

Quark-lepton color-flavor unification

Antonio Delgado^{*} and Seth Koren[†]

Department of Physics and Astronomy, *University of Notre Dame*,
Notre Dame, Indiana 46556, USA



(Received 9 June 2026; accepted 14 September 2026; published 14 October 2026)

We present an $SU(12) \times SU(2)_L \times U(1)_R$ model unifying $SU(9)$ quark color-flavor with $SU(3)$ lepton flavor as a flavorful alternative to conventional theories of unification. We begin in the ultraviolet with a single Yukawa shared by the unified up-type quarks and neutrinos and no further new fermions. We show that gauged quark color-flavor and lepton-flavor instantons dynamically generate the bottom and tau Yukawas, which implements a massless quark solution to the strong CP problem and sets up a flavored type-I seesaw mechanism. Only two new scalar irreps are needed for the symmetry-breaking steps, which include quark color-flavor deconstruction and then infrared reunification, and the Standard Model (SM) gauge group in this theory emerges as $G_{\text{SM}} = \frac{SU(3)_C \times SU(2)_L \times U(1)_Y \times \mathbb{Z}_{18}^X}{\mathbb{Z}_3 \times \Gamma \times \mathbb{Z}_3}$, where $\Gamma \in \{1, \mathbb{Z}_2\}$ is the electroweak global structure and there is a discrete gauge symmetry $X = B - 3(L_i + L_j - L_k)$ which brings additional $\mathbb{Z}_3$ global structure to the SM. This gauge symmetry acts as a flavorful upgrade of the $\mathbb{Z}_{18}^{B+L}$ anomaly-free global symmetry of the SM and stabilizes the proton absolutely. Noninvertible chiral symmetry breaking is crucial to our model, and we discuss the rich spectrum of emergent generalized symmetries and topological defects appearing at various stages. In the infrared, the novel shared quotient between continuous and discrete groups links the one-form and two-form global symmetries of the Standard Model.

DOI: [10.1103/9h7r-2zpr](https://doi.org/10.1103/9h7r-2zpr)

I. UNIFICATION

A major lesson of the past couple hundred years of scientific inquiry has been the rampant success of reductionism: As we have probed shorter and shorter distances, we have found simpler and simpler descriptions of the world and its remarkable emergent phenomena which require fewer degrees of freedom and fewer free parameters. With such large Bayesian update factors for the proposition of simpler small-distance theories, there is great *circumstantial* evidence toward *further* unification and an even simpler theory at distances far below the Standard Model phase.

However, since the proposal to unify the gauge forces of Georgi-Glashow in 1974 [1], particle physicists have sadly seen no discoveries of the concrete beyond the Standard Model (BSM) microphysics involved in

unification models.¹ A general concern for theories that unify the SM gauge symmetries into a simple group is the lack of observed proton decay, despite exquisitely sensitive searches occurring over decades [5]. Of course it is possible the lesson is merely that unification happens at a higher scale than expected, but whenever there is conventional wisdom which is under stress from empirical observations, it is interesting to think contrarily and in particular to consider maximalist interpretations of the data.

Maybe unification, despite the tremendous appeal, is just not realized in our Universe: It turns out that if unification is *not* correct, we could definitively learn this at low energies by discovering fractionally charged particles (FCPs) [6,7], whose very existence is incompatible with some or all minimal grand unified theories (GUTs), depending on the charge. The possibility of learning such stark, far-ultra-violet information from simple infrared measurements is enabled here because the existence of fractionally charged particles probes a *topological* aspect of the field theory,

^{*}Contact author: adelgad2@nd.edu

[†]Contact author: skoren@nd.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

¹Nor have we seen direct evidence of the proposals to unify gravity and SM forces through extra dimensions going back to Nördstrom in 1914 [2], followed later by Kaluza in 1921 [3] and Klein in 1926 [4].

so does not depend on energy scale at all. Further contemplation on how to identify and make use of such *rigid* quantum field theory data contained in our theories of particle physics is warranted, and indeed we will make further use of topological data below. Toward ruling out GUTs, more work by both phenomenologists and experimentalists is needed to develop a robust search program for these signatures at colliders.²

On the other hand, discarding unification entirely based on the nonobservation of proton decay is perhaps too rash, and we ought also examine our theoretical biases regarding the link between unification and proton decay. Rather than giving up the prospect of a simpler theory at smaller distances, maybe there is a different approach to unification that leads to different conclusions about proton stability. After all, the proton of the Standard Model is *stable* due to a zero-form discrete global symmetry that results nontrivially from the fact we have multiple generations of fermions [18,19]. Here we follow to the maximum this hint that the multiplicity of generations is deeply tied to the fundamental behavior of the Standard Model particles and find that nontrivially incorporating gauged flavor into our unification paradigm can absolutely stabilize the proton even with quarks and leptons unified in the ultraviolet.

It is well known that models of unification where the UV group is nonsimple need not introduce proton decay, with Pati-Salam [20] and trinification [21] the paradigmatic and minimal examples, although pursuing this requires care regarding which scalar representations are included and can complicate the model building. However these models are often seen as unsatisfying since gauge unification is “not complete.” Is there a version of these that offers comparable virtues to well-loved $SU(5)$ [1] and $SO(10)$ [22]?

There are various ways in which theories of unification may simplify the ultraviolet description. Historically, the near-unification of SM gauge couplings has been important theoretical evidence [23], though more-precise measurements later revealed that the couplings do not unify with the SM matter [24–26]. A conceptually separate aspect of unified theories is matter unification and particularly the unification of the fermions of the SM. In the familiar “vertical” GUTs which focus on gauge coupling unification, the fermions unify into no less than three multiplets, as the generation structure merely goes along for the ride. Is there a way to achieve complete unification of the SM fermions?

Attempts to add flavor to this picture include Georgi’s program to enlarge $SU(5)$ to $SU(N > 5)$ and find the generations from different representations to explain the

breaking of flavor symmetry [27]. Berezhiani instead proposed vertical $\times$ horizontal theories like $SO(10) \times \mathcal{H}$ [28,29] which does unify the SM fermions into one multiplet. But either of these theories requires many new fermions, including perhaps new chiral states, and the proton anyway remains unstable. Furthermore, gauging flavor in this way requires having a shared flavor symmetry between the quarks and the leptons, as you might think is necessary if quarks and leptons are unified.

The “flocinaucinihilipilification” paper [30] of Allanach, Gripaio, and Tooby-Smith, betraying its name, uncovered a small number of ways in which gauged flavor can be intertwined nontrivially with the gauge structure of the SM. Davighi and Tooby-Smith constructed a fantastic theory of $SU(4)_C \times Sp(6)_L \times Sp(6)_R$ electroweak flavor unification [31] (which had originally been noticed in 1984 by Kuo and Nakagawa [32,33] but without much exploration of the theory). This theory forbids perturbative proton decay but the proton inevitably decays through $Sp(6)$ instantons [34]. The other sort of nontrivially intertwined gauge-flavor unification pointed out in that work is $SU(12) \times SU(2)_L \times SU(2)_R$ quark-lepton color-flavor unification. Note importantly that there exist versions of each of these theories which include a semidirect product with a $\mathbb{Z}_2$ factor that charge conjugates the color group while switching L and R .³ In such a theory *all* the SM fermions have truly been unified into a single irreducible representation, with no extras aside from the right-handed neutrinos.

It is color-flavor unification that will be the topic of this work. Remarkably, despite manifestly unifying the leptons and the quarks, this will allow separate gauged flavor symmetries for quarks and leptons at intermediate scales. Not only does this provide a broad motivation for why quark- and lepton-flavor physics is so different, but it also allows remarkable nonperturbative effects from the gauged flavor instantons [36–39] which work specifically because the quarks and leptons are separated—and also specifically because we have $N_g = 3$ generations, making this one of few active inroads toward the generation puzzle.

We have previously explored these sectors individually, having understood their effects from a bottom-up strategy using noninvertible symmetries which pointed us to interesting physics in these ultraviolet theories of unification. In those works, the neutrino Yukawas are generated from the charged lepton Yukawas and an instanton of $SU(3)_\ell$ lepton-flavor unification, and the down-type Yukawas are generated from the up-type Yukawas and an instanton of $SU(9)$ quark color-flavor unification. Here we further unify these two separate uses of noninvertible symmetry in

²See also [8,9] on other aspects of FCPs. We can learn similar information by discovering an axion and making measurements sensitive to its anomaly coefficients [10–17] though to our knowledge the full map of what set of measurements would be required to rule out which global structures has not been described.

³In the context of Pati-Salam this is sometimes referred to as “D parity,” and the correct discrete structure (semidirect product, *not* direct) has only seen little discussion, for example Appendix B of the recent [35].

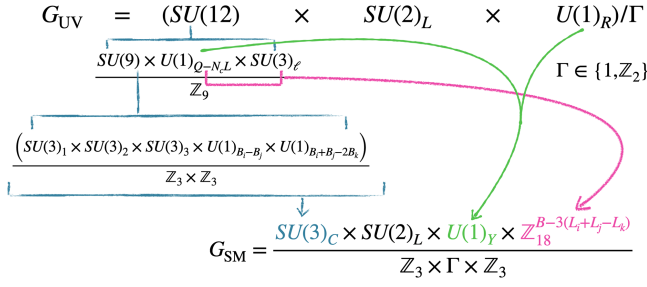


FIG. 1. The breaking chain from our ultraviolet theory down to the Standard Model. Some of these steps may happen simultaneously; or contrarily there may be additional intermediate steps.

the quark and lepton sectors. However, there is an obstacle to directly cutting and pasting the previous uses together: The heaviest quark and lepton Yukawas are reversed with respect to each other. That is, the “up-type” lepton is the *neutrino*, which has the smaller of the two lepton masses, whereas the largest quark Yukawa is the top, so the two mechanisms do not immediately unify. Here we will propose one model of notable dynamics coming from quark-lepton color-flavor unification with the expectation that this is one of many interesting phenomenologies that could be borne out as this ultraviolet theory flows to the infrared Standard Model. The model we will pursue in this work will preserve the solution to the strong CP problem while still using noninvertible symmetry breaking to interesting effect in the lepton sector in conjunction with a flavored type-I seesaw that ultimately leads to Majorana neutrino masses.

Our main interest will be in starting with solely one Yukawa and generating the various masses for the third-generation fermions through interplay between tree-level effects, one-loop effects, and nonperturbative effects. The penalty we pay with this approach is that quark-lepton unification will take place at very high energy scales, for reasons unrelated to the stability of the proton. After starting with the same Yukawa for the top quark and the neutrino, we will generate the bottom quark and tau Yukawas from gauged flavor instantons when those gauge symmetries are broken. But the breaking of this lepton-flavor symmetry coincides with the generation of Majorana neutrino masses, which for a successful seesaw mechanism must occur at the usual high scale, necessitating very high energy dynamics. See Fig. 1 for a schematic breakdown of the various symmetry-breaking steps from the ultraviolet group down to the SM phase.

As we flow down in energies and break the flavor symmetries, we will have enough freedom to generate the full structure of the masses and the Cabibbo-Kobayashi-Maskawa (CKM) and Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrices, but producing a beautiful model of flavor will be left for future work. As is standard in GUT model building, in this first pass we will not attempt to write down a full scalar potential but rather assume the appropriate scalars condense at the energies we want them to.

Nevertheless we will *not* land on the familiar SM gauge group in the far IR—our Standard Model will have an additional discrete gauge theory. The full group will be

$$G_{SM} = \frac{SU(3)_C \times SU(2)_L \times U(1)_Y \times \mathbb{Z}_{18}^X}{Z_3 \times \Gamma \times Z_3}, \quad (1)$$

where $X = B - 3(L_i + L_j - L_k)$ will nonperturbatively protect the proton from decay and $\Gamma \in \{1, \mathbb{Z}_2\}$ is the electroweak global structure.⁴ This discrete symmetry factor was essentially identified by Babu, Gogoladze, and Wang [40] long ago and shown to appear from this theory in [41]. The additional global structure appearing between the continuous SM groups and a gauged discrete symmetry is novel here. This version of the Standard Model has the same continuous degrees of freedom as normal, and local-in-field-space effects are all the same as normal, but the spectrum of topological degrees of freedom differs and their coupling to the SM has important phenomenological implications.

In Sec. II we will review the logic of noninvertible naturalness and the noninvertible symmetries appearing in this work. In Sec. III we will give the ultraviolet theory of quark-lepton color-flavor unification. In Sec. IV we will discuss the spontaneous breaking which separates the quarks and leptons and generates the noninvertible symmetry-breaking instanton effects. In Sec. V we will discuss the breaking of quark color-flavor unification, and in Sec. VI we will discuss the breaking of lepton flavor and the preservation of a discrete gauge symmetry. At each of these stages we will discuss both the perturbative phenomenology and nonperturbative aspects of the physics including instantons, topological defects, and emergent generalized symmetries. In Sec. VII we will match the ultraviolet and infrared gauge couplings, and finally in Sec. VIII we will discuss the higher-form symmetries of the Standard Model on which we land in the infrared. We conclude in Sec. IX.

II. NONINVERTIBLE NATURALNESS

The quark-lepton color-flavor unified theory is the setting for our work. In Sec. III we will introduce this theory and thereafter transition down in energy scales while

⁴The X quantum number is really best understood as $Q - 9(L_i + L_j - L_k) = 3[B - 3(L_i + L_j - L_k)]$, where Q is quark number under which the quarks have unit charge. We have used “ B ” above because it is the familiar convention and because two symmetries which differ only by normalization are the same symmetry, so it is not technically incorrect. But it is far less confusing to consider topological effects in the normalization where all particles have integer charge, so that a rotation by 2π is the identity and all anomaly coefficients and indices are integers, and in the rest of this work we will use unit normalizations. (That is, except for Sec. VII when we will switch to conventional normalization so that our results can be compared clearly to others.)

explaining various symmetry-breaking steps. However, the construction of this theory, and the bottom-up understanding of its core mechanism, has been enabled by our understanding of noninvertible symmetry breaking in flavored Z' theories, and so we will first here explain the relevant concepts and applications. Readers eager to see the top-down model can skip ahead to Sec. III on a first pass, returning to this section when needed for additional perspective on various symmetry-breaking effects.

The basic model-building benefit of analyzing noninvertible chiral symmetries is to systematically learn from the infrared about possible ultraviolet instanton effects that may generate certain couplings in the infrared theory. This guidance goes past 't Hooft naturalness [42] in not only providing assurance that the small size of a given coupling may be preserved upon RG flow, but also in pointing to a particular ultraviolet mechanism for generating that small coupling in the first place. This strategy has thus been termed “noninvertible naturalness.” Having this signposted direction for instanton effects is especially useful, as they provide one of few known ways to generate exponential, parametric hierarchies between parameters or scales.

In particular these types of noninvertible symmetries can occur when our gauge group admits more instantons on a general manifold than on $\mathbb{R}^4$ (really its one-point compactification S^4 , as is familiar from instanton studies). The crucial aspect of general manifolds is their ability to support magnetic fluxes on compact cycles, and in fact we can skip considering exotic manifolds entirely if we instead study our theory on $\mathbb{R}^4$ with 't Hooft lines inserted [43–46]. For simplicity we find the former criterion easier to state.

That is, say we are given an infrared H gauge theory with field strength $H^{\mu\nu}$. As normal, to understand the dynamical effects of instantons in breaking chiral symmetries in the infrared H theory, we study the possible values of

$$I_{\text{IR}}(S^4) = \int_{S^4} \frac{\text{Tr}(H\tilde{H})}{16\pi^2}, \quad (2)$$

the spectrum of instantons that appear around empty flat space. These instantons break chiral symmetries by generating 't Hooft vertices in our theory on $\mathbb{R}^4$. However, as theorists, we are free to study our theory on whatever sort of background we wish and in particular to consider the instantons which appear on a more general manifold

$$I_{\text{IR}}(\mathcal{M}^4) = \int_{\mathcal{M}^4} \frac{\text{Tr}(H\tilde{H})}{16\pi^2}. \quad (3)$$

To be perfectly clear, we are studying these configurations *not* because we inherently care about what our theory does on $\mathcal{M}_4$ but because such auxiliary calculations can give us as effective field theorists more information about our H gauge theory that can be useful to understand the theory on $\mathbb{R}^4$. In particular, one may often find that as sets,

$$I_{\text{IR}}(\mathcal{M}^4) \supset I_{\text{IR}}(S^4), \quad (4)$$

and there are more instantons allowed on a general manifold than dynamically break chiral symmetries in our IR theory on $\mathbb{R}^4$. The most familiar example is $U(1)$ gauge theory, where $I_{\text{IR}}(S^4) = 0$ and there are no Abelian instantons in flat space, but $I_{\text{IR}}(\mathcal{M}^4) = \mathbb{Z}$ meaning in general there are integer-valued instantons on a spacetime with a 2-cycle (for example, the torus T^4).⁵ Another pertinent example is the quotient group $SU(N)/\mathbb{Z}_N$, where $I_{\text{IR}}(S^4) = \mathbb{Z}$ being the familiar integer instantons of $SU(N)$, but $I_{\text{IR}}(\mathcal{M}^4) = \mathbb{Z}/N$ and in general “fractional” instantons are allowed due to the identification of the $\mathbb{Z}_N$ center of $SU(N)$ with the identity. See e.g. [6] for some review on the basics of thinking about gauge theories with quotient groups and e.g. [11,37,47,48] for insight into fractional instantons.

The technology of generalized symmetries tells us that in these scenarios the instantons which live only on exotic manifolds have the effect on the theory in $\mathbb{R}^4$ of converting some chiral symmetries into noninvertible symmetries [rather than simply breaking them, as the $I_{\text{IR}}(S^4)$ instantons do] [49,50]. Furthermore, the technical construction of noninvertible chiral symmetries teaches us something special about their explicit breaking. A noninvertible chiral symmetry relies on the IR theory having a magnetic one-form global symmetry and admitting magnetic monopole defects. As the magnetic one-form symmetry of the H gauge theory is essential, this implies that if we go to a UV G gauge theory which breaks the magnetic one-form symmetries of the H gauge theory [for example, by $\pi_2(G/H) \neq 1$ supplying the correct magnetic monopoles of the H theory as dynamical objects, not simply defects], then the noninvertible chiral symmetries of the H theory *must* be broken in the G theory.

Systematic consideration of these symmetries leads us to the following model-building strategy called noninvertible naturalness. When we go to the UV G theory that breaks the magnetic one-form symmetry we will find

$$I_{\text{UV}}(S^4) \supset I_{\text{IR}}(\mathcal{M}^4), \quad (5)$$

and we have upgraded the instantons which did not live on empty flat space in the H gauge theory into instantons which do in the G gauge theory. Then in the UV theory these now have the familiar dynamical effects of chiral symmetry breaking. While discovering in your infrared theory that a coupling is technically natural teaches you that *if* that coupling were small in the ultraviolet it would stay small into the infrared, discovering that a coupling is noninvertibly natural guides you to ultraviolet theories that *explain* why that coupling is small to begin with.

⁵For $U(1)$ since there is no trace over generators we actually mean $I_{\text{IR}} = \int F\tilde{F}/(16\pi^2)$.

At a basic level the lesson—that you should look for an ultraviolet theory which supplies the “missing” instantons—could have been guessed long ago. However it seems that historically it was not, as we have been able to use this strategy to discover useful instanton effects in minimal extensions of the SM. Furthermore, quite generally the point of understanding the symmetry statement is to learn the systematics, which teaches us concrete strategies for using this symmetry information. For technical details we refer to earlier papers e.g. [36,37,49]. More generally than just this concept, in recent years many works have appeared using various concepts from generalized global symmetries to learn new particle physics e.g. [51–69].

A. Noninvertible symmetry breaking in the quark sector

Models of Z' s, in which an additional $U(1)$ symmetry is gauged, offer a natural bottom-up strategy to search for profitable noninvertible symmetries leading to ultraviolet completions in which small couplings are generated by instantons. The most minimal Z' models are those which do not require additional matter, being simply a gauging of a nonanomalous (approximate) symmetry of the SM itself, a subgroup of the continuous global flavor symmetry $SU(3)^5 \times U(1)_{B-L}$ [where of the other four $U(1)$'s, one has become hypercharge and three have Adler-Bell-Jackiw anomalies under the three factors of the SM, though some discrete factors are preserved [18]].

In the lepton sector, the guidance from these noninvertible symmetries was used to find ultraviolet completions of lepton-flavor Z' models in which either Dirac or Majorana neutrino masses are generated by lepton-flavor instantons and so automatically exponentially suppressed [36]. In particular, in the Dirac case, the simplest model to implement noninvertible symmetry breaking is an $SU(3)_\ell$ gauge theory of lepton-flavor unification. In the quark sector, the guidance from these noninvertible symmetries pointed us to an overlooked theory of $SU(9)$ quark color-flavor unification in which the simplest solutions to the strong CP problem are revived in a flavor-symmetric way (the massless quark solution [37], the visible Peccei-Quinn-Weinberg-Wilczek (PQWW) axion [38], and the invisible Dine-Fischler-Srednicki-Zhitnitsky (DFSZ) axion [39], all of which face problems in their SM versions). Here we will unify the quark- and lepton-flavor effects.

In the quark sector this work will implement the massless quark solution to the strong CP problem via noninvertible symmetry breaking, and we refer to our earlier [37] for a pedagogical introduction to both the problem and this style of solution. But we note an important difference in the

noninvertible symmetry-breaking strategy from our previous uses. In prior work, proceeding from the bottom up, we have considered the theory of the SM extended by gauging

$$(SU(3)_C \times U(1)_{B_i+B_j-2B_k})/\mathbb{Z}_3, \quad (6)$$

where B_i refers to baryon number of the i th generation and, interestingly, the crucial quotient is only possible because $N_c = N_g$ in the SM. In [37] we first discovered the theory with the quotient can have a noninvertible, flavor-universal $\mathbb{Z}_3 \subset U(1)$ Peccei-Quinn (PQ) symmetry which acts as, for example,

$$Q_i \rightarrow \omega Q_i, \quad \bar{u}_i \rightarrow \omega^{-1} \bar{u}_i, \quad \bar{d}_i \rightarrow \omega \bar{d}_i, \quad (7)$$

where $\omega = e^{2\pi i/3}$ is the $\mathbb{Z}_3$ generator. This noninvertible symmetry exists firstly because the integer instantons from $I_{\text{IR}}(S^4)$ do not fully break the chiral $U(1)$ symmetries as their anomaly coefficients are a multiple of $3 = N_g$ and secondly due to “fractional instantons” present in the theory from $I_{\text{IR}}(\mathcal{M}^4)$ which wind through both factors as a result of the nontrivial quotient and as a result have anomaly coefficients without the factor of N_g . Then this $\mathbb{Z}_3$ factor is not broken, but it is converted by the fractional instantons into a noninvertible symmetry. The physical effect of this symmetry, in the case with only the SM matter fields, is that it protects all of the down-quark Yukawas and turns them into spurions for a noninvertible symmetry.

We have used this infrared guidance for model building through flavor unification to color-flavor unification

$$\frac{SU(3) \times U(1)}{\mathbb{Z}_3} \subset \frac{SU(3) \times SU(3)}{\mathbb{Z}_3} \subset SU(9), \quad (8)$$

where the noninvertible $\mathbb{Z}_3$ chiral symmetry persists in the intermediate stage. It is not until the color-flavor unification that the color-flavor magnetic monopoles are introduced from $\pi_2(SU(9)/[SU(3)^2/\mathbb{Z}_3]) = \mathbb{Z}_3$ and noninvertible symmetry breaking takes place. From the ultraviolet perspective this noninvertible symmetry breaking occurs via the instantons of the $SU(9)$ theory. Then in this UV theory we may start with solely the up-type Yukawa and a good PQ symmetry, and let the $SU(9)$ instantons generate the down-type Yukawas, bearing out the massless quark solution to strong CP .

In the present case, we can make use of the theoretical data of this noninvertible symmetry in a different way to effectively supply the missing instantons that lower the anomaly coefficient by a factor of 3: deconstruction of color by flavor. Here we consider a different intermediate completion of the quark flavored Z' . The infrared theory can arise from

$$\frac{SU(3)_1 \times SU(3)_2 \times SU(3)_3 \times U(1)_{B_1-B_2} \times U(1)_{B_1+B_2-2B_3}}{\mathbb{Z}_3 \times \mathbb{Z}_3}, \quad (9)$$

where the quarks of the i th generation are charged under $SU(3)_i$.⁶ Deconstruction also forbids mixing between generations at this stage, as did turning on the non-Abelian flavor symmetry above, which must be generated upon Higgsing back to the flavor-universal $SU(3)_C$.

The theory with nontrivial “global structure” can be seen to result from taking the $SU(3)^3 \times U(1)^2$ theory and quotienting the gauge group by a pair of $\mathbb{Z}_3$ elements, each of which lies in the diagonal between one of the Abelian groups and the non-Abelian groups under which those charged quarks transform. This imposes two equivalence relations between elements of $SU(3)^3 \times U(1)^2$:

$$\mathbb{Z}_3: 1 \sim (\omega 1_1, \omega^{-1} 1_2, 1_3, \omega^{-1} 1, 1), \quad (10)$$

$$\mathbb{Z}'_3: 1 \sim (\omega 1_1, \omega 1_2, \omega 1_3, 1, \omega^{-1} 1). \quad (11)$$

These quotients modify the one-form global symmetries of the theory, as will be discussed further in Sec. V: The charged matter content is restricted, and the magnetic monopoles of the theory now include both Abelian and non-Abelian fluxes. This type of structure fits into a family of gauge groups $(SU(N)^N \times U(1)^{N-1})/\mathbb{Z}_N^{N-1}$, which in general can be UV completed into $SU(N^2)$. Again, the possibility of the SM bearing this out is a special consequence of the fact that $N_g = N_c$.

In this case, by going to the deconstruction phase, the noninvertible symmetry will be broken but not due to the inclusion of the relevant magnetic monopoles. As discussed above, the noninvertible symmetry here relied on the fact that $SU(3)_C$ is flavor universal, such that mixed anomalies with it are proportional to N_g . Deconstruction allows a novel way to find noninvertible symmetry breaking by explicitly separating the degrees of freedom into N_g different gauge groups. Then the different $SU(3)_i$ integer instantons directly have smaller anomaly coefficients by a factor $1/N_g$ and so accomplish more breaking of the global symmetries than did the flavor-universal $SU(3)_C$. In this theory we still find $I_{\text{intermediate}}(\mathcal{M}^4) \supset I_{\text{intermediate}}(S^4)$, and the fractional instantons are still missing. However, now the integer instantons of $SU(3)_i$ break $U(1)_{\text{PQ}}$ entirely, with no part of it remaining as a noninvertible symmetry. This novel strategy to break noninvertible symmetries will be discussed further in [70].

Very interestingly, despite new monopoles not appearing at this ultraviolet scale, there is a sense in which this mechanism for breaking noninvertible symmetries by deconstruction still does add in fractional instantons. This is due to

⁶In fact in this presentation there is an additional $\mathbb{Z}_2$ identification between $U(1)_{B_1-B_2}$ and $U(1)_{B_1+B_2-2B_3}$. This factor can be trivialized by instead using the reparametrization of the $U(1)$'s as $U(1)_{B_1-B_2} \times U(1)_{B_2-B_3}$, which is equivalent, but here we want to keep explicit the connection to the further infrared group in which the noninvertible symmetry analysis was performed.

the nontrivial matching of the gauge couplings along this breaking:

$$\frac{1}{\alpha_s} = \sum_{i=1}^3 \frac{1}{\alpha_i}. \quad (12)$$

Thus in the flavor-symmetric limit the instanton action *does* jump by a factor of 3:

$$S_{\text{inst}}^{(\text{intermediate})} = \frac{1}{3} S_{\text{inst}}^{(\text{IR})}. \quad (13)$$

We find it intriguing that despite noninvertible symmetry breaking having not technically been accomplished by the existence of true fractional instantons (with fractional topological charge), in practice it is nonetheless accomplished by instantons with fractional action, just like one would find from instantons with fractional topological charge. It would be very interesting to further understand the relationship between these approaches.

III. THE ULTRAVIOLET $SU(12) \times SU(2)_L \times U(1)_R$ THEORY

We begin with $SU(3)_\ell$ lepton flavor unified with $SU(9)$ quark color-flavor into $SU(12)$. Gauged $U(1)_{Q-N_c L}$ appears automatically as a subgroup, and relatedly the gauged $U(1)$ symmetry starts as $U(1)_R$ as is familiar from Pati-Salam models. Here we choose not to gauge the full $SU(2)_R$ because a key ingredient will be a PQ symmetry which distinguishes up and down.

This means the SM fermions are in three irreducible representations (see Table I), and we introduce two quark-lepton color-flavor scalar irreps as well in addition to the Higgs (see Table II).

- (i) The *complex* adjoint Σ_b^a which spontaneously breaks the $SU(12)$. We outline this in stages in Fig. 1 for conceptual benefit but hierarchical breaking is not necessary, and indeed our simple benchmark will join most of the steps. This is simply an organization of two real adjoints and the $U(1)_\Sigma$ need not be respected—indeed, to generate the weak CP -violating angles Σ should spontaneously and/or explicitly violate this symmetry.
- (ii) The two-index symmetric ϕ^{ab} which is the only new scalar with an electroweak quantum number and can interact with the SM fermions at tree level. Remarkably, this plays multiple distinct roles: spontaneously breaking lepton flavor and providing right-handed neutrino Majorana masses, spontaneously breaking $U(1)_R$ and $U(1)_{Q-N_c L}$ down to hypercharge, preserving a discrete lepton-flavored gauge symmetry which prevents proton decay, and also dynamically distinguishing up-type vs down-type quarks to enable the generation of the CKM phase.

TABLE I. Gauge and global symmetry charges of fermions in the UV phase.

	$SU(12)$	$SU(2)_L$	$U(1)_R$	$U(1)_{PQ}$	$U(1)_{Q+L}$
Ψ	12	2	0	0	+1
ψ_u	$\overline{12}$	$\cdots$	+1	0	-1
ψ_d	$\overline{12}$	$\cdots$	-1	+1	-1

TABLE II. Gauge and global symmetry charges of scalars in the UV phase.

	$SU(12)$	$SU(2)_L$	$U(1)_R$	$U(1)_{PQ}$	$U(1)_{Q+L}$
H	$\cdots$	2	-1	0	0
Σ_b^a	143	$\cdots$	0	0	0
ϕ^{ab}	78	$\cdots$	-2	0	+2

We start with an extremely simple Lagrangian:

$$\mathcal{L} = y_t H \Psi \psi_u + \frac{1}{2} \lambda \phi \psi_u \psi_u + V(H, \Sigma, \phi), \quad (14)$$

with y_t and λ two free parameters which are naturally $\mathcal{O}(1)$. We do not fully specify the potential through which the scalars interact but take it as given that the scalars can condense in the directions we dictate, as is usual in initial work on theories of unification.

The theory of Eq. (14) additionally has two global symmetries which act nontrivially on the fermions, both of which are anomalous. The absence of interactions for ψ_d is enforced by a classical symmetry effecting $\psi_d \rightarrow e^{i\alpha} \psi_d$, which at the quantum level becomes a Peccei-Quinn symmetry and will be crucial for the instanton effects we will find. As to the four fields which do interact in Eq. (14), the free theory of these fields would have $U(1)^4$, two directions of which are broken by the two interactions, leaving both the gauged $U(1)_R$ and a global $U(1)_{Q+L}$ in addition to the $U(1)_{PQ}$. In Table I we take a useful choice of basis for the global symmetries in which we identify $U(1)_{PQ}$ as a direction with solely $SU(12)$ non-Abelian mixed anomaly and $U(1)_{Q+L}$ as a direction with solely $SU(2)_L$ non-Abelian mixed anomaly. We refer to our recent [39] for detailed discussion on the abstract space of global symmetries and the scalar field space in theories with PQ symmetries.

The direction $U(1)_{Q+L}$ is quark plus lepton number, which parallels the phenomenology of baryon plus lepton number in the Standard Model. This symmetry is “accidental” in that any renormalizable Lagrangian must classically respect it due to the gauge charges. The $SU(2)_L$ instantons already give the familiar SM ’t Hooft vertex between 12 fermions $\mathcal{L} \subset e^{a_1 \dots a_{12}} \Psi_{a_1} \dots \Psi_{a_{12}}$ which allow processes that exchange nine quarks for three leptons and

breaks $U(1)_{Q+L}$ into the conserved $\mathbb{Z}_{12}^{Q+L}$ which overlaps with the $SU(12)$ gauge symmetry.

The $SU(12)$ instantons fully violate $U(1)_{PQ}$, and as these effects grow in magnitude into the infrared they will lead to the ’t Hooft vertices and noninvertible symmetry breaking we need to generate the missing Yukawas and solve strong CP . This holds true because the $SU(12)$ gauge coupling is asymptotically free at $\beta_{12} = -109/3$. On the other hand the $U(1)_R$ factor has $\beta_R = 362/3$ as the result of the large scalar 78 representation, implying that a UV completion must be relatively close. Here this will not impose any requirements—as we will discuss further in the next section, the scale of our dynamics will necessarily be quite high already, so that M_{pl} will cut off the $U(1)_R$ theory before the Landau pole (see Fig. 7).

The matter content of this theory is very minimal, and we will end up finding that some colored scalar spectator fields are needed to modify the running and consistently achieve the many things we are demanding from this minimal setup. What matters for us is their contribution to the beta function, so we are agnostic about their identities and here assume they have no interactions with the SM fields. However, we mention one well-motivated way to add colored degrees of freedom is to explore upgrading the Higgs sector from a color-flavor singlet to include $SU(12)$ singlet + adjoint H_u and H_d , which can achieve the modified running, can help facilitate full UV gauge coupling unification, and can also help generate the full Yukawa structure. We postpone analyzing this structure to future work and in Sec. VII just take as a free parameter an additional contribution $\beta_{\text{spectator}}$ coming from the spectator fields.

A. Global structure and generalized symmetries

The ultraviolet theory contains a continuous Abelian factor, so this theory possesses a $U(1)_m^{(1)}$ magnetic one-form symmetry under which the ’t Hooft lines are charged. This is due to the absence of dynamical magnetic monopoles for $U(1)_R$, and this magnetic one-form symmetry will match on to the $U(1)_Y$ hypercharge magnetic one-form symmetry⁷ and then the $U(1)_{\text{EM}}$ electromagnetic one-form symmetry. This means that this theory itself does not suffer a monopole problem, as electromagnetic monopoles are not present. We will see below that other magnetic monopoles will emerge at intermediate stages, but these will all get confined by strings before the SM phase is reached. In a given further ultraviolet theory—for example embedding $U(1)_R \subset SU(2)_R$ or embedding the whole theory in a five-dimensional orbifold—there will be extra topology which

⁷In the Standard Model phase this hypermagnetic one-form symmetry is intertwined in an approximate two-group with the SM flavor symmetries [53]. In that paper the connection with traditional, “vertical” GUTs was explored; it would be very interesting to understand how that structure emerges from a flavorful GUT like this one.

will also provide dynamical electromagnetic monopoles and break the one-form symmetry explicitly. But this is not the only consideration for the one-form symmetries of this UV theory.

There are two possible versions of this theory which possess either a $\mathbb{Z}_2^{(1)}$ electric one-form symmetry or a modified $U(1)^{(1)}$ magnetic one-form symmetry. Consider a simultaneous transformation by $\mathbb{Z}_2$ in each of the three gauge factors—this diagonal $\mathbb{Z}_2$ acts trivially on all the fields listed in Tables I and II. Then we are allowed to consider two possibilities for the gauge group.

Option 1: $G_{UV} = SU(12) \times SU(2)_L \times U(1)_R$. This theory has a $\mathbb{Z}_2^{(1)}$ electric one-form symmetry under which the charged Wilson lines are those that cannot be ended on the matter included in this theory so far—that is, any Wilson line which picks up a negative sign under that shared $\mathbb{Z}_2$ transformation. In the far infrared, this maps on to Wilson lines with fractional electric charge $(1/2 + k)e$, with $k \in \mathbb{Z}$. The minimal field content above does not include any such fields, being only a theory of the chiral SM matter. But for this gauge group, with general matter there will be additional complex scalars or Dirac fermions on which those Wilson lines can end because they do transform with a negative sign under that $\mathbb{Z}_2$. The masses of these fields can be entirely separate from the scales generated for the SM fields in this theory, and in particular any mass for such a fermion is technically natural. If they happen to be light, these FCPs may be searched for at the LHC [6,7]. Quantum gravity suggests higher-form global symmetries should not be absolute, so if this is the gauge group one should generically expect such additional fields to be present. This gauge theory will flow in the infrared to the SM electroweak gauge group with no global structure $G_{EW} = SU(2)_L \times U(1)_Y$.

Option 2: $G'_{UV} = (SU(12) \times SU(2)_L \times U(1)_R)/\mathbb{Z}_2$. This theory may be constructed by gauging the $\mathbb{Z}_2^{(1)}$ electric one-form symmetry of the theory above. As the fractionally charged Wilson lines are charged under this one-form symmetry, gauging it removes them from the spectrum of gauge-invariant line operators and as a result FCPs cannot exist in this theory. Group theoretically, the magnetic one-form symmetry remains $U(1)_m^{(1)}$ as $\pi_1(G'_{UV}) \simeq \pi_1(G_{UV})$, but now the closed paths through the gauge group are changed. The fact that the minimal winding path now goes from the origin to $(-\mathbb{1}_{12}, -\mathbb{1}_2, e^{i\pi})$ means that the R -magnetic monopoles with odd charges are now dressed also with half-fractional non-Abelian magnetic fluxes. In the infrared this theory will flow to the electroweak group $G_{EW} = (SU(2)_L \times U(1)_Y)/\mathbb{Z}_2$.

The physics that we discuss in the rest of this work will not be affected by the choice of $\mathbb{Z}_2^{(1)}$ symmetry. However, connecting this theory to a further ultraviolet completion may demand one or the other choice.

We comment briefly on the existence of noninvertible symmetries in this phase: There are not any, despite that

there are “missing” instantons $I(S^4) \subset I(\mathcal{M}^4)$. On a general manifold, there are integer instanton backgrounds of $U(1)_R$ which break the $U(1)_{PQ}$ Peccei-Quinn symmetry to $\mathbb{Z}_{12}^{PQ}$, but as this symmetry is already fully anomalous in $SU(12)$ integer instanton backgrounds, this does not imply a noninvertible symmetry unless we were to turn off the $SU(12)$ gauge coupling. A more subtle possibility to check is that in the theory with the $\mathbb{Z}_2$ quotient there are *fractional* instantons which have correlated, fractional instanton numbers under all three groups. However, these do not violate $U(1)_{Q+L}$ [which can be understood due to their inherent left-right symmetry inherited from the possible $SU(2)_R$ UV completion] so also do not lead to noninvertible symmetries in this theory. In a theory that introduced more fields and so more global symmetries, it is possible that these Abelian integer instantons or fractional instantons would lead to interesting noninvertible symmetries, pointing to nonperturbative physics effects in a further full unification into a simple group. But here we make no use of these instantons.

IV. BREAKING $SU(12)$ SEPARATES QUARK AND LEPTON FLAVOR

One structural mystery of the Standard Model is that the global quark- and lepton-flavor symmetries are broken in very different ways, from the relative sizes of the masses to the mixing patterns. If we are to upgrade flavor to a gauge symmetry in the context of a theory that unifies quarks and leptons, this seems to be an obstruction.

While it is possible to break $SU(12) \rightarrow SU(4)_C \times SU(3)_F$ leading to a Pati-Salam(esque) theory with gauged flavor, it turns out to be more profitable to break “in the Pati-Salam direction” first:

$$SU(12) \rightarrow \frac{SU(9)_{\text{quarks}} \times SU(3)_{\text{leptons}} \times U(1)_{Q-N_c L}}{\mathbb{Z}_9}, \quad (15)$$

where the $\mathbb{Z}_9$ is generated by $(\zeta^{-1}\mathbb{1}_9, \omega\mathbb{1}_3, \zeta)$, with $\zeta = e^{2\pi i/9}$, $\omega = \zeta^3$. This leads to a theory of *separate* gauged flavor symmetries for the quarks and the leptons, which is crucial to the interesting phenomenologies we will find. The nontrivial global structure can be seen from the fact that the generator of $Q - N_c L$ is the familiar T^{15} of $SU(4)$ in Pati-Salam but now extended across $SU(12)$ as the direct product $\mathbb{1}_3 \otimes T^{15}$, with $\mathbb{1}_3$ the 3×3 identity matrix. After breaking this theory down to the SM, the $\mathbb{Z}_3$ subgroup of this quotient will eventually become the $\mathbb{Z}_3$ quotient of the SM, $(SU(3)_C \times U(1)_Y)/\mathbb{Z}_3$, as is familiar from Pati-Salam theories.

This breaking can be implemented by the adjoint Σ^a_b condensing with vacuum expectation value (VEV)

$$\langle \Sigma^a_b \rangle = \Lambda_{12} \text{diag}(3\mathbb{1}_3, -\mathbb{1}_9). \quad (16)$$

TABLE III. Fermions in the $SU(9) \times SU(3)$ phase and their gauge representations.

	$SU(9)_q$	$SU(3)_\ell$	$SU(2)_L$	$U(1)_R$	$U(1)_{Q-N_cL}$
Ψ_q	9	$\dots$	2	0	+1
Ψ_ℓ	$\dots$	3	2	0	-3
ψ_{qu}	$\bar{9}$	$\dots$	$\dots$	+1	-1
ψ_{qd}	$\bar{9}$	$\dots$	$\dots$	-1	-1
$\psi_{\ell u}$	$\dots$	$\bar{3}$	$\dots$	+1	+3
$\psi_{\ell d}$	$\dots$	$\bar{3}$	$\dots$	-1	+3

One of the off-diagonal “bifundamentals” $(9, \bar{3})_{+4}$ is eaten by the gauge bosons and becomes heavy. We show in Tables III and IV the charges of the matter under the remaining gauge symmetries.

The Lagrangian decomposes to

$$\mathcal{L} = y_l H(\Psi_q \psi_{qu} + \Psi_\ell \psi_{\ell u}) + \lambda \left(\frac{1}{2} \phi^{ab} \psi_{qu,a} \psi_{qu,b} + \phi^{a\alpha} \psi_{qu,a} \psi_{\ell u,\alpha} + \frac{1}{2} \phi^{\alpha\beta} \psi_{\ell u,\alpha} \psi_{\ell u,\beta} \right), \quad (17)$$

and higher-dimension operators are generated upon integrating out the heavy gauge bosons such as

$$\Lambda_{12}^2 \mathcal{L}_6 \simeq (\psi_{qd} \sigma^\mu \psi_{\ell d}^\dagger)(\psi_{\ell u} \sigma_\mu \psi_{qu}^\dagger) + \dots \quad (18)$$

arising from the currents coupling to the heavy $(9, \bar{3})_{+4}$ vector boson.

Naively the quarks and leptons may enjoy distinct vector-like global symmetries for quark and lepton number, but these are linked by the gauged $U(1)_{Q-N_cL}$. We may continue to consider quark plus lepton number which the $SU(2)_L$ anomaly continues to break to $\mathbb{Z}_{12}^{Q+L}$. We could alternatively consider, for example, $Q + N_cL = 3(B + L)$ which is broken by the anomaly instead to $\mathbb{Z}_{18}^{Q+N_cL}$ as happens in the low-energy Standard Model as well and protects the proton [18]. There are also PQ symmetries acting separately on the right-handed down quarks or charged leptons, each of which is fully broken by $SU(9)$ or $SU(3)_\ell$ mixed anomalies. The irrelevant operator in $\mathcal{L}_6$ couples them together and breaks them down to their diagonal $U(1)_{\text{PQ},q+\ell}$, which has the effect of placing a lower bound on the amount of Peccei-Quinn breaking in one sector with respect to the other (see footnote 9 below).

In this phase we also have an emergent $U(1)^{(1)}$ magnetic one-form global symmetry due to the Abelian factor of $U(1)_{Q-N_cL}$. The quotient between $Q - N_cL$, $SU(3)$, and $SU(9)$ again modifies the structure of this one-form symmetry, as the minimal winding path now goes from the origin to $(\zeta^{-1} \mathbb{1}_9, \omega \mathbb{1}_3, \zeta)$. Then among the $\mathbb{Z}$ spectrum of 't Hooft lines, those not at $3\mathbb{Z}$ units of $Q - N_cL$ magnetic flux have additional compensating one-third fractional $SU(3)$ -magnetic flux, and those not at $9\mathbb{Z}$ units

TABLE IV. Complex scalars in the $SU(9) \times SU(3)$ phase. The fields in bold are not needed at lower energies and could receive large $SU(12)$ -breaking sized masses. $a, b = 1 \dots 9$ are $SU(9)$ indices and $\alpha, \beta = 1 \dots 3$ are $SU(3)$ indices.

	$SU(9)_q$	$SU(3)_\ell$	$SU(2)_L$	$U(1)_R$	$U(1)_{Q-N_cL}$
H	$\dots$	$\dots$	2	-1	0
Σ_b^a	80	$\dots$	$\dots$	0	0
Σ_β^α	$\dots$	8	$\dots$	0	0
$(\Sigma)_\beta^a$	9	$\bar{3}$	$\dots$	0	+4
(Σ)	$\dots$	$\dots$	$\dots$	0	0
ϕ^{ab}	45	$\dots$	$\dots$	-2	+2
$\phi^{a\alpha}$	9	3	$\dots$	-2	-2
$\phi^{\alpha\beta}$	$\dots$	6	$\dots$	-2	-6

of $Q - N_cL$ magnetic flux have additional compensating fractional $SU(9)$ -magnetic fluxes as well. If $SU(12)$ is dynamically broken in the early Universe, these 't Hooft lines are brought to life as magnetic monopoles which form necessarily due to causality [71]. It would be very interesting to understand if these monopoles have a Callan-Rubakov-like effect [72,73] due to the mixed anomaly of $SU(9)$ with the PQ symmetry. Such monopoles could potentially play a role in baryogenesis [67] with the additional benefit that at a lower scale when $U(1)_{Q-N_cL}$ is Higgsed, they will become end points of cosmic strings whose tension can cause them to annihilate.

A. Beta functions and matching

The running of the gauge couplings in this stage depends on which degrees of freedom get mass at the scale Λ_{12} , so there is a range of possible beta functions for each factor. Some of the scalars could receive a large mass $\sim \Lambda_{12}$ and be removed from the spectrum, while others are needed at subsequent breaking steps. The adjoints are needed to spontaneously break the flavor gauge groups, as is the two-index symmetric lepton-flavored $\phi^{a\beta}$, while $\phi^{a\alpha}$ is used in the flavor scheme to distinguish up vs down. The others are not strictly necessary, and our benchmark model decouples them.

Across the $SU(12)$ -breaking scale the non-Abelian gauge couplings match as they have trivial index of embedding, $g_{12} = g_9 = g_\ell$. However for the Abelian factor the sensible infrared quantization of the couplings often deviates from the ultraviolet normalization of $\text{Tr}(T^2) = \frac{1}{2}$ over the fundamental. We have the nontrivial matching condition $g_{12} = \sqrt{72} g_{Q-N_cL}$.⁸ In terms of running, the beta

⁸Note this large factor is because we are normalizing $Q - N_cL$ in the unbroken phase with the quarks having charge 1, which is thrice their $B - L$ charge. In the standard normalization this would be $g_{12} = \sqrt{8} g_{B-L}$. Note also that this differs by a factor of root 3 from the relation in Pati-Salam theories, $g_4 = g_3 = \sqrt{8/3} g_{B-L}$, as the gauge groups are thrice as big.

functions for both gauge flavor symmetries remain asymptotically free at $\beta_9 = -169/6$ for $SU(9)$ and $\beta_\ell = -19/3$ for $SU(3)_\ell$, which means as we flow down toward the IR, g_9 grows with respect to g_ℓ . This is crucial as it allows the different running of couplings to be responsible for generating the difference between the bottom and the tau Yukawa (along possibly with different scales for the breaking of these symmetries) as we will discuss in the next section. For the $U(1)$ factors, our benchmark has $\beta_{Q-N_c L} = 204$ and $\beta_R = 182/3$. Note that in general one should also keep track of the kinetic mixing induced by the bicharged fields above, but in our benchmark this will not be important.

B. Instanton effects and noninvertible symmetry breaking

From the perspective of the ultraviolet theory, we have simply the breaking of anomalous symmetries by instantons, which for an asymptotically free theory dominantly occurs in the infrared. The infrared noninvertible symmetries are broken in the $SU(9)$ and $SU(3)_\ell$ gauge theories by quark color-flavor and lepton-flavor instantons, which generate separately both of y_b and y_τ . In principle we would like the small hierarchy y_τ/y_b to result from slight differences in the evolution of the two sectors following the breaking of the unified group. See Fig. 2 showing the generation of the down-type Yukawas from the 't Hooft vertices.

The 't Hooft vertices of $SU(3)_\ell$ have four Fermi zero modes, one each of $\psi_{\ell u} = \bar{\nu}_\alpha$ and $\psi_{\ell d} = \bar{e}_\alpha$ and two for $\Psi_\ell = L^\alpha$. Using the tree-level instanton action $8\pi^2/g^2$ evaluated at the scale of the largest breaking, after the up-type Yukawa is used to contract the neutrino legs the instantons generate, at the lowest order of approximation

$$\mathcal{L} \supset y_t H L^\alpha \bar{\nu}_\alpha + y_t^* e^{-\frac{2\pi}{a_\ell(\Lambda_\ell)}} \tilde{H} L^\alpha \bar{e}_\alpha, \quad (19)$$

where Λ_ℓ is the scale at which $SU(3)_\ell$ breaks, the needed coupling to produce y_τ is large enough that it is not clear the dilute instanton gas approximation can be trusted, and precise predictions for the needed α_ℓ coupling in this theory will require lattice simulation. To be clear, it is only the size of the coefficient which is uncertain—the generation of this operator is a robust prediction due to the Adler-Bell-Jackiw anomaly structure.

The 't Hooft vertices of $SU(9)$ analogously have four Fermi zero modes as well, one each of $\psi_{qu} = \bar{u}_a$ and $\psi_{qd} = \bar{d}_a$ as well as two for $\Psi_q = Q^a$. We again contract the up-type legs with the up-type Yukawa, now producing a down-type Yukawa of the form

$$\mathcal{L} \supset y_t H Q^a \bar{u}_a + y_t^* e^{-\frac{2\pi}{a_9(\Lambda_9)}} \tilde{H} Q^a \bar{d}_a, \quad (20)$$

where Λ_9 is the scale at which $SU(9)$ breaks and the needed coupling here is even larger, so again lattice simulations

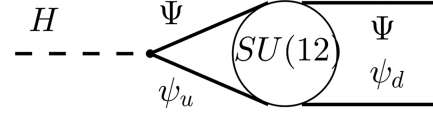


FIG. 2. An $SU(12)$ instanton violates PQ by one unit and generates a down-type Yukawa from an up-type Yukawa. The largest contributions come after $SU(12)$ breaks and there are separate $SU(9)$ instantons for the quarks and $SU(3)_\ell$ instantons for the leptons, but the form remains the same.

will be required to determine the necessary size of α_9 to generate the correct Yukawa, just as they were needed to understand the original massless quark solution.^{9,10}

The quark and lepton theories begin with equal gauge couplings $\alpha_\ell(\Lambda_{12}) = \alpha_9(\Lambda_{12})$, but the different multiplicity factors lead to different one-loop running. If we focus on the ratio y_τ/y_b , then the boundary condition at Λ_{12} drops out and this ratio depends on the possible mild amount of scale separation between the quark- and lepton-flavor-breaking scales. Proceeding with the naive estimate

$$y_\tau/y_b = \exp 2\pi(\alpha_9^{-1}(\Lambda_9) - \alpha_\ell^{-1}(\Lambda_\ell)) \quad (21)$$

and since

$$\alpha_9^{-1}(\Lambda_9) \simeq \alpha^{-1}(\Lambda_{12}) - \frac{\beta_9}{2\pi} \log\left(\frac{\Lambda_9}{\Lambda_{12}}\right), \quad (22)$$

$$\alpha_\ell^{-1}(\Lambda_\ell) \simeq \alpha^{-1}(\Lambda_{12}) - \frac{\beta_\ell}{2\pi} \log\left(\frac{\Lambda_\ell}{\Lambda_{12}}\right), \quad (23)$$

then we can find the needed lepton-breaking scale as

$$\frac{\Lambda_\ell}{\Lambda_{12}} = \left(\frac{y_\tau}{y_b}\right)^{1/\beta_\ell} \left(\frac{\Lambda_9}{\Lambda_{12}}\right)^{\beta_9/\beta_\ell}. \quad (24)$$

The structure of our model has type-I seesaw Majorana masses generated at $SU(3)_\ell$ breaking which fixes the overall scale, $\Lambda_\ell \sim 10^{15}$ GeV. If we choose the two

⁹From here we can see the effect of the higher-dimensional operators in Eq. (18) is to generate the PQ-violating leptonic operator from the $SU(9)$ 't Hooft vertex but suppressed by $\Lambda_9^2/\Lambda_{12}^2$. This could be a problem if we wanted to generate only $y_\tau \ll y_b$, but in fact they are not so different and so the higher-dimensional operators are not dangerous.

¹⁰An important difference from the original massless quark solution and from our earlier massless quark solution in quark color-flavor unification [37] is that the dynamics here take place at a high scale, as they are tied to the seesaw scale. While the massless quark solution can be automatically safe from the quality problem if all the scales are $\lesssim 10^8$ GeV, here our model does suffer from the quality problem. This is one reason to search for a quark-lepton flipped model from extradimensional dynamics in which one begins with y_t and y_τ and generates y_b and y_ν , which could take place at much lower scales. Or we must add a module to solve the quality problem, but we leave study of this important problem to future work.

flavor-breaking scales equal, then all three scales must be nearly degenerate $\Lambda_\ell = \Lambda_9 \simeq 0.96\Lambda_{12}$. If we want a separation of scales, the most we can do is raise $SU(12)$ breaking to around the Planck scale $\Lambda_\ell/\Lambda_{12} \sim 10^{-3}$ which requires $\Lambda_9/\Lambda_{12} \sim 1/5$. Any choice works fine at the level we are considering, but for simplicity in Sec. VII we will consider the case where all the breaking occurs at just around the same scale.

While the α dependence drops out of this estimate and so one may hope this prediction is more robust than the absolute values of y_τ, y_b , there is much physics they ignore. These estimates could be made more precise by accounting for the instanton density in further detail (see e.g. [74,75] for recent approaches). As our purpose here is understanding the overall structure, we will not belabor a more-precise instanton computation. Furthermore this simple estimate has not accounted for the running of the Yukawas themselves. We hope this interesting scenario, along with previously studied possibilities for these $SU(9)$ instantons [37–39], motivates lattice simulations of this gauge theory.

V. BREAKING QUARK COLOR-FLAVOR

Here we consider further breaking of $SU(9)$ by the adjoint VEV

$$\langle \Sigma_a^b \rangle = \Lambda_9 \text{diag}(+\mathbb{1}_3, 0 \times \mathbb{1}_3, -\mathbb{1}_3), \quad (25)$$

which breaks

$$SU(9) \rightarrow \frac{SU(3)_1 \times SU(3)_2 \times SU(3)_3 \times U(1)_{T^3} \times U(1)_{T^8}}{\mathbb{Z}_3 \times \mathbb{Z}_3} \quad (26)$$

performing color-flavor deconstruction.¹¹ Under this decomposition we have the branchings

$$\begin{aligned} 9 &\rightarrow (3, 1, 1)_{(+1,+1)} \oplus (1, 3, 1)_{(-1,+1)} \oplus (1, 1, 3)_{(0,-2)}, \\ 80 &\rightarrow 2(1, 1, 1)_{(0,0)} \\ &\oplus (8, 1, 1)_{(0,0)} \oplus (1, 8, 1)_{(0,0)} \oplus (1, 1, 8)_{(0,0)} \\ &\oplus (3, \bar{3}, 1)_{(+2,0)} \oplus (1, 3, \bar{3})_{(-1,+3)} \oplus (\bar{3}, 1, 3)_{(-1,-3)} \\ &\oplus (\bar{3}, 3, 1)_{(-2,0)} \oplus (1, \bar{3}, 3)_{(+1,-3)} \oplus (3, 1, \bar{3})_{(+1,+3)}. \end{aligned}$$

¹¹In prior work we have considered the symmetry-breaking pattern $SU(9) \rightarrow SU(3)^2/\mathbb{Z}_3$ which may be implemented by the three-index symmetric 165 of $SU(9)$. This pattern is viable here as well, with the most simple possibility having the 165 arise from the 364 of $SU(12)$. This results in a 165 with nontrivial $Q - N_c L$ charge, so it should also carry $U(1)_R$ charge in the right combination such that its condensation preserves $U(1)_Y$. While this is a possibility, the condensation of such a field will break an interesting discrete gauge symmetry and prevent some of the physics we will discuss in the next section, so we will think about another possibility. Producing a 165 which is neutral under $Q - N_c L$ requires such a large $SU(12)$ representation that we do not pursue this possibility.

Here the subscripts denote the charges under $U(1)_{T^3} \times U(1)_{T^8} = U(1)_{B_1-B_2} \times U(1)_{B_1+B_2-2B_3}$, and all representations can be constructed from the fundamental, but we give the adjoint for convenience. The labels come from writing the generators as $\mathbb{I}_3 \otimes T^i$ with T^i a diagonal generator of $SU(3)$.¹²

The structure of the quotients within $SU(9)$ has been discussed already in Sec. II, but we recall that due to further unification our representations also transform nontrivially under $U(1)_{Q-N_c L}$, and the deconstruction phase inherits another $\mathbb{Z}_9$ quotient from the $SU(9)$ quotient. In this phase the quotient is generated by

$$\mathbb{Z}_9: \mathbb{1} \sim (\omega^{-1}\mathbb{1}_3, \zeta^{-1}, \omega\mathbb{1}_3, \zeta\mathbb{1}) \quad (27)$$

as an element of $SU(3)_3 \times U(1)_{T^8} \times SU(3)_\ell \times U(1)_{Q-N_c L}$, with action on the other factors by the identity.

Then the group structure at this stage is actually

$$\left(\left[\frac{SU(3)^3 \times U(1)^2}{\mathbb{Z}_3 \times \mathbb{Z}_3} \right] \times SU(3)_\ell \times U(1)_{Q-N_c L} \right) / \mathbb{Z}_9. \quad (28)$$

Furthermore, we note that there can be an approximate global S_3 rearrangement symmetry in this infrared group, which is perhaps more obvious if we choose a different basis of the $U(1)$ groups such as $B_1 - B_2$ and $B_1 - B_3$. Note also that at this intermediate stage there is a further emergent $[U(1)^{(1)}]^2$ magnetic one-form symmetry in the quark color-flavor sector as a result of the $U(1)^2$ gauged quark flavor difference symmetries. There are 't Hooft lines labeled by (m_1, m_2, m_3) such that $\sum_i m_i = 0$, but their structure is also sensitive to the quotient. In particular if all $m_i \in 3\mathbb{Z}$, then these have purely Abelian magnetic fluxes, but a general monopole also have fractional non-Abelian magnetic fluxes of the three $SU(3)_i$ groups compensating for $m_i/3$ fluxes. Of course the $SU(9)$ theory explicitly breaks the magnetic one-form symmetries in the UV by providing the dynamical magnetic monopoles on which the 't Hooft lines can end, meaning that in the early Universe this breaking leads to further production of magnetic monopoles past those already produced in the $SU(12)$ breaking. These monopoles will also become confined by cosmic strings once the deconstructed theory is broken down further to the strong sector of the SM.

In the breaking from $SU(9)$, half of the bifundamental scalars are eaten by the off-diagonal vector bosons, but one complex $(3, \bar{3}, 1) \oplus (1, 3, \bar{3}) \oplus (\bar{3}, 1, 3)$ remains and we assume these stay light to accomplish the next stage of breaking; see Fig. 3. The condensation of these

¹²Note also that the $SU(3)_i$ labels need not map directly to what we normally think of as the i th quark generation. It would also be interesting to understand the possibility of a mismatch in quark and lepton generation numbers arising out of this framework.

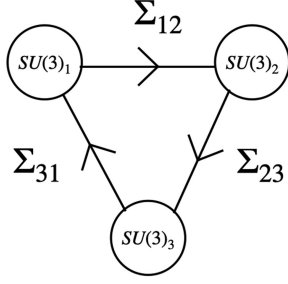


FIG. 3. A quiver diagram depicting the non-Abelian parts of the gauge theory as well as the bifundamental link fields which will break the deconstructed symmetry.

“link fields” as

$$\langle \Sigma_{b_j}^{a_i} \rangle = \Lambda_3 \delta_{b_j}^{a_i} \quad (29)$$

breaks the deconstructed gauge theory back to the flavor reunified color group $SU(3)^3 \rightarrow SU(3)_C$. The $U(1)$ factors will be broken simultaneously as a result of their Abelian charges, and resultingly only the $\mathbb{Z}_3$ subgroup of the $\mathbb{Z}_9$ quotient remains. This step of breaking must happen at a high scale $\gtrsim 1000$ TeV to avoid constraints on flavor-changing neutral currents, though a fuller study of avoiding such constraints here is left to future work. We note in recent years there has been interesting work on models of flavor deconstruction where the link fields get VEVs at very different scales e.g. [76,77] and it would be interesting to explore that idea in this context, such that new fields charged under the third-generation factor could be accessible at colliders.

If S_3 rearrangements are an approximate global symmetry and the quark flavor adjoint enjoyed an approximate $U(1)$ phase symmetry such that a combination of global and gauge transformation preserves the vacuum, then it is possible a type of “Alice strings” [78–80] may exist which have an additional global symmetry twist [81–83].¹³ We defer a full characterization of their existence to future work, but note it would be interesting to think about the interplay of these strings with both particles and magnetic monopoles, potentially using generalized symmetry technology e.g. [84–86].

A. Generating CKM

In [37] we showed that quark color-flavor $SU(9)$ unification could implement a flavor-symmetric massless quark solution to strong CP , and we refer there for more detail. In addition to generating the down-quark Yukawas from the up-quark Yukawas, a crucial component was the

¹³It is also possible in a modified model that a gauged S_3 could be preserved in this phase. This could involve a VEV for $(\Sigma^\dagger)^a_b \Sigma^c_d$ without a VEV for Σ^a_b itself, which is possible, and it would be interesting to pursue this possibility in detail.

demonstration that general quark Yukawas—and in particular a large CP -violating CKM angle—could be generated without upsetting the massless quark solution. The proof of principle scheme developed there worked as follows.

- (i) Fields arising from the $SU(9)$ complex adjoint Σ_b^a spontaneously break the intermediate color-flavor symmetries down to $SU(3)_C$. These same fields also spontaneously or explicitly violate CP through the scalar potential, e.g. in the simplest case

$$V_{\mathcal{CP}}(\Sigma) \supset \eta_1 \text{Tr}(\Sigma^4) + \eta_2 \text{Tr}(\Sigma^2)^2 + \dots + \text{H.c.}, \quad (30)$$

where we stress the mechanism works no matter how the CP violation appears in the potential.

- (ii) The symmetries and spectrum (including in particular the absence of vectorlike quarks) dictate that the breaking from Σ is communicated to the quarks only at loop level, which generates complex Yukawas that are automatically Hermitian (in the basis where the tree-level Yukawas are real, which always exists because of the UV PQ symmetry and the down-type Yukawas having been generated by instantons). This results from the fact that for every diagram with an insertion of a $V_{\mathcal{CP}}$ interaction on the Σ leg there is a corresponding diagram with $V_{\mathcal{CP}}^\dagger$ inserted on the $\Sigma^\dagger$ leg.

Then the Yukawas generated have a structure

$$(y_d)^a_b \sim y_b \left(1^a_b + \frac{\alpha}{4\pi} \frac{\{\Sigma, \Sigma^\dagger\}^a_b}{\Lambda_3^2} + \dots + \frac{\alpha}{(4\pi)} \frac{\eta_1^\dagger (\Sigma^{\dagger 4})^a_b + \eta_2^\dagger \text{Tr}(\Sigma^{\dagger 2})(\Sigma^{\dagger 2})^a_b}{\Lambda_3^4} + \dots + \text{H.c.} \right), \quad (31)$$

where the first line contains terms with manifestly real coefficients and the second line contains CP -violating terms with complex coefficients but which come always with a Hermitian conjugate term. This means we can generate nontrivial δ_{CKM} while manifestly preserving $\bar{\theta} = 0$. See Fig. 4.

- (iii) Loop communication through the heavy gluons alone would generate the same Yukawa structure for the up and down Yukawas, which would result in vanishing CKM phase, so another pair of fields must be present that distinguishes up and down quarks by communicating additional violation only to one type, again at loop level. See Fig. 5.

In the $SU(9)$ theory, in order to find contributions distinguishing between up and down we needed to add new

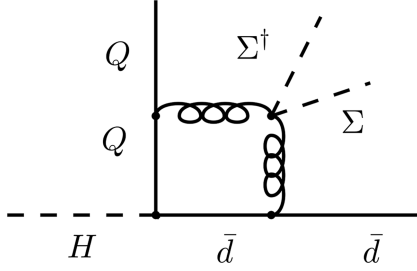


FIG. 4. Flavor violation communicated to the down-type quarks through a loop of heavy gluons upon $SU(3)^3 \rightarrow SU(3)$. This yields complex, yet Hermitian Yukawas.

fields to the theory accomplishing this. However, here we already have ϕ^{aa} and $\psi_{\ell u}$, which bear out the exact phenomenology needed. ϕ^{aa} has a cubic or quartic coupling to Σ_b^a and connects only the up-type quarks to the right-handed neutrinos via the $\lambda \phi^{aa} \psi_{qu} \psi_{\ell u}$ interaction. This will generate contributions like

$$(y_u)^a{}_b \sim \frac{y_t}{y_b} (y_d)^a{}_b + \frac{y_t |\lambda|^2 c}{(4\pi)^2} \frac{(\Sigma^\dagger \Sigma)^a{}_b}{M_\phi^2} + \dots + \text{H.c.}, \quad (32)$$

where λ is the coupling of ϕ to $\psi_u \psi_u$, c is a quartic coupling, and these contributions skew y_u and y_d apart and allow the generation of a nontrivial V_{CKM} . If $M_N \gg M_\phi$, it is this scale that appears in the denominator, but in our benchmark $M_N \sim M_\phi \sim \langle \Sigma \rangle$. This structure appears for free simply from the embedding of $SU(9)$ color-flavor unification into $SU(12)$ unification along with our strategy to generate Majorana neutrino masses by breaking lepton flavor.

While in previous work the intermediate group has been the flavor-unified $SU(3)^2/\mathbb{Z}_3$, here the intermediate group has undergone flavor deconstruction, but this does not affect the workings of this mechanism. In our case mixing will be generated when the bifundamental link fields Σ_{ab} of Fig. 3 spontaneously break $SU(3)^3 \rightarrow SU(3)_C$.¹⁴ The scalar potential in full generality contains many cubic and quartic terms, giving in principle enough freedom to generate the full CKM matrix by tuning the various coefficients to control the various entries of Eq. (32). The development in this theory of a mechanism which would automatically generate a SM-like structure is an important goal for future work.

¹⁴As an interesting connection we note Agrawal and Howe [87,88] considered a massless quark solution starting with $SU(3)_i^3$ and generating each generation's smallest Yukawa from instantons of that generation. In this case they generate flavor breaking using the link fields connected to the SM through dimension-five operators e.g. $H Q_i \Sigma_j^i \bar{u}^j$. It would be interesting to determine if their model can be embedded in $SU(9)$ color-flavor unification through their suggested UV completion which adds vectorlike quarks.

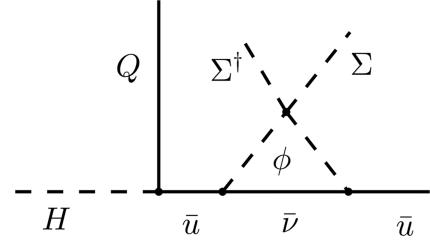


FIG. 5. The up-type Yukawas receive gluon loop corrections and additional loop-level flavor-violating corrections from ϕ^{aa} and $\bar{\nu}$ interactions which are Hermitian. The resulting Yukawas still have real determinant.

VI. BREAKING LEPTON FLAVOR TO HYPERCHARGE AND DISCRETE GAUGE THEORY

Now we wish to break lepton flavor. We will condense not only the adjoint Σ , but also the two-index symmetric ϕ . As in previous sections we will discuss the breaking in multiple stages, though these can occur at the same scale.

Step 1. The complex adjoint Σ_β^a gets a VEV which preserves solely the lepton-flavor direction $T^8 = U(1)_{L_i+L_j-2L_k} \subset SU(3)_\ell$, of the form

$$\langle \Sigma_\beta^a \rangle = \Lambda_\ell \begin{pmatrix} a_1 & c_1 & 0 \\ c_2 & a_2 & 0 \\ 0 & 0 & -(a_1 + a_2) \end{pmatrix}, \quad (33)$$

where a_i and c_i are complex and the upper 2×2 block is non-normal. Under $SU(3) \rightarrow U(1)_{T^8}$ each real adjoint decomposes into four real singlets and two complex charge 3, and among them three singlets and two complex scalar modes are eaten by the heavy gauge bosons.

Step 2. The two-index symmetric ϕ^{ab} gets a generic VEV, which necessarily breaks the two continuous flavor-independent $U(1)$ directions down to hypercharge $Y = (Q - N_c L) - 3R$. Under the lepton-flavored $U(1)_{T^8}$, the two-index symmetric has decomposed as $6 \rightarrow 2(-1) + 3(+2) + 1(-4)$, written as multiplicity (charge). By writing the most general transformation of these under $U(1)_{T^8} \times U(1)_{Q-N_c L}$, one may find that hypercharge is not the only preserved symmetry, as under the linear combination $Q - N_c L - 6T^8$, the two-index symmetric components have charges $2(0) + 3(-18) + 1(+18)$. As a result the condensation preserves

$$\mathbb{Z}_{18} \subset U(1)_{Q-N_c L-6(L_i+L_j-2L_k)}, \quad (34)$$

which we hereon refer to as $\mathbb{Z}_{18}^X \subset U(1)_X$ for convenience.

Furthermore, the global structure appearing after this breaking and color reunification is

$$\frac{G_{\text{color}} \times U(1)_Y \times \mathbb{Z}_{18}^X}{\mathbb{Z}_3 \times \mathbb{Z}_3} \quad (35)$$

with both possible diagonal $\mathbb{Z}_3$ subgroups being set equal to the identity. The $\mathbb{Z}_3$ quotient between color and hypercharge arises out of the quotient between $Q - N_c L$ and the color group, which above we saw appeared at the first step of breaking. Then since $Y = (Q - N_c L) - 3R$, we have $Y(\text{mod } 3) = Q - N_c L(\text{mod } 3)$, meaning the $\mathbb{Z}_3$ subgroup is blind to the difference and hypercharge inherits this overlap. The same too with X .

A. Proton stability and trinucleon decay

What are the consequences of this discrete symmetry? We may rewrite it as

$$Q - 9(L_i + L_j - L_k) = 3[B - 3(L_i + L_j - L_k)], \quad (36)$$

where the right-hand side, divided through by three, is the faithfully acting symmetry after confinement. As noted in [40,41], this has the remarkable consequence of forbidding proton decay. Under this infrared symmetry the baryons have charge $+1$ whereas the leptons have ± 3 depending on their generation. Thus the proton is the lightest particle with nonzero charge (mod 3) and is exactly stable. We recall that already in the Standard Model the proton is stable as a result of an exact, anomaly-free $\mathbb{Z}_{2N_g}^{B+L}$ global symmetry in the infrared, in addition to the fact that we have $N_g > 1$ generations. We see that in quark-lepton color-flavor unification, this same fact of proton stability being

guaranteed by the existence of multiple generations is upgraded into a gauge symmetry which is now respected even by quantum gravity.¹⁵

The baryon-number-violating interactions consistent with this symmetry change baryon number by three, for example at leading operator dimension $3\mathcal{N} \rightarrow \bar{\ell}_i$, where $\mathcal{N}$ is a nucleon and ℓ_i is an appropriate charged or neutral antilepton of the i th flavor. Inclusive searches for trinucleon decays have been performed, for example in germanium [89], with claimed lower bounds on the partial lifetimes of order 10^{26} years. In our model the rate is extremely suppressed to the point of unobservability, as this process comes from a dimension-15 operator when written in terms of quarks and leptons, and our symmetry-breaking dynamics occurs at quite high scales. Still it would be interesting to better understand the constraints on all possible trinucleon decay channels and the extent to which this can teach us about low-energy BSM physics which is consistent with other constraints [90].

B. Lepton masses and mixings

The loop-induced mechanism for quark flavor breaking works in the same way in the lepton sector, though here the leptonic gluons transfer flavor-breaking effects from both Σ and ϕ to the lepton Yukawas, and we get a schematic structure

$$(y_e)_{\beta}^{\alpha} \sim y_{\tau} \left(\mathbb{1}_{\beta}^{\alpha} + \frac{\alpha}{4\pi} \frac{\{\Sigma, \Sigma^{\dagger}\}_{\beta}^{\alpha} + (\phi^{\dagger}\phi)_{\beta}^{\alpha}}{\Lambda_{\ell}^2} + \dots + \frac{\alpha}{(4\pi)} \frac{\eta^{\dagger}(\Sigma^{\dagger 4})_{\beta}^{\alpha} + a^{\dagger}(\phi^{\dagger}\Sigma^2\phi)_{\beta}^{\alpha} + \dots}{\Lambda_{\ell}^4} + \dots + \text{H.c.} \right). \quad (37)$$

Here again the presence of CP -violating couplings in the scalar potential allows CP violation to transfer to the Yukawas, and we emphasize there are many terms in the $V(\Sigma, \phi)$ potential which can all contribute to the generated Yukawas.

However, ϕ also generates Majorana masses for the right-handed neutrinos, which communicates additional flavor breaking through the flavorful type-I seesaw mechanism. In concert with the large y_t -sized Dirac neutrino Yukawas, this generates Majorana masses for the active neutrinos. Since the leptons have generation-independent charge 9 under the discrete $\mathbb{Z}_{18}^X$ symmetry, the right-handed Majorana mass matrix is completely allowed. This comes from the structure

$$\mathcal{L} \supset \frac{1}{2} \lambda \phi^{\alpha\beta} \bar{\nu}_{\alpha} \bar{\nu}_{\beta}, \quad (38)$$

and we see that if all the components of $\phi^{\alpha\beta}$ condense, then we fully populate the Majorana neutrino mass matrix $M^{\alpha\beta} = \lambda \langle \phi \rangle^{\alpha\beta}$.

Together these implement the type-I seesaw. Taking a general Dirac Yukawa $y_{\nu\beta}^{\alpha}$ (since $y_t \delta_{\beta}^{\alpha}$ also gets corrections as in Fig. 6), the resulting active neutrino masses are

$$m_{\nu\alpha\beta} = v^2 (y_{\nu})_{\alpha}^{\gamma} (M^{-1})_{\gamma\delta} (y_{\nu})_{\beta}^{\delta}, \quad (39)$$

where $(M^{-1})_{\gamma\delta} M^{\delta\sigma} = \delta_{\gamma}^{\sigma}$ and $v \simeq 174$ GeV. This seesaw effect can skew the PMNS matrix into anarchy even if the lepton Dirac Yukawas have a similar structure to the quarks, and we have enough freedom in the Σ, ϕ VEVs to accommodate arbitrary masses and mixings. The model can also accommodate a large ‘‘Dirac phase,’’ which is the

¹⁵Note that such a discrete symmetry stabilizing the proton also requires only $N_g > 1$ in the generalization of this model to any $SU(N_g(N_c + 1))$ quark-lepton color-flavor unification. That is, while the noninvertible symmetry effects in the quark sector work specially because $N_g = N_c$ and in the lepton sector work specially because $N_g > 2$, the ability of this model to stabilize the proton with a discrete gauge symmetry does not provide further insight into the generation puzzle.

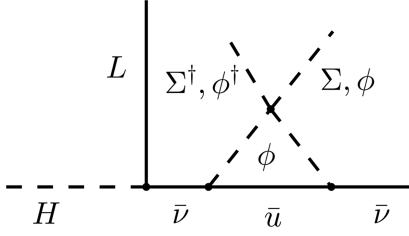


FIG. 6. The Dirac neutrino Yukawas receive leptonic gluon loop corrections and additional loop-level flavor-violating corrections from ϕ^{aa} and $\bar{u}$ interactions.

phase probed by neutrino oscillations, and arises from the appropriate Jarlskog invariant

$$J_{\text{PMNS}} \propto \text{Im det} [m_e m_e^\dagger, m_\nu m_\nu^\dagger], \quad (40)$$

where $m_e = y_e v$, having a parallel structure to the CKM phase. With both Σ, ϕ the general potential will have both of their VEVs complex, providing additional sources of CP violation compared to the quarks, meaning there is no problem attaining large CP violation in the neutrino sector if need be.

VII. MATCHING SM GAUGE COUPLINGS

Here we will consider the one-loop evolution of the gauge couplings of this model all together to ensure we can match on to the SM gauge couplings at electroweak scales. In this section alone we will consider the conventional normalizations for the gauge couplings, such that e.g. the right-handed down quark has $Y = 1/3$, $B = -1/3$, and $R = -1/2$.

We will consider the simplest case where all around the Majorana mass scale we have

$$SU(12) \times U(1)_R \rightarrow \frac{G_{\text{color}} \times U(1)_Y \times \mathbb{Z}_{18}^X}{\mathbb{Z}_9 \times \mathbb{Z}_3}, \quad (41)$$

and at this stage the color group is $(SU(3)^3 \times U(1)^2)/\mathbb{Z}_3^2$. That is, $SU(12)$ breaking, $SU(9)$ quark color-flavor breaking, and $SU(3)$ lepton-flavor breaking happen nearly simultaneously at the scale Λ_{12} , as does the breaking of $U(1)_R \times U(1)_{Q-NcL} \rightarrow U(1)_Y$. We also consider $G_{\text{color}} \rightarrow SU(3)_C$ all at once at a lower scale Λ_3 from the link fields Σ_i . It is possible for Λ_3 to also be quite high, but a low Λ_3 has the benefit that flavor-changing effects may be observable in the near term.

Our model has not added any $SU(2)_L$ -charged matter, so g_L evolves as in the SM at all energies. Below the scale Λ_{12} there are also no new states charged under hypercharge, so the hypercharge coupling also evolves as in the SM up to this scale. In the usual normalizations, we have $Y = (B - L)/2 - R$, so at this scale the hypercharge coupling matches to the $R, Q - NcL$ couplings as

$$\alpha_Y^{-1} = \alpha_R^{-1} + \frac{1}{4} \alpha_{B-L}^{-1}, \quad (42)$$

and at this scale $\alpha_{B-L} = \alpha_{12}/8$. Then we can determine the α_R needed as

$$\alpha_R^{-1} = \alpha_Y^{-1} - 2\alpha_{12}^{-1}. \quad (43)$$

Since α_R is a free parameter, this works as long as $\alpha_Y^{-1} > 2\alpha_{12}^{-1}$ at this scale, which will easily be the case. The size of α_{12} is set by the requirement of generating the correct down-type Yukawas—this can only be accurately calculated on the lattice, so we simply take $\alpha_{12}(\Lambda_{12}) = 1$ as a benchmark.

At the electroweak scale we measure (see the PDG [91] for more precision)

$$\alpha_Y^{-1} \simeq 98, \quad \alpha_L^{-1} \simeq 30, \quad \alpha_C^{-1} \simeq 8.5, \quad (44)$$

where we note we are of course not using the modified hypercharge gauge coupling by a factor $3/5$ as is done in studies of unification into $SU(5)$ or $SO(10)$.

In Fig. 7 we plot an example case with $\Lambda_{12} = 10^{15}$ GeV, a representative scale for the seesaw sector, and $\Lambda_3 = 1.4 \times 10^6$ GeV, such that we might avoid constraints on flavor-changing neutral currents. As mentioned above, we require some colored spectator fields and we assume additional $SU(12)$ -charged, scalar matter giving $\beta_{\text{spectator}} = 31/3$ to help turn around the $SU(3)_i$ gauge couplings with the same contribution to each $SU(3)_i$ so they are no longer asymptotically free between these two scales, while we assume these extra degrees of freedom decouple at Λ_3 . For fixed breaking scales, a different number of these would simply modify $\alpha_{12}(\Lambda_{12})$.

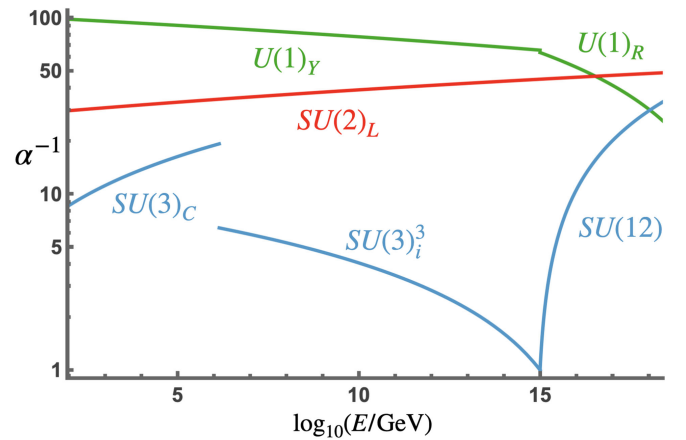


FIG. 7. The couplings as a function of scale in the simple benchmark model where most of the breaking steps occur simultaneously at $\Lambda_{12} = 10^{15}$ GeV, accounting only for one-loop running and tree-level matching at thresholds. Nontrivial matching occurs in $U(1)_Y$ emerging out of $U(1)_R$ and $SU(12)$, and at $SU(3)^3 \rightarrow SU(3)$ breaking at $\Lambda_3 \simeq 1.4 \times 10^6$ GeV.

Since the ultraviolet group is not simple, no gauge coupling unification is demanded, but it is interesting if it is achievable nonetheless as it points to the possibility of further ultraviolet unification. In this example we see that the $U(1)_R$ and $SU(2)_L$ couplings meet up around 3.3×10^{16} GeV, and their unification would be consistent with the ultraviolet completion mentioned in the introduction of $(SU(2)_R \times SU(2)_L) \rtimes \mathbb{Z}_2$. Alternatively we could take note of $U(1)_R$ and $SU(12)$ meeting at $\sim 10^{18}$ GeV and think about adding electroweak charged matter to find a full unification below the Planck scale. One possibility, as mentioned in the introduction, is to upgrade the Higgs sector to $SU(12)$ adjoint + singlet fields and perhaps to include H_u and H_d separately. These fields could help with the running as well as having useful roles to play in generating the quark masses and mixings with less tuning required. We leave for future work the possibility of gauge coupling unification in models like this.

VIII. ONE- AND TWO-FORM SYMMETRIES OF THE STANDARD MODEL

An interesting feature of the infrared of this theory is that we do not land on the familiar Standard Model *group*. Recently attention has been called to the fact that the SM has not been uniquely defined by data or theory. The so-far explored family of field theories differ in the “global structure” of the gauge group or in gauge-invariant language in the one-form global symmetry of the theory [6,10–12,92]. However, this is far from the only family of field theories which have the same local-in-field-space phenomenology of the familiar SM, and the studied possibilities do not even exhaust the possible global structures of the SM.

In particular, there are interesting versions of the SM which have additional discrete gauge theory factors. In gauge-invariant language, this leads to two-form global symmetries of the Standard Model. These factors may appear simply out of well-motivated UV completions, as we have seen from the dynamics above which have led to the SM group in this theory being

$$G_{\text{SM}} = \frac{SU(3)_C \times SU(2)_L \times U(1)_Y \times \mathbb{Z}_{18}^X}{\mathbb{Z}_3 \times \mathbb{Z}_3 \times \Gamma}, \quad (45)$$

with $\Gamma \in \{1, \mathbb{Z}_2\}$ the global structure we began with in the ultraviolet which has descended to the electroweak global structure.

If we were to instead study the Standard Model with $\mathbb{Z}_{18}^X$ gauged and no quotient between X and Y , that theory would have a $\mathbb{Z}_3^{(1)}$ electric one-form symmetry. This would correspond to the absence of particles with $Y \neq X \pmod{3}$, and resultingly that Wilson lines in such representations, despite being allowed in this theory, are not “endable.” The theory under consideration here gauges

that $\mathbb{Z}_3^{(1)}$ symmetry (discussed further below) and this forbids such particles from existing, which makes searching for such “fractionally X -onic” particles a possible low-energy probe of this model.¹⁶

To our knowledge this is the first time a particle physics model has been identified that contains a quotient shared between continuous and discrete gauge groups. The systematic construction of Standard Models with two-form symmetries, and their interplay with the possible one-form symmetries, is being considered in ongoing work. In the rest of this section we just make a few further comments on the expected physics, saving detailed analysis to be presented in the near future.

There have only been few investigations of versions of the Standard Model which have discrete gauge symmetries.¹⁷ The most studied possibility, once right-handed neutrinos are included, is a gauged $\mathbb{Z}_N$ discrete subgroup of $B-L$. The case of $N=2$ appears, for example, from $SO(10)$ models bearing out the type-I seesaw, with the neutrino modes first discussed in [96]. For larger N to our knowledge the only works considering the cosmic strings are our earlier [97,98].

When this arises from a gauged $U(1)_{B-L}$ which is spontaneously broken, for general N this bears out a charge- N Abelian Higgs theory and at low energies one arrives at a “BF theory,” which describes the universal infrared dynamics of pure $\mathbb{Z}_N$ discrete gauge theory.¹⁸ The fields consist of a one-form gauge field A_μ^X and a two-form gauge field $B_{\mu\nu}$ which is the dual of the pseudoscalar phase of the Higgs ϕ , the continuous part of which has become the longitudinal mode of A^X . They interact via the Lagrangian

$$\mathcal{L} \supset \frac{N}{4\pi} \partial_\mu A_\nu^X B_{\rho\sigma} \epsilon^{\mu\nu\rho\sigma} + J^\mu A_\mu^X, \quad (46)$$

where J^μ is the X current inherited from the theory with continuous gauge symmetry. In the farthest IR, having integrated out the charged particles and the cosmic strings, this theory has emergent one- and two-form global symmetries $\mathbb{Z}_N^{(1)} \times \mathbb{Z}_N^{(2)}$. The charged objects are Wilson lines $\exp i q \oint A^X$ (electric particle worldlines) and Wilson

¹⁶In spirit this is just the same as how fractionally electrically charged particles probe the usual Standard Model structure, although here it is much more difficult to test at a collider as it requires determining the X charge of a discovered particle. Since we have no massless X vector, this is not generically possible, but one can write down models where discovering new fields and measuring multiple couplings would allow one to reconstruct the X charge. If a particle with $Y \neq X \pmod{3}$ is discovered, this model is flatly ruled out.

¹⁷Discrete gauge symmetries have been considered more often in the context of the Minimal Supersymmetric Standard Model (MSSM), starting from [93–95].

¹⁸See e.g. [99,100] for derivation and further discussion.

surfaces $\exp ik \oint B$ (cosmic string world sheets) which interact only through topological linking.

In our theory, the condensation of $\phi^{\alpha\beta}$ in the early Universe produces cosmic strings around which both $\phi^{\alpha\beta}$ and the $U(1)_X$ gauge field A^X wind. These strings all have magnetic fluxes running through their cores. Those strings with fluxes in units of the Dirac fluxes of $U(1)_{Q-N_c L}$ and $U(1)_{L_i+L_j-2L_k}$ (taking into account the relevant quotients) are unstable to the pair production of magnetic monopoles. As a result of the nontrivial $N = 18$ charge of ϕ , there are additional $\mathbb{Z}_6$ stable cosmic strings which have further fractional magnetic fluxes that cannot be discharged by monopole-antimonopole production, often called “BF strings.”

However, in relation to the general charge- N Abelian Higgs our theory with $X = Q - 9(L_i + L_j - L_k)$ becomes more complicated in two different ways, both of which mean that the BF theory is only schematic in this case. The first complication is that the condensation of $\phi^{\alpha\beta}$ not only breaks $U(1)_X$, but also produces a large Majorana mass for right-handed neutrinos. This issue also arises in models of gauged $U(1)_{B-L}$ spontaneously broken to $\mathbb{Z}_2$ [96,101,102]. The large Majorana mass appears everywhere apart from the cores of cosmic strings, where $\langle \phi^{\alpha\beta} \rangle = 0$ can be restored and the right-handed neutrino modes can become massless. Furthermore, the topological structure of the Dirac operators in these twisted gauge field backgrounds necessitates the existence of zero-energy solutions [103,104]. Together these effects lead to localized $\bar{\nu}$ zero modes on strings, at least before further electroweak symmetry breaking. Technically, this extra coupling complicates the dualization of the longitudinal mode of ϕ to $B_{\mu\nu}$. Physically, the Fermi zero modes change the physics of the cosmic strings, allowing them to superconduct lepton-flavor currents, and the far-infrared theory must be adjusted in some manner to account for this. To our knowledge understanding the correct far-infrared description remains an important open question.

Furthermore, in the case of a BF theory the two-form symmetry is connected to the SM only through the one-form gauge field’s coupling to a SM current, which physically leads to the effect of the SM particles picking up Aharonov-Bohm phases around the cosmic strings. Such cosmic strings from flavored symmetries have received only little attention [105,106]. However as we have seen in the example here, there can also be a subtler connection between the new two-form symmetry and the dynamics of the SM. In this case the zero-form gauge group has a quotient between subgroups of color, hypercharge, and the discrete X group, which means that the one-form symmetry structure affects the two-form symmetry structure by correlating the magnetic fluxes.

For now we ignore the former issue regarding zero modes. Then we can construct the $(U(1)_Y \times \mathbb{Z}_{18}^X)/\mathbb{Z}_3$ theory by starting with the $(U(1)_Y \times \mathbb{Z}_{18}^X)$ theory and implementing the quotient by gauging the $\mathbb{Z}_3^{(1)}$ electric

one-form symmetry which acts on both the hypercharge and the discrete gauge sectors as a shift of the gauge fields $A_\mu^X \rightarrow A_\mu^X + \Lambda_\mu$, $A_\mu^Y \rightarrow A_\mu^Y - \Lambda_\mu$ with $\partial_{[\nu} \Lambda_{\mu]} = 0$. This construction can be thought of in two steps: First, we couple the theory to a background two-form gauge field $C_{\mu\nu}$ transforming as $C_{\mu\nu} \rightarrow C_{\mu\nu} + \partial_\mu \Lambda_\nu$ in such a way to make the theory formally invariant under arbitrary one-form transformations $\partial_{[\nu} \Lambda_{\mu]} \neq 0$. Then we gauge this one-form symmetry by summing over $C_{\mu\nu}$ in the path integral, which requires including discrete dynamics for $C_{\mu\nu}$ by giving it a one-derivative kinetic term in the form of a BF theory with a new one-form field E_μ :

$$\mathcal{L} = \frac{1}{4g_Y^2} (\partial_\mu A_\nu^Y + C_{\mu\nu})^2 + \frac{18}{4\pi} \left(\partial_\mu A_\nu^X - \frac{1}{2} C_{\mu\nu} \right) B_{\rho\sigma} \epsilon^{\mu\nu\rho\sigma} + \frac{3}{4\pi} \partial_\mu E_\nu C_{\rho\sigma} \epsilon^{\mu\nu\rho\sigma}. \quad (47)$$

Here the first term is the familiar kinetic term for hypercharge coupled to the new two-form gauge field $C_{\mu\nu}$ so it is invariant under the gauged one-form symmetry, the second term is the discrete kinetic term (BF term) for A_ν^X after having been made invariant under the gauged one-form symmetry and which enforces that A_ν^X gauges a $\mathbb{Z}_{18}^{(0)} \subset U(1)^{(0)}$ symmetry, and the final term is a BF term for $C_{\rho\sigma}$ itself which enforces that it gauges a $\mathbb{Z}_3^{(1)} \subset U(1)^{(1)}$ symmetry. Gauge invariance under $B_{\mu\nu} \rightarrow B_{\mu\nu} + \partial_\mu \beta_\nu$ requires that E also transforms as $E_\mu \rightarrow E_\mu + 6\beta_\mu$. As appropriate, the BF terms also enforce that the discrete gauge fields are restricted to be flat (have vanishing field strength) as a discrete field has no propagating degrees of freedom.

We will return in future work to study this low-energy theory in further detail, along with more generally the possible two-form symmetries of the Standard Model. One physical takeaway from this section is that the cosmic strings produced in this theory interact with the SM degrees of freedom not only through Aharonov-Bohm scattering, but also in a subtler way described (as yet obscurely) by Eq. (47), and it would be very interesting to understand their effects on, for example, the cosmic neutrino background. The applicable constraints on these cosmic strings depend sensitively on their interactions with the plasma and their emission spectra and require further study. A broader conceptual point that is exemplified here is that there are various interesting ways to couple a topological quantum field theory to the SM,¹⁹ and the phenomenology of these possibilities has hardly been explored.

¹⁹Note that in the second and third terms spacetime indices are contracted solely with the Levi-Civita pseudotensor, so there is manifestly no metric dependence in these Topological QFT (TQFT) factors. Note also that Juven Wang and collaborators have compiled much topological data on the SM and GUTs [19,107–114].

IX. CONCLUSION

In this work we have proposed quark-lepton color-flavor unification as a novel vista in the landscape of ultraviolet theories where microphysics gets simpler at small distances, providing a radical conceptual departure from conventional vertical theories of unification. The generation structure of the Standard Model is nontrivially incorporated into the familiar forces simply by gauging some of the approximate global symmetries of the SM fermions, providing multiple structural benefits. The SM matter is unified into few representations without requiring new fermions. The starkly disparate phenomenologies of quarks and leptons nonetheless emerge from a unified, highly symmetric ultraviolet origin. Most strikingly the proton is absolutely stable as the result of a discrete gauge symmetry that emerges naturally in the infrared, which is the only possibility for absolute proton stability in four-dimensional unified theories without additional fermions [115].

Here we have established this framework and exhibited interesting phenomenologies arising from both perturbative and nonperturbative aspects of this gauge theory, as it breaks down from the ultraviolet theory to the Standard Model phase. In the ultraviolet we started with a single Yukawa coupling shared between the up-type quarks and the neutrinos. In the quark sector, instantons generated the bottom Yukawa and resolved the strong CP problem in a flavor-symmetric massless quark solution. In the lepton sector, instantons break a noninvertible symmetry to generate the tau Yukawa and explain the hierarchy between the heaviest quarks and the heaviest leptons, which also sets up a flavored type-I seesaw mechanism that flips the charged vs neutral mass hierarchy. The symmetries have special structure because $N_g = 3$ specifically, providing one of few potential inroads toward the generation puzzle.

Intense study of the full scope of phenomenology in this framework is surely merited, and we have only begun to explore the structure. An obvious and important target is a better handle on the scalar sector, including how easily achievable the required VEVs are and how to generate a realistic flavor structure without requiring fine-tuning. As to the extra hierarchy problems, we note that the usual solutions (which seem not be realized at the weak scale for the electroweak hierarchy problem) still work fine to stabilize these flavor scales, and it would be interesting to study e.g. a supersymmetric version of this theory. We have suggested that upgrading the SM Higgs to a Two-Higgs-doublet model (2HDM) with color-flavor adjoints + singlets may be a useful direction here. Getting a handle on these aspects will reveal the most testable aspects of this theory in precision experiments and possibly at colliders, for example the related [41,116], though many more near-term signatures would be available in a variant of this theory that lowered the scale of lepton-flavor symmetry breaking.

There is also much interesting physics in the cosmology of this theory, where the various defects we have

seen appear at intermediate scales will be brought to life, especially if the Higgsing steps are more spread out than in our benchmark model. As the flavor gauge symmetries are broken, a wide variety of magnetic monopoles appear: flavor-universal $Q - N_c L$ monopoles, both with and without non-Abelian color-flavor and lepton-flavor magnetic fluxes; flavored $Q_i - Q_j$ quark family difference monopoles, which may also have flavor-specific color fluxes; and lepton family difference monopoles, all of which can annihilate at later stages when the $U(1)$ symmetries are broken and the monopoles become the end points of cosmic strings. Still other strings are produced when the infrared discrete X symmetry emerges, and these are stabilized by fractional magnetic fluxes and might superconduct lepton-flavor currents. These dynamical topological defects deserve dedicated further study and may have interesting nonperturbative effects on the SM plasma in the early universe. They will also produce a distinct spectrum of gravitational waves between the appearances, interactions, and disappearances of these varied types of nonperturbative objects, which will provide unique, testable signatures of this model in future gravitational wave observatories. Further fruitful topics in the cosmology of this model include identifying ways in which additional required BSM physics is naturally incorporated into this framework. The potential mechanisms for baryogenesis are manifold, between leptogenesis [117], the possibility of first-order flavor-breaking phase transitions and bubble wall dynamics, flavored topological defects, and the possible inclusion of axions [118]. This model also provides a top-down framework for the existence of dark matter candidates with SM flavor quantum numbers.

Finally, the infrared appearance of an additional discrete structure coupled to the Standard Model degrees of freedom further heralds a new program of studying phenomenological implications of topological quantum field theory factors. Much of the recent work on generalized symmetries in particle physics motivates that such topological degrees of freedom can have important phenomenological effects. Indeed the many dynamical topological defects which appear at high energies in this model can leave behind important, subtle effects in the infrared, as we have seen with the noninvertible chiral symmetries which depend on the structure of magnetic one-form symmetries. Understanding more fully the generalized symmetry structure at each phase of this theory may reveal additional striking phenomenology.

By considering a maximalist interpretation of the experimental data, we have found a novel alternative to the types of unification on which decades of study has been spent. In this framework the surprising stability of the proton does not portend epicycles on old theories but is rather a harbinger of the deep intertwining of the Standard Model with topological degrees of freedom.

ACKNOWLEDGMENTS

We thank Adam Martin and Fengwei Yang for helpful conversations and Maximilian Berbig and Adam Martin for comments on a draft of this manuscript. We have made use of FeynGame [119] to create figures, and throughout this work we have found LieART [120] to be a helpful resource. During revision we used ChatGPT-6 Astra for technical cross-checks, calculations for Fig. 7, and general editing;

we take full responsibility for the manuscript. This work is partially supported by the National Science Foundation under Grant No. PHY-2412701.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

-
- [1] H. Georgi and S. L. Glashow, Unity of all elementary particle forces, *Phys. Rev. Lett.* **32**, 438 (1974).
 - [2] G. Nordstrom, On the possibility of unifying the electromagnetic and the gravitational fields, *Phys. Z.* **15**, 504 (1914).
 - [3] T. Kaluza, Zum Unitätsproblem der physik, *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1921**, 966 (1921).
 - [4] O. Klein, Quantum theory and five-dimensional theory of relativity. (in German and English), *Z. Phys.* **37**, 895 (1926).
 - [5] K. Abe *et al.* (Super-Kamiokande Collaboration), Search for proton decay via $p \rightarrow e^+ \pi^0$ and $p \rightarrow \mu^+ \pi^0$ in 0.31 megaton · years exposure of the Super-Kamiokande water Cherenkov detector, *Phys. Rev. D* **95**, 012004 (2017).
 - [6] S. Koren and A. Martin, Fractionally charged particles at the energy frontier: The SM gauge group and one-form global symmetry, *SciPost Phys.* **18**, 004 (2025).
 - [7] S. Koren and A. Martin, Phenomenology of fractionally charged particles: Two reps are better than one, *Nucl. Phys. B* **1021**, 117196 (2025).
 - [8] R. Alonso, D. Dimakou, and M. West, Fractional-charge hadrons and leptons to tell the Standard Model group apart, *Phys. Lett. B* **863**, 139354 (2025).
 - [9] H.-L. Li and L.-X. Xu, Understanding the SM gauge group from SMEFT, *J. High Energy Phys.* **07** (2024) 199.
 - [10] C. Cordova, S. Hong, and L.-T. Wang, Axion domain walls, small instantons, and non-invertible symmetry breaking, *J. High Energy Phys.* **05** (2024) 325.
 - [11] M. Reece, Axion-gauge coupling quantization with a twist, *J. High Energy Phys.* **10** (2023) 116.
 - [12] Y. Choi, M. Forslund, H. T. Lam, and S.-H. Shao, Quantization of axion-gauge couplings and noninvertible higher symmetries, *Phys. Rev. Lett.* **132**, 121601 (2024).
 - [13] P. Agrawal, M. Nee, and M. Reig, Axion couplings in grand unified theories, *J. High Energy Phys.* **10** (2022) 141.
 - [14] P. Agrawal, M. Nee, and M. Reig, Axion couplings in heterotic string theory, *J. High Energy Phys.* **02** (2025) 188.
 - [15] P. Agrawal, M. Nee, and M. Reig, Axion couplings in Orbifold GUTs, *J. High Energy Phys.* **10** (2026) 032.
 - [16] J. N. Benabou, K. Fraser, M. Reig, and B. R. Safdi, String theory and grand unification suggest a submicroelectron-volt QCD axion, *Phys. Rev. D* **112**, 066003 (2025).
 - [17] M. Reig and T. Weigand, Testing the heterotic string with the axion-photon coupling, *J. High Energy Phys.* **01** (2026) 006.
 - [18] S. Koren, A note on proton stability in the Standard Model, *Universe* **8**, 308 (2022).
 - [19] J. Wang, Z. Wan, and Y.-Z. You, Proton stability: From the standard model to beyond grand unification, *Phys. Rev. D* **106**, 025016 (2022).
 - [20] J. C. Pati and A. Salam, Lepton number as the fourth color, *Phys. Rev. D* **10**, 275 (1974); *Phys. Rev. D* **11**, 703(E) (1975).
 - [21] A. De Rújula, H. Georgi, and S. Glashow, Trinification of all elementary particle forces, in *Fifth workshop on grand unification*, Vol. 7 (World Scientific, Singapore, 1984).
 - [22] H. Fritzsch and P. Minkowski, Unified interactions of leptons and hadrons, *Ann. Phys. (Amsterdam)* **93**, 193 (1975).
 - [23] H. Georgi, H. R. Quinn, and S. Weinberg, Hierarchy of interactions in unified gauge theories, *Phys. Rev. Lett.* **33**, 451 (1974).
 - [24] U. Amaldi, W. de Boer, and H. Furstenau, Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP, *Phys. Lett. B* **260**, 447 (1991).
 - [25] C. Giunti, C. W. Kim, and U. W. Lee, Running coupling constants and grand unification models, *Mod. Phys. Lett. A* **06**, 1745 (1991).
 - [26] P. Langacker and M.-x. Luo, Implications of precision electroweak experiments for M_t , ρ_0 , $\sin^2 \theta_W$ and grand unification, *Phys. Rev. D* **44**, 817 (1991).
 - [27] H. Georgi, Towards a grand unified theory of flavor, *Nucl. Phys. B* **156**, 126 (1979).
 - [28] Z. G. Berezhiani, Horizontal symmetry and quark—lepton mass spectrum: The $SU(5) \times SU(3)$ -h model, *Phys. Lett. B* **150B**, 177 (1985).
 - [29] Z. G. Berezhiani, Grand unification of fermion masses, in *Summer Institute: From Particle Physics to Cosmology* (1993).arXiv:hep-ph/9312222.
 - [30] B. C. Allanach, B. Gripaios, and J. Tooby-Smith, Semi-simple extensions of the Standard Model gauge algebra,

- Phys. Rev. D **104**, 035035 (2021); **106**, 019901(E) (2022).
- [31] J. Davighi and J. Tooby-Smith, Electroweak flavour unification, *J. High Energy Phys.* **09** (2022) 193.
 - [32] T.-K. Kuo and N. Nakagawa, The generation problem and the symplectic group, *Nucl. Phys.* **B250**, 641 (1985).
 - [33] T.-K. Kuo and N. Nakagawa, An $Sp(6)$ —Left $X U(1)$ hypercharge extension of the electroweak theory, *Phys. Rev. D* **30**, 2011 (1984).
 - [34] S. Koren, A. Martin, and P. Sesma, Proton stability in electroweak flavour unification (to be published).
 - [35] K. S. Babu, B. Bajc, and V. Susič, Trinification from E_6 symmetry breaking, *J. High Energy Phys.* **07** (2023) 011.
 - [36] C. Cordova, S. Hong, S. Koren, and K. Ohmori, Neutrino masses from generalized symmetry breaking, *Phys. Rev. X* **14**, 031033 (2024).
 - [37] C. Cordova, S. Hong, and S. Koren, Noninvertible Peccei-Quinn symmetry and the massless quark solution to the strong CP problem, *Phys. Rev. X* **15**, 031011 (2025).
 - [38] A. Delgado and S. Koren, Non-invertible Peccei-Quinn symmetry, natural 2HDM alignment, and the visible axion, *J. High Energy Phys.* **02** (2025) 178.
 - [39] G. Choi, S. Hong, and S. Koren, Global structure, non-invertible PQ symmetry, and the DFSZ domain wall problem (2026). [arXiv:2603.26867](https://arxiv.org/abs/2603.26867).
 - [40] K. S. Babu, I. Gogoladze, and K. Wang, Gauged baryon parity and nucleon stability, *Phys. Lett. B* **570**, 32 (2003).
 - [41] J. Davighi, A. Greljo, and A. E. Thomsen, Leptoquarks with exactly stable protons, *Phys. Lett. B* **833**, 137310 (2022).
 - [42] G. 't Hooft, Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking, *NATO Sci. Ser. B* **59**, 135 (1980).
 - [43] J. Fan, K. Fraser, M. Reece, and J. Stout, Axion mass from magnetic monopole loops, *Phys. Rev. Lett.* **127**, 131602 (2021).
 - [44] C. Csáki, R. Ovadia, O. Telem, J. Terning, and S. Yankielowicz, Abelian instantons and monopole scattering, *J. High Energy Phys.* **11** (2024) 165.
 - [45] I. García García, M. Kongsore, and K. Van Tilburg, Dyon loops and Abelian instantons, *J. High Energy Phys.* **05** (2026) 244.
 - [46] I. García García and E. Maderazo, Black holes and Abelian instantons, *J. High Energy Phys.* **10** (2026) 029.
 - [47] M. M. Anber, S. Hong, and M. Son, New anomalies, TQFTs, and confinement in bosonic chiral gauge theories, *J. High Energy Phys.* **02** (2022) 062.
 - [48] S. Hong, H. Kim, S. M. Lee, and D. Seo, Multi-instantons, multi-axions, and non-invertible symmetries in 4d QFT, *J. High Energy Phys.* **03** (2026) 204.
 - [49] C. Cordova and K. Ohmori, Noninvertible chiral symmetry and exponential hierarchies, *Phys. Rev. X* **13**, 011034 (2023).
 - [50] Y. Choi, H. T. Lam, and S.-H. Shao, Noninvertible global symmetries in the Standard Model, *Phys. Rev. Lett.* **129**, 161601 (2022).
 - [51] T. D. Brennan, S. Hong, and L.-T. Wang, Coupling a cosmic string to a TQFT, *J. High Energy Phys.* **03** (2024) 145.
 - [52] T. D. Brennan, A new solution to the Callan Rubakov effect, *J. High Energy Phys.* **11** (2024) 170.
 - [53] C. Córdova and S. Koren, Higher flavor symmetries in the Standard Model, *Ann. Phys. (Berlin)* **535**, 2300031 (2023).
 - [54] N. Craig and M. Kongsore, High-quality axions from higher-form symmetries in extra dimensions, *Phys. Rev. D* **111**, 015047 (2025).
 - [55] D. Sehayek and N. Craig, Generalized symmetries and emergence in axion effective field theories, [arXiv:2604.11877](https://arxiv.org/abs/2604.11877).
 - [56] D. Aloni, E. García-Valdecasas, M. Reece, and M. Suzuki, Spontaneously broken (-1) -form $U(1)$ symmetries, *SciPost Phys.* **17**, 031 (2024).
 - [57] C. Cheung, M. Derda, J.-H. Kim, V. Nevoa, I. Rothstein, and N. Shah, Generalized symmetry in dynamical gravity, *J. High Energy Phys.* **10** (2024) 007.
 - [58] M. van Beest, P. Boyle Smith, D. Delmastro, Z. Komargodski, and D. Tong, Monopoles, scattering, and generalized symmetries, *J. High Energy Phys.* **03** (2025) 014.
 - [59] M. van Beest, P. Boyle Smith, D. Delmastro, R. Mouland, and D. Tong, Fermion-monopole scattering in the Standard Model, *J. High Energy Phys.* **08** (2024) 004.
 - [60] J. Davighi, B. Gripaios, and N. Lohitsiri, Global anomalies in the Standard Model(s) and Beyond, *J. High Energy Phys.* **07** (2020) 232.
 - [61] J. Davighi, A. Greljo, and N. Selimovic, Topological portal to the dark sector, *Phys. Rev. Lett.* **134**, 111804 (2025).
 - [62] J. Davighi and N. Lohitsiri, WZW terms without anomalies: Generalised symmetries in chiral Lagrangians, *SciPost Phys.* **17**, 168 (2024).
 - [63] J. Davighi, S. Moldovsky, H. Murayama, C. Scherb, and N. Selimovic, Topological freeze-out by semi-annihilation, *J. High Energy Phys.* **02** (2026) 133.
 - [64] S. Chen, A. Cherman, G. Choi, and M. Neuzil, Cheshire θ terms, Aharonov-Bohm effects, and axions, *Phys. Rev. D* **112**, 045012 (2025).
 - [65] C. P. Burgess and F. Quevedo, Dual is different: EFTs, axions and nonpropagating form fields in cosmology, [arXiv:2509.11340](https://arxiv.org/abs/2509.11340).
 - [66] M. Suzuki and R. Yokokura, Lazarides-Shafi axion models as Dijkgraaf-Witten theories, *Phys. Rev. D* **113**, L091902 (2026).
 - [67] T. D. Brennan, L.-T. Wang, and H. Xiao, Monopole Catalyzed Baryogenesis with a θ angle, [arXiv:2412.14239](https://arxiv.org/abs/2412.14239).
 - [68] N. Kaloper, General relativity on the multiverse and nature's hierarchies, *Phys. Rev. D* **106**, 044023 (2022).
 - [69] E. García-Valdecasas, M. Reece, and M. Suzuki, Monopole breaking of Chern-Weil symmetries, *SciPost Phys.* **18**, 162 (2025).
 - [70] S. Hong, S. Kim, and S. Koren, Non-invertible symmetry, dirac neutrino masses, and deconstructing B-L (to be published).
 - [71] T. W. B. Kibble, Topology of cosmic domains and strings, *J. Phys. A* **9**, 1387 (1976).
 - [72] C. G. Callan, Jr., Dyon-fermion dynamics, *Phys. Rev. D* **26**, 2058 (1982).

- [73] V. A. Rubakov, Adler-Bell-Jackiw anomaly and fermion number breaking in the presence of a magnetic monopole, *Nucl. Phys.* **B203**, 311 (1982).
- [74] C. Csáki, R. T. D'Agnolo, E. Kuflik, and M. Ruhdorfer, Instanton NDA and applications to axion models, *J. High Energy Phys.* **04** (2024) 074.
- [75] P. Sesma, A functional treatment of small instanton-induced axion potentials, *J. High Energy Phys.* **03** (2025) 026.
- [76] M. Bordone, C. Cornella, J. Fuentes-Martin, and G. Isidori, A three-site gauge model for flavor hierarchies and flavor anomalies, *Phys. Lett. B* **779**, 317 (2018).
- [77] J. Davighi and G. Isidori, Non-universal gauge interactions addressing the inescapable link between Higgs and flavour, *J. High Energy Phys.* **07** (2023) 147.
- [78] A. S. Schwarz, Field theories with no local conservation of the electric charge, *Nucl. Phys. B* **208**, 141 (1982).
- [79] M. G. Alford, K. Benson, S. R. Coleman, J. March-Russell, and F. Wilczek, The interactions and excitations of nonabelian vortices, *Phys. Rev. Lett.* **64**, 1632 (1990); *Phys. Rev. Lett.* **65**, 668(E) (1990).
- [80] M. Bucher, H.-K. Lo, and J. Preskill, Topological approach to Alice electrodynamics, *Nucl. Phys.* **B386**, 3 (1992).
- [81] G. R. Dvali and G. Senjanovic, Semiglobal Alice strings, *Phys. Rev. D* **51**, 7148 (1995).
- [82] R. Sato, F. Takahashi, and M. Yamada, Unified origin of axion and monopole dark matter, and solution to the domain-wall problem, *Phys. Rev. D* **98**, 043535 (2018).
- [83] C. Chatterjee, T. Higaki, and M. Nitta, Note on a solution to domain wall problem with the Lazarides-Shafi mechanism in axion dark matter models, *Phys. Rev. D* **101**, 075026 (2020).
- [84] B. Heidenreich, J. McNamara, M. Montero, M. Reece, T. Rudelius, and I. Valenzuela, Non-invertible global symmetries and completeness of the spectrum, *J. High Energy Phys.* **09** (2021) 203.
- [85] J. McNamara and M. Reece, Reflections on parity breaking, [arXiv:2212.00039](https://arxiv.org/abs/2212.00039).
- [86] G. Arias-Tamargo and D. Rodriguez-Gomez, Non-invertible symmetries from discrete gauging and completeness of the spectrum, *J. High Energy Phys.* **04** (2023) 093.
- [87] P. Agrawal and K. Howe, A flavorful factoring of the strong CP problem, *J. High Energy Phys.* **12** (2018) 035.
- [88] P. Agrawal and K. Howe, Factoring the strong CP problem, *J. High Energy Phys.* **12** (2018) 029.
- [89] M. Agostini *et al.* (GERDA Collaboration), Search for trinucleon decays of ^{76}Ge in GERDA, *Eur. Phys. J. C* **83**, 778 (2023).
- [90] S. Koren and A. Martin, Alchemy at the energy and underground Frontiers (to be published).
- [91] S. Navas *et al.* (Particle Data Group), Review of particle physics, *Phys. Rev. D* **110**, 030001 (2024).
- [92] D. Tong, Line operators in the Standard Model, *J. High Energy Phys.* **07** (2017) 104.
- [93] M. C. Bento, L. J. Hall, and G. G. Ross, Generalized matter parities from the superstring, *Nucl. Phys.* **B292**, 400 (1987).
- [94] L. E. Ibanez and G. G. Ross, Discrete gauge symmetry anomalies, *Phys. Lett. B* **260**, 291 (1991).
- [95] L. E. Ibanez and G. G. Ross, Discrete gauge symmetries and the origin of baryon and lepton number conservation in supersymmetric versions of the standard model, *Nucl. Phys.* **B368**, 3 (1992).
- [96] R. Jeannerot, A New mechanism for leptogenesis, *Phys. Rev. Lett.* **77**, 3292 (1996).
- [97] N. Craig, I. Garcia Garcia, and S. Koren, Discrete gauge symmetries and the weak gravity conjecture, *J. High Energy Phys.* **05** (2019) 140.
- [98] S. Koren, Cosmological lithium solution from discrete gauged B-L, *Phys. Rev. Lett.* **131**, 091003 (2023).
- [99] T. D. Brennan and S. Hong, Introduction to generalized global symmetries in QFT and particle physics, [arXiv:2306.00912](https://arxiv.org/abs/2306.00912).
- [100] T. Banks and N. Seiberg, Symmetries and strings in field theory and gravity, *Phys. Rev. D* **83**, 084019 (2011).
- [101] N. Sahu, P. Bhattacharjee, and U. A. Yajnik, B—L cosmic strings and baryogenesis, *Phys. Rev. D* **70**, 083534 (2004).
- [102] P.-H. Gu and H. Mao, Constraint on B—L cosmic string from leptogenesis with degenerate neutrinos, *Phys. Lett. B* **619**, 226 (2005).
- [103] C. G. Callan, Jr. and J. A. Harvey, Anomalies and fermion zero modes on strings and domain walls, *Nucl. Phys.* **B250**, 427 (1985).
- [104] E. Witten, Superconducting strings, *Nucl. Phys.* **B249**, 557 (1985).
- [105] J. L. Chkareuli, Flavored cosmic strings and monopoles, *Phys. Lett. B* **272**, 207 (1991).
- [106] T. M. Bibilashvili and G. R. Dvali, Superconducting cosmic strings in models with spontaneously broken family symmetry, *Phys. Lett. B* **248**, 259 (1990).
- [107] Z. Wan and J. Wang, Beyond standard models and grand unifications: Anomalies, topological terms, and dynamical constraints via cobordisms, *J. High Energy Phys.* **07** (2020) 062.
- [108] J. Wang, Anomaly and cobordism constraints beyond the Standard Model: Topological force, [arXiv:2006.16996](https://arxiv.org/abs/2006.16996).
- [109] J. Wang, Anomaly and cobordism constraints beyond grand unification: Energy hierarchy, [arXiv:2008.06499](https://arxiv.org/abs/2008.06499).
- [110] J. Wang, Ultra unification, *Phys. Rev. D* **103**, 105024 (2021).
- [111] J. Wang and Y.-Z. You, Gauge enhanced quantum criticality beyond the Standard Model, *Phys. Rev. D* **106**, 025013 (2022).
- [112] J. Wang, Z. Wan, and Y.-Z. You, Cobordism and deformation class of the Standard Model, *Phys. Rev. D* **106**, L041701 (2022).
- [113] J. Wang and Y.-Z. You, Gauge enhanced quantum criticality between grand unifications: Categorical higher symmetry retraction, [arXiv:2111.10369](https://arxiv.org/abs/2111.10369).
- [114] P. Putrov and J. Wang, Categorical symmetry of the standard model from gravitational anomaly, *Phys. Rev. D* **110**, 125028 (2024).
- [115] S. Koren, A. Martin, and P. Sesma, Completing our understanding of proton decay in minimal unified theories (to be published).

- [116] A. Greljo, P. Stangl, A. E. Thomsen, and J. Zupan, On $(g - 2)_\mu$ from gauged $U(1)_X$, [J. High Energy Phys.](#) **07** (2022) 098.
- [117] A. Greljo, X. Ponce Díaz, and A. E. Thomsen, Insights on the cosmic origin of matter from proton stability, [J. Cosmol. Astropart. Phys.](#) **11** (2025) 043.
- [118] M. Berbig and S. Koren, Rise of the lepto-axion (to be published).
- [119] L. Bündgen, R. V. Harlander, S. Y. Klein, and M. C. Schaaf, FeynGame 3.0, [Comput. Phys. Commun.](#) **314**, 109662 (2025).
- [120] R. Feger, T. W. Kephart, and R. J. Saskowski, LieART 2.0—A Mathematica application for Lie algebras and representation theory, [Comput. Phys. Commun.](#) **257**, 107490 (2020).