Standard model effective field theory up to mass dimension 12

R. V. Harlander, T. Kempkens, and M. C. Schaaf

Institute for Theoretical Particle Physics and Cosmology, RWTH Aachen University, 52056 Aachen, Germany

(Received 13 June 2023; accepted 14 August 2023; published 21 September 2023)

We present a complete and nonredundant basis of effective operators for the Standard Model effective field theory up to mass dimension 12 with three generations of fermions. We also include operators coupling to gravity via the Weyl tensor. The results are obtained by implementing the algorithm of Li *et al.* and are provided in the form of Supplemental Material.

DOI: 10.1103/PhysRevD.108.055020

I. INTRODUCTION

Owing to the lack of discoveries of particles beyond the Standard Model (SM), one of the main research directions over the past decade has been effective field theories (EFTs). Given a renormalizable low-energy theory $\mathcal{L}$ like the SM, EFTs parametrize possible new heavy physics beyond that theory in terms of operators of mass dimension larger than four. The Wilson coefficients compensate for the surplus mass dimension through inverse powers of a generic new-physics scale Λ which is assumed to be much larger than the energy at which this physics is probed. The operators are composed of the fields of $\mathcal{L}$ and obey the same gauge symmetries as $\mathcal{L}$.

In order to uniquely constrain the Wilson coefficients by experiment, a number of possible redundancies among the operators need to be taken into account. On the one hand, these arise because operators which vanish by equations of motion (EoMs) can be eliminated through a redefinition of the field variables in the path integral [1,2]. Moreover, operators differing by a total derivative can be identified. Finally, operators which are related by algebraic identities associated with the group structure of the underlying internal or Lorentz symmetry can have dependencies.

The EFT constructed from the SM is usually referred to as Standard Model effective field theory (SMEFT). At mass dimension five, it contains only the lepton-number violating Weinberg operator (up to generation multiplicities). Currently, the nonredundant basis of operators is known up to mass dimension nine [3–10]. Beyond that, only the *number* of independent operators that form a basis[1] is currently known, albeit broken up into operators with a specific field content. This number can be obtained from a suitable Hilbert series [11,12] or by explicitly considering the field quantum numbers and permutation symmetries [13]. The fact that it grows roughly exponentially with the mass dimension implies that, beyond a certain order in $1/\Lambda$, it is necessary to give the task of explicitly constructing the basis to a computer.

In fact, in Refs. [3,4], partly building on concepts developed in Ref. [14], a fully algorithmic approach was used to determine the SMEFT basis up to mass dimensions eight and nine. Utilizing various functions of Refs. [15,16],[2] its implementation has been published as a MATHEMATICA package [17], but its application currently appears to be restricted to mass dimensions equal to or less than nine.

An automated approach to constructing EFT bases is also desirable from the point of view that as-of-yet undiscovered light particles might still exist which could couple to the SM via effective operators. Examples for this are sterile neutrinos [18] or axionlike particles [19]. In addition, one should expect that the low-energy limit of some theory that includes both the SM interactions as well as gravity will be described by an effective Lagrangian which extends SMEFT by gravitational fields, resulting in general relativity $\oplus$ SMEFT (GRSMEFT) [20].

Finally, an efficient automated approach to constructing EFTs will allow one to study operator bases at higher mass dimension, and thus operators with a richer structure. For example, it is only starting at mass dimension ten that $B-L$ can be violated at $\Delta(B-L)=4$, where B and L are the baryon and lepton number [21,22]. Furthermore, studying higher-dimensional operators, or specific subsets

[1]From now on, a *basis* shall always denote a complete, nonredundant basis, unless stated otherwise.

[2]Thanks to R. Fonseca and the authors of Ref. [17] for clarifications on this issue.

thereof, may give insight on a possible all-order structure of such operators; see Ref. [23], for example.

In this paper, we report on a reimplementation of the algorithm of Refs. [3,4,17] into SageMath [24] which is a free open-source mathematics software system, and the symbolic manipulation system FORM [25,26], both licensed under the GNU General Public License (GPL) [27]. Using this code, named AutoEFT, we have evaluated the (on-shell) SMEFT basis as well as the GRSMEFT basis up to mass dimension 12.

A description of AutoEFT, its usage, and the program itself will be deferred to a separate publication.[3] The main purpose of the present paper is to make the SMEFT and GRSMEFT operators of mass dimensions 10, 11, and 12 accessible to the public.[4] However, the number of operators is too large for listing them in this paper. Instead, we provide them in electronic form, using a notation which is both rather compact, but also straightforward to interpret.

The remainder of the paper is structured as follows. In Sec. II we briefly review the algorithm of Refs. [3,4,17] on which the results of this paper are based. This includes the treatment of scalar fields, chiral fermions, vector bosons, as well as gravitons. The majority of Sec. III is devoted to describing the format in which we encode the operators. The results themselves are provided in the form of Supplemental Material [28] which accompanies this paper. Section IV contains our conclusions and an outlook.

II. CONSTRUCTING EFT OPERATOR BASES

In this section, the concepts introduced in Refs. [3,4] for the systematic construction of an EFT operator basis are briefly reviewed. Contrary to the formal representation in Refs. [3,4], only the central ideas are illustrated.

A. Field representations

In a standard construction of SMEFT operators, as it was done for the Warsaw basis at mass dimension six [6], for example, the fundamental building blocks are taken to be complex scalars ϕ, Dirac spinors $\Psi_{\mathrm{L/R}}$, field strength tensors $F_{\mu\nu}$, and the derivative[5] D_μ. It turns out that for the operator construction at higher mass dimensions, it is more convenient to characterize the fields and derivatives by the irreducible representation (j_l, j_r) in which they transform under the Lorentz group $\mathrm{SL}(2,\mathbb{C}) \simeq \mathrm{SU}(2)_l \times \mathrm{SU}(2)_r$. This corresponds to adopting a chiral basis, where the SM fields are scalars, two-component Weyl spinors, or chiral field-strength tensors:

$$\begin{aligned} &\phi \in (0,0), \\ &\psi_\alpha \in (1/2,0), \qquad \psi^{\dagger\dot{\alpha}} \in (0,1/2), \\ &F_{\mathrm{L}\alpha\beta} \in (1,0), \qquad F_{\mathrm{R}}^{\dot{\alpha}\dot{\beta}} \in (0,1). \end{aligned} \tag{1}$$

Here, $\alpha, \beta, \ldots$ and $\dot{\alpha}, \dot{\beta}, \ldots$ denote indices of the fundamental representation of $\mathrm{SU}(2)_l$ and $\mathrm{SU}(2)_r$, respectively. Note that each of the fields in (1) has a unique helicity value

$$h = j_r - j_l. \tag{2}$$

The derivative transforms nontrivially under both $\mathrm{SU}(2)_l$ and $\mathrm{SU}(2)_r$,

$$D_\alpha^{\dot{\alpha}} \in (1/2, 1/2). \tag{3}$$

Thus, the derivative of a field has the same helicity as the field itself. This definition of the fields is equivalent to the conventional representation of the SM fields. The translation between the two notations is given by

$$\begin{aligned} &F^{\mu\nu} = \frac{i}{4}(F_{\mathrm{L}}^{\alpha\beta}\sigma^{\mu\nu}_{\alpha\beta} - F_{\mathrm{R}}^{\dot{\alpha}\dot{\beta}}\bar{\sigma}^{\mu\nu}_{\dot{\alpha}\dot{\beta}}),\ D_\mu = -\frac{1}{2}D_\alpha^{\dot{\alpha}}(\sigma_\mu)^\alpha_{\dot{\alpha}} \\ &\Psi_{\mathrm{L}} = \begin{pmatrix}\psi_\alpha \\ 0\end{pmatrix},\ \bar{\Psi}_{\mathrm{L}} = (0, \psi^\dagger_{\dot{\alpha}}),\ \Psi_{\mathrm{R}} = \begin{pmatrix}0 \\ \psi_{\mathbb{C}}^{\dagger\dot{\alpha}}\end{pmatrix},\ \bar{\Psi}_{\mathrm{R}} = (\psi^\alpha_{\mathbb{C}}, 0), \\ &F_{\mathrm{L}\alpha\beta} = \frac{i}{2}F_{\mu\nu}\sigma^{\mu\nu}_{\alpha\beta},\ F_{\mathrm{R}}^{\dot{\alpha}\dot{\beta}} = -\frac{i}{2}F^{\mu\nu}\bar{\sigma}^{\dot{\alpha}\dot{\beta}}_{\mu\nu},\ D_\alpha^{\dot{\alpha}} = D_\mu(\sigma^\mu)_\alpha^{\dot{\alpha}}, \end{aligned} \tag{4}$$

where $\psi_{\mathbb{C}}$ denotes a charge conjugated spinor, and the σ matrices are given by

$$\begin{aligned} &\sigma^{\mu\nu} = \frac{i}{2}(\sigma^\mu\bar{\sigma}^\nu - \sigma^\nu\bar{\sigma}^\mu), \qquad \bar{\sigma}_{\mu\nu} = \frac{i}{2}(\bar{\sigma}_\mu\sigma_\nu - \bar{\sigma}_\nu\sigma_\mu), \\ &\sigma^\mu_{\alpha\dot{\alpha}} = (I, \vec{\sigma}), \qquad \bar{\sigma}^{\mu\dot{\alpha}\alpha} = (I, -\vec{\sigma}). \end{aligned} \tag{5}$$

Here, I is the 2×2 identity matrix and $\vec{\sigma} = (\sigma^1, \sigma^2, \sigma^3)$ denotes the Pauli matrices.

B. Construction of the operators

The operators can be classified into *families*,[6] characterized by the tuple

$$(n_{F_{\mathrm{L}}}, n_\psi, n_\phi, n_{\psi^\dagger}, n_{F_{\mathrm{R}}}; n_D), \tag{6}$$

where n_Φ equals the number of fields with helicity h_Φ, and n_D is the number of derivatives. It is further useful to define

$$n_l = n_{F_{\mathrm{L}}} + \frac{1}{2}n_\psi + \frac{1}{2}n_D, \quad n_r = n_{F_{\mathrm{R}}} + \frac{1}{2}n_{\psi^\dagger} + \frac{1}{2}n_D, \tag{7}$$

which correspond to the sum of $j_{l/r}$ of each field and derivative as defined in (1). Lorentz invariance and a given

[3] A β version of the code is accessible upon request from the authors.

[4] The lower-dimensional operators will be provided as well.

[5] Throughout this paper, *derivative* denotes the gauge covariant derivative, unless stated otherwise.

[6] In Ref. [3] the term *subclass* is used instead.

mass dimension d of the operators put constraints on the n_i such as

$$N + n_l + n_r = d, \tag{8}$$

where N is the total number of fields, and thus restrict the set of allowed families.

Since Hermitian conjugation exchanges the two SU(2) representations of the Lorentz group, taking the conjugate of the operators of a particular family generates the *conjugate family*

$$(n_{F_\mathrm{L}}, n_\psi, n_\phi, n_{\psi^\dagger}, n_{F_\mathrm{R}}; n_D)^\dagger \equiv (n_{F_\mathrm{R}}, n_{\psi^\dagger}, n_\phi, n_\psi, n_{F_\mathrm{L}}; n_D). \tag{9}$$

Hence, one can identify *real* families that satisfy

$$(n_{F_\mathrm{L}}, n_\psi, n_\phi, n_{\psi^\dagger}, n_{F_\mathrm{R}}; n_D)^\dagger = (n_{F_\mathrm{L}}, n_\psi, n_\phi, n_{\psi^\dagger}, n_{F_\mathrm{R}}; n_D). \tag{10}$$

All operators in a real family are either Hermitian, or their conjugate operator is part of the same family. The remaining families are all *complex*, such that any operator features a distinct Hermitian conjugate version, which is part of the conjugate family.

All operators in a specific family can be further characterized by their *type*, which corresponds to a specific multiset of fields from a given model. Types of the same field content are identified by ordering the fields by increasing helicity and sorting fields of the same helicity alphanumerically. For example, the dimension-five Weinberg operator, consisting of two lepton doublets and two Higgs doublets, belongs to the family $(0,2,2,0,0;0)$ and the type L^2H^2.

C. Lorentz structure

In this section, only the Lorentz structure of an operator will be considered, while all internal symmetry and generation indices of the fields will be neglected. All operators of a particular family can thus be identified for the purpose of this discussion.

Since any operator $\mathcal{O}$ of the EFT will be constructed from the objects of (1) and (3), it will be of the form

$$\mathcal{O}^{\mathrm{Lorentz}} = (T^{\mathrm{Lorentz}})^{\boldsymbol{\alpha}_1\ldots\boldsymbol{\alpha}_N}_{\dot{\boldsymbol{\alpha}}_1\ldots\dot{\boldsymbol{\alpha}}_N} \prod_{i=1}^{N} (D^{n_i}\Phi_i)^{\dot{\boldsymbol{\alpha}}_i}_{\boldsymbol{\alpha}_i}, \tag{11}$$

where $\boldsymbol{\alpha}_i = (\alpha_i^{(1)}, \ldots, \alpha_i^{(m_i)})$ and $\dot{\boldsymbol{\alpha}}_i = (\dot{\alpha}_i^{(1)}, \ldots, \dot{\alpha}_i^{(\tilde{m}_i)})$ are multi-indices with $m_i = n_i + 2j_{l,i}$ and $\tilde{m}_i = n_i + 2j_{r,i}$, where n_i is the number of derivatives acting on the field Φ_i, and $(j_{l,i}, j_{r,i})$ defines its representation according to (1). The indices $\alpha_i^{(1)}, \ldots \alpha_i^{(n_i)}$ and $\dot{\alpha}_i^{(1)}, \ldots, \dot{\alpha}_i^{(n_i)}$ are to be associated with the derivatives acting on Φ_i, while the remaining indices are associated with the field itself.

D. Redundancies

One of the most challenging aspects when constructing EFTs is the elimination of redundancies, i.e., operators that are related to other operators by certain identities. As stated in Ref. [3], for the chiral convention of the fields, the structure of T^{Lorentz} must be a polynomial in $\epsilon^{\alpha\beta}$ and $\epsilon_{\dot{\alpha}\dot{\beta}}$, where ϵ is totally antisymmetric, and $\epsilon^{12} = \epsilon_{21} = \epsilon^{\dot{1}\dot{2}} = \epsilon_{\dot{2}\dot{1}} = 1$. This already eliminates all redundancies from Fierz identities (up to Schouten identities). Furthermore, redundancies due to EoMs or the identity $i[D_\mu, D_\nu] = \sum F_{\mu\nu}$, where the sum is over all gauge groups, can be eliminated by substituting

$$(D^{n_i}\Phi_i)^{\dot{\boldsymbol{\alpha}}_i}_{\boldsymbol{\alpha}_i} \to (D^{n_i}\Phi_i)^{(\dot{\boldsymbol{\alpha}}_i)}_{(\boldsymbol{\alpha}_i)} \tag{12}$$

in Eq. (11), where $(\boldsymbol{\alpha}_i) = (\alpha_i^{(1)}, \ldots, \alpha_i^{(m_i)})_{\mathrm{sym}}$ and $(\dot{\boldsymbol{\alpha}}_i) = (\dot{\alpha}_i^{(1)}, \ldots, \dot{\alpha}_i^{(\tilde{m}_i)})_{\mathrm{sym}}$ denote totally symmetric multi-indices. In the following, we refer to the combined object on the rhs of (12) as a *building block*.

The remaining redundancies to be considered are due to integration by parts and Schouten identities. These can be avoided by introducing an auxiliary $\mathrm{SU}(N)$ group, where N equals the total number of fields [cf. Eq. (8)]. Under this group, the (dotted) spinor indices of T^{Lorentz} transform as the (anti-)fundamental representation, i.e.,

$$\boldsymbol{\alpha}_i \to \sum_{j=1}^{N} U_{ij}\boldsymbol{\alpha}_j, \qquad \dot{\boldsymbol{\alpha}}_i \to \sum_{j=1}^{N} U^\dagger_{ij}\dot{\boldsymbol{\alpha}}_j, \tag{13}$$

with $\boldsymbol{\alpha}_i$, $\dot{\boldsymbol{\alpha}}_i$ defined in Eq. (11), and U $(U^\dagger)$ group elements of $\mathrm{SU}(N)$ in the (anti-)fundamental representation. For a given family, T^{Lorentz} contains n_l ϵ tensors with undotted and n_r ϵ tensors with dotted indices. Consequently, T^{Lorentz} must transform in the representation

$$\mathcal{R} = \overline{\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}}^{\,\otimes n_r} \otimes \begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}^{\,\otimes n_l} \tag{14}$$

under the auxiliary $\mathrm{SU}(N)$ group. According to the Littlewood-Richardson (LR) rule, the two factors can be decomposed as

$$\mathcal{R} = \Bigg(\overbrace{\overline{\begin{array}{|c|c|c|}\hline & \cdots & \\ \hline & \cdots & \\ \hline \end{array}}}^{n_r} \oplus \;\ldots\Bigg) \otimes \Bigg(\overbrace{\begin{array}{|c|c|c|}\hline & \cdots & \\ \hline & \cdots & \\ \hline \end{array}}^{n_l} \oplus \;\ldots\Bigg), \tag{15}$$

where dropping all diagrams represented by the dots eliminates the Schouten identities. The further decomposition of Eq. (15) into irreducible representations results in a Young diagram of shape

$$\lambda = N-2\left\{\begin{array}{l}\overbrace{\square\cdots\square}^{n_r}\overbrace{\square\cdots\square}^{n_l}\\ \square\cdots\square\,\square\cdots\square\\ \vdots\qquad\vdots\\ \square\cdots\square\end{array}\right. , \qquad (16)$$

plus diagrams that contain at least one column with $N-1$ entries which will be discussed below. A basis of tensors that transform under this irreducible representation can be constructed from semistandard Young tableaux (SSYTx) of shape λ and content[7] $\mu = [n_r - 2h_1, \ldots, n_r - 2h_N]$ [cf. Eqs. (2), (7), and (8)] as follows: Replace the first n_r columns through the SU(N) relation (no summation implied)

$$\begin{array}{c}\boxed{a}\\ \boxed{b}\\ \vdots\\ \boxed{x}\end{array} = \mathcal{E}^{ab\ldots xyz}\,\overline{\begin{array}{c}\boxed{y}\\ \boxed{z}\end{array}}, \qquad (17)$$

where $\mathcal{E}$ denotes the N-dimensional Levi-Civita symbol, and $a, b, \ldots, x, y, z$ is some permutation of $1, 2, \ldots, N$. Subsequently identify each column with an ϵ tensor according to

$$\overline{\begin{array}{c}\boxed{i}\\ \boxed{j}\end{array}} \sim \epsilon_{\dot{\alpha}_i\dot{\alpha}_j} \qquad \text{and} \qquad \begin{array}{c}\boxed{i}\\ \boxed{j}\end{array} \sim \epsilon^{\alpha_i\alpha_j}. \qquad (18)$$

Since the set of all SSYTx forms a basis in the vector space of a particular representation, performing these steps for every such SSYTx leads to a set of Lorentz tensors T^{Lorentz} which is nonredundant. To see that it is also complete, consider a Young diagram other than Eq. (16) in the decomposition of Eq. (15). As pointed out above, it contains at least one column of length $N-1$. Such columns correspond to tensors proportional to $\sum_{j=1}^{N} \epsilon^{\alpha_i\alpha_j}\epsilon_{\dot{\alpha}_j\dot{\alpha}_k}$. The index pair $(\alpha_j, \dot{\alpha}_j)$ implies a derivative acting on the jth field, so that the sum corresponds to a total derivative.

E. Internal symmetries

After identifying all independent Lorentz structures, one can continue in a similar manner with the remaining symmetries of the underlying low-energy theory. In the following, the expression *internal symmetry* refers to a global or a local U(1) or SU(n) symmetry. In analogy to Sec. II. C, all Lorentz and generation indices will be neglected in this section, and only the transformation properties of the fields under the internal symmetry group are considered.

Concerning Abelian internal symmetries, it is required that the total charge of an operator under this symmetry vanishes. In the SM, for example, this is achieved by only considering combinations of fields such that the sum of their hypercharges equals zero, hence forming a U(1) invariant operator. However, one may also allow for the breaking of a certain U(1) symmetry, specified by the amount the associated charge of the operators may deviate from zero. This can be useful, for example, if one is only interested in operators that violate baryon number conservation up to a certain degree ΔB.

For each non-Abelian symmetry, the *modified LR rule* [3] is used to construct all sets of independent tensors. To apply this method, the fields of the low-energy theory must be characterized in terms of fundamental SU(n) indices only. Consider, for example, the non-Abelian part of the SM, i.e., SU(3) ⊗ SU(2). In terms of Young diagrams, the fundamental representations are identified as $\mathbf{3} = \square$ for SU(3), and $\mathbf{2} = \square$ for SU(2). For the SM fields that transform under the fundamental representations, the symmetry of their indices can be identified with the states

$$Q_{a\,i} \sim \boxed{a} \otimes \boxed{i}, \qquad q_a \sim \boxed{a}, \qquad L_i \sim \boxed{i}, \quad \text{and} \qquad H_i \sim \boxed{i}, \qquad (19)$$

where Q and L denote a left-handed quark and lepton doublet, respectively; q the right-handed quark fields; and H the Higgs doublet. Here, $a, b, c, \ldots$ and $i, j, k, \ldots$ denote fundamental indices of SU(3) and SU(2), respectively.

The fields that do not transform under the fundamental representation of the internal symmetry group can nevertheless be written in terms of quantities with fundamental indices only. In particular, for the SM,

$$\begin{aligned} G_{abc} &= \epsilon_{acd}(\lambda^A)^d{}_b G^A, \quad W_{ij} = \epsilon_{jk}(\tau^I)^k{}_i W^I, \\ Q^\dagger_{abi} &= \epsilon_{abc}\epsilon_{ij}(Q^\dagger)^{cj}, \quad q^\dagger_{ab} = \epsilon_{abc}(q^\dagger)^c, \\ L^\dagger_i &= \epsilon_{ij}(L^\dagger)^j, \quad H^\dagger_i = \epsilon_{ij}(H^\dagger)^j, \end{aligned} \qquad (20)$$

where A and I are adjoint indices of SU(3) and SU(2). The λ^A are the Gell-Mann matrices for SU(3), and the τ^I are the SU(2) Pauli matrices.

Considering the Young diagrams for the adjoint representation of SU(3) and SU(2), given by $\mathbf{8} = \begin{array}{l}\square\square\\ \square\end{array}$ and $\mathbf{3} = \square\square$, one can identify the symmetry of the fundamental symmetry group indices of G_{abc} and W_{ij} with the states

[7] A Young tableau with content (or weight) $\mu = [i, j, k, \ldots]$ has i entries of the number 1, j entries of the number 2, k entries of the number 3, and so on. This requirement ensures that the resulting tensors have the correct indices corresponding to the fields specified in the family. A Young tableau is called semistandard if the entries weakly increase along each row and strictly increase down each column.

$$G_{abc} \sim \begin{array}{|c|c|}\hline a & b \\ \hline c \\ \cline{1-1}\end{array} \qquad \text{and} \qquad W_{ij} \sim \begin{array}{|c|c|}\hline i & j \\ \hline\end{array}, \tag{21}$$

meaning that G_{abc} is symmetrized in $a \leftrightarrow b$ and subsequently antisymmetrized in $a \leftrightarrow c$, while W_{ij} is totally symmetric in $i \leftrightarrow j$. Equivalently, the antifundamental representations of SU(3) and SU(2) are given by $\bar{\mathbf{3}} = \overline{\square} = \begin{array}{|c|}\hline \\ \hline \\ \hline\end{array}$ and $\bar{\mathbf{2}} = \overline{\square} = \square$, respectively. The symmetry of the fundamental symmetry group indices can be identified with

$$Q^{\dagger}_{ab\,i} \sim \begin{array}{|c|}\hline a \\ \hline b \\ \hline\end{array} \otimes \begin{array}{|c|}\hline i \\ \hline\end{array}, \qquad q^{\dagger}_{ab} \sim \begin{array}{|c|}\hline a \\ \hline b \\ \hline\end{array}, \qquad L^{\dagger}_{i} \sim \begin{array}{|c|}\hline i \\ \hline\end{array}, \quad \text{and} \qquad H^{\dagger}_{i} \sim \begin{array}{|c|}\hline i \\ \hline\end{array}, \tag{22}$$

where the only nontrivial symmetry is given by $Q^{\dagger}_{abi}$ and $q^{\dagger}_{ab}$ which are antisymmetric under the exchange $a \leftrightarrow b$.

From the discussion above, it is clear that one can write a generic operator as

$$\mathcal{O}^{\mathrm{SU}(n)} = (T^{\mathrm{SU}(n)})^{\boldsymbol{I}_1 \ldots \boldsymbol{I}_N} \prod_{i=1}^{N} (\Phi_i)_{\boldsymbol{I}_i}, \tag{23}$$

where $\boldsymbol{I}_i$ is a multi-index containing $\mathrm{SU}(n)$ fundamental indices only. The independent set of tensors $T^{\mathrm{SU}(n)}$ is constructed by combining the Young tableaux of each field in the operator according to the modified LR rule, and subsequently identifying each column of the resulting tableau with an $\mathrm{SU}(n)$ ϵ tensor.

F. Permutation symmetries

It is important to note that, up to this point, the algorithm considers all fields that occur in an operator as distinct. But if an operator contains several copies of the same field, new redundancies can occur. This happens because in expressions like Eqs. (11) and (23), the sum over indices is not explicitly carried out. Therefore, the Lorentz and internal symmetry groups can no longer be treated independently as soon as the operator contains identical fields.

As a generalization of this, a theory may contain several copies of fields that transform identically under the Lorentz and internal symmetry groups and are thus indistinguishable for our considerations. In the spirit of the SM, we will refer to these copies as *generations*.[8] Owing to the Lorentz and internal symmetry of the operator, not all combinations of generations are independent of one another, in general.

Both of these problems can be treated in the same way when introducing *generation indices* for *all* fields, even for those that occur only in a single generation. The expression *repeated fields* denotes the product of fields which at most differ by their generation index. The general strategy is to decompose any operator into a sum of terms with specific permutation symmetry λ of the generation indices for repeated fields [14].[9]

Following the philosophy of Ref. [3], the permutation symmetry is not inscribed on the operator by symmetrizing the generation indices themselves, but at the level of the Lorentz and internal symmetry tensors they are multiplied with. Consider a specific *type*, i.e., the set of operators which contain a certain set of fields. Up to now, the algorithm has generated a set of tensors $T^{\mathrm{Lorentz}} = \{T_1^{\mathrm{Lorentz}}, \ldots, T_l^{\mathrm{Lorentz}}\}$ for the Lorentz symmetry, and $T^{\mathrm{SU}(n_k)} = \{T_1^{\mathrm{SU}(n_k)}, \ldots, T_{l_k}^{\mathrm{SU}(n_k)}\}$ for each internal symmetry; see Secs. II. C and II. E. These tensors are now combined in such a way that they reflect the permutation symmetries, labeled by λ:

$$\mathcal{T}_j^{\lambda} = \sum_i \mathcal{K}_{ji}^{\lambda} (T^{\mathrm{Lorentz}} \otimes T^{\mathrm{SU}(n_1)} \otimes \cdots \otimes T^{\mathrm{SU}(n_k)})_i, \tag{24}$$

where the $\mathcal{K}$ are obtained from the plethysm technique and inner-product decomposition (for details, see Ref. [3]). The redundancies due to the permutation symmetries are reflected in the fact that $\mathcal{K}$ is an $n \times m$ matrix with $n \leq m = l \cdot l_1 \cdots l_k$ in general. The set of independent operators is then given by contracting the fields with $\mathcal{T}_j^{\lambda}$ for each $j = 1, \ldots, n$ and each permutation symmetry λ, and enumerating the generation indices according to the set of SSYTx. Explicit examples will be given below.

G. Including gravity

The algorithm described above can be extended to theories which include fields with higher spin [17,20,29]. In particular, gravitational interactions can be taken into account by considering the Weyl tensor $C_{\mu\nu\rho\sigma}$ as an additional building block in the EFT construction [20]:

$$C_{\mu\nu\rho\sigma} = R_{\mu\nu\rho\sigma} - (g_{\mu[\rho}R_{\sigma]\nu} - g_{\nu[\rho}R_{\sigma]\mu}) + \frac{1}{3} g_{\mu[\rho}g_{\sigma]\nu}R, \tag{25}$$

where $R_{\mu\nu\rho\sigma}$ denotes the Riemann tensor, $R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}$ the Ricci tensor, $R = R^{\mu}{}_{\mu}$ the Ricci scalar, and $g_{\mu\nu}$ the metric

[8] Ref. [3] uses the term *flavors* instead.

[9] In general, the symbol λ can denote multiple irreducible representations of the symmetric group, one for each set of repeated fields.

tensor, and the indices between square brackets are to be antisymmetrized. In analogy to Sec. II. A, we write the Weyl tensor in terms of left- and right-handed components that transform under irreducible representations (j_l, j_r) of the Lorentz group:

$$C^{\mu\nu\rho\sigma} = \frac{1}{64}\left(C_{\mathrm{L}}^{\alpha\beta\gamma\delta}\sigma^{\mu\nu}_{\alpha\beta}\sigma^{\rho\sigma}_{\gamma\delta} + C_{\mathrm{R}}^{\dot{\alpha}\dot{\beta}\dot{\gamma}\dot{\delta}}\bar{\sigma}^{\mu\nu}_{\dot{\alpha}\dot{\beta}}\bar{\sigma}^{\rho\sigma}_{\dot{\gamma}\dot{\delta}}\right), \tag{26}$$

where

$$C_{\mathrm{L}\alpha\beta\gamma\delta} = C_{\mu\nu\rho\sigma}\sigma^{\mu\nu}_{\alpha\beta}\sigma^{\rho\sigma}_{\gamma\delta} \in (2,0), \quad \text{and}$$
$$C_{\mathrm{R}}^{\dot{\alpha}\dot{\beta}\dot{\gamma}\dot{\delta}} = C^{\mu\nu\rho\sigma}\bar{\sigma}^{\dot{\alpha}\dot{\beta}}_{\mu\nu}\bar{\sigma}^{\dot{\gamma}\dot{\delta}}_{\rho\sigma} \in (0,2). \tag{27}$$

Denoting by $n_{C_{\mathrm{L/R}}}$ the number of left-/right-handed Weyl tensors in the operators, we extend the families (6) to

$$(n_{C_{\mathrm{L}}}, n_{F_{\mathrm{L}}}, n_{\psi}, n_{\phi}, n_{\psi^\dagger}, n_{F_{\mathrm{R}}}, n_{C_{\mathrm{R}}}; n_D). \tag{28}$$

While all previously considered fields satisfy the relation

$$[\Phi] = 1 + |h_\Phi|, \tag{29}$$

with $[\cdot]$ denoting the mass dimension in four space-time dimensions and h_Φ the helicity, this relation is violated for the Weyl tensor, which satisfies

$$[C_{\mathrm{L/R}}] = |h_{C_{\mathrm{L/R}}}| = 2, \tag{30}$$

instead. Therefore, the algorithm needs to account for the actual mass dimension of the Weyl tensor. For example, instead of Eq. (8), we get

$$N + n_l + n_r - n_{\mathrm{GR}} = d, \tag{31}$$

where $n_{\mathrm{GR}} \equiv n_{C_{\mathrm{L}}} + n_{C_{\mathrm{R}}}$ can take values from $0, \ldots, \min(d/2, N)$, and the definitions in Eq. (7) are modified, such that

$$n_l = 2n_{C_{\mathrm{L}}} + n_{F_{\mathrm{L}}} + \frac{1}{2}n_\psi + \frac{1}{2}n_D,$$
$$n_r = 2n_{C_{\mathrm{R}}} + n_{F_{\mathrm{R}}} + \frac{1}{2}n_{\psi^\dagger} + \frac{1}{2}n_D. \tag{32}$$

An example for an operator type of dimension-12 in GRSMEFT will be given below.

III. APPLICATION TO SMEFT AND GRSMEFT

According to the algorithm described in Sec. II, we construct on-shell operator bases for SMEFT and GRSMEFT up to mass dimension 12, using our implementation AutoEFT. To be consistent with the existing literature, we adopt the all-left chirality convention for the SM fields. The EFT field content is given in Table I.

TABLE I. SM field content in the all-left chiral convention (cf. Sec. II. A). Spinor indices are denoted by α and β, while a, b, c and i, j denote fundamental indices of the gauge groups SU(3) and SU(2), respectively. Each fermion carries an additional generation index g. The Young tableaux of the SU(n) groups shown in columns three and four have been introduced in Sec. II. E. The `Symbol` denotes the characters that represent the corresponding field in the Supplemental Material [28]. Hermitian conjugated fields are denoted by a trailing "+". The covariant derivative is represented by the character "`D`".

Field	h	SU(3)	SU(2)	U(1)	Generations	`Symbol`
$(G_{\mathrm{L}\alpha\beta})_{abc}$	-1	[a b] / [c]		0	1	`GL`
$(W_{\mathrm{L}\alpha\beta})_{ij}$	-1		[i j]	0	1	`WL`
$B_{\mathrm{L}\alpha\beta}$	-1			0	1	`BL`
$(Q^g_\alpha)_{ai}$	$-1/2$	[a]	[i]	$1/6$	3	`Q`
$(u^g_{\mathbb{C}\alpha})_{ab}$	$-1/2$	[a] / [b]		$-2/3$	3	`uC`
$(d^g_{\mathbb{C}\alpha})_{ab}$	$-1/2$	[a] / [b]		$1/3$	3	`dC`
$(L^g_\alpha)_i$	$-1/2$		[i]	$-1/2$	3	`L`
$e^g_{\mathbb{C}\alpha}$	$-1/2$			1	3	`eC`
H_i	0		[i]	$1/2$	1	`H`
$C_{\mathrm{L}\alpha\beta\gamma\delta}$	-2			0	1	`CL`

These fields are related to the conventional notation by Eqs. (4), (20), (26), and (27). For example, the SM Dirac spinors are given by

$$Q_{\mathrm{L}} = \begin{pmatrix} Q_\alpha \\ 0 \end{pmatrix}, \qquad u_{\mathrm{R}} = \begin{pmatrix} 0 \\ u^{\dagger\dot{\alpha}}_{\mathbb{C}} \end{pmatrix}, \qquad d_{\mathrm{R}} = \begin{pmatrix} 0 \\ d^{\dagger\dot{\alpha}}_{\mathbb{C}} \end{pmatrix},$$
$$L_{\mathrm{L}} = \begin{pmatrix} L_\alpha \\ 0 \end{pmatrix}, \qquad e_{\mathrm{R}} = \begin{pmatrix} 0 \\ e^{\dagger\dot{\alpha}}_{\mathbb{C}} \end{pmatrix}, \tag{33}$$

where all internal symmetry and generation indices are suppressed.

For a given mass dimension, the operators are grouped into families (6), and in each family, the operators are classified by their type (i.e., the field content). For each of these types, an output file is generated in YAML format.[10] It contains the relevant information to construct all operators belonging to this type. We found this way of presenting the results a reasonable compromise which allows one to encode the large number of operators in a rather compact form without leaving too much calculational effort to the user.

In the following, four representative cases are displayed to show how the output files can be interpreted in terms of Sec. II.

[10] See https://yaml.org/.

A. No repeated fields

As a first example, let us consider the dimension-ten operators of type

$$\mathcal{Q}_{\mathrm{NRF}} = W_{\mathrm{L}} Q u_{\mathbb{C}} H W_{\mathrm{L}}^{\dagger} D^2, \tag{34}$$

encoded in the file

`operators/5/1FL_2psi_1phi_1FR_2D/1WL_1Q_1uC_1H_1WL+_2D.yml`

included in `smeft_10.tar.xz`. They contain a weak gauge field and its conjugate (`1WL`, `1WL+`), a quark doublet (`1Q`), a charge conjugate up-type quark singlet (`1uC`), a Higgs field (`1H`), and two derivatives (`2D`). This also illustrates the naming scheme of the output files of AutoEFT. The content of this file is displayed in Listing 1. We deliberately include redundant information in the output files in order to facilitate their interpretation and to allow for consistency checks when constructing the explicit operators.

```
# '1WL_1Q_1uC_1H_1WL+_2D.yml' generated by AutoEFT 1.0.0
version: 1.0.0
type:
- {WL: 1, Q: 1, uC: 1, H: 1, WL+: 1, D: 2}
- complex
generations: {WL: 1, Q: 3, uC: 3, H: 1, WL+: 1}
n_terms: 10
n_operators: 90
invariants:
  Lorentz:
    O(Lorentz,1): +eps(1_1,3_1)*eps(1_2,3_2)*eps(2_1,3_3)*eps(3_1~,5_1~)*eps(3_2~,5_2~) * WL(1_1,1_2)*Q(2_1)*(D^2 uC)(3_1,3_2,3_3,3_1~,3_2~)*H*WL+(5_1~,5_2~)
    O(Lorentz,2): -eps(1_1,3_1)*eps(1_2,3_2)*eps(2_1,4_1)*eps(3_1~,5_1~)*eps(4_1~,5_2~) * WL(1_1,1_2)*Q(2_1)*(D uC)(3_1,3_2,3_1~)*(D H)(4_1,4_1~)*WL+(5_1~,5_2~)
    O(Lorentz,3): +eps(1_1,3_1)*eps(1_2,4_1)*eps(2_1,4_2)*eps(4_1~,5_1~)*eps(4_2~,5_2~) * WL(1_1,1_2)*Q(2_1)*uC(3_1)*(D^2 H)(4_1,4_2,4_1~,4_2~)*WL+(5_1~,5_2~)
    O(Lorentz,4): -eps(1_1,2_1)*eps(1_2,3_1)*eps(3_2,4_1)*eps(3_1~,5_1~)*eps(4_1~,5_2~) * WL(1_1,1_2)*Q(2_1)*(D uC)(3_1,3_2,3_1~)*(D H)(4_1,4_1~)*WL+(5_1~,5_2~)
    O(Lorentz,5): +eps(1_1,2_1)*eps(1_2,4_1)*eps(3_1,4_2)*eps(4_1~,5_1~)*eps(4_2~,5_2~) * WL(1_1,1_2)*Q(2_1)*uC(3_1)*(D^2 H)(4_1,4_2,4_1~,4_2~)*WL+(5_1~,5_2~)
  SU3:
    O(SU3,1): +eps(2_1,3_1,3_2) * WL*Q(2_1)*uC(3_1,3_2)*H*WL+
  SU2:
    O(SU2,1): +eps(1_1,4_1)*eps(1_2,5_1)*eps(2_1,5_2) * WL(1_1,1_2)*Q(2_1)*uC*H(4_1)*WL+(5_1,5_2)
    O(SU2,2): +eps(1_1,2_1)*eps(1_2,5_1)*eps(4_1,5_2) * WL(1_1,1_2)*Q(2_1)*uC*H(4_1)*WL+(5_1,5_2)
permutation_symmetries:
- vector: Lorentz * SU3 * SU2
- symmetry: {WL: [1], Q: [1], uC: [1], H: [1], WL+: [1]}
  n_terms: 10
  n_operators: 90
  matrix: |-
    [1 0 0 0 0 0 0 0 0 0]
    [0 1 0 0 0 0 0 0 0 0]
    [0 0 1 0 0 0 0 0 0 0]
    [0 0 0 1 0 0 0 0 0 0]
    [0 0 0 0 1 0 0 0 0 0]
    [0 0 0 0 0 1 0 0 0 0]
    [0 0 0 0 0 0 1 0 0 0]
    [0 0 0 0 0 0 0 1 0 0]
    [0 0 0 0 0 0 0 0 1 0]
    [0 0 0 0 0 0 0 0 0 1]
```

Listing 1 Content of the file `1WL_1Q_1uC_1H_1W+_2D.yml`, which encodes the operators of Eq. (34).

The file is structured by certain keywords which we explain in the following:

1. version: The version of AutoEFT which was used to produce the output file.
2. type: A list[11] of two elements. The first element (line 4) specifies the multiplicities of the fields and derivatives in the operator. It is equivalent to the specification in Eq. (34). The second element (line 5) states that the operator type is complex, (i.e., the operators are not Hermitian). The Hermitian conjugate type is contained in a different file. In the present case, this would be 1WL_1H+_1Q+_1uC+_1WL+_2D.yml. For Hermitian operators, the second entry of type is real.
3. generations: Provides the number of generations for each field. In this case, it specifies that there is only a single generation of WL, WL+, and H, but three generations of quarks Q and uC.
4. n_terms: The total number of operators with independent Lorentz and SU(n) index contractions. It does not take into account the number of generations though. In Listing 1, there are five independent ways of contracting the Lorentz indices of the operators while there is only one option for SU(3), and two for SU(2), so the number of terms n_terms is $5 \cdot 1 \cdot 2 = 10$.
5. n_operators: The total number of independent operators, taking into account the independent values the generation indices can take. For Eq. (34), there are $3 \cdot 3 = 9$ independent combinations of the Q and $u_{\mathbb{C}}$ generation indices for each of the ten terms; hence n_operators is $10 \cdot 9 = 90$.
6. invariants: The list of invariant index contractions of the fields in the operator. The contraction of indices for each internal symmetry group and for the Lorentz group is listed separately, as indicated by the subkeywords Lorentz, SU(3), and SU(2) in lines 10, 16, and 18. Each independent contraction is labeled by O(<G>,<m>), where <G> is the name of the Lorentz or internal symmetry group, and <m> enumerates the contractions.

   The indices are denoted by <i>_<j>, where <i> is the position of the field that carries this index, and <j> is the position of the index *on* the field. Per invariant contraction, each index appears exactly twice, and summation is implied. For the Lorentz group, <i>_<j>~ denotes dotted indices. Furthermore, the indices are associated with the building blocks of (12), rather than with the fields. Note that the dotted and undotted indices of the building blocks are understood to be (separately) symmetrized (cf. Sec. II. D).

   The symbol eps denotes the ϵ tensor with eps(1,2)=eps(2~,1~)=1 for the Lorentz group, and eps(1,2,…,n)=1 for any internal SU(n) group. All indices not associated with the symmetry group in question are suppressed on the fields.[12]
7. permutation_symmetries: The list of permutation symmetries according to Sec. II. F. Since the operator type of Eq. (34) does not involve repeated fields, this entry (lines 22–36) is redundant and is included only for consistency. Details will be discussed in Sec. III. B for a nontrivial example.

In Listing 1, the invariant contractions are given by [cf. Eqs. (11) and (23)]

$$
\begin{aligned}
\text{line 11}\colon\ \mathcal{O}_1^{\text{Lorentz}} &= T_1^{\text{Lorentz}} \circ \mathcal{Q}_{\text{NRF}} = \epsilon^{\alpha_1\gamma_1}\epsilon^{\alpha_2\gamma_2}\epsilon^{\beta_1\gamma_3}\epsilon_{\dot\gamma_1\dot\eta_1}\epsilon_{\dot\gamma_2\dot\eta_2} W_{\text{L}\alpha_1\alpha_2} Q_{\beta_1} (D^2 u_{\mathbb{C}})^{(\dot\gamma_1\dot\gamma_2)}_{(\gamma_1\gamma_2\gamma_3)} H W_{\text{L}}^{\dagger\,\dot\eta_1\dot\eta_2},\\
\text{line 12}\colon\ \mathcal{O}_2^{\text{Lorentz}} &= T_2^{\text{Lorentz}} \circ \mathcal{Q}_{\text{NRF}} = -\epsilon^{\alpha_1\gamma_1}\epsilon^{\alpha_2\gamma_2}\epsilon^{\beta_1\delta_1}\epsilon_{\dot\gamma_1\dot\eta_1}\epsilon_{\dot\delta_1\dot\eta_2} W_{\text{L}\alpha_1\alpha_2} Q_{\beta_1} (D u_{\mathbb{C}})^{\dot\gamma_1}_{(\gamma_1\gamma_2)} (DH)^{\dot\delta_1}_{\delta_1} W_{\text{L}}^{\dagger\,\dot\eta_1\dot\eta_2},\\
\text{line 13}\colon\ \mathcal{O}_3^{\text{Lorentz}} &= T_3^{\text{Lorentz}} \circ \mathcal{Q}_{\text{NRF}} = \epsilon^{\alpha_1\gamma_1}\epsilon^{\alpha_2\delta_1}\epsilon^{\beta_1\delta_2}\epsilon_{\dot\delta_1\dot\eta_1}\epsilon_{\dot\delta_2\dot\eta_2} W_{\text{L}\alpha_1\alpha_2} Q_{\beta_1} u_{\mathbb{C}\gamma_1} (D^2 H)^{(\dot\delta_1\dot\delta_2)}_{(\delta_1\delta_2)} W_{\text{L}}^{\dagger\,\dot\eta_1\dot\eta_2},\\
\text{line 14}\colon\ \mathcal{O}_4^{\text{Lorentz}} &= T_4^{\text{Lorentz}} \circ \mathcal{Q}_{\text{NRF}} = -\epsilon^{\alpha_1\beta_1}\epsilon^{\alpha_2\gamma_1}\epsilon^{\gamma_2\delta_1}\epsilon_{\dot\gamma_1\dot\eta_1}\epsilon_{\dot\delta_1\dot\eta_2} W_{\text{L}\alpha_1\alpha_2} Q_{\beta_1} (D u_{\mathbb{C}})^{\dot\gamma_1}_{(\gamma_1\gamma_2)} (DH)^{\dot\delta_1}_{\delta_1} W_{\text{L}}^{\dagger\,\dot\eta_1\dot\eta_2},\\
\text{line 15}\colon\ \mathcal{O}_5^{\text{Lorentz}} &= T_5^{\text{Lorentz}} \circ \mathcal{Q}_{\text{NRF}} = \epsilon^{\alpha_1\beta_1}\epsilon^{\alpha_2\delta_1}\epsilon^{\gamma_1\delta_2}\epsilon_{\dot\delta_1\dot\eta_1}\epsilon_{\dot\delta_2\dot\eta_2} W_{\text{L}\alpha_1\alpha_2} Q_{\beta_1} u_{\mathbb{C}\gamma_1} (D^2 H)^{(\dot\delta_1\dot\delta_2)}_{(\delta_1\delta_2)} W_{\text{L}}^{\dagger\,\dot\eta_1\dot\eta_2},\\
\text{line 17}\colon\ \mathcal{O}_1^{\text{SU}(3)} &= T_1^{\text{SU}(3)} \circ \mathcal{Q}_{\text{NRF}} = \epsilon^{b_1c_1c_2} W_{\text{L}} Q_{b_1} u_{\mathbb{C}c_1c_2} H W_{\text{L}}^{\dagger},\\
\text{line 19}\colon\ \mathcal{O}_1^{\text{SU}(2)} &= T_1^{\text{SU}(2)} \circ \mathcal{Q}_{\text{NRF}} = \epsilon^{k_1m_1}\epsilon^{k_2n_1}\epsilon^{l_1n_2} W_{\text{L}k_1k_2} Q_{l_1} u_{\mathbb{C}} H_{m_1} W^{\dagger}_{\text{L}n_1n_2},\\
\text{line 20}\colon\ \mathcal{O}_2^{\text{SU}(2)} &= T_2^{\text{SU}(2)} \circ \mathcal{Q}_{\text{NRF}} = \epsilon^{k_1l_1}\epsilon^{k_2n_1}\epsilon^{m_1n_2} W_{\text{L}k_1k_2} Q_{l_1} u_{\mathbb{C}} H_{m_1} W^{\dagger}_{\text{L}n_1n_2},
\end{aligned}
\tag{35}
$$

[11] In YAML, this is called a *sequence*. Its elements are marked by the leading dashes in lines 4 and 5 of Listing 1.

[12] This means that, if the operators contain only fields that are singlets under a particular symmetry group G (i.e., if there is no index to be contracted), the corresponding entry G: contains just a single element +1 multiplied by the fields without indices.

with $\mathcal{Q}_{\mathrm{NRF}}$ from Eq. (34). The operation $T \circ \mathcal{Q}$ implies that all indices other than those of T are suppressed in $\mathcal{Q}$. Since only contractions with the Lorentz tensors determine the positions of the derivatives, we omit the latter in the contractions with tensors of the internal symmetries.

The complete set of independent operators $\boldsymbol{T} \cdot \mathcal{Q}_{\mathrm{NRF}}$ is thus given by all possible combinations of the Lorentz and $\mathrm{SU}(n)$ tensors:

$$\boldsymbol{T} \equiv T^{\mathrm{Lorentz}} \otimes T^{\mathrm{SU}(3)} \otimes T^{\mathrm{SU}(2)} = \begin{pmatrix} T_1^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_1^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \\ T_2^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_2^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \\ T_3^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_3^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \\ T_4^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_4^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \\ T_5^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_5^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \end{pmatrix}. \tag{36}$$

Three examples for the ten independent terms (cf. line 7 in Listing 1) are given by

$$\begin{aligned} (T_1^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)}) \circ \mathcal{Q}_{\mathrm{NRF}} &= \epsilon^{\alpha_1\gamma_1}\epsilon^{\alpha_2\gamma_2}\epsilon^{\beta_1\gamma_3}\epsilon_{\dot\gamma_1\dot\eta_1}\epsilon_{\dot\gamma_2\dot\eta_2}\epsilon^{b_1c_1c_2}\epsilon^{k_1m_1}\epsilon^{k_2n_1}\epsilon^{l_1n_2} \\ &\quad \times W_{\mathrm{L}\alpha_1\alpha_2k_1k_2}Q_{\beta_1b_1l_1}(D^2u_{\mathbb{C}c_1c_2})^{(\dot\gamma_1\dot\gamma_2)}_{(\gamma_1\gamma_2\gamma_3)}H_{m_1}W^{\dagger\,\dot\eta_1\dot\eta_2}_{\mathrm{L}n_1n_2}, \\ (T_1^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)}) \circ \mathcal{Q}_{\mathrm{NRF}} &= \epsilon^{\alpha_1\gamma_1}\epsilon^{\alpha_2\gamma_2}\epsilon^{\beta_1\gamma_3}\epsilon_{\dot\gamma_1\dot\eta_1}\epsilon_{\dot\gamma_2\dot\eta_2}\epsilon^{b_1c_1c_2}\epsilon^{k_1l_1}\epsilon^{k_2n_1}\epsilon^{m_1n_2} \\ &\quad \times W_{\mathrm{L}\alpha_1\alpha_2k_1k_2}Q_{\beta_1b_1l_1}(D^2u_{\mathbb{C}c_1c_2})^{(\dot\gamma_1\dot\gamma_2)}_{(\gamma_1\gamma_2\gamma_3)}H_{m_1}W^{\dagger\,\dot\eta_1\dot\eta_2}_{\mathrm{L}n_1n_2}, \\ (T_2^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)}) \circ \mathcal{Q}_{\mathrm{NRF}} &= -\epsilon^{\alpha_1\gamma_1}\epsilon^{\alpha_2\gamma_2}\epsilon^{\beta_1\delta_1}\epsilon_{\dot\gamma_1\dot\eta_1}\epsilon_{\dot\delta_1\dot\eta_2}\epsilon^{b_1c_1c_2}\epsilon^{k_1m_1}\epsilon^{k_2n_1}\epsilon^{l_1n_2} \\ &\quad \times W_{\mathrm{L}\alpha_1\alpha_2k_1k_2}Q_{\beta_1b_1l_1}(Du_{\mathbb{C}c_1c_2})^{\dot\gamma_1}_{(\gamma_1\gamma_2)}(DH_{m_1})^{\dot\delta_1}_{\delta_1}W^{\dagger\,\dot\eta_1\dot\eta_2}_{\mathrm{L}n_1n_2}. \end{aligned} \tag{37}$$

Up to this point, we have suppressed the generation indices of the quarks. Since the quark and antiquark transform differently under the symmetry groups, we can simply attach a generation index to each of the two quark fields in Eq. (34).

B. Repeated fields

The operator Eq. (34) treated in Sec. III. A is special in the sense that it does not contain repeated fields, i.e., all fields in this type transform differently under the symmetry group of the theory. As mentioned in Sec. II. F, if repeated fields are present, new redundancies can arise. The output file thus requires information on the permutation symmetry. As an example, we consider the dimension-11 operators of type

$$\mathcal{Q}_{\mathrm{RF}} = L^2 d_{\mathbb{C}} e_{\mathbb{C}} u_{\mathbb{C}}^2 H^2, \tag{38}$$

contained in the file

`operators/8/6psi_2phi/2L_1dC_1eC_2uC_2H.yml`

that is included in `smeft_11.tar.xz`. Its content is displayed in Listing 2.

```
1  # '2L_1dC_1eC_2uC_2H.yml' generated by AutoEFT 1.0.0
2  version: 1.0.0
3  type:
4  - {L: 2, dC: 1, eC: 1, uC: 2, H: 2}
5  - complex
6  generations: {L: 3, dC: 3, eC: 3, uC: 3, H: 1}
7  n_terms: 5
8  n_operators: 891
9  invariants:
10   Lorentz:
11     O(Lorentz,1): +eps(1_1,4_1)*eps(2_1,5_1)*eps(3_1,6_1) * L(1_1)*L(2_1)*dC(3_1)*
     ↪ eC(4_1)*uC(5_1)*uC(6_1)*H*H
12     O(Lorentz,2): +eps(1_1,3_1)*eps(2_1,5_1)*eps(4_1,6_1) * L(1_1)*L(2_1)*dC(3_1)*
     ↪ eC(4_1)*uC(5_1)*uC(6_1)*H*H
13     O(Lorentz,3): +eps(1_1,3_1)*eps(2_1,4_1)*eps(5_1,6_1) * L(1_1)*L(2_1)*dC(3_1)*
     ↪ eC(4_1)*uC(5_1)*uC(6_1)*H*H
14     O(Lorentz,4): +eps(1_1,2_1)*eps(3_1,5_1)*eps(4_1,6_1) * L(1_1)*L(2_1)*dC(3_1)*
     ↪ eC(4_1)*uC(5_1)*uC(6_1)*H*H
15     O(Lorentz,5): +eps(1_1,2_1)*eps(3_1,4_1)*eps(5_1,6_1) * L(1_1)*L(2_1)*dC(3_1)*
     ↪ eC(4_1)*uC(5_1)*uC(6_1)*H*H
16   SU3:
17     O(SU3,1): +eps(3_1,3_2,5_2)*eps(5_1,6_1,6_2) * L*L*dC(3_1,3_2)*eC*uC(5_1,5_2)*
     ↪ uC(6_1,6_2)*H*H
18   SU2:
19     O(SU2,1): +eps(1_1,7_1)*eps(2_1,8_1) * L(1_1)*L(2_1)*dC*eC*uC*uC*H(7_1)*H(8_1)
20     O(SU2,2): +eps(1_1,2_1)*eps(7_1,8_1) * L(1_1)*L(2_1)*dC*eC*uC*uC*H(7_1)*H(8_1)
21 permutation_symmetries:
22 - vector: Lorentz * SU3 * SU2
23 - symmetry: {L: [1, 1], dC: [1], eC: [1], uC: [2], H: [2]}
24   n_terms: 2
25   n_operators: 324
26   matrix: |-
27     [ 8 -4  0  0 -4  2 -4  2  6 -3]
28     [ 0  0  8 -4 -4  2 -4  2  2 -1]
29 - symmetry: {L: [1, 1], dC: [1], eC: [1], uC: [1, 1], H: [2]}
30   n_terms: 1
31   n_operators: 81
32   matrix: |-
33     [ 0  0  0  0  4 -2  0  0 -2  1]
34 - symmetry: {L: [2], dC: [1], eC: [1], uC: [2], H: [2]}
35   n_terms: 1
36   n_operators: 324
37   matrix: |-
38     [ 0  0  0  0  0  0  4 -2 -2  1]
39 - symmetry: {L: [2], dC: [1], eC: [1], uC: [1, 1], H: [2]}
40   n_terms: 1
41   n_operators: 162
42   matrix: |-
43     [ 0  0  0  0  0  0  0  0 -2  1]
```

Listing 2: Content of file `2L_1dC_1eC_2uC_2H.yml`, which encodes the operators of Eq. (38).

The meaning of the first 20 lines in this file was explained in the previous section. If all fields in Eq. (38) were different, one would arrive at ten terms, according to the $5 \cdot 1 \cdot 2 = 10$ independent Lorentz and SU(n) tensors:

$$\boldsymbol{T} \equiv T^{\mathrm{Lorentz}} \otimes T^{\mathrm{SU}(3)} \otimes T^{\mathrm{SU}(2)} = \begin{pmatrix} T_1^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_1^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \\ T_2^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_2^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \\ T_3^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_3^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \\ T_4^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_4^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \\ T_5^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_1^{\mathrm{SU}(2)} \\ T_5^{\mathrm{Lorentz}} \otimes T_1^{\mathrm{SU}(3)} \otimes T_2^{\mathrm{SU}(2)} \end{pmatrix}. \quad (39)$$

Attaching generation indices to the fields, the number of operators would be $10 \cdot 3^6 = 7290$, because there are six fields with three generations each. This, however, neglects the permutation symmetries arising from the repeated fields. In this example, there are three sets of repeated fields, each of which occurs twice, as can be seen from the first element of `type` (line 4). This reduces the actual number of terms and operators to those given in lines 7 and 8. As described in Sec. II. F, this is because, on the one hand, the permutation symmetries induce dependencies between the terms in Eq. (39), reducing the number of independent terms. On the other hand, a particular permutation symmetry restricts the number of independent combinations of generation indices in the field monomials.

Consider, for example, the second item (marked by the character "-" in the first column) under `permutation_symmetries`, i.e., lines 23–28 of Listing 2. The entry for the keyword `symmetry` should be interpreted as a set of Young tableaux, representing the permutation symmetry for each kind of field. In this case[13]

$$\begin{aligned} &\texttt{L:[1,1]} &&\hat{=} && \lambda_L = \boxed{\begin{matrix} i \\ j \end{matrix}}, \\ &\texttt{dC: [1]} &&\hat{=} && \lambda_{d_{\mathbb{C}}} = \boxed{k}, \\ &\texttt{eC: [1]} &&\hat{=} && \lambda_{e_{\mathbb{C}}} = \boxed{l}, \\ &\texttt{uC: [2]} &&\hat{=} && \lambda_{u_{\mathbb{C}}} = \boxed{m\,|\,n}, \\ &\texttt{H: [2]} &&\hat{=} && \lambda_H = \boxed{r\,|\,s}. \end{aligned} \quad (40)$$

The fillings are the generation indices, i.e., $i, j, k, l, m, n \in \{1,2,3\}$, and $r = s = 1$, because there are three generations of leptons and quarks, and only one of Higgs bosons.

[13] I.e. `[i,j,…]` in the `symmetry` entry denotes a Young diagram with `i` boxes in the first row, `j` boxes in the second, etc.

For this permutation symmetry, only the contractions with two combinations of the ten tensors in Eq. (39) are independent. They are given by [cf. Eq. (24)]

$$\boldsymbol{\mathcal{T}} = \mathcal{K}\boldsymbol{T}, \quad (41)$$

where

$$\mathcal{K} = \begin{pmatrix} 8 & -4 & 0 & 0 & -4 & 2 & -4 & 2 & 6 & -3 \\ 0 & 0 & 8 & -4 & -4 & 2 & -4 & 2 & 2 & -1 \end{pmatrix} \quad (42)$$

is the matrix listed under the keyword `matrix` in lines 26–28 of Listing 2. The rank of this matrix is equal to the number of terms `n_terms` in line 24. Its form depends on the order of the factors in the Kronecker product[14] in Eq. (39). For clarity, this order is added to the AutoEFT output file under the keyword `vector`; see line 22 in Listing 2.

Considering the permutation symmetries of the fields, only such values of generation indices are independent for which the Young tableaux in (40) are semistandard. For the case at hand, we thus have

$$\begin{aligned} (i,j) &\in \{(1,2),(1,3),(2,3)\} \equiv \mathcal{S}_{1,1}^{(3)}, \\ (m,n) &\in \{(1,1),(1,2),(1,3),(2,2),(2,3),(3,3)\} \equiv \mathcal{S}_2^{(3)}, \\ (r,s) &\in \{(1,1)\} \equiv \mathcal{S}_2^{(1)}, \end{aligned} \quad (43)$$

where $\mathcal{S}_{\lambda_\Phi}^{(n_g)}$ is the set of independent generation indices for the repeated field Φ with permutation symmetry λ_Φ and n_g generations. The generation indices of fields that appear only once are still independent and are therefore not restricted further, i.e., $k, l \in \{1,2,3\}$. In total, out of the 3^6 combinations of the generation indices, only $3 \cdot 6 \cdot 1 \cdot 3^2 = 162$ need to be considered for this particular representation of the permutation symmetry. There are thus $2 \cdot (3 \cdot 6 \cdot 1 \cdot 3^2) = 324$ independent operators of this type, in agreement with line 25.

There are three more permutation symmetries:

1. lines 29–33: $\lambda_L = \boxed{\begin{matrix} i \\ j \end{matrix}}$, $\lambda_{d_{\mathbb{C}}} = \boxed{k}$, $\lambda_{e_{\mathbb{C}}} = \boxed{l}$, $\lambda_{u_{\mathbb{C}}} = \boxed{\begin{matrix} m \\ n \end{matrix}}$, $\lambda_H = \boxed{r\,|\,s}$, in which case there is only one independent combination (cf. line 30) of tensor contractions, with

$$\mathcal{K} = \begin{pmatrix} 0 & 0 & 0 & 0 & 4 & -2 & 0 & 0 & -2 & 1 \end{pmatrix}. \quad (44)$$

The independent values of generation indices are given by $(i,j), (m,n) \in \mathcal{S}_{1,1}^{(3)}$, $(r,s) \in \mathcal{S}_2^{(1)}$, and $k, l \in \{1,2,3\}$. This results in $1 \cdot (3 \cdot 3 \cdot 1 \cdot 3^2) = 81$ independent operators, as indicated in line 31.

[14] The Kronecker product $A \otimes B$ of an $m \times n$ matrix A with a $p \times q$ matrix B is a $pm \times qn$ matrix which is obtained by replacing all entries of A by their product with the matrix B. Note that $A \otimes (B \otimes C) = (A \otimes B) \otimes C \equiv A \otimes B \otimes C$.

2. lines 34–38: $\lambda_L = \boxed{i}\boxed{j}$, $\lambda_{d_\mathbb{C}} = \boxed{k}$, $\lambda_{e_\mathbb{C}} = \boxed{l}$, $\lambda_{u_\mathbb{C}} = \boxed{m}\boxed{n}$, $\lambda_H = \boxed{r}\boxed{s}$, in which case the only independent term is given by the matrix

$$\mathcal{K} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 4 & -2 & -2 & 1 \end{pmatrix}, \tag{45}$$

and with $(i,j),(m,n) \in \mathcal{S}_2^{(3)}$, $(r,s) \in \mathcal{S}_2^{(1)}$, and $k,l \in \{1,2,3\}$, resulting in $1 \cdot (6 \cdot 6 \cdot 1 \cdot 3^2) = 324$ operators.

3. lines 39–43: $\lambda_L = \boxed{i}\boxed{j}$, $\lambda_{d_\mathbb{C}} = \boxed{k}$, $\lambda_{e_\mathbb{C}} = \boxed{l}$, $\lambda_{u_\mathbb{C}} = \begin{array}{|c|}\hline m \\ \hline n \\ \hline \end{array}$, $\lambda_H = \boxed{r}\boxed{s}$, in which case the only independent term is given by the matrix

$$\mathcal{K} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 1 \end{pmatrix}, \tag{46}$$

and with $(i,j) \in \mathcal{S}_2^{(3)}$, $(m,n) \in \mathcal{S}_{1,1}^{(3)}$, $(r,s) \in \mathcal{S}_2^{(1)}$, and $k,l \in \{1,2,3\}$, resulting in $1 \cdot (6 \cdot 3 \cdot 1 \cdot 3^2) = 162$ operators.

In total, the file `2L_1dC_1eC_2uC_2H.yml` thus encodes $324 + 81 + 324 + 162 = 891$ operators, as stated in line 8 of Listing 2. The number of terms `n_terms` in line 7 indicates that there are five independent combinations, given by Eqs. (42) and (44) to (46).

C. SMEFT at mass dimension 12

The complete basis of operators up to mass dimension 12 for SMEFT is provided, in the format described in Secs. III. A and III. B, in the Supplemental Material [28] together with this paper. It took a few seconds to generate the operator basis at mass dimensions 5, 6, and 7, a few minutes for mass dimension 8 and 9, and a few hours for mass dimension 10 and 11. At mass dimension 12, the largest amount of time was spent by AutoEFT on incorporating the permutation symmetry of the operators with six gluonic field strength tensors, which takes of the order of 10^4 CPU hours. For example, for the type G_L^6, the 2175 contractions of the field strength tensors with the 15 independent Lorentz and 145 independent SU(3) tensors is reduced to only eight independent operators. Their explicit expressions are too long to display them in this paper. They are encoded in the file

`operators/6/6FL/6GL.yml`

which is about 140 KB in size. Instead, as an example of a dimension-12 operator, we display the following type of eight-fermion operators in Listing 3:

$$L^3 e_\mathbb{C} L^\dagger d_\mathbb{C}^{\dagger 2} u_\mathbb{C}^\dagger, \tag{47}$$

contained in the file

`operators/8/4psi_4psi+/`
`3L_1eC_1L+_2dC+_1uC+.yml`

that is included in `smeft_12.tar.xz`.

```
# '3L_1eC_1L+_2dC+_1uC+.yml' generated by AutoEFT 1.0.0
version: 1.0.0
type:
- {L: 3, eC: 1, L+: 1, dC+: 2, uC+: 1}
- complex
generations: {L: 3, eC: 3, L+: 3, dC+: 3, uC+: 3}
n_terms: 6
n_operators: 4617
invariants:
  Lorentz:
    O(Lorentz,1): +eps(1_1,3_1)*eps(2_1,4_1)*eps(5_1~,7_1~)*eps(6_1~,8_1~) * L(1_1)
    ↪ *L(2_1)*L(3_1)*eC(4_1)*L+(5_1~)*dC+(6_1~)*dC+(7_1~)*uC+(8_1~)
    O(Lorentz,2): +eps(1_1,3_1)*eps(2_1,4_1)*eps(5_1~,6_1~)*eps(7_1~,8_1~) * L(1_1)
    ↪ *L(2_1)*L(3_1)*eC(4_1)*L+(5_1~)*dC+(6_1~)*dC+(7_1~)*uC+(8_1~)
    O(Lorentz,3): +eps(1_1,2_1)*eps(3_1,4_1)*eps(5_1~,7_1~)*eps(6_1~,8_1~) * L(1_1)
    ↪ *L(2_1)*L(3_1)*eC(4_1)*L+(5_1~)*dC+(6_1~)*dC+(7_1~)*uC+(8_1~)
    O(Lorentz,4): +eps(1_1,2_1)*eps(3_1,4_1)*eps(5_1~,6_1~)*eps(7_1~,8_1~) * L(1_1)
    ↪ *L(2_1)*L(3_1)*eC(4_1)*L+(5_1~)*dC+(6_1~)*dC+(7_1~)*uC+(8_1~)
  SU3:
    O(SU3,1): +eps(6_1,7_1,8_1) * L*L*L*eC*L+*dC+(6_1)*dC+(7_1)*uC+(8_1)
  SU2:
    O(SU2,1): +eps(1_1,3_1)*eps(2_1,5_1) * L(1_1)*L(2_1)*L(3_1)*eC*L+(5_1)*dC+*dC+*
    ↪ uC+
    O(SU2,2): +eps(1_1,2_1)*eps(3_1,5_1) * L(1_1)*L(2_1)*L(3_1)*eC*L+(5_1)*dC+*dC+*
    ↪ uC+
```

```
20 permutation_symmetries:
21 - vector: Lorentz * SU3 * SU2
22 - symmetry: {L: [1, 1, 1], eC: [1], L+: [1], dC+: [2], uC+: [1]}
23   n_terms: 1
24   n_operators: 162
25   matrix: |-
26     [ 2 -1  2 -1 -1  2 -1  2]
27 - symmetry: {L: [2, 1], eC: [1], L+: [1], dC+: [2], uC+: [1]}
28   n_terms: 1
29   n_operators: 1296
30   matrix: |-
31     [ 0  1  0  1  1 -1  1 -1]
32 - symmetry: {L: [3], eC: [1], L+: [1], dC+: [2], uC+: [1]}
33   n_terms: 1
34   n_operators: 1620
35   matrix: |-
36     [ 0  1  0  1 -1  0 -1  0]
37 - symmetry: {L: [1, 1, 1], eC: [1], L+: [1], dC+: [1, 1], uC+: [1]}
38   n_terms: 1
39   n_operators: 81
40   matrix: |-
41     [ 2 -1 -2  1 -1  2  1 -2]
42 - symmetry: {L: [2, 1], eC: [1], L+: [1], dC+: [1, 1], uC+: [1]}
43   n_terms: 1
44   n_operators: 648
45   matrix: |-
46     [ 0  1  0 -1  1 -1 -1  1]
47 - symmetry: {L: [3], eC: [1], L+: [1], dC+: [1, 1], uC+: [1]}
48   n_terms: 1
49   n_operators: 810
50   matrix: |-
51     [ 0  1  0 -1 -1  0  1  0]
```

Listing 3: Content of file `3L_1eC_1L+_2dC+_1uC+.yml`, which encodes the operators of Eq. (47).

In this example, there are two sets of repeated fields, L^3 and $d_{\mathbb{C}}^{\dagger 2}$. Lines 22, 27, 32, 37, 42, and 47 specify the permutation symmetries of the generation indices i, j, k of the three lepton doublets

$$\lambda_L \in \left\{ \begin{array}{|c|}\hline i \\ \hline j \\ \hline k \\ \hline \end{array}, \begin{array}{|c|c|}\hline i & j \\ \hline k \\ \cline{1-1} \end{array}, \begin{array}{|c|c|c|}\hline i & j & k \\ \hline \end{array} \right\}. \tag{48}$$

The corresponding sets of independent generation indices are thus given by

$$\begin{aligned} \mathcal{S}^{(3)}_{1,1,1} &\equiv \{(1,2,3)\}, \\ \mathcal{S}^{(3)}_{2,1} &\equiv \{(i,j,k) | 1 \le i \le j \le 3 \wedge i < k \le 3\}, \\ \mathcal{S}^{(3)}_{3} &\equiv \{(i,j,k) | 1 \le i \le j \le k \le 3\}, \end{aligned} \tag{49}$$

containing 1, 8, and 10 elements, respectively. The permutation symmetries of the generation indices m, n of the two down-type quarks are given by

$$\lambda_{d_{\mathbb{C}}} \in \left\{ \begin{array}{|c|c|}\hline m & n \\ \hline \end{array}, \begin{array}{|c|}\hline m \\ \hline n \\ \hline \end{array} \right\}, \tag{50}$$

and thus the corresponding sets of generation indices are $\mathcal{S}^{(3)}_{2}$ and $\mathcal{S}^{(3)}_{1,1}$ with 6 and 3 elements, respectively [see Eq. (43)].

In Listing 3, all combinations of the symmetries Eqs. (48) and (50) are present, and each combination is given by exactly one term. Therefore, including the 3^3 generation multiplicities of $e_{\mathbb{C}}$, $L^{\dagger}$, and $u_{\mathbb{C}}^{\dagger}$, the total number of independent operators is given by

$$(1 + 8 + 10) \cdot (6 + 3) \cdot 3^3 = 4617, \tag{51}$$

in agreement with line 8 of Listing 3.

D. GRSMEFT at mass dimension 12

The complete basis of operators up to mass dimension 12 for GRSMEFT is provided, in the format described in Secs. III. A and III. B, in the Supplemental Material [28] together with this paper. As an example, we present in Listing 4 the operators of type

$$C_{\mathrm{L}} L^2 e_{\mathbb{C}} d_{\mathbb{C}}^{\dagger\,2} u_{\mathbb{C}}^{\dagger} D. \tag{52}$$

```
# '1CL_2L_1eC_2dC+_1uC+_1D.yml' generated by AutoEFT 1.0.0
version: 1.0.0
type:
- {CL: 1, L: 2, eC: 1, dC+: 2, uC+: 1, D: 1}
- complex
generations: {CL: 1, L: 3, eC: 3, dC+: 3, uC+: 3}
n_terms: 7
n_operators: 1539
invariants:
  Lorentz:
    O(Lorentz,1): -eps(1_1,2_1)*eps(1_2,3_1)*eps(1_3,3_2)*eps(1_4,4_1)*eps(3_1~,6_1
    ↪ ~)*eps(5_1~,7_1~) * CL(1_1,1_2,1_3,1_4)*L(2_1)*(D L)(3_1,3_2,3_1~)*eC(4_1)*
    ↪ dC+(5_1~)*dC+(6_1~)*uC+(7_1~)
    O(Lorentz,2): -eps(1_1,2_1)*eps(1_2,3_1)*eps(1_3,3_2)*eps(1_4,4_1)*eps(3_1~,5_1
    ↪ ~)*eps(6_1~,7_1~) * CL(1_1,1_2,1_3,1_4)*L(2_1)*(D L)(3_1,3_2,3_1~)*eC(4_1)*
    ↪ dC+(5_1~)*dC+(6_1~)*uC+(7_1~)
    O(Lorentz,3): +eps(1_1,2_1)*eps(1_2,3_1)*eps(1_3,4_1)*eps(1_4,4_2)*eps(4_1~,6_1
    ↪ ~)*eps(5_1~,7_1~) * CL(1_1,1_2,1_3,1_4)*L(2_1)*L(3_1)*(D eC)(4_1,4_2,4_1~)*
    ↪ dC+(5_1~)*dC+(6_1~)*uC+(7_1~)
    O(Lorentz,4): +eps(1_1,2_1)*eps(1_2,3_1)*eps(1_3,4_1)*eps(1_4,4_2)*eps(4_1~,5_1
    ↪ ~)*eps(6_1~,7_1~) * CL(1_1,1_2,1_3,1_4)*L(2_1)*L(3_1)*(D eC)(4_1,4_2,4_1~)*
    ↪ dC+(5_1~)*dC+(6_1~)*uC+(7_1~)
    O(Lorentz,5): -eps(1_1,2_1)*eps(1_2,3_1)*eps(1_3,4_1)*eps(1_4,5_1)*eps(5_1~,6_1
    ↪ ~)*eps(5_2~,7_1~) * CL(1_1,1_2,1_3,1_4)*L(2_1)*L(3_1)*eC(4_1)*(D dC+)(5_1,5
    ↪ _1~,5_2~)*dC+(6_1~)*uC+(7_1~)
    O(Lorentz,6): +eps(1_1,2_1)*eps(1_2,3_1)*eps(1_3,4_1)*eps(1_4,6_1)*eps(5_1~,6_1
    ↪ ~)*eps(6_2~,7_1~) * CL(1_1,1_2,1_3,1_4)*L(2_1)*L(3_1)*eC(4_1)*dC+(5_1~)*(D
    ↪ dC+)(6_1,6_1~,6_2~)*uC+(7_1~)
    O(Lorentz,7): -eps(1_1,2_1)*eps(1_2,3_1)*eps(1_3,4_1)*eps(1_4,7_1)*eps(5_1~,7_1
    ↪ ~)*eps(6_1~,7_2~) * CL(1_1,1_2,1_3,1_4)*L(2_1)*L(3_1)*eC(4_1)*dC+(5_1~)*dC
    ↪ +(6_1~)*(D uC+)(7_1,7_1~,7_2~)
  SU3:
    O(SU3,1): +eps(5_1,6_1,7_1) * CL*L*L*eC*dC+(5_1)*dC+(6_1)*uC+(7_1)
  SU2:
    O(SU2,1): +eps(2_1,3_1) * CL*L(2_1)*L(3_1)*eC*dC+*dC+*uC+
permutation_symmetries:
- vector: Lorentz * SU3 * SU2
- symmetry: {CL: [1], L: [2], eC: [1], dC+: [2], uC+: [1]}
  n_terms: 3
  n_operators: 972
  matrix: |-
    [ 0  0  1  1 -1 -1  2]
    [ 0  0  2  2  0  0  0]
    [ 0  0  0  0  2  2  0]
- symmetry: {CL: [1], L: [2], eC: [1], dC+: [1, 1], uC+: [1]}
  n_terms: 2
  n_operators: 324
  matrix: |-
    [ 0  0  1 -1 -1  1  0]
    [ 0  0  2 -2  0  0  0]
- symmetry: {CL: [1], L: [1, 1], eC: [1], dC+: [2], uC+: [1]}
  n_terms: 1
  n_operators: 162
  matrix: |-
    [ 2  2 -1 -1  1  1 -2]
- symmetry: {CL: [1], L: [1, 1], eC: [1], dC+: [1, 1], uC+: [1]}
  n_terms: 1
  n_operators: 81
  matrix: |-
    [ 2 -2 -1  1  1 -1  0]
```

Listing 4: Content of the file `1CL_2L_1eC_2dC+_1uC+_1D.yml`, which encodes the operators of (52).

As indicated in Table I, the symbol `CL` denotes the left-handed Weyl tensor C_{L} defined in Eq. (27). From the examples above, the reader should by now be able to reconstruct all operators from this listing.

TABLE II. SMEFT families at mass dimension 12. The complete table can be found in smeft_12.tar.xz as table12.pdf.

Family	Types	Terms	Operators
${F_\mathrm{L}}^2\phi^2 F_\mathrm{R} D^4 + \mathrm{H.c.}$	22	422	422
${F_\mathrm{R}}^5 D^2 + \mathrm{H.c.}$	22	78	78
$\phi^4 F_\mathrm{R} D^6 + \mathrm{H.c.}$	4	96	96
⋮	⋮	⋮	⋮
$F_\mathrm{L}\psi^4 {F_\mathrm{R}}^2 + \mathrm{H.c.}$	128	1590	61398
$\psi^4\phi^2 F_\mathrm{R} D^2 + \mathrm{H.c.}$	58	4504	161772
257	11942	472645	75577476

TABLE III. GRSMEFT⊖SMEFT families at mass dimension 12. The complete table can be found in grsmeft_12.tar.xz as table12.pdf.

Family	Types	Terms	Operators
$C_\mathrm{L}{F_\mathrm{L}}^2\phi^2 D^4 + \mathrm{H.c.}$	8	160	160
$C_\mathrm{L}F_\mathrm{L}{F_\mathrm{R}}^3 D^2 + \mathrm{H.c.}$	24	96	96
${C_\mathrm{L}}^2 F_\mathrm{L} F_\mathrm{R} C_\mathrm{R} D^2 + \mathrm{H.c.}$	6	24	24
⋮	⋮	⋮	⋮
${C_\mathrm{L}}^3\phi^2 {F_\mathrm{R}}^2 + \mathrm{H.c.}$	8	8	8
${F_\mathrm{L}}^2\phi\psi^{\dagger 2} F_\mathrm{R} C_\mathrm{R} + \mathrm{H.c.}$	94	150	1350
460	6097	70528	3936965

E. Ancillary files

The complete set of SMEFT and GRSMEFT operators up to mass dimension 12 with three generations of fermions are published as Supplemental Material [28] together with this paper. In addition, we provide a table with the numbers of *types*, *terms*, and *operators* contained in a given family (see Tables II and III as an example) as well as the associated Hilbert series for each mass dimension. This information can be extracted solely from the attached operator files.

The operators are provided in the format described in Secs. III. A and III. B. The files are included in the archives smeft_<ndim>.tar.xz and grsmeft_<ndim>.tar.xz, where <ndim> is the mass dimension.[15] They are collected in directories as

```
operators/<N>/<family>/<type>.yml
```

where <N>, <family>, and <type> are the number of fields N, the family and the type of the operator as defined in Sec. II. B. Concrete examples are given in Secs. III. A to III. D. This format makes it simple to access specific operators in the set.

For SMEFT up to mass dimension 12, the size of the files amounts to about 442 MB. This number seems to grow roughly with the number of operators, from which we infer that it will reach 1 TB at mass dimension 20, and 1 PB at mass dimension 26.

The Hilbert series provides the number of operators per type and thus constitutes a helpful check on our results. We have compared our numbers for SMEFT and GRSMEFT up to dimension 12 to the results for the Hilbert series as obtained by ECO [12] and found full agreement. We have further performed a number of consistency checks on our results, for example that no two $\mathcal{K}$ matrices for a particular operator type contain linearly dependent rows, etc. An immediate comparison of the results at lower mass dimension to the existing literature is highly nontrivial though, due to different representations of the final basis [3,4,9].[16] A general conversion tool which would allow such comparisons is currently under development [30].

IV. CONCLUSIONS

We have evaluated the SMEFT and GRSMEFT operator basis with three generations of fermions up to mass dimension 12. They were obtained by reimplementing the algorithm of Refs. [3,4,17] into a noncommercial software. Aside from the results obtained in this paper, we also confirm the completeness of the algorithm up to mass dimension 12. The operators are provided in the form of searchable and compact Supplemental Material [28].

In future work, we plan to extend the capabilities of AutoEFT in various respects, for example to keep operators that vanish by equations of motion [31], as they are needed for the renormalization of the operators. Furthermore, we plan to implement general basis transformations in the spirit of Refs. [32,33], which would allow the fully automated matching of EFTs to an ultraviolet complete theory, for example by combining AutoEFT with methods like UOLEA [34–37] (see also Ref. [38]).

ACKNOWLEDGMENTS

We would like to thank Svenja Diekmann, Jakob Linder, and Maximilian Rzehak for constructive discussions. This research was supported by the Deutsche Forschungsgemeinschaft (DFG) under Grant No. 400140256—GRK 2497: The physics of the heaviest particles at the LHC, and Grant No. 396021762—TRR 257: P3H—Particle Physics Phenomenology after the Higgs Discovery.

[15] Note that the file smeft_<ndim>.tar.xz contains the complete set of SMEFT operators for mass dimension <ndim> while grsmeft_<ndim>.tar.xz contains only operators including the Weyl tensor. The full GRSMEFT basis is recovered by the union of smeft_<ndim>.tar.xz and grsmeft_<ndim>.tar.xz.

[16] This is even true for the comparison with the earlier results based on the same algorithm, because further manipulations have been applied to the final results in Refs. [3,4] in order to present them in a more compact form.

[1] C. Arzt, Reduced effective lagrangians, Phys. Lett. B **342**, 189 (1995).

[2] J. C. Criado and M. Pérez-Victoria, Field redefinitions in effective theories at higher orders, J. High Energy Phys. 03 (2019) 038.

[3] H.-L. Li, Z. Ren, J. Shu, M.-L. Xiao, J.-H. Yu, and Y.-H. Zheng, Complete set of dimension-eight operators in the standard model effective field theory, Phys. Rev. D **104**, 015026 (2021).

[4] H.-L. Li, Z. Ren, M.-L. Xiao, J.-H. Yu, and Y.-H. Zheng, Complete set of dimension-nine operators in the standard model effective field theory, Phys. Rev. D **104**, 015025 (2021).

[5] W. Buchmüller and D. Wyler, Effective lagrangian analysis of new interactions and flavour conservation, Nucl. Phys. **B268**, 621 (1986).

[6] B. Grzadkowski, M. Iskrzyński, M. Misiak, and J. Rosiek, Dimension-six terms in the Standard Model Lagrangian, J. High Energy Phys. 10 (2010) 085.

[7] L. Lehman, Extending the standard model effective field theory with the complete set of dimension-7 operators, Phys. Rev. D **90**, 125023 (2014).

[8] Y. Liao and X.-D. Ma, Renormalization group evolution of dimension-seven baryon- and lepton-number-violating operators, J. High Energy Phys. 11 (2016) 043.

[9] C. W. Murphy, Dimension-8 operators in the Standard Model Effective Field Theory, J. High Energy Phys. 10 (2020) 174.

[10] Y. Liao and X.-D. Ma, An explicit construction of the dimension-9 operator basis in the standard model effective field theory, J. High Energy Phys. 11 (2020) 152.

[11] B. Henning, X. Lu, T. Melia, and H. Murayama, 2, 84, 30, 993, 560, 15456, 11962, 261485, …: Higher dimension operators in the SM EFT, J. High Energy Phys. 08 (2017) 016; 09 (2019) 019.

[12] C. B. Marinissen, R. Rahn, and W. J. Waalewijn, …, 83106786, 114382724, 1509048322, 2343463290, 27410087742, … efficient Hilbert series for effective theories, Phys. Lett. B **808**, 135632 (2020).

[13] R. M. Fonseca, The Sym2Int program: Going from symmetries to interactions, J. Phys. Conf. Ser. **873**, 012045 (2017).

[14] R. M. Fonseca, Enumerating the operators of an effective field theory, Phys. Rev. D **101**, 035040 (2020).

[15] R. M. Fonseca, GroupMath: A MATHEMATICA package for group theory calculations, Comput. Phys. Commun. **267**, 108085 (2021).

[16] R. M. Fonseca, Calculating the renormalisation group equations of a SUSY model with Susyno, Comput. Phys. Commun. **183**, 2298 (2012).

[17] H.-L. Li, Z. Ren, M.-L. Xiao, J.-H. Yu, and Y.-H. Zheng, Operators for generic effective field theory at any dimension: On-shell amplitude basis construction, J. High Energy Phys. 04 (2022) 140.

[18] H.-L. Li, Z. Ren, M.-L. Xiao, J.-H. Yu, and Y.-H. Zheng, Operator bases in effective field theories with sterile neutrinos: $d \leq 9$, J. High Energy Phys. 11 (2021) 003.

[19] A. M. Galda, M. Neubert, and S. Renner, ALP—SMEFT interference, J. High Energy Phys. 06 (2021) 135.

[20] M. Ruhdorfer, J. Serra, and A. Weiler, Effective field theory of gravity to all orders, J. High Energy Phys. 05 (2020) 083.

[21] A. Kobach, Baryon number, lepton number, and operator dimension in the Standard Model, Phys. Lett. B **758**, 455 (2016).

[22] A. Helset and A. Kobach, Baryon number, lepton number, and operator dimension in the SMEFT with flavor symmetries, Phys. Lett. B **800**, 135132 (2020).

[23] E. E. Jenkins and A. V. Manohar, Algebraic structure of lepton and quark flavor invariants and CP violation, J. High Energy Phys. 10 (2009) 094.

[24] The Sage Developers, SageMath, the Sage Mathematics Software System (Version 9.7), 2022, https://www.sagemath.org.

[25] J. A. M. Vermaseren, New features of FORM, arXiv:math-ph/0010025.

[26] J. Kuipers, T. Ueda, J. A. M. Vermaseren, and J. Vollinga, FORM version 4.0, Comput. Phys. Commun. **184**, 1453 (2013).

[27] GNU General Public License, version 3, http://www.gnu.org/licenses/gpl.html, 2007, Last retrieved 2020-01-01.

[28] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.108.055020 for the full set of SMEFT and GRSMEFT operators up to mass dimension 12.

[29] G. Durieux and C. S. Machado, Enumerating higher-dimensional operators with on-shell amplitudes, Phys. Rev. D **101**, 095021 (2020).

[30] R. V. Harlander, J. W. Linder, M. Rzehak, and M. C. Schaaf (to be published).

[31] Z. Ren and J.-H. Yu, A complete set of the dimension-8 Green's basis operators in the standard model effective field theory, arXiv:2211.01420.

[32] A. Falkowski, B. Fuks, K. Mawatari, K. Mimasu, F. Riva, and V. Sanz, Rosetta: An operator basis translator for standard model effective field theory, Eur. Phys. J. C **75**, 583 (2015).

[33] F. Sannino, P. Stangl, D. M. Straub, and A. E. Thomsen, Flavor physics and flavor anomalies in minimal fundamental partial compositeness, Phys. Rev. D **97**, 115046 (2018).

[34] A. Drozd, J. Ellis, J. Quevillon, and T. You, The universal one-loop effective action, J. High Energy Phys. 03 (2016) 180.

[35] B. Summ and A. Voigt, Extending the universal one-loop effective action by regularization scheme translating operators, J. High Energy Phys. 08 (2018) 026.

[36] M. Krämer, B. Summ, and A. Voigt, Completing the scalar and fermionic universal one-loop effective action, J. High Energy Phys. 01 (2020) 079.

[37] J. Fuentes-Martín, M. König, J. Pagès, A. E. Thomsen, and F. Wilsch, A proof of concept for matchete: An automated tool for matching effective theories, Eur. Phys. J. C **83**, 662 (2023).

[38] A. Carmona, A. Lazopoulos, P. Olgoso, and J. Santiago, Matchmakereft: Automated tree-level and one-loop matching, SciPost Phys. **12**, 198 (2022).