PHYSICAL REVIEW D **82,** 055010 (2010)

New approach to flavor symmetry and an extended naturalness principle

S. M. Barr

Department of Physics and Astronomy University of Delaware Newark, Delaware 19716 USA

(Received 20 July 2010; published 13 September 2010)

A class of nonsupersymmetric extensions of the standard model is proposed in which there is a multiplicity of light scalar doublets in a multiplet of a nonabelian family group with the standard model Higgs doublet. Anthropic tuning makes the latter light, and consequently the other scalar doublets remain light because of the family symmetry. The family symmetry greatly constrains the pattern of flavor-changing neutral-current interactions (FCNC) and p decay operators coming from scalar-exchange. Such models show that useful constraints on model-building can come from an extended naturalness principle when the electroweak scale is anthropically tuned.

DOI: 10.1103/PhysRevD.82.055010 PACS numbers: 12.10.Dm

I. INTRODUCTION

In models with low-energy supersymmetry (SUSY), the supersymmetry allows the existence at low energy of many scalar fields in a way consistent with the old "naturalness principle" [1], because of the nonrenormalization theorems. Supersymmetry also allows great leeway in the structure of the superpotential without violating the naturalness principle. The consequent flexibility in model building is both an advantage and a disadvantage. One disadvantage is that a huge number of terms are permitted whose coefficients are not constrained by any principle (unless somewhat *ad hoc* symmetries are imposed). Many of these terms violate baryon number, flavor or CP. Thus, the very flexibility of low-scale SUSY seems to undo several of the greatest successes of the standard model, which were due to the fact that the symmetries of the standard model greatly restrict the couplings that can be written down.

If low-energy supersymmetry is abandoned there is one large price to pay, namely, the fine-tuning of the mass-squared parameter (μ^2) of the standard model (SM) Higgs doublet. That is not necessarily a disaster, however, as such a tuning might be anthropically accounted for [2–4], and there is a corresponding gain, namely, more constrained model-building. Without low-energy supersymmetry, the possibilities for new particles at or near the electroweak-scale are more limited and their small masses must either be dynamically generated or protected by conventional (i.e. non-SUSY) symmetries, such as chiral symmetry and gauge invariance, which may place strong conditions on their interactions. (One sees this, of course, in technicolor models, which are so highly constrained that it is difficult to construct realistic models based on the idea.)

In this paper, we consider models without low-energy supersymmetry, and propose an extension of the old naturalness principle, which we call the "extended naturalness principle". According to this principle, there should be no parameters in a model that are very small (or otherwise take very special values) unless this has *either* a conventional "natural" explanation in terms of symmetry principles and dynamical mechanisms, *or* can be accounted for by "anthropic tuning." We will see, by considering a certain class of models, that this extended naturalness principle can greatly constrain model building (as the old naturalness principle did) and lead to interesting and testable scenarios.

The kind of model that will be considered in this paper is characterized by having a multiplicity of light scalar doublets (instead of just one, as in the standard model). One reason for considering such a possibility is that the presence of several (six or even five) light scalar doublets can lead to satisfactory unification of gauge couplings without supersymmetry [5,6]. But to make a multiplicity of scalar doublets light in a way consistent with the extended naturalness principle requires new symmetries. The basic idea assumed here is that these symmetries are family symmetries. The "extra" light scalar doublets will be assumed to be in a multiplet of a nonabelian group G_F with the SM Higgs doublet. That means that when the mass of the SM Higgs doublet is made small by anthropic tuning, so will the entire multiplet of scalar doublets.

The idea that the tiny (and negative) mass-squared of the SM Higgs may be the result of anthropic tuning in what is now called a "multiverse" was proposed in [2]. Some reasons to regard this possibility as plausible have been stated by S. Weinberg: "If the electroweak-scale is anthropically fixed, then we can give up the decades long search for a natural solution to the hierarchy problem. This is a very attractive prospect, because none of the "natural" solutions that have been proposed, such as technicolor or low-energy supersymmetry, were ever free of difficulties. In particular, giving up low-energy supersymmetry can restore some of the most attractive features of the non-supersymmetric standard model: automatic conservation of baryon and lepton number in interactions up to dimension 5 and 4 respectively; natural conservation of flavors in neutral currents; and a small neutron electric dipole moment" [4].

If the SM Higgs doublet is in a nontrivial multiplet of G_F and gives mass to the known quarks and leptons through renormalizable $d = 4$ Yukawa operators, then the known quarks and leptons would also have to be in nontrivial G_F multiplets. In that case, G_F would be a "family symmetry". This is an attractive possibility as there would be a single explanation for both the multiplicity of the light quarks and leptons and the multiplicity of the light scalar doublets, namely, that they were multiplets of a family symmetry. (Moreover, having "families" of both the fermions and the scalars seems less unbalanced than the standard model, where there are large numbers of quarks and leptons, but only one scalar multiplet.) The most obvious possibilities for G_F, given the fact that there are three light families of quarks and leptons, are $SU(3)$ and $SO(3)$, with the light families being in a 3 in either case. If $G_F = SU(3)$, there could be six light scalar doublets either in $\mathbf{3} + \mathbf{3}$ or in 6. If $G_F = SO(3)$, there could be six light scalar doublets in $\mathbf{3} + \mathbf{3}$ or five of them in 5. We shall consider these possibilities briefly later, and shall discuss a toy model based on $SO(3)_F$ in Section III, but we shall find it easier to construct a realistic model based on $SO(4)_F$, which will be discussed in Section IV. (There is a long history of models with nonabelian continuous family symmetries. See, for example, the papers in [7] and references cited therein. Recently, nonabelian discrete family groups have been used in models with multiple scalar doublets, as in [8].)

The extra light scalar doublets must have masses of at least several TeV to avoid excessive flavor violation in neutral current processes. Thus, there must be splitting of the G_F multiplet of scalar doublets by an amount at least an order of magnitude greater than the mass of the SM Higgs doublet. The idea is that the *overall* mass-squared of the G_F multiplet of scalar doublets (i.e. the G_F-invariant part) "scans" among the "domains" or "subuniverses" of the universe and is anthropically set to have a value (in our subuniverse or domain) such that the *lightest* member of the G_F multiplet (i.e. the SM Higgs doublet) has a mass-squared that is negative and of order -(100 GeV)^2. That means that all the other members of the G_F multiplet will have masses-squared that are of order M_F^2, where $M_F^2 \gg 1$ TeV is the scale of splitting of the multiplet. (Of course, it must be only the lightest scalar doublet whose mass-squared is pushed negative, since if more than one is, the one with most negative mass-squared will get a vacuum expectation value (VEV) of order M_F, which is not anthropically viable [2].) The breaking of G_F can be dynamical, and thus M_F can naturally be much smaller than the unification scale, and, in particular, in the TeV range. Figure 1 illustrates the idea schematically.

An obvious issue if the unification scale is around 10^{14} GeV (as in the non-SUSY SM with six scalar doublets) is rapid proton decay. In [6] this was avoided by assuming the unified group to be the "trinification" group $SU(3)_c \times SU(3)_L \times SU(3)_R \times S_3$, where the S_3 cyclically permutes the three $SU(3)$ factors. We shall use the same group. (For an outline of the trinification scheme assumed in this paper, see Appendix A.) This eliminates proton decay mediated by gauge boson exchange. However, since some of the extra light scalar doublets in our models shall necessarily couple to light quarks and leptons with couplings of order 1 (as will be seen), proton decay by the exchange of their superheavy colored partners becomes an issue. It is shown in Appendix B that there exist terms in the scalar potential that can give the dangerous colored scalars masses large enough to avoid excessive proton decay. The proton decay that does exist, however, will have characteristic branching ratios determined by the symmetry G_F, as will be discussed later.

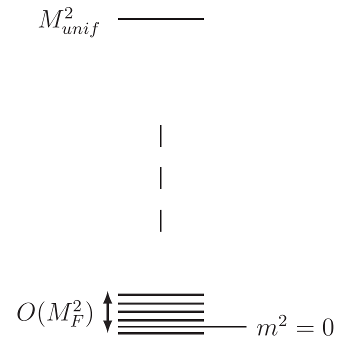

FIG. 1. The overall (G_F-invariant) mass-squared of the G_F multiplet of scalar doublets is tuned so the lightest ($H_{\rm SM}$) has small negative mass-squared. The others then have positive mass-squared from G_F breaking.

Another possibility besides trinification is unification based on a simple group, such as $SU(5)$ or $SO(10)$, broken by orbifold compactification [9]. If the light quarks and leptons live on a brane where there is only the SM gauge group, proton decay can be suppressed, since the gauge bosons of $G_{\rm GUT}/G_{\rm SM}$ will vanish on the SM brane by the orbifold boundary conditions, and similarly for the colored scalars that would mediate proton decay. On the other hand, gauge kinetic terms in the 4D Lagrangian on the SM brane would not respect $G_{\rm GUT}$ and would affect the gauge unification. If these effects were small for some reason, then unification could still take place with 5 or 6 light scalar doublets. We shall not pursue this possibility here, but will henceforth assume the unified group to be $SU(3)_c \times SU(3)_L \times SU(3)_R \times S_3$, which will be denoted G_U for short. The irreducible multiplet that contains a family, namely $(3, \bar{3}, 1) + (\bar{3}, 1, 3) + (1, 3, \bar{3})$ will be denoted henceforth as F.

In the later sections it will be seen that the interrelated set of assumptions we have made (no low-energy SUSY, unification of gauge couplings by a multiplicity of light scalar doublets whose mass is related to that of the SM Higgs by a family symmetry) leads to tightly constrained possibilities for model building and interesting and novel phenomenological consequences. While anthropic tuning in general, and of the electroweak scale in particular, cannot be tested directly, this shows that when it is combined with the extended naturalness principle and other attractive assumptions, predictive schemes can result. The rest of the paper will be organized as follows. In Section II, some basic points about the anthropic tuning of the electroweak scale will be discussed. In Section III, a simple $SO(3)_F$ toy model will be presented and its inadequacies pointed out. In Section IV, a realistic model based on $SO(4)_F$ will be described and analyzed. Some details having to do with the trinification group and its breaking will be discussed in the appendices.

II. ANTHROPIC TUNING OF THE ELECTROWEAK SCALE

The idea that the electroweak scale might be anthropically determined in the context of what is now called the multiverse scenario was proposed in [2]. There it was noted that the cosmological constant $\Lambda(\approx 10^{-120}M_{P\ell}^4)$ and the Higgs mass parameter of the standard model $\mu^2(\approx 10^{-34}M_{P\ell}^2)$ are the two smallest parameters in our current theory (in natural units) as well as the ones whose smallness has proven most difficult to explain in conventional ways. (The next smallest parameter $\bar{\theta}$ is only bounded to be less than about 10^{-10}, and several viable natural explanations exist for its smallness [10]. Many technically natural symmetry explanations have also been proposed for the small quark and lepton Yukawa couplings of the lighter families. And the explanation of the smallness of $\Lambda_{\rm QCD}/M_{P\ell}$ in terms of the logarithmic running of α_s and dimensional transmutation is perfectly natural.) It was therefore suggested in [2] that Λ and μ^2 are the most plausible candidates for anthropic tuning. This was also suggested by S. Weinberg [4]: "The most optimistic hypothesis is that the only constants that scan are the few whose dimensionality is a positive power of mass: the vacuum energy and whatever mass or masses set the scale of electroweak symmetry breaking."

The reason that the magnitude of μ^2 must be at or lower than $(100\ \text{GeV})^2$ for the evolution of life to be at all plausible is that larger (negative) values of μ^2 lead to larger v ($\equiv \langle H_{\rm SM}\rangle$), and therefore larger quark masses. This leads to larger pion masses and shorter range of the nucleon-nucleon potential. This makes nuclei more unstable, beginning with the all-important deuteron, which becomes unbound if v is larger than its observed value by a factor of about 1.4 to 2.7 [2]. The case of positive μ^2 require separate arguments, given in [2] (for $\Lambda = 0$) and [3] (for $\Lambda \neq 0$).

An extension of the analysis of [2] to the case of two doublets H_u and H_d, which transform under the standard model gauge group as $(1, 2 + \frac{1}{2})$ and $(1, 2, -\frac{1}{2})$ respectively (and which couple, respectively, to the up quarks and to the down quarks and charged leptons) was give in [3]. This is relevant to the present work, since such a pair of doublets exists in many unified schemes, including trinification. With two doublets, there is a 2×2 mass-squared matrix that can be written as

$$(H_u, H_d^*)\begin{pmatrix} M_u^2 & \Delta^2 \\ \Delta^{2*} & M_d^2 \end{pmatrix}\begin{pmatrix} H_u^* \\ H_d \end{pmatrix}. \tag{1}$$

Only one fine-tuning is required to make the SM doublet light, namely, of the determinant of the mass-squared matrix, with all the elements of the matrix remaining of the unification scale. The light doublet would be

$$\begin{aligned} &H_{\rm SM} = \cos\theta_H H_u + e^{i\alpha}\sin\theta_H H_d^*, \\ &\tan 2\theta_H \equiv 2|\Delta^2|/(M_u^2 - M_d^2) = O(1), \qquad \alpha \equiv \arg\Delta^2. \end{aligned} \tag{2}$$

The other mass eigenstate, which remains superheavy, is $H_h = -e^{-i\alpha}\sin\theta_H H_u + \cos\theta_H H_d^*$. In typical unified models, where H_u couples to up quarks and H_d couples to down quarks and charged leptons, this will be reflected in the Yukawa couplings of the SM Higgs field at low energy, which will be proportional to $\cos\theta_H$ for the up quarks and proportional to $\sin\theta_H$ for the down quarks and charged leptons. If at least two independent combinations of the parameters M_u^2, M_d^2, and Δ^2 "scan" among the domains, then both μ^2 (the mass-squared parameter of the SM Higgs) and $\tan\theta_H$ scan. As shown in [3], not only do anthropic considerations set μ^2 to be near -(100 GeV)^2, but they also set $\tan\theta_H$ such that the d-quark mass is comparable to, but slightly larger than, the u-quark mass, as indeed observed. (See also [11].)

In our case, the scalar doublets are in a representation R of the family group G_F. Denoting them H_u^A and H_d^A, where $A = 1, \ldots, R$, one has

$$(H_u^A, H_{dA}^*)\begin{pmatrix} M_u^2\delta_A^B & (\Delta^2)_{AB} \\ (\Delta^{2*})^{AB} & M_d^2\delta_B^A \end{pmatrix}\begin{pmatrix} H_{uB}^* \\ H_d^B \end{pmatrix}. \tag{3}$$

Note that if $G_F = SU(N)$ and R is an N or other complex representation, then $(\Delta^2)_{AB}$ breaks G_F and must be near the electroweak scale. In that case, a single tuning to make the determinant of the matrix small is not sufficient. Such a tuning would make *either* M_u^2 *or* M_d^2 to be of order the weak scale, which would result in the light Higgs doublets being almost purely of the type H_u or of the type H_d. That would not give electroweak-scale masses to all the quarks and leptons. In order for all the light quarks and leptons to obtain realistic masses, *both* M_u^2 *and* M_d^2 would separately have to be tuned to be of order the weak scale.

On the other hand, if $G_F = SO(N)$ and R is an N or other real representation, then the matrix can have the form

$$(H_u^A, H_d^{*A})\begin{pmatrix} M_u^2 & \Delta^2 \\ \Delta^{2*} & M_d^2 \end{pmatrix}\delta^{AB}\cdot\begin{pmatrix} H_u^{*B} \\ H_d^B \end{pmatrix}. \tag{4}$$

In this case, all the elements of the mass matrix can be of order the unification scale and a *single* tuning of the determinant is enough to make a multiplet of scalar doublets light that is a mixture of both H_u and H_d type. Specifically, what will be made light from a single tuning is an R-multiplet of scalar doublets $H^A \equiv (\cos\theta_H H_u^A + e^{i\alpha}\sin\theta_H H_d^{*A})$, where $A = 1, \ldots, R$.

Not only do models with $G_F = SO(N)$ require fewer fine-tunings, in this sense, but the problem of canceling G_F anomalies does not arise. We shall therefore consider only models with orthogonal family groups.

III. AN $SO(3)_F$ TOY MODEL

Several key characteristics of the kind of model being proposed in this paper can be understood in a simple toy model with $G_F = SO(3)$. (A realistic model with $G_F = SO(4)$ will be discussed in Section IV.) The gauge group of the model is

$$\begin{aligned} G &= G_U \times G_F \times G_{\rm DSB} \\ G_U &= SU(3)_c \times SU(3)_L \times SU(3)_R \times S_3 \qquad (5) \\ G_F &= SO(3)_F \qquad G_{\rm DSB} = SU(N)_{\rm DSB}. \end{aligned}$$

The confining group $SU(N)_{\rm DSB}$ plays the role of dynamically breaking the family group $SO(3)_F$. The three pieces of the G_F multiplet F, namely $(3, \bar{3}, 1)$, $(\bar{3}, 1, 3)$, and $(1, 3, \bar{3})$, will be denoted by the subscripts q, $\bar{q}$, and ℓ, respectively. The three families are in an $(F, 3, 1)$ of $G_U \times G_F \times G_{\rm DSB}$ that will be denoted $\psi^i = \psi_q^i + \psi_{\bar{q}}^i + \psi_\ell^i$, where i is an $SO(3)_F$ vector index. The light scalar doublets are in an $(F, 5, 1)$ of $G_U \times G_F \times G_{\rm DSB}$, which will be denoted $\Phi^{(ij)} = \Phi_q^{(ij)} + \Phi_{\bar{q}}^{(ij)} + \Phi_\ell^{(ij)}$. Note that $\Phi_\ell^{(ij)}$ is a rank-2 symmetric traceless tensor of $SO(3)_F$ and therefore contains 5 doublets $H_u^{(ij)}$ and 5 doublets $H_d^{(ij)}$. The Yukawa terms of the quarks and leptons are therefore just

$$\begin{aligned} \mathcal{L}_{\rm Yukawa} = y[(\psi_q^i \psi_{\bar{q}}^j)\Phi_\ell^{(ij)} + \text{cyclic}] \\ + y'[(\psi_\ell^i \psi_\ell^j)\Phi_\ell^{(ij)} + \text{cyclic}]. \qquad (6) \end{aligned}$$

The expression "+ *cyclic*" refers to the S_3 permutations of the three $SU(3)$ trinification groups that take $q \to \bar{q} \to \ell \to q$.

The $\Phi_\ell^{(ij)}$ has a nonzero VEV, but $\Phi_q^{(ij)}$ and $\Phi_{\bar{q}}^{(ij)}$ clearly do not, as color is unbroken, so only the terms in the cyclic permutation that are explicitly written out in Eq. (6) contribute to quark and lepton masses. The $\Phi_\ell^{(ij)}$ contains, among other components, $H_u^{(ij)}$ and $H_d^{(ij)}$. (See Appendix A.) These have G_F-invariant mass-squared terms of the form

$$\begin{aligned} M_u^2 H_u^{(ij)*} H_u^{(ij)} + M_d^2 H_d^{(ij)*} H_d^{(ij)} + \Delta^2 H_u^{(ij)} H_d^{(ij)} \\ + \Delta^{2*} H_u^{(ij)*} H_d^{(ij)*}. \qquad (7) \end{aligned}$$

As already explained in Section II, anthropic fine-tuning will make light a 5 of light doublets, $H^{(ij)} = \cos\theta_H H_u^{(ij)} + e^{i\alpha}\sin\theta_H H_d^{(ij)*}$. The terms in Eq. (7) are G_F-invariant, and so leave these 5 light doublets degenerate. (Their degeneracy will be lifted when G_F is dynamically broken.) As shown in Appendix A, the terms in Eq. (7) come from the terms

$$\begin{aligned} M_\Phi^2 \Phi_\ell^{(ij)*}\Phi_\ell^{(ij)} + [M_\Delta \Phi_\ell^{(ij)}\Phi_\ell^{(ij)}\Phi_\ell + \text{H.c.}] \\ + \sigma\,\text{Tr}(\Phi_\ell^{(ij)*}\Phi_\ell^{(ij)})\,\text{Tr}(\Phi_\ell^*\Phi_\ell) + \rho\,\text{Tr}(\Phi_\ell^{(ij)*}\Phi_\ell^{(ij)}\Phi_\ell^*\Phi_\ell), \\ (8) \end{aligned}$$

where $\Phi = \Phi_q + \Phi_{\bar{q}} + \Phi_\ell$ is a singlet under $SO(3)_F$ and gets superlarge VEVs in the SM-singlet components of Φ_ℓ. (The traces in Eq. (8) refer to G_U indices.) Of course, along with the terms in Eq. (8) come those that result from the S_3 permutations.

The sector that dynamically breaks $SO(3)_F$ consists of fermions in the fundamental and antifundamental representations of a confining gauge group $SU(N)_{\rm DSB}$: $\chi^{\mu i} = (1, 3, N)$ and $\bar{\chi}_{\mu I} = (1, 1, \bar{N})$, where I is just a label runs from 1 to 3. These form condensates, which without loss of generality can be written $\langle \chi^{\mu i}\bar{\chi}_{\mu I}\rangle = (f_N)^3 \delta_I^i$ and which break $SO(3)_F$ completely. There is also a set of several real $SO(3)_F$ triplet scalars distinguished by the label J: $\eta_J^i = (1, 3, 1)$. These are "messenger fields" that communicate the breaking of $SO(3)_F$ to the standard model fields. These have an explicit positive mass-squared term, with mass that is naturally superheavy, and also a Yukawa coupling to the $\chi^{\mu i}$, $\bar{\chi}_{\mu I}$:

$$\mathcal{L}_{\rm DSB} = \frac{1}{2}(M_\eta^2)_{JK}\eta_J^i \eta_K^i + [y_{IJ}(\chi^{\mu i}\bar{\chi}_{\mu I})\eta_J^i + \text{H.c.}]. \quad (9)$$

Note that we adopt the summation convention that repeated indices are summed over, for indices of all types, including labels like I, J, and K above. When the $\chi^{\mu i}$ and $\bar{\chi}_{\mu I}$ form a condensate, it gives a linear term for the η_J^i that induces a VEV

$$\langle \eta_J^i \rangle = (f_N)^3 (M_\eta^2)_{JK}^{-1} y_{iK}. \qquad (10)$$

M_η^2 is naturally superlarge, while the scale f_N is dynamically generated by the $SU(N)_{\rm DSB}$ interactions and can naturally be of any magnitude. If f_N is at an intermediate scale, then the VEV of the η_J^i can naturally be near the weak scale. For example, with $M_\eta^2 \sim (10^{15}\ \text{GeV})^2$ and $f_N \sim 10^{11}$ GeV, one has $\langle \eta_J^i \rangle \sim 1$ TeV.

Since f_N is the scale of the breaking of the local $SO(3)_F$ family symmetry, the family gauge bosons have mass of an intermediate scale and produce negligible flavor-changing neutral current (FCNC) interactions. The gauge symmetries of the model do not allow any direct renormalizable couplings of the fields $\chi^{\mu i}$ and $\bar{\chi}_{\mu I}$ to the light fields (i.e. to the quarks and leptons and light scalar doublets). The breaking of $SO(3)_F$ is communicated to the light fields by the η_J^i. (Note that the η_J^i are superheavy, even though their VEVs are near the weak scale. The group $SO(3)_F$, again, is broken at high scales by the dynamical condensate.)

In particular, it is easily seen that there is only one renormalizable term that gives $SO(3)_F$-breaking splittings of the multiplet of light scalar doublets, namely

$$\mathcal{L}'_\Phi = \lambda_{KI}(\Phi_\ell^{(ij)*}\Phi_\ell^{(jk)}\eta_K^k\eta_I^i). \tag{11}$$

If we define the Hermitian matrix m^2, which we shall call the "master matrix", by $(m^2)^{ki} \equiv \eta_K^k\lambda_{KI}\eta_I^i$ (remembering the summation convention), then the terms in Eq. (11) give $SO(3)_F$-breaking masses to the 5 light scalar doublets $H^{(ij)}$ of the form

$$H^{(ij)*}H^{(jk)}(m^2)^{ki} = \mathrm{Tr}[H^*Hm^2]. \tag{12}$$

As explained in Section II, anthropic tuning will set the mass-squared of the lightest of the five scalar doublets to be negative and of order -(100 GeV)^2. The other four scalar doublets will then have mass-squared of order $m^2 \sim \langle\eta_J^i\rangle^2$. Since these other four scalar doublets will mediate FCNC processes, their masses must be at least several TeV. On the other hand, they should not be much larger than this, as else they will not give unification of gauge couplings.

Which linear combination of the five $H^{(ij)}$ is the lightest (i.e. which linear combination is the SM Higgs doublet $H_{\rm SM}$) directly determines the "textures" of the Yukawa couplings of $H_{\rm SM}$ to the light quarks and leptons. For example, if $H_{\rm SM}$ were purely $H^{(23)} \subset \Phi_\ell^{(23)}$, then by Eqs. (2) and (6) one sees that there would be terms $y\ \cos\theta_H(\bar{u}_2u_3 + \bar{u}_3u_2)H_{\rm SM} + ye^{i\alpha}\ \sin\theta_H(\bar{d}_2d_3 + \bar{d}_3d_2)H_{\rm SM} + y'e^{i\alpha}\ \sin\theta_H(\ell_2^+\ell_3^- + \ell_3^+\ell_2^-)H_{\rm SM}$, so that the textures would all be of the same form, having only nonvanishing 23 and 32 elements. In general, however, the master matrix $(m^2)^{ki} = \eta_K^k\lambda_{KI}\eta_I^i$ is a nontrivial 3×3 Hermitian matrix. Therefore, one expects $H_{\rm SM}$ to be a linear combination of all 5 of the $H^{(ij)}$ and the quark and lepton mass textures to have all their elements nonzero. In particular, if $H_{\rm SM} = \sum_{ij}a_{ij}H^{(ij)}$, then the textures will simply be proportional to a_{ij}.

One of the interesting features of this kind of model, therefore, is that there is a direct connection between the spectrum of the light scalar doublets and the pattern of Yukawa couplings of the standard model Higgs field. The master matrix determines the pattern of masses and mixings of the scalar doublets, which in turn determines which linear combination is the standard models Higgs field, which then in turn determines the quark and lepton textures. A hierarchy among the splittings of the scalar doublets leads to a hierarchy among the quark and lepton masses. Suppose, for example, that the master matrix m^2 has the hierarchical form

$$m^2 = m_0^2\begin{pmatrix} 1+\delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{12} & \delta_{22} & \delta_{23} \\ \delta_{13} & \delta_{23} & \delta_{33} \end{pmatrix}, \qquad \delta_{ij} \ll 1, \qquad m_0^2 > 0. \tag{13}$$

Then one can see that the term $[H^{(ij)*}H^{(jk)}(m^2)^{ki}]$ will give mass-squared contributions to those $H^{(ij)}$ which have one or more indices equal to 1 (namely $H^{(12)}$, $H^{(13)}$, and $H'^{(11)} \equiv (2H^{(11)} - H^{(22)} - H^{(33)})/\sqrt{6})$) that are of order m_0^2, while it gives mass-squared contributions to the others ($H^{(23)}$ and $H'^{(22)} \equiv (H^{(22)} - H^{(33)})/\sqrt{2}$) that are only $O(\delta_{ij})m_0^2$. The lightest scalar doublet is the standard model Higgs $H_{\rm SM}$, whose mass-squared is pushed slightly negative by anthropic fine-tuning. We will call the next lightest the "lightest extra scalar doublet (LESD)" $H_{\rm LESD}$. These two lightest doublets, $H_{\rm SM}$ and $H_{\rm LESD}$, are approximately linear combinations of $H'^{(22)}$ and $H^{(23)}$ and are therefore only split from each other by $O(\delta_{ij})m_0^2$. The other three scalar doublets are split from these two by $O(m_0^2)$. The pattern is schematically shown in Fig. 2.

The splittings shown in Fig. 2 come from diagonalizing the explicit form of the mass matrix of the scalar doublets, which (from Eq. (13)) is to leading order in δ_{ij} given by

$$[H'^{(22)}H^{(23)}H'^{(11)}H^{(12)}H^{(13)}] \times \begin{bmatrix} \frac{\delta_{22}+\delta_{33}}{2} & \frac{\delta_{23}^*-\delta_{23}}{2} & \frac{\delta_{33}-\delta_{22}}{2\sqrt{3}} & \frac{\delta_{12}}{2} & -\frac{\delta_{13}}{2\sqrt{2}} \\ \frac{\delta_{23}-\delta_{23}^*}{2} & \frac{\delta_{22}+\delta_{33}}{2} & \frac{\delta_{23}^*-\delta_{23}}{2\sqrt{3}} & \frac{\delta_{13}}{2\sqrt{2}} & \frac{\delta_{12}}{2} \\ \frac{\delta_{33}-\delta_{22}}{2\sqrt{3}} & \frac{\delta_{23}-\delta_{23}^*}{2\sqrt{3}} & 2/3 & \frac{2\delta_{12}^*-\delta_{12}}{2\sqrt{3}} & -\frac{\delta_{13}}{2\sqrt{6}} \\ \frac{\delta_{12}^*}{2} & \frac{\delta_{13}^*}{2\sqrt{2}} & \frac{2\delta_{12}-\delta_{12}^*}{2\sqrt{3}} & 1/2 & \frac{\delta_{23}^*}{2} \\ -\frac{\delta_{13}^*}{2\sqrt{2}} & \frac{\delta_{12}^*}{2} & -\frac{\delta_{13}^*}{2\sqrt{6}} & \frac{\delta_{23}}{2} & 1/2 \end{bmatrix} \times m_0^2\begin{bmatrix} H'^{(22)} \\ H^{(23)} \\ H'^{(11)} \\ H^{(12)} \\ H^{(13)} \end{bmatrix}, \tag{14}$$

Therefore, the lightest scalar doublet $H_{\rm SM}$ is predominantly a linear combination of $H'^{(22)} = (H^{(22)} - H^{(33)})/\sqrt{2}$ and $H^{(23)}$, with $O(\delta_{ij})$ admixtures of the others. (Note that this depends on the sign of the largest elements in m^2, which were chosen in Eq. (13) to be positive.) Specifically, one finds $H_{\rm SM}$ to be of the form

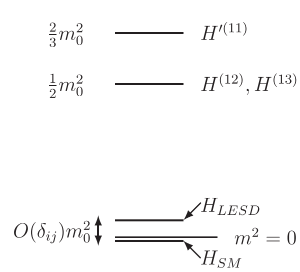

FIG. 2. A schematic plot of the mass-squared spectrum of the 5 of scalar doublets in the toy $SO(3)_F$ model.

$$H_{SM} \cong \cos\gamma((H^{(22)} - H^{(33)})/\sqrt{2} + \sin\gamma H^{(23)} + O(\delta_{12}, \delta_{13})H^{(12)} + O(\delta_{12}, \delta_{13})H^{(13)} + O((\delta_{22} - \delta_{33}), \mathrm{Im}(\delta_{23}))(2H^{(11)} - H^{(22)} - H^{(33)})/\sqrt{6}). \tag{15}$$

That implies that the quark and lepton textures have the form

$$M_{q,\ell} \sim \begin{pmatrix} O((\delta_{22}-\delta_{33}), Im(\delta_{23})) & O(\delta_{12}, \delta_{13}) & O(\delta_{12}, \delta_{13}) \\ O(\delta_{12}, \delta_{13}) & \cos\gamma & \sin\gamma \\ O(\delta_{12}, \delta_{13}) & \sin\gamma & -\cos\gamma \end{pmatrix} \langle H_{SM} \rangle. \tag{16}$$

Note that the largest (smallest) elements of these textures correspond to the lightest (heaviest) scalar doublets. Moreover, the ratios of splittings within the 5 of scalar doublets is closely related to the ratios of quark and lepton masses. The splitting between the two lightest scalar doublets turns out to be (from Eq. (14)) $O(\mathrm{Im}(\delta_{23}), (\delta_{22} - \delta_{33})^2)m_0^2$, while the splitting that separates these two doublets from the three heavier doublets is $O(1)m_0^2$. See Fig. 2. Compare this to the ratio of masses of the first family of fermions to the masses of the heavier families, which is $O((\delta_{22} - \delta_{33}), \mathrm{Im}(\delta_{23}))$, as can be seen from Eq. (16).

One sees from Eq. (16) that this $SO(3)_F$ model is not realistic, because the quark and lepton mass matrices have to be traceless. (The 5 of $SO(3)$ is a traceless tensor.) One cannot therefore have a threefold fermion mass hierarchy: if one of the families is made very light, the tracelessness forces the other two families to have nearly equal and opposite masses to each other.

Another unrealistic feature of this toy model is that the up quark, down quark, and charged lepton mass matrices (which shall be denoted M_U, M_D, M_L) are all proportional to the same matrix, namely $\langle H^{(ij)} \rangle$. Consequently there is no Cabibbo-Kobayashi-Maskawa quark-mixing matrix (CKM) mixing. The distinction between the different types of fermions (up quarks, down quarks, and charged leptons) comes from the breaking of the unified group $SU(3)_c \times SU(3)_L \times SU(3)_R \times S_3$. The superlarge VEVs that do this breaking must be invariant under $SO(3)_F$ (otherwise $SO(3)_F$ would be broken at superlarge scales). Therefore, the breaking of $SO(3)_F$, which generates the nontrivial quark and lepton textures, does not depend on these G_U-breaking VEVs and the textures do not "know" that the unified group is broken, and hence the textures have the same form for the different types of fermions.

A third difficulty of the $SO(3)_F$ model is that there are only five light scalar doublets, and as Fig. 2 shows, three of them are orders of magnitude heavier than 1 TeV. (If $O(\delta_{ij})m_0^2 > 1$ TeV, as must be the case if H_{LESD} does not mediate excessive flavor-changing neutral-current interactions (FCNC) processes, then the other extra light scalar doublets, whose masses are $O(m_0^2)$, must be several orders of magnitude heavier than a TeV.) This does not give a good unification of gauge couplings.

Some of these difficulties can be overcome in the context of $SO(3)_F$ by making the model more complicated. However, we shall find in Section IV that they can be overcome in a very simple way by going to $SO(4)_F$. There will then be four light families; one of these, however, can be made heavy by "mating" and getting mass with an $SO(4)_F$-singlet mirror family. The full mass matrices of the fermions would then contain a 4×4 block for the four families in the 4 of $SO(4)_F$. There can be a threefold hierarchy among the eigenvalues of such a matrix. Tracelessness will then cause the two largest eigenvalues to be nearly equal and opposite. That will not matter, however, because one of these large eigenvalues can be of the family that mates with the mirror family. Nevertheless, one will find that the near degeneracy of these two (largest) eigenvalues is related to a near degeneracy of the two lightest scalar doublets—just as in the $SO(3)_F$ toy model (Fig. 2). This is a general feature of this kind of model, and it is interesting phenomenologically because it means that one of the extra scalar doublets (the LESD) will dominate over all the others in low-energy phenomenology. That makes these models much more predictive than they would otherwise be.

Since the LSED is split from the SM Higgs doublet by an amount much smaller than the other splittings within the G_F multiplet, most of the extra scalar doublets have to be several orders of magnitude heavier than a TeV. In an $SO(4)_F$ model this can compensate for the fact that there are nine scalar doublets, rather than five or six, and give a good unification of gauge couplings.

IV. A REALISTIC $SO(4)_F$ MODEL

The $SO(4)_F$ model is quite similar to the $SO(3)_F$ toy model except that in addition to the families in a vector of the family group, there is a mirror family that is a singlet of the family group. The quarks and leptons are therefore in two multiplets of $G_U \times SO(4)_F \times SU(N)_{DSB}$, namely $(F, 4, 1) = \psi^i = \psi^i_q + \psi^i_{\bar{q}} + \psi^i_\ell$, where $i = 1, \ldots, 4$, and $(\bar{F}, 1, 1) = \bar{\psi} = \bar{\psi}_q + \bar{\psi}_{\bar{q}} + \bar{\psi}_\ell$. The $SO(4)_F$ -singlet mirror family will mate with one of the families leaving three families light. The quark and lepton Yukawa terms are (Equation (6))

$$\mathcal{L}_{\rm Yuk} = y[\psi^i_q \psi^j_{\bar{q}} \Phi^{(ij)}_\ell + \text{cyclic}] + y'[\psi^i_\ell \psi^j_\ell \Phi^{(ij)}_\ell + \text{cyclic}] + y_J[\psi^i_q \bar{\psi}_q \eta^i_J + \text{cyclic}], \tag{17}$$

where the η^i_J are now in (1, 4, 1) of the full gauge group and couple as in Eq. (9) to the $\chi^{\mu i}$, $\bar{\chi}_{\mu I}$, which are now in $(1, 4, N)$ and *four* $(1, 1, \bar{N})$.

The quark and lepton textures consequently have the form

$$(f_1 f_2 f_3 f_4 \bar{f}^c) \begin{pmatrix} a_f^{11} & a_f^{12} & a_f^{13} & a_f^{14} & b_f^1 \\ a_f^{12} & a_f^{22} & a_f^{23} & a_f^{24} & b_f^2 \\ a_f^{13} & a_f^{23} & a_f^{33} & a_f^{34} & b_f^3 \\ a_f^{14} & a_f^{24} & a_f^{34} & a_f^{44} & b_f^4 \\ b_{f^c}^1 & b_{f^c}^2 & b_{f^c}^3 & b_{f^c}^4 & 0 \end{pmatrix} \begin{pmatrix} f_1^c \\ f_2^c \\ f_3^c \\ f_4^c \\ \bar{f} \end{pmatrix}, \quad (18)$$

where f stands for u, d, or ℓ^-, and f^c stands for u^c, d^c, and ℓ^+. The 4×4 block is symmetric and traceless and is given by (see Eqs. (2) and (17))

$$a_u^{ij} = \cos\theta_H y \langle H^{(ij)} \rangle, \qquad a_d^{ij} = e^{i\alpha} \sin\theta_H y \langle H^{(ij)} \rangle,$$
$$a_\ell^{ij} = e^{i\alpha} \sin\theta_H y' \langle H^{(ij)} \rangle \Rightarrow a_u^{ij} \propto a_d^{ij} \propto a_\ell^{ij}. \quad (19)$$

From the third term in Eq. (17) it appears that the entries $b_f^i = b_{f^c}^i = \Sigma_J y_J \langle \eta_J^i \rangle$ and that these are the same for $f = u, d, \ell$. However, as discussed in Appendix C, there can be $d > 4$ effective operators (involving the superlarge VEVs that break the unification group G_U) which have the effect at low-energy of making the Yukawa couplings y_J in Eq. (17) different for different fermion types. Thus the third term in Eq. (17) should really be written $y_J^Q Q^i \bar{Q} \eta_J^i + y_J^{u^c} (u^c)^i \bar{u}^c \eta_J^i + y_J^{d^c} (d^c)^i \bar{d}^c \eta_J^i + y_J^L L^i \bar{L} \eta_J^i + y_J^{\ell^c} \bar{\ell}^c \eta_J^i$. So that one has $b_f^i = \Sigma_J y_J^f \langle \eta_J^i \rangle$ and $b_{f^c}^i = \Sigma_J y_J^{f^c} \langle \eta_J^i \rangle$, where $f = u, d, \ell^-$, and $f^c = u^c, d^c, \ell^+$. It turns out, as explained in Appendix C, that in simple situations $y_J^Q = y_J^{u^c} = y_J^{\ell^c}$, but $y_J^{d^c}$ and y_J^L can be different. (This ultimately stems from the fact that in the representation F of G_U there are extra superheavy fermions in each family with the same SM charges as d^c and L and their conjugates. These extra fermions are called D^c and L' in the Appendices.) So

$$b_u^i = b_d^i = b_{u^c}^i = b_{\ell^c}^i \neq b_\ell^i \neq b_{d^c}^i. \quad (20)$$

It follows that the full 5×5 matrices of the different fermion types (u, d, ℓ) are no longer simply proportional to each other, and so CKM mixing can occur. Moreover, the tracelessness of the $a^{(ij)}$ is no longer a problem. For example, suppose that there is a hierarchy in $a^{(ij)}$ such that its first and second rows and columns are very small (with the first much smaller than the second), and suppose that $b_u^i = b_{u^c}^i$ points in the $i = 4$ direction, then the mass matrix of the up quarks has the form

$$M_U = \begin{pmatrix} a^{11} & a^{12} & a^{13} & a^{14} & 0 \\ a^{12} & a^{22} & a^{23} & a^{24} & 0 \\ a^{13} & a^{23} & a^{33} & a^{34} & 0 \\ a^{14} & a^{24} & a^{34} & -a'^{33} & B \\ 0 & 0 & 0 & B & 0 \end{pmatrix}, \quad (21)$$

where $a'^{33} = a^{33} + a^{11} + a^{22} \cong a^{33}$. Since $B \sim \langle \eta \rangle \gg$ TeV, while the a^{ij} are proportional to the electroweak breaking VEV ~100 GeV, what happens with the form in Eq. (21) is that the three observed light families of up quarks, u, c, and t, are approximately those with $i = 1, 2, 3$, while the $i = 4$ up quark gets a mass much greater than a TeV with the mirror family up quark. The 3×3 mass matrix of the observed up quarks is then approximately

$$\tilde{M}_U \cong \begin{pmatrix} a^{11} & a^{12} & a^{13} \\ a^{12} & a^{22} & a^{23} \\ a^{13} & a^{23} & a^{33} \end{pmatrix}, \quad (22)$$

which is unconstrained by the tracelessness condition of the 4×4 matrix a^{ij} and can have a realistic hierarchy.

As in the $SO(3)_F$ toy model, the hierarchy in the quark and lepton textures is closely connected to a hierarchy in the spectrum of the scalar doublets, of which there are 9 in the $SO(4)_F$ model. Suppose, for example, the master matrix $(m^2)^{ki} \equiv \eta_K^k \lambda_{KI} \eta_I^i$ (which is now 4×4, of course) has $(m^2)^{11} \gg (m^2)^{22} \gg$ the other elements. Then the four scalar doublets with an index 1 (namely, $H^{(12)}$, $H^{(13)}$, $H^{(14)}$, and $H'^{(11)} \equiv (3H^{(11)} - H^{(22)} - H^{(33)} - H^{(44)})/\sqrt{12}$) will be much heavier than the three scalar doublets without an index 1 but with an index 2 (namely, $H^{(23)}$, $H^{(24)}$, and $H'^{(22)} \equiv (2H^{(22)} - H^{(33)} - H^{(44)})/\sqrt{6}$), which in turn will be much heavier than the two scalar doublets that have neither a 1 nor a 2 index (namely $H'^{(33)} \equiv (H^{(33)} - H^{(44)})/\sqrt{2}$ and $H^{(34)}$). Moreover, these two lightest doublets will have a relatively small splitting, as shown schematically in Fig. 3. (Figure 3 is not drawn to scale. $M_{\rm unif}$ is supposed to be many of orders of magnitude larger than m_0^2. $H'^{(11)}$ is supposed to be one or 2 orders of magnitude heavier than $H'^{(22)}$, and so forth.)

There is mixing among these scalars, of course, so that the lightest scalar doublet (the SM Higgs doublet) is predominantly a linear combination of $(H^{(33)} - H^{(44)})/\sqrt{2}$ and $H^{(34)}$, but with small admixtures of the others. Corresponding to this hierarchy, as seen in the last section, the standard model Higgs doublet will have its largest Yukawa couplings in the 33, 34, 43, 44 elements, the next largest in the 22, 23, 32, 24, 42 elements, and the smallest in the 11, 12, 21, 13, 31, 14, 41 elements. Moreover, the ratios of the splittings in the scalar multiplet are closely related to the ratios of the elements of the 4×4 block of the fermion mass matrices.

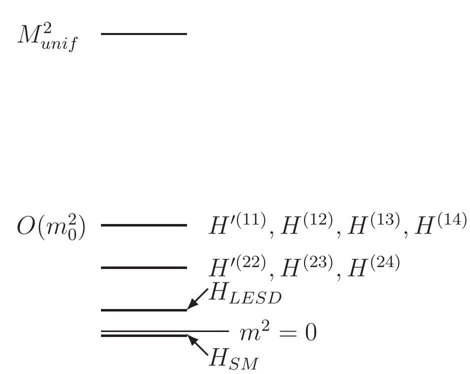

FIG. 3. A schematic plot of the mass-squared spectrum of the 9 of scalar doublets in the $SO(4)_F$ model.

As explained below, the mass matrices of up quarks, down quarks, and charged leptons can, without loss of generality, be brought to the form

$$M_U = \begin{pmatrix} a^{11} & a^{12} & a^{13} & a^{14} & 0 \\ a^{12} & a^{22} & a^{23} & a^{24} & 0 \\ a^{13} & a^{23} & a^{33} & a^{34} & 0 \\ a^{14} & a^{24} & a^{34} & -a'^{33} & B \\ 0 & 0 & 0 & B & 0 \end{pmatrix}, \qquad (23)$$

$$M_D = t\begin{pmatrix} a^{11} & a^{12} & a^{13} & a^{14} & 0 \\ a^{12} & a^{22} & a^{23} & a^{24} & 0 \\ a^{13} & a^{23} & a^{33} & a^{34} & 0 \\ a^{14} & a^{24} & a^{34} & -a'^{33} & B/t \\ C_1 & 0 & 0 & C_4 & 0 \end{pmatrix}, \qquad (24)$$

$$M_L = r\begin{pmatrix} a^{11} & a^{12} & a^{13} & a^{14} & D_1 \\ a^{12} & a^{22} & a^{23} & a^{24} & D_2 \\ a^{13} & a^{23} & a^{33} & a^{34} & 0 \\ a^{14} & a^{24} & a^{34} & -a'^{33} & D_4 \\ 0 & 0 & 0 & B/r & 0 \end{pmatrix}. \qquad (25)$$

These forms are achieved as follows: By the freedom to choose an $SO(4)_F$ basis, one can simultaneously do the same $SO(4)_F$ rotation to all the f^i and $(f^c)^i$. Under this, of course, the matrices a^{ij} retain their symmetric form and the vectors $b^i_u = b^i_d = b^i_{u^c} = b^i_{\ell^c}$ can be brought to the form $(0, 0, 0, B)$. The vectors $b^i_{d^c}$ and b^i_ℓ, being different in general, will not be brought to a special form by this rotation. However, by another change of $SO(4)_F$ basis that involves rotating only in the 1,2,3 directions, one leaves the form of $(0, 0, 0, B)$ unchanged and the vector $b^i_{d^c}$ can be brought to the form $(C_1, 0, 0, C_4)$. Finally, by a third change of basis that involves only the 2,3 directions, one can bring b^i_ℓ to the form $(D_1, D_2, 0, D_4)$. The parameters t and r are defined by $t \equiv \tan\theta_H$, and $r \equiv y'/y$, where y and y' are Yukawa couplings appearing in Eq. (17).

From Eqs. (23)–(25) it can be seen that the effective mass matrices of the three light families of up quarks, down quarks, and charged leptons, $\tilde{M}_U$, $\tilde{M}_D$, and $\tilde{M}_L$, depend on 14 parameters: a^{ij}, t, r, C_1/C_4, D_1/D_4, and D_2/D_4. These must fit 12 observables: six quark masses, three charged lepton masses, and the three CKM parameters V_{us}, V_{cb}, and V_{ub}. (The neutrino masses can arise in several ways, as discussed in Appendix C, and depend on several other parameters.) If one considers just the quarks, there are 11 parameters to fit 9 quantities. Realistic fits can be obtained, and will be presented in another place.

The model can fit, but does not predict, the quark and lepton masses and mixing angles, but by fitting those quantities, one determines enough parameters of the model to allow one to calculate in terms of only a few unknown parameters the flavor violation mediated by the extra scalar doublets—which in fact is dominated by the exchange of the lightest extra scalar doublet (LESD), as well as all the proton-decay branching ratios.

To illustrate how predictive the $SO(4)_F$ model is, consider for simplicity the case where CP is conserved and all parameters are real. The traceless part of the 4×4 symmetric matrix $(m^2)^{ij}$, which has 9 parameters, determines the complete mass spectrum of the 9 light scalar doublets (except for the overall mass of the 9-plet, which is determined anthropically, and is known once the mass-squared of the SM Higgs doublet is measured directly). That means that m^2 determines which linear combinations of $H^{(ij)}$ the SM Higgs doublet is and therefore the entries a^{ij} in the mass matrices in Eq. (23). Therefore the 9 parameters in the "master matrix" m^2, together with the two parameters C_1/C_4 and t in Eq. (24), determine 8 measurable quantities: the six quark masses and two CKM mixings. (In this CP conserving case, the mixing V_{ub} is a real number, and therefore not realistic.) However, far more than that is also determined. The master matrix m^2 determines the masses of *all nine* of the light scalar doublets and which linear combinations of $H^{(ij)}$ they are. Consequently, it determines also their Yukawa coupling matrices to the quarks and therefore all the flavor-changing amplitudes at low energy, which involves the coefficients of many four-fermion operators. To put it another way, just fitting the quark masses and the CKM angles leaves *three* undetermined parameters in terms of which all the FCNC amplitudes in the quark sector can be calculated. (Actually, there are only two undetermined parameters, since the unification of gauge couplings gives one constraint on the mass spectrum of the scalar doublets.)

If one considers also the charged leptons, there is even more predictivity. Three additional model parameters, namely r, D_1/D_4, and D_2/D_4, allow one to determine the textures of the charged leptons, and thus the masses m_e, m_μ, and m_τ). The net effect, therefore, is that without any more undetermined parameters being brought in the couplings of the lepton sector and many more observable quantities can be computed. Among these are the coefficients of all the flavor-violating four-fermion operators that involve charged leptons, of which there are many ($\mu^+_R e^-_L \bar{s}_R d_L$, $\mu^-_R e^+_L e^-_R e^+_L$, etc.).

Finally, in terms of just a few more parameters, one can also predict all the proton decay branching ratios. Proton decay is mediated by the exchange of $\tilde{D}^{(ij)}$ and $\tilde{D}^{c(ij)}$. (Actually, these mix with $\tilde{d}^{(ij)}$ and $\tilde{d}^{c(ij)}$. See Appendix A.) These are superheavy, and so the G_F-breaking splittings among their masses can be neglected in computing proton decay. The masses of these colored scalars are dominated by two terms, which in the notation of Appendix B are $M'_\Delta(\langle\tilde{S}\rangle\tilde{D}^{c(ij)} + \langle\tilde{N}^c\rangle\tilde{d}^{c(ij)})\tilde{D}^{(ij)}$. The couplings of these colored scalars are completely known, since they are part of the same $SO(4)_F$ multiplet with the SM Higgs doublet, and are controlled by the same Yukawa terms (the first two in Eq. (17)). For example, in the basis of Eqs. (23)–(25), $\tilde{D}^{(ij)}$ just couples in the $ij(ji)$ direction. There is also proton decay mediated by the G_F-singlet colored scalars $\tilde{D}$, $\tilde{D}^c$. These amplitudes would depend on an additional parameter. (Again, the pattern of the Yukawa couplings of these G_F-singlet colored scalars is completely known, since they couple as $\psi_q^i \psi_{\bar{q}}^i \Phi_\ell$ plus similar terms.)

In sum, fitting the quark and charged lepton masses and the CKM angles, leaves only a small number of undetermined parameters in terms of which the coefficients of many flavor-violating four-fermion operators and all the proton branching ratios can be computed.

The counting is different if complex phases are taken into account. On the one hand, there are more model parameters (the complex phases), but most of these can be "absorbed" by field redefinitions, and there are also more quantities that in principle can be measured (the coefficients of CP-violating operators). We leave the fitting of the quark and lepton masses, and the predictions of the patterns of FCNC and proton-decay amplitudes to future work.

V. CONCLUSIONS

In this paper it is shown how a realistic model can be constructed in which there is a multiplicity of light scalar doublets (one of which is the standard model Higgs doublet), just as there is a multiplicity of fermion families. The multiplicity of light scalar doublets can give satisfactory gauge coupling unification [6]. The multiplicity of both light fermions and light scalars is due to their forming multiplets of a nonabelian family group. This family group protects the lightness of the "extra" scalar doublets by tying their masses to that of the standard model Higgs doublet. The mass of the standard model Higgs doublet is "anthropically tuned" to be small [2–4].

While the anthropic tuning of the scalar masses cannot be tested, there are many consequences of the nonabelian family symmetry that can be tested. In particular, the couplings of the scalars are all related to each other by the family symmetry. Knowledge of the quark and lepton masses and mixings therefore gives much information about the pattern of couplings of all the scalars. In this way many predictions of the patterns of the flavor-changing mediated by the extra light scalar doublets and of proton decay mediated by the superheavy colored partners of the light scalars can in principle be extracted. The number of parameters is enormously restricted by the family symmetry.

Here it has been shown that it is relatively easy to construct realistic models based on orthogonal family groups, and, in particular, one based on $SO(4)_F$ has been described in detail. It may be possible to use many other kinds of family symmetries, such as $SU(N)$ or nonabelian discrete symmetries.

The models discussed here are meant to illustrate the utility in guiding model-building of an extension of the old "naturalness principle", which is called here the "extended naturalness principle." This extended principle forbids apparent tunings of parameters that are not justified either by symmetry principles and dynamical mechanisms (as required by the original naturalness principle), or by anthropic considerations. Whereas anthropic tuning of a parameter is not something that can be directly tested, the requirement that tunings be either anthropically justified or be natural in the usual sense can constrain model building and lead to testable scenarios. The kinds of models presented here, which can be highly predictive, only make sense (it would seem) in the context of an anthropically tuned electroweak scale.

APPENDIX A

The group used for unification in this paper is $G_U = SU(3)_c \times SU(3)_L \times SU(3)_R \times S_3$, where S_3 permutes the three $SU(3)$ factors cyclically. A multiplet that is used both for a family and for the Higgs field that breaks the electroweak symmetry is $(3, \bar{3}, 1) + (\bar{3}, 1, 3) + (1, 3, \bar{3})$, which is denoted F throughout this paper. The following table shows how standard model fields are contained in this multiplet. Our convention in the appendices is that unprimed indices refer to $SU(3)_c$, primed indices refer to $SU(3)_L$, barred indices refer to $SU(3)_R$, $a = 1, 2, 3$ is a color $SU(3)$ index, λ' is a weak $SU(2)$ index, and a tilde over a field means that it is a boson with the same SM charges as the fermion field denoted by the same letter. The SM hypercharge is given by $Y/2 = -\frac{1}{3}\lambda'_8 - \frac{1}{3}\bar{\lambda}_3 + \bar{\lambda}_8$, where $\lambda_3 = \mathrm{diag}(\frac{1}{2}, -\frac{1}{2}, 0)$ and $\lambda_8 = \mathrm{diag}(\frac{1}{2}, \frac{1}{2}, -1)$.

Rep	*Fermions*	*Bosons*	*Y/2*
$(3, \bar{3}, 1)$	$(\psi_q)^a_{\lambda'} = Q = \begin{pmatrix} u \\ d \end{pmatrix}$	$(\Phi_q)^a_{\lambda'} = \tilde{Q} = \begin{pmatrix} \tilde{u} \\ \tilde{d} \end{pmatrix}$	1/6
	$(\psi_q)^a_{3'} = D$	$(\Phi_q)^a_{3'} = \tilde{D}$	−1/3
$(\bar{3}, 1, 3)$	$(\psi_{\bar{q}})^{\bar{1}}_a = d^c$	$(\Phi_{\bar{q}})^{\bar{1}}_a = \tilde{d}^c$	1/3
	$(\psi_{\bar{q}})^{\bar{2}}_a = u^c$	$(\Phi_{\bar{q}})^{\bar{2}}_a = \tilde{u}^c$	−2/3
	$(\psi_{\bar{q}})^{\bar{3}}_a = D^c$	$(\Phi_{\bar{q}})^{\bar{3}}_a = \tilde{D}^c$	1/3
$(1, 3, \bar{3})$	$(\psi_\ell)^{\lambda'}_{\bar{1}} = L'$	$(\Phi_\ell)^{\lambda'}_{\bar{1}} = H_d$	−1/2
	$(\psi_\ell)^{\lambda'}_{\bar{2}} = \bar{L}'$	$(\Phi_\ell)^{\lambda'}_{\bar{2}} = H_u$	1/2
	$(\psi_\ell)^{\lambda'}_{\bar{3}} = L$	$(\Phi_\ell)^{\lambda'}_{\bar{3}} = \tilde{L}$	−1/2
	$(\psi_\ell)^{3'}_{\bar{1}} = N^c$	$(\Phi_\ell)^{3'}_{\bar{1}} = \tilde{N}^c$	0
	$(\psi_\ell)^{3'}_{\bar{2}} = e^+$	$(\Phi_\ell)^{3'}_{\bar{2}} = \tilde{e}^+$	1
	$(\psi_\ell)^{3'}_{\bar{3}} = S$	$(\Phi_\ell)^{3'}_{\bar{3}} = \tilde{S}$	0

(A1)

There are two scalar multiplets that transform as F under G_U, $\Phi^{(ij)}$ and Φ. The former is rank-2 symmetric traceless tensor under the family group G_F, the latter a singlet under G_F. $\Phi^{(ij)} = \Phi_q^{(ij)} + \Phi_{\bar{q}}^{(ij)} + \Phi_\ell^{(ij)}$ and $\Phi = \Phi_q + \Phi_{\bar{q}} + \Phi_\ell$. These couple to the fermion families, which also transform as F under G_U, but as vectors under G_F: $\psi^i = \psi_q^i + \psi_{\bar{q}}^i + \psi_\ell^i$.

There are the following types of Yukawa couplings:

$$\begin{aligned} &y[(\psi_q^i\psi_{\bar{q}}^j)\Phi_\ell^{(ij)} + (\psi_{\bar{q}}^i\psi_\ell^j)\Phi_q^{(ij)} + (\psi_\ell^i\psi_q^j)\Phi_{\bar{q}}^{(ij)}] \\ &\quad + y'[(\psi_q^i\psi_q^j)\Phi_q^{(ij)} + (\psi_{\bar{q}}^i\psi_{\bar{q}}^j)\Phi_{\bar{q}}^{(ij)} + (\psi_\ell^i\psi_\ell^j)\Phi_\ell^{(ij)}] \\ &\quad + Y[(\psi_q^i\psi_{\bar{q}}^i)\Phi_\ell + (\psi_{\bar{q}}^i\psi_\ell^i)\Phi_q + (\psi_\ell^i\psi_q^i)\Phi_{\bar{q}}] \\ &\quad + Y'[(\psi_q^i\psi_q^i)\Phi_q + (\psi_{\bar{q}}^i\psi_{\bar{q}}^i)\Phi_{\bar{q}} + (\psi_\ell^i\psi_\ell^i)\Phi_\ell] \end{aligned} \tag{A2}$$

Suppressing the family indices these contain the following kinds of terms

$$\begin{array}{lll} (\psi_q\psi_{\bar{q}})\Phi_\ell & +(\psi_{\bar{q}}\psi_\ell)\Phi_q & +(\psi_\ell\psi_q)\Phi_{\bar{q}} \\ = (Qd^c)H_d & +(d^cL')\tilde{Q} & +(L'Q)\tilde{d}^c \\ +(Qu^c)H_u & +(u^c\bar{L}')\tilde{Q} & +(\bar{L}'Q)\tilde{u}^c \\ +(QD^c)\tilde{L} & +(D^cL)\tilde{Q} & +(LQ)\tilde{D}^c \\ +(Dd^c)\tilde{N}^c & +(d^cN^c)\tilde{D} & +(N^cD)\tilde{d}^c \\ +(Du^c)\tilde{e}^+ & +(u^ce^+)\tilde{D} & +(e^+D)\tilde{u}^c \\ +(DD^c)\tilde{S} & +(D^cS)\tilde{D} & +(SD)\tilde{D}^c, \end{array} \tag{A3}$$

and

$$\begin{array}{lll} (\psi_\ell\psi_\ell)\Phi_\ell & +(\psi_q\psi_q)\Phi_q & +(\psi_{\bar{q}}\psi_{\bar{q}})\Phi_{\bar{q}} \\ = (Le^+)H_d & +(QQ)\tilde{D} & +(d^cu^c)\tilde{D}^c \\ +(L'L)\tilde{e}^+ & +(QD)\tilde{Q} & +(u^cD^c)\tilde{d}^c \\ +(e^+L')\tilde{L} & & +(D^cd^c)u^c \\ +(LN^c)H_u & & \\ +(\bar{L}'L)\tilde{N}^c & & \\ +(N^c\bar{L})\tilde{L} & & \\ +(\bar{L}'S)H_d & & \\ +(L'\bar{L}')\tilde{S} & & \\ +(SL')H_u & & \end{array} \tag{A4}$$

In the scalar multiplets $\Phi^{(ij)}$ and Φ, only the parts $\Phi_\ell^{(ij)}$ and Φ_ℓ have components with nonzero VEVs, as otherwise color would be broken. Φ_ℓ contains superlarge VEVs in $\tilde{S}(= \Phi_{\bar{3}}^{3'})$ and $\tilde{N}^c(= \Phi_{\bar{1}}^{3'})$, which help break G_U down to the standard model group (which, of course, also means breaking S_3). These large VEVs get rid of the extra fermions in each family by giving mass to the $D, D^c, L', \bar{L}'$, as can be seen from Eqs. (A3) and (A4). (If $\langle\tilde{N}^c\rangle = 0$, then D mates purely with D^c, and $\bar{L}'$ mates purely with L'. With $\langle\tilde{N}^c\rangle \neq 0$, however, D mates partly with d^c, so that the light right-handed down quarks are linear combinations of D^c and d^c. Similarly, $\bar{L}'$ mates partly with L, so that the light lepton doublets are linear combinations of L and L'.)

The $\Phi_\ell^{(ij)}$ contains the doublets $H_u^{(ij)}$, $H_d^{(ij)*}$, and $\tilde{L}^{(ij)*}$. Anthropic fine tuning makes one linear combination of these light (as discussed below) which we call $H^{(ij)}$. The lightest of the $H^{(ij)}$ is the standard model Higgs doublet H_{SM}.

The G_F-invariant masses of the scalar doublets get contributions from the terms

$$\begin{aligned} &M_\Phi^2\Phi_\ell^{(ij)*}\Phi_\ell^{(ij)} \\ &\quad +[M_\Delta\Phi_\ell^{(ij)}\Phi_\ell^{(ij)}\Phi_\ell + H.c.] \\ &\quad +\sigma\, Tr(\Phi_\ell^{(ij)*}\Phi_\ell^{(ij)})\, Tr(\Phi_\ell^*\Phi_\ell) + \rho\, Tr(\Phi_\ell^{(ij)*}\Phi_\ell^{(ij)}\Phi_\ell^*\Phi_\ell), \end{aligned} \tag{A5}$$

where "Tr" in the last two terms refers to traces over the $SU(3)_c \times SU(3)_L \times SU(3)_R$ indices. There are other terms that are related to those in Eq. (A5) by S_3 permutations. There are also other quartic terms that do not contribute to the superlarge masses of the scalar doublets. The terms in Eq. (A3) give a mass-squared matrix for the scalar doublets of the form

$$(H_u, H_d^*, \tilde{L}^*)\begin{pmatrix} M^2 & M_\Delta\langle\tilde{S}\rangle & M_\Delta\langle\tilde{N}^c\rangle \\ M_\Delta^*\langle\tilde{S}\rangle^* & M^2 + \rho|\langle\tilde{N}^c\rangle|^2 & \rho\langle\tilde{S}\rangle^*\langle\tilde{N}^c\rangle \\ M_\Delta^*\langle\tilde{N}^c\rangle^* & \rho\langle\tilde{N}^c\rangle^*\langle\tilde{S}\rangle & M^2 + \rho|\langle\tilde{S}\rangle|^2 \end{pmatrix} \times \begin{pmatrix} H_u^* \\ H_d \\ \tilde{L} \end{pmatrix}, \tag{A6}$$

where $M^2 = M_\Phi^2 + \sigma(|\langle\tilde{S}\rangle|^2 + |\langle\tilde{N}^c\rangle|^2)$.

In the discussion in the main text, only the mixing of H_u and H_d were considered, not $\tilde{L}$, for ease of discussion. Including the mixing with $\tilde{L}$ does not qualitatively affect the conclusions reached in the text. Note that in general there are three unequal eigenvalues of this matrix. Also, by having two parameters scan, such as M_Φ^2 and M_Δ, both the lightest eigenvalue μ^2 and the parameter that was called $\tan\theta_H$ in the text will scan.

APPENDIX B

From the Yukawa couplings shown in Appendix A one sees that $\tilde{D}$, $\tilde{D}^c$ can mediate proton decay. The terms $(d^cN^c)\tilde{D}$, $(u^ce^+)\tilde{D}$ conserve B and L only if $\tilde{D}$ has $B = \frac{1}{3}$, $L = 1$; whereas the term $(QQ)\tilde{D}$ conserves B and L only if $\tilde{D}$ has $B = -\frac{2}{3}$, $L = 0$. Since both kinds of terms are present, $\tilde{D}$ exchange mediates proton decay. Similar arguments apply to $\tilde{D}^c$. (On the other hand, the exchange of $\tilde{Q}$ does not cause dangerous proton decay. The terms $(d^cL')\tilde{Q}$ and $(D^cL)\tilde{Q}$ conserve B and L if $\tilde{Q}$ has $B = \frac{1}{3}$ and $L = -1$. Then the term $(QD)\tilde{Q}$ violates B and L, but this term contains the purely superheavy quark D, and so does not produce rapid proton decay.)

The question is whether the $\tilde{D}$ and $\tilde{D}^c$ can be made heavy enough to avoid rapid proton decay, while leaving G_U unbroken down to the scale 10^{14} GeV. This can be done by the terms $(M'_\Delta\Phi_q^{(ij)}\Phi_{\bar{q}}^{(ij)}\phi_\ell + \text{cyclic})$ and $M''_\Delta\Phi_q\Phi_{\bar{q}}\Phi_\ell + \text{cyclic})$. These give masses

$(M'_\Delta \langle \tilde{S} \rangle)\tilde{D}^{(ij)}\tilde{D}^{c(ij)}$ and $(M''_\Delta \langle \tilde{S} \rangle)\tilde{D}\tilde{D}^c$. Since G_U invariance allows M'_Δ and M''_Δ to be arbitrarily large, the masses of the dangerous colored scalars can be much larger than 10^{14} GeV. Note that these are different terms than the $(M_\Delta \Phi_\ell^{(ij)}\Phi_\ell^{(ij)}\phi_\ell + \text{cyclic})$ that are responsible for the δ^2 term in the scalar doublet mass matrix (Equation (8) of the text), and unrelated to it by S_3. Making the dangerous scalars heavy to suppress proton decay does create split multiplets that give threshold corrections to the running of the gauge couplings above the scale 10^{14} GeV.

APPENDIX C

As discussed after Eq. (19) in the text, the Yukawa couplings $y_J[\psi_q^i \bar{\psi}_q^i \eta_J^i + \text{cyclic}]$ must be different for the different types of fermions (up quarks, down quarks, and charged leptons) in order to get realistic mass matrices. If they are not different, then all the 3×3 mass matrices of the observed quarks and leptons, $\tilde{M}_U$, $\tilde{M}_D$, and $\tilde{M}_L$ will be proportional to each other, giving no CKM mixing and unrealistic mass relations. To make the couplings different, the superlarge VEVs that break the unified group must come into the low energy Yukawa couplings. This can happen in a simple way if superheavy fields that get mass from these VEVs are "integrated out" to give effective $d > 4$ Yukawa terms.

Suppose, for example, that in the $SO(4)_F$ model of Section IV there is an additional superheavy family-mirror family pair, $\psi' + \bar{\psi}'$ The complete set of quarks and leptons is thus $(F, 4, 1) = \psi^i$, $(\bar{F}, 1, 1) = \bar{\psi}$, $(F, 1, 1) = \psi'$, $(\bar{F}, 1, 1) = \bar{\psi}'$. Then the G_F-singlet family ψ' will "mate" with some linear combination of the two G_F-singlet mirror families $\bar{\psi}$ and $\bar{\psi}'$ to get a superlarge mass, leaving one mirror family light. Superlarge VEVs that break the unified group can also contribute to these superlarge masses. Thus, the mirror family that remains light will consist of linear combinations of $\bar{\psi}$ and $\bar{\psi}'$ that know about the breaking of the unified group.

Consider the following terms;

$$y_J(\psi^i \bar{\psi})\eta_J^i + y'_J(\psi^i \bar{\psi}')\eta_J^i. \tag{C1}$$

Suppose that for a type of fermion f, the "light" $\bar{f}$ (the one that does not get a superlarge mass) is a combination $\bar{f} = \alpha_f \bar{f}_{(\psi)} + \beta_f \bar{f}_{(\bar{\psi}')}$. Then, the above term gives

$$[\alpha_f(y_J\langle \eta_J^i \rangle) + \beta_f(y'_J\langle \eta_J^i \rangle)]f^i \bar{f}. \tag{C2}$$

Since, in general, $y_J\langle \eta_J^i \rangle$ and $y'_J\langle \eta_J^i \rangle$ point in different directions in $SO(4)_F$ space, and α_f and β_f can be different for different fermion types f, one sees that the desired difference in the textures can result: the 1×4 and 4×1 blocks can be different in M_U, M_D, and M_L.

The superheavy fermion mass terms that are relevant are

$$\begin{aligned}
&(M_1 \psi' \bar{\psi} + M_2 \psi' \bar{\psi}')_{[qq+\text{cyclic}]} \\
&\quad +(\lambda_1 \psi \psi \Phi + \lambda_2 \psi' \psi' \Phi + \lambda_3 \psi \psi' \Phi)_{[\ell\ell\ell+\text{cyclic}]} \\
&\quad +(\rho_1 \psi \psi \Phi + \rho_2 \psi' \psi' \Phi + \rho_3 \psi \psi' \Phi)_{[q\bar{q}\ell+\text{cyclic}]} \\
&\quad +(\lambda_4 \bar{\psi}\, \bar{\psi}\, \Phi^* + \lambda_5 \bar{\psi}' \bar{\psi}' \Phi^* + \lambda_6 \bar{\psi}\, \bar{\psi}' \Phi^*)_{[\ell\ell\ell+\text{cyclic}]} \\
&\quad +(\rho_4 \bar{\psi}\, \bar{\psi}\, \Phi^* + \rho_5 \bar{\psi}' \bar{\psi}' \Phi^* + \rho_6 \bar{\psi}\, \bar{\psi}' \Phi^*)_{[q\bar{q}\ell+\text{cyclic}]}.
\end{aligned} \tag{C3}$$

The complete problem will not be analyzed explicitly here, but the significant points that emerge from such an analysis will be indicated. If the superheavy VEVs of Φ_ℓ (namely $\langle \tilde{S} \rangle$ and $\langle \tilde{N}^c \rangle$) are neglected, one only has the terms $M_1 \psi' \bar{\psi} + M_2 \psi' \bar{\psi}')_{[qq+\text{cyclic}]}$, which treat all types of fermions the same. The distinction between the different fermion types comes from the superlarge VEVs of Φ_ℓ, which break the unified group G_U. But these superlarge VEVs give mass only to the species D, D^c/d^c, L'/L, and $\bar{L}'$. That is why it is only the particle types d^c and L that get distinguished from the others, as stated in Eq. (20).

Neutrino masses can arise in several ways. Perhaps the simplest is to introduce fermions that are singlets under G_U and vectors under the family group, to play the role of "right-handed neutrinos" in the type-I seesaw mechanism. For example, in the $SO(4)_F$ model they would transform under $G_U \times SO(4)_F \times SU(N)_{\text{DYN}}$ as (1, 4, 1). Denote these by N^i. Then the following couplings would be allowed:

$$\begin{aligned}
&\mathcal{Y}_1[(\psi_\ell^i N^j)\Phi_\ell^{*(ij)} + \text{cyclic}] \\
&\quad +\mathcal{Y}_2[(\psi_\ell^i N^i)\Phi_\ell^* + \text{cyclic}] \\
&\quad +M_N(N^i N^i) \\
&\gamma[(\Phi_\ell^* \Phi_\ell^{(ij)})(\Phi_\ell^* \Phi_\ell^{(ij)}) + \text{cyclic}] + H.c.
\end{aligned} \tag{C4}$$

Integrating out the N^i would give a tree-level contribution to the light neutrino masses of the type-I form, namely $M_\nu = -M_{\text{Dirac}} M_R^{-1} M_{\text{Dirac}}^T$, with $(M_{\text{Dirac}})^{ij} = \mathcal{Y}_1 \langle H_d^{(ij)} \rangle^*$ and $(M_R)^{ij} = M_N \delta^{ij}$. This by itself would give unrealistic neutrino masses, as they would have a strong hierarchy of masses. There would also be loop contributions to the neutrino masses that went as $(M_\nu)^{ij} = \delta^{ij} \frac{1}{16\pi^2} \gamma(\mathcal{Y}_2)^2 \times \langle H_u^{(k\ell)} \rangle^2 / \mathcal{M}$, where $\mathcal{M}$ is a combination of superheavy masses that arises from the momentum integral of the loop. Other contributions to the light neutrino mass matrix, both tree-level and from loops are also possible, depending on what fields and couplings are present at high scales. Thus, such models are not predictive of neutrino properties.

[1] L. Susskind, Phys. Rev. D **20**, 2619 (1979); G. 'tHooft, in *Recent Developments in Gauge Theories*, edited by G. 'tHooft *et al.* NATO Advanced Study Institutes, Ser. B., 59 (Plenum Press, New York, 1980), 135.
[2] V. Agrawal, S. M. Barr, J. F. Donoghue, and D. Seckel, Phys. Rev. D **57**, 5480 (1998); Phys. Rev. Lett. **80**, 1822 (1998).
[3] S. M. Barr and A. Khan, Phys. Rev. D **76**, 045002 (2007).
[4] S. Weinberg, in *Universe or Multiverse*, edited by B. Carr (Cambridge University Press, Cambridge, England, 2006), p. 29.
[5] S. Willenbrock, Phys. Lett. B **561**, 130 (2003).
[6] J. Sayre, S. Wiesenfeldt, and S. Willenbrock, Phys. Rev. D **73**, 035013 (2006).
[7] Z. G. Berezhiani and M. Yu. Khlopov, Yad. Fiz. **51**, 1157 (1990) [Sov. J. Nucl. Phys. **51**, 739 (1990)]; Yad. Fiz. **51**, 1479 (1990) [Sov. J. Nucl. Phys. **51**, 935 (1990)].
[8] N. Kifune, J. Kubo, and A. Linz, Phys. Rev. D **77**, 076010 (2008); K. Kawashima, J. Kubo, and A. Linz, Phys. Lett. B **681**, 60 (2009); J. Kubo and A. Linz, arXiv:1007.0680.
[9] Y. Kawamura, Prog. Theor. Phys. **105**, 691 (2001); **105**, 999 (2001).
[10] R. Peccei and H. Quinn, Phys. Rev. Lett. **38**, 1440 (1977); A. E. Nelson, Phys. Lett. B **136**, 387 (1984);S. M. Barr, Phys. Rev. D **30**, 1805 (1984); Phys. Rev. Lett. **53**, 329 (1984).
[11] C. Hogan, Rev. Mod. Phys. **72**, 1149 (2000); Phys. Rev. D **74**, 123514 (2006).