

Scalar doublet models confront τ and b anomalies

James M. Cline*

*Department of Physics, McGill University, 3600 Rue University, Montréal, Québec, Canada H3A 2T8
and Niels Bohr International Academy and Discovery Center, Niels Bohr Institute,
University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*

(Received 22 December 2015; published 7 April 2016)

There are indications of a possible breakdown of the standard model, suggesting that τ lepton interactions violate flavor universality, particularly through B meson decays. *BABAR*, *Belle*, and *LHCb* report high ratios of $B \rightarrow D^{(*)}\tau\nu$. There are long-standing excesses in $B \rightarrow \tau\nu$ and $W \rightarrow \tau\nu$ decays and a deficit in inclusive τ to strange decays. We investigate whether two Higgs doublet models with the most general allowed couplings to quarks, and a large coupling to τ leptons, can explain these anomalies while respecting other flavor constraints and technical naturalness. Fits to $B \rightarrow D^{(*)}\tau\nu$ data require couplings of the new Higgs doublet to down-type quarks, opening the door to many highly constrained flavor-changing neutral current processes. We confront these challenges by introducing a novel ansatz that relates the new up- and down-type Yukawa couplings, and demonstrate viable values of the couplings that are free from fine-tuning. LEP and LHC searches for new Higgs bosons decaying via $H^0 \rightarrow \tau^+\tau^-$ and $H^\pm \rightarrow \tau^\pm\nu$ allow a window of masses $m_H = [100\text{--}125]$ GeV and $m_\pm \sim 100$ GeV that is consistent with the predictions of our model. Contamination of the $W^+ \rightarrow \tau^+\nu$ signal by $H^+ \rightarrow \tau^+\nu$ decays at LEP could explain the apparent $W \rightarrow \tau\nu$ excess. We predict that the branching ratio for $B_s \rightarrow \tau^+\tau^-$ is not far below its current limit of several percent. An alternative model with decays of $B \rightarrow D^{(*)}\tau\nu_s$ to a sterile neutrino is also argued to be viable.

DOI: 10.1103/PhysRevD.93.075017

I. INTRODUCTION

The origin of the Standard Model (SM) flavor structure is a mystery, and any model predicting new patterns of flavor violation must confront very strong experimental bounds. This has given rise to the minimal flavor violation (MFV) paradigm [1–4] as a guide for constructing new physics beyond the SM that has been highly influential in recent years. MFV is extremely effective for suppressing flavor-changing neutral currents (FCNCs). In this work we confront some hints of new physics for which MFV seems generally too strong to accommodate the observed deviations. We are thus motivated to consider an alternative that can allow for larger nonstandard flavor effects.

Several recent experiments indicate possible deviations from the SM in some flavor-specific observables involving τ leptons. *BABAR*, *Belle*, and *LHCb* report the ratios $R(D)$ and $R(D^*)$, defined as

$$R(X) = \frac{\mathcal{B}(\bar{B} \rightarrow X\tau\bar{\nu})}{\mathcal{B}(\bar{B} \rightarrow X\ell\bar{\nu})}, \quad (1)$$

where $\ell = e, \mu$. The summary of the SM predictions and the measurements is shown in Table I. The reported measurements are consistent with each other and with previously reported results [5–7]. The measurements are also consistent with e/μ universality. However, the naively combined experimental value for the ratio $R(D^{(*)})$ differs from the SM prediction by more than 3σ .

There have been other hints of a breakdown of lepton flavor universality between τ and e/μ . The measured decay rate of $B \rightarrow \tau\nu$ displays some tension with the SM prediction. Although a recent measurement by *Belle* [8] has reduced the discrepancy to the level of 1.7σ , the current world average measurement remains a factor of 1.5 higher than the SM prediction (see Ref. [9] for a recent review). The observed rate of $W \rightarrow \tau\nu$ is also in tension with the standard model predictions: the LEP measurement is $\sim 10\%$ above the SM value, at 2.4σ significance. The inclusive decays of τ to strange quarks yield a value of the CKM matrix element V_{us} significantly lower than that required for unitarity [10].

A number of authors have studied $B \rightarrow D^{(*)}\tau\nu$ in the context of type-III two Higgs doublet models (2HDMs), in which the most general couplings of fermions to both doublets are allowed, as well as model-independent analyses that include this framework [11–17]. There are two

TABLE I. Summary of experimental and predicted values for $R(D)$ and $R(D^*)$.

| ... | $R(D)$ | $R(D^*)$ |
|----------------------|-----------------------------|-----------------------------|
| SM | 0.297 ± 0.017 | 0.252 ± 0.005 |
| <i>Belle</i> [6] | $0.375 \pm 0.064 \pm 0.026$ | $0.293 \pm 0.038 \pm 0.015$ |
| <i>BABAR</i> [5] | $0.440 \pm 0.058 \pm 0.042$ | $0.332 \pm 0.024 \pm 0.018$ |
| <i>LHCb</i> [7] | | $0.336 \pm 0.027 \pm 0.030$ |
| Experimental average | 0.408 ± 0.050 | 0.321 ± 0.021 |

*jcline@physics.mcgill.ca

possible operators contributing to the hadronic part of these processes, mediated by charged Higgs exchange, proportional to $C_{S_R}^{cb} \bar{c}_L b_R$ and $C_{S_L}^{cb} \bar{c}_R b_L$, respectively. (For simplicity we assume that the coefficients are real in the present work.) Some studies [11,13–15,17] found that $C_{S_L}^{cb}$ by itself is sufficient to get a good fit to the observed decay rates. However, several recent analyses [16,18] obtain best-fit regions requiring $C_{S_R}^{cb} \sim -C_{S_L}^{cb}$. In particular, these studies use not only the total rates but also the differential decay distributions as inputs to their fits, finding that $C_{S_L}^{cb}$ by itself does not fit the decay spectra.

This difference is crucial for model building, since having $C_{S_L}^{cb} \neq 0$ only requires that the new up-type Yukawa matrix ρ_u (which couples mainly to the nonstandard Higgs doublet) is important, while keeping the down-type couplings $\rho_d \cong 0$. If $C_{S_R}^{cb}$ is also large, then $\rho_d \sim \rho_u$, making it much more challenging to satisfy constraints on FCNCs. The purpose of this paper is to see how far one can go toward overcoming these challenges, within the context of 2HDMs, if the indication for $C_{S_R}^{cb} \sim -C_{S_L}^{cb}$ persists in future analyses.

We will show that some of the flavor challenges can be addressed if ρ_u and ρ_d are related to each other in a particular way that involves the CKM matrix. This is a new ansatz for helping to give flavor protection to type III 2HDMs, which might be of interest more generally than for the particular applications that motivated us here. It is quite different from MFV, yet it appears to facilitate adequate control over FCNCs to make the theory viable, especially in the down-quark sector where the constraints are strongest.

The model is strongly constrained by LEP and LHC searches for the new charged Higgs decaying into $\tau\nu$ and the neutral one decaying to $\tau^+\tau^-$. We find a window $\sim[100\text{--}125]$ GeV of allowed masses for the new scalars that passes the collider constraints while allowing for an explanation of the B decay anomalies. Scalars of these masses are just beyond the kinematic reach of LEP, while being in a region of low efficiency for LHC searches, if their couplings to quarks are sufficiently small.

The outline of the paper is as follows. In Sec. II we define the model. In Sec. III we derive constraints on the new Yukawa couplings ρ_u, ρ_d arising from fits to $R(D)$ and $B \rightarrow \tau\nu$, and a few key flavor-sensitive decays. Section IV examines the collider constraints determining the allowed mass range of the new ~ 100 GeV Higgs bosons. In Sec. V we present a novel ansatz relating ρ_u and ρ_d that allows these constraints to be satisfied in a controlled way. It is a linear relation involving the Cabibbo-Kobayashi-Maskawa (CKM) matrix V , a diagonal unitary matrix U , and an $O(1)$ parameter η : $\rho_u^\dagger V = \eta UV \rho_d$.

In Sec. VI we calculate observables from meson oscillations that most strongly constrain the scenario, while in Sec. VII we show that rare decay processes that might challenge it are within the experimental limits. Section VIII

obtains a numerical fit to the couplings ρ_d^{ij} that determine ρ_u through our ansatz. In Sec. IX we estimate the size of loop contributions to the nonstandard Yukawa and Higgs couplings to establish technical naturalness of the model. In Sec. X we outline a microscopic model that naturally implements the “charge transformation” mechanism for relating ρ_u and ρ_d in the manner of our ansatz. We outline an alternative version of the model in Sec. XI, where the leptonic coupling ρ_e is replaced by a coupling ρ_ν to neutrinos, assuming a light sterile neutrino in the anomalous decays of B , rather than ν_τ . This model is less constrained by LHC searches for the neutral Higgs. Conclusions are given in Sec. XII. Details of the sterile neutrino version of the model are given in the Appendix.

II. THE MODEL

We begin with the most general two Higgs doublet model, where H_1 and H_2 are the doublets, each coupling to all the fermions. They have the conventional decomposition

$$\begin{aligned} H_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}H_1^+ \\ v + H_1^r + iH_1^i \end{pmatrix}, \\ H_2 &= \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}H_2^+ \\ H_2^r + iH_2^i \end{pmatrix} \end{aligned} \quad (2)$$

in terms of the real and imaginary parts of the neutral components. The Yukawa coupling Lagrangian is

$$\begin{aligned} \mathcal{L}_Y &= -\bar{Q}_L \hat{y}_u \tilde{H}_1 u_R - \bar{Q}_L \hat{y}_d H_1 d_R - \bar{L}_L \hat{y}_e H_1 e_R \\ &\quad - \bar{Q}_L \hat{\rho}_u \tilde{H}_2 u_R - \bar{Q}_L \hat{\rho}_d H_2 d_R - \bar{L}_L \hat{\rho}_e H_2 e_R + \text{H.c.}, \end{aligned} \quad (3)$$

where flavor, color, and $SU(2)_L$ indices have been suppressed, and $\tilde{H}_i^a = \epsilon_{ab} H_i^{b*}$. The scalar Lagrangian is given by

$$\mathcal{L}_S = |D^\mu H_1|^2 + |D^\mu H_2|^2 - V(H_1, H_2), \quad (4)$$

where the potential is defined as

$$\begin{aligned} V &= \lambda \left(H_1^\dagger H_1 - \frac{v^2}{2} \right)^2 + m_2^2 (H_2^\dagger H_2) \\ &\quad + (m_{12}^2 (H_1^\dagger H_2) + \text{H.c.}) + \lambda_1 (H_1^\dagger H_1) (H_2^\dagger H_2) \\ &\quad + \lambda_2 (H_1^\dagger H_2) (H_2^\dagger H_1) + [\lambda_3 (H_1^\dagger H_2)^2 + \text{H.c.}] \\ &\quad + [\lambda_4 (H_1^\dagger H_2) (H_2^\dagger H_2) + \lambda_5 (H_2^\dagger H_1) (H_1^\dagger H_1) + \text{H.c.}] \\ &\quad + \lambda_6 (H_2^\dagger H_2)^2. \end{aligned} \quad (5)$$

In this basis of fields, H_2 has no vacuum expectation value, requiring the condition $m_{12}^2 + \lambda_5^* v^2/2 = 0$. This is just a choice of field coordinates, which in general can always be achieved by doing a rotation (conventionally denoted by angle β as well as a possible rephasing) between H_1 and H_2 ; however, in Sec. X we will argue that the Yukawa couplings were generated directly in this basis, so that the \hat{y} and $\hat{\rho}$

matrices can naturally have very different magnitudes and structures.

For simplicity we will assume that the potential (5) is CP -conserving, so that there is no mixing between scalars and the pseudoscalar. The rotation between the Higgs basis fields and the CP -even mass eigenstates is

$$\begin{pmatrix} H_2^r \\ H_1^r \end{pmatrix} = \begin{pmatrix} c_{\beta\alpha} & -s_{\beta\alpha} \\ s_{\beta\alpha} & c_{\beta\alpha} \end{pmatrix} \begin{pmatrix} h \\ H \end{pmatrix}. \quad (6)$$

Here we have used notation that is conventional in 2HDMs, such that for $s_{\beta\alpha} \cong 1$, the SM-like Higgs boson h is mostly H_1^r . For small $|c_{\beta\alpha}|$, the mixing angle is approximately determined by

$$c_{\beta\alpha} \cong \frac{\lambda_5 v^2}{2(m_h^2 - m_H^2)}, \quad (7)$$

where the SM-like, new neutral, and charged Higgs boson masses are, respectively,

$$\begin{aligned} m_h^2 &\cong 2\lambda v^2, \\ m_H^2 &\cong m_2^2 + \frac{1}{2}(\lambda_1 + \lambda_2 + 2\lambda_3)v^2, \\ m_A^2 &\cong m_H^2 - 2\lambda_3 v^2, \\ m_{\pm}^2 &= m_2^2 + \frac{1}{2}\lambda_1 v^2. \end{aligned} \quad (8)$$

These approximations are valid for small mixing. We will also require that the splittings between masses of the neutral scalars H, A (CP -even and CP -odd, respectively) are small, so that they can be regarded as components of a complex neutral field for most purposes. This not only simplifies the model but also proves useful for suppressing some FCNC effects as we will show. Small splittings are consistent with $|c_{\beta\alpha}| \ll 1$ and $|\lambda_3| \ll 1$, since it can be shown (without any approximation) that

$$m_H^2 - m_A^2 = c_{\beta\alpha}^2(m_H^2 - m_h^2) + 2\lambda_3 v^2 \quad (9)$$

(see, for example, [19]). We will therefore assume that $|\lambda_3| \ll 1$ in addition to $|c_{\beta\alpha}| \ll 1$. Although λ_2 could *a priori* be relatively large, in this work we will be interested in masses of order $m_H \lesssim m_h$ and $m_{\pm} \sim 100$ GeV, corresponding to $\lambda_2 \lesssim 0.2$. Electroweak precision data [see Eq. (10.26) of Ref. [20]] would allow for larger splittings, with m_H as large as 175 GeV for $m_{\pm} \sim 100$ GeV. The couplings λ_4, λ_6 play no direct role for our predictions but can be relevant for understanding the expected size of radiative corrections to the λ_i couplings, as we will discuss in Sec. IX. Vacuum stability requires that $\sqrt{\lambda\lambda_6/2} > -\lambda_2$ if $\lambda_2 < 0$.

As usual, biunitary rotations on the quark fields in \mathcal{L}_Y diagonalize $\hat{y}_u, \hat{y}_d, \hat{y}_e$, with the scalar doublets still in the

Higgs basis. The subsequent rotation (6) then brings \mathcal{L}_Y to the form

$$\mathcal{L}_Y = -\frac{1}{\sqrt{2}} \sum_{\substack{\phi=h,H,A \\ f=u,d,e}} y_{\phi ij}^f \bar{f}_i \phi P_R f_j + \text{H.c.}, \quad (10)$$

$$y_{hij}^f = s_{\beta\alpha} \frac{\sqrt{2}m_f^i}{v} \delta_{ij} + c_{\beta\alpha} \rho_f^{ij}, \quad (11)$$

$$y_{Hij}^f = c_{\beta\alpha} \frac{\sqrt{2}m_f^i}{v} \delta_{ij} - s_{\beta\alpha} \rho_f^{ij}, \quad (12)$$

$$y_{Aij}^f = \rho_f^{ij} \times \begin{cases} +i, & f = u \\ -i, & f = d, e \end{cases}, \quad (13)$$

where $P_R = (1 + \gamma_5)/2$ is the usual chiral projector and $v \simeq 246$ GeV (see, for example, the discussion in [21]). The matrices ρ_f^{ij} with $f = e, u, d$ are in general complex and can induce tree-level FCNCs. They are given explicitly by

$$\rho_u = L_u^\dagger \hat{\rho}_u R_u,$$

$$\rho_d = L_d^\dagger \hat{\rho}_d R_d,$$

$$\rho_e = L_e^\dagger \hat{\rho}_e R_e,$$

where the unitary matrices transform between the weak and mass eigenstates, and determine the CKM matrix $V = L_u^\dagger L_d$. The charged scalars couple to the fermions as

$$\begin{aligned} \mathcal{L} &= -\bar{\nu}(U_\nu^\dagger \rho_e) H^+ P_R e \\ &\quad - \bar{u}(V \rho_d P_R - \rho_u^\dagger V P_L) H^+ d + \text{H.c.}, \end{aligned} \quad (14)$$

where $U_\nu = L_\nu^\dagger L_e$ is the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) neutrino mixing matrix. Since neutrino oscillations are unimportant in the processes under consideration, we henceforth replace $U_\nu \nu \rightarrow \nu$ with the understanding that ν refers to the initially emitted flavor eigenstate.

III. EXPLAINING THE ANOMALIES

Our primary motivation is to present a framework that is able to simultaneously explain the excess signals in processes with final state τ leptons: $B \rightarrow D^{(*)} \tau \nu$, $B \rightarrow \tau \nu$, and $W \rightarrow \tau \nu$. In addition, we consider the hint of a deficit in $\tau \rightarrow K^- \nu$ decays. In this section we will show how these can come about at tree level due to exchange (or decay) of the charged Higgs H^\pm , for appropriate choices of the new Yukawa couplings in ρ_e, ρ_d , and ρ_u . The decays of B, B_s , and h into $\tau^+ \tau^-$ provide immediate constraints on the scenario, which we therefore also consider in this section.

A. $B \rightarrow D\tau\nu$, $B \rightarrow D^{(*)}\tau\nu$

New contributions to $B \rightarrow D\tau\nu$ can be mediated by the tree-level exchange of the charged Higgs H^\pm if $\rho_e^{i\tau}$ is nonzero, as can be seen from Eq. (14). The matrix element $\rho_e^{\tau\tau}$ turns out to be the optimal choice for satisfying the combined constraints from LHC searches for the neutral boson H^0 and rare leptonic decays of B and B_s mesons. We will therefore assume that $\rho_e^{\tau\tau} \neq 0$, while the remaining entries in ρ_e^{ij} are very small or vanishing.

Integrating out the H^\pm then produces the effective Hamiltonian

$$\begin{aligned}\mathcal{H} &= \frac{\rho_e^{\tau\tau*}}{m_\pm^2} [\bar{\tau} P_L \nu_\tau] [\bar{c} (V\rho_d P_R - \rho_u^\dagger V P_L)^{cb} b] \\ &\equiv \frac{1}{\Lambda^2} (\bar{\tau} P_L \nu_\tau) [C_{S_R}^{cb} (\bar{c} P_R b) + C_{S_L}^{cb} (\bar{c} P_L b)]\end{aligned}\quad (15)$$

that is relevant for $b \rightarrow c\tau\nu$ at the quark level. Reference [16] performed a fit to the $B \rightarrow D^{(*)}\tau\nu$ rates and decay spectral using the two operators in (15), which interfere with the standard model contributions. Two viable solutions for the Wilson coefficients were found there, of which the smaller ones correspond to

$$\frac{C_{S_R}^{cb}}{\Lambda^2} \cong \frac{(\rho_e^{\tau\tau})^* (V\rho_d)^{cb}}{m_\pm^2} \cong \frac{1.25}{\text{TeV}^2}, \quad (16)$$

$$\frac{C_{S_L}^{cb}}{\Lambda^2} \cong -\frac{(\rho_e^{\tau\tau})^* (\rho_u^\dagger V)^{cb}}{m_\pm^2} \cong -\frac{1.02}{\text{TeV}^2}. \quad (17)$$

There is an intriguing relationship between the couplings,

$$(V\rho_d)^{cb} \cong (\rho_u^\dagger V)^{cb}, \quad (18)$$

about which we will say more below.

B. $B \rightarrow \tau\nu$

The contribution of the new charged Higgs to $B^+ \rightarrow \tau^+\nu$ decay modifies the branching ratio (BR) as [13,22,23]

$$\begin{aligned}\mathcal{B}(B^+ \rightarrow \tau^+\nu) &= \frac{G_F^2 |V_{ub}|^2}{8\pi} m_\tau \tau_B f_B m_B \left(1 - \frac{m_\tau^2}{m_B^2}\right)^2 \\ &\times \left|1 + \frac{m_B^2}{\bar{m}_b m_\tau} \frac{(C_{S_R}^{ub} - C_{S_L}^{ub})}{C_{SM}^{ub}}\right|^2,\end{aligned}\quad (19)$$

where \bar{m}_b is the b -quark $\overline{\text{MS}}$ mass, $C_{S_R}^{ub}$ and $C_{S_L}^{ub}$ are defined analogously to $C_{S_R}^{cb}$ and $C_{S_L}^{cb}$ in Eq. (15), and $C_{SM}^{ub} = 4G_F V_{ub}/\sqrt{2}$.

To estimate the possible allowable size of the new physics (NP) contribution, we take the enhancement factor in the second line of (19) to be less than 2.6, the ratio between the 3σ maximum allowed value of the world

average measurement $(1.14 \pm 0.27) \times 10^{-4}$, and the CKMfitter prediction [24] 0.76×10^{-4} of the BR. This gives the bounds

$$-\frac{0.02}{\text{TeV}^2} \lesssim \frac{\rho_e^{\tau\tau}}{m_\pm^2} ((V\rho_d)^{ub} + (\rho_u^\dagger V)^{ub}) \lesssim \frac{0.05}{\text{TeV}^2}. \quad (20)$$

Curiously, this suggests a relation similar to (18), but with the opposite sign,

$$(V\rho_d)^{ub} \cong -(\rho_u^\dagger V)^{ub}. \quad (21)$$

In Sec. V we will present an ansatz that combines these two conditions in a concise way.

If the indication for a factor of 1.5 excess in the observed versus SM predicted partial width is interpreted as evidence for new physics, then the condition (21) is not a strict equality, and we should replace (20) with the condition

$$\left| \frac{(V\rho_d)^{ub} + (\rho_u^\dagger V)^{ub}}{(V\rho_d)^{cb} + (\rho_u^\dagger V)^{cb}} \right| \cong 8 \times 10^{-3}, \quad (22)$$

where we used (16) and (17) to eliminate $\rho_e^{\tau\tau}/m_\pm^2$.

C. $W \rightarrow \tau\nu$

The branching ratios for $W \rightarrow \ell\nu$ were measured by the LEP experiments for the individual lepton flavors, from the production of W^+W^- pairs. The averaged results are [20]

$$\begin{aligned}B(W \rightarrow e\nu) &= (10.71 \pm 0.16)\%, \\ B(W \rightarrow \mu\nu) &= (10.63 \pm 0.15)\%, \\ B(W \rightarrow \tau\nu) &= (11.38 \pm 0.21)\%.\end{aligned}\quad (23)$$

The ratio of decays to $\tau\nu$ versus the first two generations is

$$R_{\tau/\ell} = \frac{B(W \rightarrow \tau\nu)}{\frac{1}{2}[B(W \rightarrow e\nu) + B(W \rightarrow \mu\nu)]} = 1.066 \pm 0.028, \quad (24)$$

which deviates from 1 by 2.35σ . It was suggested by Refs. [25,26] that the excess could be due to contamination of the W decay signals by charged Higgs bosons with mass close to m_W decaying to $\tau\nu$. In a detailed reanalysis of data reported by DELPHI, it was found that the discrepancy could be reduced to 1.03 for a charged Higgs mass of $m_\pm = 81$ GeV and $B(H \rightarrow \tau\nu) = 0.7$, $B(H \rightarrow qq) = 0.3$, values ruled out by the more recent LEP study [27]. Reference [25] found that for $B(H \rightarrow \tau\nu) = 1$, which is the appropriate limit for our model, the observed $R_{\tau/\ell}$ could be explained if $m_\pm = 94_{-3}^{+4}$ GeV in the 2σ region. This is marginally compatible with the combined LEP limit of $m_\pm < 94$ GeV [27].

More recent LHC measurements of $W \rightarrow \tau\nu$ [28] do not see evidence for any excess, but this does not contradict

having an observable effect at LEP since the production mechanism for H^\pm in this case depends upon its coupling to quarks, which is very small in our model. The same remark applies for the Tevatron, where no such effect was observed either.

We do not attempt to reanalyze the LEP signal here, but point out one potentially important difference between our model and those considered previously in this context. We will require a sizable coupling $\rho_e^{\tau\tau} \cong 2.5$ leading to a large width for the charged Higgs,

$$\Gamma_{H^\pm} = \frac{|\rho_e^{\tau\tau}|^2}{16\pi} m_\pm \cong 12 \text{ GeV} \quad (25)$$

for $m_\pm = 100 \text{ GeV}$. This could allow for a greater effect on $R_{\tau/\ell}$ with $m_\pm > 95 \text{ GeV}$ than would be possible in the usually assumed case where H^\pm width effects are ignored.

D. $\tau \rightarrow K^- \nu$

Decays of τ into strange particles are among the processes used to determine the CKM matrix element V_{us} . The HFAG collaboration recently noted that the inclusive decays of this kind lead to a determination that is 3.4σ below the value consistent with unitarity, while the average from all τ decays is 2.9σ too low [10]. This discrepancy could be explained if the contributions from charged Higgs exchange interfere destructively with the SM amplitudes.

Focusing on the specific decay $\tau \rightarrow K^- \nu$, the new contribution to the amplitude is given by

$$\mathcal{M}_{\tau \rightarrow K \nu} = \frac{\rho_e^{\tau\tau} (V \rho_d + \rho_u^\dagger V)_{us} f_K m_K^2}{2m_\pm^2 (m_u + m_s)} (\bar{u}_\tau P_L u_\nu). \quad (26)$$

Taking the central values of $|V_{us}| = 0.2211 \pm 0.0020$ determined from $\tau \rightarrow K \nu$ and $|V_{us}| = 0.2255 \pm 0.0010$ from CKM unitarity [10], we estimate that

$$(V \rho_d + \rho_u^\dagger V)_{us} \cong -4.2 \times 10^{-4} \quad (27)$$

(assuming fiducial values $\rho_e^{\tau\tau} = 2.5$ and $m_\pm = 100 \text{ GeV}$ that will be preferred below). The minus sign is necessary to get destructive interference between W^\pm and H^\pm exchange, given that the relative signs of $\rho_e^{\tau\tau}$ and $V \rho_d$ are fixed by requiring constructive interference in $B \rightarrow D^* \tau \nu$ decays.

We will find that there is some mild tension between (27) and other observables (notably $b \rightarrow s \gamma$) in the numerical fit to be described in Sec. VIII, so that we do not insist on this potential anomaly in our fits. However, it can plausibly fit into the general pattern of deviations in τ interactions that are addressed by our model.

E. Constraint from $h \rightarrow \tau^+ \tau^-$

Because of mixing in the scalar sector, the SM-like Higgs acquires small additional couplings $c_{\beta\alpha} \rho_f$ to fermions. They contribute to the partial width of h as

$$\Gamma(h \rightarrow f_i \bar{f}_j + f_j \bar{f}_i) = N_c \frac{m_h}{16\pi} (|s_{\beta\alpha} y_{ij} + c_{\beta\alpha} \rho_{ij}|^2 + \{i \leftrightarrow j\}) \quad (28)$$

for all kinematically accessible final states (with N_c colors), where $y_{ij} = \sqrt{2} m_i \delta_{ij} / v$. In our model, $\rho_e^{\tau\tau}$ is the largest new coupling.

New contributions to the decays into $\tau^+ \tau^-$ are constrained by ATLAS and CMS observations [29]. Deviations from the SM expectation are characterized by a coupling modifier $\kappa_\tau = |1 + c_{\beta\alpha} \rho_e^{\tau\tau} / y_\tau| \in [0.64, 1.14]$ at 2σ , where y_τ is the SM Yukawa coupling. We get the least restrictive constraint if $c_{\beta\alpha} \rho_e^{\tau\tau}$ is negative, in which case there are two solutions,

$$\begin{aligned} |c_{\beta\alpha} \rho_e^{\tau\tau}| &< 3.7 \times 10^{-3}, \\ -0.022 &< c_{\beta\alpha} \rho_e^{\tau\tau} < -0.017, \end{aligned} \quad (29)$$

at 95% confidence level. Later we will adopt a fiducial value of $\rho_e^{\tau\tau} = 2.5$. In that case the central value of (29) implies $c_{\beta\alpha} = -7.8 \times 10^{-3}$. This region corresponds to the amplitude for $h \rightarrow \tau^+ \tau^-$ having the opposite sign relative to the SM value.

F. Constraints from $B, B_s, \Upsilon \rightarrow \tau^+ \tau^-$

The considerations leading to (16) fix only a linear combination of the ρ_d couplings, namely $V_{cd} \rho_d^{db} + V_{cs} \rho_d^{sb} + V_{cb} \rho_d^{bb}$. It is useful at this point to notice that ρ_d^{db} is strongly constrained by the upper limit of 4.1×10^{-3} [20] on the BR for $B^0 \rightarrow \tau^+ \tau^-$. In our model, this decay is mediated by H^0 exchange, with rate

$$\Gamma(B \rightarrow \tau^+ \tau^-) \cong \frac{|\rho_e^{\tau\tau} \rho_d^{db}|^2 f_B^2 m_B^2 (m_B^2 - 2m_\tau^2)^{3/2}}{64\pi \bar{m}_b^2 m_H^4}, \quad (30)$$

where $f_B = 0.19 \text{ GeV}$ is the B decay constant [30]. If we tried to satisfy the constraint in (16) with only ρ_d^{db} non-vanishing, anticipating that $m_\pm \sim m_H \sim 100 \text{ GeV}$ (see Sec. IV), Eq. (30) would then imply that $\Gamma(B \rightarrow \tau^+ \tau^-)$ is as large as the total measured width of B^0 . We find the upper limit

$$|\rho_d^{db}| < 1.0 \times 10^{-3} \left(\frac{m_H}{100 \text{ GeV}} \right)^2 \left| \frac{2.5}{\rho_e^{\tau\tau}} \right|. \quad (31)$$

Similarly, trying to use ρ_d^{sb} to saturate (16) results in a branching ratio of 0.1 for $B_s \rightarrow \tau^+ \tau^-$. However, the current limits on this decay channel are very weak,

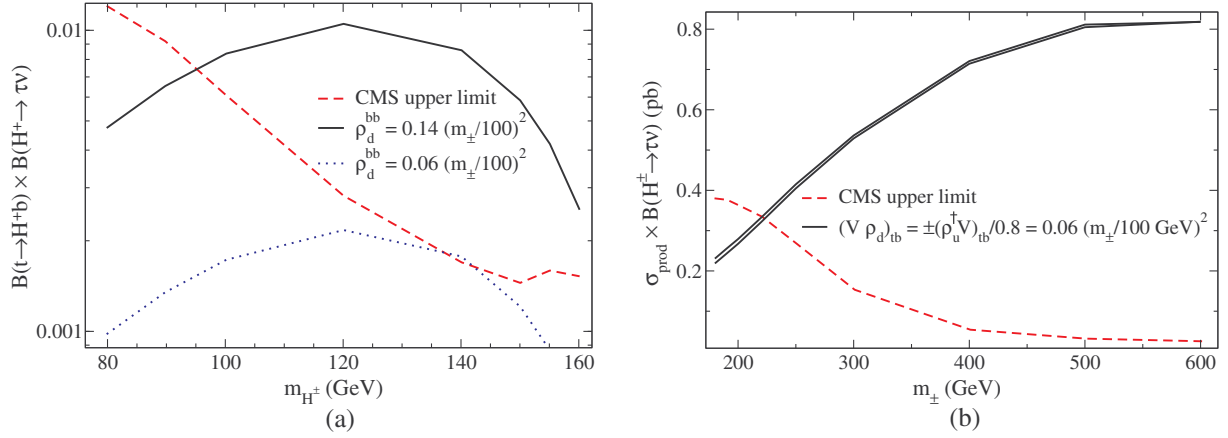


FIG. 1. Left (a): CMS constraints [34] on a charged Higgs decaying to $\tau\nu$ in the $m_{\pm} < m_t$ and $m_{\pm} > m_t$ regions. Left (a): upper limit on $B(t \rightarrow bH^+) \cdot B(H^+ \rightarrow \tau^+\nu)$ versus m_{\pm} in the low mass region. Right (b): upper limit on production cross section times $B(H^+ \rightarrow \tau^+\nu)$ for $m_{\pm} > 180$ GeV. Assumed values of the ρ_d^{bb} coupling from Eq. (33) are indicated.

$B(B_s \rightarrow \tau^+\tau^-) < 5\%$ [31,32]. Using $f_{B_s} = 0.225$ GeV, this gives a bound on $|\rho_d^{sb}|$ of

$$|\rho_d^{sb}| < 2.8 \times 10^{-3} \left(\frac{m_H}{100 \text{ GeV}} \right)^2 \left| \frac{2.5}{\rho_e^{\tau\tau}} \right| \left(\frac{B(B_s \rightarrow \tau^+\tau^-)}{5 \times 10^{-2}} \right)^{1/2}. \quad (32)$$

It follows that we must rely upon ρ_d^{bb} to provide at least part of the contribution to the Wilson coefficient C_{S_R} , if we insist on its central value from (16). This gives the constraint

$$|\rho_d^{bb}| \left(\frac{100 \text{ GeV}}{m_{\pm}} \right)^2 \left| \frac{\rho_e^{\tau\tau}}{2.5} \right| \cong [0.05 - 0.12], \quad (33)$$

where the lowest value in the interval corresponds to saturating the limit (32).

With $\rho_d^{bb} \neq 0$, neutral H^0 exchange also leads to the decays of the $b\bar{b}$ bound state $\chi_{b0} \rightarrow \tau^+\tau^-$ at tree level. In the SM such decays are dominantly electromagnetic, which greatly suppresses the BR of the H^0 -mediated process. No bounds on leptonic decay modes of χ_{b0} are given by the Particle Data Group. For Υ the branching ratio for $\tau^+\tau^-$ final states is $(2.60 \pm 0.10)\%$ but since it is a vector, H^0 cannot mediate the decay at tree level. Rather it proceeds at one loop with virtual H^0 and photon exchange. We estimate that it contributes less than 10^{-14} to the branching ratio.

IV. COLLIDER CONSTRAINTS

We next consider LEP and LHC searches for charged and neutral Higgs bosons with decays principally into τ leptons, as predicted in our model. The charged Higgs can also have an indirect signature through its effect on the $h \rightarrow \gamma\gamma$ partial width.

A. Charged Higgs searches

ATLAS [33] and CMS [34] have recently reported on searches for charged Higgs particles decaying into $\tau\nu$, which is the principal decay channel of H^{\pm} in our model. These searches constrain our scenario in the mass ranges $m_{\pm} \in [80, 160]$ GeV and $[180, 1000]$ GeV, complementing previous LEP studies that excluded $m_{\pm} < 95$ GeV [27] for H^+ decaying with a branching ratio of 100% into $\tau^+\tau^-$ as is the case in our model.

At low masses $m_{\pm} < 160$ GeV, the product of branching ratios $B(t \rightarrow H^+b) \cdot B(H^+ \rightarrow \tau^+\nu)$ is bounded, since the dominant production process is through top quark decays into $H^{\pm}b$. Our model predicts that

$$B(t \rightarrow H^+b) \cong \frac{(1 + \eta^2)(\rho_d^{bb})^2}{32\pi m_t^3 \Gamma_t} (m_t^2 - m_{\pm}^2)^2, \\ B(H^+ \rightarrow \tau\nu) = (1 + 3|\rho_d^{bb}/\rho_e^{\tau\tau}|^2)^{-1} \cong 1, \quad (34)$$

ignoring m_b , where $\Gamma_t = 1.4$ GeV is the measured width of the top quark.¹ The prediction is plotted along with the CMS limit in Fig. 1(a), using the upper and lower values of ρ_d^{bb} consistent with $R(D)$ from Eq. (33). For the smaller value of ρ_d^{bb} , there is almost no restriction on the allowed mass m_{\pm} . The D0 collaboration finds a much weaker limit on $B(t \rightarrow H^+b) \cdot B(H^+ \rightarrow \tau^+\nu) \lesssim 0.2$ in this mass range [35]. CDF obtains the stronger limit of 0.06 [36], which, however, is still not competitive, and we do not show the Tevatron limits on the plot.

At higher masses $m_{\pm} > 180$ GeV, the H^{\pm} is produced by its coupling to tb , either through $gg \rightarrow H^+ \bar{t}b$ or $g\bar{b} \rightarrow H^+ \bar{t}$. In Fig. 1(b) we compare the CMS bound to the model predictions taking the lower value of ρ_d^{bb}

¹The NP contribution to Γ_t must be less than 1% in the experimentally allowed region.

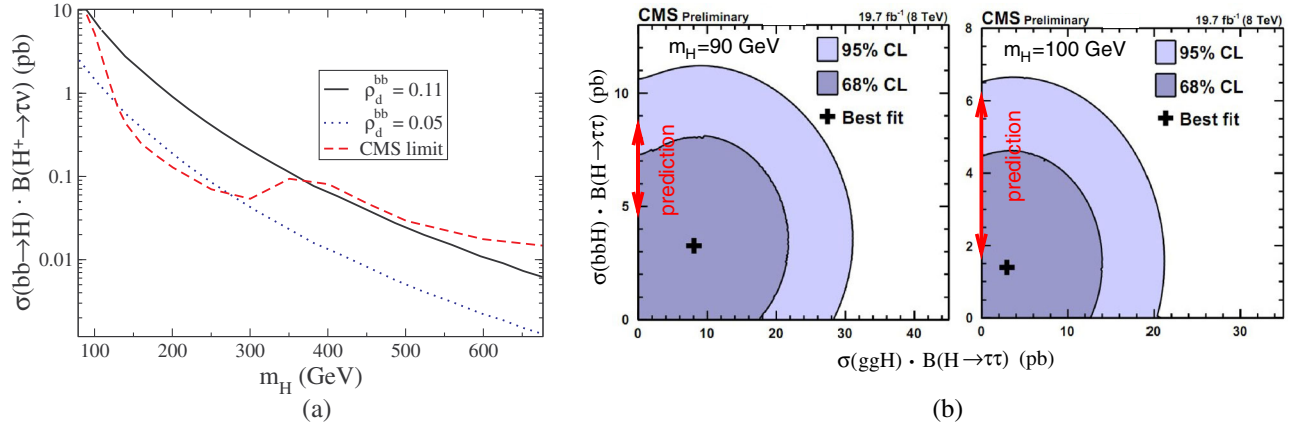


FIG. 2. Left (a): predicted bb fusion cross section versus neutral Higgs mass for $\rho_d^{bb} = 0.05$ and 0.11 , and CMS preliminary upper limit [39]. Right (b): allowed regions in the plane of bb fusion versus gg fusion cross sections, for $m_H = 90, 100$ GeV. Theoretical predictions for $0.07(0.05) < \rho_d^{bb} < 0.10$ are indicated for $m_H = 90(100)$ GeV, having negligible $\sigma(ggH)$. The figure was adapted from Ref. [39].

indicated in Eq. (33).² A charged Higgs mass up to 220 GeV could be consistent with this search.

B. Neutral Higgs searches

ATLAS and CMS searches for the neutral H^0 decaying to $\tau^+\tau^-$ [37–39] put much stronger constraints on our model, forcing us to consider low values of both m_H and m_\pm . Since the ρ_d^{bb} coupling scales as m_\pm^2 to fit $R(D)$ [Eq. (33)], H^0 typically has a larger coupling to b quarks than does the SM Higgs boson. As a result, neutral H^0 production by gluon-gluon fusion [40–42], which is the dominant process for the SM Higgs, can be small compared to bb fusion, leading to strong constraints on the $\sigma(bbH)$ cross section. These limits are weakest at low m_H , and also at low m_\pm due to the scaling of $\rho_d^{bb} \propto m_\pm^2$.

In Fig. 2(a) we plot the predictions of our model for $\sigma(bbH)$ versus m_H [note that $B(H \rightarrow \tau^+\tau^-) = 1$ to a very good approximation], using the values $\rho_d^{bb} = 0.05$ and 0.11 suggested by Eq. (33). To compute $\sigma(bbH)$, we rescaled the cross sections obtained in Ref. [43] (which are computed for a range of m_H) by the more accurate recent results (computed at a few values of m_H) in Ref. [44]. Only for the lower value of ρ_d^{bb} are there any regions consistent with low m_H . Large values of m_H cannot be reconciled with ρ_d^{bb} as small as assumed (~ 0.1) because of the $\rho_d^{bb} \propto m_\pm^2$ scaling, and the need to keep $|m_\pm - m_H| \lesssim 75$ GeV to respect electroweak precision constraints. In the optimistic case of $\rho_d^{bb} = 0.05$, we find an upper limit of $m_H < 125$ GeV.

The CMS search has marginal evidence for excess events at $m_H \cong 90\text{--}100$ GeV, as shown in Fig. 2(b). There is a slight preference for nonzero values of the two production cross sections $\sigma(ggH)$ and $\sigma(bbH)$. Our model predicts

very small values of the former, ~ 0.1 pb, but significant values of $\sigma(bbH)$. We show the range of predictions corresponding to $\rho_d^{bb} = 0.05$ to 0.10 by the vertical arrows. The lower value corresponds to saturating the limit on $B_s \rightarrow \tau^+\tau^-$ in Eq. (32) in order to make ρ_d^{bb} as small as possible in Eq. (33). The higher value corresponds to a branching ratio for $B_s \rightarrow \tau^+\tau^-$ of $5\%/\sqrt{2} = 3.5\%$. There is a strong correlation between $B(B_s \rightarrow \tau^+\tau^-)$ and the possibility to satisfy the CMS constraint, leading to our prediction that $B(B_s \rightarrow \tau^+\tau^-)$ cannot be much smaller, unless the evidence for C_{S_R} from $R(D)$ [Eq. (16)] becomes weaker.³

LEP also constrained the neutral Higgs boson mass in the case of interest for our model, where $H \rightarrow \tau^+\tau^-$ almost exclusively [45]. The statistic S_{95} is defined as the 95% C.L. upper bound on the production cross section of H^0 pairs, in units of the theoretical cross section for $e^+e^- \rightarrow Z^* \rightarrow HH^*$. In the LEP analysis it was assumed that both neutral Higgs bosons are nearly degenerate, which is the same assumption that we make in our model. The pair production cross section is model independent, since it depends only upon the SU(2) gauge interactions of the extra scalar doublet. S_{95} versus m_H is listed in Table II, showing that m_H must be greater than 95 GeV. For $m_H \geq 100$ GeV, the allowed cross section is more than 5 times greater than predicted, allowing for overlap between the LEP- and CMS-allowed regions.

CDF constrained the gluon fusion cross section $\sigma(ggH)$ times $B(H \rightarrow \tau^+\tau^-)$ to be less than 1.7 pb for $m_H = 115$ GeV [46], while the SM prediction is $\sigma(ggH)_{\text{SM}} = 1.07$ pb [47]. In our model $\sigma(ggH)/\sigma(ggH)_{\text{SM}} = (\rho_u^{tt}/y_t)^2 \cong 2 \times 10^{-3}$, far below the CDF limit. Similar to LHC, searches

²We thank Grace Dupuis for computing this production cross section using MADGRAPH.

³At $m_H = 125$ GeV, with $\rho_e^{\tau\tau} = 2.5$, we can obtain $\rho_d^{bb} = 0.07$ with $B(B_s \rightarrow \tau^+\tau^-) = 2\%$. Lower values of $m_H = 125$ require larger $B(B_s \rightarrow \tau^+\tau^-)$.

TABLE II. LEP limits [45] on the production cross section of neutral Higgs bosons from pair production $e^+e^- \rightarrow HH^* \rightarrow \tau^+\tau^-\tau^+\tau^-$, as a function of their mass m_H in GeV. S_{95} is the 95% C.L. upper limit on the ratio of the observed cross section to the predicted one.

| m_H | 90.0 | 92.5 | 95.0 | 97.5 | 100.0 |
|----------|------|------|------|------|-------|
| S_{95} | 0.39 | 0.70 | 1.07 | 2.88 | 5.29 |

for $H \rightarrow \tau^+\tau^-$ in association with b quarks are more sensitive. D0 constrained $\sigma(bbH) \cdot B(H \rightarrow \tau^+\tau^-) < 0.8$ pb for $m_H = 115$ GeV [48]. The SM cross section is $\sigma(bbH) \sim 6$ fb near this mass [49], which gets scaled by $(\rho_d^{bb}/y_b)^2 \cong 5$ in our model, again much smaller than the limit.

In summary, $m_{\pm} \cong 100$ GeV and $m_H \cong [100-125]$ GeV are the favored mass ranges for satisfying the combined limits from LEP and LHC, subject to the constraints from flavor physics discussed in Sec. III. A large value $\rho_e^{\tau\tau} = 2.5$ is also needed, which we will show is allowed by lepton flavor universality of $Z \rightarrow \ell\ell$ decays, in Sec. VII A. Larger values of m_{\pm} require larger values of ρ_d^{bb} to explain $R(D)$, making it more difficult to respect searches for the neutral Higgs.

C. Charged Higgs contribution to $h \rightarrow \gamma\gamma$

The charged Higgs contributes to $h \rightarrow \gamma\gamma$ at one loop, with an amplitude that is proportional to

$$A \sim 3Q_t^2 A_{1/2}(\tau_t) + A_1(\tau_W) + g_{\pm} A_0(\tau_{\pm}),$$

$$A_0(\tau) = -\tau^{-2}(\tau - \arcsin(\sqrt{\tau}))^2 \quad (35)$$

(see, for example, Ref. [50]) where the first two terms are from the top quark and W boson loop, giving -6.5 , while $\tau_{\pm} = (m_h/2m_{\pm}^2)$ and $g_{\pm} = \lambda_1 v^2/(2m_{\pm}^2)$ in the last term. The effective coupling strength is therefore $\kappa_{\gamma} = |A/6.5| \in [0.72, 1.14]$ using constraints from ATLAS and CMS [29]. We then find that the Higgs potential coupling λ_1 is bounded by

$$-0.7 < \lambda_1 < 1.4 \quad (36)$$

for $m_{\pm} = 100$ GeV.

V. CHARGE TRANSFORMATION FLAVOR ANSATZ

Two Higgs doublet models have had a long history of proposed mechanisms to control FCNCs, starting with that of Glashow and Weinberg [51], where up and down quarks are restricted to couple to different Higgs doublets. More recent ideas include the Cheng-Sher texture [52], MFV [1–3], and alignment [14,53]. The ansatz we suggest is distinct from these and takes the form

$$\rho_u^{\dagger} V = \eta U V \rho_d, \quad (37)$$

where $\eta \lesssim 1$ and U is a diagonal unitary matrix, whose first element is $U_{11} = -1$, while the second is $U_{22} = +1$. The third element U_{33} is not yet determined by experimental constraints. We note that if $U_{33} \cong -1$, then U would be special unitary. The structure (37) must, of course, be supplemented by a choice of entries for ρ_d from which ρ_u can be computed, or vice versa.

Let us comment on the general utility of this ansatz for the B decays of interest. The signs chosen for U_{11} and U_{22} are such that the relations (18) and (21) are satisfied. The effect of the sign difference between U_{11} and U_{22} can be understood by using (37) to eliminate $\rho_u^{\dagger} V$ from the charged current interactions of the quarks with H^+ , which then take the form

$$\mathcal{L} = -H^+ \bar{u}_i (P_R - \eta U_{ii} P_L) (V \rho_d)_{ij} d_j + \text{H.c.} \quad (38)$$

If $\eta U_{11} = -1$, the projection operators combine as $P_R + P_L = 1$ for the coupling to up quarks, forming a pure scalar $\bar{u}b$ that cannot interpolate between the pseudoscalar B^+ meson and the vacuum, hence giving no contribution to $B \rightarrow \tau\nu$ decay. For $\eta U_{22} = +1$, the combination $P_R - P_L = \gamma_5$ is pure pseudoscalar, in agreement with the sign difference in the fit result $C_{S_R}^{cb} \cong -C_{S_L}^{cb}$ [16] for $B \rightarrow D\tau\nu$.

The value of η is independently determined by either of the two anomalous measurements. Using (16), (17), and (22), respectively, we find that

$$\eta \cong \begin{cases} 0.78, & R(D^{(*)}) \\ 0.83, & B \rightarrow \tau\nu \end{cases}. \quad (39)$$

It is encouraging that these two estimates are consistent within the experimental errors. We adopt the compromise $\eta = 0.8$ in the following. We note that in the limit $\eta = 1$, the excess in $B \rightarrow D^{(*)}\tau\nu$ is completely in the vector D^* channel and absent from the D final states, because of the parity of the $\bar{\tau}\gamma_5 b$ pseudoscalar coupling. By letting $\eta \neq 1$, this charged current coupling acquires a scalar component interpolating to pseudoscalar D final states as well. The current data are consistent with most of the anomaly being in the D^* channel since the error bars are smaller there (see Table I).

The relation (37) at first sight looks peculiar, since it relates two flavor symmetry breaking effects, associated with quarks of opposite charges. For convenience we give it the name of “charge transformation” (CT) mechanism. In Sec. X we will show that such a structure can reasonably arise from a more fundamental theory of flavor. For now we will take it as a working hypothesis and check whether it is sufficient to help control FCNCs, in conjunction with some specific choices of ρ_d couplings.

In Sec. VIII we will allow for all elements of ρ_d to be nonzero, consistent with a wide variety of experimental constraints. Here we make the approximation of real-valued ρ_d (as well as CKM matrix) so that there are only nine parameters in ρ_d . The fact that there exists a solution that can satisfy many more than nine constraints (not all of which are upper bounds because of the anomalies) is striking. Moreover, we will show that there is no need for fine-tuning of the parameters.

VI. FCNC CONSTRAINTS: MESON MIXING

Although the anomalies in question can be accounted for with only the ρ_d^{bb} element dominating in ρ_d (and $\rho_d^{sb} \sim 0.02\rho_d^{bb}$), naturalness demands that we consider non-vanishing values of the other entries. Neutral meson mixing ($K^0 - \bar{K}^0$, $D^0 - \bar{D}^0$, $B^0 - \bar{B}^0$, $B_s - \bar{B}_s$) provides strong constraints on their sizes. In this section we determine the tree-level and one-loop predictions of the model in the presence of general couplings.

A. Neutral meson mixing: Generalities

The new Higgs bosons induce contributions to neutral meson oscillations. At the quark level, they can be described by an effective Hamiltonian in which the bosons have been integrated out. In general it can contain a number of operators with different Lorentz and color structure. Even though tree-level exchanges only produce two of these operators, at one loop two additional ones are also generated.

The most general effective Hamiltonian for neutral meson mixing is

$$H = \sum_{ij} \left(\sum_{k=1,5} c_k^{ij} Q_k^{ij} + \sum_{k=1,3} \tilde{c}_k^{ij} \tilde{Q}_k^{ij} \right), \quad (40)$$

where the flavor indices run over $ij = sd, cu, bs, bd$ (also denoted K, D, B_s, B_d , respectively) and the operators are

$$Q_1^{ij} = (\bar{q}_{L,i}^\alpha \gamma^\mu q_{L,j}^\alpha)(\bar{q}_{L,i}^\beta \gamma_\mu q_{L,j}^\beta), \quad (41)$$

$$Q_2^{ij} = (\bar{q}_{R,i}^\alpha q_{L,j}^\alpha)(\bar{q}_{R,i}^\beta q_{L,j}^\beta), \quad (42)$$

$$Q_3^{ij} = (\bar{q}_{R,i}^\alpha q_{L,j}^\beta)(\bar{q}_{R,i}^\beta q_{L,j}^\alpha), \quad (43)$$

$$Q_4^{ij} = (\bar{q}_{R,i}^\alpha q_{L,j}^\alpha)(\bar{q}_{L,i}^\beta q_{R,j}^\beta), \quad (44)$$

$$Q_5^{ij} = (\bar{q}_{R,i}^\alpha q_{L,j}^\beta)(\bar{q}_{L,i}^\beta q_{R,j}^\alpha). \quad (45)$$

Here α, β are color indices, and \tilde{Q}_k are related to Q_k by taking $R \leftrightarrow L$. The coefficients of the latter are experimentally constrained at the same level as the ones without tildes.

Integrating out the neutral scalars, we obtain the coefficients

$$C_2^{ij} = \sum_{\phi=h,H,A} \frac{(y_\phi^\dagger)_{ij}^2}{2m_\phi^2}, \quad (46)$$

$$\tilde{C}_2^{ij} = \sum_{\phi=h,H,A} \frac{(y_\phi)_{ij}^2}{2m_\phi^2}, \quad (47)$$

$$C_4^{ij} = \sum_{\phi=h,H,A} \frac{(y_\phi^\dagger)_{ij}(y_\phi)_{ij}}{2m_\phi^2}. \quad (48)$$

If H and A are nearly degenerate as we envision in our model, then there is strong destructive interference between their contributions to C_2 , \tilde{C}_2 . This can be understood in terms of the original complex fields $H \pm iA$ from the fact that the $\langle (H \pm iA)^2 \rangle$ propagator vanishes when H and A are degenerate. On the other hand, there is no such cancellation for C_4 , so it provides the most stringent constraints, unless one of $(y_\phi)_{ij}$ or $(y_\phi^\dagger)_{ij}$ vanishes. However, naturalness favors roughly symmetric Yukawa matrices, as we will show, so that C_4^{ij} is not suppressed in this way.

B. Tree-level constraints on mixing

New tree-level contributions to neutral meson mixing are mediated by the neutral Higgs bosons. The C_2^{ij} coefficients get contributions of opposite signs from the CP -even and odd boson exchanges. Using the mass relation (9), they can be reorganized into the form

$$\begin{aligned} \tilde{C}_2^{ij} &= \frac{(\rho_d^{ij})^2}{2} \left[\frac{c_{\beta\alpha}^2}{m_h^2} - \frac{c_{\beta\alpha}^2}{m_A^2} + \frac{c_{\beta\alpha}^2 m_h^2 - 2\lambda_3 v^2}{m_A^2 m_H^2} \right], \\ C_2^{ij} &= \frac{(\rho_d^{ji*})^2}{2} \left[\frac{c_{\beta\alpha}^2}{m_h^2} - \frac{c_{\beta\alpha}^2}{m_A^2} + \frac{c_{\beta\alpha}^2 m_h^2 - 2\lambda_3 v^2}{m_A^2 m_H^2} \right]. \end{aligned} \quad (49)$$

The terms proportional to $c_{\beta\alpha}^2$ are negligible if λ_3 is not too small. Later we will find that $\lambda_3 \cong 10^{-3}$ can be consistent with technical naturalness, so we adopt this value in what follows. For $m_H \cong m_A \cong 100$ GeV, the constraints on the coefficients become [54,55]

$$\begin{aligned} \bar{\rho}_d^{sd} &< \begin{Bmatrix} 1.3 \times 10^{-4} \\ 9.3 \times 10^{-6} \end{Bmatrix}, & \bar{\rho}_d^{cu} &< \begin{Bmatrix} 5.1 \times 10^{-4} \\ 1.2 \times 10^{-4} \end{Bmatrix}, \\ \bar{\rho}_d^{bd} &< 1.1 \times 10^{-3}, & \bar{\rho}_d^{bs} &< 9.6 \times 10^{-3}, \end{aligned} \quad (50)$$

where $\bar{\rho}_d^{ij}$ stands for either $|\rho_d^{ji}|$ or $|\rho_d^{ij}|$. For K (sd) and D (cu), we show the separate limits from the real (upper) and imaginary (lower) parts of C_2^{ij} . In our fits we will impose the more stringent ones. These limits scale as $(m_H/100 \text{ GeV})(\lambda_3/10^{-3})^{-1/2}$.

In the limit of small Higgs mixing and nearly degenerate H and A , the C_4^{ij} coefficients take the form

$$C_4^{ij} \cong \frac{\rho_q^{ij} \rho_q^{ji*}}{m_H^2} \quad (51)$$

where $q = d$ for K, B_d, B_s and $q = u$ for D mixing. Unlike C_2 , they are not suppressed by $c_{\beta\alpha}$ or λ_3 . Again for $m_H = 100$ GeV, we have the upper limits

$$\begin{aligned} \sqrt{|\rho_d^{ds} \rho_d^{sd}|} &< \left\{ 6.0 \times 10^{-6} \right\}, \quad \sqrt{|\rho_u^{uc} \rho_u^{cu}|} < \left\{ 2.4 \times 10^{-5} \right\}, \\ \sqrt{|\rho_d^{db} \rho_d^{bd}|} &< 4.6 \times 10^{-5}, \quad \sqrt{|\rho_d^{sb} \rho_d^{bs}|} < 4.0 \times 10^{-4}. \end{aligned} \quad (52)$$

They scale as $m_H/(100 \text{ GeV})$. By comparison of (50) and (52) we see that C_4 gives more stringent constraints than C_2 if the coupling matrices ρ_q^{ij} are symmetric, which will be approximately true in our later determination.

C. One-loop contributions to mixing

At one loop, the charged and neutral Higgs bosons give contributions to neutral meson mixing that are higher order in $\rho_{u,d}$, except for loops involving W^\pm exchange. We start with box diagrams involving the exchange of two scalars, followed by exchange of one H^\pm and a W boson. For mesons containing down-type quarks we find

$$\begin{aligned} C_1^{ij} &= \frac{1}{128\pi^2} \left(\frac{(\rho_d^\dagger \rho_d)_{ij}^2}{m_\pm^2} + \frac{(\rho_d \rho_d^\dagger)_{ij}^2}{m_H^2} \right), \\ \tilde{C}_1^{ij} &= \frac{(\rho_d^\dagger \rho_d)_{ij}^2}{128\pi^2} \left(\frac{1}{m_\pm^2} + \frac{1}{m_H^2} \right), \\ C_2^{ij} &= \tilde{C}_2^{ij} = \frac{1}{4} C_4^{ij} = [(V\rho_d)_{it}^\dagger (V\rho_d)_{ij}]^2 \frac{m_i^2 (\ln \frac{m_\pm^2}{m_i^2} - 2)}{64\pi^2 m_\pm^4}, \\ C_5^{ij} &= \frac{(\rho_d^\dagger \rho_d)_{ij}}{32\pi^2} \left(\frac{(\rho_d \rho_d^\dagger)_{ij}}{m_H^2} + \frac{(\rho_d^\dagger \rho_d)_{ij}}{m_\pm^2} \right), \end{aligned} \quad (53)$$

while for D^0 mesons

$$\begin{aligned} C_1^{uc} &= \frac{1}{128\pi^2} \left(\frac{(V\rho_d \rho_d^\dagger V^\dagger)_{uc}^2}{m_\pm^2} + \frac{(V\rho_d^\dagger \rho_d V^\dagger)_{uc}^2}{m_H^2} \right), \\ \tilde{C}_1^{uc} &= \frac{(V\rho_d \rho_d^\dagger V^\dagger)_{uc}^2}{128\pi^2} \left(\frac{1}{m_\pm^2} + \frac{1}{m_H^2} \right), \\ C_5^{uc} &= -\frac{(V\rho_d \rho_d^\dagger V^\dagger)_{uc}^2}{32\pi^2 m_\pm^2} - \frac{(V\rho_d \rho_d^\dagger V^\dagger)_{uc} (V\rho_d^\dagger \rho_d V^\dagger)_{uc}}{32\pi^2 m_H^2}. \end{aligned} \quad (54)$$

We omit the $C_{2,4}^{uc}$ coefficients that are suppressed by m_b^2 .

The box diagrams containing $H^\pm W^\mp$ exchange give rise to

$$\begin{aligned} C_1^{ij} &= \frac{g_2^2 V_{ij} V_{ii}^* (V\rho_d)_{ij} (V\rho_d)_{ii}^*}{128\pi^2 m_\pm^4} m_i^2 \left(\ln \frac{m_\pm^2}{m_i^2} - 2 \right), \\ C_4^{ij} &= \frac{g_2^2 (\rho_d^\dagger)_{ij} (\rho_d)_{ij}}{64\pi^2 m_\pm^2} \end{aligned} \quad (55)$$

for down-quark type mesons, while for D^0 mesons

$$C_4^{uc} = \frac{g_2^2 (V\rho_d^\dagger V^\dagger)_{uc} (V\rho_d V^\dagger)_{uc}}{64\pi^2 m_\pm^2}, \quad (56)$$

and we neglect the m_b^2 -suppressed contribution to C_1^{uc} .

These loop contributions turn out to be much smaller than the tree-level ones previously considered; in the numerical fit of Sec. VIII they are at most a factor of $\sim 10^{-3}$ below the upper limits. But since they depend upon different combinations of the ρ couplings, which are hierarchical, it was not *a priori* obvious that they should be negligible.

VII. RARE DECAYS AND $(g-2)_\tau$

Our model predicts a variety of rare decays beyond those already considered in Sec. III, and a new contribution to the anomalous magnetic moment of the τ . Although they are potentially constraining, most of them turn out to be less so than tree-level meson mixing. The loop enhancement of $Z \rightarrow \tau^+ \tau^-$ (and $Z \rightarrow \nu_\tau \bar{\nu}_\tau$) is the most important of these since it sets the limit on how large the $\rho_e^{\tau\tau}$ coupling can be, which is central to the explanation of the τ anomalies. The prediction for $\tau \rightarrow \pi^- \nu$ is close to the experimental limit for this decay, while the NP amplitude of $b \rightarrow s \gamma$ is the next most significant, a factor of 4 below the experimental limit.

A. $Z \rightarrow \tau^+ \tau^-$

The coupling $\rho_e^{\tau\tau}$ introduces lepton universality violation in $\Gamma(Z \rightarrow \bar{\ell} \ell)$, when comparing $\ell = \tau$ to $\ell = e, \mu$. Such deviations are constrained by LEP, which has reported [20]

$$R_{\tau/e} = \frac{\Gamma(Z \rightarrow \tau^+ \tau^-)}{\Gamma(Z \rightarrow e^+ e^-)} = 1.0019 \pm 0.0032. \quad (57)$$

The new contributions from exchange of the heavy charged and neutral scalars are shown in Fig. 3. These one-loop diagrams give the effective interaction term for the right-handed component of τ coupling to Z ,

$$\begin{aligned}
\mathcal{L}_{\text{eff},\tau_R} &= -(g_{\tau_R} + \delta g_{\tau_R})(\bar{\tau}_R \gamma_\mu \tau_R) Z^\mu, \\
g_{\tau_R} &= g_Z s_W^2, \quad g_Z = \frac{e}{c_W s_W}, \\
\delta g_{\tau_R} &= -\frac{|\rho_e^{\tau\tau}|^2 g_Z}{32\pi^2} F_{\tau_R}, \\
F_{\tau_R} &= \frac{1}{2}(F_a^0 - (1 - 2s_W^2)F_b^0 + s_W^2 F_c^0) \\
&\quad + \frac{1}{2}(-(1 - 2s_W^2)F_a^\pm + F_b^\pm + s_W^2 F_c^\pm), \quad (58)
\end{aligned}$$

where $c_W = \cos \theta_W$ and $s_W = \sin \theta_W$. In the limit of vanishing lepton masses, the loop integrals involving the neutral Higgs H^0 are given by

$$\begin{aligned}
F_a^0 &= 2 \int_0^1 dx \int_0^{1-x} dy \left(\ln \frac{(x+y)m_H^2 - xym_Z^2}{\mu^2} \right), \\
F_b^0 &= 2 \int_0^1 dx \int_0^{1-x} dy \left(\ln \frac{M_b^2}{\mu^2} + \frac{xym_Z^2}{M_b^2} + 1 \right) \\
&\quad [\text{with } M_b^2 = (1-x-y)m_H^2 - xym_Z^2], \\
F_c^0 &= 1 - 2 \ln \frac{m_H^2}{\mu^2}, \quad (59)
\end{aligned}$$

while those involving H^\pm (denoted by F_i^\pm) are the same but with $m_H^2 \rightarrow m_\pm^2$. We have neglected terms of order $c_{\beta\alpha}^2$ in the Higgs mixing. Dependence on the renormalization scale μ drops out in F_{τ_R} and F_{τ_L} (below).

The analogous expressions to (58) for the couplings of τ_L are given by

$$\begin{aligned}
\mathcal{L}_{\text{eff},\tau_L} &= -(g_{\tau_L} + \delta g_{\tau_L})(\bar{\tau}_L \gamma_\mu \tau_L) Z^\mu, \\
g_{\tau_L} &= -\frac{1}{2}g_Z(1 - 2s_W^2), \\
\delta g_{\tau_L} &= -\frac{|\rho_e^{\tau\tau}|^2 g_Z}{32\pi^2} F_{\tau_L}, \\
F_{\tau_L} &= -\frac{1}{2}F_a^0 + s_W^2 F_b^0 - \frac{1}{4}(1 - 2s_W^2)F_c^0. \quad (60)
\end{aligned}$$

Writing $R_{\tau/e} = 1 + \Delta R_{\tau/e}$, the predicted value of the deviation is

$$\begin{aligned}
\Delta R_{\tau/e} &= 2 \frac{g_{\tau_R} \delta g_{\tau_R} + g_{\tau_L} \delta g_{\tau_L}}{g_{\tau_R}^2 + g_{\tau_L}^2} \\
&= \frac{|\rho_e^{\tau\tau}|^2}{8\pi^2} \left(\frac{-2s_W^2 F_{\tau_R} + (1 - 2s_W^2)F_{\tau_L}}{4s_W^4 + (1 - 2s_W^2)^2} \right). \quad (61)
\end{aligned}$$

We take the 1σ experimental upper limit which gives

$$-0.0013 < \Delta R < 0.0051 \quad (62)$$

to obtain the upper bounds on $|\rho_e^{\tau\tau}|$ shown in Fig. 4. (A similar calculation for more general boson masses was carried out in Ref. [56].)

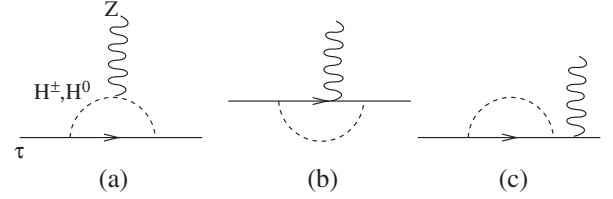


FIG. 3. Diagrams contributing to $Z \rightarrow \tau^+ \tau^-$.

B. $Z \rightarrow \nu_\tau \bar{\nu}_\tau$

Similar diagrams to those in Fig. 3 contribute to the amplitude for $Z \rightarrow \nu_\tau \bar{\nu}_\tau$ decay. We find that the perturbation to the tree-level coupling in analogy to (58) and (60) is given by

$$\begin{aligned}
\mathcal{L}_{\text{eff},\nu_\tau} &= -(g_{\nu_\tau} + \delta g_{\nu_\tau})(\bar{\nu}_\tau \gamma_\mu \nu_\tau) Z^\mu, \\
g_{\nu_\tau} &= \frac{1}{2}g_Z, \\
\delta g_{\nu_\tau} &= -\frac{|\rho_e^{\tau\tau}|^2 g_Z}{32\pi^2} F_{\nu_\tau}, \\
F_{\nu_\tau} &= \frac{1}{2}(1 - 2s_W^2)F_a^\pm + s_W^2 F_b^\pm + \frac{1}{4}F_c^\pm. \quad (63)
\end{aligned}$$

The branching ratio to invisible decays is changed by

$$\frac{\Delta \Gamma_{\text{inv}}}{\Gamma_{\text{inv}}} = \frac{2g_{\nu_\tau} \delta g_{\nu_\tau}}{3g_{\nu_\tau}^2} = -\frac{|\rho_e^{\tau\tau}|^2 F_{\nu_\tau}}{24\pi^2}. \quad (64)$$

For the fiducial values $\rho_e^{\tau\tau} = 2.5$, $m_\pm = 100$ GeV that we will adopt, this leads to an increase $\Delta \Gamma_{\text{inv}} = 1.1$ MeV in the invisible width of the Z. This is close to but consistent with the combined LEP upper limit $\Delta \Gamma_{\text{inv}} < 2$ MeV [57] at 95% C.L.

C. $W \rightarrow \tau \nu$

Distinct from the tree-level H^\pm decays that might have faked $W \rightarrow \tau \nu$ events at LEP, discussed in Sec. III C, there is an actual perturbation to the amplitude for $W \rightarrow \tau \nu$ from loops analogous to those in Fig. 4. In this case, diagrams of type (b) do not contribute because they require a chirality flip leading to suppression by m_τ , since W couples only to left-handed particles. The remaining diagrams give a fractional correction to the $W\tau\nu$ coupling of

$$\frac{\delta g_{W\tau\nu}}{g_{W\tau\nu}} = -\frac{|\rho_e^{\tau\tau}|^2}{64\pi^2} \left[2F_a^\pm + \frac{1}{2}(F_c^0 + F_c^\pm) \right] \quad (65)$$

with the loop functions in brackets evaluating to ~ -0.1 for the Higgs boson masses of interest. For $\rho_e^{\tau\tau} = 2.5$, this leads to a fractional increase in the branching ratio of 0.2%, which is the same as the experimental error [20]. This contribution, while not enough by itself, goes in the right

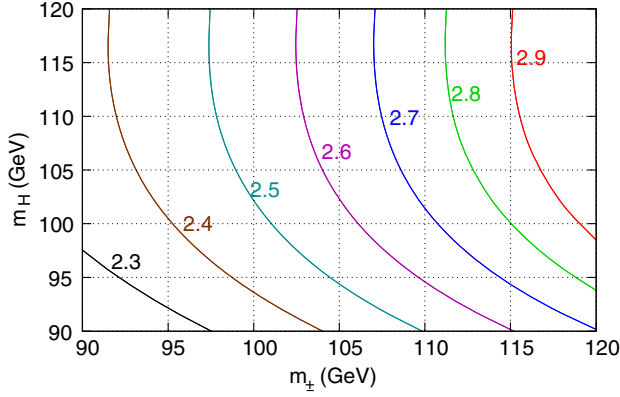


FIG. 4. Contours of upper bound on $|\rho_e^{\tau\tau}|$ from lepton flavor universality of $Z \rightarrow \ell^+ \ell^-$ decays, in the plane of neutral versus charged Higgs boson masses.

direction and could work in combination with the H^\pm decays to explain the observed excess.

D. τ anomalous magnetic moment

The anomalous magnetic moment of the τ is at present only weakly constrained, $-0.052 < \Delta a_\tau < 0.013$ [20]. At one loop, the leading contribution in our model comes from neutral H exchange (see, for example, Ref. [58]),

$$\Delta a_\tau = \frac{|\rho_e^{\tau\tau}|^2 m_\tau^2}{8\pi^2 m_H^2} F\left(\frac{m_\tau}{m_H}\right) \cong 2 \times 10^{-5} \left(\frac{100 \text{ GeV}}{m_H}\right)^2, \quad (66)$$

where the loop function evaluates to $F \cong 7$. The analogous contribution from charged Higgs exchange has a much smaller loop function $\cong -0.2$.

Frequently the dominant contribution to such processes in 2HDMs is the two-loop Barr-Zee (or Bjorken-Weinberg) [59,60] diagram with a top quark or other particle in one of the loops. We find that indeed the contribution from the top quark loop exceeds the one-loop contribution, giving

$$\Delta a_\tau = \frac{\alpha \rho_e^{\tau\tau} \rho_u^{tt}}{8\pi^3} \mathcal{F}^{(1)}[(m_t/m_H)^2] \lesssim 7 \times 10^{-4}, \quad (67)$$

where $\mathcal{F}^{(1)}[(m_t/m_H)^2] = -1.14$ for $m_H = 100$ GeV, using the notation of Ref. [58], and we took $\rho_u^{tt} \lesssim 0.08$ from Eqs. (33) and (39). Although this is much smaller than the current experimental bound, it is 2 orders of magnitude larger than the SM prediction [61]. We find that the other Barr-Zee diagrams are smaller, contributing 10^{-6} from the τ loop analogous to (67) and 10^{-5} from the diagrams with t, b, ν, H^\pm, W^\pm in the loops ($\Delta a_\tau^{(4)}$ in the notation of [58]).

E. Hadronic decays $\tau \rightarrow \pi^- \nu$, $K^- \nu$ and $D_s \rightarrow \tau \nu$

Charged Higgs exchange contributes to hadronic decays of the τ , the simplest of which are $\tau \rightarrow \pi^- \nu$ with branching ratio $B = (10.83 \pm 0.06)\%$ and $\tau \rightarrow K^- \nu$ with

$B = (7.00 \pm 0.10) \times 10^{-3}$ [20]. The amplitude for $\tau \rightarrow K^- \nu$ was already given in Eq. (26). Using the CT flavor ansatz (37) we have $(V\rho_d + \rho_u^\dagger V)_{us} = (1 - \eta)(V\rho_d)_{us}$. For $\tau \rightarrow \pi^- \nu$, one replaces $(V\rho_d)_{us} \rightarrow (V\rho_d)_{ud}$, $m_s \rightarrow m_d$, $m_K \rightarrow m_\pi$, and $f_K \rightarrow f_\pi$.

In both decays, the NP contribution interferes with that of the SM. If we assume that the hint for new physics in $\tau \rightarrow K \nu$ discussed in Sec. III D is just due to a statistical fluctuation, then by demanding that the extra contribution to the branching ratio does not exceed the experimental error, we find the constraints

$$|(V\rho_d)_{us}| < 7 \times 10^{-4}, \quad |(V\rho_d)_{ud}| < 1 \times 10^{-3}, \quad (68)$$

assuming that $\rho_e^{\tau\tau} = 2.5$ and $m_\pm = 100$ GeV. We note that this constrains couplings ρ_d^{qd} and ρ_d^{qs} different from those (ρ_d^{sb} and ρ_d^{bb}) required to explain the B decay anomalies. In the fit to be described below (Sec. VIII), we obtain $(V\rho_d)_{us} = 2 \times 10^{-4}$, $(V\rho_d)_{ud} = 9.5 \times 10^{-4}$. The NP contribution to $\tau \rightarrow \pi^- \nu$ is therefore close to the limit.

The matrix element for $D_s \rightarrow \tau \nu$ is

$$\mathcal{M}_{D_s \rightarrow \tau \nu} = (1 + \eta) \frac{\rho_e^{\tau\tau} (V\rho_d)_{cs} f_{D_s} m_{D_s}^2}{2m_\pm^2 (m_c + m_s)} (\bar{u}_\tau P_L u_\nu), \quad (69)$$

where $f_{D_s} = 0.248$ GeV [62]. Using the observed branching ratio $(5.55 \pm 0.24)\%$ [20] we obtain the bound

$$|(V\rho_d)_{cd}| < 1 \times 10^{-3}. \quad (70)$$

The value from our fit, $(V\rho_d)_{cd} = -2 \times 10^{-4}$, is consistent.

F. $b \rightarrow s \gamma$

The off-diagonal couplings in ρ_d and $V\rho_d$ introduce new contributions to $b \rightarrow s \gamma$ at one loop, which is encoded by the effective Hamiltonian [63]

$$H = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* (C_7 \mathcal{O}_7 + C_7' \mathcal{O}_7'), \quad (71)$$

$$\mathcal{O}_7(\mathcal{O}_7') = \frac{em_b}{16\pi^2} \bar{s}(\sigma_{\mu\nu} P_{R(L)}) b F^{\mu\nu}.$$

Because of operator mixing, one should also consider the analogous operators $\mathcal{O}_8, \mathcal{O}_8'$ for the chromomagnetic moments. These have been computed for type I and II 2HDMs [64,65] but not (as far as we can tell) for a general type III model. A full computation for this case might be interesting for future study. However, most of the contributions appearing in our model can be inferred from the earlier calculations by transcribing the right- and left-handed couplings of the charged Higgs from the type I/II models,

$$\mathcal{L} = (4G_F/\sqrt{2})^{1/2} \bar{u}(\xi m_u V P_L - \xi' V m_d P_R) d, \quad (72)$$

where $m_{u,d}$ are the quark mass matrices. The dominant contributions to the one-loop charged Higgs diagrams in our model can be estimated by taking

$$\begin{aligned}\xi^2 &\rightarrow \frac{1}{2m_t^2 G_F} \frac{(\rho_u^\dagger V)_{ts}}{V_{ts}} \frac{(\rho_u^\dagger V)_{tb}}{V_{tb}} \\ &= 45 U_{tt} (V\rho_d)_{ts} (V\rho_d)_{tb},\end{aligned}\quad (73)$$

$$\begin{aligned}\xi_{\xi'}^{\xi'} &\rightarrow \frac{1}{2m_t m_b G_F} \frac{(\rho_u^\dagger V)_{ts}}{V_{ts}} \frac{(V\rho_d)_{tb}}{V_{tb}} \\ &= 2.3 \times 10^3 (V\rho_d)_{ts} (V\rho_d)_{tb}\end{aligned}\quad (74)$$

in the Wilson coefficients [65,66]

$$\begin{aligned}C_7 &= \xi_{\xi'}^{\xi'} \left(\frac{-3y^2 + 2y}{6(y-1)^3} \ln y + \frac{3y - 5y^2}{12(y-1)^2} \right) \\ &\quad + \xi^2 \left(\frac{-3y^3 + 2y^2}{12(y-1)^4} \ln y + \frac{-8y^3 - 5y^2 + 7y}{72(y-1)^3} \right), \\ C_8 &= \xi_{\xi'}^{\xi'} \left(\frac{y}{2(y-1)^3} \ln y + \frac{y^2 - 3y}{4(y-1)^2} \right) \\ &\quad + \xi^2 \left(\frac{y^2}{4(y-1)^4} \ln y + \frac{-y^3 + 5y^2 + 2y}{24(y-1)^3} \right),\end{aligned}\quad (75)$$

where $y = (m_t/m_\pm)^2$. In (73) we have indicated the expressions following from our flavor ansatz (37) that involve the undetermined sign $U_{tt} = \pm 1$. In the type I/II models, C_7' is smaller than C_7 by a factor of m_s/m_b , but we do not expect that in our model since there is no suppression of the right-handed couplings by m_d . Instead, the primed coefficients are given by (75) after interchanging $\rho_u^\dagger V \leftrightarrow V\rho_d$ in (73). With our flavor ansatz (37), this implies that ξ^2 becomes larger by the factor $1/\eta^2 = 1.56$ while $\xi_{\xi'}$ remains the same.

Recent constraints on C_7 and C_7' (by which we always mean the NP contributions) at the scale of m_b have been determined by Ref. [67],

$$\begin{aligned}C_7(m_b) &\in [-0.055, 0.02], \\ C_7'(m_b) &\in [-0.03, 0.065]\end{aligned}\quad (76)$$

at 2σ . The coefficients (75) evaluated at the weak scale must be run down to m_b [64],

$$\begin{aligned}C_7(m_b) &= \eta^{16/23} C_7 + \frac{8}{3} (\eta^{14/23} - \eta^{16/23}) C_8 \\ &\cong 0.6 C_7 + 0.1 C_8,\end{aligned}\quad (77)$$

at leading order in QCD corrections, where $\eta = \alpha_s(m_W)/\alpha_s(m_b) \cong 0.5$. The primed coefficients run in the analogous way. The numerical fit of Sec. VIII yields $C_7(m_b) \cong C_7'(m_b) \cong 4.9 \times 10^{-3}$, 4 times below the limit for C_7 .

There are also Barr-Zee two-loop contributions that we find to be much smaller. For example, the diagram with a top quark loop and neutral H^0 exchange generates [19,68]

$$C_7 = \frac{\sqrt{2} N_c Q_b Q_t^2 \alpha \rho_d^{sb} \rho_u^{tt}}{8\pi m_b m_t G_F V_{tb} V_{ts}} f(m_t^2/m_H^2) \lesssim 10^{-4}, \quad (78)$$

where the loop function $f(m_t^2/m_H^2) \cong -1$.

G. $s \rightarrow d\gamma$

For the radiative decays of lighter quarks, it is not necessarily a good approximation to assume that the top quark contribution in the loop dominates, because the relevant coupling $(V\rho_d)_{td}$ is CKM suppressed, and for $c \rightarrow u\gamma$ the dominant graph is from an internal b quark. For these decays we content ourselves with an estimate based upon the analogous treatment of leptonic processes $\tau \rightarrow \mu\gamma$ studied in 2HDMs [68], which obtains the separate contributions from neutral as well as charged Higgs exchange. Defining the operator coefficients in the effective Hamiltonian as

$$\mathcal{H}_{\text{eff}} = \frac{Q_s e m_s}{2} \bar{s} \sigma_{\mu\nu} (A_R^{sd} P_R + A_L^{sd} P_L) b F^{\mu\nu}, \quad (79)$$

we find

$$\begin{aligned}A_R^{sd} &= (A_L^{sd})^* = \sum_{q=u,c,t} \frac{Q_q m_q}{Q_s m_s} \frac{(V\rho_d)_{qd}^* (V\rho_d)_{qs}}{16\pi^2 m_\pm^2} f\left(\frac{m_\pm}{m_q}\right), \\ A_R^{sd}[A_L^{sd}] &= \sum_{q=d,s,b} \frac{m_q \rho_d^{dq} \rho_d^{qs} [(\rho_d^{qd} \rho_d^{qs})^*]}{m_s 16\pi^2 m_H^2} f\left(\frac{m_H}{m_q}\right), \\ A_R^{sd}[A_L^{sd}] &= \frac{\rho_d^{ds} y_s c_{\beta\alpha} [\rho_d^{sd*} y_s c_{\beta\alpha}]}{16\pi^2 m_h^2} f\left(\frac{m_h}{m_s}\right),\end{aligned}\quad (80)$$

where the loop function is $f(x) \cong \ln x^2 - 3/2$. Our numerical fit values of the couplings implies $|A_{L,R}^{sd}| \cong -2 \times 10^{-5} \text{ TeV}^{-2}$.

The dipole operator gives rise to a hadronic matrix element

$$\langle \pi^0 | \bar{s} \sigma^{\mu\nu} d | K^0 \rangle = (p_\pi^\mu p_K^\nu - p_K^\mu p_\pi^\nu) \frac{\sqrt{2} f_T^{K\pi}}{m_K + m_\pi} \quad (81)$$

with $f_T^{K\pi} = 0.4$ [69]. It vanishes for on-shell photons in the decay $K \rightarrow \pi\gamma$, but gives a nonvanishing contribution to leptonic modes mediated by the off-shell photon. Because $A_L^{sd} = A_R^{sd*}$, it does not contribute to the CP -violating decay $K_L \rightarrow \pi \ell^+ \ell^-$, but it does contribute to $K_S \rightarrow \pi \ell^+ \ell^-$ whose measured branching ratio is $(3 \pm 1.5) \times 10^{-9}$ [20].

Adapting results of Ref. [70] for K_L decay, we find

$$\frac{\Delta B(K_S \rightarrow \pi^0 e^+ e^-)}{B(K^+ \rightarrow \pi^0 e^+ \nu_e)} = \frac{(2\zeta e^2 Q_s m_s \tilde{B}_T \text{Re}(A_R^{sd})^2 \tau(K_S))}{V_{us} G_F m_K} \frac{\tau(K_S)}{\tau(K^+)}, \quad (82)$$

where $\tilde{B}_T = 1.2$ and ζ accounts for the renormalization of A_R^{sd} between the scale m_H and $\mu = 2$ GeV where the lattice matrix elements are computed. Assuming that the chromomagnetic moment $g_s \bar{d} \sigma_{\mu\nu} G^{\mu\nu} s$ gets generated with the same coefficient as the electromagnetic one $Q_s e \bar{d} \sigma_{\mu\nu} F^{\mu\nu} s$ at the scale m_H and accounting for the mixing of these operators under renormalization, $\zeta = \eta^2(1 - 3 \times 8(1 - \eta^{-1})) = 2.7$, where $\eta = (\alpha_s(m_H)/\alpha_s(m_b))^{2/23}(\alpha_s(m_b)/\alpha_s(\mu))^{2/25} = 0.9$. Equation (82) then gives the new physics contribution $\Delta B(K_S \rightarrow \pi^0 e^+ e^-) = 3 \times 10^{-13}$, far below the measured value.

H. $c \rightarrow u\gamma$

Proceeding similar to the case of $s \rightarrow d\gamma$, the dipole operators for $c \rightarrow u\gamma$ get contributions to their coefficients given by

$$\begin{aligned} A_R^{cu} &= (A_L^{cu})^* = \sum_{q=d,s,b} \frac{Q_q m_q}{Q_c m_c} \frac{(V\rho_d)_{uq}(V\rho_d)_{cq}^*}{16\pi^2 m_\pm^2} f\left(\frac{m_\pm}{m_q}\right), \\ A_R^{bs}[A_L^{cu}] &= \sum_{q=u,c,t} \frac{m_q}{m_c} \frac{\rho_u^{uq} \rho_u^{qc} [(\rho_u^{cq} \rho_u^{qu})^*]}{16\pi^2 m_H^2} f\left(\frac{m_H}{m_q}\right), \\ A_R^{cu}[A_L^{cu}] &= \frac{\rho_u^{uc} y_c c_{\beta\alpha} [\rho_u^{cu*} y_c c_{\beta\alpha}]}{16\pi^2 m_h^2} f\left(\frac{m_h}{m_c}\right). \end{aligned} \quad (83)$$

The second of these (mediated by H^0 in the loop) is the largest, contributing $A_R^{cu} \cong 2 \times 10^{-4} \text{ TeV}^{-2}$. It is difficult to put precise constraints on this quantity because of highly uncertain long-distance contributions to the observable amplitudes. Here we content ourselves with a comparison to the SM short-distance contribution, estimated to be $A_{\text{SM}}^{cu} = 0.02 G_F V_{us} V_{cs} / (2\sqrt{2}\pi^2 Q_c) \cong 2 \times 10^{-3} \text{ TeV}^{-2}$ [71,72]. On this basis the new contribution appears to be sufficiently small, especially since the observed $\Delta c = 1$ decays are dominated by the long-distance contributions.

VIII. NUMERICAL DETERMINATION OF COUPLINGS

We now demonstrate numerically that it is possible to find values of the parameters consistent with all observables. We continue to assume that $m_H \cong m_A$ for the new neutral Higgs bosons, and we adopt the benchmark choice $m_H = 115$ GeV, while taking $m_\pm = 100$ GeV and $\rho_e^{\tau\tau} = 2.5$, consistent with a Higgs mixing angle $c_{\beta\alpha} \cong -8 \times 10^{-3}$ from Eq. (29). These values also satisfy collider constraints as long as ρ_d^{bb} , ρ_u^{tt} are sufficiently small, as we will verify, and $Z \rightarrow \ell\ell$ universality.

The best-fit values of the quark couplings ρ_d are determined using a χ^2 statistic that incorporates the most constraining observables (in addition to the anomalies we set out to address), namely the tree-level contributions to meson mixing. We minimize χ^2 with respect to the elements of ρ_d , with ρ_u determined by the CT ansatz (37), requiring that the upper limits on the Wilson coefficients C_4^M not be exceeded for any meson M . Minimizing χ^2 leaves some degeneracy in the fit with respect to products of the form $\rho_q^{ij} \rho_q^{ji}$, which generally must be small to satisfy the mixing constraints. We partially resolve this degeneracy by trying to enforce $|\rho_q^{ij}| \sim |\rho_q^{ji}|$ as much as possible, to avoid having matrix elements that are unnaturally small, as we will discuss in Sec. IX.

We make the simplifying approximation of real-valued ρ_d and ρ_u .⁴ This requires ignoring the phase of the CKM matrix as well since $\rho_u = \eta V \rho_d^\dagger V^\dagger U$ according to (37), where we take $\eta = 0.8$ and $U = \text{diag}(-1, 1, -1)$ for definiteness. We therefore approximate V as an SO(3) matrix using Eqs. (12.3) and (12.4) of [20] with the replacement $\bar{\rho} + i\bar{\eta} \rightarrow (\bar{\rho}^2 + \bar{\eta}^2)^{1/2}$, leaving for future work to incorporate phases into the analysis.

Using this fitting procedure, an example of couplings that are consistent with all constraints is

$$\rho_d = \begin{pmatrix} 8.3 \times 10^{-4} & 1.1 \times 10^{-7} & 1.3 \times 10^{-4} \\ 3.8 \times 10^{-7} & 7.7 \times 10^{-4} & 2.8 \times 10^{-3} \\ -1.2 \times 10^{-5} & -2.1 \times 10^{-5} & 5.5 \times 10^{-2} \end{pmatrix}, \quad (84)$$

$$\rho_u = \begin{pmatrix} -6.6 \times 10^{-4} & 3.5 \times 10^{-6} & 1.4 \times 10^{-4} \\ -2.1 \times 10^{-5} & 7.8 \times 10^{-4} & 1.8 \times 10^{-3} \\ -7.6 \times 10^{-4} & 4.0 \times 10^{-3} & 4.4 \times 10^{-2} \end{pmatrix}, \quad (85)$$

$$V\rho_d = \begin{pmatrix} 8.1 \times 10^{-4} & 1.7 \times 10^{-4} & 9.5 \times 10^{-4} \\ -1.9 \times 10^{-4} & 7.5 \times 10^{-4} & 5.0 \times 10^{-3} \\ -6.8 \times 10^{-6} & -5.3 \times 10^{-5} & 5.5 \times 10^{-2} \end{pmatrix}. \quad (86)$$

Recall that only ρ_d is independent; ρ_u is determined, and the charged Higgs couplings $V\rho_d$ are shown for convenience. Other solutions can be found with smaller values of the matrix elements not needed for the B decay anomalies (ρ_d^{bb} and ρ_d^{sb}); we have allowed the former to be nearly as large as is consistent with meson mixing constraints.

In Fig. 5 we show the predicted values versus experimental limits on the magnitude of the C_4^{K,D,B,B_s} Wilson coefficients corresponding to the tree-level contributions to

⁴However, we do not assume that phases are small when applying the limits on Wilson coefficients from meson mixing, where the bounds on imaginary parts can be orders of magnitude stronger than on the real parts. We allow for the possibility that the phases are $O(1)$ for the interpretation of these bounds, by imposing the more stringent imaginary part limits.

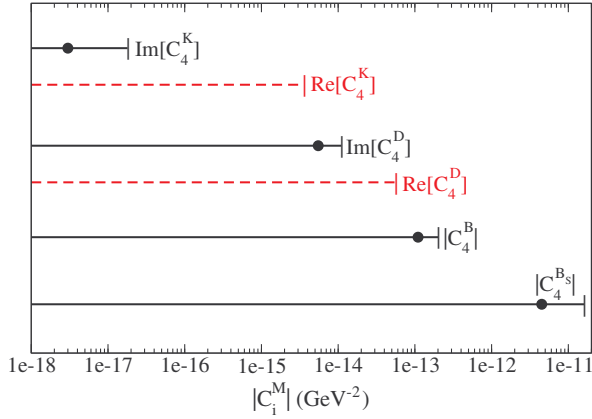


FIG. 5. Summary of meson mixing limits and predictions (dots) for the Wilson coefficients that are most constraining for our model. Allowed ranges for $\text{Im}(C_4^j)$ are shown as solid lines, while those for $\text{Re}(C_4^j)$ are dashed lines, for the cases where there is a distinction. The new neutral Higgs boson mass is assumed to be 115 GeV.

meson mixing from the H exchange. For K^0 and D^0 we satisfy the more stringent constraints on the imaginary part of C_4 , noting that $\text{Im}(C_4^D)$ comes from the phase of $(V\rho_d^\dagger V^\dagger)_{uc}(V\rho_d^\dagger V^\dagger)_{cu} \cong (V_{ub}V_{cs}^*|\rho_d^{bb}|)^2$, which is of the same order as the real part.

For the values of ρ_d^{bb} and ρ_d^{sb} given in (84), the cross section $\sigma(bbH) = 1.2$ pb for production of H by bb fusion, not far below the CMS upper limit of 1.8 pb, while the branching ratio for $B_s \rightarrow \tau^+\tau^-$ is predicted to be 2.9%, close to the current upper limit of 5%.

A. Including $\tau \rightarrow K\nu$ deficit

In the preceding fit we did not try to obtain the negative value of $(V\rho_d)_{us}$ favored by Eq. (27) for explaining the low $\tau \rightarrow K\nu$ determination of V_{us} . Doing so introduces some tension with the limit on $b \rightarrow s\gamma$. We are able to obtain $(V\rho_d)_{us} = -1.8 \times 10^{-3}$, so that $(V\rho_d + \rho_u^\dagger V)_{us} = (1 - \eta)(V\rho_d)_{us} = -3.6 \times 10^{-4}$, close to the target value of (27), while respecting all other constraints except for a marginal violation of the 2σ limit on $b \rightarrow s\gamma$. The fit gives $C_7' = -0.036$, which is still in the 3σ allowed region of Ref. [67].

IX. ONE-LOOP CORRECTIONS TO COUPLINGS

A texture present in the ρ matrices at tree level gets modified by loops involving products of $\rho_{u,d}$ as well as the CKM matrix V . Rather than estimating all possible loop corrections, it is more efficient to use a spurion analysis in which the Yukawa matrices are taken to transform under the full $\text{SU}(3)_u \times \text{SU}(3)_d \times \text{SU}(3)_Q$ flavor symmetries, constructing all combinations that transform in the same way as the couplings of interest. This generates a large

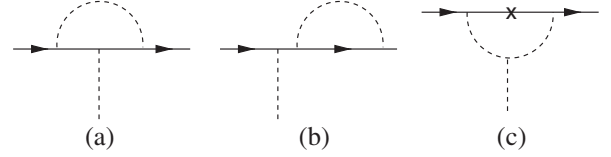


FIG. 6. (a), (b) One-loop corrections to couplings captured by the spurion analysis. (c) Correction requiring a chirality flip, which is not explicitly included in the spurion analysis.

subset of the complete set of flavor structures that should arise from the loop corrections.

The procedure captures the contributions from loops carrying momenta between the fundamental scale down to the scale of electroweak symmetry breaking. In particular, it accounts for one-loop diagrams of the type shown in Figs. 6(a) and 6(b). Diagrams of the type 6(c) require a mass insertion in the fermion line, which needs a more detailed computation. We defer such a study to the future, hoping that the terms included are reasonably representative of the full corrections. It is also possible that they give an overestimate of the true corrections, as the example of $Z \rightarrow \tau^+\tau^-$ in Sec. VII A showed. In that process, the perturbation to the vertex turns out to be considerably smaller than a naive estimate of the loop diagrams suggested.

It is clearest to think initially in the unbroken phase, using the couplings $\hat{y}_{u,d,e}$ and $\hat{\rho}_{u,d,e}$ of the Lagrangian (3) in the original field basis before diagonalizing $\hat{y}_{u,d,e}$. The spurion transformation properties of the Yukawa matrices under the flavor symmetries are

$$\begin{aligned} (\hat{y}_u, \hat{\rho}_u) &\rightarrow V_Q^\dagger (\hat{y}_u, \hat{\rho}_u) V_u, \\ (\hat{y}_d, \hat{\rho}_d) &\rightarrow V_Q^\dagger (\hat{y}_d, \hat{\rho}_d) V_d, \\ (\hat{y}_e, \rho_e) &\rightarrow V_L^\dagger (\hat{y}_e, \rho_e) V_e, \end{aligned} \quad (87)$$

where V_i denotes an element of the $\text{SU}(3)_i$ flavor subgroup.

At one loop, corrections that are cubic in the couplings are generated. Purely on the basis of the symmetries, we see that the following matrix structures would be allowed (now considering only the quark couplings):

$$\begin{aligned} \delta(\hat{y}_u, \hat{\rho}_u) &\sim \left\{ \begin{aligned} &[(\hat{y}_u, \hat{\rho}_u) \cdot (\hat{y}_u^\dagger, \hat{\rho}_u^\dagger)] \cdot (\hat{y}_u, \hat{\rho}_u) \\ &+ [(\hat{y}_d, \hat{\rho}_d) \cdot (\hat{y}_d^\dagger, \hat{\rho}_d^\dagger)] \cdot (\hat{y}_u, \hat{\rho}_u) \end{aligned} \right\}, \\ \delta(\hat{y}_d, \hat{\rho}_d) &\sim \left\{ \begin{aligned} &[(\hat{y}_u, \hat{\rho}_u) \cdot (\hat{y}_u^\dagger, \hat{\rho}_u^\dagger)] \cdot (\hat{y}_d, \hat{\rho}_d) \\ &+ [(\hat{y}_d, \hat{\rho}_d) \cdot (\hat{y}_d^\dagger, \hat{\rho}_d^\dagger)] \cdot (\hat{y}_d, \hat{\rho}_d) \end{aligned} \right\}. \end{aligned} \quad (88)$$

Hence flavor symmetry alone allows 16 possible combinations as corrections to each kind of coupling. In practice, not all of these are realized by the diagrams in Fig. 6, as we will explicitly check. Moreover, for a coupling to a given external Higgs field, half of these are suppressed by the

small mixing angle $c_{\beta\alpha}$, since, for example, the product of the two vertices associated with a loop involving H goes like $\hat{\rho}^2 + c_{\beta\alpha}\hat{\rho}\hat{y} + c_{\beta\alpha}^2\hat{y}^2$, while those connected to h give $\hat{y}^2 + c_{\beta\alpha}\hat{\rho}\hat{y} + c_{\beta\alpha}^2\hat{\rho}^2$. Finally, it is convenient once the appropriate structures are identified to transform to the basis where the fermion mass matrices are diagonalized, and express the results in terms of y_i (the diagonalized version of \hat{y}_i) and ρ_i . This introduces factors of the CKM matrix V wherever there is a mismatch between u - and d -type indices. These come from diagrams with charged Higgs exchange.

A. Corrections to quark couplings

Beyond tree level, we can no longer characterize the nonstandard couplings by just ρ_d and ρ_u because the simple relation between the nonstandard couplings of the light Higgs h and the couplings of the heavy Higgs bosons is not preserved. The nonstandard Yukawa couplings get corrections of the form

$$\delta\mathcal{L}_q = -\sum_{q=u,d} \bar{q}_L \left(\frac{h}{\sqrt{2}} \delta y_{q,h} + H \delta \rho_{q,H} \right) q_R - H^\dagger \bar{u} (\delta(V\rho_d)_\pm P_R - \delta(\rho_u^\dagger V)_\pm P_L) d + \text{H.c.}, \quad (89)$$

where $H = (H^0 + iA^0)/\sqrt{2}$. From examination of the diagrams in Figs. 6(a) and 6(b), we estimate the corrections

$$\begin{aligned} \delta y_{u,h} = & \epsilon_u^1 y_u y_u^\dagger y_u + \epsilon_u^2 \rho_u \rho_u^\dagger y_u + \epsilon_u^3 y_u \rho_u^\dagger \rho_u \\ & + \epsilon_u^4 \rho_u y_u^\dagger \rho_u + \epsilon_u^5 V \rho_d y_d^\dagger V^\dagger \rho_u + \epsilon_u^6 V \rho_d \rho_d^\dagger V^\dagger y_u \\ & + c_{\beta\alpha} \{ \epsilon_u^7 \rho_u \rho_u^\dagger \rho_u + \epsilon_u^8 y_u \rho_u^\dagger y_u + \epsilon_u^9 \rho_u y_u^\dagger y_u \\ & + \epsilon_u^{10} y_u y_u^\dagger \rho_u + \epsilon_u^{11} V \rho_d \rho_d^\dagger V^\dagger \rho_u \}, \end{aligned} \quad (90)$$

$$\begin{aligned} \delta y_{d,h} = & \epsilon_d^1 y_d y_d^\dagger y_d + \epsilon_d^2 \rho_d \rho_d^\dagger y_d + \epsilon_d^3 y_d \rho_d^\dagger \rho_d \\ & + \epsilon_d^4 \rho_d y_d^\dagger \rho_d + \epsilon_d^5 V^\dagger \rho_u y_u^\dagger V \rho_d + \epsilon_d^6 V^\dagger \rho_u \rho_u^\dagger V y_d \\ & + c_{\beta\alpha} \{ \epsilon_d^7 \rho_d \rho_d^\dagger \rho_d + \epsilon_d^8 y_d \rho_d^\dagger y_d + \epsilon_d^9 \rho_d y_d^\dagger y_d \\ & + \epsilon_d^{10} y_d y_d^\dagger \rho_d + \epsilon_d^{11} V^\dagger \rho_u \rho_u^\dagger V \rho_d \}, \end{aligned} \quad (91)$$

$$\begin{aligned} \delta y_{u,H} = & c_{\beta\alpha} \{ \eta_u^1 y_u y_u^\dagger y_u + \eta_u^2 \rho_u \rho_u^\dagger y_u + \eta_u^3 y_u \rho_u^\dagger \rho_u \\ & + \eta_u^4 \rho_u y_u^\dagger \rho_u + \eta_u^5 V \rho_d y_d^\dagger V^\dagger \rho_u + \eta_u^6 V \rho_d \rho_d^\dagger V^\dagger y_u \} \\ & + \eta_u^7 \rho_u \rho_u^\dagger \rho_u + \eta_u^8 y_u \rho_u^\dagger y_u + \eta_u^9 \rho_u y_u^\dagger y_u \\ & + \eta_u^{10} y_u y_u^\dagger \rho_u + \eta_u^{11} V \rho_d \rho_d^\dagger V^\dagger \rho_u, \end{aligned} \quad (92)$$

$$\begin{aligned} \delta y_{d,H} = & c_{\beta\alpha} \{ \eta_d^1 y_d y_d^\dagger y_d + \eta_d^2 \rho_d \rho_d^\dagger y_d + \eta_d^3 y_d \rho_d^\dagger \rho_d \\ & + \eta_d^4 \rho_d y_d^\dagger \rho_d + \eta_d^5 V^\dagger \rho_u y_u^\dagger V \rho_d + \eta_d^6 V^\dagger \rho_u \rho_u^\dagger V y_d \} \\ & + \eta_d^7 \rho_d \rho_d^\dagger \rho_d + \eta_d^8 y_d \rho_d^\dagger y_d + \eta_d^9 \rho_d y_d^\dagger y_d \\ & + \eta_d^{10} y_d y_d^\dagger \rho_d + \eta_d^{11} V^\dagger \rho_u \rho_u^\dagger V \rho_d, \end{aligned} \quad (93)$$

$$\begin{aligned} \delta(V\rho_d)_\pm = & \zeta_d^1 \rho_u \rho_u^\dagger V \rho_d + \zeta_d^2 y_u \rho_u^\dagger V y_d + \zeta_d^3 y_u y_u^\dagger V \rho_d \\ & + \zeta_d^4 V \rho_d \rho_d^\dagger \rho_d + \zeta_d^5 V \rho_d y_d^\dagger y_d \\ & + c_{\beta\alpha} \{ \zeta_d^6 y_u \rho_u^\dagger V \rho_d + \zeta_d^7 \rho_u \rho_u^\dagger V y_d + \zeta_d^8 \rho_u y_u^\dagger V \rho_d \\ & + \zeta_d^9 V \rho_d \rho_d^\dagger y_d + \zeta_d^{10} V \rho_d y_d^\dagger \rho_d \}, \end{aligned} \quad (94)$$

$$\begin{aligned} \delta(\rho_u^\dagger V)_\pm = & \zeta_u^1 \rho_u^\dagger V \rho_d \rho_d^\dagger + \zeta_u^2 y_u^\dagger V \rho_d y_d^\dagger + \zeta_u^3 \rho_u^\dagger V y_d y_d^\dagger \\ & + \zeta_u^4 \rho_u^\dagger \rho_u \rho_u^\dagger V + \zeta_u^5 y_u^\dagger y_u \rho_u^\dagger V \\ & + c_{\beta\alpha} \{ \zeta_u^6 \rho_u^\dagger V \rho_d y_d^\dagger + \zeta_u^7 y_u^\dagger V \rho_d \rho_d^\dagger + \zeta_u^8 \rho_u^\dagger V y_d \rho_d^\dagger \\ & + \zeta_u^9 y_u^\dagger \rho_u \rho_u^\dagger V + \zeta_u^{10} \rho_u^\dagger y_u \rho_u^\dagger V \}. \end{aligned} \quad (95)$$

Here the coefficients ϵ_f^i , η_f^i , ζ_f^i are all assumed to be of order $1/16\pi^2$. The contributions that are suppressed by $c_{\beta\alpha}$ can be understood as having an odd or even number of ρ or y insertions, respectively. For completeness, we include two corrections that exist also within the standard model, namely $\epsilon_{u,d}^1$. We omit them from the following analysis since they do not involve the new physics we are investigating.

To test the degree of tuning required by our numerical fit, we have computed the maximum of each of these estimates using the numerical values of the couplings in (86). The magnitude of correction to each coupling, relative to its tree-level value, and the correction responsible for the largest effect in each matrix, is given by

$$\left| \frac{\delta \rho_{d,H}}{\rho_d} \right| = \begin{pmatrix} 10^{-8} & 10^{-6} & 10^{-4} \\ 10^{-5} & 10^{-7} & 10^{-5} \\ 10^{-5} & 10^{-4} & 10^{-5} \end{pmatrix}, \quad \eta_d^8, \quad (96)$$

$$\left| \frac{\delta \rho_{u,H}}{\rho_u} \right| = \begin{pmatrix} 10^{-7} & 10^{-4} & 10^{-2} \\ 10^{-4} & 10^{-5} & 10^{-2} \\ 10^{-2} & 10^{-2} & 10^{-2} \end{pmatrix}, \quad \eta_u^{8-10}, \quad (97)$$

$$\left| \frac{\delta y_{d,h}}{c_{\beta\alpha} \rho_d} \right| = \begin{pmatrix} 10^{-5} & 0.2 & 0.2 \\ 0.01 & 10^{-4} & 10^{-2} \\ 0.1 & 0.6 & 0.2 \end{pmatrix}, \quad \epsilon_d^5, \quad (98)$$

$$\left| \frac{\delta y_{u,h}}{c_{\beta\alpha} \rho_u} \right| = \begin{pmatrix} 10^{-3} & 0.9 & 2 \\ 0.3 & 0.05 & 0.8 \\ 0.2 & 0.2 & 0.4 \end{pmatrix}, \quad \epsilon_u^6, \quad (99)$$

$$\left| \frac{\delta(V\rho_d)_\pm}{V\rho_d} \right| = \begin{pmatrix} 10^{-9} & 10^{-7} & 10^{-5} \\ 10^{-7} & 10^{-7} & 10^{-5} \\ 10^{-2} & 10^{-2} & 10^{-2} \end{pmatrix}, \quad \zeta_d^3, \quad (100)$$

$$\left| \frac{\delta(\rho_u^\dagger V)_\pm}{\rho_u^\dagger V} \right| = \begin{pmatrix} 10^{-7} & 10^{-5} & 10^{-5} \\ 10^{-6} & 10^{-5} & 10^{-5} \\ 10^{-2} & 10^{-2} & 10^{-2} \end{pmatrix}, \quad \zeta_u^5. \quad (101)$$

The most potentially worrisome elements are the corrections to $y_{d,h}^{bs}$ and $y_{u,h}^{uc}$, which can increase the tree-level contributions to D and B_s mixing mediated by light Higgs exchange. The relatively large corrections to $y_{u,h}$, namely $\delta y_{u,h}^{ut}, \delta y_{u,h}^{ct} \sim 1$, are harmless since they only affect flavor-changing decays of the top quark, which are weakly constrained by observations. The other corrections can perturb the predictions for the C_2 mixing coefficients in Eq. (49) by factors of at most $O(1)$. But these coefficients are less constraining than the C_4 's in our fit. The one that comes closest is $C_2^{B_s}$ which is 0.06 of the experimental limit. Thus there is plenty of room for the tree-level couplings to receive corrections of the order we find without violating any experimental constraints.

B. Lepton couplings

Unlike for the quark couplings, naturalness does not require us to turn on any significant off-diagonal elements

$$\begin{aligned} \delta\lambda_1 &\sim \epsilon \left[(\lambda + \lambda_6)(6\lambda_1 + 2\lambda_2) + 2\lambda_1^2 + \lambda_2^2 + 2\lambda_4^2 + \lambda_1 \left((\rho_e^{\tau\tau})^2 + 3y_t^2 - \frac{9}{2}g^2 \right) + \frac{9}{8}g^4 - 6y_t^2(\rho_u^{\tau\tau})^2 \right] \sim 3 \times 10^{-2}, \\ \delta\lambda_2 &\sim \epsilon \left[2\lambda_2(\lambda + \lambda_6) + 4\lambda_1\lambda_2 + 2\lambda_2^2 + 5\lambda_4^2 + \lambda_2 \left((\rho_e^{\tau\tau})^2 + 3y_t^2 - \frac{9}{2}g^2 \right) - 6y_t^2(\rho_u^{\tau\tau})^2 \right] \sim 9 \times 10^{-3}, \\ \delta\lambda_3 &\sim \epsilon \left[\frac{5}{2}\lambda_4^2 - 3y_t^2(\rho_u^{\tau\tau})^2 \right] \sim 1 \times 10^{-3}, \\ \delta\lambda_4 &\sim \epsilon \left[\lambda_4 \left(12\lambda_6 + 3\lambda_1 + 4\lambda_2 - \frac{9}{2}g^2 + \frac{3}{2}y_t^2 + \frac{1}{2}(\rho_e^{\tau\tau})^2 \right) - 6y_t(\rho_u^{\tau\tau})^3 \right] \sim 2 \times 10^{-2}, \\ \delta\lambda_5 &\sim \epsilon [\lambda_4(3\lambda_1 + 2\lambda_2) - 6y_t^3\rho_u^{\tau\tau}] \sim 4 \times 10^{-4}, \\ \delta\lambda_6 &\sim \epsilon \left[12\lambda_6^2 + \lambda_1^2 + \lambda_1\lambda_2 + \frac{1}{2}\lambda_2^2 + 6\lambda_4^2 + \frac{9}{16}g^4 - \frac{9}{2}\lambda_6g^2 - 6(\rho_u^{\tau\tau})^4 - 2(\rho_e^{\tau\tau})^4 \right] \sim -0.5, \end{aligned} \quad (103)$$

where g is the $SU(2)_L$ gauge coupling and we have ignored terms involving g' [the $SU(1)$ hypercharge] and the small λ_3 and λ_5 couplings. We have included the effect of $\rho_e^{\tau\tau}$ where it is not suppressed by powers of the SM tau Yukawa coupling. To obtain the numerical estimates, we chose fiducial values of the other couplings that are consistent with the assumed mass spectrum $m_\pm = 100$ GeV, $m_H = 115$ GeV,

$$\lambda_1 = 0.3, \quad \lambda_2 = 0.1, \quad \lambda_4 = 0.3, \quad \lambda_6 = 0.7. \quad (104)$$

The potentially worrisome corrections are those for the smallest couplings, $\lambda_5 \cong -6 \times 10^{-4}$ using $c_{\beta\alpha} = -8 \times 10^{-3}$ [see Eq. (29)] and Eq. (7), and $\lambda_3 \cong 10^{-3}$. Comparison with the estimates in (103) indicates that these values are relatively stable. Our choice of couplings in (104) allows

in ρ_e^{ij} . In the absence of neutrino masses, these are not generated by loops. Charged Higgs exchange generates an off-diagonal coupling of order

$$\delta y_{e,h}^{\mu\tau} \sim \frac{\lambda_1 \rho_e^{\tau\tau} m_\mu m_\nu}{16\pi^2 m_\pm^2}, \quad (102)$$

which is negligible. This conclusion would also remain true if we allowed for nonvanishing ρ_e^{ee} and $\rho_e^{\mu\mu}$ entries (with smaller values than $\rho_e^{\tau\tau}$). We do not pursue a more complete exploration of the allowed leptonic couplings here.

C. Higgs potential couplings

We can estimate the size of corrections to the Higgs potential couplings λ_i more definitely than those for the quark couplings since the beta functions are known; see, for example, Ref. [73]. Our scenario requires that $\lambda_3 \ll 1$ and $\lambda_5 \ll 1$, whereas the other λ_i could be larger. Taking $\epsilon = 1/16\pi^2$ (which ignores possible logarithmic enhancements), the dominant contributions to the one-loop corrections are of order

for some accidental cancellation in $\delta\lambda_5$ between the bosonic and fermionic loops. Even without such a cancellation, the contribution from the top quark by itself is $\sim 3\lambda_5$ which requires only a mild coincidence between tree and loop contributions to obtain the desired value. Although the correction to λ_6 is relatively large, the phenomenology of the model is largely insensitive to its value.

D. Landau pole

The large coupling $\rho_e^{\tau\tau} \cong 2.5$ may be expected to give rise to a Landau pole at a relatively low scale, indicating that further new physics will be required to achieve a UV complete description. To estimate this scale we consider the renormalization group equations that depend most sensitively on $\rho_e^{\tau\tau}$,

$$\begin{aligned}\frac{d\lambda_6}{d\ln\mu^2} &\cong \frac{1}{16\pi^2}(12\lambda_6 - 2(\rho_e^{\tau\tau})^4), \\ \frac{d\rho_e^{\tau\tau}}{d\ln\mu^2} &\cong \frac{1}{16\pi^2}(\rho_e^{\tau\tau})^3.\end{aligned}\quad (105)$$

Numerically solving using the initial conditions $\lambda_6 = 0.7$ [see Eq. (104)] and $\rho_e^{\tau\tau} \cong 2.5$ at the scale $\mu = 100$ GeV, we find that the couplings diverge at $\mu \cong 55$ TeV.

X. MICROSCOPIC ORIGIN OF CT ANSATZ

As an example of what kind of physics could give rise to the CT ansatz (37), we construct a model where the $SU(3)_Q \times SU(3)_u \times SU(3)_d$ flavor symmetry is spontaneously broken by bifundamentals $\Phi_u, \Phi_d, \Phi, M_u, M_d$, coupling to heavy $SU(2)$ -singlet quarks $U_{R,L}$ and $D_{R,L}$. The charges of the fields under the flavor symmetries are shown in Fig. 7. As in (3), H_1 is the SM-like Higgs field and H_2 is the new doublet, before mixing of the neutral mass eigenstates, and we take the Lagrangian at the high scale to be

$$\begin{aligned}\mathcal{L} = & \frac{1}{\Lambda}(H_1 \bar{Q}_L \Phi_u u_R + \tilde{H}_1 \bar{Q}_L \Phi_d d_R) \\ & + H_2 \bar{Q}_L U_R + \tilde{H}_2 \bar{Q}_L D_R \\ & + \bar{U}_L \Phi^\dagger u_R + \bar{D}_L \Phi^\dagger d_R \\ & + \bar{U}_L M_u U_R + \bar{D}_L M_d D_R,\end{aligned}\quad (106)$$

which respects the full flavor symmetry. In Table III we show the charge assignments under a Z_4 symmetry that allows the interactions in (106) while forbidding those with H_1 and H_2 interchanged. This symmetry gets spontaneously broken by Vacuum expectation value (VEVs) of the bifundamental fields $M_{u,d}$, allowing for subsequent generation of the terms in the Higgs potential (5) that break the symmetry (i.e., the terms with coefficients λ_4 and λ_5). The

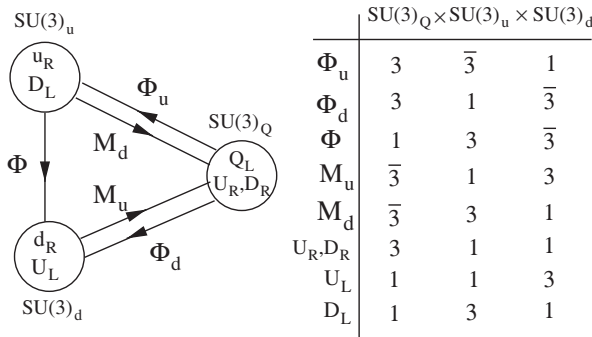


FIG. 7. Moose diagram indicating the transformation properties of the SM quark fields u_R, d_R, Q_L , the heavy singlet quarks U, D , and the bifundamental $\Phi, \Phi_u, \Phi_d, M_u, M_d$ under the $SU(3)_Q \times SU(3)_u \times SU(3)_d$ flavor symmetry.

TABLE III. Z_4 charge assignments needed for the allowed terms in the Lagrangian (106).

| H_1 | H_2 | Φ | $\Phi_{u,d}$ | $M_{u,d}$ | Q_L | (u_R, d_R) | (U_L, D_L) | (U_R, D_R) |
|-------|-------|--------|--------------|-----------|-------|--------------|--------------|--------------|
| 1 | -1 | 1 | 1 | -1 | i | i | i | $-i$ |

$m_{12}^2 H_1^\dagger H_2$ term that breaks it softly can be allowed from the outset, to avoid cosmological problems from domain walls.

In (106) we have not specified a fully renormalizable Lagrangian, but merely assumed that the SM-like Yukawa couplings arise from the VEVs, $\hat{y}_{u,d} = \langle \Phi_{u,d} \rangle / \Lambda$ with some large mass scale Λ . Our main interest is in the origin of the new Yukawa couplings $\hat{\rho}_{u,d}$. Assuming the simple symmetry-breaking pattern $\langle M_d \rangle = \langle M_u \rangle / \eta = M$ times the unit matrix in flavor space, after integrating out the heavy U, D quarks the new Yukawas are given by $\hat{\rho}_u = \eta \langle \Phi^\dagger \rangle / M$, $\hat{\rho}_d = \langle \Phi \rangle / M$. However, this is in the basis where $\hat{y}_{u,d}$ are not yet diagonalized. As usual, we must transform $u_R \rightarrow R_u^\dagger u_R$, $d_R \rightarrow R_d^\dagger d_R$, $u_L \rightarrow L_u^\dagger u_L$, $d_L \rightarrow L_d^\dagger d_L$. In the quark mass basis, $\rho_u = \eta L_u \langle \Phi^\dagger \rangle R_u^\dagger / M$, $\rho_d = L_d \langle \Phi \rangle R_d^\dagger / M$.

With these results, we can now explain the origin of the ansatz (37) by computing the two sides of that relation,

$$\begin{aligned}\eta UV \rho_d &= \eta U L_u \frac{\langle \Phi \rangle}{M} R_d^\dagger, \\ \rho_u^\dagger V &= \eta R_u \frac{\langle \Phi \rangle}{M} L_d^\dagger,\end{aligned}\quad (107)$$

where we have used $V = L_u L_d^\dagger$. Equality of $\eta UV \rho_d$ and $\rho_u^\dagger V$ follows from taking

$$R_u = U L_u, \quad R_d = L_d. \quad (108)$$

One recognizes the condition $R_d = L_d$ as that which would arise if \hat{y}_d is a symmetric matrix. The other relation $R_u = U L_u$ implies that \hat{y}_u is symmetric. This means that \hat{y}_u splits into two pieces, one symmetric and the other antisymmetric, having no nonvanishing elements in common. For example, if $U = \text{diag}(-1, 1, -1)$, then \hat{y}_u has the structure

$$\hat{y}_u = \begin{pmatrix} a & 0 & b \\ 0 & c & 0 \\ b & 0 & d \end{pmatrix} + \begin{pmatrix} 0 & e & 0 \\ -e & 0 & f \\ 0 & -f & 0 \end{pmatrix}. \quad (109)$$

We imagine that it is possible to find a potential $V(\Phi_u)$ whose minimum has this form. Then our ansatz, which at first sight appears contrived, can be a simple consequence of the SM-like Yukawa matrix \hat{y}_d being symmetric in the underlying theory of flavor, while \hat{y}_u has the pattern (109), along with the charge transformation bifundamental Φ whose VEV gives rise to both ρ_u and ρ_d simultaneously.

Our focus has been to explain the initially peculiar-looking relation between quark couplings ρ_u and ρ_d

proposed herein, rather than the leptonic couplings ρ_e . From the flavor perspective the ansatz that ρ_e is dominated by the single element $\rho_e^{\tau\tau}$ is natural since no dangerous FCNCs arise even for large values of $\rho_e^{\tau\tau}$. However, it may be possible to accommodate ρ_e into the UV model in the obvious manner in parallel to the quark couplings, by adding fields with interactions

$$H_2 \bar{L}_L E_R + \bar{E}_L \Phi E_R + \bar{E}_L M_e E_R. \quad (110)$$

We then predict that

$$\rho_e = \frac{\langle \Phi \rangle}{M_e}, \quad (111)$$

assuming that the mass matrix M_e of the heavy vectorlike leptons is proportional to the unit matrix. If $\langle \Phi \rangle$ is strongly dominated by the 33 elements, to explain the hypothesized leptonic couplings, this would lead to quark couplings that are also dominated by the 33 elements, and with other elements generated from these by mixing with the approximate structure of the CKM matrix. Although we do not pursue this quantitatively here, the values given in (84) and (85) appear to be roughly consistent with this expectation.

XI. ALTERNATIVE MODEL

It is possible to design a similar model that is less constrained by collider searches, if a light sterile neutrino ν_s exists. The $R(D)$ anomaly could then be explained by the new process $B \rightarrow D^{(*)} \nu_s$ contributing to the observed decays. Similarly B would get the new decay channel $B \rightarrow \tau \nu_s$, and $H^\pm \rightarrow \tau \nu_s$ could contaminate the $W \rightarrow \tau \nu$ signal at LEP. However, the apparent rate for $\tau \rightarrow K \nu$ could only be increased in this model because of the lack of interference with the SM amplitude. This same absence would also change the fits to the Wilson coefficients for explaining $R(D)$: we estimate that

$$(C_{S_R}^{cb}, C_{S_L}^{cb}) = (2.14, -1.41) \left(\frac{\Lambda}{\text{TeV}} \right)^2 \quad (112)$$

by fitting to the decay rates, which are larger than (16) and (17) to compensate for the lack of interference (see Appendix A). It would require a dedicated analysis to check whether this choice of coefficients significantly degrades the agreement with the decay spectra.

This scenario has the advantage that the stringent collider constraints from searches for $H \rightarrow \tau^+ \tau^-$ are evaded, since now H decays almost exclusively into $\nu_s \nu_\tau$. In Appendix A we estimate that m_H can become as large as 175 GeV, although it is still preferable to keep m_\pm close to 100 GeV to keep the new quark couplings small so that the predicted branching ratio for $B \rightarrow \nu_s \nu_\tau$ remains reasonably small. Even though this decay mode is expected to be less constrained than that for $B \rightarrow \tau^+ \tau^-$, a theoretical

understanding of the total width for B in the SM compared to the experimental value limits how large it can be. We give further details about this alternative model in the Appendix.

XII. CONCLUSIONS

The model we have presented is admittedly unlikely, requiring coincidences of several new particles and decay modes that are just below the threshold of detection. It is much more likely that some of the experimental anomalies that motivated the model will disappear. However, if the $R(D)$ anomaly proves to be real and needs both Wilson coefficients $C_{S_R}^{cb}$, $C_{S_L}^{cb}$ as indicated by several fits to the data, then our scenario seems to be the only kind of two Higgs doublet models that can be compatible with the observations. The simultaneous explanation of the other tentative anomalies in $B \rightarrow \tau \nu$, $W \rightarrow \tau \nu$ (and possibly $\tau \rightarrow K \nu$) is an added bonus that requires little extra model-building input.

The most striking prediction is that new Higgs bosons of mass ~ 100 – 125 GeV that may have been just beyond the kinematic reach of LEP have couplings to b quarks that put the neutral one just below the current sensitivity of CMS searches. We expect that more data should soon reveal the existence of the neutral H in the $\tau^+ \tau^-$ channel at the LHC. A further prediction is that $B_s \rightarrow \tau^+ \tau^-$ will be observed with a surprisingly high branching ratio of several percent. The coupling of the SM-like Higgs boson to τ should be smaller than the SM expectation, possibly having the wrong sign. Higher precision tests of $Z \rightarrow \ell \ell$ universality should start to reveal an excess in $Z \rightarrow \tau^+ \tau^-$, and in the invisible Z width due to extra $Z \rightarrow \nu_\tau \bar{\nu}_\tau$ decays.

The framework also suggests that other observables could be on the edge of revealing new physics: $\tau \rightarrow \pi \nu$, $b \rightarrow s \gamma$, and the neutral meson mixing amplitudes. These are less definite predictions, since we have allowed the new flavor-violating Yukawa couplings $\rho_{u,d}^{ij}$ [apart from those directly involved in explaining $R(D)$] to be nearly as large as possible while remaining consistent with experimental constraints. It is possible that they are smaller, even though there is no fundamental reason that they should be. By studying the expected size of loop contributions to the new couplings, we found that they could indeed be smaller in many cases without requiring any fine-tuning.

Even if all the hints of new physics that motivated this study should disappear, some of the ideas presented here could still be of value. First, the flavor ansatz (37) reduces the arbitrariness of the new couplings, allowing us to parametrize everything in terms of ρ_d alone. In the absence of anomalous $R(D)$, the need for sizable ρ_d^{sb} would disappear and make a symmetric ansatz for ρ_d^{ij} possible, further reducing the number of independent new Yukawa couplings. The primary motivation for our ansatz was to give a more definite flavor structure to the charged Higgs couplings than is generic if ρ_u and ρ_d are independent.

Second, we have shown that it need not be a disaster to allow generic new Yukawa couplings in two Higgs doublet models, even in the absence of any particular mechanism for suppressing FCNCs. It could be that the dangerous couplings are simply small, even though there is no symmetry principle to explain their smallness. Our model presents a counterexample to the usual concern that small values require fine-tuning. We estimated that the relative corrections from loops to the Yukawa couplings of the new Higgs fields are all less than 10^{-2} , Eqs. (96), (97), (100), and (101). The corresponding corrections to the nonstandard couplings of the SM-like Higgs (98) and (99) can in some cases exceed their tree-level values, but this does not lead to any significant FCNCs from h exchange, since they are still small enough to remain well below constraints from meson oscillations, as long as the Higgs potential coupling λ_3 that controls the splitting between m_H and m_A is $\lesssim 10^{-3}$.

A variant model where the charged Higgs couples to $\bar{\nu}_s \tau_L$ (where ν_s is a light sterile neutrino) instead of $\bar{\tau}_R \nu_\tau$ is outlined in Sec. XI and Appendix A. It has greater freedom in the allowed boson masses and couplings to quarks, making it harder to rule out. Whether it can provide as good a fit to $R(D)$ requires further study.

Our original intent was to use a large $\rho_e^{\mu\tau}$ coupling instead of $\rho_e^{\tau\tau}$ to explain the hint of $h \rightarrow \mu^\pm \tau^\mp$ decays of the SM Higgs seen by CMS and ATLAS [74,75]. This turns out to be much more difficult because of the $H \rightarrow \mu^\pm \tau^\mp$ decay (with 100% branching ratio) of the new neutral boson. Even though no formal limits on m_H with this decay channel have been published, we believe it would have been seen in the searches for $h \rightarrow \mu\tau$ of the SM Higgs, ruling out this model. Hence a further prediction of the present model is that $h \rightarrow \mu\tau$ events will prove to be a statistical fluctuation.

ACKNOWLEDGMENTS

I thank M. Trott for inspiring and initially collaborating on this project, N. Arkani-Hamed, B. Battacharya, G. D'Ambrosio, G. Dupuis, A. Ferrari, T. Gershon, J. Griffiths, K. Kainulainen, Z. Ligeti, D. London, A. McCarn, S. Robertson, R. Sato, S. Sekula, W. Shepherd, J. Shigemitsu, and B. Vachon for useful correspondence or discussions, and Mila Huskat for encouragement. I am grateful to the NBIA for its generous hospitality during this work, which is also supported by NSERC (Canada).

APPENDIX: STERILE NEUTRINO MODEL

Here we provide some details relating to the alternative model with the primary leptonic coupling of the new Higgs fields being to the left-handed lepton doublets and sterile neutrinos ν_R ,

$$\mathcal{L}_Y \ni -\bar{L}_L \hat{Y}_\nu \tilde{H}_1 \nu_R - \bar{L}_L \hat{\rho}_\nu \tilde{H}_2 \nu_R + \text{H.c.}$$

There is also a Majorana mass term for the sterile neutrinos,

$$\frac{1}{2}(\bar{\nu}_R M_R \nu_R^c + \bar{\nu}_R^c M_R^* \nu_R). \quad (\text{A1})$$

Without loss of generality, we can work in the basis where M_R is diagonal,

$$M_R = \text{diag}(M_1, M_2, m_s), \quad (\text{A2})$$

giving a sterile neutrino ν_s with Majorana mass m_s , assumed to be negligibly small for having an observable effect on the decays of B mesons. For simplicity we will assume that $\hat{Y}_\nu^{i3} = 0$ so that ν_s gets no Dirac mass from the VEV of H_1 . After electroweak symmetry breaking, when H_1 has obtained the VEV $v/\sqrt{2}$ and when the heavy states are integrated out, a Majorana mass matrix is generated for the light neutrinos through the seesaw mechanism,

$$m_\nu = \frac{1}{2} v^2 \hat{Y}_\nu^T \hat{M}_s^{-1} \hat{Y}_\nu, \quad (\text{A3})$$

where \hat{M}_s is the submatrix of M_s containing the large eigenvalues. Despite having only two heavy sterile neutrinos, the seesaw mechanism works as usual to explain the small masses of the active neutrinos. The neutrino mass matrix m_ν gets diagonalized by the unitary transformation $\nu_L \rightarrow L_\nu \nu_i$ where ν_i denotes the mass eigenstates.

In the mass eigenbasis for the fermions and scalars, the new Yukawa couplings of neutral scalars to neutrinos take the form

$$\begin{aligned} \mathcal{L}_{Y_\nu} &= -\frac{1}{\sqrt{2}} \sum_{\phi=h,H,A} y_{\phi i}^\nu \bar{\nu}_i \phi P_R \nu_s + \text{H.c.}, \\ y_{hi}^\nu &= c_{\beta\alpha} \rho_\nu^i, \\ y_{Hi}^\nu &= -s_{\beta\alpha} \rho_\nu^i, \\ y_{Ai}^\nu &= i\rho_\nu^i, \end{aligned} \quad (\text{A4})$$

where $\rho_\nu = L_\nu^\dagger \hat{\rho}_\nu$ is a vector in the neutrino flavor space. The charged Higgs couples to the neutrinos via

$$\mathcal{L} = -H^+ [\bar{\nu}(U_\nu^\dagger \rho_e) P_R e - \bar{\nu}_s (\rho_\nu^\dagger U_\nu) P_L e], \quad (\text{A5})$$

where $U_\nu = L_\nu^\dagger L_e$ is the PMNS neutrino mixing matrix. We note that if \hat{Y}_e was originally diagonal for some reason, then $\rho_\nu^\dagger U_\nu = \hat{\rho}_\nu$, so that having $\hat{\rho}_\nu^i \cong \bar{\rho}_\nu \delta_{i\tau}$ in the original field basis would explain why $(\rho_\nu^\dagger U_\nu)^\tau$ is the dominant component of $(\rho_\nu^\dagger U_\nu)$ despite the large mixing angles in U_ν . We will make this assumption in the following, and for simplicity $\rho_e = 0$, so that the charged Higgs coupling to leptons reduces to $\bar{\rho}_\nu H^+ \bar{\nu}_s \tau_L + \text{H.c.}$

1. Refitting $R(D)$

To fit the $R(D)$ observations with this model, we must take into account that the new amplitude for $B \rightarrow D^{(*)}\tau\nu_s$ no longer interferes with the SM contribution due to the different neutrino flavor. To make the proper adjustment it is useful to parametrize the interference effect that occurs in the original model, where the ratios depend upon $x_{\pm} = (C_{S_R}^{cb} \pm C_{S_L}^{cb})/C_{SM}^{cb}$ with $C_{SM}^{cb} = 2\sqrt{2}G_F V_{cb}$, as [13]

$$R(D) = R(D)_{SM}(1 + 1.5x_+ + 1.0x_+^2),$$

$$R(D^*) = R(D^*)_{SM}(1 + 0.12x_- + 0.05x_-^2). \quad (A6)$$

The fit (16) and (17) of Ref. [16] corresponds to $(x_+, x_-) = (0.17, 1.66)$. In the sterile neutrino model, (A6) is modified by omitting the terms linear in x_{\pm} . We find that $(x_+, x_-) \rightarrow (0.53, 2.59)$ to compensate for this change, leading to the Wilson coefficients (112).

This rescaling ignores the effect of having no interference on the decay spectra, where the NP contribution to the amplitude multiplies q^2 , the invariant lepton pair mass squared [76]. Therefore the decay distribution will have a larger q^4 contribution and smaller q^2 , hardening the spectrum. We leave to future work to quantify the effect of this on the fits. Here it is mainly important that $C_{S_R}^{cb}$ remains relatively large, which was the motivation for this study.

Then Eq. (15) implies

$$\frac{C_{S_R}^{cb}}{\Lambda^2} \cong \frac{\bar{\rho}_\nu^*(V\rho_d)^{cb}}{m_{\pm}^2} \cong \frac{2.1}{\text{TeV}^2}, \quad (A7)$$

$$\frac{C_{S_L}^{cb}}{\Lambda^2} \cong -\frac{\bar{\rho}_\nu^*(\rho_u^\dagger V)^{cb}}{m_{\pm}^2} \cong -\frac{1.4}{\text{TeV}^2}. \quad (A8)$$

Comparison with (16) and (17) implies that $|\bar{\rho}_\nu \rho_{u,d}|$ must be larger than in our previous determination by a factor of 1.7, while the parameter η becomes smaller, $\eta \cong 0.67$.

2. $Z \rightarrow \tau^+\tau^-$ and $Z \rightarrow \nu_\tau \bar{\nu}_\tau$

In contrast to the case of the $\rho_e^{\tau\tau}$ coupling, there are only two diagrams contributing to $Z \rightarrow \tau^+\tau^-$ with the $\bar{\rho}_\nu$ coupling, shown in Fig. 8. They involve only an exchange of H^\pm , giving the effective interaction term

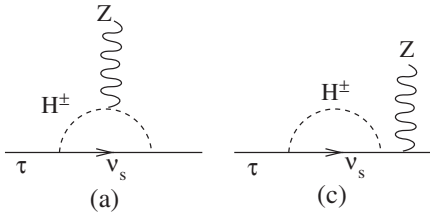


FIG. 8. Diagrams contributing to $Z \rightarrow \tau^+\tau^-$ in sterile neutrino model. Diagrams (a) and (c) correspond to the analogous ones in Fig. 3, while the analog of (b) is not present.

$$\mathcal{L}_{\text{eff},\tau} = -(g_{\tau_L} + \delta g_{\tau_L})(\bar{\tau}_L \gamma^\mu \tau_L) Z^\mu$$

$$\text{with } g_{\tau_L} = -\frac{1}{2}g_Z(1 - 2s_W^2), \quad g_Z = \frac{e}{c_W s_W},$$

$$\delta g_{\tau_L} = -\frac{|\bar{\rho}_\nu|^2 g_Z}{32\pi^2} F_{\tau_L},$$

$$F_{\tau_L} = -\frac{1}{2}(1 - 2s_W^2) \left(F_a^\pm + \frac{1}{2} F_c^\pm \right). \quad (A9)$$

Applying the limit (61) and (62), we find

$$|\bar{\rho}_\nu| < 2.9 \left(\frac{m_{\pm}}{m_Z} \right)^2, \quad (A10)$$

which is less restrictive than the analogous bound on $\rho_e^{\tau\tau}$. For $m_{\pm} = 100$ GeV, we find $|\bar{\rho}_\nu| < 3.5$, and in general the bound is closely numerically fit by $|\bar{\rho}_\nu| < 2.9(m_{\pm}/m_Z)^2$. Combining this with (A7) puts a lower bound on the ρ_d couplings,

$$(V\rho_d)^{cb} > 6.1 \times 10^{-3}. \quad (A11)$$

For the new contribution to $Z \rightarrow \nu_\tau \bar{\nu}_\tau$, the relevant expressions are as in (63) with the replacements $\rho_e^{\tau\tau} \rightarrow \bar{\rho}_\nu$ and

$$F_{\nu_\tau} = \frac{1}{2}F_a^0 + \frac{1}{4}F_c^0. \quad (A12)$$

This evaluates to be 0.5–0.6 times smaller than F_{ν_τ} in the original model, leading to a smaller contribution to the invisible Z width. In the alternative model, there is also a new contribution $Z \rightarrow \nu_s \nu_s$, but it does not interfere with any SM amplitude so it is negligible.

3. Collider constraints

With H^0 decaying nearly 100% to $\nu_s \nu_\tau$, LHC constraints from neutral Higgs searches are essentially removed. We are free to take $m_H \cong 175$ GeV for example. Such a value is just compatible with Electroweak precision data (EWPd) constraints (the ρ parameter) on the $m_H - m_{\pm}$ mass splitting if $m_{\pm} = 100$ GeV. Then (32) is marginally satisfied (taking it to now apply to $B_s \rightarrow \nu\nu$ decays) with a value that is compatible with (A11) even if $\rho_d^{bb} = 0$. In this limiting case we have

$$\bar{\rho}_\nu \rho_d^{sb} = 0.021. \quad (A13)$$

Then ρ_d^{sb} is the dominant coupling in ρ_d and does the job of providing a large enough $C_{S_R}^{cb}$ coefficient for fitting $R(D)$. With negligible ρ_d^{bb} , the production of H^0 is nullified by any other means than the subdominant electroweak $Z^* \rightarrow HA$ or $W^* \rightarrow H^\pm H$ processes. Moreover with ρ_d^{bb} very small, the branching ratio $B(t \rightarrow H^+ b)$ becomes

negligible, since $\rho_d^{bb} \rightarrow V_{ts}\rho_d^{sb}$ in Eq. (34), circumventing the search for charged Higgs bosons.

The combination $\bar{\rho}_\nu c_{\beta\alpha}$ is constrained by the invisible width of the Higgs boson due to $h \rightarrow \nu_s \nu_\tau$, limited to $B(h \rightarrow \nu_s \nu_\tau) < 36\%$ by CMS [77]. This implies $|\bar{\rho}_\nu| < 1.7 \times 10^{-2}/|c_{\beta\alpha}|$. A more stringent bound comes from the degradation of the total Higgs signal strength $\mu \cong 1.1 \pm 0.1$ [29] by invisible decays (not compensated by any increase in production), which implies $\Delta B(h \rightarrow \nu\nu) \cong 1 - \mu = -0.1 \pm 0.1$ [78]. The 2σ upper bound implies

$$|\bar{\rho}_\nu| < \frac{9 \times 10^{-3}}{|c_{\beta\alpha}|}.$$

Since (A13) is compatible with both $B_s \rightarrow \nu\nu$ and $R(D)$, we are free to take a larger mixing angle and a smaller $\bar{\rho}_\nu$, for example, $c_{\beta\alpha} = 10^{-2}$ and $\bar{\rho}_\nu = 0.9$, alleviating the mild naturalness tension for keeping $c_{\beta\alpha}$ very small, that we encountered in the model with $\rho_e^{\tau\tau}$.

We have the freedom to deviate from the limiting case (A13) by turning on ρ_d^{bb} again, such that (A7) is fulfilled by a linear combination of ρ_d^{sb} and ρ_d^{bb} . This reduces the branching ratio $B(B_s \rightarrow \nu\nu)$ and increases that of $t \rightarrow H^\pm b$ so that CMS searches for $H^\pm \rightarrow \tau\nu$ apply. The resulting

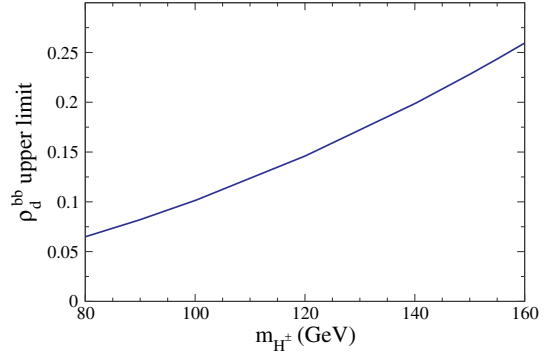


FIG. 9. Upper limit on ρ_d^{bb} inferred from CMS search [34] for $H^\pm \rightarrow \tau\nu$ as a function of charged Higgs mass m_\pm , for $B(H^\pm \rightarrow \tau\nu) = 1$.

constraint on ρ_d^{bb} , plotted in Fig. 9, prevents us from attributing more than 70% of $(V\rho_d)^{cb}$ (controlling the Wilson coefficient $C_{S_R}^{cb}$) to the contribution from $V_{cb}\rho_d^{bb}$. In this other limiting case, the branching ratio for $B_s \rightarrow \nu\nu$ is reduced to the level of 0.5%. Thus while the alternative version of the model is less constrained, it still predicts a significant contribution to the invisible width of B_s .

-
- [1] R. S. Chivukula and H. Georgi, *Phys. Lett. B* **188**, 99 (1987).
 - [2] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia, *Nucl. Phys. B* **645**, 155 (2002).
 - [3] V. Cirigliano, B. Grinstein, G. Isidori, and M. B. Wise, *Nucl. Phys. B* **728**, 121 (2005).
 - [4] A. L. Kagan, G. Perez, T. Volansky, and J. Zupan, *Phys. Rev. D* **80**, 076002 (2009).
 - [5] J. P. Lees *et al.* (BABAR Collaboration), *Phys. Rev. Lett.* **109**, 101802 (2012).
 - [6] M. Huschle *et al.* (Belle Collaboration), *Phys. Rev. D* **92**, 072014 (2015).
 - [7] R. Aaij *et al.* (LHCb Collaboration), *Phys. Rev. Lett.* **115**, 111803 (2015); **115**, 159901 (2015).
 - [8] I. Adachi *et al.* (Belle Collaboration), *Phys. Rev. Lett.* **110**, 131801 (2013).
 - [9] A. Soffer, *Mod. Phys. Lett. A* **29**, 1430007 (2014).
 - [10] A. Lusiani, M. Chruszcz, K. Hayasaka, H. Hayashii, J. M. Roney, B. Shwartz, and S. Banerjee, *Nucl. Part. Phys. Proc.* **260**, 32 (2015).
 - [11] S. Fajfer, J. F. Kamenik, I. Nisandzic, and J. Zupan, *Phys. Rev. Lett.* **109**, 161801 (2012).
 - [12] A. Datta, M. Duraisamy, and D. Ghosh, *Phys. Rev. D* **86**, 034027 (2012); M. Duraisamy and A. Datta, *J. High Energy Phys.* **09** (2013) 059.
 - [13] A. Crivellin, C. Greub, and A. Kokulu, *Phys. Rev. D* **86**, 054014 (2012).
 - [14] A. Celis, M. Jung, X. Q. Li, and A. Pich, *J. High Energy Phys.* **01** (2013) 054.
 - [15] M. Tanaka and R. Watanabe, *Phys. Rev. D* **87**, 034028 (2013).
 - [16] M. Freytsis, Z. Ligeti, and J. T. Ruderman, *Phys. Rev. D* **92**, 054018 (2015).
 - [17] A. Crivellin, J. Heeck, and P. Stoffer, *Phys. Rev. Lett.* **116**, 081801 (2016).
 - [18] J. P. Lees *et al.* (BABAR Collaboration), *Phys. Rev. D* **88**, 072012 (2013).
 - [19] D. Aristizabal Sierra and A. Vicente, *Phys. Rev. D* **90**, 115004 (2014).
 - [20] K. A. Olive *et al.* (Particle Data Group), *Chin. Phys. C* **38**, 090001 (2014).
 - [21] Y. Omura, E. Senaha, and K. Tobe, *J. High Energy Phys.* **05** (2015) 028.
 - [22] W. S. Hou, *Phys. Rev. D* **48**, 2342 (1993).
 - [23] S. Fajfer, J. F. Kamenik, and I. Nisandzic, *Phys. Rev. D* **85**, 094025 (2012).
 - [24] J. Charles *et al.*, *Phys. Rev. D* **91**, 073007 (2015).
 - [25] R. Dermisek, arXiv:0807.2135.
 - [26] J.-h. Park, *J. High Energy Phys.* **10** (2006) 077.
 - [27] G. Abbiendi *et al.* (ALEPH and DELPHI and L3 and OPAL and LEP Collaborations), *Eur. Phys. J. C* **73**, 2463 (2013).
 - [28] Z. Cyczula (ATLAS Collaboration), *Nucl. Phys. B, Proc. Suppl.* **253–255**, 180 (2014).
 - [29] ATLAS and CMS Collaborations, Report No. ATLAS-CONF-2015-044.
 - [30] H. Na, C. J. Monahan, C. T. H. Davies, R. Horgan, G. P. Lepage, and J. Shigemitsu, *Phys. Rev. D* **86**, 034506 (2012).

- [31] Y. Grossman, Z. Ligeti, and E. Nardi, *Phys. Rev. D* **55**, 2768 (1997).
- [32] T. Blake, T. Gershon, and G. Hiller, *Annu. Rev. Nucl. Part. Sci.* **65**, 113 (2015).
- [33] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **03** (2015) 088.
- [34] V. Khachatryan *et al.* (CMS Collaboration), *J. High Energy Phys.* **11** (2015) 018.
- [35] V. M. Abazov *et al.* (D0 Collaboration), *Phys. Lett. B* **682**, 278 (2009).
- [36] T. A. Aaltonen *et al.* (CDF Collaboration), *Phys. Rev. D* **89**, 091101 (2014).
- [37] V. Khachatryan *et al.* (CMS Collaboration), *J. High Energy Phys.* **10** (2014) 160.
- [38] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **11** (2014) 056.
- [39] CMS Collaboration, Report No. CMS-PAS-HIG-14-029.
- [40] S. Dawson, *Nucl. Phys.* **B359**, 283 (1991).
- [41] D. de Florian and M. Grazzini, *Phys. Lett. B* **674**, 291 (2009).
- [42] LHC Higgs Cross Section Working Group, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>.
- [43] M. Trott and M. B. Wise, *J. High Energy Phys.* **11** (2010) 157.
- [44] M. Bonvini, A. S. Papanastasiou, and F. J. Tackmann, *J. High Energy Phys.* **11** (2015) 196.
- [45] S. Schael *et al.* (ALEPH and DELPHI and L3 and OPAL and LEP Working Group for Higgs Boson Searches Collaborations), *Eur. Phys. J. C* **47**, 547 (2006).
- [46] T. Aaltonen *et al.* (CDF Collaboration), *Phys. Rev. Lett.* **108**, 181804 (2012).
- [47] J. Baglio and A. Djouadi, *J. High Energy Phys.* **10** (2010) 064.
- [48] S. Chakrabarti (D0 and CDF Collaborations), [arXiv:1110.2421](https://arxiv.org/abs/1110.2421).
- [49] S. Dittmaier, M. Kramer, and M. Spira, *Phys. Rev. D* **70**, 074010 (2004).
- [50] P. Posch, *Phys. Lett. B* **696**, 447 (2011).
- [51] S. L. Glashow and S. Weinberg, *Phys. Rev. D* **15**, 1958 (1977).
- [52] T. P. Cheng and M. Sher, *Phys. Rev. D* **35**, 3484 (1987).
- [53] A. Pich and P. Tuzon, *Phys. Rev. D* **80**, 091702 (2009).
- [54] M. Bona *et al.* (UTfit Collaboration), *J. High Energy Phys.* **03** (2008) 049.
- [55] O. Gedalia, Y. Grossman, Y. Nir, and G. Perez, *Phys. Rev. D* **80**, 055024 (2009).
- [56] T. Abe, R. Sato, and K. Yagyu, *J. High Energy Phys.* **07** (2015) 064.
- [57] S. Schael *et al.* (ALEPH and DELPHI and L3 and OPAL and SLD and LEP Electroweak Working Group and SLD Electroweak Group and SLD Heavy Flavour Group Collaborations), *Phys. Rep.* **427**, 257 (2006).
- [58] V. Ilisie, *J. High Energy Phys.* **04** (2015) 077.
- [59] J. D. Bjorken and S. Weinberg, *Phys. Rev. Lett.* **38**, 622 (1977).
- [60] S. M. Barr and A. Zee, *Phys. Rev. Lett.* **55**, 2253 (1985).
- [61] F. Burger, G. Hotzel, K. Jansen, and M. Petschlies, [arXiv:1501.05110](https://arxiv.org/abs/1501.05110).
- [62] C. T. H. Davies, C. McNeile, E. Follana, G. P. Lepage, H. Na, and J. Shigemitsu, *Phys. Rev. D* **82**, 114504 (2010).
- [63] E. Lunghi and J. Matias, *J. High Energy Phys.* **04** (2007) 058.
- [64] B. Grinstein, R. P. Springer, and M. B. Wise, *Nucl. Phys.* **B339**, 269 (1990).
- [65] B. Grinstein, M. J. Savage, and M. B. Wise, *Nucl. Phys.* **B19**, 271 (1989).
- [66] T. M. Aliev, G. Hiller, and E. O. Iltan, *Nucl. Phys.* **B515**, 321 (1998).
- [67] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, [arXiv:1510.04239](https://arxiv.org/abs/1510.04239).
- [68] S. Davidson and G. J. Grenier, *Phys. Rev. D* **81**, 095016 (2010).
- [69] I. Baum, V. Lubicz, G. Martinelli, L. Orifici, and S. Simula, *Phys. Rev. D* **84**, 074503 (2011).
- [70] D. Becirevic, V. Lubicz, G. Martinelli, and F. Mescia (SPQcdR Collaboration), *Phys. Lett. B* **501**, 98 (2001).
- [71] C. Greub, T. Hurth, M. Misiak, and D. Wyler, *Phys. Lett. B* **382**, 415 (1996).
- [72] S. Fajfer, A. Prapotnik, S. Prelovsek, P. Singer, and J. Zupan, *Nucl. Phys. B, Proc. Suppl.* **115**, 93 (2003).
- [73] J. M. Cline, K. Kainulainen, and M. Trott, *J. High Energy Phys.* **11** (2011) 089.
- [74] V. Khachatryan *et al.* (CMS Collaboration), *Phys. Lett. B* **749**, 337 (2015).
- [75] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **11** (2015) 211.
- [76] J. F. Kamenik and F. Mescia, *Phys. Rev. D* **78**, 014003 (2008).
- [77] CMS Collaboration, Report No. CMS-PAS-HIG-15-012.
- [78] J. R. Espinosa, M. Muhlleitner, C. Grojean, and M. Trott, *J. High Energy Phys.* **09** (2012) 126.