Zee model with quasidegenerate neutrino masses and where to find it

R. Primulando$^{1,a}$, J. Julio$^{2,b}$, P. Uttayarat$^{3,c}$

$^1$ Center of Theoretical Physics, Department of Physics, Parahyangan Catholic University, Jl. Ciiumbuleuit 94, Bandung 40141, Indonesia
$^2$ National Research and Innovation Agency, Kompleks Puspiptek Serpong, South Tangerang 15314, Indonesia
$^3$ Department of Physics, Srinakharinwirot University, 114 Sukhumvit 23 Rd., Wattana, Bangkok 10110, Thailand

Received: 12 January 2022 / Accepted: 7 March 2022 / Published online: 24 March 2022
© The Author(s) 2022

Abstract We present a Zee model with a family dependent $Z_2$ symmetry for radiative neutrino masses. Our motivation is to get a model that correctly describes neutrino oscillation phenomena, while at the same time offers definite predictions. The imposed $Z_2$ symmetry greatly reduces the number of free parameters in the model. These parameters are then determined from the neutrino data, from which one can study its outcomes. Our setup only admits quasidegenerate neutrino masses with the sum of neutrino masses between 100 and 130 meV, the effective Majorana mass between 20 and 40 meV, and the effective electron neutrino mass between 48 and 53 meV. The ratio of the vacuum expectation values of the Higgs doublets, $\tan \beta$, is found to be $\tan \beta \lesssim 0.5$ and $\tan \beta \gtrsim 10$. The former is ruled out by lepton flavor violation (LFV) processes, such as $\mu \rightarrow e\gamma$ and $\mu \rightarrow e$ conversion, which are determined up to two loops. For the latter, these LFV processes are within reach of the next generation of experiments. Moreover, for $\tan \beta \gtrsim 10$, the couplings of heavy neutral scalars to dimuon are significant. If they are sufficiently light, i.e., $\lesssim 200$ GeV, collider search for their decays into muon pair provides a stronger constraint on most parts of the model parameter space than the LFV ones.

1 Introduction

The mounting evidence of neutrino oscillations has led to the fact that neutrinos must be massive and mix. This indicates that the Standard Model (SM) must be considered as an effective theory; it cannot explain how neutrinos gain masses, let alone why their masses are so tiny compared to those of other SM fermions. There is a variety of models in literature attempting to explain this phenomenon. The simplest example is the (type-I) seesaw mechanism, in which neutrinos get their masses at tree level via an exchange with electroweak singlet right-handed neutrinos [1–5]. The correct neutrino masses of order 0.1 eV will require the right-handed neutrino masses to be very large, i.e., of the order $10^{14}$ GeV. With such a large scale, it is then difficult to test this mechanism at laboratories.

There is also another alternative of neutrino mass generation, namely the radiative mechanism. Here neutrino masses are induced at loop level, see [6–8] for pioneers of this class of models. Within this mechanism, due to loop and, possibly, chirality suppression factors, the new physics scale can be much lower; in some cases it can be even at TeV scale, placing it within reach of the Large Hadron Collider (LHC), see for example Refs. [7–12]. One of the well-studied models within this class is the Zee model [6]. Considered as an extension of the two-Higgs-doublet model by a singly-charged scalar, in its general form, this model induces tree-level flavor-changing neutral currents (FCNCs) in both lepton and quark sectors. Usually, such dangerous processes can be avoided by imposing the natural flavor conservation [13,14], by which only one Higgs doublet can couple to right-handed fermions of the same charge. However, if one applies this to neutrino mass generation [15], it will result in vanishing of all diagonal elements of neutrino mass matrix, leading to its exclusion by solar neutrino and KamLAND data. It should be noted, however, that the general Zee model is still compatible with oscillation data, but owing to its large number of parameters, it is hard to make a prediction out of it. For earlier phenomenological studies of the general Zee model, see Refs. [16–25].

In the past, an attempt to make a viable and predictive Zee model was done by introducing a family-dependent $Z_4$ symmetry [22]. Within this scenario, tree-level FCNCs in the lepton sector were allowed, but their rates were found to be
below experimental limits even for very light scalar mediators. Thanks to the reduced number of parameters, the model was very predictive: it allowed for the inverted ordering of neutrino mass with no CP violation in the lepton sector. This result was in an excellent agreement with neutrino data for some time, but now it has been disfavored by the T2K result [26], which has reported the evidence of leptonic CP violation at more than 3σ for the BO case. Although other experiment such as NOvA may not severely exclude the CP conserving region [27,28], it is nevertheless useful to seek a model that can naturally allow for CP violation.

In this paper, we modify the Zee model by imposing a family-dependent $Z_2$ symmetry. This allows a nontrivial mixing in the lepton sector, hence tree-level FCNCs. The group itself is smaller than $Z_4$, so one may expect that the current scenario contains more parameters than those using $Z_4$. In fact, here we have an additional complex parameter compared to that of Ref. [22]. It makes the neutrino mass matrix complex, which in turn allows for nonvanishing CP violation. In total, this model contains 6 real parameters describing the neutrino mass and mixing. All can be fitted into neutrino oscillation data well, with the prediction that the present model accommodates only the quasidegenerate neutrino masses. As a consequence, the result of this model is very sensitive to future experiments aiming on measuring neutrino mass. The sum of neutrino masses and the effective mass for neutrino oscillation, $\sum m_i$, independent of $\beta$, is clustered into two regimes, namely, $\tan \beta \lesssim 0.5$ and $\tan \beta \gtrsim 10$, independent of the detail of the scalar phenomenology. The LFV processes, such as $\mu \rightarrow e\gamma$ and $\mu \rightarrow e$ conversion in nuclei, evaluated up to the two-loop Barr–Zee diagrams [31], further rule out $\tan \beta \lesssim 0.5$ regime for wide range of scalar masses (i.e., $\lesssim 10$ TeV). This is partly because the corresponding $e\mu$ coupling is significantly enhanced in this case. Meanwhile, the Higgs couplings to top quarks, relevant for such two–loop diagrams (see [32,33]), remain sizable. This is in contrast to the case of $\tan \beta \gtrsim 10$, where the corresponding LFV rates are found to be well below the current bounds, and thus can only be probed by future generation experiments, such as Mu2e [34] or MEGII [35]. But what we find interesting is the importance of collider searches. Specifically we find that in one of the two possible ways to couple quarks to the Higgses, the decay of a heavy Higgs into a dimuon final state gives a stronger constraint on the model than the present LFV transitions, provided that the heavy Higgs is 200 GeV or lighter. Using the CMS partial data set of 35.9 fb$^{-1}$ on this particular $\mu^+\mu^-$ channel [36], we can rule out a great portion of the model parameter space. The CMS full data set, once available, will certainly be able to falsify this scenario.

The rest of the paper is organized as follows. In Sect. 2, we shall discuss the detail of the model. In Sect. 3, we discuss how this model describes the neutrino mass and mixing. This is followed by the discussion of relevant lepton flavor violation processes in Sect. 4 and collider phenomenology in Sect. 5. We conclude and discuss our finding in Sect. 6.

2 The model

The model presented here is an extension of the SM. The scalar sector is expanded to contain two Higgs doublets $\Phi_a$ with $a = 1, 2$ and a singly-charged scalar singlet $\eta^+$. A discrete $Z_2$ symmetry that is flavor dependent is imposed on the lepton and the scalar sectors as follows

$$L_i \equiv \begin{pmatrix} \nu_i \\ \ell_i \end{pmatrix}: (-1, 1, 1), \quad e_R i: 1, \quad \Phi_1 : 1, \quad \Phi_2 : -1, \quad \eta^+ : -1,$$

where $L$ and $e_R$ denote the SM lepton doublet and singlet, respectively, and $i$ is a lepton-family index $i = e, \mu, \tau$. Family-dependent $Z_2$ charge assignment in the lepton sector will lead to LFV processes. We will consider LFV constraints in Sect. 4.

With the charge assignment given in Eq. (1), the Yukawa interactions in lepton sector can be written as

$$
L_{\text{Yuk}}^{\text{lepton}} = \sum_{i=\nu, \mu, \tau} \sum_{a=\mu, \tau} Y_{ai} \bar L_a e_R i \Phi_1 \\
+ \sum_{i=\nu, \mu, \tau} Y_{iC} \bar L_i e_R i \Phi_2 \\
+ \sum_{a=\mu, \tau} f_{ea} L_e^\gamma C (i\sigma^2) L_a \eta^+ + \text{h.c.},
$$

where $C$ is the charge conjugation matrix and the $SU(2)_L$ indices are contracted by the antisymmetric tensor $i\sigma^2$. Due to the Fermi statistics, the coupling matrix $f$ is antisymmetric in flavor space. In addition to the aforementioned Yukawa interactions, there also exists a gauge-invariant term that is part of the scalar potential

$$V \supset \mu \Phi_1^T (i\sigma_2) \Phi_2 \eta^- + \text{h.c.}
$$

The non-vanishing $f$ and $\mu$ break lepton number by two units, leading to the generation of Majorana-type neutrino masses.

After the electroweak symmetry breaking, each Higgs doublet acquires vacuum expectation value (vev), denoted
as $\langle \Phi_a \rangle \equiv v_a/\sqrt{2}$. It is also customary to rotate the two Higgs doublet into the so-called Higgs basis

$$H_1 = \cos \beta \Phi_1 + \sin \beta \Phi_2$$

$$H_2 = \sin \beta \Phi_1 - \cos \beta \Phi_2,$$

(4)

with $\tan \beta \equiv v_2/v_1$, such that only one Higgs doublet has vev $v = \sqrt{v_1^2 + v_2^2} = 246$ GeV. In a more explicit form we have

$$H_1 = \left( \frac{v + h_1 + iG}{\sqrt{2}} \right), \quad H_2 = \left( \frac{H^+}{h_2 + iA} \frac{G^+}{\sqrt{2}} \right).$$

Here $G^+$ and $G$ are the would-be Goldstone bosons eaten by the $W^+$ and $Z$, while $H^+$ and $A$ are the massive charged scalar and CP-odd scalar, respectively. The two CP-even scalars, denoted by $h_1$ and $h_2$, in principle, can mix with the CP-odd one. However, throughout this paper, we shall assume that CP be a good symmetry of the scalar sector, so there is no mixing between CP-even and CP-odd scalar fields. Furthermore, based on current measurements of the 125 GeV Higgs couplings at the LHC [37–39], which in agreement with the SM expectations, it is instructive to adopt the decoupling limit, in which $h_1$ couples to fermions and gauge bosons as in the SM. Throughout this paper, we shall identify $h_1$ with $h$, the 125 GeV Higgs boson, and $h_2$ with $H$, the heavy Higgs. It is straightforward to generalize our result to a case where $h_1$ can mix with $h_2$.

It is worth noting that, due to the trilinear coupling of Eq. (3), there is also a mixing between $H^+$ and $\eta^+$ characterized by

$$\begin{pmatrix} H^+ \\ \eta^+ \end{pmatrix} = \begin{pmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{pmatrix} \begin{pmatrix} H^+_{12} \\ \eta^+ \end{pmatrix},$$

where $H^+_{12}$ are mass eigenstates and $\sin 2\gamma = \sqrt{2} \mu v/(m^2_{H^+1} - m^2_{\eta^+}).$ However, such mixing is constrained by neutrino mass to be rather small, see Fig. 1. Therefore, in practice, we can treat both charged scalars as nearly physical fields.

In the Higgs basis, the Yukawa interactions involving the Higgs doublets read

$$L^\text{lepton}_{\text{Yuk}} = \sum_{i=e,\mu,\tau} \sum_{\alpha=\mu,\tau} \bar{Y}_{ai} \ell_i e_{Ri} \left[ \cos \beta H_1 + \sin \beta H_2 \right]$$

$$+ \sum_{i=e,\mu,\tau} \bar{Y}_i \ell_i e_{Ri} \left[ \sin \beta H_1 - \cos \beta H_2 \right],$$

(7)

The flavor structures of these two Yukawa matrices, $Y_i$ and $Y_{ai}$, bear similarity to the ones presented in Ref. [22]. The latter can be considered as a $3 \times 3$ matrix with the vanishing first row. We can bring such a matrix into a diagonal form, with $Y_{\mu\mu}$ and $Y_{\tau\tau}$ being the only nonzero entries, by performing a unitary field redefinition on the $L_\mu, \tau$ and $e_{Ri}$ fields. Note that such field redefinitions affect neither the form of the antisymmetric coupling matrix $f$ nor the form of gauge charged currents. The former keeps its antisymmetric form

$$f = f_{e\mu} \begin{pmatrix} 0 & 1 & re^{i\theta} \\ -1 & 0 & 0 \\ -r & 0 & 0 \end{pmatrix},$$

(8)

with $re^{i\theta} \equiv f_{e\tau}/f_{e\mu}$ because the field redefinitions are performed on the $\mu-\tau$ sector.

Using the basis where $Y_{ai}$ is diagonal, the lepton mass matrix can be expressed as

$$M_{\ell} = \frac{v}{\sqrt{2}} \begin{pmatrix} Y_e \sin \beta & Y_{\mu} \sin \beta & Y_{\tau} \sin \beta \\ 0 & Y_{\mu\mu} \cos \beta & 0 \\ 0 & 0 & Y_{\tau\tau} \cos \beta \end{pmatrix}.$$}

(9)

Similar to the case of Ref. [22], all complex phases in the above leptonic mass matrix can be removed, leaving us with a real and positive matrix. We, then, diagonalize $M_{\ell}$ by a bi-orthogonal transformation

$$O^T_{L} M_{\ell} O_R = \text{diag}(m_\ell, m_{\mu}, m_{\tau}),$$

(10)

where $O_L$ and $O_R$ are orthogonal matrices used to rotate the $\ell_i$ and $e_{Ri}$ fields into their mass eigenstates. They can be written as

$$O_{L} = \begin{pmatrix} c_{\theta_1} & -s_{\theta_1} & 0 \\ s_{\theta_1} & c_{\theta_1} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_{\theta_2} & 0 & -s_{\theta_2} \\ 0 & 1 & 0 \\ s_{\theta_2} & 0 & c_{\theta_2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{\theta_3} & -s_{\theta_3} \\ 0 & s_{\theta_3} & c_{\theta_3} \end{pmatrix},$$

(11)

where $c_{\theta_i}$ ($s_{\theta_i}$) stands for $\cos \theta_i$ ($\sin \theta_i$). For $O_R$, one just needs to make a replacement $\theta_i \to \alpha_i$. Now, out of 5 real parameters in $M_{\ell}$, 3 of them will become charged lepton masses, indicating that only two mixing angles turn out to be physical. There is no sacred recipe of how to choose these physical angles. One can, for example, use the four zero entries of $M_{\ell}$ in Eq. (10) to eliminate four angles in favor of the two physical ones. These two angles, say, $\theta_1$ and $\alpha_3$, will later be determined through neutrino oscillation data. We give the complete expression in Appendix. Note that the procedure outlined above, in conjunction with taking the decoupling limit, will make leptonic interactions with $h$ flavor diagonal. Their strengths are proportional to lepton masses, as in the SM.

Having diagonalized the charged lepton mass matrix, we now turn to flavor-changing Yukawa interactions, which in
the present case are given by
\[
L_{\text{Yuk}}^{\text{lepton}} \ni f_{ij} L_i^T C (i \sigma^2) L_j \eta^+ + \sqrt{2} Y_{ij} \bar{L}_i e_R^j H_2 + \text{h.c.} \tag{12}
\]

The explicit form of \( f \) can be seen in Eq. (8), while the matrix coupling \( Y \) is defined as
\[
Y = \frac{1}{\nu} \begin{pmatrix} -\cot \beta & 0 & 0 \\ 0 & \tan \beta & 0 \\ 0 & 0 & \tan \beta \end{pmatrix} M_L. \tag{13}
\]

Following the rotation of the leptonic fields into their mass eigenbasis, the two matrices \( f \) and \( Y \) get transformed into \( \tilde{f} = O_L^T f O_L \) and \( \tilde{Y} = O_L^T Y O_R \), respectively. Such rotation induces the (2,3) entries in the antisymmetric coupling \( \tilde{f} \), which were zero in the original basis. As for \( \tilde{Y} \), Eq. (13) implies that it has a hierarchical structure, namely \( \tilde{Y}_{ij} \tilde{Y}_{ji} \). Now with all this in hand, we can write down the Yukawa interactions in lepton sector as follows
\[
L_{\text{Yuk}}^{\text{lepton}} = 2 \sum_{i<j} \tilde{f}_{ij} \bigg( v_i^T C \ell_j - v_j^T C \ell_i \bigg) \eta^+ + \frac{m_{\ell_i}}{\nu} \delta_{ij} \tilde{\eta}_i e_R^j h \\
+ \sqrt{2} \tilde{Y}_{ij} \nu_i e_R^j H^+ + i \tilde{Y}_{ij} \tilde{\eta}_i e_R^j A \\
+ \tilde{Y}_{ij} \tilde{\eta}_i e_R^j H + \text{h.c.}. \tag{14}
\]

The same argument also applies for the quark sector. However, in order to avoid the notorious tree-level flavor-changing neutral currents (FCNC), it is desirable to adopt natural flavor conservation [13,14], that is, a \( Z_2 \) charge assignment is chosen such that right-handed quarks of the same charge couple only to a single Higgs doublet \( \Phi_d \). Mathematically,
\[
L_{\text{Yuk}}^{\text{quarks}} = Y_{ij}^u \bar{Q}_i \Phi_d u_R^j + Y_{ij}^d \bar{Q}_i \Phi_d d_R^j + \text{h.c.}, \tag{15}
\]
where \( \Phi_d \equiv i \sigma_2 \Phi_d^* \). After rotating the Higgs doublets using Eq. (4), we can write the Lagrangian as [40,41]
\[
L_{\text{Yuk}}^{\text{quark}} = \frac{m_{\nu_i}}{\nu} \bigg( \tilde{Q}_i q_j h + \xi_{\nu_i}^d \tilde{Q}_i q_j H - i e_A^\nu \tilde{\xi}_{\nu_i}^d \tilde{\eta}_i \gamma_5 q_j A \bigg) \\
- \frac{\sqrt{2}}{\nu} \tilde{Q}_i \left( \xi_{\nu_i}^d V_{ij} m_{d_j} P_R + \xi_{\nu_i}^u m_{u_i} V_{ij} P_L \right) d_j H^+ + \text{h.c.}, \tag{16}
\]
where \( V \) denotes the Cabibbo–Kobayashi–Maskawa (CKM) mixing matrix and \( P_{L,R} = \frac{1}{2} (1 \mp \gamma_5) \) are projection operators. In the decoupling limit, we have \( \xi_{\nu_i}^H = \xi_{\nu_i}^d \) and \( \xi_{\nu_i}^d = -\xi_{\nu_i}^u \). It is then straightforward to see that, in the case of quarks coupling universally to \( \Phi_1, \Phi_2 \), \( \xi_{\nu_i}^u = \{ \tan \beta, -\cot \beta \} \). Similarly, in the case where only \( \Phi_1 \) couples to \( d_R \) and \( \Phi_2 \) to \( u_R \), as in the type-II 2HDM, we have \( \xi_{\nu_i}^u = -1/\xi_{\nu_i}^d = -\cot \beta \).

Although the quark sector does not seem related to the neutrino sector, neutrino oscillation data will play an important role in determining whether the quarks couple to \( \Phi_1 \) or \( \Phi_2 \) for relatively light scalars. As we will see later in the next section, fitting the model into neutrino data will result in the allowed values of \( \tan \beta \). The very same parameter is also known to be constrained from rare \( B \) decays, in particular the \( B \to X_s \gamma \) transitions. From Ref. [42], we can see that if quarks couple to \( \Phi_2 \) there exists a lower bound for \( \tan \beta \), albeit charged scalar mass dependent. For example, if \( m_{H^+} = 130 \) GeV, \( \tan \beta \gtrsim 3 \). This bound loosens to \( \tan \beta \gtrsim 1.3 \) for \( m_{H^+} = 800 \) GeV. Similarly, in the case where quarks couple to \( \Phi_1 \), this bound will translate to \( \tan \beta \lesssim 0.33 \) \( \tan \beta \lesssim 0.77 \) for \( m_{H^+} = 130 \) GeV \( (800 \) GeV). This suggests, in the case where quarks couple universally to one Higgs doublet and the scalars are light, quarks are dictated to couple to \( \Phi_1 (\Phi_2) \) for low (large) \( \tan \beta \). Although such \( B \to X_s \gamma \) transitions only constrain the charged Higgs mass, whereas in the present paper we will mostly discuss the neutral Higgses, the bound is still relevant since the masses of exotic scalars cannot differ so greatly, or else there will be inconsistencies with the electroweak \( \rho \) parameter constraint and perturbativity, see Refs. [41,43].

In the case where \( \Phi_1 \) couples to \( d_R \) and \( \Phi_2 \) to \( u_R \), the bounds from \( B \to X_s \gamma \) require that \( m_{H^+} \gtrsim 600 \) GeV for \( \tan \beta \gtrsim 2 \). At the same time, \( \tan \beta \) is bounded from above by other rare \( B \) decay processes. The strongest upper bound is provided by \( B_s \to \mu^+ \mu^- \), with \( \tan \beta \lesssim 17 \) for \( m_{H^+} = 600 \) GeV. The bound loosens to \( \tan \beta \lesssim 22 \) for \( m_{H^+} = 800 \) GeV. If, instead, one considers the case where \( \Phi_1 \) couples to \( u_R \) and \( \Phi_2 \) to \( d_R \), the bounds from \( B \to X_s \gamma \) would translate to \( m_{H^+} \gtrsim 600 \) GeV for \( \tan \beta \lesssim 0.5 \). The corresponding constraints from \( B_s \to \mu^+ \mu^- \) read \( \tan \beta \lesssim 0.06 \) \( (0.05) \) for \( m_{H^+} = 600 \) \( (800) \) GeV. Improved measurements on \( B \to X_s \gamma \) and \( B_s \to \mu^+ \mu^- \) will put more constraints on this allowed parameter space.

3 Neutrino phenomenology

In this section we study the neutrino phenomenology in our model. Neutrino masses are induced at the one-loop level by Feynman diagram shown in Fig. 1. In addition to this, there is also another diagram whose internal particles are replaced by their charge conjugates, which is made possible by Majorana property of neutrinos. The sum of the two diagrams yields a symmetric neutrino mass matrix
\[
M_{\nu} = \kappa (\hat{f} M_{\nu}^{\text{diag}} \hat{\nu}^T + \hat{Y} M_{\nu}^{\text{diag}} \hat{f}^T). \tag{17}
\]
where \( 16 \pi^2 \kappa = \sin 2 \gamma \ln (M^2_{H_1^\pm} / M^2_{H_2^\pm}) \). Due to the fact that \( \hat{f} \) contains a complex phase, the neutrino mass matrix is in general complex. It can be diagonalized by a unitary trans-
neutrino mixing angles can be written as
\[ U^T M_i U = \text{diag.}(m_1, m_2, m_3), \]  
where \( U \) is the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix. This transformation relates neutrinos in the flavor eigenbasis to those in the mass eigenbasis via
\[ v_i = U_{i\alpha} v_\alpha. \]  
Here \( i \) denotes family indices (\( e, \mu, \tau \)) and \( \alpha = 1, 2, 3. \)

Note that, from Eq. (18), the PMNS matrix \( U \) will be parameterized by 3 mixing angles and 3 phases. The three neutrino mixing angles can be written as
\[ s_{12}^2 = \frac{|Ue_2|^2}{1 - |Ue_3|^2}, \quad s_{13}^2 = |Ue_3|^2, \quad s_{23}^2 = \frac{|U\mu_3|^2}{1 - |Ue_3|^2}, \]  
with \( s_{ij} (c_{ij}) \) standing for \( \sin \theta_{ij} (\cos \theta_{ij}) \). One of the phases, known as the Dirac CP phase, is inferred from the Jarlskog invariance
\[ J = \text{Im} \left[ U_{i\alpha} U_{\rho\beta}^* U_{\epsilon\gamma} U_{\epsilon\gamma}^* \right]. \]  
The other two phases, unmeasured so far, are called Majorana phases.

Within this model, the neutrino mass matrix is expressed in terms of 6 free parameters. We will make a scan over 5 of them, which are two mixing angles of the charged lepton rotation matrices (chosen to be \( \theta_1 \) and \( \alpha_1 \)), \( r \) and \( \theta \) of the \( \tilde{f} \) couplings, and \( \tan \beta \). Then, by using Eqs. (17), (18), (20), and (21), we fit them to 5 neutrino observables, i.e., \( s_{12}^2, s_{23}^2, s_{13}^2, J, \) and \( R \equiv \Delta m_{\text{sol}}^2 / \Delta m_{\text{atm}}^2 \), whose numerical values are presented in Table 1. We demand that the aforementioned neutrino observables lie within their 2\( \sigma \) ranges from their central values. To avoid the unknown Majorana phases, we will not use Eq. (18) directly, but instead diagonalize the Hermitian matrix \( M^T_\nu M_\nu \), through \( U^T M_\nu U = \text{diag.}(m_1^2, m_2^2, m_3^2). \) Once we find a fit, the absolute neutrino mass can be determined using \( \kappa f e_\mu \) with the help of neutrino mass splittings.

We first analyze the neutrino mass matrix in Eq. (17) analytically. Note the mass matrix is approximately in the form
\[ (M_\nu)_{ij} \approx \kappa [m_\epsilon \tilde{f}_{ij} \tilde{Y}_{j3} + \tilde{f}_{j3} \tilde{Y}_{i3}] + m_\mu (\tilde{f}_{i2} \tilde{Y}_{j2} + \tilde{f}_{j2} \tilde{Y}_{i2}) \]
\[ \equiv (M^{(0)}_\nu + M^{(1)}_\nu)_{ij}. \]  
Since the Yukawa coupling \( \tilde{Y} \) is hierarchical with \( \tilde{Y}_{ij} \sim m_{\epsilon j} \), we see that the leading neutrino mass matrix is parametrically of order \( m_\epsilon^2 \) with a correction of order \( m_\mu^2 \). \( M^{(0)}_\nu \) is a rank 2 matrix with a vanishing (3,3) element. This implies \( \theta_{23} \) must lie in the first octant, see, e.g., [22,44]. Moreover, since \( M^{(0)}_\nu \) contains one vanishing eigenvalue, one would expect from the full mass matrix two sizable eigenvalues and one eigenvalue being suppressed by \( m_\mu^2 / m_\epsilon^2 \). These observations, taken at face value, seem to favor inverted ordering (IO) over normal ordering (NO) neutrinos with \( \theta_{23} < \pi / 4 \). However, we note that if the parameters of the model conspire to make \( M^{(1)}_\nu \) becomes comparable to \( M^{(0)}_\nu \), we could get a quasidegenerate neutrino mass compatible with both NO and IO. In this scenario, we could also accommodate \( \theta_{23} > \pi / 4 \).

We next analyze the neutrino mass matrix numerically by performing a scan on the model parameter space. In our scan, we will consider both of NO and IO. For the solar mass splitting, we use the usual definition \( \Delta m_{\text{sol}}^2 = m_2^2 - m_1^2 \), but for the atmospheric mass splitting, we adopt the convention used by Ref. [45], that is, \( \Delta m_{\text{atm}}^2 = m_3^2 - m_1^2 \) for NO and \( \Delta m_{\text{atm}}^2 = m_2^2 - m_3^2 \) for IO. Moreover, to reflect the current status of \( \theta_{23} \), which can lie either in the first or second octant, we also consider its two global fit values. From our scan, we find that neutrino masses are quasidegenerate with \( 0.6 \lesssim m_1 / m_3 \lesssim 0.8 \) for NO and \( 2 \lesssim m_1 / m_3 \lesssim 15 \) for IO.

Since the neutrino masses are quasidegenerate, the parameter space is further constrained by the bounds on neutrino masses. The most stringent constraint is provided by cosmological observations. Assuming the standard CDM model of the universe, Planck 2018 data places an upper bound
\[ \sum m_\nu < 0.26 \text{ eV at 95\% CL} \]  
[47]. The bound can even be stronger if one combines Planck data with baryonic acoustic oscillation (BAO) measurement, which puts bound on \( \sum m_\nu < 0.13 \text{ eV at 95\% CL} \) [47]. In Fig. 2, we present the viable parameter space consistent with neutrino oscillation parameters and the neutrino mass bound, i.e., \( \sum m_\nu < 0.13 \text{ eV} \). The viable parameter space is separated into two distinct regions, \( \tan \beta \lesssim 0.5 \) and \( \tan \beta \gtrsim 10 \). We will refer to them as the small and the large \( \tan \beta \), respectively. Note that the viable parameter space corresponds to the inverted neutrino masses ordering; the NO one is excluded by the Planck data, which is partly due to quasidegeneracy property of neutrino masses. Moreover, both octants of \( \theta_{23} \) are allowed. In Fig. 2, allowed regions corresponding to the first and second octants are depicted by red and blue points, respectively.

We conclude this section by determining the effective neutrino masses, relevant for laboratory-based neutrino experiments, for our viable parameter space. First, we discuss the total neutrino masses, which is estimated in this model to

|     | Normal ordering | Inverted ordering |
|-----|----------------|-----------------|
| \( s_{12}^2 \) | 0.304 ± 0.012 | 0.304 ± 0.013 |
| \( s_{23}^2 \) | 0.573 ± 0.018 | 0.575 ± 0.018 |
| \( s_{13}^2 \) | 0.02219 ± 0.00063 | 0.02238 ± 0.00063 |
| \( \Delta m_{\text{sol}}^2 / 10^{-5} \text{ eV}^2 \) | 7.42 ± 0.20 | 7.42 ± 0.20 |
| \( \Delta m_{\text{atm}}^2 / 10^{-3} \text{ eV}^2 \) | 2.517 ± 0.027 | 2.498 ± 0.028 |
| \( J_{\text{CP}} \) | -0.0084 ± 0.0143 | -0.0327 ± 0.00066 |

Table 1: Central values and the 1\( \sigma \) range for neutrino oscillation parameters obtained from Ref. [45] (also see [46]). In our analysis, we also consider the case of which \( \theta_{23} \) lies in the first octant.
be between 100 and 130 meV. This range is within reach of the Simon Observatory, which has projection of sensitivity up to 40 meV [29]. Second, experiments measuring the end point in beta decay spectrum use the effective electron neutrino mass, defined as $(m^{\text{eff}}_\nu)^2 = \sum_{i=1}^3 |U^2_{ei}| m^2_i$. For our viable parameter space, we have 48 meV $\lesssim m^{\text{eff}}_\nu < 56$ meV. This range is about 2 orders of magnitude below the current experimental limit of $m^{\text{eff}}_\nu < 1100$ meV at 90% CL reported by the KATRIN experiment [49]. Future experiments, like HOLMES [50], will push this bound to 200 meV, which is still one order of magnitude higher than our estimate. Third, the neutrinoless double beta decay $(0\nu\beta\beta)$ effective mass $m_{ee}$, defined as the $(1,1)$ component of the neutrino mass matrix $m_{ee} = |\sum_{i=1}^3 U^2_{ei} m_i|$, lies in a range of 20 meV $\lesssim m_{ee} < 40$ meV, see Fig. 3. This is just below the best experimental limit $m_{ee} < 61 - 165$ meV at 90% CL provided by the KamLAND-ZEN experiment [51], which searches for $0\nu\beta\beta$ in $^{136}\text{Xe}$. The interval in upper bound reflects the uncertainty in the nuclear matrix elements used in extracting the limit. This range of values should be testable by phase-II LEGEND experiment [30], able to probe value between 13 and 29 meV.

4 Constraints from lepton flavor violations

The lepton sector of the model contains tree-level FCNC. As a result, the parameter space of the model is strongly constrained by LFV measurements provided the extra Higgs bosons are not too heavy. In this section, we give analytic results for LFV observables. Some of the observables considered, such as the semileptonic decays, depend also on the quarks Yukawa couplings, albeit being flavor diagonal. To achieve this, we simply implement the natural flavor conser-
viation in the quark sector, hence forbidding tree-level FCNC of this sector.

4.1 LFV in leptonic decays

The LFV coupling matrix $\hat{Y}$ of the $H$ and $A$ lead to tree-level LFV decays of the lepton, $\ell \to \ell' \ell'' \ell'''$, via a Feynman diagram in Fig. 4. This decay can be easily described in term of effective 4-fermion operators [52]

$$\mathcal{L}_{\text{eff}} = \epsilon_{\ell \ell'' \ell''' L} P_X \ell (\bar{\ell} j P_Y \ell_k),$$

(23)

where $P_X, P_Y = P_L, P_R$ are the projection operators. The partial decay width for $\ell \to \ell_i \ell_j \ell_k$ is then given by

$$\Gamma_{\ell \to \ell_i \ell_j \ell_k} = \frac{m_\ell^5}{512(1 + 8 j)\pi^3} \left( \frac{7}{768} |\epsilon_{\ell i j k}|^2 + \frac{19}{768} |\epsilon_{\ell X Y Z}|^2 - \frac{5}{384} \text{Re} \left[ \epsilon_{\ell X Y Z}^* \epsilon_{\ell X Y Z} \right] \right),$$

(24)

where we have ignored the final state lepton masses. In the above expression, the sum over projections is implicit. That is $|\epsilon_{\ell X Y Z}|^2 = |\epsilon_{\ell X Y Z}|^2 + |\epsilon_{\ell X Y Z}|^2 + |\epsilon_{\ell X Y Z}|^2 + |\epsilon_{\ell X Y Z}|^2 + |\epsilon_{\ell X Y Z}|^2$. In the case that all leptons in the final state are of the same family, the above expression reduces to

$$\Gamma_{\ell \to \bar{\ell} \ell} = \frac{m_\ell^5}{1024\pi^3} \left( \frac{21}{m_H^4} + \frac{21}{m_A^4} - \frac{10}{m_H^4 m_A^4} \right) + \mathcal{O} \left( \frac{m_\ell^4}{m_\ell^4} \right).$$

(25)

Current experimental limits are $Br(\mu \to 3e) < 1.0 \times 10^{-12}$, $Br(\tau \to 3e) < 2.7 \times 10^{-8}$ and $Br(\tau \to 3\mu) < 2.1 \times 10^{-8}$ at 90% CL [53]. On the viable parameter space of the model, these constraints from tree-level leptonic LFV decays are much weaker compared to radiative LFV decays.

4.2 LFV in semileptonic decays

The LFV couplings of the $A$ also lead to a semi-leptonic decays $\tau \to \ell P$, where $P$ is a neutral pseudoscalar meson. The partial decay width, ignoring the final state lepton mass, is given by [52,54]

$$\Gamma_{\tau \to \ell \pi} = \frac{(m_\tau^2 - m_\ell^2)^2}{256\pi^3 m_\tau^3} \frac{m_A^4}{v^2 m_\ell^2} \left( \frac{\xi_A^2 - \xi_{\ell\ell'}^2}{4} \right) \left( \hat{Y}_{\ell\ell'}^2 + \hat{Y}_{\ell\ell}^2 \right),$$

(26)

$$\Gamma_{\tau \to \ell \eta^{(c)}} = \frac{(m_\tau^2 - m_\ell^2)^2}{256\pi^3 m_\tau^3} \frac{1}{v^2 m_\ell^2} \times \left( \frac{\xi_A^2 + \xi_{\ell\ell'}^2}{4} \right) \left( \hat{Y}_{\ell\ell'}^2 + \hat{Y}_{\ell\ell}^2 \right),$$

(27)

where $\xi_A^2$ is the coefficient defined just right after Eq. (16) and $f_\pi \simeq (130.2 \pm 0.8) \text{ MeV}$ [55] is the pion decay constant and

$$h_A^q = h_{Aq}^q = 0.001 \pm 0.003 \text{ GeV}^3,$$

$$h_A^q = -0.055 \pm 0.003 \text{ GeV}^3,$$

(28)

$$h_A^q = 0.068 \pm 0.005 \text{ GeV}^3,$$

are the hadronic matrix elements [56]. Since the decays proceed only via the pseudoscalar $A$, one can use limits on the branching ratio to constrain the LFV coupling $\hat{Y}$ as a function of $m_A$. We find that the decay $\tau \to \ell \eta^{(c)}$ gives the strongest bounds. Using $Br(\tau \to e \eta^{(c)}) < 9.2 \times 10^{-8}$ and $Br(\tau \to \mu \eta^{(c)}) < 6.5 \times 10^{-8}$ [53], the upper limits are

$$|\hat{Y}_{\ell\ell'}| \lesssim \frac{0.97}{|\xi_{\ell\ell'}^2|} \left( \frac{m_A}{130 \text{ GeV}} \right)^2,$$

(29)

$$|\hat{Y}_{\ell\ell}^c| \lesssim \frac{0.81}{|\xi_{\ell\ell'}^2|} \left( \frac{m_A}{130 \text{ GeV}} \right)^2,$$

(30)

where we have dropped the suppressed coupling $\hat{Y}_{\ell\ell}$ and the hadronic matrix element $h_A^q$. Notice that the bounds only depend on the strange quark coupling, $|\xi_{\ell\ell'}^2|$. As an illustrative example of the above bounds, consider the scenario where $m_A$ is around the weak scale and $0.3 \lesssim \tan \beta \lesssim 50$. If strange quark couples to $\Phi_2$, we have $|\xi_{\ell\ell'}^2| = \cot \beta$. In this case, the strongest upper bound on the LFV coupling $|\hat{Y}_{\ell\ell}|$ and $|\hat{Y}_{\ell\ell'}|$ is around $O(10^{-1})$. If instead, strange quark couples to $\Phi_1$ so that $|\xi_{\ell\ell'}^2| = \tan \beta$, the strongest upper bound becomes $O(10^{-2})$. In both cases, the strongest upper bound is well above the typical value of the couplings in our viable parameter space.

4.3 LFV in radiative decays

Equation (14) can also induce radiative LFV decays. For a general $\ell \to \ell' \gamma^*$ transition, its effective Lagrangian takes the form of (see, e.g., [57])

$$\mathcal{L}_{\ell\to\ell'\gamma^*} = \frac{e m_{\ell'}^2}{4\pi^2} \zeta_{\ell\ell'} \gamma_{\mu \nu} v_{\ell \ell'} (aL \ell + cR \ell P) A_\mu + cR \ell P \ell A^\nu \epsilon \gamma^2 \gamma_{\mu \nu} + \text{h.c.,}$$

(31)

where $q \equiv p_{\ell'} - p_{\ell}$ is the momentum transfer of the photon. Note the vanishing of the non-dipole term when the photon is on-shell. The form of Eq. (31) also guarantees the fulfillment of Ward’s identity. For $\ell \to \ell' \gamma$ process, its decay width is given by

$$\Gamma(\ell \to \ell' \gamma) = \frac{\alpha e_{\ell\ell} m_\ell^5}{64\pi^4} \left( |\epsilon_{\ell L}|^2 + |\epsilon_{\ell R}|^2 \right),$$

(32)
where \( \alpha_{em} \equiv e^2/4\pi \) is the QED fine structure constant. At one loop, such decay proceeds through penguin diagrams mediated by neutral and charged scalars. Its contributions to those Wilson’s coefficients are found to be [58,59]

\[
c_{L}^{1{\text{-loop}}} = -\frac{\hat{Y}_{\ell a}\hat{Y}_{\ell' a}}{24} \sum_{\phi = H,A,H^+} \frac{(-1)Q_{\phi}}{m_{\phi}^2} + \frac{\hat{Y}_{\ell a}\hat{Y}_{\ell' a}}{8} \sum_{\phi = H,A} \frac{(-1)C_{P}}{m_{\phi}^2} (3 + 2 \ln \frac{m_{\phi}^2}{m_{\phi}^2}) ,
\]

\[
c_{R}^{1{\text{-loop}}} = -\frac{\hat{Y}_{\ell a}^{\dagger}\hat{Y}_{\ell' a}}{24} \sum_{\phi = H,A} \frac{1}{m_{\phi}^2} + \frac{\hat{Y}_{\ell a}^{\dagger}\hat{Y}_{\ell' a}}{8} \sum_{\phi = H,A} \frac{(-1)C_{P}}{m_{\phi}^2} (3 + 2 \ln \frac{m_{\phi}^2}{m_{\phi}^2}) ,
\]

where \( Q_{\phi} \) and \( C_{P} \) denote the electric charge and charge-parity eigenvalue of the respected scalar. The first term in each equation above denotes the case where the chirality flip occurs in the external leg of the decaying lepton, while the second term denotes the flip in the internal leptons. The 2-loop contributions to \( c_{L,R} \) arise via Barr-Zee diagrams. This type of diagrams, especially those containing top-quark or \( W \) boson loop, may be equally or even more important than those of the 1-loop, due to the 2-loop suffering less chirality suppression and containing fewer \( \hat{Y}_{ab} \), which is typically smaller than the top-Yukawa or \( SU(2) \) gauge couplings [31,32,60,61]. In the present case, we adopt the decoupling limit, so the \( W \)-loop vanishes. This makes top loop the sole contribution to \( c_{L,R}^{2\text{-loop}} \), i.e.,

\[
c_{L}^{2\text{-loop}} = -\frac{\alpha_{em}}{\pi} \frac{\hat{Y}_{\ell \ell'} \hat{Y}_{\ell' \ell}}{v m_{\ell} m_{\ell'}} \left( \frac{2}{3} h_s + g(z_{1H}) - g(z_{1A}) \right) + \frac{1 - 4 s_{W}^2}{16 s_{W}^2 c_{W}^2} \left[ \frac{f(z_{1H},z_{1Z})}{\bar{f}(z_{1H},z_{1A})} - g(z_{1A},z_{1Z}) \right] ,
\]

\[
c_{R}^{2\text{-loop}} = -\frac{\alpha_{em}}{\pi} \frac{\hat{Y}_{\ell \ell'} \hat{Y}_{\ell' \ell}}{v m_{\ell} m_{\ell'}} \left( \frac{2}{3} h_s + g(z_{1H}) + g(z_{1A}) \right) + \frac{1 - 4 s_{W}^2}{16 s_{W}^2 c_{W}^2} \left[ \frac{f(z_{1H},z_{1Z})}{\bar{f}(z_{1H},z_{1A})} \right],
\]

where \( z_{ab} \equiv m_{a}^2/m_{b}^2 \). The first term in brackets is the photon-exchange contribution with loop functions expressed as

\[
f(z) = \frac{z}{2} \int_{0}^{1} dx \frac{1 - 2x(1 - x)}{x(1 - x) - z} \ln \frac{x(1 - x)}{z} ;
\]

\[
g(z) = \frac{z}{2} \int_{0}^{1} dx \frac{1}{x(1 - x) - z} \ln \frac{x(1 - x)}{z} ,
\]

whereas the second term denotes the \( Z \) contribution with functions defined as \( h(x, y) = (x h(y) - y h(x)) / (x - y) \). Due to the factor of \( 1 - 4 s_{W}^2 \), one should quickly recognize that the \( Z \) contribution is much suppressed compared to that of photon.

For the photon exchange, the functions \( f(z) \) and \( g(z) \) differ only on \( 1 - 2x(1 - x) \), which is a positive definite function. Therefore, the two functions will always come with the same sign. In the case of \( c_{L}^{2\text{-loop}} \), this leads to a destructive interference between \( H \) and \( A \) loops. For a certain mass set up, it could even lead to a complete cancelation, leaving us with the suppressed \( Z \) contribution. Along with the fact that \( |\hat{Y}_{\ell \ell'} / \hat{Y}_{\ell' \ell}| \sim m_{\ell} / m_{\ell'} \), we have \( c_{L}^{2\text{-loop}} \ll c_{R}^{2\text{-loop}} \).

For our viable parameter space obtained in Sect. 3, see Fig. 2, \( \mu \to e\gamma \) measurement places the strongest constraint on it. If the extra Higgs boson are not too heavy, \( \mu \to e\gamma \) decay generically rules out the part of parameter space with small tan \( \beta \). This is because the LFV coupling \( \tilde{Y}_{\ell \ell'} \) for the low tan \( \beta \) region is enhanced while its counterpart in the high tan \( \beta \) region is suppressed.

For concreteness, we consider benchmark scenario where only one neutral scalar is light with \( m_{\phi} = 130 \text{ GeV} \), where \( \phi = H / A \), while other Higgs are heavy. This choice of benchmark is relevant for collider search to be discussed in Sect. 5. To be specific, the heavier Higgs bosons are taken to be 650 GeV. The mass splitting in the Higgs sector chosen in this benchmark is consistent with perturbativity and the electroweak \( \rho \)-parameter constraints on the scalar quartic couplings [41]. Moreover, since our tan \( \beta \) are large, constraints from \( B \) physics discussed in Sect. 2 dictate that top quark couples to \( \Phi_2 \). The constraints from \( \mu \to e\gamma \) transition for our benchmark scenario is shown in Fig. 5. Again, red (blue) points correspond to \( \theta_{23} \) being in the first (second) octant. Note that the branching ratio \( \mu \to e\gamma \) decreases with increasing \( m_{\phi} \), as expected.
4.4 $\mu \rightarrow e$ conversion in nuclei

The process of $\mu \rightarrow e$ conversion in atomic nuclei also provides a stringent constraint on the LFV coupling, in particular $Y_{\mu e}$ and $Y_{e\nu}$. The formalism of such process has been discussed intensively in Refs. [57, 62]. We will focus on the coherent transition, where the nucleus final state does not change from its initial state during the conversion. Because of that, the effective Lagrangian at the quark level takes the form of

$$L_{\text{eff}} = \frac{e m_{\mu}}{8\pi^2} c_L \xi_{\alpha \beta} a_{\mu} F_{\alpha \beta} - \frac{1}{2} \sum_q \left[ (g_{\mu e}^q \bar{p} R \mu)(\bar{q} q) + (g_{L e}^q \bar{q} p \nu)(\bar{q} \gamma \nu q) \right]$$

$$+ (L \leftrightarrow R) + \text{h.c.} \quad (38)$$

This quark-level Lagrangian get converted into the nucleon-level one by evaluating the corresponding nuclear matrix elements. That is, $(N|q \bar{q} \Gamma_K q|N) \equiv G_{K}(q,N) \tilde{N} \tilde{\Gamma}_K \tilde{N}$ with $\tilde{\Gamma}_K = \{1, \gamma_\alpha\}$ and $N = p, n$. Other types of operators such as $\tilde{q} \gamma_5 q$, $\tilde{q} \gamma_\alpha \gamma_5 q$, and $\tilde{q} \sigma_{\alpha \beta} q$ have no contribution to the coherent process, hence they are omitted from Eq. (38).

The couplings of each operator can be directly determined. For vector operators, which are induced by the monopole term of Eq. (31), they are given by $g_{L.V}^{q} = -8\pi a_{em} Q_q a_{L,R}$, with $Q_q$ symbolizing the quark electric charge whereas $a_{L,R}$ being loop functions [58, 59]

$$a_L = -\frac{1}{144\pi^2} \tilde{Y}_{\mu e} \tilde{Y}_{\alpha e} \sum_{\phi=H, A} \frac{1}{m_\phi^2} \left(4 + 3 \ln \frac{m_q^2}{m_\phi^2}\right),$$

$$a_R = -\frac{1}{144\pi^2} \tilde{Y}_{\mu e} \tilde{Y}_{\alpha e} \sum_{\phi=H, A} \frac{1}{m_\phi^2} \left(4 + 3 \ln \frac{m_q^2}{m_\phi^2} - \frac{1}{m_H^2}\right) \quad \text{h.c.} \quad (39)$$

The scalar operators arise via tree-level exchange of $H$, which gives

$$g_{LS}^q = \left(\frac{2}{m_H^2}\right)(\bar{q} \gamma \nu q)(\bar{q} \gamma \nu q),$$

$$g_{RS}^q = \left(\frac{2}{m_H^2}\right)(\bar{q} \gamma \nu q)(\bar{q} \gamma \nu q) \quad (40)$$

Now the conversion rate can be calculated by using the formula given in Ref. [62] (see also Ref. [33]), that is,

$$\Gamma(\mu \rightarrow e \text{ conv.}) = \left| -\frac{e}{16\pi^2} c_L D + \tilde{g}_{LS}^q S(p) + \tilde{g}_{LS}^q S(n) + \tilde{g}_{L V}^q V(p) \right|^2$$

$$+ \left| -\frac{e}{16\pi^2} c_R D + \tilde{g}_{RS}^q S(p) + \tilde{g}_{RS}^q S(n) + \tilde{g}_{R V}^q V(p) \right|^2 \quad \text{(41)}$$

where $D$, $S^{(p,n)}$, and $V^{(p)}$ are the corresponding overlap integrals for each operator. Their definitions and numerical values (in the unit of $m_{\mu}^{-2}$) for various nuclei have been tabulated in [62]. The coefficients $\tilde{g}_s$’s stand for effective couplings to nucleon $N$, which are defined as

$$\tilde{g}_{L.K,RK}^q = \sum_{q} G_{K}^{(q,N)} s_{L.K,RK}^q \quad (42)$$

One should note that there is a subtle difference between the vector and scalar operators. The vector operators characterize the difference between the number of quarks and antiquarks. Thus the summation runs over the valence quarks with $G_{V}^{(u,p)} = G_{V}^{(d,n)} = 2$ and $G_{V}^{(u,s)} = G_{V}^{(d,p)} = 1$. This leads to the vanishing effective neutron couplings $g_{LS}^{(u,n)} = g_{RS}^{(u,n)} = 0$. The scalar operators, on the other hand, quantify each quark contribution to nucleon mass via $G_{S}^{(u,N)} = f^{(q,N)} m_N/m_q$. Therefore, the sum runs for all quark flavors, including those of heavy flavors ($Q = c, b, t$) mediating the gluon exchange. For light quarks $q = u, d, s$, the parameters $f^{(q,N)}$ have been determined by lattice calculations and are presented in Table 2, while for heavy quarks, it is determined through

$$f^{(Q,N)} = \frac{2}{27} \left(1 - \sum_{q} f^{(q,N)}\right) \quad (43)$$

It should be noted that in both operators $G_{S,V}^{(u,q)} = G_{S,V}^{(d,q)}$ and $G_{S,V}^{(u,q)} = G_{S,V}^{(d,q)}$, reflecting the isospin invariance of the nucleon.

For our purpose, we seek the strongest constraint, coming from $\mu \rightarrow e$ conversion on gold nucleus. The branching fraction of the rate is $\Gamma(\mu \rightarrow e \text{ conv.}/\Gamma(\text{captured}) < 7 \times 10^{-13}$
The new contribution is given by 
\[ \text{Eq. (32)} \]

physics, which, in present model, arises via scalar exchanges. The discrepancy can be attributed to a possible presence of new physics. In this section, we study the LHC phenomenology of the new neutral scalars, i.e., \( \hat{H} \) and \( \hat{A} \). For light \( \hat{H} \) and \( \hat{A} \) within the collider reach, the parameter space consistent with both neutrino data and LFV constraints corresponds to \( \tan \beta \gtrsim 10 \). Firstly, we will discuss the case of the quarks coupled universally to one of the scalar doublet. Since only high value of \( \tan \beta \) is allowed, constraints from \( B \) physics \([42]\), such as \( B_d \to \mu \nu \), restrict the case where quarks couple only to \( \Phi_2 \) provided that \( m_{H^+} \lesssim 1 \text{ TeV} \) \(^1\). Since we are working in the decoupling limit, the main production channels for both the \( H \) and the \( A \) are gluon fusion via the top quark loop. They can also be produced in the association with the top pair or the bottom pair, but we found that these channels are insignificant compared with gluon fusion.

The decay cross-section can be calculated from the would-be SM Higgs boson by a rescaling the quark Yukawa couplings with \( \xi_{H^0} = m_{H^0} \). In our analysis, we use the would-be SM Higgs boson cross-sections provided in Ref. \([69]\). On the other hand, the pseudoscalar-gluon fusion cross-sections is larger compared to that of \( H \) with the same mass, with a loop factor of

\[ \frac{\sigma_{ggF}^A}{\sigma_{ggF}^H} \sim \frac{F(x)}{1 + (1 - x)F(x)}, \]

where \( x = 4m_t^2/m_{H^0}^2 \) and

\[ F(x) = \begin{cases} \sin^{-1} \sqrt{1/x}, & x \geq 1, \\ -\frac{1}{4} \left[ \ln \left( \frac{1 + \sqrt{1 - x}}{1 - \sqrt{1 - x}} \right) - i\pi \right]^2, & x < 1. \end{cases} \]

Since \( \tan \beta \) is large over the parameter space of interest, both the \( H \) and the \( A \) production cross-sections and decay rates to quarks are suppressed. However, the Yukawa couplings of \( H \) and \( A \) to the leptons can be enhanced. The couplings \( \hat{Y}_{\mu\mu} \) and \( \hat{Y}_{\tau\tau} \) in the case of allowed parameter space are shown in Fig. 6. From the figure we see that \( \hat{Y}_{\mu\mu} \) is enhanced compared with the SM Higgs Yukawa coupling. The value of \( \hat{Y}_{ee} \) also gets comparable enhancements compared with its

---

\(^1\) The flavor phenomenology of quarks that couple universally to \( \Phi_1 \) with high value of \( \tan \beta \) would be the same with the type-I 2HDM \([41]\) with low \( \tan \beta \), hence excluded by the \( B \) physics constraints.
would be SM-value. Moreover, since this region is safe from the LFV constraints, in this region the LFV couplings (\(\hat{Y}_{ee}, \hat{Y}_{e\tau}, \hat{Y}_{\mu\tau}\)) are relatively small. As a result, the relevant LHC signature \(H/A \rightarrow \mu^+\mu^-\) is considerably large if the scalars are within the LHC reach. On the other hand, while \(\hat{Y}_{ee}\) gets some enhancements, the branching fractions of \(H/A \rightarrow e^+e^-\) are still relatively small, hence less relevant in constraining the model.

At the 13-TeV LHC, CMS has searched for a supersymmetric Higgs boson decaying to \(\mu^+\mu^-\). Based on 35.9 fb\(^{-1}\) data, CMS reports the 95\% upper limit on the gluon fusion production cross-section \((\sigma)\) times branching ratio (BR) into \(\mu^+\mu^-\), ranging from 44.7 fb at \(m_H = 130\) GeV to 17.5 fb at \(m_H = 200\) GeV [36]. We translate the CMS bounds into the constraints on our parameter space in the case \(m_{H/A} = 130\) GeV and \(m_{H/A} = 200\) GeV, see Fig. 7. In the two plots we independently show bounds for each \(H\) and \(A\). One can see that if the pseudoscalar \(A\) is the lightest, the LHC bounds become more stringent due to larger cross section. In either case, the lower value of \(\tan\beta\), corresponding to the red points (see Fig. 2), is disfavored. As shown in Fig. 6, the blue points have larger values of \(\hat{Y}_{\mu\mu}\) hence larger branching fraction to the \(\mu^+\mu^-\). However, the production cross-section is suppressed by \(1/\tan\beta^2\), which weakens the collider bound. The LHC bounds start to be irrelevant once other decay channels, such as \(A \rightarrow hZ\) and \(H \rightarrow hh\), open up and dominate the branching fractions.

For light enough \(H\) or \(A\), collider bounds are stronger than the LFV bounds with only little parts of allowed parameter space left. We expect that these bounds will get stronger once the full LHC Run II data are analyzed with possibility of ruling out the whole region at the end of LHC runs.

Since only large \(\tan\beta\) allowed from LFV bounds, there is only one more additional scenario for light scalar that satisfies the bounds from rare \(B\) decays. This scenario is when \(\Phi_1\) couples to \(d_Rj\) and \(\Phi_2\) to \(u_Rj\), as in type-II 2HDM. In this scenario, the Yukawa coupling to the down-type quark is enhanced by a factor of \(\tan\beta\) while the coupling to the up-type quark is reduced by the same factor. Therefore, while \(\hat{Y}_{HH}\) is increased to the would-be SM coupling, the value of the Yukawa coupling is still smaller than the value of \(\hat{Y}_{b\tilde{b}}\). The branching fractions of \(H\) and \(A\) are then dominated by the decay to \(b\tilde{b}\). For low value of \(m_H\), CDF and D0 presented combined analysis for this channel for \(90\) GeV < \(m_H < 300\) GeV [70]. Their bounds are rather weak, e.g. \(\tan\beta \lesssim 40\) for \(m_H = 130\) GeV. CMS and ATLAS also performed some analysis in this search channel [71,72]. However, the collaborations only show the bounds for \(m_H \gtrsim 300\) GeV. An analysis at lower mass will be interesting on probing this scenario of quark-scalar couplings.

**6 Conclusions and discussions**

We have analyzed the Zee model for radiative neutrino mass generation with a family dependent discrete \(Z_2\) symmetry imposed on the left-handed lepton doublets. The \(Z_2\) symmetry dictates the resulting neutrino masses, consistent with oscillation data, to be quasidegenerate with \(\tan\beta\) predicted to be in a restrictive range—\(\tan\beta \lesssim 0.5\) (small \(\tan\beta\)) and \(\tan\beta \gtrsim 10\) (large \(\tan\beta\)), see Fig. 2. Combining oscillation data with constraints from cosmology on the sum of neutrino masses, we found that the viable parameter space supports only the inverted neutrino masses ordering (IO).

The viable parameter space of the model considered here is well placed to be probed by both the ongoing and the planned experiments. On the neutrino mass front, see Fig 3, our scenario predicts the sum of the three active neutrino mass, \(\sum m_\nu\), to lie in the range \(0.10\) eV \(\lesssim \sum m_\nu \lesssim 0.13\) eV. This is well within reach of the Simon Observatory, which
projects to measure the neutrino mass sum with a sensitivity of 40 meV [29]. The effective Majorana mass, \( m_{\nu_{e}} \), is predicted to be in a range of 20–40 meV, putting it within reach of the planned phase-II of the LEGEND experiment, which is projected to constrain \( m_{\nu_{e}} \lesssim 13–29 \) meV [30]. The effective electron neutrino mass, \( m_{\nu_{e}}^{eff} \), on the other hand, lies between 48 and 53 meV, which is an order of magnitude below the projection \( m_{\nu_{e}}^{eff} < 200 \) meV at 90% CL of both the KATRIN and the HOLMES experiments [49,50].

The family dependent \( Z_2 \) symmetry in the lepton sector give rise to flavor violating couplings for the neutral scalars \( H \) and \( A \). In the decoupling limit, the LFV couplings of \( H \) and \( A \) are given by \( \tilde{Y} \) described in Sect. 2. In particular, the LFV coupling \( \tilde{Y}_{e\mu} \) is enhanced in the low tan \( \beta \) part of the viable parameter space identified by neutrino data. Hence, provided that \( H \) and \( A \) are a few TeV or lighter, \( \mu \rightarrow e\gamma \) and \( \mu \rightarrow e \) conversion measurements firmly rule out the low tan \( \beta \) region. For the high tan \( \beta \) region, next generation experiments are capable of probing it. The Mu2e experiment is expected to reach a sensitivity of \( 3 \times 10^{-17} \) on the \( \mu \rightarrow e \) conversion rate [34], which is sensitive enough to probe part of our parameter space, see Fig. 5. On the other hand, the MEGII experiment is projected to reach \( 6 \times 10^{-14} \) sensitivity on the branching ratio \( Br(\mu\rightarrow e\gamma) \) [35], just above the expected branching ratio in our scenario.

In the case where only \( \Phi_2 \) couples to quarks, especially when \( H \) and \( A \) are light enough, e.g., \( m_{H/A} \lesssim 200 \) GeV, LHC searches on \( H/A \rightarrow \mu^+\mu^- \) provide a more stringent constraint on the parameter space of interest than those of LFV, thanks to the enhanced \( \tilde{Y}_{\mu\mu} \) coupling, see Fig. 6. By recasting the CMS search for a Higgs decaying to \( \mu^+\mu^- \) in supersymmetric context, we found that LHC data typically exclude part of the relevant parameter space, see Fig. 7, which corresponds to region having somewhat lower tan \( \beta \) (i.e., red points, see Fig. 2). It is interesting to note that the CMS search has been performed on a small subset of Run II data. In the event that CMS analyzes the full dataset, they could rule out a light Higgs in this particular scenario, or even discover it.

Finally, we note that the neutrino mass scale is related to the charged scalar mass. Parametrically, \( m_{\nu} \) scales as \( \mu \tilde{f} m_{\nu}^{2}/m_{H^{+}}^{2} \), where \( m_{\ell} \) and \( m_{H^{+}} \) are the charged lepton and charged scalar masses. The tri-linear coupling \( \mu \) and the Yukawa coupling \( \tilde{f} \) cannot be arbitrary large since the former leads to fine-tuning on the light Higgs boson mass, while the latter is constrained by perturbativity. Specifically, taking \( m_{\nu} \sim 10 \) meV, \( \mu \lesssim 1.5 \) TeV [73] and \( |\tilde{f}| < 4\pi \), we obtain \( m_{H^{+}} \lesssim 100 \) TeV.

**Acknowledgements** The work of RP was supported by the Parahyangan Catholic University under Grant no. III/LPPM/2022-0271/P. The work of JJ was supported in part by the Indonesian Institute of Sciences (which later became the National Research and Innovation Agency) through Research Support Facility Program. PU would like to thank P. Ongmongkolkul for helpful discussions on the iminuit package [74]. PU also acknowledges the National Science and Technology Development Agency; National e-Science Infrastructure Consortium, Chulalongkorn University, and the Chulalongkorn Academic Advancement into Its 2nd Century Project (Thailand) for providing computing infrastructure that has contributed to the results reported within this paper. The work of PU was supported in part by the Srinakharinwirot University under Grant no. 035/2565 and the Thailand Toray Science Foundation.

**Data Availability Statement** This manuscript has no associated data or the data will not be deposited. [Authors’ comment: This is a theoretical study and there is no experimental data.]

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Funded by SCOAP3.

**Appendix: Diagonalization of the charged lepton mass matrix**

Here we present the expression for parameters of mixing matrices \( O_{L,R} \) that diagonalize the charged-lepton mass matrix, as in Eq. (10). Those parameters are defined according to the decomposition given in Eq. (11), with \( \theta_{i} \rightarrow \alpha_{i} \) for \( O_{R} \). As mentioned in the main text, four of the mixing parameters can be expressed in terms of the other two, chosen to be \( \theta_{1} \) and \( \alpha_{3} \). Their expressions are given below.

\[
\begin{align}
\sin \theta_{3} &= \frac{s_{\alpha_{3}} m_{\ell_{i}} m_{\ell_{i}} - c_{\alpha_{3}} s_{\alpha_{2}} s_{\theta_{1}} c_{\theta_{1}} (m_{\ell_{i}}^{2} - m_{\mu}^{2})}{\sqrt{c_{\alpha_{3}}^{2} (c_{\theta_{1}}^{2} m_{\mu}^{2} + s_{\theta_{1}}^{2} m_{\ell_{i}}^{2}) (s_{\alpha_{2}}^{2} (c_{\theta_{1}}^{2} m_{\ell_{i}}^{2} + s_{\theta_{1}}^{2} m_{\mu}^{2}) + c_{\alpha_{3}}^{2} m_{\mu}^{2}) (s_{\alpha_{3}} m_{\ell_{i}} m_{\ell_{i}} - c_{\alpha_{3}} s_{\alpha_{2}} s_{\theta_{1}} c_{\theta_{1}} (m_{\ell_{i}}^{2} - m_{\mu}^{2}))}} \tag{A1} \\
\sin \theta_{2} &= \frac{m_{\mu} s_{\theta_{1}} (c_{\alpha_{3}} s_{\alpha_{2}} c_{\alpha_{1}} - c_{\alpha_{3}} s_{\alpha_{1}}) + m_{\ell_{i}} s_{\theta_{1}} (c_{\alpha_{3}} s_{\alpha_{2}} s_{\alpha_{1}} + s_{\alpha_{3}} c_{\alpha_{1}})}{\sqrt{(m_{\mu} s_{\theta_{1}} (c_{\alpha_{3}} s_{\alpha_{2}} c_{\alpha_{1}} - c_{\alpha_{3}} s_{\alpha_{1}}) + m_{\ell_{i}} s_{\theta_{1}} (c_{\alpha_{3}} s_{\alpha_{2}} s_{\alpha_{1}} + s_{\alpha_{3}} c_{\alpha_{1}}))^{2} + c_{\alpha_{3}}^{2} c_{\alpha_{2}}^{2} m_{\ell_{i}}^{2}}} \tag{A2}
\end{align}
\]
\[ s_{a_1} = \frac{m_{\mu} \mu}{\sqrt{c_{a_1}^2 m_\mu^2 + m_1^2 c_{a_1}^2}}, \]
\[ s_{a_2} = -\frac{c_{a_1} s_{a_1} c_0 n_2 (m_2^2 - m_1^2)}{s_{a_1} n_1 m_\tau (c_{a_1}^2 m_2^2 + s_{a_1}^2 m_\mu^2 - m_\tau^2)).} \]
67. B. Abi et al., Muon g-2. Phys. Rev. Lett. 126, 141801 (2021). arXiv:2104.03281 [hep-ex]

68. S. Borsanyi et al., Nature 593, 51 (2021). arXiv:2002.12347 [hep-lat]

69. J.R. Andersen et al. (LHC Higgs Cross Section Working Group), (2013), 10.5170/CERN-2013-004. arXiv:1307.1347 [hep-ph]

70. T. Aaltonen et al., CDF, D0. Phys. Rev. D 86, 091101 (2012). arXiv:1207.2757 [hep-ex]

71. G. Aad et al. (ATLAS), Phys. Rev. D 102, 032004 (2020). arXiv:1907.02749 [hep-ex]

72. A.M. Sirunyan et al., CMS. JHEP 08, 113 (2018). arXiv:1805.12191 [hep-ex]

73. Y. Cai, J. Herrero-García, M.A. Schmidt, A. Vicente, R.R. Vlkas, Front. Phys. 5, 63 (2017). arXiv:1706.08524 [hep-ph]

74. H. Dembinski, P. Ongmongkolkul, C. Deil, D.M. Hurtado, H. Schreiner, M. Feickert, Andrew, C. Burr, J. Watson, F. Rost, A. Pearce, L. Geiger, B.M. Wiedemann, C. Gohlke, Gonzalo, J. Drotleff, J. Eschle, L. Neste, M.E. Gorelli, M. Baak, O. Zapata, odidev, scikit-hep/iminuit: v2.8.4 (2021)