

Naturally large Yukawa hierarchies

Enrico Nardi*

*INFN, Laboratori Nazionali di Frascati, Via Enrico Fermi 40, I-00044 Frascati, Italy,
 Departamento de Física Teórica, C-XI, Facultad de Ciencias, Universidad Autónoma de Madrid,
 C.U. Cantoblanco, 28049 Madrid, Spain,
 and Instituto de Física Teórica UAM/CSIC, Nicolas Cabrera 15, C.U. Cantoblanco, 28049 Madrid, Spain*
 (Received 15 May 2011; published 11 August 2011)

The spontaneous breaking of the $SU(3)^5$ quark/lepton flavor symmetry by means of three multiplets of scalar “Yukawa fields” admits vacua with one $\mathcal{O}(1)$ and two vanishing vacuum expectation values for each multiplet. If the number of generations is equal to three, and only in this case, the vanishing vacuum expectation values are lifted to exponentially suppressed entries by the inclusion of symmetry invariant logarithmic terms. A strong hierarchy for the Yukawa couplings and a quark mixing matrix that approaches a diagonal form are obtained in a natural way from $\mathcal{O}(1)$ parameters. This scenario provides a concrete realization of the minimal flavor violation hypothesis.

DOI: 10.1103/PhysRevD.84.036008

PACS numbers: 11.30.Hv, 11.30.Qc, 12.15.Ff

I. INTRODUCTION

The standard model (SM) provides an accurate description of particle physics phenomena. Particle interactions are derived from local symmetries and are explained at a fundamental level by the gauge principle. Myriads of experimental tests have confirmed the correctness of this picture. However, the SM cannot explain the values of the particle masses and mixing angles. Even if the seed of fermion masses is eventually identified, as it might happen soon at the LHC, a new theory is required for explaining the puzzling features of the observed pattern of Yukawa couplings: (1) For each value of the electric charge $Q = -1, -\frac{1}{3}, +\frac{2}{3}$, there is a threefold replica of fermions that are characterized by the same set of (known) quantum numbers, and as a result, the Yukawa couplings are arranged into generic 3×3 matrices. We have no clue about the origin of fermion family replication. (2) Unless new quantum charges are postulated, fermions belonging to different generations but with the same quantum numbers under the SM gauge group are indistinguishable at the fundamental level. It is then puzzling that their masses are instead arranged in a strong hierarchical structure. (3) The Yukawa matrices for the up and down quarks constitute two sets of mutually independent parameters. It is then surprising that in the basis in which the down-quark Yukawa matrix is diagonal and with a given ordering of its entries e.g. from small to large, the Yukawa matrix for the up quarks, when ordered in the same way, is also approximately diagonal. (4) Additionally, the theoretical prejudice that there is new physics not too far above the electroweak scale brings in one more puzzle: why new physics effects are not seen in flavor violating processes. We believe that the scenario we are going to discuss can shed some light on all these issues.

II. SYMMETRY AND RENORMALIZABLE INVARIANTS

Looking at the SM gauge sector one can readily recognize that fermions are arranged into triplets of states with the same gauge quantum numbers. It is then natural to postulate some symmetry group that commutes with the SM gauge group and has three-dimensional representations. The symmetry, however, is not realized in the spectrum, and generally this signals a noninvariant ground state yielding spontaneous symmetry breaking (SSB). To pursue further these simple considerations, we need to identify the symmetry group, and make an ansatz about the way it is broken. A brief review of known properties of the SM can guide us in carrying out this task.

The group of symmetry transformations of the SM quarks' and leptons' gauge invariant kinetic terms is [1] $\mathcal{G} = \mathcal{G}^{(q)} \times \mathcal{G}^{(l)}$ with $\mathcal{G}^{(q)} = U(3)_Q \times U(3)_u \times U(3)_d$ and $\mathcal{G}^{(l)} = U(3)_\ell \times U(3)_e$, where Q and ℓ denote the quark and lepton $SU(2)$ doublets, and u, d , and e the quark and lepton $SU(2)$ singlets. In the SM \mathcal{G} is broken explicitly by the fermion Yukawa couplings; however, some $U(1)$ factors are left unbroken. In the quark sector $U(1)_Y \times U(1)_B$ of hypercharge and baryon number remain good symmetries, and in the lepton sector $U(1)_Y$ remains unbroken as well. The SM lepton sector is, however, incomplete since it cannot accommodate massive neutrinos, and it is likely, although experimentally not yet confirmed, that unlike baryon number, $U(1)_L$ of lepton number is broken. Whether this is true (Majorana neutrinos) or not (Dirac neutrinos) is not relevant for our discussion; thus in the following we assume that the broken subgroup of \mathcal{G} is $\mathcal{G}_B = \mathcal{G}_B^{(q)} \times \mathcal{G}_B^{(l)}$ with

$$\mathcal{G}_B^{(q)} = SU(3)_Q \times SU(3)_u \times SU(3)_d \times U(1)_d, \quad (1)$$

$$\mathcal{G}_B^{(l)} = SU(3)_\ell \times SU(3)_e \times U(1)_e. \quad (2)$$

*enrico.nardi@lnf.infn.it

The Abelian factors $U(1)_{d,e}$ correspond to phase rotation of the $SU(2)$ singlet d quarks and e leptons. For the quarks, we could have equally well chosen the Abelian factor to be $U(1)_u$; however, assuming that the Yukawa couplings of the d quarks (and of the leptons [2,3]) break an additional symmetry with respect to the Yukawa couplings of the u quarks can provide a simple justification for the suppression for the bottom (and tau) mass with respect to the mass of the top. We will exploit this symmetry argument in what follows.

Guided by the previous considerations, we assume a fundamental symmetry that contains $\mathcal{G}_B = \mathcal{G}_B^{(q)} \times \mathcal{G}_B^{(l)}$ as a subgroup that gets spontaneously broken, and that under the quark Eq. (1) and lepton Eq. (2) factors, the SM fermions transform, respectively, as

$$Q = (3, 1, 1)_0, \quad u = (1, 3, 1)_0, \quad d = (1, 1, 3)_1, \quad (3)$$

$$\ell = (3, 1)_0, \quad e = (1, 3)_1. \quad (4)$$

As regards the way \mathcal{G}_B is dynamically broken, the simplest choice is to interpret the SM explicit breaking as the result of SSB. That is, we assume that the Yukawa couplings of the SM quarks and leptons correspond to vacuum expectation values (vev) of scalar fields Y_u, Y_d, N_d and Y_e, N_e that are coupled to the fermions in a symmetry invariant way via nonrenormalizable operators:

$$-\mathcal{L}_Y = \frac{1}{\Lambda} \bar{Q} Y_u u H + \frac{1}{\Lambda^2} N_d \bar{Q} Y_d d \tilde{H} + \frac{1}{\Lambda^2} N_e \bar{\ell} Y_e e \tilde{H}, \quad (5)$$

where H is the Higgs field (with $\tilde{H} = i\sigma_2 H^*$) and Λ is a large scale where the effective operators in Eq. (5) arise. Invariance of \mathcal{L}_Y under $\mathcal{G}_B^{(q)}$ and $\mathcal{G}_B^{(l)}$ fixes the following assignments:

$$Y_u = (3, \bar{3}, 1)_0, \quad Y_d = (3, 1, \bar{3})_0, \quad N_d = (1, 1, 1)_{-1}, \quad (6)$$

$$Y_e = (3, \bar{3})_0, \quad N_e = (1, 1)_{-1}. \quad (7)$$

A more economical choice than the three multiplets of scalars in Eq. (6) and the two in Eq. (7) is also possible. In fact, by assigning $U(1)_{d,e}$ charges to $Y_{d,e}$ we would not need to introduce the complex scalars $N_{d,e}$. It is, however, more convenient to keep a clear distinction between the hierarchy between the top and bottom/tau masses from the Yukawa hierarchy between quarks and leptons of the same type. For simplicity, we will describe the first one by means of two ‘‘Abelian spurions’’ $\eta_{N_{d,e}} \equiv \langle N_{d,e} \rangle / \Lambda$, and we take $\eta_{N_{d,e}} \approx m_{b,\tau} / m_t$ as given numbers. We will briefly comment on the SSB of $U(1)_{d,e}$ only at the end of the paper. The hierarchy between generations is instead explained via SSB of the $SU(3)^5$ quark/lepton flavor symmetry, that is, by the dynamical selection of vevs with the required structure.

We note, in passing, that the introduction of the Abelian spurion η_{N_d} implies that $U(1)_d$ invariant operators like $\bar{Q} Y_d Y_d^\dagger Q$ are not suppressed with respect to $\bar{Q} Y_u Y_u^\dagger Q$.

This is different from what is commonly assumed in minimal flavor violation (MFV) extensions of the SM [4], in which effective operators involving Y_d are always suppressed, and it resembles more MFV extensions of two Higgs doublets models like supersymmetry [4] in which $\tan\beta$ plays basically the role of $\eta_{N_{d,e}}$. The formal difference is that, in the present case, the absence of suppression factors for $Y_d Y_d^\dagger$ follows from a symmetry argument.

The assignments in Eqs. (3) and (4) imply that all the fermions with the same quantum numbers under the SM gauge group are characterized by the same quantum numbers also under \mathcal{G}_B , and therefore the different generations contain exact replica of the same set of states. As we will see, the fact that for each triplet of identical fermions two Yukawa couplings in first approximation vanish, while the third one is $\mathcal{O}(1)$, corresponds precisely to a nonsymmetric ground state that yields a symmetry breaking pattern in qualitative agreement with observations. Note that this picture is fundamentally different from assuming new symmetries under which fermions with the same SM quantum numbers transform differently; that is, for example, the basic ingredient of the popular Froggatt-Nielsen mechanism [5]. In that case the hierarchy of the Yukawa couplings follows from a dimensional hierarchy in the corresponding effective Yukawa operators, which is obtained by assigning to the lighter generations larger values of new Abelian charges. The fact that fermion families appear to replicate would then be just an illusory feature due to our incomplete knowledge of the fundamental symmetries, and not a fundamental property of the SM fermions.

Of course, the idea of promoting the Yukawa couplings to vevs of scalar fields is not new. A scenario proposed recently that is based on this idea is, for example, Koide’s ‘‘yukawaon’’ model (see e.g. [6,7]). In that case a unique flavor symmetry group $O(3)$ is assumed, with symmetric matrices of Yukawa fields (the yukawaons) transforming in the reducible $(\mathbf{3} \times \mathbf{3})_{\text{symm}} = \mathbf{1} + \mathbf{5}$ of the group. The Yukawa fields for the different fermion species are then distinguished by introducing additional $U(1)$ factors, under which the fermions and yukawaons transform with suitable charges. Clearly, since the flavor symmetry is very different with respect to the one we are considering, Eqs. (1) and (2), the scalar potential for the yukawaons also has little resemblance to the one studied in this paper. A framework which is more similar to ours is discussed in Ref. [8]. This paper contains a careful analysis of the broken and unbroken Abelian symmetries, and also introduces a nice set of correspondences between physical observables (fermion masses and mixing angles) and invariants of the scalar potential. Given the large number of observables, this set includes several nonrenormalizable terms. In contrast, we only consider the renormalizable potential, which contains enough information to explore qualitatively the fermion Yukawa hierarchies but, as we will see, not enough to also pin down the quark mixing angles. Another difference is

that Ref. [8] avoids the introduction of Abelian spurions by assigning Abelian charges to the Yukawa fields, with the result that some terms that are allowed for us are forbidden in their case by the $U(1)$ symmetries. A study of the scalar potential for the Yukawa fields, including higher order operators, has also been presented in [9], which also puts forth the possibility that the Yukawa couplings are generated from pairs of vevs of scalar fields that transform, respectively, only under $SU(3)_Q$ and $SU(3)_{u,d}$. This scenario can be simply realized by replacing, for example, Y_u in Eq. (5) by a pair of fields $\chi_u^L \cdot \chi_u^R / \Lambda$. Altogether, while the issue of constructing a symmetry invariant scalar potential for the Yukawa fields, including higher dimension invariant terms, has already been addressed, for example, in the quoted Refs. [8,9], to our knowledge the possible effects of quantum corrections to the effective potential, that in our approach play a crucial role, have not been considered before.

Symmetry invariants and the T, A, \mathcal{D} parametrization

To carry out a basis independent analysis of the SSB of the $SU(3)^5$ flavor symmetry, it is convenient to write the multiplets of scalar ‘‘Yukawa fields’’ $Y_{u,d,e}$ in their singular value decomposition:

$$\begin{aligned} Y_u &= \mathcal{V}_u^\dagger \chi_u \mathcal{U}_u, & Y_d &= \mathcal{V}_d^\dagger \chi_d \mathcal{U}_d, \\ Y_e &= \mathcal{V}_e^\dagger \chi_e \mathcal{U}_e, \end{aligned} \quad (8)$$

where the matrices \mathcal{V} and \mathcal{U} are unitary while the matrices χ are diagonal with nonnegative real entries:

$$\begin{aligned} \chi_u &= \text{diag}(u_1, u_2, u_3), & \chi_d &= \text{diag}(d_1, d_2, d_3), \\ \chi_e &= \text{diag}(e_1, e_2, e_3). \end{aligned}$$

Note that while the matrices of *singular values* χ are unique (modulo reordering of their entries) \mathcal{V} and \mathcal{U} are not, and can be redefined according to $\mathcal{V} \rightarrow \Phi \mathcal{V}$, $\mathcal{U} \rightarrow \Phi \mathcal{U}$, where Φ is a diagonal matrix of phases. This can be used to remove three phases, for example, in \mathcal{V} .

We now require invariance of the scalar potential for the Yukawa fields under the special bi-unitary transformations $Y_{u,d} \rightarrow V_Q Y_{u,d} U_{u,d}^\dagger$ and $Y_e \rightarrow V_\ell Y_e U_e^\dagger$ with $\det V_{Q,\ell} = \det U_{u,d,e} = +1$. Assuming three generations, for the u -quark sector we can write the following invariants:

$$T_u = \text{Tr}(Y_u Y_u^\dagger) = \sum_i u_i^2, \quad (9)$$

$$A_u = \text{Tr}[\text{adj}(Y_u Y_u^\dagger)] = \frac{1}{2} \sum_{i \neq j} u_i^2 u_j^2, \quad (10)$$

$$\mathcal{D}_u = \det(Y_u) = e^{i\delta_u} \prod_i u_i \equiv e^{i\delta_u} D, \quad (11)$$

where $\delta_u = \arg[\det(\mathcal{V}_u^\dagger \mathcal{U}_u)]$, and a similar relation holds for $\mathcal{D}_u^* = \det(Y_u^\dagger)$ with $\delta_u \rightarrow -\delta_u$. The invariance of the

trace T_u and of the determinant \mathcal{D}_u under a $SU(3)_Q \times SU(3)_u$ transformation $Y_u \rightarrow V_Q Y_u U_u^\dagger$ is obvious. The second invariant in Eq. (10) is the trace of the *adjugate*, that is, the trace of the transpose of the matrix of cofactors. Its invariance can be proved as follows: Laplace’s formula applied to the product of two $n \times n$ matrices A and B reads $(AB) \cdot \text{adj}(AB) = \det(AB) = \det(A) \cdot \det(B)$ from which the product rule $\text{adj}(AB) = \text{adj}(B) \cdot \text{adj}(A)$ is easily derived. Moreover, for a special unitary matrix $\text{adj}(V) = V^\dagger$. Invariance of A_u under $Y_u Y_u^\dagger \rightarrow V_Q Y_u Y_u^\dagger V_Q^\dagger$ then follows straightforwardly. For general $n \times n$ matrices, $T_{uu} = \text{Tr}(Y_u Y_u^\dagger Y_u Y_u^\dagger) = \sum_i u_i^4$ is also a renormalizable invariant. However, in the case of 3×3 matrices it is not an independent one: $T_{uu} = T_u^2 - A_u$. Two other sets of T, A, \mathcal{D} invariants completely similar to Eqs. (9)–(11) can be written also for Y_d and Y_e , and thus all the results that we will derive for Y_u apply equally well to them, too. Therefore, in the following we will drop wherever possible the subscript u , and we will refer to the parametrization of the T, A, \mathcal{D} invariants in Eqs. (9)–(11) as the TAD parametrization.

III. SYMMETRY BREAKING

In this section we study the general potential for the Yukawa fields $Y_{u,d,e}$ invariant under \mathcal{G}_B , and we classify the different minima that yield SSB. The renormalizable potential constructed from the T, A, \mathcal{D} invariants, Eqs. (9)–(11), reads¹

$$\hat{V} = \Lambda^4 V = \Lambda^4 (V_T + V_A + V_D), \quad (12)$$

$$V_T = \lambda \left[T - \frac{m^2}{2\lambda} \right]^2, \quad (13)$$

$$V_A = \tilde{\lambda}' A, \quad (14)$$

$$V_D = \tilde{\mu} D + \tilde{\mu}^* \mathcal{D}^* = 2\mu \cos(\phi + \delta) D. \quad (15)$$

In the first equation we have factored out a large scale Λ that, for simplicity, can be identified with the effective scale in Eq. (5), so that all the parameters and fields in $V = V_T + V_A + V_D$ have no dimensions; for definiteness, we also assume that all the Lagrangian parameters are renormalized at this same scale. V_T in Eq. (13) contains the two invariants constructed from the trace $V_T = \lambda T^2 - m^2 T$ plus an irrelevant constant. We require $\lambda > 0$ and $m^2 > 0$ in order to have a potential bounded from below and to trigger SSB. The last equality in Eq. (15) is obtained by defining $\tilde{\mu} = \mu e^{i\phi}$, that is, $\mu \equiv |\tilde{\mu}| = |\tilde{\mu}^*|$. The parameter $\tilde{\lambda}'$ that multiplies A can be either positive or negative,

¹According to our simplification of trading the scalars $N_{d,e}$ for the two spurions $\eta_{N_{d,e}}$, we initially omit writing $NN^\dagger T$ terms. A coupling with the Higgs $HH^\dagger T$ can also be omitted as long as $\langle HH^\dagger \rangle / \Lambda^2 \ll m^2$.

and we need to consider both possibilities. In the following we will refer to its absolute value by simply dropping the tilde, $\lambda' \equiv |\tilde{\lambda}'|$.

By exploiting the TAD parametrization introduced in Eqs. (9)–(11) the hierarchy of the SM Yukawa couplings can be conveniently described in terms of vevs of invariants, and corresponds to minima that satisfy the condition

$$\langle D \rangle^{1/3} \ll \langle A \rangle^{1/4} \ll \langle T \rangle^{1/2}. \quad (16)$$

Of course, in our dimensionless approach the exponents can be equally well dropped. There is a simple correspondence between the TAD vevs and the vevs of the field components $\langle u_i \rangle$. For example, assuming that a vev $\langle \chi \rangle$ with a large hierarchy for its components has been found, by labeling its entries according to $\langle u_1 \rangle \ll \langle u_2 \rangle \ll \langle u_3 \rangle$ we have

$$\frac{\langle A \rangle}{\langle T \rangle^2} \approx \frac{\langle u_2^2 \rangle}{\langle u_3^2 \rangle}, \quad \frac{\langle D \rangle^2}{\langle T \rangle \langle A \rangle} \approx \frac{\langle u_1^2 \rangle}{\langle u_3^2 \rangle}. \quad (17)$$

Since the value of the top Yukawa coupling fixes $\langle T_u \rangle \approx \langle u_3^2 \rangle \approx 1$ and naturalness suggests $\langle T_{d,e} \rangle \approx \mathcal{O}(1)$ as well, it follows that vacua characterized in the first approximation by $\langle D \rangle = \langle A \rangle = 0$ and $\langle T \rangle \approx 1$ are well suited for generating the Yukawa hierarchies.

From Eq. (13) we immediately see that V_T is minimized for values on the spherical surface in $\langle u_i \rangle$ space $\langle T \rangle = \frac{m^2}{2\lambda}$. Note that while we must require $\frac{m^2}{2\lambda} \approx 1$, a perturbative $\lambda < 1$ implies $m^2 < 1$, and thus in the first approximation contributions of nonrenormalizable operators to V can be neglected. As regards A and D , they are both maximized for symmetric vacua $\langle \chi \rangle = (u_s, u_s, u_s)$ and minimized when they vanish. To ensure $\langle D \rangle = 0$ either $\langle \delta \rangle = \pm \frac{\pi}{2} - \phi$ or at least one entry in χ must vanish, while for $\langle A \rangle = 0$ two entries must vanish. Which particular minimum on the surface $\langle T \rangle = \text{const}$ is selected then depends on the signs and values of $\tilde{\lambda}'$ and of $\mu \cdot \cos(\phi + \delta)$. Below we classify the different types of minima yielding SSB, and we show that vacua with the required property $\langle D \rangle = \langle A \rangle = 0$ indeed occur.

Case 1.— $\tilde{\lambda}' = -\lambda' < 0$. V_A is negative and its absolute value is maximized for symmetric vacua $\langle \chi \rangle = (u_s, u_s, u_s)$. A negative value of V_D further lowers the minimum, which fixes $\langle \delta \rangle = \pi - \phi$, while D is also maximized for symmetric vacua. The potential is bounded if $\lambda' < 3\lambda$, in which case we obtain

$$u_s = \frac{\mu}{4(3\lambda - \lambda')} \left[1 + \sqrt{1 + 8(3\lambda - \lambda') \frac{m^2}{\mu^2}} \right]. \quad (18)$$

This case yields a phenomenologically uninteresting non-hierarchical pattern $\langle T \rangle \approx \langle D \rangle \approx \langle A \rangle \approx 1$.

Case 2.— $\tilde{\lambda}' = \lambda' > 0$. In this case $V_A = \lambda A$ is positive and is minimized when $A = 0$, which favors vacua with two vanishing entries $\langle \chi \rangle = (0, 0, u_t)$. The value of V_D is extremized in either of the following cases.

Case 2a.— $V_D = -2\mu D$, which occurs for $\langle \delta \rangle = \pi - \phi$. In this case $\langle \chi \rangle$ acquires a symmetric structure (u_s, u_s, u_s) that maximizes D .

Case 2b.— $\langle \chi \rangle$ has at least one vanishing entry yielding $V_D = 0$. In this case $\langle \delta \rangle$ is left undetermined. $\langle \chi \rangle = (0, 0, u_t)$ is in fact favored over a single vanishing entry because it also ensures $V_A = 0$. The two vevs are

$$u_s = \frac{\mu}{4(3\lambda + \lambda')} \left[1 + \sqrt{1 + 8(3\lambda + \lambda') \frac{m^2}{\mu^2}} \right], \quad (19)$$

$$u_t = \frac{m}{\sqrt{2\lambda}}. \quad (20)$$

Since $V(u_t) = 0$, vacua with two vanishing entries are selected if $V(u_s)$ is positive, which occurs for

$$\frac{\mu^2}{m^2} < 2\lambda \left[\left(4 + \frac{\lambda'}{\lambda} \right)^{3/2} - \left(8 + 3 \frac{\lambda'}{\lambda} \right) \right]. \quad (21)$$

In this case $\langle T \rangle \approx 1$ while $\langle D \rangle = \langle A \rangle = 0$, which represents an interesting first approximation to Eq. (16).

A natural question to ask is what the effect could be on the hierarchies of vevs of adding to the tree level potential nonrenormalizable terms like, for example, $T_m^n = (T[(YY^\dagger)^m])^n$, $A_m^n = (A[(YY^\dagger)^m])^n$, $D_m^n = \mathcal{D}[Y^n Y^{\dagger m}]$. One can easily convince himself that no qualitative differences would arise. This is because after imposing the SSB condition $\langle T \rangle = \text{const} \neq 0$ all these terms are maximized for symmetric vacua $\langle \chi \rangle = (u_s, u_s, u_s)$ and minimized when one or two entries vanish, which still gives the same types of vacua found with the renormalizable potential. In particular, there is no way to lift the vanishing vevs by introducing higher order operators, and no hierarchical patterns can be generated starting from $\mathcal{O}(1)$ parameters (that is, without resorting to fine-tuning).

A. Lifting the vanishing vevs

In case 2b the vacua are characterized by $\langle D \rangle = \langle A \rangle = 0$, which is a good starting condition to generate minima satisfying Eq. (16). Of course, to have an acceptable phenomenology we must lift these two vevs to appropriately small, but nonvanishing values. In order to achieve this let us add to V an invariant term proportional to the logarithm of the three-point interaction:

$$V_D \rightarrow V_D + V_D^{(1)} = 2\mu \cos(\phi + \delta) D(1 + c_D \log D), \quad (22)$$

where c_D is a small numerical factor. Here, at the cost of some arbitrariness, we justify the introduction of $V_D^{(1)}$ only on the basis of symmetry properties and of naive dimensional analysis; however, we can expect that terms of this type will be generated by quantum corrections to the effective potential, in which case the loop coefficient c_D would be computable. The effect of $V_D^{(1)}$ is that of generating a large negative log that shifts the term $D(1 + c_D \log D)$

towards a nonvanishing negative value. The minimum is then obtained for $\langle \delta \rangle = -\phi$ and for a nonvanishing, but exponentially suppressed value of D :

$$\langle D \rangle = e^{-((1/c_D)+1)} \equiv \epsilon_D. \quad (23)$$

This gives an elegant realization of an old remark by Nambu [10]. Given that in the unperturbed solution $\langle \chi \rangle = (0, 0, u_t)$ we have $u_t \approx 1$, Eq. (23) implies $\langle u_1 \rangle \cdot \langle u_2 \rangle \equiv u_u \cdot u_c \approx \epsilon_D$. Note that $\log D$ is unable to differentiate between u_u and u_c , and that terms involving higher powers of $\log D$ are also ineffective for distinguishing these two entries. However, adding a $\log A$ term makes the magic that an exponential hierarchy is induced also between the two suppressed entries u_u and u_c .

$$V_A \rightarrow V_A + V_A^{(1)} = \tilde{\lambda}' A (1 + c_A \log A) \quad (24)$$

in fact yields

$$\langle A \rangle = e^{-((1/c_A)+1)} \equiv \epsilon_A. \quad (25)$$

Then, if we assume $c_D < c_A$ the hierarchical pattern in Eq. (16) is realized. In terms of vevs of component fields, setting, for simplicity, $u_t = 1$, Eqs. (23) and (25) give a system of two equations,

$$u_u \cdot u_c = \epsilon_D, \quad u_u^2 + u_c^2 = \epsilon_A - \epsilon_D^2, \quad (26)$$

that has real solutions for $\epsilon_D \leq \epsilon_A/2$. At first order in ϵ_D/ϵ_A the solutions are

$$u_u^2 = \frac{\epsilon_D^2}{\epsilon_A}, \quad u_c^2 = \epsilon_A, \quad (27)$$

where the particular labeling u and c is of course arbitrary. Equivalently, inserting (23) and (25) into Eq. (17) we obtain

$$\frac{u_c^2}{u_t^2} \approx e^{-((1/c_A)+1)}, \quad \frac{u_u^2}{u_c^2} \approx e^{-2((1/c_D)-(1/c_A))}. \quad (28)$$

Without introducing any unnaturally small parameters, we thus obtain a pattern of Yukawa couplings that is characterized by an exponentially strong hierarchy. For the up-quark sector, for example, we need $c_A \sim -(\log \langle A_u \rangle)^{-1} \approx 0.10$ and $c_D \sim -(\log \langle D_u \rangle)^{-1} \approx 0.06$.² Finally, we should also add to the λT^2 interaction a $\log T$ correction. This, however, has no particular consequences since, in any case, the minimization of V_T fixes $\langle T \rangle \approx 1$.

B. Coupling the up- and down-quark sectors

As long as the up- and down-quark sectors are treated separately, the relative hierarchical ordering of the vevs of the component fields $\langle u_i \rangle$ and $\langle d_i \rangle$ is irrelevant, and we have a set of $3! \times 3!$ equivalent vacua. This large degeneracy is partially reduced when the renormalizable

invariants that couple the two sectors are included. There are two possible terms:

$$T_u \cdot T_d = \text{Tr}(Y_u Y_u^\dagger) \cdot \text{Tr}(Y_d Y_d^\dagger) = \sum_{ij} u_i^2 d_j^2, \quad (29)$$

$$T_{ud} = \text{Tr}(Y_u Y_u^\dagger Y_d Y_d^\dagger) = \text{Tr}(V^\dagger \chi_u^2 V \chi_d^2), \quad (30)$$

where V is a unitary matrix of fields that, in terms of the bi-unitary parametrization Eq. (8), is given by $V = \mathcal{V}_u \mathcal{V}_d^\dagger$. It is now convenient to replace T_{ud} by the invariant combination

$$A_{ud} \equiv T_u \cdot T_d - T_{ud}. \quad (31)$$

Note that A_{ud} is not related to adjugate matrices, and its invariance follows solely from its definition. The contribution to the scalar potential from coupling the u and d sectors is

$$V_{ud} = 2\tilde{\lambda}_{ud} T_u \cdot T_d + \tilde{\lambda}'_{ud} A_{ud} + \dots, \quad (32)$$

where the dots stand for logarithmic terms, that we expect could be relevant (see below) but whose effects in this case are difficult to analyze analytically, and thus, as a first approximation, we leave them out. We now show that the first term in V_{ud} yields the same vacuum conditions implied by Eq. (13), that is, that the minimum of the potential is localized on the surface of two three-spheres of constant $\langle T_u \rangle$ and $\langle T_d \rangle$. We start by shifting the couplings of the terms linear and quadratic in $T_{u,d}$ [see Eq. (13)] according to

$$m_{u,d}^2 \rightarrow m_{u,d}^2 + \tilde{m}_{u,d}^2, \quad (33)$$

$$\lambda_{u,d} \rightarrow \lambda_{u,d} + \tilde{\lambda}_{u,d}, \quad (34)$$

where $\tilde{m}_{u,d}^2$ in the first line is an arbitrary constant of the same sign as $\tilde{\lambda}_{u,d}$ that we will fix below. With a little algebra, and omitting irrelevant constants, we can rewrite this as

$$\begin{aligned} V_{T(u+d)} &\equiv V_{T_u} + V_{T_d} + 2\tilde{\lambda}_{ud} T_u \cdot T_d \\ &= \lambda_u \left[T_u - \frac{m_u^2}{2\lambda_u} \right]^2 + \lambda_d \left[T_d - \frac{m_d^2}{2\lambda_d} \right]^2 \\ &\quad + \tilde{\lambda}_{ud} \left[T_u + T_d - \frac{m_{ud}^2}{2\lambda_{ud}} \right]^2. \end{aligned} \quad (35)$$

Note that the ratio $\tilde{m}_{ud}^2/\tilde{\lambda}_{ud}$ that would appear in the last set of square brackets is always positive by construction, which justifies omitting the tilde on both parameters. We can now fix m_{ud}^2 to satisfy

$$\frac{m_{ud}^2}{\lambda_{ud}} = \frac{m_u^2}{\lambda_u} + \frac{m_d^2}{\lambda_d}, \quad (36)$$

where $\lambda_{u,d}$ and $m_{u,d}^2$ are the redefined parameters appearing in the right-hand side of Eqs. (33) and (34). If $\tilde{\lambda}_{ud} > -\frac{\lambda_u \cdot \lambda_d}{\lambda_u + \lambda_d}$ SSB occurs, with the minimum of the potential

²Because the number of field components in Y is large, $N = 18$, a typical loop factor $\frac{N}{64\pi^2} \approx 0.03$ is of about the correct size.

located on the two surfaces $\langle T_u \rangle = \frac{m_u^2}{2\lambda_u}$ and $\langle T_d \rangle = \frac{m_d^2}{2\lambda_d}$. Note that the mixed $T_{u,d} \cdot T_e$ terms can also be “reabsorbed” in a similar way. Genuinely new effects, and, in particular, the relative ordering of the Yukawa hierarchies of the u and d sectors, come from the second term in Eq. (32) that can be written more explicitly as³

$$\tilde{\lambda}'_{ud} A_{ud} = \tilde{\lambda}'_{ud} \sum_{ij} (1 - |V_{ij}|^2) u_i^2 d_j^2. \quad (37)$$

Since V is unitary, $|V_{ij}|^2 \leq 1$, and A_{ud} then is the sum of positive semidefinite terms that cannot all vanish, which ensures $A_{ud} > 0$. For $\tilde{\lambda}'_{ud} = \lambda'_{ud} > 0$ the contribution to the potential is then minimized when A_{ud} is at its minimum, and, as we will now show, this occurs when $\langle V_{ij} \rangle$ approaches a diagonal form. Let us consider the vacuum configurations obtained in case 2b and label the two largest component vevs as $u_t \approx d_b \sim \mathcal{O}(1)$. All the other vevs are exponentially suppressed and generically of $\mathcal{O}(\epsilon)$. We can then write

$$A_{ud} = (1 - |V_{tb}|^2) \times \mathcal{O}(1) + \sum_{(ij) \neq (tb)} (1 - |V_{ij}|^2) \times \mathcal{O}(\epsilon). \quad (38)$$

Small values $A_{ud} \sim \mathcal{O}(\epsilon)$ can result only when the modulus of the field V_{tb} that couples to the entries in χ_u and χ_d that have the largest vevs is exponentially close to 1, and thus the minimum must occur around this configuration. Unitarity then implies that the off-diagonal vevs $\langle |V_{ij}|^2 \rangle$ ($j \neq b$) and $\langle |V_{jb}|^2 \rangle$ ($j \neq t$) are accordingly suppressed. In this way a “third generation structure” emerges, in the sense that t and b get almost decoupled from the other quarks. For the two lighter generations it is more difficult to carry out this argument: we would need to confront various different contributions to A_{ud} that are all of $\mathcal{O}(\epsilon)$, and thus we are not allowed to neglect either $V_{D,A}^{(1)}$ that are also of $\mathcal{O}(\epsilon)$, or a $\log T_{ud}$ term that can be expected to induce corrections of a similar size. However, it is also clear that at this level we do not have enough information to determine univocally the structure of $\langle V \rangle$. For example, the fact that V_{ij} enters Eq. (38) only through its modulus square implies that there is no information on complex phases.

To summarize, we have seen that once the component vevs in $\langle \chi_u \rangle$ and $\langle \chi_d \rangle$ are conventionally ordered in the same way, e.g. with increasing size, minimizing the coupling term A_{ud} pushes $\langle V \rangle$ to approach a diagonal form, a dynamical behavior that can provide an explanation for the most puzzling feature of the Cabibbo Kobayashi Maskawa matrix. Although only qualitative, this result is certainly nontrivial; in fact *a priori* nothing could have guaranteed that the largest mixings will not occur between the $\mathcal{O}(1)$

³Using unitarity $\sum_j |V_{ij}|^2 = \sum_j |V_{ji}|^2 = 1$, we can also write $A_{ud} - \sum_{i \neq j} u_i^2 d_j^2 = \sum_{ij} u_i^2 (d_i^2 - d_j^2) |V_{ij}|^2 = \sum_{ij} d_i^2 (u_i^2 - u_j^2) |V_{ji}|^2$ which puts in evidence that if $\langle d_i \rangle = \langle d_j \rangle$ or $\langle u_j \rangle = \langle u_i \rangle$ the corresponding $\langle V_{ij} \rangle$ remains undetermined.

and the $\mathcal{O}(\epsilon)$ components in $\langle \chi_{u,d} \rangle$, and actually something like this would occur for $\tilde{\lambda}' = -\lambda' < 0$.

IV. WHY THREE GENERATIONS?

Our approach to explain the fermion mass hierarchy does not provide an explanation for why there are three families; however, it does imply an interesting connection between the number of generations n_g and the qualitative features of the Yukawa couplings’ hierarchy. Since this section contains mainly simple arguments based on dimensional analysis, we simplify the discussion by considering only the case of a real determinant $\mathcal{D} = D$.

For $n_g > 4$, A and D have $\dim(A, D) > 4$ and thus cannot appear in the renormalizable potential. However, $T_{uu} = \text{Tr} \chi_u^4$ is now independent of T_u and A_u and provides a new renormalizable invariant term. The potential can be conveniently written as $V = \lambda(T - \eta^2)^2 + \tilde{\lambda}' \sum_{i \neq j} u_i^2 u_j^2$, where the first term yields $\langle T \rangle = \eta^2$ while (for $\tilde{\lambda}' = \lambda' > 0$) minimization of the second term implies that $n_g - 1$ components have vanishing vevs. We see that the absence of additional renormalizable interactions does not allow either lifting the vanishing determinant or generating a hierarchical pattern for the $n_g - 1$ vanishing couplings.

If $n_g = 4$, then $\dim(D) = 4$ and we can arrange for an exponentially suppressed $\langle D \rangle \neq 0$. However, $\dim(A) = 6$, and then besides T_u^2 , T_{uu} , and D , there are no other renormalizable interactions. Therefore, also in this case it is not possible to remove all the degeneracies between the suppressed vevs and obtain a fully hierarchical pattern.

The case $n_g = 2$ is a bit more involved and requires a more detailed discussion. For 2×2 matrices $A_u = T_u$ and $T_{uu} = T_u^2 - 2D_u^2$, so we just have the two basic invariants T and D . We can write the renormalizable potential as

$$\hat{V} = \Lambda^4 V = \Lambda^4 (V_T + V_D), \quad (39)$$

$$V_T = \lambda \left(T - \frac{m^2}{2\lambda} \right)^2, \quad V_D = \tilde{\lambda}' \left(D + \frac{\tilde{\mu}^2}{2\tilde{\lambda}'} \right)^2. \quad (40)$$

SSB occurs for $\lambda > 0$ and $m^2 > 0$ and, if $\tilde{\lambda}' = \lambda' > 0$ and $\tilde{\mu}^2 = \mu^2 > 0$, the determinant vanishes at the minimum, implying a vanishing vev for one component field. We can try to lift $\langle D \rangle = 0$ by adding to the $\lambda' D^2$ interaction a logarithm:

$$V_D \rightarrow V_D + V_D^{(1)} = \mu^2 D + \lambda' D^2 (1 + c_D \log D). \quad (41)$$

Varying with respect to D and equating to zero gives the condition

$$\log \langle D \rangle = - \left(\frac{1}{c_D} + \frac{1}{2} \right) - \frac{\mu^2}{2\lambda' c_D} \frac{1}{\langle D \rangle} \equiv -\alpha - \frac{\beta}{\langle D \rangle}. \quad (42)$$

We have a minimum if $\beta \leq \langle D \rangle \ll 1$, where the second inequality, which implies $\frac{1}{c_D} \ll \frac{2\lambda'}{\mu^2} \sim \mathcal{O}(1)$, follows from

requiring a large hierarchy. Let us seek a solution of the form

$$\langle D \rangle = e^{-(\alpha - W)} \quad (43)$$

with W a suitable function. Substituting this formal solution into Eq. (42) yields

$$We^W = -\beta e^\alpha. \quad (44)$$

The function W defined implicitly through $W(x)e^{W(x)} = x$ is known in mathematics as the Lambert W function [11]. On the negative real axis W is real and two-valued over the interval $-1/e \leq x \leq 0$, with $W(-1/e) = -1$, and the two branches W_0 and W_{-1} are identified according to $W_0(x) \geq -1$ and $W_{-1}(x) \leq -1$. To suppress $\langle D \rangle$ adequately we need to make $\alpha - W$ sufficiently large. Given that $\frac{1}{c_D} \ll \mathcal{O}(1)$, $\alpha = \frac{1}{c_D} + \frac{1}{2}$ is never large. A large and negative W is, in principle, possible given that $W_{-1}(x) \rightarrow x \rightarrow 0 - \infty$. However, for small negative values of its argument, $|W_{-1}|$ grows only logarithmically, $W_{-1}(-x) \approx \log x - \log(-\log x)$, and therefore an exponential suppression of $\langle D \rangle$ would require an exponentially small β . Therefore, for two generations and natural values of the parameters, no strong hierarchy can result. We can then conclude that in our scenario a scalar potential invariant under a $SU(n_g)^5$ (flavor) symmetry can naturally yield a fully hierarchical pattern of Yukawa couplings only when $n_g = 3$.

V. SPONTANEOUS BREAKING OF THE ABELIAN SYMMETRIES

What we have done until now can be easily generalized to include the fields $N_{d,e}$ responsible for the breaking of the Abelian subgroups $U(1)_{d,e}$. Let us define for $i = d, e$ the $U(1)_i$ invariant bilinears $T_{N_i} = N_i N_i^\dagger$. These two invariants can couple in a renormalizable way only among themselves or to the traces $T_{u,d,e}$. Omitting A, D , and logarithmic terms, by generalizing the derivation of Eq. (35) we can write the scalar potential for the quadratic invariants $\mathcal{T} = (T_u, T_d, T_e, T_{N_d}, T_{N_e})$ as

$$V_{\mathcal{T}} = \sum_{IJ} \lambda_{IJ} [(\mathcal{T}_I - \eta_I^2) + (\mathcal{T}_J - \eta_J^2)]^2, \quad (45)$$

where $\lambda_{IJ} = \lambda_{JI}$. If all the eigenvalues of the Hessian $\partial_I \partial_J V_{\mathcal{T}}$ are positive, the potential is bounded and SSB occurs, resulting in the vevs $\langle \mathcal{T}_I \rangle = \eta_I^2$. There is no need to give the explicit relations between λ_{IJ} , η_I^2 and the parameters λ, m^2 that appear in the potential when written in a more familiar form; the important qualitative point is that one expects all $\eta_I^2 \sim \mathcal{O}(1)$. While, on the one hand, this means that the inclusion of SSB of $U(1)_{d,e}$ leaves the results of the previous sections unaffected, on the other hand the ratios between the bottom/tau masses and the

mass of the top suggest instead $\eta_N \sim 10^{-2}$. It would be rather unpleasant, after reaching so far, abandoning at this point our dogma about naturalness and assuming *ad hoc* values for the fundamental parameters in order to reproduce these two small numbers. However, there are other possible ways out. For example, we can assign to the $SU(2)$ singlets d and e Abelian charges larger than 1, which would imply a stronger dimensional suppression of their effective Yukawa operators. Taking, for example, $\eta_N \sim 0.3$, that we can still regard as a natural value, then a charge of 4 would yield the required small factor $\eta_N^4 \sim 10^{-2}$. Alternatively, we could simply assume that the heavy messengers carrying $U(1)_{d,e}$ charges needed to generate the effective Yukawa operators for d and e [see Eq. (5)] have a mass scale Λ' about 2 orders of magnitude larger than the scale Λ of the neutral messengers. In this way $\eta_N \approx (\Lambda/\Lambda') \eta_T \approx 10^{-2}$ is easily obtained.

VI. CONCLUSIONS

Let us recap what we have done and what we have achieved. In the SM, the fermions' gauge invariant kinetic terms are characterized by a $SU(5)^5$ quark/lepton flavor symmetry that is broken *explicitly* by the Yukawa couplings. It is a big puzzle why the values of these symmetry breaking parameters span 6 orders of magnitude, with no apparent regularity besides the fact that their strong hierarchy is qualitatively similar for all three types of fermions. Following the same approach proposed in recent works [8,9], we have promoted this symmetry to an exact one that is broken *spontaneously* by the ground state of a symmetry invariant scalar potential. We have chosen the multiplets of scalar fields in such a way that their vevs correspond precisely to the Yukawa couplings so that, among other things, by construction the SM Yukawa couplings are the only source of flavor violation. This promotes the spurion technique (widely used in connection with the MFV hypothesis) to a concrete piece of fundamental physics.

We have introduced a slight variant with respect to the most popular MFV scenarios by factorizing out the breaking of the two Abelian factors $U(1)_{d,e}$ from the breaking of the non-Abelian $SU(5)^5$ flavor group. This allowed us to treat the lepton, and up- and down-quark sectors in a similar way. We have first considered the three sectors uncoupled, and we have parametrized the respective potentials in terms of the three TAD invariants, which renders the classification of the structures of the symmetry breaking vacua intuitively simple. We have found that for a large part of the parameter space the ground state is characterized by one component for each multiplet of Yukawa fields with an $\mathcal{O}(1)$ vev, while the vevs of the other components vanish. This can be considered as a first interesting result of our analysis.

To each three-point and four-point scalar interaction we have then added a symmetry invariant logarithmic

correction, and we have shown that this has the effect of lifting the vanishing vevs to exponentially suppressed values. In this way a Yukawa hierarchy that is qualitatively similar for all three types of fermions arises quite naturally. Highlighting how the fermion mass hierarchies could be generated in a natural way by quantum corrections to the effective scalar potential of the Yukawa fields is probably the most interesting observation that stems from our analysis.

As a further step we have included renormalizable terms that couple the up- and down-quark sectors, and we have found that the corresponding contributions to the scalar potential are minimized when the heaviest (top and bottom) quarks are almost decoupled from the lighter ones. It is precisely the suppression of light-heavy mixings that gives rise to the family structure of the quarks, and this suppression naturally occurs in our approach.

We have also considered briefly what would happen for a generic number of generations, and we have concluded that for any number different from 3 a fully hierarchical pattern cannot result. This can be restated in a stronger way by saying that, within our scenario, the observed hierarchy of the Yukawa couplings implies precisely three generations. Finally, we have argued that accounting for the SSB of the two $U(1)_{d,e}$ Abelian factors does not modify the previous results.

Clearly, several issues related to this work deserve further study, and some are listed below.

Nambu-Goldstone bosons (NGB).—SSB implies the presence of NGB, for which strong constraints exist from unseen hadron decays, astrophysics, and from flavor violating processes. A standard solution is gauging the symmetry [12,13]. A very large scale suppressing the NGB

couplings to ordinary particles could also provide a way out.

Leptons.—We have only considered the SM lepton sector, which is known to be incomplete. It would be interesting to extend this scenario to the most popular models for massive neutrinos, and see if something could be said about neutrino masses and mixings.

Cabibbo Kobayashi Maskawa matrix.—We have shown how the two heaviest quarks naturally decouple from the lighter ones, hinting at the emergence of a family structure for the quarks. However, we have also concluded that the full structure of $\langle V \rangle$ remains undetermined because there is not enough information in the renormalizable scalar potential. Clearly this point deserves further study.

Effective potential.—What we consider by far the most important issue is verifying if the logarithmic terms that play such a crucial role in our construction are effectively generated by quantum corrections, and with reasonable values of the coefficients. We anticipate that, because of the several different types of interactions and of the large number of component fields, this appears to be a nontrivial task. However, once this program is carried out, if the required logarithms appear then we believe that a quite interesting candidate for a theory of the SM Yukawa sector will be found.

ACKNOWLEDGMENTS

It is a pleasure to thank Alberto Casas for encouraging comments and Jose Ramon Espinosa for pointing out the complex nature of \mathcal{D} . Rodrigo Alonso, Belen Gavela, Luca Merlo, and Stefano Rigolin are warmly acknowledged for several conversations that stimulated this work.

-
- [1] R. S. Chivukula and H. Georgi, *Phys. Lett. B* **188**, 99 (1987).
 - [2] A. S. Joshipura, K. M. Patel, and S. K. Vempati, *Phys. Lett. B* **690**, 289 (2010).
 - [3] R. Alonso, G. Isidori, L. Merlo, L. A. Munoz, and E. Nardi, *J. High Energy Phys.* **06** (2011) 037.
 - [4] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia, *Nucl. Phys.* **B645**, 155 (2002).
 - [5] C. D. Froggatt and H. B. Nielsen, *Nucl. Phys.* **B147**, 277 (1979).
 - [6] Y. Koide, *Phys. Rev. D* **78**, 093006 (2008).
 - [7] Y. Koide, *Phys. Rev. D* **79**, 033009 (2009).
 - [8] T. Feldmann, M. Jung, and T. Mannel, *Phys. Rev. D* **80**, 033003 (2009).
 - [9] R. Alonso, M. B. Gavela, L. Merlo, and S. Rigolin, *J. High Energy Phys.* **07** (2011) 012.
 - [10] Y. Nambu, "Spontaneous Symmetry Breaking and the Origin of Mass," in Proceedings of the International Conference on Fluid Mechanics and Theoretical Physics in Honor of Professor Pei-Yuan Chou's 90th Anniversary, Beijing, China, Report No. EFI-92-37.
 - [11] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, *Adv. Comput. Math.* **5**, 329 (1996).
 - [12] M. E. Albrecht, T. Feldmann, and T. Mannel, *J. High Energy Phys.* **10** (2010) 089; M. E. Albrecht, Ph.D. thesis, Technische Universität Munich, 2010, <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:vbv:91-diss-20100818-981978-1-7>.
 - [13] B. Grinstein, M. Redi, and G. Villadoro, *J. High Energy Phys.* **11** (2010) 067.