Radiative effects in gauge extensions of the Standard Model of particle interactions

Michal Malinský

Habilitation thesis
2020
The effort to understand the universe is one of the very few things that lifts human life a little above the level of farce, and gives it some of the grace of tragedy.

Steven Weinberg
The first three minutes

Dedication

I dedicate this thesis to the people that make my life meaningful – to my wife whose patience I often stretch to limits with no effect on her unflattering support, to my children to whom I owe a lot, to my parents who gave me all they could, and to all friends whose enthusiasm keeps me going.

Professionally, I would like to thank to three men that, over the years, formed my way of thinking about physics – Jiří Hořejší for the attitude, Stefano Bertolini for the method and Goran Senjanović for the passion.

In Prague, February 2020
# Contents

1 Introduction

2 Perturbative baryon and lepton number violation in gauge extensions of the Standard Model

## 2.1 Neutrino masses in simple extensions of the Standard Model

### 2.1.1 RH singlet fermions and Dirac neutrino masses

### 2.1.2 Dirac neutrinos and the SM hypercharge de-quantization

### 2.1.3 Majorana fermions

### 2.1.4 Massive Majorana neutrino phenomenology

## 2.2 Majorana neutrino mass as a $d = 5$ operator in the Standard Model

## 2.3 Standard Model at $d = 6$ level and perturbative baryon number violation

### 2.3.1 $B$ and $L$ – violating $d = 6$ operators and their phenomenology

### 2.3.2 Tree-level renormalizable openings of the $d = 6$ BLNV operators

## 2.4 High-energy convergence of the SM gauge couplings

### 2.4.1 Renormalized gauge couplings and their evolution in the SM

## 2.5 Grand unification of the SM interactions

### 2.5.1 Evolution of gauge couplings in the BLNV extensions of the SM

### 2.5.2 The grand unification hypothesis and its implications

### 2.5.3 Calculability of proton lifetime in unified models

### 2.5.4 Topological defects

## 2.6 Towards a potentially realistic grand unified theory

### 2.6.1 The minimal $SU(5)$ Georgi-Glashow model

### 2.6.2 Phenomenology of the minimal $SU(5)$ GUT

### 2.6.3 The failure of the minimal $SU(5)$ GUT

### 2.6.4 The minimal supersymmetric $SU(5)$ GUT

### 2.6.5 The issues of the minimal supersymmetric $SU(5)$ GUT
### 3 Radiative effects in potentially realistic unified gauge models

- **3.1** Rank=5 extended gauge models ............................................. 50
  - **3.1.1** Left-right symmetric models ............................................. 51
  - **3.1.2** The Pati-Salam model ............................................. 52
  - **3.1.3** SO(10) grand unification ............................................. 54
- **3.2** Minimal SO(10) GUTs ............................................. 55
  - **3.2.1** Proton decay in SO(10) GUTs ............................................. 57
  - **3.2.2** Yukawa sector of simple SO(10) GUTs ............................................. 58
- **3.3** The spectacular failure of the minimal SUSY SO(10) ............................................. 58
  - **3.3.1** The structure of the the minimal SUSY SO(10) GUT ............................................. 59
  - **3.3.2** The neutrino challenge to SUSY GUTs ............................................. 61
- **3.4** Quantum salvation of the minimal SO(10) GUT ............................................. 62
  - **3.4.1** The cons and pros of the non-SUSY SO(10) GUTs ............................................. 62
  - **3.4.2** The tree-level vacuum of the minimal SO(10) Higgs models ............................................. 62
  - **3.4.3** The minimal SO(10) Higgs model(s) at the loop level ............................................. 65
  - **3.4.4** The minimal potentially realistic and testable GUTs ............................................. 67
- **3.5** Aspects of renormalization group evolution in theories with more than a single U(1) gauge factor ............................................. 70
  - **3.5.1** Renormalization of Abelian gauge theories (QED) ............................................. 71
  - **3.5.2** Renormalization of QED squared ............................................. 72
  - **3.5.3** The link between $Z_A$ and $\delta \xi$ in different renormalization schemes ............................................. 74
  - **3.5.4** Beta-functions in schemes with matrix gauge couplings ............................................. 76
  - **3.5.5** One-loop matching in schemes with $U(1)$ mixing ............................................. 78
- **3.6** Towards the minimal renormalizable theory of perturbative baryon number violation ............................................. 80
  - **3.6.1** Witten’s loop in the SO(10) GUTs ............................................. 81
  - **3.6.2** Witten’s loop in the flipped SU(5) unification ............................................. 83
  - **3.6.3** The minimal renormalizable theory of perturbative $B$ violation ............................................. 85
  - **3.6.4** Calculating Witten’s loop in the flipped SU(5) context ............................................. 88

### 4 Conclusions and outlook

- **5 Supplementary material** ............................................. 93
  - **5.1** The novel SO(10) seesaw mechanism ............................................. 93
  - **5.2** The minimal renormalizable SUSY SO(10) no-go ............................................. 99
  - **5.3** The minimal SO(10) GUT resurrection ............................................. 115
  - **5.4** Two-loop SUSY RGEs with U(1) mixing ............................................. 139
  - **5.5** The minimal theory of perturbative $B$ violation ............................................. 167
Chapter 1

Introduction

The quest for a simple and accurate description of the elementary constituents of matter and their interactions is arguably one of the biggest intellectual adventures of all times. Based on the revolutionary insights of Maxwell, Einstein, Planck and many other giants of modern physics, our understanding of electromagnetic, strong and weak nuclear forces governing the quantum world matured over the last century into what is nowadays claimed to be the most accurate physics theory ever, the Standard Model (SM) of particles and their interactions. Fifty years of an enormous experimental scrutiny, crowned recently by the discovery of its last missing ingredient – the Higgs boson – makes us believe that the SM, indeed, provides a perfectly valid description of the microcosmos at energies stretching up to, at least, the TeV scale.

Yet, there are now clear signals that the Standard Model in its original Glashow-Salam-Weinberg (GSW) formulation \[1-3\] does not encompass all observable particle physics phenomena – the discovery of neutrino flavour oscillations \[4-6\] strongly indicates that neutrinos, the most elusive of all SM particles, are not massless as assumed in the GSW construction. Moreover, their unprecedented lightness, as implied either directly by the beta-decay data (see, e.g., \[7\]) or indirectly by various cosmological limits \[8\], suggests that masses of neutrinos may have a rather different dynamical origin than those of all other SM matter fields.

Indeed, the most plausible explanation of their sub-eV mass scale has to do with the interesting option that neutrinos – as the only electrically neutral matter fermions – may be in fact Majorana spinors whose mass could be associated (in the SM picture) to the unique $d = 5$ lepton-number-violating operator $LLHH/\Lambda \ [9]$; in that case, the smallness of the neutrino masses would be attributed to the largeness of the so-called seesaw scale $\Lambda$ typically assumed to be in the $10^{12-14}$ GeV ballpark. By definition, such a Majorana mass term violates all additive quantum numbers the neutrinos carry; in particular, it breaks the global rephasing invariance of the SM Lagrangian associated
to the lepton number \((L)\). Although, for a long time, \(L\) has been considered a perfect empirical symmetry of the low-energy world, there is actually nothing sacred about the SM lepton number conservation – it is indeed violated by non-perturbative effects associated to the so-called triangle anomalies \([10,11]\) of the corresponding quantum currents. On the other hand, at low temperature these effects are so small that there is essentially no way to look for them in laboratory-based searches; nevertheless, they might have played a crucial role in the hot early Universe \([12]\). In this respect, the seesaw realisation of neutrino masses can be seen as a first indication of a perturbative lepton number violation (LNV) in beyond-Standard-Model (BSM) physics; as such, it is subject to enormous experimental efforts stretching form the neutrinoless double-beta decay activities at the intensity frontier (see, e.g., \([13,14]\)) to the LNV collider searches (cf. \([15]\)) on the high-energy side.

Another piece of observation the SM fails to account for is the value of the baryon-to-photon number density ratio in the early Universe, a crucial parameter governing the abundances of the light nuclei created during Weinberg’s “first three minutes” of its thermal history. Indeed, the measurements of the intensity of the Deuterium absorption lines in the light of high-redshift quasars yield values which are some 8 orders of magnitude above the \(10^{-18}\) SM estimate based on the assumption of an exact baryon-number \((B)\) symmetry of the initial conditions.

Since the SM baryon-number current features the same anomaly structure as the leptonic one and, as such, it may be viewed as just the other facet of essentially the same coin, it is very natural to ask whether there are any indications that it may also be perturbatively violated in the BSM physics. In the effective SM picture this is analogous to asking at which scale the \(d=6\) baryon-number-violating (BNV) operators \([9,16,17]\) mediating “classically” forbidden processes like, e.g., proton decay, BNV neutron decays etc., are generated and whether there is any chance to get a firm grip on them.

From that perspective, the profound idea of grand unification of strong and electroweak interactions \([18]\) as a possible theory of perturbative baryon number violation comes about as a perfectly natural and logical continuation of this line of thoughts\(^1\). Indeed, if the physics at some high energy scale was governed by a gauge theory based on a simple gauge group \(G\) [such as, e.g., \(SU(5)\) or \(SO(10)\)] containing the SM gauge symmetry \(G_{SM} = [SU(3) \otimes SU(2) \otimes U(1)]/Z_6\) as a subgroup, the quarks and leptons would occupy common irreducible representations of \(G\); consequently, the gauge bosons associated to the coset \(G/G_{SM}\) (and the dynamical remnants of the Higgs multiplets triggering the relevant spontaneous symmetry breakdown) will mediate \(B\) and \(L\) vio-

---

\(^1\)Needless to say, this is not the original reasoning of Georgi and Glashow who were rather motivated by the uniqueness of the SM gauge group embedding into the \(SU(5)\), c.f. \([18]\).
lating quark-lepton transitions. The \( d = 6 \) (and higher-order) BNV operators are then obtained by integrating these heavy degrees of freedom out of the low-energy effective theory and, in any specific scheme, their structure will closely reflect the details of the underlying dynamics. Besides that, such a unified-gauge-theory realisation of the \( d = 6 \) SM BNV operators has one great virtue: the relevant energy scale \( M \) (defined essentially as the mass of the mediators of the BNV transitions) is fixed by the requirement that the running gauge couplings in the effective theory converge to a common value close to \( M \). Remarkably enough, this is exactly what happens in the SM at around \( 10^{15-16} \) GeV (focusing on the non-abelian sector and assuming, for simplicity, no new physics between \( M \) and the electroweak scale). Hence, unlike for the seesaw scale \( \Lambda \) governing the \( d = 5 \) neutrino mass operator (which can be anywhere below about \( 10^{15} \) GeV as long as the coefficient of the Weinberg operator is suppressed accordingly), in grand unified theories (GUTs) the scale of perturbative \( d = 6 \) BNV is subject to a strong constraint.

Although, obviously, the extreme remoteness of \( M \) makes any direct collider scrutiny of the GUT paradigm inconceivable, there are several reasons to believe that the basic gauge unification idea can be to a high degree testable even with the current technology. Besides proton decay as a hallmark of BNV, the spontaneous breakdown of a unified gauge-group \( G \) to the \( SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \) of the Standard Model can give rise to massive topological defects (monopoles, cosmic strings, domain walls) corresponding to stable extended gauge/Higgs field configurations with very peculiar features. These, among other things, include the so-called Callan-Rubakov effect \([19,20]\) corresponding to, e.g., a strong enhancement of the BNV rates in the core of such objects. Although the relevant monopole fluxes are nowadays likely to be exponentially suppressed by the early cosmological inflation, the discovery of monopoles would provide such a robust evidence for grand unification that they are still subject to dedicated experimental efforts (cf. \([21]\)) typically along with the “classical” proton decay searches. In this respect, Nature seems to be (as often) rather generous because the raw estimate of \( M \) in the \( 10^{15-16} \) GeV ballpark yields proton lifetime in the range of about \( 10^{30-36} \) years, right at the verge of feasibility; let us note that these numbers correspond roughly to one decay in about a ton (\( 10^{30} \) years) to a megaton (\( 10^{36} \) years) of material in about a year.

This, obviously, represents an enormous experimental challenge and, thus, the progress has been relatively slow (amounting to roughly one order of magnitude improvement per decade). To name just few of the most important early searches, let us mention the NUSEX \([22]\), FREJUS \([23]\), SOUDAN \([24]\) and Kamiokande \([25]\) experiments from 1980’s and the 3rd phase of IMB \([26]\) at the beginning of the 1990’s
which probed the proton lifetime up to about $10^{33}$ years. The current best limits from Super-Kamiokande (SK) \cite{27} reach up to $2 \times 10^{34}$ years in the $p^{+} \rightarrow e^{+} \pi^{0}$ channel and to about $10^{33}$ years for $p^{+} \rightarrow K^{+} \nu$. On the other hand, these numbers suggest that we may be just on the brink of really observing the first experimental signal of this kind. In any case, the close complementarity between the methodology of the proton decay searches and that of the very lively and extensive neutrino-physics programme (which is central to most of the upcoming large-volume facilities like DUNE or Hyper-K) ensures a bright future for the experimental searches for the baryon number violation for at least the next three decades. Hence, we should cross our fingers and stay tuned.

In this thesis we shall review the crucial points along the lines of reasoning sketched above and provide a brief account of the candidate’s contribution to the evolution of this thrilling field of the high-energy physics research. In doing so we shall necessarily be rather selective as to which of its many aspects shall be entertained and which – with all due respect to their pioneers – shall be suppressed (or even entirely omitted). Let us also note that the selected publications enclosed in Chapter 5 represent just a small fragment of candidate’s achievements; their complete list is available as a part of his professional CV enclosed in the file.
Chapter 2

Perturbative baryon and lepton number violation in gauge extensions of the Standard Model

2.1 Neutrino masses in simple extensions of the Standard Model

The overwhelming evidence that at least two out of the three known neutrino states are massive clearly calls for a generalisation of the original GSW formulation of the Standard Model (SM) where they were, by construction, two-components Weyl fermions.

At first glance, this exercise has a trivial solution: postulating the existence of new spin-$\frac{1}{2}$ quantum fields $N_R$ transforming as full singlets under the $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ SM gauge symmetry does not seem to be a big deal, the more that the presence of such fields (if provided in three copies) nicely restores the symmetry among the numbers of the left- and right-handed (LH, RH) fermions of the theory.

2.1.1 RH singlet fermions and Dirac neutrino masses

Indeed, with the extra RH leptons at hand one can speculate about a leptonic analogue of the up-type Yukawa coupling structure $\overline{Q}_L Y_u u_R \tilde{H}$ in the form

$$\mathcal{L} \ni \overline{L}_L Y_\nu N_R \tilde{H} + h.c. ,$$  \hspace{1cm} (2.1)

producing, in the broken phase with $\langle \tilde{H} \rangle = \frac{1}{\sqrt{2}} (v, 0)^T$, a mass term of the same (Dirac) type as those encountered in the SM for the charged fermions. Let us note that,
assuming the standard hypercharge assignment of all the SM charged matter fields\footnote{At first glance this is the most natural thing to do; however, this assumption may not be fully justified in extensions of the Standard Model featuring new chiral fermions, cf. Sect. 2.1.2.}

\begin{align}
Y_{QL} &= +\frac{1}{6}, \quad Y_{u_R} = +\frac{2}{3}, \quad Y_{d_R} = -\frac{1}{3}, \quad Y_{L_L} = -\frac{1}{2}, \quad Y_{e_R} = -1,
\end{align}

\(N_R\) is enforced to be a complete SM gauge singlet and, as such, it would not feel any of the SM gauge interactions. In this respect, adding such a new field into the play does not upset the basic SM phenomenology in any way. Second, requiring \(m_\nu = Y_\nu v/\sqrt{2}\) to be in the sub-eV ballpark (as indicated by beta decay experiments \footnote{Numerically, formula (2.4) yields \(\text{BR}(\mu \to e\gamma)\) in the \(10^{-46}\) ballpark. Note that there is no chance to see a signal like that as the sensitivity of the current experiments (such as MEG \footnote{MEG} \footnote{MEG}) does not exceed about \(4 \times 10^{-13}\).}) the Yukawa matrix \(Y_\nu\) is sentenced to be very small \(|Y_\nu| \lesssim 10^{-11}\), almost at the verge of negligibility. Note, however, that in principle there is nothing wrong about the relative smallness of \(Y_\nu\) as compared to the other Yukawa matrices in the SM; indeed, Yukawa couplings are self-renormalized and, hence, a small one remains small even at the loop level.

Let us turn to the phenomenology though. As simple as the SM extension with (2.1) looks, the physical consequences of the presence of even a single extra term along with the traditional SM interactions are paramount:

1. **Lepton flavour violation.** With a “complete” set of RH leptons in the game the flavour structure of the leptonic charged-current (CC) interaction Lagrangian fully resembles the situation in the quark sector. In the mass basis, a \(3 \times 3\) unitary matrix (usually called Pontecorvo-Maki-Nakagawa-Sakata, PMNS) encompasses all the flavour-changing charged-current (CC) interactions involving leptons:

\begin{equation}
\mathcal{L}^{CC}_{\text{lept.}} = \frac{g}{\sqrt{2}} \bar{e}_L \gamma^\mu U_{\text{PMNS}} \nu_i W^-_\mu + h.c.
\end{equation}

Besides neutrino oscillations, loop-induced flavour-changing neutral current processes (FCNC) such as \(\mu \to e\gamma\), \(\tau \to 3e\) will be allowed, albeit with immeasurably small rates, for instance \footnote{29} \footnote{30}

\begin{equation}
\Gamma(\mu \to e\gamma) \sim \frac{G_F^2 m_\mu^5}{192\pi^3} \frac{3\alpha}{128\pi} \sin^2 2\theta \left( \frac{m_2^2 - m_1^2}{M_W^2} \right)^2,
\end{equation}

where a convenient two-flavour approximation (reflected by the presence of a single leptonic mixing angle and only two masses) has been employed.

2. **Leptonic CP violation.** Another phenomenon intimately connected to the non-trivial mixing in the lepton sector CC’s (2.3) is the presence of a new source of
CP violation related to the (thus far unknown) complex structure of the PMNS matrix. Besides its potentially striking effects in the neutrino oscillation phenomena it would also contribute to other leptonic CP-odd observables such as the electric dipole moment (EDM) of the electron $d_e$. To this end, let us note that $d_e \sim 10^{-38}\text{e.cm}$ obtained in the SM [31] (to be compared with the experimental limit $d_e^{\text{exp}} \lesssim 10^{-29}\text{e.cm}$) receives its first contribution only at the four-loop level due to the need to transfer the CP phase from the quark sector Cabibbo-Kobayashi-Maskawa (CMK) matrix into the leptons; with the extra CP violation in the PMNS matrix the same observable can be generated already at two loops [33] (though hardly enhanced; the purely leptonic contribution to $d_e$ would not exceed $d_e \lesssim 10^{-43}\text{e.cm}$ either).

2.1.2 Dirac neutrinos and the SM hypercharge de-quantization

So far, the introduction of the RH singlet(s) into the SM field list and the addition of the relevant Yukawa term (2.1) to the SM Lagrangian was merely beneficial. Unfortunately, at the quantum level, this simple scheme suffers from a serious pathology when it comes to one of the most beautiful features of the original SM, namely, the mechanism of the (hyper)charge quantization by the requirement of a complete cancellation of chiral anomalies.

To this end, let us first recall that unlike for the non-Abelian charges (i.e., the eigenvalues of the $T^3_c$, $T^3_e$ and $T^3_L$ Cartans of the $SU(3)_c \otimes SU(2)_L$ part of the SM gauge group) whose quantization is ensured by the properties of the special unitary groups’ representations, Abelian charges like, e.g., the eigenvalues of the hypercharge operator $Y$ (the generator of $U(1)_Y$ of the SM) can be, from the group theory perspective, arbitrary real numbers. From this point of view, the hypercharges of the SM matter fields in Eq. (2.2) may be (at the classical level) viewed as rather phenomenological quantities which, however, happen to be carefully selected to yield the desired pattern of electric charges following the Gell-Mann-Nishijima relation

$$Q = T^3_L + Y.$$  \hspace{1cm} (2.5)

The precision with which this game must be played is rather unprecedented; for instance, the current neutron neutrality constraints [34] are such that the hypercharges of the up and down quarks must conspire to the level of at least 21 significant digits!

Cancellation of chiral anomalies in the SM

From this perspective, it is great that the SM provides a clear rationale for this “fine-tuning” (and, in turn, for the electric neutrality of atomic matter) at the quantum
level. The argument has to do with the delicacy of the quantum structure of chiral gauge theories that, in general, are prone to a certain pathological behaviour which consists in the possible dependency of some of the physical amplitudes on the selection of the gauge in the corresponding calculations. As such, these so-called chiral anomalies \cite{10,11} must be absent if a theory under consideration is to play the role of an internally consistent and calculable model of Nature.

The requirement of absence of such anomalies may thus be used as a powerful discriminator among different settings and, in particular, among different sets of charges chiral fermions can attain in specific scenarios. The point is that the core anomaly structure is purely algebraic,

$$A^{abc} \propto \text{Tr}_{RL} \left( \left\{ T^a_{RL}, T^b_{RL} \right\} T^c_{RL} \right) - \text{Tr}_{RR} \left( \left\{ T^a_{RR}, T^b_{RR} \right\} T^c_{RR} \right), \quad (2.6)$$

where $T^{a,b,c}$ stand for the generators of all the gauge symmetries at play and the traces (Tr) are taken over all their representations $R_{L,R}$ accommodating left-handed (L) and right-handed (R) Weyl spinors, respectively. As such, demanding $A^{abc} = 0$ for all $a$, $b$ and $c$ translates into algebraic equations for the charges involved (especially the Abelian ones).

Remarkably enough, playing this game for the hypercharges\footnote{Note that the logic of the game here is different from the usual calculations performed in order to confirm that the SM with hypercharges selected as in Eq. (2.2) is indeed anomaly free. In the current case, we take all $Y_Q,u,d,L,e$ as free parameters and search for all non-trivial (i.e., non-zero) solutions of the relevant system of algebraic equations.} of five different types of matter fields encountered in each of the three SM generations, one reveals the traditional hypercharge configuration \footnote{Note for completeness that there is in principle another (almost trivial) solution with $Y_Q = Y_L = Y_e = 0$ and $Y_u = -Y_d$ which, however, is rather bizarre and may be discarded either on phenomenological grounds or by elevating the SM gauge group to a full-fledged left-right symmetry (cf. Sect. 3.1.1); for more information see, e.g., \cite{35}.} \footnote{Note for completeness that there is in principle another (almost trivial) solution with $Y_Q = Y_L = Y_e = 0$ and $Y_u = -Y_d$ which, however, is rather bizarre and may be discarded either on phenomenological grounds or by elevating the SM gauge group to a full-fledged left-right symmetry (cf. Sect. 3.1.1); for more information see, e.g., \cite{35}.} as the only non-trivial one \footnote{Note for completeness that there is in principle another (almost trivial) solution with $Y_Q = Y_L = Y_e = 0$ and $Y_u = -Y_d$ which, however, is rather bizarre and may be discarded either on phenomenological grounds or by elevating the SM gauge group to a full-fledged left-right symmetry (cf. Sect. 3.1.1); for more information see, e.g., \cite{35}.} (modulo the overall normalization). As this is one of the central points of the discussion here, let us elaborate on these lines a little bit further. In fact, there is a neat trick which makes the inspection of the entire set of anomalies unnecessary; it can be shown \cite{35,36} that the solution of the general problem is equivalent to considering just the $SU(2)_L \otimes U(1)_Y$ anomalies together with the constraints emerging from the requirement of the gauge invariance of the up and down quark and charged leptons’ Yukawa terms in the SM Lagrangian. Working this out, one reveals the following set of equations for the hypo-

\[ A^{abc} \propto \text{Tr}_{RL} \left( \left\{ T^a_{RL}, T^b_{RL} \right\} T^c_{RL} \right) - \text{Tr}_{RR} \left( \left\{ T^a_{RR}, T^b_{RR} \right\} T^c_{RR} \right), \quad (2.6) \]
theoretically unknown values of $Y_Q$, $Y_u$, $Y_d$, $Y_L$ and $Y_e$ (and $Y_H$ of the Higgs field):

\begin{align}
U(1)^3 \text{ anomaly: } & 6Y_Q^3 + 2Y_L^3 - 3Y_u^3 - 3Y_d^3 - Y_e^3 = 0, \quad (2.7) \\
SU(2)^2 \otimes U(1) \text{ anomaly: } & 3Y_Q + Y_L = 0, \quad (2.8) \\
\text{up-type-quark Yukawa: } & -Y_Q + Y_u - Y_H = 0, \quad (2.9) \\
\text{down-type-quark Yukawa: } & -Y_Q + Y_d + Y_H = 0, \quad (2.10) \\
\text{charged-lepton Yukawa: } & -Y_L + Y_e + Y_H = 0. \quad (2.11)
\end{align}

This comprises 5 equations for 6 unknowns; taking $Y_H$ as a parameter the system is readily solved (barring the quasi-trivial option):

\begin{align}
Y_Q = +\frac{1}{3}Y_H, \quad Y_u = +\frac{4}{3}Y_H, \quad Y_d = -\frac{2}{3}Y_H, \quad Y_L = -Y_H, \quad Y_e = -2Y_H. \quad (2.12)
\end{align}

Hence, the anomaly-free hypercharges compatible with the SM data are homogeneous in $Y_H$ and, thus, their ratios are fixed to be fractions of small integers!

**Cancellation of chiral anomalies in the SM+3 RH $SU(3)_c \otimes SU(2)_L$ singlets**

Naïvely, one would expect little change in the game above if a RH neutrino field $N_R$ (i.e., a full singlet with respect to the $SU(3)_c \otimes SU(2)_L$ of the SM) with an a-priori unknown hypercharge $Y_N$ is admitted into the play; indeed, with an extra unknown there is also a new constraint stemming from the existence of the corresponding Yukawa term \(^5\) so the situation should not change qualitatively at all. But it does! The point is that the resulting system of 6 equations, namely, (2.8)-(2.11) which remain intact, together with a slightly modified version of Eq. (2.7) and an extra Yukawa-sector condition

\begin{align}
6Y_Q^3 + 2Y_L^3 - 3Y_u^3 - 3Y_d^3 - Y_e^3 - Y_N^3 = 0, \quad (2.13)
\end{align}

and an extra Yukawa-sector condition

\begin{align}
-Y_L + Y_N - Y_H = 0, \quad (2.14)
\end{align}

is no longer independent. This may be readily seen by substituting into (2.13) for $Y_Q$ from (2.8) and for $Y_u$, $Y_d$, and $Y_e$ from (2.9)-(2.11) which yields

\begin{align}
(Y_H + Y_L)^3 - Y_N^3 = 0. \quad (2.15)
\end{align}

This, however, is trivially fulfilled if (2.14) holds. Hence, the general solution here is parametrised by 2 leftover quantities, e.g., $Y_H$ and $Y_N$, and it turns out to be inhomogeneous in either of the two. Normalizing the entire set so that $Y_H = +\frac{1}{2}$ one

\footnote{For simplicity, from now on we shall drop the chirality labels of the matter fields $f_{L,R}$ in $Y_f$'s.}
reveals
\[ Y_Q = +\frac{1}{6} - \frac{1}{3}Y_N, \ Y_u = \frac{2}{3} - \frac{1}{3}Y_N, \ Y_d = -\frac{1}{3} - \frac{1}{3}Y_N, \ Y_L = -\frac{1}{2} + Y_N, \ Y_e = -1 + Y_N \]
(2.16)

with no constraint of \( Y_N \)! Thus, there is a one-parametric class of admissible hyper-charges out of which only some (those with \( Y_N \in \mathbb{Z} \)) correspond to a quantized set.

Needless to say, this results has profound implications for the understanding of the 1 per 10\(^{21}\) level of neutrality of neutrons \[34\] – in the SM extension with 3 generations of RH neutrinos advocated in Sect. 2.1.1 the rationale for this to be the case is simply gone!

\section*{U(1)\(_{B-L}\) as a gaugeable symmetry in the 3 RH neutrino context}

The peculiar qualitative change in the behaviour of solutions to the SM chiral anomaly cancellation conditions in presence of three RH SU(3)c\(\otimes\)SU(2)\(_L\) singlets may be understood rather simply by inspecting the linear dependency of the “dequantized” hyper-charges \( (Y'_f) \) of Eq. (2.16) on the (apriori unknown) real parameter \( Y_N \). Interestingly, one can write \( Y'_f \)'s as
\[ Y'_f = Y_f + x(B - L)_f \]
(2.17)

where \( Y_f \) stands for the SM solution (2.12), \( x = -Y_N \) is a real number and \( (B - L)_f \) is the difference of the baryon and lepton number charges of the \( f \)-type fermions. Since, however, \( U(1)_{B-L} \) can be promoted to a full-fledged gauge symmetry in the presence of 3 RH neutrinos\(^6\), any linear combination of the non-anomalous \( Y_f \) of the SM and another non-anomalous and potentially local \( B - L \) charge as in Eq. (2.17) is also a candidate for a non-anomalous \( U(1) \) gauge symmetry generator that may play the role of an alternative SM hypercharge. Again, the beautiful and profound mechanism of the SM (hyper)charge quantization by means of the chiral anomaly cancellation requirements is lost in the SM extension with three RH SU(3)c \(\otimes\) SU(2)\(_L\) fermionic singlets.

\subsection*{2.1.3 Majorana fermions}

There are two basic approaches to dealing with this rather unpleasant situation:

1. In principle, one does not need to add three copies of the fermionic SU(3)c \(\otimes\) SU(2)\(_L\) RH singlets. Hence, \( B - L \) would not be anomaly free as a potentially local charge and the hypercharge redefinition freedom (2.17) would not exist. This

\(^6\)Note that in the SM \( B \) as well as \( L \) are individually anomalous but \( B - L \) is a non-anomalous global symmetry. However, the \( (B - L)_3 \) current is still anomalous and, hence, \( B - L \) can not be promoted to a local symmetry in the SM. The presence of one extra SU(3)c \(\otimes\) SU(2)\(_L\) RH gauge singlet with \( L = 1 \) per generation ensures that \( A_{(B-L)_3} \) also vanishes.
is the case of, e.g., the two minimal alternatives to the canonical type-I seesaw based on the scalar or fermionic $SU(2)_L$ triplets (see also Sect. 2.2.2) which, by construction, deprive the (local) $B - L$ anomaly cancellation mechanism from a crucial ingredient.

2. Alternatively, one can entirely dismiss any thoughts about “gaugeability” of $B - L$ by writing down a Majorana mass term for the $SU(3)_C \otimes SU(2)_L$ RH fermionic singlets $N_R$ which breaks this symmetry explicitly.

The possibility to have a massive spinor with only two dynamical components was first noticed by Ettore Majorana in 1937 [37]. It can be heuristically understood by counting the degrees of freedom: A massive charged spin-$1/2$ particle is described by a four-component object because the associated quantum field (Dirac spinor) must be able to describe the two opposite-charge states of the particles and antiparticles, together with the two helicity degrees of freedom for each. This number can be reduced from four to two by only two means; either giving up half of the helicity states (reducing the original Dirac to a Weyl spinor describing a massless particle) or giving up all charges and, thus, any distinction between a particle and an antiparticle (a Majorana spinor).

**Majorana mass term for RH neutrinos**

Assuming that the extra RH spinors discussed above are completely $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ neutral the free part of the relevant Majorana Lagrangian can be written as

$$\mathcal{L}_M = i N_R \gamma^\mu \partial_\mu N_R - \frac{1}{2} M_M N_R^T C N_R + h.c.$$  \hspace{1cm} (2.18)

where $M_M$ is, in general, a complex mass parameter (a symmetric complex matrix if more than a single $N_R$ is considered), $C$ stands for the spinorial charge conjugation matrix and the extra factor of $1/2$ in the second (Hermitean) term is there in order to compensate for the double counting due to the omnipresent $+ h.c.$ associated with all other mass/Yukawa terms in the SM Lagrangian. A couple of comments may be worth here:

1. It is trivial to see that the asymptotic states created from the vacuum by the rising operators in $N_R$ obey the standard relativistic free-particle dispersion relation of the form $p^2 - M_M^2 = 0$.

---

Scenarios in which $U(1)_{B-L}$ is broken spontaneously are discussed at length in Chapter 3.

More precisely, the internal symmetry representation under which such an object transforms should be real, admitting, for instance, a triplet of $SU(2)$ or an octet of $SU(3)$.
Due to the symmetry of \((\text{the matrix version of})\ M_M\) and the equality of the associated unitary matrices \(U\) acting on it from the left and right sides there is no residual freedom in the phase redefinition of \(U\) that would preserve the real diagonal form of \(M_M\). This, in turn, makes it impossible here to absorb all 5 out of 6 complex phases of the “raw form” of \(U^{\text{PMNS}}\) in \((2.3)\) as in the Dirac case of Sect. 2.1.1 but only 3. Thus, there are in principle three irremovable (i.e., physical) CP-violating phases left in the game (for three generations of \(N_R\)) which may play a role in situations/processes related to the Majorana nature of \(N_R\) (such as lepton number violation, see below).

**Type-I seesaw mechanism and the typical RH neutrino mass scale**

In practice, i.e., in case of the simplest SM extensions with three \(N_R\)’s, the Majorana mass term (2.18) should not be alone in the full Lagrangian otherwise the SM neutrinos will still be massless. This potential issue is readily resolved by the addition of the Yukawa term (2.1); the complete structure of the \(N_R\)-Lagrangian then reads

\[
\mathcal{L} \ni -Y_\nu L_L N_R \bar{H} - \frac{1}{2} M_M N_R^T C N_R + h.c. \quad (2.19)
\]

Obviously, the key is the mixing of \(N_R\) with the neutral components of the SM \(L_L\) provided by the Yukawa term in (2.19). In the broken phase, the non-derivative part of the electrically neutral sector of the structure above can be readily rewritten as

\[
\mathcal{L}^{\text{mass}} = -\frac{1}{2} n_L^T C M n_L + h.c., \quad (2.20)
\]

where \(n_L^T = (\nu_L, N_R^c)^T\) and

\[
M \equiv \begin{pmatrix} 0 & m_D \\ m_D^T & M_M \end{pmatrix} \quad (2.21)
\]

is a \(6 \times 6\) complex symmetric matrix which, besides the Majorana part \(M_M\), includes a general complex \(3 \times 3\) Dirac-type mass matrix defined as

\[
m_D \equiv \frac{1}{\sqrt{2}} Y_\nu v. \quad (2.22)
\]

The asymptotic neutrino states are obtained by a suitable diagonalisation of this matrix which, unlike in the Dirac case\(^{[10]}\) yields in general 6 different (in size, not only in sign) eigenvalues describing 6 independent Majorana spinors!

\(^9\)Here we assume three generations of \(N_R\) being added to the SM in order to treat all its three generations symmetrically. However, from the phenomenology point of view this is not strictly necessary as one can be well off even with just two copies of \(N_R\).

\(^{10}\)Note that for Dirac neutrinos \(M_M = 0\) and the same diagonalisation procedure yields pairs of two-component eigenstates with eigenvalues of the same size but opposite CP parities which is a situation equivalent to having 3 four-component objects.
Dynamically, these eigenstates can be in most cases organised into two triplets representing a light and a heavy part of the neutrino sector. This is a consequence of the following simple argument:

- Unlike for \( m_D \) whose magnitude is tightly connected to the electroweak symmetry breaking, \( M_M \) is an explicit singlet mass term that has nothing to do with the electroweak scale \( v \) and, as such, it may be significantly larger\(^{11}\) than \( v \).

- For \( |m_D| \ll |M_M| \) the diagonalisation of (2.21) proceeds in two steps, the first of which (a diagonalisation into \( 3 \times 3 \) blocks) yields

\[
\mathbf{M} \equiv \begin{pmatrix} m_\nu & 0 \\ 0 & M_N \end{pmatrix},
\]

with

\[
m_\nu \equiv -m_D M_M^{-1} m_D^T + O(|m_D|^3 M_M^{-2})
\]

and

\[
M_N \equiv M_M + O(|m_D|),
\]

respectively. Thus, one ends up with three \textit{light} SM-like neutrino mass eigenstates, dominated by the \( SU(2)_L \)-doublet components \((\nu_L)\) with just a tiny admixture of the singlet \((N_R)\) ones, with a mass term in the form

\[
\mathcal{L}^{\text{light}} = -\frac{1}{2} \nu^T C m_\nu \nu + h.c.,
\]

and three heavy eigenstates with masses driven by

\[
\mathcal{L}^{\text{heavy}} = -\frac{1}{2} N^T C M_N N + h.c.,
\]

which are singlet-dominated and, thus, practically sterile with respect to the SM gauge interactions\(^{12}\).

Note that the stipulated hierarchy between the Dirac and the RH Majorana mass terms \( |m_D| \ll |M_M| \) makes it possible to attribute the smallness of \( m_\nu \) to the \( M_M^{-1} m_D^T \) factor in (2.24) and, thus, unload the enormous \( |Y_\nu| \lesssim 10^{-11} \) suppression imposed on the Yukawa couplings in the Dirac neutrino case of Sect. 2.1.1. In principle, even \( O(1) \) entries in \( Y_\nu \) are admissible\(^{13}\) as long as \( M_M \) falls into the ballpark of

\[
M_M \sim 10^{12-13}\text{GeV}.
\]

\(^{11}\)To this end, it is usually assumed that \( M_M \) is generated by the breaking of a higher gauge symmetry encompassing that of the SM in such a way that \( N_R \) transforms non-trivially under its action.

\(^{12}\)It is worth emphasising here that this does not mean that the heavy Majorana neutrinos do not entertain any interactions - one should not forget about the Yukawa ones!

\(^{13}\)This assumption is actually easy to justify as in most SM extensions discussed in Chapter. 3 the neutrino Yukawa couplings are naturally correlated with those or the up-type quarks by the extended symmetries of the underlying Lagrangians.
Let us also mention that the inverse proportionality of $m_{\nu}$ to $M_M$ earns this scheme the traditional name of “seesaw mechanism” \[38\]. However, one should bear in mind that, as simple as it is, this is by far not the only dynamical way to devise naturally light massive neutrinos in simple extensions of the SM, cf. Sect. 2.2.

### 2.1.4 Massive Majorana neutrino phenomenology

Remarkably enough, even with just a tiny mixing among the $SU(2)_L$ doublet and singlet components in the physical neutrino spectrum (driven by the off-diagonal $O(|m_D|/|M_M|)$ factors in the unitary transformation bringing the defining-basis mass matrix (2.22) into the block-diagonal form (2.23)) the low-energy phenomenology of the seesaw schemes can differ drastically from that of the SM:

- **Perturbative lepton number violation:** The first obvious difference is the perturbative non-conservation of the lepton number owing to the shape of Eq. (2.26) which, in the SM, is a good accidental global symmetry, at least at the perturbative level\[15\]. This may, in principle, exhibit itself in extremely-low-background processes where even a minuscule effect can be deciphered.

- **Leptonic CP violation:** Another common aspect of the stipulated Majorana nature of the light neutrinos is the presence of new sources of CP violation owing to the shape of the corresponding neutrino mass term (2.26). As we argued, there are in principle 3 CP phases in the leptonic sector; one of these comes as a full analogy of the CKM phase $\delta_{\text{CKM}}$ and, as such, it is called the Dirac CP phase; the other two (which can be transferred back and forth between $U_{\text{PMNS}}$ and the eigenvalues of $M_M$) are called Majorana.

Both these types of BSM effects may find rather spectacular incarnation in the laboratory experiments and in cosmology.

#### Neutrinoless double beta decay

Perhaps the most famous of such observables is the hypothetical neutrinoless double beta decay that should be undergone by some even-even nuclei for which a sequence of

\[14\text{Note that both terms are needed to claim LNV: without the first one } L(N_R) = 0 \text{ leads to a conserved } L \text{ while without the second term the same happens for } L(N_R) = +1. \text{ Needless to say, } L(\tilde{H}) \neq 0 \text{ is not an option because } H \text{ triggers the electroweak symmetry breaking.}\]

\[15\text{It is well known that this is not true at the non-perturbative level where topologically non-trivial extended field configurations (instantons, sphalerons) may be relevant in the path integration, thus leading to both baryon and lepton number non-conservation \[12\text{39].}\]
“standard” single beta decay transitions is kinematically forbidden. The most prominent of these systems are the isotopes $^{48}\text{Ca}$, $^{76}\text{Ge}$, $^{82}\text{Se}$, $^{100}\text{Mo}$, $^{130}\text{Te}$ and $^{136}\text{Xe}$ which are known to undergo a “standard” double beta decay (i.e., with two neutrinos in the final state) into their $X^A_{Z+2}$ counterparts (or $X^A_{Z-2}$ in the double $\beta^+$ channel) with enormous lifetimes and, at the same time, they can be found in the Earth’s crust in non-negligible amounts. This makes it possible to speculate about the observability of an analogous process in which the pair of Majorana neutrinos “annihilate” in the core rather than escape the nucleus (cf. Fig. 2.1) which, in turn, may be identified by its very specific final state kinematics.

Interestingly, in most of the parameter space of the simple seesaw scheme above (as well as in other scenarios) the relevant effective leptonic matrix element $\langle m_{ee} \rangle$ exhibits a relatively universal lower limit which makes us hope that, with a bit of luck (and a steady support from the funding bodies) the relevant signal may be detected within the next few decades, see, e.g., [40] and references therein. Needless to say, this would be a phenomenal achievement as it would provide the first evidence of a glitch in the assumed perturbative-level symmetry structure of the Standard Model.

Let us finally note that this claim is often further supported by the so called “Scheckter-Valle theorem” [41] which stipulates that the observation of the neutrinoless double decay implies the Majorana nature of neutrinos, thus claiming their mutual equivalence! This can be heuristically understood by “dressing” the effective $d = 9$ “$0\nu2\beta$ blackbox” operator by the SM fields which gives rise to a contribution to the light neutrino propagator of the Majorana structure, see Fig. 2.2.
Baryogenesis through leptogenesis

It is well known that the standard ΛCDM cosmology, married with the SM of particles and their interactions, fails miserably with the prediction of the baryon to photon number density ratio $\eta_B$, an all-important parameter for the modelling of the primordial nucleosynthesis during the “first three minutes” of the early Universe. Indeed, its measured value of roughly $6 \times 10^{-9}$ (see e.g. [42] and references therein) is about 10 orders of magnitude above the SM upper limit based on the simple (and plausible) assumption of the baryon-antibaryon symmetry of the post-inflationary cosmic plasma.

The only reasonable way to reconcile the theory with the measurement is to revoke this assumption and, instead, look for a dynamical origin of the approximate $10^9 + 1$ to $10^9$ baryon-to-antibaryon number density ratio that was to be established before the freeze-out of the baryonic populations in the early Universe. This, according to Sacharov [43], may have happened if:

1. Baryon number was not an exact symmetry of Nature;

2. C and CP were not good symmetries of the early Universe dynamics;

3. The system underwent an out-of-equilibrium (OOE) epoch.

The first of these criteria is trivial as one can cook a net baryon number from zero only if there are interactions capable of changing it. As for the second, this deals with the discrete symmetries correlating the pace of processes of a net $B$ creation in one chirality channel with that of the $B$ destruction in the parity-conjugated one. The third condition then ensures that the creation of $B$ at some point is not entirely undone by a reverse process occurring at the same time at some other place.
Figure 2.3: The leptonic CP-asymmetry generation in the out-of-equilibrium decay of the heavy neutral fermions in the SM extensions with RH neutrinos. The leading effect emerges from the interference between the tree-level and 1-loop diagrams.

Remarkably enough, the SM seems to fulfil all these conditions, at least qualitatively: 1. For $T > M_Z$ the sphaleron processes run fast and $B$ violation is a standard feature of the hot SM plasma; 2. Both $C$ and $CP$ are, indeed, violated by the weak interactions and 3. the Universe may have fulfilled the OOE Sacharov’s condition if the electroweak phase transition was strongly first order.\(^\text{16}\)

However, the situation turns out to be way less optimistic as neither condition 2. nor 3. are met in the Standard Model on the quantitative level. As for the former, CP violating dynamics turns out to be subject to strong screening in the hot and dense plasma, see e.g. \cite{46}. Moreover, for $m_H \sim 125$ GeV the SM electroweak phase transition was very likely 2nd order and, hence, the yield of the asymmetry generated through the associated electroweak baryogenesis process is expected to be parametrically smaller than the desired $1:10^9$.

Interestingly, the heavy Majorana sector of the seesaw extension of the SM discussed in Sect. 2.1.3 offers a very elegant way out of this conundrum \cite{47}. The basic idea is that the actual baryogenesis process may have been preceded by an epoch of spontaneous leptogenesis in which a net lepton number was created in the $C$ and $CP$-asymmetric out-of-equilibrium decays of the heavy Majorana neutrinos \cite{2.27} and, only later on, it was (partially) transferred into the baryons. A careful analysis of the central quantity of interest, namely, the CP asymmetry of the (lightest in most cases) heavy Majorana neutrino decays, see Fig. 2.3

\[ \varepsilon_{CP} = \frac{\Gamma(N \rightarrow LH) - \Gamma(N \rightarrow LH^*)}{\Gamma(N \rightarrow LH) + \Gamma(N \rightarrow LH^*)}, \] \(^2.29\)

reveals an upper limit \cite{48,49} on the CP asymmetry of the SM lepton production in

\(^{16}\)In such a case the baryogenesis processes would take place at the surface of the asymmetric phase bubbles expanding at almost the speed of light \cite{44,45} into the still symmetric false vacuum environment.
the form
\[ \varepsilon_{CP} \leq \frac{1}{8\pi} \frac{M_1(m_3 - m_2)}{v^2}, \] (2.30)
which, for the desired \(\varepsilon_{CP} \gtrsim 10^{-7}\) yields a lower limit on the mass of the lightest heavy neutrino in the form
\[ M_1 \gtrsim 10^9 \text{GeV}. \] (2.31)
This, indeed, is perfectly compatible with the “seesaw” picture of the light neutrino mass generation discussed above, cf. Eq. (2.28).

### 2.2 Majorana neutrino mass as a \(d = 5\) operator in the Standard Model

The arguments of the previous subsection suggest that the characteristic energy scale of the new dynamics that may be behind the light neutrino masses and other effects discussed in Sect. 2.1.4 is likely to be rather high, actually many orders of magnitude above the electroweak scale. This, at one hand, diminishes its potentially harmful impact to other precision data but, on the other hand, makes it generally difficult to test any specific shape of such a new dynamics in any but few particular channels.

#### 2.2.1 Theories with vastly different scales

Technically, this has to do with one of the most general features of quantum physics often phrased as the “independence of the low-energy\(^{18}\) observables on any ‘new dynamics’ with a parametrically larger characteristic energy scale”. It is sometimes (not very rigorously) justified in the language of the uncertainty principle which, in principle, makes it possible for an “extremely heavy virtual particle pair”\(^{19}\) to be “created” out of the quantum vacuum but, as a price for its “large mass”, such an “object” may “exist” for only a very short “time” and, as such, it can not affect any observable at a substantial level.

\(^{17}\)Note that the stipulated lower limit on \(\varepsilon_{CP}\) is slightly more conservative than the experimental constraint of \(\eta_B\). The reason here is the presence of the so called “washout” effects corresponding to the LNV re-scattering processes undergone by the light leptons and antileptons during the asymmetry generation era which can effectively wipe out a significant portion of the asymmetric decay yield.

\(^{18}\)This statement should perhaps be put into a better perspective by reminding the reader about how the notion of “low-energies” evolved over the last century - from Lord Rutherford’s few MeV standpoint the CERN’s LHC with its 14 TeV centre of mass energy would have likely been a machine beyond imagination, yet it may soon become a mere pre-accelerator for a yet more powerful monster - the FCC [50].

\(^{19}\)The quotation marks here indicate the care required for the proper interpretation of the terms enclosed.
The Appelquist-Carazzone theorem

In the realm of the quantum field theory these expectations find their formal incarnation in the so called decoupling theorem by Thomas Appelquist and James Carazzone [51] which, in one of its possible formulations, stipulates the inverse-power dependency of the renormalized Green’s functions of a low-energy sector of a certain theory (calculated in momentum renormalization schemes) on the masses characterising the heavy part of its spectrum. Several comments are perhaps worth making here:

1. The main scope of the theorem is to deal with UV-divergent loop graphs where the interchangeability of the integration and limit operations is not guaranteed. It is trivial though for finite graphs such as tree-level ones or loops with negative degrees of divergence.

2. The need for the renormalizability of the low-energy sector alone is easy to understand on a qualitative basis: In such a case, every singlet mass parameter of the complete theory may be viewed as a mere cut-off \( \Lambda \) in the momentum integration within the low-energy theory graphs; as such, it should disappear from the physics of the renormalizable low-energy sector in the \( \Lambda \to \infty \) limit.

3. On the formal side, the amplitudes (and/or other quantities of interest such as masses, if such a distinction is desired\(^{20}\)) of the low-energy theory entertain an explicit decoupling behaviour only in physical schemes, i.e., when the low-energy observables are parametrised in terms of physical quantities (or, at least, quantities defined in some of the momentum schemes). In schemes in which the definition of the counterterms is more or less ad-hoc (such as minimal subtraction etc.) with little reference to the underlying dynamics the relevant formulae do not need to exhibit any apparent decoupling of the heavy sector.

4. One more subtlety is perhaps worth pointing out here: it is implicitly understood that the dimensionless couplings between the light and heavy sectors do not grow more than logarithmically (as suggested by the renormalization group) with the heavy sector masses. Along with the renormalizability argument, this is the reason why, for instance, the top quark can not be formally decoupled from the Standard Model – indeed, its Yukawa coupling grows with \( m_t \).

As an illustration of especially point 3 above consider a “light” \( \phi^4 \) theory

\[
L_\phi = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - \lambda \phi^4
\]  

\(^{20}\)Personally, the author prefers to talk about renormalized Green’s functions instead as they provide a unified language for all the aforementioned physical aspects.
coupled to a heavy version of the same
\[ \mathcal{L}_\Phi = \frac{1}{2} \partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2} M^2 \Phi^2 - \rho \Phi^4 \quad (2.33) \]
through a trilinear interaction term
\[ \mathcal{L}_{\phi \Phi} = \kappa \phi \Phi^2, \quad (2.34) \]
where \( \kappa \) is a dimensionfull parameter. The light-sector two-point 1PI Green’s function in dimensional regularisation before imposing renormalization conditions then reads (displaying explicitly only the important heavy sector contribution)
\[ \Gamma_{\phi \phi}(p^2) \propto p^2 - m^2 + \frac{\kappa^2}{16\pi^2} \left( \frac{1}{\epsilon} - \int_0^1 dx \log \frac{M^2 - x(1-x)p^2}{4\pi\mu^2} \right) + \ldots + \delta Z_\phi p^2 + \delta m^2, \quad (2.35) \]
where \( \delta Z_\phi \) and \( \delta m^2 \) are the “light” theory counterterms. These are in general different in different renormalization schemes:

- The on-shell scheme renormalization conditions
  \[ \Gamma(p^2 = m^2) = 0, \quad \frac{d}{dp^2} \Gamma(p^2 = m^2) = 1 \quad (2.36) \]
  which, among other things, immediately identify the physical mass of the light scalar with \( m \), yield
  \[ \delta Z_\phi = -\frac{\kappa^2}{16\pi^2} \log \frac{x(1-x)}{M^2 - x(1-x)m^2}, \]
  \[ \delta m^2 = -\frac{\kappa^2}{16\pi^2} \left[ \frac{1}{\epsilon} - \int_0^1 dx \left\{ \log \frac{M^2 - x(1-x)m^2}{4\pi\mu^2} - \log \frac{m^2 x(1-x)}{M^2 - x(1-x)m^2} \right\} + \ldots, \]
  and thus
  \[ \Gamma_{\phi \phi}^{OS}(p^2) \propto p^2 - m^2 - \frac{\kappa^2}{16\pi^2} \int_0^1 dx \log \frac{M^2 - x(1-x)p^2}{M^2 - x(1-x)m^2} + \frac{(p^2 - m^2)x(1-x)}{M^2 - x(1-x)m^2} + \ldots. \quad (2.37) \]

It is trivial to see from here that in the \( M \to \infty \) limit the contribution of the heavy sector entirely drops from \( \Gamma^{OS} \), i.e., the decoupling of the heavy sector from the light one is apparent.

---

21Note that the specific form (2.34) of the cross-talk between the light and heavy sectors of the scheme has been chosen only for the calculational simplicity; other choices (such as \( \phi^2 \Phi^2 \)) lead to the same conclusions.

22Here and in what follows we shall often use “rationalised natural units” in which all the multiplicative factors irrelevant for the merit of the argument are set to 1 and the additive ones to 0.
On the contrary, in the minimal subtraction (MS) scheme \( \delta Z_\phi = 0 \) and
\[
\delta m^2 = -\frac{\kappa^2}{16\pi^2} \frac{1}{\varepsilon}.
\] (2.38)

Hence, the renormalized two-point 1PI Green’s function in the MS scheme reads
\[
\Gamma_{\phi\phi}^{MS}(p^2) \propto p^2 - m^2 - \frac{\kappa^2}{16\pi^2} \int_0^1 dx \log \frac{M^2 - x(1-x)p^2}{4\pi\mu^2} + \ldots
\] (2.39)

For \( M^2 \gg p^2, m^2 \) the root of this function (corresponding to the one-loop physical mass of \( \phi \)) is clearly far from \( p^2 \sim m^2 \); actually, in the \( M \to \infty \) limit it can be readily isolated at around
\[
p^2 \sim m^2 + \frac{\kappa^2}{16\pi^2} \log \frac{M^2}{4\pi\mu^2} \equiv m_{\text{phys}}^2
\] (2.40)
which, formally, grows with \( M \). Note, however, that the decoupling theorem still works here, only in a slightly different manner: writing the RHS of Eq. (2.39) in terms of \( m_{\text{phys}}^2 \) rather than \( m^2 \) one is taken to the situation which (up to a redefinition of certain parameters) is qualitatively identical to the on-shell setting above.

**Low-energy effective theories and their symmetries**

From a wider perspective, the A-C theorem suggests a very practical approximative approach to the theories with vastly different scales. The inverse-power proportionality to the heavy sector masses in the low-energy Green’s functions of the light sector (as observed in (2.37) with the log expanded) justifies an effective parametrisation of the low-energy amplitudes in terms of local operators with positive powers of \( M \) in denominators. These, due to the gauge-singlet nature of the contractions corresponding to the complete-theory internal heavy field propagators must respect all the symmetries of the full theory. These operators are equipped with generic numerical coefficients to be determined by matching such an effective Lagrangian \(^2\) to the complete theory amplitudes in the low-energy regime.

In general, such a matching procedure consists in integrating out the heavy degrees of freedom from the complete theory in its path-integral formulation\(^2\). This, besides providing relations between the fundamental and the effective couplings also highlights those classes of effective operator structures which are “available” within any specific scheme of interest.

\(^2\)In the canonical formalism, this is usually approached by solving the (classical) equations of motion for the heavy fields and substituting the result back into the Lagrangian. The error committed by working at the classical level here can be shown to be comparable with the effects of higher orders in the perturbative expansion, cf. \(^\).
2.2.2 Standard Model as a low-energy effective theory

From this perspective, any perturbative high-energy extension of the SM can be mapped – in the low-energy regime – to (a specific subset of) the complete list of \( d > 4 \) effective operators. At each level of dimensionality, such a list contains a number of structures which are supposedly independent with respect to the transformations that should leave the physics contents of the theory intact (e.g., total-derivative completion, use of the equations of motion, Fierz transformation etc.). A systematic classification of such “independent operator bases” for various dimensions can be found in, e.g., [54–56].

\[ \text{d=5 Weinberg’s operator and its renormalizable tree-level openings} \]

As an example of the practical application of these principles let us turn our attention back to the situation discussed in Sect. 2.1.3, namely, to the low-energy behaviour of the SM with three RH neutrinos with masses way above the EW scale, cf. (2.28) or (2.31). The relevant part of the Lagrangian density

\[
\mathcal{L}_{NR} = i \bar{N}_R \gamma^\mu \partial_\mu N_R - \frac{1}{2} M_M N^T_R C N_R - Y_\nu L_L N_R \tilde{H} + \text{h.c.} \quad (2.41)
\]

yields the Euler-Lagrange equations of motion for \( N_R \) in the low-energy regime (i.e., ignoring the kinetic term contribution) in the form\(^{24}\)

\[
N_R = -M_M^{-1} Y^T_\nu (H^T i \sigma_2 L_L)^c + \ldots \quad (2.42)
\]

This, substituted back into Eq. (2.41), provides a \( d = 5 \) effective structure

\[
\mathcal{L}_{\text{eff}} = \kappa (L_L i \sigma_2 H)^T C (L_L i \sigma_2 H) + \text{h.c.}, \quad (2.43)
\]

where the brackets encompass the \( SU(2) \) doublet contractions (with the transposition acting in the spinorial space) and \( \kappa \equiv -Y_\nu M_M^{-1} Y^T_\nu / 2 \). Equation (2.43) corresponds to the famous non-renormalizable operator identified by Weinberg in his seminal paper [9].

Remarkably enough, in the same study it was also shown that this structure is \textit{unique} at the \( d = 5 \) level of the higher-dimensional operator ladder. This, however, means that in the broken phase one should recover the \( d = 3 \) Majorana operator for the light neutrinos in the form (2.26) which, given (2.22) and (2.24), is indeed the case.

\(^{24}\)Here it turns out to be way more convenient to rewrite all the Dirac-type contractions into the Majorana formalism.
Seesaw type I: From this perspective one can view the SM extension with the RH neutrinos as a specific example of the high-scale renormalizable dynamics that, in the low-energy regime, exhibits itself (predominantly) in the form of the $d = 5$ interaction (2.43) which, among other things, generates a Majorana mass term for the light neutrinos. Historically, this scheme comes under the name of the “Type-I seesaw mechanism” [38, 57]. It is interesting (though perhaps not surprising) that the type-I

Seesaw type II: Instead of a singlet fermion of Sect. 2.1.3 one can employ a scalar [58, 59] in a “t-channel” type of structure as depicted in Fig. 2.6. An $SU(2)_L$ scalar triplet $\Delta_L$ with hypercharge +1 can be coupled to the pair of $L_L$’s via a Yukawa term like (neglecting generation structure for simplicity)

$$\mathcal{L}_{\Delta_L}^Y = Y_{\Delta_L} L_L^T C i \sigma_2 \bar{\sigma} L_L \Delta_L + h.c., \quad (2.44)$$

which, together with a super-renormalizable coupling to the Higgs doublet pair of the form

$$\mathcal{L}_{\Delta_L}^H = \mu H^T i \sigma_2 \bar{\sigma} H \Delta_L + h.c. \quad (2.45)$$

Figure 2.4: A Feynman diagram visualisation of the $d = 5$ Weinberg’s operator.

Figure 2.5: The "microscopic" structure of the light neutrino mass within the type-I seesaw scheme. The $N_R$ field is an $SU(2)_L$-singlet fermion with hypercharge 0. The seesaw dynamics is not the only way to realize the Weinberg’s $d = 5$ operator in Fig. 2.4 within a renormalizable high-energy dynamical scheme. At the level of its tree-level “openings” there are actually two more options.
and the corresponding kinetic and mass terms

$$L^{\text{free}}_{\Delta L} = \partial^\mu \Delta_L^\dagger \partial_\mu \Delta_L - M_\Delta^2 \Delta_L^\dagger \Delta_L,$$

yield again a structure like that in Eq. (2.43) with

$$c = \mu Y_\Delta^2 / M_\Delta^2. \quad (2.46)$$

Figure 2.6: The "microscopic" structure of the light neutrino mass within the type-II seesaw scheme. The $\Delta_L$ field is an $SU(2)_L$-triplet scalar with hypercharge +1.

It is worth noting that the triplet nature of $\Delta_L$ with respect to the $SU(2)_L$ is enforced by the doublet nature of $L$ and $H$ in (2.44) and (2.45) - the antisymmetry of singlet contractions is incompatible with the gauge structure of either of these terms.

Another comment concerning the role of different scales in the game is perhaps worth here: unlike in the type-I case the type-II seesaw effective light neutrino mass is driven by two mass parameters ($\mu$ and $M_\Delta$). This, for $O(1)$ couplings, makes it possible to get $m_\nu$ in the eV ballpark even if $M_\Delta$ is relatively light (TeV-scale) as long as $\mu$ is appropriately small. Hence, the type-II seesaw mechanism may work in a mode in which it leaves behind interesting collider signatures! An interested reader can find more information in, e.g., [60] and references therein.

Seesaw type III: Interestingly, yet another tree-level renormalizable opening of the Weinberg’s operator (2.43) can be devised with an $SU(2)_L$-triplet fermion [61] instead of the type-I singlet. This so called type-III seesaw scheme (see Fig. 2.7) differs from the type-I variant in two aspects only:

- The “mediator” is a gauge non-singlet field and, as such, its excitations can be produced at the colliders if they are light enough (i.e., for $M_F$ in the TeV ballpark). This, however, comes for the price of very tiny Yukawa couplings.

- With a perturbatively conserved lepton number assigned to $F$ the $B - L$ local anomaly does not vanish. Hence, $U(1)_{B-L}$ cannot be gauged in this scheme.

\footnote{Note that, to this end, the Pauli matrices in (2.44) and (2.45) can be eliminated by means of the relevant completeness relations.}
Naturalness of perturbative lepton number violation in BSM physics

From a wider perspective, it is very interesting that the evidence of non-zero neutrino masses – the first solid experimental signal of physics beyond the Standard Model – together with the desire to preserve the beautiful mechanism of charge quantization by anomalies, point straight towards the Weinberg’s operator (2.43), the lowest-dimensional structure in the effective-operator ladder of the SM. In this respect, the perturbative lepton number violation, as arcane as it looks at the renormalizable SM level, seems to be absolutely natural for its simplest phenomenological extensions.

2.3 Standard Model at $d = 6$ level and perturbative baryon number violation

Without much of exaggeration one can perhaps even say that the observation of the LFV effects as the first signal of the BSM physics was, in fact, expected as it corresponds to the least suppressed $d = 5$ layer revealed upon opening up the Pandora’s box of the SM effective operators. There is no reason, however, for the story to stop there, the more that the number of independent gauge-invariant effective structures grows exponentially with $d$. Actually, there are almost 60 different operator types in the SM at the $d = 6$ level, about 1000 at $d = 8$ and so on [56].

However, in what follows we shall dive just one more level beneath the Weinberg operator discussed so far, i.e., we shall discuss (a carefully selected subset of) the $d = 6$ SM effective operators only. For our current needs it will be sufficient to divide them into three basic classes:

1. Operators generating effects which are strictly forbidden (at the perturbative level) in the SM. With essentially no SM background these structures have a great

---

26 A full discussion of all classes of the SM $d = 6$ operator and the specific structures within would go way beyond the scope of this thesis; an interested reader can find further information in, e.g., [51,53].
potential to provide “clean” signals in the experimental environment of any kind at any of the three major fronts of the HEP research (energy, intensity, precision). As we shall see, they typically break accidental global symmetries of the SM (such as baryon and lepton numbers). Actually, it is quite remarkable that this set is not empty!

2. Operators corresponding to effects which are tree-level forbidden in the SM. In this class there are typically operators triggering, e.g., flavour-changing neutral currents or CP effects (such as electric dipole moments) beyond those in the Standard Model. As for the former, these correspond to effective 4-fermion interactions while the latter come in the form similar to the Pauli operator in QED (modulo the extra Higgs field necessary to satisfy the $SU(2)_L \otimes U(1)_Y$ symmetry requirements). The associated effects are often heavily suppressed in the SM and, hence, the backgrounds in the corresponding experimental searches are likely to be under control.

3. Operators corresponding to effects with significant SM backgrounds. This is the richest of the three sets. However, from the perturbative expansion perspective, the operators here typically yield higher-order corrections to the SM amplitudes and, hence, their effects are usually buried under a significant burden of the SM background.

For practical reasons, in the rest of this writing we shall fully focus on the first class above. The point is that the absence of any perturbative backgrounds makes it possible to look for the associated physics even if the relevant suppression scale is very far above the electroweak scale, perhaps even close to the Planck mass.

As for the others, let us just note that the second type is a typical encounter in supersymmetric theories where the effective suppression scale is, for various reasons, expected to be in the TeV ballpark (and, hence, often becomes a problem rather than a feature of interest). Finally, almost any new physics with coloured heavy fields contributes to the third class.

2.3.1 $B$ and $L$ – violating $d = 6$ operators and their phenomenology

There are basically four different types of $d = 6$ baryon and lepton number violating structures quoted in the recent classification studies, see Table 2.1. From their structure, it is clear that these operators can have non-zero matrix elements between

\[27\text{Recall, for instance, that FCNC's appear at the one-loop level in the SM while the leptonic EDMs} \]
Table 2.1: One of many possible representations of the gauge structure of the four different types of BLNV terms one can write into the SM effective Lagrangian at the $d = 6$ level. For simplicity, all fields have been written in the left-handed chirality convention and all indices have been suppressed; $C$ is the charge conjugation matrix. Note that the above structures may be recast in many different ways using, e.g., Fierz identities or utilising the Majorana instead of the Dirac language; for instance, $O_1$ is identical to $Q_{duq}$ in [55] etc.

Proton decay as a spectacular signal of $d = 6$ perturbative BLNV

On the phenomenology side, this selection rule suggests that baryons can transmute into leptons (+ objects with zero $B$ and $L$) at the $d = 6$ level of the effective operator ladder! The most spectacular of such processes is then the decay of the lightest of (color-singlet) baryons, the proton, into an antilepton and a meson. While the first is required by the statistics and phase space considerations, the latter is needed for the energy-momentum conservation. Kinematically, the processes with the maximum phase space available for the products are those with $\pi^+ \nu$ and $\pi^0 e^+$ in the final states, but there is a plethora of other options stretching up to $K^* \mu^+$ or even $\eta' e^+$ two-body decays. Note also that proton decay into $\tau +$ anything is impossible!

---

27 can be generated at four loops only, cf. Sect. 2.1.1.

28 Note that, until recently, the situation concerning the correct counting of individual operators, their “types”, or the numbers of independent terms needed for encompassing all of their effects at the level of Lagrangians was far from clear even for such a low dimensionality as $d = 6$; in this respect, the steady progress on new techniques such as those based on the Hilbert series (see, e.g., [65] and references therein) entertained in the last decade eventually lead to a full reconciliation among different claims.

29 It is interesting to compare this with the $\Delta B = 0$ and $\Delta L = 2$ structure of the $d = 5$ Weinberg operator.
Proton decay searches and lifetime limits

The history of (so far unsuccessful) proton decay searches and the corresponding lower limits on its lifetime is extremely rich and there is not enough room here to pass through it. Instead, in what follows, we shall recapitulate just the main phases of this undertaking which stretch from the first simple (essentially table-top) experiments up to the current monstrous facilities such as Super-Kamiokande [27].

Perhaps the first solid quantitative information about proton lifetime came in 1950’s from Maurice Goldhaber who noticed that an overly fast proton disintegration in living tissues would have devastating effects. His estimate of \( \tau_p \gtrsim 10^{19} \) years (based on the contemporary estimate for lethal dose for humans at the level of about \( 10^{14-15}\) MeV absorbed per year) made it clear that backgrounds (radioactive as well as cosmic) would play a significant role in any dedicated \( p \)-decay searches; subsequently, experimentalists were “driven under the ground” for decades. Among the first dedicated searches one should mention namely the experiment by F. Reines and M. Crouch [66] from early 1970’s which took place in a 3200m deep “East Rand” Gold Mine in South Africa and pushed the naïve Goldhaber’s estimate up to about \( 10^{30} \) years [67].

With such an enormous lifetime limit the only chance for progress became multi-ton-scale experiments in which, optimally, the target body is transparent enough to admit full-volume detection. Hence, the era of kiloton-scale water-Cherenkov (WC) detectors begun (with KamiokaNDe [25], IMB [26], Super-K [27] and the upcoming Hyper-K [68] representing just few of these gargantuous instruments) and, to date, we are still harvesting the results obtained with this class of machines [30]. Besides price considerations, it is namely their universality (all of them are great neutrino detectors) and relative simplicity of the \( p \)-decay final state discrimination which makes this technology the primary method of experimental research in the field.

For illustration, consider the WC signature of the “golden channel” \( p \rightarrow \pi^0 e^+ \) process schematically depicted in Fig. 2.8 with almost a GeV of energy available \( e^+ \) is produced highly relativistic and a characteristic Cherenkov cone points in the direction of its flight before it stops and annihilates with an electron in water, only to produce a pair of delayed gammas. On the other side, \( \pi^0 \) decays almost instantaneously into a pair of hard \( \gamma \)’s which produce two more (fuzzy) cones from the associated electromagnetic cascades. It means that the event is in principle fully contained within the detector volume and the parent mass can be reconstructed with a good precision. This is

\[ ^{30} \text{It is perhaps worth mentioning that recently the liquid Argon time-projection chambers (LArTPC) were put forward (e.g., LBNE/LBNF/DUNE [69]) as viable and affordable alternative to water Cherenkov detectors with several distinct features which make them even better in some channels (especially those involving charged Kaons in the final state).} \]
important namely for the suppression of the backgrounds dominated by atmospheric neutrino effects.

The current best limits from the Super-K searches are at the level of $10^{34}$ years for $p \rightarrow \pi^0e^+$ and about an order of magnitude milder for those with neutrinos in the final state ($p \rightarrow \pi^0\nu$, $p \rightarrow K^+\nu$ etc.); the projected sensitivity of the future almost megaton-scale facilities (such as Hyper-K) should about ten times better.

These numbers can be readily translated into very stringent upper limits on the dimensionfull coefficients of the $d = 6$ operators of Table 2.1 which turn out to be typically in the $(10^{15}\text{GeV})^{-2}$ ballpark. Hence, the searches for baryon number violation signals in particle decays represents a great opportunity to learn about possible new dynamics at unprecedentedly large scales.

2.3.2 Tree-level renormalizable openings of the $d = 6$ BLNV operators

What are the implications of the numbers above for the structure and characteristic scale(s) of the SM extensions that can possibly generate the $B$ and $L$ violating operators of Table 2.1? To understand this, one has to look at their renormalizable openings as we did it with the Weinberg’s operator in Sect. 2.2.2.

1. $O_1$: In Table 2.1 $O_1$ was written in the form of a product of two vector currents so it is natural to attempt to open it “in the middle”, i.e., work with $j_1^\mu \equiv \overline{u}_L\gamma^\mu Q_L$ and $j_2^\mu \equiv d_L\gamma^\mu L_L$ which, as far as their gauge quantum numbers are concerned, transform as: $[j_1] = (\mathbf{3} \otimes \mathbf{3}, \mathbf{2}, +\frac{5}{6})$ and $[j_2] = (\mathbf{3}, \mathbf{2}, -\frac{5}{6})$. Hence, at the renormalizable level, they can both couple to a vector boson $V_1$ transforming like\footnote{Here, in the anticipation of a future embedding of $V_1$ into an adjoint representation of a gauge group $G$ of a renormalizable Yang-Mills theory we write both components of its stipulated decomposition with respect to the SM subgroup of $G$.} $(\mathbf{3}, \mathbf{2}, +\frac{5}{6}) \oplus (\mathbf{3}, \mathbf{2}, -\frac{5}{6})$. However, this is not the only way to generate $O_1$ at the
tree level – performing a Fierz transformation the same operator can be written as
\[ u^T_C d^T_L \ Q^T_L C L_L \] and the same logic as above leads to a scalar colour triplet mediator \( S_1 = (3, 1, -\frac{1}{3}) \). Actually, there is even a third option corresponding to a swapping of \( Q_L \) and \( L_L \) in the scalar contraction above (thus working with \( u^T_L C d^T_L \ L^T_L C Q_L \) instead) which does not change anything about the scalar mediator but, on the Fierz-equivalent side, opens another option for a vector mediator, namely, \( V_2 = (\bar{3}, 2, -\frac{1}{3}) \oplus (3, 2, +\frac{1}{3}) \). Thus, the single effective structure we begun with can be obtained in three different ways!

2. \( O_2 \): Again, looking first at its vector \( \times \) vector form one can identify \( V_1 \) as a possible mediator and, on the Fierz-transformed side, \( \Delta^c \) as its scalar counterpart corresponding to the reshuffled structure like \( u^T_L C e^c_L \ Q^T_L C Q_L \). There is, however, no second vector option as with \( O_1 \) because the permutation of the doublets therein does not bring anything new.

3. \( O_3 \): Since this structure is composed of only two different fields there is not much of a room for any Fierz metamorphosis here and, thus, the only structures to look at are \( Q^T_L C Q_L \) and \( Q^T_L C L_L \). Since the former transforms like \((3 \otimes 3, 2 \otimes 2, +\frac{1}{3})\) while the latter as \((3, 2 \otimes 2, +\frac{1}{3})\) there are only two types of scalar mediators that can generate it, namely, the notorious \( S_1 = (3, 1, -\frac{1}{3}) \) and a new \( S_2 = (3, 3, -\frac{1}{3}) \).

4. \( O_4 \): Here, finally, we encounter a structure that can not be Fierz-transformed in any way but which offers permutation of the three fields within. Dividing it into \( u^T_L C u^c_L \) and \( d^T_L C e^c_L \) offers \( S_3 = (3, 1, -\frac{4}{3}) \) as a possible mediator while in the permuted case with \( u^T_L C d^c_L \) and \( u^T_L C e^c_L \) one again recovers \( S_1 \).

All this information is comprehensively covered in Table 2.2. At low energies these fields would play the role of the internal lines of the Feynman diagrams corresponding to the quark-level proton decay hard processes such as

\[
\begin{array}{cc}
\begin{array}{ccc}
q & \bar{q} & V \\
q & \bar{q} & S
\end{array}
\end{array}
\quad (2.47)
\]

These, at the level of amplitudes, yield a universal structure like \( A \sim k^2/M^2 \), where \( k \) stands for the gauge or Yukawa couplings and \( M \) denotes the corresponding gauge or scalar mediator mass. At the hadronic level, this eventually boils down to

\[
\Gamma(p \to \text{antilepton} + \text{meson}) \sim \frac{k^4}{M^4} m_p^5, \quad (2.48)
\]
Table 2.2: Possible tree-level mediators of the $d = 6$ BLNV effective operators of Table 2.1 in renormalizable extensions of the SM. The symbol in the brackets in the first column corresponds to a more traditional notation we shall conform to in the rest of the text. The words “gauge” and/or “Yukawa” refers to the type of the coupling to the matter currents the mediator entertains.

<table>
<thead>
<tr>
<th>mediator</th>
<th>spin</th>
<th>quantum numbers</th>
<th>operators generated</th>
<th>couplings</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_1 (X)$</td>
<td>vector</td>
<td>$\left(3, 2, \frac{5}{6}\right) \oplus \left(3, 2, -\frac{5}{6}\right)$</td>
<td>$\mathcal{O}_1, \mathcal{O}_2$</td>
<td>gauge</td>
</tr>
<tr>
<td>$V_2 (Y)$</td>
<td>vector</td>
<td>$\left(3, 2, -\frac{1}{6}\right) \oplus \left(3, 2, \frac{1}{6}\right)$</td>
<td>$\mathcal{O}_1$</td>
<td>gauge</td>
</tr>
<tr>
<td>$S_1 (\Delta_c)$</td>
<td>scalar</td>
<td>$\left(3, 1, \frac{1}{3}\right)$</td>
<td>$\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3, \mathcal{O}_4$</td>
<td>Yukawa</td>
</tr>
<tr>
<td>$S_2 (\Delta'_c)$</td>
<td>scalar</td>
<td>$\left(3, 3, -\frac{1}{3}\right)$</td>
<td>$\mathcal{O}_3$</td>
<td>Yukawa</td>
</tr>
<tr>
<td>$S_3 (\Delta''_c)$</td>
<td>scalar</td>
<td>$\left(3, 1, -\frac{4}{3}\right)$</td>
<td>$\mathcal{O}_4$</td>
<td>Yukawa</td>
</tr>
</tbody>
</table>

where the proton mass with the fifth power emerges on the dimensional grounds. Thus, for $\mathcal{O}(1)$ couplings$^{32}$ the experimental lower limit of the order of $10^{33–34}$ years on the proton lifetime quoted above translate to a generic lower limit for the masses of the mediators underpinning the $d = 6$ BLNV effective operators at the level of about $10^{15}$ GeV!

2.4 High-energy convergence of the SM gauge couplings

With such an enormous scale, residing some 12–13 orders of magnitude above what is currently accessible at the Earth-based accelerator experiments, there seems to be essentially no way to probe the relevant BLNV dynamics directly (at least for those schemes whose dominant effects exhibit themselves at the level of $d = 6$ operators$^{33}$).

However, this does not mean that there is no way to say anything qualified about the likelihood of observability of the BLNV processes like $p$-decay or about the shape of the underlying interactions. Remarkably enough, the humongous scale of $10^{15}$ GeV identified from the lower limits on $p$ decay in the preceding section is not that far from the stipulated $10^{13}$ GeV ballpark where the seesaw dynamics “prefers” to reside, cf. (2.28), and it is certainly compatible with the Davidson-Ibarra limit (2.31).

$^{32}$Obviously, the lower limit on $M$ quoted here can be reduced considerably if small couplings are involved.

$^{33}$It is perhaps worth noting that in the case that the BLNV effects were due to $d > 6$ operators the lower limits on the associated effective scale(s) get reduced considerably. However, one should then provide a good argument for the absence/irrelevance of the $d = 6$ effects.
2.4.1 Renormalized gauge couplings and their evolution in the SM

Remarkably enough, there is a hint on such a large scale even in the very structure of the old good Standard Model itself. It is related to the evolution of its running couplings which, indeed, exhibit a nice convergence feature at about $10^{15-16}$ GeV!

The simplest way to see this is to write down and solve the renormalization group equations for the two non-abelian$^{33}$ gauge couplings $g_3 \equiv g_s$ and $g_2 \equiv g$

$$\frac{d}{d\mu} g_i = \beta_i, \quad i = 2, 3,$$  \hspace{1cm}(2.49)

with $\beta_i$ denoting the relevant $\beta$-functions, together with the corresponding electroweak-scale initial conditions (imposed at $\mu = M_Z$)

$$g_3(M_Z) = \sqrt{4\pi \alpha_s(M_Z)}, \quad g_2(M_Z) = \frac{e(M_Z)}{\sin \theta_W(M_Z)}.$$ \hspace{1cm}(2.50)

In the formulae above $\alpha_s$ encodes the QCD coupling strength, $e$ is the electric charge and $\theta_W$ is the weak mixing angle. The $\beta$-functions are, at the lowest non-trivial order, given by the infamous formula of Gross, Wilczek and Politzer $^{70,71}$

$$\beta_G = \frac{1}{16\pi^2} \left[ -\frac{11}{3} C_G^2 + \frac{2}{3} \sum_{fW} T_2^G(R_{fW}) + \frac{1}{3} \sum_{sc} T_2^G(R_{sc}) \right] g^3 + \ldots = \frac{1}{16\pi^2} b_G g^3 + \ldots$$ \hspace{1cm}(2.51)

where $C_G^2$ stands for the quadratic Casimir of the group factor $G$ and $T_2^G(R)$ denotes the index of the $G$-representations $R_{fW}$ and $R_{sc}$ hosting the Weyl fermions and complex scalars, respectively.

The system (2.49) receives a particularly simple form in the logarithmic coordinates $t = \frac{1}{2\pi} \log \mu/M_Z$. Defining $\alpha_i^{-1} \equiv 4\pi/g_i^2$ one arrives at

$$\frac{d}{dt} \alpha_i^{-1} = -b_i, \quad i = 2, 3,$$ \hspace{1cm}(2.52)

with a solution

$$\alpha_i^{-1}(t) = \alpha_i^{-1}(t_0) - (t - t_0)b_i$$ \hspace{1cm}(2.53)

which, for $b_3 = -7$ and $b_2 = -\frac{19}{6}$ calculated theoretically and $\alpha_3^{-1}(M_Z) \sim 8.6$ and $\alpha_2^{-1}(M_Z) \sim 29.9$ from the experiment, yields a picture like in Fig. 2.9. This, however, is highly interesting for at least two reasons:

$^{33}$For the time being we shall ignore the abelian (hypercharge) SM coupling $g'$, the reason is that unlike for the non-abelian generators there is no natural normalization scheme for the abelian one and, thus, from the pure SM perspective, $g'$ can be arbitrarily rescaled along with the hypercharge generator $Y$ as long as their product is preserved.
Figure 2.9: Evolution of the (inverse squared) running non-abelian gauge couplings in the Standard Model assuming no new dynamics kicking in throughout the evolution. The strengths of the two interactions converge to the same value at \( t \sim 5.5 \) which corresponds to roughly \( \mu \sim 10^{16} \text{ GeV} \).

- The “effective strengths” of the gauge interactions governed by the \( SU(2)_L \) and \( SU(3)_c \) group factors tend to converge at high energies and turn out to be the same at a certain very large scale which is practically identical with the lower limit on the mass of the \( d = 6 \) BLNV mediators discussed above.

- Note that this feature should not be viewed as a trivial consequence of different slopes of two straight lines stretching in a plane; actually, in order to get an intersection in the desired semi-plane (for \( t > 0 \)), let alone in the very interesting ballpark of \( \mu \equiv M_G \sim 10^{16} \text{ GeV} \), the initial conditions (i.e., the experimental data) and the values of the relevant \( \beta \)-functions (encoded in the structure of the SM) must conspire to a high degree!

Nevertheless, what is the real physics content of Fig 2.9? Indeed, extending the two curves beyond \( \mu \sim M_G \) makes them diverge again; from this perspective their intersection at about \( 10^{16} \text{ GeV} \) would have to be interpreted as a mere coincidence. However, what if their slopes change at the point of intersection due to a change in the field content of the theory such that the two curves continue as (optically) a single one for \( \mu > M_G \)? Would this correspond to a next step towards the eternal dream of a fully unified description of particles and their interactions? And, if affirmative, what does it have to do with the baryon and lepton number violating operators and their phenomenology discussed at length in the preceding sections?
2.5 Grand unification of the SM interactions

In order for the unification picture sketched in the previous paragraph to make any sense one should, eventually, realise these ideas in a fully dynamical scheme. This, however, requires several ingredients:

1. A certain specific set of new fields would have to be added into the game in such a way that the two running coefficients \( b_2 \) and \( b_3 \), as different as they are in the SM, become identical at and above \( M_G \).

2. Such a change in the slopes has to occur at a very specific scale – at around \( M_G \sim 10^{16} \) GeV.

3. If vector fields were to be employed for that sake one should eventually aim at a complete gauge framework with a gauge group \( G \) that would contain the entire SM one as a substructure, i.e., \( G \supset SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \); in such a scheme the gluons together with the SM A and B fields would span just a part of the adjoint representation of \( G \) along with another set of vector fields that, indeed, may play the role of the extra degrees of freedom of point 1.

From the low-energy perspective, it is like a miracle that the simplest complete picture of this kind can be obtained with just a pair of extra fields with their quantum numbers corresponding to the \( V_1 \) and \( S_1 \) mediators of the \( d = 6 \) BLNV effective operators in the SM!

2.5.1 Evolution of gauge couplings in the BLNV extensions of the SM

With \( V_1 \) and \( S_1 \) at play at and above \( \mu \sim M_G \) there are extra contributions to the one-loop \( b_i \) coefficients in Eqs. (2.52) driving the running gauge couplings from their intersection point onwards:

\[
\Delta \begin{pmatrix} b_2 \\ b_3 \end{pmatrix}_{V_1} = \begin{pmatrix} -11 \\ -\frac{22}{3} \end{pmatrix}, \quad \Delta \begin{pmatrix} b_2 \\ b_3 \end{pmatrix}_{S_1} = \begin{pmatrix} 0 \\ \frac{1}{6} \end{pmatrix}.
\]

Remarkably, this is all that is needed for a “homogenisation” of the SM \( b_i \)’s which, at and above \( \mu = M_G \), change to a common value of \(-85/6\). The gauge coupling evolution in this situation is depicted in Fig. 2.10.

\footnote{This simple phrasing is essentially consistent at the one-loop level of the RG analysis; however, one has to be way more careful if higher order corrections are to be taken into account, cf. \[72,73\].}
Figure 2.10: Evolution of the (inverse squared) running non-abelian gauge couplings in the Standard Model (lower part) into which a pair of fields V_1 and S_1 from Table 2.2 is added at the scale of the $\alpha_{2,3}^{-1}$ confuence.

2.5.2 The grand unification hypothesis and its implications

Hence, a simple conspiracy of the Standard Model dynamics with just a couple of extra fields from the Table 2.2 of possible tree-level mediators of the $d=6$ BLNV effective operators suggests a potentially renormalizable framework including a unified description of the $SU(3)_c$ and $SU(2)_L$ gauge interactions together with a new force which leads to potentially testable BLNV signals, very clear from any SM background! Hence, one encounters a canonical incarnation of a physical theory in the best Popperian sense – the “old” SM is fully encompassed, some of its shortcomings can be addressed ($L$ is not a sacred perturbative-level symmetry anymore) and there is a new dynamics predicted at a very specific scale with a load of spectacular new phenomena not far beyond the current experimental limits.

In some sense, the profound beauty and simplicity of such a picture may be even disturbing. Indeed, the proximity of $M_G \sim 10^{16}$ GeV and the $\sim 10^{15}$ GeV lower limit on the suppression scale associated to the $d = 6$ effective operators probed in the current proton decay searches (cf. Sect. 2.3.2) raises questions about the fragility of the whole scheme whose defining feature is the presence of the “big desert” between the electroweak scale and $M_G$. Is it really the case that no new dynamics is to be expected anywhere between $10^2$ and $10^{16}$ GeV?

---

35

36It is difficult to overemphasise the rarity of this feature in the swampland of the SM extensions (even among the most popular ones) where the scale of the new dynamics and, thus, the size of the associated new physics effects, is typically an unknown external parameter.
To this end, it is perhaps interesting to comment briefly on the possible changes to this scheme in presence of the additional degrees of freedom associated with the renormalizable realisations of the \( d = 5 \) Weinberg operator in the three types of the \textit{“tree-level”} seesaw mechanism discussed in Sect. 2.2.

1. As for the RH neutrinos of type-I seesaw these are complete gauge singlets and, as such, they do not affect the calculated \( b \)-coefficients above in any way, at least not at the one-loop level. From this perspective, the gauge unification pattern in Fig. 2.10 remains practically intact in this case.

2. With an extra \( \Delta_L \equiv (1, 3, +1) \) scalar at some \( 10^{13} \text{ GeV} \) only small changes can be inflicted on the gauge unification pattern in type-II seesaw. However, if the triplet is to be pushed down, perhaps into the vicinity of the electroweak scale, one has to be more careful as, in such a case, the slope of the \( \alpha_2^{-1} \) curve gets shallower for a significant portion of the running and the intersection point with \( \alpha_3^{-1} \) is pushed well below \( 10^{16} \text{ GeV} \).

3. The type-III seesaw setting with its fermionic \( SU(2)_L \) triplets at a large scale is somewhere in between the two scenarios above.

**The running \( U(1)_Y \) gauge coupling and the GUT-compatible hypercharge**

Another very important aspect of the story is the fate of the abelian gauge coupling \( g' \) associated to the SM hypercharge. If the unification of the gauge interactions of the SM is to be complete (or \textit{“grand”}) then also \( g' \) should conform the \( g \)-confluence constraint at \( M_G \) as do the non-abelian couplings. On one hand, given the aforementioned arbitrariness of the SM hypercharge operator overall normalization – and thus also that of \( g' \), cf. Sect. 2.4.1 – this may seem like a trivial requirement (and, in fact, it is so as long as we look at the unification from the low-energy perspective only). However, there is still something very interesting happening to the SM \( g' \) when \( V_1 \) and \( S_1 \) are integrated in at \( M_G \). Indeed, at the level of the individual classes of contributions to the \( b \)-coefficients corresponding to the three different types of terms in Eq. (2.51)

\[
\begin{pmatrix}
b' \\
b_2 \\
b_3
\end{pmatrix}_{\text{SM+} V_1 + S_1} = -\frac{11}{3} \left( \begin{array}{c}
\frac{25}{3} \\
5 \\
5
\end{array} \right)_{\text{vectors}} + \frac{2}{3} \left( \begin{array}{c}
\frac{10}{3} \\
2 \\
2
\end{array} \right)_{\text{matter}} + \frac{1}{3} \left( \begin{array}{c}
\frac{5}{6} \\
\frac{1}{2} \\
\frac{1}{2}
\end{array} \right)_{\text{scalars}}
\]

(2.55)

one can see that there is a \textit{universal scaling factor} of \( \frac{5}{3} \) that, if somehow stripped from the first row above (corresponding to a redefinition of \( Y \) and \( g' \)) would equalize not

\[\text{Note that the contributions to } b' \text{ in formula (2.55) have been calculated in the usual SM hypercharge normalization fixed by Eq. (2.3).}\]
only the resulting b’s but even the individual contributions associated to the vectors, matter fields and scalars, respectively! Thus, there is a very strong indication of what the “canonical” hypercharge normalization should be in scenarios in which one may eventually wish to talk about a unified description of not only the interaction strengths but also the effective field content behind; see Sect. 2.6.2.

The final comment concerns the fate of the other SM couplings, especially those in the Yukawa sector. Another likely consequence of the effect of a “unified” description of matter and scalar fields stipulated by the observation (2.55) is a more concise account for the flavour structure within the high-energy theory which, as a by-product, may provide correlations between different effective SM Yukawa matrices and, thus, shed at least some light on the deep issue of the SM quark and lepton masses and mixing.

2.5.3 Calculability of proton lifetime in unified models

Perhaps the most important implication of the simple unification picture sketched above is the fact that it tells us a lot about the size of the inherent new physics effects (in particular the BLNV ones) which, from the SM effective theory perspective of Sect. 2.3.1 are almost entirely out of control. Indeed, with the mediator scale set by (and calculable from) the requirement of the gauge coupling unification\(^{38}\) and the strong correlations among the relevant couplings governing the hard BLNV amplitudes of Fig. 2.47 one may even attempt to provide a numerical estimate of the proton lifetime. To this end, let us reiterate that the Nature has apparently been very generous to us (again!) – with the \(10^{16}\) GeV ballpark value of \(M_G\) and for \(O(1)\) couplings the rough proton lifetime estimates fall just to the (logarithmic) vicinity of the current experimental limits discussed in Sect. 2.3.1!

On the practical side, however, this may be a very formidable task of a high degree of complexity. Barring the notorious difficulties associated with the “translation” of the perturbative quark-level amplitudes of Fig. (2.47) into the hadronic ones\(^{39}\) there is namely the exponential sensitivity of the position of the unification scale \(M_G\) (entering in 4th power into the decay width) on most of the ingredients of the gauge running analysis including, e.g., the \(M_Z\)-scale initial conditions, the “threshold effects” associ-

---

\(^{38}\) As we said, the unification should take place not only within the gauge couplings but also in other structures such as Yukawas etc. However, the rigidity of the gauge sector of a generic gauge theory as compared to the model-dependence of its other sectors, together with the technical simplicity of the RG evolution of the same (at least at the 1-loop level) makes it natural to determine \(M_G\) primarily from there.

\(^{39}\) Note that there was a significant progress in the last decade in the lattice calculations of the B-violating hadronic matrix elements\(^{74}\) which made it possible to inhibit the associated theoretical uncertainties to such a degree that they are no longer a real concern.
ated to the possible splitting of the masses of the high-scale multiplets\textsuperscript{75} (relevant namely for higher-order calculations) and the need to account for the flavour structure of the $B$- and $L$-violating matter currents coupled to the heavy mediators. Last, but not least, there is the issue of the proximity of $M_G$ and the Planck scale $M_{Pl}$\textsuperscript{76,78} which, among other things, may inflict significant effects in the matching of the unified theory to the SM which, very often, are out of any control; see also Sect. 3.4.4. Thus, there are very few (if any) good-quality proton lifetime estimates in the existing literature.

### 2.5.4 Topological defects

The need to enhance the SM gauge symmetry to a higher-symmetry structure $G$ in presence of new vector fields as stipulated in Sect. 2.5 calls, at least in the most conservative approach, for a “repetition” of the classical Higgs trick (associated, traditionally, to the phenomenon of spontaneous electroweak symmetry breaking in the SM) in the unified scenario where $G$ must be eventually broken to $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$. Besides the scale at which this should occur ($M_G$ vs. $M_Z$) the main difference between the unified and the electroweak symmetry breaking is the topological structure of the corresponding coset spaces which, in the SM case, yields no stable topological defects. This, however, does not necessarily happen for the breaking of $G$, especially if the associated gauge group happens to be simple (i.e., when the unification is “grand”). In such a case the second homotopy class $\pi_2$ of the coset $G/SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ may be non-trivial and, hence, a topologically stable particle-like monopole solutions may exist\textsuperscript{79–81} in the spectrum of such models. This, in fact, is even inevitable if $G$ is also simply connected; then $\pi_2[G/SU(3)_c \otimes SU(2)_L \otimes U(1)_Y] = \pi_1[SU(3)_c \otimes SU(2)_L \otimes U(1)_Y] = \mathbb{Z}$.

Besides the topologically ensured stability (and a relatively easy production\textsuperscript{82,83} at the “GUT-epoch” of the early Universe some $10^{-37}$ s after the Big Bang) such objects would have very interesting properties. These, among other things, include:

- **Very high ionisation energy loss in matter**: Indeed, the Dirac quantization condition\textsuperscript{84,85} for the magnetic charge $Q_M$ in the form

$$Q_E Q_M = 2\pi n$$

yields an effective “monopole fine structure constant”

$$\alpha_M \equiv \frac{Q_M^2}{4\pi} = \frac{n^2}{4\alpha}$$

which means that the monopole electromagnetic interactions with matter are characterised by a dimensionless coupling that, even for $n = 1$, is $1/4\alpha^2 \sim 4700$
times larger than the fine structure constant relevant for the QED interactions of the SM leptons or hadrons.

- Catalysis of a rapid non-perturbative baryon number violation (so called Callan-Rubakov effect [19, 20]): In the centre of the t’Hooft-Polyakov [79, 80] monopole solution the VEV of the Higgs field responsible for the breaking of $G$ is effectively zero and, hence, the effects of the coset gauge fields (i.e., those carrying lepto-quark charges like, e.g., $V_1$ of Table 2.2) are not, at least locally, inhibited by their large mass parameters. This makes the cross-section of their BLNV interactions proportional to the square of the geometrical size of this region which, for a monopole with mass $\sim M_G$, scales like $M_G^{-2}$; this is to be compared with the $M_G^{-4}$ scaling of the perturbative effects discussed in Sect. 2.3.2. If a significant number of such monopoles is captured in the interior of compact astronomical objects such as white dwarfs a non-negligible increase in their luminosity can be in principle measurable. To this end, it is quite interesting that the non-observation of such effects provide orders-of-magnitude better constraints [86] on the monopole flux than direct searches based on the ionisation or magnetic effects [87].

Having already mentioned the Dirac condition (2.56) it is perhaps worth making one more conceptual comment here. Indeed, the classical Dirac’s argument [84] is that if a magnetic monopole exists then the flux tube which, in the Maxwelian electrodynamics, must extend from it along a semiaxis pointing towards an antimonopole somewhere far away would be unobservable if and only if all electric charges in the entire Universe were quantized. This, in fact, is perfectly compatible with the behaviour of the irreducible representations of simple compact Lie groups as typical incarnations of the unified symmetry structure $G$ whose generators have, necessarily, discrete spectra.

2.6 Towards a potentially realistic grand unified theory

The general concepts discussed above call for specific examples. These can be roughly classified by:

1. The Lie group/algebra of the unified symmetry $G$ which, by definition, should contain that of the Standard Model (i.e., $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \equiv H$) as a substructure. At the same time, it would be nice if $G$ was not “far larger than $H$”, because otherwise the symmetry breaking mechanism may get rather complicated – note that, for the lowest-rank $G$’s it can make a difference whether we demand the $H \subset G$ embedding at the level of groups or algebras - for instance, at the
group level, $SU(3) \otimes SU(2) \otimes U(1)$ is formally not a subgroup of $SU(5)$ but it is so as long as their algebras are concerned. \footnote{As a matter of fact, the true gauge symmetry of the SM is not $SU(3) \otimes SU(2) \otimes U(1)$ but $SU(3) \otimes SU(2) \otimes U(1)/Z_6$ and only the latter does admit an embedding into $SU(5)$ even at the group level \footnote{As much as, for instance, the $\lambda \phi^4$ theory does within the realm of interacting QFT’s.}}

2. The structure of the matter sector. There is often no need to add anything extra on top of the SM matter fermions (like, e.g., in $SU(5)$, cf. Sect. \ref{sec:gg-model}) or the extensions may be very mild and well motivated (as in the $SO(10)$ models of Chapter \ref{chap:so10}).

3. The complexity/reality of the irreducible representations of $G$ – the chiral structure of the SM calls for the existence of complex irreps of $G$ - the reason is the need to maintain the sharp distinction between the LH and RH fields entertained in the SM, at least at low energies.

4. The structure of the scalar sector - this is by far the least constrained part of the spectrum and, hence, a defining feature of any specific unified model. Perturbativity and non-tachyonicity constraints are among the most stringent requirements imposed on its choice.

Besides these, minimality will often be our main guiding principle (because there is hardly anything better). Interestingly, as trivial as it sounds there is no universal definition of what it means - should it be the number of fields or the number of free parameters of a model? In what follows we shall entertain the second option as, indeed, the main concern of natural philosophy should be predictivity.

2.6.1 The minimal $SU(5)$ Georgi-Glashow model

The first work on the grand unification ever published \footnote{As much as, for instance, the $\lambda \phi^4$ theory does within the realm of interacting QFT’s.} was written in 1974 by H. Georgi and S. Glashow. Though it was mainly mathematical in scope (focusing on the identification of rank=4 groups that may play the role a grand-unified gauge symmetry) it defines a paradigmatic minimal framework\footnote{As a matter of fact, the true gauge symmetry of the SM is not $SU(3) \otimes SU(2) \otimes U(1)$ but $SU(3) \otimes SU(2) \otimes U(1)/Z_6$ and only the latter does admit an embedding into $SU(5)$ even at the group level \footnote{As much as, for instance, the $\lambda \phi^4$ theory does within the realm of interacting QFT’s.}} based on an $SU(5)$ gauge group. This, back in the mid of 1970’s, looked like a very interesting candidate for superseding the just-born electroweak theory and QCD.

The basic structure of the Georgi-Glashow (G-G) model is extremely simple indeed. First, there is the adjoint irrep. $24$ hosting the gauge fields which, under the SM, decomposes into

$$24_V = (8, 1, 0) \oplus (1, 3, 0) \oplus (1, 1, 0) \oplus (3, 2, -\frac{5}{6}) \oplus (\bar{3}, 2, +\frac{5}{6}). \quad (2.58)$$
Needless to say, the first three factors above carry the gauge quantum numbers of the SM gluons and the $A$ and $B$ fields, respectively. Notice also that, indeed, the last two terms in (2.58) correspond exactly to the $V_1$ vector ($+h.c.$) identified in Sect. 2.3.2 that was crucial later in Sect. 2.5.1 to attain the “unified” shape of the trans-$M_G$ beta functions. In this sense, baryon and lepton number violation (BLNV) is an intrinsic feature of the G-G model.

Second, the five independent irreps hosting each family of matter fermions in the SM are embedded into just a pair of $SU(5)$ representations - a 5-dimensional vector and a 10-dimensional 2-index antisymmetric tensor:

\[
\bar{5}_F = (\bar{3}, 1, +\frac{1}{3}) \oplus (1, \bar{2}, -\frac{1}{3}) = d_R^c \oplus L_L ,
\]

\[
10_F = (\bar{3}, 1, -\frac{2}{3}) \oplus (3, 2, +\frac{1}{3}) \oplus (1, 1, +1) = u_R^c \oplus Q_L \oplus e_R^c .
\]

Two comments are perhaps worth making here:

- Note that all the right-handed SM matter spinors have been conveniently written in terms of their charge-conjugated counterparts which are left-handed. First, this is a must in order to put different chirality fields into the same multiplet of a higher symmetry which is supposed to commute with the Lorentz group; second, the same transformation flips the $-\frac{1}{3}$ hypercharge of $d_R$ into $+\frac{1}{3}$ of $d_L^c$ which, unlike for the former, matches the $-\frac{1}{2}$ hypercharge of $L_L$ and, thus, also the zero trace condition imposed on the $SU(5)$ generators in any representation.

- The fact that $\bar{5}$ is used above for accommodating $L_L$ and $d_L^c$ rather than the “optically simpler” 5 is a pure convention which is set by the shape of the 24 defining generators of the natural 5-dimensional representation. This is usually chosen in such a way that the Gell-Mann matrices defining their upper-left $3 \times 3$ sub-blocks enter there without extra complex conjugation. Note that in the opposite case one would have to work with $\bar{10}_F$ instead of $10_F$ so a bar appears in either case.

Finally, there are 2 irreps including the two Higgs fields necessary for breaking the rank=4 $SU(5)$ gauge symmetry down to that of the $SU(3)_c \otimes U(1)_Q$ of QCD$\otimes$QED via an intermediate $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ of the SM, namely:

\[
24_S = (8, 1, 0) \oplus (1, 3, 0) \oplus (1, 1, 0) \oplus (3, 2, -\frac{5}{6}) \oplus (\bar{3}, 2, +\frac{5}{6})
\]

and

\[
5_S = (1, 2, +\frac{1}{2}) \oplus (3, 1, -\frac{1}{2}) .
\]
As for the former (adjoint scalar), the singlet in its SM decomposition, i.e., the 3rd term in (2.61), justifies its choice as an agent to perform the first symmetry breaking step. On the same footing, the 5S provides the SM Higgs doublet, along with the S1 scalar of Sects. 2.3.2 and 2.5.1. In this sense, the crucial extra fields V_1 and S_1 are supplied in the most natural way in the G-G model!

Let us also comment in brief on the possible role of the extra degrees of freedom that were not considered in Sect. 2.5.1 but which are there in the minimal SU(5) theory, namely, the contents of 24_S. First, not all components of the decomposition (2.61) are propagating: indeed, the (3, 2, −5/6) ⊕ (3, 2, +5/6) scalars play the role of the Goldstone modes to be “eaten” in the unitary gauge by V_1 in order to give them O(M_G) masses so only the first three will be in the physical spectrum of the theory. Second, the effect of all these fields in the (Feynman-gauge, one loop) argument of Sect. 2.5.1 is practically irrelevant because these degrees of freedom come as components of an entire irrep of the unified SU(5) group. As such, they must contribute homogeneously to all three beta-functions for the effective SM couplings (if the SU(3) ⊗ SU(2) ⊗ U(1) language is preferred) and, thus, leave the scale of their intersection intact. This expectation is easy to verify by direct calculation.

The last remark concerns the Higgs mechanism triggered by a large VEV of the SM singlet in (2.61). The hypothesis that it generates masses only for the desired V_1 = (3, 2, −5/6) ⊕ (3, 2, +5/6) vector can be verified readily by spanning the complete set of the SU(5) vector fields onto the natural basis of its adjoint representation, namely, the very generator matrices T^a:

\[ |V^a⟩ = V^a T^a \text{ (no summation over } a), a = 1 \ldots 24. \] (2.63)

Recall that the action of the generators of the adjoint representation (T_{adj}^b) on such vectors can be written in terms of the commutator of a pair of fundamental generators

\[ T_{adj}^b |V^a⟩ \propto V^a \left[ T^a, T^b \right]. \] (2.64)

In this formalism, the singlet component which should bear the SU(5)-breaking but SU(3)_c ⊗ SU(2)_L ⊗ U(1)_Y-conserving VEV must thus be spanned over the extra Cartan operator (T^{24}) that commutes with the three SM ones. In the conventional basis in which the SU(3)_c generators occupy the upper-left 3 × 3 corner of the corresponding 5 × 5 Hermitian traceless matrix space and the SU(2)_L ones are accommodated in the

\footnote{Notice that the first step of the symmetry breaking, i.e., SU(5) → SU(3) ⊗ SU(2) ⊗ U(1), preserves the rank of the initial algebra. Thus, the possible choices of the irreps which can be used for that is very limited and, to this end, the adjoint is a natural candidate.}
lower-right $2 \times 2$ sector $T^{24}$ reads
\[
T^{24} \propto \left( \begin{array}{cc}
\frac{1}{3} & 1_{3 \times 3} \\
-\frac{1}{2} & 1_{2 \times 2}
\end{array} \right),
\]
and so does the “ket” corresponding to the SM singlet: $|1, 1, 0\rangle = V_{24} T^{24}$. Hence, the mass matrix for the vector fields
\[
M_{ab}^2 = g^2 \langle 0| T^a_{\text{adj}} T^b_{\text{adj}} |0\rangle = \text{Tr} \ g^2 V_G^2 [T^{24}, T^a] [T^b, T^{24}]
\]
can receive non-zero contributions only for $a$ and $b$ which correspond to the fields spanned over the fundamental generators carrying both $SU(3)_c$ and $SU(2)_L$ indices, i.e., $(3, 2, -\frac{5}{6}) \oplus (\bar{3}, 2, +\frac{5}{6})$.

It can also be shown that the vacuum manifold defined by the scalar potential
\[
V = m^2 \text{Tr} 24^2_S + \lambda (\text{Tr} 24^2_S)^2 + \rho \text{Tr} 24^4_S
\]
hosts only three extrema out of which only the original $SU(5)$ and the final $SU(3) \otimes SU(2) \otimes U(1)$-symmetric one are (local) minima.

### 2.6.2 Phenomenology of the minimal $SU(5)$ GUT

Besides the expected new physics in the form of baryon and lepton number violation the unified theories often provide insights into the structure of their low-energy effective descendants in the form of correlations among independent parameters of the SM. To this end, there are two key observations to be made in the G-G context:

1. **The value of the SM weak mixing angle is predicted.** This has to do with the structure of the hypercharge operators in the $SU(5)$ irreps hosting (not only) the SM matter. In order to maintain simplicity all generators of simple Lie groups are conventionally orthonormalized by conditions like
\[
\text{Tr} T^i_R T^j_R = T_2(R) \delta^{ij}
\]
where $T_2(R)$ is a universal real number (called index) specific to the irrep $R$. The normalization is usually fixed so that $T_2$ of the vector irreps of $SU(N)$’s is $\frac{1}{2}$. Looking at $5_F$ in $(2.59)$ it is easy to verify that this choice indeed coincides with the usual normalization of the SM $SU(3)_c$ (anti)triplet generators (halves of the Gell-Mann matrices, $\lambda^a/2$) as well as with those of the $SU(2)_L$ doublets (halves

\[43\text{We do not consider the 5}_S \text{ component here as its VEV necessarily reduces the rank of the gauge group and, as such, it can only be responsible for the subsequent electroweak symmetry breaking with } v \ll V_G.\]
of the Pauli matrices, \( \sigma^k/2 \). On the other hand, the normalization of the SM hypercharge \( Y_{SM} \) as of Eq. (2.5) must be adjusted into the GUT-natural form \( Y_G = nY_{SM} \) so that

\[
\sum_{5_F} Y_G^2 = n^2 \sum_{5_F} Y_{SM}^2 = n^2 \left( 3 \times \frac{1}{9} + 2 \times \frac{1}{4} \right) = \frac{5}{6} n^2 = \frac{1}{2}, \tag{2.69}
\]

which yields

\[
n = \pm \sqrt{\frac{3}{5}}. \tag{2.70}
\]

Choosing, conventionally, the positive solution the GUT-adjusted hypercharge reads

\[
Y_G = \sqrt{\frac{3}{5}} Y_{SM} \tag{2.71}
\]

and the proper decomposition of \( \mathbf{5}_F \) is thus

\[
\mathbf{5}_F = (\mathbf{3}, 1, +\sqrt{\frac{1}{15}}) \oplus (1, \mathbf{2}, -\sqrt{\frac{3}{20}}), \tag{2.72}
\]

and similarly for the other irreps. The change \( Y_{SM} \rightarrow Y_G \) in the structure of the covariant derivative of the theory inflicts a change in the associated gauge coupling too: On the SU(5) side one has \( D_\mu^{SU(5)} = \partial_\mu - ig_5 Y_G A^\mu_Y + \ldots \) which is to be matched to the SM structure \( D_\mu^{SM} = \partial_\mu - ig' Y_{SM} B_\mu + \ldots \); this can be done if and only if

\[
g_5 \sqrt{\frac{3}{5}} \equiv g' \tag{2.73}
\]

at the matching scale, i.e., at \( M_G \). Hence, there are two different languages in which one can describe the same hypercharge dynamics, either in the language of the SM with \( g' \) and \( Y_{SM} \) everywhere or with \( g_5 \) and \( Y_G \). It is only in the latter case though that the associated gauge coupling (i.e., \( g_5 \)) should be universal at \( M_G \), not with the former \( g' \). Thus, in the beta-functions’ analysis of Sect. 2.5.2 one should have worked with \( g_1 \equiv \sqrt{\frac{5}{3}} g' \) and the associated \( b_1 \) from scratch instead of \( b' \) calculated from \( Y_{SM} \)! This, indeed, supplies the extra multiplicative factor of \( \frac{3}{5} \) to the first row of Eq. (2.55) and brings it into a fully universal form

\[
\begin{pmatrix}
  b_1 \\
  b_2 \\
  b_3 
\end{pmatrix}_{SM+V_1+S_1} = -\frac{11}{3} \begin{pmatrix}
  5 \\
  5 \\
  5 
\end{pmatrix}_{\text{vectors}} + \frac{2}{3} \begin{pmatrix}
  2 \\
  2 \\
  2 
\end{pmatrix}_{\text{matter}} + \frac{1}{3} \begin{pmatrix}
  \frac{1}{3} \\
  \frac{1}{3} \\
  \frac{1}{3} 
\end{pmatrix}_{\text{scalars}}. \tag{2.74}
\]

Hence, the unified nature of the 3 gauge interactions in presence of \( V_1 \) and \( S_1 \) (barring the universal extra contribution from \( 24_S \) in the full G-G setting) becomes obvious (even tautologic) in this language!
Finally, the gauge unification condition $g_1 = g_2 = g_3$ at $M_G$, rewritten for $g$ and $g'$ of the SM as $g = \sqrt{\frac{2}{3}} g'$ at $M_G$, makes it possible to calculate the (tree-level) SM weak mixing angle at $M_G$:

$$\sin^2 \theta_W(M_G) \equiv \frac{g'^2}{g^2 + g'^2} = \frac{1}{\frac{5}{3} + 1} = \frac{3}{8} = 0.375.$$  \hspace{1cm} (2.75)

This, at one hand, looks way bigger than the measured value $\sin^2 \theta_W(M_Z) \sim 0.232$ but, at the same time, it can not be discarded right away because of the large effects of running from $M_G$ to $M_Z$; indeed, $b_2 = -\frac{19}{6}$ and $b' = +\frac{41}{6}$ below $M_G$ and, thus, the RHS of (2.75) will get reduced towards $M_Z$. In any case, the final decision requires a more detailed analysis.

2. Some of the SM Yukawa couplings are strongly correlated. This is another example of a feature that provides an interesting insight into the (conceptually) free parameters of the SM, namely, its flavour structure. It has to do with the shape of the (renormalizable) Yukawa Lagrangian of the minimal $SU(5)$ which, in the most economical case of just one “Yukawa-active” scalar representation, can be written like

$$\mathcal{L}_Y = Y_{ij}^5 5_F^{iT} C 10_F^{J} 5_S^* + Y_{ij}^{10} 10_F^{iT} C 10_F^{J} 5_S^*,$$  \hspace{1cm} (2.76)

where $i$ and $j$ are family indices, $T$ denotes transposition in the spinorial space and $C$ is the associated charge conjugation matrix. In the asymmetric phase, i.e., with non-zero VEV $v$ of the $SU(2)_L$ doublet in $5_S$ one can readily identify the mass matrices of the SM fermions (at $M_G$):

$$M_d^T = M_l = Y_5 \frac{v}{\sqrt{2}}, \quad M_u = M_u^T = Y_{10} \frac{v}{\sqrt{2}}.$$  \hspace{1cm} (2.77)

Note that there is a strong correlation among the down-quark and charged-lepton sectors in the minimal $SU(5)$ GUT and that the up-quark Yukawa coupling is symmetric. This, in turn, reduces the number of independent SM Yukawa couplings considerably.

\[\text{44} \text{ It is possible to work with only a single scalar } 5 \text{ because it (or its complex conjugate) can be found in both relevant tensor products}\]

\[\begin{align*}
5 \otimes 10 &= 5 \oplus 45 \\
10 \otimes 10 &= 5 \oplus 45 \oplus 50
\end{align*}\]
2.6.3 The failure of the minimal $SU(5)$ GUT

As beautiful and natural as it looks this picture actually does not withstand a more thorough scrutiny. First, the aforementioned flavour correlations lead to a direct prediction for the (running) masses of down-quarks and charged-lepton which should – generation by generation – coincide at $M_G$. While this works at least to some extent for $m_b$ and $m_\tau$ it fails miserably for the first two families.

As a matter of fact, this per se does not need to be fatal for the G-G model because the relevant Lagrangian (2.76) may be subject to large corrections associated to the $d > 4$ Planck-scale induced operators such as those discussed in [91]. Indeed, since $M_G/M_{Pl}$ ratio is around 1% one can expect significant corrections to the two smallest eigenvalues of both $M_d$ and $M_l$ which may be sufficient to correct the “unrealistic” first- and second-generation $d = 4$ predictions.

The G-G scenario, however, suffers from way more severe drawbacks than that:

1. It can not (in its minimal form) account for non-zero neutrino masses. The point is the absence of the necessary structure to incarnate the Dirac option or any of the three basic seesaw variants of the neutrino mass generation of Sect. 2.3.2. The obvious solution, i.e., the addition of RH neutrinos in the form of $SU(5)$ singlets is not very satisfactory as it brings in a new Yukawa matrix which does not entertain correlations with any other. Perhaps even more importantly, the natural scale of the associated gauge-singlet Majorana mass term, given its $SU(5)$-singlet nature, exceeds $M_G$ which is difficult to reconcile with the neutrino oscillations data.

2. The weak mixing angle (evolved from $M_G$ to $M_Z$) turns out wrong. Performing the simple exercise discussed below Eq. (2.75) one recovers $\sin^2 \theta_W(M_Z) \sim 0.195$ which is many (tens of) standard deviations off the measured value. In the early days of the subject [90] this has not been an issue due to the lack of good data; however, the situation has changed drastically with the advent of LEP and its precision weak sector measurements.

Hence, the minimal $SU(5)$ grand unified model by Georgi and Glashow is nowadays considered dead and, as such, it enjoys the status of a benchmark scenario rather than a full-fledged physical theory.

2.6.4 The minimal supersymmetric $SU(5)$ GUT

There are, of course, many proposals aiming at overcoming the limitations of the original G-G model such as, for instance, settings with extra matter fermions [92] or extra scalars [93]. These, if pushed deep enough into the “GUT desert” may change the gauge running pattern considerably.
However, by far the most popular of these is the idea of TeV-scale supersymmetry (SUSY) which, with the onset of the precision electroweak era in the early 1980’s, even became a BSM model-building paradigm. In this framework, the main drawbacks of the Standard Model discussed in Chapter 1 and of the minimal \(SU(5)\) GUT of Sect. 2.6.3 are addressed as follows:

- The wrong value of \(\sin^2 \theta_W\) in the Georgi-Glashow model: The “doubling of degrees of freedom” (together with the need to add another Higgs doublet to tame the issues with holomorphy of the superpotential as well as with potential gauge anomalies) in the minimal supersymmetric extension of the SM (MSSM) bends the RG running of the SM gauge couplings so that the experimental value of \(\sin^2 \theta_W(M_Z) \sim 0.232\) is “miraculously” attained.

- The absence of a suitable dark matter candidate in the SM. It is well known that there is no room in the Standard Model to account for the cold dark matter (CDM) particle candidate favoured by the current ΛCDM Standard Model of cosmology. The light neutrinos, the only stable enough neutral particles around, could account for the observed DM critical density fraction if and only if the sum of their masses was at the level of about 10 eV, at odds with the oscillation and \(\beta\)-decay data. Moreover, their “hot relic” nature leads to contradictions with the small-scale structure formation data if they were to represent the dominant DM component. To this end, the low-energy SUSY can come to rescue if \(R\)-parity, one of the empirical extra symmetry requirements often imposed on its superpotential, is exact – in such a case the neutralino (if it happens to be the lightest supersymmetric particle, LSP) provides an almost ideal particle-like CDM candidate.

- The apparent need for order-by-order fine-tunning among the bare and the high-scale (cut-off) contributions in the formula for the electroweak VEV. This issue is often quoted as one of the facets of the notorious “hierarchy problem” which was stirring the HEP community\(^{45}\) for at least the last 30 years. In this respect, the low-energy SUSY brings a partial relief to the issue by the radiative stabilisation of any mechanism that is eventually employed at the tree level.

In this respect, the minimal SUSY \(SU(5)\) GUT \(^{96, 97}\) represents the most economical grandunified completion of the MSSM. On the technical side, barring the standard elements of the SUSY model building like the need to embed all the standard \(SU(5)\)

\(^{45}\)Not the author though whose renegade attitude to this conundrum is well documented by his recent works \(^{94, 95}\).
quantum fields into superfields (chiral or vector), the presence of the soft-SUSY breaking sector etc., the basic structure of the minimal SUSY $SU(5)$ GUT closely resembles that of the Georgi-Glashow scheme of Sect. 2.6.1. The main difference is the presence of two different copies of 5-dimensional Higgs multiplets (as with the doublets in the MSSM) and, in particular, the simplicity of the $d = 3$ renormalizable superpotential which relates the higgsino couplings with matter-matter currents to the ordinary Yukawa terms. These, in turn, play an important role in the description of the BLNV phenomena like proton decay which, unlike in the non-SUSY theories, can proceed through triplet-higgsino-mediated effective operators already at the $d = 5$ level (in the effective MSSM language), see Fig. 2.11.

Figure 2.11: The basic structure of the hard-process Feynman graphs behind proton decay in TeV-scale SUSY models. The diagrams with gluino, bino and neutral higgsino dressings are typically suppressed by the first and second generation Yukawa couplings governing the underlying $d = 5$ effective operators. The flavour structure displayed corresponds to the dominant $p$-decay channel which, in SUSY, is usually $p^+ \to K^+ \nu$.

2.6.5 The issues of the minimal supersymmetric $SU(5)$ GUT

This, however, indicates a potential problem with the minimal supersymmetric $SU(5)$ GUT, namely, the overly fast proton decay. Indeed, comparing the structure of the amplitudes of Fig. 2.11 with those of the non-supersymmetric variant depicted in Fig. 2.47 (especially the leading vector-mediated graph therein) the two are related by a multiplicative factor of roughly

$$ R \equiv A_{\text{SUSY}}/A_{\text{non-SUSY}} \sim \frac{1}{16\pi^2} \frac{Y^2}{g^2} \frac{M_G}{m_{\text{SUSY}}}, $$

(2.78)

where the first piece on the RHS correspond to the SUSY loop suppression, the second to the Yukawa vs. gauge domination of the leading-order contributions in SUSY and non-SUSY scenarios and the third to the bosonic vs. fermionic nature of the relevant
high-scale mediator assisted on the SUSY side by the gaugino/higgsino propagator effect of the order of the soft-SUSY breaking scale $m_{\text{SUSY}}$.

Given the enormous hierarchy $M_G/m_{\text{SUSY}} \sim 10^{13}$ attained within TeV-scale SUSY scenarios it looks like there is no way whatsoever to get the proton lifetime anywhere near the desired (non-supersymmetric) limit of $10^{34}$ years corresponding to $R \sim 10^9$. This conclusion is, however, premature as the other two contributions in $(2.78)$ provide further suppression ( $10^{-2.5}$ for the loop factor and some $10^{-8}$ due to the presence of the first generation Yukawa couplings). Hence, the situation of the the minimal SUSY $SU(5)$ scenario is rather severe but there is, technically, still not a strict no-go; for more details the reader is kindly deferred to $[98]$ and references therein. Note also that the minimal SUSY $SU(5)$ fully shares the drawbacks of its non-SUSY version with accounting for the neutrino masses, see Sect. $2.6.3$.

Hence, none of the two canonical versions of the minimal grand unification seems fully realistic and there is a good reason to look for better realisations of the gauge unification paradigm outside the realm of the very restrictive rank = 4 models.
Chapter 3

Radiative effects in potentially realistic unified gauge models

3.1 Rank=5 extended gauge models

Relaxing the rank = 4 requirement imposed in the previous Chapter on the defining symmetry of the minimal gauge extensions of the Standard Model the number of model-building options grows rapidly.

Perhaps the first thing that comes to mind when an extra gauge generator is admitted into the game along with the 4 Cartans of the SM is the $U(1)_{B-L}$ symmetry discussed briefly in Sect. 2.1.2 which, in the presence of 3 copies of RH neutrinos (each equipped with a unit of lepton number) becomes a good candidate for a gauge charge. There is even more to that though:

- The spectrum of such a RH-neutrino-extended variant of the SM (3NSM) becomes very symmetric as the RH neutrinos nicely “fill” the apparent vacancy in the sector of the RH leptons.

- The hypercharges of the RH sector of the 3NSM exhibit an intriguing pattern in which the two members of the “natural pairs”, i.e., $u_R$ and $d_R$ or $N_R$ and $e_R$ differ by exactly the same amount ($\Delta Y_{u_R,d_R} = \Delta Y_{N_R,e_R} = +1$).

- This closely resembles the situation of the electric charges $Q$ in the LH sector whose differences within the $SU(2)_L$ multiplets obeys the same $\Delta Q = +1$ rule formalised in the concept of isospin and the Gell-Mann-Nishijima formula (2.5).

- In attempt to replay the same game with $Y$ instead of $Q$ the RH variant of the weak isospin is an obvious hypothesis to put forward (with the corresponding
gauge factor denoted by $SU(2)_R$ and the 3 associated generators by $T_R^i$. This, however, calls for an extra $U(1)$ charge $X$ such that
\[ Y = T_R^3 + X. \quad (3.1) \]

Remarkably enough, such an additional $X$ charge turns out to be nothing but $\frac{1}{2}(B-L)$; thus, one is left with an intriguingly symmetric relation
\[ Q = T_L^3 + T_R^3 + \frac{B-L}{2}, \quad (3.2) \]
which not only connects the apparent doublet structure of the RH sector to the profound concept of the $B-L$ symmetry but it is also very aesthetically appealing. Hence, it is more than natural to take this observation as a starting point at the quest for the rank=5 gauge extensions of the SM and consider the
\[ G_{LR} \equiv SU(3)_c \otimes SU(2)_L \otimes SU(2)_R \otimes U(1)_{B-L} \quad (3.3) \]
group as the first of its milestones.

### 3.1.1 Left-right symmetric models

Among the most appealing features of the gauge models based on the $SU(3)_c \otimes SU(2)_L \otimes SU(2)_R \otimes U(1)_{B-L}$ symmetry is the phenomenon of the high-scale parity restoration \[58\]. Indeed, in the unbroken phase the LH and RH fermions (including $N_R$) enjoy the same attention of their respective $SU(2)$ gauge factors and also their additional $U(1)$ charges proportional to $B-L$ are the same, see Table 3.1; the chiral structure of the SM is then revealed in the asymmetric phase only, i.e., after the symmetry breaking of
\[ SU(3)_c \otimes SU(2)_L \otimes SU(2)_R \otimes U(1)_{B-L} \rightarrow SU(3)_c \otimes SU(2)_L \otimes U(1)_Y. \quad (3.4) \]
This is usually triggered either by a scalar field which is an $SU(2)_R$ doublet with $B-L = \pm 1$ (thus following the familiar SM pattern) or by an $SU(2)_R$ triplet $(1, 1, 3, \pm 2)$. From the phenomenology point of view the latter option is particularly interesting:

1. **In the models with LR-symmetric scalar sector the type-II seesaw mechanism for neutrino mass generation is naturally in operation.** The point is that the $SU(2)_R \otimes U(1)_{B-L}$ breaking provided by the VEV of $\Delta_R = (1, 1, 3, +2)$ gets projected onto its LH companion $\Delta_L = (1, 3, 1, +2)$ of type-II seesaw (cf. Sect. 2.2.2) through the pair of the electroweak VEVs carried by the Higgs bi-doublet $H = \cdots$
Table 3.1: Minimal fermionic contents of the left-right symmetric extensions of the SM (single SM matter generation + RH neutrino displayed).

<table>
<thead>
<tr>
<th>Field/Chirality</th>
<th>3221 quantum numbers</th>
<th>3RHNSM contents</th>
<th>$Y = T_R^3 + (B - L)/2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_L$</td>
<td>$(3, 2, 1, +\frac{1}{2})$</td>
<td>$(u_L, d_L)$</td>
<td>$+\frac{1}{6}$</td>
</tr>
<tr>
<td>$Q_R$</td>
<td>$(3, 1, 2, +\frac{1}{3})$</td>
<td>$u_R, d_R$</td>
<td>$+\frac{2}{3}, -\frac{1}{3}$</td>
</tr>
<tr>
<td>$L_L$</td>
<td>$(1, 2, 1, -1)$</td>
<td>$(\nu_L, e_L)$</td>
<td>$-\frac{1}{2}$</td>
</tr>
<tr>
<td>$L_R$</td>
<td>$(1, 1, 2, -1)$</td>
<td>$N_R, e_R$</td>
<td>$0, -1$</td>
</tr>
</tbody>
</table>

(1, 2, 2, 0) by their $\Delta_L^\dagger H^2 \Delta_R$ interaction in the scalar potential. Schematically, the vacuum condition

$$\left\langle \frac{\partial V}{\partial \Delta_L^\dagger} \right\rangle = M_\Delta^2 \langle \Delta_L \rangle + \lambda \langle H \rangle^2 \langle \Delta_R \rangle + \ldots = 0$$

implies that $\Delta_L$ receives an “induced” VEV of the order of $\langle H \rangle^2 \langle \Delta_R \rangle / M_\Delta^2 \propto \langle H \rangle^2 / \langle \Delta_R \rangle$. Hence, the seesaw suppression of also the type-II light neutrino mass contribution is connected to the scale of the $SU(2)_R \otimes U(1)_{B-L}$ symmetry breaking as in the type I seesaw.

2. In the supersymmetric versions of the LR models the $R$-parity is automatically conserved in the triplet-breaking scenarios. The point is that the even $B - L$ charge of $\Delta_R$ makes the vacuum neutral with respect to the $R$-parity defined as

$$R = (-1)^{2S + 3(B - L)}$$

(with $S = 0$ for spinless fields) and, hence, $R$ remains a conserved quantum number even in the broken phase. Consequently, the neutralino LSP is a natural WIMP dark matter candidate.

Besides all this, the SM fermions entertain a high degree of correlations among their effective Yukawa couplings, which may, e.g., provide a rationale for the smallness of the CKM mixing and so on.

### 3.1.2 The Pati-Salam model

Remarkably enough, the $SU(3)_c$ and $U(1)_{B-L}$ factors of the LR gauge group of Section 3.1.1 can be further combined into a full-fledged $SU(4)$ symmetry with a Cartan

1 This, in the simplest situation of the single bi-doublet (1, 2, 2, 0) scenario, is even pathological, one typically needs extra multiplets.
set including the two diagonal Gell-Mann matrices (extended by zeros to a $4 \times 4$ structure) and the $4 \times 4$ traceless $B - L$ generator with $\{\pm \frac{1}{3}, \pm \frac{1}{3}, \pm \frac{1}{3}, -1\}$ on the diagonal. These, together with the associated rising and lowering operators of the $SU(4)$ act naturally on a pair of 4-dimensional vector representations whose upper three components host the three colour eigenstates of quarks with the corresponding lepton on the lowest position. Thus, the gauge symmetry is further enhanced into what is usually called the Pati-Salam gauge group

$$G_{PS} \equiv SU(4)_C \otimes SU(2)_L \otimes SU(2)_R$$

and a very appealing quark-lepton unification is achieved, with “lepton number as a fourth colour” as its trademark \[99\].

Besides encompassing virtually all the interesting features of the aforementioned LR models (corresponding to one of its symmetry breaking chains, see below) there are other aspects of the Pati-Salam scheme worth mentioning here:

1. **There are leptoquark-type of fields in the spectrum.** This is a consequence of the extended gauge symmetry whose $(\mathbf{15}, \mathbf{1}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{3}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{1}, \mathbf{3})$ adjoint representation (especially, its first factor) decomposes under the $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ subgroup as

$$\mathbf{(8, 1, 0)} \oplus (\mathbf{1, 1, 0}) \oplus (\mathbf{3, 1}, -\frac{2}{3}) \oplus (\mathbf{3, 1}, +\frac{2}{3}) \oplus \ldots .$$

Indeed, the last two components above (to be called $X_{\mu} + h.c.$) are perfectly suited to provide a link between the same-chirality quarks and leptons:

$$\mathcal{L}_{PS} \supset \frac{g_4}{\sqrt{2}} \left( \bar{L}_L \gamma_{\mu} Q^i_L + \bar{d}_R \gamma_{\mu} u^i_R + \bar{e}_R \gamma_{\mu} d^i_R \right) X_{\mu i} + h.c.$$

Clearly, the number of SM baryons and leptons is not conserved at the level of elementary vertices in the Pati-Salam types of models. However, this does not mean that there should be any $B$ and $L$ violating effects observed at colliders – indeed, the $B$ and $L$ “leaking” from one side of a Feynman diagram containing a vertex from (3.9) is repaid on the other side of the graph where $X$ gets transformed back to the relevant matter fermion pair. In other words, $B$ and $L$ are not really broken; they are just being carried around by $X$.

It may also be worth noting that the Pati-Salam scenario discussed here is actually not the most minimal framework where the $X$ vectors can emerge – the $SU(4)_C \otimes SU(2)_L \otimes U(1)_R$ gauge theory, very popular recently \[100\], also contains these fields.

---

\[2\]Strictly speaking, the diagonal of the $B - L$ should read $\sqrt{\frac{3}{8}}\{+\frac{1}{3}, +\frac{1}{3}, +\frac{1}{3}, -1\}$ otherwise it does not conform the same normalization as the other operators of the $SU(4)_C$ algebra.
2. The Pati-Salam theory has a monopole. To see this one should study the second homotopy class of the $G_{PS}/SM$ quotient group

$$\pi_2 \left( \frac{SU(4) \otimes SU(2) \otimes SU(2)}{SU(3) \otimes SU(2) \otimes U(1)} \right) = \pi_2 \left( SU(4) \otimes SU(2)/SU(3) \otimes U(1) \right), \quad (3.10)$$

which, thanks to the fact that

$$\pi_1 (SU(4) \otimes SU(2)) = \pi_1 (SU(4)) \oplus \pi_1 (SU(2)) = 0, \quad (3.11)$$

i.e., that $SU(4) \otimes SU(2)$ is simply connected, yields

$$\pi_2 \left( SU(4) \otimes SU(2)/SU(3) \otimes U(1) \right) = \pi_1 \left( SU(3) \otimes U(1) \right) = \mathbb{Z}. \quad (3.12)$$

The mass of this monopole should be somewhat above the scale of the Pati-Salam symmetry breaking. For more information including, e.g., cosmological constraints on its early-Universe production see [101] and references therein.

3. The Yukawa pattern naturally accommodates the Georgi-Jarlskog texture [102]. This concerns an interesting coincidence between the phenomenological low-energy relation $m_\mu/m_s \sim 3$ and one of the features of the electroweak symmetry-breaking pattern available in the Pati-Salam type of models. Suppose that the Higgs field of the SM contains a significant component from the $(15, 2, 2)$ Pati-Salam scalar. Since the VEV of this multiplet must be colour neutral, in the $SU(4)_C$ adjoint representation space it must be proportional to the last Cartan operator (i.e., on the $U(1)_{B-L}$ generator in this case; see also formula (2.65) of Sect. 2.6.1) which, in the traditional basis with $SU(3)_C$ spanning over the upper-left $3 \times 3$ sector of the first 8 $SU(4)_C$ generators, receives the form

$$T^{15} \propto \begin{pmatrix} \frac{1}{3} & \mathbb{1}_{3 \times 3} \\ -\mathbb{1} & \end{pmatrix}. \quad (3.13)$$

Hence, the fraction of the electroweak VEV carried by this multiplet will naturally generate 3 times larger contributions to the effective mass matrices of leptons than to the quark ones.

3.1.3 SO(10) grand unification

With all this at hand, there is just a small final step to be made in order to identify the simplest grandunified scenario featuring the LR symmetry [3.3] as its potential low-energy descendant. Indeed, given the algebraic isomorphisms

$$su(4) \approx so(6) \quad \text{and} \quad su(2) \oplus su(2) \approx so(4) \quad (3.14)$$
it is clear that the entire algebra of the Pati-Salam symmetry is contained in the algebra
of the simple and compact $SO(10)$ group. Thus, $SO(10)$ is often claimed to be the most
natural unified rank = 5 gauge extension of the SM and, as a matter of fact, it can also
be shown to be the only potentially realistic candidate for grand unification\footnote{For a very interesting alternative rank=5 unified theory of baryon and lepton number violation which, however, is not “grand”, the reader is deferred to Sect. 3.6.} at this
level. Hence, in what follows, we shall focus primarily on the class of $SO(10)$ unified
scenarios and, in particular, on their minimal SUSY and non-SUSY realisations.

One more comment concerning the anomalies in gauge extensions of the SM is
perhaps worth here. Unlike in theories based on the $SU(N)$ gauge structure (or direct
products of several such factors) in which the requirement of the absence of gauge
anomalies provides strong constraints on their matter content (cf. Sect 2.1.2), the
$SO(N)$ gauge structure is automatically anomaly free for all $N > 2$ and $N \neq 6$.
This follows from the generic impossibility to construct a fully antisymmetric tensor
from a product of three antisymmetric $SO(N)$ generators. The case of $N = 6$ is
singular because $SO(6)$ is locally isomorphic to $SU(4)$; on a more technical level, the
antisymmetry of the orthogonal groups’ generators makes it possible to construct a
non-vanishing anomaly structure like

$$\text{Tr} \left( \{ M^{ij}, M^{kl} \} M^{mn} \right) \propto \varepsilon^{ijklmn}, \quad (3.15)$$

with a completely antisymmetric tensor of the $SO(6)$ on the right hand side. This
observation makes it also clear why the Georgi-Glashow model with matter in $10 \oplus 5$
is anomaly free: adding a full (harmless) $SU(5)$ singlet to each generation of matter
the field content of a full 16-dimensional spinor of $SO(10)$ is attained.

### 3.2 Minimal $SO(10)$ GUTs

Perhaps the most appealing feature of the $SO(10)$ GUTs is the fact that an entire
generation of the SM matter fields can be accommodated in its single spinorial irrep,

$$16 = (3, 2, +1/6) \oplus (1, 2, -1/2) \oplus (\overline{3}, 1, -2/3) \oplus (\overline{3}, 1, +1/3) \oplus (1, 1, 0) \oplus (1, 1, +1) \quad (3.16)$$

and, at the same time, one RH neutrino per generation is inevitable. This makes the
implementation of the type-I seesaw mechanism very natural, the more that the RH
neutrino Majorana mass is forbidden in the unbroken phase (there is no $SO(10)$ singlet
Figure 3.1: The most common breaking chains of $SO(10)$ down to the $SU(3)_c \otimes U(1)_Q$ gauge symmetry at low energies. The associated scalar irreps capable of triggering the indicated transitions are indicated along each of the lines by the symbols of Table 3.2.

in $16 \otimes 16$) and, as such, it naturally emerges below the GUT scale as a consequence of the spontaneous $SO(10)$ breaking.

This, however, can be triggered by many different scalar fields and may proceed through various intermediate symmetry stages, see Table 3.2 and Fig. 3.1. Barring the very exotic sequence triggered by $144$, cf. [103] at least two different scalar irreps must be employed in order to get from $SO(10)$ down to the SM; the minimal options are $45 \oplus 16$ (or $45 \oplus 126$) capable of passing through the $B - F$ chain or $210 \oplus 16$ (or $210 \oplus 126$) passing through the $B - F$, $A - D - F$ or $A - E$ sequences. Naturally, these also identify the basic model building strategies found in the literature.

The last common ingredient of all $SO(10)$ unified models is the set of gauge fields which is hosted by the 45-dimensional adjoint representation decomposing as

\[
45 = (8, 1, 0) \oplus (1, 3, 0) \oplus (1, 1, 0) \oplus (3, 2, +\frac{5}{6}) \oplus (3, 2, -\frac{5}{6}) \oplus (3, 2, -\frac{1}{6}) \oplus (3, 2, +\frac{1}{6})
\]

under the $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ subgroup.

\[\text{Note that this is not the case of the simple extensions of the Georgi-Glashow $SU(5)$ GUTs in which the RH neutrinos enter as full gauge singlets.}\]
### Table 3.2: The “symmetry breaking power” of various scalar $SO(10)$ irreps up to dimension 210. The quantum numbers of the submultiplets correspond to the Pati-Salam subgroup $SU(4)_C \otimes SU(2)_L \otimes SU(2)_R$. The letters denoting the different breaking steps are those used in Fig. 3.1.

<table>
<thead>
<tr>
<th>$SO(10)$ irrep</th>
<th>PS sub-multiplet</th>
<th>Symmetry breaking steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>$(1, 2, 2)_{10}$</td>
<td>$H$</td>
</tr>
<tr>
<td>16</td>
<td>$(4, 2, 1)_{16}$</td>
<td>$H$</td>
</tr>
<tr>
<td></td>
<td>$(\bar{4}, 1, 2)_{16}$</td>
<td>$C', G'; E, F$</td>
</tr>
<tr>
<td>45</td>
<td>$(1, 3, 1)_{45}$</td>
<td>$C, G'; F'$</td>
</tr>
<tr>
<td></td>
<td>$(15, 1, 1)_{45}$</td>
<td>$B, C; D$</td>
</tr>
<tr>
<td>54</td>
<td>$(1, 1, 1)_{54}$</td>
<td>$A, G$</td>
</tr>
<tr>
<td>120</td>
<td>$(1, 2, 2)_{120}$</td>
<td>$G'; H$</td>
</tr>
<tr>
<td></td>
<td>$(15, 2, 2)_{120}$</td>
<td>$G'; H$</td>
</tr>
<tr>
<td>126</td>
<td>$(15, 2, 2)_{126}$</td>
<td>$H$</td>
</tr>
<tr>
<td></td>
<td>$(10, 1, 3)_{126}$</td>
<td>$E, F; G'$</td>
</tr>
<tr>
<td>144</td>
<td>$(4, 2, 1)_{144}$</td>
<td>$H$</td>
</tr>
<tr>
<td></td>
<td>$(\bar{4}, 1, 2)_{144}$</td>
<td>$C', G, G'; E, F$</td>
</tr>
<tr>
<td>210</td>
<td>$(1, 1, 1)_{210}$</td>
<td>$A; C', G, G'$</td>
</tr>
<tr>
<td></td>
<td>$(15, 1, 1)_{210}$</td>
<td>$B, D; C', G, G'$</td>
</tr>
<tr>
<td></td>
<td>$(10, 2, 2)_{210}$</td>
<td>$H$</td>
</tr>
<tr>
<td></td>
<td>$(\bar{10}, 2, 2)_{210}$</td>
<td>$H$</td>
</tr>
</tbody>
</table>

#### 3.2.1 Proton decay in $SO(10)$ GUTs

Note that the first five components of decomposition (3.17) correspond to the gauge fields of the Georgi-Glashow model and there is just one new vector field $Y \oplus \bar{Y}$ present in the $SO(10)$ case. In this respect, the gauge-driven baryon number violation phenomenology of the $SO(10)$ GUTs naturally encompasses all that has been said about it in the G-G context. The only extra effect due to the presence of $Y \oplus \bar{Y}$ consists in their capacity to provide an extra contribution to the amplitudes corresponding to the $O_1$ effective $d = 6$ operator of Table 2.1 (see also Table 2.2), and, hence, loosen the correlations between the $O_1$- and $O_2$-driven effects characteristic to the $SU(5)$ settings. The presence of a second $B$ and $L$ “active” vector multiplet also brings in a second suppression scale which, in principle, can be quite different from the mass of $X$, especially along the symmetry breaking directions featuring an intermediate $SU(5)$ or flipped $SU(5)$ stage, cf. Sect. 3.6.2.
Concerning the minimal SUSY variants of the SO(10) GUTs one would naively expect that they must suffer from the same drawbacks related to the insufficient suppression of the $d = 5$ Higgsino-driven BLNV amplitudes like the supersymmetric $SU(5)$ GUTs discussed at length in Sect. 2.6.4. However, this issue is typically less severe in the SO(10) context as there is often more than a single $\Delta_c$-type (cf. Table 2.2) triplet Higgsino in the theory spectrum. This, in turn, yields more room to arrange their mixing in such a way that the lightest of the corresponding mass eigenstates (i.e., the field whose contribution should kinematically dominate the BLNV amplitudes) has, for instance, suppressed couplings to the first generation (s)quarks and (s)leptons.

### 3.2.2 Yukawa sector of simple SO(10) GUTs

In general, the proliferation of the $\Delta_c$-type triplets in the SO(10) models has to do with the complete unification of the SM matter families within SO(10) spinors. As beautiful as this is, it makes it essentially impossible to accommodate all the distinct features of the SM matter spectrum and mixing in a model with just a single “Yukawa-active” irrep coupled to the matter bilinear $16 \otimes 16$ (at the renormalizable level). The corresponding decomposition\(^5\)

$$16 \otimes 16 = 10 \oplus 120 \oplus 126 \quad (3.18)$$

indicates that, along with the minimal choice of $10$, the most economical renormalizable models are those in which the rank reduction (i.e., $U(1)_{B-L}$ breaking) is triggered by the scalar $126$ rather than $16$ because the former can help also with attaining a rich-enough flavour structure, at least at the renormalizable level. It is also worth noting that out of the three options of (3.18) it is again only $126$ which contains a SM singlet and whose VEV can thus source the Majorana mass term for the RH neutrinos.

Therefore, in what follows, we shall stick to the renormalizable SO(10) grandunified models with the powerful $126$ tensor high in their field-contents list.

### 3.3 The spectacular failure of the minimal SUSY SO(10)

To this end, in the current and in the next sections we shall first discuss the minimal supersymmetric SO(10) GUT and only later on switch to the non-SUSY variant. This approach closely follows the historical development of the field which, on the non-SUSY side, was hindered by the early observation of tachyonic instabilities in the spectrum of

\(^5\)Algebraically, the three factors correspond to a one-, three- and five-index fully antisymmetric tensors where, for the last one, only the self-dual component of the full 252-dimensional maximally antisymmetric tensor is taken.
the minimal model, cf. Sect. 3.4.3; at the same time (i.e., in mid 1980’s) SUSY became so widely popular\footnote{In this respect it is interesting to note that the rocketing popularity of the low-energy SUSY in the middle of 1980’s was partly fuelled by the need to find a successor of the minimal SU(5) which failed miserably on the prediction of the weak mixing angle, cf. Sect. 2.6.3. Remarkably, the mainstream went for further complication (marrying SU(5) with SUSY) rather than re-thinking the unification basics.} that there was almost no impetus to explore such issues any further, cf. Sect. 3.4.

3.3.1 The structure of the minimal SUSY $SO(10)$ GUT

With what has been said so far, one can right away formalize the hypothesis for the structure of the (renormalizable) Yukawa part of the minimal SUSY $SO(10)$ superpotential, namely,

$$W_Y = (16_M^T Y_{10} 16_M) 10_H + (16_M^T Y_{126} 16_M) 126_H,$$

(3.19)

where $16_M$ stands for a vector of three generations of matter, $Y_{10}$ and $Y_{126}$ denote the $3 \times 3$ Yukawa matrices (both symmetric due to the structure of the $SO(10)$ contractions) and $10_H$ and $126_H$ are the two Higgs multiplets identified above.

Before writing the remaining parts of the superpotential involving, among other things, the Higgs field self-interactions (and, thus, implicitly, the scalar potential of the model), one has to choose very carefully the set of the GUT-symmetry breaking fields\footnote{We already know that 126 is not enough as the SM singlet within is also a singlet of SU(5) that would remain unbroken.}.

This is subject to several important requirements:

1. SUSY should not be broken along with the GUT symmetry. The reason is that GUT-scale $F$- or $D$-terms make it very hard to obtain the soft SUSY breaking scale in the desired TeV-scale ballpark. Since, however, a VEV of the complex $126$ inevitably leads to $\langle D_{126} \rangle \neq 0$ the only way to ensure this is to include also the complex conjugated multiplet $126$ whose contribution would cancel $\langle D_{126} \rangle$.

2. The “SU(5) trap”, i.e., an intermediate stage looking very much like the problematic minimal SUSY SU(5) theory of Sect. 2.6.4 should be avoided by adding a Higgs superfield whose SM singlets are not simultaneously singlets of the SU(5). There are several such options offered in Table 3.2 like, e.g., 45, 54 and 210. Interestingly, 45 does not work because the $F$-terms align its VEVs with that of $126_H$ \footnote{The SM Higgs boson should be spanned on the doublet components of both 10 and 126, otherwise one is back to the overly rigid situation of only one effective} and SU(5) remains unbroken.

3. The SM Higgs boson should be spanned on the doublet components of both 10 and 126, otherwise one is back to the overly rigid situation of only one effective
Yukawa structure at play. Hence, there should be a term in the superpotential providing their mixing; out of the list above this can be done (at the renormalizable level) only by $210$.

Thus, the complete Higgs sector of the minimal potentially realistic SUSY $SO(10)$ [105–111] contains 4 irreps, namely, $10 \oplus 126 \oplus \overline{126} \oplus 210$ with the corresponding superpotential of the form

$$W_H = M_{10}^{10H} + M_{126}^{126H} \overline{126}H + M_{210}^{210H} +$$

$$+ \lambda^{210} + \eta^{210}126H \overline{126}H + \alpha^{10}126H210H + \beta^{10} \overline{126}H210H.$$  \hspace{1cm} (3.20)

In spite of the complexity of $W_H$ above several groups succeeded in calculating its spectrum in the SUSY limit and, thus, a complete analysis of the GUT-scale thresholds – a crucial ingredient of any gauge running study – became possible, see e.g. [111–113]. The scalar and Higgsino masses can be written in terms of four independent VEVs (three real ones in $210$ and one complex in $126 \oplus \overline{126}$) usually denoted by

$$\langle (1,1,1)210 \rangle \equiv p, \quad \langle (15,1,1)210 \rangle \equiv a, \quad \langle (15,1,3)210 \rangle \equiv \omega, \quad \langle (10,1,3)_{\overline{126}} \rangle \equiv \sigma. \quad (3.21)$$

Notice that $(10,1,3)_{\overline{126}}$ contains the $\Delta_R$ scalar of the LR setting discussed in Sect. 3.1.1 and, thus, $\sigma$ drives the scale of the type-I+II seesaw contributions to the netrino masses in this model. Moreover, with $126$ at play, the flavour structure of both the type-I+II contributions are intimately related to the same Yukawa matrix $Y_{126}$.

From the flavor perspective, such a seesaw structure has also got other interesting properties:

1. The type-II contribution to the light neutrino masses is correlated with the charged sector Yukawa matrices via

$$Y_d v_d = Y_{10}^{10v_d} + Y_{126}^{126v_d} \equiv M_d \quad (3.22)$$

$$Y_l v_d = Y_{10}^{10v_d} - 3Y_{126}^{126v_d} \equiv M_l$$

(with $v_d^{10}$ and $v_d^{126}$ denoting the projections of the MSSM down-type Higgs doublet VEV onto the defining components in $10_H$ and $\overline{126}_H$), which is a consequence of the simplicity of $W_Y$ of (3.19). In particular, one has

$$Y_{126} \propto M_d - M_l. \quad (3.23)$$

The absolute size of this neutrino mass contribution is then proportional to $\sigma$ times the product of the relevant projections of $(1,2,2)_{10}$ and $(\overline{10},2,2)_{210}$ onto

---

*In the SUSY limit the Higgs and Higgsino masses are identical and, thus, it is sufficient to calculate the latter.*
the physical MSSM Higgs doublets. Note that all these factors are in principle calculable within a complete model.

2. If the type-II contribution dominates the seesaw formula one obtains an intriguing correlation between the size of the atmospheric neutrino mixing and the stipulated GUT-scale convergence of the $b$-quark and $\tau$-lepton Yukawa couplings \[114\]

\[
\tan^2 \theta_A \sim \frac{\sin 2 \theta_q}{2 \sin^2 \theta_q - \left(1 - \frac{y_\tau}{y_b}\right)},
\]

(3.24)

(with $\theta_q$ corresponding to the 2–3 CKM mixing) justifying its tendency towards maximality indicated by the experiment. Besides that, a relatively large reactor mixing angle is strongly preferred in the physically viable parts of the parameter space \[115\].

3. Also the flavour structure of the RH neutrino Majorana mass is proportional to $Y_{126}$ which, in turn, enters the type-I seesaw formula as an inverse. The overall size of this contribution to the light neutrino mass matrix is, however, inverse proportional to $\sigma$.

3.3.2 The neutrino challenge to SUSY GUTs

As promising as the initial observations look, the minimal SUSY $SO(10)$ scheme turns out to be terminally ill when it comes to the global analysis of its flavour structure together with the constraints from the gauge unification. The devil is, as always, in detail (cf. \[116\] and \[117\]):

1. It turns out that no viable complete flavour fits exist if type-I seesaw contribution to the light neutrino masses is suppressed (i.e., for large $\sigma$).

2. The fits in which type-I contribution is significant thus require the $B - L$ breaking scale $\sigma$ well below the GUT scale and a very specific pattern of the MSSM VEV projections onto the defining doublets in $10$ and $\overline{126}$.

3. This, however, pushes the model into a regime in which a set of pseudo-Goldstone modes develop several orders of magnitude below the GUT scale and, hence, ruin completely the “too good to be true” MSSM gauge coupling convergence pattern.

Hence, the minimal potentially realistic supersymmetric incarnation of the $SO(10)$ grand unification paradigm of Sect. 3.3.1 has been decisively futilised in \[116\] and \[117\]. Though there have still been later attempts to save the situation by proposing minor amendments to the original scheme (see, e.g., \[118–120\]) the model has been to a large
degree abandoned by the community and it is no longer considered as a viable route towards a complete theory of perturbative baryon number violation.

3.4 Quantum salvation of the minimal $SO(10)$ GUT

With the strict no-go for the minimal SUSY $SO(10)$ revealed in the previous section it is more than natural to turn one’s attention back to the minimal non-supersymmetric version of the $SO(10)$ GUT, the more that TeV-scale SUSY becomes less and less appealing with the latest null results of (not only) the relevant LHC searches.

3.4.1 The cons and pros of the non-SUSY $SO(10)$ GUTs

Naïvely, life gets only more complicated in the non-SUSY context. With less symmetry imposed one has to deal with, e.g., higher-order operators governing the minimal defining structures ($d = 4$ interaction Hamiltonian vs. $d = 3$ superpotential at the renormalizable level), a more complicated Yukawa sector including non-holomorphic couplings, the issues related to the need for intermediate stages in the symmetry breaking pattern (as only one of the aspects of a generally more complicated quantum structure of the theory) etc.

On the other hand, with less assumptions on the theory shape the stakes are generally higher as it is way more straightforward to learn a lesson from its eventual failure than in the SUSY context. The non-supersymmetric unifications are also way easier to treat perturbatively in the vicinity of the GUT scale as the number of degrees of freedom to be integrated over in loop diagrams is reduced considerably.

3.4.2 The tree-level vacuum of the minimal $SO(10)$ Higgs models

One of the most delicate questions to be addressed in the framework of non-SUSY unifications are those concerning the structure of their vacuum. As there is no guarantee of the scalar potential convexity in its extrema the calculation of the scalar spectrum (not possessing any fermionic counterpart) is always connected to the issues of its positivity. Hence, with the non-supersymmetric GUTs, any analysis of their viability must begin with a careful inspection of the corresponding Higgs sector.

As unlikely as it sounds, a full analysis of even the minimal $SO(10)$ Higgs model may be in fact a rather formidable. Following the reasoning of Sect.3.2 its field content can be identified readily – either it is spanned on $45 \oplus 16$ or on $45 \oplus 126$, with different

---

9Recall that in supersymmetric theories the supersymmetric minimum is always the global one.
components therein playing similar roles (i.e., avoiding the \(SU(5)\) “trap”, triggering the rank reduction etc.) as in the SUSY context.

In what follows we shall consider both the \(45 \oplus 16\) and \(45 \oplus 126\) options along with each other as their main features (especially those relevant for the operation of the symmetry breaking mechanism) are to a large degree similar in both these settings. Despite the proximity of \(M_G\) to the Planck scale we shall stick to the renormalizable version(s) of these minimal model(s) as, in most cases, the contributions of the higher order operators should play a sub-leading role (if not entirely negligible, see Sect. 3.4.4).

**The tree-level scalar potentials and masses**

The \(45 \oplus 16\) variant of the minimal \(SO(10)\) Higgs model is defined by the scalar potential of the form

\[
V_{45\oplus 16} = V_{45} + V_{16} + V_{45-16}, \tag{3.25}
\]

where, following the definitions given in Appendix A of [121] (and omitting any subscripts distinguishing between different types of fields\(^{[17]}\)),

\[
V_{45} = -\frac{\mu^2}{4} \text{Tr} 45^2 + \frac{a_1}{4} (\text{Tr} 45^2)^2 + \frac{a_2}{4} \text{Tr} 45^4, \tag{3.26}
\]

\[
V_{16} = -\frac{\nu^2}{2} 16^\dagger 16 + \frac{\lambda_1}{4} (16^\dagger 16)^2 + \frac{\lambda_2}{4} (16^\dagger \Gamma_j 16_-)(16^\dagger \Gamma_j 16_+),
\]

and

\[
V_{45-16} = \alpha (16^\dagger 16) \text{Tr} 45^2 + \beta 16^\dagger 45^2 16 + \tau 16^\dagger 45 16. \tag{3.27}
\]

In case of the \(45 \oplus 126\) variant the scalar potential shares the universal \(V_{45}\) part (3.26) with that of (3.25) but differs in the other two terms, namely

\[
V_{45\oplus 126} = V_{45} + V_{126} + V_{45-126}, \tag{3.28}
\]

with

\[
V_{126} = -\frac{\nu^2}{5!} (126^* 126)_0
\]

\[
+ \frac{\lambda_0}{(5!^2)} (126^* 126)_0 (126^* 126)_0 + \frac{\lambda_2}{(4!)^2} (126^* 126)_2 (126^* 126)_2
\]

\[
+ \frac{\lambda_4}{(3!)^2 (2!)^2} (126^* 126)_4 (126^* 126)_4 + \frac{\lambda'_4}{(3!)^2} (126^* 126)_4 (126^* 126)_4
\]

\[
+ \frac{\eta_2}{(4!)^2} (126 126)_2 (126 126)_2 + \frac{\eta'_2}{(4!)^2} (126^* 126^*)_2 (126^* 126^*)_2,
\]

\(^{[10]}\)Note that, unlike in the SUSY context, all fields involved here are Lorentz scalars.
and

\[
V_{45-126} = \frac{i\tau}{4!}(45)_2(126^*126)_2 + \frac{\alpha}{2 \cdot 5!}(45_045)(126^*126)_0 \tag{3.30}
\]
\[
+ \frac{\beta_4}{4 \cdot 3!}(45_445)(126^*126)_4 + \frac{\beta_4'}{3!}(45_445)(126^*126)_4'
\]
\[
+ \frac{\gamma_2}{4!}(45_245)(126126)_2 + \frac{\gamma_2}{4!}(45_245)(126^*126^*)_2.
\]

The brackets above correspond to the following $SO(10)$ covariant structures (with or without complex conjugation):

\[
(126^*126)_0 \equiv 126^*_{ijklm}126_{ijklm}, \tag{3.31}
\]
\[
(126^*126)_2 \equiv (126^*126)_{mn} \equiv 126^*_{ijklm}126_{ijkln},
\]
\[
(126^*126)_4 \equiv (126^*126)_{lmno} \equiv 126^*_{ijklm}126_{ijlko},
\]

where all the latin indices run from 1 to 10 and their pairs are summed over. The contractions of these terms in Eqs. (3.29) and (3.30) are obvious with the only exceptions of the $4'$ brackets that read

\[
(126^*126)_4'(126^*126)_4' \equiv (126^*126)_{lmno}(126^*126)_{lnmo}, \tag{3.32}
\]
\[
(45_445)_4'(126^*126)_4' \equiv 45_{lm45no}(126^*126)_lnmo.
\]

The vacuum structure of the minimal $SO(10)$ models

The adjoint $45$ of $SO(10)$ decomposes under the $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ of the SM as

\[
45 = (3, 2, -5) \oplus (\overline{3}, 2, +5) \oplus (8, 1, 0) \oplus (1, 3, 0) \oplus (3, 1, +1) \oplus (\overline{3}, 1, -1) \oplus (1, 1, +1) \oplus (1, 1, -1) \oplus (1, 1, 0). \tag{3.33}
\]

Hence, there are two SM singlets in this multiplet which, in realistic scenarios, may possess non-zero VEVs. Note, however, that these two singlets are not equivalent from the viewpoint of the intermediate symmetries: one of them descends from the $(15, 1, 1)$ Pati-Salam component of $45$ while the other one resides in $(1, 1, 3)$. Hence, the VEV of the former (to be called $\omega_R$) preserves the $SU(2)_R$ subgroup of the $SO(10)$ while the latter ($\omega_{BL}$) leaves intact the $U(1)_{B-L}$ factor.

Concerning the complex irrep of the models (i.e., $16$ or $126$) one of their common features is the presence of one SM singlet which is capable of breaking the $U(1)_{B-L}$ symmetry and, thus, set the seesaw scale. At the same time, this field is also a singlet of the $SU(5)$ subgroup of $SO(10)$ and, thus, its VEV (to be denoted $\sigma$) is not sufficient to provide the entire $SO(10)$ symmetry breaking, cf. Sect. 3.3.1.
The masses of the scalar triplet and octet of the 45

In what follows we shall focus on the masses of the fields in the first line of the decomposition (3.33) which, due to their transformation properties, do not mix with any of the fields in $16$ or $126$. This means that the minimisation of the scalar potential yields the same simple formulae in either of the two settings, namely,

\begin{align*}
M^2_{(1,3,0)} &= 2a_2(\omega_R - \omega_{BL})(\omega_{BL} + 2\omega_R), \\
M^2_{(8,1,0)} &= 2a_2(\omega_{BL} - \omega_R)(\omega_R + 2\omega_{BL}).
\end{align*}

Remarkably enough, these expressions are simultaneously non-negative if and only if

\begin{equation}
 a_2 > 0 \quad \text{and} \quad -2 < \frac{\omega_{BL}}{\omega_R} < -\frac{1}{2}.
\end{equation}

Note, however, that the required proximity (up to a sign) of $\omega_R$ and $\omega_{BL}$ prefers a very specific shape of the vacuum manifold corresponding to a symmetry breaking pattern passing through the vicinity of the so-called flipped $SU(5) \otimes U(1)$ intermediate stage (corresponding to the $\omega_R = -\omega_{BL}$ situation).

However, unlike in the settings discussed below in Sect. 3.6.2 (in which only the non-abelian gauge couplings of the SM are required to unify) this observation represents a serious problem in the $SO(10)$ GUT context. Qualitatively, the issue is somewhat similar to that encountered in Sect. 3.3.2: on one hand, there is a need to have the seesaw scale $\sigma$ well below $M_G \sim 10^{16}$ GeV to conform the neutrino sector constraints\footnote{Actually, the gauge unification constraints on $\sigma$ from the consistency of the gauge unification pattern in non-SUSY settings\cite{122,125} prefer it in the vicinity of $10^{19}$ GeV, i.e., even lower than the scale favoured by the seesaw!} but, on the other hand, the flipped $SU(5) \otimes U(1)$ intermediate symmetry must be broken (by the same $\sigma$) in the proximity of $M_G$ to avoid the issues with overly fast proton decay. Hence, the non-SUSY $SO(10)$ GUTs with the first symmetry breaking step driven by the adjoint 45 irrep are very unlikely to conform even the basic phenomenological requirements including the non-tachyonicity of the scalar spectrum in the potentially realistic symmetry-breaking chains (i.e., those avoiding an intermediate-scale $SU(5)$-like gauge dynamics).

3.4.3 The minimal $SO(10)$ Higgs model(s) at the loop level

Interestingly, most of what is written above was understood already in the early 1980’s\cite{126,129} and, since then, it was generally assumed that the field of renormalizable $SO(10)$ GUTs broken by either $45 \oplus 16$ or $45 \oplus 126$ scalar fields is sterile. To this end, the situation changed completely in 2010 with the author’s study\cite{130} where it was...
shown that the tachyonicity of the scalar spectrum along the potentially realistic symmetry breaking chains (i.e., those passing through either the $SU(4)_C \otimes SU(2)_L \otimes U(1)_R$ or the left-right symmetric $SU(3)_C \otimes SU(2)_L \otimes SU(2)_R \otimes U(1)_{B-L}$ intermediate stages) is an artefact of the tree-level analysis and can be avoided if radiative corrections are taken into account.

The merit of [130] consists in the observation that the tree-level mass formulae (3.34) and (3.35) are unexpectedly simple given the fact that, in principle, all terms in $V_{45}$ and $V_{45-16/126}$ (besides the $\tau$ piece therein) should contribute already at the tree level. A thorough investigation of this phenomenon reveals that this is all due to the pseudo-Goldstone nature of the $(1,3,0)$ and the $(8,1,0)$ fields which, in the limit of “moduli-only” terms kept in the potential, become exact Goldstone modes of the spontaneously broken enhanced global symmetry; hence, their masses should be proportional to just a limited set of parameters which break this global symmetry explicitly. In the same perspective, the presence of only the $a_2$-proportional contribution in (3.34) may be attributed to the particular shape of all the other potentially relevant tree-level contractions in $V_{45-16/126}$. Since, however, this does not necessarily apply to the radiative corrections, the formulae (3.34) are expected to receive different types of loop contributions which, in turn, may be capable of resolving the tachyonicity conundrum of Sect. 3.4.2.

This expectation if further justified by the fact that, in the perturbative regime, $a_2$ is almost automatically pushed well below 1 by one of the minimisation conditions. This, in the $45 \oplus 126$ setting, reads

$$\tau = 2\beta_4'(3\omega_{BL} + 2\omega_R) + a_2 \frac{\omega_{BL}\omega_R}{|\sigma|^2}(\omega_{BL} + \omega_R),$$

(3.37)

which, given the need to have the seesaw scale $|\sigma|$ well below $M_G$ corresponding to the leading $\omega$ (and assuming that the other one is not parametrically smaller), can be fulfilled in the perturbative regime (i.e., with loops playing a sub-leading role here) only for $a_2 \lesssim |\sigma|^2/\omega_R\omega_{BL}$.

---

12 This expectation reflects the presence of at least two powers of 45 in each of these terms which, in the broken phase, should generate bilinears for the fields in 45.

13 For instance, the absence of the contribution from the $\beta$-proportional part of $V_{45\oplus16}$ can be understood by noticing that the relevant term has got the same group structure as the corresponding piece in the mass matrix for the gauge fields. Since, however, the VEVs in the scalar 45 can not contribute to the masses of the gluons or the SM $A$-fields, the same must happen (at least at the tree level) to their scalar counterparts.

14 It can be objected that the specific form of the tadpole equation (3.37) corresponds to the tree level approximation only; it should, however, still be the leading piece even at higher loops if the theory is perturbative and, thus, the tree-level results should represent a good approximation to the full case.

15 This is likely to be so because in the opposite case it is typically difficult to get a completely non-tachyonic rest of the scalar spectrum.
Let us illustrate this by writing down the one-loop contributions to \( M^2_{(1,3,0)} \) and \( M^2_{(8,1,0)} \) from the Feynman diagrams including the gauge degrees of freedom – these corrections are indeed universal to both the \( 45 \oplus 16 \) and \( 45 \oplus 126 \) models. A thorough analysis of the relevant piece of the one-loop effective potential à la Coleman and Weinberg [131] reveals [132]

\[
\Delta^1\text{-loop}_\text{gauge} M^2_{(1,3,0)} = \frac{g^4}{16\pi^2} (19\omega^2_{BL} + \omega_{BL} \omega_R + 16\omega_R^2) + \Delta^\log_{(1,3,0)},
\]

\[
\Delta^1\text{-loop}_\text{gauge} M^2_{(8,1,0)} = \frac{g^4}{16\pi^2} (22\omega^2_{BL} + \omega_{BL} \omega_R + 13\omega_R^2) + \Delta^\log_{(8,1,0)},
\]

where

\[
\Delta^\log_{(1,3,0)} = \frac{3g^4}{16\pi^2} \left\{ 8\omega_{BL} (|\sigma|^2 + \omega_{BL}^2) \log \left[ \frac{2g^2 (|\sigma|^2 + \omega_{BL}^2)}{\mu^2} \right] \right. \\
- 4\omega_R (|\sigma|^2 + \omega_R^2) \log \left[ \frac{2g^2 (|\sigma|^2 + \omega_R^2)}{\mu^2} \right] + 2\omega' \log \left[ \frac{1}{2} g^2 \omega^2 \right] \\
- \left[ 4|\sigma|^2 (2\omega_{BL} - \omega_R) + \omega^2 (5\omega_{BL} - 4\omega_R) \right] \log \left[ \frac{2g^2 (|\sigma|^2 + \omega_{BL}^2)}{\mu^2} \right] \right\},
\]

\[
\Delta^\log_{(8,1,0)} = \frac{3g^4}{32\pi^2} \left\{ 4 \left[ |\sigma|^2 (3\omega_{BL} + \omega_R) + \omega^2_{BL} (\omega_{BL} + 3\omega_R) \right] \log \left[ \frac{2g^2 (|\sigma|^2 + \omega_{BL}^2)}{\mu^2} \right] \\
- 8\omega_R (|\sigma|^2 + \omega_R^2) \log \left[ \frac{2g^2 (|\sigma|^2 + \omega_R^2)}{\mu^2} \right] + \omega' \log \left[ \frac{1}{2} g^2 \omega^2 \right] \\
- \left[ 4|\sigma|^2 (3\omega_{BL} - \omega_R) + \omega^2 (7\omega_{BL} - 5\omega_R) \right] \log \left[ \frac{2g^2 (|\sigma|^2 + \omega_{BL}^2)}{\mu^2} \right] \right\},
\]

provided

\[
\omega \equiv \omega_R + \omega_{BL} \quad \text{and} \quad \omega' \equiv \omega_R - \omega_{BL}.
\]

It is not difficult to show that there is indeed a lot of points in the \((\omega_{BL}, \omega_R, |\sigma|)\) space corresponding to the phenomenologically preferred symmetry breaking chains (i.e., those avoiding intermediate SU(5) stages) where both expressions (3.38) and (3.39) are positive.

### 3.4.4 The minimal potentially realistic and testable GUTs

Hence, after almost 30 years in oblivion, the minimal \( SO(10) \) Higgs model was brought back to life [133–136] as a seed of a potentially realistic theory which, however, is inherently of a quantum nature. The natural question is then whether such a theory
can, in some of its parameter space point(s), accommodate all the low-energy data and, if affirmative, what would be its predictions for the new physics signals such as proton decay, leptonic CP violation, absolute neutrino mass scale etc.

**Planck-scale effects in the GUT-scale determination**

It is remarkable that these efforts can be further justified by another rather unique feature the minimal SO(10) Higgs models possess, namely, their particular robustness with respect to the Planck-scale-induced effects in the gauge unification analysis. To see this, let us recall that whenever any unified gauge symmetry is being broken by a multiplet (to be called $\Phi$) which admits a $d = 5$ coupling to the relevant gauge-kinetic form

$$L^{(5)} \ni \frac{\rho}{\Lambda} F_{\mu\nu} \Phi F^{\mu\nu},$$

with $\rho$ denoting a dimensionless (presumably $O(1)$) effective coupling and $\Lambda$ standing for the effective cut-off scale, one gets a non-canonical gauge-kinetic form in the broken phase

$$L_{\text{kin}} \ni -\frac{1}{4} \left(1 - 4\rho \frac{\langle \Phi \rangle}{\Lambda} \right) F_{\mu\nu} F^{\mu\nu}.$$  \hspace{1cm} (3.44)

For the sake of retaining the standard perturbative expansion, the gauge fields must be first canonically normalized by a suitable rescaling transformation which, however, depends (on the ratio of) two, in principle unknown, quantities $\rho$ and $\Lambda$. This, in general, induces inhomogeneous shifts in the definitions of the three effective SM couplings in terms of the unified one and, hence, uncertainties in the corresponding matching between the GUT and any lower-energy theory. Surprisingly, even for as small as 1% effects of this kind (corresponding to the very natural choice of $\Lambda = M_{Pl}$ and $\rho \sim 1$ with $\langle \Phi \rangle \sim 10^{16}$ GeV), these errors can have serious impact on the precision calculations of $M_G$ as they enter the relevant formulae exponentially. In practice, the resulting uncertainty can easily “smear” thus obtained $M_G$ into a domain stretching over more than an order of magnitude! This, however, sets a limit on the accuracy attainable in most proton lifetime calculations (typically at the level of several orders of magnitude) which, in turn, renders all attempts to discriminate among different models on the basis of (non)observation of proton decay essentially meaningless.

**Taming the leading Planck-scale effects in the minimal SO(10) GUT**

Interestingly, the irreducible theoretical uncertainties of the kind described above are absent from the minimal SO(10) GUT. The reason is that its unified-symmetry-breaking VEV resides in the scalar transforming as the adjoint $45$ whose coupling to the $F_{\mu\nu} F^{\mu\nu}$ bilinear is identically zero due to the antisymmetry of $45^{ab}$ in the group indices, i.e.,
\[ F^{\mu \nu} \Phi^{ab} F_{b \mu \nu} = 0. \] Note that this is not the case of the other popular \( SO(10) \) symmetry breaking models utilising either 54 or 210 because both these multiplets are present in the symmetric product of two adjoints:

\[ [45 \otimes 45]_{\text{sym}} = 54 \oplus 210 \oplus 770. \] (3.45)

In this sense, the \( SO(10) \) models with the GUT-scale symmetry breaking triggered by the adjoint are arguably very unique concerning their proton-decay predictive potential. This, together with the ongoing construction of the new generation of large-scale detectors such as Hyper-K \[68\] or DUNE \[137\] (fuelled mainly by their potential to serve as very powerful neutrino telescopes), is the main reason why this specific class of the minimal \( SO(10) \) scenarios has been receiving so much attention recently.

The minimal potentially realistic and testable \( SO(10) \) GUT

The first steps along the lines of constructing and working out the potentially realistic and testable \( SO(10) \) GUTs of this kind have been attempted recently in studies \[133\] \[136\]. The salient features of these settings are:

- **Scalar sector containing 45 \( \oplus \) 126**: The choice of the 5-index antisymmetric (anti-) self-dual tensor of the \( SO(10) \) rather than the spinorial 16 discussed in Sect. 3.4.2 is motivated mainly by the preference of a renormalizable (and, thus, potentially predictive) Yukawa sector. For that sake, 126 is almost ideal as it contributes not only to the effective Yukawa couplings of the charged SM matter fermions but it also generates a large mass term for the RH neutrinos (and, thus, a natural type-I seesaw contribution\[16\] to the light neutrino masses). On top of that, the \( U(1)_{B-L} \) subgroup of \( SO(10) \) broken by the SM singlet of 126 leaves behind a residual \( Z_2 \) symmetry which behaves like a matter parity and, thus, can stabilize fermionic dark-matter candidates of various kinds \[138\]. Note also that the Yukawa sector of the minimal potentially realistic renormalizable model must be equipped with one more scalar irrep in order to smear the effective Yukawa degeneracies across different matter sectors; this is usually taken to be 10, partly for its minimality and partly for the fact that it does not interfere with any of the findings above due to the absence of any SM singlets within.

- **Complicated vacuum structure**: The purely quantum nature of the models under consideration bring a notorious difficulty with the classification of the shapes of

\[16\] Recall that there is typically also a type-II seesaw piece emerging from the induced sub-electroweak-scale VEV of the scalar \( SU(2)_L \) triplet in 126, cf. Sect. 3.1.1.
the scalar spectrum conforming the conditions of non-tachyonicity and perturbativity which are obvious prerequisites of any sensible attempts of the GUT-scale determination as the precursor of the subsequent proton lifetime calculations. The effective potential methods are typically used for this purpose (cf. Sect. 3.4.3) but even with these at hand a complete chart of the regions of the model’s parameter space conforming these constraints is still subject of an intensive research.

- **Intermediate scales**: Irrespective of the details of the high-energy spectrum in the fully realistic settings there are features that can be expected already at the current level of understanding. The most prominent is perhaps the need to push (at least) one of the naturally heavy scalar fields well below the GUT scale, otherwise there would be no realistic gauge unification pattern supporting a $B - L$ breaking (seesaw) scale in its phenomenologically preferred ballpark of about $10^{12-13}$ GeV. Pushed to the extreme, it is even conceivable to get some of these states relatively close to the LHC domain, cf. [133](#).

For more information the reader is kindly deferred to the aforementioned studies [133](#), [136](#) and a recent review article [132](#).

### 3.5 Aspects of renormalization group evolution in theories with more than a single U(1) gauge factor

In the $SO(10)$ GUTs of Sect. 3.1.3 or their left-right-symmetric descendants discussed in Sect. 3.1.1 one often encounters a situation in which an intermediate-scale effective gauge theory features two Abelian factors (like, e.g., $U(1)_R \otimes U(1)_{B-L}$ of the breaking chains in Fig. 3.1 passing along the $F'$ branch there). As innocent as such a situation looks the occurrence of multiple $U(1)$ gauge factors has rather non-trivial consequences for the quantum structure of the theory including a spectacular proliferation of couplings required for their formal renormalizability and also for the consistency of the overall physical picture.

In this section we shall pass through the basics of the renormalization procedure in such settings starting with a short review of the situation in the spinorial QED

---

*To this end, it is perhaps worth noting that there are just two candidate multiplets in the minimal $45 \oplus 126$ $SO(10)$ Higgs model that can support viable symmetry breaking patterns without invoking any other field in the GUT desert: the $(8, 2, +\frac{1}{2})$ and $(6, 3, +\frac{1}{3})$ scalars. Remarkably enough, the same fields have been identified recently [139](#) in the low-energy SM effective field theory approach to the peculiar $B$-decay anomalies observed by Belle and LHCb as the most promising candidates for the new degrees of freedom underpinning the BSM dynamics behind these effects.*

*The spinorial version of QED has been chosen only for illustration properties; the same principles
followed by a thorough discussion of the peculiarities encountered in its simplest $U(1) \otimes U(1)$ extension.

### 3.5.1 Renormalization of Abelian gauge theories (QED)

Let us start with the standard QED bare Lagrangian (with the charge of the Dirac spinor $\psi$ normalized to $Q_\psi = 1$)

$$
L_B = \overline{\psi}_B (\partial - m_B) \psi_B - ie_B \overline{\psi}_B A_B \psi_B - \frac{1}{4} F_{B \mu \nu} F_{B}^{\mu \nu},
$$

which is conveniently redefined in terms of the renormalized quantities as $L_B = L + \delta L$

where

$$
L = \overline{\psi}(\partial - m) \psi - ie \overline{\psi} A \psi - \frac{1}{4} F_{\mu \nu} F_{\mu \nu}.
$$

and

$$
\delta L = \overline{\psi}_B (\partial - m_B) \psi_B - ie_B \overline{\psi}_B A_B \psi_B - \frac{1}{4} F_{B \mu \nu} F_{B}^{\mu \nu} - \overline{\psi}(\partial - m) \psi - ie \overline{\psi} A \psi + \frac{1}{4} F_{\mu \nu} F_{\mu \nu}.
$$

Defining the renormalized fields in the standard manner, i.e., $\psi_B = Z_{\psi}^{1/2} \psi$ and $A_B = Z_A^{1/2} A$, the fit to the “traditional form” of the counterterm Lagrangian

$$
\delta L = \delta Z_{\psi} \overline{\psi} \partial \psi - \delta m \overline{\psi} \psi - i \delta e \overline{\psi} A \psi - \delta Z_A \frac{1}{4} F_{\mu \nu} F_{\mu \nu}
$$

yields

$$
\delta Z_{\psi} = Z_{\psi} - 1 \quad \text{and} \quad \delta Z_A = Z_A - 1.,
$$

(equation 3.50)

together with

$$
m_B = Z_{\psi}^{-1} (m + \delta m) \equiv Z_{\psi}^{-1} Z_{m} m, \quad (3.51)
$$

$$
e_B = Z_{\psi}^{-1} Z_A^{-1/2} (e + \delta e) \equiv Z_{\psi}^{-1} Z_A^{-1/2} Z_e e, \quad (3.52)
$$

provided

$$
m + \delta m \equiv Z_{m} m \quad \text{and} \quad e + \delta e \equiv Z_e e. \quad (3.53)
$$

Hence, the fundamental Lagrangian in terms of the renormalized quantities reads

$$
L_B = \overline{\psi}(\partial - m) \psi - ie \overline{\psi} A \psi - \frac{1}{4} F_{\mu \nu} F_{\mu \nu} + \delta Z_{\psi} \overline{\psi} \partial \psi - \delta Z_m m \overline{\psi} \psi - i \delta Z_e e \overline{\psi} A \psi - \delta Z_A \frac{1}{4} F_{\mu \nu} F_{\mu \nu},
$$

with $\delta Z_e = Z_e - 1$ and $\delta Z_m = Z_m - 1$; this, subsequently, leads to the standard Feynman rules one can find in the textbooks. In the QED, all the factors above are numbers apply also in the scalar version of the theory.
and all UV divergences in all orders of the perturbative expansion can be absorbed into redefinitions of $\delta Z_\psi, \delta Z_A, \delta Z_m$ and $\delta Z_e$.

Formally, the QED beta-function is then obtained by taking a derivative with respect to the renormalization scale in Eq. (3.52) and, similarly, the running mass is governed by Eq. (3.51). The relevant Ward-Takahashi identity $Z_\psi = Z_e$ which holds to all orders in perturbation theory then ensures that, for the sake of the beta function calculation, it is sufficient to compute $\delta Z_A$ in any given renormalization scheme

$$e_B = Z_A^{-1/2} e .$$

In order to be able to interpret the renormalization scale $\mu$ as an energy at which a certain process is considered one should take the (logarithmic) derivative of $\delta Z_A$ in the class of momentum schemes; this, up to two loops, is however identical to taking the log-$\mu$-derivative of its (much simpler) MS or $\overline{\text{MS}}$ form with respect to the relevant UV-divergence structure ($\varepsilon^{-1}$ in dimensional regularisation with $d = 4 - 2\varepsilon$).

### 3.5.2 Renormalization of QED squared

In QED squared, i.e., in the $U(1) \otimes U(1)$ gauge theory one should assume that there are at least 2 different fields $\psi_i, i = 1, 2$ (otherwise there is not point in talking about more than a single gauge factor) with charges $Q_{ij}$ under the $j$-th $U(1)$ with an associated coupling $e_j$. Assuming that $\psi_1$ and $\psi_2$ carry different $Q$-charges (and, thus, they are not identical and can not mix) the renormalization factors $\delta Z_\psi$ and $\delta Z_m$ defined above should be just doubled, i.e., $\delta Z_\psi \rightarrow \delta Z_{\psi_1}$, $\delta Z_m \rightarrow \delta Z_{m_1}$. Analogously, each group factor receives its own gauge coupling $e_j$ and, hence, $\delta Z_e \rightarrow \delta Z_{e_j}$.

**The naïve definition of the $Z_A$ counterterm in QED squared**

Similarly, there are 2 gauge fields $A_j$ so, naïvely, one is tempted to do the same, namely $\delta Z_A \rightarrow \delta Z_{A_j}$ hoping that it would suffice. *But it does not*. The reason is that there are in general three different UV-divergent diagrams corresponding to the one-loop gauge propagator corrections (a.k.a. vacuum polarisation), namely, the diagonal $A_1 - A_1$ and $A_2 - A_2$ ones whose divergences can be absorbed in $\delta Z_{A_{1,2}}$, but also a third one, $A_1 - A_2 = A_2 - A_1$ for which there is no counterterm left unless one introduces the off-diagonal $F_{1\mu\nu} F_{2\mu\nu}$ piece into the Lagrangian.\(^{19}\) The presence of such a term should be, in fact, even expected; unlike for non-Abelian field strength tensors which, by definition, carry group indices, the Abelian-case $F_{\mu\nu}$ structure is not only gauge covariant but

\(^{19}\)Strictly speaking, this is true only if at least one matter field is charged under both groups, otherwise the theory trivially decays into two non-communicating sectors.
even invariant and, thus, a Lorentz contraction like \( F_{1\mu\nu}F_{2}^{\mu\nu} \) qualifies as a Lagrangian density contribution.

The “correct” definition of counterterms in QED squared

Thus, from scratch, the gauge-kinetic part of the QED squared Lagrangian calls for a matrix structure:

\[
\mathcal{L}_B = \bar{\psi}_B (\hat{\theta} - m_B) \psi_B - i Q e_B \bar{\psi}_B A B \psi_B - \frac{1}{4} F_{B\mu\nu} \xi_B F_B^{\mu\nu}.
\]  

As usual, \( \mathcal{L}_B \) can be decomposed into \( \mathcal{L}_B = \mathcal{L} + \delta \mathcal{L} \) where

\[
\mathcal{L} = \bar{\psi}(\hat{\theta} - m)\psi - i\bar{\psi}QeA\psi - \frac{1}{4} F_{\mu\nu}\xi F^{\mu\nu},
\]

\[
\delta \mathcal{L} = \bar{\psi}(Z\psi - 1)\hat{\theta}\psi - \bar{\psi}(Z^{1/2}m_B Z^{1/2}\psi - m)\psi - i\bar{\psi}(Z^{1/2}Q e_B Z^{1/2}_A Z^{1/2}_B - Qe)A\psi

- \frac{1}{4} F_{\mu\nu}(Z^{1/2}_A \xi B Z^{1/2}_A - \xi) F^{\mu\nu},
\]

with “matrix redefinitions” of the fields

\[
\psi_B = Z^{1/2}_\psi \psi \quad \text{and} \quad A_B = Z^{1/2}_A A.
\]

Needless to say, everything is in principle a vector or a matrix now, and \( \xi \) in particular. Since, however, the matter fields do not mix, \( Z_\psi, m_B \) and \( m \) can be taken diagonal and one can clump these factors and define \( \delta Z_\psi \) and \( \delta Z_m \) as in Eqs. (3.50) and (3.53) (with all relevant quantities replaced by diagonal matrices) and get

\[
\delta \mathcal{L} = \bar{\psi} \delta Z_\psi \hat{\theta} \psi - \bar{\psi} \delta Z_m m \psi - i\bar{\psi}(Z_\psi Q e_B Z^{1/2}_A - Qe)A\psi - \frac{1}{4} F_{\mu\nu}(Z^{1/2}_A \xi B Z^{1/2}_A - \xi) F^{\mu\nu}.
\]

For the time being, we shall retain a generic \( \xi \), i.e., the kinetic terms of the gauge fields shall not be canonically normalized. To proceed, one should define

\[
Q \delta e \equiv Z_\psi Q e_B Z^{1/2}_A - Qe \quad \Leftrightarrow \quad e + \delta e = Q^{-1} Z_\psi Q e_B Z^{1/2}_A,
\]

in full analogy with (3.53) and

\[
\delta \xi \equiv Z^{1/2}_A \xi B Z^{1/2}_A - \xi.
\]

At this point, it may not be a-priori clear how to define the multiplicative counterterm for the gauge coupling: indeed, \( e \) has got two qualitatively different indices and one can in principle multiply from any side. However, it is much more convenient to do it from the left, i.e.,

\[
e + \delta e \equiv Z e,
\]
because then one reveals
\[ e_B = Q^{-1}Z^{-1}_\psi QZeZ_A^{-1/2} \]  
(3.63)

and, thanks to the Ward identity \[ Q^{-1}Z^{-1}_\psi QZe = 1 \], one obtains a matrix version of the equation (3.55) in the form
\[ e_B = eZ_A^{-1/2} \]  
(3.64)

However, a multiplicative renormalization does not make much sense for the gauge kinetic term of Eq. (3.61), so for that one we rather stick to the additive convention. In any case, the counterterm to the gauge kinetic term reads simply
\[ -\frac{1}{4}F_{\mu\nu}\delta\xi F^{\mu\nu} \]  
(3.65)

and the full counterterm Lagrangian receives the final form
\[ \delta\mathcal{L} = \bar{\psi}\delta Z_{\psi}\psi - \bar{\psi}\delta Z_m\gamma\psi - i\bar{\psi}Q\delta Z_eA\psi - \frac{1}{4}F_{\mu\nu}\delta\xi F^{\mu\nu}. \]  
(3.66)

3.5.3 The link between \( Z_A \) and \( \delta\xi \) in different renormalization schemes

This is all right in principle but not very practical yet – on one side one can relatively easily determine \( \delta\xi \) from the structure of the UV divergences of the gauge field propagators but, on the other hand, its correspondence to the central quantity of our interest, namely, \( Z_A \) in Eq. (3.64), is non-linear, cf. (3.61). It is, however, relatively easy to trade \( Z_A \) for \( \delta\xi \) and, thus, connect the running of \( e \) to a quantity at hand. To this end, consider the \( e_B\xi_B^{-1}e_B^T \) which, due to (3.61), receives a simple form
\[ e_B\xi_B^{-1}e_B^T = eZ_A^{-1/2}\xi_B^{-1}Z_A^{-1/2}e^T = e(\xi + \delta\xi)^{-1}e^T, \]  
(3.67)

or, even more conveniently,
\[ (e_B\xi_B^{-1}e_B^T)^{-1} = (e^T)^{-1}(\xi + \delta\xi)e^{-1}, \]  
(3.68)

which holds to all orders in the perturbative expansion. Note that since the LHS of (3.68) is independent of \( \mu \) so must be the RHS; this, in turn, correlates (in a relatively simple way) the evolution of \( e \) and \( \xi \) and the form of \( \delta\xi \).

Besides that, Eq. (3.68) provides the key for the understanding of different approaches to the renormalization of such theories adopted in the literature.

---

\(^{20}\)In what follows, we shall adopt a convention in which all real vectors are treated as column matrices; thus, their dot product can always be written as \( x^Ty \) while \( xy^T \) corresponds to the outer (tensor) product.
Retaining $e$ diagonal and keeping a non-trivial $\xi$ at play

First, there is an option to work with $e$ entirely diagonal but then $Z_A$ must be diagonal too in order to retain this property throughout its running; c.f. Eqs. (3.64). However, in such a case Eq. (3.61) can be fulfilled (for a matrix-like $\delta \xi$) if and only if a non-trivial $\xi$ is kept in the play (note that for a fixed $\xi_B$ and a diagonal $Z_A$ the structure $Z_A^{1/2} \xi_B Z_A^{1/2}$ therein clearly does not complement any initial and hypothetically constant matrix $\xi$ to fulfil (3.61) for a specific $\delta \xi$ on its LHS for all $\mu$).

Equivalently, without dynamically changing $\xi$ one can not retain the RHS of Eq. (3.68) constant for diagonal $e$ - note that there are 3 independent evolving combinations of running parameters in $\delta \xi$ while only 2 parameters are ready in $e$ to compensate for their $\mu$ dependence!

In other words, one can na"ively take the renormalization-scale derivative of Eq. (3.68) and require that it vanishes; an attempt to solve such a linear system with respect to individual derivatives of $e_1$ and $e_2$ fails unless there is also an extra derivative of a component from $\xi$ at one’s disposal. Note that, in the current approach, this is also one way to derive the RGE for the $\xi$ parameter in practice, cf. [140].

Getting rid of $\xi$ at the expense of a matrix-like $e$ and a matrix-like $Z_A$

Alternatively, one can arrange things in such a way to live without $\xi$ altogether. This amounts to redefining first the bare gauge fields in Eq. (3.56) in order to absorb $\xi_B$ therein, namely,

$$A_B \rightarrow \tilde{A}_B = \sqrt{\xi_B} A_B,$$  \hspace{1cm} (3.69)

where the square root of the symmetric $\xi_B$ matrix is defined in the standard manner. This changes the bare Lagrangian (3.56) into

$$\mathcal{L}_B = \overline{\psi}_B (\partial - m_B) \psi_B - i Q \overline{\psi}_B \tilde{e}_B \tilde{A}_B \psi_B - \frac{1}{4} \tilde{F}_{B\mu\nu} \tilde{F}^{\mu\nu}_B$$  \hspace{1cm} (3.70)

where the new set of gauge couplings

$$\tilde{e}_B = e_B \xi_B^{-1/2}$$  \hspace{1cm} (3.71)

constitutes a non-diagonal matrix.

Performing the same redefinitions at the level of the renormalized Lagrangian like before one obtains formulae identical to Eqs. (3.57) and (3.58) with the following replacements:

$$e \rightarrow \tilde{e}, \quad A \rightarrow \tilde{A}, \quad F \rightarrow \tilde{F}, \quad Z_A \rightarrow Z_{\tilde{A}}, \quad \text{and, in particular,} \quad \xi \rightarrow 1, \quad \xi_B \rightarrow 1.$$  \hspace{1cm} (3.72)
The rest of the analysis, i.e., Eqs. (3.59)-(3.68), follows along the same lines as above. Renaming (for optical reasons) $\delta \xi$ to $\delta Z$ in the analogue of the former Eq. (3.61) one ends up with

$$\delta \mathcal{L} = \overline{\psi} \delta \mathcal{Z} \phi \psi - \overline{\psi} \delta \mathcal{Z} m \psi - \overline{\psi} Q \delta \mathcal{Z} \tilde{e} \tilde{A} \psi - \frac{1}{4} \tilde{F}_{\mu \nu} \delta \mathcal{Z} \tilde{F}^{\mu \nu}$$

(3.73)

where $\delta \mathcal{Z}$ is actually identical with the former $\delta Z$ because it is a multiplicative renormalization factor; in other words, the transformation

$$e \rightarrow \tilde{e} = e^{\xi/2}$$

(3.74)

acts homogeneously in formula (3.62) and thus the relevant multiplicative factors ($Z$ in the lagrangian before reabsorption and $\tilde{Z}$ after that) are the same in both cases. This also ensures the validity of the Ward identity necessary to bring the analogue of formula (3.63) into the simple form

$$\tilde{e} B = \tilde{e} Z A^{-1/2}.$$  

(3.75)

This scheme looks even better suited for practical purposes because the renormalization parameter $Z A$ entering the equation above is directly connected to the gauge kinetic counterterm and, thus, easily accessible without any need to go for a non-linear formula like Eq. (3.68).

To recapitulate, we managed to get rid of the non-canonical form of the gauge kinetic term at the expense of a non-diagonal matrix of couplings $\tilde{e}$ (rather than just their pair as in $e$) and a matrix-like counterterm $\delta Z A$.

### 3.5.4 Beta-functions in schemes with matrix gauge couplings

The simplicity of the latter scheme can be readily appreciated in a sample calculation of the gauge-couplings’ evolution in QED-squared. The relevant analogue of formula (3.68), given (3.72) and (3.75), reads here

$$(\tilde{e} B \tilde{e} T)^{-1} = (\tilde{e} T)^{-1} Z A^{-1} \tilde{e} T = (\tilde{e} T)^{-1} (1 + \delta Z A) \tilde{e}^{-1} = (\tilde{e} e T)^{-1} + (\tilde{e} T)^{-1} \delta Z A \tilde{e}^{-1}.$$  

(3.76)

Taking the logarithmic $\mu$-derivative of both sides above with respect to the renormalization scale $\mu$ in momentum schemes (or, equivalently, with respect to the UV-pole structure in MS or $\overline{\text{MS}}$) one obtains

$$\mu \frac{d}{d \mu} (\tilde{e} e T)^{-1} = -(\tilde{e} T)^{-1} \left( \frac{d}{d \log \mu} \delta Z A \right) \tilde{e}^{-1} + \ldots$$

(3.77)
where the ellipsis stands for higher order terms. The structure of $\delta Z_\tilde{A}$ is simple to obtain from the diagrams of the type

which yield

$$\delta Z_{\text{MS}}^{\tilde{A}} = \frac{1}{8\pi^2} \tilde{e}^T \gamma \tilde{e} \frac{1}{\varepsilon} \quad \text{with} \quad \gamma \equiv \frac{2}{3} QQ^T, \quad (3.78)$$

and, thus,

$$\frac{d}{d\log \mu} \delta Z_{\text{MS}}^{\tilde{A}} = \frac{d}{d\varepsilon^{-1}} \delta Z_{\text{MS}}^{\tilde{A}} = \frac{1}{8\pi^2} \tilde{e}^T \gamma \tilde{e}. \quad (3.79)$$

Combining all this together one finally receives

$$\mu \frac{d}{d\mu} (\tilde{e} \tilde{e})^{-1} = -\frac{1}{8\pi^2} \gamma + \ldots. \quad (3.80)$$

In what follows, we shall stick to the standard notation used in the literature \cite{141,142} (i.e., replace $\tilde{e}$ by $G$) and work with $A \equiv GG^T/4\pi$ and $t = \frac{1}{2\pi} \log \mu/\mu_0$. In these coordinates, the evolution equation \ref{eq:3.79} can be recast in a particularly simple form

$$\frac{d}{dt} A^{-1} = -\gamma + \ldots \quad (3.81)$$

which clearly resembles the situation encountered in the standard spinorial QED, see also \ref{eq:2.52}.

**Alternative derivation of the same result**

Note that the same result can be obtained in a less fancy way by inspecting the $\mu$-dependence of the three-body gauge amplitudes of the form\footnote{Note that the Ward identities connecting the vertex corrections to the matter propagator counter-term ensuring the flavour-blindness of the charge renormalization are at work here in the same manner as in the ordinary QED so, as before, all that’s needed are just the gauge propagator corrections.}

\begin{align}
\frac{d}{dt} A^{-1} &= -\gamma + \ldots \quad (3.81)
\end{align}

where, for the sake of simplicity, we have been focusing on the renormalization of the $g_{AB}$ coupling, i.e., the external gauge-field leg corresponds to $A_B$ and the matter current

\begin{align}
A_B &\quad A_C &\quad Q_A
\end{align}
is assumed to carry solely the $U(1)_A$ charge. The corresponding evolution equation for $G$ defined as

$$ G \equiv \begin{pmatrix} g_{AA} & g_{AB} \\ g_{BA} & g_{BB} \end{pmatrix} $$

(3.83)

can be written in a simple matrix form as

$$ \mu \frac{d}{d\mu} G = \frac{1}{(4\pi)^2} G(G^T \gamma G). $$

(3.84)

Note that the configuration of various “building blocks” on the RHS above reflects the transformation properties of $G$, $Q$ and $A$ (with $Q^T = (Q_A, Q_B)$ and $A^T = (A_A, A_B)$) necessary for preserving the form of the covariant derivative (for $\psi$) $D \ni Q^T G A$, namely,

$$ G \rightarrow O_1 G O_2^T, $$

(3.85)

$$ Q \rightarrow O_1 Q, $$

(3.86)

$$ A \rightarrow O_2 A, $$

(3.87)

where $O_{1,2}$ are independent real orthogonal matrices. Given this, the matrix shape of the RHS of Eq. (3.84) is practically enforced up to an overall numerical factor which, however, can be obtained trivially by matching to the known structure of the single-$U(1)$ case. Note that this is exactly the method used in the study [141] complementing the seminal results of Martin and Vaughn [143], see Sect. 5.4.

Finally, taking into account the symmetry properties of $\gamma$, Eq. (3.84) implies

$$ \mu \frac{d}{d\mu} (GG^T) = \frac{1}{8\pi^2} (GG^T) \gamma (GG^T), $$

(3.88)

or, equivalently

$$ (GG^T)^{-1} \left[ \mu \frac{d}{d\mu} (GG^T) \right] (GG^T)^{-1} = \frac{1}{8\pi^2} \gamma. $$

(3.89)

which, using the general identity for regular matrix functions $A^{-1} (\frac{d}{d\tau} A) A^{-1} = -\frac{d}{d\tau} A^{-1}$ and the definitions above yields again the desired result (3.81).

### 3.5.5 One-loop matching in schemes with $U(1)$ mixing

So far, we have been discussing just the shapes of the renormalization group equations in different renormalization schemes traditionally adopted in theories with multiple $U(1)$ gauge factors. However, in practical calculations, these would be useless without the corresponding initial conditions obtained by matching the structure of the high and low-energy Lagrangians and, in particular, the coupling within.

In general, such a procedure closely resembles the recipe one follows even in the Standard Model case of matching the higher-energy $SU(2)_L \otimes U(1)_Y$ gauge structure to that of the effective low-energy $U(1)_Q$ one:
1. First, one should write down the high-scale (+) and low scale (−) covariant derivatives in terms of the relevant high- and low-scale theory charges \((Q_+ \text{ and } Q_-\), respectively) and the corresponding gauge fields \((A_+ \text{ and } A_-)\), i.e. \(D_+ \ni Q^T G_+ A_+ \) and \(D_- \ni Q^T G_- A_- \) (with all Lorentz factors suppressed for simplicity).

2. Second, one should express the \(A_+\) fields in terms of \(A_-\) and a suitable orthogonal matrix: \(A_+ = O A_-\).

3. Subsequently, the identification of \(D_-\) and \(D_+\) makes it possible to get all \(Q_-\)'s as linear combinations of \(Q_+\)'s (by looking at the \(A_-\)'s there) with coefficients corresponding to the entries of the specific “light” columns of the \(O\) matrix.

4. Next, one should combine this information with the model definitions of the \(Q_-\) charges in terms of the \(Q_+\) ones (typically given as \(Q_- = P Q_+\) where \(P\) is a suitable rectangular matrix, often with more columns than rows; in what follows we shall denote its individual rows by \(p^T\) where \(p\)'s will be column vectors of coordinates of the individual operators of \(Q_-\) in the basis of operators in \(Q_+\)) and equate the coefficients of the individual \(Q_+\) factors. This is justified by the fact that the action of the covariant derivatives must match on all fields in the model and there should be enough such fields in order to distinguish among all the charges.

5. Hence, one gets a set of equations for the “light columns” \(O_l\) of the total \(O\) matrix (i.e., those corresponding to the still massless gauge bosons) in the form \(G_+ O_l = P^T G_-\).

6. The last step is to solve for the \(O_l\) matrix \(O_l = G_+^{-1} P^T G_-\) and use its orthogonality \(O_l^T O_l = 1\) (which, however, applies only from one side!) to get

\[
1 = (G_-)^T P (G_+^{-1})^T G_-^T P^T G_- = (G_-)^T P (G_+ G_+^T)^{-1} P^T G_-.
\] (3.90)

The desired matching condition between \(G_-\) and \(G_+\) then follows readily:

\[
(G_- G_-^T)^{-1} = P (G_+ G_+^T)^{-1} P^T.
\] (3.91)

As an example, consider the simplest case of such a setting corresponding to the \(U(1)_Q \times U(1)_Q\) gauge theory parametrised by a 2×2 matrix of gauge couplings (3.83) which gets broken down to a \(U(1)_q\) scheme with the corresponding gauge coupling \(g\).

Suppose that the low-energy charge \(q\) is expressed in terms of the high-energy ones \(Q_A, Q_B\) as

\[
q = p_A Q_A + p_B Q_B,
\] (3.92)
then \( p^T = (p_A, p_B) \) so that \( q = pTQ \) with \( Q = (Q_A, Q_B)^T \). Then the matching formula (3.91) yields
\[
g^{-2} = p^T(GG^T)^{-1}p. \tag{3.93}
\]

**Matching in the “canonical” \( U(1)_R \otimes U(1)_{B-L} \rightarrow U(1)_Y \) case**

Let us apply the prescription above to the canonical example of the LR-symmetry breaking (cf. Sect. 3.1.1) passing through a \( U(1)_R \otimes U(1)_{B-L} \) stage which is eventually broken to the SM hypercharge \( U(1)_Y \). With \( Y = \sqrt{3/5}R + \sqrt{2/5}X \) (note that \( Y \) here stands for the GUT-compatible hypercharge defined by (2.71) and \( X \) denotes the properly normalized \( B - L \) charge connected to the “empirical” one via the \( \sqrt{3/5} \) factor, see Sect. 3.1.2) one has
\[
p^T = (\sqrt{3/5}, \sqrt{2/5}), \tag{3.94}
\]
and the procedure above yields
\[
g_Y^{-2} = \frac{3}{5}(g_{XX}^2 + g_{XR}^2) + \frac{2}{5}(g_{RR}^2 + g_{RX}^2) - \frac{2}{5}\sqrt{6}(g_{RR}g_{XR} + g_{RX}g_{XX})\left(\frac{g_{RR}g_{XX} - g_{RX}g_{XR}}{(g_{RR}g_{XX} - g_{RX}g_{XR})^2}\right). \tag{3.95}
\]
Note that forgetting about off-diagonals, it trivially reduces to the naïve (and wrong) relation \( \alpha_Y^{-1} = \frac{3}{5}\alpha_R^{-1} + \frac{2}{5}\alpha_X^{-1} \) which, however, often appears in the literature. Nevertheless, in most cases of physical interest the initial condition is such that the off-diagonal couplings are absent by definition (because \( U(1)_R \) typically descends from a non-abelian factor) and, thus, the error due to their omission is relatively small with only negligible effects to the final results (unless the double-\( U(1) \) gauge running stage is “long”\(^{22} \)).

### 3.6 Towards the minimal renormalizable theory of perturbative baryon number violation

The renormalizable \( SO(10) \) GUT studied in the preceding sections is arguably the minimal potentially realistic and calculable grand unified model and, as such, it is certainly worth a thorough scrutiny along the lines sketched above. However, it is obviously not a simple setting to deal with - the need for a detailed account for even small details of its quantum structure makes it utterly complex at the technical level, cf. [132]; hence, the number of its phenomenological analysis is very limited. The undertaking is further complicated by the presence of the humongous (anti-selfdual part of the) 5-index antisymmetric tensor whose main purpose (besides rank reduction)

\(^{22}\)An extreme example of the size of the error committed by omitting the off-diagonal factors from the running can be found in the supplementary material of Sect. 5.4.
is to provide a renormalizable and tree-level Majorana mass for the RH neutrinos in the matter 16M.

Assume for the moment that the U(1)B−L subgroup of SO(10) is rather broken by the minimal scalar multiplet with this capacity, namely, the scalar 16S. It is well known that, in such a case, one can still obtain a contribution to the RH neutrino masses from a d = 5 contractions of the type

$$\mathcal{L}^{(5)} \ni \frac{\lambda_{ij}}{\Lambda} 16^i_M 16^j_M 16^2_S,$$

where i, j are generation indices, λ is a 3×3 symmetric matrix of dimensionless couplings and Λ denotes an effective cut-off scale. In the asymmetric phase this operator yields $M_R \sim \lambda \langle 16_S \rangle^2 / \Lambda$ which, assuming $\Lambda \sim M_{Pl}$, gives $M_R$ in the desired $10^{12}$ − $10^{13}$ GeV ballpark for $\langle 16_S \rangle \sim 10^{16}$ GeV. At first glance, this is exactly what one needs in order to obtain the correct light neutrino masses via type-I seesaw; however, a closer inspection reveals at least two serious drawbacks of this scenario:

1. **One gets no information about the flavour structure of $M_R$.** Let us recall that this was one of the great benefits of having 126S rather than 16S which, indeed, made the minimal SO(10) GUTs of Sections 3.3 and 3.4.4 (potentially) testable and, thus, so interesting.

2. **There are issues with the GUT-scale unification of the gauge couplings.** With $\langle 16_S \rangle \sim 10^{16}$ GeV one typically ends up with a situation in which the SO(10) symmetry is broken completely in the vicinity of $10^{16}$ GeV and, hence, the theory looks entirely like the SM below $M_G$. Note, however, that this does not need to be a problem if the model is supersymmetric down to the TeV scale!

However, in the early days of GUTs, none of these two points was taken very seriously because nothing was known about neutrino masses in point 1., and there were no good electroweak-scale data to appreciate the severity of the tension in point 2.

**3.6.1 Witten’s loop in the SO(10) GUTs**

To this end, a very nice trick was pulled out by Witten [144] in 1980 which made it possible to generate the RH neutrino masses in the SO(10) GUT context with 16S instead of 126S without the need to resort to the non-renormalizable mechanism above. The key consists in giving up the tree-level origin of $M_R$. Indeed, the two-loop diagram in Fig. 3.2 provides the necessary contraction of the fermionic matter bilinear

\[23\text{Note that } \Lambda \sim M_{Pl} \text{ is the most natural assumption here as } M_{Pl} \text{ is the only scale of BSM physics which is almost universally accepted to exist.}\]
Figure 3.2: A sample two-loop Feynman diagram contributing to the Majorana masses for the RH neutrinos (residing in $16_M$) in SO(10) GUTs. Note that the graph corresponds to just one of many two-loop contributions in the broken-phase perturbative expansion with VEVs kept in the interaction part of the Hamiltonian.

$16_M \otimes 16_M$ to a pair of $16_S$ VEVs (arranged in a way to resemble an effective 5-index antisymmetric tensor structure) through a very specific contraction of the internal lines of the relevant Feynman graph. The consistency of this picture can be verified readily:

- There are, indeed, two units of $B-L$ carried into the vacuum by the VEVs of $16_S$ in Fig. 3.2.
- The algebraic structure of the loop propagators ($10_S \otimes 45_G \otimes 45_G$) attached to the fermionic line does admit a fully antisymmetric contraction and, thus, can mimic the 5-index antisymmetric tensor structure of $126$.
- The matrix of Yukawa coupling of $10_S$ is fully symmetric in the flavour space and so is also the $M_R$ generated through the graph in Fig. 3.2.

This looks like a perfect alternative to both the tree-level $M_R$ generation entertained in the models with $126_S$ and the non-renormalizable mechanism one would naively have to resort to in scenarios with $16_S$ instead.

**Witten’s mechanism SO(10) model building issues**

On the practical side, however, the beautiful Witten’s mechanism of the preceding section has never been implemented in a fully compelling unified setting. There are two main reasons having to do with point 2. in the second paragraph of Sect. 3.6.2, namely, the difficulty to obtain a potentially viable gauge unification pattern:

- In the traditional TeV-scale SUSY scenarios in which $\langle 16_S \rangle \sim 10^{16}$ GeV is not a problem for gauge unification the scale of $M_R$ turns out to be strongly suppressed
with respect to the natural expectation of \( M_R \sim \lambda(16_S)^2/16\pi^2 M_G \sim M_G/16\pi^2 \) by the \( m_{\text{soft}}/M_G \) extra factor due to the SUSY non-renormalization theorems \(^{145, 146}\).

- In non-SUSY theories the gauge running with \( \langle 16_S \rangle \sim 10^{16} \) GeV cries for an intermediate scale. However, since the residual symmetry left intact by \( \langle 16_S \rangle \) is \( SU(5) \) there is not much of a room for such a scale to emerge from the subsequent gauge symmetry breaking and, hence, the only potentially viable scenarios are those featuring extra fine-tunning(s). In the extreme case, one can consider the split-SUSY \(^{147}\) variant of the Witten’s scenario \(^{148, 149}\) in which the squarks and sleptons are maximally diverted from the TeV scale where the gauginos and higgsinos ensure the MSSM-like gauge unification.

### 3.6.2 Witten’s loop in the flipped \( SU(5) \) unification

In what was written above there is a clear indication that a remedy may eventually come from relaxing the stringent grand unification constraints imposed (among other sectors) on the gauge couplings of the model. This, however, calls for a “non-grand” unified scenarios, i.e., those based on non-simple symmetry groups. At the same time, it would be very welcome if the predictive power of such models was not entirely ruined concerning the effects of our main interest, i.e., perturbative baryon and lepton number violation.

Remarkably, there is indeed a very interesting gauge model just half way between the fragility of the \( SO(10) \) GUTs and baryon number triviality of its simple descendants such as Pati-Salam or LR-symmetric models discussed in Sects. 3.1.2 and 3.1.1. It is based on the maximal \( SU(5) \otimes U(1) \) subgroup of \( SO(10) \) and, as such, it stands out of the usual mantra of the need for the LR-symmetrisation of the gauge symmetry in presence of the RH neutrinos. It is not similar to the \( SU(5) \) GUTs of Sect. 2.6.1 either as the SM hypercharge is not fully contained in the \( SU(5) \) gauge factor.

**Flipped \( SU(5) \) unification overview**

Let’s just recapitulate the salient features of the traditional approach to the “marriage” between the \( SU(5) \) and \( U(1) \) gauge symmetries (the latter to be from now on called \( U(1)_Z \)): Assigning a unit of the \( Z \) charge to the \( 10 \) of \( SU(5) \) the following set of charges is anomaly free:

\[
(10, +1) \oplus (\overline{5}, -3) \oplus (1, +5) .
\]

Hence, 16 matter fields of each of the SM generations + 3 RH neutrinos can be accommodated in such a (reducible) representation of \( SU(5) \otimes U(1) \). However, the trivial
matter embedding along the lines of the minimal \( SU(5) \) à la Georgi and Glashow discussed in Sect. 2.6.1, namely, \( L \) and \( d^c \) in \((\mathbf{5}, -3)\), \( Q, u^c \) and \( e^c \) in \((\mathbf{10}, +1)\) and \( N^c \) in \((\mathbf{1}, +5)\) with the hypercharge generator identified with (a properly normalized) \( T_{24} \) of \( SU(5) \), is not what we are after here.

Interestingly, there is a second option \([150, 151]\) corresponding to swapping \( u^c \) with \( d^c \) and \( \nu^c \) with \( e^c \) that, however, requires a non-trivial (“flipped”) embedding of \( Y \) into the full \( SU(5) \otimes U(1)_Z \), namely

\[
Y = \frac{1}{5} (Z - T_{24}). \tag{3.98}
\]

This also means that one can use \( \mathbf{10}_S = (\mathbf{10}, +1) \) to break the symmetry down to the SM instead of the larger adjoint \( \mathbf{24} \) of the standard \( SU(5) \) and, thus, save a number of scalar degrees of freedom \([25]\).

The “flipped” hypercharge embedding also induces important changes in the baryon number violation phenomenology. Indeed, the leptoquark degrees of freedom in the gauge sector spanning over \((\mathbf{24}, 0) \oplus (\mathbf{1}, 0)\) transform as \((\mathbf{3}, \mathbf{2}, +\frac{1}{6}) \oplus (\overline{\mathbf{3}}, \mathbf{2}, -\frac{1}{6})\) rather than \((\mathbf{3}, \mathbf{2}, -\frac{5}{6}) \oplus (\overline{\mathbf{3}}, \mathbf{2}, +\frac{5}{6})\) of the Georgi-Glashow model and, hence, different pattern of \( d = 6 \) BLNV operators is generated at the SM effective theory level.

The devil, as usual, is in detail. Unlike in the “standard” \( SU(5) \) where the RH neutrino mass underpinning the type-I seesaw is trivially introduced as a direct mass term in the relevant \( \mathbf{1}_M \mathbf{1}_M \) matter singlet bilinear, one needs a large 50-dimensional two-index symmetric tensor transforming as \((\mathbf{50}, -2)\) to do the same in the current scenario, thus enlarging the so far limited number of degrees of freedom by about a factor of 4. Hence, in spite of its overall simplicity, the flipped \( SU(5) \) framework does not seem to provide more insight into the flavour aspects of lepton number violation than its “standard” \( SU(5) \) counterpart.

The minimal renormalizable flipped \( SU(5) \) unification

The important observation made\([26]\) in \([154]\) was that the drawbacks of the original Witten’s loop mechanism in the \( SO(10) \) context and the baroqueness of the minimal

\[ \text{Note that the simplified symmetry breaking mechanism based on } \mathbf{10}_S = (\mathbf{10}, +1) \text{ is in the core of the “missing partner” doublet-triplet splitting mechanism that brought the flipped } SU(5) \text{ scenario a lot of attention in 1980’s.} \]

\[ \text{The associated simplification of the scalar sector necessary for the proper symmetry breaking down to the SM is one of the typical benefits of the “flipped” scenarios; to this end, one can quote the situation in the flipped SUSY } SO(10) \text{ scenarios (see, e.g., [152]) in which the notoriously cumbersome } \mathbf{210} \oplus \mathbf{126} \oplus \overline{\mathbf{126}} \text{ minimal Higgs sector (cf. Sect. 3.3.1) or its } \mathbf{45} \oplus \mathbf{54} \oplus \mathbf{126} \oplus \overline{\mathbf{126}} \text{ variant can be replaced with just } \mathbf{45} \oplus 2 \times (\mathbf{16} \oplus \overline{\mathbf{16}}). \]

\[ \text{For the sake of completeness let us note that a similar scheme was considered in the string theory context already in 1991 in [153].} \]
potentially realistic renormalizable flipped $SU(5)$ model sketched above can cure each other. Indeed, the Feynman diagram in Fig. 3.3 can be viewed as that of Witten in Fig. 3.2 minimally adopted to the flipped $SU(5)$ context. The main point is that the VEV of the (Hermitian conjugate of) $(10, +1)$ sticking out of the graph can be as large as $M_G$ with no tension in the gauge running because the $g_Z$ coupling associated to the $U(1)_Z$ factor is not required to unify with that of the $SU(5)$ part where only $g_3$ and $g_2$ of the SM are fully contained. Note also that, algebraically, the effect of 50 is mimicked by the tensor product of $27 \otimes 24 \otimes 24$ though perhaps not as clearly as it was for the 126 within $10 \otimes 45 \otimes 45$ of SO(10).

3.6.3 The minimal renormalizable theory of perturbative $B$ violation

Hence, one arrives at a very attractive scenario which shares the nice features of both worlds, namely, the BLNV sector’s predictive power of the minimal SO(10) GUTs and the relative simplicity of the $SU(5)$ settings, with the extra benefit of providing a potentially realistic framework for the implementation of the beautiful Witten’s idea.

In the minimal version the high-scale spectrum and the symmetry breaking pattern of the model is encoded in the scalar potential

$$V = \frac{1}{2} m_{T_0}^2 \text{Tr}(10_S^\dagger 10_S) + m_{S_5}^2 5_S^\dagger 5_S + \frac{1}{8} (\mu \varepsilon_{ijklm} 10_S^ij 10_S^kl 5_S^m + \text{h.c.})$$

$$+ \frac{1}{4} \lambda_1 [\text{Tr}(10_S^\dagger 10_S)]^2 + \frac{1}{4} \lambda_2 \text{Tr}(10_S^\dagger 10_S 10_S^\dagger 10_S) + \lambda_3 (5_S^\dagger 5_S)^2$$

$$+ \frac{1}{2} \lambda_4 \text{Tr}(10_S^\dagger 10_S)(5_S^\dagger 5_S) + \lambda_5 5_S^\dagger 10_S 10_S^\dagger 5_S,$$

Note that the quantum numbers of the 5-dimentional scalar generating the Yukawa contraction of the type $10_M 10_M 5_S$ are trivially $(5, -2)$.
where $\lambda_{1,\ldots,5}$ are real numerical couplings and $\mu$ provides the necessary mixing between $10_S$ and $5_S$, cf. Fig. 3.3. In the broken phase (triggered by a unification-scale VEV of the SM singlet component of $10_S$ accompanied by the electroweak VEV of $5_S$) the scalar spectrum of the theory comprises of 16 real Goldstone modes (corresponding to $25 - 9 = 16$ massive vector bosons of $SU(5) \otimes U(1)$ broken down to the low-scale $SU(3)_c \otimes U(1)_Q$), a light Higgs boson, a heavy SM singlet $S$ from $10_S$ and a pair of heavy SM colour-triplet leptoquarks $\Delta_{1,2}$ admixed (via $\mu$) from the two $(3, 1, -\frac{1}{3})$ components in $5_S \oplus 10_S$. Besides that, the physical heavy spectrum contains a vector leptoquark $X^\mu$ with the SM quantum numbers $(3, 2, +\frac{1}{6}) + h.c.$ and a heavy SM singlet. Note also that the unified symmetry breaking occurs in one-step, i.e., the $SU(5) \otimes U(1)$ is broken directly into the $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ of the Standard Model and one can expect the characteristic scale of the heavy degrees of freedom to be in the $10^{16}$ GeV ballpark. For further details, an interested reader is kindly deferred to the original work [154].

**Phenomenology of the flipped SU(5) unification with Witten’s loop**

In what follows, the central quantity of interest will be namely the Yukawa Lagrangian which, in the minimal case, reads

$$\mathcal{L} \ni Y_{10} 10_M^{\dagger} 10_{M} + Y_{5} 10_{M}^{\dagger} 5_{S}^{\dagger} (5_{S}^{-2})^{\ast} + Y_{1} 5_{M}^{\dagger} 1_{M}^{\dagger} 5_{S}^{\dagger} + h.c.,$$

(3.100)

where $Y_{10}$, $Y_{5}$ and $Y_{1}$ are (complex) $3 \times 3$ matrices of Yukawa couplings ($Y_{10}$ is symmetric). It yields the following expressions and correlations for the (high-scale) effective matter mass/Yukawa matrices:

$$M^d = (M^d)^T \propto Y_{10},$$
$$M^u = (M^u)^T \propto Y_{5},$$
$$M^e \propto Y_{1},$$

(3.101)

$$M^\nu_M = 0.$$

Note that, as expected, no Majorana mass matrix for the RH neutrinos is generated at the tree level. Moreover, it is the down-quark mass matrix that turns out to be symmetric rather than $M_u$ in the standard $SU(5)$ case, cf. (2.77). Instead, $M_u$ is tightly correlated to the Dirac neutrino mass matrix; this is actually the seed of the model’s significant predictive power in the flavour sector.

As anticipated, the Majorana RH neutrino mass is generated at the quantum level by the loops of the kind depicted in Fig. 3.3. A short inspection reveals that its flavour structure must be driven by $Y_{10}$ as it is the only Yukawa therein and its overall scale
should be governed by two powers of $\langle 10_S \rangle$. Moreover, the relevant expression should vanish for $\mu \to 0$; hence, on purely dimensional grounds, one expects

$$M_M^\nu \propto \frac{1}{16\pi^2} Y_{10} g^4 \mu \frac{\langle 10_S \rangle^2}{M_X^2}, \quad (3.102)$$

where $M_X$ is the scale of the heavy gauge boson mass.

With this at hand, the type-I seesaw formula combined with the non-tachyonicity conditions for the scalar spectrum (for details see Sect. 5.5) yields

$$D_u U^\dagger_\nu (m^\nu_\text{diag.})^{-1} U^*_\nu D_u \leq \frac{\alpha_U}{64\pi^4} \sqrt{2\lambda_5} |Y_{10}| V_G F, \quad (3.103)$$

where $D_u$ is the diagonal form of the up-quark mass matrix, $U_\nu$ is the unitary transformation diagonalising the light neutrino masses $m^\nu$ (i.e., $m^\nu = U^T \nu m^\nu_\text{diag.} U_\nu$), $\alpha_U$ is the $SU(3) \otimes SU(2)$-unification-scale value of the associated “generalised fine structure constant” of the model and $F$ is an $O(1)$ factor calculable from the relevant Feynman graphs, see next section.

Note that for a fixed light neutrino spectrum (conveniently parametrised by the mass of the lightest neutrino $m_1$) and with the assumption of perturbativity of the $\lambda_{2,5}$ and $Y_{10}$ couplings\footnote{Perturbativity constraints are generally very tricky. For the sake of simplicity, all that we shall assume here is that none of the dimensionless couplings involved exceeds $4\pi$.} in formula (3.103) one obtains a strong constraint on the possible shapes of $U_\nu$. For instance, for small enough $m_1$ large 1–3 angle in $U_\nu$ is clearly forbidden as it would propagate the big 11 element of the $(m^\nu_\text{diag.})^{-1}$ matrix to the 33-element of $U^\dagger_\nu (m^\nu_\text{diag.})^{-1} U_\nu$ which would, subsequently, pick up the pair of the top quark masses in (3.103) and, hence, violate the desired inequality there by orders of magnitude. At the same time, heavier light neutrino spectrum will be preferred as it would also have the tendency to alleviate this issue. Both these observation are extremely welcome as they provide non-trivial constraints on both the lepton and baryon number violating phenomena in this setting.

**Lepton number violation phenomenology**

As for $L$ violation, some first (unification-scale dependent) lower limits on the absolute scale of the light neutrino Majorana masses were derived in [155], with a clear preference of relatively heavy light neutrino spectra. Besides that, scans through the parameter space (including CP phases) for the spots which do support thermal leptogenesis (see Sect. 2.1.3) impose further limits for the absolute light neutrino scale. This is work in progress, to appear soon.
Baryon number violation phenomenology

Concerning perturbative $B$ violation, $U_{\nu}$ is a structure governing practically all the leading order gauge-leptoquark contributions to the two-body proton decay amplitudes in the current scenario. This is due to the flavour structure of the $d = 6$ BLNV operator $\bar{F}Q\bar{w}L$ (a Fierz-transform of $O_1$ in Table 2.1) which, barring its strongly suppressed companion involving heavy RH neutrinos, is the only relevant piece of $\mathcal{L}_{\text{int}}$ here

$$\mathcal{L}_{\text{int}} \supset (U_{\bar{d}})_{a_1}(U_{e}^\dagger)_{b_1}(\bar{w}_{(1)}^{r\mu})_{(a)}\gamma_{\mu}e_{(b)} + (U_{d}U_{\nu}^\dagger)_{b_0}(U_{e}^\dagger)_{c_1}(\bar{d}_{(1)})_{(a)}\gamma_{\mu}d_{(b)}\gamma_{\mu}\nu_{(c)}. \tag{3.104}$$

In (3.104) the bracketed indices label generations, $U_f$ and $U_{f^c}$ for $f = u, d, e$ are defined as $M_f = U_f^T D_f U_{f^c}$ and all is written in the basis in which the top-quark mass matrix is diagonal. From here it is clear that the amplitudes for the $p$-decay neutral-meson final states factorize $(U_{e}^\dagger)_{b}$ (where $b$ stands for the flavour of the final state lepton) and $U_{d^c} = U_d$ following from (3.101) is nothing but the CKM matrix here. Thus, the ratios of the same-lepton partial decay widths are fully calculable in the minimal flipped SU(5) scenario! Moreover, for $U_{d^c} = U_d$ the flavour structure in the second term above, i.e., the piece governing the decays into charged mesons+neutrinos, depends only on $U_{\nu}$. Hence, the minimal flipped SU(5) à la Witten represents a framework in which the gauge contributions to the two-body proton decay amplitudes are very strongly correlated in all channels. As such, the model may be viewed as a genuine theory of perturbative baryon number violation.

3.6.4 Calculating Witten’s loop in the flipped SU(5) context

Needless to say, a decisive numerical analysis of this attractive scenario\(^{29}\) requires a detailed calculation of the $F$-factor in formula (3.103). This was recently performed in the study \cite{155} (enclosed as a Supplementary material in Sect. 5.5). Perhaps the most interesting aspect of the calculation worth elaborating on here is the IR- and UV-divergence structure of the relevant Feynman graphs which can be conveniently written in the massive perturbation theory employing the unitary gauge, see Fig. 3.4.

There are few observations one can make right away:

- The sum of the two two-loop topologies depicted in Fig. 3.4 should be UV-finite.

This is clear as there is no tree-level counterterm in the model that can tame any UV divergence.

\(^{29}\)It is worth noting that in order to make the model potentially realistic a second copy of the scalar $\mathbf{5}_S$ has to be added. The main reason is that, in the minimal setting, $Y_{10}$ in (3.103) is overly constrained by Eq. (3.101) and a realistic hierarchy of light neutrino masses can not be attained, cf. \cite{154}. Remarkably enough, even with the second copy of $\mathbf{5}_S$ all the features underpinning the predictivity of the model are preserved: $M^d$ remains symmetric, $M^u$ is still equal to $M^d_D$ and $M^e$ is as unconstrained as before.
Figure 3.4: The complete set of unitary-gauge Feynman graphs generating the RH neutrino masses in the minimal flipped $SU(5)$ setting in the “massive” perturbative expansion (i.e., with the VEVs absorbed into the massive propagators of all the relevant fields). Note that this approach minimises the total number of graphs for the price of their higher complexity.

- There should be no UV sub-divergences either that would require insertion of the first-order counterterms anywhere. The point is that none of the graphs with the trilinear local counterterms (replacing any of the loops involved) exists because the first non-trivial contraction corresponding to the given configuration of the external lines emerges only at two loops.

- One should be free to set the light matter fermion masses in Fig. 3.4 to zero as these should play no role in such loops and all the IR divergences potentially emerging in such a limit should eventually disappear.

This is, indeed, what eventually happens when the calculations are performed in detail. For each of the two $\Delta$-scalars the UV-divergent structure of the diagram on the left in Fig. 3.4 reads (in $d = 4 - 2\varepsilon$)

$$\Sigma_{1UV}(0) = -\frac{1}{(4\pi)^4} \left[ \frac{3}{2\varepsilon} - \frac{m_{\Delta}^4}{2m_X^4} \left( \frac{1}{2\varepsilon^2} + \frac{3}{2\varepsilon} - \frac{1}{\varepsilon} \log \frac{m_{\Delta}^2}{M^2} \right) \right], \quad (3.105)$$

(with $M$ denoting the regularisation scale) while that of the graph on the right turns out to be

$$\Sigma_{2UV}(0) = \frac{1}{(4\pi)^4} \left[ \frac{3}{4\varepsilon} + \frac{m_{\Delta}^4}{4m_X^4} \left( \frac{1}{2\varepsilon^2} + \frac{3}{2\varepsilon} - \frac{1}{\varepsilon} \log \frac{m_{\Delta}^2}{M^2} \right) \right]. \quad (3.106)$$

Since, however, these contributions sum with a relative factor of 2 in the total integral

$$I_{\Delta} \equiv -(4\pi)^4 [\Sigma_1(0) + 2\Sigma_2(0)] \quad (3.107)$$

the overall cancellation of the UV divergences follows for each cluster of graphs. As one can expect, the finite part of $I_{\Delta}$ which should be only a function of $s \equiv \frac{m_{\Delta}^2}{m_X^2}$.
behaves regularly\textsuperscript{30} even in the asymptotic limits with $s$ large or small, namely

\[
I_\Delta \sim -3 + O(s^{-1} \log^2 s) \quad \text{for } s \to \infty ,
\]

\[
I_\Delta \sim 3 + s \left(3 \log s + \pi^2 - \frac{15}{2}\right) + O(s^2 \log^2 s) \quad \text{for } s \to 0 .
\]

In conclusion, the absolute value of the $F$ factor in Eq. (3.103) which is just a weighted sum of $I_\Delta$'s for the two $\Delta_{1,2}$ scalars, is bounded from above. This, in turn, provides a clear rationale for the phenomenological expectations spelled out in the previous section.

\textsuperscript{30}Interestingly, $I_\Delta$ is even a monotonic function of $s$, see Sect. 5.5
Chapter 4

Conclusions and outlook

After almost 50 years since the discovery of an internally consistent framework for
description of the elementary constituents of matter and their interactions – the Stan-
dard Model – the Higgs boson, the last missing piece of the underlying structure, has
eventually been discovered. As beautiful and soothing the picture became there is still
a number of indications that this has not been the final word the elementary particle
physics had to say about the world around. Indeed, in the years to come our confidence
is likely to be challenged by:

- Clear experimental signals of physics that the SM can not account for; among
  the most prominent of these are the neutrino flavour oscillation effects.

- Theoretical issues related to the structural aspects of the SM such as flavour and
  other things; for instance, we have no clue on what is behind the peculiar matter
  generation pattern, why there is so little CP violation in the strong interactions
  etc. Remarkably enough, we do not know even such a basic thing as whether the
  very SM vacuum is stable or not.

Even more trouble is likely to be encountered if the Standard Model is to be eventually
married with gravity:

- On the observational side, the SM provides no clue for the nature and origin of the
  peculiar matter-like but invisible gravitating component of the Universe’s energy
  density budget (dark matter), let alone its negative-pressure counterpart (often
  called dark energy) which, together, overwhelm the part accounted for by the
  SM by almost a factor of 20. We do not understand the number of baryons left
  behind the initial annihilation inferno if the initial conditions were B-symmetric;
  if the were not, we do not know why.
- As far as the theory is concerned, there is no realistic (even potentially) and calculable framework that can consistently account for the intrinsically quantum nature of the SM together with the general covariance of the Einstein gravity. We have no clue about what is behind the vast disparity between the electroweak and Planck scales; actually, we do not even know whether there is anything to be understood there at the first place.

The discovery of neutrino masses with all its basic implications as described in this thesis paves a particularly interesting (yet pretty conservative) avenue for a further systematic exploration of the Terra Incognita slowly emerging on the horizons of namely the intensity frontier experimental activities. In this struggle, the effects of baryon and lepton number violation are likely to play a prominent role either as the very objectives of the research or at least as irreducible elements of the undertaking.

From this perspective, the candidate’s research activities on the frontier of perhaps the most natural theoretical approach to the conundrum – the BSM theories with extended gauge symmetries – represent a valuable and relevant contribution to the current understanding of especially the quantum structure of unified gauge models. Among the most important of these one should perhaps mention the series of works in which the paradigmatic minimal SUSY GUT model has been decisively refuted, the original idea that the tachyonic vacuum instabilities of the minimal $SO(10)$ GUT may be lifted by quantum effects or the full logical completion of the two-loop renormalization-group programme set out by Martin and Vaughn [4] for softly broken supersymmetric theories has been achieved.

It is always the time that eventually decides on whether anything of this would be relevant for the future of the beautiful subject at stakes. At the moment, one may only hope that the quest for the unified description of matter and its interactions, underlying all efforts elaborated on in this thesis, will be one day rewarded by a clear signal of baryon and/or lepton number violation. If we were ready and lucky this may even be the case within the upcoming generation of experimental facilities.
Chapter 5

Supplementary material

Here we shall present and comment upon six selected candidate’s publications reflecting the extent of his contribution to the subject of Grand unified theories and their quantum structure.

5.1 The novel $SO(10)$ seesaw mechanism


This article is enclosed as the candidate’s most cited work (having earned about 280 citations within the inspirehep.net database as of January 2020). It elaborates on the unexpected possibility to disentangle the seesaw scale from the scale of the $SU(2)_R \otimes U(1)_{B-L}$ breaking in the SUSY GUT context exploiting the interesting features of the inverse seesaw scheme put forward in [156, 157]. The key to a viable and internally consistent setting is the minimality of the $SO(10)$ symmetry breaking pattern passing through an intermediate $U(1)_R \otimes U(1)_{B-L}$ symmetry stage (see Section 3.5 of this thesis for details) which is, subsequently, broken by a single scalar field that does not contain any SM-charged components. Thus, the $B-L$ breaking scale is not only decoupled from the $SU(2)_R$ one but it becomes essentially free from the gauge unification perspective. At the same time, the $B-L$ breaking scale drops from the seesaw formula (at the leading order) due to an interplay between the size of the induced $SU(2)_L \otimes U(1)_Y$-breaking VEV of an additional $SU(2)_L$ doublet in the scalar 16 of $SO(10)$ and the shape of the relevant linear-seesaw contribution to the light neutrino masses.

Besides having written the major part of the manuscript the candidate contributed to the study by two central ingredients, namely, by noticing the presence of the induced VEV of the $SU(2)_L$ doublet in the scalar 16 enforced by the SUSY $F$-flatness conditions and by providing the entire tedium of the renormalization group analysis.
Supersymmetric SO(10) Seesaw Mechanism with Low B-L Scale

M. Malinsky*
Scuola Superiore di Studi Avanzati, Via Beirut 4, I-34014, and INFN, Sezione di Trieste, Italy

J.C. Romão†
Departamento de Física and CFTP, Instituto Superior Técnico, Avenue Rovisco Pais 1, 1049-001 Lisboa, Portugal

J. W. F. Valle‡
AHEP Group, Instituto de Física Corpuscular—CSIC/Universitat de València Edificios Institutos de Paterna, Apt 22085, E-46071 Valencia, Spain

(Received 7 July 2005; published 14 October 2005)

We propose a new seesaw mechanism for neutrino masses within a class of supersymmetric SO(10) models with broken D parity. It is shown that in such scenarios the B-L scale can be as low as TeV without generating inconsistencies with gauge coupling unification nor with the required magnitude of the light neutrino masses. This leads to a possibly light new neutral gauge boson as well as relatively light quasi-Dirac heavy leptons. These particles could be at the TeV scale and mediate lepton flavor and CP violating processes at appreciable levels.

DOI: 10.1103/PhysRevLett.95.161801 PACS numbers: 12.60.Jv, 12.10.Dm, 11.30.Fs, 98.80.Cq

M. Malinsky*
Scuola Superiore di Studi Avanzati, Via Beirut 4, I-34014, and INFN, Sezione di Trieste, Italy

J.C. Romão†
Departamento de Física and CFTP, Instituto Superior Técnico, Avenue Rovisco Pais 1, 1049-001 Lisboa, Portugal

J. W. F. Valle‡
AHEP Group, Instituto de Física Corpuscular—CSIC/Universitat de València Edificios Institutos de Paterna, Apt 22085, E-46071 Valencia, Spain

(Received 7 July 2005; published 14 October 2005)

We propose a new seesaw mechanism for neutrino masses within a class of supersymmetric SO(10) models with broken D parity. It is shown that in such scenarios the B-L scale can be as low as TeV without generating inconsistencies with gauge coupling unification nor with the required magnitude of the light neutrino masses. This leads to a possibly light new neutral gauge boson as well as relatively light quasi-Dirac heavy leptons. These particles could be at the TeV scale and mediate lepton flavor and CP violating processes at appreciable levels.

DOI: 10.1103/PhysRevLett.95.161801 PACS numbers: 12.60.Jv, 12.10.Dm, 11.30.Fs, 98.80.Cq

The origin of neutrino masses is the most well-kept secret of modern elementary particle physics. The basic dimension-five operator which leads to neutrino masses [1] can arise from physics at vastly different scales. One popular alternative is the seesaw mechanism, in which case the small neutrino masses are induced by the exchange of superheavy neutral fermions [2–5] or super-heavy scalars, or both [6–8]. The light neutrino masses are given as

\[ M_\nu \approx -v^2 YM^{-1} Y^T. \]  

(1)

An alternative inverse seesaw scheme has been suggested [9] for theories which lack the representation required to implement the canonical seesaw, as happens in a class of string inspired models.

In addition to the normal neutrinos \( \nu \), such an inverse seesaw mechanism employs two sequential SU(3) \( \otimes \) SU(2) \( \otimes \) U(1) singlets, \( v^c \) and \( S \) (these are all left-handed two-component spinors [6]). The effective neutrino mass matrix has the following form:

\[
M_\nu = \begin{pmatrix}
0 & Yv & 0 \\
Y^T v & 0 & M \\
0 & M & \mu
\end{pmatrix} 
\]  

(2)

in the basis \( \nu_L, v^c_L, S_L \). Here \( Y \) is the Yukawa matrix parametrizing the \( Y^{ij} v^{cj} C^{-1} \) term + H.c. interactions, while \( M \) and \( \mu \) are SU(3) \( \otimes \) SU(2) \( \otimes \) U(1) invariant mass entries. When \( \mu \to 0 \), a global lepton-number symmetry is exactly conserved and all three light neutrinos are strictly massless. Yet it has been shown [10,11] that lepton flavor and CP can be violated at appreciable levels in the absence of supersymmetry, provided the scale \( M \) is sufficiently low. When \( \mu \neq 0 \), the mass matrix for the light eigenstates is given by

\[ M_\nu \approx -v^2 (YM^{-1}) \mu (M^{-1} Y^T). \]

(3)

One sees that neutrinos can be made very light, as required by oscillation data [12], even if \( M \) is very low, far below the unified scale \( MG(M \ll M_G) \), provided \( \mu \) is very small, \( \mu \ll M \). This scheme has a very rich and interesting phenomenology, since no new scales need to be added to generate the small neutrino masses, instead a small parameter \( \mu \) is added. Note that in such a SU(3) \( \otimes \) SU(2) \( \otimes \) U(1) inverse seesaw the smallness of \( \mu \) is natural, in the ’t Hooft sense [13], as the symmetry enhances when \( \mu \to 0 \). However, there is no dynamical understanding of this smallness.

Here we provide an alternative inverse seesaw realization consistent with a realistic unified SO(10) model. Such embedding brings several issues: (i) The \( v^c S \) entry \( M \) [generated by the vacuum expectation value (VEV) of a Higgs multiplet \( \chi_R \) with SU(3), SU(2)_L, SU(2)_R \( \otimes \) U(1)\_\( B-L \) quantum numbers (1, 1, 2, −1)] breaks the \( B-L \) symmetry, now gauged. The corresponding scale \( \langle \chi_R \rangle \) must be compatible with gauge coupling unification. Together with the requirement of low-energy supersymmetry to stabilize the hierarchies, this places rather strong constraints on how we must fill the “desert” of particles below \( M_G \). (ii) The need to justify the magnitude of the singlet \( \mu \) \( SS \) mass. (iii) It implies a nonzero \( v S \) entry in Eq. (2), proportional to the VEV of the \( L-R \) partner of \( \chi_R \), namely, \( \chi_L \equiv (1, 2, 1, +1) \). Let us now discuss one by one these three points and show that there indeed exists a supersymmetric SO(10) model that addresses all these conditions in a satisfactory way and offers a new way to understand the smallness of neutrino masses.

First, note that there are several mechanisms that could be used to get rid of the \( SS \) term in Eq. (2). For example, we

\[ \text{(3)} \]
can treat the SO(10) embedding into E6 where the fermionic singlet could be a member of a 27-dimensional irreducible representation with the familiar SO(10) \( \otimes \) U(1)\(_x\) decomposition,

\[
27_F = 1^4_F \otimes 16_F \otimes 10_F^2.
\]

If at the \( E_6 \) scale there is no 35' Higgs representation, the U(1)-charge of the 1\(_F\)1\(_F\) matter bilinear is so large that it is very hard to saturate it. Thus, as long as the corresponding U(1) is unbroken we have \( \mu = 0 \). Even if we break the U(1) symmetry at some lower scale, it could be rather complicated to generate an effective SS entry, which brings further suppression, even at the level of effective operators. From now on we will neglect \( \mu \).

Now consider the \( \nu S \) term. A typical SO(10) superpotential contains the following terms:

\[
W \ni M_{16}^H 16_H \bar{T}_6 + \rho 16_H 16_H 10_H + \text{H.c.}
\]

The fact that in "standard" supersymmetric SO(10) models there is a small induced VEV generated for the neutral component of the first entry of SO(10) (that break the SM) the requirement comes from this graph, up to transposition. The neutrino mass is suppressed by the unification scale, not by the B-L breaking scale, which can be low.

Concerning the stability of the texture zeros at the \( \nu^c \nu^c \)  and SS entries in formula (6) it can be protected as long as the U(1)\(_x\) and the U(1)\(_y\) of E\(_6\) \( \otimes \) SO(10) \( \otimes \) U(1)\(_x\) in E\(_6\) inspired setups are exact. Indeed, U(1)\(_x\) must be broken at the \( v_R \) scale \( [16_H \otimes \bar{T}_6] \) always has a U(1)\(_x\) charge. However, the charge of \( \nu_R \) is such that the relevant operators arise only at higher orders and may be neglected.

Now we turn to gauge coupling unification. As was shown by Deshpande et al. [16] the scale at which the SU(2)\(_R\) \( \otimes \) U(1)\(_{B-L}\) symmetry is broken to U(1)\(_V\) can be arbitrarily low if we populate properly the desert from \( M_Z \) to \( M_G \). In their case, this is achieved by putting three copies of \( (1, 1, 2, +1) \otimes (1, 1, 2, -1) \) coming from \( 16_H \otimes \bar{T}_6 \) right at the \( v_R \) scale. Then the (one-loop) minimal supersymmetric standard model (MSSM) running of the \( \alpha_y^{-1} \) can be "effectively" extended above the \( v_R \) scale (\( \alpha_y^{-1} = \frac{3}{2} \alpha_{B-L}^{-1} + \frac{3}{2} \alpha_{B-L}^{-1} \)) by a conspiracy between the running of \( \alpha_{B-L}^{-1} \) and \( \alpha_y^{-1} \). However, such a scheme is rather \( ad \ hoc \) as we need to push three identical copies of a Higgs multiplet to a very low scale, at odds with the "minimal fine-tuning."

Here we present a more compelling scheme in which the SU(2)\(_R\) breaking scale \( v_R \) is separated from the low U(1)\(_{B-L}\) \( \otimes \) U(1)\(_y\) breaking scale \( v_R \), \( v_R \gg v_R \) in the chain SU(2)\(_R\) \( \otimes \) U(1)\(_{B-L}\) \( \rightarrow \) U(1)\(_R\) \( \otimes \) U(1)\(_{B-L}\) \( \rightarrow \) U(1)\(_y\). At each step we assume just those multiplets needed to break the relevant symmetry. The first step is achieved by a light admixture of the (1, 1, 3, 0) multiplets living in 45 and 210, while the second stage is driven by the light component of the (1, 1, \( +\frac{1}{2} \), -1) \( \otimes \) (1, 1, \( -\frac{1}{2} \), +1) scalars [in SU(3) \( \otimes \) SU(2) \( \otimes \) U(1)\(_R\) \( \otimes \) U(1)\(_{B-L}\) notation] of 16\(_H\) \( \otimes \bar{T}_6 \). Note that to allow for such a L-R asymmetric setup the D parity of SO(10) must be broken.

Let us further specify the ingredients of our supersymmetric SO(10) model needed to implement the mechanism

\[
F = F_{\nu}^c + \sum_{i} F_{\nu}^{i} \chi_{i}^L + \text{H.c.}
\]
described above. As usual, we use three copies of $16^f$ to accommodate the SM fermions and for each of them we add a singlet fermion $1^f$ to play the role of $S_L$. A realistic fermionic spectrum requires more than one copy of $10_H$ Higgs multiplet. Moreover, we assume one (or more) copy of $16_H \oplus T_6H$ to implement our new supersymmetric seesaw mechanism. To prevent fast proton decay via dimension 4 operators, we assign the matter fermions in $16^f$ and $1^f$ with a discrete matter parity that forbids the mixing of $16^f$ and $16_H$. Finally, we add a $45_H$ and $210_H$ to trigger the proper symmetry breaking pattern with no $D$ parity below the unified scale [17–19]. The SO(10) invariant Yukawa superpotential then reads

$$W_Y = Y_{0ij} 16^f_i 16^f_j 10^f_H + F_{ijk} 16^f_i T_6^k H_T.$$  \hspace{1cm} (8)

We do not impose other discrete symmetries to reduce the number of parameters that might, however, be welcome in connection with the doublet-triplet splitting problem in a more detailed analysis. The Higgs superpotential is

$$W_H = M_{10}^f 16^f_i T_6^j H_T^i + M_{210}^f 16^f_i 10^f_H + M_{45}^f 16^f_i 45^f_H$$
$$+ M_{210}^f 210^f_H 210^f_H + 3 \rho_{klm}^{16} 16^f_i 16^f_j 10^f_H$$
$$+ \rho_{klm}^{45} 16^f_i T_6^j H_T^i 45^f_H$$
$$+ \lambda H^2 16^f_i T_6^j H_T^i 210^f_H$$
$$+ \lambda^2 H^3 210^f_H + \varepsilon 210^f_H.$$  \hspace{1cm} (9)

The components of $210^f_H$ and $45^f_H$ that receive unified-scale VEVs and trigger the breaking of SO(10) to SU(3)$_c$ $\otimes$ SU(2)$_L$ $\otimes$ SU(2)$_R$ $\otimes$ U(1)$_{B-L}$ are (in Pati–Salam language) $210^f_H \equiv (15, 1, 1) \oplus (1, 1, 1)$ and $45^f_H \equiv (15, 1, 1)$. As shown in [17–19], this pattern can accommodate the desired $D$-parity breaking allowing for an intermediate $L$-$R$ symmetric group with an asymmetric particle content, leading to distinct $g_L$ and $g_R$ below $M_G$. The subsequent SU(2)$_R$ $\rightarrow$ U(1)$_R$ breaking at $v_R$ is induced by the VEV of a light superposition of $(1, 1, 3, 0)_{210}$ and $(1, 1, 3, 0)_{45}$ that can mix below the unified scale. Next, the U(1)$_R$ $\otimes$ U(1)$_{B-L}$ is broken down at $v_R$ by the VEVs of the light component of type $(1, 1, +\frac{1}{2}, -1) \oplus (1, 1, -\frac{1}{2}, +1)$ coming from $(1, 1, 2, -1) \oplus (1, 1, 2, 1)$ of $16^f_H \oplus T_6^j H_T^i$. The final SM breaking step is as usual provided by the VEVs of the $(1, 2, 2, 0)$ bidoublet components. Note that unlike the example given in [16] there is no artificial redundancy in the number of light states living at intermediate scales.

Let us finally inspect the one-loop gauge coupling unification. Using the normalization convention $2\pi f(\mu) = \ln(\mu/M_Z)$ we have (for $M_A < M_H$)

$$\alpha^{-1}(M_A) = \alpha^{-1}(M_H) + b(t_B - t_A)$$

in the ranges $[M_3, M_\lambda^k]$, $[M_3, v_R]$, and $[V_R, M_{GUT}]$, $M_\lambda^k$ is the supersymmetry breaking scale taken at $\sim 1$ TeV. Between $v_R$ and $V_R$ the two U(1) factors mix and the running of $\alpha^{-1}_L$ and $\alpha^{-1}_{B-L}$ requires separate treatment. The Cartans obey the traditional formula (with “physically” normalized $B$-$L$ and $Y_W$)

$$Y_W = 2T^R_L + (B - L).$$

Note that the SO(10) normalization of $b_{B-L}$ is $b_{B-L} = \frac{3}{8} b_{B-L}$. Once the $D$ parity is broken below $M_G$, we have $g_L \neq g_R$. The Higgs sector in the stage down to $v_R$ is as follows: $1 \times (1, 1, 3, 0)$, $1 \times (1, 1, 2, +1) \oplus (1, 1, 2, -1)$, and $1 \times (1, 2, 2, 0)$. This gives rise to the $b$ coefficients $b_3 = -3$, $b_L = 1$, $b_R = 4$, and $b_{B-L} = 20$.

At the subsequent stage from $v_R$ to $v_R$ we keep only the weak scale bidoublet $(1, 2, 2, 0)$ that below $v_R$ splits into a pair of $L$ doublets with the quantum numbers $(1, 2, +\frac{1}{2}, 0) \oplus (1, 2, -\frac{1}{2}, 0)$ under the SU(3)$_c$ $\otimes$ SU(2)$_L$ $\otimes$ U(1)$_R$ $\otimes$ U(1)$_{B-L}$ group and a part of $(1, 1, 2, +1) \oplus (1, 1, 2, -1)$ that is needed to break U(1)$_R$ $\otimes$ U(1)$_{B-L}$ to U(1)$_R$ $\otimes$ a pair of the $\chi_R$ fields $(1, 1, +\frac{1}{2}, -1) \oplus (1, 1, -\frac{1}{2}, +1)$. Since these fields are neutral with respect to all SM charges, the position of the $v_R$ scale does not affect the running of the “effective $\alpha^{-1}_L$” (given by the appropriate matching condition) and the only effects arise from the absence of the right-handed $W$ bosons at this stage. Using the SU(2)$_R$ normalization of the U(1)$_R$ charge the matching condition at $v_R$ is trivial. The relevant $b$ coefficients of SU(3)$_c$ $\otimes$ SU(2)$_L$ and the matrix of anomalous dimensions of the mixed U(1)$_R$ $\otimes$ U(1)$_{B-L}$ couplings are $b_3 = -3$, $b_L = 1$, and

$$\begin{pmatrix}
\gamma_{11} & \gamma_{12} \\
\gamma_{21} & \gamma_{22}
\end{pmatrix} = \begin{pmatrix}
\frac{15}{2} & -1 \\
-1 & 18
\end{pmatrix}. \hspace{1cm} (10)
$$

Below the $v_R$ scale the model is the ordinary MSSM with the $b$ coefficients $b_3 = -3$, $b_L = 1$, and $b_Y = 33/5$, and finally, the $b$ coefficients for the SM stage below the $M_S$ scale are $b_3 = -7$, $b_L = -3$, and $b_Y = 21/5$. The $v_R$-scale matching condition reads

$$\alpha^{-1}_L(v_R) = \frac{3}{2} \alpha^{-1}_R(v_R) + \frac{5}{2} \alpha^{-1}_{(B-L)}(v_R).$$

Recalling that

$$\alpha^{-1}_L(M_Z) = \frac{3}{5} \left(1 - \sin^2\theta_W \right) \alpha^{-1}(M_Z) \hspace{0.5cm} \text{and} \hspace{0.5cm} \alpha^{-1}_R(M_Z) = \sin^2\theta_W \alpha^{-1}(M_Z)$$

the initial condition (for central values

\begin{center}
\begin{tabular}{c|c|c}
$\gamma_{11}$ & $\gamma_{12}$ & $\gamma_{21}$ \\
$\gamma_{22}$ & & \\
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{c|c|c}
$\gamma_{11}$ & $\gamma_{12}$ & $\gamma_{21}$ \\
$\gamma_{22}$ & & \\
\end{tabular}
\end{center}

FIG. 2. The one-loop gauge coupling unification in the model described in the text. The $D$ parity is broken at $M_G$, and the intermediate scales $v_R$ and $v_R$ correspond to SU(2)$_R$ $\rightarrow$ U(1)$_R$ and U(1)$_R$ $\otimes$ U(1)$_{B-L}$ $\rightarrow$ U(1)$_Y$ breaking, respectively.
of the input parameters) is $\alpha_{1}^{-1}(M_{Z}) \approx 59.38$, $\alpha_{2}^{-1}(M_{Z}) \approx 29.93$, and $\alpha_{3}^{-1}(M_{Z}) \approx 8.47$ [20].

Inspecting the results of the numerical analysis (Figs. 2 and 3) one confirms that the $v_{R}$ scale does not affect the predicted value of $\alpha_{1}^{-1}(M_{Z})$ and remains essentially free at the one-loop level. Thus, the unification pattern is fixed entirely by the interplay of $M_{S}$ and $v_{R}$. The lower bound $v_{R} \approx 10^{14}$ GeV is consistent with the standard minimally fine-tuned supersymmetry SO(10) behavior; see for instance [21].

In summary, we have proposed a variant supersymmetric SO(10) seesaw mechanism, which involves a dynamical scale $v_{R}$ that can be rather low, as it is not strongly restricted either by gauge coupling unification or neutrino masses. The smallness of neutrino masses coexists with a light $B-L$ gauge boson, possibly at the TeV scale, that can be produced at the Large Hadron Collider, by the Drell-Yan process. Moreover, the “heavy” neutrinos involved in the seesaw mechanism [see Eq. (6)] get masses at $v_{R}$ and can therefore be sufficiently light as to bring a rich set of testable implications. For example, their exchange can induce flavor violating processes, such as $\mu \rightarrow e \gamma$ with potentially very large rates, similar to the inverse seesaw model of Eq. (3) [10,11]. We conclude that, even in an unified seesaw context, the dynamics underlying neutrino masses may have observable effects at accelerators and in the flavor sector.

This work was supported by Spanish Grant No. BFM2002-00345 and by the EC Human Potential Program RTN network MRTN-CT-2004-503369. We thank Rabi Mohapatra and Goran Senjanović for a useful discussion.

Note Added.—Recently, we came across related papers by Steve Barr and collaborators and by Fukuyama et al. [22]. These indeed have strong elements in common with the early work in Refs. [9,10]. However, the mechanism we now propose differs crucially from all of these in that our $B-L$ scale can be very low, in contrast to theirs. We show how our key new feature not only accounts for the observed neutrino mass scale, but also fits with the gauge unification condition in SO(10).

*Electronic address: malinsky@sissa.it
\textsuperscript{2}Electronic address: jorge.romao@ist.utl.pt
\textsuperscript{3}Electronic address: valle@iftc.uv.es

5.2 The minimal renormalizable SUSY SO(10) no-go


This article is enclosed as the candidate’s first game-changing contribution to the subject of grand unified theories.

As described in Sect. 3.3.1 at the beginning of the first decade of this century the minimal renormalizable SUSY SO(10) model [105–111] was widely recognised as one of the most promising potentially realistic supersymmetric grand unified theories on the market. Besides its unprecedented simplicity owing to the low number of Higgs-type chiral supermultiplets it was known for providing interesting predictions for the flavour structure of its low-energy effective descendants (MSSM in most cases), namely, i) an automatic conservation of $R$-parity, ii) a rationale for the near-maximality of the atmospheric mixing in the lepton sector (attributed to the phenomenological high-energy convergence of the tau-lepton and bottom-quark Yukawa couplings) and iii) a lower limit for the reactor mixing angle (close to its recently measured value).

Originally, the latter two observations were made under a simplifying assumption of the dominance of the type-II seesaw contribution to the light neutrino masses which, however, could not be justified without a particularly tedious calculation of the induced VEVs of the $SU(2)_L$ triplets of $126 \oplus \overline{126}$ (as a by-product of a thorough investigation of the high-energy spectrum of the theory). The situation, luckily, changed with the publication of the studies [111, 158] which provided the missing information. Thus, all of a sudden, it was possible to perform a thorough scrutiny of the minimal model including, in particular, a complete $\chi^2$-analysis of the matter fermion mass, mixing and CP patterns and a comprehensive study of the relevant gauge unification chains (revealing, in particular, the positions of the seesaw scale), all that in connection with the detailed understanding of the heavy part of the model’s spectrum.

With all this at hand, it was decisively concluded (cf. Sect. 3.3.2 of this thesis) that no point in (the perturbative part of) the parameter space exists that would support all the imposed phenomenological constraints. The importance of this result has been widely acknowledged by the HEP community, earning to the article about 140 citations (in the inspirehep.net database as of 1/2020).

Besides having written a significant part of the manuscript the candidate’s contribution to the study consisted of providing all the necessary information related to the shape of the high-energy spectrum and the complete renormalization group running analysis in the cases of main interest.
Fermion masses and mixings in $SO(10)$ models and the neutrino challenge to supersymmetric grand unified theories

Stefano Bertolini* and Thomas Schwetz†

Scuola Internazionale Superiore di Studi Avanzati, via Beirut 4, I-34014 Trieste and INFN, Sezione di Trieste, Italy

Michal Malinsky‡

School of Physics and Astronomy, University of Southampton, SO16 1BJ Southampton, United Kingdom and CERN Theory Division, CH-1211 Geneva 23, Switzerland

(Received 9 May 2006; published 29 June 2006)

We present a detailed study of the quark and lepton mass spectra in a $SO(10)$ framework with one $10_H$ and one $126_H$ Higgs representations in the Yukawa sector. We consider in full generality the interplay between type I and type II seesaw for neutrino masses. We first perform a $\chi^2$ fit of fermion masses independent on the structure of the GUT Higgs potential and determine the regions of the parameter space that are preferred by the fermion mass sum rules. We then apply our study to the case of the minimal renormalizable SUSY $SO(10)$ GUT with one $10_H$, one $126_H$, one $126_L$, and one $210_H$ Higgs representations. By requiring that proton decay bounds are fulfilled we identify a very limited area in the parameter space where all fermion data are consistently reproduced. We find that in all cases gauge coupling unification in the supersymmetric scenario turns out to be severely affected by the presence of lighter than GUT (albeit $B-L$ conserving) states. We then conclusively show that the minimal supersymmetric $SO(10)$ scenario here considered is not consistent with data. The fit of neutrino masses with type I and type II seesaws within a renormalizable $SO(10)$ framework strongly suggests a non-SUSY scenario for gauge unification.

DOI: 10.1103/PhysRevD.73.115012

PACS numbers: 12.10.-g, 12.15.Ff, 12.60.Jv, 14.60.Pq

I. INTRODUCTION

Understanding the pattern of fermion masses and mixings is one of the longstanding challenges in particle physics. In this respect grand unified theories (GUTs) do provide an appealing and powerful tool to address the multiplicity of matter states, by involving stringent relations among the known fermions (and possibly implying a natural enlargement of the minimal sector). Very appealing candidates for a GUT are models based on the $SO(10)$ gauge group [1]. All the known fermions plus right-handed neutrinos fit into three 16-dimensional spinorial representations of $SO(10)$, and hence the model naturally leads to neutrino masses based on the seesaw mechanism [2,3].

In this work we concentrate on a minimal renormalizable version of the supersymmetric $SO(10)$ GUT, where fermion mass matrices are obtained from the Yukawa couplings of the matter fields to 10 and 126 dimensional Higgs representations [4]. The $126_H$ representation contains the scalar multiplets $(10, 3, 1) \oplus (\overline{10}, 1, 3)$ under the Pati-Salam subgroup [5] $SU(4)_C \otimes SU(2)_L \otimes SU(2)_R$, whose vacuum expectation values (VEVs) generate mass matrices for the left- and right-handed neutrino fields, respectively. Hence, in general there are contributions to the effective mass matrix of the light neutrinos from the type I [2] as well as type II [3] seesaw mechanisms. Thanks to its minimality the model leads to constrained relations among quark and lepton mass matrices [6,7]:

\begin{align}
M_d &= v_d^{10}Y_{10} + v_L^{126}Y_{126}, \\
M_u &= v_u^{10}Y_{10} + v_u^{126}Y_{126}, \\
M_\ell &= v_\ell^{10}Y_{10} - 3v_d^{126}Y_{126}, \\
M_D &= v_u^{10}Y_{10} - 3v_u^{126}Y_{126}, \\
M_L &= v_L^{126}, \\
M_R &= v_R^{126}. \\
\end{align}

Here $M_D, M_L, M_R$ denote the Dirac neutrino mass matrix, the mass matrix of the light left-handed neutrinos, and the mass matrix of the heavy right-handed neutrinos, respectively. The indices $d, u$ refer to down- and up-quarks, $\ell$ denotes the charged leptons, while $Y_{10}$ and $Y_{126}$ are two complex symmetric matrices related to the $10_H$ and $126_H$ Yukawa interactions. The various $v$'s denote the VEVs of the relevant Higgs multiplets, with $v_{u,d}^{10,126}$ standing for the $SU(2)_L$ doublet components giving rise to the light MSSM-like Higgses, while $v_R(v_L)$ are the $SU(2)_L$ singlet (triplet) VEVs entering the type I (type II) seesaw formulas for the light neutrino masses. The effective mass matrix for the light neutrinos is then written as

\begin{align}
M_{\nu} &= M_L - M_D M_R^{-1} M_D^T, \\
\end{align}

where the symmetry of $M_D$ is used. In Eq. (2) the first and second terms correspond to type II and type I seesaws, respectively. Hence, all fermion mass matrices are pre-
dicted in terms of two complex symmetric Yukawa matrices and six VEVs.

Quite an amount of activity has been devoted recently to the study of the minimal renormalizable SUSY $SO(10)$ model—see for instance Refs. [8–26]. This interest has been triggered to some extent by the experimental progress in neutrino physics, and the observation of Refs. [10,11], that within the minimal $SO(10)$ model with type II seesaw dominance $b - \tau$ unification allows naturally for a large lepton mixing due to the destructive interference between the 33 entries of the down-quark and charged lepton mass matrices. Analyses of proton decay within this setting are found, e.g., in Refs. [27,28], while various extensions of the model have been considered in view of relaxing some of the tensions with the data, by adding nonrenormalizable terms [17], by minimally extending the Higgs sector [18], or by including Yukawa couplings of the fermion fields with an additional 120-dimensional Higgs representation [23,24,26].

Most of the previously quoted analyses consider the dominance of one type of seesaw and assume that the vacuum gives the correct neutrino mass scale. Very recently two studies appeared showing an intrinsic antagonism between seesaw neutrino masses and coupling unification in the minimal SUSY $SO(10)$ scenario [29,30], pointing out a critical tension between the needed seesaw neutrino scale and the detailed spectrum arising from the minimal $SO(10)$ breaking GUT vacuum [15,19,31].

The aim of the present paper is twofold. First, in Secs. II and III we study in detail the implications of the fermion mass sum rules that emerge from $10_H$ and $126_H$ Yukawa terms in a $SO(10)$ GUT framework. The analysis is performed in full generality, allowing for complex Yukawa couplings and VEVs, and considering the neutrino mass matrix as originating from an admixture of type I and type II seesaw as given by Eq. (2). In the second part of the paper, Sec. IV, we specialize our study to the minimal renormalizable SUSY $SO(10)$ model, by including the constraints coming from the detailed structure of the GUT symmetry breaking vacuum and performing an exhaustive study of the parameter space by optimizing a $\chi^2$ function. This approach is complementary to the random parameter searches applied in previous studies, since the algorithm converges to an optimal solution (if it exists), and isolated solutions (allowed by an acceptable fine-tuning of the parameters) can be found, which are easily missed in a random parameter scan.

We show conclusively that the minimal renormalizable $SO(10)$ supersymmetric scenario does not allow for a consistent fit of the fermion mass spectrum, the model being unable to reproduce the correct neutrino mass scale via type I and/or type II seesaw. Our negative result can be cast in more general (and simple) terms considering that, while (i) the present constraints on proton decay force any GUT scale to lie above $10^{16}$ GeV and (ii) SUSY gauge unification works very well considering a desert scenario between the weak and the GUT scales, the neutrino mass coming from type I and/or type II seesaws, being proportional via Yukawa and Higgs potential couplings to $m_{\text{weak}}^{\text{GUT}}$, requires for $m_{\text{weak}} = 200$ GeV, $M < 10^{15}$ GeV. As a consequence (barring a strong interacting sector in the theory), intermediate mass scales appear which are bound to affect SUSY gauge coupling unification at the least.

All this applies to the minimal renormalizable setup. As it was pointed out very recently [32], nonminimal realizations of the renormalizable Yukawa sector such as those containing an additional 120-dimensional Higgs multiplet may provide a way out of the issues: the neutrino mass scale can be enhanced by the type I seesaw with right-handed neutrino masses below the GUT scale due to tiny Yukawa couplings associated to the $126_H$ multiplet, whose role in the charged matter sector can be taken over by the new $120_H$ multiplet. Although the simplest realization of this programme (where the charged matter sector is dominated by $10_H$ and $120_H$ and the Yukawas of $126_H$ are neglected) seems to fail [33], this attempt does provide a direction for further studies. Alternatively, Planck induced nonrenormalizable operators may also provide the scale suppression needed by the neutrino sector; see for instance Ref. [34] and references therein.

On the other hand, the simplicity and the high level of predictivity that makes the minimal renormalizable $SO(10)$ GUT extremely appealing is lost in both extensions. Before calling the minimal setup a dead end, an extensive analysis of potential loopholes in the above mentioned arguments is due. This aim we pursue with the present paper to the end.

II. DESCRIPTION OF THE ANALYSIS

A. The parameterization

In order to investigate whether Eqs. (1) and (2) allow for fermion masses and mixing in agreement with the data we proceed as follows. It turns out to be convenient to express the $Y_{10}$ and $Y_{126}$ Yukawa matrices in terms of $M_{u}$ and $M_{d}$, and substitute them in the expressions for $M_{u}$, $M_{D}$, and $M_{\nu}$:

$$M_{u} = f_{u}[(3 + r)M_{d} + (1 - r)M_{\ell}], \quad (3)$$

$$M_{D} = f_{u}[3(1 - r)M_{d} + (1 + 3r)M_{\ell}], \quad (4)$$

where

$$f_{u} = \frac{1}{4} \frac{v_{u}^{10}}{v_{d}^{10}}, \quad r = \frac{v_{d}^{10}}{v_{u}^{10}} \frac{v_{u}^{126}}{v_{d}^{126}}. \quad (5)$$

The neutrino mass matrix is obtained as

$$M_{\nu} = f_{\nu} \left[(M_{d} - M_{\ell}) + \frac{M_{D}}{f_{u}^{2}} (M_{d} - M_{\ell})^{-1} \right] \frac{M_{D}}{f_{u}}, \quad (6)$$

with
\[ f_r = \frac{1}{4} \frac{v_L}{v_d^{125}}, \quad \xi = -\frac{(4f_r v_d^{126})^2}{v_L v_R}. \]  

The parameter \( |\xi| \) controls the relative importance of the type I and type II seesaw terms: For \( |\xi| \to 0 \) one obtains pure type II seesaw, whereas \( |\xi| \to \infty \) (with \( f_r, |\xi| \approx \text{const} \)) corresponds to type I seesaw. For \( |\xi| \approx 1 \) both contributions are comparable.

In what follows we denote diagonal mass matrices by \( \hat{m}_i, x = u, d, \ell, v \), with eigenvalues corresponding to the particle masses, i.e., being real and positive. We choose a basis where the down-quark matrix is diagonal: \( M_d = \hat{m}_d \).

In this basis \( M_d \) is a general complex symmetric matrix, that can be written as \( M_d = W^\dagger \hat{m}_d W \), where \( W \) is a general unitary matrix. Without loss of generality \( f_u \) and \( f_r \) can be taken to be real and positive. Hence, the independent parameters are given by three down-quark masses, three charged lepton masses, three angles, and six phases in \( W, f_u, f_r \), together with two complex parameters \( r \) and \( \xi \): 21 real parameters in total, among which are eight phases.

Using Eqs. (3), (4), and (6) all observables [six quark masses, three Cabibbo-Kobayashi-Maskawa matrix (CKM) angles, one CKM phase, three charged lepton masses, two neutrino mass-squared differences, the mass (CKM) angles, one CKM phase, three charged lepton masses, three angles, and six phases] can be calculated in terms of the lightest neutrino, three Pontecorvo-Maki-Nakagawa-Sakata (PMNS) angles, and three leptonic CP phases: 22 quantities altogether can be calculated in terms of these input parameters. Although the number of parameters is larger than the number of observables presently measured, the system is sensibly over-constrained due to the nonlinear structure of the problem.

Since we work in a basis where the down-quark mass matrix is diagonal, the CKM matrix is given by the unitary matrix diagonalizing the up-quark mass matrix up to diagonal phase matrices:

\[ \hat{m}_u = W_u M_u W_u^\dagger \]  

with

\[ W_u = \text{diag}(e^{i\beta_1}, e^{i\beta_2}, e^{i\beta_3})V_{\text{CKM}} \text{diag}(e^{i\alpha_1}, e^{i\alpha_2}, 1), \]  

where \( \alpha_i, \beta_i \) are unobservable phases at low energy. The neutrino mass matrix given in Eq. (6) is diagonalized by \( \hat{m}_\nu = W'_\nu M'_\nu W'_\nu^\dagger \), and the PMNS matrix is determined by \( W'_\nu W'_\nu^\dagger = \hat{D}_1 V_{\text{PMNS}} \hat{D}_2 \), where \( \hat{D}_1 \) and \( \hat{D}_2 \) are diagonal phase matrices similar to those in Eq. (9).

### B. Input data and \( \chi^2 \) analysis

As input data we use quark and lepton masses and mixing angles evaluated at the GUT scale, based on the renormalization group equation (RGE) analysis of Ref. [35]. As a typical example we consider a SUSY scale \( M_{\text{SUSY}} = 1 \text{ TeV} \), \( \tan \beta = 10 \), and a GUT scale \( M_{\text{GUT}} = 2 \times 10^{16} \text{ GeV} \). Since charged lepton masses are known with an accuracy of better than \( 10^{-3} \), we do not consider them as free parameters but fix them to the (GUT scale) central values from Ref. [35]:

\[ m_e = 0.3585 \text{ MeV}, \quad m_\mu = 75.67 \text{ MeV}, \quad m_\tau = 1292.2 \text{ MeV}. \]  

The remaining data are listed in Table I. For the heavy quark masses \( m_c, m_t, m_t \) we adopt the values and uncertainties used in Ref. [35], whereas for the light quarks we update the data to the ranges given in Ref. [36]: \( m_u = 1.5-4 \text{ MeV}, \quad m_d = 4-8 \text{ MeV}, \quad m_s = 80-130 \text{ MeV} \) (MS masses at 2 GeV). The corresponding values at the GUT scale are given in Table I, together with the CKM angles and phase [24].

Concerning the neutrino parameters, we note that in the setup under consideration the neutrino mass spectrum is generally normal and hierarchical with \( m_1 < m_2 < m_3 \). In this case the RGE running of the PMNS angles [37,38] is of order \( 10^{-5} \) \((1 + \tan^2 \beta) / 10^{-3} \), and therefore negligible to a good approximation. The running of the neutrino masses is small as well [a running effect common to all three neutrino masses can be absorbed in the free parameter \( f_r \), compare Eq. (6)]. In our analysis we use the low energy neutrino parameters, as obtained from recent global fits to neutrino oscillation data [39,40]. We do not include any constraint on the lightest neutrino mass \( m_1 \), since the values that are obtained are much below the sensitivity of the present data.

Let us denote the central values and errors of the observables by \( O_i \) and \( \sigma_i \), for \( i = 1, \ldots, 15 \) runs over all the quantities listed in Table I. As described above, the predictions of these observables, \( P_i \), depend on the parameter values.

### Table I. Sample of GUT scale input data used in this work

<table>
<thead>
<tr>
<th>Observables</th>
<th>Input data</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_d ) [MeV]</td>
<td>1.24 \pm 0.41</td>
</tr>
<tr>
<td>( m_t ) [MeV]</td>
<td>21.7 \pm 5.2</td>
</tr>
<tr>
<td>( m_b ) [GeV]</td>
<td>1.06^{+0.14}_{-0.09}</td>
</tr>
<tr>
<td>( m_u ) [MeV]</td>
<td>0.55 \pm 0.25</td>
</tr>
<tr>
<td>( m_e ) [MeV]</td>
<td>210 \pm 0.21</td>
</tr>
<tr>
<td>( m_\nu ) [GeV]</td>
<td>82.4^{+5.0}_{-14.8}</td>
</tr>
<tr>
<td>( \sin^2 \theta_{13} )</td>
<td>0.0351 \pm 0.0013</td>
</tr>
<tr>
<td>( \sin^2 \theta_{12} )</td>
<td>0.0032 \pm 0.0005</td>
</tr>
<tr>
<td>( \delta_{12} )</td>
<td>0.2243 \pm 0.0016</td>
</tr>
<tr>
<td>( \Delta m^2_{21} ) [eV^2]</td>
<td>( 7.9 \pm 0.3 \times 10^{-5} )</td>
</tr>
<tr>
<td>(</td>
<td>\Delta m^2_{31}</td>
</tr>
</tbody>
</table>
rameters $x_\alpha$, where $\alpha = 1, \ldots, 18$ runs over three down-quark masses, nine real parameters in $W_\ell$, $f_u$, $f_\nu$, $|r|$, $\arg(r)$, $|\xi|$, $\arg(\xi)$. Then a $\chi^2$ function is constructed as

$$\chi^2(x_\alpha) = \sum_{i=1}^{15} \frac{(P(x_\alpha) - O_i)^2}{\sigma_i^2}. \quad (11)$$

The data are fitted by minimizing this function with respect to the parameters $x_\alpha$. The minimization of the scale factors $f_u$ and $f_\nu$ can be done analytically. The remaining 16 dimensional minimization is performed with an algorithm based on the so-called downhill simplex method [41].

Let us note that the minimization is technically rather challenging. The problem involves parameters which differ by many orders of magnitude, and some of the solutions are extremely fine-tuned, which leads to very steep valleys in the $\chi^2$ landscape. Moreover, due to the high dimensionality and the nonlinearity of the problem there is a large number of local minima. In the numerical analysis much effort has been devoted to find the absolute minimum, involving random methods and dedicated investigations to each particular issue under consideration. However, as it is well known, it is difficult to assess by numerical methods with absolute confidence that an absolute minimum has been found. Although each minimum has been carefully tested for possible improvements, one should always keep in mind the possibility that a better solution might exist somewhere in the parameter space.

III. GENERAL FIT OF FERMION MASSES AND MIXINGS

For the discussion of the obtained fits, we classify the solutions by the values of the parameter $|\xi|$ which controls the relative weight of the type I and type II seesaw mechanism. In Fig. 1 the $\chi^2$ minimum is shown as a function of this parameter. This means that for fixed $|\xi|$, $\chi^2(x_\alpha)$ is minimized with respect to all the other parameters $x_\alpha$ with $\alpha \neq |\xi|$. In Table II some parameter values and the predictions for the data are given for four sample points.

One of the main results of this work is clearly visible in Fig. 1: We find a pronounced minimum for $|\xi| \approx 0.36$ which corresponds to a mixture of type I and type II seesaw with a comparable size of both terms. Such a “mixed” solution provides an excellent fit to the data with $\chi^2 = 0.35$. From the results given in Table II in the column labeled “Mixed,” it is seen that all observables are fitted within $\pm 0.4\sigma$. In particular, all the lepton mixing angles and the neutrino mass-squared differences are very close to their experimental values. Also CKM CP violation is described correctly by the value $\delta_{\text{CKM}} = 61^\circ$ obtained in this solution. We conclude that a scenario with both seesaw

$^1$Since we fix the charged lepton masses to the values given in Eq. (10) they are neither included in the set of observables $O_i$ nor are they among the parameters $x_\alpha$. terms of comparable size allows for an excellent description of fermion masses and mixings, confirming the results of Ref. [25]. Whether this solution is viable still depends on the detailed study of the global vacuum of the given $SO(10)$ model.

The $\chi^2$ increases by increasing $|\xi|$, and it approaches a value of $\chi^2 = 4.3$ for $|\xi| \approx 10$, when the neutrino mass matrix becomes dominated by the type I seesaw term. Also in the case of complete type I dominance the fit is very good, with most observables within $\pm 0.3\sigma$, with the sole exception of the down-quark mass $m_d$ which shows a $-1.87\sigma$ deviation from its prediction (see Table II, column “Type I”). As we will discuss in more detail later, a low value of $m_d$ is required for a valuable type I seesaw fit. Let us note that all neutrino parameters are in excellent agreement with the observations, and, in particular, the correlation between $\theta_{12}^{\text{PMNS}}$, $\theta_{23}^{\text{PMNS}}$, and the ratio $\Delta m^2_{31}/\Delta m^2_{21}$ found in Ref. [25] seems not to apply here. This solution is not plagued by the problem of accommodating the correct $\delta_{\text{CKM}}$ phase found in Ref. [17] and discussed on general grounds in Ref. [24].

For a pure type II solution with $|\xi| = 0$ there is more tension in the fit. Although a $\chi^2 \approx 14.5$ might be acceptable for 15 data points from a statistical point of view, several observables are 1 to $2\sigma$ away from the central value. Among all, one needs a large value of the strange-quark mass, namely, $m_s \approx 31$ MeV that is $1.85\sigma$ too large, while the PMNS angle $\theta_{23}$ is too small by $2.2\sigma$. Furthermore, $m_d$, the CKM phase, and the PMNS angle $\theta_{12}$ show a pull greater than $1\sigma$. These results agree with previous analyses of pure type II solutions [14,24,25].

The dashed line in Fig. 1 corresponds to an interesting variant of a solution with comparable type I and type II contributions. Formally this solution has a rather small
TABLE II. Parameter values and predictions in four example solutions corresponding to different terms dominating the neutrino mass matrix: type I, type II, or both contributions of comparable size (mixed and mixed'). In the column “Pred.” the predicted values \( P_i \) for the observables are given, the column “Pull” shows the number of standard deviations from the observations, \( (P_i - \bar{O}_i)/\sigma_i \), using the data and errors from Table I. Deviations of more than 1 \( \sigma \) are highlighted in boldface. The final \( \chi^2 \) is the sum of the squares of the numbers in the pull column. See the text for comments on the values of the leptonic CP phases.

<table>
<thead>
<tr>
<th>Observable</th>
<th>Type II</th>
<th>Mixed'</th>
<th>Mixed</th>
<th>Type I</th>
</tr>
</thead>
<tbody>
<tr>
<td>(</td>
<td>\bar{\xi}</td>
<td>)</td>
<td>0.0024</td>
<td>0.0024</td>
</tr>
<tr>
<td>(\arg(\bar{\xi}))</td>
<td>0.866π</td>
<td>1.018π</td>
<td>1.318π</td>
<td>1.318π</td>
</tr>
<tr>
<td>(</td>
<td>r</td>
<td>)</td>
<td>0.3278</td>
<td>1.9977</td>
</tr>
<tr>
<td>(</td>
<td>\bar{r}</td>
<td>)</td>
<td>0.408π</td>
<td>1.849π</td>
</tr>
<tr>
<td>(f_u)</td>
<td>16.62</td>
<td>11.51</td>
<td>18.77</td>
<td>19.23</td>
</tr>
<tr>
<td>(f_v)</td>
<td>1.671×10^{-10}</td>
<td>4.519×10^{-10}</td>
<td>8.732×10^{-10}</td>
<td>3.613×10^{-17}</td>
</tr>
</tbody>
</table>

\(\chi^2\) | 14.5 | 4.1 | 0.35 | 4.3 |

The value of \(|\bar{\xi}|\), which would signal type II dominance. However, in this case one eigenvalue of \((M_d - M_e)^{1/2}\) is very small, i.e., this matrix is close to singular, which implies that its inverse has large entries. As a consequence the second term in Eq. (6) turns out to be comparable to the first term, in spite of the small value of \(|\bar{\xi}|\). Let us denote the type II term in Eq. (6) by \(M^{11}_{dd}\) and the type I term by \(M^{10}_{3}\). Then for the solution denoted mixed’ in Fig. 1 and Table II we find that the matrix entries are of a similar size: 0.2 \(\leq |\bar{M}_{ij}/M_{ij}^{10}| \leq 1.5\) for all \( i, j = 1, 2, 3 \). Obviously this solution involves a very precise tuning between the \(M_d\) and \(M_e\) matrices such that the difference becomes close to singular. Changing the input values for the down-quark masses and the charged lepton mixing matrix \(W_e\) by a factor of \(1 + 10^{-4}\) destroys in general the fit and leads to \(\chi^2\) values of order 100.\(^2\) Let us add that we obtain mixed’ type solutions by extrapolating from \(|\bar{\xi}| = 10^{-4}\) down to very small values of \(|\bar{\xi}| \approx 10^{-9}\). However, at some point these results are likely to become unreliable, since the numerical inversion of a nearly singular matrix is limited by the accuracy of the algorithm.

**A. The role of \(m_d\)**

The down-quark mass plays a relevant role in obtaining a good fit to the data. As it is visible from Table II the type II, mixed’, and type I solutions require values of \(m_d\) which are about 1.2\(\sigma\), 1.8\(\sigma\), and 1.9\(\sigma\) below the preferred value given in Table I. In Fig. 2 we illustrate how the fits become worse when \(m_d\) is increased. Especially the very good fits of the type I and mixed’ solutions with \(\chi^2 \approx 4\) are strongly affected by increasing \(m_d\), and values of \(m_d \approx 1.5\) MeV at \(M_{GUT}\) lead to \(\chi^2 \approx 15\) and 20, respectively.
Also for the pure type II case the fit soon becomes unacceptable if \( m_d \) is increased. Only the mixed solution has enough freedom to accommodate larger values of the down-quark mass, and in this case also a very good fit is possible for \( m_d \approx 2 \) MeV.

As a matter of fact, it is quite easy to understand these results on an analytical basis. As it was pointed out in Refs. [14,24], the quality of the charged sector fit is gauged by the tiny electron mass that obeys an approximate formula of the form [24]:

\[
|k'\tilde{m}_e|e^{i\phi} = -|\bar{\rho}|e^{i\beta_1}F_d\Lambda^4 + e^{i\alpha_2}F_c\Lambda^6 - A^2\Lambda^2e^{i\alpha_1}\Lambda^6\frac{|\bar{\rho}|}{e^{i\alpha_2} - |\bar{\rho}|} + \mathcal{O}(\Lambda^7),
\]

where \( \tilde{m}_e \equiv m_e/m_t \), \( \Lambda \equiv 1 - \rho - i\eta \) (\( \rho, \eta \) being the Wolfenstein CKM parameters and \( \lambda \) the Cabibbo angle), while \( F_d \equiv m_d/m_t/\Lambda^4 \), \( F_c \equiv m_c/m_t/\Lambda^4 \) are \( \mathcal{O}(1) \) factors. The CKM phases \( \alpha_i, \beta_i \) are defined in Ref. [24]. The parameters \( \tilde{r} \) and \( k' \) are given by

\[
|\tilde{r}| = \frac{m_s}{m_t}f_d|r - 1|, \quad |k'| = \frac{m_s}{m_t}f_d|r + 3|.  \tag{13}
\]

It is clear that in order to get near the physical value \( \tilde{m}_e \sim 2.5 \times 10^{-4} \) for \( |k'| \sim 0.25 \) (as suggested by the relevant trace identities), the dominant first term on the right-hand side of Eq. (12) must be either strongly suppressed (leading to the observed effect of low \( m_d \) preference) or cancelled to a large extent by the subleading ones (often at odds with the CKM phase in the first quadrant, c.f. [14,24]).

Therefore, for low values of \( m_d \) the available portion of the parameter space is larger and allows for a better global fit of the remaining physical parameters. It is perhaps worth mentioning that there is another pattern in our data (though by far much weaker than the low \( m_d \) preference) that can be justified on the same grounds, namely, a generic drift towards lower \( m_t \) and higher \( ms \) regions. One verifies that in such a case the subdominant term proportional to \( F_c \) gets larger and allows for a better “screening” of the dominant first term on the right-hand side of formula (12).

B. Predictions for the neutrino parameters

In this section we discuss in some detail the predictions of the studied setup for the neutrino sector. Our main results are summarized in Fig. 3, where we show how the \( \chi^2 \) changes if \( \theta_{23}^{\text{PMNS}}, \theta_{13}^{\text{PMNS}} \), and the mass of the lightest neutrino \( m_1 \) are varied. Technically this analysis is performed in the following way: to test the variation of an observable \( O_k \), the term with \( i = k \) is removed from the \( \chi^2 \) given in Eq. (11) (this ensures that the only constraint on the observable comes from the mass sum rules and not from the input data). Then, in order to test a certain value \( O^* \) for the observable \( O_k \) a term is added to the \( \chi^2 \) with a very small error of 1%, to confine the fit: \( (P_k(x^*_0) - O^*)^2/(0.01O^*)^2 \). After the minimization this term is removed, and the \( \chi^2 \) is evaluated at the point obtained in the minimization.

Figure 3 (left) shows the constraint on the PMNS mixing angle \( \theta_{23}^{\text{PMNS}} \). One can see that for the type I, mixed, and mixed’ solutions, there is no definite prediction for this angle and very good fits are possible with values of \( \theta_{23}^{\text{PMNS}} \) in the whole range allowed by the data. In contrast a scenario with pure type II seesaw shows a clear preference for small values of \( \sin^2\theta_{23}^{\text{PMNS}} \). In particular, maximal mixing \( \sin^2\theta_{23}^{\text{PMNS}} = 0.5 \) is disliked with respect to the best fit value \( \sin^2\theta_{23}^{\text{PMNS}} = 0.36 \) with \( \Delta \chi^2 = 11 \). Hence, the pure type II seesaw model predicts sizable deviations from maximal mixing within the reach of upcoming neutrino oscillation experiments, see, e.g., Refs. [40,42]. This result is in agreement with Refs. [24,25]. Let us stress, however, that deviations from maximal mixing are not a general prediction of the \( SO(10) \) model under consideration; it holds only for the pure type II case.

Concerning the mixing angle \( \theta_{13}^{\text{PMNS}} \), for all solutions the best fit point predicts values close to the present upper bound and clearly within the reach of upcoming neutrino oscillation experiments [40,43]. Also in this case the pure type II solution gives the most stringent prediction. However, if the fit is stretched to some degree, the type I and mixed solutions allow also for smaller values of \( \theta_{13}^{\text{PMNS}} \) that might be difficult to detect in the next round of experiments. For instance, in the mixed solution case a fit with \( \chi^2 \sim 7 \) is possible for \( \sin^2\theta_{13}^{\text{PMNS}} \approx 2 \times 10^{-3} \). An interesting feature of the mixed' solution is that even for very tiny values of \( \sin^2\theta_{13}^{\text{PMNS}} \leq 10^{-4} \) a reasonable fit can be ob-
tained with $\chi^2 = 12.6$. The main contributions to the $\chi^2$ in this case are deviations of $-1.7\sigma$ for $m_{d1}$, $-1.2\sigma$ for $m_1$, $-2.0\sigma$ for $m_2$, and $+1.4\sigma$ for $\sin^2\theta_{13}^{\text{PMNS}}$. The existence of this solution shows that if no signal of $\theta_{13}^{\text{PMNS}}$ is detected by the upcoming experiments it might be still possible to construct models with viable predictions for fermion masses and mixings, although the amount of fine-tuning will increase.

Another interesting piece of information the setup under consideration provides concerns the shape of the neutrino mass spectrum. The most solid prediction is the normal mass ordering, which means that $m_1 < m_2 < m_3$; no viable solution has been found for inverted ordering ($\Delta m^2_{31} < 0$) which is equally allowed by present oscillation data. The predictions for the absolute neutrino mass scale are given in terms of the ratio of the lightest neutrino mass $m_1$ to the square root of the “solar” mass-squared difference:

$$R \equiv \frac{m_1}{\sqrt{\Delta m^2_{21}}}.$$  

The best fit values for this ratio given in Table II for the four example solutions are in the range $0.28 \leq R \leq 0.48$. These values show that there is only a modest hierarchy in the neutrino masses. For example, a value of $R = 0.3$ implies that $m_2 \approx 3.5m_1$, i.e., $m_1$ and $m_2$ are of the same order of magnitude. From the plot in the right panel of Fig. 3 one can infer the allowed ranges for the ratio $R$. We find reasonable fits for values in the range

$$0.2 \leq R \leq 2,$$

where the upper bound implies $m_2 \approx 1.1m_1$. For the pure type II solution the ratio $R$ is stronger constrained to values around the best fit point of 0.28, whereas the constraint is weakest for the mixed solution. Note that a quasidegenerate

The fact that the fit always gives a hierarchical neutrino mass spectrum justifies a posteriori the neglect of the running of the neutrino parameters [37,38].
Moving $\delta_{\text{PMNS}}$ from the preferred value to $\pm 60^\circ$ increases the $\chi^2$ by 10 units.

IV. THE MINIMAL SUSY SO(10) GUT

Having analyzed the implications on the fermion mass fit of the $10_H$ plus $\overline{126}_H$ SO(10) Yukawa sector, we focus now on the study of the so-called minimal renormalizable SUSY SO(10) GUT [13]. The model is characterized (in addition to $10_H$ and $\overline{126}_H$) by the presence of the $126_H$ and $210_H$ representations in the Higgs sector. The $126_H$ is needed to preserve the GUT scale $D$ flatness while $210_H$ plays the dual role of triggering the spontaneous SO(10) gauge symmetry breaking and provides the necessary mixing among the $10_H$ and $\overline{126}_H$ weak doublet components needed to achieve (after electroweak symmetry breaking) a realistic fermion spectrum. The tiny VEV of the left-handed triplet component of $\overline{126}_H$ responsible for the type II seesaw is induced via $210_H$ couplings as well. The model is minimal in the number of parameters, 26 altogether (soft SUSY breaking aside). The same number is found in the minimal supersymmetric standard model (MSSM) with right-handed neutrinos and it is much less than the number of parameters in the corresponding SUSY SU(5) GUT.

Since none of the $126_H$ nor $210_H$ Higgs multiplets can couple to the SO(10) matter bilinear $16_m \otimes 16_m$ at the renormalizable level, the Yukawa superpotential reads

$$W_Y = 16_m(Y_{1010_H} + Y_{126\overline{126}_H})16_m,$$

while the SO(10) Higgs sector is described by

$$W_H = \frac{M_{210}}{4!} 210_H^2 + \frac{\lambda}{4!} 210_H^3 + \frac{M_{126}}{5!} 126_H \overline{126}_H$$

$$+ \frac{\eta}{5!} 126_H 210_H \overline{126}_H + M_{10} 10_H^2$$

$$+ \frac{1}{4!} 210_H 10_H (\alpha 126_H + \hat{\alpha} \overline{126}_H).$$

The minimization of the scalar potential has been analyzed in great detail in Refs. [15,29,31]. The need of a careful study of the GUT scale potential in order to perform consistent predictions of the fermion mass textures is emphasized in these papers. The light MSSM-like Higgs doublets entering the low energy Yukawa potential are in general a superposition of corresponding doublet components of all the Higgs multiplets. The set of VEVs which describes the vacuum of the model is governed by one complex parameter $x$ [15]:

$$\langle 1, 1, 1 \rangle_{210} = - \frac{M_{210}}{\lambda} \frac{x(1 - 5x^2)}{(1 - x^2)},$$

$$\langle 15, 1, 1 \rangle_{210} = - \frac{M_{210}}{\lambda} \frac{(1 - 2x - x^2)}{(1 - x)},$$

$$\langle 15, 1, 3 \rangle_{210} = \frac{M_{210}}{\lambda} x,$$

$$\langle \overline{10}, 1, 3 \rangle_{126}(10, 1, 3)_{\overline{126}} = \frac{2M_{210}}{\eta \lambda} \frac{x(1 - 3x)(1 + x^2)}{(1 - x)^2}.$$ (20)

Avoiding $D$-term SUSY breaking requires $\langle \overline{10}, 1, 3 \rangle_{126} = (10, 1, 3)_{\overline{126}} = v_R$. In Fig. 4 we depict the relevant SO(10) breaking patterns together with the related VEVs.

The various mass parameters in the GUT superpotential can be written as functions of $x$ as well. For instance

$$M_{126} = \frac{2 \eta \lambda}{(x - 1)^2}.$$ (21)

The mass parameter $M_{10}$ is determined by the minimal fine-tuning condition [13,15] which drives the mass of two Higgs doublets down to the weak scale. The relevant formula reads

$$M_{10} = \frac{M_{210}}{2 \eta \lambda} \frac{\alpha \hat{\alpha}}{(x - 1)p_3 p_5},$$

where $p_n$ are polynomials of $x$; see Refs. [15,29]. The

![FIG. 4. Symmetry breaking patterns and related VEVs in the minimal SUSY SO(10). We do not display breaking chains involving intermediate SU(5) symmetries, at odds with the proton decay constraints. Pati-Salam notation is used.](image-url)

115012-8
SU(2)_L triplet mass, relevant for the type II seesaw, can be written as
\[ M_T = M_{210}^{\eta} \frac{x(4x^2 - 3x + 1)}{(x - 1)^2}, \] (23)
while the tiny induced SU(2)_L triplet VEV is given by [13,31]
\[ v_L \equiv \langle 10, 3, 1 \rangle_{26} = \frac{(\alpha v_u^{10} + \sqrt{6} \eta v_u^{10}) v_u^{10}}{M_T}. \] (24)

The electroweak VEV components \( v_u \)'s are determined in terms of the electroweak scale and the \( x \) parameter; c.f. Refs. [15,29].

In the first part of the paper we learned that the best fit of the present data on fermion masses and mixing is obtained by a mixed contribution of type I and type II seesaws. The model independent analysis based only on the form of the Yukawa potential in Eq. (18) assumed the vacuum state to provide the required set of parameters, in particular, the correct neutrino mass scale. In the following we examine thoroughly the question whether such a vacuum exists in the minimal renormalizable SUSY SO(10) framework. The crucial issues are gauge coupling unification and the phase convention of Ref. [29]. After minimal fine-tuning, the Higgs potential can be described in terms of the electroweak scale and the parameter \( \lambda \).

For our purpose (fermion masses and symmetry breaking) without loss of generality we can set \( \Phi = 0 \). As emphasized in Refs. [15,29], the parameter \( \lambda \) provides a systematic and effective way of describing the SO(10) symmetry breaking patterns and the related fermion mass scales.

In the minimal SUSY SO(10) framework under consideration, the VEVs appearing in the fermion sum rules Eq. (1) are no longer free parameters (as assumed in the numerical analysis of Sec. III) but are functions of the Higgs potential parameters; c.f. Eq. (25). Using the results and notation of Ref. [29], one obtains the following expressions for the VEV combinations appearing in the fermion fit as defined in Eqs. (5) and (7):

\[ f_u = \frac{1}{4} \tan \beta \frac{N_u}{N_d}, \]
\[ r = \frac{2x q_5}{p_2 p_5} + 1, \]
\[ f_v = \frac{1}{M_{210}} \tan \beta \sin \alpha \sqrt{\frac{|\lambda| N_u^2}{|\eta| N_d}} |f_{ll}(x)|, \]
\[ \xi = \frac{1}{16} f_1(x). \]

Here \( N_u, N_d \) are functions of the parameters in Eq. (25) given explicitly in the appendix of Ref. [29]; \( \nu = 174 \text{ GeV} \) is the electroweak scale, while

\[ f_1 = \frac{2x q_5}{p_2 p_5}, \quad f_{ll} = \frac{(x - 1)(4x - 1)p_3 q_5^2 \sigma}{2x p_2 q_2}, \] (27)

where

\[ \sigma = \frac{2x(1 - 3x)(1 + x^2)}{(1 - x)^2}. \] (28)

and \( p_5, q_5 \) are polynomials of \( x \) of order \( n \) (see Table I in Ref. [29]).

We now address the question of whether a realistic fermion fit is possible, given the constraints from the Higgs sector encoded in Eqs. (26) which together with Eqs. (3), (4), and (6) lead to a highly constrained system of relations. First we note that not much freedom is left in adjusting the parameters \( \alpha, \tilde{\alpha}, \eta, \lambda \). This follows from the functional form of the VEVs in Eqs. (20), of the potential mass parameters, and of \( f_v, N_u, N_d, \) as well as from the requirement of perturbativity of the Higgs potential. Hence, in what follows we set \( \alpha = \tilde{\alpha} = \eta = \lambda = 1 \) [29]. Similar considerations suggest also \( |x| \sim O(1) \), and therefore we first restrict our search to the range \( 0.1 \leq |x| \leq 10 \) (deferring a discussion on the small and large \( x \) regimes to the end of the section).

In order to identify candidate solutions we perform a scan in the complex \( x \) plane by first fixing \( \alpha = \tilde{\alpha} = \eta = \lambda = 1 \). The \( x \) dependence of \( r \) and \( \xi \) is taken into account according to Eq. (26), leaving however the neutrino mass normalization \( f_v \) as a free parameter to be determined from the fit. Also \( f_u \) becomes a free parameter, if \( \tan \beta \) is allowed to vary.

In Fig. 5 we show the \( \chi^2 \) contours of the fermion mass fit in the complex \( x \) plane, where darker areas represent better fits of the fermion masses and mixings according to the \( \chi^2 \) test.
values given. As is visible from the figure, we find large regions in the x plane where a very good fit with $\chi^2 < 10$ can be obtained. The crucial question is whether within these regions it is possible to obtain also a consistent value of $f_\nu$. From Eq. (6) one finds that $f_\nu$ or $f_\nu |\xi|$ has to be of order $\sqrt{\Delta m_{31}^2/m_b} \sim 5 \times 10^{-11}$ to provide a neutrino mass scale required by oscillation data, in agreement with the values reported in Table II. In contrast, from Eq. (26) one finds a natural size of

$$f_\nu \sim \nu \tan \beta / M_{210} \sim 5 \times 10^{-13} \left(\tan \beta / 55\right)$$  \hspace{1cm} (29)$$

for $M_{210} = 2 \times 10^{16}$ GeV. This estimate illustrates that in the framework under consideration a gap of about 2 orders of magnitude exists between the generic prediction of the neutrino masses and the scale required by the data. We have verified this expectation numerically and for the most part of the complex x plane shown in Fig. 5, the value of $f_\nu(x)$ according to Eq. (26) is 2 to 4 orders of magnitude smaller than the value of $f_\nu$ required by the fit [denoted by $f_\nu(\text{fit})$].

The only way left to obtain the needed neutrino mass scale is to consider special points in the x plane, where either $|f_\nu(x)|$ or $|f_\nu(\xi)|$ becomes large because of roots in the denominators of Eq. (27). These singularities are marked in Fig. 5 by the crosses. The closed black contours correspond to the regions where $f_\nu(x)/f_\nu(\text{fit}) > 0.01$. We find that only close to the singularities $f_\nu$ can be large enough to provide the required neutrino mass scale. On the other hand, it appears that most of the singularities drop into regions where the fermion mass fit is poor. We carefully checked numerically the regions close to the relevant singularities and came to the conclusion that only in the near neighborhood of $x = \pm i$ a realistic fit of the matter sector is possible. Helpful discussions of specific points and limiting cases in the x complex plane based on analytical considerations can be found in Refs. [15,29].

In Table III we report the best fit solution we found. The absolute value and phase of x are tuned at the level of $10^{-7}$ in the close neighborhood of i. The “symmetric” solution

TABLE III. Parameter values and outcomes for the best fit solution, $x \approx i$. The reference mass parameter $M_{210}$ is set at a typical MSSM GUT scale. The charged lepton masses are given in Eq. (10). Deviations above 2$\sigma$ in the fermion mass data are highlighted in boldface. The final $\chi^2$ is the sum of the squares of the numbers in the pull column.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Best fit</th>
<th>$\lambda, \alpha, \tilde{\alpha}$ fixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>x</td>
<td>$</td>
</tr>
<tr>
<td>$\arg(x)/\pi$</td>
<td>0.4999999781</td>
<td>0.4999999815</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.3191</td>
<td>0.1741</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.5642</td>
<td>1</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>2.9455</td>
<td>1</td>
</tr>
<tr>
<td>$\tilde{\alpha}$</td>
<td>4.0332</td>
<td>1</td>
</tr>
<tr>
<td>$M_{210}$ [GeV]</td>
<td>$2.000 \times 10^{16}$</td>
<td>$2.000 \times 10^{16}$</td>
</tr>
<tr>
<td>$\tan \beta$</td>
<td>41.15</td>
<td>45.88</td>
</tr>
<tr>
<td>$</td>
<td>\xi</td>
<td>$</td>
</tr>
<tr>
<td>$\arg(\xi)/\pi$</td>
<td>$-0.60492$</td>
<td>$-0.53116$</td>
</tr>
<tr>
<td>$</td>
<td>r</td>
<td>$</td>
</tr>
<tr>
<td>$\arg(r)/\pi$</td>
<td>$-0.2568$</td>
<td>$-0.2568$</td>
</tr>
<tr>
<td>$f_\nu$</td>
<td>12.28</td>
<td>11.47</td>
</tr>
<tr>
<td>$(1, 1, 1)$ [GeV]</td>
<td>$3.836 \times 10^{16}$</td>
<td>$6.000 \times 10^{16}$</td>
</tr>
<tr>
<td>$(15, 1, 1)$ [GeV]</td>
<td>$2.557 \times 10^{16}$</td>
<td>$4.000 \times 10^{16}$</td>
</tr>
<tr>
<td>$(15, 1, 3)$ [GeV]</td>
<td>$1.279 \times 10^{16}$</td>
<td>$2.000 \times 10^{16}$</td>
</tr>
<tr>
<td>$(10, 1, 3)$ [GeV]</td>
<td>$3.423 \times 10^{13}$</td>
<td>$3.009 \times 10^{13}$</td>
</tr>
<tr>
<td>$M_{\nu}$ [GeV]</td>
<td>$4.135 \times 10^{16}$</td>
<td>$2.043 \times 10^{16}$</td>
</tr>
<tr>
<td>$M_T$ [GeV]</td>
<td>$8.655 \times 10^{15}$</td>
<td>$1.354 \times 10^{16}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observable</th>
<th>Pred.</th>
<th>Pull</th>
<th>Pred.</th>
<th>Pull</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_d$ [MeV]</td>
<td>0.3435</td>
<td>$-2.2$</td>
<td>0.3373</td>
<td>$-2.2$</td>
</tr>
<tr>
<td>$m_u$ [MeV]</td>
<td>27.91</td>
<td>1.2</td>
<td>26.36</td>
<td>0.90</td>
</tr>
<tr>
<td>$m_b$ [MeV]</td>
<td>1065</td>
<td>0.038</td>
<td>1112</td>
<td>0.37</td>
</tr>
<tr>
<td>$m_c$ [MeV]</td>
<td>0.5596</td>
<td>0.038</td>
<td>0.5639</td>
<td>0.056</td>
</tr>
<tr>
<td>$m_t$ [MeV]</td>
<td>212.8</td>
<td>0.15</td>
<td>213.1</td>
<td>0.16</td>
</tr>
<tr>
<td>$m_l$ [MeV]</td>
<td>82090</td>
<td>$-0.21$</td>
<td>79550</td>
<td>$-0.19$</td>
</tr>
<tr>
<td>$\sin \phi^{CKM}_{23}$</td>
<td>0.0351</td>
<td>$-0.039$</td>
<td>0.0350</td>
<td>$-0.065$</td>
</tr>
<tr>
<td>$\sin \phi^{CKM}_{13}$</td>
<td>0.00325</td>
<td>0.10</td>
<td>0.00329</td>
<td>0.18</td>
</tr>
<tr>
<td>$\sin \phi^{CKM}_{12}$</td>
<td>0.2244</td>
<td>0.033</td>
<td>0.2244</td>
<td>0.057</td>
</tr>
<tr>
<td>$\delta^{CKM}[^\circ]$</td>
<td>64.32</td>
<td>0.31</td>
<td>70.10</td>
<td>0.72</td>
</tr>
<tr>
<td>$\sin \theta^{PMNS}_{23}$</td>
<td>0.4893</td>
<td>$-0.16$</td>
<td>0.5002</td>
<td>0.0035</td>
</tr>
<tr>
<td>$\sin \theta^{PMNS}_{13}$</td>
<td>0.0133</td>
<td>0.86</td>
<td>0.01085</td>
<td>0.70</td>
</tr>
<tr>
<td>$\sin \theta^{PMNS}_{12}$</td>
<td>0.3007</td>
<td>$-0.37$</td>
<td>0.2953</td>
<td>$-0.59$</td>
</tr>
<tr>
<td>$\Delta m_{31}^2 [10^{-5} \text{ eV}^2]$</td>
<td>7.875</td>
<td>$-0.082$</td>
<td>7.860</td>
<td>$-0.13$</td>
</tr>
<tr>
<td>$\Delta m_{32}^2 [10^{-3} \text{ eV}^2]$</td>
<td>2.199</td>
<td>$-0.002$</td>
<td>2.274</td>
<td>0.20</td>
</tr>
<tr>
<td>$m_1/\sqrt{\Delta m_{32}^2}$</td>
<td>0.3351</td>
<td>0.3289</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\chi^2$ 7.25 7.30
whereas in the third column we have taken $f_4$ implies dominated by type I seesaw since $\xi \rightarrow \infty$, and hence $C_2$ see Eqs. (26) and (27). From the data in the table one finds $f_4(\xi) \sim 10^{-11}$, which is the correct value to provide the required neutrino mass scale. The table shows that a very good fit of all quark and lepton mass and mixing parameters (including the neutrino sector) is obtained, with just a 2.2$\sigma$ pull for $m_d$ that is acceptable in view of the systematic uncertainties entering the light quark mass determinations.

B. Gauge couplings unification and proton decay for $x \approx \pm i$

What remains to be checked is whether the best fit solution satisfies unification and proton decay constraints. For $x = \pm i$ the $SO(10)$ gauge symmetry is broken to $G_{3211} = SU(3)_C \times SU(3)_L \times U(1)_R \times U(1)_{R-L}$ by the VEV of $(15, 1, 3)_{210}$ (compare Fig. 4 and Refs. [15,29]). The calculation of the particle spectrum shows that in this limit a set of unmixed states that transform as $(8, 3, 0, 0)$ under $G_{3211}$ or $(8, 3, 0)$ under the SM gauge group remains light. In the exact limit $x = \pm i$ these Goldstone states are massless because of the spontaneous breaking of accidental global symmetries of the Higgs potential. Upon breaking the $G_{3211}$ gauge symmetry to the SM via the VEVs, $\langle 10, 1, 3 \rangle_{126} = \langle 10, 1, 3 \rangle_{126} = v_R (\equiv M_{3211})$ they acquire a mass $M_{PS}$ of the order of $M_{3211}/M_{PS}$, where $M_{PS} = \langle 1, 1, 1 \rangle_{210}$ is the Pati-Salam scale. According to Ref. [15] the expression for the pseudo-Goldstone mass is

$$M_{PS} = -4M_{210} \frac{(2x - 1)(x^2 + 1)}{(x - 1)^2},$$

which is exactly zero for $x = \pm i$. In our best fit case reported in Table III we find $M_{3211} \sim 10^{13}$ GeV, while $M_{PS} \sim 10^{16}$ GeV. One therefore expects $M_{PS} \sim 10^{10}$ GeV, as we consistently find (see table).

In spite of the fact that these states are much lighter than the GUT scale, they do not affect proton decay because of their zero $B - L$ charge. However, they do affect heavily the running of the gauge couplings [they transform as adjoint representations of $SU(3)_C$ and $SU(2)_L$]. Figure 6 shows how dramatically the running is affected. The unification of the gauge couplings is completely spoiled, with the coupling constant of $SU(2)_L$ diverging below the GUT scale. Hence, the only solution found providing a realistic fit to all fermion masses and mixing parameters has to be discarded because of the dramatic failure of the gauge couplings unification. The minimal renormalizable SUSY $SO(10)$ scenario seems, therefore, failing to provide a realistic description of the low energy world. In general, if right-handed neutrinos transform nontrivially under the GUT gauge symmetry, the minimum neutrino mass scale required by oscillation data implies via type I (or type II) seesaw the presence of states at intermediate mass scales [29,30], that spoil the successful unification within the MSSM.

FIG. 6. Running gauge coupling constants in the setup of Table III. We have $M_{PS} = 1$ TeV, $M_{PS} = M_{PS} \approx (10, 1, 3), \quad M_{3211} = (15, 1, 3)_{210}, \quad M_{PS} = (1, 1, 1), \quad M_{PS} = (1, 1, 1),$ and $M_{PS} = M_{PS} = 2 \times 10^{16}$ GeV.

In passing let us stress that the example discussed above shows quite clearly that the naïve argument of setting all relevant particle states at the scale of the symmetry breaking step may be far from correct. One should always be aware that much lighter states may appear in the spectrum as a consequence of the spontaneous breaking of accidental (would-be) symmetries of the scalar potential [4].

C. Large/small $|x|$ regime

Before conclusively rejecting the minimal renormalizable SUSY $SO(10)$ scenario, we must make sure that the domain $0.1 \leq |x| \leq 10$ considered in the previous analysis covers the physically relevant region. Indeed, the shape of the $\chi^2$ contours in Fig. 5 suggests that all the matter fermion data might be well reproduced in areas outside the considered region. Let us give simple semianalytic arguments against this option.

---

6We neglect the tiny hierarchy among the VEVs of $(1, 1, 1)_{210}, (15, 1, 3)_{126},$ and $(15, 1, 3)_{210}$ as all of them are confined for $x = \pm i$ within an interval of less than 1 order of magnitude; c.f. Eq. (20) and Table III.

7More precisely, the quantum numbers of the states responsible for this symmetry breaking step are $(1, 1, \pm 1, \mp 2)^{3211} \in (1, 1, 3, \mp 2)^{3211}$ and $(10, 1, 3)_{126} \oplus H.c.$
By observing that in the large $x$ regime all $210_H$ VEVs in Eq. (20) scale as $M_{210|x}/\lambda$ and proton decay forces us to maintain most of the GUT states above $M_{\text{GUT}} \approx 10^{16}$ GeV, we must require

$$\frac{M_{210|x}}{\lambda} \approx 10^{16} \text{ GeV.} \quad (31)$$

In the large $x$ limit (and for large $\tan \beta$) from Eqs. (26) and (27) and the expressions for $N_\nu$, $N_d$ given in Ref. [29], one finds for the neutrino mass scales

$$f_\nu[\xi] \approx \frac{v}{M_{210|x}} \tan \beta, \quad f_\nu \approx \frac{v}{M_{210|x}} \frac{\tan \beta}{|x|}, \quad (32)$$

in the type I and type II seesaw cases, respectively. By requiring $M_{210|x} \approx \text{const}$ [see Eq. (31)] these relations show that also for large $x$ the neutrino mass scale cannot be efficiently enhanced over its natural value $v^2/M_{\text{GUT}}$, and accordingly our numerical calculation yields values of $f_\nu$ of about 4 orders of magnitude too small in the region $|x| > 10$. In passing let us remark that in the $|x| \gg 1$ regime there appear a number of states with masses around $M_{210} < M_{\text{GUT}}$ that again spoil gauge unification, albeit not affecting proton decay (see Ref. [15]).

In the small $|x|$ regime one finds for both seesaw types

$$f_\nu, f_\nu[\xi] \approx \frac{v}{M_{210|x}} \tan \beta \sqrt{|x|}. \quad (33)$$

Once again taking into account Eqs. (20) and (31) and the fact that scalar couplings are bound to vary in a narrow range by perturbativity on one side and mass scale splittings on the other, no substantial enhancement of the neutrino mass can be expected. Our $\chi^2$ fit does confirm numerically these expectations.

We therefore conclusively show, independently confirming the conclusions of Ref. [30], that the minimal renormalizable SUSY $SO(10)$ setup in spite of noteworthy and impressive features in reproducing the observed flavor textures fails in reproducing the neutrino mass scale. Generally, in a renormalizable GUT framework, type I and type II seesaws call for intermediate scales that spoil (SUSY) gauge unification.

V. CONCLUSIONS

In the present paper we studied in great detail the fit of quark and lepton masses and mixing data within a class of supersymmetric $SO(10)$ grand unified models with a minimal renormalizable Yukawa sector based on one $10_H$ and one $\overline{126}_H$ Higgs representations. A systematic optimization of the relevant $\chi^2$ function has been performed, complementary to random parameter searches applied in previous studies. We have shown that for a comparable size of the type I and type II seesaw terms, an excellent fit to the fermion data can be obtained, where all observables (including the CKM CP phase) are fitted with less than 0.4 standard deviations. Solutions based on pure type I and type II seesaw have been discussed as well, and we have identified a new class of possible solutions corresponding to a singular behavior of the type II Majorana mass matrix. The corresponding predictions for the neutrino parameters have been investigated in detail.

In the second part of the paper these general results were confronted with the additional restrictions emerging in the minimal renormalizable SUSY $SO(10)$ scenario from the model vacuum as well as from proton decay and unification constraints. We identified a very limited (and fine-tuned) area in the parameter space where, given the constraints from the minimization of the Higgs potential, a fit to the quark and lepton masses and mixing parameters is possible. However, all solutions providing a realistic fit to the matter fermion data had to be discarded because of the failure in reproducing the strength of the low energy gauge couplings. In GUTs that embed nontrivially right-handed neutrinos, the absolute neutrino mass scale required by oscillation data generally implies via type I and/or type II seesaws the presence of states at intermediate mass scales, that are likely to spoil the successful unification achieved in the MSSM. Our analysis did in fact exclude, in the case of the minimal SUSY $SO(10)$ model, the existence of (fine-tuned) exceptional solutions.

The minimal renormalizable SUSY $SO(10)$ scenario based on one $10_H$, one $\overline{126}_H$, one $126_H$, and one $210_H$ Higgs representations is conclusively rejected. Our discussion suggests that either one attempts to enlarge the scalar (Yukawa) potential in order to soften some of the parameter correlations on the vacuum, or nonrenormalizable terms are included providing effectively the needed depletion of the seesaw scale. Alternatively, one is let to consider the intriguing option of a nonsupersymmetric GUT scenario, where intermediate scales, preferred by neutrino data, are needed anyway by gauge coupling unification.

ACKNOWLEDGMENTS

We thank W. Grimus for useful discussions, C. S. Aulakh and R. N. Mohapatra for comments on the manuscript, and T. Shindou for providing $C++$ routines used in this analysis. The work of T. S. is supported by a Marie Curie Intra-European Fellowship within the 6th European Community Framework Program. The work of S. B. is partially supported by MIUR and the RTN European Program MRTN-CT-2004-503369. M. M. is grateful to SISSA and CERN for the hospitality during the preparation of the manuscript.
FERMION MASSES AND MIXINGS IN $SO(10)$ MODELS . . .


5.3 The minimal $SO(10)$ GUT resurrection


In this article the candidate’s hypothesis that quantum corrections can resolve the notorious tachyonicity issues of the minimal adjoint $SO(10)$ Higgs model (see Sect. 3.4.3) has been worked out. This textbook-level result has brought back to life an entire class of minimal grand unified models which were considered non-realistic for almost quarter of a century.

As described in Sect. 3.4 of this thesis a serious issue with the minimal implementation of the Higgs mechanism in the realm of the $SO(10)$ gauge theories has been revealed at the beginning of 1980’s. In sharp contrast to the naïve expectation based on the contemporary studies of other models of spontaneous symmetry breaking the minimal $SO(10)$ Higgs scheme has been shown to possess no stable classical vacuum state that could support potentially realistic symmetry breaking chains.

The merit of the enclosed study consists in the observation that the surprising formal simplicity of the mathematical expressions for certain scalar masses underpinning this issue can be attributed to the pseudo-Goldstone nature of these fields. Several sets of spurion couplings defining various accidental global symmetries of the classical potential have been identified and it was argued that these symmetries should be explicitly broken by higher-order corrections, thus providing a room for the radiative vacuum stabilisation mechanism à la Coleman and Weinberg. A representative set of leading quantum corrections generated by the gauge fields has been calculated and a potentially realistic locally stable quantum vacuum configuration has been found.

These ground-breaking results sparked a renaissance of the minimal $SO(10)$ GUT with many subsequent works following the logic outlined in the article. The paper and its follow-up studies co-authored by the candidate have, to date, earned in total over 160 citations in the inspirehep.net database.

Besides formulating the central idea of understanding the critical states in terms of the Goldstone theorem language the candidate has worked out many details of the calculation including, e.g., a convenient expansion for the second derivatives of the effective potential, identification of the mechanism of cancellation of spurious IR divergences in the quantum-level pole masses etc. He has also contributed by writing a significant part of the manuscript.
I. INTRODUCTION

The grand unified theories (GUTs) [1] qualify among the most appealing physics scenarios beyond the standard model (SM) of electroweak and strong interactions. Though being under scrutiny for about 35 years, they still attract a lot of attention across the high-energy community due to their intrinsic predictivity and to their potential for understanding the origin of our low-energy world texture. Apart from offering definite experimental motivations for, e.g., proton decay or monopole searches, GUTs typically give rise to nontrivial correlations among observables associated to different SM sectors. The most prominent of these is the consistent determination of the weak-mixing angle and the strong coupling arising from the gauge coupling unification in a weak scale supersymmetric scenario.

In recent years, an extra boost to the field was triggered by the discovery of nonzero neutrino masses in the sub-eV region. Within the grand-unified scenarios this discovery translates into constraints on the intermediate scales (typically well separated from the unification scale $M_G \sim 10^{16}$ GeV) underpinning some variant of the seesaw mechanism [2,3]. Furthermore, the observed peculiarity of the lepton mixing pattern [4] challenges the flavor structure of the simplest models due to the strong correlations in the Yukawa sector. In this respect, the requirement of minimality, that stands for the simplicity of the relevant Higgs sector, is a valuable guiding principle for model building.

On this basis, it has been argued recently that the minimal supersymmetric $SO(10)$ model [5–7] is indeed incompatible with the electroweak flavor constraints [8]. The minimal supersymmetric setting suffers from an inherent proximity of the GUT and the seesaw scales, at odds with the lower bound on the neutrino mass scale implied by the oscillation phenomena. The proposed ways out (resorting, e.g., to a nonminimal Higgs sector [9] or invoking split supersymmetry [10]) hardly pair the appeal of the minimal setting.

Were a large (GUT scale) breaking of global supersymmetry at play (a possible LHC test of this hypothesis has been recently put forward in Ref. [11]), then baryon number violating $d = 5$ operators decouple from our low-energy world and gauge unification exhibits naturally the required splitting between the seesaw and the GUT scales [12–15]. Nevertheless, devising a realistic and simple enough $SO(10)$ GUT along these lines remains a rather nontrivial task.

The main reason has to do with the structure of the minimal Higgs sector of nonsupersymmetric $SO(10)$ models. A full breaking of the GUT symmetry down to the SM can be achieved via a pair of Higgs multiplets: one 45-dimensional adjoint representation, $45_H$, and one 16-dimensional representation, $16_L$. Such a setting, suitable for a supersymmetric gauge unification, requires an extended matter sector, including 45 and 120 multiplets, in order to accommodate realistic fermion masses [17].
dimensional spinorial representation, 16_H (or one 126-dimensional tensor representation 126_H). A SM preserving breaking pattern is controlled by two 45_H vacuum expectation values (VEVs) and one 16_H (or 126_H) VEV. Different configurations of the two adjoint VEVs preserve different SO(10) subalgebras, namely, 4c_2L_1R, (shorthand notation for SU(4)_c \otimes SU(2)_L \otimes U(1)_R, \ 3_c 2L 2R 1_{B-L}, \ 3_c 2L 1_R 1_{B-L}, \ and \ the \ flipped \ or \ standard \ SU(5) \otimes U(1)). Except for the latter case, the subsequent breaking to the SM is obtained via the standard SU(5) conserving 16_H (or 126_H) VEV.

Remarkably enough, a consistent SO(10) gauge symmetry breaking in the usual low-scale supersymmetric context requires minimally 45_H \otimes 54_H \ [18] \ (or 210_H \ [6] \ in the renormalizable variant), in addition to 10_H \otimes 16_H (or 126_H \otimes 126_H).

The phenomenologically favored scenarios allowed by gauge coupling unification correspond minimally to a two-step breaking along one of the following directions [15]:

\[
SO(10) \rightarrow 4c_3L_2R_{1-B-L}^{M_4} \rightarrow \text{SM},
\]

(1)

\[
SO(10) \rightarrow 4c_4L_2_1R^{M_4} \rightarrow \text{SM},
\]

(2)

where the first breaking stage is driven by the 45_H VEVs, while the breaking to the SM at the intermediate scale M_4, several orders of magnitude below the unification scale M_G, is controlled by the 16_H (or 126_H) VEV. One of the two 45_H VEVs may also contribute to the second step (see the discussion on the required intermediate scale Higgs multiplets in Ref. [15] and in Sec. V F).

Gauge unification, even without proton decay limits, excludes any intermediate SU(5)-symmetric stages. On the other hand, a series of studies in the early 1980s of the 45_H \otimes 16_H model [19–21] indicated that the only intermediate stages allowed by the scalar sector dynamics were the flipped SU(5) \otimes U(1) for leading 45_H VEVs or the standard SU(5) GUT for dominant 16_H VEV.

This observation excluded the simplest SO(10) Higgs sector from realistic consideration.

In this paper we show that the exclusion of the breaking patterns in Eqs. (1) and (2) is an artifact of the tree-level potential. As a matter of fact, some entries of the scalar Hessian are accidentally overconstrained at the tree level. A number of scalar interactions that, by a simple inspection of the relevant global symmetries and their explicit breaking, are expected to contribute to these critical entries, are not effective at the tree level.

On the other hand, once quantum corrections are considered, contributions of \( O(M_H^2/16\pi^2) \) induced on these entries open in a natural way all group-theoretically allowed vacuum configurations. Remarkably enough, the study of the one-loop effective potential can be consistently carried out just for the critical tree-level Hessian entries (that correspond to specific pseudo-Goldstone boson masses). For all other states in the scalar spectrum, quantum corrections remain perturbations of the tree-level results and do not affect the discussion of the vacuum pattern.

Our conclusions apply to any Higgs setting where the first step of the SO(10) gauge symmetry breaking is driven by the 45_H VEVs, while the other Higgs representations control the intermediate and weak scale stages. The results presented here and in Ref. [15] do open the path towards a realistic nonsupersymmetric SO(10) unification. A detailed study of minimal setups will be the subject of a future work.

The paper is organized as follows. The study of the tree-level scalar potential and the related scalar mass spectrum are concisely reviewed in Secs. II and III. A detailed understanding of the mass textures is developed in Sec. IV in terms of a systematic discussion of the accidental global symmetries and the associated pseudo-Goldstone bosons. In Sec. V we calculate the relevant quantum corrections by means of the one-loop effective potential, and we prove the existence of the new vacua. The main results and the prospects for further developments are summarized in Sec. VI. Most of the technical aspects of the work are deferred to Appendices A, B, C, and D.
for the scenarios in which $B - L$ is broken for instance by a 126-dimensional $SO(10)$ tensor.

We shall therefore study the structure of the vacua of a $SO(10)$ Higgs potential with only the $45_H \oplus 16_H$ representation at play. Following the common convention, we define $16_H \equiv \chi$ and denote by $\chi^+$ and $\chi^-$ the multiplets transforming as positive and negative chirality components of the reducible 32-dimensional $SO(10)$ spinor representation, respectively. Similarly, we shall use the symbol $\Phi$ (or the derived $\phi$, c.f. Appendix A) for the adjoint Higgs representation $45_H$ (or its components in the natural basis).

The minimal $SO(10)$ GUT accommodates the SM matter in three copies of $SO(10)$ spinors $16'_L$, $(i = 1, 2, 3)$. The fermions (and their Yukawa interactions) do not play any role in the GUT-scale dynamics and will not be considered further (we assume the masses of the right-handed neutrinos to be small with respect to the unification scale). The detailed study of realistic Higgs and Yukawa sectors will be the subject of a forthcoming paper.

A. The tree-level Higgs potential

The most general renormalizable tree-level scalar potential which can be constructed out of $45_H$ and $16_H$ reads (see for instance Refs. [22,23]):

$$V_0 = V_\Phi + V_\chi + V_{\Phi \chi},$$  (3)

where, according to the notation in Appendix A,

$$V_\Phi = -\frac{\mu^2}{2} \text{Tr} \Phi^2 + \frac{a_1}{4} (\text{Tr} \Phi^2)^2 + \frac{a_2}{4} \text{Tr} \Phi^4,$$

$$V_\chi = -\frac{\nu^2}{2} \chi^+ \chi + \frac{\lambda_1}{4} (\chi^+ \chi)^2 + \frac{\lambda_2}{4} (\chi^+ \Gamma_j \chi^-)(\chi^+ \Gamma_j \chi^-),$$  (4)

and

$$V_{\Phi \chi} = \alpha (\chi^+ \chi) \text{Tr} \Phi^2 + \beta \chi^+ \Phi^2 \chi + \gamma \chi^+ \Phi \chi. $$  (5)

The mass terms and coupling constants above are real by Hermiticity. Linear and cubic $\Phi$ self-interactions are absent due the zero trace of the $SO(10)$ adjoint representation. For the sake of simplicity, all tensorial indices have been suppressed.

B. The symmetry breaking patterns

1. The SM singlets

There are in general three SM singlets in the $45_H \oplus 16_H$ representation of $SO(10)$. Using $(B - L)/2 \equiv X$ and labeling the field components according to $3_c, 2_L, 2_R 1_X$, the SM singlets reside in the $(1, 1, 1, 0)$ and $(1, 1, 2, +\frac{1}{2})$ submultiplets of $45_H$ and in the $(1, 1, 2, +\frac{1}{2})$ component of $16_H$. We denote their VEVs as

$$\langle (1, 1, 1, 0) \rangle = \omega_Y, \quad \langle (1, 1, 2, +\frac{1}{2}) \rangle = \omega_R,$$  (6)

where $\omega_{Y,R}$ are real and $\chi_R$ can be taken real by a phase redefinition of the $16_H$. Different VEV configurations trigger the spontaneous breakdown of the $SO(10)$ symmetry into a number of subgroups. Namely, for $\chi_R = 0$ one finds

$$\omega_R = 0, \quad \omega_Y \neq 0:\quad 3_c, 2_L, 2_R 1_X$$

$$\omega_R \neq 0, \quad \omega_Y = 0:\quad 4_c 2_L 1_R$$

$$\omega_R \neq 0, \quad \omega_Y \neq 0:\quad 3_c, 2_L 1_R 1_X$$

$$\omega_R = -\omega_Y \neq 0:\quad \text{flipped} 5', 1_Z$$

$$\omega_R = \omega_Y \neq 0:\quad \text{standard} 5 1_Z$$

with $5 1_Z$ and $5', 1_Z$ standing for the two different embedding of the $SU(5)$ subgroup into $SO(10)$, i.e. standard and “flipped”, respectively, (see the discussion at the end of the section).

When $\chi_R \neq 0$ all intermediate gauge symmetries are spontaneously broken down to the SM group, with the exception of the last case, which maintains the standard $SU(5)$ subgroup unbroken and will not be further considered.

The classification in Eq. (7) depends on the phase conventions used in the parametrization of the SM singlet subspace of $45_H \oplus 16_H$. The statement that $\omega_R = \omega_Y$ yields the standard $SU(5)$ vacuum while $\omega_R = -\omega_Y$ corresponds to the flipped setting defines a particular basis in this subspace (see Sec. II B 3). The consistency of any chosen framework is then verified against the corresponding Goldstone boson spectrum.

The decomposition of the $45_H$ and $16_H$ representations with respect to the relevant $SO(10)$ subgroups is detailed in Tables I and II.

2. The L-R chains

According to the analysis in Ref. [15], the potentially viable breaking chains fulfilling the basic gauge unification constraints (with a minimal $SO(10)$ Higgs sector) correspond to the settings:

$$\omega_Y \gg \omega_R \gg \chi_R: \quad SO(10) \rightarrow 3_c, 2_L, 2_R 1_X \rightarrow 3_c, 2_L 1_R 1_X \rightarrow 3_c, 2_L 1_Y$$  (8)

$$\omega_R \gg \omega_Y \gg \chi_R: \quad SO(10) \rightarrow 4_c 2_L 1_R \rightarrow 3, 2_L 1_R 1_X \rightarrow 3, 2_L 1_Y.$$  (9)
As remarked in [15], the cases $\chi_R \sim \omega_R$ or $\chi_R \sim \omega_Y$ lead to effective two-step $SO(10)$ breaking patterns with a nonminimal set of surviving scalars at the intermediate scale. On the other hand, a truly two-step setup can be recovered (with a minimal fine-tuning) by considering the cases where $\omega_R$ or $\omega_Y$ exactly vanish. Only the explicit study of the scalar potential determines which of the textures are allowed.

We have verified that in all cases the GUT threshold effects related to the relevant pseudo-Goldstone mass patterns obtained in the present analysis fully comply with the unification constraints in Ref. [15]. Furthermore, the lower bounds on the position of the $B-L$ scale are consistently increased, hence improving the prospects for a successful model building.

### 3. Standard $SU(5)$ versus flipped $SU(5)$

There are in general two distinct SM-compatible embeddings of $SU(5)$ into $SO(10)$ [24,25]. They differ in one generator of the $SU(5)$ Cartan algebra and therefore in the $U(1)_2$ cofactor.

In the “standard” $SU(5)$ embedding, the weak hypercharge operator $Y = T_R^{(5)} + T_X$ belongs to the $SU(5)$ algebra and the orthogonal Cartan generator $Z$ (obeying $[T_i, Z] = 0$ for all $T_i \in SU(5)$) is given by $Z = -4T_R^{(5)} + 6T_X$.

In the flipped $SU(5)'$ case, the right-handed isospin assignment of quark and leptons into the $SU(5)'$ multiplets is turned over so that the flipped hypercharge generator reads $Y' = -T_R^{(5)} + T_X$. Accordingly, the additional $U(1)'_Y$ generator reads $Z' = 4T_R^{(5)} + 6T_X$, such that $[T_i, Z'] = 0$ for all $T_i \in SU(5)'$. Weak hypercharge is then given by $Y = (Z'/Y')/5$.

Tables I and II show the standard and flipped $SU(5)$ decompositions of the spinorial and adjoint $SO(10)$ representations, respectively.

The two $SU(5)$ vacua in Eq. (7) differ by the texture of the adjoint representation VEVs: in the standard $SU(5)$ case they are aligned with the $Z$ operator while they match the $Z'$ structure in the flipped $SU(5)'$ setting (see Appendix A4 for an explicit representation).

### III. THE CLASSICAL VACUUM

#### A. The stationarity conditions

By substituting Eq. (6) into Eq. (3) the vacuum manifold reads

<table>
<thead>
<tr>
<th>$4C_{2L}2_R$</th>
<th>$4C_{2L}1_R$</th>
<th>$3C_{2L}2_R1_X$</th>
<th>$3C_{2L}1_R1_X$</th>
<th>$3C_{2L}1_Y$</th>
<th>$5$</th>
<th>$5'1_{X'}$</th>
<th>$1_{Y'}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 1, 3)</td>
<td>(1, 1, +1)</td>
<td>(1, 1, +1, 0)</td>
<td>(1, 1, +1)</td>
<td>(1, 1, 0)</td>
<td>10</td>
<td>(10, -4)</td>
<td>+1</td>
</tr>
<tr>
<td></td>
<td>(1, 1, 0)</td>
<td>(1, 1, 0, 0)</td>
<td>(1, 1, 0)</td>
<td>(1, 1, 0)</td>
<td>1</td>
<td>(0, 0)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(1, 1, -1)</td>
<td>(1, 1, -1, 0)</td>
<td>(1, 1, -1)</td>
<td>(1, 1, -1)</td>
<td>10</td>
<td>(10, -4)</td>
<td>-1</td>
</tr>
<tr>
<td>(1, 3, 1)</td>
<td>(1, 3, 0)</td>
<td>(1, 3, 0, 0)</td>
<td>(1, 3, 0)</td>
<td>(1, 3, 0)</td>
<td>24</td>
<td>(24, 0)</td>
<td>0</td>
</tr>
<tr>
<td>(6, 2, 2)</td>
<td>(6, 2, +½)</td>
<td>(3, 2, 2, -½)</td>
<td>(3, 2, +½)</td>
<td>(3, 2, +½)</td>
<td>10</td>
<td>(24, 0)</td>
<td>-½</td>
</tr>
<tr>
<td></td>
<td>(6, 2, -½)</td>
<td>(3, 2, -½, -½)</td>
<td>(3, 2, -½)</td>
<td>(3, 2, -½)</td>
<td>24</td>
<td>(10, -4)</td>
<td>+½</td>
</tr>
<tr>
<td>(15, 1, 1)</td>
<td>(15, 1, 0)</td>
<td>(1, 1, 0, 0)</td>
<td>(1, 1, 0)</td>
<td>(1, 1, 0)</td>
<td>24</td>
<td>(24, 0)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(3, 1, 1, +½)</td>
<td>(3, 1, 0, +½)</td>
<td>(3, 1, +½)</td>
<td>(3, 1, +½)</td>
<td>10</td>
<td>(10, -4)</td>
<td>-½</td>
</tr>
<tr>
<td></td>
<td>(3, 1, 1, -½)</td>
<td>(3, 1, 0, -½)</td>
<td>(3, 1, -½)</td>
<td>(3, 1, -½)</td>
<td>10</td>
<td>(10, -4)</td>
<td>-2½</td>
</tr>
</tbody>
</table>
\[ V_0 = -2\mu^2(2\omega_\rho^2 + 3\omega_Y^2) + 4a_1(2\omega_\rho^2 + 3\omega_Y^2)^2 \]
\[ + \frac{a_2}{4}(8\omega_\rho^4 + 21\omega_Y^2 + 36\omega_\rho^2\omega_Y^2) - \frac{\nu^2}{2} \chi_R^2 \]
\[ + \frac{\lambda_1}{4} \chi_R^4 + 4\alpha \chi_R^2(2\omega_\rho^2 + 3\omega_Y^2) \]
\[ + \frac{\beta}{4} \chi_R^2(2\omega_\rho + 3\omega_Y)^2 - \frac{\tau}{2} \chi_R^2(2\omega_\rho + 3\omega_Y). \]  

The corresponding three stationary conditions can be conveniently written as
\[ \frac{1}{8} \left( \frac{\partial V_0}{\partial \omega_R} - \frac{2}{3} \frac{\partial V_0}{\partial \omega_Y} \right) = \left[ -\mu^2 + 4a_1(2\omega_\rho^2 + 3\omega_Y^2) \right. \]
\[ + \frac{a_2}{4}(4\omega_\rho^2 + 7\omega_Y^2 - 2\omega_\rho\omega_Y) \]
\[ \left. + 2\alpha \chi_R^2 \right] (\omega_R - \omega_Y) = 0 \]
\[ (11) \]
\[ \omega_Y \frac{\partial V_0}{\partial \omega_R} - \omega_R \frac{2}{3} \frac{\partial V_0}{\partial \omega_Y} = \left[ -4a_2(\omega_R + \omega_Y) \omega_R \omega_Y \right. \]
\[ + \beta \chi_R^2(2\omega_\rho + 3\omega_Y) \]
\[ \left. - \tau \chi_R^2(\omega_R - \omega_Y) \right] = 0 \]
\[ (12) \]
\[ \frac{\partial V_0}{\partial \chi_R} = \left[ -\nu^2 + \lambda_1 \chi_R^2 + 8\alpha(2\omega_\rho^2 + 3\omega_Y^2) \right. \]
\[ + \frac{\beta}{2}(2\omega_\rho + 3\omega_Y)^2 - \tau(2\omega_\rho + 3\omega_Y) \left. \right] \chi_R = 0. \]

We have chosen linear combinations that factor out the uninteresting standard SU(5) \( \otimes U(1)_L \) solution, namely \( \omega_R = \omega_Y \).

In summary, when \( \chi_R = 0 \), Eqs. (11) and (12) allow for four possible vacua:

(i) \( \omega = \omega_R = \omega_Y \) (standard 512)

(ii) \( \omega = \omega_R = -\omega_Y \) (flipped 5’1’2)

(iii) \( \omega_R = 0 \) and \( \omega_Y \neq 0 \) (32L2R1X)

(iv) \( \omega_R \neq 0 \) and \( \omega_Y = 0 \) (4C2L1R).

As we shall see, the last two options are not tree-level minima. Let us remark that for \( \chi_R \neq 0 \), Eq. (12) implies naturally a correlation among the 45_H and 16_H VEVs, or a fine-tuned relation between \( \beta \) and \( \tau \), depending on the stationary solution. In the cases \( \omega_R = -\omega_Y, \omega_R = 0 \) and \( \omega_Y = 0 \) one obtains \( \tau = \beta \omega, \tau = 3\beta \omega_Y \) and \( \tau = 2\beta \omega_R \) respectively. Consistency with the scalar mass spectrum must be verified in each case.

**B. The tree-level spectrum**

The gauge and scalar spectra corresponding to the SM vacuum configuration (with nonvanishing VEVs in 45_H \( \otimes 16_H \)) are detailed in Appendix C.

The scalar spectra obtained in various limits of the tree-level Higgs potential, corresponding to the appearance of accidental global symmetries, are derived in Appendices C 2 a, C 2 b, C 2 c, C 2 d, and C 2 e. The emblematic case \( \chi_R = 0 \) is scrutinized in Appendix C 2 f.

**C. Constraints on the potential parameters**

The parameters (couplings and VEVs) of the scalar potential are constrained by the requirements of boundedness and the absence of tachyonic states, ensuring that the vacuum is stable and the stationary points correspond to physical minima.

Necessary conditions for vacuum stability are derived in Appendix B. In particular, on the \( \chi_R = 0 \) section one obtains
\[ a_1 > -\frac{13}{80}a_2. \]

Considering the general case, the absence of tachyons in the scalar spectrum yields among else
\[ a_2 < 0, \quad -2 < \omega_Y/\omega_R < -\frac{1}{2}. \]

The strict constraint on \( \omega_Y/\omega_R \) is a consequence of the tightly correlated form of the tree-level masses of the (8, 1, 0) and (1, 3, 0) submultiplets of 45_H, labeled according to the SM (3, 2_L 1_Y) quantum numbers, namely
\[ M^2(1, 3, 0) = 2a_2(\omega_Y - \omega_R)(\omega_Y + 2\omega_R), \]
\[ M^2(8, 1, 0) = 2a_2(\omega_R - \omega_Y)(\omega_R + 2\omega_Y), \]
that are simultaneously positive only if Eq. (15) is enforced. For comparison with previous studies, let us remark that in the \( \tau = 0 \) limit (corresponding to an extra Z_2 symmetry \( \Phi \to -\Phi \)) the intersection of the constraints from Eq. (12), Eqs. (16) and (17) and the mass eigenvalues of the (1, 1, 1) and (3, 2, +1/6) states, yields
\[ a_2 < 0, \quad -1 \leq \omega_Y/\omega_R \leq -\frac{2}{3}, \]
thus recovering the results of Refs. [19–21].

In either case, one concludes by inspecting the scalar mass spectrum that flipped SU(5) \( \otimes U(1)_L \) is for \( \chi_R = 0 \) the only solution admitted by Eq. (12) consistent with the constraints in Eq. (15) (or Eq. (18)). For \( \chi_R \neq 0 \), the fine-tuned possibility of having or \( \omega_Y/\omega_R \sim -1 \) such that \( \chi_R \) is obtained at an intermediate scale fails to reproduce the SM couplings [15]. Analogous and obvious conclusions hold for \( \omega_Y \sim \omega_R \sim \chi_R \sim M_G \) and for \( \chi_R \gg \omega_{R,Y} \) (standard SU(5) in the first stage).

This is the origin of the common knowledge that non-supersymmetric SO(10) settings with the adjoint VEVs driving the gauge symmetry breaking are not phenomenologically viable. In particular, a large hierarchy between the 45_H VEVs, that would set the stage for consistent unification patterns, is excluded.
The key question is: why are the masses of the states in Eqs. (16) and (17) so tightly correlated? Equivalently, why do they depend on \( a_2 \) only?

**IV. UNDERSTANDING THE SCALAR SPECTRUM**

A detailed comprehension of the patterns in the scalar spectrum may be achieved by understanding the correlations between mass textures and the symmetries of the scalar potential. In particular, the appearance of accidental global symmetries in limiting cases may provide the rationale for the dependence of mass eigenvalues on specific couplings. To this end we classify the most interesting cases, providing a counting of the would-be Goldstone bosons (WGB) and pseudo-Goldstone bosons (PGB) for each case. A side benefit of this discussion is a consistency check of the explicit form of the mass spectra.

**A. 45 only with \( a_2 = 0 \)**

Let us first consider the potential generated by \( 45_H \), namely \( V_0 \) in Eq. (3). When \( a_2 = 0 \), i.e. when only trivial \( 45_H \) invariants (built off moduli) are considered, the scalar potential exhibits an enhanced global symmetry: \( O(45) \). The spontaneous symmetry breaking (SSB) triggered by the \( 45_H \) VEV reduces the global symmetry to \( O(44) \). As a consequence, 44 massless states are expected in the scalar spectrum. This is verified explicitly in Appendix C2a. Considering the case of the \( SO(10) \) gauge symmetry broken to the flipped \( SU(5)' \otimes U(1)_{X'} \), \( 45 - 25 = 20 \) WGB, with the quantum numbers of the coset \( SO(10)/SU(5)' \otimes \langle 1 \rangle_{X'} \) algebra, decouple from the physical spectrum, while \( 44 - 20 = 24 \) PGB remain, whose masses depend on the explicit breaking term \( a_2 \).

**B. 16 only with \( \lambda_2 = 0 \)**

We proceed in analogy with the previous discussion. Taking \( \lambda_2 = 0 \) in \( V_\chi \) enhances the global symmetry to \( O(32) \). The spontaneous breaking of \( O(32) \) to \( O(31) \) due to the \( 16_H \) VEV leads to 31 massless modes, as it is explicitly seen in Appendix C2b. Since the gauge \( SO(10) \) symmetry is broken by \( \chi_R \) to the standard \( SU(5) \), \( 45 - 24 = 21 \) WGB, with the quantum numbers of the coset \( SO(10)/SU(5) \) algebra, decouple from the physical spectrum, while \( 31 - 21 = 10 \) PGB do remain. Their masses depend on the explicit breaking term \( \lambda_2 \).

**C. A trivial 45-16 potential \( (a_2 = \lambda_2 = \beta = \tau = 0) \)**

When only trivial invariants (i.e. moduli) of both \( 45_H \) and \( 16_H \) are considered, the global symmetry of \( V_0 \) in Eq. (3) is \( O(45) \otimes O(32) \). This symmetry is spontaneously broken into \( O(44) \otimes O(31) \) by the \( 45_H \) and \( 16_H \) VEVs yielding \( 44 + 31 = 75 \) GB in the scalar spectrum (see Appendix C2d). Since in this case, the gauge \( SO(10) \) symmetry is broken to the SM gauge group, \( 45 - 12 = 33 \) WGB, with the quantum numbers of the coset \( SO(10)/SM \) algebra, decouple from the physical spectrum, while \( 75 - 33 = 42 \) PGB remain. Their masses are generally expected to receive contributions from the explicitly breaking terms \( a_2, \lambda_2, \beta \) and \( \tau \).

**D. A trivial 45-16 interaction \( (\beta = \tau = 0) \)**

Turning off just the \( \beta \) and \( \tau \) couplings still allows for independent global rotations of the \( \Phi \) and \( \chi \) Higgs fields. The largest global symmetries are those determined by the \( a_2 \) and \( \lambda_2 \) terms in \( V_0 \), namely \( O(10)_{a_5} \) and \( O(10)_{\lambda_6} \), respectively. Consider the spontaneous breaking to global flipped \( SU(5)' \otimes U(1)_{X'} \) and the standard \( SU(5) \) by the \( 45_H \) and \( 16_H \) VEVs, respectively. This setting gives rise to \( 20 + 21 = 41 \) massless scalars. The gauged \( SO(10) \) symmetry is broken to the SM group so that \( 33 \) WGB decouple from the physical spectrum. Therefore, \( 41 - 33 = 8 \) PGB remain, whose masses receive contributions from the explicit breaking terms \( \beta \) and \( \tau \). All of these features are readily verified by inspection of the scalar mass spectrum in Appendix C2e.

**E. A tree-level accident**

The tree-level masses of the crucial \( (1, 3, 0) \) and \( (8, 1, 0) \) multiplets belonging to the \( 45_H \) depend only on the parameter \( a_2 \), but not on the other parameters expected (c.f. Sec. IV C), namely \( \lambda_2, \beta \) and \( \tau \).

While the \( \lambda_2 \) and \( \tau \) terms cannot obviously contribute at the tree level to \( 45_H \) mass terms, one would generally expect a contribution from the \( \beta \) term, proportional to \( \chi_R^2 \). Using the parametrization \( \Phi = \sigma_{ij} \phi_{ij}/\sqrt{4} \), where the \( \sigma_{ij} \) (i, j \( \in \{1, \ldots, 10\} \)) matrices represent the \( SO(10) \) algebra on the 16-dimensional spinor basis (c.f. Appendix A), one obtains a \( 45_H \) mass term of the form

\[
\frac{\beta}{16} \chi_R^2 (\sigma_{ij})_{16}(\sigma_{kl})_{16}(\sigma_{ij})_{16}(\sigma_{kl})_{16}.
\]

The projection of the \( \phi_{ij} \) fields onto the \( (1, 3, 0) \) and \( (8, 1, 0) \) components lead, as we know, to vanishing contributions.

This result can actually be understood on general grounds by observing that the scalar interaction in Eq. (19) has the same structure as the gauge boson mass from the covariant-derivative interaction with the \( 16_H \), c.f. Eq. (C4). As a consequence, no tree-level mass contribution from the \( \beta \) coupling can be generated for the \( 45_H \) scalars carrying the quantum numbers of the standard \( SU(5) \) algebra.

This behavior can be again verified by inspecting the relevant scalar spectra in Appendix C2.

The above considerations provide a clear rationale for the accidental tree-level constraint on \( \omega_{Y}/\omega_{X} \), that holds independently on the size of \( \chi_R \).

On the other hand, we should expect the \( \beta \) and \( \tau \) interactions to contribute \( O(M_Z/4\pi) \) terms to the masses of \( (1, 3, 0) \) and \( (8, 1, 0) \) at the quantum level.
Similar contributions should also arise from the gauge interactions, that break explicitly the independent global transformations on the $45_U$ and $16_U$ discussed in the previous subsections.

The typical one-loop self-energies, proportional to the $45_U$ VEVs, are diagrammatically depicted in Fig. 1. While the exchange of $16_U$ components is crucial, the $\chi_R$ is not needed to obtain the large mass shifts. In the phenomenologically allowed unification patterns it gives actually negligible contributions.

It is interesting to notice that the $\tau$-induced mass corrections do not depend on the gauge symmetry breaking, yielding an $SO(10)$ symmetric contribution to all scalars in $45_U$.

One is thus led to the conclusion that any result based on the particular shape of the tree-level $45_U$ vacuum is drastically affected at the quantum level. Let us emphasize that although one may in principle avoid the $\tau$-term by means of, e.g., an extra $Z_2$ symmetry, no symmetry can forbid the $\beta$-term and the gauge loop contributions.

In case one resorts to $126_H$, in place of $16_H$, for the purpose of $B - L$ breaking, the more complex tensor structure of the $126_H 45^2_U 126_H$ quartic invariants in the scalar potential may admit tree-level contributions to the states $(1, 3, 0)$ and $(8, 1, 0)$ proportional to $\langle 126_H \rangle$. On the other hand, as mentioned above, whenever $\langle 126_H \rangle$ is small on the unification scale, the same considerations apply, as for the $16_H$ case.

**F. The $\chi_R = 0$ limit**

From the previous discussion it is clear that the answer to the question whether the non-$SU(5)$ vacua are allowed at the quantum level is independent on the specific value of the $B - L$ breaking VEV ($\chi_R \ll M_G$ in potentially realistic cases).

In order to simplify the study of the scalar potential beyond the classical level it is therefore convenient (and sufficient) to consider the $\chi_R = 0$ limit.

When $\chi_R = 0$ the mass matrices of the $45_U$ and $16_H$ sectors are not coupled. The stationary equations in Eqs. (11) and (12) lead to the four solutions

(i) $\omega = \omega_R = \omega_Y (5 \, 1_2^2)$

(ii) $\omega = \omega_R = - \omega_Y (5 \, 1_2^2)$

(iii) $\omega_R = 0$ and $\omega_Y \neq 0 (3, 2_1^2, 2_R^3)$

(iv) $\omega_R \neq 0$ and $\omega_Y = 0 (4_1^1, 2_1^2, 1_R^3)$

In what follows, we will focus our discussion on the last three cases only.

It is worth noting that the tree-level spectrum in the $\chi_R = 0$ limit is not directly obtained from the general formulas given in Appendix C 2 c, since Eq. (13) is trivially satisfied for $\chi_R = 0$. The corresponding scalar mass spectra are derived and discussed in Appendix C 2 f. Yet again, it is apparent that the non-$SU(5)$ vacuum configurations exhibit unavoidable tachyonic states in the scalar spectrum.

**V. THE QUANTUM VACUUM**

**A. The one-loop effective potential**

We shall compute the relevant one-loop corrections to the tree-level results by means of the one-loop effective potential (effective action at zero momentum) [26]. We can formally write

$$V_{\text{eff}} = V_0 + \Delta V_s + \Delta V_f + \Delta V_g,$$

where $V_0$ is the tree-level potential and $\Delta V_{s,f,g}$ denote the quantum contributions induced by scalars, fermions and gauge bosons, respectively. In dimensional regularization with the modified minimal subtraction (MS) and in the Landau gauge, they are given by

$$\Delta V_s(\phi, \chi, \mu) = \frac{\eta}{64 \pi^2} \text{Tr} \left[ W^4(\phi, \chi) \left( \log \frac{W^2(\phi, \chi)}{\mu^2} - \frac{3}{2} \right) \right],$$

$$\Delta V_f(\phi, \chi, \mu) = \frac{\kappa}{64 \pi^2} \text{Tr} \left[ M^4(\phi, \chi) \left( \log \frac{M^2(\phi, \chi)}{\mu^2} - \frac{3}{2} \right) \right].$$

**FIG. 1.** Typical one-loop diagrams that induce, for $\langle \chi \rangle = 0$, $O(\tau / 4 \pi, \beta(\phi) / 4 \pi, g^2(\phi) / 4 \pi)$ renormalization to the mass of $45_U$ fields at the unification scale. They are relevant for the PGB states, whose tree-level mass is proportional to $a_2$.**
\[
\Delta V_{g}(\phi, \chi, \mu) = \frac{3}{64\pi^{2}} \text{Tr} \left[ \mathcal{M}^{4}(\phi, \chi) \left( \log \frac{\mathcal{M}^{2}(\phi, \chi)}{\mu^{2}} - \frac{3}{2} \right) \right].
\]

(23)

with \( \eta = 1(2) \) for real (complex) scalars and \( \kappa = 2(4) \) for Weyl (Dirac) fermions. \( W, M \) and \( \mathcal{M} \) are the functional scalar, fermion and gauge boson mass matrices, respectively, as obtained from the tree-level potential.

In the case at hand, we may write the functional scalar mass matrix \( W^{2}(\phi, \chi) \) as a 77-dimensional Hermitian matrix, with a Lagrangian term

\[
\frac{1}{2} \psi^{\dagger} W^{2} \psi,
\]

(24)
defined on the vector basis \( \psi = (\phi, \chi, \chi^{\ast}) \). More explicitly, \( W^{2} \) takes the block form

\[
W^{2}(\phi, \chi) = \begin{pmatrix}
V_{\phi \phi} & V_{\phi \chi} & V_{\phi \chi^{\ast}} \\
V_{\chi \phi} & V_{\chi \chi} & V_{\chi \chi^{\ast}} \\
V_{\chi^{\ast} \phi} & V_{\chi^{\ast} \chi} & V_{\chi^{\ast} \chi^{\ast}}
\end{pmatrix},
\]

(25)

where the subscripts denote the derivatives of the scalar potential with respect to the set of fields \( \phi, \chi \) and \( \chi^{\ast} \). In the one-loop part of the effective potential \( V \equiv V_{0} \).

We neglect the fermionic component of the effective potential since there are no fermions at the GUT scale (we assume that the right-handed (RH) neutrino mass is substantially lower than the unification scale).

The functional gauge boson mass matrix, \( \mathcal{M}^{2}(\phi, \chi) \) is given in Appendix C, Eqs. (C3) and (C4).

### B. The one-loop stationary equations

The first derivative of the one-loop part of the effective potential, with respect to the scalar field component \( \psi_{a} \), reads

\[
\frac{\partial \Delta V_{\psi}}{\partial \psi_{a}} = \frac{1}{64\pi^{2}} \text{Tr} \left[ \{ W_{\psi_{a}}, W^{2} \} \left( \log \frac{W^{2}}{\mu^{2}} - \frac{3}{2} \right) + W^{2} W_{\psi_{a}} \right]
\]

(26)

where the symbol \( W_{\psi_{a}}^{2} \) stands for the partial derivative of \( W^{2} \) with respect to \( \psi_{a} \). Analogous formulas hold for \( \partial \Delta V_{\phi, \chi} / \partial \psi_{a} \). The trace properties ensure that Eq. (26) holds independently on whether \( W^{2} \) does commute with its first derivatives or not.

The calculation of the loop-corrected stationary equations due to gauge bosons and scalar exchange is straightforward (for \( \chi_{R} = 0 \) the 45\(_{L} \) and 16\(_{H} \) blocks decouple in Eq. (25)). On the other hand, the corrected equations are quite cumbersome, and we do not explicitly report them here. It is enough to say that the quantum analogue of Eq. (12) admits analytically the same solutions as we had at the tree level. Namely, these are \( \omega_{R} = \omega_{V} \), \( \omega_{R} = -\omega_{V} \), \( \omega_{R} = 0 \) and \( \omega_{V} = 0 \), corresponding, respectively, to the standard 51\(_{L} \), flipped 51\(_{V} \), 3_c \( 2_{L} \) \( 2_{R} \) \( 1_{X} \) and 4_c \( 2_{L} \) \( 1_{R} \) preserved subalgebras.

### C. The one-loop scalar mass

In order to calculate the second derivatives of the one-loop contributions to \( V_{\text{eff}} \) it is in general necessary to take into account the commutation properties of \( W^{2} \) with its derivatives that enter as a series of nested commutators.

The general expression can be written as

\[
\frac{\partial^{2} \Delta V_{\psi}}{\partial \psi_{a} \partial \psi_{b}} = \frac{1}{64\pi^{2}} \text{Tr} \left[ W_{\psi_{a}}^{2} W^{2} + W^{2} W_{\psi_{b}}^{2} \right] \left( \log \frac{W^{2}}{\mu^{2}} - \frac{3}{2} \right) + \sum_{m=1}^{\infty} \frac{1}{m} \sum_{k=1}^{m} \left( \begin{array}{c}
m \\
k
\end{array} \right) \left[ W^{2}, W_{\psi_{a}}^{2} \right]_{-1} \left[ W^{2}, W_{\psi_{b}}^{2} \right]_{-1} \left[ W^{2} \right]_{-1} \left[ W^{2} \right]_{-1}
\]

(27)

where the commutators in the last line are taken \( k - 1 \) times. Let us also remark that, although not apparent, the RHS of Eq. (27) can be shown to be symmetric under \( a \leftrightarrow b \), as it should be. In specific cases (for instance when the nested commutators vanish or they can be rewritten as powers of a certain matrix commuting with \( W \)) the functional mass evaluated on the vacuum may take a closed form.

#### 1. Running and pole mass

The effective potential is a functional computed at zero external momenta. Whereas the stationary equations allow for the localization of the new minimum (provided the VEVs are translationally invariant), the mass shifts obtained from Eq. (27) define the running masses \( m_{ab}^{2} \)

\[
\tilde{m}_{ab}^{2} \equiv \left. \frac{\partial^{2} V_{\text{eff}}(\phi)}{\partial \psi_{a} \partial \psi_{b}} \right|_{(\phi)} = m_{ab}^{2} + \Sigma_{ab}(0),
\]

(28)

where \( m_{ab}^{2} \) are the renormalized masses and \( \Sigma_{ab}(p^{2}) \) are the \( \overline{\text{MS}} \) renormalized self-energies. The physical (pole) masses \( M_{ab}^{2} \) are then obtained as a solution to the equation

\[
\det[p^{2} \delta_{ab} - (\tilde{m}_{ab}^{2} + \Delta \Sigma_{ab}(p^{2}))] = 0,
\]

(29)

where

\[
\Delta \Sigma_{ab}(p^{2}) = \Sigma_{ab}(p^{2}) - \Sigma_{ab}(0).
\]

(30)

For a given eigenvalue,

\[
M_{a}^{2} = \tilde{m}_{a}^{2} + \Delta \Sigma_{a}(M_{a}^{2})
\]

(31)

gives the physical mass. The gauge and scheme dependence in Eq. (28) is canceled by the relevant contributions from Eq. (30). In particular, infrared divergent terms in Eq. (28) related to the presence of massless WGB in the Landau gauge cancel in Eq. (31).

Of particular relevance is the case when \( M_{a} \) is substantially smaller than the (GUT-scale) mass of the particles that contribute to \( \Sigma(0) \). At \( \mu = M_{G} \), in the \( M_{a}^{2} \ll M_{G}^{2} \)
In this case the running mass computed from Eq. (28) contains the leading gauge independent corrections. As a matter of fact, in order to study the vacua of the potential in Eq. (20), we need to compute the zero momentum mass corrections just to those states that are tachyonic at the tree level and whose corrected mass turns out to be of the order of $M_G/4\pi$.

We may safely neglect the one-loop corrections for all other states with masses of order $M_G$. It is remarkable, as we shall see, that for $\chi_R = 0$ the relevant corrections to the masses of the critical PGB states can be obtained from Eq. (27) with vanishing commutators.

D. One-loop PGB masses

The stringent tree-level constraint on the ratio $\omega_Y/\omega_R$, coming from the positivity of the $(1, 3, 0)$ and $(8, 1, 0)$ masses, follows from the fact that some scalar masses depend only on the parameter $a_2$. On the other hand, the discussion on the would-be global symmetries of the scalar potential shows that in general their mass should depend on other terms in the scalar potential, in particular $\tau$ and $\beta$.

A set of typical one-loop diagrams contributing $O((\phi)/4\pi)$ renormalization to the masses of $45_H$ states is depicted in Fig. 1. As we already pointed out the $16_H$ VEV does not play any role in the leading GUT-scale corrections (just the interaction between $45_H$ and $16_H$, or with the massive gauge bosons is needed). Therefore we henceforth work in the strict $\chi_R = 0$ limit, that simplifies substantially the calculation. In this limit the scalar mass matrix in Eq. (25) is block diagonal (c.f. Appendix C 2f) and the leading corrections from the one-loop effective potential are encoded in the $V_{\chi' \chi}$ sector.

More precisely, we are interested in the corrections to those $45_H$ scalar states whose tree-level mass depends only on $a_2$ and have the quantum numbers of the preserved non-Abelian algebra (see Sec. IVA and Appendix C 2f). It turns out that focusing to this set of PGB states the functional mass matrix $W^2$ and its first derivative do commute for $\chi_R = 0$ and Eq. (27) simplifies accordingly. This allows us to compute the relevant mass corrections in a closed form.

The calculation of the EP running mass from Eq. (27) leads for the states $(1, 3, 0)$ and $(8, 1, 0)$ at $\mu = M_G$ to the mass shifts

$$\Delta M^2(1, 3, 0) = \frac{1}{4\pi^2}[\tau^2 + \beta^2(2\omega_R^2 - \omega_R\omega_Y + 2\omega_Y^2) + g^4(16\omega_R^2 + \omega_Y\omega_R + 19\omega_Y^2)],$$

where the subleading (and gauge dependent) logarithmic terms are not explicitly reported. For the vacuum configurations of interest we find the results reported in Appendix D. In particular, we obtain

(i) $\omega = \omega_R = -\omega_Y (5' 1_x)$:

$$M^2(24, 0) = -4a_2\omega_Y^2 + \frac{\tau^2 + (5\beta^2 + 34g^4)\omega_Y^2}{4\pi^2},$$

(ii) $\omega_R = 0$ and $\omega_Y \neq 0 (3_c 2_l 2_R 1_x)$:

$$M^2(1, 3, 0) = M^2(1, 1, 3, 0) = 2a_2\omega_Y^2 + \frac{\tau^2 + (2\beta^2 + 19g^4)\omega_Y^2}{4\pi^2},$$

(iii) $\omega_R \neq 0$ and $\omega_Y = 0 (4_c 2_l 1_R)$:

$$M^2(1, 3, 0) = -4a_2\omega_R^2 + \frac{\tau^2 + (2\beta^2 + 16g^4)\omega_Y^2}{4\pi^2},$$

In the effective theory language Eqs. (35)–(39) can be interpreted as the one-loop GUT-scale matching due to the decoupling of the massive $SO(10)/G$ states where $G$ is the preserved gauge group. These are the only relevant one-loop corrections needed in order to discuss the vacuum structure of the model.

It is quite apparent that a consistent scalar mass spectrum can be obtained in all cases, at variance with the tree-level result.

In order to fully establish the existence of the non-$SU(5)$ minima at the quantum level one should identify the regions of the parameter space supporting the desired vacuum configurations and estimate their depths. We shall address these issues in the next section.
E. The one-loop vacuum structure

1. Existence of the new vacuum configurations

The existence of the different minima of the one-loop effective potential is related to the values of the parameters $a_2$, $\beta$, $\tau$ and $g$ at the scale $\mu = M_G$. For the flipped $5'/1_x'$ case it is sufficient, as one expects, to assume the tree-level condition $a_2 < 0$. On the other hand, from Eqs. (36)–(39) we obtain

(i) $\omega_R = 0$ and $\omega_Y \neq 0$ ($3c_2L_2R_1\chi$):

$$- 8\pi^2 a_2 < \frac{\tau^2}{\omega_Y^2} + 2\beta^2 + 19g^4,$$

(ii) $\omega_R \neq 0$ and $\omega_Y = 0$ ($4c_2L_1\chi$):

$$- 8\pi^2 a_2 < \frac{\tau^2}{\omega_R^2} + \beta^2 + 13g^4.$$

Considering for naturalness $\tau \sim \omega_Y$, Eqs. (40) and (41) imply $|a_2| < 10^{-2}$. This constraint remains within the natural perturbative range for dimensionless couplings. While all PGB states whose mass is proportional to $-a_2$ receive large positive loop corrections, quantum corrections are numerically irrelevant for all of the states with GUT-scale mass. On the same grounds we may safely neglect the multiplicative $a_2$ loop corrections induced by the 45$_H$ states on the PGB masses.

2. Absolute minimum

It remains to show that the non-$SU(5)$ solutions may actually be absolute minima of the potential. To this end it is necessary to consider the one-loop corrected stationary equations and calculate the vacuum energies in the relevant cases. Studying the shape of the one-loop effective potential is a numerical task. On the other hand, in the approximation of neglecting at the GUT scale the logarithmic corrections, we may reach nondetailed but definite conclusions. For the three relevant vacuum configurations we obtain:

(i) $\omega = \omega_R = -\omega_Y$ ($5'/1_x'$)

$$V(\omega, \chi_R = 0) = -\frac{3\nu^4}{16\pi^2} + \frac{5\alpha\nu^2}{\pi^2} + \frac{5\beta\nu^2}{16\pi^2} + \frac{5\tau^2}{16\pi^2}$$

$$+ \left(-100a_1 - \frac{55a_2}{4} + \frac{600a_2^2}{\pi^2}\right)$$

$$+ \frac{45a_1a_2}{\pi^2} - \frac{645a_2^2}{32\pi^2} + \frac{100a_2^2}{\pi^2}$$

$$+ \frac{25\alpha\beta}{2\pi^2} + \frac{65\beta^2}{64\pi^2} - \frac{5g^4}{2\pi^2},$$

(ii) $\omega_R = 0$ and $\omega_Y \neq 0$ ($3c_2L_2R_1\chi$)

$$V(\omega_Y, \chi_R = 0) = -\frac{3\nu^4}{16\pi^2} + \left(\frac{3\alpha\nu^2}{\pi^2} + \frac{3\beta\nu^2}{16\pi^2} - \frac{3\tau^2}{16\pi^2}\right)$$

$$+ \left(-36a_1 - \frac{21a_2}{4} + \frac{216a_2^2}{\pi^2}\right)$$

$$+ \frac{33a_1a_2}{\pi^2} + \frac{45a_2^2}{32\pi^2} + \frac{36\alpha^2}{\pi^2}$$

$$+ \frac{9\alpha\beta}{2\pi^2} + \frac{21\beta^2}{64\pi^2} - \frac{15g^4}{16\pi^2}.$$

(iii) $\omega_R \neq 0$ and $\omega_Y = 0$ ($4c_2L_1\chi$)

$$V(\omega_R, \chi_R = 0) = -\frac{3\nu^4}{16\pi^2} + \left(\frac{2\alpha\nu^2}{\pi^2} + \frac{\beta\nu^2}{8\pi^2} - \frac{\tau^2}{8\pi^2}\right)$$

$$+ \left(-16a_1 - 2a_2 + \frac{96a_2^2}{\pi^2}\right)$$

$$+ \frac{42a_1a_2}{\pi^2} + \frac{147a_2^2}{32\pi^2} + \frac{16a_2^2}{\pi^2}$$

$$+ \frac{2\alpha\beta}{\pi^2} + \frac{\beta^2}{8\pi^2} - \frac{7g^4}{16\pi^2}.$$ 

A simple numerical analysis reveals that for natural values of the dimensionless couplings and GUT mass parameters any of the qualitatively different vacuum configurations may be a global minimum of the one-loop effective potential in a large domain of the parameter space.

This concludes the proof of existence of all of the group-theoretically allowed vacua. Nonsupersymmetric $SO(10)$ models broken at $M_G$ by the 45$_H$ SM-preserving VEVs do exhibit at the quantum level the full spectrum of intermediate symmetries. This is crucially relevant for those chains that, allowed by gauge unification, are accidentally excluded by the tree-level potential.

F. The extended survival hypothesis

In a realistic $SO(10)$ unification setup, throughout all the stages of the symmetry breaking one usually assumes that the scalar spectrum obeys the so-called extended survival hypothesis (ESH) that reads [27]: “at every stage of the symmetry breaking chain only those scalars are present that develop a VEV at the current or the subsequent levels of the spontaneous symmetry breaking.”
The ESH is equivalent to performing the minimal number of fine-tunings imposed onto the scalar potential so that all the symmetry breaking steps are obtained at the desired scales. On the technical side one must identify all the Higgs multiplets needed by the breaking pattern and tune their masses according to the gauge symmetry down to the scale of their VEVs. The effects of the presence of these states at intermediate scales has been considered in our recent analysis of nonsupersymmetric SO(10) unification patterns [15], up to one exception that we shall now shortly comment upon.

The relevant patterns preserve at the first stage the group $3, 2_L, 2_R \chi_X$ (for $\omega_R = 0$) and $4_c, 2_L \chi_R$ (for $\omega_Y = 0$). The breaking to the SM gauge group $3_c, 2_L 1_Y$ is achieved by means of the VEV $\chi_R$, constrained to stay at an intermediate scale by gauge unification. Minimally, one must therefore maintain at this scale either of the $16_H$ multiplets $(1, 1, 2, +\frac{1}{2})$ and $(4, 1, -\frac{1}{2})$, in the $3, 2_L, 2_R \chi_X$ and $4_c, 2_L \chi_R$ cases, respectively.

As one can see from the scalar spectrum given in Appendix C2f, in the $3, 2_L, 2_R \chi_X$ vacuum, the scalars $(1, 2, 1, -\frac{1}{2})$ and $(1, 1, 2, +\frac{1}{2})$ receive a mass contribution that is linear in the D-odd VEV $\omega_Y$ and that breaks their degeneracy. Thus, just the RH doublet $(1, 1, 2, +\frac{1}{2})$, which contains the field acquiring the VEV $\chi_R$, may be minimally fine-tuned at that mass scale.

Turning on $\omega_Y \neq 0$ or $\omega_R \neq 0$ at the $\chi_R$ scale leads to a nonminimal set of Higgs states at the intermediate scale [15], namely, the $45_H$ multiplets $(1, 1, 3, 0)$ and $(15, 1, 0)$ in the $3, 2_L, 2_R \chi_X$ and in the $3, 2_L, 2_R \chi_R$ setting respectively (these are the accidentally tachyonic states at the tree level). Inspection of the one-loop mass spectra shows that the needed minimal fine-tuning can be indeed performed.

It is worth noticing that although in the $3, 2_L, 2_R \chi_X$ stage D-parity is broken by $\omega_Y$, the masses of the states $(1, 1, 3, 0)$ and $(1, 1, 3, 0)$, depending quadratically on $\omega_Y$ (see Appendix C2f), do remain degenerate and are both tuned at the scale $\omega_R \sim \chi_R$, where the LR symmetry is broken. The presence of the additional LH triplet $(1, 3, 1, 0)$ at the intermediate scale $\omega_R$ has a welcome impact on the gauge coupling running. Compared to the results given in [15] for such a breaking pattern, the intermediate $B - L$ scale is raised by almost 1 order of magnitude (to about $10^{11}$ GeV), while the GUT scale is slightly lowered to about $10^{16}$ GeV. Detailed threshold effects can be considered once the model dependent scalar spectrum is fully worked out.

A final comment is in order. All of the states exchanged in the relevant mass loop corrections in Sec. VD have natural GUT masses. On the other hand, the ESH requires tuning the masses of some of these states at a much lower scale. In a realistic setting, this involves some of the $16_H$ submultiplets. The fine-tuning apparently generates an infrared divergence problem in the one-loop corrections. However, in analogy to our discussion of the WGB contributions to the effective potential in Sec. VC, the infrared divergent terms appearing in the one-loop zero momentum mass corrections disappear when considering the corrections to the physical pole masses. Thus, they can be safely discarded.

VI. SUMMARY AND OUTLOOK

In this paper, we have scrutinized the long-standing result that the class of the minimal nonsupersymmetric \( \text{SO}(10) \) unified models, with the GUT symmetry broken by the VEVs of the 45-dimensional adjoint representation, cannot provide a successful gauge unification. This common knowledge was based on the observation that the tree-level dynamics of the minimal scalar sector allowed only for \( "\text{SU}(5)" \) breaking patterns. This, in turn, clashes with the intermediate symmetry patterns required by nonsupersymmetric \( \text{SO}(10) \) unification, that enforce an intermediate threshold well below the GUT scale.

We argued that the old result is an artifact of the tree-level Higgs potential and showed that quantum corrections have a dramatic impact. The minimization of the one-loop effective potential in the paradigmatic $\chi_R = 0$ limit shows that the simplest \( \text{SO}(10) \) model with a $45_H \otimes 16_H$ Higgs sector allows for any of the intermediate symmetry patterns available to the pair of the SM-preserving VEVs in $45_H$. In particular, the $\text{SU}(4)_C \otimes \text{SU}(2)_L \otimes U(1)_R$ and $\text{SU}(3)_C \otimes \text{SU}(2)_L \otimes \text{SU}(2)_R \otimes U(1)_{B-L}$ chains are supported. Our result generally applies to any Higgs sector where the vacuum is dominated by the $45_H$ VEVs.

This observation opens the option of reconsidering the minimal nonsupersymmetric \( \text{SO}(10) \) model as a reference framework for model building. Extending the Higgs sector to include one $10_H$ (together with either one $126_H$ or one $16_H$) provides the playground for exploring the possibility of a realistic and predictive GUT, along the lines of the recent efforts in the supersymmetric context.

ACKNOWLEDGMENTS

S.B. acknowledges support by the MRTN European Program MRTN-CT-2006-035863. The work of M. M. is supported by the Royal Institute of Technology (KTH), Contract No. SII-56510.

APPENDIX A: \( \text{SO}(10) \) ALGEBRA REPRESENTATIONS

We briefly recall here for convenience the basics of \( \text{SO}(10) \) algebra representations. For a general introduction see Refs. [28,29].

1. Tensorial representations

The Hermitian and antisymmetric generators of the fundamental representation of \( \text{SO}(10) \) are given by
where \( a, b, i, j = 1, \ldots, 10 \) and the square bracket stands for antisymmetrization. They satisfy the \( SO(10) \) commutation relations

\[
[\epsilon_{ij}, \epsilon_{kl}] = -i(\delta_{jk}\epsilon_{il} - \delta_{ik}\epsilon_{jl} - \delta_{jl}\epsilon_{ik} + \delta_{il}\epsilon_{jk}),
\]

(A2) with normalization

\[
\text{Tr} \, \epsilon_{ij} \epsilon_{kl} = 2\delta_{[kl}\delta_{j]}.
\]

(A3) and Dynkin index 2.

The fundamental (vector) representation \( \phi_a (a = 1, \ldots, 10) \) transforms as

\[
\phi_a \rightarrow \phi_a - \frac{i}{2} \lambda_{ij} (\epsilon_{ij} \phi)_a,
\]

(A4) where \( \lambda_{ij} \) are the infinitesimal parameters of the transformation.

The adjoint representation is then obtained as the antisymmetric part of the 2-index \( 10_a \otimes 10_b \) tensor \( \phi_{ab} \) \((a, b = 1, \ldots, 10)\) and transforms as

\[
\phi_{ab} \rightarrow \phi_{ab} - \frac{i}{2} \lambda_{ij} [\epsilon_{ij}, \phi]_{ab}.
\]

(A5) Notice that \( [\epsilon_{ij}, \phi]^T = -[\epsilon_{ij}, \phi] \) and \( [\epsilon_{ij}, \phi]^T = [\epsilon_{ij}, \phi] \).

2. Spinorial representations

Following the notation of Ref. [21], the \( SO(10) \) generators \( S_{ij} \) \((i, j = 0, \ldots, 9)\) acting on the 32-dimensional spinor \( \Xi \) are defined as

\[
S_{ij} = \frac{1}{4!} [\Gamma_p, \Gamma_{ij}], \quad [\Gamma_p, \Gamma_{ij}] = 2\delta_{ij},
\]

(A6) with an explicit representation given by

\[
\Gamma_0 = \begin{pmatrix} 0 & I_{16} \\ I_{16} & 0 \end{pmatrix}, \quad \Gamma_p = \begin{pmatrix} 0 & is_p \\ -is_p & 0 \end{pmatrix}, \quad p = 1, \ldots, 9,
\]

(A7) where the \( s_p \) matrices are defined as \((k = 1, \ldots, 3)\)

\[
s_k = \eta_k \rho_3, \quad s_{k+3} = \sigma_k \rho_1, \quad s_{k+6} = \tau_k \rho_2.
\]

(A8) The matrices \( \sigma_k, \tau_k, \eta_k \), and \( \rho_k \), are given by the following tensor products of 2 \( \times \) 2 matrices:

\[
\sigma_k = I_2 \otimes I_2 \otimes I_2 \otimes \Sigma_k, \quad \tau_k = I_2 \otimes I_2 \otimes \Sigma_k \otimes I_2,
\]

\[
\eta_k = I_2 \otimes \Sigma_k \otimes I_2 \otimes I_2, \quad \rho_k = \Sigma_k \otimes I_2 \otimes I_2 \otimes I_2,
\]

(A9) where \( \Sigma_k \) stand for the ordinary Pauli matrices. Defining

\[
s_{pq} = \frac{1}{2i} [s_p, s_q]
\]

(A10) for \( p, q = 1, \ldots, 9 \), the algebra (A6) is represented by

\[
S_{p0} = \frac{1}{2} \begin{pmatrix} s_p & 0 \\ 0 & -s_p \end{pmatrix}, \quad S_{pq} = \frac{1}{2} \begin{pmatrix} s_{pq} & 0 \\ 0 & s_{pq} \end{pmatrix}.
\]

(A11)

The Cartan subalgebra is spanned over \( S_{03}, S_{12}, S_{45}, S_{78} \) and \( S_{69} \). One can construct a chiral projector \( \Gamma_\chi \), that splits the 32-dimensional spinor \( \Xi \) into a pair of irreducible 16-dimensional components:

\[
\Gamma_\chi = 2^{-5} S_{03} S_{12} S_{45} S_{78} S_{69} = \begin{pmatrix} 0 & I_{16} \\ I_{16} & 0 \end{pmatrix}.
\]

(A12) It is readily verified that \( \Gamma_\chi \) has the following properties:

\[
\Gamma_\chi^2 = I_{32}, \quad [\Gamma_{\chi}, \Gamma_i] = 0 \quad \text{and hence} \quad [\Gamma_{\chi}, S_{ij}] = 0.
\]

Introducing the chiral projectors \( P_\pm = \frac{1}{2} (I_{32} \pm \Gamma_\chi) \), the irreducible chiral spinors are defined as

\[
\chi^+ = P_+ \Xi \equiv \begin{pmatrix} \chi \\ 0 \end{pmatrix}, \quad \chi^- = P_- \Xi \equiv \begin{pmatrix} 0 \\ \chi^c \end{pmatrix},
\]

(A13) where \( \chi^c = C \chi^* \) and \( C \) is the \( SO(10) \) charge conjugation matrix (see next subsection). Analogously, we can use the chiral projectors to write \( S_{ij} \) as

\[
S_{ij} = P_+ S_{ij} P_+ + P_- S_{ij} P_- = \frac{1}{2} \begin{pmatrix} \sigma_{ij} & 0 \\ 0 & \tilde{\sigma}_{ij} \end{pmatrix},
\]

(A14) where the properties \([P_\pm, S_{ij}] = 0, P_{\pm}^2 = P_\pm \) and \( P_+ + P_- = I_{32} \) were used.

Finally, matching Eq. (A14) with Eq. (A11), one identifies the Hermitian generators \( \sigma_{ij}/2 \) and \( \tilde{\sigma}_{ij}/2 \) acting on the \( \chi \) and \( \chi^c \) spinors, respectively, as

\[
\sigma_{p0} = s_p, \quad \sigma_{pq} = s_{pq}, \quad \tilde{\sigma}_{p0} = -s_p, \quad \tilde{\sigma}_{pq} = s_{pq}.
\]

(A15)

From their normalization

\[
\frac{1}{4} \text{Tr} \sigma_{ij} \sigma_{kl} = \frac{1}{4} \text{Tr} \tilde{\sigma}_{ij} \tilde{\sigma}_{kl} = 4\delta_{ijkl}\delta_{jl},
\]

(A16) we recover the Dynkin index 4 of the 16-dimensional spinorial representation.

It is convenient to trace out the \( \sigma \)-matrices in the invariants built off the adjoint representation in the natural basis \( \Phi \equiv \sigma_{ij} \phi_{ij}/4 \). From the traces of two and four \( \sigma \)-matrices one obtains

\[
\text{Tr} \, \Phi^2 = -2 \text{Tr} \phi^2,
\]

(A17)

\[
\text{Tr} \, \Phi^4 = \frac{3}{4} (\text{Tr} \phi^2)^2 - \text{Tr} \phi^4.
\]

(A18)

In order to maintain a consistent notation, from now on we shall label the indices of the spinorial generators from 1 to 10, and use the following mapping from the basis of Ref. [21] into the basis of Ref. [7] for vectors and tensors:

\[
\{0312457869\} \rightarrow \{12345678910\}.
\]

3. The charge conjugation \( C \)

According to the notation of the previous subsection, the spinor \( \chi \) and its complex conjugate \( \chi^c \) transform as
\[ \chi \rightarrow \chi - \frac{i}{4} \lambda_{ij} \sigma_{ij} \chi, \quad \chi^* \rightarrow \chi^* + \frac{i}{4} \lambda_{ij} \sigma_{ij}^T \chi^*. \]  

(A19)

The charge conjugated spinor \( \chi^c = C \chi^* \) obeys

\[ \chi^c \rightarrow \chi^c - \frac{i}{4} \lambda_{ij} \tilde{\sigma}_{ij} \chi^c, \]  

(A20)

and thus \( C \) satisfies

\[ C^{-1} \tilde{\sigma}_{ij} C = - \sigma_{ij}^T. \]  

(A21)

Taking into account Eq. (A9), a formal solution reads

\[ C = \sigma_{2} \tau_{2} \eta_{2} \rho_{2}, \]  

(A22)

which in our basis yields

\[ C = \text{antidiag}(+1, -1, -1, +1, +1, +1, -1, -1, +1, +1), \]  

(A23)

and hence \( C = C^* = C^{-1} = C^T = C^\dagger \).

### 4. The Cartan generators

It is convenient to write the five \( SO(10) \) Cartan generators in the \( 3 \times 2 \times 2 \times 1 \) basis \( (X = (B - L)/2) \), where the physical interpretation is obvious. For the spinorial representation we have

\[ T_R^{(3)} = \frac{1}{4}(\sigma_{12} + \sigma_{34}), \quad \tilde{T}_R^{(3)} = \frac{1}{4}(-\sigma_{12} + \sigma_{34}), \]  

\[ T_L^{(3)} = \frac{1}{4}(\sigma_{34} - \sigma_{12}), \quad \tilde{T}_L^{(3)} = \frac{1}{4}(\sigma_{34} + \sigma_{12}), \]  

\[ T_c^{(3)} = \tilde{T}_c^{(3)} = \frac{1}{4}(\sigma_{56} - \sigma_{78}), \]  

(A24)

\[ T_c^{(8)} = \tilde{T}_c^{(8)} = \frac{1}{4\sqrt{3}}(\sigma_{56} + \sigma_{78} - 2 \sigma_{910}), \]  

\[ T_X = \tilde{T}_X = -\frac{2}{3}(\sigma_{56} + \sigma_{78} + \sigma_{910}). \]

While the \( T \)'s act on \( \chi \), the \( \tilde{T} \)'s (characterized by a sign flip in \( \sigma_{12} \)) act on \( \chi^c \). The normalization of the Cartan generators is chosen according to the usual SM convention. A GUT-consistent normalization across all generators is obtained by rescaling \( T_X \) (and \( T_X \)) by \( \sqrt{3/2} \).

In order to obtain the physical generators acting on the fundamental representation it is enough to replace \( \sigma_{ij}/2 \) in Eq. (A24) by \( \epsilon_{ij} \).

With this information at hand, one can identify the spinor components of \( \chi \) and \( \chi^c \):

\[ \chi = (\nu, u_1, u_2, u_3, l, d_1, d_2, d_3, -d_3, d_2^*, d_1^*, \nu^*, -u_3^*, -u_2^*, u_1^*), \]  

(A25)

and

\[ \chi^c = (\nu^*, u_1^*, u_2^*, u_3^*, \nu^*, d_1^*, d_2^*, d_3^*, -d_3, d_2, d_1, -l, u_3, -u_2^*, -u_1^*, \nu^*), \]  

(A26)

where a self-explanatory SM notation has been naturally extended into the scalar sector. In particular, the relative signs in Eqs. (A25) and (A26) arise from the charge conjugation of the \( SO(6) \sim SU(4)_C \) and \( SO(4) \sim SU(2)_L \otimes SU(2)_R \) components of \( \chi \) and \( \chi^c \).

The standard and flipped embeddings of \( SU(5) \) commute with two different Cartan generators, \( Z \) and \( Z' \) respectively:

\[ Z = -4T_R^{(3)} + 6T_X, \quad Z' = 4T_R^{(3)} + 6T_X. \]  

(A27)

Given the relation \( \text{Tr}(T_R^{(3)})^2 = \frac{3}{2} \text{Tr}T_X^2 \) one obtains

\[ \text{Tr}(YZ) = 0, \quad \text{Tr}(YZ') \neq 0, \]  

(A28)

where \( Y = T_R^{(3)} + T_X \) is the weak hypercharge generator.

As a consequence, the standard \( SU(5) \) contains the SM group, while \( SU(5)' \) has a subgroup \( SU(3)_c \otimes SU(2)_L \otimes U(1)' \), with

\[ Y' = -T_R^{(3)} + T_X. \]  

(A29)

In terms of \( Z' \) and of \( Y' \) the weak hypercharge reads

\[ Y' = \frac{3}{2}(Z' - Y'). \]  

(A30)

Using the explicit form of the Cartan generators in the vector representation one finds

\[ Z' \propto \text{diag}(-1, -1, +1, +1, +1) \otimes \Sigma_2, \]  

(A31)

\[ Z \propto \text{diag}(+1, +1, +1, +1, +1) \otimes \Sigma_2. \]  

(A32)

The vacuum configurations \( \omega_R = -\omega_Y \) and \( \omega_R = \omega_Y \) in Eq. (7) are aligned with the \( Z' \) and the \( Z \) generator respectively, thus preserving \( SU(5)' \otimes U(1)' \) and \( SU(5) \otimes U(1)_Z \), respectively.

### APPENDIX B: VACUUM STABILITY

The boundedness of the scalar potential is needed in order to ensure the global stability of the vacuum. The requirement that the potential is bounded from below sets nontrivial constraints on the quartic interactions. We do not provide a fully general analysis for the whole field space, but limit ourselves to the constraints obtained for the given vacuum directions.

1. \( (\omega_R, \omega_Y, \chi_R) \neq 0 \)

From the quartic part of the scalar potential \( V_0^{(4)} \) one obtains
\[ 4a_1(2\omega_R^2 + 3\omega_Y^2)^2 + \frac{a_2}{4}(8\omega_R^2 + 21\omega_Y^4 + 36\omega_R^2\omega_Y^2) \\
+ \frac{\lambda_1}{4}\chi_R^4 + 4\alpha\chi_R^2(2\omega_R^2 + 3\omega_Y^2) + \frac{\beta}{4}\chi_R^2(2\omega_R + 3\omega_Y)^2 \\
- \frac{\tau}{2}\chi_R^2(2\omega_R + 3\omega_Y) > 0. \] (B1)

Notice that the \( \lambda_2 \) term vanishes along the \( 16_H \) vacuum direction.

2. \( \omega_R = \omega_Y = 0, \chi_R \neq 0 \)

Along this direction the quartic potential \( V_0^{(4)} \) reads

\[ V_0^{(4)} = \frac{1}{4}\lambda_1\chi_R^4, \] (B2)

which implies

\[ \lambda_1 > 0. \] (B3)

3. \( \omega = \omega_R = -\omega_Y, \chi_R = 0 \)

From now on, we focus on the \( \chi_R = 0 \) case, c.f. Sec. IV.F. On this orbit the quartic part of the scalar potential reads

\[ V_0^{(4)} = \frac{3}{4}\omega^4(80a_1 + 13a_2). \] (B4)

Taking into account that the scalar mass spectrum implies \( a_2 < 0 \), we obtain

\[ a_1 > -\frac{13}{88}a_2. \] (B5)

4. \( \omega_R = 0, \omega_Y \neq 0, \chi_R = 0 \)

At the tree level this VEV configuration does not correspond to a minimum of the potential. It is nevertheless useful to inspect the stability conditions along this direction. Since

\[ V_0^{(4)} = \frac{3}{2}(48a_1 + 7a_2)\omega_Y^4, \] (B6)

boundedness is obtained, independently on the sign of \( a_2 \), when

\[ a_1 > -\frac{7}{48}a_2. \] (B7)

5. \( \omega_R \neq 0, \omega_Y = 0, \chi_R = 0 \)

In analogy with the previous case we have

\[ V_0^{(4)} = 2(8a_1 + a_2)\omega_R^4, \] (B8)

which implies the constraint

\[ a_1 > -\frac{1}{8}a_2. \] (B9)

In the case \( a_2 < 0 \) the constraint in Eq. (B5) provides the global lower bound on \( a_1 \).
\[M^2(1, 1, +1) = g^2 \chi_R^2,\]
\[M^2_\lambda(3, 1, -\frac{2}{3}) = g^2 \chi_R^2,\]
\[M^2_\lambda(3, 2, 0) = 0,\]
\[M^2_\lambda(8, 1, 0) = 0,\]
\[M^2_\lambda(3, 2, -\frac{5}{6}) = 0,\]
\[M^2_\lambda(3, 2, +\frac{1}{6}) = g^2 \chi_R^2,\]
\[M^2_\lambda(1, 1, 0) = \begin{pmatrix} \frac{2}{3} \sqrt{\frac{2}{1}} \\ \sqrt{\frac{2}{1}} \end{pmatrix} g^2 \chi_R^2,\]
\[\text{where the last matrix is again spanned over } (\psi_{24}^{15}, \psi_1^{15}),\]
yielding
\[\text{Det } M^2_\lambda(1, 1, 0) = 0,\]
\[\text{Tr } M^2_\lambda(1, 1, 0) = \frac{5}{2} g^2 \chi_R^2.\]

The number of vanishing entries corresponds to the dimension of the SU(5) algebra preserved by the 16_H VEV \( \chi_R \).

Summing together the 45_H and 16_H contributions, we recognize 12 massless states, that correspond to the SM gauge bosons.

### 2. Anatomy of the scalar spectrum

In order to understand the dependence of the scalar masses on the various parameters in the Higgs potential we detail the scalar mass spectrum in the relevant limits of the scalar couplings, according to the discussion on the accidental global symmetries in Sec. IV.

#### a. 45 only

Applying the stationary conditions in Eqs. (11) and (12), to the flipped 5' 1_H vacuum with \( \omega = \omega_R = -\omega_Y \), we find
\[M^2(24, 0) = -4a_2\omega^2, \quad M^2(10, -4) = 0,\]
\[M^2(1, 0) = 2(80a_1 + 13a_2)\omega^2,\]
and, as expected, the spectrum exhibits 20 WGB and 24 PGB whose mass depends on \( a_2 \) only. The required positivity of the scalar masses gives the constraints
\[a_2 < 0 \quad \text{and} \quad a_1 > -\frac{13}{80} a_2,\]
where the second equation coincides with the constraint coming from the stability of the scalar potential (see Eq. (B5) in Appendix B).

#### b. 16 only

When only the 16_H part of the scalar potential is considered the symmetry is spontaneously broken to the standard SU(5) gauge group. Applying the stationary Eq. (13) we find
\[M^2(\bar{5}) = 2\lambda_2\chi_R^2, \quad M^2(10) = 0,\]
\[M^2(1) = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \lambda_1 \chi_R^2.\]
in the \((\psi_{16}^{16}, \psi_1^{16})\) basis, that yields
\[\text{Det } M^2(1) = 0, \quad \text{Tr } M^2(1) = \lambda_1 \chi_R^2,\]
and, as expected, we count 21 WGB and 10 PGB modes whose mass depends on \( \lambda_2 \) only. The required positivity of the scalar masses leads to
\[\lambda_2 > 0 \quad \text{and} \quad \lambda_1 > 0,\]
where the second equation coincides with the constraint coming from the stability of the scalar potential (see Eq. (B3) in Appendix B).

#### c. Mixed 45-16 spectrum (\( \chi_R \neq 0 \))

In the general case the unbroken symmetry is the SM group. Applying first the two stationary conditions in Eq. (11) and (13) we find the spectrum below. The 2 \( \times \) 2 matrices are spanned over the \((\psi_{24}^{45}, \psi_{16}^{45})\) basis whereas the 4 \( \times \) 4 SM singlet matrix is given in the \((\psi_{24}^{45}, \psi_{16}^{45}, \psi_1^{16}, \psi_1^{16})\) basis.

\[M^2(1, 1, +1) = \begin{pmatrix} \beta \chi_R^2 + 2a_2\omega_Y(\omega_R + \omega_Y) & \chi_R(\tau - 3\beta\omega_Y) \\ \chi_R(\tau - 3\beta\omega_Y) & 2\omega_R(\tau - 3\beta\omega_Y) \end{pmatrix},\]
\[M^2(3, 1, -\frac{2}{3}) = \begin{pmatrix} \beta \chi_R^2 + 2a_2\omega_R(\omega_R + \omega_Y) & \chi_R(\tau - \beta(2\omega_R + \omega_Y)) \\ \chi_R(\tau - \beta(2\omega_R + \omega_Y)) & 2\omega_Y(\tau - \beta(2\omega_R + \omega_Y)) \end{pmatrix},\]
\[M^2(1, 3, 0) = 2a_2(\omega_Y - \omega_R)(\omega_R + 2\omega_Y), \quad M^2(8, 1, 0) = 2a_2(\omega_R - \omega_Y)(\omega_R + 2\omega_Y), \quad M^2(3, 2, -\frac{5}{6}) = 0.\]
By applying the remaining stationary condition in Eq. (12) one obtains

$$M^2(1,1,0) = \begin{pmatrix}
\frac{1}{2}(3\beta\chi^2_R + 4(a_2\omega_R + a_2\omega_Y + (48a_1 + 7a_2)\omega^2_Y)) & \sqrt{6}(\frac{2\chi^2_R}{2} + 2(16a_1 + 3a_2)\omega_R\omega_Y) \\
\sqrt{6}(\frac{2\chi^2_R}{2} + 2(16a_1 + 3a_2)\omega_R\omega_Y) & \beta\chi^2_R + 2(4(8a_1 + a_2)\omega^2_R + a_2\omega_Y\omega_R + a_2\omega^2_Y)
\end{pmatrix},$$

$$(C17)$$

In Eqs. (C14)–(C18) we recognize the 33 WGB with the quantum numbers of the coset $SO(10)/SM$ algebra.

In using the stationary condition in Eq. (12), we paid attention not to divide by $(\omega_R + \omega_Y)$, since the flipped vacuum $\omega = \omega_R = -\omega_Y$ is an allowed configuration. On the other hand, we can freely put $\omega_R$ and $\omega_Y$ into the denominators, as the vacua $\omega_R = 0$ and $\omega_Y = 0$ are excluded at the tree level. The coupling $a_2$ in Eq. (C18) is understood to obey the constraint

$$\frac{4a_2(\omega_R + \omega_Y)\omega_R\omega_Y + \beta\chi^2_R(2\omega_R + 3\omega_Y) - \tau\chi^2_R = 0.}$$

$$(C19)$$

d. A trivial 45-16 potential ($a_2 = \lambda_2 = \beta = \tau = 0$)

It is interesting to study the global symmetries of the scalar potential when only the moduli of 45$_H$ and 16$_H$ are present. In order to correctly count the corresponding PGB, the $(1, 1, 0)$ mass matrix in the limit of $a_2 = \lambda_2 = \beta = \tau = 0$ needs to be scrutinized. We find in the $(\psi^{45}_{14}, \psi^{45}_{15}, \psi^{16}_{14}, \psi^{16}_{15})$ basis,

$$M^2(1,1,0) = \begin{pmatrix}
96a_1\omega_Y^2 & 32\sqrt{6}a_1\omega_R\omega_Y & 8\sqrt{3}\alpha\chi_R\omega_Y \\
32\sqrt{6}a_1\omega_R\omega_Y & 64a_1\omega_R^2 & 8\sqrt{2}\alpha\chi_R\omega_R \\
8\sqrt{3}\alpha\chi_R\omega_Y & 8\sqrt{2}\alpha\chi_R\omega_R & \frac{1}{2}\lambda_1\chi^2_R \\
8\sqrt{3}\alpha\chi_R\omega_Y & 8\sqrt{2}\alpha\chi_R\omega_R & \frac{1}{2}\lambda_1\chi^2_R
\end{pmatrix},$$

$$(C20)$$

with the properties

$$\text{Rank } M^2(1,1,0) = 2, \quad \text{Tr}M^2(1,1,0) = 64a_1\omega_R^2 + 96a_1\omega_Y^2 + \lambda_1\chi^2_R.$$
As expected from the discussion in Sec. IV, Eqs. (C14)–(C20) in the $a_2 = \lambda_2 = \beta = \tau = 0$ limit exhibit 75 massless modes out of which 42 are PGB.

**e. A trivial 45-16 interaction ($\beta = \tau = 0$)**

In this limit, the interaction part of the potential consists only of the $\alpha$ term, which is the product of $45_H$ and $16_H$ moduli. Once again, in order to correctly count the massless modes we specialize the $(1, 1, 0)$ matrix to the $\beta = \tau = 0$ limit. In the $(\psi_2^{15}, \psi_1^{15}, \psi_1^{16}, \psi_1^{18})$ basis, we find

$$M^2(1, 1, 0, 0) = \begin{pmatrix}
2(a_2 \omega_R^2 + a_2 \omega_Y \omega_R + (48a_1 + 7a_2) \omega_Y^2) & 2\sqrt{6}(16a_1 + 3a_2)\omega_R \omega_Y & 8\sqrt{3}\alpha \chi_R \omega_Y & 8\sqrt{3}\alpha \chi_Y \omega_Y \\
2\sqrt{6}(16a_1 + 3a_2)\omega_R \omega_Y & 2(4(8a_1 + a_2)\omega_R^2 + a_2 \omega_Y \omega_R + a_2 \omega_Y^2) & 8\sqrt{2}\alpha \chi_R \omega_R & 8\sqrt{2}\alpha \chi_Y \omega_R \\
8\sqrt{3}\alpha \chi_R \omega_Y & 8\sqrt{2}\alpha \chi_R \omega_R & \frac{1}{2}\lambda_1 \chi_R^2 & \frac{1}{2}\lambda_1 \chi_R^2 \\
8\sqrt{3}\alpha \chi_Y \omega_Y & 8\sqrt{2}\alpha \chi_Y \omega_R & \frac{1}{2}\lambda_1 \chi_R^2 & \frac{1}{2}\lambda_1 \chi_R^2
\end{pmatrix},$$

(C22)

with the properties

$$\text{Rank} M^2(1, 1, 0, 0) = 3,$$

$$\text{Tr} M^2(1, 1, 0, 0) = 2((32a_1 + 5a_2)\omega_R^2 + 8(6a_1 + a_2)\omega_Y^2 + 2a_2 \omega_R \omega_Y) + \lambda_1 \chi_R^2.$$  

(C23)

According to the discussion in Sec. IV, upon inspecting Eqs. (C14)–(C18) in the $\beta = \tau = 0$ limit, one finds 41 massless scalar modes of which 8 are PGB.

**f. The 45-16 scalar spectrum for $\chi_R = 0$**

The application of the stationary conditions in Eqs. (11) and (12) (for $\chi_R = 0$, Eq. (13) is trivially satisfied) leads to four different spectra according to the four vacua: standard $\mathcal{5}_2^1$, flipped $\mathcal{5}^1_{1_2}$, $\mathcal{3}_2^1 \mathcal{2}_L \mathcal{1}_Y$ and $\mathcal{4}_C \mathcal{2}_L \mathcal{1}_Y$. We specialize our discussion to the last three cases.

The mass eigenstates are conveniently labeled according to the subalgebras of $SO(10)$ left invariant by each vacuum. With the help of Tables I and II one can easily recover the decomposition in the SM components. In the limit $\chi_R = 0$ the states $45_H$ and $16_H$ do not mix. All of the WGB belong to the $45_H$, since for $\chi_R = 0$ the $16_H$ preserves $SO(10)$.

Consider first the case: $\omega = \omega_R = -\omega_Y$ (which preserves the flipped $\mathcal{5}^1_{1_2}$ group). For the $45_H$ components we obtain:

$$M^2(24, 0) = -4a_2 \omega^2, \quad M^2(10, -4) = 0, \quad M^2(1, 0) = 2(80a_1 + 13a_2)\omega^2.$$  

(C24)

Analogously, for the $16_H$ components we get:

$$M^2(10, +1) = \frac{1}{4}(\omega^2(80\alpha + \beta) + 2\tau \omega - 2\nu^2),$$

$$M^2(5, -3) = \frac{1}{4}(\omega^2(80\alpha + 9\beta) - 6\tau \omega - 2\nu^2),$$

$$M^2(1, +5) = \frac{1}{4}(5\omega^2(16\alpha + 5\beta) + 10\tau \omega - 2\nu^2).$$

(C25)

Since the unbroken group is the flipped $\mathcal{5}^1_{1_2}$ we recognize, as expected, $45 - 25 = 20$ WGB. When only trivial $45_H$ invariants (moduli) are considered the global symmetry of the scalar potential is $O(45)$, broken spontaneously by $\omega$ to $O(44)$. This leads to 44 GB in the scalar spectrum. Therefore $44 - 20 = 24$ PGB are left in the spectrum. On general grounds, their masses should receive contributions from all of the explicitly breaking terms $a_2, \beta$ and $\tau$. As it is directly seen from the spectrum, only the $a_2$ term contributes at the tree level to $M^2(24, 0)$. By choosing $a_2 < 0$ one may obtain a consistent minimum of the scalar potential. Quantum corrections are not relevant in this case.

Consider then the case $\omega_R = 0$ and $\omega_Y \neq 0$ which preserves the $\mathcal{3}_2^1 \mathcal{2}_L \mathcal{2}_R \mathcal{1}_Y$ gauge group. For the $45_H$ components we obtain:

$$M^2(1, 3, 1, 0) = 2a_2 \omega_R^2, \quad M^2(1, 1, 3, 0) = 2a_2 \omega_Y^2, \quad M^2(8, 1, 1, 0) = -4a_2 \omega_Y^2, \quad M^2(3, 2, 2, -\frac{1}{3}) = 0,$$

$$M^2(3, 1, 1, -\frac{2}{3}) = 0,$$

$$M^2(1, 1, 1, 0) = 2(48a_1 + 7a_2)\omega_Y^2.$$  

(C26)
Analogously, for the $16_H$ components we get:

\[
M^2 \left(3, 2, 1, \frac{1}{6} \right) = \frac{1}{4} \left(\omega_5^2 (48\alpha + \beta) - 2\tau \omega_Y - 2\nu^2\right),
\]

\[
M^2 \left(\bar{3}, 1, 2, -\frac{1}{6} \right) = \frac{1}{4} \left(\omega_5^2 (48\alpha + \beta) + 2\tau \omega_Y - 2\nu^2\right),
\]

\[
M^2 \left(1, 2, 1, -\frac{1}{2} \right) = \frac{1}{4} \left(\omega_5^2 (48\alpha + 9\beta) + 6\tau \omega_Y - 2\nu^2\right),
\]

\[
M^2 \left(1, 1, 2, +\frac{1}{2} \right) = \frac{1}{4} \left(\omega_5^2 (48\alpha + 9\beta) - 6\tau \omega_Y - 2\nu^2\right).
\]

(C27)

Worth noting is the mass degeneracy of the $(1, 3, 1, 0)$ and $(1, 1, 3, 0)$ multiplets which is due to the fact that for $\omega_R = 0$ D-parity is conserved by even $\omega_Y$ powers. On the contrary, in the $16_H$ components the D-parity is broken by the $\tau$ term that is linear in $\omega_Y$.

Since the unbroken group is $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3$, there are $45 - 15 = 30$ WGB, as it appears from the explicit pattern of the scalar spectrum. When only trivial invariants (moduli terms) of $45_H$ are considered the global symmetry of the scalar potential is $O(45)$, broken spontaneously to $O(44)$, thus leading to 44 GB in the scalar spectrum. As a consequence $44 - 30 = 14$ GB are left in the spectrum. On general grounds, their masses should receive contributions from all of the explicitly breaking terms $a_2$, $\beta$ and $\tau$. As it is directly seen from the spectrum, only the $a_2$ term contributes at the tree level to the mass of the 14 GB, leading unavoidably to a tachyonic spectrum. This feature is naturally lifted at the quantum level.

Let us finally consider the case $\omega_R \neq 0$ and $\omega_Y = 0$ (which preserves the $4^C_2 \times 1^L$ gauge symmetry). For the $45_H$ components we find:

\[
M^2(15, 1, 0) = 2a_2 \omega_R^2,
\]

\[
M^2(1, 3, 0) = -4a_2 \omega_R^2,
\]

\[
M^2(6, 2, +\frac{1}{2}) = 0, \quad M^2(6, 2, -\frac{1}{2}) = 0,
\]

\[
M^2(1, 1, +1) = 0, \quad M^2(1, 1, 0) = 8(8a_1 + a_2) \omega_R^2.
\]

(C28)

For the $16_H$ components we obtain:

\[
M^2(4, 2, 0) = 8\alpha \omega_R^2 - \frac{1}{2} \nu^2,
\]

\[
M^2(\bar{4}, 1, +\frac{1}{2}) = \omega_R^2 (8\alpha + \beta) + \tau \omega_Y - \frac{1}{2} \nu^2, \quad (C29)
\]

The unbroken gauge symmetry in this case corresponds to $4^C_2 \times 1^L$. Therefore, one can recognize $45 - 19 = 26$ WGB in the scalar spectrum. When only trivial (moduli $45_H$ invariants are considered the global symmetry of the scalar potential is $O(45)$, which is broken spontaneously by $\omega_R$ to $O(44)$. This leads globally to 44 massless states in the scalar spectrum. As a consequence, $44 - 26 = 18$ PGB are left in the $45_H$ spectrum, that should receive mass contributions from the explicitly breaking terms $a_2$, $\beta$ and $\tau$. At the tree level only the $a_2$ term is present, leading again to a tachyonic spectrum. This is an accidental tree-level feature that is naturally lifted at the quantum level.

**APPENDIX D: ONE-LOOP MASS SPECTRA**

We have checked explicitly that the one-loop corrected stationary Eq. (12) maintains in the $\chi_R = 0$ limit the four tree-level solutions, namely, $\omega_R = \omega_Y$, $\omega_R = -\omega_Y$, $\omega_R = 0$ and $\omega_Y = 0$, corresponding, respectively, to the standard $5^L_1$, flipped $5^L_{1'}$, $3 \times 2^L_1 \times 1_R$ and $4^C_2 \times 1^L$ vacua.

In what follows we list, for the last three cases, the leading one-loop corrections, arising from the gauge and scalar sectors, to the critical PGB masses. For all other states the loop corrections provide only subleading perturbations of the tree-level masses, that are irrelevant to the present discussion.

### 1. Gauge contributions to the PGB mass

Before focusing to the three relevant vacuum configurations, it is convenient to write the gauge contribution to the $(1, 3, 0)$ and $(8, 1, 0)$ states in the general case.
One can easily recognize the (tree-level) masses of the gauge bosons in the log’s arguments and cofactors (see Appendix C 1 a). Note that only the massive states do contribute to the one-loop correction. (see Sec. V C).

Let us now specialize to the three relevant vacua. First, for the flipped $5'_{1_{L}}$ case $\omega = \omega_{R} = -\omega_{Y}$ one has:

$$\Delta M^2(24, 0) = \frac{17g^4\omega^2}{2\pi^2} + \frac{3g^4\omega^2}{2\pi^2} \log \left(\frac{4g^2\omega^2}{\mu^2}\right).$$  \hspace{1cm} (D3)

Similarly, for $\omega_R = 0$ and $\omega_Y \neq 0$ (3, 2$_L$ 2$_R$ 1$_X$):

$$\Delta M^2(1, 3, 1, 0) = \Delta M^2(1, 1, 3, 0) = \Delta M^2(1, 1, 3, 0)$$
$$= \frac{19g^4\omega^2_Y \beta^2}{4\pi^2} + \frac{2g^4\omega^2_Y \beta^2}{4\pi^2} \log \left(\frac{g^2\omega^2_Y}{\mu^2}\right)$$
$$- \frac{24g^4\omega^2_Y \beta^2}{4\pi^2} \log \left(\frac{4g^2\omega^2_Y}{\mu^2}\right).$$  \hspace{1cm} (D4)

$$\Delta M^2(8, 1, 1, 0) = \frac{11g^4\omega^2_Y \beta^2}{2\pi^2} + \frac{3g^4\omega^2_Y \beta^2}{2\pi^2} \log \left(\frac{4g^2\omega^2_Y}{\mu^2}\right).$$  \hspace{1cm} (D5)

Finally, for $\omega_{R} \neq 0$ and $\omega_{Y} = 0$ (4, 2$_L$ 1$_R$):

$$\Delta M^2(15, 1, 0) = \frac{13g^4\omega^2_{R} \beta^2}{4\pi^2} + \frac{9g^4\omega^2_{R} \beta^2}{4\pi^2} \log \left(\frac{g^2\omega^2_{R}}{\mu^2}\right)$$
$$- \frac{12g^4\omega^2_{R} \beta^2}{4\pi^2} \log \left(\frac{4g^2\omega^2_{R}}{\mu^2}\right).$$  \hspace{1cm} (D7)

2. Scalar contributions to the PGB mass

Since the general formula for the SM vacuum configuration is quite involved, we give directly the corrections to the PGB masses on the three vacua of our interest. We consider first the case $\omega = \omega_{R} = -\omega_{Y}$ (flipped $5'_{1_{L}}$):

$$\Delta M^2(24, 0) = \frac{\tau^2 + 5\beta^2\omega^2}{4\pi^2} + \frac{1}{128\pi^2 \omega} \left[ (-5\beta - \tau)(5\omega(16\alpha + 5\beta - 2\tau) - 2\nu^2) \right.$$
$$\times \log \left(\frac{5\omega(16\alpha + 5\beta - 2\tau)}{4\mu^2}\right) + (\omega(3\tau + 8\alpha + 3\beta + \beta\omega^2(27\beta - 400\alpha) - 10\tau^2)$$
$$+ \nu^2(10\beta - 6\tau) \log \left(\frac{\omega^2(80\alpha + 9\beta - 6\tau - 2\nu^2)}{4\mu^2}\right) + 2(\tau(33\beta - 80\alpha\omega)$$
$$+ \beta\omega^2(400\alpha + 17\beta + 10\tau^2) + 2\nu^2(\tau - 5\beta\omega)) \log \left(\frac{\omega^2(80\alpha + 17\beta + 2\tau^2 \omega - 2\nu^2)}{4\mu^2}\right) \right].$$  \hspace{1cm} (D8)

For $\omega_{R} = 0$ and $\omega_{Y} \neq 0$ (3, 2$_L$ 2$_R$ 1$_X$), we find:

$$\Delta M^2(1, 3, 1, 0) = \Delta M^2(1, 1, 3, 0) = \Delta M^2(1, 1, 3, 0)$$
$$= \frac{\tau^2 + 2\beta^2\omega^2_Y}{4\pi^2} + \frac{1}{16\pi^2 \omega} \left[ -(\tau - 3\beta\omega_Y)(-3\omega^2_Y(16\alpha + 3\beta) + 6\tau\omega_Y + 2\nu^2) \right.$$
$$\times \log \left(\frac{\omega^2_Y(48\alpha + 9\beta - 6\tau\omega_Y - 2\nu^2)}{4\mu^2}\right) - (\beta\omega_Y + \tau)(\omega^2_Y(48\alpha + \beta + 2\tau\omega_Y - 2\nu^2)$$
$$\times \log \left(\frac{\omega^2_Y(48\alpha + \beta + 2\tau\omega_Y - 2\nu^2)}{4\mu^2}\right) + (3\tau\omega^2_Y(16\alpha - 11\beta) + \beta\omega^3_Y(240\alpha + 17\beta)$$
$$\times \log \left(\frac{\omega^2_Y(48\alpha + 9\beta - 6\tau\omega_Y - 2\nu^2)}{4\mu^2}\right) + (\omega^2_Y(9\beta\tau - 48\alpha\omega) + 3\beta\omega^3_Y(9\beta - 16\alpha) + 2\omega_Y(\beta\nu^2 - \tau^2) + 2\nu^2\tau)$$
$$\times \log \left(\frac{\omega^2_Y(48\alpha + 9\beta + 6\tau\omega_Y - 2\nu^2)}{4\mu^2}\right) \right].$$  \hspace{1cm} (D9)
\[ \Delta M^2(8, 1, 1, 0) = \frac{\tau^2 + 3 \beta^2 \omega_8^2}{4 \pi^2} + \frac{1}{64 \pi^2 \omega_Y} \left[ -(\tau - 3 \beta \omega_Y)(-3 \omega_8^2(16\alpha + 3\beta) + 6\tau \omega_Y + 2\nu^2) \right. \\
\times \log \left( \frac{\omega_8^2(48\alpha + 9\beta) - 6\tau \omega_Y - 2\nu^2}{4\mu^2} \right) + (\omega_Y^2(21\beta \tau - 48\alpha \tau) + \beta \omega_Y^2(144\alpha + 11\beta) \right. \\
+ \omega_Y(6\tau^2 - 6\beta \nu^2) + 2\nu^2 \tau) \log \left( \frac{\omega_8^2(48\alpha + \beta) + 2\tau \omega_Y - 2\nu^2}{4\mu^2} \right) - (3\beta \omega_Y + \tau)(\omega_8^2(48\alpha + 9\beta) \\
+ 6\tau \omega_Y - 2\nu^2) \log \left( \frac{\omega_8^2(48\alpha + 9\beta) + 6\tau \omega_Y - 2\nu^2}{4\mu^2} \right) + (3\tau \omega_Y^2(16\alpha - 7\beta) + \beta \omega_Y^2(144\alpha + 11\beta) \\
+ \omega_Y(6\tau^2 - 6\beta \nu^2) - 2\nu^2 \tau) \log \left( \frac{\omega_8^2(48\alpha + \beta) - 2\tau \omega_Y - 2\nu^2}{4\mu^2} \right) \bigg]. \tag{D10} \]

Finally, for \( \omega_R \neq 0 \) and \( \omega_Y = 0 \), we have:

\[ \Delta M^2(1, 3, 0) = \frac{\tau^2 + 2 \beta^2 \omega_8^2}{4 \pi^2} + \frac{1}{64 \pi^2 \omega_R} \left[ 16\omega_R(16\alpha \beta \omega_R^2 - \beta \nu^2 + \tau^2) \log \left( \frac{8\alpha \omega_R^2 - \frac{\tau^2}{2}}{\mu^2} \right) \right. \\
- 4(\tau - 2\beta \omega_R)(-2\omega_R^2(8\alpha + \beta) + 2\tau \omega_R + \nu^2) \log \left( \frac{\omega_R^2(8\alpha + \beta) - \tau \omega_R - \frac{\nu^2}{2}}{\mu^2} \right) \\
- 4(2\beta \omega_R + \tau)(2\omega_R^2(8\alpha + \beta) + 2\tau \omega_R - \nu^2) \log \left( \frac{\omega_R^2(8\alpha + \beta) + \tau \omega_R - \frac{\nu^2}{2}}{4\mu^2} \right) \bigg]. \tag{D11} \]

\[ \Delta M^2(15, 1, 0) = \frac{\tau^2 + 3 \beta^2 \omega_8^2}{4 \pi^2} + \frac{1}{64 \pi^2 \omega_R} \left[ 8\omega_R(16\alpha \beta \omega_R^2 - \beta \nu^2 + \tau^2) \log \left( \frac{8\alpha \omega_R^2 - \frac{\tau^2}{2}}{\mu^2} \right) \right. \\
- 4(2\beta \omega_R^2(8\alpha - \beta) - 16\alpha \tau \omega_R^2 + \omega_R(\tau^2 - \beta \nu^2) + \nu^2 \tau) \log \left( \frac{\omega_R^2(8\alpha + \beta) - \tau \omega_R - \frac{\nu^2}{2}}{\mu^2} \right) \bigg] \\
+ 4(2\beta \omega_R^2(\beta - 8\alpha) - 16\alpha \tau \omega_R^2 + \omega_R(\beta \nu^2 - \tau^2) + \nu^2 \tau) \log \left( \frac{\omega_R^2(8\alpha + \beta) + \tau \omega_R - \frac{\nu^2}{2}}{\mu^2} \right). \tag{D12} \]

Also in these formulas we recognize the (tree-level) mass eigenvalues of the 16H states contributing to the one-loop effective potential (see Appendix C 2 f). Notice that the singlets with respect to each vacuum, namely, (1, 0), (1, 1, 1, 0) and (1, 1, 0), for the flipped \( S^1 Z \), \( 3_2 2_1 L \) and \( 4_2 2_1 L \) vacua, respectively, receive a tree-level contribution from both \( a_1 \) as well as \( a_2 \) (see Appendix C 2 f). The \( a_1 \) term leads the tree-level mass and radiative corrections can be neglected.

One may verify that in the limit of vanishing VEVs the one-loop masses vanish identically on each of the three vacua, as it should be. This is a nontrivial check of the calculation of the scalar induced corrections.


5.4 Two-loop SUSY RGEs with U(1) mixing


The main scope of this article was to fill an unpleasant “hole” in the literature on the renormalization group techniques in supersymmetric gauge field theories in which the corresponding gauge symmetry contains more than a single Abelian group factor, cf. Sect. 3.5 of this thesis. By doing that, the programme set out by the seminal works of Y. Yamada and S. Martin and M. Vaughn in the early 1990’s has been finally completed and a self-contained toolbox facilitating the construction of two-loop beta-functions for any softly broken $\mathcal{N} = 1$ renormalizable supersymmetric gauge field theory has been provided.

Perhaps the most interesting aspect of the study consists in the method by which the desired results have been obtained. Indeed, the obvious brute-force approach which, on one hand, was guaranteed to yield the result but, on the other hand, represented an enormous tedium with only limited cross-checking options, was almost completely avoided by exploiting the symmetries of the covariant derivative, the basic building element of the theory, cf. Sect. 3.5.4. With this technique, one can attempt to construct the two-loop beta-functions from the covariant blocks including gauge couplings organised into matrices and the other structures (group invariants, soft terms etc.) by a mere matching to the known results for the theories with a single $U(1)$ gauge factor. Remarkably, this method admitted to resolve almost all ambiguities up to a very limited set of residual Feynman graphs’ topologies which still required the old-fashioned treatment. Besides providing the final completion of the “vocabulary” of Martin and Vaughn’s Physical Review D 50 (1994) 2282 the article elaborates also on the issue of the matching for the gauge couplings and gaugino masses and illustrates the basic principles on a set of specific examples.

The importance of these results has been widely recognised by the HEP community; this can be clearly demonstrated on the fact that the study itself and its sequel Physics Letters B726, 882 (2013) have, to date, attracted over 100 citations in the inspirehep.net database.

The candidate’s main contribution to the study was the crucial observation that the symmetry properties of the gauge-coupling- and charge matrices, together with matching to the existing single-$U(1)$ results, provide a very powerful method for circumventing the swampland of the ab-initio approach. Besides that, he contributed by the “Methods” of Sect. 2, examples of Sect. 4 and by Appendix A therein.
Running soft parameters in SUSY models with multiple $U(1)$ gauge factors

Renato M. Fonseca$^a$, Michal Malinský$^b$, Werner Porod$^c$, Florian Staub$^c$,*

$^a$ Centro de Física Teórica de Partículas, Instituto Superior Técnico, Universidade Técnica de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal

$^b$ AHEP Group, Instituto de Física Corpuscular – C.S.I.C./Universitat de València Edificio de Institutos de Paterna, Apartado 22085, E-46071 València, Spain

$^c$ Institut für Theoretische Physik und Astronomie, Universität Würzburg Am Hubland, 97074 Wuerzburg, Germany

Received 20 July 2011; accepted 17 August 2011
Available online 24 August 2011

Abstract

We generalize the two-loop renormalization group equations for the parameters of the softly broken SUSY gauge theories given in the literature to the most general case when the gauge group contains more than a single Abelian gauge factor. The complete method is illustrated at two-loop within a specific example and compared to some of the previously proposed partial treatments.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Since the advent of the renormalization group (RG) techniques [1–4], a lot of effort has been put into the calculation of the $\beta$-functions and anomalous dimensions of specific theories. For instance, full-fledged two-loop formulae for non-supersymmetric gauge models became available as early as in 1984 thanks to the seminal works by Machacek and Vaughn [5–7]. In the context of supersymmetry (SUSY), the need to adopt the existing machinery for the soft SUSY-
breaking sector postponed the arrival of the first generic two-loop results for about ten years [8–15]. Since then, there have even been attempts to go beyond two loops in the literature, cf., [16,17].

For the sake of simplicity, in many of the pioneering works the gauge group was assumed to contain at most one Abelian gauge factor. The point is that with more than a single gauged \( U(1) \) in play, a new qualitative feature requiring a dedicated treatment emerges. This is due to the fact that Abelian field tensors \( F_{\mu\nu} \) are not only gauge-covariant but rather gauge-invariant quantities and, thus, unlike the non-Abelian ones, they can contract among each other without violating gauge invariance, giving rise to off-diagonal kinetic terms [18,19].

Moreover, even if such terms happen to be absent from the tree level Lagrangian at a certain scale, they are in general re-introduced by the renormalization-group evolution [20,21]. The reason is that the anomalous dimension \( \gamma \) driving the relevant renormalization group equations (RGEs) are in general non-diagonal symmetric matrices in the gauge-field space, thus giving rise to off-diagonal corrections to the gauge boson propagators. These, in turn, require extra counterterms in order to retain renormalizability.

Actually, there are several exceptions to this basic rule. For instance, it can be that all the relevant \( U(1) \) couplings originate from a common gauge factor and, thus, barring threshold effects, all of them happen to be equal at a certain scale. In such a case, accidentally, the charges and the gauge fields can be simultaneously rotated at the one-loop level so that no off-diagonalsities pop up in \( \gamma \) [10,20] and one can use the simple form of the RGEs for individual gauge couplings. This is relatively easy to implement in the non-SUSY case where only the gauge sector has to be taken into account; the only price to be paid is the presence of continuous charges in the game.

In supersymmetry, the \( U(1) \) gaugino soft masses can also mix, and thus one has to deal with the non-diagonalities in the gaugino sector too. Again, the rotated basis can be helpful if both, gaugino masses and gauge couplings, unify at the same scale. However, this method is consistent only at the one-loop level where the evolution equations for the gauge couplings and gaugino soft masses essentially coincide. At two loops, Yukawa couplings and trilinear soft SUSY breaking couplings enter and the relevant algebraic structures are independent of each other which, in turn, renders this approach useless.

For the non-SUSY gauge theories, the full generalization of the original two-loop results for gauge groups with at most a single \( U(1) \) factor to the case with multiple \( U(1) \)'s has been formulated relatively recently, see, e.g., [22] and dedicated two-loop studies focusing on such effects in the context of, e.g., grand unified theories (GUTs) are available [23]. However, for the softly-broken SUSY gauge theories, the general two-loop evolution equations for the soft-breaking parameters in presence of the \( U(1) \)-mixing effects have not yet been given.\(^1\)

In this study, we aim to fill this gap by presenting a set of substitution rules which generalize the results of [10,11] to the case where the gauge group involves more than a single Abelian gauge factor.

The practical applications of these results are manifold. For instance, in SUSY GUTs featuring an extended intermediate \( U(1)_R \times U(1)_{B-L} \) stage, see e.g. [25], the \( U(1) \)-mixing effects can shift the effective MSSM bino soft mass by several per cent with respect to the

\(^1\) In Ref. [24] the effect of the mixings of several \( U(1) \)'s has been taken into account in the anomalous dimensions of the superfields and in the beta-functions of the gauge couplings which serve as basis for the corresponding parts in the RGEs of the soft SUSY breaking parameters.
naive estimate where such effects are neglected. In principle, this can have non-negligible
effects for the low-energy phenomenology. In this respect, let us just mention that the
theories with a gauged $U(1)_{B-L}$ surviving down to the proximity of the soft SUSY-breaking
scale have become rather popular recently due to their interesting implications for the R-
parity and the mechanism of its spontaneous violation [26–28], for Leptogenesis [29,30], etc.

This work is organized as follows: In Section 2 we recapitulate the salient features of gauge
theories with several Abelian gauge factor focusing namely on the different renormalization con-
ventions. A specific scheme in which the desired generalization of [10] can be carried out in
a particularly efficient way is identified. In Section 3 the relevant substitution rules upgrading
those in [10] to the most general form are given and the methods for resolving some ambiguities
emerging throughout their derivation are briefly commented upon. In Section 4 we discuss illustrate
the importance of the kinetic mixing effects in a pair of specific models, focusing namely
on the comparison between the “rotated basis” method advocated in [10] and the full-fledged
two-loop treatment. Then we conclude. For the sake of completeness, we add a set of append-
dices: some technical details of the renormalization scheme definition are given in Appendix A;
the basic formulae of [10] for a simple gauge group and their generalization to the case with
product groups can be found in Appendix B. In Appendix C, the interested reader can find de-
tails of the derivation of our main results presented in Section 3. Finally, Appendix D is devoted
to several remarks on the gauge and gaugino matching in theories with multiple $U(1)$ gauge
factors.

2. Methods

As mentioned in the introduction, going from a single-$U(1)$ to the multiple-$U(1)$ case is not
straightforward as it generally amounts to a qualitative change in the Lagrangian. In particular,
there is a need for an extra set of counterterms which, in simple words, keep the renormalized
off-diagonal two-point Green’s functions in the gauge sector finite. This, however, implies that
the renormalized Lagrangian must contain a structure connecting the field tensors associated to
different $U(1)$’s in the gauge-kinetic terms, namely

$$\mathcal{L}_{\text{kin.}} \ni - \frac{1}{4} F^{\mu \nu} \xi F_{\mu \nu}$$

(1)

where the different field tensors have been grouped into an $n$-dimensional vector $F^{\mu \nu}$ (with $n$
denoting the number of independent gauged $U(1)$ factors) and $\xi$ is an $n \times n$ real and symmet-
ric matrix. This amounts to $\frac{1}{2} n(n - 1)$ extra dynamical parameters. These quantities are then
governed by a new set of evolution equations which have to be added to those governing the
individual gauge couplings and other relevant parameters such as Yukawas etc. This, indeed, is
the method adopted in some of the first studies of the subject, see, e.g., [22].

Alternatively, one can work in a renormalization scheme in which the $\xi$-term in Eq. (1) is
transformed out by a suitable redefinition of the gauge fields, namely,

$$A \to \xi^{1/2} A$$

(2)

which also leads to the canonical normalization of the gauge fields. This, indeed, affects the
interaction part of the covariant derivative

$$Q^T_i \tilde{G} A \to Q^T_i \tilde{G} \xi^{-1/2} A,$$

(3)
where \( \tilde{G} \) is the original diagonal matrix\(^2\) of \( n \) individual gauge couplings associated to the \( n \) Abelian gauge factors and \( Q_i \) is the vector\(^3\) of the relevant \( U(1) \) charges. Similarly, the gauge-kinetic counterterm is transformed

\[
Z_B^{1/2} \xi B Z_A^{1/2} - \xi \to \xi^{-1/2} Z_B^{1/2} \xi B Z_A^{1/2} \xi^{-1/2} - 1 \equiv \delta Z_A, 
\]

(4)

where the subscript \( B \) denotes bare quantities and \( Z_A^{1/2} \) is the original (diagonal) gauge-field renormalization factor \( A_B = Z_A^{1/2} A \). Hence, the \( \xi^{-1/2} \) factor can be subsumed into a new set of \( \frac{1}{2} n(n-1) \) “effective” gauge couplings whose combinations populate the off-diagonal entries of an “extended gauge-coupling matrix”

\[
G \equiv \tilde{G} \xi^{-1/2},
\]

(5)

and a suitably redefined gauge-kinetic counterterm.

Thus, in this scheme, the off-diagonality in the gauge-kinetic part of the renormalized Lagrangian is absorbed by the covariant derivative, while the gauge-kinetic counterterm \( \delta Z_A \) is naturally off-diagonal in order to absorb the divergences in the off-diagonal two-point functions. Moreover, the simple QED-like relation between the bare and renormalized Abelian gauge coupling matrices (omitting the tildes)

\[
G_B = G Z_A^{-1/2}
\]

(6)

remains intact because the relevant Ward identities that lead to the cancellation of \( Z_\psi \) and the \( Z_G \) factors, cf., Eq. (A.9), follow from the gauge invariance. Therefore, it is sufficient to work with a matrix-like gauge-coupling structure forgetting entirely about the \( \xi \)-origin of its off-diagonal entries.

This strategy, which is entirely equivalent to the former one with a dynamical \( \xi \), is much more suitable for our task because it essentially amounts to replacing all the polynomials including individual gauge couplings in [10] by the relevant matrix structures, with no need\(^4\) to deal with the evolution equations for the \( \xi \) matrix not discussed here.

This, however, is not entirely straightforward in practice. Indeed, the commutativity of c-numbers has been widely used in [10] in order to cast their results in a compact form. Thus, one has to be very careful to avoid ambiguities stemming from the generic non-commutativity of the matrix-like \( G \)’s. Furthermore, also the Abelian gaugino soft masses have to be arranged into a matrix structure \( M \), which brings in an extra complication.

In doing so, an invaluable key is provided by some of the residual reparametrization symmetries of the renormalized Lagrangian. In particular,

\[
Q_i \to O_1 Q_i, 
\]

(7)

\[
G \to O_1 G O_2^T, 
\]

(8)

\[
A \to O_2 A, 
\]

(9)

\(^2\) With indices in the group and gauge-field spaces, respectively.

\(^3\) With a lower index assigning the corresponding matter-field.

\(^4\) Obviously, no information is lost so one can obtain the relevant RGEs for \( \xi \) components from the ones with the matrix-like gauge couplings. Indeed, the number of the off-diagonal entries in \( \xi \) is the same like the number of independent physical parameters governing the off-diagonal entries of \( G \); here one has to take into account the freedom to bring \( G \) into a triangular form by a suitable redefinition of the \( U(1) \) charges.
where \( O_1 \) and \( O_2 \) are arbitrary orthogonal matrices acting in the group and gauge-field spaces, respectively, leave the interaction part of the covariant derivative \( Q_i^T G A \) invariant. Under the same set of transformations, the gaugino mass matrix is rotated to
\[
M \rightarrow O_2 M O_2^T.
\]
(10)

Naturally, these symmetries must be reflected at the RGE level.

Thus, for instance, only those combinations \( C \) of \( G \) and \( \gamma \propto \sum_i Q_i Q_i^T \) that transform as \( C \rightarrow O_1 C O_2^T \) are allowed to enter the right-hand side of the renormalization group equation for \( G \). However, at one-loop level, there is only one structure involving a third power of \( G \) and one power of \( \gamma \) that can come up from a matter-field loop in the gauge propagator, namely \( G G^T \gamma G \), so one immediately concludes that
\[
\tilde{\beta}_G^{1\text{loop}} \propto G G^T \gamma G.
\]
(11)

The proportionality coefficient is trivially obtained by matching this to the single-\( U(1) \) case. This also illustrates that it is more convenient to work in the scheme with off-diagonal \( G \) than in the scheme with a non-trivial \( \xi \), simply because the transformation properties of \( G \) (which is a general real matrix) are more restrictive than the transformation properties of \( \xi \) (which is symmetric).

However, at two loop-level this becomes more complicated because then, for instance, all gauge couplings including those corresponding to the semi-simple part of the total gauge group mix among each other and/or with the relevant gaugino masses. Next, different Feynman-graph topologies can be subsumed under the same specific term in [10,11] and, hence, ambiguities must be resolved, which often require some amount of a “reverse engineering”.

Nevertheless, as we shall demonstrate in the next section, all such ambiguities, if properly traced back to the original diagrams, can be sorted out and a clear and elegant picture emerges.

3. Results

In this section, we shall describe the generic method of constructing the fully general two-loop RGEs for softly-broken supersymmetric gauge theories out of the results of [10,11] relevant to the case of at most a single Abelian gauge-group factor. For the sake of completeness, the relevant formulae for the cases of (i) a simple gauge group and (ii) the product of several simple factors with at most a single \( U(1) \) are reiterated in Appendices B.1 and B.2, respectively. The computation has been done using the \( DR' \) scheme defined in [13].

3.1. Notation and conventions

The gauge group is taken to be \( G_A \otimes G_B \otimes \cdots \otimes U(1)^n \), where the \( G_X \)'s are simple groups. We shall use uppercase indices for simple group-factors only; lowercase indices are used either for all groups or, in some specific cases, for \( U(1) \) groups only.\(^5\) As mentioned before, the \( U(1) \) sector should be treated as a whole and described in terms of a general real \( n \times n \) gauge-coupling matrix \( G \), an \( n \times n \) symmetric soft-SUSY breaking gaugino mass-matrix \( M \) and a column vector

---

\(^5\) This will be evident from the context; we follow as closely as possible [10] and when quoting results contained therein, the \( a \) and \( b \) indices go over all groups (simple and \( U(1) \) groups). On other occasions, when referring to particular components of the \( U(1) \)-related \( G \) and \( M \) matrices and vectors, \( a \) and \( b \) stretch over the \( U(1) \) groups only.
of charges $Q_i$ for each chiral supermultiplet $\Phi_i$. Notice, however, that $V_i \equiv G^T Q_i$ for each $i$ are the only combinations of $Q_i$ and $G$ which appear in the Lagrangian and, thus, all the general RGEs can be, in principle, written in terms of $V$’s and $M$ only. We shall follow this convention with a single exception of the evolution equations for the gauge couplings which are traditionally written in terms of $dG/d\log t$ rather than $dV/d\log t$ — indeed, in this case we shall adhere to the usual practice. As a consequence, we expect an isolated $G$ popping up in these equations.

Before proceeding any further we shall define some of the expressions that are used in the RGEs:

- $C_a(i)$: Quadratic Casimir invariant of the representation of superfield $\Phi_i$ under the group $G_a$;
- $C(G_a)$: Quadratic Casimir invariant of the adjoint representation of group $G_a$;
- $S_a(i)$: Dynkin index of the representation of superfield $\Phi_i$ under the group $G_a$;
- $d_a(i)$: Dimension of the representation of $\Phi_i$ under the group $G_a$;
- $d(G_a)$: Dimension of group $G_a$;
- $S_a(R)$: Dynkin index of group $G_a$ summed over all chiral supermultiplets — $S_a(R) = \sum_i S_a(i)$;
- $S_a(R)C_b(R)$: Defined as $\sum_i \frac{S_a(i)C_b(i)}{d_a(i)}$;
- $S_a(R)V_R^T V_R$: Defined as $\sum_i \frac{S_a(i)V_R^T V_i}{d_a(i)}$;
- $S_a(R)V_R^T M V_R$: Defined as $\sum_i \frac{S_a(i)V_R^T M V_i}{d_a(i)}$.

In addition, sometimes one has to deal with the explicit representation matrices of the gauge groups (denoted in [10] by $t_i^{AJ}$). Notice that here $A$ is not a group index but rather a coordinate in the adjoint representation of the corresponding Lie algebra (e.g., $A = 1, \ldots, 3$ in $SU(2)$, $A = 1, \ldots, 8$ in $SU(3)$, etc.). Naturally, whenever we refer to results of Refs. [10,11] for a simple gauge group (collected in Appendix B.1), the $a$ and $b$ indices will be omitted. In all cases, repeated indices are not implicitly summed over.

3.2. Constructing the general substitution rules

Let us now sketch in more detail the general strategy for upgrading the “product” substitution rules of Section III in Ref. [10] to the most general case of an arbitrary gauge group. For sake of simplicity, we shall focus on a limited number of terms here; the interested reader can find a more elaborate exemplification of the basic procedure in Appendix C.

Let us begin with, e.g., the term $g^2 C(r)$ appearing for instance in Eq. (B.3) and, subsequently, in the substitution rules of [10] for product groups, Eq. (B.34). It is clear that this has to be replaced by $\sum_A g_A^2 C_A(r) + ‘U(1) part’. For a single $U(1)$, $g^2 C(r) = g^2 y_r^2 \sim V_r V_r$ so this ‘$U(1)$ part’ can only take the form$^6$: $V_r^T V_r = Q_r^T G G^T Q_r$. There is no other way to obtain a number from two vectors $V_r$. Remarkably, this expression sums automatically the contributions of all the $U(1)$’s.

$^6$ If there is a single Abelian factor group, we denote by $y_i$ the (hyper)charge of chiral superfield $\Phi_i$, which is just a number.
Similarly, $Mg^2 C(r)$ (in Eq. (B.15) for example) is replaced by $\sum_A M_A g^2_A C_A(r) + 'U(1) part'; the ingredients for the construction of the 'U(1) part' are two vectors $V_r$ and the gaugino mass matrix $M$. Only $V_r^T M V_r$ forms a number.

In fact, this simple procedure allows us to generalize many of the terms in the RGEs of [10, 11], Section II (and/or Appendix B.1). As a more involved example, consider for instance the $g^4 t^{A\bar{A}}_i \text{Tr}[t^A C(r)m^2]$ structure popping up in Eq. (B.25). It is not difficult to see that all terms where the representation matrices $t^A$ appear explicitly are only relevant for the Abelian groups. Hence, if for a single $U(1)$ one has $g^4 t^{A\bar{A}}_i \text{Tr}[t^A C(r)m^2] = g^4 \delta^i_3 y_i \sum_p y^3_p (m^2)_p$, it can be immediately deduced that, in the general case, $g^4 t^{A\bar{A}}_i \text{Tr}[t^A C(r)m^2] \rightarrow \delta^i_3 \sum_p (V^T_p V_p)(V^T_p V_p)(m^2)_p$.

The RGEs of $G$ and $M$ represent a bigger challenge, because they are matrix equations (i.e., the gauge indices remain open). On the other hand, this should be viewed as an advantage because all the relevant equations must then respect the reparametrization symmetries (7)–(10). In this respect, let us reiterate Eqs. (7)–(9) which imply that $V$’s transform as $V_i \rightarrow O_2 V_i$. These symmetries are especially powerful in the $\beta$-functions for the gauge couplings which, due to Eq. (8), inevitably take the generic form $GV_i(\cdots)V^T_j$ for some chiral indices $i, j$. For example, $g^3 S(R) \sim g^3 \sum_p y^2_p$ can only take the form $G \sum_p V_p V^T_p$.

Concerning the gaugino soft masses $M$, let us for instance take a look at the $2g^2 S(R) M$ term appearing in Eq. (B.21). Its generalized variant should be, obviously, built out of a pair of $V_p$ vectors and the $M$ matrix. However, there are only two combinations of these objects that transform correctly under $O_2$, namely, $M V_p V^T_p$ and $V_p V^T_p M$. Thus, due to the symmetry of $M$, one obtains $2g^2 S(R) M \rightarrow M \sum_p V_p V^T_p + \sum_p V_p V^T_p M$.

Another important ingredient of the analysis is provided by the existing substitution rules linking the case of a simple gauge group (Section II in [10] and/or Appendix B.1) to the settings with group products (Section III in [10] and/or Appendix B.2). Consider, for example, the $g^2 S(R) C(R)$ term in Eq. (B.6) which, according to [10], gets replaced by $\sum_b g^2_a g^2_b S_a(R) C_b(R)$, see formula (B.31) for the product groups. Let us recall that the expression $S(R) C(R)$ has a very particular meaning — it is the sum of the Dynkin indices weighted by the quadratic Casimir invariant, so $\sum_b g^2_a g^2_b S_a(R) C_b(R) = \sum_{b,p} g^3_a g^2_b S_{(p)}(C_{(p)}) \frac{d_a(p)}{d_a(p)}$. With this in mind, whenever $a$ refers to the Abelian part of the gauge group, one should replace $g^3_a S_a(p) \rightarrow GV_p V^T_p$, $\sum_b g^2_b C_b(p) \rightarrow \frac{\sum_b g^2_b C_b(p) + V^T_p V_p}{d_a(p)}$ and $d_a(p) = 1$. Therefore, for the Abelian sector, $g^2 S(R) C(R) \rightarrow \sum_p GV_p V^T_p [\sum_b g^2_b C_b(p) + V^T_p V_p]$.

However, sometimes even a detailed inspection of the underlying expressions does not admit for an unambiguous identification of its generalized form. Then, a careful analysis of the structure of the contributing Feynman diagrams is necessary. Remarkably, the number of such singular cases is rather limited and can be carried out rather efficiently, as shown in Appendix C.

3.3. List of substitution rules

Depending on the group sector (Abelian or simple), we get different RGEs for the gauge couplings and the gaugino masses. The parameters are then either the matrices $G$, $M$ or the numbers $g_A$, $M_A$. For the Abelian sector, one obtains:

$$C(G) \rightarrow 0,$$

$$g^3 S(R) \rightarrow G \sum_p V_p V^T_p,$$
\[ g^5 S(R) C(R) \rightarrow \sum_p G V_p V_p^T \left[ \sum_B g_B^2 C_B(p) + V_p^T V_p \right], \] (14)

\[ \frac{g^3 C(k)}{d(G)} \rightarrow G V_k V_k^T, \] (15)

\[ 2g^2 S(R) M \rightarrow M \sum_p V_p V_p^T + \sum_p V_p V_p^T V_k, \] (16)

\[ g^2 C(k) \rightarrow V_k V_k^T, \] (17)

\[ 2g^2 C(k) M \rightarrow M V_k V_k^T + V_k V_k^T M, \] (18)

\[ 16g^4 S(R) C(R) M \rightarrow \sum_p \left\{ 4(M V_p V_p^T + V_p V_p^T M) \left[ \sum_B g_B^2 C_B(p) + V_p^T V_p \right] \right. \]
\[ + 8V_p V_p^T \left[ \sum_B M_B g_B^2 C_B(p) + V_p^T M V_p \right] \}. \] (19)

For a simple group factor \( G_A \), the substitution rules of [10] do not need to be changed except for two cases:

\[ g^5 S(R) C(R) \rightarrow g_A^3 S_A(R) \left[ \sum_B g_B^2 C_B(R) + V_R^T V_R \right], \] (20)

\[ 16g^4 S(R) C(R) M \rightarrow 8M_A S_A(R) \left[ \sum_B g_B^2 C_B(R) + V_R^T V_R \right] \]
\[ + 8S_A(R) \left[ \sum_B M_B g_B^2 C_B(R) + V_R^T M V_R \right]. \] (21)

As for the rest of the parameters in a SUSY model, the relevant substitution rules read:

\[ g^2 C(r) \rightarrow \sum_A g_A^3 C_A(r) + V_r^T V_r, \] (22)

\[ M g^2 C(r) \rightarrow \sum_A M_A g_A^3 C_A(r) + V_r^T M V_r, \] (23)

\[ M^* g^2 C(r) \rightarrow \sum_A M_A^* g_A^3 C_A(r) + V_r^T M^* V_r, \] (24)

\[ M M^* g^2 C(r) \rightarrow \sum_A M_A M_A^* g_A^3 C_A(r) + V_r^T M M^* V_r, \] (25)

\[ g^4 C(r) S(R) \rightarrow \sum_A g_A^4 C_A(r) S_A(R) + \sum_p (V_r^T V_p)^2, \] (26)

\[ M g^4 C(r) S(R) \rightarrow \sum_A M_A g_A^4 C_A(r) S_A(R) + \sum_p (V_r^T M V_p)(V_r^T V_p), \] (27)

\[ g^4 C^2(r) \rightarrow \sum_{A,B} g_A^2 g_B^2 C_A(r) C_B(r) + 2 \sum_A g_A^2 C_A(r) (V_r^T V_r) + (V_r^T V_r)^2, \] (28)

\[ M g^4 C^2(r) \rightarrow \sum_{A,B} M_A g_A^2 g_B^2 C_A(r) C_B(r) \]
\[
+ \sum_{A} g_{A}^{2} C_{A}(r) [M_{A}(V_{r}^{T} V_{r}) + (V_{r}^{T} M V_{r})] + (V_{r}^{T} M V_{r})(V_{r}^{T} V_{r}), \tag{29}
\]

\[
g_{A}^{4} C (G) C (r) \rightarrow \sum_{A} g_{A}^{4} C (G_{A}) C_{A}(r), \tag{30}
\]

\[
M g_{A}^{4} C (G) C (r) \rightarrow \sum_{A} M_{A} g_{A}^{4} C (G_{A}) C_{A}(r), \tag{31}
\]

\[
MM^{*} g_{A}^{4} C (G) C (r) \rightarrow \sum_{A} M_{A} M^{*}_{A} g_{A}^{4} C (G_{A}) C_{A}(r), \tag{32}
\]

\[
g_{A}^{2} t_{i}^{A j} \text{Tr}(t^{A} m^{2}) \rightarrow \delta_{i}^{j} \sum_{p} (V_{i}^{T} V_{p})(m^{2})_{p}^{p}, \tag{33}
\]

\[
g_{A}^{2} t_{i}^{A j} \text{Tr}(t^{A} m^{2})_{r} \rightarrow \delta_{i}^{j} (V_{i}^{T} V_{i})(m^{2})_{r}^{r}, \tag{34}
\]

\[
g_{A}^{2} t_{i}^{A j} \text{Tr}(t^{A} C(r) m^{2}) \rightarrow \delta_{i}^{j} \sum_{p} (V_{i}^{T} V_{p})(V_{p}^{T} V_{p})(m^{2})_{p}^{p}, \tag{35}
\]

\[
g_{A}^{2} C (i) \text{Tr}[S(r) m^{2}] \rightarrow \sum_{A} g_{A}^{4} C_{A}(i) \text{Tr}[S_{A}(r) m^{2}] + \sum_{p} (V_{i}^{T} V_{p})^{2} (m^{2})_{p}^{p}, \tag{36}
\]

\[
24 g_{A}^{4} MM^{*} C (i) S(R) \rightarrow 24 \sum_{A} g_{A}^{4} M_{A} M^{*}_{A} C_{A}(i) S_{A}(R)
+ 8 \sum_{p} \left[ (V_{i}^{T} M V_{p})(V_{i}^{T} M^{\dagger} V_{p}) + (V_{i}^{T} M M^{\dagger} V_{p})(V_{i}^{T} V_{p}) \right]
+ (V_{i}^{T} M^{\dagger} M V_{p})(V_{i}^{T} V_{p}), \tag{37}
\]

\[
48 g_{A}^{4} MM^{*} C(r)^{2} \rightarrow \sum_{A, B} g_{A}^{2} g_{B}^{2} C_{A}(r) C_{B}(r) \left[ 32 M_{A} M^{*}_{A} + 8 M_{A} M^{*}_{B} + 8 M_{B} M^{*}_{A} \right]
+ \sum_{A} g_{A}^{2} C_{A}(r) \left[ 32 M_{A} M^{*}_{A}(V_{r}^{T} V_{r}) + 16 M_{A}(V_{r}^{T} M^{\dagger} V_{r}) \right]
+ 16 M^{*}_{A}(V_{r}^{T} M V_{r}) + 32 (V_{r}^{T} M M^{\dagger} V_{r})
+ [32 (V_{r}^{T} M M^{\dagger} V_{r})(V_{r}^{T} V_{r}) + 16 (V_{r}^{T} M V_{r})(V_{r}^{T} M^{\dagger} V_{r})]. \tag{38}
\]

4. Comparison with other methods to include $U(1)$-mixing

4.1. General discussion

So far, several approaches to the SUSY $U(1)$-mixing conundrum have been proposed in the literature. Let us take a brief look at some of them and comment on their limitations as compared to the complete two-loop treatment advocated in this work.

As we have already mentioned in the Introduction, one can attempt to choose a convenient pair of bases in the $U(1)$-charge and gauge-field spaces for which the situation might simplify [10,20]. For instance, it is always possible to diagonalize the one-loop anomalous dimensions

\[
\gamma = \sum_{i} Q_{i} Q_{i}^{T} \tag{39}
\]

by means of a suitable $O_{1}$ rotation $Q_{i} \rightarrow O_{1} Q_{i} = Q_{i}^{T}$, see (7), so that $\gamma' = O_{1} \gamma O_{1}^{T}$ is diagonal. This, of course, inflicts a change on the gauge-coupling matrix $G \rightarrow O_{1} G$. However, if all the
relevant $U(1)$ gauge couplings happen to emanate from a single point, i.e., $G \propto 1$ at some scale, $O_1$ can be passed through $G$ and absorbed by a suitable redefinition of the gauge fields (9) where now $O_2 = O_1$. This way, the one-loop evolution of $G$ is driven by a diagonal $\gamma'$ and the initial condition $G \propto 1$ remains intact. Thus, no off-diagonalities emerge in this case and it is consistent to work with the usual RGEs for individual gauge couplings, one per each $U(1)$ factor.

This approach, however, is generally limited to the evolution with a complete $U(1)$ unification. This is very often not the case in practice, in particular in the GUTs in which the hypercharge is a non-trivial linear combination of the relevant Cartans, such as in left-right models based on the $SU(2)_L \otimes SU(2)_R \otimes U(1)_{B-L}$ gauge group, see Section 4.2. Moreover, not only gauge couplings but also the $U(1)$ gaugino soft masses should coincide at the unification scale otherwise the method fails in the soft sector already at the one-loop level. The point is that only then the generalized one-loop correlation between the gauge couplings and the gaugino masses

$$GM^{-1}G^T = \text{const.} \quad (40)$$

ensures the gaugino mass diagonality along the unification trajectory.

At the two-loop level more complicated structures such as higher powers of charges, gauge couplings, Yukawas, etc., enter the anomalous dimensions and, in general, there is no way to diagonalize simultaneously all the evolution equations. Though there is still a trick one can implement in the gauge sector if the $U(1)$ couplings do not unify [20], there is no general way out in the supersymmetric case for the gauginos as also discussed in [31]. Thus, a full-fledged two-loop approach as presented in this work is mandatory and, in fact, it turns out to be even technically indispensable if there happen to be more than two Abelian gauge groups as, for instance, in [21, 32] and many string-inspired constructions.

4.2. Simple illustrations

Let us illustrate the importance of the kinetic mixing effects in a couple of simple scenarios which exhibits all the salient features discussed above.

4.2.1. One-loop effects

4.2.1.1. Gauge couplings We shall consider the one-loop evolution of the gauge couplings in the SUSY $SO(10)$ model of Ref. [25] in which the unified gauge symmetry is broken down to the MSSM in three steps, namely, $SO(10) \rightarrow SU(3)_c \otimes SU(2)_L \otimes SU(2)_R \otimes U(1)_{B-L} \rightarrow SU(3)_c \otimes SU(2)_L \otimes U(1)_R \otimes U(1)_{B-L} \rightarrow$ MSSM; the corresponding breaking scales shall be denoted by $M_G$, $M_R$ and $M_{BL}$, respectively. Further details including the field contents at each of the symmetry breaking stages can be found in Ref. [25].

For our purposes, it is crucial that in this model the ratio $M_R/M_{BL}$ can be as large as $10^{10}$ and, hence, the $U(1)$ mixing effects become important. Note that even a short $SU(3)_c \otimes SU(2)_L \otimes SU(2)_R \otimes U(1)_{B-L}$ stage is sufficient to split the $g_R$ and the $g_{B-L}$ gauge couplings such that the extended gauge-coupling matrix $G$ at the $M_R$ scale is rather far from being proportional to the unit matrix. Thus, there is no way to choose the $O_1$ and $O_2$ rotation matrices such that both $G$ and $\gamma$

$$\gamma = N \begin{pmatrix} 15/2 & -1 \\ -1 & 18 \end{pmatrix} N \quad (41)$$
are simultaneously diagonalized. Here $N = \text{diag}(1, \sqrt{3/8})$ ensures the canonical normalization of the $B - L$ charge within the $SO(10)$ framework. Therefore, the one-loop evolution equation relevant to the $U(1)_R \otimes U(1)_{B-L}$ stage has to be matrix-like and reads in the Abelian sector

$$\frac{d}{dt} A^{-1} = -\gamma,$$

(42)

where $A^{-1} = 4\pi (G G^T)^{-1}$ and $t = \log(\mu/\mu_0)/2\pi$.

The reason that the $U(1)_R \otimes U(1)_{B-L}$ stage can be so long has to do with the fact that this gauge symmetry is broken by neutral components of an $SU(2)_R$ doublet pair, namely, $(1, 1, +\frac{1}{2}, -1) \oplus (1, 1, -\frac{1}{2}, +1) = \chi_R^0 \oplus \bar{\chi}_R^0$ which are full SM singlets, and as such they do not affect the low-energy value of $\alpha_Y^{-1}$. Indeed, the would-be change inflicted on $\alpha_Y^{-1}$ by the presence or absence of $\chi_R^0 \oplus \bar{\chi}_R^0$ is given by

$$\Delta \alpha_Y^{-1} = p_Y^T \Delta A^{-1}(M_{BL}) p_Y \propto p_Y^T \Delta \gamma p_Y = 0 \quad \text{(43)}$$

where $p_Y^T = (\sqrt{3/5}, \sqrt{2/5})$ are the coordinates of the MSSM hypercharge in the $U(1)_R \otimes U(1)_{B-L}$ algebra and $\Delta \gamma$ denotes the relevant change of the $\gamma$ matrix. Therefore, at the one-loop level, the position of the $M_{BL}$ scale is not constrained by the low-energy data and, hence, barring other phenomenological constraints, it can be pushed as close to the MSSM scale $M_S$ as desired.

However, this simple argument works only if the $U(1)$-mixing effects are properly taken into account. Remarkably, if they are simply neglected, $\Delta \gamma$ receives only diagonal entries and $\alpha_Y^{-1}(M_Z)$ becomes a function of $M_{BL}$. Moreover, stretching the $M_{BL}-M_R$ range to maximum, the erroneous shift inflicted on $\alpha_Y^{-1}(M_Z)$ can become as large as 4 percent as can be seen by comparing Figs. 1 and 2. Alternatively, in order to retain the desired value of $\alpha_Y^{-1}(M_Z)$, one would have to re-adjust $M_R$ by several orders of magnitude, cf., Fig. 2. This, however, could
Fig. 2. The same like in Fig. 1 but without the kinetic mixing effects taken into account. With the GUT-scale boundary condition and $M_R$ as above, the low-energy value of $\alpha^{-1}_Y$, namely, $\alpha^{-1}_Y(M_Z) = 62.51$ (black solid lines) differs from the one obtained in the full-fledged calculation by as much as 4 percent. Alternatively, if one attempts to obtain the right value of $\alpha^{-1}_Y(M_Z)$ by adjusting the $SU(2)_R$-breaking scale, the new $M'_R$ scale must be shifted with respect to the correct $M_R$ by as much as 4 orders of magnitude (in blue dashed lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

have a large impact on, e.g., the MSSM soft spectrum [33], and, in more general constructions, also on $M_G$ and $\alpha_G$, with ramifications for $d = 6$ proton decay etc.

Finally, let us note that the “rotated-basis” method discussed in brief in Section 4.1 is only partially successful because the $g_R$ and $g_{B-L}$ gauge couplings do not coincide at the $M_R$. Indeed, the value of $\alpha^{-1}_Y(M_Z)$ obtained this way, namely, $\alpha^{-1}_Y(M_Z) = 60.93$, is closer to the correct value than that received with no mixing at all, but still some 2% off the correct value.

4.2.1.2. Gaugino masses In order to fully appreciate the method advocated in this work, we should look at the interplay between the gauge and the soft sector. For example, at one loop-level, a simple illustration is by Eq. (40) which ties the gauge couplings $G$ together with the gaugino soft masses $M$. Consequently, the bino mass obeys at the scale $M_S$

$$M_Y(M_S) = \frac{\alpha_Y(M_S)}{\alpha_G} p_Y m_{1/2} p_Y,$$  \hspace{1cm} (44)

where $m_{1/2}$ is the GUT-scale gaugino soft mass matrix. From Eq. (44) we see that the ratio $M_Y(M_S)/\alpha_Y(M_S)$ depends on whether one includes the mixing effects or not as already noticed in Ref. [34]. Note that with non-universal initial conditions, i.e. $m_{1/2}$ not being proportional to the unit matrix, the $p_Y^T m_{1/2} p_Y$ term mixes up all entries of $m_{1/2}$. Moreover, in the special case that the Abelian gauge couplings unify, even the one-loop gaugino sector evolution can be fully accounted for by the “rotated-basis” trick.

4.2.2. Two-loop effects

At two-loop level our method becomes already important in cases with gauge coupling unification at a certain scale. We illustrate this by taking as an example the model presented in Ref. [27] where an intermediate $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \otimes U(1)_{B-L}$ gauge symmetry is assumed to originate from a grand-unified framework. We assume two cases: (i) full gauge
coupling unification at $2 \times 10^{16}$ GeV and (ii) a small difference of 5% between the two $U(1)$ couplings caused by possible GUT-scale threshold effects. In the gaugino sector we assume universal boundary conditions in both cases, but the effect gets even stronger if one considers in addition threshold effects in the gaugino sector as well.

The results are given in Table 1. Remarkably, besides the expected equivalence of the “rotated-basis” method and the full-fledged calculation at the one-loop level, the relevant effective hypercharge gauge coupling turns out to be identical to the one obtained even at two loop-level if exact gauge coupling unification is assumed. The reason is, that all additional states not present in the MSSM are charged only with respect to $U(1)_{B-L}$ but are neutral under the MSSM gauge group. In the gaugino sector the first deviations show up already in this case which however are only at the per-mile level. In case that one includes also threshold corrections at the GUT-scale the effects are at the percent level leading to shifts in the masses potentially measurable already at the LHC.

Last but not least we remark, that the effects would be even larger if the $U(1)_{Y}$ would result from the breaking of $U(1)_{R} \otimes U(1)_{B-L}$ as discussed in the previous example.

5. Conclusions and outlook

In this work, we have discussed the structure of the renormalization group equations in softly-broken supersymmetric models with more than a single Abelian gauge group. Indeed, with multiple $U(1)$ gauge factors at play, the effects of kinetic mixing among the Abelian gauge fields must be taken into account in order to keep the theory renormalizable.

Though, formally, the evolution equations available in the literature do not exhibit any obvious pathologies if such subtleties are not taken into account, the calculations based on these formulas are in general incomplete and, thus, the results are internally inconsistent. This is even more
pronounced in the context of SUSY models because it affects also the evolution of the soft SUSY parameters, in particular the evolution of the gaugino mass parameters.

Remarkably enough, the issue of the $U(1)$ mixing in the softly-broken supersymmetric gauge theories has never been addressed in full generality, even at one loop. The main aim of the current study was to fill this gap and provide a fully self-consistent method for dealing with the renormalization group evolution of the gauge couplings and the soft SUSY-breaking parameters up to the two-loop level.

To this end, we have studied in detail the existing two-loop renormalization group equations valid for the case of at most a single Abelian gauge factor at play given in Ref. [10,11] and extended these to account for the most general case of a gauge group with any number of $U(1)$ factors.

In particular, we have argued that all the $U(1)$ mixing effects can be consistently included if the gauge couplings and the soft SUSY-breaking gaugino masses associated to the individual Abelian gauge-group factors are generalized to matrices and these are then substituted into the formulae in [10,11] in a specific manner. This, however, is a highly non-trivial enterprise, mainly due to the non-commutativity of the relevant matrix-like structures, and a number of ambiguities had to be resolved. In this respect, the residual reparametrization invariance of the covariant derivative associated to the redefinition of the Abelian gauge fields turned out to be a very useful tool, yet in many cases one had to resort to a detailed analysis of the relevant Feynman diagrams.

The general method has been illustrated for two cases: (i) at one-loop level where due to an breaking of the original group two $U(1)$ factors emerge with different gauge couplings and (ii) at the two-loop-level in a model where gauge coupling unifications occurs but where threshold corrections are taken into account. In both case we obtain effects in the percent range and we remark, that none of the previously proposed partial treatments can account for the full effects.

Last but not least, let us stress again that our results are completely generic and, as such, they do not require any specific assumptions about the charges of the chiral multiplets in the theory and/or the boundary conditions applied to the relevant gauge couplings. This makes the framework very suitable for implementation into computer algebraic codes calculating two-loop renormalization group equations in softly-broken supersymmetric gauge theories such as SARAH [35–37] and Susyno [38].

**Acknowledgements**

We thank M. Hirsch for fruitful discussion and F. Braam and J. Reuter for interesting insights about the different approaches to the kinetic mixing problem. In addition, we thank B. O’Leary for reading through the manuscript.

The work of M.M. was supported by the Marie Curie Intra European Fellowship within the 7th European Community Framework Programme FP7-PEOPLE-2009-IEF, contract number PIEF-GA-2009-253119, by the EU Network grant UNILHC PITN-GA-2009-237920, by the Spanish MICINN grants FPA2008-00319/FPA and MULTIDARK CAD2009-00064 (Consolider-Ingenio 2010 Programme) and by the Generalitat Valenciana grant Prometeo/2009/091. W.P. and F.S. have been supported by the German Ministry of Education and Research (BMBF) under contract No. 05H09WWE. R.F. acknowledges the support provided by the Fundação para a Ciência e a Tecnologia under the grant SFRH/BD/47795/2008.
Appendix A. Renormalization of QED $\otimes$ QED

In this appendix, we comment in more detail on renormalization of abelian gauge theories, focusing on the simplest non-trivial case exhibiting the effects of kinetic mixing, namely the “QED-squared” scenario featuring two independent abelian gauge groups $U(1) \otimes U(1)$.

Let us start with the basic bare Lagrangian of QED including an explicit kinetic-mixing term

$$\mathcal{L}_B = \bar{\psi}_{i,B}(i\partial - m_i)\psi_{i,B} - \bar{\psi}_{i,B}Q_i^T G_A \bar{A}_B \psi_{i,B} - \frac{1}{4} F_{B\mu\nu}^2 F_B^{\mu\nu}. \quad (A.1)$$

Here $\psi_i$ are the relevant matter fields (whose number must be in general equal to or greater than the number of the abelian gauge factors otherwise there is no way to distinguish among all stipulated $U(1)$ factors), $A$ stands for a 2-component vector (in the group space) comprising the gauge fields associated to different $U(1)$ factors, $G$ stands for a (so far formal, i.e., diagonal) $2 \times 2$ matrix containing the relevant pair of gauge couplings and $\xi$ is a symmetric and real $2 \times 2$ matrix parametrizing the gauge-kinetic form.

A.1. Scheme A: A non-canonical gauge propagator and diagonal gauge couplings

Leaving $\xi_B$ in the game, one defines the renormalized and the counterterm Lagrangians as

$$\mathcal{L} = \bar{\psi}_i(i\partial - m_i)\psi_{i} - \bar{\psi}_i Q_i^T G_A \psi_{i} - \frac{1}{4} F_{\mu\nu}^2 F_{\mu\nu}, \quad (A.2)$$

$$\delta \mathcal{L} = i \bar{\psi}_i \delta Z_{\psi_i}/\partial \psi_{i} - \bar{\psi}_i \delta Z_m m_i \psi_{i} - \bar{\psi}_i Q_i^T \delta Z_G G_B \psi_{i} - \frac{1}{4} F_{\mu\nu} \delta \xi F_{\mu\nu}, \quad (A.3)$$

where

$$Q_i^T \delta Z_G G = Z_{\psi_i} Q_i^T G_B Z_A^{-1/2} - Q_i^T G, \quad (A.4)$$

$$\delta \xi = Z_A^{-1/2} \xi_B Z_A^{-1/2} - \xi, \quad (A.5)$$

and $\delta Z_m$ and $\delta Z_{\psi_i}$ are unimportant for our considerations. These counterterms are fixed by the renormalization conditions so that they render the renormalized Green’s functions of the theory UV-finite. For a diagonal $Z_A$ and for any fixed $\xi_B$, the off-diagonal entries in $\delta \xi$ cannot be matched by the right-hand side of Eq. (A.5) unless $\xi$ is a dynamical quantity. Remarkably, in this scheme the gauge coupling can be retained in a diagonal form throughout the RG evolution. This is because the relation between the bare and the renormalized couplings $G_B = G Z_A^{-1/2}$ can be brought into the form (trading $Z_A$ for $\xi$ and $\delta \xi$)

$$G_B \xi_B^{-1} G_T = G Z_A^{-1/2} \xi_B^{-1} Z_A^{-1/2} G_T = G(\xi + \delta \xi)^{-1} G^T \quad (A.6)$$

(which holds to all orders in perturbation theory) from where it is clear that any non-diagonal entry of the RHS of the evolution equation for $G$ can be absorbed into $\xi$. So, in this scheme, $\xi$ is a dynamical quantity while $G$ can be kept diagonal.

A.2. Scheme B: A canonical gauge propagator and non-diagonal gauge couplings

If, instead, $\xi$ is absorbed by a suitable gauge-field redefinition $A \rightarrow \xi^{1/2} A \equiv \tilde{A}$ into matrix for the coupling constants, one is left with
\[ \mathcal{L} = \bar{\psi}_i (i/\partial - m_i) \psi_i - \bar{\psi}_i Q_i^T \tilde{G} / \tilde{A} \psi_i - \frac{1}{4} \tilde{F}_{\mu\nu} \tilde{F}^{\mu\nu}, \]
\[ \delta \mathcal{L} = i \bar{\psi}_i \delta Z \psi_i / \partial \psi_i - \bar{\psi}_i \delta Z m_i \psi_i - \bar{\psi}_i Q_i^T \delta Z \tilde{G} / \tilde{A} \psi_i - \frac{1}{4} \tilde{F}_{\mu\nu} \delta Z \tilde{A} \tilde{F}^{\mu\nu}, \] (A.7)

where
\[ Q_i^T \delta Z \tilde{G} = Z \psi_i Q_i^T \tilde{G} B Z_\tilde{A}^{1/2} - Q_i^T \tilde{G}, \] (A.8)
\[ \delta Z \tilde{A} = (Z_\tilde{A}^{1/2})^T Z_\tilde{A}^{1/2} - 1, \] (A.9)
with \( Z_\tilde{A}^{1/2} = \xi^{-1/2} Z_\tilde{A}^{1/2} \xi^{1/2} \) and, as before,
\[ \tilde{G}_B = \tilde{G} Z_\tilde{A}^{-1/2}. \] (A.10)

It is again clear that the non-diagonality inflicted on \( Z \) by the renormalization conditions renders the RHS of the gauge-coupling evolution equation non-diagonal. However, in this scheme, \( \xi \) has been swallowed by the gauge-field renormalization counterterm and, as such, does not need to be treated as an extra dynamical quantity. In other words, the whole effect is accounted for by the off-diagonal form of the generalized gauge coupling \( \tilde{G} \).

**Appendix B. Recapitulation of the two-loop RGEs for simple groups and their products with at most one \( U(1) \)**

**B.1. Case A: Simple gauge group**

For completeness we display here the RGEs in the case of a simply gauge group based on [10–12]. For a general \( N = 1 \) supersymmetric gauge theory with superpotential
\[ W(\Phi) = L_i \Phi_i + \frac{1}{2} \mu^{ij} \Phi_i \Phi_j + \frac{1}{6} Y^{ijk} \Phi_i \Phi_j \Phi_k, \] (B.1)
the soft SUSY-breaking scalar terms are given by
\[ V_{\text{soft}} = \left( S^{ij} \phi_i + \frac{1}{2} b^{ij} \phi_i \phi_j + \frac{1}{6} h^{ijk} \phi_i \phi_j \phi_k + \text{c.c.} \right) + (m^2)^{ij} \phi_i \phi_j^* + \frac{1}{2} \lambda a \lambda a. \] (B.2)

Here we will follow [10] and assume that repeated indices are summed over. Note also that lowered indices imply conjugation (e.g., \( Y^{ijk} \equiv Y^{ijk*} \)). In the notation defined in Section 3, the anomalous dimensions of the chiral superfields are given by
\[ \gamma_i^{(1)} j = \frac{1}{2} Y_{i pq} Y^{jpq} - 2 \delta_j^i g^2 C(i), \] (B.3)
\[ \gamma_i^{(2)} j = g^2 Y_{i pq} Y^{jpq} \left[ 2 C(p) - C(i) \right] - \frac{1}{2} Y_{imn} Y^{npq} Y_{pqr} Y^{mrj} + 2 \delta_j^i g^4 \left[ C(i) S(R) + 2 C(i)^2 - 3 C(G) C(i) \right], \] (B.4)
and the \( \beta \)-functions for the gauge couplings are given by
\[ \beta_g^{(1)} = g^3 \left[ S(R) - 3 C(G) \right], \] (B.5)
\[ \beta_g^{(2)} = g^5 \left\{ -6 \left[ C(G) \right]^2 + 2 C(G) S(R) + 4 S(R) C(R) \right\} - g^3 Y^{ijk} Y_{ijk} C(k)/d(G). \] (B.6)

The corresponding RGEs are defined as
\[
\frac{dg}{dt} = \frac{1}{16\pi^2} \beta_g^{(1)} + \frac{1}{(16\pi^2)^2} \beta_g^{(2)}.
\]  

(B.7)

Here, we used \( t = \ln Q \), where \( Q \) is the renormalization scale. The \( \beta \)-functions for the superpotential parameters can be obtained by using superfield technique. The obtained expressions are

\[
\beta_Y^{ij} = Y_{ij} \left[ \frac{1}{16\pi^2} Y_{p}^{(1)k} + \frac{1}{(16\pi^2)^2} Y_{p}^{(2)k} \right] + (k \leftrightarrow i) + (k \leftrightarrow j),
\]

(B.8)

\[
\beta_{\mu}^{ij} = \mu_{ij} \left[ \frac{1}{16\pi^2} Y_{p}^{(1)j} + \frac{1}{(16\pi^2)^2} Y_{p}^{(2)j} \right] + (j \leftrightarrow i),
\]

(B.9)

\[
\beta_L = L_{ij} \left[ \frac{1}{16\pi^2} Y_{p}^{(1)i} + \frac{1}{(16\pi^2)^2} Y_{p}^{(2)i} \right].
\]

(B.10)

The expressions for trilinear, soft-breaking terms are

\[
\frac{d}{dt} h_{ij}^{(1)jk} = \frac{1}{16\pi^2} \left[ \beta_h^{(1)} \right]_{ij}^{jk} + \frac{1}{(16\pi^2)^2} \left[ \beta_h^{(2)} \right]_{ij}^{jk},
\]

(B.11)

with

\[
\left[ \beta_h^{(1)} \right]_{ij}^{jk} = \frac{1}{2} h_{ij} Y_{lmn} Y_{mnp} - Y_{ijl} Y_{lmn} h_{mnp}
- 2(h_{ij} - 2MY_{ij}) g^2 C(k) + (k \leftrightarrow i) + (k \leftrightarrow j),
\]

(B.12)

\[
\left[ \beta_h^{(2)} \right]_{ij}^{jk} = -\frac{1}{2} h_{ij} Y_{lmn} Y_{mnp} Y_{pqr} Y_{mrr} - Y_{ijl} Y_{lmn} Y_{mnp} Y_{pqr} h_{mrr}
- Y_{ijl} Y_{lmn} h_{mnp} Y_{pqr} Y_{mrr} + (h_{ij} Y_{ljq} Y_{pjk} + 2Y_{ijl} Y_{lpq} h_{pjk}
- 2MY_{ij} Y_{lpq} Y_{pjk}) g^2 [2C(p) - C(k)] + (2h_{ijk} - 8MY_{ijk})
\times g^4 [C(k)S(R) + 2C(k) - 3C(G)C(k)] + (k \leftrightarrow i) + (k \leftrightarrow j).
\]

(B.13)

For the bilinear soft-breaking parameters, the expressions read

\[
\frac{d}{dt} b_{ij} = \frac{1}{16\pi^2} \left[ \beta_b^{(1)} \right]_{ij} + \frac{1}{(16\pi^2)^2} \left[ \beta_b^{(2)} \right]_{ij},
\]

(B.14)

with

\[
\left[ \beta_b^{(1)} \right]_{ij} = \frac{1}{2} b_{ij} Y_{lmn} Y_{mnj} + \frac{1}{2} Y_{ijl} Y_{lmn} b_{mn} + \mu_{ij} Y_{lmn} h_{mijn}
- 2(b_{ij} - 2M\mu_{ij}) g^2 C(i) + (i \leftrightarrow j),
\]

(B.15)

\[
\left[ \beta_b^{(2)} \right]_{ij} = -\frac{1}{2} b_{ij} Y_{lmn} Y_{pqn} Y_{pqr} Y_{mri} - \frac{1}{2} Y_{ijl} Y_{lmn} \mu_{mr} Y_{pqr} h_{pqn}
- \mu_{ij} Y_{lmn} h_{mpq} Y_{pqr} Y_{mri} - \mu_{ij} Y_{lmn} Y_{pqn} Y_{pqr} h_{mri}
- \frac{1}{2} Y_{ijl} Y_{lmn} b_{mr} Y_{pqr} Y_{pqn}
+ 2Y_{ijl} Y_{lpq} (b_{pq} - \mu_{pq} M) g^2 C(p) + (b_{ij} Y_{lpq} Y_{pqj} + 2\mu_{ij} Y_{lpq} h_{pqj}
- 2\mu_{ij} Y_{lpq} Y_{pqj} M) g^2 [2C(p) - C(i)] + (2b_{ij} - 8\mu_{ij} M)
\times g^4 [C(i)S(R) + 2C(i) - 3C(G)C(i)] + (i \leftrightarrow j).
\]

(B.16)

Finally, the RGEs for the linear soft-breaking parameters are
\[ \frac{d}{dt} S^i = \frac{1}{16\pi^2} [\beta_S^{(1)}]^i + \frac{1}{(16\pi^2)^2} [\beta_S^{(2)}]^i, \quad (B.17) \]

with
\[ [\beta_S^{(1)}]^i = \frac{1}{2} Y_{ipln} S^n + L^p Y_{plhn} + \mu^i k^n l + 2 Y^{ikp} (m^2)^i_p \mu_{kkl} + h^{ikl} b_{kl}, \quad (B.18) \]
\[ [\beta_S^{(2)}]^i = 2 g^2 C(l) Y^{ikl} Y_{plk} S^n - \frac{1}{2} Y^{ikq} Y_{qst} Y_{plk} S^n - 4 g^2 C(l) (Y^{ikl} M - h^{ikl}) Y_{plk} L^p \]
\[ - [Y^{ikq} Y_{qst} h^{ilst} Y_{plk} + h^{iql} Y_{qst} Y_{lkl} Y_{plk}] L^p - 4 g^2 C(l) Y_{jln} (\mu^{lnt} B - h^{lnt}) \mu^{ij} \]
\[ - [Y_{jln} h^{qst} Y_{lkl} Y_{plk} + Y_{jln} Y_{qst} Y_{lkl} Y_{plk}] L^p + 4 g^2 C(l) (2 Y^{ikl} \mu_{kkl} M) \]
\[ - Y^{ikl} b_{kl} M - h^{ikl} \mu_{kkl} M^* + h^{ikl} b_{kl} + Y^{ipl} (m^2)_{p} \mu_{kkl} + Y^{ikp} (m^2)_{p} \mu_{kkl} \]
\[ Y^{iqp} (m^2)^k Y_{qst} Y_{lkl} + Y^{ikq} Y_{qst} Y_{lkl} (m^2)^l Y_{p} \mu_{kkl} + Y^{ikp} (m^2)^q M Y_{qst} Y_{lkl} \]
\[ + 2 Y^{ikq} Y_{qst} (m^2)^l Y_{lkl} + Y^{ikq} h_{qst} h^{lkl} \mu_{kkl}. \quad (B.19) \]

With these results, the list of the \( \beta \)-functions for all couplings is complete. Now, we turn to the RGEs for the gaugino masses, squared masses of scalars and vacuum expectation values. The result for the gaugino masses is
\[ \frac{d}{dt} M = \frac{1}{16\pi^2} [\beta_M^{(1)}] + \frac{1}{(16\pi^2)^2} [\beta_M^{(2)}], \quad (B.20) \]

with
\[ \beta_M^{(1)} = g^2 [2 S(R) - 6 C(G)] M, \quad (B.21) \]
\[ \beta_M^{(2)} = g^4 \left\{ -24 [C(G)]^2 + 8 C(G) S(R) + 16 S(R) C(R) \right\} M \]
\[ + 2 g^2 [h^{ijk} - M Y^{ijk}] Y_{ijkl} C(k)/d(G). \quad (B.22) \]

The one- and two-loop RGEs for the scalar mass parameters read
\[ \frac{d}{dt} m_{ij} = \frac{1}{16\pi^2} [\beta_{m_{ij}}] + \frac{1}{(16\pi^2)^2} [\beta_{m_{ij}}], \quad (B.23) \]

with
\[ [\beta_{m_{ij}}]^j_i = \frac{1}{2} Y_{ipq} Y^{pqn} (m^2)_{n}^j + \frac{1}{2} Y^{ipq} Y_{pqn} (m^2)_{i}^n + 2 Y_{ipq} Y^{ipr} (m^2)^q_r + h_{ipq} h^{ipq} \]
\[ - 8 \delta_{ij} M M^* g^2 C(i) + 2 g^2 \text{Tr}[A^2 m^2], \quad (B.24) \]
\[ [\beta_{m_{ij}}]^j_i = - \frac{1}{2} (m^2)^l_i Y_{lmn} Y_{nr} Y_{pqr} Y_{pqn} - \frac{1}{2} (m^2)^l_i Y_{lmn} Y_{nnr} Y_{pqr} Y_{pqn} \]
\[ - h_{ilm} Y^{jin} Y_{npq} h_{mpq} - h_{ilm} Y^{jin} (m^2)^l_i Y_{npq} Y_{pqm} \]
\[ - Y_{ilm} Y^{jin} h_{npq} h_{mpq} - 2 Y_{ilm} Y^{jin} Y_{npq} Y_{mpr} (m^2)^q_r - h_{ilm} Y^{jin} Y_{npq} Y_{mpq} \]
\[ - Y_{ilm} Y^{jin} (m^2)^l_i Y_{npq} Y_{pqm} - Y_{ilm} Y^{jin} h_{npq} Y_{mpq} \]
\[ + [(m^2)^l_i Y_{ipq} Y^{ipq} + Y_{ipq} Y_{pqm} (m^2)^j_i + 4 Y_{ipq} Y^{ipl} (m^2)^q_i + 2 h_{ipq} h^{ipq} \]
\[ - 2 h_{ipq} Y^{ipq} M - 2 Y_{ipq} h^{ipq} M^* + 4 Y_{ipq} Y^{ipq} M M^* g^2 C(p) + C(q) - C(i)] \]
\[-2g^2 t^A_i (t^A m^2)^l_r Y_{lpq} Y^{rpq} + 8g^4 t^A_i \text{Tr}[t^A C(r)m^2] \]
\[+ \delta^i_j g^4 M M^\dagger [24C(i)S(R) + 48C(i)^2 - 72C(G)C(i)] \]
\[+ 8\delta^i_j g^4 C(i) (\text{Tr}[S(r)m^2] - C(G)MM^\dagger). \tag{B.25} \]

The RGEs for a VEV \( v^i \) are proportional to the anomalous dimension of the chiral superfield whose scalar component receives the VEV

\[
\frac{d}{dt} v^i = \nu^i \left[ \frac{1}{16\pi^2} \gamma^{(1)}_p + \frac{1}{(16\pi^2)^2} \gamma^{(2)}_p \right]. \tag{B.26} \]

B.2. Product groups with at most one \( U(1) \)

To generalize the formulas above to the case of a direct product of gauge groups, the following substitution rules are needed [10]. Note, we give these replacements here only for completeness and they are not sufficient in the case of several \( U(1) \) gauge groups, see Section 3 for the necessary extensions. For the \( \beta \) functions of gauge couplings and gauginos the rules are

\[
g^3 C(G) \rightarrow g^3_a C(G_a), \tag{B.27} \]
\[
g^3 S(R) \rightarrow g^3_a S_a(R), \tag{B.28} \]
\[
g^5 C(G)^2 \rightarrow g^5_a C(G_a)^2, \tag{B.29} \]
\[
g^5 C(G)S(R) \rightarrow g^5_a C(G_a)S_a(R), \tag{B.30} \]
\[
g^5 S(R)C(G) \rightarrow \sum_b g^3_a g^2_b S_a(R)C_b(R), \tag{B.31} \]
\[
16g^4 S(R)C(R)M \rightarrow 8 \sum_b g^2_a g^2_b S_a(R)C_b(R)(M_a + M_b). \tag{B.32} \]
\[
g^3 C(k)/d(G) \rightarrow g^3_a C_a(k)/d(G_a). \tag{B.33} \]

For all the other \( \beta \) functions, we need

\[
g^2 C(r) \rightarrow \sum_a g^2_a C_a(r), \tag{B.34} \]
\[
g^4 C(r)S(R) \rightarrow \sum_a g^4_a C_a(r)S_a(R), \tag{B.35} \]
\[
g^4 C(r)C(G) \rightarrow \sum_a g^4_a C_a(r)C(G_a), \tag{B.36} \]
\[
g^4 C(r)^2 \rightarrow \sum_a \sum_b g^2_a g^2_b C_a(r)C_b(r), \tag{B.37} \]
\[
48g^4 MM^\dagger C(i)^2 \rightarrow \sum_a \sum_b g^2_a g^2_b C_a(i)C_b(i)(32M_a M^\dagger_b + 8M_a M^\dagger_a + 8M_b M^\dagger_b), \tag{B.38} \]
\[
g^2 t^A_i \text{Tr}(t^A m^2) \rightarrow \sum_a g^2_a (t^A_i)\dagger \text{Tr}(t^A_a m^2), \tag{B.39} \]
\[
g^2 t^A_i (t^A m^2)_{l_r} Y_{lpq} Y^{rpq} \rightarrow \sum_a g^2_a (t^A_i)\dagger (t^A_a m^2)_{l_r} Y_{lpq} Y^{rpq}. \tag{B.40} \]
\[ g^4 t_i^A \text{Tr}[t^A C(r)m^2] \rightarrow \sum_a \sum_b g_a^2 g_b^2 (t^A)^i_j \text{Tr}[t_a^A C_b(r)m^2], \quad (B.41) \]
\[ g^4 C(i) \text{Tr}[S(r)m^2] \rightarrow \sum_a g_a^4 C_a(i) \text{Tr}[S_a(r)m^2]. \quad (B.42) \]

**Appendix C. Obtaining the substitution rules**

In this appendix, we illustrate in more detail the methods used throughout the derivation of the substitution rules given in Section 3.

**C.1. The role of the \( V_i \) vectors and the \( M \) matrix**

As mentioned in the text, the \( U(1) \) gauge coupling matrix \( G \) and the charge vectors \( Q_i \) of the chiral superfields \( \Phi_i \) appear always through the combination \( V_i = G^T Q_i \). The only exception are the RGEs of \( G \), where there should be a leading free \( G \). For example, the \( \psi^\dagger \psi A^\mu_\alpha \) vertex is proportional to \( V_i^a \) (component \( a \) of the vector \( V_i \)),

\[ \psi_i \quad \Downarrow \quad V_i^a \quad \overbrace{A^\mu_\alpha}^{\text{C}} \quad \psi_i \]

Similarly the vertices \( \phi^\dagger \phi A^\mu_\alpha \), \( \phi^\dagger \phi A^\mu_\alpha A^\nu_\beta \), \( \phi^\dagger \psi \lambda^a \) and the Yukawa independent part of \( \phi^\dagger \phi \phi^\dagger \phi \) are proportional to \( V_i^a \), \( V_i^a V_i^b \), \( V_i^a \) and \( V_i^T V_j \) respectively. In addition, there is to consider the \( U(1) \) gaugino mass matrix \( M \),

\[ M_{ab} \]

where \( M_{ab} \) is the \( a, b \) component of \( M \).

**C.2. RGEs with no \( U(1) \) indices**

Only diagrams underlying the RGEs for \( G \) and \( M \) contain external \( U(1) \) gauge bosons/gauginos. As such, in all other equations, while vectors \( V_i \) and the matrix \( M \) may be present, they must be in combinations that are scalars with no free \( U(1) \) indices.

Consider \( M g^2 C(i) \) appearing in the one-loop RGE of the bilinear scalar soft terms \( b^{ij} \), which is to be replaced by \( M A g^2 A^\dagger C(i) + V_i^T M V_i \). The simple groups contribution does not interest us though, so we shall neglect it. We can see that \( V_i^T M V_i \) is the only structure that can generalize the expression \( M g^2 C(i) = M g^2 y_i^2 \) for one \( U(1) \) group only. Observe also the contraction of the \( U(1) \) indices in the expression — it comes from the possibility of having any of the \( U(1) \) gauginos in the internal lines of the contributing diagram,
The amplitude is proportional to $\sum_{a,b} V^a_i M_{ab} V^b_j \mu^{ij} = \mu^{ij} V^T_i M V_j$. Note that, for any pair of values $i, j$ the gauge symmetry forces $\mu^{ij} = 0$ unless $V_i + V_j = 0$ which means that $\mu^{ij} V_j = -\mu^{ij} V_i$ so the amplitude of the diagram is indeed proportional to $V^T_i M V_i$.

This requirement that expressions with $V$’s and $M$’s must form scalars is enough to derive Eqs. (20)–(25), (28)–(35) and (38) from the existing substitution rules for gauge groups with multiple factors. We are left with the terms $g^4 C(r) S(R)$, $M g^4 C(r) S(R)$, $g^4 C(i) \text{Tr}[S(r) m^2]$ and $24 g^4 M M^\star C(i) S(R)$. Note that one can write $S(R)$ as $\text{Tr}[S(r)]$ in the notation of Ref. [10], so in all four cases there is a sum over field components of chiral superfields. For diagrams with up to two-loops and with no external gauginos nor gauge bosons, the factors $S(R)$ and $\text{Tr}[S(r) m^2]$ can only come from the following sub-diagrams:

![Sub-diagrams](image)

(C.4)

Take for example $g^4 C(r) S(R) \sim \sum_p g^4 y^2_r y^2_p$. The reason why one cannot immediately generalize this expression to include $U(1)$ mixing effects is because in theory it could take the form $\sum_p (V^T_r V_p)(V^T_r V_p)$ or $\sum_p (V^T_r V_r)(V^T_p V_p)$. But looking at the above diagrams, such ambiguities go away because in all cases the $V$’s which are summed over (the $V_p$’s) do not contract with each other. They contract with something else at the other ends of the gauge boson/gaugino lines. As such, there are no $V^T_p V_p$’s in these expressions and with this piece of information, combined with the known rules for a gauge group with multiple factors, Eqs. (26), (27) and (36) follow. The substitution rule given in Eq. (37) for $24 g^4 M M^\star C(r) S(R)$ appearing in the two-loop equation of the soft scalar masses is more complicated since the placement of the $M$, $M^\dagger$ gaugino mass matrices between these $V$’s is relevant. Nevertheless, from the following diagrams we can calculate that the $U(1)$’s contribution to this term is $8 \sum_p [(V^T_i M V_p)(V^T_i M^\dagger V_p) + (V^T_i M M^\dagger V_p)(V^T_i V_p) + (V^T_i M^\dagger M V_p)(V^T_i V_p)]$:

![Diagram](image)

(C.5)

---

7 It is conceivable that they could come also from diagrams with one $\phi^* \phi^* \phi \phi$ vertex, but we may choose an appropriate gauge, the Landau gauge, where these are 0 because an external scalar line always couples to a gauge bosons at a three point vertex.
C.3. RGEs with U(1) indices

The RGEs for \( G \) and \( M \) are the only ones with free U(1) indices. For the beta functions of the gaugino masses, we will be interested in looking at diagrams with two incoming gauginos. As for the coupling constant, due to the Ward identities, the contributing diagrams are those with two external gauge bosons. From the amplitude of these diagrams we still have to add a \( G \) factor in order to obtain \( \beta_G \).

\[
\sum_{a'} G_{aa'} \times A_{a'}^\mu \times \frac{d}{dt} A_b^\nu \tag{C.8}
\]

Note that all the terms in \( \beta_G \) must be of the form \( GV_i^T (\cdots) V_j \) for some \( i, j \) as mentioned in the text. These qualitative considerations suffice for our purposes.

Keep also in mind that

(i) the RGEs are invariant under the set of transformations \( G \rightarrow O_1 G O_2^T \), \( V_i \rightarrow O_2 V_i \), \( M \rightarrow O_2 M O_2^T \) for any orthogonal matrices \( O_1, O_2 \);

(ii) \( M \) is a symmetric matrix and so must be \( \frac{dM}{dt} \).

Taken together, these considerations allow us to deduce Eqs. (13)–(18) (Eq. (12) is trivial).

We shall exemplify this for the case of \( 16g^4S(R)C(R)M \) which for multiple factor groups is replaced by \( 8 \sum g_a^2 g_b^2 S_a(R)C_b(R)(M_a + M_b) \) in the RGEs of \( M_a \). This is the same as \( 8 \sum g_a^2 g_b^2 S_a(p)C_b(p)(M_a + M_b) \). Groups \( a \) and \( b \) are independent so the expressions \( \sum_b g_b^2 C_b(p), \sum_b M_b g_b^2 C_b(p) \) are decoupled from \( g_a^2 S_a(p), M_a g_a^2 S_a(p) \). Inclusion of U(1) mixing effects in the first pair of expressions is easy because there are no free U(1) indices: \( \sum_b g_b^2 C_b(p) \rightarrow \sum_B g_B^2 C_B(p) + V_p^T V_p \), and \( \sum_b M_b g_b^2 C_b(p) \rightarrow \sum_B M_B g_B^2 C_B(p) + V_p^T M V_p \).

If the group \( a \) is an U(1), then in case of a single U(1) this corresponds to \( g_a^2 S_a(p) \rightarrow V_p V_p^T \). Similarly, the only symmetric matrix expression which respects the \( O_2 \) symmetry that can generalize \( M_a g_a^2 S_a(p) \) is \( \frac{1}{2}(M V_p V_p^T + V_p V_p^T M) \).
Assembling these pieces gives Eq. (19) for $16 g^4 S(R) C(R) M$. The structure of the final expression is verifiable by looking at the relevant diagrams:

(C.9)

Appendix D. Gauge-coupling and gaugino-mass matching

In this appendix we comment on yet another key ingredient of any practical application of the methods advocated in this work, namely, on the matching between the symmetric and asymmetric phases of spontaneously broken models with multiple $U(1)$ gauge groups, where the latter is described by an effective theory with a reduced dimensionality of the Abelian sector.

D.1. Gauge couplings

Whenever a charged field develops a VEV, the original $U(1)_a \otimes U(1)_b$ gauge symmetry gets broken down to a single $U(1)_c$ spanned over the unbroken combination of the original generators

$$Q'_i = p_a Q^a_i + p_b Q^b_i \quad \text{(no summation).} \quad \text{(D.1)}$$

A gauge coupling associated to this residual symmetry is then in general a function of the original gauge couplings $g_a$ and $g_b$, the $p$-coefficients above, and, at higher loop order, also other quantities such as group Casimirs etc. Let us recapitulate in brief how the matching between the two regimes determines $g_c$, focusing mainly on the “tree-level matching” as needed when working with the one-loop RGEs. We comment on the changes for two-loop evolution at the end of this section.

It is convenient to address the issue in two steps. First, one can consider the matching between two QED$^2$ scenarios which are connected just by a pair of $O$-rotations as in Eqs. (7)–(9). In particular, any specific choice of $O_1$ represents a transition from the original set of the $U(1)$ charges $Q_i = (Q^a_i, Q^b_i)^T$ to a new set $Q'_i = (Q^c_i, Q^d_i)^T$ where $Q'_i = O_1 Q_i$. This, in turn, transforms the relevant (matrix of) gauge couplings as $G \to G' = O_1 G$ so that the interaction part of the covariant derivative remains intact, $Q^T_i G A = Q'^T_i G'A$. Regardless of whether simultaneously an $O_2$ rotation has been performed on the gauge fields, i.e., $A \to A' = O_2 A$ (inducing $G \to G' = O_1 G O_2^T$), one always has $G'^T G' = O_1 G G'^T O_1^T$, or, equivalently,

$$(G'^T G')^{-1} = O_1 (G G'^T)^{-1} O_1^T \quad \text{(D.2)}$$

which, as we shall see below, is more useful in practice. Hence, for any specific choice of $O_1$, this equation yields the link between the specific non-linear combinations of $g_{aa}$, $g_{ab}$, $g_{ba}$ and $g_{bb}$ entries of $G$ and the $g_{cc}$, $g_{cd}$, $g_{dc}$ and $g_{dd}$ entries of $G'$ and, as such, provides the desired matching condition. Note that, given the $p_a$ and $p_b$ coefficients in Eq. (D.1), the first row of $O_1$ is fixed and its second row is determined from orthogonality up to a global sign.

However, $O_2$ plays an important role if one, e.g., needs to get a coupling associated to a specific gauge field, for instance the one associated to the “residual” $U(1)_c$ that survives the
$U(1)_a \otimes U(1)_b$ breakdown. This gauge boson corresponds to the massless eigenstate of the gauge-boson mass matrix

$$M^2_A = G^T \langle H \rangle^\dagger Q^*_H \bar{Q}^T_H \langle H \rangle G$$

(D.3)

where $\langle H \rangle$ is the $U(1)_a \otimes U(1)_b$-breaking VEV. Since the conserved charge $Q^c_H$ in Eq. (D.1) annihilates this VEV, it is convenient to go into the primed basis where $Q^T_H \langle H \rangle = (0, V)$. Thus,

$$M^2_A = G^T O^T_1 \langle H \rangle^\dagger Q^*_H \bar{Q}^T_H \langle H \rangle O_1 G = G^T O^T_1 \begin{pmatrix} 0 & 0 \\ 0 & V^2 \end{pmatrix} O_1 G.$$  

(D.4)

A convenient gauge-field transformation $A \to A' = O_2 A$ would bring this mass matrix into a diagonal form

$$M^2_{A'} = G'^T \begin{pmatrix} 0 & 0 \\ 0 & V^2 \end{pmatrix} G'$$

(D.5)

(using $G' = O_1 G O^T_2$) if and only if $G'$ is upper-triangular, i.e.,

$$G' = \begin{pmatrix} g'_{cc} & g'_{cd} \\ 0 & g'_{dd} \end{pmatrix}.$$  

(D.6)

The zero in the 12 entry has a clear physical implication: for the couplings of the surviving (massless) $A'_c$ gauge boson only $g_{cc} \equiv g'_{cc}$ is relevant in the effective theory whereas one does need to care the $g'_{cd}$ and $g'_{dd}$ couplings of the heavy $A'_d$ which is integrated out except for the calculation of higher-dimensional effective operators.

From here it is also easy to understand why Eq. (D.2) is better suited for practical purposes than the one without inverse: Indeed, the 1–1 entry of $(G'G'^T)^{-1}$ reveals $g'_{cc}$ in a very simple way, namely,

$$(G'G'^T)^{-1} = \begin{pmatrix} g'^{-2}_{cc} & -g'_{cd}g'^{-1}_{cd}g'^{-2}_{cc} \\ (g'^2_{cc} + g'^2_{cd})g'^{-2}_{dd}g'^{-2}_{cc} \end{pmatrix}$$

(D.7)

(the dotted term follows from symmetry) so that it is sufficient to look at the 1–1 entry of the RHS of Eq. (D.2), whilst without the inverse

$$G'G'^T = \begin{pmatrix} g'^2_{cc} + g'^2_{cd} & g'_{cd}g'_{dd} \\ g'_{cd}g'_{dd} & g'^2_{dd} \end{pmatrix}$$

(D.8)

and in order to extract $g'_{cc}$ one has to solve a non-linear system involving all three independent entries of the RHS of Eq. (D.2), namely, $O_1 G G^T O_1^T$.

At higher-loop orders, the situation becomes slightly more involved, especially if the $U(1) \otimes U(1)$ gauge structure is tensored with a semi-simple gauge factor $G_X$. For example gauge-boson canonical normalization effects have to be considered at the two-loop level [39,40], which yields for example extra group-Casimir factors associated to $G_X$ entering formulae like Eq. (D.2), see, e.g., [23]. The specific shape of these terms is, however, renormalization scheme dependent.

### D.2. SUSY and the gaugino masses

Concerning the gaugino masses, the situation is only slightly more involved here. Sticking to the simplest $U(1)_a \otimes U(1)_b \to U(1)_c$ case as before, the gaugino($\lambda$)–higgsino($\bar{H}$) mass matrix
reads schematically (in the $O_2$-rotated basis bringing the gauge boson mass matrix to a block-diagonal form, cf., Eq. (D.5), and with $O_1$ rotation imposed on charges)

\[
M_{\lambda_{i}', H} = \begin{pmatrix}
M' & \sqrt{2}G'^T Q'_H (H)^* \\
\cdot & W_{HH} \\
\cdot & W_{H'H'}
\end{pmatrix}.
\]  

(D.9)

Here $M' = O_2M O_2^T$ is the gaugino soft mass matrix, $W$ is the superpotential; the subscripts of $W$ denote the derivatives of $W$ with respect to the superfield $H$ and its charge conjugate $H^c$. As before, due to the specific choice of the primed basis one has $Q^T_H (H) = (0, V)$ and $Q^T_H (H^c) = (0, V^*)$. Furthermore, the triangular shape of $G'$, see Eq. (D.6), ensures that the $\lambda_c$ gaugino corresponding to the first row/column receives no mass contribution due to the spontaneous symmetry breaking.

Thus, one concludes that the effective soft mass of the gaugino associated to the surviving gauge group $U(1)_c$ is given by the $c–c$ entry of the $M'$ mass matrix

\[
M_c = M'_{cc}.
\]  

(D.10)

Note that, in principle, there is no need to calculate the $O_2$ matrix because one can trade it for the gauge couplings and the (known) $O_1$:

\[
O_2 M O_2^T = G'^T O_1 G^{-1} M G^{-1} O_1^T G'
\]  

(D.11)

and thus

\[
G'^{-1} M' G'^{-1} = O_1 G^{-1} M G^{-1} O_1^T.
\]  

(D.12)

The main advantage of this formula is that, again, the LHS reveals $(M')_{cc}$ in a particularly simple manner, namely

\[
G'^{-1} M' G'^{-1} = \begin{pmatrix}
M'_{cc}/g'^2_{cc} & R_{12} \\
\cdot & R_{22}
\end{pmatrix},
\]  

(D.13)

with $R_{12} \equiv -g'_{cd} M'_{cc} + g'_{cc} M'_{cd} g'_{dd}/g'^2_{cc}$ and $R_{22} \equiv (g'^2_{cc} M'_{dd} - 2g'_{cc} g'_{cd} M'_{cd} + g'^2_{cd} M'_{cc})/g'^2_{dd} g'^2_{cc}$, and, thus, in combination with (D.7), admits for a simple extraction of the surviving gaugino effective soft mass. Moreover, due to the one-loop RGE invariance of the $G^{-1} M G^{-1}$ combination, the RHS of Eq. (D.12) is directly connected to the initial condition.

References

5.5 The minimal theory of perturbative B violation

C. Arbeláez Rodríguez, H. Kolešová, M. Malinský, Physical Review D89 (2014) 055003, DOI: 10.1103/PhysRevD.89.055003

The first of these articles represents a ground-breaking research of the candidate and his students in the field of minimal models of perturbative $B$ violation; the second study is a sequel including mainly technical details not given in the first one.

The Witten’s mechanism for the radiative generation of neutrino masses in the $SO(10)$ GUTs, as simple and beautiful as it is (cf. Sect. [3.6.1]), has been for a long time taken as a mere curiosity without much of a practical use in the realm of potentially realistic unified models. This was mainly due to the dichotomy between the gauge unification constraints in the non-supersymmetric settings and the need to keep the $B-L$ symmetry-breaking VEVs of the relevant scalars in the vicinity of the GUT scale in order for the effective type-I seesaw scale to be generated in the desired $10^{12−13}$ GeV ballpark.

The central point of the study below was the observation that there exist more economical frameworks in which the original Witten’s idea can be naturally realised without the need to conform the draconic grand-unification constraints of the $SO(10)$ settings. It was argued that the flipped $SU(5)$ gauge framework equipped with the Witten’s loop is not only way simpler than any other renormalizable flipped $SU(5)$ scenario considered before but, thanks to the high degree of flavour correlations within, it may be even viewed as the most minimal model of perturbative baryon number violation ever conceived as it has something to say about virtually all the gauge-induced two-body decay channels of nucleons. Besides that, it has a potential to provide fundamental insights into the absolute light neutrino mass scale and the early-Universe baryon number asymmetry generation via the leptogenesis mechanism.

The candidate’s contribution to this research was central as he was not only the inventor of the ideas sketched above (and in Sect. [3.6.2]) but also a driving force of the subsequent research conducted on this ground. Besides most of the manuscript writing, he contributed by the explicit calculation and discussion of the two-loop diagrams studied in the second paper and by deriving the $p$-decay partial widths in the representative case presented therein. At the moment of writing he is supervising the preparation of a third study in which a complete numerical synthesis of the existing proton lifetime, absolute neutrino mass scale, gauge running and leptogenesis constraints should be accomplished, with possible predictions for the leptonic CP and the associated neutrinoless double-beta-decay observables.
Witten’s mechanism in the flipped SU(5) unification

Carolina Arbeláez Rodríguez*
Centro de Física Teórica de Partículas, Instituto Superior Técnico, Universidade Técnica de Lisboa,
Avenida Rovisco Pais 1, 1049-001, Lisboa Portugal and AHEP Group,
Instituto de Física Corpuscular–C.S.I.C./Universitat de València Edificio de Institutos de Paterna,
Apartado 22085, E–46071 València, Spain

Helena Kolešová‡
Institute of Particle and Nuclear Physics, Faculty of Mathematics and Physics,
Charles University in Prague, V Holešovičká 2, 180 00 Praha 8, Czech Republic and
Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague,
Břehová 7, 115 19 Praha 1, Czech Republic

Michal Malinský‡
Institute of Particle and Nuclear Physics, Faculty of Mathematics and Physics,
Charles University in Prague, V Holešovičká 2, 180 00 Praha 8, Czech Republic
(Received 3 October 2013; published 3 March 2014)

We argue that Witten’s loop mechanism for the right-handed Majorana neutrino mass generation identified originally in the SO(10) grand unification context can be successfully adopted to the class of the simplest flipped SU(5) models. In such a framework, the main drawback of the SO(10) prototype—in particular, the generic tension among the gauge unification constraints and the absolute neutrino mass scale—is alleviated, and a simple yet potentially realistic and testable scenario emerges.

DOI: 10.1103/PhysRevD.89.055003 PACS numbers: 12.10.-g, 12.10.Kt, 14.80.-j

I. INTRODUCTION

The apparent absence of supersymmetry in the sub-TeV domain indicated by the current LHC data reopens the question whether the unprecedented smallness of the absolute neutrino mass scale may be ascribed to a loop suppression with the underlying dynamics in the TeV ballpark rather than the traditional seesaw [1–6] picture featuring a very high scale, typically far beyond our reach. Recently, there has been a lot of activity in this direction with, e.g., dedicated studies of the Zee [7], Zee-Babu [8–10], and other models (cf. [11,12] and references therein) focusing on their distinctive low-energy phenomenology and, in particular, their potential to be probed at the LHC and other facilities, see, e.g., [13–17].

With the upcoming generation of megaton-scale experiments [18–20] dedicated, besides precision neutrino physics, to the search of perturbative baryon number violating (BNV) processes such as proton decay, the same question can be readdressed from the high-energy perspective. In principle, there can be high-scale loop diagrams behind the right-handed (RH) neutrino masses underpinning the seesaw mechanism rather than a direct low-scale $LL$ contraction, with possible imprints in the BNV physics.

Among such options, a prominent role is played by Witten’s scheme [21] in the framework of the $SO(10)$ grand unification (GUT) where a pair of lepton-number violating vacuum expectation values (VEVs) is tied to the leptonic sector at two loops. Its main beauty consists in the observation that the RH neutrino masses are generated at the renormalizable level even in the simplest realization of $SO(10)$ with just the minimal scalar contents sufficient for the desired spontaneous symmetry breaking (i.e., $10 \oplus 16 \oplus 45$, cf. [22] and references therein); hence, there is in principle no need to invoke large scalar representations for that sake.

In practice, however, Witten’s mechanism has never found a clearly natural realization as a basis for a potentially realistic model building. Among the possible reasons there is, namely, the dichotomy between the gauge unification constraints and the absolute size of Witten’s loop governed by the position of the $B - L$ breaking scale $M_{B-L}$ which is required to be around the GUT-scale $M_G$, due to the $(\alpha/\pi)^2$ suppression factor, in order to yield the “correct” seesaw scale $M_R \sim (\alpha/\pi)^2 M_{B-L}/M_G$ in the $10^{13}$ GeV ballpark. On one hand, this is exactly the situation encountered in supersymmetric (SUSY) GUTs where the one-step breaking picture characterized by a close proximity of $M_{B-L}$ and $M_G$ is essentially inevitable; at the same time, however, the low-scale supersymmetry makes the $F$-type loops at the GUT scale entirely academic due to the large cancellation involved. On the other hand, non-SUSY GUTs generally require $M_{B-L} \ll M_G$ in order to account for the gauge unification constraints for which Witten’s mechanism yields contribution much below the desired $M_R \sim 10^{13}$ GeV.
In this respect, the beginning of the 1980s, when the low-energy SUSY was not yet mainstream and the lack of detailed information about the standard model (SM) gauge coupling evolution as well as the absolute light neutrino mass scale obscured the issue with the too low Witten’s $M_R$ in non-SUSY scenarios, was the only time when this business really flourished.\footnote{This can be seen at the citation counts of the original study [21] as about 70% of its today’s total dates back to before 1985.} For a more recent attempt to implement such ideas in a simple, yet potentially realistic scenario the reader is deferred to, e.g., the works [23,24] where the split supersymmetry scheme supports both $M_{b-L} \sim M_G$ and very heavy scalar superpartners for which, in turn, the GUT-scale $F$-type Witten’s loop is not entirely canceled.

In this work we approach this conundrum from a different perspective; in particular, we stick to the core of Witten’s loop while relaxing, at the same time, the strict gauge unification constraints. For that sake, we depart from the canonical realization of Witten’s mechanism in a full-fledged $SO(10)$ GUT to its “bare-bone” version which, as we point out, can be sensibly implemented within its simpler cousin, namely, the flipped $SU(5)$ [25–27].

Indeed, the strict full gauge unification constraints inherent to the $SO(10)$ GUTs are relaxed in such a scenario [owing to the nonsimple structure of its $SU(5) \otimes U(1)$ gauge group] which, in turn, makes it possible to have the rank-reducing vacuum expectation value (VEV) governing Witten’s loop in the $10^{16}$ GeV ballpark even if the theory is nonsupersymmetric.

The reason we are focusing just on the flipped $SU(5)$ framework is twofold: First, the baryon-number violating observables such as the $d = 6$ proton decay [28] may still be used to constrain specific scenarios even if the underlying dynamics is as high as at $10^{16}$ GeV, as we will comment upon in the following. This virtue is obviously lost if one picks any of the “smaller” subgroups of $SO(10)$ such as Pati-Salam\footnote{Let us recall that proton decay in Pati-Salam requires a conspiracy in the Higgs sector as it does not run solely through the gauge interactions.} [29], let alone the number of left-right symmetric (LR) settings based on the $SU(3) \otimes SU(2)_L \otimes SU(2)_R \otimes U(1)_{b-L}$ gauge symmetry. Second, the flipped variant of $SU(5) \otimes U(1) \subset SO(10)$ is the only one for which a radiative generation of the RH neutrino masses makes sense because in the standard $SU(5)$ the RH neutrinos are gauge singlets and as such they receive an explicit singlet mass term.

Besides this, the flipped scenario has got other virtues: the proton decay estimates\footnote{For a nice discussion on how to use BNV observables to distinguish between the standard and the flipped $SU(5)$ see, e.g., [35].} may be under better control than in the standard $SU(5)$ because the leading theoretical uncertainties in the GUT-scale calculation (namely, the few-percent ambiguities in the GUT-scale matching of the running gauge couplings due to the Planck-induced effects [30–33]) are absent. Furthermore, the flipped scenario offers better perspectives for a solution of the doublet-triplet splitting problem (if desired; see, e.g., [34]) and, unlike in the “standard” $SU(5)$, there is no monopole problem in the flipped case either.

On top of that, the proposed scenario is in a certain sense even simpler than the standard approach to the minimal\footnote{Minimality here refers to models without extra matter fields; for an alternative approach including, for instance, extra singlet fermions see, e.g., [37].} renormalizable flipped $SU(5)$ where the seesaw scale is associated to the VEV of an extra scalar representation transforming as a 50-dimensional four-index tensor under $SU(5)$ coupled to the fermionic $10 \otimes 10$ bilinear (see, e.g., [36]); indeed, such a large multiplet is not necessary in the flipped $SU(5)$ à la Witten; as we shall argue, the two models can even be distinguished from each other if rich-enough BNV physics is revealed at future facilities. In particular, we observe several features in the typical ranges predicted for the $\Gamma(p \rightarrow \pi^0 e^+)$ and $\Gamma(p \rightarrow \pi^0 \mu^+)$ partial widths [as well as for those related by the isospin symmetry such as $\Gamma(p \rightarrow \eta e^+$) etc.] that are trivially absent in the model with $50_H$ in the scalar sector. Remarkably enough, this makes it even possible to obtain rather detailed information about all kinematically allowed $d = 6$ nucleon decay channels in large portions of the parameter space where the theory is stable and perturbative.

The work is organized as follows: In Sec. II, after a short recapitulation of the salient features of the standard and flipped $SU(5)$ models and the generic predictions of the partial proton decay widths therein, we focus on the Witten’s loop as a means to constrain the shape of the (single) unitary matrix governing the proton decay channels into neutral mesons in the flipped case. In Sec. III we perform a detailed analysis of the simplest scenario in which a set of interesting partial proton decay widths are revealed with their strengths governed by the absolute size of Witten’s diagram. In Sec. IV, we adopt this kind of analysis to the minimal potentially realistic scenario. Then we conclude.

II. $SU(5) \otimes U(1) \subset SO(10)$ à LA WITTEN

Let us begin with the basics of the flipped $SU(5)$ scheme and a short account of the $d = 6$ proton decay in the $SU(5)$-based unifications focusing, namely, on the minimal versions of the standard and flipped scenarios and the potential to discriminate experimentally among them if proton decay would be seen in the future.

A. The flipped $SU(5)$ basics

The quantum numbers of the matter multiplets in the $SU(5) \otimes U(1)_X$ extensions of the canonical $SU(5)$
framework are dictated (up to an overall normalization factor) by the requirement of the gauge anomaly cancellation:

$$5_M \equiv (5, -3), \quad 10_M \equiv (10, +1), \quad 1_M \equiv (1, +5).$$  (1)

Besides the “standard” $SU(5)$ assignment there is a second “flipped” embedding of the standard model (SM) hypercharge into the corresponding algebra, namely,

$$Y = \frac{1}{5}(X - T_{24}),$$  (2)

where the $SU(5)$ generator $T_{24}$ is in this case understood to conform the SM normalization (i.e., $Y = T_{24}$ and $Q = T_L^3 + Y$ in the standard case). This swaps $u^c \leftrightarrow d^c$ and $\nu^c \leftrightarrow e^c$ with respect to the standard $SU(5)$ field identification and, hence, the RH neutrinos fall into $10_M$ rather than $5_M$. This also means that a VEV of a scalar version of $\nu$ can spontaneously break the $SU(5) \otimes U(1)_X$ gauge symmetry down to the SM.$^6$

Besides that, the scheme benefits from several nice features not entertained by the “standard” $SU(5)$ scenario, namely: (i) The Yukawa Lagrangian

$$\mathcal{L} \supset Y_{10}10_M5_H + Y_{10}5_M5_H + Y_{10}5_M15_H + \text{H.c.},$$  (3)

including the 5-dimensional scalar $5_H = (5, -2)$ hosting the SM Higgs doublet, yields $M_d = M_d^I$, $M_e$ arbitrary and, in particular, $M_d^D = M_d^I$, none of which is in a flagrant conflict with the observed quark and lepton flavor structure as it is the case for $M_d = M_d^I$ in the “standard” $SU(5)$.

(ii) The gauge unification is in a better shape than in the “standard” $SU(5)$ case because only the two non-Abelian SM couplings are required to unify (which, indeed, do at around $10^{16}$ GeV, cf. Appendix C)—note that the SM hypercharge is a “mixture” of the $T_{24}$ and $X$ charges (2) and, thus, the SM coupling $g'$ obeys a nontrivial matching condition including an unknown coupling $g_X$ associated to the extra $U(1)_X$ gauge sector. Hence, there is no need to invoke the TeV-scale supersymmetry for the sake of the gauge unification here as in the “standard” $SU(5)$ case.

(iii) Remarkably, the issue with the out-of-control Planck-scale induced shifts of the effective gauge couplings (and thus induced large uncertainties in the $M_G$ determination [30–33]) is absent at the leading order because there is no way to couple the $10_H$ as the carrier of the large-scale VEV to the pair of the gauge field tensors $F_{\mu\nu}$. Thus, the prospects of getting a reasonably good grip on the proton lifetime in the flipped $SU(5)$ are much better than in the ordinary $SU(5)$ model.

The main drawback of such a scenario is the fact that the simplest “conservative” mechanism for generating a Majorana mass term for the RH neutrinos at the tree level requires an extra 50-dimensional scalar field $50_H \equiv (50, -2)$ whose large VEV in the $10_M 10_M 50_H$ contraction picks just the desired components.$^7$ Obviously, one pays a big price here (i.e., 100 real degrees of freedom which further reduce the effective Planck scale [30,38–40]) and there is not much insight into the neutrino mass generation that this may provide (as, e.g., there is no grip on the flavor structure). Hence, this approach is not optimal as it totally ignores the bounty of the recent high-precision neutrino data.

B. Proton decay in the standard and flipped $SU(5)$

Since the new dynamics associated to the rich extra gauge and scalar degrees of freedom of the flipped $SU(5)$ scenario takes place at a very high scale the most promising observables it can find its imprints in are those related to the perturbative baryon number violation, namely, proton decay.

To this end, the flipped version of the $SU(5)$ unification is in a better shape than its “standard” cousin as it provides a relatively good grip [26,28] on the partial proton decay widths to neutral mesons and charged leptons, whereas there is usually very little one can say on general grounds about these in the standard $SU(5)$ where those are the charged meson plus antineutrino channels which are typically under better theoretical control. Needless to say, this is very welcome as the observability of the charged leptons in the large-volume liquid scintillator [18]/water-Cherenkov [19]/liquid Argon [20] experiments boosts the expected signal to background ratio and, hence, provides a way better sensitivity (by as much as an order-of-magnitude) in these channels than in those with the unobserved final-state antineutrino.

Let us just note that this has to do, namely, with the hypercharge of the heavy $d = 6$ proton-decay-generating gauge colored triplets which under the SM transform as $(3, 2, -\frac{5}{6})$ in the standard $SU(5)$ case and as $(3, 2, +\frac{1}{2})$ in the flipped $SU(5)$, respectively.$^8$ As for the former, the relevant $d = 6$ effective BNV operators are [28] of the $O^I \propto u^c Q \bar{c}^c \bar{Q}$ and $O^{II} \propto u^c \bar{Q} d^c L$ type while for the latter these are $O^{III} \propto \bar{d}^c Q u^c L$ and $O^{IV} \propto \bar{d}^c Q \bar{d}^c \bar{c}^c$ where “pairing” is always between the first two and the last two fields therein. Hence, the neutral meson + charged lepton decays in the standard $SU(5)$ receive contributions from both $O^I$ and $O^{II}$ while it is only $O^{III}$ that drives it in the

$^5$Recall that in the standard $SU(5)$ $Q$, $u^c$, and $e^c$ are in $10_M$, $d^c$ and $L$ in $5_M$, and $\nu^c$ in $1_M$.

$^6$This is the observation in the core of the “missing partner” doublet-triplet splitting mechanism (mainly relevant to supersymmetry) that brought a lot of interest to the flipped $SU(5)$ scenario in the 1980s [34].

$^7$As does $(\overline{26}_H)$ coupled to $16_M 16_M$ in $SO(10)$.

$^8$This is also reflected by the classical notation where the $SU(2)$ components of the former are called $X$ and $Y$ while for the latter these are usually denoted by $X'$ and $Y'$. 

055003-3
flipped scenario.\textsuperscript{9} On the other hand, the situation is rather symmetric in the charged meson + neutrino channels which in both cases receive sizeable contributions from only one type of a contraction [$\mathcal{O}^{HI}$ in $SU(5)$ and $\mathcal{O}^{III}$ in its flipped version]. Let us also note that the predictivity for these channels is further boosted by the coherent summation over the (virtually unmeasurable) neutrino flavors; hence, the inclusive decay widths to specific charged mesons are typically driven by the elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix. For instance, in a wide class of simple $SU(5)$ GUTs (namely, those in which the up-type quark mass matrix is symmetric) the $p$-decay widths to $\pi^+$ and $K^+$ can be written as

\begin{align}
\Gamma(p \to \pi^+ \bar{\nu}) &= F_1 |(V_{\text{CKM}})_{11}|^2 M, \\
\Gamma(p \to K^+ \bar{\nu}) &= |F_2|(V_{\text{CKM}})_{11}|^2 + F_3 |(V_{\text{CKM}})_{12}|^2 M. \tag{4}
\end{align}

where $F_{1,2,3}$ are calculable numerical factors and $M$ is a universal dimensionful quantity governed by the parameters of the underlying “microscopic” theory such as the GUT scale, the gauge couplings, etc. This feature is yet more pronounced in the simple flipped scenarios (namely, those in which the down-type quark mass matrix is symmetric\textsuperscript{10}) ; there one even obtains a sharp prediction

\begin{equation}
\Gamma(p \to K^+ \bar{\nu}) = 0 \tag{5}
\end{equation}

which is a clear smoking gun of the flipped $SU(5)$ unification. For more details an interested reader is referred to the dedicated analysis \cite{35}.

Coming back to the neutral meson channels in the simplest flipped $SU(5)$ scenarios (i.e., assuming symmetry of the down quark mass matrix), the partial widths of our main interest may be written in the form

\begin{align}
\frac{\Gamma(p \to \pi^0 e^+_{\alpha})}{\Gamma(p \to \pi^+ \bar{\nu})} &= \frac{1}{2} |(V_{\text{CKM}})_{11}|^2 |(V_{\text{PMNS}} U_{\nu})_{\alpha}|^2, \tag{6} \\
\frac{\Gamma(p \to \eta e^+_{\alpha})}{\Gamma(p \to \pi^+ \bar{\nu})} &= \frac{C_2}{C_1} |(V_{\text{CKM}})_{11}|^2 |(V_{\text{PMNS}} U_{\nu})_{\alpha}|^2, \tag{7} \\
\frac{\Gamma(p \to K^0 e^+_{\alpha})}{\Gamma(p \to \pi^+ \bar{\nu})} &= \frac{C_3}{C_1} |(V_{\text{CKM}})_{12}|^2 |(V_{\text{PMNS}} U_{\nu})_{\alpha}|^2, \tag{8}
\end{align}

where $V_{\text{PMNS}}$ stands for the Pontecorvo-Maki-Nakagawa-Sakata leptonic mixing matrix and $U_{\nu}$ is the unitary matrix diagonalizing the light neutrino masses.\textsuperscript{11} Note that $V_{\text{PMNS}} U_{\nu} = U_{\nu}^T$ is the LHS diagonalization matrix in the charged lepton sector [see Eq. (A9)]; we write it in such a “baroque” way because $V_{\text{PMNS}}$ is measurable and, as will become clear, $U_{\nu}$ is constrained in the model under consideration. The absolute scale in Eqs. (6)–(8) is set by

\begin{equation}
\Gamma(p \to \pi^+ \bar{\nu}) = C_1 \left(\frac{g_G}{M_G}\right)^4, \tag{9}
\end{equation}

where $g_G$ is the $SU(5)$ gauge coupling and the numerical factors

\begin{align}
C_1 &= \frac{m_p}{8\pi f_e^2} A_2 \alpha^2 (1 + D + F)^2, \tag{10} \\
C_2 &= \frac{m_p^2 - m_\pi^2}{48\pi f_e^2} A_2^2 \alpha^2 (1 + D - 3F)^2, \tag{11} \\
C_3 &= \frac{m_p^2 - m_K^2}{8\pi f_e^2} A_2^2 \alpha^2 \left[1 + \frac{m_p}{m_B} (D - F)\right]^2 \tag{12}
\end{align}

are obtained by chiral Lagrangian techniques, see \cite{28} and references therein, \cite{35} and Appendix A. From Eqs. (6)–(8), the theory’s predictive power for the proton decay to neutral mesons (especially for the “golden” $p \to \pi^0 e^+$ channel), in particular, its tight correlation to neutrino physics, is obvious as the only unknown entry in Eqs. (5)–(8) is the unitary matrix $U_{\nu}$.

In what follows we shall exploit the extra constraints on the lepton sector flavor structure emerging in the flipped $SU(5)$ model with Witten’s loop in order to obtain constraints on the admissible shapes of the $U_{\nu}$ matrix and, hence, get a grip on the complete set of proton decay observables. Let us note that this is impossible in the models in which the RH neutrino masses are generated in the “standard” way (e.g., by means of an extra 50\textsubscript{H}) where, due to the entirely new type of a contraction entering the lepton sector Lagrangian, $U_{\nu}$ typically remains unconstrained.

\section*{C. Witten’s mechanism in the flipped SU(5)}

The main benefit of dealing with a unification which is not “grand” (i.e., not based on a simple gauge group) is the absence of the strict limits on the large-scale symmetry breaking VEVs from an overall gauge coupling convergence at around $10^{16}$ GeV. Indeed, unlike in the $SO(10)$ GUTs which typically require the rank-breaking VEV (e.g., that of 16- or 126-dimensional scalars) to be several orders of magnitude below $M_G$ \cite{43-46} and, hence, too low for

\textsuperscript{9}In fact, $\mathcal{O}^{IV}$ is almost always irrelevant as it yields a left-handed antineutrino in the final state with typically (in the classical seesaw picture) a very tiny projection on the light neutrino mass eigenstates.

\textsuperscript{10}This, in fact, is the prominent case when the flipped-$SU(5)$ proton decay is robust, i.e., cannot be rotated away, cf. \cite{28,35,41}; for a more recent account of the same in a flipped-$SU(5)$ scenario featuring extra matter fields see, e.g., \cite{42}.

\textsuperscript{11}Let us anticipate that Eqs. (6)–(8) are written in the basis in which the up-type quark mass matrix is diagonal and real; needless to say, the observables of our interest are all insensitive to such a choice. For more details see Appendix B.
Witten’s loop to account for the “natural” $10^{12-14}$ GeV RH neutrino masses domain, no such issue is encountered in the $SU(5) \otimes U(1)$ scenario due to its less restrictive partial unification pattern. In particular, only the non-Abelian SM gauge couplings are supposed to converge toward $M_G$ which, in turn, should be compatible with the current proton lifetime limits; no other scale is needed. Furthermore, the $SU(5) \otimes U(1)$-breaking VEV $V_G = \langle 10_H \rangle$ is perfectly fit from the point of view of the gauge structure of Witten’s type of a diagram in this scenario.

1. Witten’s loop structure

As in the original $SO(10)$ case the gauge and loop structure of the relevant graphs (cf., Fig. 1) conforms\(^\text{12}\) to several basic requirements: (i) there should be two $V_G$’s sticking out of the diagram so that the correct “amount” of the $U(1)_X$ breaking is provided for the desired RH neutrino Majorana mass term; (ii) the interactions experienced by the fermionic current must mimic the $10_\mu \cdot 10_M \cdot 50_\mu$ coupling of the renormalizable models in which the RH neutrino mass is generated at the tree level; (iii) only the minimal set of scalars required for the spontaneous symmetry breaking should be employed. Given that, the structure depicted in Fig. 1 turns out to be the simplest option;\(^\text{13}\) indeed, $5 \otimes 24 \otimes 24$ (where 24 stands for the gauge fields) is the minimum way to devise the desired $50$. Note also that the $U(1)_X$ charge of the gauge $24_G$’s is zero and, thus, the two units of $X$ are delivered to the leptons via their Yukawa interaction with $5_H$. We have checked by explicit calculation that, indeed, the gauge structure of the graph yields a nonzero contribution for, and only for, the RH neutrino.

2. The right-handed neutrino mass matrix

Following the standard Feynman procedure the RH neutrino mass matrix can be written in the form\(^\text{14}\)

$$M^M_\nu = \left(\frac{1}{16\pi^2}\right)^2 g_G^2 Y_{10\mu} \langle 10_H \rangle^2 \frac{M_G}{M^2} \times O(1),$$

(13)

where $g_G$ is the (unified) gauge coupling corresponding to that of the $SU(5)$ part of the gauge group, $\mu$ is the (dimensionful) trilinear scalar coupling among $10_\mu$’s and $5_H$, cf. Eq. (15), $Y_{10}$ is the Yukawa coupling of $5_H$ to the

\(^{12}\)Note that the quantum numbers of the submultiplets under the $SU(5)$ subgroup of $SO(10)$ indicated in Witten’s original work [21] are irrelevant here as the RH neutrinos in the flipped scenario reside in 10 of $SU(5)$ rather than in a singlet.

\(^{13}\)Note that minimality in this context depends on the specific construction of the perturbation expansion as, e.g., one diagram in the broken phase approach with massive propagators corresponds to a tower of graphs in the unbroken-phase theory when the VEVs are included in the interaction Hamiltonian.

\(^{14}\)Note that due to the symmetry of $Y_{10}$ the algebraic structure of the “permuted” graphs is the same as the one in Fig. 1 and, hence, all contributions are covered by expression (13).

\(^{15}\)Obviously, there are several equivalent approaches to the evaluation of the momentum integrals involved in the $O(1)$ factor: one can either work in the mass basis in which the propagators are diagonal and the couplings contain the rotations from the defining to the mass basis or vice versa; in principle, one may even work in a massless theory with VEVs in the interaction part of the Lagrangian.
perturbative consistency constraints following from the requirements of the SM vacuum stability and general perturbativity which, together with the above-mentioned bounds on $K$, impose rather strict limits on the absolute scale of the RH neutrino masses.

3. Constraints from the SM vacuum stability

Here we attempt to identify the parameter-space domains that may support a stable SM vacuum, i.e., those for which there are no tachyons (i.e., no negative-sign eigenvalues of the relevant scalar mass-squared matrix) in the spectrum.

Tree-level scalar potential.—Let us parametrize the tree-level scalar potential as

$$V_0 = \frac{1}{2} m_{10}^2 \text{Tr}(10_H^t 10_H) + m_{5_H}^2 5_H^t 5_H$$

$$+ \frac{1}{8} (\mu_{ijkl} 10_H^i 10_H^j 10_H^k 10_H^l + \text{H.c.}) +$$

$$+ \frac{1}{4} \lambda_1 |\text{Tr}(10_H 10_H)|^2 + \frac{1}{4} \lambda_2 \text{Tr}(10_H^t 10_H 10_H^t 10_H)$$

$$+ \lambda_3 (5_H^t 5_H)^2 + \frac{1}{2} \lambda_4 \text{Tr}(10_H^t 10_H)(5_H^t 5_H)$$

$$+ \lambda_5 5_H^t 10_H 10_H^t 5_H,$$

where $10_H$ and $5_H$ are conveniently represented by a $5 \times 5$ complex antisymmetric matrix and a 5-component complex column vector, respectively, and the normalization factors in the interaction terms have been chosen such that they ensure simplicity of the resulting Feynman rules and, hence, of the results below. Note that we choose a basis in which the GUT-scale VEV $V_G$ and the electroweak VEV $v$ are accommodated in the following components:

$$(10^{15}) = -(5^{44}) = V_G, \quad (5^4) = v. \quad (16)$$

The SM vacuum.—The SM vacuum stationarity conditions read

$$V_G[m_{10}^2 + V_G^2 (2\lambda_1 + \lambda_2) + v^2 (\lambda_4 + \lambda_5)] = 0,$$

$$v [m_3^2 + 2v^2 \lambda_3 + V_G^2 (\lambda_4 + \lambda_5)] = 0. \quad (17)$$

There are in general four solutions to this system, namely,

$$V_G = v = 0 : SU(5) \otimes U(1),$$

$$V_G \neq 0, v = 0 : SU(3) \otimes SU(2) \otimes U(1),$$

$$V_G \neq 0, v \neq 0 : SU(3) \otimes U(1),$$

$$V_G = 0, v \neq 0 : SU(4) \otimes U(1),$$

with the preserved symmetry indicated on the right; the first three then correspond to consecutive steps in the physically relevant symmetry breaking chain.

The scalar masses.—As long as only the signs of the scalar mass-squares are at stake one can work in any basis. Using the “real field” one, i.e., $\Psi = \{10^i, 10^j, 5^i, 5^j\}$, the mass matrix $M^2 \equiv \langle \partial^2 V / \partial \Psi^* \partial \Psi \rangle$ evaluated in the SM vacuum has the following system of eigenvalues (neglecting all subleading terms):

$$m_{G_{1, \ldots, 16}}^2 = 0 \quad (18)$$

$$m_H^2 = \left[ 4\lambda_3 - \frac{2(\lambda_4 + \lambda_5)^2}{2\lambda_1 + \lambda_2} \right] v^2, \quad (19)$$

$$m_3^2 = 2(2\lambda_1 + \lambda_2)V_G^2, \quad (20)$$

$$m_{2, \lambda_2}^2 = -\frac{1}{2}(\lambda_2 + \lambda_5)V_G^2 - \frac{1}{2} V_G \sqrt{(\lambda_2 - \lambda_5)^2 V_G^2 + 4\mu^2},$$

$$m_{2, \lambda_2}^2 = -\frac{1}{2}(\lambda_2 + \lambda_5)V_G^2 + \frac{1}{2} V_G \sqrt{(\lambda_2 - \lambda_5)^2 V_G^2 + 4\mu^2}. \quad (21)$$

A few comments are worth making here:

(i) The 16 zeroes in Eq. (18) correspond to the Goldstone bosons associated to the spontaneous breakdown of the $SU(5) \otimes U(1)$ symmetry to the $SU(3)_c \otimes U(1)_Q$ of the low-energy QCD $\otimes$ QED.

(ii) $m_H$ is the mass of the SM Higgs boson. Let us note that the recent ATLAS [48] and CMS [49] measurements of $m_H$ indicate that the running effective quartic Higgs coupling at around $M_G$, i.e., the parenthesis in Eq. (19), should be close to vanishing, see, e.g., [50] and references therein.

(iii) $m_3$ is the mass of the heavy singlet in $10_H$.

(iv) The remaining eigenvalues correspond to the leftover mixture of the colored triplets with the SM quantum numbers $(3, 1,-1/3)$ from $5_H \otimes 10_H$ (6 real fields corresponding to each eigenvalue).

Absence of tachyons.—Clearly, there are no tachyons in the scalar spectrum as long as

$$2\lambda_1 + \lambda_2 > 0, \quad (22)$$

$$2\lambda_3 (2\lambda_1 + \lambda_2) > (\lambda_4 + \lambda_5)^2, \quad (23)$$

$$\lambda_2 + \lambda_5 < 0, \quad (24)$$

and, in particular,

$$|\lambda_2 + \lambda_5| V_G^2 > \sqrt{(\lambda_2 - \lambda_5)^2 V_G^2 + 4\mu^2}, \quad (25)$$

which may be further simplified to $\mu^2 < \lambda_2 \lambda_5 V_G^2$. Combining this with (24) one further concludes that both $\lambda_2$ and $\lambda_5$ must be negative. This also means that $\lambda_1$ must be

055003-6
positive and obey $2\lambda_1 > |\lambda_2|$ and, at the same time $\lambda_3$ must be positive. To conclude, the $\mu$ factor in formula (14) is subject to the constraint

$$|\mu| \leq \sqrt{\lambda_2\lambda_3} V_G$$

(26)

in all parts of the parameter space that can, at the tree level, support a (locally) stable SM vacuum.

4. Perturbativity constraints

Let us briefly discuss the extra constraints on the RHS of Eq. (14) implied by the requirement of perturbativity of the couplings therein. Since the graph in Fig. 1 emerges at the GUT scale it is appropriate to interpret these couplings as the running parameters evaluated at $M_G$. Note that the effective theory below this threshold is the pure SM and, thus, one may use the known qualitative features of the renormalization group evolution of the SM couplings to assess their behavior over the whole domain from $\mu$ to $V_G$. In general, one should assume that for all couplings perturbativity is not violated at $M_G$ and below $M_G$ the same holds for the “leftover” parameters of the effective theory. To that end, one should consider several terms in the perturbative expansion of all amplitudes in the relevant framework and make sure the (asymptotic) series thus obtained does not exhibit pathological growth of higher-order contributions (to a certain limit). This, in full generality, is clearly a horrendous task so we shall as usual adopt a very simplified approach. In particular, we shall make use of the fact that the running of all dimensionless couplings in the SM is rather mild so, in the first approximation, it is justified to consider their values at only one scale and assume the running effects will not parametrically change them. Hence, in what follows we shall assume that

$$|\lambda_i| \leq 4\pi \quad \forall i$$

(27)

for all the couplings in the scalar potential.

5. Resulting bounds on the $U_\nu$ matrix

With this in hand one can finally derive the desired constraints on the $U_\nu$ matrix governing the proton decay channels to neutral mesons (6)–(8). Indeed, using the seesaw formula, one can trade $M^\nu_\mu$ in Eq. (14) for the physical light neutrino mass matrix $m_{LL}$ and the Dirac part of the full $6 \times 6$ seesaw matrix

$$M^\nu_\mu = -M^D_\nu (m_{LL})^{-1}(M^D_\nu)^T$$

which, due to the tight link between $M^D_\nu$ and the up-type quark mass matrix in the simplest scenarios, $M^D_\nu = M^D_u$, yields $M^{\nu}_\mu = -M^D_\nu (m_{LL})^{-1}M_u$. Furthermore, the basis in the quark sector can always be chosen such that the up-quark mass matrix is real and diagonal (see Appendix B); at the same time, one can diagonalize $m_{LL} = U_\nu^T D_\nu U_\nu$ and obtain

$$M^{\nu}_\mu = -D_\nu U_\nu^T D^{-1}_\nu U_\nu^T D_\nu u$$

(28)

Combining this with formula (14) and implementing the vacuum stability constraint (26) one obtains

$$|D_\nu U_\nu^T D^{-1}_\nu U_\nu^T D_\nu u| \leq \frac{\alpha_G}{64\pi} \sqrt{\lambda_2\lambda_3} |Y_{10}| V_G K,$$

(29)

where we denoted $\alpha_G = g^2_G/4\pi$. Finally, assuming $\max_{i,j \in \{1,2,3\}} |(Y_{10})_{ij}| = 1$ and saturating the perturbativity constraints (27) we have

$$\max_{i,j \in \{1,2,3\}} |(D_\nu U_\nu^T D^{-1}_\nu U_\nu^T D_\nu u)_{ij}| \leq \frac{\alpha_G}{16\pi^2} V_G K,$$

(30)

which provides a very conservative global limit on the allowed form of $U_\nu$ and, hence, on the proton decay partial widths (6)–(8).

6. Unification constraints

Let us finish this preparatory section by discussing in brief the constraints from the requirement of the convergence of the running $SU(3)_c$ and $SU(2)_L$ gauge couplings at high energy which shall provide basic information about the scales involved, in particular, the approximate value of the $V_G$ parameter. Given (16), the $SU(2)_L$ doublet of the proton-decay-mediating colored triplet gauge fields ($X', Y'$) has mass $M_G = \frac{1}{2} g_G V_G$ while the mass of the heavy $U(1)_{T_3} \otimes U(1)_X$ gauge boson (i.e., the one orthogonal to the surviving massless SM $B$-field associated to hypercharge) reads $M_\nu = 2 \sqrt{\frac{1}{2} g^2_G + g^2_3} V_G$ in the units in which the $U(1)_X$ generator is normalized as in Eqs. (1) and (2).

Let us note again that in the flipped scenario of our interest the $M_G$ parameter corresponds to the scale at which the $(X', Y')$ are integrated into the theory in order to obey the $SU(3)_c$ and $SU(2)_L$ unification constraints. The specific location of this point and, thus, the absolute size of the proton decay width, however, depends also on the position of the other thresholds due to the extra scalars to be integrated in at around $M_G$, in particular, the $SU(5) \otimes U(1)_X / SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ Goldstone bosons (18), the heavy singlet (20), and the heavy colored triplets (21). Rather than going into further details here we defer a dedicated analysis of the situation in Appendix C and, in what follows, we shall stick to just a single reference scale of $M_G = 10^{16.5}$ GeV which corresponds to the lower limit obtained therein. This, in turn, yields $\Gamma^{-1}(p \rightarrow \pi^+\bar{\nu})$ of the order of $10^{38.5}$ years, cf. Fig. 8. Remarkably enough, there is also an upper limit of the order of $10^{42}$ years which, however, is attained only in a “fine-tuned” region where the inequality (26) is saturated.
III. A SAMPLE MODEL ANALYSIS

In order to exploit formula (30), it is convenient to begin with its thorough inspection which, as we shall see, will provide a simple analytic information on the potentially interesting regions of the parameter space which will, subsequently, feed into the analysis of the BNV observables. Later on, we shall compare the analytics with results of a dedicated numerical analysis.

A. Parameter space

1. CP conserving setup

For the sake of simplicity, we shall start with $U_v$ real orthogonal which shall be parametrized by three CKM-like angles $\omega_{12}$, $\omega_{23}$, and $\omega_{13}$:

$$U_v = U_{2-3}(\omega_{23}) U_{1-3}(\omega_{13}) U_{1-2}(\omega_{12})$$

where $U_{i-j}(\omega_{ij})$ stands for a rotation in the $i$-$j$ plane by an angle $\omega_{ij}$, e.g.

$$U_{2-3}(\omega_{23}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \omega_{23} & \sin \omega_{23} \\ 0 & -\sin \omega_{23} & \cos \omega_{23} \end{pmatrix}.$$  

Assuming normal neutrino hierarchy we parametrize the (diagonal) neutrino mass matrix $D_{\nu} = \text{diag}(m_1, m_2, m_3)$ by the (smallest) mass $m_1$ of the mostly electronlike eigenstate. The other two masses are then computed from the oscillation parameters $|\Delta m_{12}^2| = 2.43 \times 10^{-3}$ eV$^2$, $|\Delta m_{23}^2| = 7.54 \times 10^{-5}$ eV$^2$ [51,52] and, for the sake of this study, we mostly ignore the uncertainties in these observables. Let us note that for the inverted hierarchy the analysis is technically similar but physically less interesting, see below.

As long as the ratios of $m_1$'s are all below $m_1/m_{\nu}$, i.e., for $m_1 \gtrsim 10^{-4}$ eV (which we shall assume in the simple analysis below), the LHS of Eq. (30) is maximized for $(D_{\nu} U_{2-3}^\dagger D_{\nu} U_{1-3})_{33} = m_1^2 (U_{2-3}^\dagger U_{1-3})_{33}$. Hence, Eq. (30) gets reduced to (using $V_G = 2M_G/g_{G}$)

$$(U_{2-3}^\dagger U_{1-3})_{33} \leq \frac{g_{G}}{32\pi^2 m_1^2} \times 10^{16.5} \text{GeV} \approx K \times 3 \text{ eV}^{-1},$$

where we have taken $g_{G} = 0.5$. Besides that, one gets

$$(U_{2-3}^\dagger U_{1-3})_{33} = \frac{\sin^2 \omega_{13}}{m_1} + \cos^2 \omega_{13} \left( \frac{\sin^2 \omega_{23}}{m_2} + \frac{\cos^2 \omega_{23}}{m_3} \right),$$

which shows that the CKM-like parametrization of $U_v$ is very convenient because $\omega_{12}$ drops entirely from Eq. (33).

For further insight, let us consider the extreme cases first. For $\omega_{13} = \omega_{23} = 0$ (and for arbitrary $\omega_{12}$) one has $(U^\dagger_{2-3} D_{\nu} U_{1-3})_{33} = m_1^{-3}$, whereas for $\omega_{13} = \omega_{23} = \frac{\pi}{2}$ the same equals to $m_1^{-1}$. While $m_3^{-1}$ ranges from 11 eV$^{-1}$ to 20 eV$^{-1}$ for all $m_1$'s lower than the current Planck and large-scale-structure limit of about $18 \times 0.08$ eV [53], $m_1^{-1}$ may range in principle from 12 eV$^{-1}$ to infinity. This explains why the latter setting may not be allowed by (32) if $m_1$ and $K$ are small enough.

For the general case it is convenient to notice that the RHS of Eq. (33) is a convex combination of the inverse neutrino masses. Thus, for $m_1^{-1} \leq K \times 3 \text{ eV}^{-1}$ the inequality (32) is satisfied trivially. This can be clearly seen in Fig. 2 where the allowed parameter space is depicted: for $m_1 \gtrsim 3 K^{-1}$ eV, i.e., in the lower part of the plot, all $\omega_{23}$ and $\omega_{13}$ are allowed. On the other hand, if $m_1^{-1} \approx 11 \text{ eV}^{-1} > K \times 3 \text{ eV}^{-1}$, i.e., if $K \lesssim 4$, (32) is never fulfilled. There are two different regimes in the nontrivial region $m_1^{-1} \gtrsim K \times 3 \text{ eV}^{-1} > m_1^{-3}$: if $m_1^{-1} \gtrsim K \times 3 \text{ eV}^{-1} \gtrsim m_1^{-3}$ then for small enough $\omega_{13}$ any $\omega_{23}$ is allowed. More interestingly, for $m_1^{-1} \gtrsim K \times 3 \text{ eV}^{-1} \gtrsim m_1^{-3}$, the allowed domain is confined to bounded regions around $\omega_{13} = \omega_{23} = 0$. This fully justifies the “chimneylike” shape in Fig. 2 for $m_1^{-1} \gtrsim K \times 3 \text{ eV}^{-1}$. It also follows that the allowed region becomes wider in the $\omega_{23}$ direction as $K$ grows, see again Fig. 2. For $K$ above a certain critical value, the chimney would be wide open in the $\omega_{23}$ direction.

This is also why the results are less interesting for the inverted hierarchy—there the two heavier neutrino masses are much closer to each other and, hence, the interesting region where $\omega_{13}$ and $\omega_{23}$ are constrained turns out to be too narrow.

2. CP violation

Second, let us discuss the case when $U_v$ is an arbitrary unitary matrix. In the CKM-like parametrization

$$U_v = P_L U_{2-3}(\omega_{23}) U_{1-3}(\omega_{13}, \sigma) U_{1-2}(\omega_{12}) P_R,$$

where, as usual, $P_L = \text{diag}(e^{i\rho_1}, e^{i\rho_2}, e^{i\rho_3})$ and $P_R = \text{diag}(1, e^{i\rho_4}, e^{i\rho_5})$ are pure phase matrices, $U_{2-3}(\omega_{23})$ and $U_{1-2}(\omega_{12})$ are as above, cf. Eq. (31), and $U_{1-3}(\omega_{13}, \sigma)$ contains an extra Dirac-like phase $\sigma$ analogous to the $CP$ phase in the CKM matrix:

$$U_{1-3}(\omega_{13}, \sigma) = \begin{pmatrix} \cos \omega_{13} & 0 & \sin \omega_{13} e^{-i\sigma} \\ 0 & 1 & 0 \\ -\sin \omega_{13} e^{i\sigma} & 0 & \cos \omega_{13} \end{pmatrix}.$$ 

It is clear that $\rho_4$ and $\rho_5$ drop from the combination in the decay rates (6)–(8) and, hence, they do

---

17Note that the value corresponds to the Planck + BAO limit [54] quoted in [53], i.e., $\sum m_\nu < 0.23$ eV at 95% C.L.

18Note that the RHS of Eq. (33) is $\pi$-periodic.
not need to be considered. Since the analytics gets too complicated here let us just note that $\rho_1$, $\rho_2$, and $\rho_3$ play a very minor role in shaping the allowed parameter space and, thus, the only important phase in the game is $\sigma$; for $\sigma$ close to maximal, the strict bounds on $\omega_{23}$ can be lost for much lighter $m_1$ than in the $CP$ conserving case. As one can see in Fig. 3, for significant $\sigma$’s the $\omega_{23}$ parameter is typically out of control unless $m_1$ is taken to be very tiny [assuming again, for simplicity, the dominance of the 33 element of the RH neutrino mass matrix (28)].

B. Observables

Since there is no $U_e$ in the partial proton decay widths to charged meson and the rates (7)–(8) differ from (6) only by calculable numerical factors let us focus here solely to $\Gamma(\nu\rightarrow\pi^0\ell^+) \equiv \Gamma_{\ell}$ for $\ell = e, \mu$.

It is not difficult to see that if $\omega_{23}$ can be arbitrary (such as in the lower parts of the allowed regions in Figs. 2 and 3) there is no control over $\Gamma_{\ell}$. However, if both $\omega_{13}$ and $\omega_{23}$ are restricted, there may be an upper bound on $|\langle V_{\text{PMNS}} U_\nu \rangle_{21}|$ and, hence, on $\Gamma_{\mu}$, while there is no such feature observed in $\Gamma_{\nu}$. On the other hand, there is a strong correlation among $\Gamma_{e}$ and $\Gamma_{\mu}$ which is clearly visible in the sum of the two decay rates; indeed, there is instead a lower bound on $\Gamma_{e} + \Gamma_{\mu}$. Hence, in what follows we shall stick to these two independent observables and note that very similar features can be seen in the decay rates to $K^0$ and $\eta$ related to these by the isospin symmetry.

To proceed, one also has to take into account that both $\Gamma_{\mu}$ and $\Gamma_{e} + \Gamma_{\mu}$ in general depend on $\omega_{12}$. Since, however, these relations are linear one can derive analytic expressions for “optimal” $\omega_{12}$’s in each case such that $\Gamma_{\mu}$ is maximized and $\Gamma_{e} + \Gamma_{\mu}$ is minimized for any given values of $\omega_{13}$ and $\omega_{23}$. Focusing, for simplicity, on the $CP$ conserving case one has ($V$ stands for the PMNS matrix)

$$
\tan \omega_{12}^{\text{opt}} = \frac{V_{23} \sin \omega_{23} - V_{22} \cos \omega_{23}}{V_{21} \cos \omega_{13} - \sin \omega_{13} (V_{23} \cos \omega_{23} + V_{22} \sin \omega_{23})}
$$

for the maximal value of $\Gamma_{\mu}$ (given $\omega_{13}$ and $\omega_{23}$), whereas $\Gamma_{e} + \Gamma_{\mu}$ is (for given $\omega_{13}$ and $\omega_{23}$) minimized for

$$
\tan \omega_{12}^{\text{opt}} = \frac{V_{33} \sin \omega_{23} - V_{32} \cos \omega_{23}}{V_{31} \cos \omega_{13} - \sin \omega_{13} (V_{33} \cos \omega_{23} + V_{32} \sin \omega_{23})}.
$$

FIG. 2. The shape of the allowed parameter space ($\omega_{23}$ and $\omega_{13}$ governing $U_e$ on the horizontal axes and the minus log of the lightest neutrino mass $m_1$ on the vertical; note that $m_1$ decreases from bottom to top) in the $CP$ conserving setting discussed in Sec. III A for $K = 10$ in the upper and $K = 30$ in the lower panel, respectively. The allowed points are all those in the interior of the depicted structure. The straight cut in the lower part corresponds to the current cosmology limit on the lightest neutrino mass $m_1 \lesssim 8 \times 10^{-2}$ eV [53], see the discussion in the text.

FIG. 3. The same as in Fig. 2 for the $CP$ violating setting with the “Dirac” phase in $U_e$ set to $\sigma = \pi/2$ and $K = 20$. The net effect of a nonzero $\sigma$ is that $\omega_{23}$ remains unconstrained unless $m_1$ is really tiny [for which case the dominance of the 33 element in the RH neutrino mass formula (28) is assumed]. The effects of the “outer” phases of $U_e$ in the observables discussed in Sec. III A are small so we conveniently fixed all of them to zero.
In Fig. 4, the solid contours in the upper two panels denote $\Gamma_\mu$ in units of $0.8 \times 10^{-2}$ $\Gamma(p \to \pi^+\bar{\nu})/(|V_{\text{CKM}}|_{11})^2$, which is disfavored by cosmology [53].

![Fig. 4](image_url)

**Fig. 4.** Contour plots of the $\omega_{12}$-extremes (cf. Sec. III B) of the partial widths $\Gamma(p \to p^0\mu^+)$ (upper panels, decreasing with darkening color) and $\Gamma(p \to p^0\mu^+)+\Gamma(p \to p^0\mu^+)$ (lower panels) superimposed with the (dashed) boundaries of the regions allowed by Eq. (32) evaluated for $m_1 = 0.8 \times 10^{-2}$ eV (left), and $m_1 = 0.8 \times 10^{-3}$ eV (right), respectively. In all the plots the innermost and outermost dashed contours correspond to $K = 7$ and $K = 30$, respectively.

In Fig. 4, the solid contours in the upper two panels denote $\Gamma_\mu$ in units of $0.8 \times 10^{-2}$ $\Gamma(p \to \pi^+\bar{\nu})/(|V_{\text{CKM}}|_{11})^2$ (see Appendix C) evaluated at the point $\omega_{12}(\omega_{23}, \omega_{13}, \omega_{23}, \omega_{13})$, i.e., at its upper limits for each $\omega_{23}$ and $\omega_{13}$; similarly, the lower limits on $\Gamma_\mu$ are displayed in the lower panels (the color code is such that the decay rates decrease in darker regions). At the same time, the dashed lines are boundaries of the regions allowed by (32) for different $K$'s, i.e., the “horizontal cuts” through different “chimneys” such as those in Fig. 2 at a constant $m_1$.

Remarkably enough, if $K$ is not overly large, there is a global upper limit on $\Gamma_\mu$, and a global lower limit on $\Gamma_\mu$ on the boundaries of the relevant allowed regions. Sticking to the $(-\pi/2, +\pi/2)$ interval for both $\omega_{13}$ and $\omega_{23}$, which is fully justified by the symmetry properties of the relevant formulas, the precise position of such a maximum (minimum) could be found numerically or well approximated by taking $\omega_{13} = 0$ and the relevant $\omega_{23}$ on the boundary:

$$\cos^2 \omega_{23} = \frac{m_2^{-1} - 3 \text{K eV}^{-1}}{m_2^{-1} - m_3^{-1}}.$$  

(35)

This formula holds for both observables, i.e., for the maximum of $\Gamma_\mu$ as well as for the minimum of $\Gamma_\mu$.

In what follows, we shall focus on a pair of observables $X_\mu$ and $X_{e+\mu}$ defined conveniently as

$$X_\mu \equiv \frac{\Gamma(p \to \pi^+\mu^+)}{\frac{1}{2}\Gamma(p \to \pi^+\bar{\nu})/(|V_{\text{CKM}}|_{11})^2},$$  

(36)

$$X_{e+\mu} \equiv \frac{\Gamma(p \to \pi^0\nu^+)}{\frac{1}{2}\Gamma(p \to \pi^0\bar{\nu})/(|V_{\text{CKM}}|_{11})^2};$$  

(37)

their normalization (besides the trivial $(|V_{\text{CKM}}|_{11})^2$ piece) is fully governed by the size of the $\Gamma(p \to \pi^+\bar{\nu})$ factor studied in detail in Appendix C.

**Fig. 5** (color online). The global upper limits on $X_\mu$ (upper plot) and the global lower limits on $X_{e+\mu}$ (lower plot), cf. Eqs. (36) and (37), as functions of the lightest neutrino mass (in the normal hierarchy case). The lowermost line on the upper plot and the uppermost line on the lower plot correspond to $K = 7$, with every consecutive contour for $K$ increased by 2. The dots represent an independent numerical calculation of the same decay rates for $K = 7$ with randomly chosen real $U_{\mu}$'s; only those points satisfying (30) are permitted. The hatched area corresponds to $m_1 > 0.08$ eV which is disfavored by cosmology [53].

one just has to choose $\omega_{23} \in (0, \pi/2)$ for the former and $\omega_{23} \in (-\pi/2, 0)$ for the latter, respectively.

### C. Results

In what follows, we shall focus on a pair of observables $X_\mu$ and $X_{e+\mu}$ defined conveniently as

$$X_\mu \equiv \frac{\Gamma(p \to \pi^+\mu^+)}{\frac{1}{2}\Gamma(p \to \pi^+\bar{\nu})/(|V_{\text{CKM}}|_{11})^2},$$  

(36)

$$X_{e+\mu} \equiv \frac{\Gamma(p \to \pi^0\nu^+)}{\frac{1}{2}\Gamma(p \to \pi^0\bar{\nu})/(|V_{\text{CKM}}|_{11})^2};$$  

(37)

their normalization (besides the trivial $(|V_{\text{CKM}}|_{11})^2$ piece) is fully governed by the size of the $\Gamma(p \to \pi^+\bar{\nu})$ factor studied in detail in Appendix C.
1. CP conserving case

If $U_{\nu}$ is real and orthogonal, both analytic and numerical analyses are tractable so it is interesting to see how these compare. In the upper plot in Fig. 5, the solid lines indicate the analytic upper bounds on $X_{\mu}$ for a set of different $K$’s whereas the lower plots depict the corresponding lower bounds on $X_{e+\mu}$, respectively. The points superimposed on both plots represent the results of a numerical analysis. For that sake, $m_1$ and the three CKM-like angles $\omega_{12}$, $\omega_{23}$, and $\omega_{13}$ were chosen randomly and we fixed $K = 7$; only those points satisfying the inequality (30) are allowed in the plot. We can see that, in spite of the simple $\omega_{13} = 0$ assumption on the extremes of $X$’s, the analytic curves fit fairly well with the numerics. The agreement is slightly worse for larger $m_1$ which, however, is the case when the $\omega_{13} = 0$ approximation becomes rather rough.\footnote{It is clear from Fig. 4 that the approximation of reaching the minimum at $\omega_{13} = 0$ is more accurate for smaller $m_1$ (plots on the right-hand side) where the allowed regions are very narrow in the $\omega_{13}$ direction.}

Concerning the physical interpretation of the results there are several options of how to read Fig. 5 and similar plots given in the next section. For instance, for a fixed $K$ (assuming, e.g., one can learn more about the high-scale structure of the theory from a detailed renormalization group analysis) a measurement of $X_{\mu}$ imposes a lower limit on the mass of the lightest neutrino (e.g., $K = 7$ and $X_{\mu} \sim 0.8$ is possible if and only if $m_1 \gtrsim 10^{-2}$ eV, etc.) Alternatively, for a given $K$ and a measured value of $m_1$ one gets a prediction for $X_{\mu}$ (for example, if $K = 7$ and $m_1 \sim 10^{-2}$ eV then $X_{\mu}$ is required to be below about 0.8). Obviously, a similar reasoning can be applied to $X_{e+\mu}$.

2. CP violation

The numerical analysis for a complex $U_{\nu}$ is far more involved and, besides that, there is no simple analytics that it can be easily compared to. We allowed the three CKM-like angles and all the CP phases to vary arbitrarily within their domains and also $m_1$ was scanned randomly on the logarithmic scale. For $\sigma$ close to zero one obtains similar features in $X_{\mu}$ and $X_{e+\mu}$ as in the CP conserving case.
regardless of the other three phases $\rho_1, \rho_2, \rho_3$, see Fig. 6. If, however, also $\sigma$ is varied randomly, then both of these effects can be seen only for tiny $m_1 \lesssim 10^{-6}$ eV, cf. Fig. 7. This, at least for the case of a dominant 33 element of the RH neutrino mass formula (28), can be easily understood from the shape of the allowed parameter space depicted on Fig. 3—there is no restriction on $\omega_{23}$ for moderate $m_1$ while for $m_1$ very tiny $\omega_{13}$ and $\omega_{23}$ are again restricted to a bounded area.

IV. POTENTIALLY REALISTIC SCENARIOS

A careful reader certainly noticed that, up to now, we have left aside the fact that in the most minimal model with only a single $5_H$ in the scalar sector the size of the Yukawa matrix entering Witten’s loop is further constrained by the need to reproduce the down-quark masses. Indeed, in such a case

\[ Y_{10} \sim \frac{1}{\sqrt{2}} \frac{M_d}{v}, \]  

(38)

which, barring renormalization group running, is at most of the order of $m_2/v \sim 2\%$. Hence, in the very minimal model Witten’s loop is further suppressed and the inequality (30) cannot be satisfied unless $K$ is extremely large. In this respect, the perturbativity limits implemented in the discussion above are, strictly speaking, academic.

Another issue is the $M^M_u \propto M_d$ correlation which, regardless of the size of the proportionality factor, renders the light neutrino spectrum too hierarchical. Indeed, for $m_{LL} \propto M^T_u (M_d)^{-1} M_u$ which in the $M_d$-diagonal basis reads

\[ m_{LL} \propto W_R D_u V^T_{CKM} (D_d)^{-1} V^T_{CKM} D_d W_R^T \]

(provided $V^\prime_{CKM}$ is the “raw” form of the CKM matrix including the five extra phases usually rotated away in the SM context and $W_R$ is an arbitrary unitary matrix) one typically gets $m_2 : m_1 \sim 0.001$ while the data suggest this ratio to be close to $\frac{\sqrt{m^2_5}}{\Delta m^2_{31}} \sim 0.1$. Hence, a potentially realistic generalization of the minimal scenario is necessary together with a careful analysis of the possible impacts of the extra multiplets it may contain on the results obtained in the previous sections.

There are clearly many options on how to avoid the unwanted suppression of $Y_{10}$ and get a realistic RH neutrino spectrum in more complicated models. One may, for example, add extra vectorlike fermions that may allow large $Y_{10}$ by breaking the correlation (38), heavy extra singlets, etc. However, in many cases the structure of such a generalized scheme changes so much that some of the vital ingredients of the previous analysis are lost.

In order to deal with this, let us first recapitulate the main assumptions behind the central formula (30) underpinning the emergence of all the features in the proton decay channels into neutral mesons seen in Sec. III: First, the down-type quark mass matrix $M_d$ was required to be symmetric. This is not only crucial for the sharp prediction (5) but, on more general grounds, also to avoid the option of “rotating away” the $d = 6$ gauge-driven proton decay from the flipped $SU(5)$ altogether, cf. [28,35,41]. Second, in getting a grip on the LHS of Eq. (13) we made use of the tight $M^D_u \propto M^D_d$ correlation. Obviously, both these assumptions are endangered in case one embarks on indiscriminate model building.

A. The model with a pair of scalar 5’s

Remarkably enough, the simplest conceivable generalization of all, i.e., the model with an extra 5-dimensional scalar (which resembles the two-Higgs-doublet extension of the SM), renders the scheme perfectly realistic and, at the same time, it leaves all the key prerequisites of the analysis in Sec. III intact.

1. The Yukawa sector and flavor structure

Assuming both doublets in $5_H \oplus 5_H'$ do have nonzero projections onto the light SM Higgs the extended Yukawa Lagrangian

\[ \mathcal{L} \equiv Y'_{10} 10_M 10_{5_H} + Y'_{10} 10_M 10_{5_H'} + Y' 10_{M^2} 5_{5_H} \]

+ $Y' 10_{M^2} 5_{5_H'} + Y' 10_{M^2} 5_{5_H} + H.c.$

(39)

gives rise to the following set of sum rules for the effective quark and lepton mass matrices

\[ M^D_u = M^D_d \propto Y_{5} v_5 + Y_{5'} v_{5'}, \]

(40)

\[ M_d = M^T_d = Y_{10} v_5 + Y_{10}' v_{5'}, \]

(41)

\[ M_e = Y_{1} v_5 + Y_{1}' v_{5}, \]

(42)

Naïvely, one would say that adding three extra $3 \times 3$ Yukawa matrices (symmetric $Y'_{10}$, arbitrary $Y'_{5'}$, and $Y'_{1}'$) the predictive power of the theory would be totally ruined. However, from the perspective of the analysis in Secs. II and III the only really important change is the presence of $Y'_{10}$; adding $Y'_{5'}$ and $Y'_{1}'$ does not worsen the predictive power of the minimal setting at all because, for the former, $M^D_u \propto M^D_d$ is still maintained and, for the latter, $M_e$ remains as theoretically unconstrained as before.

Indeed, the net effect of $Y'_{10}$ is just the breakdown of the unwanted $M^D_u \propto M_d$ correlation due to an extra term in the generalized version of formula (13):
\[ M^M_\nu = \left( \frac{1}{16\pi^2} \right)^2 g^4_G (Y_{10\mu} + Y'_{10\mu})^2 \frac{(10_H)^2}{M^2_G} \times \mathcal{O}(1). \] (43)

Here \( \mu' \) is the trilinear coupling of \( S'_{10} \) to the pair of \( 10_\mu \)'s analogous to the third term in formula (15); as long as \( \mu'/\mu \) is different enough from \( v'/v \) one can fit all the down-quark masses without any need for a suppression in \( Y_{10} \) and \( Y'_{10} \).

Given this, the whole analysis in Sec. III can be repeated with the only difference that Eq. (26) becomes more technically involved (but, conceptually, it remains the same) and, with that, there is essentially just an extra factor of 2 popping up on the RHS of the generalized formula (30):

\[ \max_{i,j \in \{1,2,3\}} |(D_{\alpha} U^\dagger_{\beta} D_{\gamma} U_{\delta}^* D_{\alpha})_{ij}| \leq \frac{\alpha_G}{8\pi^2} V_G K. \] (44)

Hence, all results of Sec. III can be, in first approximation, adopted to the fully realistic case by a mere rescaling of the \( K \) factor. For example, the allowed points depicted in Fig. 6 for \( K = 8 \) in the basic model are allowed in the generalized setting with \( K = 4 \) and so on.

V. CONCLUSIONS AND OUTLOOK

In this work we point out that the radiative mechanism for the RH neutrino mass generation, identified by E. Witten in the early 1980s in the framework of the simplest \( SO(10) \) grand unified models, can find its natural and potentially realistic incarnation in the realm of the flipped \( SU(5) \) theory. This, among other things, makes it possible to resolve the long-lasting dichotomy between the gauge unification constraints and the position of the \( B-L \) breaking scale governing Witten’s graph: on one side, the current limits on the absolute light neutrino mass require \( M_{B-L} \) to be close to the GUT scale which, on the other hand, is problematic to devise in the nonsupersymmetric unifications and even useless in the SUSY case where Witten’s loop is typically canceled. In this respect, the relaxed unification constraints inherent to the flipped \( SU(5) \) scheme allow not only for a natural and a very simple implementation of this old idea but, at the same time, for a rich enough GUT-scale phenomenology (such as perturbative baryon number violation, i.e., proton decay) so that the minimal model might be even testable at the near future facilities.

In particular, we have studied the minimal renormalizable flipped \( SU(5) \) model focusing on the partial proton decay widths to neutral mesons that, in this framework, are all governed by a single unitary matrix \( U_\nu \) to which one gets a grip through Witten’s loop. Needless to say, this is impossible in the usual case when the tree-level RH neutrino masses are generated by means of an extra 50-dimensional scalar and/or extra matter fields due to the general lack of constraints on the new couplings in such models. Hence, there are two benefits to this approach: the scalar sector of the theory does not require any multiplet larger than the 10-dimensional two-index antisymmetric tensor of \( SU(5) \) and, at the same time, one obtains a rather detailed information about all \( d = 6 \) proton decay channels in terms of a single and possibly calculable parameter.

To this end, we performed a detailed analysis of the correlations among the partial proton decay widths to \( \pi^0 \) and either \( e^+ \) or \( \mu^+ \) in the final state and we observed strong effects in the \( \Gamma(p \rightarrow \pi^0\mu^+) \) partial width (an upper bound) and in \( \Gamma(p \rightarrow \pi^0e^+) + \Gamma(p \rightarrow \pi^0\mu^+) \) (a lower bound) across a significant portion of the parameter space allowed by the perturbative consistency of the model, as long as normal neutrino hierarchy is assumed and the Dirac-type \( CP \) violation in the lepton sector is small. In other cases, such effects are observable only if the lightest neutrino mass is really tiny.

Concerning the strictness of the perturbativity and/or the SM vacuum stability constraints governing the size of these effects, there are several extra inputs that may, in principle, make these features yet more robust and even decisive for the future tests of the simplest models. If, for instance, proton decay would be found in the near future (at LBNE and/or Hyper-K) the implied upper limit on the unification scale (which, obviously, requires a dedicated higher-loop renormalization group analysis based on a detailed effective potential study) would further constrain the high-scale spectrum of the theory which, in turn, feeds into the computation of Witten’s loop and, thus, the \( K \) factor; this, in reality, may be subject to stronger constraints than those discussed in Sec. II with clear implications for the relevant partial widths. To this end, there are also other high-energy signals that may be at least partially useful for this sake such as the baryon asymmetry of the Universe; although the \( U_\nu \) matrix drops from the “canonical” leading order contribution to the \( CP \) asymmetry of the RH neutrino decays in leptogenesis, the size of the effective Yukawa couplings may still be constrained and, thus, also the \( K \) factor. This, however, is beyond the scope of the current study and will be elaborated on elsewhere.

ACKNOWLEDGMENTS

The work of M. M. is supported by the Marie-Curie Career Integration Grant within the 7th European Community Framework Programme FP7-PEOPLE-2011-CIG, Contract No. PCIG10-GA-2011-303565 and by the Research proposal MSM0021620859 of the Ministry of Education, Youth and Sports of the Czech Republic. The work of H. K. is supported by the Grant Agency of the Czech Technical University in Prague, Grant No. SGS13/217/OHK4/3 T/14. The work of C. A. R. is in part supported by EU Network Grant No. UNILHC PITN-GA-2009-237920 and by the Spanish MICINN Grant No. FPA2011-22975, and No. MULTIDARK CSD2009-00064 and the Generalitat Valenciana (Prometeo/2009/091). She is grateful for the hospitality of the Institute
of Particle and Nuclear Physics of the Charles University in Prague during her visits in spring 2013. We are indebted to Stefano Bertolini and Martin Hirsch for reading through the preliminary versions of the manuscript.

APPENDIX A: THE PROTON DECAY RATES

In this appendix we rederive some of the results of paper [35] and rewrite them in our notation. The proton decay partial widths to neutral mesons in the flipped SU(5) model read

\[ \Gamma(p \rightarrow e^+ e^-) = \frac{C_1}{2} |c(e^\alpha, d^C)|^2, \]

\[ \Gamma(p \rightarrow \eta e^+ e^-) = C_2 |c(e^\alpha, d^C)|^2, \]

\[ \Gamma(p \rightarrow K^0 e^+ e^-) = C_3 |c(e^\alpha, s^C)|^2, \]

with the constants \( C_1, C_2, C_3 \) defined in (10)–(12). The \( p \)-decay widths to charged mesons obey

\[ \Gamma(p \rightarrow \pi^+ \pi^-) = C_1 \sum_{i=1}^{3} |c(\nu_i, d, d^C)|^2, \]

\[ \Gamma(p \rightarrow K^+ \pi^-) = \sum_{i=1}^{3} |B_4 c(\nu_i, d, s^C) + B_5 c(\nu_i, s, d^C)|^2, \]

where

\[ B_4 = \frac{m_d^3 - m_K^3}{2 f_{\pi} \sqrt{2\pi m^3_{\pi}}} A_1 \left[ \frac{2m^2_L}{3m_B} D, \right] \]

\[ B_5 = \frac{m_d^3 - m_K^3}{2 f_{\pi} \sqrt{2\pi m^3_{\pi}}} A_1 \left[ 1 + \frac{m_p}{3m_B} (D + 3F) \right], \]

can be obtained from the chiral Lagrangian. The flavor structure of the basic contractions can be written like

\[ c(e^\alpha, d^C) = k^2 (U_d^c (U^L_u)^\dagger \beta_\alpha (U^R_{e^L})^\dagger)_{1\alpha}, \]

\[ c(\nu_i, d^\alpha, d^C) = k^2 (U_d^c (U^L_{e^C})^\dagger \beta_\alpha (U^R_{e^L})^\dagger)_{1\alpha}. \]

Here \( k = g_G/M_G \) and the unitary matrices \( U_d, U_{d^L}, U_{e^C}, \) and \( U_{e^L}^R \) provide the diagonalization of the quark and lepton mass matrices:

\[ m_{d^L} = U_d^c D_d U_e \]

\[ M_d = U_d^c D_d U_d \]

\[ M_u = U_u^c D_u U_u. \]

Note that \( M_d \) and \( m_{d^L} \) are symmetric, hence, instead of a biunitary, a single-unitary-matrix transformation can be used to diagonalize each of them. In this notation

\[ V_{CKM} \propto U_u^L U_d^L \]

\[ V_{PMNS} \propto U_e^L U_d^L \]

where the proportionality sign turns into equality once the extra phases (unphysical from the SM perspective) are stripped down. Hence, the flavor structure of the d = 6 proton decay partial widths to neutral mesons and charged leptons is governed by

\[ |c(\nu_i, d^\alpha, d^C)|^2 = k^4 |(V_{CKM})_{1\beta}|^2 |(U^R_{e^L})^\dagger|_{1\alpha}^2. \]

For a symmetric \( M_d \) another important feature of the flipped SU(5) scheme is recovered: \( c(\nu_i, d^\alpha, d^C) \propto \delta_{\alpha\beta} \); this implies \( \Gamma(p \rightarrow K^+ \pi^-) = 0 \). Moreover, considering \( \sum_{i=1}^{3} |(U^R_u U^L_u)_{1i}|^2 = 1 \) one gets

\[ \Gamma(p \rightarrow \pi^+ \pi^-) = \frac{m_p}{8\pi f^2_{\pi}} A_2^2 |a|^2 (1 + D + F)^2. \]

APPENDIX B: THE CHOICE OF \( M_u^u \)-DIAGONAL BASIS

It is convenient to choose the basis in which \( M_u^u \) is diagonal, i.e., \( U^R_u = U^R_{u^u} = 1 \). To justify this choice, we have to prove that all the quantities of our interest are independent of this choice. This concerns, in particular, the CKM and PMNS matrices and the proton decay widths (A1)–(A5), i.e., the coefficient (A10).

First, obviously, a transformation \( U_{u^u}^R \rightarrow U_{u^u}^R V \) where \( V \) is a unitary matrix must be compensated by a simultaneous change \( U_j \rightarrow U_j V \) so that the CKM matrix (A8) remains intact. Second, changing \( U^R_{e^L} - U^R_{e^L} W \) by a unitary \( W \) requires \( U^R_{e^L} \rightarrow U^R_{e^L} W \) otherwise (A10) is not preserved. On top of that, \( U^R_{e^L} \) is related to \( U_{e^C} \) via seesaw \( m_{d^L} = U_d^c D_d U_{e^C} = M_d^R (M_d^M)^{-1} \) \( m_{d^L} = -(U^R_{e^L})^\dagger D_d U_{e^C} (M_d^M)^{-1} (U^R_{e^L})^\dagger D_d U_{e^C} \), hence also \( U_{e^C} \rightarrow U_{e^C} W \) is induced. The transformations of \( U^R_{e^L} \) and \( U_{e^C} \) then act against each other so that also the PMNS matrix (A9) remains unchanged. Thus, it is possible to choose \( U^R_{u^u} = U^R_{u^u} = 1 \) without affecting any of the quantities discussed in Secs. II and III. In the \( M_u^u \)-diagonal basis the coefficient (A10) reads

\[ |c(\nu_i, d^\alpha, d^C)|^2 = k^4 |(V_{CKM})_{1\beta}|^2 |(V_{PMNS} U_{e^C})_{1\alpha}|^2. \]

APPENDIX C: SU(3)_c \( \otimes \) SU(2)_L GAUGE UNIFICATION

In order to get any quantitative grip on the absolute scale of the proton lifetime in the model(s) of interest, in
particular, on $\Gamma(p^+ \to \pi^+ \bar{\nu})$ providing the overall normalization of the results depicted in Figs. 5–7 one has to inspect thoroughly the constraints emerging from the requirement of the (partial) gauge coupling unification. Since the model is not “grand” unified in the sense that only the non-Abelian part of the SM gauge group is embedded into a simple component of the high-energy gauge group, this concerns only the convergence of the $g_3$ and $g$ couplings of the SM. Besides the “initial condition” defined by the values of $\alpha$ and $\alpha_2 \equiv g^2/4\pi = \alpha/\sin^2 \theta_W$ at the $M_Z$ scale and the relevant beta-functions the most important ingredient of such analysis is the heavy gauge and scalar spectrum shaping the evolution of $\alpha_1$ and $\alpha_2$ in the vicinity of $M_G$ [conveniently defined as the mass of the $(X', Y')$ gauge bosons] and, ultimately, their coalescence above the last of the heavy thresholds.

As a reference setting let us start with the situation corresponding to the very simplest approximation in which all these heavy fields happen to live at a single scale ($M_G$); then, $M_G$ turns out to be at $10^{16.8} - 10^{17}$ GeV at one loop where the uncertainty corresponds to the 3-$\sigma$ band for $\alpha_1(M_Z)$ and it gets reduced to about $10^{16.6} - 10^{16.8}$ GeV at two loops.

Needless to say, such a single-mass-scale assumption is oversimplified as, in fact, the masses of the heavy colored triplet scalars $\Delta_1$ and $\Delta_2$, cf. Eq. (21) and the masses of the $(X', Y')$ gauge bosons [to quote only those states that are relevant here, i.e., $SU(3)_c \otimes SU(2)_L$ nonsinglets] depend on different sets of parameters and, hence, may differ considerably; this, in particular, applies for $\Delta_1$ that may be almost arbitrarily light if the inequality (26) gets saturated. This, obviously, may lead to a significant change in the “naive” estimate above.

In what follows, we shall focus on a simplified setting in which $\Delta_2 = \lambda_2 \equiv \lambda$ reflecting the symmetry of the relevant relations (21) and (26) under their exchange and fix $g_G = 0.5$. Hence, the masses of $\Delta_1$, $\Delta_2$, and $(X', Y')$ are fully fixed given $\lambda$, $\mu$, and $V_{G}$. This also means that if one fixes $m_{\Delta_1}$, $\lambda$, and $\mu$, then $m_{\Delta_2}$ and $M_G$ are fully determined and the unification condition can be tested. In turn, it can be used to get a correlation among the unification-compatible values of, say, $m_{\Delta_2}$ and $M_G$; the resulting situation is depicted in Fig. 8. The shape of the allowed regions therein (in particular, the relatively shallow slope of the allowed bands for a fixed proportionality factor $x$ between $\mu$ and $\lambda V_{G}$) is easily understood: the effect of integrating in the $(X', Y')$ gauge bosons (plus the relevant Goldstines in the Feynman gauge) is much stronger than that of the two colored scalars $\Delta_{1,2}$ (assuming $x < 1$, i.e., $m_{\Delta_2}$ not parametrically smaller than $m_{\Delta_1}$) and, hence, a small shift in $M_G$ is enough to compensate even for significant changes in $m_{\Delta_{1,2}}$.

To conclude, the (two-loop) unification constraints limit the allowed domain of $M_G$ to the interval stretching from approximately $10^{16.5}$ GeV attained in the bulk of the parameter space up to about $10^{17.5}$ GeV if the fine-tuned configurations with $x \sim 1$ are considered.

Witten’s loop in the minimal flipped $SU(5)$ unification revisited

Dylan Harries,* Michal Malinský,† and Martin Zdráhal‡

Institute of Particle and Nuclear Physics, Faculty of Mathematics and Physics,
Charles University in Prague, V Holešovičkách 2, 180 00 Praha 8, Czech Republic

(Received 20 August 2018; published 21 November 2018)

In the simplest potentially realistic renormalizable variants of the flipped $SU(5)$ unified model the right-handed neutrino masses are conveniently generated by means of the Witten’s two-loop mechanism. As a consequence, the compactness of the underlying scalar sector provides strong correlations between the low-energy flavor observables such as neutrino masses and mixing and the flavor structure of the fermionic currents governing the baryon and lepton number violating nucleon decays. In this study, the associated two-loop Feynman integrals are fully evaluated and, subsequently, are used to draw quantitative conclusions about the central observables of interest such as the proton decay branching ratios and the absolute neutrino mass scale.

DOI: 10.1103/PhysRevD.98.095015

I. INTRODUCTION

Though not a genuine grand unified theory (GUT), the flipped $SU(5)$ gauge theory [1–3] still attracts significant attention [4–7] due to several rather unique features it exhibits. In particular, one-stage symmetry breaking down to the standard model (SM) can be achieved regardless of whether or not a TeV-scale supersymmetry is assumed. The corresponding Higgs sector can also be very small, as it is sufficient to employ just a single 10-dimensional representation to accomplish the necessary symmetry breaking. This is to be compared to the 24 of the Georgi-Glashow $SU(5)$ [8] and/or $45 \oplus 16$ (or even $45 \oplus 126$) of the minimal $SO(10)$ GUTs (see, e.g., Refs. [9,10] and references therein).

Flipped $SU(5)$ models also share several other nice features with their truly unified cousins. From the point of view of phenomenology, two such features stand out as being particularly relevant due to their immediate experimental consequences. First, as in the $SO(10)$ GUTs, 3 right-handed (RH) neutrinos are enforced in the spectrum, allowing for the use of a type-I seesaw mechanism to generate the light neutrino masses. Additionally, as in $SU(5)$ there is only one heavy gauge boson, which typically yields somewhat stronger correlations between the flavor structure of the baryon and lepton number violating (BLNV) currents and the low-energy flavor observables, and hence one can often say quite a bit about, e.g., the proton lifetime.

However, upon closer inspection one finds a certain level of tension between the practical implications of these two points. For example, in order to implement the standard type-I seesaw with the RH neutrinos at hand, a 50-dimensional four-index scalar $50_\chi$ of $SU(5)$ is typically added [11] together with a $3 \times 3$ complex symmetric Yukawa matrix $Y_{50}$ in order to generate the desired RH Majorana mass term via a renormalizable coupling such as $Y_{50}^{ij} 10_{f_i}^T C^{-1} 10_{f_j} 50_\chi$. Besides enlarging the scalar sector enormously (and, hence, disposing of the uniquely small size of the “minimal” Higgs sector noted above as one of the most attractive structural features of the framework), the extra scalar and the associated Yukawa at play reduces the value of the low-energy neutrino masses and the lepton mixing data as constraints for the proton lifetime estimates as it essentially leaves the neutrino sector on its own.

Remarkably enough, this dichotomy may be overcome by noticing [12,13] that the RH neutrino masses in flipped $SU(5)$ models may be generated even without the unpleas-ant extra $50_\chi$ at the two-loop level by means of a variant of the mechanism first identified by Witten in the $SO(10)$ context [14]. The two main features [13] of this scenario are, first, a simple relation among the seesaw and the GUT scales where the former is, roughly speaking, given by the latter times an extra two-loop suppression and, second, a rigid correlation between the flavor structures of the neutrino and charged sectors, which in most cases may be transformed into a set of strong constraints for, e.g., the proton decay partial widths and branching ratios.

To this end, the Witten’s-loop-equipped flipped $SU(5)$ may even be viewed as the most economical renormalizable theory of the BLNV nucleon decays, much simpler
than, e.g., the potentially realistic variants of the $SO(10)$ and even the $SU(5)$ GUTs.

From this perspective, it is interesting that in Ref. [13] most of the basic features of this framework may have been identified even without an explicit calculation of the graphs involved in Witten’s mechanism. In this work we intend to overcome this drawback by a careful inspection of the Feynman graphs and their evaluation which, as we shall see, will clarify several other points left unaddressed in the preceding studies. In particular, the calculation will make it very clear that the minimal potentially realistic and renormalizable incarnation of the scheme under consideration is the variant featuring a pair of 5-dimensional scalars in the Higgs sector (besides a single copy of the “obligatory” 10-dimensional $10_H$ scalar). Second, it will be shown that, in this framework, the light neutrino spectrum is always forced to be on the heavy side (actually, within the reach of the KATRIN experiment [15]), which, among other things, provides a clear smoking gun signal of the scheme.

In Sec. II we first provide a brief review of the flipped $SU(5)$ gauge theory context, identify the Feynman graphs underpinning the radiative RH neutrino mass generation in the minimal and next-to-minimal models, and exploit the seesaw formula in order to get strong constraints on their parameter space. Section III is devoted to a detailed analysis of the relevant two-loop graphs in the scenario with one copy of the 5-dimensional scalar in the Higgs sector; this setting is simple enough to allow for a complete analytic understanding of the results. In Sec. IV these findings are used for the identification of the minimal potentially realistic model of this kind, which is subsequently shown to be strongly constrained and potentially highly predictive. Most of the technical details of the lengthy calculations are deferred to a set of Appendices.

II. FLIPPED $SU(5)$ À LA WITTEN

The defining feature of the flipped $SU(5)$ unifications is the “nonstandard” embedding of the SM hypercharge operator within its $SU(5) \otimes U(1)_X$ gauge symmetry algebra, namely

$$Y = \frac{1}{5} (X - T_{24}),$$

where $T_{24}$ stands for the usual hypercharge-like generator of the standard $SU(5)$ (normalized in such a way that the electric charge obeys $Q = T^3_L + T_{24}$) and $X$ is the unique nontrivial anomaly-free generator of the additional $U(1)$ normalized in such a way that it receives integer values over the three basic irreps accommodating each generation of the SM matter,

$$\bar{5}_M = (5, -3), \quad 10_M = (10, +1), \quad 1_M = (1, +5).$$

where the first number in brackets gives the $SU(5)$ representation and the second the charge under $U(1)_X$. Compared to the standard $SU(5)$ case, the SM matter fields $\nu^c_L$ and $\bar{d}^c_L$ are swapped with respect to their usual assignments, i.e., the former is a member of $\bar{5}_M$ while the latter resides in $10_M$. Similarly, $e^c_L$ is found in the $SU(5)$ singlet and the compulsory RH neutrino $\nu^c_R$ replaces it in the 10-plet.

As for the gauge fields, the $(24, 0) \oplus (1, 0)$ adjoint of $SU(5) \otimes U(1)_X$ in this context contains a multiplet $X\mu$ transforming under $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ as $(3, \bar{2}, +\frac{1}{6})$, plus its Hermitian conjugate, rather than the traditional hypercharge-$\frac{1}{6}$ gauge bosons of the standard $SU(5)$. The remaining degrees of freedom account for the 12 SM gauge fields and one additional heavy singlet.

The minimal Higgs sector sufficient for breaking the $SU(5) \otimes U(1)$ symmetry down to the SM and, subsequently, to the $SU(3) \otimes U(1)$ of QCD + QED consists of 10 heavy scalars, one can readily write the most non-renormalizable $^2$ Yukawa Lagrangian (suppressing all flavor indices)

$$m^2 = \frac{g^2 V_{ij}^2}{2}$$

for the gauge bosons $X\mu$, where $g_5$ is the $SU(5)$ gauge coupling. The color triplet, $SU(2)_L$ singlet components of $10_H$ and $\bar{5}_H$ also mix at this stage to form a pair of massive color triplets $\Delta_{1,2}$ transforming under the SM gauge symmetry as $(3, 1, -\frac{1}{3})$, with masses $m_{\Delta_{1,2}}$. Further details regarding the tree-level scalar spectrum in this minimal flipped $SU(5)$ model are given in Appendix B.

For the above embedding of the SM matter content and minimal set of Higgs scalars, one can readily write the most general renormalizable $^2$ Yukawa Lagrangian (suppressing all flavor indices)

\[095015-2\]

---

^1It may be worth pointing out here that, due to the nonzero $U(1)_X$ charge of $10_H$ inherent to the flipped $SU(5)$ models, there is no way to build a nonrenormalizable $d = 5$ operator (presumably Planck-scale suppressed) that might, in the broken phase, affect the gauge-kinetic form and hence introduce significant theoretical uncertainties in the high-scale gauge-matching conditions and the determination of the GUT scale. As a result, one of the primary sources of irreducible uncertainties hindering the predictive power of the “standard” GUTs (such as the Georgi-Glashow $SU(5)$ or the nonminimal $SO(10)$ models with either 54 or 210 breaking the unified symmetry) is absent from this class of models.

^2Note that in nonrenormalizable settings the benefits of the scheme may be lost as the Witten’s loop contribution may be swamped by the effects of, e.g., the $d = 5$ nonrenormalizable operators of the $10_H 10_M 10_H 10_H$ type.
\[ \mathcal{L} \equiv Y_{10}10_M10_M^5H + Y_{\bar{5}}10_M\bar{5}_M^5H + Y_{1}\bar{5}_M1\bar{5}_M^5H + \text{H.c.} \]  
(4)

with \(Y_{10}, Y_{\bar{5}}, \text{and } Y_1\) denoting the relevant 3 \(\times\) 3 complex Yukawa coupling matrices; note that the first of these, unlike the latter two, is required to be symmetric in its flavor indices, i.e., \(Y_{10} = Y_{10}^\dagger\). In the broken phase, the second term in Eq. (4) yields a strong correlation among the Dirac neutrino mass matrix \(M_D\) and the up-type quark mass matrix \(M_u\), namely,

\[ M_D^T = M_u \]  
(5)

at the GUT scale. The flavor symmetric nature of \(Y_{10}\) also means that the down-type quark mass matrix satisfies \(M_d = M_d^T\), while the couplings in Eq. (4) say nothing specific about the mass matrix \(M_e\) of the charged leptons. As we shall see, these correlations will turn out to be central for the high degree of predictivity of this framework\(^3\) entertained in the following sections.

A. The RH neutrino masses and type-I seesaw

So far, we have left aside any discussion of the physical light neutrino masses in the current scenario. Obviously, Eq. (5) cannot be the whole story as it corresponds to overly large Dirac neutrino masses; the only case when this may be acceptable would be within a variant of the seesaw mechanism.

At the tree level, this could be achieved in the most natural way by employing a 50-dimensional scalar\([11]\) coupled to the 10\(C\) \(\times\) 10\(M\) fermionic bilinear; the VEV of a singlet therein then gives rise to the desired RH neutrino mass term. However, as was noted in Sec. I, the associated single-purpose extra Yukawa matrix does not bring any additional insight into the flavor structure of the model, and limits the extent to which low-energy data can be used in constraining proton decay observables. Therefore we do not adopt this option here; instead, we inspect the quantum structure of the minimal model for the desired effect.

Remarkably enough, there is no way to generate the desired RH neutrino mass in the current model\([13]\) at the one-loop level. This, however, does not mean that there is no one-loop contribution to the neutrino masses generated at all; indeed, in the presence of scalars with quantum numbers \((3, 1, -\frac{1}{2})\) and \((\bar{3}, \bar{2}, +\frac{1}{2})\) the LH Majorana mass can be devised via a “colored” variant of the notorious Zee mechanism\([16-19]\). However, this does not bring any relief to the Dirac mass issue above as, without additional

\[^3\text{To this end, it is worth noting that these relations remain intact even in models with more than a single copy of } S_H \text{ in the scalar sector; as we shall see, this (especially the symmetry of } M_D \text{) will be crucial for the construction of the minimal potentially realistic scenario identified in Sec. IV B.}\]

---

**FIG. 1.** The two nonequivalent topologies of the two-loop graphs contributing to the RH neutrino Majorana mass in the minimal flipped SU(5) model under consideration. The vector field X corresponds to the \((3, \bar{2}, +\frac{1}{2})\) component of the adjoint while the pair of Δ’s are the two mass eigenstates of the \((3, 1, -\frac{1}{2})\) colored scalars mixed from the relevant components of 10\(H\) and \(S_H\), respectively.

1. The Witten’s loop structure

The simultaneous presence of the diquark-type of interactions, mediated by the \(X_p\) and \(Δ_{1,2}\) bosons, together with their leptoquark counterparts (involving the same set of fields) in the flipped SU(5) framework implies that diagrams generating the desired RH Majorana neutrino mass can be drawn at two loops. Let us note that the corresponding pair of topologies depicted in Fig. 1 can be viewed as reduced versions of Witten’s original SO(10) graph(s)\([14]\).

In what follows we shall work in the broken phase perturbation theory with masses in the free Hamiltonian\(^4\) and in the unitary gauge so that there are no Goldstone modes around. This reduces the number of relevant graphs considerably, at the cost of making the Feynman integrals somewhat more complicated compared to other cases.

Based on the graphs in Fig. 1 that remain in this approach, it is immediately possible to make several comments on both the flavor structure and overall scale of the generated Majorana mass matrix \(M_{\mu}^H\). The flavor structure in particular plays a central role in what follows, and is governed by the Yukawa couplings appearing in each of the contributing graphs. In each of the two topologies there is only a single Yukawa coupling present, associated with the couplings of \(Δ_i\) to the fermions. These couplings involve only the \(S_H\) components of \(Δ_i\), since it is only these

\[^4\text{Hence, we are avoiding the need to sum over an infinite tower of graphs (like the one drawn in Witten’s original work [14]) with increasing numbers of VEV insertions. On the other hand, the explicit proportionality to the } μ \text{ parameter governing the mixing between the } 10_H \text{ and } S_H \text{ multiplets (see Appendix B), which is obvious in the massless perturbation theory, becomes more involved in the massive case where } μ \text{ emerges at the level of the relevant mixing matrix in the scalar sector, Eq. (B8).}\]
components that couple to the fermions through the Yukawa interactions in Eq. (4). Moreover, since all of the fermions appearing in the two graphs in Fig. 1 reside in 10M, the single relevant Yukawa coupling matrix is the symmetric Y_{10}. Hence, in the minimal model there is a tight correlation between the radiatively generated RH neutrino Majorana mass matrix and the mass matrix of the down-type quarks, making the scheme rather predictive.

The overall scale of M^U_{ν}, on the other hand, depends on both the Yukawa couplings in Y_{10} as well as the gauge couplings and the sizes of the mass parameters entering into each of the graphs. One can initially estimate it to be proportional to the dominant mass entry in the relevant graphs suppressed by the appropriate two-loop factor and the combination of gauge (entering raised to the fourth power) and Yukawa couplings.

Of the various mass parameters appearing in the evaluation of the graphs, the fermionic masses m_ν should play no role in the integrals as the singlet Majorana mass generation does not rely on the electroweak symmetry breaking. Hence, in dealing with the Feynman integration we shall work in the chiral limit with all SM fermions massless. This, in principle, may lead to spurious IR divergences in the form of, e.g., log(m_ν/Q) arising in individual partial fractions of the integrands, where Q is the renormalization scale, but as a whole M^U_{ν} should be stable in the m_ν → 0 limit.

Similarly, it is natural to expect that in the other extreme case corresponding to one of the scalars Δ_ν becoming significantly lighter with respect to the X_μ boson masses (and, hence, bringing about another practically massless propagator) M^M_{ν} should also remain regular; hence, the only mass that can make it to the denominators in the final result is m_X. This also suggests that, barring the couplings, each individual graph should be governed by powers of the m_ν/m_X ratio which, in turn, makes it merely a function of a single^5 parameter.

One more comment concerning the relative size of the aforementioned one-loop LH Majorana neutrino mass contribution is worth making here. On purely dimensional grounds, it is indeed expected to be significantly smaller than the “standard” type-I contribution due to M^M_{ν} and M^D_{ν}. First, the corresponding graphs will be inversely proportional to the relevant scalar leptoquark masses,^6 which are well above the typical seesaw scale ballpark of 10^{12−13} GeV. Second, the loop factor of 1/16π^2 will further suppress this contribution placing it, eventually, at the level of some 10^{-6} eV, which makes it negligible for the current analysis.

2. Seesaw as the key to the phenomenology

Before coming to the evaluation of the graphs in Fig. 1 it is important to stress that this is not all just an academic exercise; quite to the contrary, the information obtained in Sec. III has a profound impact on the phenomenology of the model.

The point is that, due to the seesaw formula, M^U_{ν} is related to the physical light neutrino mass matrix m_{LL} and the Dirac neutrino mass matrix M^D_{ν} via

\[ M^U_{ν} (m_{LL})^{-1} (M^D_{ν})^T = -M^M_{ν}. \]

Using Eq. (5), this can be conveniently written as

\[ W_ν ≡ D_u U_ν (m^\text{diag})^{-1} U^T_ν D_u = -M^M_{ν}. \]

where D_u is the diagonal form of the up-type quark mass matrix and U_ν is the matrix diagonalizing m_{LL}, i.e., m_{LL} = U^T_ν m^\text{diag}_ν U_ν. Note that in the derivation above we have implicitly adopted the basis in which the up-type quark mass matrix is real and diagonal, see Ref. [13] for further information.

Hence, up to an a priori unknown unitary matrix and the overall light neutrino mass scale, parametrized e.g., by the mass of the heaviest of the light neutrinos m^\text{max}_ν, the matrix W_ν defined in Eq. (7) is completely determined by the low-energy quark masses and neutrino oscillation data. This is to be compared with M^M_{ν} appearing as the right-hand side of Eq. (7), which is set by the heavy spectrum of the model (i.e., the masses of the heavy triplet scalars and gauge bosons) and the gauge and Yukawa couplings, and is therefore subject to other strong constraints. In particular, m_X, m_Δ, and g_5 must be such that the unification pattern is consistent with the low-energy data and compatible with the nonobservation of proton decay with at least 10^{34} years of lifetime [20].

Hence, demanding consistency of Eq. (7) with the data one can derive constraints on m^\text{max}_ν and, in particular, on U_ν, which is central to the BLNV phenomenology of the model. Indeed, U_ν drives all the proton decay branching ratios into neutral mesons including the “golden channel” p → π^0ν→̅ final state:

\[ \frac{Γ(p → π^0ν→̅)}{Γ(p → π^+ν)} = \frac{1}{2} |(V^\text{CKM}_{11})^2|(V^\text{PMNS}_ν U_ν)_{α1}|^2, \]

\[ \frac{Γ(p → ν_α)}{Γ(p → π^+ν)} = \frac{C_2}{C_1} |(V^\text{CKM}_{11})^2|(V^\text{PMNS}_ν U_ν)_{α1}|^2, \]

\[ \frac{Γ(p → K^0ν_α)}{Γ(p → π^+ν)} = \frac{C_3}{C_1} |(V^\text{CKM}_{12})^2|(V^\text{PMNS}_ν U_ν)_{α1}|^2, \]

Assuming, implicitly, that the renormalization scale dependence eventually disappears as a consequence of the assumed UV-finiteness of the full result.

It is perhaps worth mentioning that the scalar (S) with the SM quantum numbers (3, 2, +1/2) is formally absent in the unitary gauge as it is the would-be Goldstone mode giving mass to the X_μ vector; however, the same effect is then generated via the corresponding graphs with X_μ instead of S.
where the $C_i$’s are various low-energy factors calculable using chiral Lagrangian techniques (see, e.g., Ref. [21] and references therein) and $V_{\text{CKM}}$ and $V_{\text{PMNS}}$ are the Cabibbo-Kobayashi-Maskawa and the Pontecorvo-Maki-Nakagawa-Sakata mixing matrices, respectively.

In this sense, the minimal flipped $SU(5)$ unification equipped with the Witten’s loop mechanism can be viewed as a particularly simple (if not the most minimal of all) theory of the absolute neutrino mass scale and, at the same time, the two-body BLNV nucleon decays.

### B. Consistency constraints and implications

Let us now work out the aforementioned consistency constraints in more detail and give some basic examples of their possible implications. First, it should be noted that there is a lower limit on the largest entry of $W_\nu$ that depends on $m_{\nu}^{\text{max}}$ and the shape of $U_\nu$. Taking into account the typical 50% reduction of the running top quark Yukawa between $M_T$ and the unification scale (at around $10^{16}$ GeV) and taking, e.g., $m_{\nu}^{\text{max}} = 1$ eV and $U_\nu = 1$ one finds that the $(3,3)$ entry of $W_\nu$ is as large as about

$$|W_\nu(33)| \sim 6.4 \times 10^{12} \text{ GeV}. \quad (9)$$

The same magnitude, however, may not so easily be achieved for the $(3,3)$ entry of $M^M_\nu$ as required by Eq. (7) due to the generic $10^{-3}$ geometrical suppression in the relevant two-loop graphs and a possible further suppression associated with the Yukawa coupling $Y_{10}$; the latter may be especially problematic in the minimal scenario (4) because then $Y_{10}$ is fixed by the down-type quark masses and, thus, brings about another suppression of some $10^{-2}$ to $(M^M_\nu)_{33}$.

However, this correlation is loosened if there is more than a single copy of $5_H$ in the scalar sector. As was already indicated in Ref. [13], the additional $Y'_{10}$ associated to an extra $S'_{10}$ can conspire with the original $Y_{10}$ to do two things at once: they may partially cancel in the down-type quark mass formula to account for the moderate suppression of $M_d/M_Z$ yet their other combination governing $M^M_\nu$ (weighted by the appropriate scalar mixings) may still remain large, thus avoiding the problematic additional $10^{-2}$ suppression. In what follows, we shall model this situation by imposing a humble $|y| \lesssim 4\pi$ perturbativity criterion on all the $Y_{10}$ and $Y_{10}'$ entries.

However, even in such a case the $\sim 10^{13}$ GeV lower limit on the largest entry $(W_\nu(33))$ may still be problematic because, for $U_\nu \neq 1$, it may be further enhanced by the admixture of the yet larger (2.2) and, in particular, the (1,1) entry of $(m_\nu^{\text{diag}})^{-1}$; as a matter of fact the latter is not constrained at all given that the lightest neutrino mass eigenstate may still be extremely light. Thus, the lower bound on the magnitude of the largest element of $W_\nu$ gets further boosted over the naïve estimate of $10^{13}$ GeV whenever $U_\nu$ departs significantly from unity, which in turn constrains all of the partial widths, Eqs. (8).

Hence, a thorough evaluation of the graphs in Fig. 1 will decide several important questions, namely:

1. Can the elements of $M^M_\nu$ ever be big enough to be consistent (at least in the most optimistic scenario with $U_\nu \sim 1$) with $W_\nu$, as required by Eq. (7), in the case of the single $5_H$ scenario with its typical extra $10^{-2}$ suppression at play?

2. If not, can the two-$5_H$ scenario work? What would then be the corresponding lower limit for $m_{\nu}^{\text{max}}$ in this scenario?

3. In either case, what is the allowed domain for the entries of $U_\nu$ and, thus, for the corresponding BLNV nucleon decay rates?

This is what we turn our attention to in the remainder of this article.

### III. WITTEN’S LOOP CALCULATION

The leading contribution to the radiatively generated RH neutrino mass in the current scheme may be computed by considering the graphs in Fig. 1 evaluated at zero external momentum, see Appendix C, with the relevant interaction terms given in Appendix A. In the minimal renormalizable model containing only a single $10_H$ and one or more $5_H$ representations, no one-loop contribution to the RH neutrino mass matrix can be generated, nor do there exist any one-loop counterterm graphs. The resulting expression for the RH Majorana neutrino mass matrix in the case of a single $5_H$ multiplet reads

$$M^M_\nu = -\frac{3g^2}{4\pi^2} Y_G \sum_{i=1}^{2} (-8Y_{10}'(U_\Delta)_{ij}(U_\Delta')_{ij}I_3(s_i)), \quad (10)$$

where the scalar mixing matrix elements $(U_\Delta)_{ij}$ are given in Appendix B, and $I_3(s_i)$ is the sum of the corresponding loop integrals evaluated at zero external incoming momentum,

$$I_3(s_i) = -(4\pi)^4(\Sigma_i^\mu(0) + 2\Sigma_i^\mu(0)), \quad (11)$$

regarded as a function of $s_i = m_{\Delta_i}/m_{\Delta_j}^2$. Recall that there is an overall extra factor of 2 included in Eq. (10) related to the permutation of the two external neutral field lines (for $I = J$) or to the symmetry of $Y_{10}$ (for $I \neq J$). The integrals $\Sigma_i^\mu(0)$ and $\Sigma_i^\mu(0)$, corresponding to topology 1 and 2 respectively, are given by

$$i\Sigma_i^\mu(0) = i \int \frac{d^4p}{(2\pi)^4} \frac{d^4q}{(2\pi)^4} \frac{1}{\Delta^0_{\mu} - q^0} \frac{1}{p^0 - q^0} \times \left\{ -g^{\mu\nu} + \frac{1}{m_{\Delta_i}} p^\mu q^\nu + g^{\mu\nu} - \frac{1}{m_{\Delta_i}} q^\mu q^\nu \right\}, \quad (12)$$
\[ i \Sigma^P_2(0) = i \int \frac{d^4p}{(2\pi)^4} \int \frac{d^4q}{(2\pi)^4} \frac{1}{-q^\mu p^\mu} \frac{1}{q^2 - m^2_{\Delta_i}} \times \frac{-g^\mu_\nu + \frac{1}{m^2_{\chi}} p^\mu p^\nu - g^\mu_\nu}{(p + q)^2 - m^2_{\chi}} \times \frac{1}{p^2 - m^2_{\chi}}. \]

(13)

The integrals in Eqs. (12) and (13) are evaluated by reducing them to expressions involving (variants of) the brackets by Veltman and van der Bij [22], which may be evaluated directly [22–27]. The details of this reduction, and the resulting analytic expressions for the two-loop integrals, are given in Appendix D. In particular, using the results given in Ref. [22] and appropriate generalizations thereof, it is found that the contributing brackets are free of potential IR divergences in the limit of massless internal fermions, such that the fermion masses may safely be allowed to vanish as in Eqs. (12) and (13). On the other hand, each graph is individually UV divergent. Setting \( e = 2 - \frac{D}{2} \), where \( D \) is the spacetime dimensionality, the divergences are found to be

\[ -(4\pi)^4 \Sigma^P_1(0) = \frac{3}{2e} - \frac{m^4_{\Delta_i}}{2m^4_{\chi}} \left( 1 - \frac{3}{2e} \right) - \frac{1}{e} \frac{m^2_{\Delta_i}}{Q^2}, \]

(14)

and

\[ -(4\pi)^4 \Sigma^P_2(0) = -\frac{3}{4e} + \frac{m^4_{\Delta_i}}{4m^4_{\chi}} \left( 1 - \frac{3}{2e} \right) - \frac{1}{e} \frac{m^2_{\Delta_i}}{Q^2}. \]

(15)

It follows from Eq. (11) that the total contribution \( I_3(s_j) \) to the RH neutrino mass matrix is UV finite, as must be the case here due to the absence of the necessary counterterms.

IV. RESULTS

The behavior of the result for the purely kinematic piece of the RH neutrino mass matrix, \( I_3(s_j) \), is shown in Fig. 2. Notably, the magnitude of \( I_3(s_j) \) is bounded for all \( s_j \geq 0 \). Indeed, from the analytic result given in Eq. (D31), one has that for \( s \to 0 \),

\[ I_3(s \to 0) = 3 + s \left( 3 \log s + s^2 - \frac{15}{2} \right) + O(s^2 \log^2 s), \]

(16)

while in the opposite limit with \( s \to \infty \),

\[ I_3(s \to \infty) = -3 + O(s^{-1} \log^2 s). \]

(17)

FIG. 2. Plot of the function \( I_3(s) \) appearing in the RH neutrino mass matrix.

A. RH neutrino masses in the minimal model

With \( I_3(s_j) \) determined, we may proceed to evaluate the size of \( M^M_\nu \) in Eq. (10). Substituting in the explicit forms of the mixing matrix elements in Eq. (B11) one obtains

\[ M^M_\nu = \frac{-3g^4}{(4\pi)^4} (-8Y_{10}) V_G \tilde{I}, \]

(18)

where

\[ \tilde{I} = \sum_{i=1}^{2} \frac{2\nu^\nu (2\lambda_2 + g^2_{3i} s_i)}{4|\nu|^2 + (2\lambda_2 + g^2_{3i} s_i)^2} I_3(s_i), \]

(19)

and \( \nu = \mu/V_G \). We note that \( \tilde{I} \to 0 \) as \( \mu \to 0 \), reflecting the fact that the graphs rely on the \( 10H - 5H \) mixing. It is also clear from Eq. (19) that, since \( I_3(s) \) is bounded, \( \tilde{I} \) cannot be made arbitrarily large to compensate for the suppression factors noted in Sec. II. To develop some sense of the allowed size of \( \tilde{I} \), it is useful to substitute for \( s_j \) from Eq. (B7) and inspect \( \tilde{I} \) as a function of \( \nu, \lambda_2, \lambda_5 \), and \( g_3 \), neglecting all terms that are of order of \( \nu^2 / V_G^2 \), where \( \nu \) is the electroweak VEV, see Eq. (B2). Requiring that the tree-level vacuum be locally stable implies [13] \( \lambda_{2,5} < 0 \) and

\[ |\nu| \leq \sqrt{\lambda_2 \lambda_5}. \]

(20)

When this bound is saturated, i.e., when \( |\nu| = \sqrt{\lambda_2 \lambda_5} \), the mass \( m_{\Delta_i} \) vanishes for all values of \( \lambda_2, \lambda_5 \) while \( m^2_{\chi_i} = -(\lambda_2 + \lambda_5) V_G^2 \). The resulting value of \( \tilde{I} \) for this special case is shown in the \((\lambda_2, \lambda_5)\) plane in Fig. 3. In particular, it should be noted that the value of \( \tilde{I} \) is unchanged under the interchange \( \lambda_2 \leftrightarrow \lambda_5 \), as can be easily verified from Eqs. (19) and (B7), and \( |\tilde{I}| \leq 3 \) for all values of \( \lambda_2 \) and \( \lambda_5 \). The maximal value of \( |\tilde{I}| \) is achieved for \( \lambda_2 = \lambda_5 \), with \( |\tilde{I}| \to 3 \) as \( \lambda_2 = \lambda_5 \to -\infty \).
Qualitatively different behavior results in the more general case that \(\nu\) does not saturate the bound given in Eq. (20). In this minimal model the problem is exacerbated by the fact that \(\lambda_2\) and \(\lambda_3\). Consequently, this state never decouples and \(I\) therefore does not vanish. Technically, this arises because \(I_3(s_1) = 3\) while \(I_3(s_2) \to -3\), with the two contributions still entering \(I\) with coefficients of equal magnitude but opposite sign.

However, even in the most optimistic case with \(|I| \to 3\), the above results make it clear that there is little hope for a viable prediction of the light neutrino spectrum in the minimal scenario under consideration. For acceptable values of \(m_\chi \sim 10^{17} \text{ GeV}\), and taking \(g_s \approx 0.5\), the elements of \(M^4\) are found to be \(\lesssim 10^{12} \text{ GeV}\) after taking into account the \(\sim 10^{-2}\) suppression associated with presence of \(Y_{10}\). This is to be compared with the (optimistic) lower bound of \(\sim 10^{13} \text{ GeV}\) for the elements of the left-hand side of Eq. (7). Evidently, in the case when only a single \(S_H\) is present in the spectrum the answer to whether Eq. (7) can be satisfied is negative. In fact, in this minimal model the overall predictive power of the theory is not significantly harmed by this addition; in particular, doing so does not spoil the key Yukawa relations used in obtaining Eq. (7). With a second \(S_H\) multiplet, the Yukawa sector of the model reads

\[
\mathcal{L} \ni Y'_{10}^{10} M_{10}^{10} S_H + Y'_{10}^{10} M_{10}^{10} S_H' + Y_3 M_{10}^{5} S_H + Y_5 M_{10}^{5} S_H' + Y_1 M_{10}^{5} S_H + Y_1 M_{10}^{5} S_H' + \text{H.c.,}
\]

where \(Y'_{10}\) is of course also flavor symmetric. In this scenario, the Dirac neutrino mass matrix still remains tightly correlated with the up-type quark masses, with the GUT scale relation

\[
M_D^0 = M_u^0 \propto Y_5 v + Y_5' v'
\]

holding at tree-level, where \(v'\) is the VEV associated with the electrically neutral component of \(S_H\). By contrast, the analogous relationship between the

![FIG. 3. Contour plot of \(I\) as defined in Eq. (19), in the \((\lambda_2, \lambda_3)\)-plane, with \(g_s = 0.5\) and \(\nu = \alpha \sqrt{\lambda_2 \lambda_3}\) for \(\alpha = 1\), corresponding to the maximal value of \(|\nu|\) consistent with a locally stable SM vacuum.](image1)

![FIG. 4. Plot of the range of variation of \(I\) as a function of \(\lambda_2 = \lambda_3 = \lambda\), with \(g_s = 0.5\) and \(\mu = \alpha \sqrt{\lambda_2 \lambda_3} V_G\), for \(\alpha \in [0, 1]\). The dashed vertical line denotes the naïve perturbativity limit \(|\lambda| \leq 4\pi\).](image2)

each in Eq. (19) are equal in magnitude but of opposite sign, resulting in the two terms cancelling. Physically, this corresponds to the expected dynamical decoupling of the heavy scalar states in the \(m_{\lambda_2} \to \infty\) limit.
down-type quark masses and the generated RH neutrino Majorana masses, $M_d$, $M^M_d \propto Y_{10}$, is no longer preserved. While $M_d \propto Y_{10} \nu + Y_{10} \nu'$, the appropriate generalization of Eq. (10) reads

$$
(M^M_d)^{ij} = - \frac{3G_F}{4\pi^2} V_C \sum_{i=1}^{3} \sum_{j=2}^{3} (-8Y^{ij}_Y)(U^\Delta_{1i})(U^\Delta_{3j})J^3(s_i).
$$

(24)

where $Y_j = Y_{10}$ when $j = 2$ and $Y_j = Y'_{10}$ when $j = 3$, with $U_\Delta$ now a $3 \times 3$ mixing matrix as defined in Eq. (B16). Thus, in general, $M_d$ and $M^M_d$ are determined by different linear combinations of the Yukawa couplings $Y_{10}$ and $Y'_{10}$. In turn, this means that the generic suppression of $M^M_d$ by a factor $\propto M_d$ may be avoided in the two-$5_H$ scenario. On the other hand, it is still the case that the elements of $M^M_d$ are bounded from above, at least so long as it is required that all couplings remain perturbative.

1. Phenomenology of the minimal potentially realistic model

As the ignorance of yet higher-order effects makes any such perturbativity constraints somewhat arbitrary in general, in what follows we shall give two examples of the $M^M_d$ estimates corresponding to two different choices of the upper limits on the effective (running) SM down-quark Yukawa couplings. These, according to Eq. (A3), obey $Y_d \equiv 8Y_{10}$ and $Y'_d \equiv 8Y'_{10}$ at the matching scale. The two cases to be considered are (i) $|Y_d|$, $|Y'_d| \lesssim 1$ and (ii) $|Y_d|$, $|Y'_d| \lesssim 4\pi$. For the former case (motivated by the SM value of the top Yukawa coupling) one has the following upper limit on $M^M_d$ calculated from Eq. (24)

case i) $|M^M_d| \lesssim 6.4 \times 10^{12} \left(\frac{m_X}{10^{17} \text{ GeV}}\right)$ GeV. (25)

while for the latter one obtains

case ii) $|M^M_d| \lesssim 8.0 \times 10^{13} \left(\frac{m_X}{10^{17} \text{ GeV}}\right)$ GeV. (26)

Note that in both cases we have used the (numerical) upper limit

$$
\left| \sum_{i=1}^{3} \sum_{j=2}^{3} (U^\Delta)^{1i}(U^\Delta)^{3j}s_i \right| \lesssim 3
$$

(27)

which is completely analogous to the limit discussed in Sec. IVA for the single-$5_H$ case.

Remarkably, for the typical flipped $SU(5)$ value of $m_X = 10^{17}$ GeV (see, e.g., Ref. [13]) the case (i) limit, Eq. (25), is just on the borderline of compatibility with the optimistic lower limit in Eq. (9) on $|W_\nu|$, while the latter case (ii) in principle admits lower values of $m_X$.

This, in turn, implies that there is generally not much room for any significant admixture of the second neutrino (inverse) mass within the element $(U_\nu)_{13}$, hence, the only allowed $U_\nu$’s in Eq. (7) are those for which $(U_\nu)_{13}$ and $(U_\nu)_{23}$ are small.

To this end, the model clearly calls for a dedicated numerical analysis including a detailed calculation of the heavy spectrum that conforms to, among other things, the requirement of a significant spread of the scalar triplets in order to maximize $[\tilde{T}]$. This, however, is beyond the scope of the current study and will be elaborated on elsewhere.

At this point, let us just illustrate the typical situation by evaluating the most significant proton-decay two-body branching ratios (neglecting the kinematically suppressed vector-meson channels for simplicity) in the $(U_\nu)_{13} = (U_\nu)_{23} = 0$ limit with the 1-2 mixing angle $\theta_{12}$ therein chosen in such a way that $\Gamma(p \to X^0 \mu^+)$ is maximized (see Ref. [13] for further details):

$$
\begin{align*}
\text{Br}(p \to \pi^0 \mu^+) &\approx 80.0\%, \\
\text{Br}(p \to \pi^0 e^+) &\approx 14.2\%, \\
\text{Br}(p \to \pi^0 \mu^+) &\approx 5.5\%, \\
\text{Br}(p \to K^0 e^+) &\approx 0.1%.
\end{align*}
$$

(28)

Needless to say, for nonextremal values of $\theta_{12}$ these branching ratios may vary; in particular, $\text{Br}(p \to \pi^0 e^+)/\text{Br}(p \to \pi^0 \mu^+)$ should increase.

Finally, let us say a few words about the lower limits on the mass of the heaviest SM neutrino in the two cases (25) and (26). As for the former, one obtains

$$
m_3 \gtrsim \left(\frac{10^{17} \text{ GeV}}{m_X}\right) \text{ eV}
$$

(29)

while for the latter one has

$$
m_3 \gtrsim 0.08 \left(\frac{10^{17} \text{ GeV}}{m_X}\right) \text{ eV}
$$

(30)

which, actually, turns out to be independent on the specific form of the $U_\nu$ matrix as long as the 1-3 and 2-3 mixings therein are small (see the discussion above). With this at hand, any specific experimental upper limit on the absolute

$\tilde{T}$These, however, may not be that simple to get within potentially realistic unification chains, see Appendix C of Ref. [13].

$\tilde{T}$Given the structure of the seesaw formula in the current context (7) together with the tight constraints on the structure of the $U_\nu$ matrix we generally assume the hierarchy of the light neutrino mass eigenstates to be normal.
V. CONCLUSIONS AND OUTLOOK

The two-loop radiative RH neutrino mass generation mechanism originally identified by Witten in 1980s in the $SO(10)$ context finds a beautiful incarnation in the class of renormalizable flipped $SU(5)$ unified theories where, among other effects, it avoids the need for the 50-dimensional tensor. This, in turn, renders the simplest potentially realistic scenarios perhaps the most minimal (partially) unified gauge theories on the market, with strong implications for some of the key beyond-standard-model observables such as the absolute neutrino mass scale and proton decay.

In this work we have focused on a thorough evaluation of the relevant Feynman graphs in these scenarios paying particular attention to their analytic properties and the absolute size of the effect which turns out to be the key to the consistency of the scenario as a whole. It has been shown that there is no way to be consistent with the data with only one 5-dimensional scalar multiplet at play and, hence, the minimal potentially realistic setup must include two such irreps in the scalar sector (along with the 10-dimensional tensor).

As it turns out, such a minimal flipped $SU(5)$ model is subject to strong constraints on its allowed parameter space that lead to rather stringent limits on the absolute light neutrino mass scale as well as the BLNV two-body nucleon decays. A thorough numerical analysis of the corresponding correlations is deferred to a future study.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the Grant Agency of the Czech Republic (GAČR) under the Contract No. 17-04902S. We would like to thank Helena Kolešová, Jiří Horčíček, Jiří Novotný, Catarina Simões and Diego Aristizábal Sierra for illuminating discussions.

APPENDIX A: THE INTERACTION LAGRANGIAN

The radiative generation of the RH neutrino masses involves only a small subset of the interactions associated with the full flipped $SU(5)$ Lagrangian. Working in the $SU(5) \otimes U(1)_X$ broken phase, we extract the required interactions from the kinetic terms and general Yukawa Lagrangian, Eq. (4), making use of FELNRULES [28,29] and FEYNARTS [30,31] to verify that all terms and contributing diagrams are accounted for. As discussed in Sec. II, when the model contains only a single $\tilde{5}_H$ representation the relevant diagrams are found to be those in Fig. 1, arising from the interaction Lagrangian

$$L_{\text{int}} \supset \frac{g_3^2}{2} e_{ijk} \epsilon^{\beta \mu \nu} V_{G} X_i Y_{\mu} \hat{D}_{ij} + \frac{g_5}{\sqrt{2}} e_{ijk} X_{\mu} d_{L_j} \rho^{\nu} Q_{L_i}^\mu$$

$$+ \frac{g_5}{\sqrt{2}} e_{ijk} \epsilon_{\mu \nu} (Q_{L_i})_{ij} \rho^\nu \nu^\nu - 8 Y_{I}^{HJ} \epsilon^\nu \nu^\nu_{L_i} C^{-1} \nu^\nu T_i$$

$$- 4 Y_{I}^{HJ} e_{ijk} \epsilon_{\mu \nu} (Q_{L_i})_{ij} T^\nu C^{-1} Q_{L_j}^\mu T^k + \text{H.c.}$$

(A1)

where $i, j, k$ and $\alpha, \beta$ denote the $SU(3)_C$ and $SU(2)_L$ indices, respectively, and $e_{ijk}$ and $\epsilon_{\alpha \beta}$ are the relevant fully antisymmetric tensors with $e_{123} = -\epsilon_{12} = 1$. In this expression, $\hat{D}$ denotes the $(3, 1, + \frac{1}{2})$ components of the scalar 10$_H$, $T$ the $(3, 1, - \frac{1}{2})$ components of 5$_H$, $Q_{L_i}$ the quark doublet $(3, 2, + \frac{1}{2}) \in 10_M$, $d_{L_j}$ the down-type quark singlet $(3, 1, + \frac{1}{2}) \in 10_M$, and $\nu^\nu_{L_i}$ the $(1, 1, 0)$ components of 10$_M$. The charged vector bosons $X_{\mu}$ associated with the breaking of $SU(5) \otimes U(1)_X$ have SM quantum numbers $(3, \bar{2}, + \frac{1}{6})$.

Following the breakdown of the $SU(5) \otimes U(1)_X$ symmetry due to the non-zero VEV $V_{G}$, the scalar states $D$ and $T$ mix to form the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ eigenstates $\Delta_{1,2,3}$, as described in Appendix B.

Let us note that in deriving the central formula Eq. (10), especially the overall factor of 3 therein, the color and isospin factors in Eq. (A1) play a crucial role. It is also worth noting that the exact cancellation of the UV divergences discussed in Sec. III, which relies on the extra factor of 2 in Eq. (11), emerges from the difference of the overall numerical factors in the last two terms in Eq. (A1).

After including an additional $5'_{H}$ to arrive at the minimal realistic model discussed in Section IV B, the interaction Lagrangian remains rather similar. The additional of Yukawa couplings involving $5'_{H}$ leads to the set of interaction terms (with color indices suppressed for simplicity)

$$L_{\text{int}}^\text{THSM} = L_{\text{int}} - \left[8 Y_{I}^{HJ} \epsilon_{\alpha \beta} (Q_{L_i})_{ij} T^\nu C^{-1} Q_{L_j}^\mu T^k + \text{H.c.} \right]$$

(A2)

where $T^\nu$ denotes the additional $(3, 1, - \frac{1}{2})$ multiplet contained in $5'_{H}$, which mixes with the states $D$ and $T$ to yield a set of $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ eigenstates $\Delta_{1,2,3}$.

For the sake of completeness and matching to the SM Yukawa couplings we also present the terms involving the doublet Higgs interactions here:

$$- L_{\text{int}} \supset \left[8 Y_{I}^{HJ} \epsilon_{\beta \rho} H^\beta d_{L_i} C^{-1} Q'_{L_j} + Y_{I}^{HJ} H^\beta d_{L_i} C^{-1} Q'_{L_j} + Y_{I}^{HJ} H^\beta d_{L_i} C^{-1} Q'_{L_j} + \text{H.c.} \right]$$

(A3)

where the SM Higgs doublet $H$ consists of the components of $5_{H}$ transforming under the SM gauge group as $(1, 2, - \frac{1}{2})$, $u^\nu_{L_i}$ and $\epsilon^\nu_{L_j}$ are the components of $\tilde{5}_M$ transforming as $(3, 1, - \frac{1}{2})$ and $(1, 2, - \frac{1}{2})$ respectively, and $\epsilon^\nu_{L_i}$ denotes the single component of $1_M$, transforming as $(1, 1, +1)$. 

095015-9
APPENDIX B: TRIPLET SCALAR SPECTRUM AND MIXING

1. Model with a single 5_H representation

The tree-level scalar potential in the model with a single 5_H may be written

\[
V = \frac{1}{2} m_{10}^2 \text{Tr}(10_H^i 10_H^j) + m_5^2 5_H^i 5_H^j + \frac{1}{8} \left( \mu e_{ijklm} 10_H^i 10_H^j 5_H^k 5_H^l + \text{H.c.} \right) + \frac{1}{4} \lambda_1 \text{Tr}(10_H^i 10_H^j)^2 + \frac{1}{4} \lambda_2 \text{Tr}(10_H^i 10_H^j 10_H^k 10_H^l) + \lambda_3 (5_H^i 5_H^j)^2 + \frac{1}{2} \lambda_4 \text{Tr}(10_H^i 10_H^j) (5_H^i 5_H^j)
\]

The scalar basis is chosen such that the spontaneous breaking of SU(5) \( \otimes U(1)_X \) and the subsequent electroweak symmetry breaking takes place via the nonzero VEVs

\[
\langle 10_H \rangle^{45} = -\langle 10_H \rangle^{54} = V_G, \quad \langle 5_H \rangle^4 = v.
\]

Requiring that this corresponds to a stationary point of the scalar potential yields the conditions

\[
V_G [m_{10}^2 + V_G^2 (2 \lambda_1 + \lambda_2) + v^2 (2 \lambda_4 + \lambda_5)] = 0, \quad (B3)
\]

\[
v [m_5^2 + 2 \lambda_3 v^2 + V_G^2 (2 \lambda_4 + \lambda_5)] = 0, \quad (B4)
\]

which permit the parameters \( m_5^2, m_{10}^2 \) to be eliminated in favor of the VEVs.

After the breakdown of SU(5) \( \otimes U(1)_X \) to SU(3)_C \( \otimes \) SU(2)_L \( \otimes U(1)_Y \), the charged vector bosons \( X_\mu \) associated with the broken generators acquire masses \( m_X \) given by Eq. (3). The scalar states \( T \) and \( D \) of relevance to the generation of the RH neutrino masses mix, with the mass matrix (in the basis \( (D^\dagger, T) \))

\[
M_\Delta = \begin{pmatrix}
-\lambda_2 V_G^2 & \mu V_G \\
\mu^* V_G & m_5^2 + \lambda_4 V_G^2
\end{pmatrix}, \quad (B5)
\]

where Eq. (B3) with \( v = 0 \) has been used to eliminate \( m_{10}^2 \). This is diagonalized by a unitary matrix \( U_\Delta \) according to

\[
U_\Delta M_\Delta U_\Delta^\dagger = \begin{pmatrix}
m_{\Delta_1}^2 & 0 \\
0 & m_{\Delta_2}^2
\end{pmatrix},
\]

with

\[
m_{\Delta_1,2}^2 = \frac{1}{2} \left[ m_5^2 + (\lambda_4 - \lambda_2) V_G^2 \right] \pm \sqrt{\left[ m_5^2 + (\lambda_2 + \lambda_4) V_G^2 \right]^2 + 4 |\mu|^2 V_G^2}. \quad (B6)
\]

which, in the electroweak vacuum, simplifies into

\[
m_{\Delta_1,2}^2 = \frac{V_G^2}{2} \left\{ -(\lambda_2 + \lambda_5) \mp \sqrt{(\lambda_2 - \lambda_5)^2 + \frac{4 |\mu|^2}{V_G^2}} \right\}. \quad (B7)
\]

The elements of the mixing matrix \( U_\Delta \) read

\[
(U_\Delta)_{11} = \frac{\mu V_G}{\sqrt{\mu^2 V_G^2 + (m_{\Delta_1} + \lambda_2 V_G^2)^2}}, \quad (U_\Delta)_{12} = \frac{m_{\Delta_1}^2 + \lambda_2 V_G^2}{\sqrt{\mu^2 V_G^2 + (m_{\Delta_1} + \lambda_2 V_G^2)^2}}, \quad (U_\Delta)_{21} = \frac{\mu^* V_G}{\sqrt{\mu^2 V_G^2 + (m_{\Delta_1} + \lambda_2 V_G^2)^2}}, \quad (U_\Delta)_{22} = \frac{m_{\Delta_2}^2 + \lambda_2 V_G^2}{\sqrt{\mu^2 V_G^2 + (m_{\Delta_2} + \lambda_2 V_G^2)^2}}. \quad (B8)
\]

2. Model with two 5_H representations

In the minimal realistic model with two 5_H representations, we take the tree-level scalar potential to be given by
The field basis is again chosen such that the fields $10H$ and $5H$ acquire nonzero VEVs given by Eq. (B2), while

$$\langle S'_H \rangle^4 = v'.$$ (B10)

The corresponding conditions that must hold for this to be a stationary point of the potential are

$$f_i = 0, i = 1, 2, 3,$$ (B11)

where

$$f_1 = v_1 m_1^2 + v_2 m_2^2 + 3 v_2^2 v_1 \eta_1 + v_3^2 \eta_3 + v_3^2 v_2 \eta_2 + v_2 V_G^2 (\lambda_7 + \lambda_8) + 2 v_1^2 \lambda_3 + v_1 V_G^2 (\lambda_4 + \lambda_5) + v_1^2 (\lambda_6 + \lambda_6 + 2 \eta_2),$$ (B12)

$$f_2 = v_2 m_2^2 + v_1 m_1^2 + 3 v_1 \eta_1 + v_3^2 v_2 \eta_3 + v_1^2 (\lambda_7 + \lambda_8) + 2 v_1^2 \lambda_3 + v_2 V_G^2 (\lambda_4 + \lambda_5) + v_2^2 (\lambda_6 + \lambda_6 + 2 \eta_2),$$ (B13)

$$f_3 = V_G m_1^2 + V^3 (\lambda_4 + \lambda_2) + v_1^2 V_G (\lambda_4 + \lambda_3) + v_2^2 V_G (\lambda_4 + \lambda_3) + 2 v_1 v_2 V_G (\lambda_4 + \lambda_5).$$ (B14)

In deriving the above, and in all expressions below, we restrict our attention to the case where all couplings are real.

In the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ symmetric phase, i.e., for $V_G \neq 0, v = v' = 0$, the set of scalar color triplets that mix is extended to include the color triplet $T'$ associated with $S'_H$. The $3 \times 3$ mass matrix, in the basis $(\bar{D}', T, T')$, reads

$$M^2_\Delta = \begin{pmatrix}
-\lambda_2 V_G^2 & \mu V_G & \mu' V_G \\
\mu V_G & m_1^2 + \lambda_4 V_G^2 & m_1^2 + \lambda_7 V_G^2 \\
\mu' V_G & m_1^2 + \lambda_7 V_G^2 & m_2^2 + \lambda_4 V_G^2 \\
\end{pmatrix},$$ (B15)

where Eq. (B14) with $v = v' = 0$ has been used to eliminate the dependence on $m_0^2$. The resulting mass eigenstates $(\Delta_1, \Delta_2, \Delta_3)$ are obtained through the rotation

$$\begin{pmatrix}
\Delta_1 \\
\Delta_2 \\
\Delta_3
\end{pmatrix} = U_\Delta \begin{pmatrix}
\bar{D}' \\
T \\
T'
\end{pmatrix},$$ (B16)

where the unitary matrix $U_\Delta$ diagonalizes $M^2_\Delta$ according to

$$U_\Delta M^2_\Delta U^\dagger_\Delta = \text{diag}(m_{\Delta_1}^2, m_{\Delta_2}^2, m_{\Delta_3}^2).$$ (B17)

APPENDIX C: RADIATIVE FERMION MASS GENERATION

In general, the physical mass of a single spin-1/2 fermion is obtained as the value of $m$ for which

$$\langle k + m \rangle \Gamma^{(2)}(k) = 0 \quad \forall \; k \; \text{such that} \; k^2 = m^2,$$ (C1)

where $\Gamma^{(2)}(k)$ is the renormalized two-point 1PI Green’s function,

$$\Gamma^{(2)}(k) = Z(k) k^2 \Sigma(0).$$ (C2)

In this expression, $Z(k)$ corresponds to the wave function renormalization and $\Sigma(0)$ is the zero incoming momentum contribution to the appropriate sum of Feynman diagrams. Taken together, Eq. (C1) and Eq. (C2) imply that

$$m Z(m^2) = \Sigma(0),$$ (C3)

which generally amounts to a transcendental equation to be solved for the physical mass $m$. An expression for $m$ may be obtained perturbatively by writing $Z(m^2) = 1 + \Delta Z(m^2)$, $\Sigma(0) = m_0 + \Delta m_0$, where the first and second term in each expression correspond to the tree-level and loop corrections to each quantity, respectively. One finds the result

$$m = m_0 + \Delta m_0 + \Delta m_0 \Delta Z(m_0^2) + \ldots,$$ (C4)

where we show only the leading part of the higher-order contribution. Therefore, in the general case with $m_0 \neq 0$, a calculation of the leading higher-order contribution to the physical mass would require the evaluation of the loop corrections to both $\Sigma(0)$ and $Z(k^2)$.

However, for the case studied in this article in which the RH neutrinos are massless at tree-level, Eq. (C4) reads simply $m = \Delta m_0 = \Sigma(0)$ at leading order.

APPENDIX D: EVALUATION OF THE TWO LOOP FEYNMAN INTEGRALS

1. Veltman-Van der Bij brackets

Remarkably enough, there is an entire industry concerning the evaluation methods for the zero-external-momentum two-point 1PI graphs, see, e.g., Ref. [22] or Ref. [27] and references therein.

The principal object in these methods are the so-called Veltman-Van der Bij brackets. As the original paper uses an Euclidean metric and a different choice of dimensional regularization parameter $\epsilon$, we give here all of the relevant expressions in our particular convention, i.e., in Minkowski metric $g = \text{diag}(1, -1, -1, -1)$ and with the number of spacetime dimensions equal to $D = 4 - 2\epsilon$.  

095015-11
We introduce the brackets in the following way

\[
\{M_{11}, M_{12}, \ldots; M_{21}, \ldots; M_{31}, \ldots\} = \int \frac{d^4p}{(2\pi)^4} \int \frac{d^4q}{(2\pi)^4} \frac{1}{(p^2 - M_{11}^2)(p^2 - M_{12}^2)\ldots(q^2 - M_{21}^2)\ldots((p+q)^2 - M_{31}^2)\ldots},
\]

\[
(D1)
\]

\[
\{M_{11}, M_{12}, \ldots\} = \int \frac{d^4p}{(2\pi)^4} \frac{1}{(p^2 - M_{11}^2)(p^2 - M_{12}^2)\ldots},
\]

\[
(D2)
\]

\[
\{M_{11}, \ldots; M_{21}, \ldots; M_{31}, \ldots\}\{A(p, q)\} = \int \frac{d^4p}{(2\pi)^4} \int \frac{d^4q}{(2\pi)^4} \frac{1}{(p^2 - M_{11}^2)\ldots(q^2 - M_{21}^2)\ldots((p+q)^2 - M_{31}^2)\ldots} A(p, q).
\]

\[
(D3)
\]

With the last expression we have introduced a shorthand notation that simplifies the form of this Appendix.

Note that the brackets are invariant under the exchange of positions of the individual groups of components, which can be obtained by the change of variables \((p \leftrightarrow q)\) and \((p + q \rightarrow p, \ -q \rightarrow q)\).

By a partial cancellation of fractions we can derive various reduction formulae of the type

\[
\{M_A, M_B, m_b; M_C\}[p^2] = \{M_A; M_B, m_b; M_C\} + m_b^2\{M_A, M_B, m_b; M_C\}.
\]

\[
(D4)
\]

A similar trick using \(p^2 - M_B^2 - (p^2 - M_A^2) = M_A^2 - M_B^2\) can be used for a simplification of brackets of the type \(\{M_A; M_B; \alpha; \beta\}\)

\[
\{M_A, M_B; \alpha; \beta\} = \frac{1}{M_A^2 - M_B^2}\{M_A; \alpha; \beta\} - \{M_B; \alpha; \beta\}.
\]

\[
(D6)
\]

It is also possible to show that

\[
\{M_A; M_B; M_C\}[(p + q)^2]
\]

\[
= \{M_A\}\{M_B\} + M_C^2\{M_A; M_B; M_C\}.
\]

\[
(D7)
\]

Using all of these methods we can express the relevant two-loop integrals in terms of simple brackets \(\{M_A; M_B; M_C\}\). It is of use to rewrite them further into double brackets

\[
\{2M_A; M_B; M_C\} = \{M_A, M_A; M_B; M_C\},
\]

\[
(D8)
\]

which are dimensionless (cf. Ref. [22]). The operation transcribing simple brackets into double brackets is ’t Hooft’s \(p\)-operation [32]. In our notation it reads

\[
\{M_A; M_B; M_C\} = \frac{1}{D - 3}(M_A^2\{2M_A; M_B; M_C\}
\]

\[
+ M_B^2\{2M_B; M_C; M_A\}
\]

\[
+ M_C^2\{2M_C; M_A; M_B\}).
\]

\[
(D9)
\]

2. Topology 1

Topology 1 of Fig. 1 leads to the kinematic form (i.e., neglecting the specific form of the vertices) of the integral given in Eq. (12). By using \(D\)-dimensional gamma matrix gymnastics, it can be simplified into

\[
\Sigma_1^f(0) = -\{m_X, 0; m_X, 0; m_{\Delta}\} \left\{ (D - 4)q^2 + 4p \cdot q \right\} - \frac{p^2 + q^2}{m_X^2 - m_X^2} p \cdot q \cdot q^2 + \frac{p \cdot q}{m_X} p \cdot q.
\]

\[
(D10)
\]

The slashed product can be rewritten into \(p \cdot q = p \cdot q - iP \cdot \sigma_{\mu\nu} q^\nu\). After performing the \(p\) integration the second term would have to be of the form \(i\sigma_{\mu\nu} q^\nu\) and, due to the antisymmetry of \(\sigma_{\mu\nu}\), such a term will not contribute. After the operations given above, we obtain

\[
\Sigma_1^f(0) = \frac{m_{\Delta}^2}{2m_X^2}\{0; 0; m_{\Delta}\} - (D - 1) \left\{ \frac{1}{2m_X^2} A_0(m_{\Delta}^2)^2 
\]

\[
+ \frac{m_{\Delta}^2}{2}\{m_X, 0; m_X, 0; m_{\Delta}\} - \{m_X, 0; m_X; m_{\Delta}\} \right\}.
\]

\[
(D11)
\]

This may be rewritten in terms of the simple brackets using relations similar to those in Eq. (D6).
3. Topology 2

Neglecting the specific form of the vertices, Topology 2 of Fig. 1 leads to the second integral in Eq. (12). It can be simplified into (again making use of the antisymmetry of $\sigma_{uv}$)

$$\Sigma_2^p(0) = -\{m_X, 0; m_\Delta, 0; m_X\} \left[(2 - D) p \cdot q - \frac{2p^2 q^2}{m_X^2} - \frac{2p^2 + q^2}{m_X^2} p \cdot q + \frac{p^4 q^2}{m_X^2} + \frac{p^2 (q^2 + p^2)}{m_X^2} p \cdot q + \frac{p^2}{m_X^2} (p \cdot q)^2 \right].$$  \hspace{1cm} (D12)

The result after simplification reads

$$\Sigma_2^p(0) = \frac{2 - D}{2} \{0; m_\Delta, 0; m_X\} + \frac{3 - D}{2} \{m_X, 0; m_\Delta; m_X\} + \frac{m_X^2}{4m_X^2} \{2\{m_X; 0; m_\Delta\} - \{m_X; m_X; m_\Delta\}\} + \frac{D - 2}{2m_\Delta^2 m_X^2} A_0(m_X^2) A_0(m_\Delta^2) - \frac{1}{4m_X^2} A_0(m_X^2)^2.$$

\hspace{1cm} (D13)

4. Integrals

For the reader’s convenience, we list here the results of the integrals appearing in the expressions in our convention. As integrals $A_0(M_A^2)$ appear in the results in the second power, we need to evaluate also the term linear in $\epsilon$. This gives

$$A_0(M_A^2) = Q^{4-D} \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 - M_A^2} = -i \frac{M_A^2}{(4\pi)^2} \left[-\frac{1}{\epsilon} + L_A - \frac{\epsilon}{2} \left(L_A^2 + 1 + \frac{\pi^2}{6}\right)\right] + O(\epsilon^2),$$  \hspace{1cm} (D14)

where

$$L_A = \log \frac{M_A^2}{Q^2} - \log 4\pi + \gamma - 1,$$

with $Q$ being the renormalization scale and $\gamma$ the Euler-Mascheroni constant.

As was already stated, all of the simple brackets can be obtained from the double brackets using Eq. (D9). Therefore, we give here the result only for them. It reads

$$\{2M; M_a; M_b\} = \frac{1}{(4\pi)^4} (S(M) - f(a, b)) + O(\epsilon),$$  \hspace{1cm} (D16)

where

$$S(M) = -\frac{1}{2e^2} + \frac{1}{e} \left(L + \frac{1}{2}\right) - \left(L^2 + L + \frac{1}{2} + \frac{\pi^2}{12}\right),$$  \hspace{1cm} (D17)

$$a = \frac{M_a^2}{M^2}, \hspace{1cm} b = \frac{M_b^2}{M^2},$$  \hspace{1cm} (D18)

and the function $f(a, b)$ is given by

$$f(a, b) = -\frac{1}{2} \log a \log b + \frac{1 - a - b}{2\sqrt{q}} \left[\text{Li}_2\left(-\frac{x_2}{y_1}\right) + \text{Li}_2\left(-\frac{y_2}{x_1}\right) - \text{Li}_2\left(-\frac{x_1}{y_2}\right) - \text{Li}_2\left(-\frac{y_1}{x_2}\right)\right]$$

$$+ \text{Li}_2\left(-\frac{b - a}{x_2}\right) + \text{Li}_2\left(-\frac{a - b}{y_2}\right) - \text{Li}_2\left(-\frac{b - a}{x_1}\right) - \text{Li}_2\left(-\frac{a - b}{y_1}\right),$$  \hspace{1cm} (D19)

$$f(b, b) = -\frac{(b - 1)}{\sqrt{1 - 4b}} \left[2\text{Li}_2\left(-\frac{\sqrt{1 - 4b - 1}}{\sqrt{1 - 4b + 1}}\right) + \frac{y_2^2}{2} + \frac{1}{2} \log^2 \left(-\frac{\sqrt{1 - 4b - 1}}{\sqrt{1 - 4b + 1}}\right)\right] - \frac{1}{2} \log^2(b).$$  \hspace{1cm} (D20)

In Eq. (D19) and Eq. (D20) the quantities $q, x_{1,2}$, and $y_{1,2}$ are defined by

$$q \equiv 1 - 2(a + b) + (a - b)^2,$$  \hspace{1cm} (D21)

$$x_{1,2} \equiv \frac{1}{2}(1 + b - a \pm \sqrt{q}),$$  \hspace{1cm} (D22)

$$y_{1,2} \equiv \frac{1}{2}(1 + a - b \pm \sqrt{q}).$$  \hspace{1cm} (D23)

In addition to Eq. (D20) giving the value of $f(a, b)$ when $a = b$, it is helpful to note the other special cases

$$f(0, 0) = \frac{\pi^2}{6},$$  \hspace{1cm} (D24)

$$f(0, b) = \text{Li}_2(1 - b),$$  \hspace{1cm} (D25)

$$f(0, b^{-1}) = -\frac{1}{2} \log^2 b - f(0, b).$$  \hspace{1cm} (D26)
5. The kinematic structure of the self-energies

Rewriting Eq. (D11) and Eq. (D13) yields the expressions in terms of double brackets,

\[
\Sigma^p(0) = -\frac{1}{D-3} \frac{m^4}{2m_X^2} \{2m_\Delta;0;0\} - \frac{D-1}{2m_X^2} A_0(m_\chi^2) + \frac{D-1}{D-3} (2\{2m_X;m_\Delta\} - \{2m_X;0;m_\Delta\})
\]

\[
+ \frac{D-1}{2m_X^2} \{2m_\Delta;m_X;0\} - (2\{2m_\Delta;m_X\} - \{2m_\Delta;0;0\})
\]

\[
+ \frac{D-1}{D-3} \frac{m^2}{m_X^2} \{2m_\Delta;0;m_\Delta\} - \{2m_X;m_\Delta\} + \{2m_X;m_\Delta\} - \{2m_X;0;m_\Delta\}
\] (D27)

\[
\Sigma^p(0) = \frac{D-2}{2m_\Delta^2 m_X^2} A_0(m_\chi^2) A_0(m_\chi^2) - \frac{1}{4m_X} A_0(m_\chi^2) + \frac{D-2}{D-3} \frac{m^2}{2m_\Delta^2} \{2m_\Delta;0;0\} - \{2m_X;0;m_\Delta\}
\]

\[
+ \frac{m^2}{2m_X^2} \{2m_\Delta;m_X;0\} - (2\{2m_\Delta;m_X\} - \{2m_\Delta;0;0\})
\]

\[
+ \frac{D-2}{2(D-3)} \{2m_\Delta;m_X;0\} - \frac{1}{2} \{2m_X;m_\Delta;0\} - \{2m_X;0;m_\Delta\}
\]

\[
+ \frac{1}{D-3} \frac{m^2}{4m_X^2} \{2m_\Delta;0;m_\Delta\} - \{2m_X;m_\Delta\}.
\] (D28)

Using the explicit expression for the double brackets, Eq. (D16), \(\Sigma^p(0)\) and \(\Sigma^p(0)\) are then finally found to be given by (where \(s_i = \frac{m^2}{m_X^2}\) as above)

\[
(4\pi)^4 \Sigma(0)^p = -\frac{3}{2\epsilon} + 3L_X - 2 + \frac{s^2}{2} \left[ \frac{1}{2\epsilon^2} - \frac{1}{\epsilon} \left( L_\Delta - \frac{1}{2} \right) + \left( L_\Delta^2 - L_\Delta + \frac{3}{2} \right) \right]
\]

\[
+ 3(f(0,s_i) - 2f(1,s_i)) + \frac{3}{2} s^2[f(s_i^{-1},s_i^{-1}) - 2f(0,s_i)]
\]

\[
+ 3s_i[f(1,s_i) - f(0,s) - f(s_i^{-1},s_i^{-1}) + f(0,s_i^{-1})] + 2s^2 f(0,0).
\] (D29)

\[
(4\pi)^4 \Sigma(0)^p = \frac{3}{4\epsilon} - \frac{1}{2} \left[ L_X + 2L_\Delta - (L_\Delta - L_X)^2 - 1 \right] - \frac{s^2}{4} \left[ \frac{1}{2\epsilon^2} - \frac{1}{\epsilon} \left( L_\Delta - \frac{1}{2} \right) + \left( L_\Delta^2 - L_\Delta + \frac{3}{2} \right) \right]
\]

\[
- s_i^{-1}[f(0,0) - f(0,s_i)] + f(0,s_i^{-1}) + f(1,s_i) - \frac{1}{2} f(0,s_i)
\]

\[
- \frac{s_i}{2} [f(s_i^{-1},0) - f(s_i^{-1},s_i^{-1}) + f(0,s_i) - f(1,s_i)] - \frac{s^2}{4} [2f(0,s_i^{-1}) - f(s_i^{-1},s_i^{-1})].
\] (D30)

Note that the individual diagrams are UV divergent, with the divergent terms given by Eqs. (14) and (15). However, as noted in Sec. III, their combination appearing in Eq. (11) yielding the total contribution to the RH neutrino mass matrix is finite and compact,

\[
I_3(s) = 1 + 2 \log s + s(1 - 2s) \log^2 s + 2(s^{-1} - 1)[f(0,0)(1 + s + s^2) + 2sf(1,s)]
\]

\[
+ f(0,s)(1 + s)(1 + 2s) + s^2 f(s^{-1},s^{-1})]
\] (D31)
Bibliography


203


