On the chirality of the SM and the fermion content of GUTs

Renato M. Fonseca*

AHEP Group, Instituto de Física Corpuscular, C.S.I.C./Universitat de València
Edifício de Institutos de Paterna, Apartado 22085, E–46071 València, Spain

23 June 2015

Abstract

The Standard Model (SM) is a chiral theory, where right- and left-handed fermion fields transform differently under the gauge group. Extra fermions, if they do exist, need to be heavy otherwise they would have already been observed. With no complex mechanisms at work, such as confining interactions or extra-dimensions, this can only be achieved if every extra right-handed fermion comes paired with a left-handed one transforming in the same way under the Standard Model gauge group, otherwise the new states would only get a mass after electroweak symmetry breaking, which would necessarily be small (~ 100 GeV). Such a simple requirement severely constrains the fermion content of Grand Unified Theories (GUTs). It is known for example that three copies of the representations $\mathbf{\bar{5}} + \mathbf{10}$ of $SU(5)$ or three copies of the **16** of $SO(10)$ can reproduce the Standard Model's chirality, but how unique are these arrangements? In a systematic way, this paper looks at the possibility of having non-standard mixtures of fermion GUT representations yielding the correct Standard Model chirality. Family unification is possible with large special unitary groups — for example, the **171** representation of $SU(19)$ may decompose as $3\,(\mathbf{16}) + \mathbf{120} + 3\,(\mathbf{1})$ under $SO(10)$.

Keywords: SM chirality, fermions in GUTs, family unification, $SU(19)$ model.

---
*E-mail: renato.fonseca@ific.uv.es

1 Introduction

There is currently no explanation for the flavor structure of the Standard Model (SM) and Grand Unified Theories (GUTs) developed over the past decades have failed to shed light on this issue since particles with different flavors are usually assigned to distinct copies of a single representation of the enlarged gauge group. For example, with $SU(5)$ one considers three copies (one for each flavor) of the representations $\overline{\mathbf{5}}$ and $\mathbf{10}$, containing exactly the SM fermions [1]: three replicas of $Q = \left(\mathbf{3}, \mathbf{2}, \frac{1}{6}\right)$, $u^c = \left(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3}\right)$, $d^c = \left(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3}\right)$, $L = \left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}\right)$, and $e^c = (\mathbf{1}, \mathbf{1}, 1)$. In $SO(10)$ models, three $\mathbf{16}$'s contain all SM fermions plus three right-handed neutrinos $N^c = (\mathbf{1}, \mathbf{1}, 0)$ [2, 3]. Once the $SO(10)$ symmetry is broken, a vector (Majorana) mass $m N^c N^c$ is allowed for each of these extra fermion states, explaining why they have yet to be (directly) observed. Increasing further the size of the group, there is also the well known possibility of having three copies of the $\mathbf{27}$ in E_6-based models [4], which contain 11 additional vector particles per generation.

In order to completely explain flavor with GUTs it would be necessary to place the SM fermions in a single representation of the gauge group.[1] This idea goes by the name of *family unification* and it was attempted in the past with a variety of groups [5–18]. For instance, the spinor representation of $SO(10 + 2N')$ can be broken into $2^{N'-1}$ copies of the $\mathbf{16}$ of $SO(10)$, yet it also contains an equal amount of $\overline{\mathbf{16}}$'s. Therefore, without confining interactions [5, 6, 19, 20], extra dimensions [21–35], or some other elaborate mechanism, one cannot give a big mass to all these mirror families without making all the families super-heavy as well.[2]

In fact, mirror families are just part of a larger problem: in general it is necessary to justify why all types of exotic fermions are heavy. Take as an example the representation $\mathbf{560}$ of $SO(10)$, which is the smallest one containing all SM fermions. On top of the fact that the excess of fermions $(5Q + 4u^c + 4d^c + 3L + 3e^c)$ over mirror fermions $(1\overline{Q} + 1\overline{u^c} + 1\overline{d^c} + 1\overline{L} + 1\overline{e^c})$ is not the correct one, there are also fermions in exotic SM representations such as $\left(\overline{\mathbf{15}}, \mathbf{1}, \frac{1}{3}\right)$, and none in matching conjugate representations. Such states could only acquire an electroweak (EW) scale mass and therefore would have already been seen at the Tevatron and the LHC.

Perhaps the idea of unifying the three families in a single GUT fermion representation is too ambitious. One should then also consider models where the observed fermion states are distributed over various GUT representations [36–51]. Such models might still be quite interesting: if the GUT representations are not just mere copies of one another, the gauge symmetry alone might explain the existence of non-trivial flavor structures at low energies.

GUTs with an exotic fermion content may also have unusual features which go against what is usually taken for granted. One of them is non-standard normalizations of the hypercharge operator, which we shall now discuss. In order to see if the gauge couplings unify at a high scale in a given model, one usually takes the values of $g_1 = \sqrt{5/3}g'$, $g_2 = g$ and $g_3 = g_s$ at roughly the Z-boson mass scale and runs them with the renormalization group equations up to high energies. The explanation for the numerical factor $\sqrt{5/3}$ is simple: the Lagrangian depends on the product of the gauge coupling constant g' times the hypercharge operator Y, so the change $(g', Y) \to (n^{-1}g', nY)$ for some n is of no consequence in the SM. Comparing the three gauge coupling constants is then pointless. However, if $SU(3)_C \times SU(2)_L \times U(1)_Y$ is a remnant of a larger simple gauge group, then suddenly there is a natural value for this n parameter: one

---

[1]Rigorously speaking, when referring to a group we have in mind its algebra.

[2]The presence of confining gauge interactions could be an elegant solution to this problem, as pointed out in [5, 6]. For example, if $SO(18)$ breaks into $SO(10) \times SO(5)$ such that $\mathbf{256} \to 3\,(\mathbf{16}, \mathbf{1}) + (\mathbf{16}, \mathbf{5}) + 2\,(\overline{\mathbf{16}}, \mathbf{4})$ and if $SO(5)$ becomes non-perturbative at some high-scale, then one would expect that the only fermions which would remain light would be the three $\mathbf{16}$'s which are $SO(5)$ singlets. However, it seems difficult to drive $SO(5)$ into a non-perturbative regime at high energies.

would like Y to be one of the generators of this enlarged group, in which case its normalization must be the same as the rest of the generators: $\text{Tr}\left(T_a^2\right) = $ constant. As such, if we identify the components of the $\overline{\mathbf{5}}$ in an $SU(5)$ theory with those of the SM representations $d^c = \left(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3}n\right)$ and $L = \left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}n\right)$, then $Y = n \times \text{diag}\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -\frac{1}{2}, -\frac{1}{2}\right)$ and the third generator of $SU(2)_L$ is given by the matrix $T_{3L} = \text{diag}\left(0, 0, 0, \frac{1}{2}, -\frac{1}{2}\right)$, hence $|n| = \sqrt{3/5}$.

This hypercharge normalization factor is often mentioned as being specific to $SU(5)$ models, even though it is actually very generic. Nevertheless, it is conceivable that n might differ from $\sqrt{3/5}$: for instance, in $SU(5)$ one might try to identify d^c with the representation $\left(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3}\sqrt{3/5}\right)$ inside the $\mathbf{10}$, u^c with $\left(\overline{\mathbf{3}}, \mathbf{1}, \frac{4}{3}\sqrt{3/5}\right)$ in the $\mathbf{45}$, and so on, in which case $g_1 = -\frac{1}{2}\sqrt{5/3}g'$ would be the correct relation.[3] Another possibility would be to have, for example, an $SU(7)$ model with the fundamental representation breaking into $X \equiv \left(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3}n\right) + \left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}n\right) + (\mathbf{1}, \mathbf{1}, m) + (\mathbf{1}, \mathbf{1}, -m)$ with a non-zero m, in which case it is clear that n will not be equal to $\sqrt{3/5}$, so $g_1 \neq \sqrt{5/3}g'$ if we were to identify d^c and L with the first two SM representations. Note that X forms a unitary 7-dimensional (reducible) representation of the SM group, so it is certainly possible to make the fundamental representation of $SU(7)$ break in this way.

What about having $\mathbf{7} \rightarrow \left(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3}n\right) + \left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}n\right) + (\mathbf{1}, \mathbf{2}, 0)$? Such a scenario is even more interesting. The fundamental representation of $SU(7)$ branches into a single color anti-triplet, as usual, but now there are two doublets, which means that one should take $1/\left(2\sqrt{2}\right)$ times the Pauli matrices as the generators of $SU(2)_L$, and not half the Pauli matrices. In such a model, the correctly normalized gauge couplings would then be $g_1 = \sqrt{5/3}g'$, $g_2 = \sqrt{2}g$ and $g_3 = g_s$.[4] So, despite the widespread belief that this issue only affects abelian groups, clearly there might be potentially interesting normalization corrections to any of the gauge coupling constants in GUTs with non-standard fermion assignments.

In summary, with or without family unification, it seems appropriate to systematically study the possible ways of arranging the fermions in Grand Unified Theories. The requirement that the SM chirality must be reproduced is a simple yet very stringent constraint which can be readily used to narrow down the list of possibilities. The aim of this paper is precisely to analyze, in a comprehensive way, the fermion sector of GUTs based on different groups, checking whether or not it is possible to obtain only the observed three families of fermions plus vector particles. Importantly and in contrast to what is almost universally done in the literature, the fact that the SM group can be embedded in more than one way in a given GUT group will not be overlooked. The aim of the present work is therefore somewhat similar to the one of the papers [37, 39, 52, 53], but it is substantially broader in scope. For example, comparing with the interesting paper [53], we do not require (a) asymptotic freedom of the GUT (which severely restricts the number of fermion components allowed and consequently the group), nor (b) absence of gauge anomalies (although they get canceled automatically in almost all cases) and, above all, we do not make the (c) "bold assumption" that the embedding of the SM group is as trivial as possible.

---

[3] As explained later on, this particular example fails because one would not find anywhere the representation $\left(\mathbf{3}, \mathbf{2}, -\frac{1}{3}\sqrt{3/5}\right)$ needed for the left-handed quarks.

[4] It is amusing to consider the possibility of (almost) unifying the three gauge couplings at low energies exclusively in this way (although baryon number violation would be a concern). Conceptually, it is not very complicated: for example, if the fundamental representation of $SU(9)$ is broken into $\left(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3}n\right) + \left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}n\right) + (\mathbf{1}, \mathbf{2}, n) + (\mathbf{1}, \mathbf{2}, -n)$ with n necessarily equal to $\sqrt{3/29}$, successfully associating the d^c and L fermions of the SM with the first two representations would imply that g_1, g_2 and g_3 have almost the same value at the EW scale (up to around $\sim 10\%$).

We shall first provide some generic considerations about the method used to scan over the various GUT groups, representations and embeddings (section 2) and, following that, the results for each group are presented and discussed (section 3 supplemented by an appendix). The main conclusions are summarized at the end (section 4).

2 Framework of the analysis

2.1 The chirality of GUT models and representations

Let us briefly discuss and settle on a precise definition of (SM) chirality. Consider some embedding of the SM group G_{SM} in a bigger group G. We shall be interested in tracking the representations R^i_{SM} of G_{SM} contained in some fermion representation R of G — the so-called *branching rules* of R. Yet, since pairs of SM vector fermions are irrelevant for the present analysis (as they can be made very heavy), we may define the chirality of R to be the vector $\chi(R)$ with component i given by the number of SM representations R^i_{SM} contained in R minus the number of SM representations R^{i*}_{SM} in R:

$$\chi_i(R) \equiv \left(\#R^i_{SM} \in R\right) - \left(\#R^{i*}_{SM} \in R\right) . \tag{1}$$

For any real[5] SM representation $(R^i_{SM} = R^{i*}_{SM})$ we always get $\chi_i(R) = 0$. On the other hand, we have the relation $\chi(R^*) = -\chi(R)$ which implies that $\chi(R)$ is the null vector for a real R $(R = R^*)$. As such, SM (or GUT) real representations can be ignored completely and furthermore, concerning complex representations, the effect of having n copies of R^i_{SM}(or R) in a model is the same as subtracting from it n copies of R^{i*}_{SM}(or R^*) as far as chirality is concerned. For this reason, in this work we take $-R^i_{SM}$(or $-R$) to be the exactly the same as R^{i*}_{SM}(or R^*).

In the case of sums of representations of G, chirality is taken to be simply the sum of the chirality of each representation,

$$\chi_i\left(R^1 + \cdots + R^n\right) \equiv \chi_i\left(R^1\right) + \cdots + \chi_i\left(R^n\right) , \tag{2}$$

so we can speak of the chirality of a model. For example, in the basis where we consider only $R^i_{SM} = Q, u^c, d^c, L, e^c$ the chirality of the Standard Model is given by the vector $\chi(SM) = (3, 3, 3, 3, 3)^T$.

This definition of chirality encodes in a precise way the intuitive notion associated to this word. It counts the number of each type of SM representation, factoring out real and conjugate pairs of representations. In the following, we shall see how it allows us to turn the problem of finding GUTs with the SM chirality into solving a system of linear equations.

2.2 GUTs with the correct chirality

With the above definition, finding the chirality of a representation of a group $G \supset G_{SM}$ is a matter of decomposing it into SM representations. In order to do so, one must first know how G_{SM} is embedded in G, and it turns out that figuring all the possible ways of doing so is a complicated problem, which we shall discuss later. For now, we may assume that this embedding information

---

[5]In this work, real representations are those which are equivalent to their conjugate, so they include what is sometimes called real and pseudo-real representations in other contexts.

is known and fixed. If so, one can use computer programs such as `Susyno` [54][6] or `LieART` [55] to decompose in a systematic way any representation of the group G into those of G_{SM} (for this work, the former was used).

In turn, with the branching rules of a list of representations R_i of G, it is a rather simple exercise of linear algebra to find all integer linear combinations $\sum_i c_i R_i$ of the R_i with the SM chirality. Indeed, defining $\mathcal{M}$ to be the matrix with entries $\mathcal{M}_{ij} = \chi_i (\Psi_j)$, the vector $c = (c_1, \cdots, c_n)^T$ whose components are the integers c_i we seek is the solution to the linear system

$$\chi (SM) = \mathcal{M} \cdot c \tag{3}$$

where $\chi (SM)$ is a vector with the SM chirality, as mentioned previously. From $\chi (SM)$ and $\mathcal{M}$, one can extract c. As it is well known, the general solution of this equation is of the form

$$c = \widetilde{c} + \sum_i \alpha_i n_i \,, \tag{4}$$

where $\widetilde{c}$ is any particular solution of equation (3) and the vectors n_i are a basis of the nullspace of $\mathcal{M}$ (i.e., $\mathcal{M} \cdot n_i = 0$). The α_i are plain numbers which can take any value, as long as the components of the c vector are integer numbers. One can understand this generic form of c as follows: the vector $\widetilde{c}$ describes a particular combination of the R_j ($\sum_j \widetilde{c}_j R_j$) possessing the correct chirality, and each of the n_i describes an independent, non-trivial combination of the R_j ($\sum_j (n_i)_j R_j$) with no chirality; therefore an arbitrary number of n_i's can be added or subtracted to $\widetilde{c}$.

To clarify this approach, consider the following straightforward example. Take $SU(5)$ as the grand unified group, and its complex representations up to size 35 (we only need to consider one member of each conjugated pair): $\mathbf{5}, \mathbf{10}, \mathbf{15}, \mathbf{35}$. They decompose into the following eleven $SU(3)_C \times SU(2)_L \times U(1)_Y$ representations (plus their conjugates): $(\mathbf{1}, \mathbf{1}, 1)$, $\left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}\right)$, $(\mathbf{1}, \mathbf{3}, -1)$, $\left(\mathbf{1}, \mathbf{4}, -\frac{3}{2}\right)$, $\left(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3}\right)$, $\left(\mathbf{3}, \mathbf{2}, \frac{1}{6}\right)$, $\left(\overline{\mathbf{3}}, \mathbf{3}, -\frac{2}{3}\right)$, $\left(\mathbf{6}, \mathbf{1}, \frac{2}{3}\right)$, $\left(\mathbf{6}, \mathbf{2}, \frac{1}{6}\right)$, $(\overline{\mathbf{10}}, \mathbf{1}, 1)$, $\left(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3}\right)$. With this ordering of the SM representations, from the decomposition of $R = \mathbf{5}, \mathbf{10}, \mathbf{15}, \mathbf{35}$ we get

$$\chi (\mathbf{5}) = (0, -1, 0, 0, -1, 0, 0, 0, 0, 0, 0)^T \,; \quad \chi (\mathbf{10}) = (1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1)^T \,, \tag{5}$$

$$\chi (\mathbf{15}) = (0, 0, -1, 0, 0, 1, 0, -1, 0, 0, 0)^T \,, \quad \chi (\mathbf{35}) = (0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0)^T \,. \tag{6}$$

The chirality of the SM itself is

$$\chi (SM) = (3, 3, 0, 0, 3, 3, 0, 0, 0, 0, 3)^T \,. \tag{7}$$

So, how many copies of each of the four $SU(5)$ representations are needed in order to obtain the SM chirality? If the fermions of the GUT model are $c_1 (\mathbf{5}) + c_2 (\mathbf{10}) + c_3 (\mathbf{15}) + c_4 (\mathbf{35})$, solving

---

[6]The naming scheme for representations used by the program and in this work follows the convention laid out in the manual of the program `LieART` [55].

the system

$$
\begin{pmatrix} 3 \\ 3 \\ 0 \\ 0 \\ 3 \\ 3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \tag{8}
$$

yields a unique solution (the matrix in this equation has a trivial nullspace): $(c_1, c_2, c_3, c_4) = (-3, 3, 0, 0)$. This means that, with $SU(5)$ representations of size up to 35, the fermion fields must be $-3\,(\mathbf{5}) + 3\,(\mathbf{10})$ or equivalently $3\,(\mathbf{\bar{5}}) + 3\,(\mathbf{10})$, up to trivial variations (i.e., addition of real representations or conjugate pairs of complex ones).

Needless to say, the conclusion reached with this example is unremarkable given that the representations considered were just $R = \mathbf{5}, \mathbf{10}, \mathbf{15}, \mathbf{35}$. The analysis gets more interesting when bigger representations are considered. If we do so, how unique is the standard fermion content $3\,(\mathbf{\bar{5}}) + 3\,(\mathbf{10})$ in $SU(5)$ GUTs? This is an important question which we address in this work, noting that the normalization of the SM hypercharge (usually given by a factor $\sqrt{3/5}$) depends on its answer.

Unfortunately, the type of simple analysis just presented is complicated by the fact that the SM group may be embedded in a GUT group G in more than one way. In particular, it is not known a priori what are the valid ways of combining the multiple $U(1)$ factors inside G in order to form the SM's $U(1)_Y$.

2.3   Different ways of embedding G_{SM} in a group $G \supset G_{SM}$

A systematic study of the different ways in which the SM chirality can be achieved in a GUT based on a group G must necessarily take into account the distinct ways in which G_{SM} can be embedded in G. (In fact, we only need to care about branching rules, so we shall be pragmatic and equate *different embeddings* to *different branching rules*.) Regardless of the actual symmetry breaking chain, we can can view it as being made of two symbolic steps:

$$
G \overset{(1)}{\rightarrow} SU(3)_C \times SU(2)_L \times U(1)^m \overset{(2)}{\rightarrow} SU(3)_C \times SU(2)_L \times U(1)_Y \ . \tag{9}
$$

In the first step, G is reduced to $SU(3)_C \times SU(2)_L$ times a maximal number of $U(1)$ factors, while in a second step this abelian part of the group is reduced to $U(1)_Y$.

There is only a finite number of ways in which the first symmetry breaking step can be carried out (see table 1). Indeed, with the information in [56, 57] one can break any semi-simple Lie algebra in a step-wise manner, $G \rightarrow G' \rightarrow G'' \rightarrow \cdots \rightarrow SU(3)_C \times SU(2)_L \times U(1)^m$, such that the algebra of each group in this sequence is a maximal subalgebra of the preceding one,[7] discarding none of the $U(1)$ factors at this stage.

---

[7]It is said that G' is a maximal subalgebra of G if there is no subalgebra G'' of G such that $G' \subset G'' \subset G$ (other than the trivial cases $G'' = G$ or G').

There are two important points concerning this first symmetry breaking step. The first one is that the number of $U(1)$ factors in the end result $(SU(3)_C \times SU(2)_L \times U(1)^m)$ will depend on the chosen sequence of maximal subalgebras, as the rank of the groups may shrink. The second point is that the step-wise procedure of breaking G into $SU(3)_C \times SU(2)_L \times U(1)^m$ subgroups will in general produce a large number of repetitions. In order to verify whether two embeddings are indeed different, it suffices to check that the branching rules for the fundamental representation of G are distinct, with the exception of the $SO(2n)$ groups which also require the branching rules of the spinor representation to be distinct [57–59].

To illustrate these two remarks, consider the case of $G = SO(10)$. Its maximal subgroups are $SU(5) \times U(1)$, $SU(4) \times SU(2) \times SU(2)$, $SO(8) \times U(1)$, $SP(4)$, $SO(9)$, $SU(2) \times SO(7)$, $SP(4) \times SP(4)$ although only the first two correspond to chiral embeddings.[8] In any case, consider for example the breaking chains (A) $SO(10) \to SU(5) \times U(1) \to SU(3) \times SU(2) \times U(1)^2$, (B) $SO(10) \to SU(4) \times SU(2) \times SU(2) \to SU(3) \times SU(2) \times U(1)^2$, and (C) $SO(10) \to SU(2) \times SO(7) \to SU(3) \times SU(2) \times U(1)$. The fundamental and spinor representations of $SO(10)$ branch as follows,

$$
\binom{\mathbf{10}}{\mathbf{16}} \xrightarrow{A,B} \binom{(\mathbf{3},\mathbf{1},0,2)+(\overline{\mathbf{3}},\mathbf{1},0,-2)+(\mathbf{1},\mathbf{2},1,0)+(\mathbf{1},\mathbf{2},-1,0)}{(\mathbf{3},\mathbf{2},0,-1)+(\overline{\mathbf{3}},\mathbf{1},1,1)+(\overline{\mathbf{3}},\mathbf{1},-1,1)+(\mathbf{1},\mathbf{2},0,3)+(\mathbf{1},\mathbf{1},-1,-3)+(\mathbf{1},\mathbf{1},1,-3)}
$$

$$
\text{or} \quad \binom{(\mathbf{3},\mathbf{1},0,2)+(\overline{\mathbf{3}},\mathbf{1},0,-2)+(\mathbf{1},\mathbf{2},1,0)+(\mathbf{1},\mathbf{2},-1,0)}{(\overline{\mathbf{3}},\mathbf{2},0,-1)+(\mathbf{3},\mathbf{1},1,1)+(\mathbf{3},\mathbf{1},-1,1)+(\mathbf{1},\mathbf{2},0,3)+(\mathbf{1},\mathbf{1},-1,-3)+(\mathbf{1},\mathbf{1},1,-3)}, \quad (10)
$$

$$
\binom{\mathbf{10}}{\mathbf{16}} \xrightarrow{C} \binom{(\mathbf{3},\mathbf{1},2)+(\overline{\mathbf{3}},\mathbf{1},-2)+(\mathbf{1},\mathbf{3},0)+(\mathbf{1},\mathbf{1},0)}{(\mathbf{3},\mathbf{2},-1)+(\overline{\mathbf{3}},\mathbf{2},1)+(\mathbf{1},\mathbf{2},3)+(\mathbf{1},\mathbf{2},-3)}, \quad (11)
$$

using unnormalized $U(1)$ charges (the separation of the two $U(1)$'s is irrelevant; the $U(1)^2$ group as a whole should be the same). Paths (A) and (B) lead to the same branching rules: there are two of them, which are related to one-another by conjugation of the color quantum number. This is a trivial variation which exists for all chiral embeddings, so we may refer to 'pairs of chiral embeddings', although in the next section we shall simply focus on one member of each such pair of embeddings.

So overall we could say that there are three possible embeddings of G_{SM} in $SO(10)$ (having in mind step one of the symmetry breaking only), including one pair which is chiral. In the second symmetry breaking step, the two $U(1)$'s can be combined in any way to form $U(1)_Y$ — at least from a purely group theoretical perspective. Therefore, the branching rules of the natural and spinor representations under the symmetry breaking $SO(10) \to G_{SM}$ are (the normalization of the $U(1)$'s is irrelevant at the moment)

$$
\binom{\mathbf{10}}{\mathbf{16}} \to \binom{(\mathbf{3},\mathbf{1},2\beta)+(\overline{\mathbf{3}},\mathbf{1},-2\beta)+(\mathbf{1},\mathbf{2},\alpha)+(\mathbf{1},\mathbf{2},-\alpha)}{(\mathbf{3},\mathbf{2},-\beta)+(\overline{\mathbf{3}},\mathbf{1},\alpha+\beta)+(\overline{\mathbf{3}},\mathbf{1},-\alpha+\beta)+(\mathbf{1},\mathbf{2},3\beta)+(\mathbf{1},\mathbf{1},-\alpha-3\beta)+(\mathbf{1},\mathbf{1},\alpha-3\beta)}, \quad (12)
$$

$$
\to \binom{(\mathbf{3},\mathbf{1},2\beta)+(\overline{\mathbf{3}},\mathbf{1},-2\beta)+(\mathbf{1},\mathbf{2},\alpha)+(\mathbf{1},\mathbf{2},-\alpha)}{(\overline{\mathbf{3}},\mathbf{2},-\beta)+(\mathbf{3},\mathbf{1},\alpha+\beta)+(\mathbf{3},\mathbf{1},-\alpha+\beta)+(\mathbf{1},\mathbf{2},3\beta)+(\mathbf{1},\mathbf{1},-\alpha-3\beta)+(\mathbf{1},\mathbf{1},\alpha-3\beta)}, \quad (13)
$$

$$
\to \binom{(\mathbf{3},\mathbf{1},2\gamma)+(\overline{\mathbf{3}},\mathbf{1},-2\gamma)+(\mathbf{1},\mathbf{3},0)+(\mathbf{1},\mathbf{1},0)}{(\mathbf{3},\mathbf{2},-\gamma)+(\overline{\mathbf{3}},\mathbf{2},\gamma)+(\mathbf{1},\mathbf{2},3\gamma)+(\mathbf{1},\mathbf{2},-3\gamma)}, \quad (14)
$$

for some α,β,γ factors. It is worth stressing that the last case is not a special case of either of the first two. Also, while we did only consider three maximal subgroups of $SO(10)$ (chains A, B,

---

[8]By chiral embeddings we are referring to those cases $G \supset G'$ where the representations of G do not break only into real representations and conjugate pairs of complex representations of G'.

Group	# Embeddings	Group	# Embeddings	Group	# Embeddings
$SU(5)$	2 (1)	$SU(9)$	40 (19)	$SO(10)$	3 (1)
$SU(6)$	4 (2)	$SU(10)$	65 (30)	$SO(14)$	15 (2)
$SU(7)$	10 (5)	$SU(11)$	108 (50)	$SO(18)$	62 (5)
$SU(8)$	21 (10)	$SU(12)$	187 (86)	E_6	12 (5)

Table 1: Number of distinct embeddings of subgroups of the type $H = SU(3) \times SU(2) \times U(1)^{m \geq 1}$ in various simple groups G which have complex representations (the number of pairs of chiral embeddings is indicated in parenthesis). The counting includes only those cases where H is not contained in a bigger subgroup $H' = SU(3) \times SU(2) \times U(1)^{m'}$ with $m' > m$. Note that, for a given G, the number of $U(1)$ factors of the H subgroups in this condition does not need to be constant (consider the $SO(10)$ example in the main text).

C) it can be checked that all others cases lead to the embedding of the form (14) — doing so by hand is tedious (and even more so for bigger groups), and for that reason the `Susyno` program was used to automatically check for these repetitions.

Introducing physical considerations, not all embeddings of forms in equations (12)–(14) can be used to embed the SM in an $SO(10)$ GUT. The one in equation (14) would lead to a vector theory, so it can be excluded. As for the chiral embedding described by equation (12) ((13) is similar), if we are to obtain the SM fermions from the **16**, then α and β which describe the composition of $U(1)_Y$ in terms of the two $U(1)$'s of $SU(3)_C \times SU(2)_L \times U(1)^2$ must take specific values: $\beta = -1/6$ and $\alpha = \pm 1/2$. This leads to the standard G_{SM} embedding in $SO(10)$, yet, as it should be clear at this point, $U(1)_Y$ might conceivably be another combinations of the two $U(1)$'s if the SM fermions are placed in other $SO(10)$ representations. The normalization of the hypercharge may depend on this placement.

Unfortunately, with larger GUT groups the situation becomes even more complicated if we do not make assumptions about the GUT representations where the SM fields are embedded, since there are more $U(1)$ factors to consider. This situation is not insurmountable, but it does require adaptations to the analysis suggested in section 2, since it cannot be carried out unless we know the hypercharge y of the representations (all that is known is that $y = \sum_i \alpha_i y_i$ where y_i are the charges under $U(1)^m$, and the α_i are to be determined).

There seems to be no easy way to tackle this issue, and as a consequence the scans over GUT representations were smaller for the bigger groups. One way is to just look at the first two quantum numbers and try to match in all possible ways the SM representations with the ones of $SU(3)_C \times SU(2)_L \times U(1)^m$ obtained from some list of GUT representations — this is the intuitive approach which works very well for the **16** of $SO(10)$. Whenever this approach proved to be too demanding computationally, we used instead a modification of the analysis in section 2, where the c_i (encoding the unknown combination of the GUT representations) and the α_i (encoding the unknown combination of the $U(1)$'s) are found simultaneously as the solution of complicated equations where the α_i do not appear linearly.

3 GUTs with the SM chirality

The GUTs we wish to consider should be based on a group with complex representations, otherwise they would not give rise to an effective chiral theory. The simple Lie groups with this property and which contain G_{SM} as a subgroup are $SU(N \geq 5)$, $SO(4N'+2)$ for $N' \geq 2$, and E_6. As such, in the following we shall analyze the fermion sector of GUTs based on one of these simple groups, investigating also models with the $SU(3) \times SU(3)$ gauge group (possibly with an extra $U(1)$).

3.1 $SU(5)$

We shall start by assuming that the hypercharge of the SM particles are normalized in the usual way: $y(e^c) = \sqrt{3/5}$ for example. All 2048 (pairs of) complex representations of $SU(5)$ with size no larger than 1 million were decomposed with the `Susyno` program. A total of 29037 SM complex representations appear in these decompositions, therefore one obtains the system of linear equations in (3) where $\chi(SM)$ is a 29037-dimensional vector (with null components everywhere except for five ± 3 entries), c is 2048-dimensional vector of unknown coefficients c_i (describing the number of copies of each $SU(5)$ representation), and $\mathcal{M}$ is a 29037 by 2048 matrix. Both $\chi(SM)$ and $\mathcal{M}$ are known, so it is possible to solve for the vector c as explained in section 2. It turns out that the matrix $\mathcal{M}$ has a trivial nullspace. As such, there is a single solution to equation (3), and it corresponds to the standard, well known one: three copies of $\mathbf{\bar{5}} + \mathbf{10}$.[9,10]

This simple but effective analysis shows that the $\mathbf{\bar{5}}$ and $\mathbf{10}$ fermion representations of $SU(5)$ are extremely special. However, we did assume the standard GUT hypercharge normalization factor $n_{strd} \equiv \sqrt{\frac{3}{5}}$. The usual justification for this factor is tied to the identify the components of $\mathbf{\bar{5}}$ with those of the SM representations $L = \left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}n\right)$ and $d^c = \left(\mathbf{\bar{3}}, \mathbf{1}, -\frac{1}{3}n\right)$, as discussed in the introduction of this document. Since we do not want to assume that the SM fermions are in the $\mathbf{\bar{5}}$ and $\mathbf{10}$ representations necessarily, we must admit other values for n. Which other values can it take? Looking for $SU(5)$ representations where the left-handed quarks might be embedded, we conclude that the G_{SM} representations of the form $\left(\mathbf{3}, \mathbf{2}, \frac{1}{6}n\right)$ must have $n = (1 + 6k)n_{strd}$ for some integer k. This can be easily shown analytically with the weight projection method (the reader may wish to see for example [60]) and indeed, probing the $SU(5)$ representations of size smaller or equal to a million, one encounters all the SM representations fermions $\{Q, u^c, d^c, L, e^c\}$ with the hypercharge normalizations $n/n_{strd} = -17, -11, -5, 1, 7, 13, 19$. Crucially, each of these

---

[9]We recall here that, since chiral embedding come in pairs (see the example in equations (12)–(13)), three copies of $\mathbf{5} + \mathbf{\overline{10}}$ would work as well. This is nevertheless an obvious/trivial variation which we shall ignore in the remainder of this work.

[10]Reference [37] claims the same thing: the hope of recovering the SM with more exotic $SU(5)$ representations is not possible. The author of [37] supports this assertion with the fact that the rank of $SU(5)$ and G_{SM} are the same, however it is not clear exactly how this fact can be used to proof the statement that the only non-trivial solution with the SM chirality are three copies of $\mathbf{\bar{5}} + \mathbf{10}$.

The fact that a group and a subgroup have the same rank implies that the corresponding weight projection matrix (see [60] for details) is invertible, therefore there is (at most) one $SU(5)$ field content which can break into a given combination of G_{SM} representations. However, since we do not mind adding vector fermions to the SM, there is an infinite set of combinations of G_{SM} representations which are acceptable. Out of these, the computer scan done in the current analysis shows that, using representations with size no larger than 1 million, the only valid combination of SM fermions which can come from an $SU(5)$ theory is the one associated to three copies of $\mathbf{\bar{5}} + \mathbf{10}$ (and trivial variations of it).

Furthermore, as explained in the main text, one can conceivably take a SM hypercharge normalization distinct from the canonical one, which further complicates the use of the above group rank argument to rule out non-trivial $SU(5)$ solutions with the SM chirality.

choices yields a different chirality vector $\chi(SM)$ in equation (3), and it turns out that there is no solutions except for the standard hypercharge normalization ($n/n_{strd} = 1$).

3.2 $SO(10)$

We now repeat for $SO(10)$ the same analysis which was done for $SU(5)$. According to the discussion of subsection 2.3, there is only one chiral embedding of $SU(3) \times SU(2) \times U(1)^m$ in $SO(10)$ which is the one in equation (12) (equation (13) is similar) yet, since $m = 2$, we do have to probe all possible values of α and β which encode the relation between $U(1)_Y$ and the two $U(1)$'s which are contained in $SO(10)$. A list of possibilities can be computed by breaking all $SO(10)$ representations up to some size into those of $G' = SU(3) \times SU(2) \times U(1)^2$ and then start assigning the SM fermions (at least two) to any G' representations with the correct $SU(3) \times SU(2)$ quantum numbers. Such procedure should be compared with the one used for $SU(5)$ (see above) where, instead of two, there was only one unknown parameter (a normalization factor).

This method produces an exhaustive list of (α,β) values for which $SO(10)$ breaks into $SU(3) \times SU(2) \times U(1)$ in such a way that all the SM fermions (Q, u^c, d^c, L, e^c) can be found inside some $SO(10)$ representation, with the correct ratio of hypercharges. For each value of (α,β) it is then possible to compute an $\mathcal{M}$ matrix and solve equation (3) using it. This was done for all $SO(10)$ representations up to size 1 million, and two conclusions became clear.

Firstly, concerning the embedding of G_{SM} in $SO(10)$ and the normalization of the hypercharge, it turns out that equation (3) admits solutions only for the values $(\alpha,\beta) = \left(\pm\frac{1}{2}, -\frac{1}{6}\right)$, corresponding to

$$\mathbf{16} \rightarrow Q + d^c + u^c + L + e^c + N^c \tag{15}$$

with the standard hypercharge normalization. It is worth mentioning that even though $SO(10)$ contains $SU(5)$, which in turn contains $G_{SM} = SU(3) \times SU(2) \times U(1)$ as a subgroup, there are more such G_{SM} subgroups in $SO(10)$, and that is why the branching rules may vary depending on which one is picked. With this in mind, it is interesting to note that the only branching rule which can reproduce the SM chirality (shown above) matches the one of the G_{SM} subgroup found inside $SU(5)$. We point this feature now because in the remainder of this work we shall see that this is not a specific feature of $SO(10)$: for all simple GUT groups which were tested, in order to recover the SM chirality, the embedding $G_{SM} \subset G_{GUT}$ must be such that the branching rules match those for a G_{SM} inside a particular $SU(5)$ subgroup of G_{GUT}.[11]

The second conclusion concerns the valid combinations of $SO(10)$ fermion fields: unlike $SU(5)$, we do find non-standard solutions with the SM chirality, although they involve very large representations. To be precise, referring to the framework set forth in section 2, we have the solution $3(\mathbf{16})$ to which we can add an arbitrary number of the non-trivial combinations of $SO(10)$ repre-

---

[11]The branching rules are the same, but this does not mean that $G_{SM} \subset SU(5) \subset G_{GUT}$ necessarily. Consider the following counter example: $G_{SM} \subset SU(3) \times SU(2) \times U(1)_B \times U(1)_A \subset SU(5) \times U(1)_A \subset SO(10)$ and it is possible to form the SM hypercharge group just from the $U(1)_A$ inside $SU(5)$ or from a combination of $U(1)_A$ and $U(1)_B$ (known as the flipped $SU(5)$ scenario [61]). In either case, the branching rules $SO(10) \rightarrow G_{SM}$ are the same. Something similar happens with the E_6 group, as the hypercharge group of the SM can be made from three different combinations of the three available $U(1)$'s — see for instance [62].

sentations which have no chirality (associated to the n_i vectors of equation (4)):

$$n_1\colon \ -\mathbf{126}-\mathbf{144}-\mathbf{1200}+\mathbf{2772}+\mathbf{3696}-\mathbf{4950}-\mathbf{6930'}+\mathbf{7920}+\mathbf{8064}+\mathbf{11088}-\mathbf{15120}$$
$$-\mathbf{17280}+\mathbf{17325}+\mathbf{30800}-\mathbf{34992}-\mathbf{38016}+\mathbf{48114}+\mathbf{49280}, \tag{16}$$
$$n_2\colon \ -\mathbf{16}+\mathbf{126}-\mathbf{560}+\mathbf{8064}+\mathbf{20592}-\mathbf{20790}+\mathbf{23760}+\mathbf{25200}-\mathbf{29568}-\mathbf{48114}-\mathbf{50050}$$
$$-\mathbf{90090}-\mathbf{102960}-\mathbf{124800}-\mathbf{128700}-\mathbf{144144}+\mathbf{164736}+\mathbf{196560}-\mathbf{199017}. \tag{17}$$

We recall here that $-R$ should be interpreted whenever necessary as $\overline{R}$. The field combinations n_1 and n_2 are just two out of many with no chirality; there are more such (independent) mixtures of fields, involving even bigger $SO(10)$ representations, which we will not write down here.[12] We simply point out here that the n_1 combination does not involve the $\mathbf{16}$, so any solution of the form $3(\mathbf{16}) + kn_1$ for some $k \in \mathbb{Z}$ will still involve 3 copies of the spinor representation. However, n_2 is a linear combination of representations which does contain the $\mathbf{16}$: a GUT theory with the fermion content $3(\mathbf{16}) + 3n_2$ would have no fermions in the spinor representation and still its chirality would be correct. Therefore, it is possible to build a GUT model based on the $SO(10)$ group without spinors, even though its matter content would need to be extremely large[13] and, on the other hand, the flavor problem would persist since 3 copies (or more) of each fermion representation would still be needed.

3.3 E_6

The group E_6 has 38 pairs of complex representations with size at most 1 million. There are a total of 12 distinct ways of embedding $SU(3) \times SU(2) \times U(1)^m$ in E_6, which includes 5 pairs of chiral embeddings. For each of these, we have allowed $U(1)_Y$ to be any combination of the m $U(1)$'s. Remarkably, once this variety of representations and embeddings is fully explored, it turns out that there is a unique solution with the correct chirality. In other words, there is both a unique embedding and a unique fermion field configuration which yield the SM chirality: it is 3 copies of the $\mathbf{27}$ representation with the embedding

$$\mathbf{27} \to Q + u^c + 2d^c + 2L + e^c + (d^c)^* + L^* + 2N^c. \tag{18}$$

It is well known that this branching rule matches the one for the G_{SM} subgroup of the $SO(10)$ which is inside E_6, with the $\mathbf{27}$ breaking into $\mathbf{1} + \mathbf{10} + \mathbf{16}$ of $SO(10)$, so it is clear that the hypercharge normalization factor must be the standard one ($\sqrt{3/5}$).

3.4 $SU(N > 5)$

Subsection 2.3 contains a discussion on how to obtain all the embeddings of a subgroup H in some group G (with $H = SU(3) \times SU(2) \times U(1)^m$ in mind). It involves probing all sequences of maximal subgroups $G \to G' \to G'' \to \cdots \to H$, which can be a very time consuming process that does not provide much insight on the direct relationship between H and G. It is therefore important to realize that there is a simpler way to list all these embeddings when G is a special unitary group.

For a given $G = SU(N)$, we start by picking all possible combinations of the irreducible representations of H in order to obtain a complete list of the N-dimensional (potentially reducible)

---

[12]A total of 6 independent $\boldsymbol{n_i}$ combinations with no chirality exist involving only $SO(10)$ representations with size smaller or equal to 1 million.

[13]Quantum gravity effects are expected to become relevant in such a scenario [63, 64].

representations of the subgroup H. Since the representation matrices must be unitary (otherwise the kinetic terms would not be gauge invariant), it is obvious that there must be an embedding under which the fundamental of G breaks into any of the N-dimensional (unitary) representations of H. One only needs to ensure that none of the generators of these H representations has null/trivial generators T_a as this would imply that $\text{Tr}\left(T_a^2\right) = 0$.

For example, consider the embedding of $SU(5)$ in $SU(6)$. There are three inequivalent reducible 6-dimensional representations of $SU(5)$: $\mathbf{1}+\mathbf{1}+\mathbf{1}+\mathbf{1}+\mathbf{1}+\mathbf{1}$, $\mathbf{1}+\mathbf{5}$, and $\mathbf{1}+\overline{\mathbf{5}}$. However, the generators of the algebra of the first representation are null, so there are just two embeddings of $SU(5)$ in $SU(6)$: $\mathbf{6} \to \mathbf{1}+\mathbf{5}$ and $\mathbf{6} \to \mathbf{1}+\overline{\mathbf{5}}$. Next, take the $SU(5) \to G_{SM}$ example. The fundamental representation of $SU(5)$ must break into at least one non-trivial representation of $SU(3)$, $SU(2)$, and the hypercharge group; otherwise one or more of the generators of the algebra of G_{SM} would be null. So there are just two possibilities: $\mathbf{5} \to d^c + L$ or $\mathbf{5} \to (d^c)^* + L^*$;[14] cases such as $\mathbf{5} \to (\mathbf{1},\mathbf{5},0)$ or $\mathbf{5} \to (\mathbf{1},\mathbf{2},y) + (\mathbf{1},\mathbf{2},y') + (\mathbf{1},\mathbf{1},-2y - 2y')$ do not exist since they would require that one or more subgroup algebra generators are the null matrices.

Based on these comments, all branching rules of $SU(N)$ into $SU(3) \times SU(2) \times U(1)^m$ can be found with the following algorithm:

1. Start by listing all irreducible representations of $SU(3) \times SU(2)$ up to size N.

2. Adding together the irreducible representations of $SU(3) \times SU(2)$ in all possible combinations, we get a complete list of the N-dimensional representations of this group.

3. Exclude those N-dimensional representations of $SU(3) \times SU(2)$ which transform trivially either under $SU(3)$ or $SU(2)$. Furthermore, discard the N-dimensional irreducible representations of $SU(3) \times SU(2)$ since this would force the charge operator of any additional $U(1)$ factor group to be null.

4. Finally, one must deal with the $U(1)^m$ charges of these N-dimensional representations R of $SU(3) \times SU(2)$, which must be associated to traceless matrices. Let us assume that the dimensions of the n irreducible representations of $SU(3) \times SU(2)$ in a given R and their charges under some $U(1)$ are $d \equiv (d_1, d_2, \cdots, d_n)$ and $q \equiv (q_1, q_2, \cdots, q_n)$ respectively. Any q vector is fine as long as $d \cdot q = 0$, so immediately we see that the maximal number of $U(1)$'s, which we have been calling m for some time, is exactly $n - 1$ since this is the maximum number of independent vectors orthogonal to d. Breaking $SU(3) \times SU(2) \times U(1)^m$ further down to the SM group requires, as previously explained, forming the hypercharge group from a linear combination (any) of the $m = n - 1$ $U(1)$'s obtained in this way.

Among all these embeddings, there is one which is particularly interesting (it exists for all $N \geq 5$): under it, the anti-fundamental representation $\overline{F}$ of $SU(N)$ breaks into the SM representations $d^c + L + (N - 5)N^c$ (the number N is not to be confused with N^c, which stands for the $(\mathbf{1},\mathbf{1},0)$ representation of G_{SM}). It is easy to check that the irreducible representation of size $N(N - 1)/2$ obtained from the anti-symmetric product of two F's (let us call it K here) decomposes as

$$K \to Q + u^c + e^c + (N - 5)(d^c)^* + (N - 5)L^* + \frac{(N - 5)(N - 6)}{2}N^c, \qquad (19)$$

so the combination $-3(N - 4)F + 3K$ has the correct chirality. In fact, for N between 6 and 12, it turns out that this embedding is the only one which can reproduce the SM chirality; this is

the conclusion of a scan over all combinations of fermion representations up to some limiting size (at least 1000), considering as well in the process all possible ways of forming the hypercharge operator from the available $U(1)$'s. In appendix there are various tables with the fermion content allowed by the chirality constraint in these $SU(N)$ GUTs (the combination $-3(N-4)F + 3K$ is not unique).

As the size of the $SU(N)$ groups increase, it becomes harder to manage the problem of testing all the possible ways of forming the SM hypercharge group. Yet, we should recall that for all groups G_{GUT} where such an analysis was carried out — $SO(10)$, E_6 or $SU(12 \geq n \geq 5)$ — it turned out that there was at most one branching rule that worked for each G_{GUT}, and it matched the one obtained by considering the G_{SM} inside one of the $SU(5)$ subgroups of G_{GUT}. So, with hindsight, we could have simply looked for the possible ways of embedding $SU(5)$ in the unification groups, and then break $SU(5)$ to G_{SM}; no important embedding of G_{SM} in G_{GUT} would have been missed. Or, better yet, since an $SU(5)$ theory will only reproduce the SM's chirality with the field combination $-3(\mathbf{5}) + 3(\mathbf{10})$, we would just need to see if this $SU(5)$ chirality[15] is attainable with a given unification group.

Inspired by this observation, we have analyzed the different ways of embedding $SU(5)$ in $SU(N \geq 13)$. From an $SU(5)$ perspective, the embedding in equation (19) corresponds to the following branching rules,

$$F \rightarrow \mathbf{5} + (N-5)\mathbf{1}, \tag{20}$$

$$K \rightarrow \mathbf{10} + (N-5)\mathbf{5} + \frac{(N-5)(N-6)}{2}\mathbf{1}, \tag{21}$$

where F is the fundamental representation of $SU(N)$, and K is the one in the anti-symmetric part of the product $F \times F$ as before. The field combination $-3(N-4)F + 3K$ works from the chirality point of view, and one might add that it does not lead to gauge anomalies [65, 66]: indeed, for this embedding, the chirality condition implies that the $SU(N)$ gauge anomalies cancel, so all the configurations one can build from the tables in appendix are fine.

Going through the $SU(N)$ family of groups, we have checked that the embedding in equations (20) and (21) is the only one that works. Until $SU(15)$ is reached. For this and bigger groups, two new remarkable $SU(5)$ embeddings become possible. The first one corresponds to

$$F \rightarrow 2(\mathbf{5}) + \overline{\mathbf{5}} + (N-15)\mathbf{1}, \tag{22}$$

which means that $K \equiv (F \times F)_{\text{Ant.}}$ and $L \equiv (F \times F)_{\text{Sym.}}$ will break as follows:

$$K \rightarrow 2(\mathbf{24}) + \mathbf{15} + 3(\mathbf{10}) + \overline{\mathbf{10}} + 2(N-15)\mathbf{5} + (N-15)\overline{\mathbf{5}} + \left(\frac{244 - 31N + N^2}{2}\right)\mathbf{1}, \tag{23}$$

$$L \rightarrow 2(\mathbf{24}) + 3(\mathbf{15}) + \overline{\mathbf{15}} + \mathbf{10} + 2(N-15)\mathbf{5} + (N-15)\overline{\mathbf{5}} + \left(\frac{214 - 29N + N^2}{2}\right)\mathbf{1}. \tag{24}$$

As such, the field combination $-(N-12)F + 2K - L$ will have the correct chirality and it is anomaly free. This is not the unique configuration that works for a given $N \geq 15$; there are non-trivial combinations of the representations $SU(N)$ which have no chirality, just as in the embedding in equations (20)–(21). Nevertheless, we shall not print them here (they are fairly elaborate).

---

[15] By '$SU(5)$ chirality' we are referring to the concept of chirality as discussed in section (2), but based on the analysis of $SU(5)$ representations instead of those of $SU(3) \times SU(2) \times U(1)$.

13

The second new noteworthy embedding encountered for $N \geq 15$ is the one under which the anti-fundamental representation of $SU(N)$ breaking into exactly one SM family plus singlets:

$$\overline{F} \to \overline{\mathbf{5}} + \mathbf{10} + (N-15)\mathbf{1}. \tag{25}$$

Obviously, the field configuration $-3F$ will have the correct chirality. However, the reader will immediately notice that it leads to a theory with an $SU(N)^3$ anomaly, which should be seen as a warning that the chirality condition does not always imply absence of anomalies (see [67]). For this embedding, there are non-trivial combinations n_i of the $SU(N)$ representations with no chirality which contribute to the $SU(N)^3$ anomaly, so one could have hoped that for some c_i, the combination $-3F + \sum_i c_i n_i$ would be anomaly free. Unfortunately, no such c_i exist for the $SU(N)$ groups tested. How does one judge such models then? They can be seen as non-renormalizable effective models [68]; or maybe there is some way to cancel the anomaly (perhaps string inspired [69]). In any case, we shall not worry about these anomalies — we we simply assume that these models (or variations of them) can be made part of a consistent quantum field theory. In this spirit, we shall say a few more words about the curious embedding in equation (25). Under it,

$$K \to \mathbf{45} + \overline{\mathbf{45}} + (N-15)\mathbf{5} + \overline{\mathbf{5}} + (N-15)\overline{\mathbf{10}} + \mathbf{10} + \frac{(N-15)(N-16)}{2}\mathbf{1}, \tag{26}$$

which means that the $\overline{K}$ representation of $SU(16+N')$ contains precisely N' SM families plus vector particles. In the particular case of $SU(19)$, looking through its $SO(10)$ subgroup makes things even more clear:

$$\overline{F} \to \mathbf{16} + 3(\mathbf{1}), \tag{27}$$
$$\overline{K} \to 3(\mathbf{16}) + \mathbf{120} + 3(\mathbf{1}). \tag{28}$$

In other words, it is possible to unify the three SM families in $\overline{K}$ of $SU(19)$ (and more generally N' families in $SU(16+N')$) although the model will have gauge anomalies.

The three embeddings of $SU(5)$ in $SU(N)$ which were presented above are the only ones that work for $N \leq 20$ (representations of size up to one million were considered). For $N > 20$ one will certainly find new embeddings: for example, starting with the $SU(45)$ group it is certainly possible to fit the three families (plus vector particles) in the fundamental representation (although leading to gauge anomalies once again). Yet, with increasingly big unification groups the possibilities of embedding $SU(3) \times SU(2) \times U(1)^m$ in it grow in a seemingly exponential way (see figure 1). For this reason, one might argue that models based on very large gauge groups are not as attractive as those based on smaller ones: they contain many subgroups, therefore a significant tuning of the scalar sector parameters would likely be needed in order to have the correct symmetry breaking.

3.5 $SO(10+4N)$ with $N > 0$

In four dimensions and without confining interactions, is it possible to embed the SM in a theory based on a gauge group of the family $SO(10+4N)$ for $N > 0$? These groups do have complex representations and, furthermore, it is possible to chirally embed the SM group in them (although such embeddings are not plentiful — see table 1). However, despite these promising features, after performing computer scans, it seems impossible to reproduce the SM chirality in $SO(14)$ and $SO(18)$ — at least not with fermions representations with a size smaller than 2 million. Perhaps it has to do with the fact that, unlike $SO(10)$, it is not possible to chirally embed $SU(5)$
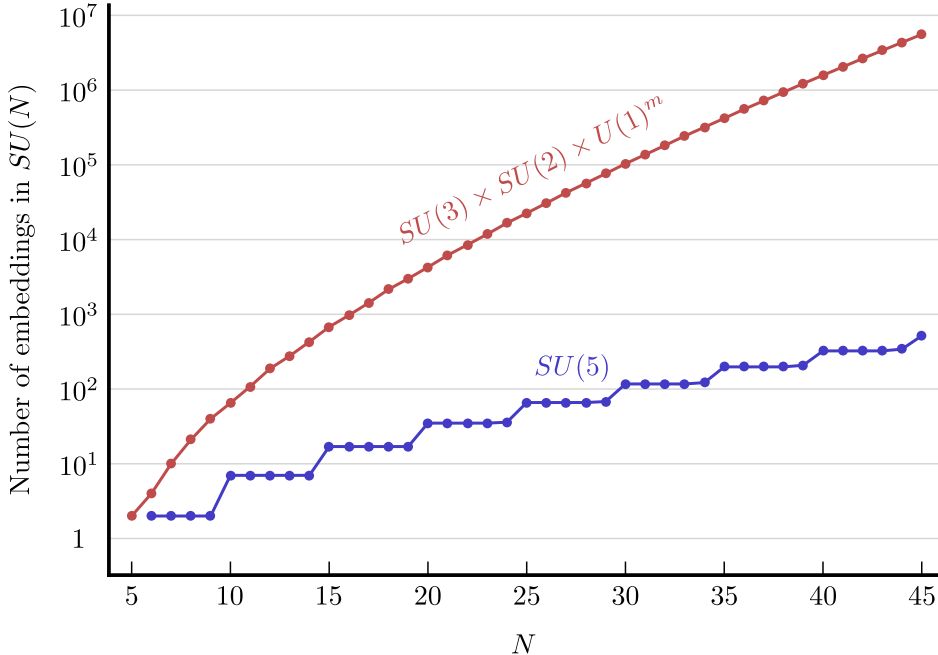
Figure 1: Number of distinct ways of embedding $SU(3) \times SU(2) \times U(1)^{m \geq 1}$ and $SU(5)$ in $SU(N)$ groups. The counting includes only those cases where $SU(3) \times SU(2) \times U(1)^m$ is embedded in such a way that it is not a subgroup of a bigger $SU(3) \times SU(2) \times U(1)^{m+1}$ also contained in $SU(N)$. In this sense, m must be maximal, yet this does not mean that its value is fixed for a given N, as explained in the main text.

in $SO(14)$ nor $SO(18)$. Interestingly, this last statement no longer holds for bigger groups of the $SO(10 + 4N)$ family.

As with the special unitary groups, it becomes hard to analyze all possible ways of embedding G_{SM} so we shall focus instead on the $SU(5)$ subgroups of $SO(10+4N)$ in the following.[16] Table 2 contains a curious piece of information: while there are two ways of embedding $SU(5)$ in $SO(10)$ — a pair of chiral embeddings — there is just one for $SO(14)$ and $SO(18)$, and under it the complex representations of these groups break into a mixture of real and pairs of complex conjugated representations of $SU(5)$. The number of embeddings is bigger for $SO(22)$ and $SO(26)$ but they too are all vector embeddings. However, in this case persistence pays off as it is possible to chirally embed $SU(5)$ in $SO(30)$.

The trouble with the $SO(10 + 4N)$ group family is that the complex representations become exponentially large with N: the spinor representation — which is the smallest one — has 4^{2+N} components. In the case of $SO(30)$, we have considered the complex representations smaller than

---

[16]The algorithm mentioned previously to find quick and easily the embeddings of a subgroup H in the special unitary groups cannot be readily adapted to the special orthogonal groups. The method consisted essentially in building all the N-dimensional representations of some $H \subset SU(N)$. With $SO(N)$, one would have to consider only the strictly real N-dimensional representations R of $H \subset SO(N)$ (i.e., exclude the pseudo-real and complex ones): for every such R there in an embedding under which the fundamental representation F of $SO(N)$ breaks into R.

But the branching rules of the spinor representation S of $SO(N)$ would be missing. One can only speculate that perhaps one can find these, up to conjugation, from the branching rules of the fundamental representation F, since F and S are related. For example, using the shorthand notation $X^{\{m\}}$ and $X^{[m]}$ to denote the completely symmetry and anti-symmetric parts of the product of m copies of a representation X, there is the relation $F^{[N]} + F^{[N-4]} + F^{[N-8]} + \cdots = S^{\{2\}} + \overline{S}^{\{2\}}$.

Groups	# Embeddings	Groups	# Embeddings
$SO(10)$	2 (1)	$SO(26)$	4 (0)
$SO(14)$, $SO(18)$	1 (0)	$SO(30)$	10 (3)
$SO(22)$	3 (0)		

Table 2: Number of district ways of embedding $SU(5)$ in the groups $SO(10+4N)$ with $0 \leq N \leq 5$. Only some of these correspond to chiral embeddings (the integers in parenthesis indicate the number of pairs of chiral embeddings in each case).

the **132562944** (there are only 7) and all the chiral embeddings of $SU(5)$ in $SO(30)$; it turns out that it is impossible to obtain three families of $\overline{\mathbf{5}} + \mathbf{10}$ plus vector particles.

As for bigger groups in the $SO(10 + 4N)$ family, they were not tested so one can only speculate about the possibility of embedding the three SM families in such models. The complex representations will be even bigger than those of $SO(30)$, each of them potentially breaking into thousands of distinct complex $SU(5)$ representations, so it seems unlikely that one could match all these sub-representations in pairs $\left(R, \overline{R}\right)$ leaving only a small excess of $\overline{\mathbf{5}}$'s and $\mathbf{10}$'s over $\mathbf{5}$'s and $\overline{\mathbf{10}}$'s. Even if it is possible, it would almost inevitably require millions of new vector particle components.[17]

3.6 $SU(3) \times SU(3)$ and $SU(3) \times SU(3) \times U(1)$

We shall consider $SU(3) \times SU(3)$ — even though it is not a simple group — because it contains G_{SM} as a subgroup, it has a minimal number of diagonal generators, and it is a group with complex representations. Besides $SU(5)$, the only other semi-simple group with these properties is $SU(3) \times SU(2) \times SU(2)$, which clearly will not yield the correct chirality as one would always have pairs of representations with opposite hypercharges. Are there $SU(3) \times SU(3)$ models with the correct chirality? Models with an extra $U(1)$ can successfully embed the SM, achieving the correct chirality, but without this extra abelian factor we shall see that this is not possible.

One of the two factors is the color group $SU(3)_C$ which must not be broken, while the other — let us call it henceforth $SU(3)_L$ — has to break into the EW group $SU(2)_L \times U(1)_Y$. We recall here that the representations of $SU(3)_L$ can be uniquely labeled by two non-negative integers $\{a, b\}$ (the Dynkin coefficients) which in terms of Young tableaux can be identified with the representation with a columns with a single row and b columns with two rows. The conjugate representation of $\{a, b\}$ is $\{b, a\}$, and their dimension is $\frac{1}{2}(a + 1)(b + 1)(a + b + 2)$ — for example, $\{0, 0\}$, $\{1, 0\}$ and $\{0, 1\}$ are the singlet, triplet (by convention), and anti-triplet representations, respectively. On the other hand, the $SU(2)_L \times U(1)_Y$ representations can be labeled by their spin j and (unnormalized) hypercharge y.

Crucially, it is possible to derive the branching rules of a generic representation of $SU(3)_L$ into those of $SU(2)_L \times U(1)_Y$: $\{a, b\}$ decomposes into representations with hypercharge $y = a - b + 3n$ ($n \in \mathbb{Z}, -a \leq n \leq b$) whose $SU(2)_L$ spins are $j = \frac{a+b-n}{2}, \frac{a+b-n}{2} - 1, \frac{a+b-n}{2} - 2, \cdots, \left|\frac{b-a-n}{2}\right|$ (for $n \geq 0$) and $j = \frac{a+b+n}{2}, \frac{a+b+n}{2} - 1, \frac{a+b+n}{2} - 2, \cdots, \left|\frac{b-a+n}{2}\right|$ (for $n < 0$).[18] The state with the biggest spin, $(j, y) = \left(\frac{a+b}{2}, a - b\right)$, is a complex $SU(2)_L \times U(1)_Y$ representation whenever the

---

[17]To have a feeling of the huge numbers involved, the three smallest complex representations of $SO(30)$ have sizes 16384, 475136, 6635520; in $SO(34)$ these numbers increase fourfold or more to 65536, 2162688, 34537472.

[18]For example, $\{3, 2\}$ decomposes into $(j, y) = \left(\frac{3}{2}, 7\right)$, $(1 \oplus 2, 4)$, $\left(\frac{1}{2} \oplus \frac{3}{2} \oplus \frac{5}{2}, 1\right)$, $(0 \oplus 1 \oplus 2, -2)$, $\left(\frac{1}{2} \oplus \frac{3}{2}, -5\right)$, and $(1, -8)$.

16

$SU(3)_L$ representation from which it originates is complex as well ($a \neq b$). On the other hand, if $a + b > 1$ this state will be an $SU(2)$ triplet or higher-dimensional representation which must have a vector mass (since it is not seen at low energies). As such, in a model where there is an $\{a,b\}$ complex representation of $SU(3)_L$ with $a + b > 1$, one needs to ensure the presence of at least another $SU(3)_L$ representation which contains the state $(j,y) = \left(\frac{a+b}{2}, b - a\right)$. Obviously this can be achieved with the representation $\{a,b\}^* = \{b,a\}$, but having both $\{a,b\}$ and its conjugate would not affect a model's chirality (see section 2) and for that reason we have stated previously that such configurations should be excluded from the analysis. Therefore, in order to obtain the $SU(2)_L \times U(1)_Y$ representation $(j,y) = \left(\frac{a+b}{2}, b - a\right)$ one must have an $SU(3)_L$ complex representation $\{a',b'\}$ with $a' + b' > a + b$. But such a $\{a',b'\}$ multiplet would decompose into complex states with spin even higher than $\frac{a+b}{2}$, presenting a renewed problem. And thus, with this circular argument it is shown that by using any $SU(3)_L$ complex fermion representation other that $\{0,0\}$ (the singlet), $\{1,0\}$ (the triplet) and $\{0,1\}$ (the anti-triplet) it will be impossible to avoid having fermions in triplet or higher dimensional representations of $SU(2)_L$ with no vector mass. And yet, with just $SU(3)_L$ singlets, triplets and anti-triplet it is not possible to reproduce the EW representations of the SM fermions, therefore one cannot embed the SM in an $SU(3)_C \times SU(3)_L$ based model.

Viable models can be built by adding an extra $U(1)_X$ factor, where the SM hypercharge Y is a combination of the T_8 generator of $SU(3)_L$ and X:

$$Y \propto \frac{\kappa}{\sqrt{3}} T_8 + X, \tag{29}$$

for some κ factor which controls the relative weight of X in Y. A known possibility is to take $\kappa = 1$, with the leptons in the representations $3\left(\mathbf{1},\overline{\mathbf{3}},-\frac{1}{3}\right) + 3(\mathbf{1},\mathbf{1},1)$ and the quarks in $2(\mathbf{3},\mathbf{3},0) + \left(\mathbf{3},\overline{\mathbf{3}},\frac{1}{3}\right) + 4\left(\overline{\mathbf{3}},\mathbf{1},-\frac{2}{3}\right) + 5\left(\overline{\mathbf{3}},\mathbf{1},\frac{1}{3}\right)$ [70]; an alternative is to take $\kappa = -3$, placing the leptons in the representations $3(\mathbf{1},\mathbf{3},0)$ and the quarks in $2\left(\mathbf{3},\overline{\mathbf{3}},-\frac{1}{3}\right) + \left(\mathbf{3},\mathbf{3},\frac{2}{3}\right) + 3\left(\overline{\mathbf{3}},\mathbf{1},-\frac{2}{3}\right) + 3\left(\overline{\mathbf{3}},\mathbf{1},\frac{1}{3}\right) + 2\left(\overline{\mathbf{3}},\mathbf{1},\frac{4}{3}\right) + \left(\overline{\mathbf{3}},\mathbf{1},-\frac{5}{3}\right)$ [71–73]. Trivial variations to the fermion content — adding real or conjugate pairs of representations of the $SU(3)_C \times SU(3)_L \times U(1)_X$ gauge group — are always possible and might even have interesting motivations (see for example [74]). But what about non-trivial variations? Just as with simple unification groups, the need to reproduce the SM chirality poses a very significant constraint. We do note that because of the extra $U(1)_X$ group factor the chirality condition does not imply automatically the absence of gauge anomalies, so these should be seen as complementary conditions on the possible fermion content of a given model.

We shall state here what are the simplest[19] non-trivial modifications which can be made to the known $SU(3)_C \times SU(3)_L \times U(1)_X$ models. If R is a complex representation of $SU(3)_C$, then it turns out that the simplest combinations of fields with no chirality and no gauge anomalies is of the form[20]

$$- 4(R,\mathbf{1},z) + 5\left(R,\mathbf{3},\frac{4}{3}\kappa + z\right) + 5\left(R,\overline{\mathbf{6}},\frac{2}{3}\kappa + z\right) - 5(R,\mathbf{8},\kappa + z)$$
$$- (R,\mathbf{10},2\kappa + z) - \left(R,\mathbf{15'},\frac{4}{3}\kappa + z\right) + \left(R,\overline{\mathbf{24}},\frac{5}{3}\kappa + z\right), \tag{30}$$

---

[19]The criteria for simplicity adopted here is based on the size of the $SU(3)_L$ representations involved: solutions with smaller representations are taken to be simpler.

[20]We stress again the content of footnote 6 concerning the naming of representations. This implies in particular that the $\overline{\mathbf{6}}$ of $SU(3)$ in this work is the $\mathbf{6}$ used in reference [60], and vice-versa.

for some arbitrary number z (κ was introduced earlier). Remarkably, this expression introduces many new $SU(3)_L$ representations going all the way up to size 24. So, for example, in the model $SU(3)_C \times SU(3)_L \times U(1)_X$ with $\kappa = 1$ [70], if we were to replace the quarks in four copies of the representation $\left(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3}\right)$ by something else, the simplest possibility would be (using equation (30) with $R = \overline{\mathbf{3}}$, $z = -\frac{2}{3}$)

$$4\left(\overline{\mathbf{3}},\mathbf{1},-\frac{2}{3}\right) \mapsto 5\left(\overline{\mathbf{3}},\mathbf{3},\frac{2}{3}\right) + 5\left(\overline{\mathbf{3}},\overline{\mathbf{6}},0\right) + 5\left(\mathbf{3},\mathbf{8},-\frac{1}{3}\right) + \left(\mathbf{3},\overline{\mathbf{10}},-\frac{4}{3}\right) + \left(\mathbf{3},\overline{\mathbf{15'}},-\frac{2}{3}\right) + \left(\overline{\mathbf{3}},\overline{\mathbf{24}},1\right). \quad (31)$$

The complexity of this mixture of representations can be interpreted as saying that the quark assignment in $SU(3)_C \times SU(3)_L \times U(1)_X$ models (for a fixed κ) is, for practical purposes, unique. Note that gauge anomaly cancellation alone would allow the replacement of $4\left(\overline{\mathbf{3}},\mathbf{1},-\frac{2}{3}\right)$ with much simpler combinations.

If R is a real representation of $SU(3)_C$, the combination in equation (30) is still valid, but it is not the simplest one. That distinction goes to

$$(R,\mathbf{3},x) + \left(R,\mathbf{3},-\frac{\kappa}{3}-x\right) - (R,\mathbf{3},y) - \left(R,\mathbf{3},-\frac{\kappa}{3}-y\right) + \left(R,\mathbf{1},\frac{\kappa}{3}-x\right) +$$
$$+ \left(R,\mathbf{1},\frac{2}{3}\kappa+x\right) - \left(R,\mathbf{1},\frac{\kappa}{3}-y\right) - \left(R,\mathbf{1},\frac{2}{3}\kappa+y\right), \quad (32)$$

where x should be different from y and $-\frac{\kappa}{3}-y$, otherwise this would be a self-conjugate combination of fields. Note also that we can drop one of the last four representations if it is made real by choosing an appropriate value for x or y. As such, referring once more to the model of [70] where $\kappa = 1$, one could replace the leptons in the three copies of the representation $\left(\mathbf{1},\overline{\mathbf{3}},-\frac{1}{3}\right)$ by three copies of the following rather more complex combination, which nevertheless does not involve $SU(3)_L$ representations bigger than (anti-)triplets:

$$\left(\mathbf{1},\mathbf{3},-\frac{2}{3}\right) + (\mathbf{1},\overline{\mathbf{3}},-y) + \left(\mathbf{1},\overline{\mathbf{3}},\frac{1}{3}+y\right) + (\mathbf{1},\mathbf{1},1) + \left(\mathbf{1},\mathbf{1},-\frac{1}{3}+y\right) + \left(\mathbf{1},\mathbf{1},-\frac{2}{3}-y\right). \quad (33)$$

In summary then, for a fixed embedding of the SM group in $SU(3)_C \times SU(3)_L \times U(1)_X$, it is fairly complicated to assign the leptons to other representations, and even more so for the quarks.

4 Concluding remarks

Pairs of right- and left-handed fermions transforming in the same way under the SM gauge group (vector fermions) can have a very high mass and therefore escape direct observation. On the other hand, unpaired ones (chiral fermions) may only get a small mass after electroweak symmetry breaking. Therefore, in Grand Unified Theories one must avoid introducing chiral fermions beyond those present in the Standard Model.

Motivated by this observation, we have analyzed in a systematic way the fermion sector of GUTs containing the SM fermions plus vector particles only. This very simple requirement on the fermion content of GUTs turns out to be a very constraining one (implying in most cases, but not all, the cancellation of all gauge anomalies). The analysis carried out assumes that there are no extra dimensions nor confining gauge interactions at high energies.

A thorough computer scan was performed over all simple groups with rank smaller than 12 (excluding $SO(22)$) and over all their representations up to some size (from a few thousands up to millions, depending on the group). A very significant part of the work consisted in tracking

down and looking at all possible ways of embedding G_{SM} in each unification group: this required cataloging the different ways that $SU(3) \times SU(2) \times U(1)^m$ can be embedded in G_{GUT} and, for each of them, to consider that the SM's $U(1)_Y$ can be a priori any linear combination of the m available $U(1)$'s.

With these simple groups, how exotic can GUTs be (the group, the embedding, and the field content)? Concerning the group, $SO(14)$ and $SO(18)$ were found not to work, leaving $SO(10)$, $SU(5 \leq N \leq 12)$ and E_6 as viable unification groups. Surprisingly, it was found that the SM gauge group must be embedded in each of these groups such that the GUT representations decompose in a unique way into SM fields. This uniqueness is far from obvious, even though model builders have been working with these groups and these embeddings for a long time. In every case, it turns out that the SM group is embedded in the GUT group in such a way that it can be viewed as going through $SU(5)$ for the calculation of the representation branching rules, $G_{GUT} \rightarrow SU(5) \rightarrow G_{SM}$, so an important consequence of this result is that the hypercharge normalization factor $\sqrt{3/5}$, usually associated to $SU(5)$ GUTs, is in fact universal.[21] Indeed, the relations $g_1 = \sqrt{5/3}g'$, $g_2 = g$ and $g_3 = g_s$ are the correct ones for all the tested cases.

As far as the field content is concerned, we dismissed the introduction of real fermion GUT representations or pairs of complex conjugate ones as trivial variations to the fermion sector of a model. It is also inconsequential to exchange the GUT representations $\sum_i R_i$ by their conjugates $\sum_i \overline{R_i}$, since one will recover the same SM representations by considering a different embedding of G_{SM} in G_{GUT}. Factoring out such variations, the standard fermion assignments $3(\overline{\mathbf{5}}) + 3(\mathbf{10})$ in $SU(5)$ and $3(\mathbf{27})$ in E_6 appear to be unique, while $3(\mathbf{16})$ in $SO(10)$ is not. In this last case, it is possible to get rid of the spinor representation, but that requires the introduction of complicated mixtures of very large representations. On the other hand, $SU(5 < N \leq 12)$ models may have a rich variety of different fermion representations, some of which have been already explored in the literature. For example, it is possible to have an $SU(11)$ model with just three fermion representations, $2(\overline{\mathbf{55}}) + \mathbf{462}$, matching in this sense the minimality of $SO(10)$ and E_6 GUTs, and perhaps exhibiting an interesting flavor structure.

Two non-simple groups were looked at as well. It was shown that no model based on the gauge group $SU(3) \times SU(3)$ can yield the SM chirality. On the other hand, viable models are known to exist with an extra $U(1)$, and in this work we commented that non-trivial changes to their fermion sector are possible, although they do need to be very elaborate.

Bigger unification groups were also considered, assuming that G_{SM} is in an $SU(5)$ subgroup of G_{GUT}. This analysis strongly suggests that $SO(10 < N \leq 30)$ are not suitable grand unified groups, in contrast to the $SU(N)$'s which, for $N \geq 15$, can actually embed G_{SM} in multiple valid ways. Interestingly, under one of these embeddings it is possible to unify N' SM families in a single representation of $SU(16 + N')$. In particular, the $\mathbf{171}$ of $SU(19)$ contains the SM fermions plus vector particles only. However, one should keep in mind that family unification with special unitary groups leads to gauge anomalies which, in a fundamental theory, need to be dealt with.

Acknowledgments

---

[21] For the extensive list of groups and representations considered at least.

Appendix

As mentioned in the text, the embedding of the SM group in $SU(5 < N \leq 12)$ characterized by the branching rule in equation (20) was the only one found to be capable of reproducing the SM chirality. The fermion content $\widetilde{c} \equiv -3(N-4)F + 3K$ is a particular example of a valid one (F being the fundamental representation of $SU(N)$ and $K \equiv (F \times F)_{\text{Ant.}}$). Other equality valid combinations of fields can be built by adding non-trivial combinations n_i of the $SU(N)$ representations with no chirality (as discussed in section (2)). This appendix contains the tables (3)–(9) which list all such independent n_i, for $5 < N \leq 12$. The biggest representation considered for the elaboration of each table was determined mainly by space considerations, given that more solutions seem to always appear if one includes bigger ones. For each group, one or more "interesting solutions" were picked in a somewhat subjective manner: they are notable for having either a reduced number of types of representations or a small number of total representations (including multiplicity).

Maximum size of representations considered	1000
Non-trivial combinations of representations with no chirality (n_i)	$-\mathbf{6} + \mathbf{21} - \mathbf{70} + \mathbf{84} - \mathbf{105}$, $-\mathbf{6} - \mathbf{15} + \mathbf{21} + \mathbf{84} + \mathbf{105'} - \mathbf{210}$, $-\mathbf{21} + \mathbf{56} + \mathbf{120} + \mathbf{210'} - \mathbf{280} - \mathbf{336}$, $-\mathbf{6} + \mathbf{21} - \mathbf{70} + \mathbf{84} - \mathbf{120} - \mathbf{210'} + \mathbf{280}$ $+\mathbf{420} - \mathbf{560} - \mathbf{840} + \mathbf{840'}$, $-\mathbf{56} - \mathbf{126} + \mathbf{315} + \mathbf{504} - \mathbf{720} + \mathbf{840''}$, $-\mathbf{6} + \mathbf{56} - \mathbf{70} + \mathbf{384} + \mathbf{420} - \mathbf{840} + \mathbf{896}$.
Interesting solutions with the SM chirality	$\widetilde{c} = -6(\mathbf{6}) + 3(\mathbf{15})$

Table 3: Information on the combinations of $SU(6)$ representations yielding the correct chirality. For example, the simplest solution, $\widetilde{c}$, together with more vector particles is discussed in [44, 46, 49].

Maximum size of representations considered	1000
Non-trivial combinations of representations with no chirality (n_i)	$-\mathbf{7} + \mathbf{21} - \mathbf{35}$, $3(\mathbf{7}) - 3(\mathbf{28}) + 2(\mathbf{112}) - \mathbf{140} + \mathbf{210}$, $4(\mathbf{7}) - 2(\mathbf{28}) + \mathbf{112} - 2(\mathbf{140}) + \mathbf{224}$, $5(\mathbf{7}) + 3(\mathbf{21}) - 4(\mathbf{28}) - 2(\mathbf{140}) - 2(\mathbf{196}) + \mathbf{490}$, $\mathbf{7} + 4(\mathbf{21}) - \mathbf{28} - 2(\mathbf{112}) - \mathbf{140} - 2(\mathbf{196}) + \mathbf{490'}$, $4(\mathbf{7}) + 4(\mathbf{21}) - 3(\mathbf{28}) - 3(\mathbf{140}) - \mathbf{196} + \mathbf{588}$, $3(\mathbf{28}) - 3(\mathbf{84}) - 2(\mathbf{189}) - 2(\mathbf{378}) + \mathbf{540} + \mathbf{756}$, $4(\mathbf{28}) - 2(\mathbf{84}) - 3(\mathbf{189}) - \mathbf{378} + 2(\mathbf{540}) + \mathbf{840}$.
Interesting solutions with the SM chirality	$\widetilde{c} = -9(\mathbf{7}) + 3(\mathbf{21})$, $\widetilde{c} - 3n_1 = -6(\mathbf{7}) + 3(\mathbf{35})$.

Table 4: Information on the combinations of $SU(7)$ representations yielding the correct chirality. For example, [45] uses $\widetilde{c}$ and [36] mentions both $\widetilde{c}$ and $\widetilde{c} - 3n_1$.

Maximum size of representations considered	1500
Non-trivial combinations of representations with no chirality (n_i)	$-3(\mathbf{8}) + 2(\mathbf{28}) - \mathbf{56}$, $6(\mathbf{8}) - 6(\mathbf{36}) + 3(\mathbf{168}) - \mathbf{216} - \mathbf{378}$, $10(\mathbf{8}) - 3(\mathbf{36}) + \mathbf{168} - 3(\mathbf{216}) + \mathbf{420}$, $15(\mathbf{8}) - 8(\mathbf{36}) + 3(\mathbf{168}) - 3(\mathbf{216}) + \mathbf{504}$, $15(\mathbf{8}) + 6(\mathbf{28}) - 10(\mathbf{36}) - 3(\mathbf{216}) + 3(\mathbf{336}) + \mathbf{1008}$, $2(\mathbf{8}) + 17(\mathbf{28}) - 3(\mathbf{36}) - 7(\mathbf{168}) - 3(\mathbf{216}) + 5(\mathbf{336}) + \mathbf{1176}$, $10(\mathbf{8}) + 10(\mathbf{28}) - 6(\mathbf{36}) - 6(\mathbf{216}) + \mathbf{336} + \mathbf{1344}$.
Interesting solutions with the SM chirality	$\widetilde{c} = -12(\mathbf{8}) + 3(\mathbf{28})$, $\widetilde{c} - 4n_1 = -5(\mathbf{28}) + 4(\mathbf{56})$.

Table 5: Information on the combinations of $SU(8)$ representations yielding the correct chirality. For example, in [47] the author considers the combination $\widetilde{c} - n_1$ which is readily seen to be the one involving the least amount of fermion components. Reference [43] considers instead the combination $\widetilde{c} - 2n_1$.

Maximum size of representations considered	1500
Non-trivial combinations of representations with no chirality (n_i)	$-6(\mathbf{9}) + 3(\mathbf{36}) - \mathbf{84},$ $-5(\mathbf{9}) + 2(\mathbf{36}) - \mathbf{126},$ $10(\mathbf{9}) - 10(\mathbf{45}) + 4(\mathbf{240}) - \mathbf{315} - \mathbf{630},$ $20(\mathbf{9}) - 4(\mathbf{45}) + \mathbf{240} - 4(\mathbf{315}) + \mathbf{720},$ $36(\mathbf{9}) - 20(\mathbf{45}) + 6(\mathbf{240}) - 4(\mathbf{315}) + \mathbf{1008},$ $45(\mathbf{9}) - 15(\mathbf{45}) + 4(\mathbf{240}) - 6(\mathbf{315}) + \mathbf{1050}.$
Interesting solutions with the SM chirality	$\begin{aligned} \widetilde{c} &= -15(\mathbf{9}) + 3(\mathbf{36}), \\ \widetilde{c} - n_1 &= -9(\mathbf{9}) + \mathbf{84}, \\ \widetilde{c} - 3n_2 &= -3(\mathbf{36}) + 3(\mathbf{126}). \end{aligned}$

Table 6: Information on the combinations of $SU(9)$ representations yielding the correct chirality. For example, the solutions $\widetilde{c} - n_1$ and $\widetilde{c} - n_1 + n_2$ are mentioned in [39] and [48] respectively.

Maximum size of representations considered	2000
Non-trivial combinations of representations with no chirality (n_i)	$-10(\mathbf{10}) + 4(\mathbf{45}) - \mathbf{120},$ $-16(\mathbf{10}) + 5(\mathbf{45}) - \mathbf{210},$ $15(\mathbf{10}) - 15(\mathbf{55}) + 5(\mathbf{330}) - \mathbf{440} - \mathbf{990},$ $35(\mathbf{10}) - 5(\mathbf{55}) + \mathbf{330} - 5(\mathbf{440}) + \mathbf{1155},$ $70(\mathbf{10}) - 40(\mathbf{55}) + 10(\mathbf{330}) - 5(\mathbf{440}) - \mathbf{1848},$ $105(\mathbf{10}) - 24(\mathbf{55}) + 5(\mathbf{330}) - 10(\mathbf{440}) + \mathbf{1980}.$
Interesting solutions with the SM chirality	$\begin{aligned} \widetilde{c} &= -18(\mathbf{10}) + 3(\mathbf{45}), \\ \widetilde{c} + 3n_1 - 3n_2 &= -3(\mathbf{120}) + 3(\mathbf{210}). \end{aligned}$

Table 7: Information on the combinations of $SU(10)$ representations yielding the correct chirality. Reference [41] mentions explicitly the solutions $\widetilde{c}$, $\widetilde{c} - n_1$ and $\widetilde{c} - 2n_1$.

Maximum size of representations considered	3000
Non-trivial combinations of representations with no chirality (n_i)	$-15(\mathbf{11}) + 5(\mathbf{55}) - \mathbf{165},$ $-35(\mathbf{11}) + 9(\mathbf{55}) - \mathbf{330},$ $-21(\mathbf{11}) + 5(\mathbf{55}) - \mathbf{462},$ $21(\mathbf{11}) - 21(\mathbf{66}) + 6(\mathbf{440}) - \mathbf{594} - \mathbf{1485},$ $56(\mathbf{11}) - 6(\mathbf{66}) + \mathbf{440} - 6(\mathbf{594}) + \mathbf{1760}.$
Interesting solutions with the SM chirality	$\begin{aligned} \widetilde{c} &= -21(\mathbf{11}) + 3(\mathbf{55}), \\ \widetilde{c} + n_1 - n_2 &= -\mathbf{11} - \mathbf{55} - \mathbf{165} + \mathbf{330}, \\ \widetilde{c} - n_3 &= -2(\mathbf{55}) + \mathbf{462}. \end{aligned}$

Table 8: Information on the combinations of $SU(11)$ representations yielding the correct chirality. The solution $\widetilde{c} + n_1 - n_2$ is mentioned in [37].

Maximum size of representations considered	4000
Non-trivial combinations of representations with no chirality (n_i)	$-21(\mathbf{12}) + 6(\mathbf{66}) - \mathbf{220}$, $-64(\mathbf{12}) + 14(\mathbf{66}) - \mathbf{495}$, $-70(\mathbf{12}) + 14(\mathbf{66}) - \mathbf{792}$, $28(\mathbf{12}) - 28(\mathbf{78}) + 7(\mathbf{572}) - \mathbf{780} - \mathbf{2145}$, $84(\mathbf{12}) - 7(\mathbf{78}) + \mathbf{572} - 7(\mathbf{780}) + \mathbf{2574}$.
Interesting solutions with the SM chirality	$\widetilde{c} = -24(\mathbf{12}) + 3(\mathbf{66})$, $\widetilde{c} - n_1 = -3(\mathbf{12}) - 3(\mathbf{66}) + \mathbf{220}$.

Table 9: Information on the combinations of $SU(12)$ representations yielding the correct chirality. The solution $\widetilde{c} + 4n_1 - 6n_2 + 4n_3$ is mentioned in [50].

References

[1] H. Georgi and S. Glashow, *Unity of all elementary-particle forces*, Phys. Rev. Lett. **32** (1974) 438–441.

[2] H. Georgi, *The state of the art — gauge theories*, AIP Conf. Proc. **23** (1975) 575–582.

[3] H. Fritzsch and P. Minkowski, *Unified interactions of leptons and hadrons*, Annals Phys. **93** (1975) 193–266.

[4] F. Gürsey, P. Ramond and P. Sikivie, *A universal gauge theory model based on E_6*, Phys. Lett. B **60** (1976) 177.

[5] M. Gell-Mann, P. Ramond and R. Slansky, *Complex spinors and unified theories*, proceedings of the Supergravity Workshop at Stony Brook (1979) 315–321, `arXiv:1306.4669 [hep-th]`.

[6] F. Wilczek and A. Zee, *Families from spinors*, Phys. Rev. D **25** (1982) 553.

[7] Y. Fujimoto, $SO(18)$ *unification*, Phys. Rev. D **26** (1982) 3183.

[8] D. Chang, T. Hubsch and R. Mohapatra, *Grand unification of three light generations*, Phys. Rev. Lett. **55** (1985) 673.

[9] J. Bagger and S. Dimopoulos, $O(18)$ *revived: splitting the spinor*, Nucl. Phys. B **244** (1984) 247.

[10] T. Hubsch and P. B. Pal, *Economical unification of three families in $SO(18)$*, Phys. Rev. D **34** (1986) 1606.

[11] J. Bagger, S. Dimopoulos, E. Masso and M. Reno, *Experimental consequences of family unification*, Phys. Rev. Lett. **54** (1985) 2199.

[12] D. Chang and R. N. Mohapatra, $SO(18)$ *unification of fermion generations*, Phys. Lett. B **158** (1985) 323.

[13] J. Bagger, S. Dimopoulos, E. Masso and M. Reno, *A realistic theory of family unification*, Nucl. Phys. B **258** (1985) 565.

[14] G. Senjanovic, F. Wilczek and A. Zee, *Reflections on mirror fermions*, Phys. Lett. B **141** (1984) 389.

[15] M. Ida, Y. Kayama and T. Kitazoe, *Inclusion of generations in $SO(14)$*, Prog. Theor. Phys. **64** (1980) 1745.

[16] A. Masiero, M. Roncadelli and T. Yanagida, *A grand unification of flavour with testable predictions*, Phys. Lett. B **117** (1982) 291.

[17] H. Sato, *Fermion masses from grand unification with $O(14)$*, Phys. Lett. B **101** (1981) 233.

[18] H. Sato, *Unification of generations and the Cabibbo angle*, Phys. Rev. Lett. **45** (1980) 1997.

[19] S. L. Adler, *Should E_8 SUSY Yang-Mills be reconsidered as a family unification model?*, Phys. Lett. B **533** (2002) 121–125, `arXiv:hep-ph/0201009`.

[20] S. L. Adler, *SU(8) family unification with boson-fermion balance*, Int. J. Mod. Phys. A **29** (2014) 1450130, `arXiv:1403.2099 [hep-ph]`.

[21] L. J. Dixon, J. A. Harvey, C. Vafa and E. Witten, *Strings on orbifolds*, Nucl. Phys. B **261** (1985) 678–686.

[22] B. R. Greene, K. H. Kirklin, P. J. Miron and G. G. Ross, *A three-generation superstring model. (I) Compactification and discrete symmetries*, Nucl. Phys. B **278** (1986) 667.

[23] B. R. Greene, K. H. Kirklin, P. J. Miron and G. G. Ross, *A three-generation superstring model. (II) Symmetry breaking and the low-energy theory*, Nucl. Phys. B **292** (1987) 606.

[24] L. E. Ibanez, J. E. Kim, H. P. Nilles and F. Quevedo, *Orbifold compactifications with three families of $SU(3) \times SU(2) \times U(1)^n$*, Phys. Lett. B **191** (1987) 282–286.

[25] L. E. Ibanez, H. P. Nilles and F. Quevedo, *Orbifolds and Wilson lines*, Phys. Lett. B **187** (1987) 25–32.

[26] K. Babu, S. Barr and B. Kyae, *Family unification in five and six dimensions*, Phys. Rev. D **65** (2002) 115008, `arXiv:hep-ph/0202178`.

[27] K. Hwang and J. E. Kim, *Orbifolded $SU(7)$ and unification of families*, Phys. Lett. B **540** (2002) 289–294, `arXiv:hep-ph/0205093`.

[28] K.-S. Choi and J. E. Kim, *Z_2 orbifold compactification of heterotic string and 6D $SO(16)$ and $E_7 \times SU(2)$ flavor unification models*, Phys. Lett. B **552** (2003) 81–86, `arXiv:hep-th/0206099`.

[29] I. Gogoladze, Y. Mimura and S. Nandi, *Unification of gauge, Higgs and matter in extra dimensions*, Phys. Lett. B **562** (2003) 307–315, `arXiv:hep-ph/0302176`.

[30] I. Gogoladze, Y. Mimura and S. Nandi, *Unity of elementary particles and forces in higher dimensions*, Phys. Rev. Lett. **91** (2003) 141801, `arXiv:hep-ph/0304118`.

[31] Z.-y. Han and W. Skiba, *Family unification on an orbifold*, Phys. Rev. D **70** (2004) 035013, `arXiv:hep-ph/0405199`.

[32] Y. Kawamura, T. Kinami and K.-y. Oda, *Orbifold family unification*, Phys. Rev. D **76** (2007) 035001, `arXiv:hep-ph/0703195`.

[33] Y. Kawamura and T. Miura, *Orbifold family unification in $SO(2N)$ gauge theory*, Phys. Rev. D **81** (2010) 075011, `arXiv:0912.0776 [hep-ph]`.

[34] K.-S. Choi and J. E. Kim, *Supersymmetric three family chiral $SU(6)$ grand unification model from F-theory*, Phys. Rev. D **83** (2011) 065016, `arXiv:1012.0847 [hep-ph]`.

[35] Y. Goto, Y. Kawamura and T. Miura, *Orbifold family unification on six dimensions*, Phys. Rev. D **88** (2013) 5 055016, `arXiv:1307.2631 [hep-ph]`.

[36] P. Frampton, *$SU(N)$ grand unification with several quark–lepton generations*, Phys. Lett. B **88** (1979) 299.

[37] H. Georgi, *Towards a Grand Unified Theory of flavor*, Nucl. Phys. B **156** (1979) 126–134.

[38] N. Baaklini, *Chiral grand unification in $SU(N > 5)$*, J. Phys. G **6** (1980) 917–931.

[39] P. Frampton, *Unification of flavor*, Phys. Lett. B **89** (1980) 352.

[40] J. E. Kim, *Model of flavor unity*, Phys. Rev. Lett. **45** (1980) 1916.

[41] S. Kashibayashi, K. Fukuma and T. Ueda, *Further constraints on $SU(N)$ flavor grand unification*, Phys. Rev. D **26** (1982) 2529–2531.

[42] P. H. Frampton and B.-H. Lee, *$SU(15)$ grand unification*, Phys. Rev. Lett. **64** (1990) 619.

[43] J. Chkareuli, *The SU(8) GUT for chiral families*, Phys. Lett. B **300** (1993) 361–366.

[44] Z. Berezhiani, *SUSY $SU(6)$: GIFT for doublet-triplet splitting and fermion masses*, Phys. Lett. B **355** (1995) 481–491, arXiv:hep-ph/9503366.

[45] J. Chkareuli, C. Froggatt, I. Gogoladze and A. Kobakhidze, *From prototype $SU(5)$ to realistic $SU(7)$ SUSY GUT*, Nucl. Phys. B **594** (2001) 23–45, arXiv:hep-ph/0003007.

[46] Q. Shafi and Z. Tavartkiladze, *Realistic supersymmetric $SU(6)$*, Phys. Lett. B **522** (2001) 102–106, arXiv:hep-ph/0105140.

[47] S. Barr, *Doubly lopsided mass matrices from unitary unification*, Phys. Rev. D **78** (2008) 075001, arXiv:0804.1356 [hep-ph].

[48] J. B. Dent, R. Feger, T. W. Kephart and S. Nandi, *Natural fermion mass hierarchy and mixings in family unification*, Phys. Lett. B **697** (2011) 367–369, arXiv:0908.3915 [hep-ph].

[49] A. Cárcamo Hernández and R. Rahman, *A SUSY $SU(6)$ GUT model with pseudo-Goldstone Higgs doublets* (2010), arXiv:1007.0447 [hep-ph].

[50] C. H. Albright, R. P. Feger and T. W. Kephart, *An explicit $SU(12)$ family and flavor unification model with natural fermion masses and mixings*, Phys. Rev. D **86** (2012) 015012, arXiv:1204.5471 [hep-ph].

[51] P. Byakti, D. Emmanuel-Costa, A. Mazumdar and P. B. Pal, *Number of fermion generations from a novel grand unified model*, Eur. Phys. J. C **74** (2014) 2730, arXiv:1308.4305 [hep-ph].

[52] M. Gell-Mann, P. Ramond and R. Slansky, *Color embeddings, charge assignments, and proton stability in unified gauge theories*, Rev. Mod. Phys. **50** (1978) 721.

[53] R. King, *Fundamental fermion representations in generalizations of the $SU(5)$ Grand Unified Gauge Theory*, Nucl. Phys. B **185** (1981) 133–156.

[54] R. M. Fonseca, *Calculating the renormalisation group equations of a SUSY model with Susyno*, Comput. Phys. Commun. **183** (2012) 2298–2306, arXiv:1106.5016 [hep-ph].

[55] R. Feger and T. W. Kephart, *LieART – a Mathematica application for Lie algebras and representation theory* (2012), arXiv:1206.6379 [math-ph].

[56] E. Dynkin, *Maximal subgroups of the classical groups*, Trans. Am. Math. Soc. **6** (1957) 245–378.

[57] E. Dynkin, *Semisimple subalgebras of semisimple Lie algebras*, Trans. Am. Math. Soc. **6** (1957) 111–244.

[58] I. Losev, *On invariants of a set of elements of a semisimple Lie algebra*, Journal of Lie Theory **20** (2010) 1 017–030, `arXiv:math/0512538 [math.RT]`.

[59] A. N. M. Minchenko, *The semisimple subalgebras of exceptional Lie algebras*, Trans. Moscow Math. Soc. **67** (2006) 225–259.

[60] R. Slansky, *Group theory for unified model building*, Phys. Rept. **79** (1981) 1–128.

[61] S. M. Barr, *A new symmetry breaking pattern for $SO(10)$ and proton decay*, Phys. Lett. B **112** (1982) 219.

[62] S. Bertolini, L. Di Luzio and M. Malinský, *Minimal flipped $SO(10) \otimes U(1)$ supersymmetric Higgs model*, Phys. Rev. D **83** (2011) 035002, `arXiv:1011.1821 [hep-ph]`.

[63] G. Dvali, *Black holes and large N species solution to the hierarchy problem*, Fortsch. Phys. **58** (2010) 528–536, `arXiv:0706.2050 [hep-th]`.

[64] X. Calmet, S. D. Hsu and D. Reeb, *Grand unification and enhanced quantum gravitational effects*, Phys. Rev. Lett. **101** (2008) 171802, `arXiv:0805.0145 [hep-ph]`.

[65] S. L. Adler, *Axial-vector vertex in spinor electrodynamics*, Phys. Rev. **177** (1969) 2426–2438.

[66] J. Bell and R. Jackiw, *A PCAC puzzle: $\pi^0 \rightarrow \gamma\gamma$ in the σ-model*, Nuovo Cim. A **60** (1969) 47–61.

[67] R. N. Cahn, *Anomalies and the particle content of Grand Unified Theories*, Phys. Lett. B **104** (1981) 282.

[68] J. Preskill, *Gauge anomalies in an effective field theory*, Annals Phys. **210** (1991) 323–379.

[69] M. B. Green and J. H. Schwarz, *Anomaly cancellation in supersymmetric $D = 10$ gauge theory and superstring theory*, Phys. Lett. B **149** (1984) 117–122.

[70] M. Singer, J. Valle and J. Schechter, *Canonical neutral-current predictions from the weak-electromagnetic gauge group $SU(3) \times U(1)$*, Phys. Rev. D **22** (1980) 738.

[71] F. Pisano and V. Pleitez, *$SU(3) \otimes U(1)$ model for electroweak interactions*, Phys. Rev. D **46** (1992) 410–417, `arXiv:hep-ph/9206242`.

[72] P. Frampton, *Chiral dilepton model and the flavor question*, Phys. Rev. Lett. **69** (1992) 2889–2891.

[73] R. Foot, O. F. Hernandez, F. Pisano and V. Pleitez, *Lepton masses in an $SU(3)_L \otimes U(1)_N$ gauge model*, Phys. Rev. D **47** (1993) 4158–4161, `arXiv:hep-ph/9207264`.

[74] S. M. Boucenna, R. M. Fonseca, F. González-Canales and J. W. F. Valle, *Small neutrino masses and gauge coupling unification*, Phys. Rev. D **91** (2015) 3 031702(R), `arXiv:1411.0566 [hep-ph]`.