi}_u} V_{\mathrm{CKM}} e^{i\widehat{\Phi}_d}, \quad V_L \to e^{-i\widehat{\Phi}_e} V_L e^{i\widehat{\Phi}_N}, \quad W \to \eta_N W e^{i\widehat{\Phi}_N}, \tag{A.7}$$

The 3×3 unitary matrix can be parameterized as

$$U_3 = e^{i\varphi} e^{i\widehat{\Psi}} U(\theta_{12}, \theta_{13}, \theta_{22}, \delta) e^{i\widehat{\Phi}}, \tag{A.8}$$

where φ is a overall phase, $\widehat{\Psi} = \mathrm{diag}(0, \psi_1, \psi_2)$ and $\widehat{\Phi} = \mathrm{diag}(0, \phi_1, \phi_2)$, and $U(\theta_{12}, \theta_{13}, \theta_{23}, \delta)$ takes the standard form

$$U(\theta_{12}, \theta_{13}, \theta_{23}, \delta) \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13} e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{A.9}$$

where $s_{ij} \equiv \sin\theta_{ij}$, $c_{ij} \equiv \cos\theta_{ij}$, and $\theta_{ij} \in [0, \pi/2], \delta \in [0, 2\pi)$.

In the quark sector, by introducing the phase redefinition in Eq. (A.7), the phases in the unitary matrix V_{CKM} can be absorbed, it takes the standard form as given in Eq. (A.9). The number of parameters are summarized as follows

Matrices	Masses	Angles	Phases
$\widehat{Y}_u$	3	0	0
$\widehat{Y}_d$	3	0	0
V_{CKM}	0	3	1
Total	6	3	1

[16] If there is a degenerate or vanishing mass spectrum, there will be a larger degrees of freedom to redefine the mixing matrix, some parameters in our current parameterization could be unphysical, and they can be removed by those redefinition. Under this enlarged symmetry, we should adopt a new parameterization, the number of parameters should exactly match the number of physical observables. On the other hand, if we adopt the most general parameterization, and set the special spectrum afterwards, some parameters in the parameterization will be redundant. In some cases, the redundant parameters could be canceled in the expression, and they never appear in the flavor invariants. In other cases, some parameters will appear together as a single polynomial in different invariants. In the latter case, the polynomial that correlates them should be considered as a single parameter in the parameterization, which should remove the redundancies.

There are a total of 10 parameters, consisting of 6 quark masses, 3 mixing angles, and 1 CP phase.

In the lepton sector, we can parameterize V_L and W as

$$\begin{aligned} V_L &= e^{i\varphi} e^{i\widehat{\Psi}} U(\theta_{12}, \theta_{13}, \theta_{23}, \delta) e^{i\widehat{\Phi}} , \\ W &= e^{i\varphi'} e^{i\widehat{\Psi}'/2} U(\theta'_{12}, \theta'_{13}, \theta'_{23}, \delta') e^{i\widehat{\Phi}'} . \end{aligned} \tag{A.10}$$

Since V_L and W share the same rephasing matrix $\widehat{\Phi}_N$, we can use this freedom to remove either $\widehat{\Phi}$ or $\widehat{\Phi}'$ in Eq. (A.10), depending on what observable we are interested in. If the flavor invariant only depends on Y_N and Y_e, then the mixing matrix W can be simply set to $W = \mathbb{1}$, since it is not an observable. We can use the phase matrices $\widehat{\Phi}_e$ and $\widehat{\Phi}_N$ to remove the phases of the unitary matrix V_L, then it takes the same form as given in Eq. (A.9). The mixing matrices in this case take the following form

$$V_L = U(\theta_{12}, \theta_{13}, \theta_{23}, \delta), \quad W = \mathbb{1} . \tag{A.11}$$

If the flavor invariant only depends on Y_N and M_N, then the mixing matrix V_L is not an observable, which can be set to identity matrix. We can use the rephasing matrix $\widehat{\Phi}_N$ to remove the phases φ' and $\widehat{\Phi}'$ in Eq. (A.10), the effect of η_N is indicated by the factor of two in the phase matrix $e^{i\widehat{\Psi}'/2}$, which limits the phases in the range of $\psi'_i/2 \in [0, \pi)$ if by convention $\psi'_i \in [0, 2\pi)$. In this case the mixing matrices will take the following form

$$V_L = \mathbb{1} , \quad W = e^{i\widehat{\Psi}'/2} U(\theta'_{12}, \theta'_{13}, \theta'_{23}, \delta') . \tag{A.12}$$

If the flavor invariant depends on the three flavor matrices Y_e, Y_N and M_N[17], then without losing of generality, we can use the rephasing phases $\widehat{\Phi}_e$ and $\widehat{\Phi}_N$ to remove the phases in V_L and the phase φ' in W in Eq. (A.10), then the mixing matrices take the following forms

$$V_L = U(\theta_{12}, \theta_{13}, \theta_{23}, \delta) , \quad W = e^{i\widehat{\Psi}'/2} U(\theta'_{12}, \theta'_{13}, \theta'_{23}, \delta') e^{i\widehat{\Phi}'} . \tag{A.13}$$

We can also use $\widehat{\Phi}_e$ and $\widehat{\Phi}_N$ to remove the phases $e^{i\varphi} e^{i\widehat{\Psi}}$ and $e^{i\varphi'} e^{i\widehat{\Phi}'}$ in V_L and W in Eq. (A.10), respectively, then the mixing matrices are given by

$$V_L = U(\theta_{12}, \theta_{13}, \theta_{23}, \delta) e^{i\widehat{\Phi}} , \quad W = e^{i\widehat{\Psi}'/2} U(\theta'_{12}, \theta'_{13}, \theta'_{23}, \delta') . \tag{A.14}$$

The third case corresponds to the most general parameterization of the mixing matrices V_L and W, we can work in this basis without loosing of generality. To be specific, we use the following parameterization

$$Y_e = \text{diag}\,(y_e, y_\mu, y_\tau) , \; Y_N = V_L \cdot \text{diag}\,(y_1, y_2, y_3) \cdot W^\dagger , \; M_N = \text{diag}\,(m_1, m_2, m_3) , \tag{A.15}$$

with V_L and W defined in Eq. (A.14). We summarize the the number of parameters in each matrix in Tab. 3.

[17] The flavor invariants can not be constructed by only Y_e and M_N, since they are disconnected objects if Y_N is not included, as shown in Fig. 2.

For a more comprehensive discussion, please refer to Ref. [26]. In this paper, they also discuss the cases for $n_N = n_f = 2$ and $(n_N, n_f) = (2, 3)$ in the lepton sector. For $n_N = n_f = 2$, all flavor matrices are 2×2 matrices, the parameterization of the 2×2 unitary matrix is

$$U_2 = e^{i\varphi} \text{diag}(1, e^{i\psi}) \cdot \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \text{diag}(1, e^{i\phi}) \,. \tag{A.16}$$

Follow a similar discussion, we can easily find the following parameterization,

$$Y_e = \text{diag}\,(y_e, y_\mu)\,, \qquad Y_N = V_L \cdot \text{diag}\,(y_1, y_2) \cdot W^\dagger, \qquad M_N = \text{diag}\,(m_1, m_2)\,, \tag{A.17}$$

with

$$V_L = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \text{diag}\left(1, e^{i\phi}\right), \quad W = \text{diag}\left(1, e^{i\varphi}\right) \cdot \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}. \tag{A.18}$$

For $(n_N, n_f) = (2, 3)$, Y_N is a 3×2 matrix, and M_N is a 2×2 matrix, as a result, the flavor matrices can be parameterized as

$$Y_e = \text{diag}\,(y_e, y_\mu, y_\tau)\,, \quad Y_N = V_L \cdot \begin{pmatrix} y_1 & 0 \\ 0 & y_2 \\ 0 & 0 \end{pmatrix} \cdot W^\dagger, \quad M_N = \text{diag}\,(m_1, m_2)\,, \tag{A.19}$$

with V_L and W defined as follows

$$V_L = U(\theta_{12}, \theta_{13}, \theta_{23}, \delta) \cdot \text{diag}\left(1, e^{i\phi}, 1\right), \quad W = \text{diag}\left(1, e^{i\varphi}\right) \cdot \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}, \tag{A.20}$$

which is analogous to the parameterization in Eq. (A.14), however the last phase in $\hat{\Phi}$ is unphysical, since it vanishes when multiplied with the zeros in the last row of the 3×2 block diagonal matrix.

Matrices	Masses	Angles	Phases
$\widehat{M}_N$	3 [2] (2)	0	0
$\widehat{Y}_N$	3 [2] (2)	0	0
$\widehat{Y}_e$	3 [2] (3)	0	0
V_L	0	3 [1] (3)	3 [1] (2)
W	0	3 [1] (1)	3 [1] (1)
Total	9 [6] (7)	6 [2] (4)	6 [2] (3)

Table 3: The number of masses, mixing angles and phases in the parameterization of the lepton sector for the case of $n_N = n_f = 3\,[n_N = n_f = 2]\,(n_N = 2, n_f = 3)$.

In the Dirac limit, M_N is set to zero, and W is not relevant, we can simply set $W = \mathbb{1}$. In the case of $n_N = n_f = 3$, the matrix V_L will become the PMNS matrix V_{PMNS}, which takes the same standard form as V_{CKM}. The parameterization is given as

$$Y_e = \text{diag}\,(y_e, y_\mu, y_\tau)\,, \quad Y_N = U(\theta_{12}, \theta_{13}, \theta_{23}, \delta) \cdot \text{diag}\,(y_1, y_2, y_3)\,. \tag{A.21}$$

Similarly, for $n_N = n_f = 2$, it can be parameterized as

$$Y_e = \mathrm{diag}\,(y_e, y_\mu)\,, \qquad Y_N = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \mathrm{diag}\,(y_1, y_2)\ , \tag{A.22}$$

where the phase of V_L in the Majorana case can be removed by the rephasing of the right-handed neutrino N. For $n_N = 2, n_f = 3$, the parameterization is given by

$$Y_e = \mathrm{diag}\,(y_e, y_\mu, y_\tau)\,, \quad Y_N = U(\theta_{12}, \theta_{13}, \theta_{23}, \delta) \cdot \begin{pmatrix} y_1 & 0 \\ 0 & y_2 \\ 0 & 0 \end{pmatrix}, \tag{A.23}$$

where the phase $e^{i\phi}$ in Eq. (A.20) can be absorbed by the field N in the Dirac case. The parameters in the Dirac limit are summarized in Tab. 4.

Matrices	Masses	Angles	Phases
$\widehat{Y}_N$	3[2](2)	0	0
$\widehat{Y}_e$	3[2](3)	0	0
V_{PMNS}	0	3[1](3)	1[0](1)
Total	6[4](5)	3[1](3)	1[0](1)

Table 4: The number of masses, mixing angles and phases in the parameterization of the lepton sector for the case of $n_N = n_f = 3\,[n_N = n_f = 2]\,(n_N = 2, n_f = 3)$ in the Dirac limit.

A.2 Algebraic parameterization

The parametrization described above is favored for its phenomenological relevance, as Yukawa matrices are factorized into eigenvalues and mixing matrices, which aligns with experimental observables. However, when exploring the algebraic structures of invariants, the inclusion of trigonometric functions introduces complexity. The use of sine and cosine as distinct entities in the polynomial expansion of flavor invariants can complicate the exploration of these structures. Consequently, alternative parameterizations that are more suitable for polynomial expressions are needed.

One possible solution is to parameterize the trigonometric functions. A frequently used parameterization for the unit circle is provided as follows:

$$x(t) = \frac{1-t^2}{1+t^2}, \quad y(t) = \frac{2t}{1+t^2}\,, \quad \text{with } t \in (-\infty, +\infty). \tag{A.24}$$

However the point $(-1, 0)$ on the unit circle can only be obtained in the limit $t \to \infty$. Another parameterization that can cover the whole circle is given by

$$x(t) = \frac{1-6t^2+t^4}{1+2t^2+t^4}, \quad y(t) = \frac{4t-4t^3}{1+2t^2+t^4}\,, \quad \text{with } t \in (-1, 1]\,. \tag{A.25}$$

Thus the sine and cosine functions in the mixing matrices of our above parameterization can be replaced with $x(t)$ and $y(t)$ respectively. However, this parameterization may introduce new complexities, as it leads to rational polynomials.

Following the parameterization in Eq. (A.15), it is possible to work with the diagonal basis of both Y_e and M_N while leaving Y_N undiagonalized. The parameterization is given as follows

$$Y_e = \mathrm{diag}\,(y_e, y_\mu, y_\tau)\,, \quad Y_N = \begin{pmatrix} r_{11} & c_{12} & c_{13} \\ r_{21} & c_{22} & c_{23} \\ r_{31} & c_{32} & c_{33} \end{pmatrix}, \quad M_N = \mathrm{diag}\,(m_1, m_2, m_3)\,, \tag{A.26}$$

where r_{ij} and c_{ij} are general labels for real and complex parameters respectively, they do not correspond to specific values, and c_{ij} can be easily written as real and imaginary parts by introducing two real parameters. The phases of the first column of Y_N are absorbed by the rephasing of charged lepton fields.[18] With this parameterization, all invariants are expressed as polynomials of simple variables, making it easier to analyze the algebraic structures of the theory. It's easy to see that the number of parameters in this parameterization is still 21, which is the same as the physical parameterization in Eq. (A.15). Similarly, for $n_N = n_f = 2$, we can parameterize it as

$$Y_e = \mathrm{diag}\,(y_e, y_\mu)\,, \qquad Y_N = \begin{pmatrix} r_{11} & c_{12} \\ r_{21} & c_{22} \end{pmatrix}, \qquad M_N = \mathrm{diag}\,(m_1, m_2)\,. \tag{A.27}$$

For $n_N = 2, n_f = 3$, the parameterization is given by

$$Y_e = \mathrm{diag}\,(y_e, y_\mu, y_\tau)\,, \quad Y_N = \begin{pmatrix} r_{11} & c_{12} \\ r_{21} & c_{22} \\ r_{31} & c_{32} \end{pmatrix}, \quad M_N = \mathrm{diag}\,(m_1, m_2)\,. \tag{A.28}$$

The mapping between these two parameterizations can be found by solving equations built from the entries of Y_N, which will not be shown here.

Under the Dirac limit, the Majorana mass term M_N is set to 0, the parameters in Yukawa matrix Y_N can be further reduced by the field redefinition of right-handed neutrino N. Starting from the Y_N in Eq. (A.26), we can absorb the phases in c_{12} and c_{13} by rephasing of N, these two real parameters can be further set to 0 by two rotations of N, i.e., $N \to R_{12}(\theta_2)R_{23}(\theta_1)N$, where R_{ij} is the rotation matrix acting on the (i, j) entries of N. The rotation R_{12} can make r_{21} and r_{31} complex, but their phases can always be absorbed by the rephasing of charged lepton fields. With these two leading zeros in the second and third columns, the rephasing of N can be used again to remove the phases in c_{22} and c_{23}, the corresponding real parameters r_{22} and r_{23} can be mixed by the rotation $N \to R_{23}(\theta_3)N$. By setting proper value of θ_3, the parameter r_{23} can be set to zero. The leading two zeros in the third column making it possible to remove the phase in c_{33} by rephasing of N, resulting in a real parameter r_{33}. The whole transformation process can be presented as follows

$$Y_N \to \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & c_{22} & c_{23} \\ r_{31} & c_{32} & c_{33} \end{pmatrix} \to \begin{pmatrix} r_{11} & 0 & 0 \\ r_{21} & c_{22} & c_{23} \\ r_{31} & c_{32} & c_{33} \end{pmatrix} \to \begin{pmatrix} r_{11} & 0 & 0 \\ r_{21} & r_{22} & r_{23} \\ r_{31} & c_{32} & c_{33} \end{pmatrix} \to \begin{pmatrix} r_{11} & 0 & 0 \\ r_{21} & r_{22} & 0 \\ r_{31} & c_{32} & c_{33} \end{pmatrix} \to \begin{pmatrix} r_{11} & 0 & 0 \\ r_{21} & r_{22} & 0 \\ r_{31} & c_{32} & r_{33} \end{pmatrix}. \tag{A.29}$$

[18] It is not necessary for the rephasing degree of freedom to target the first column, any phase in each row of Y_N can be eliminated.

There is no further field redefinition can be used to reduce the number of parameters, as a result, in the Dirac limit with $n_N = n_f = 3$, we can parameterize the flavor matrices as

$$Y_e = \text{diag}\left(y_e, y_\mu, y_\tau\right), \quad Y_N = \begin{pmatrix} r_{11} & 0 & 0 \\ r_{21} & r_{22} & 0 \\ r_{31} & c_{32} & r_{33} \end{pmatrix}. \tag{A.30}$$

We find there are exactly 10 real parameters in this parameterization as expected. Similarly, for $n_N = n_f = 2$, the flavor matrices are parameterized as

$$Y_e = \text{diag}\left(y_e, y_\mu\right), \qquad Y_N = \begin{pmatrix} r_{11} & 0 \\ r_{21} & r_{22} \end{pmatrix}. \tag{A.31}$$

For $n_N = 2, n_f = 3$, the parameterization is given by

$$Y_e = \text{diag}\left(y_e, y_\mu, y_\tau\right), \quad Y_N = \begin{pmatrix} r_{11} & 0 \\ r_{21} & r_{22} \\ r_{31} & c_{32} \end{pmatrix}. \tag{A.32}$$

B Results for graded Hilbert series and plethystic logarithm

B.1 Model with $n_N = n_f = 3$

In Eqs. (3.3,3.4), we have used the same grading for all spurions to present the results, the corresponding Hilbert series is referred to as ungraded Hilbert series. However the information encoded in the ungraded Hilbert series is compressed, which is not enough for some analysis, especially when we try to use the PL to count the numbers of basic invariants and syzygies at each order. Therefore, in this section, we present the graded Hilbert series, where the single spurion t in Eqs. (3.3,3.4) is split to the multiple spurions e, m and n, which correspond to the spurions of the flavor matrices Y_e, M_N and Y_N respectively.

The denominator of the graded Hilbert series is given as follows

$$\begin{aligned}\mathcal{D}(e,m,n) = &(1-e^2)(1-m^2)(1-n^2)(1-e^4)(1-m^4)(1-n^4)(1-e^2n^2)^2(1-m^2n^2)\\ &(1-e^6)(1-m^6)(1-n^6)(1-e^2n^4)(1-e^4n^2)(1-m^2n^4)^2(1-m^4n^2)(1-e^2m^2n^2)\\ &(1-m^2n^6)(1-m^4n^4)(1-e^2m^2n^4)(1-e^2m^4n^2)(1-e^4m^2n^2)(1-m^2n^8)\\ &(1-e^2m^2n^6)(1-e^4m^2n^4)(1-e^4m^4n^2)(1-e^4m^2n^6)(1-e^4m^4n^4)(1-e^6m^2n^4)\\ &(1-e^4m^2n^8)(1-e^8m^2n^4)(1-e^8m^4n^4),\end{aligned} \tag{B.1}$$

while there are 6582 terms in the numerator, which goes up to order $\mathcal{O}([emn]^{196})$ in total powers of spurions. Due to its length we only show the terms up to $\mathcal{O}([emn]^{26})$ below

$$\begin{aligned}\mathcal{N}(e,m,n) = &1 - e^2n^2 + 2m^4n^4 + e^4n^4 + 2e^2m^2n^4 + 2m^4n^6 + 2m^6n^4 + 4e^2m^2n^6\\ &+ 4e^2m^4n^4 + 3e^4m^2n^4 + 3m^4n^8 + 3m^6n^6 + m^8n^4 + 5e^2m^2n^8 + 7e^2m^4n^6 + 3e^2m^6n^4\\ &+ 4e^4m^2n^6 + 5e^4m^4n^4 + e^6m^2n^4 + m^4n^{10} + 3m^6n^8 + m^8n^6 + 3e^2m^2n^{10} + 9e^2m^4n^8\\ &+ 6e^2m^6n^6 + e^2m^8n^4 + 4e^4m^2n^8 + 10e^4m^4n^6 + 5e^4m^6n^4 + 3e^6m^2n^6 + 3e^6m^4n^4\end{aligned}$$

$$
\begin{aligned}
&+ m^4n^{12} + m^6n^{10} + 3m^8n^8 + e^2m^2n^{12} + 5e^2m^4n^{10} + 9e^2m^6n^8 + 2e^2m^8n^6 + e^4m^2n^{10} \\
&+ 16e^4m^4n^8 + 10e^4m^6n^6 + 2e^4m^8n^4 + 4e^6m^2n^8 + 8e^6m^4n^6 + 3e^6m^6n^4 + 2e^8m^2n^6 \\
&+ e^8m^4n^4 + m^8n^{10} + m^{10}n^8 + e^2m^4n^{12} + 6e^2m^6n^{10} + 7e^2m^8n^8 - e^2m^{10}n^6 \\
&+ 10e^4m^4n^{10} + 19e^4m^6n^8 + 4e^4m^8n^6 + e^6m^2n^{10} + 14e^6m^4n^8 + 8e^6m^6n^6 + e^6m^8n^4 \\
&+ e^8m^2n^8 + 5e^8m^4n^6 + 2e^8m^6n^4 + e^{10}m^2n^6 - m^6n^{14} + 2m^{10}n^{10} + m^{12}n^8 - 3e^2m^4n^{14} \\
&+ 2e^2m^6n^{12} + 5e^2m^8n^{10} + 2e^2m^{10}n^8 - e^2m^{12}n^6 - e^4m^2n^{14} + 6e^4m^4n^{12} + 15e^4m^6n^{10} \\
&+ 12e^4m^8n^8 - 2e^4m^{10}n^6 - e^6m^2n^{12} + 9e^6m^4n^{10} + 16e^6m^6n^8 + 2e^6m^8n^6 - 2e^8m^2n^{10} \\
&+ 9e^8m^4n^8 + 4e^8m^6n^6 + e^8m^8n^4 - 2e^{10}m^2n^8 + 2e^{10}m^4n^6 - m^6n^{16} - 2m^8n^{14} + m^{12}n^{10} \\
&- 2e^2m^4n^{16} - 4e^2m^6n^{14} + 3e^2m^8n^{12} + 2e^2m^{10}n^{10} - 2e^4m^4n^{14} + 6e^4m^6n^{12} + 7e^4m^8n^{10} \\
&- e^4m^{10}n^8 - 2e^4m^{12}n^6 - 2e^6m^2n^{14} + 2e^6m^4n^{12} + 13e^6m^6n^{10} + 8e^6m^8n^8 - 3e^6m^{10}n^6 \\
&- 3e^8m^2n^{12} + 3e^8m^4n^{10} + 8e^8m^6n^8 + 2e^8m^8n^6 - 2e^{10}m^2n^{10} - 3e^{10}m^4n^8 - e^{12}m^2n^8 \\
&- m^8n^{16} - m^{10}n^{14} - e^2m^4n^{18} - 3e^2m^6n^{16} - 6e^2m^8n^{14} - e^2m^{10}n^{12} - 3e^2m^{12}n^{10} \\
&+ e^4m^4n^{16} - 10e^4m^6n^{14} - 5e^4m^8n^{12} - 7e^4m^{10}n^{10} - 3e^4m^{12}n^8 - 7e^6m^4n^{14} + 2e^6m^6n^{12} \\
&- e^6m^8n^{10} - 2e^6m^{10}n^8 - 2e^6m^{12}n^6 - 2e^8m^2n^{14} - 4e^8m^4n^{12} + 2e^8m^6n^{10} + 3e^8m^8n^8 \\
&- 2e^8m^{10}n^6 - 2e^{10}m^2n^{12} - 6e^{10}m^4n^{10} - 3e^{10}m^6n^8 - 3e^{12}m^4n^8 - 3m^{10}n^{16} - m^{12}n^{14} \\
&- m^{14}n^{12} - e^2m^6n^{18} - 7e^2m^8n^{16} - 10e^2m^{10}n^{14} - 5e^2m^{12}n^{12} - 2e^2m^{14}n^{10} - 15e^4m^6n^{16} \\
&- 33e^4m^8n^{14} - 22e^4m^{10}n^{12} - 13e^4m^{12}n^{10} - e^4m^{14}n^8 - 2e^6m^4n^{16} - 21e^6m^6n^{14} \\
&- 21e^6m^8n^{12} - 16e^6m^{10}n^{10} - 4e^6m^{12}n^8 - 11e^8m^4n^{14} - 15e^8m^6n^{12} - 15e^8m^8n^{10} \\
&- 6e^8m^{10}n^8 - 2e^8m^{12}n^6 - e^{10}m^2n^{14} - 9e^{10}m^4n^{12} - 8e^{10}m^6n^{10} - 5e^{10}m^8n^8 - e^{10}m^{10}n^6 \\
&- 3e^{12}m^6n^8 - e^{14}m^4n^8 + \mathcal{O}([emn]^{28})\,.
\end{aligned}
\tag{B.2}
$$

It is worth noting that the multi-graded Hilbert series lacks certain properties of the ungraded Hilbert series, such as a matching number of factors in the denominator with physical observables and a palindromic form in the numerator. In addition, as already mentioned in footnote 9, there is ambiguity when determining the form of the ungraded Hilbert series, similar ambiguity can arise for the multi-graded Hilbert series. For instance, it is possible to introduce a common factor $(1+e^2n^2)$ to both the numerator and denominator, then the factors $(1-e^2n^2)^2$ in denominator become $(1-e^2n^2)(1-e^4n^4)$, the numerator will also change accordingly. The form of the Hilbert series is less constrained in the multi-graded case. This ambiguity prevent us from further exploring the invariant structures of the theory. Instead of decoding the Hilbert series, since the PL has unique forms for both ungraded and multi-graded Hilbert series, the analyzing of the PL is more helpful in our case. The invariants in our theory form a non-complete intersection ring, the PL is non-terminating as a result, according to Eq. (2.10), the PL can only be calculated up to some given order of spurions. However, as we discussed in Sec. 3.2, we should only care about the terms in PL up to the first order, where all terms are negative. We find that this occurs at order $\mathcal{O}([emn]^{26})$. The PL up to this order is given as follows

$$
\begin{aligned}
\mathrm{PL}(e,m,n) = &\left(e^2+m^2+n^2\right) + \left(e^4+m^4+n^4+e^2n^2+m^2n^2\right) + \left(e^6+m^6+n^6+e^2n^4\right. \\
&\left.+e^4n^2+2m^2n^4+m^4n^2+e^2m^2n^2\right) + \left(e^4n^4+m^2n^6+3m^4n^4+3e^2m^2n^4+e^2m^4n^2\right.
\end{aligned}
$$

$$
\begin{aligned}
&+e^4m^2n^2\big)+\big(m^2n^8+2m^4n^6+2m^6n^4+5e^2m^2n^6+4e^2m^4n^4+4e^4m^2n^4+e^4m^4n^2\big)\\
&+\big(e^6n^6+3m^4n^8+3m^6n^6+m^8n^4+5e^2m^2n^8+9e^2m^4n^6+3e^2m^6n^4+7e^4m^2n^6\\
&+6e^4m^4n^4+2e^6m^2n^4\big)+\big(m^4n^{10}+3m^6n^8+m^8n^6+3e^2m^2n^{10}+11e^2m^4n^8\\
&+8e^2m^6n^6+e^2m^8n^4+9e^4m^2n^8+14e^4m^4n^6+5e^4m^6n^4+6e^6m^2n^6+3e^6m^4n^4\\
&+e^8m^2n^4\big)+\big(m^4n^{12}+m^6n^{10}+e^2m^2n^{12}+8e^2m^4n^{10}+8e^2m^6n^8+3e^2m^8n^6\\
&+6e^4m^2n^{10}+20e^4m^4n^8+13e^4m^6n^6+2e^4m^8n^4+8e^6m^2n^8+13e^6m^4n^6\\
&+3e^6m^6n^4+3e^8m^2n^6+2e^8m^4n^4\big)+\big(2e^2m^4n^{12}+3e^4m^2n^{12}+11e^4m^4n^{10}\\
&+11e^4m^6n^8+5e^4m^8n^6+5e^6m^2n^{10}+18e^6m^4n^8+13e^6m^6n^6+e^6m^8n^4+4e^8m^2n^8\\
&+8e^8m^4n^6+2e^8m^6n^4+e^{10}m^2n^6-3m^8n^{10}-3m^{10}n^8-3e^2m^6n^{10}-4e^2m^8n^8\\
&-e^2m^{10}n^6\big)+\big(+2e^6m^6n^8+4e^6m^8n^6+9e^8m^4n^8+7e^8m^6n^6+e^8m^8n^4+3e^{10}m^4n^6\\
&-m^6n^{14}-9m^8n^{12}-8m^{10}n^{10}-4m^{12}n^8-2e^2m^4n^{14}-21e^2m^6n^{12}-32e^2m^8n^{10}\\
&-14e^2m^{10}n^8-e^2m^{12}n^6-9e^4m^4n^{12}-28e^4m^6n^{10}-18e^4m^8n^8-2e^4m^{10}n^6\\
&-e^6m^4n^{10}\big)+\big(3e^8m^8n^6+2e^{10}m^6n^6-m^6n^{16}-10m^8n^{14}-18m^{10}n^{12}-9m^{12}n^{10}\\
&-2m^{14}n^8-2e^2m^4n^{16}-34e^2m^6n^{14}-76e^2m^8n^{12}-55e^2m^{10}n^{10}-12e^2m^{12}n^8\\
&-27e^4m^4n^{14}-103e^4m^6n^{12}-109e^4m^8n^{10}-39e^4m^{10}n^8-2e^4m^{12}n^6-2e^6m^2n^{14}\\
&-45e^6m^4n^{12}-83e^6m^6n^{10}-35e^6m^8n^8-3e^6m^{10}n^6-6e^8m^2n^{12}-21e^8m^4n^{10}\\
&-9e^8m^6n^8-4e^{10}m^2n^{10}-e^{10}m^4n^8\big)+\big(e^{10}m^8n^6-e^{12}n^{12}-11m^8n^{16}-20m^{10}n^{14}\\
&-21m^{12}n^{12}-5m^{14}n^{10}-m^{16}n^8-e^2m^4n^{18}-33e^2m^6n^{16}-110e^2m^8n^{14}-116e^2m^{10}n^{12}\\
&-48e^2m^{12}n^{10}-5e^2m^{14}n^8-31e^4m^4n^{16}-174e^4m^6n^{14}-284e^4m^8n^{12}-162e^4m^{10}n^{10}\\
&-33e^4m^{12}n^8-e^6m^2n^{16}-87e^6m^4n^{14}-261e^6m^6n^{12}-226e^6m^8n^{10}-59e^6m^{10}n^8\\
&-2e^6m^{12}n^6-8e^8m^2n^{14}-94e^8m^4n^{12}-134e^8m^6n^{10}-48e^8m^8n^8-2e^8m^{10}n^6\\
&-10e^{10}m^2n^{12}-32e^{10}m^4n^{10}-15e^{10}m^6n^8-3e^{12}m^2n^{10}-2e^{12}m^4n^8\big)+\big(-5m^8n^{18}\\
&-19m^{10}n^{16}-16m^{12}n^{14}-7m^{14}n^{12}-m^{16}n^{10}-22e^2m^6n^{18}-105e^2m^8n^{16}\\
&-149e^2m^{10}n^{14}-88e^2m^{12}n^{12}-20e^2m^{14}n^{10}-e^2m^{16}n^8-21e^4m^4n^{18}-192e^4m^6n^{16}\\
&-424e^4m^8n^{14}-371e^4m^{10}n^{12}-130e^4m^{12}n^{10}-13e^4m^{14}n^8-96e^6m^4n^{16}-448e^6m^6n^{14}\\
&-603e^6m^8n^{12}-303e^6m^{10}n^{10}-45e^6m^{12}n^8-5e^8m^2n^{16}-161e^8m^4n^{14}-417e^8m^6n^{12}\\
&-314e^8m^8n^{10}-70e^8m^{10}n^8-2e^8m^{12}n^6-12e^{10}m^2n^{14}-108e^{10}m^4n^{12}-140e^{10}m^6n^{10}\\
&-40e^{10}m^8n^8-e^{10}m^{10}n^6-8e^{12}m^2n^{12}-23e^{12}m^4n^{10}-9e^{12}m^6n^8-e^{14}m^2n^{10}-e^{14}m^4n^8\big)\\
&+\mathcal{O}\left([emn]^{28}\right),
\end{aligned}
\tag{B.3}
$$

where the terms are grouped by parentheses at each order. We can see that the terms in $\mathcal{O}([emn]^{26})$ are all negative.

Under the Dirac limit, the Hilbert series can be obtained by setting $m \to 0$, which will have a very simple form as has been found for the quark sector in Ref [26]

$$
\mathcal{H}(e,n) = \frac{1+e^6n^6}{(1-e^2)\,(1-e^4)\,(1-e^6)\,(1-n^2)\,(1-n^4)\,(1-n^6)\,(1-e^2n^2)\,(1-e^4n^2)\,(1-e^2n^4)\,(1-e^4n^4)}\,. \tag{B.4}
$$

The ungraded Hilbert series is given by

$$\mathcal{H}(t) = \frac{1+t^{12}}{(1-t^2)^2(1-t^4)^3(1-t^6)^4(1-t^8)}\,. \tag{B.5}$$

The Dirac case corresponds to a complete intersection ring, and the multi-graded PL has finite number of terms, which are given as follows

$$\mathrm{PL}(e,n) = e^2+e^4+e^6+n^2+n^4+n^6+e^2n^2+e^4n^2+e^2n^4+e^4n^4+e^6n^6-e^{12}n^{12}\,. \tag{B.6}$$

The corresponding ungraded PL can be obtained by setting $e, n \to t$, which has the following form

$$\mathrm{PL}(t) = 2t^2+3t^4+4t^6+t^8+t^{12}-t^{24}\,, \tag{B.7}$$

where the positive terms correctly capture the 10 CP-even and 1 CP-odd invariants, while the negative term indicates there is a syzygy at order 24.

B.2 Model with $n_N = n_f = 2$

For completeness, we also show the Hilbert series for the case $n_N = n_f = 2$, which has already been presented in Ref. [26]. The numerator and denominator are given by

$$\begin{aligned}\mathcal{N}(e,m,n) =& 1+2e^2m^2n^4+m^4n^4+e^2m^4n^4+e^4m^4n^4+e^2m^2n^6+e^4m^2n^6-e^2m^6n^6+\\ &-e^4m^6n^6-e^2m^4n^8-e^4m^4n^8-e^6m^4n^8-2e^4m^6n^8-e^6m^8n^{12}\,,\\ \mathcal{D}(e,m,n) =& \left(1-e^2\right)\left(1-e^4\right)\left(1-m^2\right)\left(1-m^4\right)\left(1-n^2\right)\left(1-e^2n^2\right)\left(1-m^2n^2\right)\times\\ &\left(1-e^2m^2n^2\right)\left(1-n^4\right)\left(1-m^2n^4\right)\left(1-e^4m^2n^4\right)\,.\end{aligned} \tag{B.8}$$

The ungraded Hilbert series is

$$\mathcal{H}(t) = \frac{1+t^6+3t^8+2t^{10}+3t^{12}+t^{14}+t^{20}}{(1-t^2)^3(1-t^4)^5(1-t^6)(1-t^{10})}\,. \tag{B.9}$$

The multi-graded PL is given by

$$\begin{aligned}\mathrm{PL}(e,m,n) =& e^2+m^2+n^2+e^4+m^4+e^2n^2+m^2n^2+n^4+e^2m^2n^2+m^2n^4+\\ &+2e^2m^2n^4+m^4n^4+e^4m^2n^4+e^2m^4n^4+e^2m^2n^6+e^4m^4n^4+\\ &+e^4m^2n^6-e^2m^6n^6-e^2m^4n^8-\mathcal{O}([emn]^{16})\,.\end{aligned} \tag{B.10}$$

The PL also has a non-terminating series. Therefore, the theory has a non-complete intersection ring. However, the pure negative order appears at $\mathcal{O}([emn]^{14})$, the 18 lower order terms in PL are all positive, and they correspond to the basic set.

B.3 Model with $n_N = 2,\ n_f = 3$

For the case of $n_N = 2, n_f = 3$, the Hilbert series has been calculated in Ref. [30]. The numerator and denominator of the multi-graded Hilbert series is given by

$$\begin{aligned}\mathcal{N}(e,m,n) =& 1-e^2n^2+e^4n^4+2e^2m^2n^4+2e^4m^2n^4+2e^6m^2n^4+m^4n^4+e^2m^4n^4+2e^4m^4n^4+\\ &+e^6m^4n^4+e^8m^4n^4+e^2m^2n^6+e^{10}m^2n^6-e^2m^4n^6-e^4m^4n^6-3e^6m^4n^6+\end{aligned}$$

$$
\begin{aligned}
&- e^8m^4n^6 - e^{10}m^4n^6 - e^2m^6n^6 - 2e^4m^6n^6 - 2e^6m^6n^6 - 2e^8m^6n^6 - e^{10}m^6n^6 + \\
&- e^4m^2n^8 - e^6m^2n^8 - e^8m^2n^8 - e^{10}m^2n^8 - e^{12}m^2n^8 - e^2m^4n^8 - e^4m^4n^8 + \\
&- 3e^6m^4n^8 - e^8m^4n^8 - 3e^{10}m^4n^8 - e^{12}m^4n^8 - e^{14}m^4n^8 - e^4m^6n^8 + e^6m^6n^8 + \\
&- e^8m^6n^8 + e^{10}m^6n^8 - e^{12}m^6n^8 + e^6m^8n^8 + e^8m^8n^8 + e^{10}m^8n^8 - e^{10}m^2n^{10} + \\
&+ e^4m^4n^{10} + e^6m^4n^{10} + 2e^8m^4n^{10} + 2e^{12}m^4n^{10} + e^{14}m^4n^{10} + e^{16}m^4n^{10} + \\
&+ 2e^6m^6n^{10} + 3e^{10}m^6n^{10} + 2e^{14}m^6n^{10} - 2e^8m^6n^{12} - 3e^{12}m^6n^{12} - 2e^{16}m^6n^{12} + \\
&- e^6m^8n^{12} - e^8m^8n^{12} - 2e^{10}m^8n^{12} - 2e^{14}m^8n^{12} - e^{16}m^8n^{12} - e^{18}m^8n^{12} + \\
&+ e^{12}m^{10}n^{12} - e^{12}m^4n^{14} - e^{14}m^4n^{14} - e^{16}m^4n^{14} + e^{10}m^6n^{14} - e^{12}m^6n^{14} + \\
&+ e^{14}m^6n^{14} - e^{16}m^6n^{14} + e^{18}m^6n^{14} + e^8m^8n^{14} + e^{10}m^8n^{14} + 3e^{12}m^8n^{14} + \\
&+ e^{14}m^8n^{14} + 3e^{16}m^8n^{14} + e^{18}m^8n^{14} + e^{20}m^8n^{14} + e^{10}m^{10}n^{14} + e^{12}m^{10}n^{14} + \\
&+ e^{14}m^{10}n^{14} + e^{16}m^{10}n^{14} + e^{18}m^{10}n^{14} + e^{12}m^6n^{16} + 2e^{14}m^6n^{16} + 2e^{16}m^6n^{16} + \\
&+ 2e^{18}m^6n^{16} + e^{20}m^6n^{16} + e^{12}m^8n^{16} + e^{14}m^8n^{16} + 3e^{16}m^8n^{16} + e^{18}m^8n^{16} + \\
&+ e^{20}m^8n^{16} - e^{12}m^{10}n^{16} - e^{20}m^{10}n^{16} - e^{14}m^8n^{18} - e^{16}m^8n^{18} - 2e^{18}m^8n^{18} + \\
&- e^{20}m^8n^{18} - e^{22}m^8n^{18} - 2e^{16}m^{10}n^{18} - 2e^{18}m^{10}n^{18} - 2e^{20}m^{10}n^{18} - e^{18}m^{12}n^{18} + \\
&+ e^{20}m^{12}n^{20} - e^{22}m^{12}n^{22}\,,
\end{aligned}
\tag{B.11}
$$

$$
\begin{aligned}
\mathcal{D}(e,m,n) = {} & \left(1-e^2\right)\left(1-m^2\right)\left(1-n^2\right)\left(1-e^4\right)\left(1-m^4\right)\left(1-n^4\right)\left(1-e^2n^2\right)^2 \times \\
& \left(1-m^2n^2\right)\left(1-e^6\right)\left(1-e^2n^4\right)\left(1-e^4n^2\right)\left(1-m^2n^4\right) \times \\
& \left(1-e^2m^2n^2\right)\left(1-e^4m^2n^2\right)\left(1-e^4m^2n^4\right)\left(1-e^8m^2n^4\right)\,,
\end{aligned}
$$

which corresponds to the ungraded forms

$$
\begin{aligned}
\mathcal{N}(t) = {} & 1 + t^2 + t^4 + 2t^6 + 6t^8 + 10t^{10} + 18t^{12} + 23t^{14} + 28t^{16} + 31t^{18} + 34t^{20} + 32t^{22} + \\
& + 34t^{24} + 31t^{26} + 28t^{28} + 23t^{30} + 18t^{32} + 10t^{34} + 6t^{36} + 2t^{38} + t^{40} + t^{42} + t^{44}\,, \\
\mathcal{D}(t) = {} & \left(1-t^2\right)^2\left(1-t^4\right)^5\left(1-t^6\right)^4\left(1-t^8\right)\left(1-t^{10}\right)\left(1-t^{14}\right)\,.
\end{aligned}
\tag{B.12}
$$

The multi-graded PL is given by

$$
\begin{aligned}
\mathrm{PL}(e,m,n) = {} & e^2 + m^2 + n^2 + e^4 + m^4 + e^2n^2 + m^2n^2 + n^4 + e^6 + e^4n^2 + e^2m^2n^2 + e^2n^4 + \\
& + m^2n^4 + e^4m^2n^2 + e^4n^4 + 2e^2m^2n^4 + m^4n^4 + 3e^4m^2n^4 + e^2m^4n^4 + e^2m^2n^6 + \\
& + 2e^6m^2n^4 + 2e^4m^4n^4 + e^6n^6 + 2e^4m^2n^6 + e^8m^2n^4 + e^6m^4n^4 + 2e^6m^2n^6 + \\
& - e^2m^6n^6 - e^2m^4n^8 + e^8m^4n^4 + 2e^8m^2n^6 - e^6m^4n^6 - 2e^4m^6n^6 - e^6m^2n^8 + \\
& - 5e^4m^4n^8 - 2e^2m^6n^8 - m^8n^8 + e^{10}m^2n^6 - 2e^6m^6n^6 - e^8m^2n^8 - 8e^6m^4n^8 + \\
& - 6e^4m^6n^8 - e^2m^8n^8 - e^6m^2n^{10} - 2e^4m^4n^{10} - e^2m^6n^{10} - \mathcal{O}([emn]^{20})\,.
\end{aligned}
\tag{B.13}
$$

The PL is also non-terminating, thus lead to a non-complete intersection ring. The pure negative order appears at $\mathcal{O}([emn]^{20})$, and the basic set can be obtained up to $\mathcal{O}([emn]^{18})$ as shown in Ref. [30].

C Algorithms

C.1 Algorithm for the construction of invariants

The flavor invariants can be easily constructed according to the invariant graph. Taking the quark sector as an example, one can refer to the graph in the lower left panel of Fig. 2 to determine the general form of flavor invariants, denoted by $\mathrm{Tr}\big(X_u^m X_d^n X_u^k \ldots\big)$ with $X_u \equiv Y_u Y_u^\dagger$ and $X_d \equiv Y_d Y_d^\dagger$. Although this will result in an infinite number of flavor invariants, there are some identities of $n \times n$ matrices that can be utilized to reduce them, such as the Cayley-Hamilton theorem. For instance, for 3×3 matrices, the identity is given by

$$A^3 = A^2\,\mathrm{Tr}(A) - \frac{1}{2}A\left[\mathrm{Tr}(A)^2 - \mathrm{Tr}\big(A^2\big)\right] + \frac{1}{6}\left[\mathrm{Tr}(A)^3 - 3\,\mathrm{Tr}\big(A^2\big)\,\mathrm{Tr}(A) + 2\,\mathrm{Tr}\big(A^3\big)\right]\mathbb{1}_{3\times 3}\,, \tag{C.1}$$

which will reduce the power of the matrices X_u and X_d to a maximum of 3. Following the discussion in Ref. [26], another identity can be derived from the Cayley-Hamilton theorem, which is given by

$$\begin{aligned} 2\,\mathrm{Tr}(ABAC) = {} & \mathrm{Tr}(A)^2\,\mathrm{Tr}(B)\,\mathrm{Tr}(C) - \mathrm{Tr}(BC)\,\mathrm{Tr}(A)^2 - 2\,\mathrm{Tr}(AB)\,\mathrm{Tr}(A)\,\mathrm{Tr}(C) + \\ & - 2\,\mathrm{Tr}(AC)\,\mathrm{Tr}(A)\,\mathrm{Tr}(B) + 2\,\mathrm{Tr}(ABC)\,\mathrm{Tr}(A) + 2\,\mathrm{Tr}(ACB)\,\mathrm{Tr}(A) + \\ & - \mathrm{Tr}\big(A^2\big)\,\mathrm{Tr}(B)\,\mathrm{Tr}(C) + 2\,\mathrm{Tr}(AB)\,\mathrm{Tr}(AC) + \mathrm{Tr}\big(A^2\big)\,\mathrm{Tr}(BC) + \\ & + 2\,\mathrm{Tr}(C)\,\mathrm{Tr}\big(A^2B\big) + 2\,\mathrm{Tr}(B)\,\mathrm{Tr}\big(A^2C\big) - 2\,\mathrm{Tr}\big(A^2BC\big) - 2\,\mathrm{Tr}\big(A^2CB\big)\,. \end{aligned} \tag{C.2}$$

Note that by using the Cayley-Hamilton theorem, we can derive more general identities with 4 or more different matrices entering the theorem. This particular identity in Eq. (C.2) ensures that the terms present in the trace in the quark sector will not exceed a length of 4. Referring to Ref. [26] for a detailed explanation, it can be concluded that with the two identities above there are only 11 invariants without polynomial relations among them that can be constructed with the help of the invariant graphs in the quark sector of the SM. They are given by

$$\begin{aligned} & I_{2,0} = \mathrm{Tr}(X_u),\ I_{0,2} = \mathrm{Tr}(X_d),\ I_{4,0} = \mathrm{Tr}\big(X_u^2\big),\ I_{2,2} = \mathrm{Tr}(X_u X_d),\ I_{0,4} = \mathrm{Tr}\big(X_d^2\big), \\ & I_{6,0} = \mathrm{Tr}\big(X_u^3\big),\ I_{4,2} = \mathrm{Tr}\big(X_u^2 X_d\big),\ I_{2,4} = \mathrm{Tr}\big(X_u X_d^2\big),\ I_{0,6} = \mathrm{Tr}\big(X_d^3\big), \\ & I_{4,4} = \mathrm{Tr}\big(X_u^2 X_d^2\big),\ I_{6,6}^{(-)} = \mathrm{Tr}\big(X_u^2 X_d^2 X_u X_d\big) - \mathrm{Tr}\big(X_d^2 X_u^2 X_d X_u\big)\,. \end{aligned} \tag{C.3}$$

These invariants form a generating set of all quark invariants, i.e., all flavor invariants in quark sector can be written as a polynomial of these 11 invariants. Furthermore, the first 10 invariants form a primary set in the sense that they are algebraically independent invariants which capture all physical parameters in the theory.

Additionally, there exists a separate conjugate graph in the quark sector shown in the lower right panel of Fig. 2. It is worth mentioning that the two graphs are disconnected in the quark sector, which make it easy to construct invariants within separate flavor objects. Furthermore, as both matrices X_u and X_d are Hermitian, the identity $\mathrm{Tr}\big(X_u^{m*} X_d^{n*} X_u^{k*} X_d^{l*}\big) = \mathrm{Tr}\Big(X_d^{l\dagger} X_u^{k\dagger} X_d^{n\dagger} X_u^{m\dagger}\Big) = \mathrm{Tr}\big(X_d^l X_u^k X_d^n X_u^m\big)$ can be used to show that all flavor invariants constructed from the conjugate graph are included in the first graph.

If neutrinos are Dirac particles, the Majorana mass term is absent and the structure of the graph for the simplified νSM will be analogous to that of the quark sector. Then, similar flavor invariants can be easily obtained by setting $X_d \to X_e \equiv Y_e Y_e^\dagger$ and $X_u \to X_N \equiv Y_N Y_N^\dagger$ in Eq. (C.3). The inclusion of the Majorana mass term gives rise to a connected graph, as illustrated in the top panel of Fig. 2, which offers additional flavor structures for the construction of invariants. We have found that by just adding M_N to the 'Dirac-like' νSM, the number of basic invariants explodes from 11 to 459. The main difference from the quark sector is that certain invariants may become complicated and cannot be reduced through the identity presented in Eq. (C.2). It is noteworthy that this identity is applicable only when the flavor objects A, B and C can form a single trace invariant by themselves (i.e. they transform in the adjoint of one of the factors of $U(3)^3$ presented in Tab. 1), which may not always be the case for the invariants generated by the current graph. Consequently, there exists no universal form for flavor invariants in the νSM. In this regard, our objective is to exhaustively construct invariants up to order 26 by brute force, and subsequently reduce linear redundancies as much as possible by using the above two identities.

To systematically construct flavor invariants in the lepton sector, closed paths that permit the repetition of vertices and edges in a graph, known as "walks" in mathematical jargon, can be sought out. To achieve single trace invariants, these walks must be closed. As a result, our algorithm looks for closed walks in the graph. The closed walk is denoted by a directional chain of vertices that it traverses (see Eq. (3.6) for example). We have further shorten the walk notation by removing arrows (see Sec. 3.2 for detail). Therefore, all single trace flavor invariants can be presented by integers. For instance, the invariants in Eq. (3.6) will be denoted by $\mathrm{Tr}\left(Y_u Y_u^\dagger Y_d Y_d^\dagger\right) \sim 7898$ and $\mathrm{Tr}\left(Y_N M_N^* M_N M_N^* M_N Y_N^\dagger\right) \sim 234343$.

By extending one step further at each vertex, we can generate higher order walks from a given walk, for instance, extending the walk 1232 leads to

$$\{1\underline{2}1232,\ 123\underline{2}32,\ 12\underline{1}232,\ 123\underline{4}32,\ 123\underline{2}32,\ 123232,\ 123\underline{1}2\underline{2}\}\,, \tag{C.4}$$

where the underlined vertices are the ones that have been added. By using the cycling degree of freedom of the walks or of the single trace invariants, for instance, $121232 \sim 123212$, we can identify the repeated walks, and the independent walks are determine to be

$$\{121232,\ 123232,\ 123432\}\,. \tag{C.5}$$

With this algorithm, higher order walks derived from a specific walk will be exhaustively generated. As a result, starting from all possible lowest order walks, the extending algorithm is capable of constructing all walks up to a given order. The initial walks correspond to the order two flavor invariants, which are given by

$$\{12,\ 23,\ 34,\ 45,\ 56\}\,. \tag{C.6}$$

Running our algorithm up to order 26 with these initial walks results in 516101 walks. However, there are many redundancies. The first obvious redundancy stems from the graph's symmetry property. Since there are two conjugate parts in the graph, we find that the walk denoted by the labeled fields $W_1 \equiv ij\ldots kl \sim F_i F_j \ldots F_k F_l$ is always accompanied by

two walks, one with conjugate fields in reversed order $W_2 \equiv l'k'\ldots j'i' \sim F_l^* F_k^* \ldots F_j^* F_i^*$, and another one with conjugate fields $W_3 \equiv i'j'\ldots k'l' \sim F_i^* F_j^* \ldots F_k^* F_l^*$, where the primed vertex v' can be mapped to normal vertex v with $v' = 7 - v$ for the graph in the lepton sector. The walks W_1 and W_2 correspond to the same invariants according to the trace identity $\mathrm{Tr}\big(A^T\big) = \mathrm{Tr}(A)$, for instance, $\mathrm{Tr}\big(Y_e^* Y_e^T\big) = \mathrm{Tr}\Big(Y_e Y_e^\dagger\Big)$. While the walks W_1 and W_3 correspond to two conjugate invariants $\mathrm{Tr}(X)$ and $\mathrm{Tr}(X^*)$, and their linear combinations $\mathrm{Tr}(X)+\mathrm{Tr}(X^*)$ and $\mathrm{Tr}(X)-\mathrm{Tr}(X^*)$ form the CP-even and CP-odd invariants respectively. Equivalently, we can introduce the real and imaginary components of $\mathrm{Tr}(X)$ as our convention, i.e., $\mathrm{Tr}(X) = \mathcal{I} + i\,\mathcal{J} = \mathrm{Re}\,\mathrm{Tr}(X) + i\,\mathrm{Im}\,\mathrm{Tr}(X)$, where $\mathcal{I} = \mathrm{Re}\,\mathrm{Tr}(X)$ and $\mathcal{J} = \mathrm{Im}\,\mathrm{Tr}(X)$ are the CP-even and CP-odd invariants. For instance, the walks $W_1 = 23434543$ and $W_3 = 23454343$ will generate the invariants $\mathrm{Tr}(X) \equiv \mathrm{Tr}\Big(Y_N^\dagger Y_N M_N^* M_N M_N^* Y_N^T Y_N^* M_N\Big)$ and its conjugate $\mathrm{Tr}(X^*) \equiv \mathrm{Tr}\Big(Y_N^T Y_N^* M_N M_N^* M_N Y_N^\dagger Y_N M_N^*\Big)$, but the conjugate one does not provide additional information, making the real and imaginary parts of $\mathrm{Tr}(X)$ as two independent degrees of freedom for invariants. As a result, the conjugate walk W_3 will be removed from our list. These two types of redundancies can be easily identified from the structures of the walks. These reduced walks will then be converted into invariants that can be simplified by applying some matrix identities, allowing us to effectively eliminate more redundancies.

As mentioned previously, when using Eq. (C.1) and Eq. (C.2), we need to confirm that the matrices A, B and C must have the same transformation rules under the flavor group, and can form single trace invariants. We refer to such sequences of flavor matrices as adjoint objects. We note that if one flavor matrix appears consecutively in an invariant for $2n$ times ($n \geq 1$), then the corresponding flavor matrix sequence must be an adjoint object. If it appears $2n+1$ times consecutively, then there must be a $2n$ sub-sequence among them correspond to adjoint object. For example, the underlined sequences in the invariant $\mathrm{Tr}\Big(\underset{\sim}{Y_N Y_N^\dagger} Y_N \underline{M_N^* M_N} Y_N^\dagger \underset{\sim}{Y_N Y_N^\dagger} \underline{\underline{Y_e Y_e^\dagger}}\Big)$ are adjoint objects. We find two of them are identical, as a result, we can use Eq. (C.2) to eliminate this invariant as it can be represented by other invariants that are already present in our invariant list. In this invariant, the matrix products $Y_N Y_N^\dagger$, $Y_N M_N^* M_N Y_N^\dagger$ and $Y_e Y_e^\dagger$ are identified as A, B and C respectively in Eq. (C.2). It is evident that they share the same adjoint transformation properties, allowing the construction of single trace flavor invariants that accurately match the identity. If we come across an adjoint object with six or more flavor matrices in an invariant, then identity in Eq. (C.1) can be easily applied, and the invariant will be redundant. For example, we can remove the invariant $\mathrm{Tr}\Big(\underset{\sim}{Y_e Y_e^\dagger Y_e Y_e^\dagger Y_e Y_e^\dagger} \underline{Y_N Y_N^\dagger}\Big)$ because there is a sequence with six flavor matrices. The identity in Eq. (C.1) cannot eliminate $\mathrm{Tr}\big(A^3\big)$. Therefore, our invariant list will include expressions such as $\mathrm{Tr}\Big(Y_e Y_e^\dagger Y_e Y_e^\dagger Y_e Y_e^\dagger\Big)$.

By applying these two identities, our invariant list is further reduced, but there are still many more identities that can be used. For instance, setting $C \to AC$ in Eq. (C.2) yields

the following identity

$$
\begin{aligned}
&3\,\mathrm{Tr}\big(A^2BAC\big)+3\,\mathrm{Tr}\big(ABA^2C\big)=2\,\mathrm{Tr}(A)^3\,\mathrm{Tr}(B)\,\mathrm{Tr}(C)-2\,\mathrm{Tr}(A)^3\,\mathrm{Tr}(BC)+\\
&-3\,\mathrm{Tr}(A)^2\,\mathrm{Tr}(C)\,\mathrm{Tr}(AB)-3\,\mathrm{Tr}(A)^2\,\mathrm{Tr}(B)\,\mathrm{Tr}(AC)+3\,\mathrm{Tr}(A)^2\,\mathrm{Tr}(ABC)+\\
&+3\,\mathrm{Tr}(A)^2\,\mathrm{Tr}(ACB)-3\,\mathrm{Tr}\big(A^2\big)\,\mathrm{Tr}(A)\,\mathrm{Tr}(B)\,\mathrm{Tr}(C)+3\,\mathrm{Tr}\big(A^2\big)\,\mathrm{Tr}(A)\,\mathrm{Tr}(BC)+\\
&+3\,\mathrm{Tr}(A)\,\mathrm{Tr}(C)\,\mathrm{Tr}\big(A^2B\big)+3\,\mathrm{Tr}(A)\,\mathrm{Tr}(B)\,\mathrm{Tr}\big(A^2C\big)-3\,\mathrm{Tr}(A)\,\mathrm{Tr}\big(A^2BC\big)+\\
&-3\,\mathrm{Tr}(A)\,\mathrm{Tr}\big(A^2CB\big)+\mathrm{Tr}\big(A^3\big)\,\mathrm{Tr}(B)\,\mathrm{Tr}(C)-\mathrm{Tr}\big(A^3\big)\,\mathrm{Tr}(BC)+\\
&+3\,\mathrm{Tr}\big(A^2B\big)\,\mathrm{Tr}(AC)+3\,\mathrm{Tr}(AB)\,\mathrm{Tr}\big(A^2C\big),
\end{aligned}
\tag{C.7}
$$

which indicates one of $\mathrm{Tr}\big(A^2BAC\big)$ and $\mathrm{Tr}\big(ABA^2C\big)$ is redundant, but our algorithm does not eliminate either. We can further remove invariants using more complex identities derived from Cayley-Hamilton theorem at higher orders, but the identities could be nontrivial and difficult to use. Furthermore, while the identities presented in this section are for general matrices, our flavor matrices and their combinations could have special properties. For instance, the invariant may include a Hermitian matrix $Y_eY_e^\dagger$ or a symmetric matrix M_N, additional identities beyond those for general 3×3 matrices could be found in some cases. Therefore, we need a general algorithm to identify all potential identities, which should encompass both regular identities for eliminating redundant invariants and polynomial relations between non-redundant invariants (referred to as syzygies in the context of invariant). The general algorithm becomes computationally expensive when the invariant list is large, with the help of the simple matrix identities, the redundancies have been effectively removed. After we perform CP-even and CP-odd decompositions for the invariants, the number of invariants drops to 8666. Then we will pass this pre-reduced invariant set to the numerical algorithm to further remove linear redundancies and get the basic set. The whole process is presented by the flow graph in Fig. 1.

C.2 Algorithm for the construction of a generating set

As by definition any invariant in the set can be expressed as a polynomial of the generating invariants, the process of identifying the generating set entails eliminating linear redundancies from the invariant list. We have used matrix identities to reduce the invariant list in advance in the previous elimination process as we have discussed in App. C.1. In this section, we will introduce the numerical algorithm that can identify all linear relations and syzygies. The linear relations will remove all linear redundancies, and will result in a generating set of invariants, while the syzygies, alongside the linear relations, will be generated as a byproduct, and their count serves as a verification of our algorithm when compared with the PL.

Our algorithm is based on the fact that the relations of invariants follow the graded algebra, which means every term in the relations should have the same degree of the spurions. So in order to construct the linear relations or the syzygies, we need to consider all possible combinations that lead to the specific degree from our invariant list. We create a graded list $\{d_i:\{f_i\}\}$, where d_i is the degree in terms of (e,m,n), and $\{f_i\}$ is a homogeneous list of all (combinations of) invariants at degree d_i. We start from the lowest degrees, and the

graded list is given by

$$\{e^2 : \{\mathrm{Tr}\big(Y_e Y_e^\dagger\big)\},\ m^2 : \{\mathrm{Tr}(M_N M_N^*)\},\ n^2 : \{\mathrm{Tr}\Big(Y_N Y_N^\dagger\Big)\}\}\,, \tag{C.8}$$

which will be extended degree by degree. At higher degrees d_i, the homogeneous list $\{f_i\}$ can be constructed by two possible sources, which includes the invariants at order d_i from our invariant list, and all combinations of invariants $\{f_j\}$ and $\{f_k\}$ in the graded list satisfying $d_j d_k = d_i$. Once the homogeneous list $\{f_i\}$ is formed, the general equation for the polynomial identity should have the form

$$\sum_{i=1}^{n} c_i f_i = 0\,, \tag{C.9}$$

where f_i is the element in $\{f_i\}$, n is its count, and the coefficient c_i is the parameter we want to determine by numerical method. Since it is a polynomial relation, c_i can be chosen to be integer in general. The invariant (combinations) f_i is parameterized in terms of general matrices $Y_{e,N}$ and symmetric matrix M_N. In order to find the solutions to the coefficients $\{c_i\}$, a numerical algorithm is introduced. We generate n linear equations of $\{c_i\}$ with n set of random inputs for the flavor matrices[19], the coefficient matrix M is taken as input by the `Mathematica` function **NullSpace**, where the rows represent the set of random inputs and the columns correspond to the homogeneous list $\{f_i\}$. The function will give a list of vectors $V \equiv \{v_i\}$ that forms a basis for the null space of the matrix M, as a result, V is the basis solutions of $\{c_i\}$, and the identities are given by

$$\sum_{i=1}^{n} V_{ji} f_i = 0 \quad \text{for} \quad j = 1, \ldots, |V|\,. \tag{C.10}$$

The identities that involve single invariants (not products of invariants) will be called linear relations, which will be used to remove redundancies from our invariant list. But sometimes it is difficult to use the solution V to determine which invariant should be removed, since the relations of single invariants could be correlated, in such cases, the `Mathematica` function **RowReduce** can be applied to generate the row echelon form of V, which will be used to disentangle the linear relations. Here is an illustrative example: we have a homogeneous list consisting of $\{f_i\} = \{s_1, s_2, s_3, p_1, p_2\}$, with $s_{1,2,3}$ as single invariants and $p_{1,2}$ as products of invariants. Suppose they lead to a coefficient matrix M that results in a null space V and can be further reduced to the row echelon form V',

$$M = \begin{pmatrix} 2 & 1 & -3 & -1 & 1 \\ -1 & 2 & -1 & -2 & 2 \\ -3 & 3 & 0 & -3 & 3 \\ -2 & 4 & -2 & -4 & 4 \\ -4 & 5 & -1 & -5 & 5 \end{pmatrix} \xrightarrow{\textbf{NullSpace}} V = \begin{pmatrix} 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix} \xrightarrow{\textbf{RowReduce}} V' = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}. \tag{C.11}$$

[19] In order to reduce the accidental error from the random inputs, we can generate more than n equations to form a larger linear system, or increase the randomness of the inputs

From the null space matrix V, we can find following identities,

$$s_2 = p_2, \quad s_2 = -p_1, \quad s_1 = -s_2 - s_3\,. \tag{C.12}$$

In order to disentangle the dependence between them, we have to perform additional operations, while for the identities generate by V', they have clear forms,

$$s_1 = -s_3 - p_2, \quad s_2 = p_2, \quad p_1 = -p_2\,, \tag{C.13}$$

which directly determine s_1, s_2 and p_1 should be redundant, and they correspond to the columns that has leading 1 in the row echelon form V'. The first two identities tell us s_1, s_2 should be removed from our invariant list, and s_3 should be included in our generating set as a basic invariant. The first identity can also be written as $s_3 = -s_1 - p_2$, then s_1 and s_3 will change their characters, this can be understood as ambiguities in determining basic invariants. The third identity contains only products of invariants, which should be identified as a syzygy. Although this syzygy does not identify any redundant invariant, it does signify the product of invariants p_1 is redundant, and we should remove it from our homogeneous list, and result in $\{f_i\} = \{s_3, p_2\}$. It is crucial to remove p_1 from the homogeneous list in our algorithm, as keeping it would result in redundant syzygies at higher orders and prevent us from obtaining the correct number of syzygies. For instance, if p_1 is not removed, at degree $2d_i$, the syzygy $s_3 p_1 = -s_3 p_2$ will be found. Although it looks trivial in this example, in real analysis of our flavor invariants, the redundant syzygy could have a complicated form, which make it difficult to identify whether it is redundant.

Once the discussion on degree d_i is complete and the reduced graded list $\{d_i : \{f_i\}\}$ has been updated, the analysis should be repeated for all other invariants at higher degrees until we finish the scan of all invariants. All redundancies will be removed, and the invariant list will be reduced to the basic set. While scanning, it is important to keep a record of the number of basic invariants n_b and the number of syzygies n_s at each degree, the difference $n_b - n_s$ should match the corresponding term in the PL.

In the real analysis of our invariants, there are some difficulties, as already mentioned in Sec. 3.2, we will run into "redundant syzygies". They are partially removed by the homogeneous list reduction at each degree as described above, however this reduction does not cover all redundancies. Suppose we have two redundant products f_a and f_b at degree d_i, and the syzygies are denoted by the homogeneous linear function V_a and V_b on the homogeneous list $\{f_{i\neq a,b}\}$,

$$f_a = s_1 s_2^2 = V_a(\{f_{i\neq a,b}\}), \quad f_b = s_1^2 s_2 = V_b(\{f_{i\neq a,b}\})\,, \tag{C.14}$$

where s_1, s_2 are invariants at degree $d_i/3$. At higher degree, we can easily find a syzygy

$$s_1 V_a(\{f_{i\neq a,b}\}) - s_2 V_b(\{f_{i\neq a,b}\}) = 0\,, \tag{C.15}$$

since it is a linear combination of lower degree syzygies, it must be redundant. Depending on the ordering scheme of the homogeneous list, such redundant syzygies could first appear at different degrees. In our algorithm, the first redundant syzygy appears at order 26, and

we find more syzygies than expected from the PL. In order to remove such redundancies, we should record the lower degree syzygies, the linear independent redundant syzygies can be generated at a specific higher degree, the number of which should be subtracted at given degree when counting the syzygies. Although we did not initially incorporate this into our algorithm, we have verified the validity of this method by checking multiple redundant syzygies at certain degrees. Furthermore, our objective is to determine the generating set. It should not pose any issues to run our algorithm without verifying the number of syzygies.

Another issue is related to the complexity of numerical calculations. Our algorithm take random flavor matrices as inputs. When the invariants become highly complex, the matrix calculations will result in extremely large integers. This will generate a complicated linear system that will take considerable time to solve. The `Mathematica` function **RowReduce** provides a **Modulus** option, by setting it to an appropriate large prime number, the massive integers in the linear system are mapped to finite fields, which will considerably reduce the computational complexity, and our algorithm can speed up by several times. There are also some packages, such as `FiniteFlow` [56], `FiniteFieldSolve` [57], which are frameworks that can be used to solve the linear system efficiently, interested readers may refer to these packages.

C.3 Algorithm towards finding a minimal CPV set

Reducing the CP odd basic invariants to a minimal set capturing all conditions for CPV in the theory – as the Jarlskog invariant does in the SM – can become involved as the flavor structure of the theory becomes more complicated. While it is still possible to find such a set for 2 generations of fermions, this task seems bound to fail from the beginning for the 251 CP odd invariants in the case of 3 generations. In this section, we introduce a numerical algorithm to reduce the CP odd part of the generating set of a theory into a more minimal CPV set. To illustrate our algorithm, we will work with a simple example with only 2 generations of fermions in the theory defined in Sec. 2.1. We use the parameterization shown in Eqs. (A.17,A.18).

C.3.1 Mimimal CPV set for $n_f = n_N = 2$

In Ref. [26], the authors find the following 6 CP-odd invariants in the basic set

$$\begin{aligned}
J_1 &= \mathrm{ImTr}\left(M_N Y_N^\dagger Y_N Y_N^\dagger Y_e Y_e^\dagger Y_N M_N^*\right) \sim \mathcal{J}_2\\
J_2 &= \mathrm{ImTr}\left(M_N^* M_N Y_N^\dagger Y_N M_N^* Y_N^T Y_N^* M_N\right) \sim \mathcal{J}_1\\
J_3 &= \mathrm{ImTr}\left(M_N^* M_N Y_N^\dagger Y_e Y_e^\dagger Y_N M_N^* Y_N^T Y_N^* M_N\right) \sim \mathcal{J}_7\\
J_4 &= \mathrm{ImTr}\left(M_N Y_N^\dagger Y_N Y_N^\dagger Y_e Y_e^\dagger Y_N M_N^* Y_N^T Y_N^*\right) \sim \mathcal{J}_5\\
J_5 &= \mathrm{ImTr}\left(M_N M_N^* M_N Y_N^\dagger Y_e Y_e^\dagger Y_N M_N^* Y_N^T Y_e^* Y_e^T Y_N^*\right) \sim \mathcal{J}_{31}\\
J_6 &= \mathrm{ImTr}\left(M_N Y_N^\dagger Y_N Y_N^\dagger Y_e Y_e^\dagger Y_N M_N^* Y_N^T Y_e^* Y_e^T Y_N^*\right) \sim \mathcal{J}_{28}
\end{aligned} \tag{C.16}$$

which we have translated into our notation. This set is sufficient to capture all CPV effects in the 2-generation theory but not all invariants might necessarily have to be included to

Cases ↓ / Invariants →	J_1	J_2	J_3	J_4	J_5	J_6
$y_e \to 0$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$m_1 \to 0$	$\emptyset$	0	0	$\emptyset$	0	$\emptyset$
$y_1 \to y_2$	0	0	0	0	$\emptyset$	0
$\theta \to 0$	0	$\emptyset$	$\emptyset$	0	$\emptyset$	0
$\theta \to \frac{\pi}{2}$	0	$\emptyset$	$\emptyset$	0	$\emptyset$	0
$\phi \to 0$	0	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 5: Part of the unsorted full zero table for the given example corresponding to the cases given in Eq. (C.17).

capture all CPV effects. This is easy to check for a given physical spectrum, but if we want a set that works for all zero masses, texture zeros and all other special cases of the spectrum which increases the symmetry this becomes more involved.

The first step is to build all possible cases that can simplify the mass spectrum, for instance

$$\{y_e \to 0, m_1 \to 0, y_1 \to y_2, \theta \to 0, \theta \to \frac{\pi}{2}, \phi \to 0, \dots\}\,, \tag{C.17}$$

including also the combination of two or more of them applied at the same time. These cases can be captured by the conditions lead to vanishing or unphysical phases, which are discussed in Sec. H. We will then build what we call a *zero table*, which has an entry '0' when the invariant vanishes for the given case and an entry '$\emptyset$' if the invariant is non-zero which we check by plugging in random numbers for the parameters appearing in the invariants. The result for the given cases above can be found in Tab. 5.

In a second step we take all rows from this table which have at least one non-vanishing entry and combine the case in this row with another case from the full list of special cases in Eq. (C.17). We repeat this procedure until the spectrum is sufficiently simplified, such that all invariants vanish in all cases of the given iteration. In the given example this will happen when 6 cases of the list in Eq. (C.17) are combined which will lead to a zero table with only zeros for all invariants in all possible cases. Finally, we combine all zero tables into one big table.

Once we have obtained those zero tables we can analyze the data. Our strategy is as follows. We start with an empty set which acts as the candidate set for the minimal CPV set and then go through all cases in the zero table in a systematic way to check which invariant has to be added in order to capture all the physical phases in each simplified case.

In order to do this in a systematic way, we first sort the rows, i.e. the cases, of the zero table such that the rows with the least amount of non-zero entries come first in the table. These are the cases which are most likely to allow us to decide if an invariant should be included or not. For instance, if there is only one non-zero invariant in a given case and there exists one physical phase in the simplified spectrum, then this invariant for sure has to be added to our minimal set in order to capture all CPV effects.

Furthermore, we sort the columns, i.e. the invariants, of the zero table such that the

Cases ↓ / Invariants →	J_5	J_2	J_3	J_6	J_4	J_1
$y_1 \to y_2$	$\emptyset$	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$y_e \to 0,\ y_\mu \to 0$	0	$\emptyset$	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$m_2 \to m_1$,$\alpha \to 0$	0	0	0	$\emptyset$	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\alpha \to 0$	$\emptyset$	0	0	$\emptyset$	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$m_1 \to m_2$	0	0	0	$\emptyset$	$\emptyset$	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 6: Part of the sorted full zero table for the given example. All dots correspond to entries for different cases which give the same zero pattern as the previous case. The last case in the table is the last case with only 2 non-vanishing with a unique zero pattern. After that all entries can be explained by the previous ones. Hence, we can find a minimal set.

columns with the most non-zero entries come first. This allows us to add the invariant which is non-zero in the most other cases to the candidate set first, making it more likely that this invariant will also capture physical phases in many other subsequent cases and that we will eventually find the most minimal CPV set.

After the sorting is done we go through the table row by row and perform the following procedure. First, we calculate the rank of the Jacobian of the invariants w.r.t. the CP-odd variables in the theory in a given case, giving us the number of physical phases in that particular case. Then, we also calculate the rank of the Jacobian of the invariants in the candidate set w.r.t. the CP-odd variables in the theory for the given case. If they are the same we go to the next row. If they are unequal we check which invariants are non-zero in the given case and are not already part of the candidate set. From the invariants obtained in this way we first add the one which appears in the most left column according to our previous sorting as it is the most likely to increase the rank of the Jacobian of the candidate set in many subsequent cases. If adding this invariant increases the rank we keep it, otherwise we go to the next invariant which is non-vanishing in the case and not already in our candidate set. After an invariant is found which increases the rank of the set we check again if the rank of the Jacobian of the candidate set is equal to the one of the CP-odd part of the basic set, both evaluated on the given case. For equal ranks, we can continue examining the next row. Otherwise we have to add an (several) other invariant(s) to the candidate set until equality of both ranks is reached in the given special case.

Following this algorithm allows us to reduce the set of CP-odd flavor invariants in the simple example from 6 to only 4. The minimal CPV set we find is

$$\mathcal{S}_{\text{min}}^{\text{CPV}} = \{J_2, J_4, J_5, J_6\}\,, \tag{C.18}$$

i.e. this is the smallest set that has one non-vanishing CP-odd invariant for each physical phase in the theory, for all possible values of the flavorful couplings in the theory.

There are several ways to make simplifications in this brute force method. For instance, as soon as one combination of cases leads to a row of only zeroes in the zero table, one no longer has to consider it in the next iterations where all possible combinations of this case are built with all cases in Eq. (C.17). This is because these combinations will of course also only lead to zero entries. This can help reduce the large combinatorics of the problem.

There exists however also a caveat to our approach. Some of the CP-odd invariants in the neutrino models are extremely complicated polynomials that have complicated roots that do not correspond to points that increases the exact flavor symmetry of the theory. Hence, these cases are not covered by the symmetry-based zero table approach. In every algorithm that tries to reduce a CP-odd generating set to a (more) minimal CPV set, one therefore has to check if there are still enough non-zero invariants in the minimal CPV set on these points to capture all physical phases in the theory. This is of course extremely unpractical for complicated theories like the νSM, where this can happen already for two generations of fermions. However, these points are single, extremely specific points in parameter space. Thus, we are convinced that they can be ignored in most analyses and that our algorithm can still be helpful. One such case for the νSM with 2 generations of right-handed neutrinos is the last case in Eq. (H.2), which cannot be explained by an enlarged flavor symmetry group.

In the following, we will give two more examples of applying the zero table algorithm to the SM extended with 2 generations of sterile neutrinos and the 2HDM discussed in Ref. [27]. To our knowledge this is only possible by solving the CP-odd invariants of the theory explicitly and examining each of the roots one by one.

C.3.2 Minimal CPV set for $n_N = 2,\ n_f = 3$

We can also apply our algorithm to the case of only 2 generations of sterile neutrinos which has been treated in Ref. [30]. There, the authors find a set of 38 generating flavor invariants out of which 18 are CP-odd (c.f. Eq. (D.3)). Running our algorithm on this CP-odd generating set allows us to reduce the set to only 7 CP-odd invariants which capture all sufficient conditions for CP violation in all degenerate cases. In the notation introduced in Eq. (D.3), the minimal set we find is given by

$$\mathcal{S}_{\text{min}}^{\text{CPV}} = \{\mathcal{J}_1, \mathcal{J}_5, \mathcal{J}_{10}, \mathcal{J}_{28}, \mathcal{J}_{31}, \mathcal{J}_{132}, \mathcal{J}_{195}\}\,. \tag{C.19}$$

Our algorithm is however not able to find a set of invariants that captures only the necessary conditions of CP violation. In App. C.3, we discuss an example where we can show that our algorithm is not able to fully reduce the CP-odd invariants of the generating set to a minimal CP-odd set. We have also determined a minimal CPV set for the case where all fermions just come in two generations. We treat this case in App. C.3 to illustrate how the algorithm that reduces a CP-odd generating set to a minimal CPV set works.

D List of invariants

In this section, we enumerate the 459 basic invariants comprising the generating set. We use the walk-based notation to represent these invariants as introduced in Sec. C.1. Therefore,

a single trace invariant is denoted by an integer. A prefix "R@" or "I@" means that the real or imaginary part of the trace should be taken. For pure numbers, the single trace itself is inherently real. Consequently, invariants with "I@" prefixes are identified as CP-odd invariants, while all the other remeaning ones are categorized as CP-even. There are in total 208 CP-even and 251 CP-odd invariants. The complete list we found reads as follows

CP-even set: {12, 34, 23, 1212, 3434, 2323, 1232, 2343, 121212, 343434, 232323, 123232, 121232, 234543, 232343, 234343, 123432, 12123232, 23234543, 23234343, R@23434543, 12345432, R@12323432, 12343432, 12123432, 2323454543, R@2323434543, R@2343434543, 1234545432, R@1232345432, R@1232323432, R@1234345432, R@1232343432, 1212345432, 1234565432, R@1212323432, 1212343432, R@232343454543, R@232343434543, R@123234545432, R@123232345432, R@123434545432, R@123234345432, R@123234543432, R@123232343432, R@123434345432, 121234545432, R@121232345432, R@121232323432, R@123234565432, R@121234345432, R@121232343432, R@123434565432, 121234565432, R@23234343454543, R@12323234545432, R@12323434545432, R@12323454543432, R@12323234345432, R@12323234543432, R@12343434545432, R@12323434345432, R@12323454343432, R@12123234545432, R@12123232345432, R@12323454565432, R@12323234565432, R@12123434545432, R@12123234345432, R@12123234543432, R@12123232343432, R@12323434565432, R@12323456543432, R@12123434345432, R@12343434565432, R@12123456545432, R@12123234565432, R@12123434565432, 12123456565432, R@1232323434545432, R@1232323432345432, R@1232323454543432, R@1232343434545432, R@1232345454343432, R@1232323434345432, R@1232343434543432, R@1212323234545432, R@1232323454565432, R@1212323434545432, R@1212323454543432, R@1212323234345432, R@1212323234543432, R@1232343454565432, R@1232345434565432, R@1232323434565432, R@1232323456543432, R@1212343434545432, R@1212323434345432, R@1212323454343432, R@1232343434565432, R@1232345654343432, R@1212323456545432, R@1212323454565432, R@1212323234565432, R@1212343456545432, R@1212343454565432, R@1212321234345432, R@1212323434565432, R@1212323456543432, R@1212343434565432, R@1212323456565432, R@1212343456565432, R@123232343234545432, R@123232343434545432, R@123234343454543432, R@123232345654543232, R@121232323434545432, R@121232323432345432, R@121232323454543432, R@123232343454565432, R@123232343234565432, R@123232345434565432, R@121232343434545432, R@121232345454343432, R@121232323434345432, R@123234343454565432, R@123234543434565432, R@123232343434565432, R@121232343434543432, R@123234343456543432, R@121232323456545432, R@121232323454565432, R@121232123434545432, R@121232343456545432, R@121232343454565432, R@121232343212345432, R@121232345654345432, R@121232345434565432, R@121232321234345432, R@121232323434565432, R@121232323456543432, R@121234343456545432, R@121234343454565432, R@121232123434345432, R@121232343434565432, R@121232343212343432, R@121232345654343432, R@121232123456545432, R@121232323456565432, R@121232123434565432, R@121232343456565432, R@121232345656543432, R@121234343456565432, R@12123232343234545432, R@12323234323454565432, R@12123232343434545432, R@12323234343454565432, R@12123234343454543432, R@12323434345434565432, R@12123232345654543232, R@12123212345434545432, R@12123232123434545432, R@12123232343454565432, R@12123232343234565432, R@12123232345654345432, R@12123232345434565432, R@12123212343434545432, R@12123234343454565432, R@12123234343212345432, R@12123234565434345432, R@12123234543434565432, R@12123232123434345432, R@12123232343434565432, R@12123434345654345432, R@12123212343434543432, R@12123234343456543432, R@12123212345654545432, R@12123212343456545432, R@12123212343454565432, R@12123212345654345432, R@12123232123434565432, R@12123232343456565432, R@12123232345656543432, R@12123212343434565432, R@12123234343456565432, R@12123234565654343432, R@12123212343456565432, R@1212323234323454565432, R@1212321234543434545432, R@1212323234343454565432, R@1212321234343454543432, R@1212323434345434565432, R@1212321234345654545432, R@1212321234345454565432, R@1212321234343456545432, R@1212321234343454565432, R@1212321234565434345432, R@1212323212343434565432, R@1212323234343456565432, R@1212321234343456543432, R@1212323434345656543432, R@1212321234345656545432, R@1212321234565654345432, R@1212321234343456565432, R@121232123434345654345432, R@121232123434345656545432, R@121232123456565434345432, R@121232123434345656543432} ,

CP-odd set: {I@23434543, I@12323432, I@2323434543, I@2343434543, I@1232345432,
I@1232323432, I@1234345432, I@1232343432, I@1212323432, I@121232123232, I@232343454543,
I@232343234543, I@232343434543, I@232343234343, I@234343454343, I@123234545432,
I@123232343232, I@123232345432, I@123434545432, I@123432345432, I@123234345432,
I@123234543432, I@123232343432, I@123434345432, I@123432343432, I@121232345432,
I@121232323432, I@123234565432, I@121234345432, I@121232343432, I@123434565432,
I@121232123432, I@23234323454543, I@23234343454543, I@23234343234543, I@23234343454343,
I@12323234543232, I@12323234545432, I@12343234545432, I@12343232345432, I@12323434545432,
I@12323454543432, I@12323234343232, I@12323234345432, I@12323234543432, I@12343434545432,
I@12343432345432, I@12343232343432, I@12323434345432, I@12323454343432, I@12343434543432,
I@12123234545432, I@12123232343232, I@12123232345432, I@12323454565432, I@12323234565432,
I@12123434545432, I@12123432345432, I@12123234345432, I@12123234543432, I@12123232343432,
I@12343234565432, I@12323434565432, I@12323456543432, I@12123434345432, I@12123432343432,
I@12343434565432, I@12123456545432, I@12123212345432, I@12123234565432, I@12123232123432,
I@12123434565432, I@12123212343432, I@2323454323454543, I@2323434323454543, I@2323434345454343,
I@1232323454543232, I@1234543234545432, I@1232345434545432, I@1232323434545432,
I@1232323432345432, I@1232323454543432, I@1234343234545432, I@1234343232345432,
I@1232343434545432, I@1232345454343432, I@1232323434345432, I@1234343454543432,
I@1232343434543432, I@1212323234543232, I@1212323234545432, I@1232323456543232,
I@1232323454565432, I@1212343234545432, I@1212343232345432, I@1212323434545432,
I@1212323454543432, I@1212323234343232, I@1212323234345432, I@1212323234543432,
I@1234565432345432, I@1232343454565432, I@1232345654345432, I@1232345434565432,
I@1232323434565432, I@1232323456543432, I@1212343434545432, I@1212343432345432,
I@1212343232343432, I@1212323434345432, I@1212323454343432, I@1234343234565432,
I@1232343434565432, I@1232345654343432, I@1212343434543432, I@1234343456543432,
I@1212321234545432, I@1212323456545432, I@1212323454565432, I@1212323212345432,
I@1212323234565432, I@1212343456545432, I@1212343454565432, I@1212343212345432,
I@1212343234565432, I@1212321234345432, I@1212323434565432, I@1212323456543432,
I@1212323212343432, I@1212343434565432, I@1212343212343432, I@1212321234565432,
I@1212323456565432, I@1212343456565432, I@123454323234545432, I@123232343234545432,
I@1232345434345454 32, I@123232343434545432, I@123234343454543432, I@121232323454543232,
I@123232345654543232, I@121234543234545432, I@121232345434545432, I@121232323434545432,
I@121232323432345432, I@121232323454543432, I@123456543232345432, I@123232343454565432,
I@123232343234565432, I@123232345434565432, I@121234343234545432, I@121234343232345432,
I@121232343434545432, I@121232345454343432, I@121232323434345432, I@123234343454565432,
I@123234565434345432, I@123234543434565432, I@123232343434565432, I@121234343454543432,
I@121232343434543432, I@123234343456543432, I@121232321234545432, I@121232323456543232,
I@121232323456545432, I@121232323454565432, I@121234321234545432, I@121234565432345432,
I@121232123434545432, I@121232343456545432, I@121232343454565432, I@121232345654345432,
I@121232345654543432, I@121232345434565432, I@121232321234345432, I@121232323434565432,
I@121232323456543432, I@121234343456545432, I@121234343454565432, I@121234343212345432,
I@121234343234565432, I@121232123434345432, I@121232343434565432, I@121232345654343432,
I@121234343456543432, I@121232123456545432, I@121232321234565432, I@121232323456565432,

I@121234321234565432, I@121234323456565432, I@121232123434565432, I@121232343456565432, I@121232345656543432, I@121234343456565432, I@121232123456565432, I@12123454323234545432, I@12123232343234545432, I@12323234323454565432, I@12123234543434545432, I@12123232343434545432, I@12123234343454543432, I@12123232345654543232, I@12123456543234545432, I@12123456543232345432, I@12123232123434545432, I@12123232343454565432, I@12123232343234565432, I@12123232345654345432, I@12123232345434565432, I@12123434321234545432, I@12123212343434545432, I@12123234343454565432, I@12123234565434345432, I@12123234565454343432, I@12123234543434565432, I@12123232343434565432, I@12123434345654345432, I@12123234343456543432, I@12123212345654545432, I@12123232345656543232, I@12123456565432345432, I@12123456543212345432, I@12123212343456545432, I@12123212343454565432, I@12123234565654345432, I@12123232123434565432, I@12123232343456565432, I@12123232345656543432, I@12123434321234565432, I@12123434323456565432, I@12123212343434565432, I@12123234343456565432, I@12123234565654343432, I@12123434345656543432, I@12123212345656545432, I@12123432123456565432, I@12123212343456565432, I@12123232343234565432, I@12123456565432321234565432, I@12123212343456545432, I@12123212343434565432, I@12123234565434345432, I@12123232343434565432, I@12123212343454565432, I@12123456565432321234565432, I@12123212343434565432, I@12123234565654343432, I@12123212343456565432, I@12123232343434565432, I@12123456565432321234565432, I@12123212343456565432, I@12123456565432321234565432}.

The generating set is a union of CP-even set and CP-odd set. We have introduced the notation $\mathcal{I}_i(\mathcal{J}_i)$ in the main text to represent the ith invariant in the CP-even(CP-odd) set. The invariants in above lists can be easily converted to `Mathematica` expression with the **convert** function defined below

```
convert[walk_]:=Block[{map,head,num,c,t,ct},
  map={12->ct@Ye,21->Ye,23->Yn,32->ct@Yn,34->c@Mn,43->Mn,54->c@Yn,45->t@Yn,
  65->t@Ye,56->c@Ye}/.{c->Conjugate,t->Transpose,ct->ConjugateTranspose};
  head=Switch[Head@walk,R,Re@Tr@#&,I,Im@Tr@#&,_,Tr];
  num=IntegerDigits@Cases[{walk},_Integer,2][[1]];
  head[Dot@@FromDigits/@Partition[num,2,1,1]/.map]];
```

For example, evaluating **convert**[R@23434543] in `Mathematica` will generate the expression of a CP-even flavor invariant $\mathrm{Re}\,\mathrm{Tr}\Big(Y_N M_N^* M_N M_N^* Y_N^T Y_N^* M_N Y_N^\dagger\Big)$.

We have summarized the number of basic invariants, divided into the number of CP-even and CP-odd basic invariants, as well as the number of syzygies at each degree and total order in Tabs. (7, 8). Although the numbers in these tables at every degree are fixed, there are still degrees where ambiguity arises in determining the form of the invariants. This ambiguity occurs when our algorithm identifies multiple linear invariants within a single linear relation. Such linear relation can simply arise from the Cayley-Hamilton theorem as shown in Eqs. (C.2, C.7). We also discuss the ambiguity in App. C.2 around Eq. (C.13). Let us now present a specific example. At degree e^4n^4, according to the summary table, there is only one basic invariant. However, there exists another invariant $\mathcal{I} = \mathrm{Tr}\Big(Y_e Y_e^\dagger Y_N Y_N^\dagger Y_e Y_e^\dagger Y_N Y_N^\dagger\Big)$, which is not included in the basic set. We can find following linear relation

$$2\mathcal{I} + 4\mathcal{I}_{18} - \mathcal{I}_1^2\mathcal{I}_3^2 + \mathcal{I}_1^2\mathcal{I}_6 + 4\mathcal{I}_1\mathcal{I}_3\mathcal{I}_7 - 4\mathcal{I}_1\mathcal{I}_{12} - 2\mathcal{I}_7^2 + \mathcal{I}_3^2\mathcal{I}_4 - \mathcal{I}_4\mathcal{I}_6 - 4\mathcal{I}_3\mathcal{I}_{13} = 0\,. \tag{D.1}$$

In this relation, two linear invariants, namely $\mathcal{I}$ and $\mathcal{I}_{18}$, are observed. Both of these invariants are of degree e^4n^4. The ambiguity arises because it is possible to select either of them as the basic invariant, rendering the other one redundant.

We find that our generating set contains some subsets that can be used as generating set of other theories. For example, for the theory with $n_N = n_f = 2$, the generating set can be formed with the invariants

$$\text{Gen.}(n_N = n_f = 2): \{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, \mathcal{I}_4, \mathcal{I}_5, \mathcal{I}_6, \mathcal{I}_7, \mathcal{I}_8, \mathcal{I}_{14}, \mathcal{I}_{17}, \mathcal{I}_{22}, \mathcal{I}_{35}, \mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_5, \mathcal{J}_7, \mathcal{J}_{28}, \mathcal{J}_{31}\}, \tag{D.2}$$

which is also shown in Ref. [26] with a different convention. For the theory with $n_N = 2, n_f = 3$, the generating set is given by

$$\begin{aligned}\text{Gen.}(n_N = 2, n_f = 3): \{&\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, \mathcal{I}_4, \mathcal{I}_5, \mathcal{I}_6, \mathcal{I}_7, \mathcal{I}_8, \mathcal{I}_9, \mathcal{I}_{12}, \mathcal{I}_{13}, \mathcal{I}_{14}, \mathcal{I}_{17}, \mathcal{I}_{18}, \mathcal{I}_{22}, \mathcal{I}_{25}, \\ &\mathcal{I}_{34}, \mathcal{I}_{35}, \mathcal{I}_{54}, \mathcal{I}_{79}, \mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_5, \mathcal{J}_7, \mathcal{J}_9, \mathcal{J}_{10}, \mathcal{J}_{26}, \mathcal{J}_{28}, \mathcal{J}_{29}, \mathcal{J}_{31}, \mathcal{J}_{32}, \mathcal{J}_{68}, \mathcal{J}_{70}, \mathcal{J}_{72}, \\ &\mathcal{J}_{132}, \mathcal{J}_{133}, \mathcal{J}_{134}, \mathcal{J}_{195}\}.\end{aligned} \tag{D.3}$$

The above generating set has already been shown in Ref. [30], but the commutation notation is equivalently represented by taking imaginary part in our notation. We can also easily identify that the primary invariants shown in Eq. (3.13) correspond to the invariants

$$\text{Primary set}: \{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, \mathcal{I}_5, \mathcal{I}_6, \mathcal{I}_7, \mathcal{I}_8, \mathcal{I}_9, \mathcal{I}_{12}, \mathcal{I}_{13}, \mathcal{I}_{15}, \mathcal{I}_{23}, \mathcal{I}_{25}, \mathcal{I}_{34}, \mathcal{I}_{35}, \mathcal{I}_{47}, \mathcal{I}_{50}, \mathcal{I}_{54}, \mathcal{I}_{65}, \mathcal{I}_{79}, \mathcal{I}_{91}\}. \tag{D.4}$$

In the Dirac limit, the generating set will be reduced to have only 11 invariants, in which only one flavor invariant is CP-odd, and plays the same role as the Jarlskog invariant in quark sector. The invariants are given by

$$\text{Gen.(Dirac limit)}: \{\mathcal{I}_1, \mathcal{I}_3, \mathcal{I}_4, \mathcal{I}_6, \mathcal{I}_7, \mathcal{I}_9, \mathcal{I}_{11}, \mathcal{I}_{12}, \mathcal{I}_{13}, \mathcal{I}_{18}, \mathcal{J}_{10}\}. \tag{D.5}$$

Ord.	Degree	n_b		n_s
		n_e	n_o	
2	e^2	1	0	0
	m^2	1	0	0
	n^2	1	0	0
4	e^4	1	0	0
	m^4	1	0	0
	n^4	1	0	0
	e^2n^2	1	0	0
	m^2n^2	1	0	0
6	e^6	1	0	0
	m^6	1	0	0
	n^6	1	0	0
	e^2n^4	1	0	0
	e^4n^2	1	0	0
	m^2n^4	2	0	0
	m^4n^2	1	0	0
	$e^2m^2n^2$	1	0	0
8	e^4n^4	1	0	0
	m^2n^6	1	0	0
	m^4n^4	2	1	0
	$e^2m^2n^4$	2	1	0
	$e^2m^4n^2$	1	0	0
	$e^4m^2n^2$	1	0	0
10	m^2n^8	1	0	0
	m^4n^6	1	1	0
	m^6n^4	1	1	0
	$e^2m^2n^6$	3	2	0
	$e^2m^4n^4$	2	2	0
	$e^4m^2n^4$	3	1	0
	$e^4m^4n^2$	1	0	0
12	e^6n^6	0	1	0
	m^4n^8	1	2	0
	m^6n^6	1	2	0
	m^8n^4	0	1	0
	$e^2m^2n^8$	2	3	0
	$e^2m^4n^6$	4	5	0
	$e^2m^6n^4$	1	2	0
	$e^4m^2n^6$	4	3	0
	$e^4m^4n^4$	3	3	0
	$e^6m^2n^4$	1	1	0

Ord.	Degree	n_b		n_s
		n_e	n_o	
14	m^4n^{10}	0	1	0
	m^6n^8	1	2	0
	m^8n^6	0	1	0
	$e^2m^2n^{10}$	1	2	0
	$e^2m^4n^8$	4	7	0
	$e^2m^6n^6$	3	5	0
	$e^2m^8n^4$	0	1	0
	$e^4m^2n^8$	4	5	0
	$e^4m^4n^6$	6	8	0
	$e^4m^6n^4$	2	3	0
	$e^6m^2n^6$	2	4	0
	$e^6m^4n^4$	1	2	0
	$e^8m^2n^4$	1	0	0
16	m^4n^{12}	0	1	0
	m^6n^{10}	0	1	0
	m^8n^8	0	1	1
	$e^2m^2n^{12}$	0	1	0
	$e^2m^4n^{10}$	3	5	0
	$e^2m^6n^8$	3	5	0
	$e^2m^8n^6$	1	2	0
	$e^4m^2n^{10}$	2	4	0
	$e^4m^4n^8$	8	13	1
	$e^4m^6n^6$	5	8	0
	$e^4m^8n^4$	0	2	0
	$e^6m^2n^8$	3	5	0
	$e^6m^4n^6$	5	8	0
	$e^6m^6n^4$	1	2	0
	$e^8m^2n^6$	1	2	0
	$e^8m^4n^4$	1	1	0
18	m^8n^{10}	0	0	3
	$m^{10}n^8$	0	0	3
	$e^2m^4n^{12}$	1	2	1
	$e^2m^6n^{10}$	1	2	6
	$e^2m^8n^8$	1	1	6
	$e^2m^{10}n^6$	0	0	1
	$e^4m^2n^{12}$	1	2	0
	$e^4m^4n^{10}$	6	9	4
	$e^4m^6n^8$	6	9	4
	$e^4m^8n^6$	2	3	0

Ord.	Degree	n_b		n_s
		n_e	n_o	
18	$e^6m^2n^{10}$	2	4	1
	$e^6m^4n^8$	9	11	2
	$e^6m^6n^6$	6	7	0
	$e^6m^8n^4$	0	1	0
	$e^8m^2n^8$	2	3	1
	$e^8m^4n^6$	3	5	0
	$e^8m^6n^4$	1	1	0
	$e^{10}m^2n^6$	0	1	0
20	m^6n^{14}	0	0	1
	m^8n^{12}	0	0	9
	$m^{10}n^{10}$	0	0	8
	$m^{12}n^8$	0	0	4
	$e^2m^4n^{14}$	0	0	2
	$e^2m^6n^{12}$	0	0	21
	$e^2m^8n^{10}$	0	0	32
	$e^2m^{10}n^8$	0	0	14
	$e^2m^{12}n^6$	0	0	1
	$e^4m^4n^{12}$	2	3	14
	$e^4m^6n^{10}$	2	2	32
	$e^4m^8n^8$	2	1	21
	$e^4m^{10}n^6$	0	0	2
	$e^6m^2n^{12}$	1	1	2
	$e^6m^4n^{10}$	6	7	14
	$e^6m^6n^8$	7	7	12
	$e^6m^8n^6$	3	2	1
	$e^8m^2n^{10}$	1	2	3
	$e^8m^4n^8$	6	8	5
	$e^8m^6n^6$	3	5	1
	$e^8m^8n^4$	0	1	0
	$e^{10}m^2n^8$	0	1	1
	$e^{10}m^4n^6$	1	2	0
22	m^6n^{16}	0	0	1
	m^8n^{14}	0	0	10
	$m^{10}n^{12}$	0	0	18
	$m^{12}n^{10}$	0	0	9
	$m^{14}n^8$	0	0	2
	$e^2m^4n^{16}$	0	0	2
	$e^2m^6n^{14}$	0	0	34
	$e^2m^8n^{12}$	0	0	76

Table 7: Number of basic invariants n_b, divided into the number of CP-even and CP-odd basic invariants (n_e and n_o respectively), as well as the number of syzygies n_s at each degree and total order. Note that the difference $n_b - n_s$ can be mapped to the coefficient of the graded PL at each degree up to order 26.

Ord.	Degree	n_b		n_s
		n_e	n_o	
22	$e^2m^{10}n^{10}$	0	0	55
	$e^2m^{12}n^8$	0	0	12
	$e^4m^6n^{12}$	0	0	103
	$e^4m^4n^{14}$	0	0	27
	$e^4m^8n^{10}$	0	0	109
	$e^4m^{10}n^8$	0	0	39
	$e^4m^{12}n^6$	0	0	2
	$e^6m^4n^{12}$	1	1	47
	$e^6m^2n^{14}$	0	0	2
	$e^6m^6n^{10}$	2	0	85
	$e^6m^8n^8$	2	0	37
	$e^6m^{10}n^6$	0	0	3
	$e^8m^2n^{12}$	0	0	6
	$e^8m^4n^{10}$	2	3	26
	$e^8m^6n^8$	5	4	18
	$e^8m^8n^6$	2	1	0
	$e^{10}m^2n^{10}$	0	0	4
	$e^{10}m^4n^8$	2	2	5
	$e^{10}m^6n^6$	1	2	1
24	$e^{12}n^{12}$	0	0	1
	m^8n^{16}	0	0	11
	$m^{10}n^{14}$	0	0	20
	$m^{12}n^{12}$	0	0	21
	$m^{14}n^{10}$	0	0	5
	$m^{16}n^8$	0	0	1
	$e^2m^4n^{18}$	0	0	1
	$e^2m^6n^{16}$	0	0	33

Ord.	Degree	n_b		n_s
		n_e	n_o	
24	$e^2m^8n^{14}$	0	0	110
	$e^2m^{10}n^{12}$	0	0	116
	$e^2m^{12}n^{10}$	0	0	48
	$e^2m^{14}n^8$	0	0	5
	$e^4m^4n^{16}$	0	0	31
	$e^4m^6n^{14}$	0	0	174
	$e^4m^8n^{12}$	0	0	284
	$e^4m^{10}n^{10}$	0	0	162
	$e^4m^{12}n^8$	0	0	33
	$e^6m^2n^{16}$	0	0	1
	$e^6m^4n^{14}$	0	0	87
	$e^6m^6n^{12}$	0	0	261
	$e^6m^8n^{10}$	0	0	226
	$e^6m^{10}n^8$	0	0	59
	$e^6m^{12}n^6$	0	0	2
	$e^8m^2n^{14}$	0	0	8
	$e^8m^4n^{12}$	0	0	94
	$e^8m^6n^{10}$	0	0	134
	$e^8m^{10}n^6$	0	0	2
	$e^8m^8n^8$	1	0	49
	$e^{10}m^6n^8$	2	0	17
	$e^{10}m^8n^6$	1	0	0
	$e^{10}m^2n^{12}$	0	0	10
	$e^{10}m^4n^{10}$	0	0	32
	$e^{12}m^2n^{10}$	0	0	3
	$e^{12}m^4n^8$	0	1	3

Table 8: Tab. 7 continued. We only show the results up to order 24. There is no basic invariant at higher orders, only syzygies can be found. At order 26, the numbers of syzygies at each degree can be read off from the graded PL shown in Eq. (B.3). At higher orders, the PL loses its ability to explain the correct number of basic invariants and the number of syzygies. Please refer to the text around Eq. (3.12) for further details.

E CPV and CPC in the 2HDM example

E.1 Minimal CPV set of the 2HDM

Another interesting scenario for the application of basis invariants is the potential of a 2HDM

$$V(\Phi) = \Phi_a^\dagger Y^a{}_b \Phi^b + \Phi_a^\dagger \Phi_b^\dagger Z^{ab}{}_{cd} \Phi^c \Phi^d \tag{E.1}$$

where the two Higgs doublets transform into each other through global transformations in $SU(2)$ Higgs "flavor" space. This case was previously considered in Ref. [27], and the author found that the generating invariants are composed of 11 CP-even invariants and 8 CP-odd invariants. In this paper, we will adopt all the conventions and expressions of the invariants in Ref. [27].[20] The generating set is given as

$$\begin{aligned}\{&\mathcal{I}_{0,0,2},\, \mathcal{I}_{0,2,0},\, \mathcal{I}_{0,1,1},\, \mathcal{I}_{2,0,0},\, \mathcal{I}_{1,2,0},\, \mathcal{I}_{1,0,2},\, \mathcal{I}_{3,0,0},\, \mathcal{I}_{1,1,1},\, \mathcal{I}_{2,2,0},\, \mathcal{I}_{2,0,2},\, \mathcal{I}_{2,1,1},\\ &\mathcal{J}_{1,1,2},\, \mathcal{J}_{1,2,1},\, \mathcal{J}_{2,2,1},\, \mathcal{J}_{2,1,2},\, \mathcal{J}_{3,0,3},\, \mathcal{J}_{3,3,0},\, \mathcal{J}_{3,2,1},\, \mathcal{J}_{3,1,2}\}\,,\end{aligned} \tag{E.2}$$

where $\mathcal{I}_{i,j,k}$ ($\mathcal{J}_{i,j,k}$) represents the CP-even (CP-odd) invariant of degree (i,j,k). The grading is based on the symbols (q,y,t), which are used in the Hilbert series for the building blocks Y_3, Z_3, and Z_5 respectively in Ref. [27]. The author found that the eight generating invariants can be reduced to a set of four flavor invariants that can capture the necessary and sufficient conditions for CP violation.

Eventually, by running the zero table algorithm, we find that the following set is needed in order to capture all sources of CPV for all possible cases of mass spectra

$$\mathcal{S} = \{\mathcal{J}_{1,1,2}, \mathcal{J}_{1,2,1}, \mathcal{J}_{3,0,3}, \mathcal{J}_{3,3,0}, \mathcal{J}_{3,2,1}, \mathcal{J}_{3,1,2}\}\,. \tag{E.3}$$

Note, that this does not correspond to the set found in Ref. [27], where the necessary and sufficient conditions for CPC are found to be captured by $\{\mathcal{J}_{1,1,2}, \mathcal{J}_{1,2,1}, \mathcal{J}_{3,0,3}, \mathcal{J}_{3,3,0}\}$ (see App. E.2 for details). The two additional invariants are needed because we are asking a slightly different question here. The necessary and sufficient conditions are indeed captured by the roots of only four invariants. If one is however interested in capturing all physical sources of CPV, i.e. having as many algebraically independent CP-odd invariants as there are physical phases in the theory for all possible numerical values of the parameters in the theory, one is required to add two more invariants. One such case is $q_3 \to 0, q_{i2} \to 0, q_{r2} \to 0, t \to 0, y_i \to 0, y_r \to 0$, where the only two non-vanishing invariants simplify to the following expressions

$$\begin{aligned}\mathcal{J}_{1,1,2} &= -iy(2q_{r1}t_i t_r + q_{i,1}(t_i^2 - t_r^2)),\\ \mathcal{J}_{3,1,2} &= -iy(q_{i1}^2 + q_{r1}^2)(2q_{r1}t_i t_r + q_{i,1}(t_i^2 - t_r^2))\,.\end{aligned} \tag{E.4}$$

In this case, there are evidently two physical CP-odd parameters (q_{i1}, t_i) in the theory which require two CP-odd invariants to capture them. However, note that the two invariants only differ by a real factor. Therefore, even though the phases are independent, they can in principle be expressed by one CP-even and one CP-odd invariant (including the signs of the phases). We will further investigate this in a future publication.

[20] Due to the length of the invariants, we refer the reader to Ref. [27] for their explicit expressions.

E.2 Minimal CPC set of the 2HDM

In Ref. [27], the author found four syzygies that are relevant for determining the minimal CPC set. Two of them are given at order 8 with CP-odd terms

$$\begin{aligned} &3\mathcal{J}_{2,2,1}\mathcal{I}_{1,2,0} - \mathcal{J}_{3,2,1}\mathcal{I}_{0,2,0} + 3\mathcal{J}_{3,3,0}\mathcal{I}_{0,1,1} + \mathcal{J}_{1,2,1}\mathcal{I}_{2,2,0} = 0\,, \\ &3\mathcal{J}_{2,1,2}\mathcal{I}_{1,0,2} - \mathcal{J}_{3,1,2}\mathcal{I}_{0,0,2} + 3\mathcal{J}_{3,0,3}\mathcal{I}_{0,1,1} + \mathcal{J}_{1,1,2}\mathcal{I}_{2,0,2} = 0\,. \end{aligned} \tag{E.5}$$

Another two are given at order 10 with CP-even terms

$$\begin{aligned} &3\mathcal{J}_{2,2,1}^2 + 3\mathcal{J}_{1,1,2}\mathcal{J}_{3,3,0} - \mathcal{J}_{3,2,1}\mathcal{J}_{1,2,1} - \mathcal{I}_{2,0,0}\mathcal{J}_{1,2,1}^2 = 0\,, \\ &3\mathcal{J}_{2,1,2}^2 + 3\mathcal{J}_{1,2,1}\mathcal{J}_{3,0,3} - \mathcal{J}_{3,1,2}\mathcal{J}_{1,1,2} - \mathcal{I}_{2,0,0}\mathcal{J}_{1,1,2}^2 = 0\,. \end{aligned} \tag{E.6}$$

Note that these syzygies are symmetric under $y \leftrightarrow t$. The order-10 syzygies can already be presented in the form that follows Hilbert's Nullstellensatz, i.e., $\mathcal{J}_{2,2,1}$ and $\mathcal{J}_{2,1,2}$ are vanishing given that $\{\mathcal{J}_{1,2,1}, \mathcal{J}_{1,1,2}\} = 0$. Plug this in Eq. (E.5), by further requiring $\mathcal{J}_{3,3,0} = \mathcal{J}_{3,0,3} = 0$, one can find that $\mathcal{J}_{3,2,1}\mathcal{I}_{0,2,0} = \mathcal{J}_{3,1,2}\mathcal{I}_{0,0,2} = 0$. If $\mathcal{I}_{0,2,0} \neq 0 \neq \mathcal{I}_{0,0,2}$, we will have $\mathcal{J}_{3,2,1} = \mathcal{J}_{3,1,2} = 0$. If $\mathcal{I}_{0,2,0} = \mathcal{I}_{0,0,2} = 0$, by analyzing the spectrum of the parameter space, one can show that $\mathcal{J}_{3,2,1} = \mathcal{J}_{3,1,2} = 0$ is already implied. Therefore, the final minimal CPC set is chosen to be $\{\mathcal{J}_{1,2,1}, \mathcal{J}_{1,1,2}, \mathcal{J}_{3,3,0}, \mathcal{J}_{3,0,3}\}$, and the CP-odd invariants $\{\mathcal{J}_{2,2,1}, \mathcal{J}_{2,1,2}, \mathcal{J}_{3,2,1}, \mathcal{J}_{3,1,2}\}$ are vanishing given that the minimal CPC set is vanishing. This is the main results shown in Ref. [27].

The tricky part of the discussion in Ref. [27] is about the spectrum analysis, and it relies on the vanishing conditions of some invariants, such as $\mathcal{I}_{0,2,0} = \mathcal{I}_{0,0,2} = 0$, although it's easy to obtain in the 2HDM, it could be nontrivial for other theories. Hilbert's Nullstellensatz provides guidelines on the form of syzygies that can be used to eliminate invariants. Solving the vanishing conditions of specific invariants is not necessary, leading to a spectrum-free analysis. Following the numerical algorithm introduced in Ref. [27] or similar algorithm in App. C.2,[21] we can find following syzygies

$$\begin{aligned} 3\mathcal{J}_{2,2,1}^2 &= \mathcal{I}_{2,0,0}\mathcal{J}_{1,2,1}^2 - \mathcal{J}_{3,2,1}\mathcal{J}_{1,2,1} - 3\mathcal{J}_{1,1,2}\mathcal{J}_{3,3,0}\,, \\ 3\mathcal{J}_{2,1,2}^2 &= \mathcal{I}_{2,0,0}\mathcal{J}_{1,1,2}^2 - \mathcal{J}_{3,1,2}\mathcal{J}_{1,1,2} - 3\mathcal{J}_{1,2,1}\mathcal{J}_{3,0,3}\,, \\ \mathcal{J}_{3,2,1}^2 &= -\mathcal{I}_{2,0,0}\mathcal{J}_{1,2,1}\mathcal{J}_{3,2,1} + 2\mathcal{I}_{2,0,0}^2\mathcal{J}_{1,2,1}^2 - 3\mathcal{I}_{2,0,0}\mathcal{J}_{1,1,2}\mathcal{J}_{3,3,0} \\ &\quad - 18\mathcal{I}_{3,0,0}\mathcal{J}_{1,2,1}\mathcal{J}_{2,2,1} + 3\mathcal{J}_{3,3,0}\mathcal{J}_{3,1,2}\,, \\ \mathcal{J}_{3,1,2}^2 &= -\mathcal{I}_{2,0,0}\mathcal{J}_{1,1,2}\mathcal{J}_{3,1,2} + 2\mathcal{I}_{2,0,0}^2\mathcal{J}_{1,1,2}^2 - 3\mathcal{I}_{2,0,0}\mathcal{J}_{1,2,1}\mathcal{J}_{3,0,3} \\ &\quad - 18\mathcal{I}_{3,0,0}\mathcal{J}_{1,1,2}\mathcal{J}_{2,1,2} + 3\mathcal{J}_{3,0,3}\mathcal{J}_{3,2,1}\,, \end{aligned} \tag{E.7}$$

where the first two syzygies are the same as Eq. (E.6). Furthermore, we have found two additional syzygies that are distinct from Eq. (E.5). These four syzygies, as expected, can be used directly to show that the invariants $\{\mathcal{J}_{2,2,1}, \mathcal{J}_{2,1,2}, \mathcal{J}_{3,2,1}, \mathcal{J}_{3,1,2}\}$ are vanishing on the minimal CPC set $\{\mathcal{J}_{1,2,1}, \mathcal{J}_{1,1,2}, \mathcal{J}_{3,3,0}, \mathcal{J}_{3,0,3}\}$. However, we should note that the two

[21] When employing this numerical method, instead of constructing all possible monomials of invariant products at a given degree (q, t, y) to form the elements of the syzygy, we should only construct the products that include the element in the candidate minimal CPC set, along with the term of $\mathcal{J}$ to some power that we aim to verify vanishing.

new syzygies arise at order 12, which is higher than the order in Eq. (E.5). This should be considered as a necessary trade-off.

The above discussion is based on the syzygy approach. However, another interesting approach to consider is the use of ideals in the polynomial ring. As explained in App. G, the syzygy approach originates from the inclusion problem of a polynomial in an ideal defined in a polynomial ring. In the context of the 2HDM, the problem can be considered as to determine whether a power exists that would make the polynomials in the set $\mathcal{P} \equiv \{\mathcal{J}_{2,2,1}, \mathcal{J}_{2,1,2}, \mathcal{J}_{3,2,1}, \mathcal{J}_{3,1,2}\}$ belong to the ideal defined by $\mathcal{J} \equiv \langle \mathcal{J}_{1,2,1}, \mathcal{J}_{1,1,2}, \mathcal{J}_{3,3,0}, \mathcal{J}_{3,0,3} \rangle$. This can be easily checked with the **isSubset** function of the `Macaulay2` [52] package, or **PolynomialReduce** function in `Mathematica`.

F The Hironaka decomposition

As defined in Sec. 2.2, the basic set $\{\mathcal{I}_1, ..., \mathcal{I}_m\}$ allow us to write any invariant $\mathcal{I}'$ of the ring as a polynomial of the basic invariant:

$$\mathcal{I}' = P\left(\mathcal{I}_1, ..., \mathcal{I}_m\right) . \tag{F.1}$$

The Hironaka decomposition refines this last equation. This decomposition comes from the Cohen-Macaulay property [44, 45, 58] which only holds for reductive groups. In this cases it is possible to construct two finite sets of invariants: the set of primary invariants $\{\theta_1, \ldots, \theta_k\}$ which is algebraically independent, and the set of secondary invariants $\{\eta_1, \ldots, \eta_r\}$, such that any invariant $\mathcal{I}'$ can be written as

$$\mathcal{I}' = \sum_{i=1}^{r} \eta_i \, P_i(\theta_1, \ldots, \theta_k) \, , \tag{F.2}$$

where $P_i(\theta_1, \ldots, \theta_k)$ is a polynomial in the primary invariants. This decomposition is related to the HS in the following way. If z_j is the degree of θ_j, and s_i the degree of η_i, then

$$\mathcal{H}(q) = \frac{\mathcal{N}(q)}{\mathcal{D}(q)} = \frac{\sum\limits_{i=1}^{r} q^{s_i}}{\prod\limits_{j=1}^{k} \left(1 - q^{z_j}\right)} . \tag{F.3}$$

Note that, secondary invariants can be a product of several basic invariants. Furthermore, for a given group the Hironaka decomposition is not unique and the degrees of the primary and secondary invariants can be different.

G Hilbert's Nullstellensatz

Hilbert's Nullstellensatz, a fundamental result in algebraic geometry, establishes a profound connection between polynomial equations and the geometry of algebraic varieties. The traditional formulation of Hilbert's Nullstellensatz often involves a polynomial ring and its associated ideals. Consider the polynomial ring $R = k[x_1, x_2, \ldots, x_n]$ in n variables over the field k (a mathematical structure that generalizes the concept of numbers). This

ring consists of polynomials in the variables $x_1, x_2, \ldots, x_n$ with coefficients in k. We will now introduce the fundamental mathematical concepts required for presenting Hilbert's Nullstellensatz.

- Ideal

  An ideal I in the polynomial ring R is a subset of polynomials such that, for any polynomial f in the ideal I and any polynomial g in the ring, their product fg is also in I. Mathematically, an ideal I satisfies:

$$f \in I, \quad g \in R \implies fg \in I\,. \tag{G.1}$$

- Variety

  Given an ideal I, the variety $V(I)$ is the set of common zeros of all polynomials in I. Formally, a point $(a_1, a_2, \ldots, a_n)$ lies in the variety $V(I)$ if and only if every polynomial in I evaluates to zero at that point:

$$V(I) = \{(a_1, a_2, \ldots, a_n) \mid f(a_1, a_2, \ldots, a_n) = 0 \text{ for all } f \in I\}\,. \tag{G.2}$$

- Radical of an Ideal

  The radical of an ideal I, denoted by $\sqrt{I}$, is the set of all polynomials g such that some power of g belongs to I. Mathematically, $\sqrt{I}$ is defined as:

$$\sqrt{I} = \{g \mid g^k \in I \text{ for some } k \geq 1\}\,. \tag{G.3}$$

Hilbert's Nullstellensatz asserts that for any algebraically closed field k, there is a bijective correspondence between the points of a variety $V(I)$ and the radical ideals $\sqrt{I}$ defining that variety. Formally, this correspondence is expressed as:

$$\text{Ideal}(V(I)) = \sqrt{I}\,, \tag{G.4}$$

where $\text{Ideal}(V(I))$ denotes the ideal of polynomials vanishing on the variety $V(I)$.

In a more polynomial-centric language, Hilbert's Nullstellensatz can also be formulated differently. If a polynomial p vanishes on the variety $V(I)$, it belongs to $\text{Ideal}(V(I))$, and, by Hilbert's Nullstellensatz as shown in Eq. (G.4), it also belongs to $\sqrt{I}$. According to the definition of $\sqrt{I}$, there exists $k \geq 1$ such that $p^k \in I$, thus can be expressed as:

$$p^k = f_1 p_1 + f_2 p_2 + \cdots + f_m p_m\,, \tag{G.5}$$

where $f_i \in R$ and p_i are the defining polynomials of I. This equation essentially states that if p is vanishing under the common zeros of the defining polynomials of the ideal I, then the k-th power of the polynomial p can be expressed as a combination of these defining polynomials.

H CPC conditions for $n_N = n_f = 2$

The simplified model with two generations of fermions serves as a good example for the algebraic studies. There are in total 6 CP-odd generating invariants, which are shown in Eq. (C.16). Although the generating set is small, it's still difficult to find the common zero solutions of polynomials based on usual methods. However, these invariants can also be considered as ideals in the polynomial ring of the theory. In this section, we will analyze these ideals with the software package `Macaulay2` [52] based on the parameterization in Eq. (A.27). To simplify the notation, we take $r_{11} = a, r_{21} = c, c_{12} = b + p\,i$ and $c_{22} = d + q\,i$. Therefore, the polynomial ring is defined as $R := \mathbb{Q}[y_e, y_\mu, m_1, m_2, a, b, c, d, p, q]$, all CPV effects can be characterized by the ideal I defined by the six CP-odd invariants, i.e., $I \equiv \langle J_1, \ldots, J_6 \rangle$. The vanishing set denoted by $V(I)$ captures all the CPC conditions. The problem of finding common zero solutions is equivalent to finding the irreducible components of the ideal.

According to Hilbert's Nullstellensatz, the ideal of all polynomials that vanish on the common zero set $V(I)$ is the radical of the ideal $\sqrt{I}$, which can be calculated by the **radical** function in `Macaulay2`, and the CPC conditions will be captured by the minimal primes of the radical, which are shown as

$$
\begin{aligned}
&\{\langle q, p\rangle,\ \langle q, a\rangle,\ \langle p, c\rangle,\ \langle c, a\rangle,\ \langle c, b\rangle,\ \langle d, a\rangle,\ \langle d, b\rangle,\ \langle m_1, a\rangle,\ \langle m_1, c\rangle,\ \langle m_2, a\rangle,\ \langle m_2, c\rangle,\\
&\langle m_2, m_1\rangle,\ \langle m_1, y_e - y_\mu\rangle,\ \langle m_1, y_e + y_\mu\rangle,\ \langle m_2, y_e - y_\mu\rangle,\ \langle m_2, y_e + y_\mu\rangle,\\
&\langle m_1, d\,p - b\,q\rangle,\ \langle m_2, d\,p - b\,q\rangle,\ \langle m_1 - m_2, y_e - y_\mu\rangle,\ \langle m_1 - m_2, y_e + y_\mu\rangle,\\
&\langle m_1 + m_2, y_e - y_\mu\rangle,\ \langle m_1 + m_2, y_e + y_\mu\rangle,\ \langle y_e - y_\mu, a\,b + c\,d\rangle,\ \langle y_e - y_\mu, a\,p + c\,q\rangle,\\
&\langle y_e + y_\mu, a\,b + c\,d\rangle,\ \langle m_1 - m_2, a\,b\,c^2 - a^2\,c\,d + b^2\,c\,d - a\,b\,d^2 + c\,d\,p^2 - a\,b\,q^2\rangle,\\
&\langle y_e + y_\mu, a\,p + c\,q\rangle,\ \langle m_1 + m_2, a\,c^2\,p - a\,d^2\,p - a^2\,c\,q + b^2\,c\,q + c\,p^2\,q - a\,p\,q^2\rangle\}\,.
\end{aligned}
\tag{H.1}
$$

The CPC conditions can be obtained by setting the generators to 0, for instance, the first ideal in above set indicates there is one condition $p = q = 0$ that can lead to CPC, which is just the trivial solution of vanishing phases. However, there is no assumption about the variables in the ring, some unphysical solutions may arise, for example the ideal $\langle m_1, y_e + y_\mu\rangle$ indicates $m_1 = 0, y_e + y_\mu = 0$ is a CPC condition, but $y_e + y_\mu = 0$ will lead to $y_e = y_\mu = 0$ with non-negative mass assumption. This is already included in the ideal $\langle m_1, y_e - y_\mu\rangle$. After removing these unphysical ideals, the complete solutions of CPC conditions are given by following ideals

$$
\begin{aligned}
&\{\langle q, p\rangle,\ \langle q, a\rangle,\ \langle p, c\rangle,\ \langle c, a\rangle,\ \langle c, b\rangle,\ \langle d, a\rangle,\ \langle d, b\rangle,\ \langle m_1, a\rangle,\ \langle m_1, c\rangle,\ \langle m_2, a\rangle,\ \langle m_2, c\rangle,\\
&\langle m_2, m_1\rangle,\ \langle m_1, y_e - y_\mu\rangle,\ \langle m_2, y_e - y_\mu\rangle,\ \langle m_1, d\,p - b\,q\rangle,\ \langle m_2, d\,p - b\,q\rangle,\\
&\langle m_1 - m_2, y_e - y_\mu\rangle,\ \langle y_e - y_\mu, a\,b + c\,d\rangle,\ \langle y_e - y_\mu, a\,p + c\,q\rangle,\\
&\langle m_1 - m_2, a\,b\,c^2 - a^2\,c\,d + b^2\,c\,d - a\,b\,d^2 + c\,d\,p^2 - a\,b\,q^2\rangle\}\,.
\end{aligned}
\tag{H.2}
$$

In the algebraic geometry picture, all these conditions are fundamental objects, and they correspond to points, lines, surfaces, etc. In addition, each of these conditions has a connection to the special spectrum and enlarged symmetry of the theory. There are some

nontrivial conditions in the solution list, for instance the last one shows that the mass degeneracy of $m_1 = m_2$ and a vanishing of a very special combination of the matrix elements of Y_N can lead to CPC.

By analyzing the polynomial ring, it's also possible to generate the conditions lead to special spectrum, which is not necessary the CPC conditions. For instance, the conditions lead to unphysical phase of p or q can be obtained by the elimination of variables, and the relevant function in `Macaulay2` is called **eliminate**. By eliminating the CP-odd variable p or q, one can find the following conditions

$$\begin{aligned} &\text{unphysical } p\text{: } \{\langle q\rangle,\ \langle c\rangle,\ \langle m_1\rangle,\ \langle m_2\rangle,\ \langle m_1 - m_2\rangle,\ \langle y_e - y_\mu\rangle,\ \langle d, b\rangle,\ \langle d, a\rangle\}\,, \\ &\text{unphysical } q\text{: } \{\langle p\rangle,\ \langle a\rangle,\ \langle m_1\rangle,\ \langle m_2\rangle,\ \langle m_1 - m_2\rangle,\ \langle y_e - y_\mu\rangle,\ \langle d, b\rangle,\ \langle c, b\rangle\}\,, \end{aligned} \tag{H.3}$$

where the unphysical conditions such as $\langle m_1 + m_2\rangle$ is removed. These conditions can be used to analyze the special spectrum of the theory, such as the spectrum listed in the zero table algorithm. The above conditions can also be calculated with the more physical parameterization in Eq. (A.17), they are given as follows

$$\begin{aligned} \text{unphysical } \phi\text{: } &\{\langle \sin\varphi\rangle,\ \langle \sin\varphi - 1\rangle,\ \langle \sin\varphi + 1\rangle,\ \langle \sin\alpha\rangle,\ \langle \sin\alpha - 1\rangle,\ \langle \sin\alpha + 1\rangle, \\ &\langle m_1\rangle,\ \langle m_2\rangle,\ \langle m_1 - m_2\rangle,\ \langle y_1 - y_2\rangle\}\,, \\ \text{unphysical } \varphi\text{: } &\{\langle \sin\phi\rangle,\ \langle \sin\alpha\rangle,\ \langle \sin\alpha - 1\rangle,\ \langle \sin\alpha + 1\rangle,\ \langle \sin\theta\rangle,\ \langle \sin\theta - 1\rangle, \\ &\langle \sin\theta + 1\rangle,\ \langle y_e - y_\mu\rangle,\ \langle m_1 - m_2\rangle,\ \langle y_1\rangle,\ \langle y_2\rangle,\ \langle y_1 - y_2\rangle\}\,. \end{aligned} \tag{H.4}$$

By exploring these special conditions and their combinations, one can obtain all of the special spectra with enlarged symmetries in the theory.

(mention that `Macaulay2` only work for $n_N = n_f = 2$ in our theory, but it's still a good tool to study some other simple theories, such as 2HDM.)

References

[1] Y. Fukuda *et al.* (Super-Kamiokande), *Measurements of the solar neutrino flux from Super-Kamiokande's first 300 days*, Phys. Rev. Lett. **81** (1998) 1158, [Erratum: Phys.Rev.Lett. 81, 4279 (1998)], arXiv:hep-ex/9805021.

[2] Q. R. Ahmad *et al.* (SNO), *Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (2001) 071301, arXiv:nucl-ex/0106015.

[3] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, Phys. Rev. Lett. **43** (1979) 1566.

[4] P. Minkowski, *$\mu \to e\gamma$ at a Rate of One Out of 10^9 Muon Decays?*, Phys. Lett. B **67** (1977) 421.

[5] T. Yanagida, *Horizontal gauge symmetry and masses of neutrinos*, Conf. Proc. C **7902131** (1979) 95.

[6] M. Gell-Mann, P. Ramond, and R. Slansky, *Complex Spinors and Unified Theories*, Conf. Proc. C **790927** (1979) 315, arXiv:1306.4669 [hep-th].

[7] S. L. Glashow, *The Future of Elementary Particle Physics*, NATO Sci. Ser. B **61** (1980) 687.

[8] R. N. Mohapatra and G. Senjanovic, *Neutrino Mass and Spontaneous Parity Nonconservation*, Phys. Rev. Lett. **44** (1980) 912.

[9] T. Asaka, S. Blanchet, and M. Shaposhnikov, *The nuMSM, dark matter and neutrino masses*, Phys. Lett. B **631** (2005) 151, arXiv:hep-ph/0503065.

[10] T. Asaka and M. Shaposhnikov, *The νMSM, dark matter and baryon asymmetry of the universe*, Phys. Lett. B **620** (2005) 17, arXiv:hep-ph/0505013.

[11] T. Asaka, M. Shaposhnikov, and A. Kusenko, *Opening a new window for warm dark matter*, Phys. Lett. B **638** (2006) 401, arXiv:hep-ph/0602150.

[12] T. Asaka, M. Laine, and M. Shaposhnikov, *Lightest sterile neutrino abundance within the nuMSM*, JHEP **01** (2007) 091, [Erratum: JHEP 02, 028 (2015)], arXiv:hep-ph/0612182.

[13] T. Asaka, M. Laine, and M. Shaposhnikov, *On the hadronic contribution to sterile neutrino production*, JHEP **06** (2006) 053, arXiv:hep-ph/0605209.

[14] S. Davidson, E. Nardi, and Y. Nir, *Leptogenesis*, Phys. Rept. **466** (2008) 105, arXiv:0802.2962 [hep-ph].

[15] R. L. Workman *et al.* (Particle Data Group), *Review of Particle Physics*, PTEP **2022** (2022) 083C01.

[16] C. Jarlskog, *A Basis Independent Formulation of the Connection Between Quark Mass Matrices, CP Violation and Experiment*, Z. Phys. C **29** (1985) 491.

[17] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation*, Phys. Rev. Lett. **55** (1985) 1039.

[18] J. Bernabeu, G. C. Branco, and M. Gronau, *CP Restrictions on Quark Mass Matrices*, Phys. Lett. B **169** (1986) 243.

[19] G. C. Branco and M. N. Rebelo, *Weak Basis Invariant Conditions for CP Conservation in $SU(2)_L \times SU(2)_R \times U(1)_{B-L}$ Models*, Phys. Lett. B **173** (1986) 313.

[20] G. C. Branco, L. Lavoura, and M. N. Rebelo, *Majorana Neutrinos and CP Violation in the Leptonic Sector*, Phys. Lett. B **180** (1986) 264.

[21] F. J. Botella and J. P. Silva, *Jarlskog - like invariants for theories with scalars and fermions*, Phys. Rev. D **51** (1995) 3870, arXiv:hep-ph/9411288.

[22] G. C. Branco, T. Morozumi, B. M. Nobre, and M. N. Rebelo, *A Bridge between CP violation at low-energies and leptogenesis*, Nucl. Phys. B **617** (2001) 475, arXiv:hep-ph/0107164.

[23] O. Lebedev, *CP violating invariants in supersymmetry*, Phys. Rev. D **67** (2003) 015013, arXiv:hep-ph/0209023.

[24] J. F. Gunion and H. E. Haber, *Conditions for CP-violation in the general two-Higgs-doublet model*, Phys. Rev. D **72** (2005) 095002, arXiv:hep-ph/0506227.

[25] S. Davidson and H. E. Haber, *Basis-independent methods for the two-Higgs-doublet model*, Phys. Rev. D **72** (2005) 035004, [Erratum: Phys.Rev.D 72, 099902 (2005)], arXiv:hep-ph/0504050.

[26] E. E. Jenkins and A. V. Manohar, *Algebraic Structure of Lepton and Quark Flavor Invariants and CP Violation*, JHEP **10** (2009) 094, arXiv:0907.4763 [hep-ph].

[27] A. Trautner, *Systematic construction of basis invariants in the 2HDM*, JHEP **05** (2019) 208, arXiv:1812.02614 [hep-ph].

[28] B. Yu and S. Zhou, *The number of sufficient and necessary conditions for CP conservation with Majorana neutrinos: three or four?*, Phys. Lett. B **800** (2020) 135085, arXiv:1908.09306 [hep-ph].

[29] Y. Wang, B. Yu, and S. Zhou, *Flavor invariants and renormalization-group equations in the leptonic sector with massive Majorana neutrinos*, JHEP **09** (2021) 053, arXiv:2107.06274 [hep-ph].

[30] B. Yu and S. Zhou, *Hilbert series for leptonic flavor invariants in the minimal seesaw model*, JHEP **10** (2021) 017, arXiv:2107.11928 [hep-ph].

[31] B. Yu and S. Zhou, *Sufficient and Necessary Conditions for CP Conservation in the Case of Degenerate Majorana Neutrino Masses*, Phys. Rev. D **103** (2021) 035017, arXiv:2009.12347 [hep-ph].

[32] M. P. Bento, J. P. Silva, and A. Trautner, *The basis invariant flavor puzzle*, JHEP **01** (2024) 024, arXiv:2308.00019 [hep-ph].

[33] N. Darvishi, Y. Wang, and J.-H. Yu, *Automated ring-diagram framework for classifying CP invariants*, (2023), arXiv:2311.15422 [hep-ph].

[34] Q. Bonnefoy, E. Gendy, C. Grojean, and J. T. Ruderman, *Beyond Jarlskog: 699 invariants for CP violation in SMEFT*, JHEP **08** (2022) 032, arXiv:2112.03889 [hep-ph].

[35] Q. Bonnefoy, E. Gendy, C. Grojean, and J. T. Ruderman, *Opportunistic CP violation*, JHEP **06** (2023) 141, arXiv:2302.07288 [hep-ph].

[36] B. Yu and S. Zhou, *Spelling out leptonic CP violation in the language of invariant theory*, Phys. Rev. D **106** (2022) L051701, arXiv:2203.00574 [hep-ph].

[37] B. Yu and S. Zhou, *CP violation and flavor invariants in the seesaw effective field theory*, JHEP **08** (2022) 017, arXiv:2203.10121 [hep-ph].

[38] Q. Bonnefoy, C. Grojean, and J. Kley, *Shift-Invariant Orders of an Axionlike Particle*, Phys. Rev. Lett. **130** (2023) 111803, arXiv:2206.04182 [hep-ph].

[39] A. Ibarra and G. G. Ross, *Neutrino phenomenology: The Case of two right-handed neutrinos*, Phys. Lett. B **591** (2004) 285, arXiv:hep-ph/0312138.

[40] B. Feng, A. Hanany, and Y.-H. He, *Counting gauge invariants: The Plethystic program*, JHEP **03** (2007) 090, arXiv:hep-th/0701063.

[41] L. Lehman and A. Martin, *Hilbert Series for Constructing Lagrangians: expanding the phenomenologist's toolbox*, Phys. Rev. D **91** (2015) 105014, arXiv:1503.07537 [hep-ph].

[42] B. Henning, X. Lu, T. Melia, and H. Murayama, *Operator bases, S-matrices, and their partition functions*, JHEP **10** (2017) 199, arXiv:1706.08520 [hep-th].

[43] Y. Xiao, Y.-H. He, and C. Matti, *Standard Model Plethystics*, Phys. Rev. D **100** (2019) 076001, arXiv:1902.10550 [hep-th].

[44] B. Sturmfels, *Algorithms in invariant theory* (Springer Science & Business Media, 2008).

[45] H. Derksen and G. Kemper, *Computational invariant theory* (Springer, 2015).

[46] S. Benvenuti, B. Feng, A. Hanany, and Y.-H. He, *Counting BPS Operators in Gauge Theories: Quivers, Syzygies and Plethystics*, JHEP **11** (2007) 050, arXiv:hep-th/0608050.

[47] A. Hanany, E. E. Jenkins, A. V. Manohar, and G. Torri, *Hilbert Series for Flavor Invariants of the Standard Model*, JHEP **03** (2011) 096, arXiv:1010.3161 [hep-ph].

[48] C. Grojean, J. Kley, and C.-Y. Yao, *CHINCHILLA: A Mathematica package for the construction of invariants using the Hilbert series*, To appear, 2024.

[49] E. E. Jenkins and A. V. Manohar, *Rephasing Invariants of Quark and Lepton Mixing Matrices*, Nucl. Phys. B **792** (2008) 187, arXiv:0706.4313 [hep-ph].

[50] D. Eisenbud, *Commutative algebra: with a view toward algebraic geometry*, Vol. 150 (Springer Science & Business Media, 2013).

[51] M. Atiyah, *Introduction to commutative algebra* (CRC Press, 2018).

[52] D. R. Grayson and M. E. Stillman, *Macaulay2, a software system for research in algebraic geometry*, Available at `http://www2.macaulay2.com`.

[53] W. Decker, G.-M. Greuel, G. Pfister, and H. Schönemann, SINGULAR *4-3-0 — A computer algebra system for polynomial computations*, `http://www.singular.uni-kl.de` (2022).

[54] J. Kollár, *Sharp effective nullstellensatz*, Journal of the American Mathematical Society **1** (1988) 963.

[55] I. P. Ivanov, C. C. Nishi, and A. Trautner, *Beyond basis invariants*, Eur. Phys. J. C **79** (2019) 315, arXiv:1901.11472 [hep-ph].

[56] T. Peraro, *FiniteFlow: multivariate functional reconstruction using finite fields and dataflow graphs*, JHEP **07** (2019) 031, arXiv:1905.08019 [hep-ph].

[57] J. Mangan, *FiniteFieldSolve: Exactly Solving Large Linear Systems in High-Energy Theory*, (2023), arXiv:2311.01671 [hep-th].

[58] C. Procesi, *Lie Groups: An Approach through Invariants and Representations*, Universitext (Springer New York, 2007) p. 561.