Beyond Jarlskog: 699 invariants for CP violation in SMEFT

Quentin Bonnefoy,[a] **Emanuele Gendy,**[a,b] **Christophe Grojean**[a,c]
and Joshua T. Ruderman[d]

[a]*Deutsches Elektronen-Synchrotron DESY,*
Notkestr. 85, Hamburg 22607, Germany

[b]*Institute of Theoretical Physics, Universität Hamburg,*
Hamburg 22761, Germany

[c]*Institut für Physik, Humboldt-Universität zu Berlin,*
Berlin 12489, Germany

[d]*Center for Cosmology and Particle Physics, Department of Physics, New York University,*
New York, NY 10003, U.S.A.

E-mail: quentin.bonnefoy@desy.de, emanuele.gendy@desy.de,
christophe.grojean@desy.de, ruderman@nyu.edu

ABSTRACT: As SMEFT is a framework of growing importance to analyze high-energy data, understanding its parameter space is crucial. The latter is commonly split into CP-even and CP-odd parts, but this classification is obscured by the fact that CP violation is actually a collective effect that is best captured by considering flavor-invariant combinations of Lagrangian parameters. First we show that fermion rephasing invariance imposes that several coefficients associated to dimension-six operators can never interfere with operators of dimension ≤ 4 and thus cannot appear in any physical observable at $\mathcal{O}1/\Lambda^2$. For those that can, instead, we establish a one-to-one correspondence with CP-odd flavor invariants, all linear with respect to SMEFT coefficients. We explicitly present complete lists of such linear CP-odd invariants, and carefully examine their relationship to CP breaking throughout the parameter space of coefficients of dimension ≤ 4. Requiring that these invariants all vanish, together with the Jarlskog invariant, the strong-CP phase, and the 6 CP-violating dimension-6 bosonic operators, provides $699(+1 + 1 + 6)$ conditions for CP conservation to hold in any observable at leading order, $\mathcal{O}\left(1/\Lambda^2\right)$.

KEYWORDS: CP Violation, SMEFT, Flavour Symmetries

ARXIV EPRINT: 2112.03889

Contents

1 Introduction

The predictions of the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge-invariant operators of dimension four or less built with the field content of the Standard Model, which we denote SM_4, have been remarkably confirmed by experiments. In particular, the observed pattern of CP-violation (CPV) [1] remarkably confirms the Cabibbo-Kobayashi-Maskawa (CKM) description of three fermionic generations of matter [2, 3]. Especially, two key aspects of CPV as we know it, namely the facts that it is quite suppressed despite an $\mathcal{O}(1)$ phase and that it demands the existence of at least three generations, are explained in the CKM formalism by its collective origin: CPV arises in the SM_4 from the interplay between the gauge and Yukawa sectors [4, 5]. In other words, CP is broken by the simultaneous presence of non-vanishing physical parameters, which makes the breaking accidentally small in the SM_4. The notion of collective breaking has application beyond CPV, for instance in little Higgs models [6–8].

The two notions of "simultaneous presence" and "physical parameter" in the previous paragraph suggest that a measure of CPV should be carefully defined, in particular when expressed in terms of Lagrangian parameters. An efficient way to do this is to use flavor-invariants, namely quantities which do not depend on the precise labeling of each fermion generation. As we will review below, a single quantity suffices to describe CPV in the fermionic sector of the SM_4,

$$J_4 \equiv \operatorname{Im} \operatorname{Tr} \left[Y_u Y_u^\dagger, Y_d Y_d^\dagger \right]^3 = 3 \operatorname{Im} \operatorname{Det} \left[Y_u Y_u^\dagger, Y_d Y_d^\dagger \right] . \tag{1.1}$$

where $Y_{u,d}$ are the 3×3 up and down quark Yukawa matrices in the SM_4 Lagrangian. The quantity J_4 goes under the name of Jarlskog invariant [9–11], and vanishes iff. CP is conserved. The structure of J_4 is such that it is not modified by unitary reshuffling of the quark fields, which means that it corresponds to a physical quantity. In addition, its expression shows that CPV in the SM_4 is not a feature of Y_u or Y_d alone, but a feature of the whole model which can only be assessed with the knowledge of both matrices (and in particular, of the fact whether they can be simultaneously diagonalized). This "collective" property of CPV, namely the fact that it depends on several Lagrangian parameters at once, is a key property of the SM_4, as well as of its extensions. This also holds for strong CPV, whose order parameter is given by $\theta_{\mathrm{QCD}} - \arg \det (Y_u Y_d)$, and resides simultaneously within the θ-term of QCD and within the quark Yukawa matrices (see appendix F for more details).

While J_4 is the only order parameter of CPV in the fermionic sector of the SM_4, additional ones need to be specified whenever the SM_4 is extended, see e.g. refs. [12–17] for multi-Higgs doublet models, refs. [18–20] for the case of supersymmetric extensions of the SM_4, refs. [11, 21–23] for the case of additional generations of matter, refs. [24–32] for the inclusion of neutrino masses, ref. [33] for vector-like extensions, ref. [34] for CP-violating ALP EFTs or ref. [35] for models of spontaneous CP breaking. In this paper, we assume that there are no new light degrees of freedom (d.o.f.) below, or close to, the weak scale, but we remain agnostic about the presence of heavy states. In that case, the SM_4 should be understood as the low-energy approximation of some fundamental UV dynamics, i.e. it

becomes necessary to extend it into an effective field theory (EFT). Under the assumption that a decoupling limit can be consistently taken, the adequate description is the so-called Standard Model Effective Field Theory (SMEFT) [36]. There, the dynamics of the $\mathrm{SM_4}$ d.o.f. are derived from the $\mathrm{SM_4}$ Lagrangian supplemented by a tower of higher-dimensional operators,

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM_4}} + \sum_i \frac{C_i}{\Lambda^{d_i-4}} \mathcal{O}_i \,, \tag{1.2}$$

where $\mathcal{O}_i$ is a local operator of dimension $d_i > 4$, C_i a complex coefficient (sometimes called Wilson coefficient) generically of order one (modulo possible selection rules and after taking $\hbar$ dimensions into account) and Λ is a dimensionful scale associated to heavy new physics. If large enough, Λ allows for a power expansion of any phenomenological prediction in inverse powers of itself. Complete bases of operators up to dimension 8 can be found in refs. [37–39].

The coefficients C_i are generically complex and introduce a large number of new sources of CPV in SMEFT (see the counting in ref. [40] at dimension $d_i \leq 6$). In this paper, we are interested in CPV associated to flavorful Wilson coefficients, whose analysis requires the careful extraction of basis-independent physical parameters, which account for the collective properties of CPV. Therefore, the new CPV phases should be captured by CP-odd flavor-invariants, similar to J_4. We focus here on the CPV phases found in the fermionic sector, since the bosonic sources of CPV in the dimension-six SMEFT are trivially flavor-invariant.[1] Moreover, we focus on the limit of vanishing neutrino masses throughout the whole work.

One could ask: why do we need invariant quantities? For instance, it is common to associate a complex top-quark Yukawa coupling with a new source of CPV, without referring to invariants. However a complex top Yukawa only signals CPV if one works in a given flavor basis where the top is a mass eigenstate of real mass. One may wonder how to describe CPV in SMEFT, without any specific reference to the IR physics, such as the masses or the electroweak vacuum. This picture is for instance justified if one cares about new sources of CPV which arise from the matching to a given UV model, which should be analyzed at an energy scale above the electroweak vacuum expectation value (vev) and should not depend on the details of the IR dynamics, such as specific flavor bases motivated by low energy considerations. Flavor-invariants are well-suited to answer such questions, as they allow one to capture physical and collective properties of the model, to parametrize CP-odd observables in a basis-independent way, and also to make the matching with UV models and their properties easier, by decorrelating the parametrization of CPV quantities from flavor bases connected to the IR properties of $\mathrm{SM_4}$ particles.

Let us use the aforementioned example of the top-quark complex Yukawa coupling to offer a preview of the flavor-invariants considered in this paper. As said above, a top-Higgs Lagrangian with a complex Yukawa coupling

$$-\mathcal{L} \supset m_t \bar{t}t + \frac{m_t}{v}\bar{t}(\kappa_t + i\tilde{\kappa}_t \gamma_5)th = m_t \bar{t}_L t_R + \frac{m_t}{v}\bar{t}_L(\kappa_t + i\tilde{\kappa}_t)t_R h + \mathrm{h.c.}\,, \tag{1.3}$$

---

[1]For the 6 CP-odd bosonic operators appearing in the basis in ref. [37], indeed, the condition for CP conservation is simply for their coefficient to vanish.

violates CP. It can originate in SMEFT from the dimension-four and six Yukawa couplings,[2]

$$- \mathcal{L} \supset y_t \bar{Q}_L t_R \tilde{H} + \frac{C_{tH}}{\Lambda^2} \bar{Q}_L t_R \tilde{H} |H|^2 + \text{h.c.} . \tag{1.4}$$

The above expression generalizes to three generations of matter in SMEFT upon replacing $Q_L, t_R, y_t, C_{tH} \to Q_{L,i}, t_{R,j}, Y_{u,ij}, C_{uH,ij}$. Focusing on the diagonal entries of C_{uH} in a basis where Y_u is diagonal and real (with non-degenerate entries, as is experimentally relevant), their three imaginary parts $\text{Im}\, C_{uH,ii}$ violate CP. They can be captured by three independent flavor-invariants, whose expressions read

$$L_{k=1,2,3} = \text{Im Tr}\Big(X_u^{k-1} C_{uH} Y_u^\dagger \Big) \tag{1.5}$$

where $X_u \equiv Y_u Y_u^\dagger$. In the basis where $Y_u = \text{diag}\, y_{u,i}$ is diagonal and real, one has $L_{k=1,2,3} = y_{u,i}^{2k-1} \text{Im}\, C_{uH,ii}$, with an implicit sum over i. Therefore, at fixed, non-vanishing and non-degenerate $y_{u,i}$, the set of three $L_{k=1,2,3}$ maps to that of three $C_{uH,ii=1,2,3}$ in a bijective fashion, hence those three invariants capture the three new sources of CPV associated to up quark complex Yukawa couplings, as in eq. (1.3). In the generic case however, C_{uH} has off-diagonal entries even in the basis where Y_u is diagonal, all of which can be complex, such that one needs nine flavor-invariants to capture the nine new sources of CPV in C_{uH}. Although there is no unique choice, one possible set of invariants reads

$$L_{k=1,\dots,9} = \left\{ \begin{array}{lll} \text{Im Tr}\left(C_{uH} Y_u^\dagger \right) & \text{Im Tr}\left(X_u X_d C_{uH} Y_u^\dagger \right) & \text{Im Tr}\left(X_d^2 X_u^2 C_{uH} Y_u^\dagger \right) \\[2mm] \text{Im Tr}\left(X_u C_{uH} Y_u^\dagger \right) & \text{Im Tr}\left(X_d X_u C_{uH} Y_u^\dagger \right) & \text{Im Tr}\left(X_u X_d^2 X_u^2 C_{uH} Y_u^\dagger \right) \\[2mm] \text{Im Tr}\left(X_d C_{uH} Y_u^\dagger \right) & \text{Im Tr}\left(X_u^2 X_d^2 C_{uH} Y_u^\dagger \right) & \text{Im Tr}\left(X_d X_u^2 X_d^2 C_{uH} Y_u^\dagger \right) \end{array} \right\} \tag{1.6}$$

where $X_d \equiv Y_d Y_d^\dagger$.[3]

Naively, the number of new flavor-invariants should match that of the new sources of CPV. However, observables in SMEFT are truncated at a given order in inverse powers of Λ, according to the SMEFT power counting, and it happens that not all sources of CPV contribute to physical observables at this given order as a result of non-interference. In this paper, we illustrate this fact by discussing CPV observables truncated at the leading $1/\Lambda^2$

---

[2]At order $1/\Lambda^2$ (and for one generation only) the correspondence between the different coefficients reads [41, 42]

$$m_t = \frac{y_t v}{\sqrt{2}} \left(1 + \frac{1}{2} \frac{\text{Re}\, C_{tH} v^2}{y_t \Lambda^2} \right) , \quad \kappa_t + i\tilde{\kappa}_t = 1 + \frac{\text{Re}\, C_{tH}}{y_t} \frac{v^2}{\Lambda^2} + i \frac{\text{Im}\, C_{tH}}{y_t} \frac{v^2}{\Lambda^2}$$

provided we start in a basis where y_t is already real.

[3]The careful reader may note that the former L_3 in eq. (1.5) does not appear anymore in eq. (1.6). We have removed L_3 because it is not independent of the first two L's when two up-type quark masses are degenerate. This does not happen for any of the invariants in eq. (1.6), which are therefore preferred (see the following sections for a systematic treatment).

order, to which several of the new sources of CPV at dimension-six cannot contribute.[4] We therefore carefully differentiate between the power counting for observables, which we truncate at order $1/\Lambda^2$, and that of SMEFT operators, which we only include up to dimension-six, i.e. also up to order $1/\Lambda^2$. As we will explain, not all associated SMEFT coefficients can interfere with the SM_4 contribution to a given observable, and therefore they cannot all contribute at leading $1/\Lambda^2$ order to observables. We dub those which can *primary coefficients*, while we refer to the others as *secondary coefficients*. We perform the counting of the number of (both CP-even and CP-odd) SMEFT primary coefficients. Among those, the CP-odd fermionic ones, whose number is 699, are captured by flavor-invariants linear with respect to the dimension-six SMEFT coefficients, and an explicit and complete set of such flavor invariants is built. Consequently, we present a necessary and sufficient set of flavor-invariants, such that CP is conserved at $\mathcal{O}(1/\Lambda^2)$ iff. they vanish, together with J_4, the strong-CP phase, and the 6 CP-violating dimension-6 bosonic operators, so that they form a set of $699(+1+1+6)$ order parameters of CPV.

The rest of this paper is organized as follows. In section 2, we review the collective nature of CP breaking in the SM_4 (in particular the structure of the aforementioned Jarlskog invariant), introducing concepts useful for the following, such as flavor bases and flavor invariants. We additionally emphasize that collective effects also arise beyond the SM_4, in SMEFT in particular. In section 3, we refine the counting of dimension-six parameters of SMEFT, taking into account the interplay between the SMEFT power counting and the requirement of rephasing invariance of physical quantities. This reduces the relevant parameters to the primary ones. Focusing on new sources of CPV, we then present our strategy to capture them thanks to flavor-invariants linear in the SMEFT dimension-six coefficients. That this can be achieved is shown explicitly in section 4, where we discuss explicit sets of CP-odd invariants for SMEFT at dimension-six. We emphasize that they need to be carefully chosen to capture all new sources of CPV for all parts of the SM_4 parameter space, ignoring for conciseness the case of vanishing quark masses which is subsequently treated in an appendix. We finally present conclusions and future directions in section 5. Appendix A discusses flavor symmetries of the SM_4 in all relevant parts of its parameter space, which is important to determine adequate sets of CP-odd invariants as discussed in section 4. The case of vanishing quark masses is discussed at the end of this appendix. In appendix B, we consider generic properties of flavor-invariants with three generations, in particular algebraic relations between them which are consistent with the counting of primary parameters. In appendix C, the counting of primary parameters, performed in section 3 for three generations of matter, is generalized to any number of generations. Appendices D and E then contain a full list of linear CP-odd invariants (for operators respectively bilinear and quartic in fermion fields), which map to all independent primary Lagrangian parameters. Finally, appendix F includes θ_{QCD}, which allows one to build flavor-invariants with new algebraic structures, but does not suffice to increase the number of primary coefficients.

---

[4]Such CP-odd observables are at most linear in the dimension-six SMEFT coefficients, and correspond to the interference between the SM_4 and the leading SMEFT contributions to a given amplitude. A more thorough characterization of the observables we consider can be found in section 3.1.

2 The collective nature of CP breaking in the SM(EFT)

In order to motivate why we define CP-odd invariants, it is useful to review first one important and interesting aspect of CP breaking in SMEFT: it is collective. Indeed, it relies on the simultaneous presence of several complex parameters in the Lagrangian, which cannot all be made simultaneously real, even using the freedom to redefine fields (or equivalently, to define appropriately the CP transformation). In this section we review CP violation in SM_4, in order to establish our conventions and present several of the claims related to CP violation which will be repeatedly encountered in this paper.

2.1 CP-violation and complex parameters

The usual lore is that complex parameters in the Lagrangian violate CP. At the level of the fermionic Lagrangian, this claim leaves implicit crucial subtleties related to field redefinitions. The correct statement is instead that *the Lagrangian is CP-symmetric iff. one can redefine the fields so as to make all couplings real.*[5] In the SM_4, this explains why only one phase out of the six naively contained in the CKM matrix is physical and breaks CP. For instance, were the CKM matrix equal to the following unitary matrix

$$V_{\mathrm{CKM}} = \begin{pmatrix} \frac{72-21i}{325} & \frac{4}{13} & -\frac{12i}{13} \\ -\frac{12}{13} & \frac{576+168i}{1625} & \frac{49-168i}{1625} \\ -\frac{96-28i}{325} & -\frac{57}{65} & -\frac{24i}{65} \end{pmatrix}, \qquad (2.1)$$

it would not violate CP, although it is explicitly complex. Indeed, one can write

$$\begin{pmatrix} \frac{72-21i}{325} & \frac{4}{13} & -\frac{12i}{13} \\ -\frac{12}{13} & \frac{576+168i}{1625} & \frac{49-168i}{1625} \\ -\frac{96-28i}{325} & -\frac{57}{65} & -\frac{24i}{65} \end{pmatrix} = \begin{pmatrix} \frac{3-4i}{5} & 0 & 0 \\ 0 & \frac{4-3i}{5} & 0 \\ 0 & 0 & \frac{3-4i}{5} \end{pmatrix} \begin{pmatrix} \frac{3}{13} & \frac{4}{13} & \frac{12}{13} \\ -\frac{12}{13} & \frac{24}{65} & \frac{7}{65} \\ -\frac{4}{13} & -\frac{57}{65} & \frac{24}{65} \end{pmatrix} \begin{pmatrix} \frac{4+3i}{5} & 0 & 0 \\ 0 & \frac{3+4i}{5} & 0 \\ 0 & 0 & \frac{4-3i}{5} \end{pmatrix},$$

$$\qquad (2.2)$$

and absorb all the factorized diagonal phases into the fermion fields, in order to obtain a real orthogonal CKM matrix. Such a manipulation cannot be done for the following matrix,

$$V_{\mathrm{CKM}} = \begin{pmatrix} \frac{2172-5004i}{8125} & -\frac{1784+432i}{8125} & -\frac{2427+5196i}{8125} \\ -\frac{3747+3996i}{8125} & \frac{3324+912i}{8125} & \frac{4772-1164i}{8125} \\ -\frac{308+144i}{1105} & -\frac{4389+2052i}{5525} & \frac{1848+864i}{5525} \end{pmatrix}. \qquad (2.3)$$

However, whether it yields a CPV SM_4 depends on the fermion spectrum. Indeed, were two quark masses equal, the kinetic Lagrangian would have a U(2) flavor symmetry, allowing

---

[5]This is strictly speaking only true for models with continuous internal symmetries. When discrete symmetries are present, there exists the possibility that the couplings are pseudo-real, namely related to their complex conjugates via flavor transformations. Then one would get a CP-symmetric Lagrangian iff. there exists a flavor transformation which sends all couplings to their complex conjugates at once. See ref. [43] for an example, or the section 4.3 of ref. [44] for more details and references. In this text, we focus on the bulk of the SM(EFT) parameter space where any discrete symmetry is embedded into a continuous one (this is for instance automatic for non-degenerate spectra, see section 4.2 and appendix A).

for more general fermion field redefinitions. For instance, if $m_u = m_c$, we can redefine the first two flavors of up-type quarks in order to absorb the 2×2 unitary matrix which appears at the top left of the first factor on the right-hand-side of the following equality:

$$
\begin{pmatrix}
\frac{2172-5004i}{8125} & -\frac{1784+432i}{8125} & -\frac{2427+5196i}{8125} \\
-\frac{3747+3996i}{8125} & \frac{3324+912i}{8125} & \frac{4772-1164i}{8125} \\
-\frac{308+144i}{1105} & -\frac{4389+2052i}{5525} & \frac{1848+864i}{5525}
\end{pmatrix}
=
\begin{pmatrix}
-\frac{176+468i}{625} & -\frac{9-12i}{25} & 0 \\
\frac{351-132i}{625} & \frac{16+12i}{25} & 0 \\
0 & 0 & \frac{77+36i}{85}
\end{pmatrix}
\begin{pmatrix}
\frac{3}{13} & \frac{4}{13} & \frac{12}{13} \\
-\frac{12}{13} & \frac{24}{65} & \frac{7}{65} \\
-\frac{4}{13} & -\frac{57}{65} & \frac{24}{65}
\end{pmatrix},
$$

(2.4)

obtaining again a real orthogonal CKM matrix.

As is clear from these numerical examples, and as we will repeatedly illustrate, it is more convenient to phrase the condition for CP-violation in a way which does not require scanning over all possible field redefinitions. If the theory preserves CP, the following CP transformation of the (non-degenerate) fermionic mass eigenstates ψ (together with those of bosonic fields [4]) leaves the Lagrangian invariant in some field basis

$$
(\mathcal{CP})\psi(t, \vec{x})(\mathcal{CP})^{-1} = \gamma^0 C \overline{\psi}^T(t, -\vec{x}),
$$

(2.5)

where C is the (antisymmetric) charge conjugation matrix such that $\gamma^\mu C = -C(\gamma^\mu)^T$. As we anticipated, this implies that the Lagrangian couplings are real (in this field basis). For instance, if we assume that there exists the following coupling in the theory,

$$
\mathcal{L} \supset c_{1212} \left(\overline{\psi}_1 \gamma^\mu \psi_2 \right) \left(\overline{\psi}_1 \gamma_\mu \psi_2 \right) + \text{h.c.},
$$

(2.6)

we learn from the invariance under the CP transformation in eq. (2.5) that c_{1212} is real. However, the opposite statement is that the theory violates CP iff. the transformation in eq. (2.5) is never a symmetry, whatever the field basis chosen. This is not equivalent to saying that c_{1212} is complex in some basis, but that *whatever the basis chosen, there exists at least one Lagrangian parameter which is genuinely complex*[6] (which usually depends on the basis). Therefore, the condition for CPV which we look for takes the following schematic form

$$
\boxed{\text{CPV} \iff \text{Im} \,(\text{something independent of the field basis}) \neq 0.}
$$

Such a basis-independent object precisely defines a CP-odd flavor invariant. Within the framework of SMEFT, we can define flavor invariants order-by-order in the power counting. At leading order, the condition for CPV reads:

CPV at $\mathcal{O}(1/\Lambda^2) \iff \text{Im} \left(\text{something of } \mathcal{O}(1/\Lambda^2) \text{ independent of the field basis} \right) \neq 0.$

---

[6]For pseudo-real couplings, the statement is rather that all complex couplings cannot be turned simultaneously into their conjugates via the same change of basis.

	$\mathrm{SU}(3)_Q$	$\mathrm{SU}(3)_u$	$\mathrm{SU}(3)_d$	$\mathrm{SU}(3)_L$	$\mathrm{SU}(3)_e$
Y_u	**3**	$\bar{\mathbf{3}}$	**1**	**1**	**1**
Y_d	**3**	**1**	$\bar{\mathbf{3}}$	**1**	**1**
Y_e	**1**	**1**	**1**	**3**	$\bar{\mathbf{3}}$

Table 1. Flavor transformation properties of the Yukawa matrices treated as spurions.

2.2 Flavor transformations and flavor bases

As we just discussed, a meaningful statement about CP violation in the SM(EFT) must account for the possibility of field redefinitions. In addition, the SM(EFT) Lagrangian is naturally written in the unbroken electroweak phase, which does not differentiate between the three fermionic generations. Therefore, it should be possible to characterize CP violation without referring to any specific flavor labeling, in particular without identifying which combinations correspond to the mass eigenstates.

When the Lagrangian is written in terms of the gauge multiplets relevant in the unbroken phase, the kinetic Lagrangian in the fermion sector (including the gauge couplings) is invariant under a $\mathrm{U}(3)^5 = \mathrm{U}(3)_{Q_L} \times \mathrm{U}(3)_{u_R} \times \mathrm{U}(3)_{d_R} \times \mathrm{U}(3)_{L_L} \times \mathrm{U}(3)_{e_R}$ flavor group, where each factor acts on the flavor indices of the associated fermion field (we drop the chirality indices in what follows). This group is the largest under which all SMEFT coefficients can be assigned a spurious transformation so as to leave the full SM(EFT) Lagrangian unchanged.

Similarly to our conclusion in the previous section, the Lagrangian is CP-symmetric iff. one can redefine the fields so as to make all couplings real. When redefining the fermion fields by means of a $\mathrm{U}(3)^5$ flavor transformation, the precise values of all flavored couplings in SMEFT are mixed up, with the real and imaginary parts scrambled. Consequently, any order parameter of CP breaking cannot correspond to the imaginary part of a given coefficient, but instead should map to the imaginary part of a flavor-invariant combination of coefficients.

In order to build such invariants, it is useful to notice that the flavored SMEFT couplings transform under the flavor group as spurions with transformation properties which depend on the operator they couple to. For the Yukawa couplings at dimension-four, the transformations are as listed in table 1, and each (anti)fundamental representation has a charge $(-)1$ under the associated abelian group in the decomposition $\mathrm{U}(3)_X = \mathrm{SU}(3)_X \times \mathrm{U}(1)_X$, where $X = Q, u, d, L, e$.[7] Performing field redefinitions which belong to the flavor group, this set of spurious charges allows one to easily compute the couplings in the redefined theory, and to easily identify objects which are independent of such redefinitions.

Using flavor transformations, one can reach flavor bases where the Yukawa matrices have a specific form, and which we will sometimes use to explicitly evaluate invariants. For

---

[7]Out of the 5 U(1) factors only the gauged $\mathrm{U}(1)_Y$ and the combinations $\mathrm{U}(1)_{B-L}$ are conserved at the quantum level, while the other are broken by anomalies.

instance, using the singular value decomposition, we can choose that

$$Y_u = \text{diag}(y_u, y_c, y_t), \quad Y_d = V_{\text{CKM}} \cdot \text{diag}(y_d, y_s, y_b), \quad Y_e = \text{diag}(y_e, y_\mu, y_\tau), \qquad (2.7)$$

where all y's are real and positive and V_{CKM} is the Cabibbo-Kobayashi-Maskawa (CKM) matrix. We refer to this flavor basis as the *up basis*. Similarly, there exists a *down basis* where

$$Y_u = V_{\text{CKM}}^\dagger \cdot \text{diag}(y_u, y_c, y_t), \quad Y_d = \text{diag}(y_d, y_s, y_b), \quad Y_e = \text{diag}(y_e, y_\mu, y_\tau). \qquad (2.8)$$

Fixing this shape for the Yukawa couplings exhausts all flavor transformations but some diagonal ones.[8] If, in addition, a phase choice is made on the CKM matrix (for instance imposing that all its phases are given in terms of a single one as in usual parameterizations), no flavor freedom remains but the conserved baryon and lepton number symmetries $U(1)_B \times U(1)_L$. When we make such a choice below, we use the following parametrization of the CKM matrix [45],

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & c_{13}s_{12} & s_{13}e^{-i\delta_{\text{CKM}}} \\ -c_{23}s_{12} - c_{12}s_{13}s_{23}e^{i\delta_{\text{CKM}}} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta_{\text{CKM}}} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{\text{CKM}}} & -c_{12}c_{23} - c_{23}s_{12}s_{13}e^{i\delta_{\text{CKM}}} & c_{13}c_{23} \end{pmatrix}, \qquad (2.9)$$

where $c_X, s_X = \cos(\theta_X), \sin(\theta_X)$.

2.3 The collective nature of CP breaking in the SM_4

We now review in some detail the (well known) collective nature of CP breaking in the SM_4, which is useful for our purpose.

The fact that CP breaking is collective in the SM_4 can be understood from the fact that it preserves CP with up to two fermionic generations [2, 3], so that one needs the simultaneous presence of three generations to be sensitive to CP violation. A question is then: what is the order parameter of CP-breaking in the SM_4?

In order to answer this question unambiguously, one needs to mod out the impact of flavor transformations. Besides the use of invariants, a way to exhaust all flavor transformations is to work in a well-defined flavor basis, for instance in the *up* or *down basis* defined previously. Using the remaining phase rotations allowed in such bases removes all complex parameters but one, which fully specifies the flavor basis. The only leftover complex quantity can be written in a way which is independent of the phase rotations [9–11, 21, 46, 47],

$$\mathcal{J} = \text{Im}\left(V_{\text{CKM},us}V_{\text{CKM},cb}V_{\text{CKM},ub}^*V_{\text{CKM},cs}^*\right) = s_{12}c_{12}s_{13}c_{13}^2s_{23}c_{23}\sin(\delta_{\text{CKM}}), \qquad (2.10)$$

where the last equality uses eq. (2.9). It is straightforward to check that $\mathcal{J} = 0$ for the matrix in eq. (2.1).

---

[8]Their precise form depends on the basis. In the *down basis*, they are of the form $\text{diag}\left(e^{i\alpha_X^1}, e^{i\alpha_X^2}, e^{i\alpha_X^3}\right)$, such that $\alpha_Q^i = \alpha_d^i, \alpha_L^i = \alpha_e^i$ (RH up-quark phases are unconstrained).

As we saw in section 2.1, when two masses (of same-type quarks) are degenerate, there is a larger degeneracy of *up* or *down bases*, and one can further remove the leftover complex parameter from the SM$_4$ Lagrangian. Therefore, the genuine order parameter of CP breaking in the SM, J_4, is proportional to[9]

$$J_4 \propto (m_t^2 - m_c^2)(m_t^2 - m_u^2)(m_c^2 - m_u^2)(m_b^2 - m_s^2)(m_b^2 - m_d^2)(m_s^2 - m_d^2)\mathcal{J}\,. \tag{2.11}$$

One can show that there are no additional factors to J_4 [9].

We have constructed J_4 in a specific flavor basis, but it is useful to have expressions valid in all bases. In that respect, instead of looking for complex quantities invariant under mere phase rotations, one would rather consider invariants under the full flavor group. As we anticipated in the introduction, J_4, which goes under the name of Jarlskog invariant, then corresponds to [9]

$$J_4 \equiv \operatorname{Im} \operatorname{Tr}\left[Y_u Y_u^\dagger, Y_d Y_d^\dagger\right]^3 = 3 \operatorname{Im} \operatorname{Det}\left[Y_u Y_u^\dagger, Y_d Y_d^\dagger\right]\,. \tag{2.12}$$

Defined as above, J_4 is independent of the choice of flavor basis, as can be checked from the transformations in table 1. Evaluating eq. (2.12) for instance in the *up* or *down basis*, the connection with eq. (2.11) is made obvious,

$$J_4 = 6(y_t^2 - y_c^2)(y_t^2 - y_u^2)(y_c^2 - y_u^2)(y_b^2 - y_s^2)(y_b^2 - y_d^2)(y_s^2 - y_d^2)\mathcal{J}\,. \tag{2.13}$$

It can be shown that the statement that CP is broken in the SM$_4$ is equivalent to saying that J_4 does not vanish [9], therefore it corresponds to the genuine order parameter for CP breaking in the SM$_4$.

2.4 The collective nature of CP breaking beyond the SM$_4$

The search for flavor-invariant order parameters for CP breaking beyond the SM$_4$ is subject to very similar discussions. As an example, let us consider the SM$_4$ with a single generation of fermions, extended by a dimension-six Yukawa coupling for the up-type quark:

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}_4} + \frac{C_{uH}}{\Lambda^2}|H|^2 \overline{Q}_L u_R \tilde{H} + \text{h.c.}\,. \tag{2.14}$$

It is well known that such a coupling can generate a two-loop contribution to the electron EDM, a CPV observable, which reads at $\mathcal{O}(1/\Lambda^2)$ [48–50]

$$\frac{d_e}{e} = -\frac{1}{48\pi^2}\frac{v m_e m_u}{m_h^2}\frac{\operatorname{Im}(C_{uH})}{\Lambda^2}F_1\left(\frac{m_u^2}{m_h^2}, 0\right)\,, \tag{2.15}$$

where for conciseness we only kept the dominant contribution due to photons in the loop, where m_h and v are the Higgs mass and vev, respectively, and where

$$F_1(a, 0) = \int_0^1 dx \frac{\ln\left(\frac{a}{x(x-1)}\right)}{a - x(x-1)}\,. \tag{2.16}$$

---

[9]This expression depends on differences of squared masses and not, e.g., on $m_t - m_u$. This is due to the fact that only the modulus of a fermion mass is physical: any quark mass can be made complex by an appropriate rephasing of the associated RH field without changing other SM$_4$ couplings, hence we must consider the rephasing-invariant quantity $m_\psi m_\psi^* = |m_\psi|^2$ for any fermion ψ, which reduces to the mass squared in field bases where the mass is real.

The result in eq. (2.15) may suggest that $\text{Im}(C_{uH})$ acts as an order parameter of CP breaking in this theory.[10] However, this imaginary part could be rotated away by a chiral transformation of the up quark field (for instance by redefining $u_R \to e^{-i \arg(C_{uH})} u_R$), so one could wonder if there remains an observable electron EDM. The resolution to this puzzle is due to the implicit assumption that the computation is performed in a basis where the up quark has a real mass. In a generic flavor basis, the mass is complex,

$$\mathcal{L} \supset -m_u \overline{u}_L u_R - m_u^* \overline{u}_R u_L + \frac{v^2 C_{uH}}{\sqrt{2}\Lambda^2} \overline{u}_L u_R h + \frac{v^2 C_{uH}^*}{\sqrt{2}\Lambda^2} \overline{u}_R u_L h \,, \qquad (2.17)$$

where we wrote the Lagrangian in the broken phase, and a careful evaluation of the two-loop diagram (i.e. using propagators featuring complex fermion matrices) yields

$$\frac{d_e}{e} = -\frac{1}{48\pi^2} \frac{v m_e}{m_h^2} \frac{\text{Im}(m_u^* C_{uH})}{\Lambda^2} F_1 \left(\frac{|m_u|^2}{m_h^2}, 0 \right) \,, \qquad (2.18)$$

which matches eq. (2.15) when m_u is real, as it should. This expression allows us to identify a more satisfactory order parameter of CP-breaking, $\text{Im}(m_u^* C_{uH})$, which does not depend on the phase convention for the up quark. Similarly to what was discussed for the SM_4 previously, what matters here for CP breaking is not that the Yukawa coupling has an imaginary part, but that there is an irreducible imaginary part due to the simultaneous presence of both the coupling to the Higgs of the up quark and its non-zero mass. This provides also a qualitative argument for why the result in eq. (2.15) had to be explicitly proportional to m_u.

The take-away message of this section is that real or imaginary parts of coefficients are only meaningful with respect to CP breaking when the flavor basis is completely determined. In a general basis, what matters are the imaginary parts of invariant combinations of coefficients.[11] CP-odd invariants in SMEFT, especially with three flavors, are the subject of this paper.

---

[10]In the SM4, the various contributions to the electron EDM are all proportional to $\mathcal{J}$, consistently with the collective nature of CP breaking. Short-distance quark-level perturbative contributions arise at four-loop order and allow one to identify the whole structure of J_4 (see [51] for a parametric expression and [52] for a recent appraisal identifying an additional m_b^2 factor), but the dominant source comes from long-distance hadronic contributions sensitive to CPV four-fermion operators [53]. In addition, such hadronic contributions often dominate CPV observables, e.g. in the case of paramagnetic systems (see [54] for the identification of a new contribution which increases the previous result by 5 order of magnitude).

[11]Examples of this also exist in the SM4. The quantity ϵ, which encodes indirect CP violation in kaon decays, is sometimes written [1]

$$\epsilon \approx \frac{e^{i\pi/4}}{\sqrt{2}} \frac{\text{Im}(M_{12})}{\Delta m} \,,$$

where M_{12} is associated to $K^0 \leftrightarrow \overline{K^0}$ mixing and Δm is the mass difference between kaon mass eigenstates. This formula actually assumes that $\lambda_u \equiv V_{ud} V_{us}^*$ is real. The expression which is valid independently of the phase conventions reads [4]

$$\epsilon \approx \frac{e^{i\pi/4}}{\sqrt{2}} \frac{\text{Im}(M_{12} \lambda_u^2)}{\Delta m \, |\lambda_u|^2} \,.$$

3 Characterizing CP violation at dimension six

In this section, we discuss the number of primary parameters in SMEFT, as well as the parametrization of those which are CP-odd.

3.1 Primary parameters in the SMEFT

First, we count the number of flavorful primary SMEFT parameters. We remind that they are defined to be the dimension-six SMEFT parameters which generate BSM amplitudes which can interfere with SM_4 amplitudes (other parameters being called secondary). Indeed, observables computed in SMEFT are subject to a power expansion, with respect to which we focus on the leading BSM order, i.e. *we include contributions to observables up to order* $1/\Lambda^2$. For instance, in cross-section computations, we only consider the SM_4 amplitude squared and the interference with the leading BSM amplitude. Schematically, we can express a generic amplitude as

$$\mathcal{A} = \mathcal{A}^{(4)} + \mathcal{A}^{(6)} + \dots \tag{3.1}$$

where $\mathcal{A}^{(4)}$ is the leading order amplitude built with renormalizable operators, $\mathcal{A}^{(6)}$ is the next-to-leading order one, accompanied by a $1/\Lambda^2$ suppression, and the dots indicate higher order terms that we ignore. Then, observables such as cross sections, which are proportional to the amplitude squared, will receive contributions by

$$|\mathcal{A}|^2 = |\mathcal{A}^{(4)}|^2 + 2\,\mathrm{Re}\left(\mathcal{A}^{(4)}\mathcal{A}^{(6)*}\right) + \dots. \tag{3.2}$$

Our goal in this section is then to determine the primary parameter space that characterizes the first two terms in eq. (3.2).

This counting does not lead to a mere repetition of that of ref. [40], which counts primary and secondary parameters indifferently and whose results are reviewed in the first double column of table 2, due to the fact that several of the dimension-six parameters are charged under lepton numbers, unlike SM_4 parameters. Given that physical observables cannot be charged, such parameters can only interfere with themselves (or other charged BSM parameters) to form a neutral object, and are therefore secondary according to our classification.

More precisely, the free fermionic Lagrangian in the broken phase of the SMEFT and in the mass basis has abelian flavor symmetries $U(1)_{u_i, d_i, L_i}$, acting on each mass eigenstate independently. By definition, those do not affect the spectrum of asymptotic states and therefore do not affect physical predictions. They simply correspond to irreducible flavor ambiguities in a basis where mass matrices are diagonal and real, which must be fixed by further specifications (for instance, phase prescriptions in the CKM matrix). A consequence is that any observable must be expressed in terms of quantities which are invariant under these $U(1)$ phase rotations,[12] and therefore, any coefficient which is not invariant on its

---

[12]This is a slightly different statement than the one that CP violation should be characterized in a flavor-invariant way. Although they are restricted by flavor-invariant statements, amplitudes squared with flavored external states are not flavor-invariant, but they are always invariant under phase rotations.

	Type of op.	# of ops	# real	# im.	inv. under $U(1)_{L_i} - U(1)_{L_j}$	
					# real	# im.
bilinears	Yukawa	3	27	27	21	21
	Dipoles	8	72	72	60	60
	current-current	8	51	30	42	21
	all bilinears	19	150	129	123	102
4-Fermi	LLLL	5	171	126	99	54
	RRRR	7	255	195	186	126
	LLRR	8	360	288	246	174
	LRRL	1	81	81	27	27
	LRLR	4	324	324	216	216
	all 4-Fermi	25	1191	1014	774	597
	all		1341	1143	897	699

Table 2. Number of flavorful real and imaginary parameters in SMEFT at dimension-six (see tables 7 and 8 for the explicit forms of the operators). The first double column counts the number of physical parameters, the second one (highlighted in gray) counts those which are also primary (see the text).

own must enter observables multiplied by another U(1)-charged coefficient in order to form a neutral object. This story is known to readers familiar with the notion of rephasing-invariants of the CKM matrix [55]: $V_{\mathrm{CKM},ij}$ being charged under $U(1)_{d_j} - U(1)_{u_i}$, physical predictions can only depend on the moduli and quartets,[13]

$$\left| V_{\mathrm{CKM},ij} \right|^2, \quad V_{\mathrm{CKM},ij} V_{\mathrm{CKM},kl} V^*_{\mathrm{CKM},il} V^*_{\mathrm{CKM},kj}. \tag{3.3}$$

For the SMEFT at leading order, that implies that U(1)-charged dimension-six coefficients must multiply U(1)-charged dimension-four coefficients. In the quark sector, the CKM matrix is the only such object, and there does not exist any in the lepton sector, since $U(1)_{L_i}$ is a symmetry of the SM_4 Lagrangian for each i (remember that we work in the limit of vanishing neutrino masses). Therefore, all "off-diagonal" lepton coefficients in the first double column of table 2, i.e. those which are charged under $U(1)_{L_i} - U(1)_{L_j}$, correspond to secondary parameters. This requirement reduces the number of parameters to the ones in the second double column of table 2.

A similar reasoning explains why adequate entries of the CKM matrix must multiply dimension-six coefficients charged under $U(1)_{u_i}$ and/or $U(1)_{d_i}$. Those coefficients must therefore contribute to observables with an additional suppression due to the smallness of

---

[13]Notice that sextets and monomial with more entries of the CKM matrix can be expressed in terms of moduli and quartets using the unitarity of V_{CKM}.

– 12 –

the off-diagonal CKM entries. For instance, $C_{uB,13}$ (expressed in the *up basis*) can only enter observables as

$$V^*_{\text{CKM},11}V_{\text{CKM},31}C_{uB,13} \text{ or } V^*_{\text{CKM},12}V_{\text{CKM},32}C_{uB,13}, \tag{3.4}$$

where the unitarity of the CKM matrix allows us to not consider $V^*_{\text{CKM},13}V_{\text{CKM},33}C_{uB,13}$. Due to the gauge anomalies of the abelian symmetries $\text{U}(1)_{u_i,d_i,L_i}$, the θ-angles of the different gauge factors of the SM are also charged parameters, which can interfere (non-perturbatively) with appropriate SMEFT coefficients (see appendix F for a discussion of θ_{QCD}). Nevertheless, $\text{U}(1)_{L_i} - \text{U}(1)_{L_j}$ is anomaly-free in the SM, therefore the statements made previously about SMEFT coefficients in the lepton sector also hold non-perturbatively.

In this discussion, we ignored the specific cases of observables in the neutrino sector since we work in the limit of vanishing neutrino masses. Examples of observables which we allow include electric dipole moments (EDMs) [42, 49, 50, 56] or the CPV parameters ϵ_K and ϵ' in kaon physics [57–59]. We also assume that the leading BSM contribution indeed corresponds to the interference at $\mathcal{O}(1/\Lambda^2)$, and not to a dimension-six contribution squared (or the interference between the SM_4 and a dimension-eight coefficient, etc) due to some accidental suppression of the $\mathcal{O}(1/\Lambda^2)$ term. Below, we will also study the SM_4 parameter space as a whole (i.e., beyond values relevant for phenomenology), which includes points where several entries of the CKM matrix become unphysical and can be redefined away, turning additional dimension-six coefficients in the quark sector into secondary ones. This would for instance happen if all down-type quarks were massless: barring observables which are ill-defined when $m_d \to 0$ within our leading order observables, we find a further reduction of the relevant dimension-six coefficients (see table 6 in appendix A.2). Ref. [60] performs a similar counting of primary parameters (focusing on the kinematic situation where all fermions masses, but the top and bottom quark, are neglected).

3.2 CP conservation at leading order and minimal sets of CP-odd invariants

We now turn to the characterization of CP violation in SMEFT at leading order using flavor-invariants. Specifically, in the spirit of the discussion which leads to the introduction of J_4, we ask

> which flavor invariants vanish iff. CP is conserved at leading order in SMEFT? We call such a set of CP-odd invariants of minimal cardinality a *minimal set*.

The notion of minimal cardinality implies that there are no redundancy: the vanishing of each invariant in a minimal set provides an independent condition. The number of invariants in a minimal set must be larger or equal than the number of new primary sources of CPV.[14] In the case of the SMEFT at $\mathcal{O}(1/\Lambda^2)$, we find the non-trivial result that the two numbers agree for all operators (see below).

---

[14]It may need to be larger: as we detail below, we count an invariant as independent if there exists at least one point in parameter space (in terms of fermion masses or CKM entries) where it cannot be expressed in terms of other invariants of the set.

Before going further, let us discuss one subtlety associated to our definition of a minimal set of CP-odd flavor-invariants, which has to do with the parameter space considered. In the way our definition is stated, it suggests that one aims at characterizing CP-conserving points for all possible choices of parameters, i.e. quark masses as well as mixing angles and the CKM phase δ_{CKM}. However, one could also try to characterize CP-conserving points *within a given parameter subspace*, for instance for values of the quark masses which are non-vanishing. This is the choice we make in the main body of this paper: we build flavor invariants which vanish iff. CP is conserved at leading order in SMEFT *under the assumption that quark masses are non-vanishing*. Our methods also allow one to identify minimal sets of flavor-invariants when one considers vanishing quark masses, but since the expression of the required invariants is more intricate than for the simpler case of non-vanishing quark masses, we leave the resolution of this question to appendix A. One could finally restrict to characterizing CP-conservation for a smaller set of parameters, e.g. fixing the values of the quark masses or taking them non-degenerate. Our sets of invariants work in such restricted cases, but there usually exist simpler ones (which do not correspond to minimal sets on larger sets of parameters). We will encounter explicit examples in the next section.

Due to the SMEFT power counting, the conservation of CP at zeroth order first demands that $J_4 = 0$, so that the J_4 is part of any minimal set. Then, in order to build the rest of the minimal set, we look for invariants which are *linear with respect to the SMEFT dimension-six coefficients*, consistently with the goal of characterizing CPV in observables up to the first non-leading order, i.e. up to the $\mathcal{O}(1/\Lambda^2)$ interference term in the r.h.s. of eq. (3.2). This linearity is also valuable to check that we indeed have a necessary and sufficient condition for CP conservation, as it does not suffer from subtleties associated to non-linear invariants, found e.g. in neutrino physics [28]. Indeed, the question of whether vanishing invariants really implies vanishing CP phases is trivially answered here. We found that there does exist a minimal set of invariants linear with respect to the SMEFT dimension-six coefficients. This set is presented in subsequent sections, and represents our main result.

The linearity implies that the minimal sets are decomposed in minimal sets of invariants $\mathcal{I}_a(C^{(6)})$, defined for each SMEFT dimension-six operator $\mathcal{O}^{(6)}$ and its associated matrix-valued coefficient $C^{(6)}$ independently, and where the index a labels the new primary sources of CPV present in $C^{(6)}$. Therefore, we can define the notion of minimal set at the level of each SMEFT operator independently. The invariants have the following form:

$$\mathcal{I}_a(C^{(6)}) = \text{Im}\left(\text{flavor-invariant linear in } C^{(6)}\right) = T_{ai}^R \left(\text{Re}\, C^{(6)}\right)_i + T_{ai}^I \left(\text{Im}\, C^{(6)}\right)_i \equiv \mathcal{T}_{ai} \overrightarrow{C}_i^{(6)},$$
(3.5)

where the last equality describes the result of evaluating the invariant in a given flavor basis. Here we define

$$\overrightarrow{C}_i^{(6)} \equiv \left(\left(\text{Re}\, C^{(6)}\right)_1, \left(\text{Re}\, C^{(6)}\right)_2, \ldots, \left(\text{Im}\, C^{(6)}\right)_1, \ldots\right)$$
(3.6)

as the vector in flavor space composed by the vectors $\text{Re}\,/\,\text{Im}\, C^{(6)}$, not necessarily of same

length, while the transfer matrix $\mathcal{T}_{ai}$ is defined to take the block form

$$\mathcal{T} = \begin{pmatrix} T^R & T^I \end{pmatrix}. \tag{3.7}$$

By linearity, T^R_{ai} is the imaginary part of a linear combination of products of entries of dimension-four Yukawas (which are the only flavored objects at dimension-four), while T^I_{ai} is the real part of a similar, albeit generically different, combination. Those matrices T^R and T^I depend explicitly on the operator $\mathcal{O}^{(6)}$ considered. A very convenient feature of such invariants is that they automatically project out any secondary coefficient, which cannot be arranged into invariants in a linear fashion by definition.

Showing that the set of invariants is minimal can be phrased as a condition on the matrices $T^{R/I}$:

> a set of flavor invariants is a minimal set iff. the rank of the transfer matrix $\mathcal{T}$ equals the number of new primary sources of CPV in $C^{(6)}$ when $J_4 = 0$ and never does for all sets with strictly smaller cardinality.

Note that the last part of this characterization is automatic when the number of invariants, the number of primary sources of CPV and the rank are all equal. We use this condition below to check that the sets of invariants we present below are indeed minimal. We stress that the meaning of "when $J_4 = 0$" encompasses a large subset of the whole parameter space spanned by the masses and the mixing angles, as seen from the expression in eq. (2.11): it is achieved when $\theta_{ij} = 0$ or $\pi/2$, or when $m_{u,i} = m_{u,j}$ or $m_{d,i} = m_{d,j}$, for any pair i, j. In addition, setting $J_4 = 0$ via one of these choices still leaves a large freedom for the remaining parameters. For instance, one may have $\theta_{ij} = 0$ *and* $m_{u,k} = m_{u,l}$ for some $i, j, k, l = 1, \ldots, 3$. A set of flavor invariants is a minimal set only if the rank of its transfer matrix corresponds to the number of new sources of CPV *within the whole parameter space where $J_4 = 0$* (up to the restriction of non-vanishing quark masses which we adopt in the main text of this paper and relax in appendix A). We will come back to this point in section 4.

The transfer matrix $\mathcal{T}$ acts on the flavor-space vector made out of real and imaginary entries of $C^{(6)}$ (the precise order in the labeling as well as the order between real and imaginary part is unimportant). Note that the rank does not change under the action of flavor transformations (which reshuffle real and imaginary parts, as well as the entries of $T^{R/I}$).

4 Minimal set of CP-odd invariants

In this section, we present the minimal set of leading order CP-odd invariants in SMEFT at dimension-six, under the aforementioned assumption that all fermion masses are non-vanishing, which has an impact on how many sources of CPV are expected and which invariants correctly capture them. We treat the cases of vanishing masses in appendix A.

4.1 Examples

Let us present some parts of our minimal set of invariants for SMEFT at dimension-six. As we explained previously, the linearity with respect to the Wilson coefficients of

the dimension-six CP-odd observables allows one to treat the different SMEFT operators independently. The study of all SMEFT operators proceeds along identical lines, and the full set of invariants is presented in appendices D and E.

We begin by considering SMEFT operators which are bilinear in fermion fields and hermitian, and therefore have the simplest non-trivial flavor structure. Invariants under unitary groups with bi-fundamental representations must feature the invariant tensor δ^a_b, therefore they correspond to linear combinations of traces of products of matrices, arranged so that indices of a given fundamental representation and its conjugate are contracted in the trace, as seen for instance in eqs. (1.5)–(1.6). In addition, there are relations between powers of 3×3 matrices, and/or between their traces, derived from the Cayley-Hamilton theorem, which reduce the candidate invariants to a finite set. We explicitly present such properties in appendix B. For convenience we define: $X_u \equiv Y_u Y_u^\dagger$ and $X_d \equiv Y_d Y_d^\dagger$. The relevant single-trace invariants linear with respect to a SMEFT coefficient C, for a fermion bilinear operator, take the universal form[15]

$$L_{abcd}(\tilde{C}) \equiv \operatorname{Im} \operatorname{Tr}\left(X_u^a X_d^b X_u^c X_d^d \tilde{C} \right), \text{ with } a, b, c, d = 0, 1, 2 \text{ and } a \neq c, b \neq d, \qquad (4.1)$$

where $\tilde{C} = C, CY_{f=u,d,e}^\dagger$ or $Y_f C Y_f^\dagger$, depending on the chiral structure of the operator under study (see below for explicit formulae). We first choose $C = C_{Hu}$ for definiteness, and we find that the following property holds:

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}_4} + \frac{C_{Hu,ij}}{\Lambda^2} \left(iH^\dagger \overleftrightarrow{D}_\mu H \right) \overline{u_{i,R}} \gamma^\mu u_{j,R}$$

preserves CP at $\mathcal{O}(1/\Lambda^2)$ iff.

$$J_4 = L_{1100}\left(Y_u C_{Hu} Y_u^\dagger \right) = L_{2200}\left(Y_u C_{Hu} Y_u^\dagger \right) = L_{1122}\left(Y_u C_{Hu} Y_u^\dagger \right) = 0 \,. \qquad (4.2)$$

Indeed, $J_4 = 0$ is necessary so that the leading order SM_4 contribution to any CP-odd observable vanishes. Once enforced, this makes the SM_4 Lagrangian CP-symmetric, and there remains generically three new primary sources of CPV in C_{Hu}. Indeed, C_{Hu} is a (3×3) hermitian matrix (which transforms as a $\mathbf{3} \otimes \bar{\mathbf{3}}$ representation of $\mathrm{U}(3)_u$). Therefore, the minimal set for C_{Hu} should at least contain three invariants. As we explained in section 3.2, in order to show that the three invariants in eq. (4.2) capture the three necessary conditions, it is sufficient to compute the transfer matrix $\mathcal{T}$ such that,

$$\begin{pmatrix} L_{1100} \\ L_{2200} \\ L_{1122} \end{pmatrix} = \begin{pmatrix} T^R & T^I \end{pmatrix} \begin{pmatrix} \operatorname{Re} C_{Hu,11} \\ \operatorname{Re} C_{Hu,12} \\ \dots \\ \operatorname{Im} C_{Hu,12} \\ \operatorname{Im} C_{Hu,13} \\ \operatorname{Im} C_{Hu,23} \end{pmatrix}, \qquad (4.3)$$

---

[15]As we will see, those structures are also the only ones needed for 4-Fermi operators.

and show that it has rank 3. Parametrically, and in some basis, the generic case corresponds to taking $\delta_{\text{CKM}} \to 0$ in eq. (2.9) while holding all mixing angles different from $0, \pi/2$ and all quark masses non-denegerate, which is what we assume in the current section (we treat more general cases below). Then, $T^R = 0_{3 \times 6}$ and the determinant of T^I is found to be non-vanishing. Therefore, $L_{1100} = L_{2200} = L_{1122} = 0$ implies that $\text{Im}\, C_{Hu,ij} = 0$ in the basis of eq. (2.9) (or any other basis where the Yukawa matrices are real), i.e. CP is conserved. Conversely, the conservation of CP, or equivalently $\text{Im}\, C_{Hu,ij} = 0$, implies that all L's vanish since $T^R = 0$. This proves the equivalence announced above.

One may be surprised by the fact that some simple invariants, in the sense that they feature low powers of the Yukawa matrices, are not parts of the set in eq. (4.2). For instance, for C_{Hu}, the set formed by

$$L_{1100}\left(Y_u C_{Hu} Y_u^\dagger\right) , \; L_{1200}\left(Y_u C_{Hu} Y_u^\dagger\right) , \; L_{2100}\left(Y_u C_{Hu} Y_u^\dagger\right) \tag{4.4}$$

would pass the test performed in this section, namely the associated transfer matrix would generically have rank 3. Such invariants have been studied beyond SMEFT, and exist generally for any set of three hermitian matrices in the same adjoint representation, as explained in ref. [19]. However, they would not be a valid choice of invariants, since they would not correspond to sufficient conditions whatever the values of the fermion masses and the CKM matrix. For instance, $m_t = m_c$ is another way to get $J_4 = 0$ other than that discussed above. In this situation, there remains three conditions necessary for CP to be conserved (see section 4.2 for details). Therefore, imposing that a given minimal set vanishes should be equivalent to three independent conditions also when $m_t = m_c$. However (when $m_t = m_c$) we find that

$$(m_u^2 + m_c^2)L_{1100} = L_{2100}\,, \tag{4.5}$$

where the mass-dependent factor can be expressed in terms of invariant quantities,

$$m_u^2 + m_c^2 = \frac{18 \det X_u - 3\left(\text{Tr}\, X_u\right)^3 + 7\,\text{Tr}\, X_u \,\text{Tr}\, X_u^2}{6\,\text{Tr}\, X_u^2 - 2\left(\text{Tr}\, X_u\right)^2}\,. \tag{4.6}$$

Therefore, imposing that the set in eq. (4.4) vanishes only amounts to two conditions. Instead, one can check that the vanishing of the set in eq. (4.2) yields three independent conditions,[16] even when $m_t = m_c$. We expand on what happens however $J_4 = 0$ is taken in section 4.2, and show that the set in eq. (4.2) above yields a satisfactory minimal set in all cases (as long as all quark masses are non-vanishing).

Similar reasoning applies to all SMEFT operators. Let us present the results in two more cases with slightly more complicated flavor structures, the non-hermitian bilinear operator $\mathcal{O}_{uH}$ and the hermitian symmetric 4-Fermi operator $\mathcal{O}_{uu}$. The Wilson coefficient C_{uH} contains nine new primary sources of CPV, since it is an arbitrary (3×3) complex matrix (which transforms as a $(\mathbf{3}, \bar{\mathbf{3}})$ representation of $\text{U}(3)_Q \times \text{U}(3)_u$). C_{uu} contains

---

[16]The interplay between degenerate cases and flavor-invariants has also been studied in ref. [29] for massive Majorana neutrinos.

eighteen new CPV parameters ($C_{uu,ijkl}$ is "hermitian", i.e. $C^*_{uu,ijkl} = C_{uu,jilk}$, symmetric, $C_{uu,ijkl} = C_{uu,klij}$, and it transforms in the $(\mathbf{3} \otimes \bar{\mathbf{3}})^2$ of $\mathrm{U}(3)_u$).

For $\mathcal{O}_{uH}$, we find that

$$
\begin{aligned}
\mathcal{L} = \mathcal{L}_{\mathrm{SM}_4} &+ \frac{C_{uH,ij}}{\Lambda^2} \overline{Q_{i,L}} \tilde{H} u_{j,R} |H|^2 + \text{h.c.} \\
&\text{preserves CP at } \mathcal{O}(1/\Lambda^2) \text{ iff.} \\
J_4 = L_{0000} \left(C_{uH} Y_u^\dagger \right) &= L_{1000} \left(C_{uH} Y_u^\dagger \right) = L_{0100} \left(C_{uH} Y_u^\dagger \right) \\
= L_{1100} \left(C_{uH} Y_u^\dagger \right) &= L_{0110} \left(C_{uH} Y_u^\dagger \right) = L_{2200} \left(C_{uH} Y_u^\dagger \right) \\
= L_{0220} \left(C_{uH} Y_u^\dagger \right) &= L_{1220} \left(C_{uH} Y_u^\dagger \right) = L_{0122} \left(C_{uH} Y_u^\dagger \right) = 0 \,.
\end{aligned}
\tag{4.7}
$$

We now turn to $\mathcal{O}_{uu}$. For 4-Fermi operators, we generically define

$$
\mathrm{Tr}_A \left(M^{(1)}, M^{(2)}, C \right) \equiv M^{(1)}_{ji} M^{(2)}_{lk} C_{ijkl} \,, \quad \mathrm{Tr}_B \left(M^{(1)}, M^{(2)}, C \right) \equiv M^{(1)}_{li} M^{(2)}_{jk} C_{ijkl} \,, \tag{4.8}
$$

and

$$
\begin{aligned}
A^{abcd}_{efgh}(C) &= \mathrm{Im}\,\mathrm{Tr}_A \left(X_u^a X_d^b X_u^c X_d^d, X_u^e X_d^f X_u^g X_d^h, C \right) \,, \\
B^{abcd}_{efgh}(C) &= \mathrm{Im}\,\mathrm{Tr}_B \left(X_u^a X_d^b X_u^c X_d^d, X_u^e X_d^f X_u^g X_d^h, C \right) \,.
\end{aligned}
\tag{4.9}
$$

We further define

$$
C_{\tilde{u}\tilde{u}uu,ijkl} \equiv \sum_{m,n} Y_{u,im} Y^\dagger_{u,nj} C_{uu,mnkl} \,, \tag{4.10}
$$

and similarly for $C_{\tilde{u}uu\tilde{u}}, C_{u\tilde{u}\tilde{u}u}, C_{uu\tilde{u}\tilde{u}}, C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}$, and for the down quark versions. We then find that

$$
\begin{aligned}
\mathcal{L} = \mathcal{L}_{\mathrm{SM}_4} &+ \frac{C_{uu,ijkl}}{\Lambda^2} \overline{u_{i,R}} \gamma^\mu u_{j,R} \, \overline{u_{k,R}} \gamma_\mu u_{l,R} \\
&\text{preserves CP at } \mathcal{O}(1/\Lambda^2) \text{ iff.} \\
J_4 = A^{0000}_{1100} (C_{uu\tilde{u}\tilde{u}}) &= A^{1000}_{1100} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) = A^{0100}_{1100} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) \\
= A^{0000}_{2200} (C_{uu\tilde{u}\tilde{u}}) &= A^{1100}_{1100} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) = A^{0200}_{1100} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) \\
= A^{0100}_{2200} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) &= A^{0000}_{1122} (C_{uu\tilde{u}\tilde{u}}) = A^{1100}_{2200} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) \\
= A^{1000}_{1122} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) &= A^{0100}_{1122} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) = A^{1100}_{0122} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) \\
= A^{1200}_{2200} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) &= B^{0000}_{1100} (C_{u\tilde{u}\tilde{u}u}) = B^{0100}_{1100} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) \\
= B^{0200}_{2100} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) &= A^{1200}_{1122} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) = B^{1000}_{1200} (C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}}) = 0 \,.
\end{aligned}
\tag{4.11}
$$

The proofs of these equivalences follow from the same logic as for C_{Hu}: compute T^I and check that it has maximal rank, which means that the origin in invariant space is uniquely mapped to the origin in imaginary-coefficient space. Therefore, the vanishing of the invariants is equivalent to the conservation of CP (at leading order).

Finally, let us consider the leptonic case. As we argued in section 3.1, the fact that the SM_4 Lagrangian is symmetric under the lepton numbers $\mathrm{U}(1)_{L_i}$ makes several SMEFT

dimension-six coefficients in the lepton sector secondary. For the specific example of C_{eH}, this means that all off-diagonal entries are secondary, since they are charged under $U(1)_{L_i} - U(1)_{L_j}$ for suitable i, j. Therefore, although all imaginary parts of the nine entries of C_{eH} violate CP when the full SMEFT expansion is considered, the only ones which can contribute to CP-odd observables at $\mathcal{O}(1/\Lambda^2)$ are the diagonal ones. The same reasoning applies to C_{He}, which therefore does not violate CP at $\mathcal{O}(1/\Lambda^2)$. Consequently, a minimal set for C_{eH} contains three invariants, and is empty for C_{He}. Indeed, we find that (defining $X_e \equiv Y_e Y_e^\dagger$)

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}_4} + \frac{C_{He,ij}}{\Lambda^2} \left(iH^\dagger \overleftrightarrow{D}_\mu H \right) \overline{e_{i,R}} \gamma^\mu e_{j,R} + \left(\frac{C_{eH,ij}}{\Lambda^2} \overline{L_{i,L}} H e_{j,R} |H|^2 + \mathrm{h.c.} \right)$$

$$\text{preserves CP at } \mathcal{O}(1/\Lambda^2) \text{ iff.}$$

$$J_4 = \mathrm{Im}\,\mathrm{Tr}\left(C_{eH} Y_e^\dagger \right) = \mathrm{Im}\,\mathrm{Tr}\left(X_e C_{eH} Y_e^\dagger \right) = \mathrm{Im}\,\mathrm{Tr}\left(X_e^2 C_{eH} Y_e^\dagger \right) = 0 \,. \qquad (4.12)$$

Let us end this preview by saying that the above invariant sets are not unique: there are different sets of invariants which would equally well capture the necessary and sufficient conditions for CP-conservations at order $1/\Lambda^2$. Our construction of the above sets requires that the invariants have the lowest possible degree with respect to Yukawa matrices, and that they be as large as possible when the observed values of the fermion masses and CKM entries are plugged in.

4.2 Expected ranks when $J_4 = 0$

We now discuss the rank of the transfer matrices, related to the validity of the minimal sets presented above, in non-generic cases, namely when some fermion masses are degenerate and/or when there are texture zeros in the CKM matrix (which happens when some mixing angles take special values).

To apply consistently the definition of minimal sets from section 3.2, we need to carefully determine how many new primary sources of CPV there are when $J_4 = 0$, or equivalently what is the expected rank of the transfer matrix $\mathcal{T}$, irrespective of how we take $J_4 \to 0$. In our generic situation of the previous section, $J_4 \to 0$ meant $\delta_{\mathrm{CKM}} \to 0$ in eq. (2.9) while holding all mixing angles different from $0, \pi/2$ and the quark masses non-degenerate. This ensures that there are no texture zeros in the CKM matrix, so that the number of CP-breaking quantities were identified with the number of imaginary parts (in the quark sector). However, that is not the only situation captured by the ambiguous "$J_4 = 0$" condition, as we anticipated in section 3.2. Indeed, mass degeneracies and/or texture zeros in the CKM matrix may lead to a conserved flavor symmetry of the SM_4 Lagrangian larger than $U(1)_B$, which has an impact on the number of CP-odd quantities at order $1/\Lambda^2$.

The reason is identical to that discussed in section 3.1: observables should be invariant under any symmetry of the spectrum of asymptotic states. Consequently, at order $1/\Lambda^2$, SMEFT coefficients must combine with SM_4 parameters to form invariant objects, and in particular, when the SM_4 Lagrangian has a flavor symmetry (which is therefore part of

the symmetry group of the spectrum), only SMEFT coefficients invariant under this flavor symmetry can generate amplitudes which interfere with SM$_4$ ones.

In the generic case of a CP-preserving SM$_4$ Lagrangian (i.e. taking $\delta_{\text{CKM}} \to 0$ in eq. (2.9) with generic values of mixing angles and quark masses, or said differently for a real CKM matrix without texture zeros), there is no flavor symmetry beyond the baryon and lepton numbers $\text{U}(1)_B \times \text{U}(1)_{L_i}$. Therefore, any additional B- and L_i-preserving SMEFT coupling in the Lagrangian is primary, and its imaginary part is a primary source of CPV. We leave to appendix A the systematic discussion of all flavor symmetries of the SM$_4$ and their relation to mass degeneracies and/or texture zeros in the CKM matrix, but, for the sake of illustration, we discuss here two specific cases.

In the first one, the CKM matrix is non-generic and has texture zeros:

$$V_{\text{CKM}} = \begin{pmatrix} * & 0 & * \\ 0 & * & 0 \\ * & 0 & * \end{pmatrix} . \tag{4.13}$$

This texture can be achieved from eq. (2.9) by taking $s_{23} = s_{12} = 0$ (in particular, $J_4 = 0$ for such a texture). The flavor symmetry of the dimension-four action is now enlarged to $\text{U}(1)^2$, corresponding to two independent abelian transformations of $(Q_1, Q_3, u_1, u_3, d_1, d_3)$ and (Q_2, u_2, d_2) respectively (in the *up* or *down basis*), acting as follows:

$$u_{i,R} \to e^{i\xi_{q_i}} u_{i,R} , \tag{4.14}$$

with $\xi_{q_1} = \xi_{q_3}$, and similarly for other quarks.

The second example is that of a degenerate fermion spectrum. We take for definiteness $m_t = m_c$ (which again implies $J_4 = 0$). With this degeneracy, the symmetry of the mass terms becomes non-abelian, while the phase in the CKM matrix, as well as one mixing angle, is no longer physical. Indeed, we can perform a flavor transformation (here in the *up basis*),

$$Q_L \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} Q_L , \quad u_R \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} u_R , \tag{4.15}$$

such that

$$V_{\text{CKM}} \to \begin{pmatrix} c_{13} & 0 & s_{13} \\ 0 & 1 & 0 \\ -s_{13} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{4.16}$$

Therefore, the values of θ_{23} and δ_{CKM} have no physical impact. For generic values of θ_{12} and θ_{13}, the CKM matrix in eq. (4.16) has no texture zeros, and the flavor symmetry of the SM$_4$ Lagrangian in the quark sector still corresponds to the baryon number.

All possible cases are treated similarly, and the full discussion is presented in appendix A. The summary of this analysis (assuming no vanishing mass, see appendix A for the more general case) is presented in table 3 where, for each non-generic case of interest, we present the flavor symmetry group. The discussion in the lepton sector is similar, but

Parameter values		Flavor symmetries of the SM$_4$ Lagrangian
$m_u \neq m_c \neq m_t$ $m_d \neq m_s \neq m_b$	Generic V_{CKM}	U(1)$_B$
	$\|V_{\text{CKM},i_0 j_0}\| = 1$, $V_{\text{CKM},ij_0} = V_{\text{CKM},i_0 j} = 0$ $i \neq i_0, j \neq j_0$	U(1)^2
	$\|V_{\text{CKM},i_1 j_1}\| = \|V_{\text{CKM},i_2 j_2}\| = \|V_{\text{CKM},i_3 j_3}\| = 1$ for $\begin{array}{l} i_1 \neq i_2 \neq i_3 \\ j_1 \neq j_2 \neq j_3 \end{array}$ $V_{\text{CKM},ij} = 0$ elsewhere	U(1)^3
$m_u \neq m_c = m_t$ $m_d \neq m_s \neq m_b$	Generic V_{CKM} (see eq. (4.16))	U(1)$_B$
	$\|V_{\text{CKM},i_0 j_0}\| = 1$, $V_{\text{CKM},ij_0} = V_{\text{CKM},i_0 j} = 0$ $i \neq i_0, j \neq j_0$	U(1)^2
	$\|V_{\text{CKM},i_1 j_1}\| = \|V_{\text{CKM},i_2 j_2}\| = \|V_{\text{CKM},i_3 j_3}\| = 1$ for $\begin{array}{l} i_1 \neq i_2 \neq i_3 \\ j_1 \neq j_2 \neq j_3 \end{array}$ $V_{\text{CKM},ij} = 0$ elsewhere	U(1)^3
$m_u \neq m_c \neq m_t$ $m_d = m_s \neq m_b$	Same as the previous case with $V_{\text{CKM}} \leftrightarrow V_{\text{CKM}}^\dagger$	
$m_u \neq m_c = m_t$ $m_d = m_s \neq m_b$	Generic V_{CKM}	U(1)^2
	$\|V_{\text{CKM},11}\| = \|V_{\text{CKM},22}\| = \|V_{\text{CKM},33}\| = 1$ $V_{\text{CKM},ij} = 0$ elsewhere	U(1)^3
	$\|V_{\text{CKM},13}\| = \|V_{\text{CKM},22}\| = \|V_{\text{CKM},31}\| = 1$ $V_{\text{CKM},ij} = 0$ elsewhere	U(2) × U(1)
$m_u = m_c = m_t$	$m_d \neq m_s \neq m_b$	U(1)^3
	$m_d = m_s \neq m_b$	U(2) × U(1)
	$m_d = m_s = m_b$	U(3)
$m_d = m_s = m_b$	$m_u \neq m_c \neq m_t$	U(1)^3
	$m_u \neq m_c = m_t$	U(2) × U(1)
	$m_u = m_c = m_t$	U(3)

Table 3. Flavor symmetry of the SM$_4$ Lagrangian as a function of special values for quark masses (assumed to be non-vanishing, see appendix A for the general case) and entries of the CKM matrix. Conditions on the right are understood to be imposed on top of those on their left. Here only some of the possible combinations of mass degeneracies are treated. The other mass degeneracies lead to the same flavor symmetries provided the corresponding non-generic V_{CKM} are multiplied by appropriate matrices exchanging flavor labels (see footnote 18).

features at least a U(1)^3 flavor freedom in the SM$_4$, since the PMNS matrix is given by the identity. When two or three charged lepton masses are degenerate, this U(1)^3 increases to U(2) × U(1) and U(3), respectively.

This discussion allows us to work out the number of new primary sources of CPV at order $\mathcal{O}(1/\Lambda^2)$, in each (non-)generic case for the CKM matrix. Let us focus again on the two previous examples, which generalize easily to all.

When the baryon or lepton numbers are the only flavor symmetries at dimension-four (in the *up* or *down basis*), all imaginary parts of the Wilson coefficients at dimension-six

Flavour symmetries of the quark sector of the SM	Bilinears				4-Fermi						
	C_{eH} C_{eW} C_{eB}	C_{uH} C_{uG} C_{uW} C_{uB} C_{dH} C_{dG} C_{dW} C_{dB} C_{Hud}	$C_{HL}^{1,3}$ C_{He}	$C_{HQ}^{1,3}$ C_{Hu} C_{Hd}	C_{LL} C_{ee}	C_{Le}	$C_{QQ}^{1,3}$ C_{uu} C_{dd}	$C_{LQ}^{1,3}$ C_{Qe} C_{Lu} C_{eu} C_{Ld} C_{ed}	$C_{ud}^{1,8}$ $C_{Qu}^{1,8}$ $C_{Qd}^{1,8}$	C_{LedQ} $C_{LeQu}^{1,3}$	$C_{QuQd}^{1,8}$
$U(1)_B$	3	9	0	3	0	3	18	9	36	27	81
$U(1)^2$	3	5	0	1	0	3	5	3	12	15	33
$U(1)^3$	3	3	0	0	0	3	0	0	3	9	15
$U(2) \times U(1)$	3	2	0	0	0	3	0	0	1	6	7
$U(3)$	3	1	0	0	0	3	0	0	0	3	2
Two degenerate electron-type leptons	$\times \frac{2}{3}$	$\times 1$		$\times 1$		$\times \frac{2}{3}$	$\times 1$	$\times \frac{2}{3}$	$\times 1$	$\times \frac{2}{3}$	$\times 1$
All electron-type leptons degenerate	$\times \frac{1}{3}$	$\times 1$		$\times 1$		$\times \frac{1}{3}$	$\times 1$	$\times \frac{1}{3}$	$\times 1$	$\times \frac{1}{3}$	$\times 1$

Table 4. Numbers of new primary sources of CPV contained in each dimension-six SMEFT coefficient, for each of the possible flavor groups of the quark sector of the SM_4 at dimension-4 (restricting to situations where fermion masses are non-vanishing). The last two rows indicate which multiplicative coefficient should be applied to all numbers of the same column for special values of the electron-type lepton masses.

in the quark sector (and in the lepton sector, all imaginary parts which are not charged under the lepton numbers) can interfere with the SM_4. Instead, when the flavor symmetry increases to $U(1)^2$, several SMEFT coefficients become secondary. For instance, C_{Hu} transforms as (in the *up* or *down bases*)

$$C_{Hu,ij} \to e^{i(\xi_{q_j} - \xi_{q_i})} C_{Hu,ij}, \tag{4.17}$$

where $\xi_{q_1} = \xi_{q_3}$ and arbitrary ξ_{q_2} for the texture of eq. (4.13), and similarly for other textures. Therefore, for the texture of eq. (4.13), only $C_{Hu,13}$ is primary and C_{Hu} only provides one primary source of CPV, Im $C_{Hu,13}$.[17] Thus, in this case, a single invariant in the minimal set for C_{Hu} is sufficient.

This exercise can be performed for all non-generic cases for the CKM matrix and for all Wilson coefficients. This results in a set of conditions for CP conservation at leading order, whose number is in one-to-one correspondence to the number of independent CP-odd invariants in a minimal set. As we just saw, those numbers depend on the flavor symmetry of the SM_4 Lagrangian, and are given for all SMEFT operators in table 4.

We can now come back to the statement that the set of invariants in eq. (4.4) is not a satisfying one for C_{Hu}. As seen in table 4, all its off-diagonal entries are primary when $m_t = m_c$, and all their imaginary parts violate CP (in an appropriate basis), hence we need three independent invariants to capture the conditions for CP conservation.

---

[17]One can construct non-linear invariant quantities from $C_{Hu,12/23}$, an example being $C_{Hu,12}C_{Hu,23}^*$. At leading order, however, $C_{Hu,12/23}$ cannot contribute linearly to observables.

Let us stress again at this point that the fact that we found a set of invariants of minimal size (i.e. three invariants for the case of C_{Hu}) which captures the necessary and sufficient conditions for CP conservation in all non-generic cases listed in table 3 is a non-trivial result. Nevertheless, it turns out that it can be done for all SMEFT coefficients at dimension-six, as we explicitly showed.

5 Conclusions and future directions

In this paper, we have investigated the collective properties of SMEFT at dimension-six. Their first implication which we have discussed is that only a subset of Lagrangian parameters (dubbed primary parameters) can contribute linearly to observables upon interfering with the dimension-four SM_4 amplitude. This is due to the existence of flavor transformations which leave the SM spectrum and thus any observable invariant, thereby demanding that covariant Lagrangian parameters combine to form invariant objects. This applies for instance to field-rephasings associated to mass-eigenstates, implying that Lagrangian coefficients must combine into rephasing-invariant objects. Associated to the SMEFT power counting, this implies that several coefficients related to dimension-six SMEFT operators cannot contribute to observables at order $1/\Lambda^2$, since they are charged under individual lepton numbers unlike all SM_4 parameters (in the limit of zero neutrino mass). We therefore refined the usual counting of dimension-six SMEFT parameters so as to include this effect, which resulted in the counting of primary parameters in table 2.

Then, we focused on collective effects related to new sources of CP violation in the dimension-six SMEFT, both of which are captured thanks to CP-odd flavor-invariants. To respect the SMEFT power counting, we restricted to invariants linear in dimension-six coefficients, and we presented minimal sets of invariants which map in a one-to-one basis to all new primary sources of CPV. We proved this by showing that the points in parameter space where CP is conserved (at leading BSM order) are exactly the points where our new invariants vanish. This holds for all parts of the SM_4 parameter space, including degenerate cases. A complete list of CP-odd linear invariants can be found in appendices D and E. We remind the reader that this list is not unique.

There are several directions along which our work can be extended [61]. The first one concerns parameterizations of leading order CP-odd observables. Indeed, although our characterization of new primary sources of CPV is complete, it does not necessarily describe all the relevant parameter space of those observables. The reason is that, CP being broken at dimension-four already, the source of its breaking in the SM_4 can interfere with CP-even dimension-six parameters of SMEFT. On the other hand, remembering that physics should not depend on a specific field basis, a CPV observable can always be expressed in terms of CPV flavor-invariants. Taking all this into account suggests that one should study larger sets of CP-odd invariants than our minimal ones. In such sets, only a subset of invariants capture new primary sources of CPV, but each invariant captures a new, independent primary dimension-six quantity which corresponds to, or can interfere with, a source of CPV. We call such a set a *maximal set*. Given an invariant $\mathcal{I}$ in such a set, one

can always write

$$\mathcal{I} = \sum_i c_i \mathcal{I}_i + J_4 \tilde{\mathcal{I}} \,, \tag{5.1}$$

where $\{\mathcal{I}_i\}_i$ forms a minimal set in the sense described in this paper, and $\tilde{\mathcal{I}}$ is a CP-even expression built as a ratio of polynomial flavor-invariants. Indeed, given our definition of minimal sets, all CPV observables (in particular $\mathcal{I}$) must vanish at order $\mathcal{O}(1/\Lambda^2)$ when $J_4 = \mathcal{I}_i = 0$. Nevertheless, one can check that all (CP-odd or -even) primary parameters can be captured by CP-odd invariants, thanks to a rank analysis similar to that presented in section 3.2, and consistent with the fact that all primary parameters can contribute to CPV observables, either directly or via interference with the SM$_4$ CP phase.

Relatedly, one can numerically evaluate the invariants, which encode collective effects and the suppression they induce, using the observed values of the SM$_4$ parameters. This illustrates how accidentally small the absolute strength of each new source of CPV is, or if there are hierarchies among them, etc. This in turn explains (part of) the suppression in CPV observables [62–65]. One can also similarly probe the effects of specific UV hypotheses on the SMEFT coefficients, such as textures derived from flavor symmetries. A given UV model may have its own CP-odd flavor invariants, and these can be matched onto our IR invariants, inducing possible correlations amongst the IR invariants. In other directions, it would be interesting to consider other operator bases than the Warsaw basis we use in this work, and to check which expressions our invariants map to, given that the overall number of independent sources of CPV must be conserved. One could also extend our construction beyond $\mathcal{O}(1/\Lambda^2)$, e.g. to capture squared dimension-six or interfering dimension-eight SMEFT contributions. Finally, one could consider the RG evolution of invariants [31]. We anticipate that flavor invariants, at dimension-six and beyond, are essential tools for illuminating the rich CP structure of SMEFT.

Acknowledgments

A Flavor symmetries of the SM$_4$

In this appendix, we present the details behind table 3, and identify the possible flavor symmetries of the SM$_4$ Lagrangian in terms of textures in the CKM matrix. We remind that the flavor symmetries acts on the quark fields as follows:

$$u_R \to U_u u_R \,, \tag{A.1}$$

where $U_u \in \mathrm{U}(3)_u$, and similarly for all other fermionic fields d, Q, L, e. Which matrices U lead to genuine symmetries of the SM_4 Lagrangian depends on the values of the masses and of the entries of the CKM matrix. In all cases of non-trivial flavor symmetries, we find that $J_4 = 0$, so that there exists at least one combination of CP and flavor symmetries which yield a symmetry of the SM_4 Lagrangian in any basis.

In what follows, we work for definiteness in the *up basis* of eq. (2.7).

A.1 Non-vanishing quark masses

The condition for the flavor invariance of Y_u and Y_d reads

$$Y_u = U_Q^\dagger Y_u U_u \,, \quad Y_d = U_Q^\dagger Y_d U_d \,. \tag{A.2}$$

For non-vanishing quark masses, the Yukawa matrices are full rank and one can use eq. (A.2) to solve for $U_{u,d}$ as a function of U_Q and the Yukawa matrices:

$$U_u = Y_u^{-1} U_Q Y_u \,, \quad U_d = Y_d^{-1} U_Q Y_d \,. \tag{A.3}$$

Therefore, only one matrix determines the two others, and the flavor symmetry group is at most $\mathrm{U}(3)$. Imposing that $U_u^\dagger U_u = U_d^\dagger U_d = \mathbb{1}$ implies (in the *up basis*) that

$$\left[U_Q, m_u m_u^\dagger \right] = \left[V_{\mathrm{CKM}}^\dagger U_Q V_{\mathrm{CKM}}, m_d m_d^\dagger \right] = 0 \,, \tag{A.4}$$

where $m_{u/d} \equiv \mathrm{diag}(m_{u/d_i})$. In the *up basis*, quark masses are positive and real, therefore

$$\left[U_Q, m_u \right] = \left[V_{\mathrm{CKM}}^\dagger U_Q V_{\mathrm{CKM}}, m_d \right] = 0 \tag{A.5}$$

and (using the explicit expression of the Yukawa matrices in the *up basis*)

$$U_u = U_Q \,, \quad U_d = V_{\mathrm{CKM}}^\dagger U_Q V_{\mathrm{CKM}} \,. \tag{A.6}$$

The commutation relations in eq. (A.5) are additional constraints to fulfil which depend on the spectrum, as we now explore.

A.1.1 Non-degenerate quark masses

If all up-type quarks are non-degenerate, the first condition in eq. (A.5) implies that $U_Q = \mathrm{diag}\left(e^{i\xi_i} \right)$ and the second that $U_Q = V_{\mathrm{CKM}} \mathrm{diag}\left(e^{i\tilde{\xi}_i} \right) V_{\mathrm{CKM}}^\dagger$, therefore

$$V_{\mathrm{CKM},ij} = e^{i\left(\xi_i - \tilde{\xi}_j\right)} V_{\mathrm{CKM},ij} \,.$$

Consequently, all ξ's are equal and equal to the $\tilde{\xi}$'s (i.e. the flavor symmetry is given by the baryon number $\mathrm{U}(1)_B$) unless the CKM matrix has some vanishing entries. For instance, one finds a $\mathrm{U}(1)^2$ flavor symmetry when the CKM matrix has the following texture:

$$V_{\mathrm{CKM}} = \begin{pmatrix} * & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix} \,, \tag{A.7}$$

corresponding to the constraints $\xi_2 = \xi_3 = \tilde{\xi}_2 = \tilde{\xi}_3$, for arbitrary $\xi_1 = \tilde{\xi}_1$. More generally, a U(1)^2 flavor symmetry is obtained for any texture such that, given two integers (i_0, j_0), $|V_{\mathrm{CKM}, i_0 j_0}| = 1$ and $V_{\mathrm{CKM}, ij} = 0$ for $i = i_0$ or $j = j_0$. By comparing with the explicit parametrization in eq. (2.9), one finds that a mixing angle has to be equal to 0 or $\pi/2$ in all those cases, hence $J_4 = 0$ and there exists a basis where all SM$_4$ couplings are real.

A flavor symmetry U(1)^3 is obtained for all textures of V_{CKM} such that there is a single number of unit modulus in each row and column, such as e.g. $V_{\mathrm{CKM}} = \mathbb{1}$ or

$$V_{\mathrm{CKM}} = \begin{pmatrix} * & 0 & 0 \\ 0 & 0 & * \\ 0 & * & 0 \end{pmatrix} . \tag{A.8}$$

A.1.2 Degenerate quark masses

In cases with quark mass degeneracies, $J_4 = 0$ automatically and there exists a basis where all SM$_4$ couplings are real.

$\mathbf{m_t = m_c}$. Let us start with the case of two degenerate quarks of the same type, which we take to be $m_t = m_c$ for definiteness, all other masses being non-degenerate.[18]

The first relation in eq. (A.5) implies that

$$U_Q = \begin{pmatrix} e^{i\xi_1} & 0 \\ 0 & U_u^{(2)} \end{pmatrix} \text{ with } U_u^{(2)} \in \mathrm{U}(2) , \tag{A.9}$$

while the second implies that $U_Q = V_{\mathrm{CKM}} e^{i\tilde{\xi}_i} V_{\mathrm{CKM}}^\dagger$, therefore

$$V_{\mathrm{CKM}} = \begin{pmatrix} e^{i\xi_1} & 0 \\ 0 & U_u^{(2)} \end{pmatrix} \cdot V_{\mathrm{CKM}} \cdot \mathrm{diag}\left(e^{-i\tilde{\xi}_i} \right) . \tag{A.10}$$

Upon solving this equation, one finds that, similarly to the case of non-degenerate masses, one can only obtain the flavor groups U(1)$_B$, U(1)^2, and U(1)^3. Flavor symmetries beyond

---

[18]The equivalent case where down quark masses are degenerate is treated identically after the exchange $V_{\mathrm{CKM}} \leftrightarrow V_{\mathrm{CKM}}^\dagger$. We also consider the specific case $m_t = m_c$ (and $m_s = m_d$ in the case of down quarks later on), since the formulae are simpler given our parametrization of V_{CKM}. Nevertheless, the discussion (and the parametrization) can be adapted to any other quark mass degeneracy. In particular, the remarkable textures $V_{\mathrm{CKM}}^{(ji)}$ leading to a given flavor symmetry when $m_{u_j} = m_{u_i}, i < j$ are related to those when $m_t = m_c$ by

$$V_{\mathrm{CKM}}^{(ji)} = R_{i2} R_{j3} V_{\mathrm{CKM}}^{(tc)} ,$$

where R_{ab} is the matrix which exchanges rows (or columns) a and b. Similarly, textures obtained when $m_{d_j} = m_{d_i}, i < j$ are related to those when $m_s = m_d$ by

$$V_{\mathrm{CKM}}^{(ji)} = V_{\mathrm{CKM}}^{(sd)} R_{i1} R_{j2} .$$

For instance, the texture which leads to a flavor symmetry U(2) $\times$ U(1) (see below) for $m_{u_j} = m_{u_i}, i < j$ and $m_{d_l} = m_{d_k}, k < l$ is

$$V_{\mathrm{CKM}} = R_{i2} R_{j3} \cdot \begin{pmatrix} 0 & 0 & * \\ 0 & * & 0 \\ * & 0 & 0 \end{pmatrix} \cdot R_{k1} R_{l2} .$$

baryon numbers are obtained for the textures discussed above, but the generic case for the CKM matrix when $m_t = m_c$ is given in eq. (4.16) and does not have a specific texture.

$m_t = m_c$ and $m_s = m_d$. Let us now turn to the case where $m_t = m_c$ and $m_s = m_d$, all other masses being non-degenerate. The relations in eq. (A.5) imply that

$$U_Q = \begin{pmatrix} e^{i\xi_1} & 0 \\ 0 & U_u^{(2)} \end{pmatrix} = V_{\text{CKM}} \cdot \begin{pmatrix} U_d^{(2)} & 0 \\ 0 & e^{i\tilde{\xi}_3} \end{pmatrix} \cdot V_{\text{CKM}}^\dagger \text{ with } U_u^{(2)}, U_d^{(2)} \in \text{U}(2), \qquad \text{(A.11)}$$

hence

$$V_{\text{CKM}} = \begin{pmatrix} e^{i\xi_1} & 0 \\ 0 & U_u^{(2)} \end{pmatrix} \cdot V_{\text{CKM}} \cdot \begin{pmatrix} U_d^{(2)\dagger} & 0 \\ 0 & e^{-i\tilde{\xi}_3} \end{pmatrix}. \qquad \text{(A.12)}$$

Starting from the CKM matrix in eq. (4.16), which is generic when $m_t = m_c$, we can further rotate θ_{12} away by performing

$$d_R \rightarrow \begin{pmatrix} c_{12} & -s_{12} & 0 \\ s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} d_R. \qquad \text{(A.13)}$$

Recall that we work in the *up basis*, where V_{CKM} appears in Y_d and can be affected by right-handed down-quark flavor transformations which commute with the down-quark mass matrices. We then obtain

$$V_{\text{CKM}} \rightarrow \begin{pmatrix} c_{13} & 0 & s_{13} \\ 0 & 1 & 0 \\ -s_{13} & 0 & c_{13} \end{pmatrix}, \text{ i.e. with a texture } \begin{pmatrix} * & 0 & * \\ 0 & * & 0 \\ * & 0 & * \end{pmatrix}. \qquad \text{(A.14)}$$

Therefore, the generic CKM matrix for the present case has a texture which allows for a flavor symmetry at least as large as $\text{U}(1)^2$. Possible larger flavor symmetries are $\text{U}(1)^3$ or $\text{U}(2) \times \text{U}(1)$, obtained for the following respective textures,[19]

$$\begin{pmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 & * \\ 0 & * & 0 \\ * & 0 & 0 \end{pmatrix}. \qquad \text{(A.15)}$$

---

[19]Let us present some details regarding the second case to illustrate the derivation. With such a texture, one can phase-rotate the fields so as to get

$$V_{\text{CKM}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

and one finds from eq. (A.12) that

$$\mathbb{1} = \begin{pmatrix} O U_u^{(2)} O U_d^{(2)\dagger} & 0 \\ 0 & e^{i(\xi_1 - \tilde{\xi}_3)} \end{pmatrix},$$

with $O \equiv \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix} \right)$. Therefore, one obtains $U_d^{(2)} = O U_u^{(2)} O, \tilde{\xi}_3 = \xi_1$ and no further constraint, hence the group is $\text{U}(2) \times \text{U}(1)$.

If in addition to the values in table 3:	Add to the flavor group the factor:
$m_u = 0$	$\mathrm{U}(1)_{u_R}$
$m_u = m_d = 0$	$\mathrm{U}(1)_{u_R} \times \mathrm{U}(1)_{d_R}$
$m_u = m_c = 0$	$\mathrm{U}(2)_{u_R}$
$m_u = m_c = m_d = 0$	$\mathrm{U}(2)_{u_R} \times \mathrm{U}(1)_{d_R}$
$m_u = m_c = m_d = m_s = 0$	$\mathrm{U}(2)_{u_R} \times \mathrm{U}(2)_{d_R}$
$m_u = m_c = m_t = 0$	$\mathrm{U}(3)_{u_R}$
$m_u = m_c = m_t = m_d = 0$	$\mathrm{U}(3)_{u_R} \times \mathrm{U}(1)_{d_R}$
$m_u = m_c = m_t = m_d = m_s = 0$	$\mathrm{U}(3)_{u_R} \times \mathrm{U}(2)_{d_R}$
$m_u = m_c = m_t = m_d = m_s = m_b = 0$	$\mathrm{U}(3)_{u_R} \times \mathrm{U}(3)_{d_R}$

Table 5. Additional factor to the flavor symmetry of the SM_4 Lagrangian when quark masses vanish.

$\mathbf{m_u = m_c = m_t}$. When the degeneracy is maximal, the CKM matrix can be fully absorbed by a redefinition of the RH up-quarks:

$$u_R \to V_{\mathrm{CKM}} u_R, \quad V_{\mathrm{CKM}} \to \mathbb{1}. \tag{A.16}$$

The flavor symmetry is therefore at least as large as $\mathrm{U}(1)^3$. With such a CKM matrix, one gets that $U_Q = U_u = U_d$ and the flavor group is determined by the relations in eq. (A.5): the flavor symmetry group is $\mathrm{U}(1)^3$, $\mathrm{U}(2) \times \mathrm{U}(1)$, or $\mathrm{U}(3)$, respectively when no, two, or three down-quark masses are degenerate.

A.2 Vanishing masses at dimension-four

When some masses vanish at dimension-four, the flavor symmetry can contain axial phases. This case complicates the power-counting, since dimension-six Yukawa couplings now yield the leading contribution to the masses, but it can be treated as suggested in section 3.1 for the case of neutrino masses or vanishing Y_d.

Whenever a quark mass goes to zero, the flavor symmetry is enlarged since the LH and RH components now describe independent particles. A flavor symmetry $\mathrm{U}(1)_B$ would then be ugraded to $\mathrm{U}(1)_B \times \mathrm{U}(1)_{u_{R,1}}$ when $m_u \to 0$, all other parameters being kept fixed. For non-degenerate quark masses, the flavor symmetry is always abelian, and taking one mass to zero simply adds a RH $\mathrm{U}(1)$ factor to the flavor symmetry, as discussed just above. On the other hand, when two or three masses are degenerate, taking them to zero together adds a RH factor $\mathrm{U}(2)$ or $\mathrm{U}(3)$ to the flavor symmetry (even if it was abelian for non-zero masses due to some structure in the CKM matrix which distinguishes the different flavors of LH quarks). We list in table 5 the relevant cases, and present in table 6 the associated numbers of new primary sources of CPV.[20] Those new sources of CPV can still be captured

---

[20]The counting is performed as in section 4.2, i.e. one counts how many complex linear invariants (under the SM_4 flavor symmetry) there are. Let us give some examples. In the case of a symmetry $\mathrm{U}(1)^3 \times \mathrm{U}(3)_{u_R}$ and focussing on C_{Hu}, one finds that the only linear invariant reads $\delta^{ij} C_{Hu,ij}$ (summed over i, j). However,

Flavour symmetries of the quark sector of the SM	Bilinears C_{eH} C_{eW} C_{eB}	C_{uH} C_{uG} C_{uW} C_{uB} C_{dH} C_{dG} C_{dW} C_{dB} C_{Hud}	$C_{HL}^{1,3}$ C_{He}	$C_{HQ}^{1,3}$ C_{Hu} C_{Hd}	4-Fermi C_{LL} C_{ee}	C_{Le}	$C_{QQ}^{1,3}$ C_{uu} C_{dd}	$C_{LQ}^{1,3}$ C_{Qe} C_{Lu} C_{eu} C_{Ld} C_{ed}	$C_{ud}^{1,8}$ $C_{Qu}^{1,8}$ $C_{Qd}^{1,8}$	C_{LedQ} $C_{LeQu}^{1,3}$	$C_{QuQd}^{1,8}$
$U(1)_B$	3	9	0	3	0	3	18	9	36	27	81
$U(1)_B \times U(1)_{u_R}$	3	$6^4,9^4,6$	0	3, 1, 3	0	3	18, 5, 18	$9^2,3^2,9^2$	$18^2,36$	27, 18	54
$U(1)_B \times U(1)_{u_R} \times U(1)_{d_R}$	3	$6^8,4$	0	$3,1^2$	0	3	$18,6^2$	$9^2,3^4$	$8,18^2$	18^2	36
$U(1)_B \times U(2)_{u_R}$	3	$3^4,9^4,3$	0	3, 0, 3	0	3	18, 0, 18	$9^2,0^2,9^2$	$6^2,36$	27, 9	27
$U(1)_B \times U(2)_{u_R} \times U(1)_{d_R}$	3	$3^4,6^4,2$	0	3, 0, 1	0	3	18, 0, 6	$9^2,0^2,3^2$	2, 6, 18	18, 9	18
$U(1)^2$	3	5	0	1	0	3	5	3	12	15	33
$U(1)^2 \times U(1)_{u_R}$	3	$(3 \text{ or } 4)^4,5^4,$ 3 or 4	0	1, 0 or 1, 1	0	3	5, 0 or 5, 5	$3^2, (0 \text{ or } 3)^2, 3^2$	$(5 \text{ or } 8)^2, 12$	15, 9 or 12	21 or 24
$U(1)^2 \times U(1)_{u_R} \times U(1)_{d_R}$	3	$(3 \text{ or } 4)^8,$ 2 or 4	0	1, $(0 \text{ or } 1)^2$	0	3	5, $(0 \text{ or } 5)^2$	$3^2, (0 \text{ or } 3)^4$	1 or 3 or 8, $(5 \text{ or } 8)^2$	9 or 12	12 or 13 or 16
$U(1)^2 \times U(2)_{u_R}$	3	$(1 \text{ or } 2)^4, 5,$ 1 or 2	0	1, 0, 1	0	3	5, 0, 5	$3^2,0^2,3^2$	$2^2, 12$	15, 3 or 6	12
$U(1)^2 \times U(2)_{u_R} \times U(1)_{d_R}$	3	$(1 \text{ or } 2)^4, (3 \text{ or } 4)^4,$ 0 or 1 or 2	0	1, 0, 0 or 1	0	3	5, 0, 0 or 5	$3^2,0^2, (0 \text{ or } 3)^2$	0 or 2, 2, 5 or 8	9 or 12, 3	8
$U(1)^2 \times U(2)_{u_R} \times U(2)_{d_R}$	3	$(1 \text{ or } 2)^8,$ 0 or 1	0	1, 0^2	0	3	5, 0^2	$3^2, 0^4$	0, 2^2	3 or 6	4
$U(1)^3$	3	3	0	0	0	3	0	0	3	9	15
$U(1)^3 \times U(1)_{u_R}$	3	$2^4,3^4,2$	0	0	0	3	0	0	$1^2,3$	9, 6	10
$U(1)^3 \times U(1)_{u_R} \times U(1)_{d_R}$	3	$2^8,$ 1 or 2	0	0	0	3	0	0	0 or $1,1^2$	6	7 or 8
$U(1)^3 \times U(2)_{u_R}$	3	$1^4,3^4,1$	0	0	0	3	0	0	$0^2,3$	9, 3	5
$U(1)^3 \times U(2)_{u_R} \times U(1)_{d_R}$	3	$1^4,2^4,$ 0 or 1	0	0	0	3	0	0	$0^2,1$	6, 3	3 or 4
$U(1)^3 \times U(2)_{u_R} \times U(2)_{d_R}$	3	$1^8,$ 0 or 1	0	0	0	3	0	0	0	3	2
$U(1)^3 \times U(3)_{u_R}$	3	$0^4,3^4,0$	0	0	0	3	0	0	$0^2,3$	9, 0	0
$U(1)^3 \times U(3)_{u_R} \times U(1)_{d_R}$	3	$0^4,2^4,0$	0	0	0	3	0	0	$0^2,1$	6, 0	0
$U(2) \times U(1)$	3	2	0	0	0	3	0	0	1	6	7
$U(2) \times U(1) \times U(1)_{u_R}$	3	$1^4,2^4,1$	0	0	0	3	0	0	$0^2,1$	6, 3	4
$U(2) \times U(1) \times U(1)_{u_R} \times U(1)_{d_R}$	3	1	0	0	0	3	0	0	0	3	1 or 2
$U(2) \times U(1) \times U(2)_{u_R}$	3	$1^4,2^4,1$	0	0	0	3	0	0	$0^2,1$	6, 3	3
$U(2) \times U(1) \times U(2)_{u_R} \times U(1)_{d_R}$	3	$1^8,0$	0	0	0	3	0	0	0	3	2
$U(2) \times U(1) \times U(2)_{u_R} \times U(2)_{d_R}$	3	1	0	0	0	3	0	0	0	3	1
$U(2) \times U(1) \times U(3)_{u_R}$	3	$0^4,2^4,0$	0	0	0	3	0	0	$0^2,1$	6, 0	0
$U(2) \times U(1) \times U(3)_{u_R} \times U(1)_{d_R}$	3	$0^4,1^4,0$	0	0	0	3	0	0	0	3, 0	0
$U(2) \times U(1) \times U(3)_{u_R} \times U(2)_{d_R}$	3	$0^4,1^4,0$	0	0	0	3	0	0	0	3, 0	0
$U(3)$	3	1	0	0	0	3	0	0	0	3	2
$U(3) \times U(3)_{u_R}$	3	$0^4,1^4,0$	0	0	0	3	0	0	0	3, 0	0
$U(3) \times U(3)_{u_R} \times U(3)_{d_R}$	3	0	0	0	0	3	0	0	0	0	0
Two degenerate electron-type leptons	$\times\frac{2}{3}$	$\times 1$		$\times 1$	$\times\frac{2}{3}$	$\times 1$	$\times\frac{2}{3}$	$\times 1$	$\times 1$	$\times\frac{2}{3}$	$\times 1$
All electron-type leptons degenerate	$\times\frac{1}{3}$	$\times 1$		$\times 1$	$\times\frac{1}{3}$	$\times 1$	$\times\frac{1}{3}$	$\times 1$	$\times 1$	$\times\frac{1}{3}$	$\times 1$
One vanishing electron-type mass	$\times\frac{2}{3}$	$\times 1$		$\times 1$	$\times\frac{1}{3}$	$\times 1$	$\times 1$	$\times 1$	$\times 1$	$\times\frac{2}{3}$	$\times 1$
Two vanishing electron-type masses	$\times\frac{1}{3}$	$\times 1$		$\times 1$	0	$\times 1$	$\times\frac{2}{3}$	$\times 1$	$\times 1$	$\times\frac{1}{3}$	$\times 1$
All electron-type masses vanishing	0	$\times 1$		$\times 1$	0	$\times 1$	$\times\frac{1}{3}$	$\times 1$	$\times 1$	0	$\times 1$

Table 6. Numbers of new primary sources of CPV contained in each dimension-six SMEFT coefficient. When a single number appears, it applies to all operators at the top of the concerned column. When several numbers are needed, they appear as a list, where the integer power refers to the multiplicity of a given number. An entry "i or j" means that the answer depends on the details of the flavor charges. The last five rows indicate which multiplicative coefficient should be applied to all numbers of the same column for remarkable values of the electron-type lepton masses. The situation where $m_d = 0$, relevant for approximations in high-energy observables, generically corresponds to the line $U(1)^3 \times U(3)_{u_R}$, after a suitable replacement $u \leftrightarrow d$ (note that there is a single $\mathcal{O}_{LeqQ}$ operator in the Warsaw basis, while there are two $\mathcal{O}_{LeQu}^{1,3}$ operators).

by flavor invariants, however there are subtleties to take into account when masses vanish. Namely, the sets of invariants we present in the main text and in appendices D and E have transfer matrices that do not maintain maximal rank in the limit of vanishing masses. Let us illustrate what we mean by this with an example. From table 6, one learns that the three phases in $C_{HQ}^{(1,3)}$ remain primary when $m_u = m_c = 0$. However, the invariants presented in the associated set in table 10 are of the form $\text{Im}\,\text{Tr}\left(X_u M C_{HQ}^{(1,3)}\right)$ for some matrix M built out of the Yukawas. Working in the *up basis* with a vanishing CKM phase, and focusing on the contribution proportional to $C_{HQ,12}^{(1,3)}$, we find

$$\text{Im}\,\text{Tr}\left(X_u M C_{HQ}^{(1,3)}\right) \supset \text{Im}\left(C_{HQ,12}^{(1,3)}\right)\left(m_u^2 M_{21} - m_c^2 M_{12}\right) \tag{A.17}$$

which vanishes when $m_u = m_c = 0$. Therefore, the set of invariants we consider does not allow us to capture the three phases in $C_{HQ}^{(1,3)}$ in such limits. Another example is that of C_{uH}. When $Y_d = 0$, one finds only two invariants in the associated set in table 10, whereas three sources of CPV remain as shown in table 6. One could therefore conclude that one of the invariants in the set should be replaced by the missing

$$\text{Im}\,\text{Tr}\left(X_u^2 C_{uH} Y_u^\dagger\right). \tag{A.18}$$

However, this choice would not allow us to retain a sufficient rank for the set, as one finds

$$\text{Im}\,\text{Tr}\left(X_u^2 C_{uH} Y_u^\dagger\right) = (m_u^2 + m_t^2)\text{Im}\,\text{Tr}\left(X_u C_{uH} Y_u^\dagger\right) - m_u^2 m_t^2 \text{Im}\,\text{Tr}\left(C_{uH} Y_u^\dagger\right) \tag{A.19}$$

when $m_u = m_c$ (one can use formulae like eq. (4.6) to express the mass factors in terms of invariants), whereas all nine sources of CPV in C_{uH} remain primary and independent in this case, as per tables 3 and 4. Therefore, it may seem that one needs strictly more than nine invariants to capture the nine CPV phases in C_{uH}, and more generally that the necessary and sufficient conditions presented in section 4.1 are not sufficient anymore when masses can vanish. However, this is a consequence of our assumption that invariants should correspond to traces of a monomial of degree one in SMEFT coefficients, and arbitrary degree in Yukawa matrices. Instead, one could enlarge the set of invariants and include traces of sums over monomials of various degrees. For instance, defining instead $X_u \equiv 1 + Y_u Y_u^\dagger$, and similarly for other fermions, without changing the expression of the invariants, is sufficient to ensure that the vanishing of our sets is a necessary and sufficient condition for the conservation of CP at leading order.

C_{Hu} being hermitian, this combination is real in all bases and does not violate CP. In the case of a symmetry $\text{U}(1)^2 \times \text{U}(1)_{u_R}$, the $\text{U}(1)^2$ factor indicates that, in some appropriate flavor basis, the SM₄ Lagrangian possesses two independent quark number symmetries (one of which is the usual baryon number), singling out a quark flavor. Assuming without loss of generality that this flavor is the third one and focussing on C_{QuQd}, linear $\text{U}(1)^2$-invariants read $C_{QuQd,ijkl}$, $C_{QuQd,33ij}$, $C_{QuQd,3ji3}$, $C_{QuQd,i33j}$, $C_{QuQd,ij33}$, $C_{QuQd,3333}$ (where $i, j, k, l = 1, 2$), which are 33 complex coefficients. Then, one needs to know whether the massless quark is the one which is singled out by the quark number, i.e. $u_{R,3}$ (otherwise, without loss of generality we take the massless quark to be the first flavor). If so one must discard $C_{QuQd,33ij}$, $C_{QuQd,i33j}$ (reducing the 33 coefficients to 25), if not one must instead discard $C_{QuQd,i1kl}$, $C_{QuQd,31i3}$, $C_{QuQd,i133}$ (reducing to 21 coefficients).

B   Generalities about invariants

B.1   Properties of 3 × 3 matrices

Here we discuss some properties of generic 3×3 matrices, which we use throughout the paper and will refer to later on. We will follow mostly ref. [27]. The starting point is the Cayley-Hamilton theorem, which allows one to rewrite the n-th power of a $n \times n$ matrix A in terms of the powers $< n$, and that for $n = 3$ takes the form

$$A^3 = A^2 \operatorname{Tr}(A) - \frac{1}{2} A \left[\operatorname{Tr}(A)^2 - \operatorname{Tr}\left(A^2\right) \right] + \frac{1}{6} \left[\operatorname{Tr}(A)^3 - 3 \operatorname{Tr}\left(A^2\right) \operatorname{Tr}(A) + 2 \operatorname{Tr}\left(A^3\right) \right] \mathbb{1}_{3\times 3} . \tag{B.1}$$

Multiplying by A and taking the trace results in

$$\operatorname{Tr}\left(A^4\right) = \frac{1}{6} \operatorname{Tr}(A)^4 - \operatorname{Tr}\left(A^2\right) \operatorname{Tr}(A)^2 + \frac{4}{3} \operatorname{Tr}\left(A^3\right) \operatorname{Tr}(A) + \frac{1}{2} \operatorname{Tr}\left(A^2\right)^2 . \tag{B.2}$$

Shifting $A \to A + B + C$ in eq. (B.2), with B and C some other generic 3 matrices, and taking the terms of order $A^2 BC$, one obtains

$$\begin{aligned}
0 = &\operatorname{Tr}(A)^2 \operatorname{Tr}(B) \operatorname{Tr}(C) - \operatorname{Tr}(BC) \operatorname{Tr}(A)^2 - 2 \operatorname{Tr}(AB) \operatorname{Tr}(A) \operatorname{Tr}(C) + \\
&- 2 \operatorname{Tr}(AC) \operatorname{Tr}(A) \operatorname{Tr}(B) + 2 \operatorname{Tr}(ABC) \operatorname{Tr}(A) + 2 \operatorname{Tr}(ACB) \operatorname{Tr}(A) + \\
&- \operatorname{Tr}\left(A^2\right) \operatorname{Tr}(B) \operatorname{Tr}(C) + 2 \operatorname{Tr}(AB) \operatorname{Tr}(AC) + \operatorname{Tr}\left(A^2\right) \operatorname{Tr}(BC) \\
&+ 2 \operatorname{Tr}(C) \operatorname{Tr}\left(A^2 B\right) + 2 \operatorname{Tr}(B) \operatorname{Tr}\left(A^2 C\right) - 2 \operatorname{Tr}\left(A^2 BC\right) - 2 \operatorname{Tr}\left(A^2 CB\right) - 2 \operatorname{Tr}(ABAC) .
\end{aligned} \tag{B.3}$$

This property is useful for our purpose of building sets of invariants, as it implies that we only need to draw from a finite set. Let us focus on invariants relevant for this paper, such as those related to bi-fermion SMEFT operators, which are single-trace and linear with respect to the associated Wilson coefficient. To build such invariants, flavor-invariance imposes that we only use X_u, X_d and C, where $X_{q=u,d} \equiv Y_q Y_q^\dagger$ and C is the Wilson coefficient under study (up to a specific multiplication by a Yukawa matrix for operators of LR chiral structure). In principle, any invariant of the form

$$\operatorname{Tr}\left(X_u^{a_1} X_d^{b_1} X_u^{a_2} X_d^{b_2} \dots C \right) \tag{B.4}$$

is allowed. However, the formulae above imply that one cannot find third or higher powers of X_q in the trace, and that one can find at most one occurrence of X_q and another of X_q^2. These conditions reduce the possible single-trace invariants to a finite set (see appendix B.4 for explicit examples).

Finally, we mention that the Cayley-Hamilton theorem also allows us to write the determinant of a 3×3 matrix as

$$\operatorname{Det}(A) = \frac{1}{6} \left(\operatorname{Tr}(A)^3 - 3 \operatorname{Tr}(A) \operatorname{Tr}\left(A^2\right) + 2 \operatorname{Tr}\left(A^3\right) \right) . \tag{B.5}$$

B.2   Different types of invariants

In ref. [19], the author presents a discussion of CP-violating invariants in supersymmetric models, in order to find basis independent conditions for CP violation, as done here. In

that context, three types of invariants that can be built using three 3×3 matrices A, B, and C are proposed, namely

$$J_{AB} = \operatorname{Im} \operatorname{Tr} \left([A, B]^3 \right) , \qquad\qquad K_{ABC}(p, q, r) = \operatorname{Im} \operatorname{Tr} \left([A^p, B^q] C^r \right) ,$$
$$L_C(p) = \operatorname{Im} \operatorname{Tr} \left(C^p - \text{h.c.} \right) , \tag{B.6}$$

where A and B are hermitian and C generic. These are dubbed J-, K- and L-invariants respectively. In this work we adopted a similar notation, but we only employed L-invariants for our set. However, we can show that this choice is general, as the remaining two types can be written in terms of the last one.[21] To prove this, let us start from J_{AB}. First of all, using eq. (B.5), it can be shown to be equivalent to a Jarlskog-like invariant, i.e.

$$\operatorname{Im} \operatorname{Det} ([A, B]) = \frac{1}{6} \operatorname{Im} \left(\operatorname{Tr} ([A, B])^3 - 3 \operatorname{Tr} ([A, B]) \operatorname{Tr} \left([A, B]^2 \right) + 2 \operatorname{Tr} \left([A, B]^3 \right) \right) +$$
$$= \frac{1}{3} \operatorname{Im} \operatorname{Tr} \left([A, B]^3 \right) , \tag{B.7}$$

as the trace of a commutator vanishes. This also proves eq. (1.1). Then, by expanding $[A, B]^3$ and using the cyclic property of the trace we can show

$$\operatorname{Im} \operatorname{Tr} \left([A, B]^3 \right) = 3 \operatorname{Im} \operatorname{Tr} \left(A^2 B^2 A B - B A B^2 A^2 \right) = L_{A^2 B^2 A B} . \tag{B.8}$$

Finally, we prove that any to K-invariant can also be expressed in terms of the L ones. To do this, it is enough to prove it for $K_{ABC}(1, 1, 1)$ as the other cases can be obtained by redefining A, B or C. Then, we split C in its hermitian and anti-hermitian parts, i.e.

$$C_h \equiv \frac{C + C^\dagger}{2} \qquad\qquad C_a \equiv \frac{C - C^\dagger}{2} . \tag{B.9}$$

and

$$K_{ABC}(1, 1, 1) = \operatorname{Im} \operatorname{Tr} ([A, B]C) = \operatorname{Im} \operatorname{Tr} ([A, B]C_h) + \operatorname{Im} \operatorname{Tr} ([A, B]C_a) , \tag{B.10}$$

where

$$\operatorname{Im} \operatorname{Tr} ([A, B]C_h) = \operatorname{Im} \operatorname{Tr} (ABC_h - BAC_h)$$
$$= \frac{1}{2i} \left[\operatorname{Tr} (ABC_h) - \operatorname{Tr} (ABC_h)^* - \operatorname{Tr} (BAC_h) + \operatorname{Tr} (BAC_h)^* \right]$$
$$= \frac{1}{2i} \left[\operatorname{Tr} (ABC_h) - \operatorname{Tr} (C_h BA) - \operatorname{Tr} (BAC_h) + \operatorname{Tr} (C_h AB) \right]$$
$$= \frac{1}{i} \left[\operatorname{Tr} (ABC_h) - \operatorname{Tr} (C_h BA) \right] = 2 L_{ABC_h} = 2 L_{ABC} . \tag{B.11}$$

With similar steps, one can see that the piece proportional to C_a vanishes, so that

$$K_{ABC}(1, 1, 1) = 2 L_{ABC} . \tag{B.12}$$

---

[21]Also notice that J_{AB} would not suit our scopes as it is not linear in any of the two matrices in the argument.

B.3 Invariant ring and the Hilbert series

The problem of finding a minimal set of invariants we face in this paper is reminiscent of the so called Plethystic Program [66–70] (see also refs. [16, 31] and references therein, and ref. [17] for other recent techniques to compute the Hilbert series). This has been the attempt to apply tools which originate from the study of polynomial rings to physics, and in particular where representation and group theory come into play. The most remarkable success of this program has perhaps been the addition of the Hilbert Series and its Plethystic Logarithm to the theoretical physicist's toolbox. In seeing how these tools apply to this case, we will adopt the logic followed in refs. [16, 31]. Concretely, given a set of parameters $\vec{x}$ and a symmetry group G that acts on $\vec{x}$ as some representation, i.e. $\forall g \in G$, $\exists R(g)$ such that $\vec{x} \to R(g)\vec{x}$, one can define invariants $I(\vec{x})$ as the quantities that obey

$$I(\vec{x}) = I(R(g)\vec{x}) \,. \tag{B.13}$$

As the sum and products of invariants still form an invariant, from an algebraic point of view we talk about a *ring*. Within this ring, one can find the set of invariants $\{I_1, I_2, \ldots, I_m\}$ such that any additional invariant I' can be expressed as a polynomial of those of the set,

$$I' = P'(I_1, I_2, \ldots, I_m) \,. \tag{B.14}$$

The invariants in $\{I_1, I_2, \ldots, I_m\}$ are called generators, and it can be shown that, at least for all the groups relevant to physics, their number is finite. By construction, no relation like eq. (B.14) can exist between the generators. Nevertheless, there could exist a polynomial P such that

$$P(I_1, I_2, \ldots, I_m) = 0 \,. \tag{B.15}$$

Relations of this kind are called *syzygies* in the literature, and the invariants that obey a syzygy are *algebraically dependent*. Taking all syzygies into account we can successively remove invariants until we get to the set of algebraically independent ones. In this setting, the Hilbert series provides a helping hand in finding both the generators and the basic invariants [27]. It is defined as a generating function for the linearly-independent invariants:

$$\mathcal{H}(q) = \sum_{k=1}^{\infty} c_k q^k \,, \tag{B.16}$$

where $c_0 = 1$. c_k denotes precisely the number of linearly-independent invariants at dimension k, and q is an arbitrary spurionic variable satisfying $|q| < 1$, and represents a placeholder for the building blocks of the invariants. Let us make an example. Consider a theory with a coupling m transforming under a U(1) symmetry as

$$m \to e^{i\phi_m} m \,. \tag{B.17}$$

Then the basic invariant is obviously $I = mm^*$, which has dimension 2, and all the invariants of this theory will have the form I^n. Hilbert series will thus have the form

$$\mathcal{H}(q) = 1 + q^2 + q^4 + \cdots = \frac{1}{1 - q^2} \,, \tag{B.18}$$

where q^2 corresponds to I, q^4 to I^2, and so on. It can be shown that, in the general case of a semi-simple Lie algebra, the Hilbert series has the form

$$\mathcal{H}(q) = \frac{N(q)}{D(q)} \,. \tag{B.19}$$

The numerator $N(q)$ is a polynomial of degree d_N with non-negative coefficients and with the property of being palindromic, i.e.

$$N(q) = 1 + c_1 q + c_2 q^2 + \cdots + c_{d_N - 1} q^{d_N - 1} + q^{d_N}, \tag{B.20}$$

with $c_i = c_{d_N - i}$. The denominator takes the form

$$D(q) = \prod_{r=1}^{p} (1 - q^{d_r}) \,, \tag{B.21}$$

and is thus of degree $d_D = \sum_r d_r$. The number of factors is equal to the number of parameters, i.e. of physical observables, and coincides with the number of algebraically independent invariants. Moreover, the denominator provides information on what the algebraically independent invariants look like: a factor $(1 - q^{d_r})^l$ corresponds to l algebraically independent invariants of degree d_r. In the previous example, only one factor is present, corresponding to a single basic (and algebraically independent) invariant mm^*. Indeed, we start with a complex variable m, and we can remove its phase using the U(1), bringing the observables down to 1. If we enlarged our toy-model to have two parameters, m_1 and m_2, transforming under the U(1) symmetry as

$$m_1 \to e^{i\phi_1} m_1 \qquad\qquad m_2 \to e^{i\phi_2} m_2 \,, \tag{B.22}$$

we can build an example of the so called *multi-graded* Hilbert series by assigning different spurions to m_1 and m_2, say q_1 and q_2. The invariants in this case are built as all possible products of all possible powers of $I_{1,2} \equiv m_{1,2} m_{1,2}^*$, which means that the multi-graded Hilbert series is

$$h(q_1, q_2) = (1 + q_1^2 + q_1^4 + \dots)(1 + q_2^2 + q_2^4 + \dots) = \frac{1}{(1 - q_1^2)(1 - q_2^2)} \,. \tag{B.23}$$

The multi-graded Hilbert series can give more information about the structure of the invariants, but does not have in general the properties for the numerator and denominator cited for its ungraded version. The latter can here be easily obtained by setting $q_1 = q_2 = q$, i.e. $H(q) = h(q, q)$. In the examples we showed until now, the numerator has always taken the trivial form $N(q) = 1$, and the set of generators coincided with the algebraically independent invariants. When this happens, the invariant ring is said to be *free*. However, this turns out not to be always the case for more complicated groups and representations.

As one would expect, the computation we could perform straightforwardly by hand in the simple cases above quickly becomes unfeasible when larger groups are involved. Thus a general formula to compute the Hilbert series is called for. The solution is provided by the so called Molien-Weyl formula, which, for a compact, simple Lie group G takes the form:

$$H(q) = \int [d\mu]_G \frac{1}{\det(\mathbb{1} - qR(g))} \,, \tag{B.24}$$

where $[d\mu]_G$ denotes the Haar measure of the group G. If G is connected, the integral can be reduced to an integral over the *maximal torus* of the group, i.e. its largest abelian subgroup, which is just the direct product of r_0 copies of the S^1 unit circle, with r_0 the rank of the group (see e.g. ref. [71]). Thus, the integral is reduced to the computation of residues inside said circles. The integrand $[\det(\mathbb{1} - qR(g))]^{-1}$ can be rewritten as

$$\frac{1}{\det(\mathbb{1} - qR(g))} = \exp\left(\sum_{k=1}^{\infty} \frac{q^k \chi_R(z_1^k, \ldots, z_d^k)}{k}\right) \equiv \text{PE}\left[\chi_R(z_1, \ldots, z_d)q\right], \qquad \text{(B.25)}$$

where $\chi_R(z_1, \ldots, z_d) = \sum_{j=1}^{d} z_j$ is the character function of G in the representation R, d is the dimension of the representation, and z_j (with $j = 1, \ldots, d$) are the eigenvalues of $R(g)$. Eq. (B.25) makes use of the definition of the *Plethystic Exponential* (PE), which for an arbitrary function $f(x_1, \ldots, x_n)$ looks like

$$\text{PE}\left[f(x_1, \ldots, x_n)\right] = \exp\left(\sum_{i}^{\infty} \frac{f(x_1^k, \ldots, x_n^k)}{k}\right). \qquad \text{(B.26)}$$

The generalization to a multi-graded Hilbert series is then straightforward:

$$h(q_1, \ldots, q_n) = \int [d\mu]_G \prod_{i=1}^{n} \text{PE}(z_1, \ldots, z_d; q_i). \qquad \text{(B.27)}$$

An important role in this context is played by the inverse of the PE, quite fittingly called *plethystic logarithm* (PL) and defined so that

$$f(x_1, \ldots, x_n) = \text{PE}\left[g(x_1, \ldots, x_n)\right] \Leftrightarrow g(x_1, \ldots, x_n) = \text{PL}\left[f(x_1, \ldots, x_n)\right]. \qquad \text{(B.28)}$$

It can be proved that

$$\text{PL}\left[f(x_1, \ldots, x_n)\right] = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \ln\left[f(x_1^k, \ldots, x_n^k)\right]. \qquad \text{(B.29)}$$

The value of the PL lies in the following fact: the PL of a Hilbert series is a polynomial whose leading positive terms correspond to the basic invariants, i.e. to the generators, and whose leading negative terms correspond to the syzygies between them. Remarkably, when the invariant ring is free, this polynomial has a finite number of terms. Some complications arise when the groups and the representations that appear become increasingly non-trivial[22] [31, 67]. Equipped with these outstanding tools, we wish to see them applied to our case, which is the main scope of the next section.

---

[22]In ref. [72] the author argues that some of the assumptions in ref. [31] are imprecise, as the finiteness of the generating set of invariants is a consequence of the group being reductive, which is supposedly not the case for U(n). However, as explained there and clarified by the same authors of ref. [31] in ref. [73], the final result is nonetheless correct, as at least the ring of invariants of U(n) is isomorphic to that of GL($n, \mathbb{C}$), which is itself reductive. See further discussions in [74].

B.4 Finding polynomial relations between invariants

In the main body of this work, the logic we have followed to build invariants stemmed from knowing that the relative Wilson coefficient $C^{(6)}$, in a given basis, has a certain number of phases. Then we found as many independent invariants as possible, in order to obtain a transfer matrix whose rank matched the new sources of CPV in $C^{(6)}$ when $J_4 = 0$. Now, however, we could be tempted to pursue a different line of reasoning and find the relevant invariants by applying to our case the power of the Hilbert series and its Plethystic logarithm. Let us restrict to the case of quark bilinear operators. Their Wilson coefficients are generic 3×3 complex matrices that we can multiply by an appropriate number of $Y_{u,d}^{(\dagger)}$ to retrofit them into a $\mathbf{3} \otimes \bar{\mathbf{3}}$ representation of $\mathrm{SU}(3)_Q$. We refer to this combination as $C^{(6)}$ here. Then, the building blocks of our invariants are

$$C^{(6)}, (C^{(6)})^{\dagger}, X_{u,d} \in \mathbf{3} \otimes \bar{\mathbf{3}}. \tag{B.30}$$

and we can build the multi-graded Hilbert series

$$h(c, c^{\dagger}, x_u, x_d) = \int [\mathrm{d}\mu]_{\mathrm{SU}(3)} \prod_{i=\{c,c^{\dagger},x_u,x_d\}} \mathrm{PE}(\vec{z}; c), \tag{B.31}$$

with obvious associations between a spurion and the corresponding building block. The resulting expression is quite long and not particularly illuminating, so we will refrain from presenting it here. However, we can look at its ungraded version

$$H(q) = h(q, q, q, q) = \frac{N(q)}{D(q)}, \tag{B.32}$$

with

$$\begin{aligned} N(q) = &+ q^{34} + 14q^{31} + 31q^{30} + 56q^{29} + 165q^{28} + 354q^{27} + 660q^{26} + 1256q^{25} + 2097q^{24} \\ &+ 3184q^{23} + 4720q^{22} + 6404q^{21} + 7992q^{20} + 9536q^{19} + 10510q^{18} + 10744q^{17} \\ &+ 10510q^{16} + 9536q^{15} + 7992q^{14} + 6404q^{13} + 4720q^{12} + 3184q^{11} + 2097q^{10} \\ &+ 1256q^9 + 660q^8 + 354q^7 + 165q^6 + 56q^5 + 31q^4 + 14q^3 + 1, \tag{B.33} \end{aligned}$$

and

$$D(q) = (1-q)^4 (1-q^2)^{10} (1-q^3)^{10} (1-q^4)^4. \tag{B.34}$$

We can see that the numerator has the correct palindromic structure we expected, and more importantly the denominators contain 28 factors, correctly matching the 10 observables from the Standard Model and the 18 new (9 real + 9 imaginary) observables contained in $C^{(6)}$. This is already quite remarkable. However, to this point we have neither an idea of how the algebraically independent invariants look like, nor a way to extract the ones that are linear in $C^{(6)}$, which is the subset we are really interested in. To gain some more

insight, let us look at the Plethystic logarithm of the multi-graded Hilbert series:

$$
\begin{aligned}
\mathrm{PL}\left[h(c,c^\dagger,x_u,x_d)\right] = {} & (x_u + x_d) + (x_u^2 + x_d x_u + x_d^2) + (x_u^3 + x_u^2 x_d + x_u x_d^2 + x_d^3) + x_u^2 x_d^2 \\
& + x_u^3 x_d^3 - x_u^6 x_d^6 \\
& + \left(c + c^\dagger\right)\left[1 + x_u + x_d + (x_u^2 + x_d^2 + 2x_u x_d) + (2x_u^2 x_d + 2x_u x_d^2)\right. \\
& \left. + (x_u^3 x_d + 2x_u^2 x_d^2 + x_d^3 x_u) + (x_u^3 x_d^2 + x_u^2 x_d^3)\right] \\
& + \mathcal{O}\left(c^2, (c^\dagger)^2, (c+c^\dagger)x_u^3 x_d^4, (c+c^\dagger)x_u^4 x_d^3\right).
\end{aligned}
\tag{B.35}
$$

Since we are interested in invariants that are linear in $C^{(6)}$, we stopped the expansion at $\mathcal{O}\left(c,c^\dagger\right)$. The $\mathcal{O}\left(c^0 (c^\dagger)^0\right)$ terms in this expansion correspond to the invariants one can obtained in the quark sector of the Standard Model. This case has been treated in ref. [27], and the resulting algebraically independent invariants are

$$
\begin{aligned}
I_{1,0} &= \mathrm{Tr}(X_u) & I_{0,1} &= \mathrm{Tr}(X_d) \\
I_{2,0} &= \mathrm{Tr}\left(X_u^2\right) & I_{1,1} &= \mathrm{Tr}(X_u X_d) \\
I_{0,2} &= \mathrm{Tr}\left(X_d^2\right) & I_{3,0} &= \mathrm{Tr}\left(X_u^3\right) \\
I_{2,1} &= \mathrm{Tr}\left(X_u^2 X_d\right) & I_{1,2} &= \mathrm{Tr}\left(X_u X_d^2\right) \\
I_{0,3} &= \mathrm{Tr}\left(X_d^3\right) & I_{2,2} &= \mathrm{Tr}\left(X_u^2 X_d^2\right).
\end{aligned}
\tag{B.36}
$$

Notice their number is 10, correctly matching the 6 masses + 3 angles + 1 phase of the Standard Model. The generating set contain one additional invariant,

$$
I_{3,3}^{(-)} = \mathrm{Tr}\left(X_u^2 X_d^2 X_u X_d\right) - \mathrm{Tr}\left(X_d^2 X_u^2 X_d X_u\right),
\tag{B.37}
$$

corresponding to the $x_u^3 x_d^3$ term in eq. (B.35) and which is nothing but J_4. In the language adopted here, this invariant does not contain any additional observable, and is just needed to capture the sign of the SM$_4$ phase δ. The negative term $-x_u^6 x_d^6$ at the end signals that there is a syzygy at degree 12, which in this case corresponds to the fact that $\left(I_{3,3}^{(-)}\right)^2$ can be expressed in terms of the remaining 10 invariants, as expected.

The part linear in c and $c^\dagger$ of eq. (B.35) points us at the basic invariants linear in $C^{(6)}$. We see that in this case the set of basic invariants is composed by 34 element, 17 each for $C^{(6)}$ and $(C^{(6)})^\dagger$, which is larger than the basic set. Indeed, the latter is expected to have 18 elements, corresponding to the 9 new complex observables contained in $C^{(6)}$. To try and build the invariants in the generating set, we will make use of the relations showed in section B.1. Given a generic matrix $C^{(6)} \in \mathbf{3} \otimes \bar{\mathbf{3}}$ of SU(3)$_Q$, we want to contract it with as many $X_{u,d}$'s as needed to form all the possible independent invariants. Using eq. (B.1) on $X_{u,d}$, we can show that all invariants written using $X_{u,d}^n$, with $n \geq 3$, are redundant, and we only need $X_{u,d}^2$ and $X_{u,d}$ as building blocks. Moreover, using eq. (B.3), one can eliminate any invariant where a matrix is repeated. Taking into account these simplifications, one can see that the set of possible invariants is finite, and is formed by these 29 objects:

$$
\begin{array}{lll}
\mathrm{Tr}(C^{(6)}) & \mathrm{Tr}(X_u C^{(6)}) & \mathrm{Tr}(X_d C^{(6)}) \\
\mathrm{Tr}\left(X_u^2 C^{(6)}\right) & \mathrm{Tr}\left(X_d^2 C^{(6)}\right) & \mathrm{Tr}\left(X_u X_d C^{(6)}\right)
\end{array}
$$

$$
\begin{aligned}
&\text{Tr}\left(X_u X_d^2 C^{(6)}\right) && \text{Tr}\left(X_d X_u C^{(6)}\right) && \text{Tr}\left(X_d^2 X_u C^{(6)}\right) \\
&\text{Tr}\left(X_d X_u^2 C^{(6)}\right) && \text{Tr}\left(X_u^2 X_d C^{(6)}\right) && \text{Tr}\left(X_u^2 X_d^2 C^{(6)}\right) \\
&\text{Tr}\left(X_d^2 X_u^2 C^{(6)}\right) && \text{Tr}\left(X_u X_d X_u^2 C^{(6)}\right) && \text{Tr}\left(X_d X_u X_d^2 C^{(6)}\right) \\
&\text{Tr}\left(X_u^2 X_d X_u C^{(6)}\right) && \text{Tr}\left(X_d^2 X_u X_d C^{(6)}\right) && \text{Tr}\left(X_u X_d^2 X_u^2 C^{(6)}\right) \\
&\text{Tr}\left(X_u^2 X_d^2 X_u C^{(6)}\right) && \text{Tr}\left(X_d X_u^2 X_d^2 C^{(6)}\right) && \text{Tr}\left(X_d^2 X_u^2 X_d C^{(6)}\right) \\
&\text{Tr}\left(X_u X_d X_u^2 X_d^2 C^{(6)}\right) && \text{Tr}\left(X_u X_d^2 X_u^2 X_d C^{(6)}\right) && \text{Tr}\left(X_d X_u X_d^2 X_u^2 C^{(6)}\right) \\
&\text{Tr}\left(X_d X_u^2 X_d^2 X_u C^{(6)}\right) && \text{Tr}\left(X_u^2 X_d X_u X_d^2 C^{(6)}\right) && \text{Tr}\left(X_u^2 X_d^2 X_u X_d C^{(6)}\right) \\
&\text{Tr}\left(X_d^2 X_u X_d X_u^2 C^{(6)}\right) && \text{Tr}\left(X_d^2 X_u^2 X_d X_u C^{(6)}\right) \,,
\end{aligned}
\tag{B.38}
$$

and the same for $C^{(6)} \to (C^{(6)})^\dagger$. Now, since the generating set only includes 17 elements, this means that 12 invariants of eq. (B.38) can be expressed as polynomials of the remaining ones and can thus be eliminated. To do this, we employ a numerical algorithm adapted from appendix C of ref. [31] (see also ref. [16]). The logic is as follows: by assigning a dummy dimension to the building blocks, i.e. $[X_{u,d}] = 1$ and $\left[C^{(6)}\right] = 1$, we can assign a dimension to all the invariants listed above. Then, we fix some given dimension n. Picking one of the invariants in eq. (B.38), one can then take its product with as many traces from eq. (B.36) as needed to form a monomial M_i of dimension n. Repeating this for all instances of eq. (B.38), we find the set $\{M_i\}$ of all the possible monomials that are dimension n and linear in C^6. For example, at dimension $n = 2$ one can obtain the monomials

$$
\{M_i\} = \left\{ \text{Tr}\left(C^{(6)}\right) I_{1,0}, \text{Tr}\left(C^{(6)}\right) I_{0,1}, \text{Tr}\left(X_u C^{(6)}\right), \text{Tr}\left(X_d C^{(6)}\right) \right\} .
\tag{B.39}
$$

Then we set a linear combination of these monomials to zero, i.e.

$$
\sum_i a_i M_i = 0 \,,
\tag{B.40}
$$

where the a_i's are integer coefficients. We then plug random integer values for the entries of the matrices $X_{u,d}$ and C^6. This produces a linear equation for the a_i's. Repeating this last step as many times as there are M_i's, one builds a linear system for the a_i's with zero constant term. The number of independent directions of the null space of the corresponding matrix matches the number of possible relations between the M_i's. The first nontrivial result is found at dimension 5, where we get the two relations

$$
\begin{aligned}
&\text{Tr}\left(X_d X_u X_d^2 C^{(6)}\right) + \text{Tr}\left(X_d^2 X_u X_d C^{(6)}\right) + I_{0,1}\left(\text{Tr}\left(X_d^2 X_u C^{(6)}\right) + \text{Tr}\left(X_u X_d^2 C^{(6)}\right)\right) \\
&\quad + (-I_{0,1} I_{1,0} - I_{1,1})\text{Tr}\left(X_d^2 C^{(6)}\right) - I_{0,1}^2 \left(\text{Tr}\left(X_d X_u C^{(6)}\right) + \text{Tr}\left(X_u X_d C^{(6)}\right)\right) \\
&\quad + \frac{1}{3}\left(2 I_{0,1}^3 - 3 I_{0,2} I_{0,1} + I_{0,3}\right) \text{Tr}\left(X_u C^{(6)}\right) + \left(I_{0,1}^2 I_{1,0} - I_{1,2}\right) \text{Tr}\left(X_d C^{(6)}\right) \\
&\quad + \text{Tr}(C^{(6)})\left(-\frac{2}{3} I_{0,1}^3 I_{1,0} + I_{0,1}^2 I_{1,1} + I_{0,2} I_{0,1} I_{1,0} - I_{0,1} I_{1,2} - \frac{1}{3} I_{0,3} I_{1,0}\right) = 0 \,,
\end{aligned}
\tag{B.41}
$$

and the same with $X_u \leftrightarrow X_d$, which we can use to remove $\mathrm{Tr}\left(X_d X_u X_d^2 C^{(6)}\right)$ and $\mathrm{Tr}\left(X_u X_d X_u^2 C^{(6)}\right)$ from the set. At dimension 6 we obtain two more relations, and 8 more are obtained at dimension 7. With these 12 expressions, we can reduce the set to

$$
\begin{array}{lll}
\mathrm{Tr}\left(C^{(6)}\right) & \mathrm{Tr}\left(X_u C^{(6)}\right) & \mathrm{Tr}\left(X_d C^{(6)}\right) \\
\mathrm{Tr}\left(X_u^2 C^{(6)}\right) & \mathrm{Tr}\left(X_d^2 C^{(6)}\right) & \mathrm{Tr}\left(X_u X_d C^{(6)}\right) \\
\mathrm{Tr}\left(X_u X_d^2 C^{(6)}\right) & \mathrm{Tr}\left(X_d X_u C^{(6)}\right) & \mathrm{Tr}\left(X_d^2 X_u C^{(6)}\right) \\
\mathrm{Tr}\left(X_d X_u^2 C^{(6)}\right) & \mathrm{Tr}\left(X_u^2 X_d C^{(6)}\right) & \mathrm{Tr}\left(X_u^2 X_d^2 C^{(6)}\right) \\
\mathrm{Tr}\left(X_d^2 X_u^2 C^{(6)}\right) & \mathrm{Tr}\left(X_u X_d X_u^2 C^{(6)}\right) & \mathrm{Tr}\left(X_d X_u X_d^2 C^{(6)}\right) \\
\mathrm{Tr}\left(X_u X_d^2 X_u^2 C^{(6)}\right) & \mathrm{Tr}\left(X_d X_u^2 X_d^2 C^{(6)}\right).
\end{array}
\tag{B.42}
$$

The same can be repeated for $(C^{(6)})^\dagger$. These are, quite remarkably, exactly in correspondence with the relative terms in eq. (B.35). However, we now wish to find the additional relations that help us express the 8 too many (complex) invariants we have in eq. (B.42) in terms of the 9 ones we know are sufficient to express all the physical observables, i.e. the algebraically independent ones. If we expand a bit further in eq. (B.35), we see that the next two terms are degree 8 and are negative, $-(c + c^\dagger)x_u^4 x_d^3 - (c + c^\dagger)x_u^3 x_d^4$. They should then correspond to the number of syzygies at dimension 8. To obtain them explicitly, we just run again the described algorithm at dimension 8, obtaining indeed two syzygies of the expected degree. They include 107 terms out of the possible 808 one can build at this dimension, and allow us to remove $\mathrm{Tr}\left(X_u X_d^2 X_u^2 C^{(6)}\right)$ and $\mathrm{Tr}\left(X_d X_u^2 X_d^2 C^{(6)}\right)$. Running this argument at degree 9, however, we run into a mismatch. Indeed, even though the next term in eq. (B.35) would call for 4 syzygies, we only find 1, symmetric under the exchange $X_u \leftrightarrow X_d$. This is probably due to the complications that arise when the groups and representations one has to deal with start becoming less and less trivial, as in our case, and that forbid us from reading the syzygies from the negative terms directly. For a deeper discussion of this topic, see in particular refs. [31, 67] and references therein. One thing to notice, in addition, is that requiring the building blocks to be linear in $(C^{(6)})$, although justified from a physical point of view, breaks the ring structure of the invariant ring, as obviously the set we consider is no longer closed under multiplications.

In any case, even without the Plethystic logarithm as a guide, we can just run our algorithm at increasingly higher dimensions, until no more relations are found. Indeed, upon going up to dimension $n = 13$, one manages to reduce the set down to 9 independent invariants, which we can pick to be

$$
\begin{array}{lll}
\mathrm{Tr}(C^{(6)}) & \mathrm{Tr}\left(X_u C^{(6)}\right) & \mathrm{Tr}\left(X_d C^{(6)}\right) \\
\mathrm{Tr}\left(X_u X_d C^{(6)}\right) & \mathrm{Tr}\left(X_d X_u C^{(6)}\right) & \mathrm{Tr}\left(X_d^2 X_u^2 C^{(6)}\right) \\
\mathrm{Tr}\left(X_u^2 X_d^2 C^{(6)}\right) & \mathrm{Tr}\left(X_d X_u^2 X_d^2 C^{(6)}\right) & \mathrm{Tr}\left(X_u X_d^2 X_u^2 C^{(6)}\right),
\end{array}
\tag{B.43}
$$

which, upon taking its imaginary parts, matches the minimal set in for a non-hermitian fermion bilinear operator table 10.

	Bilinears					
Label		Operator	# real entries	# imaginary entries	# primary real entries	# primary imaginary entries
Modified Yukawas	Q_{eH}	$(H^\dagger H)(\bar{L}_i e_j H) + \text{h.c.}$	9	9	3	3
	Q_{uH}	$(H^\dagger H)(\bar{Q}_i u_j \widetilde{H}) + \text{h.c.}$	9	9	9	9
	Q_{dH}	$(H^\dagger H)(\bar{Q}_i d_j H) + \text{h.c.}$	9	9	9	9
Dipole	Q_{eW}	$(\bar{L}_i \sigma^{\mu\nu} e_j)\tau^I H W^I_{\mu\nu} + \text{h.c.}$	9	9	3	3
	Q_{eB}	$(\bar{L}_i \sigma^{\mu\nu} e_j) H B_{\mu\nu} + \text{h.c.}$	9	9	3	3
	Q_{uG}	$(\bar{Q}_i \sigma^{\mu\nu} T^A u_j)\widetilde{H} G^A_{\mu\nu} + \text{h.c.}$	9	9	9	9
	Q_{uW}	$(\bar{Q}_i \sigma^{\mu\nu} u_j)\tau^I \widetilde{H} W^I_{\mu\nu} + \text{h.c.}$	9	9	9	9
	Q_{uB}	$(\bar{Q}_i \sigma^{\mu\nu} u_j)\widetilde{H} B_{\mu\nu} + \text{h.c.}$	9	9	9	9
	Q_{dG}	$(\bar{Q}_i \sigma^{\mu\nu} T^A d_j) H G^A_{\mu\nu} + \text{h.c.}$	9	9	9	9
	Q_{dW}	$(\bar{Q}_i \sigma^{\mu\nu} d_j)\tau^I H W^I_{\mu\nu} + \text{h.c.}$	9	9	9	9
	Q_{dB}	$(\bar{Q}_i \sigma^{\mu\nu} d_j) H B_{\mu\nu} + \text{h.c.}$	9	9	9	9
Current-current	$Q^{(1)}_{HL}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{L}_i \gamma^\mu L_j)$	6	3	3	0
	$Q^{(3)}_{HL}$	$(H^\dagger i \overleftrightarrow{D}^I_\mu H)(\bar{L}_i \tau^I \gamma^\mu L_j)$	6	3	3	0
	Q_{He}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{e}_i \gamma^\mu e_j)$	6	3	3	0
	$Q^{(1)}_{HQ}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{Q}_i \gamma^\mu Q_j)$	6	3	6	3
	$Q^{(3)}_{HQ}$	$(H^\dagger i \overleftrightarrow{D}^I_\mu H)(\bar{Q}_i \tau^I \gamma^\mu Q_j)$	6	3	6	3
	Q_{Hu}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{u}_i \gamma^\mu u_j)$	6	3	6	3
	Q_{Hd}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{d}_i \gamma^\mu d_j)$	6	3	6	3
	Q_{Hud}	$i(\widetilde{H}^\dagger D_\mu H)(\bar{u}_i \gamma^\mu d_j) + \text{h.c.}$	9	9	9	9

Table 7. The list of dimension-6 fermionic bilinear operators of SMEFT, as given in ref. [37], together with the number of real and imaginary entries they each contain, as well as the number of primary parameters (highlighted in gray, see the text for more details). When +h.c. is specified, the hermitian conjugate of the operator must be included, too. We indicate with i, j, k, l the flavor indices and with a, b indices in the fundamental of $\text{SU}(2)_L$. T^A, $A = 1, \ldots, 8$ are the generators of the gauge $\text{SU}(3)_c$, while $\tau^I = \frac{\sigma^I}{2}$, $I = 1, 2, 3$ are the generators of $\text{SU}(2)_L$, with σ^I the Pauli matrices.

C List of dimension-6 fermionic operators and parameter counting with generic N_f

In tables 7 and 8 we reproduce the subset of operators from ref. [37] we are interested in this work, namely dimension-6 operators in SMEFT containing fermions, split between bilinear and 4-Fermi operators. For each of the considered operators we list the number of real and imaginary entries and compare them with the number of (primary) real and imaginary parameters that can appear in observables at order $1/\Lambda^2$, as explained in the main text.

In table 9 the counting of independent primary parameters is generalized to an arbitrary number of flavors N.

4-Fermi						
	Label	Operator	# real entries	# imaginary entries	# primary real entries	# primary imaginary entries
LLLL	Q_{LL}	$(\bar{L}_i\gamma_\mu L_j)(\bar{L}_k\gamma^\mu L_l)$	27	18	9	0
	$Q_{QQ}^{(1)}$	$(\bar{Q}_i\gamma_\mu Q_j)(\bar{Q}_k\gamma^\mu Q_l)$	27	18	27	18
	$Q_{QQ}^{(3)}$	$(\bar{Q}_i\gamma_\mu\tau^I Q_j)(\bar{Q}_k\gamma^\mu\tau^I Q_l)$	27	18	27	18
	$Q_{LQ}^{(1)}$	$(\bar{L}_i\gamma_\mu L_j)(\bar{Q}_k\gamma^\mu Q_l)$	45	36	18	9
	$Q_{LQ}^{(3)}$	$(\bar{L}_i\gamma_\mu\tau^I L_j)(\bar{Q}_k\gamma^\mu\tau^I Q_l)$	45	36	18	9
RRRR	Q_{ee}	$(\bar{e}_i\gamma_\mu e_j)(\bar{e}_k\gamma^\mu e_l)$	21	15	6	0
	Q_{uu}	$(\bar{u}_i\gamma_\mu u_j)(\bar{u}_k\gamma^\mu u_l)$	27	18	27	18
	Q_{dd}	$(\bar{d}_i\gamma_\mu d_j)(\bar{d}_k\gamma^\mu d_l)$	27	18	27	18
	Q_{eu}	$(\bar{e}_i\gamma_\mu e_j)(\bar{u}_k\gamma^\mu u_l)$	45	36	18	9
	Q_{ed}	$(\bar{e}_i\gamma_\mu e_j)(\bar{d}_k\gamma^\mu d_l)$	45	36	18	9
	$Q_{ud}^{(1)}$	$(\bar{u}_i\gamma_\mu u_j)(\bar{d}_k\gamma^\mu d_l)$	45	36	45	36
	$Q_{ud}^{(8)}$	$(\bar{u}_i\gamma_\mu T^A u_j)(\bar{d}_k\gamma^\mu T^A d_l)$	45	36	45	36
LLRR	Q_{Le}	$(\bar{L}_i\gamma_\mu L_j)(\bar{e}_k\gamma^\mu e_l)$	45	36	12	3
	Q_{Lu}	$(\bar{L}_i\gamma_\mu L_j)(\bar{u}_k\gamma^\mu u_l)$	45	36	18	9
	Q_{Ld}	$(\bar{L}_i\gamma_\mu L_j)(\bar{d}_k\gamma^\mu d_l)$	45	36	18	9
	Q_{Qe}	$(\bar{Q}_i\gamma_\mu Q_j)(\bar{e}_k\gamma^\mu e_l)$	45	36	18	9
	$Q_{Qu}^{(1)}$	$(\bar{Q}_i\gamma_\mu Q_j)(\bar{u}_k\gamma^\mu u_l)$	45	36	45	36
	$Q_{Qu}^{(8)}$	$(\bar{Q}_i\gamma_\mu T^A Q_j)(\bar{u}_k\gamma^\mu T^A u_l)$	45	36	45	36
	$Q_{Qd}^{(1)}$	$(\bar{Q}_i\gamma_\mu Q_j)(\bar{d}_k\gamma^\mu d_l)$	45	36	45	36
	$Q_{Qd}^{(8)}$	$(\bar{Q}_i\gamma_\mu T^A Q_j)(\bar{d}_k\gamma^\mu T^A d_l)$	45	36	45	36
LRRL	Q_{LedQ}	$(\bar{L}_i^a e_j)(\bar{d}_k Q_{la})$ + h.c.	81	81	27	27
LRLR	$Q_{QuQd}^{(1)}$	$(\bar{Q}_i^a u_j)\epsilon_{ab}(\bar{Q}_k^b d_l)$ + h.c.	81	81	81	81
	$Q_{QuQd}^{(8)}$	$(\bar{Q}_i^a T^A u_j)\epsilon_{ab}(\bar{Q}_k^b T^A d_l)$ + h.c.	81	81	81	81
	$Q_{LeQu}^{(1)}$	$(\bar{L}_i^a e_j)\epsilon_{ab}(\bar{Q}_k^b u_l)$ + h.c.	81	81	27	27
	$Q_{LeQu}^{(3)}$	$(\bar{L}_i^a\sigma_{\mu\nu} e_j)\epsilon_{ab}(\bar{Q}_s^k\sigma^{\mu\nu} u_t)$ + h.c.	81	81	27	27

Table 8. The list of dimension-6 4-Fermi operators of SMEFT, as given in ref. [37], together with the number of real and imaginary entries they each contain, as well as the number of primary parameters (highlighted in gray, see the text for more details). When +h.c. is specified, the hermitian conjugate of the operator must be included, too. We indicate with i, j, k, l the flavor indices and with a, b indices in the fundamental of SU(2)$_L$. T^A, $A = 1,\ldots,8$ are the generators of the gauge SU(3)$_c$, while $\tau^I = \frac{\sigma^I}{2}$, $I = 1, 2, 3$ are the generators of SU(2)$_L$, with σ^I the Pauli matrices.

	Type of op.	# ops	# real	# im.
bilinears	Yuk.	3	# of entries at $\mathcal{O}(1/\Lambda^2)$	
			$3N^2$	$3N^2$
			# of primary parameters entering observables at $\mathcal{O}(1/\Lambda^2)$	
			$2N^2 + N$	$2N^2 + N$
	Dipole	8	$8N^2$	$8N^2$
			$6N^2 + 2N$	$6N^2 + 2N$
	curr-curr	8	$\frac{1}{2}N(9N + 7)$	$\frac{1}{2}N(9N - 7)$
			$N(3N + 5)$	$N(3N - 2)$
	all bilinears	19	$\frac{1}{2}N(31N + 7)$	$\frac{1}{2}N(31N - 7)$
			$N(11N + 8)$	$N(11N + 1)$
4-Fermi	LLLL	5	$\frac{1}{4}N^2\left(7N^2 + 13\right)$	$\frac{7}{4}N^2\left(N^2 - 1\right)$
			$\frac{1}{2}N^2\left(N^2 + 2N + 7\right)$	$\frac{1}{2}N^2\left(N^2 + 2N - 3\right)$
	RRRR	7	$\frac{1}{8}N\left(21N^3 + 2N^2 + 31N + 2\right)$	$\frac{1}{8}N(21N + 2)(N^2 - 1)$
			$\frac{1}{2}N\left(3N^3 + 2N^2 + 8N + 1\right)$	$\frac{1}{2}N^2\left(3N^2 + 2N - 5\right)$
	LLRR	8	$4\,N^2\left(N^2 + 1\right)$	$4N^2\left(N^2 - 1\right)$
			$\frac{1}{2}N\left(4N^3 + 3N^2 + 9N + 2\right)$	$\frac{1}{2}N\left(4N^3 + 3N^2 - 6N - 1\right)$
	LRRL	1	N^4	N^4
			N^3	N^3
	LRLR	4	$4N^4$	$4N^4$
			$2N^3(N + 1)$	$2N^3(N + 1)$
	all 4-Fermi	25	$\frac{1}{8}N\left(107N^3 + 2N^2 + 89N + 2\right)$	$\frac{1}{8}N\left(107N^3 + 2N^2 - 67N - 2\right)$
			$\frac{1}{2}N\left(12N^3 + 13N^2 + 24N + 3\right)$	$\frac{1}{2}N\left(12N^3 + 13N^2 - 14N - 1\right)$
	all	44	$\frac{1}{8}N\left(107N^3 + 2N^2 + 213N + 30\right)$	$\frac{1}{8}N\left(107N^3 + 2N^2 + 57N - 30\right)$
			$\frac{1}{2}N\left(12N^3 + 13N^2 + 46N + 19\right)$	$\frac{1}{2}N\left(12N^3 + 13N^2 + 8N + 1\right)$

Table 9. Number of flavorful real and imaginary parameters in SMEFT at dimension-six with N flavors. For each type of operator, the first line (in white) counts the number of physical parameters, while the second one (highlighted in gray) counts those which are also primary.

Wilson coefficient	Number of phases	Minimal set
$C_e \equiv \begin{cases} C_{eH} \\ C_{eW} \\ C_{eB} \end{cases}$	3	$\left\{ L_0\left(C_e Y_e^\dagger\right)\ L_1\left(C_e Y_e^\dagger\right)\ L_2\left(C_e Y_e^\dagger\right) \right\}$
$C_u \equiv \begin{cases} C_{uH} \\ C_{uG} \\ C_{uW} \\ C_{uB} \end{cases}$	9	$\left\{ \begin{array}{ccc} L_{0000}\left(C_u Y_u^\dagger\right) & L_{1000}\left(C_u Y_u^\dagger\right) & L_{0100}\left(C_u Y_u^\dagger\right) \\ L_{1100}\left(C_u Y_u^\dagger\right) & L_{0110}\left(C_u Y_u^\dagger\right) & L_{2200}\left(C_u Y_u^\dagger\right) \\ L_{0220}\left(C_u Y_u^\dagger\right) & L_{1220}\left(C_u Y_u^\dagger\right) & L_{0122}\left(C_u Y_u^\dagger\right) \end{array} \right\}$
$C_d \equiv \begin{cases} C_{dH} \\ C_{dG} \\ C_{dW} \\ C_{dB} \end{cases}$		Same with $C_u Y_u^\dagger \to C_d Y_d^\dagger$
C_{Hud}		Same with $C_u Y_u^\dagger \to Y_u C_{Hud} Y_d^\dagger$
$C_{HL}^{(1,3)}, C_{He}$	0	$\emptyset$
$C_{HQ}^{(1,3)}$	3	$\left\{ L_{1100}\left(C_{HQ}^{(1,3)}\right)\ L_{2200}\left(C_{HQ}^{(1,3)}\right)\ L_{1122}\left(C_{HQ}^{(1,3)}\right) \right\}$
C_{Hu}		Same with $C_{HQ}^{(1,3)} \to Y_u C_{Hu} Y_u^\dagger$
C_{Hd}		Same with $C_{HQ}^{(1,3)} \to Y_d C_{Hd} Y_d^\dagger$

Table 10. Minimal sets of CP-odd flavor invariants for all SMEFT dimension-six Wilson coefficients associated to operators bilinear in fermion fields. We recall that $X_u \equiv Y_u Y_u^\dagger$, and similarly for down quarks or electrons. We also recall the definition in eq. (4.1). We also defined for the leptons $L_a(\tilde{C}) \equiv \mathrm{Im}\,\mathrm{Tr}\left(X_e^a \tilde{C}\right)$ with $a = 1, 2$.

D Complete minimal set of invariants for 2-Fermi operators

We list in table 10 a valid choice of minimal sets of CP-odd flavor invariants for all dimension-six Wilson coefficients associated to operators that are bilinear in fermion fields. It can be shown that they provide independent conditions matching the numbers presented in table 4, in the generic and non-generic cases listed in table 3.

E Complete minimal set of 4-Fermi invariants

We list in tables 11, 12, and 13 a valid choice of minimal sets of CP-odd flavor invariants for all dimension-six Wilson coefficients associated to operators quartic in fermion fields. It can be shown that they provide independent conditions matching the numbers presented in table 4, in the generic and non-generic cases listed in table 3.

– 43 –

Wilson coefficient	Number of phases	Minimal set
C_{LL}, C_{ee}	0	$\emptyset$
C_{Le}	3	$\left\{ B_0^0 \left(C_{LL\tilde{e}\tilde{e}} \right) \ B_0^1 \left(C_{LL\tilde{e}\tilde{e}} \right) \ B_0^2 \left(C_{LL\tilde{e}\tilde{e}} \right) \right\}$
C_{Qe}		$\left\{ \begin{array}{ccc} A_0^{1100} \left(C_{QQee} \right) & A_1^{1100} \left(C_{QQee} \right) & A_2^{1100} \left(C_{QQee} \right) \\ A_0^{2200} \left(C_{QQee} \right) & A_1^{2200} \left(C_{QQee} \right) & A_2^{2200} \left(C_{QQee} \right) \\ A_0^{1122} \left(C_{QQee} \right) & A_1^{1122} \left(C_{QQee} \right) & A_2^{1122} \left(C_{QQee} \right) \end{array} \right\}$
C_{ed}		Same with $C_{QQee} \to C_{ee\tilde{d}\tilde{d}}$ (exchanging upper with lower indices and with $Y_e \leftrightarrow Y_e^\dagger$)
C_{eu}	9	Same with $C_{QQee} \to C_{ee\tilde{u}\tilde{u}}$ (exchanging upper with lower indices and with $Y_e \leftrightarrow Y_e^\dagger$)
$C_{LQ}^{(1,3)}$		$\left\{ \begin{array}{ccc} A_{1100}^0 \left(C_{LQ}^{(1,3)} \right) & A_{1100}^1 \left(C_{LQ}^{(1,3)} \right) & A_{1100}^2 \left(C_{LQ}^{(1,3)} \right) \\ A_{2200}^0 \left(C_{LQ}^{(1,3)} \right) & A_{2200}^1 \left(C_{LQ}^{(1,3)} \right) & A_{2200}^2 \left(C_{LQ}^{(1,3)} \right) \\ A_{1122}^0 \left(C_{LQ}^{(1,3)} \right) & A_{1122}^1 \left(C_{LQ}^{(1,3)} \right) & A_{1122}^2 \left(C_{LQ}^{(1,3)} \right) \end{array} \right\}$
C_{Ld}		Same with $C_{LQ}^{(1,3)} \to C_{LL\tilde{d}\tilde{d}}$
C_{Lu}		Same with $C_{LQ}^{(1,3)} \to C_{LL\tilde{u}\tilde{u}}$
$C_{LeQu}^{(1,3)}$	27	$\left\{ \begin{array}{ccc} A_{0000}^0 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0000}^1 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0000}^2 \left(C_{L\tilde{e}Q\tilde{u}} \right) \\ A_{1000}^0 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{1000}^1 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{1000}^2 \left(C_{L\tilde{e}Q\tilde{u}} \right) \\ A_{0100}^0 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0100}^1 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0100}^2 \left(C_{L\tilde{e}Q\tilde{u}} \right) \\ A_{1100}^0 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{1100}^1 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{1100}^2 \left(C_{L\tilde{e}Q\tilde{u}} \right) \\ A_{0110}^0 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0110}^1 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0110}^2 \left(C_{L\tilde{e}Q\tilde{u}} \right) \\ A_{2200}^0 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{2200}^1 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{2200}^2 \left(C_{L\tilde{e}Q\tilde{u}} \right) \\ A_{0220}^0 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0220}^1 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0220}^2 \left(C_{L\tilde{e}Q\tilde{u}} \right) \\ A_{1220}^0 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{1220}^1 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{1220}^2 \left(C_{L\tilde{e}Q\tilde{u}} \right) \\ A_{0122}^0 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0122}^1 \left(C_{L\tilde{e}Q\tilde{u}} \right) & A_{0122}^2 \left(C_{L\tilde{e}Q\tilde{u}} \right) \end{array} \right\}$
C_{LedQ}		Same with $C_{L\tilde{e}Q\tilde{u}} \to C_{L\tilde{e}\tilde{d}Q}$ and $A_{bcde}^a \to A_{edcb}^a$

Table 11. Minimal sets of CP-odd flavor invariants for all the SMEFT dimension-six Wilson coefficients associated to operators quartic in fermion fields (continued in tables 12, 13). We recall that $X_u \equiv Y_u Y_u^\dagger$, and similarly for down quarks or electrons. We use the generalized traces introduced in eq. (4.8), as well as the compact notations in eqs. (4.9)–(4.10). We also defined for the leptons $A_b^a(C) \equiv \mathrm{Tr}_A \left(X_e^a, X_e^b, C \right)$, $B_b^a(C) \equiv \mathrm{Tr}_B \left(X_e^a, X_e^b, C \right)$ with $a, b = 1, 2$, $A_{bcde}^f(C) \equiv \mathrm{Tr}_A \left(X_e^f, X_u^b X_d^c X_u^d X_d^e, C \right)$, $A_f^{abcd}(C) \equiv \mathrm{Tr}_A \left(X_u^a X_d^b X_u^c X_d^d, \left(Y_e^\dagger Y_e \right)^f, C \right)$ and $B_{bcde}^f(C) \equiv \mathrm{Tr}_B \left(X_e^f, X_u^b X_d^c X_u^d X_d^e, C \right)$.

Wilson coefficient	Number of phases	Minimal set
$C_{QQ}^{(1,3)}$	18	$A_{1100}^{0000}(C_{QQQQ})$ $A_{1100}^{1000}(C_{QQQQ})$ $A_{1100}^{0100}(C_{QQQQ})$ $A_{2200}^{0000}(C_{QQQQ})$ $A_{1100}^{1100}(C_{QQQQ})$ $A_{2200}^{1000}(C_{QQQQ})$ $A_{2200}^{0100}(C_{QQQQ})$ $A_{1122}^{0000}(C_{QQQQ})$ $A_{2200}^{1100}(C_{QQQQ})$ $A_{2100}^{1200}(C_{QQQQ})$ $A_{1122}^{1000}(C_{QQQQ})$ $A_{1122}^{0100}(C_{QQQQ})$ $A_{1122}^{1100}(C_{QQQQ})$ $A_{2200}^{2200}(C_{QQQQ})$ $B_{1100}^{0000}(C_{QQQQ})$ $B_{2200}^{0000}(C_{QQQQ})$ $B_{1122}^{0000}(C_{QQQQ})$ $A_{1122}^{2200}(C_{QQQQ})$
C_{uu}	18	$A_{1100}^{0000}(C_{uu\tilde{u}\tilde{u}})$ $A_{1100}^{1000}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{1100}^{0100}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{2200}^{0000}(C_{uu\tilde{u}\tilde{u}})$ $A_{1100}^{1100}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{2200}^{0200}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{2200}^{0100}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{1122}^{0000}(C_{uu\tilde{u}\tilde{u}})$ $A_{2200}^{1100}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{1122}^{1000}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{1122}^{0100}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{0122}^{0100}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{2200}^{1200}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $B_{1100}^{0000}(C_{u\tilde{u}\tilde{u}u})$ $B_{1100}^{0100}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $B_{2100}^{0200}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $A_{1122}^{1200}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$ $B_{1200}^{1000}(C_{\tilde{u}\tilde{u}\tilde{u}\tilde{u}})$
C_{dd}	18	$A_{1100}^{0000}(C_{dd\tilde{d}\tilde{d}})$ $A_{1100}^{1000}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{2200}^{0000}(C_{dd\tilde{d}\tilde{d}})$ $A_{2000}^{1100}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{1100}^{0100}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{1100}^{1100}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{2200}^{1000}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{1122}^{0000}(C_{dd\tilde{d}\tilde{d}})$ $A_{2200}^{1100}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{1122}^{1000}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{1220}^{1100}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{2110}^{1200}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{0122}^{2100}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $A_{1220}^{2200}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $B_{1100}^{0000}(C_{dd\tilde{d}\tilde{d}})$ $B_{2100}^{0100}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $B_{1100}^{1000}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$ $B_{2000}^{1200}(C_{\tilde{d}\tilde{d}\tilde{d}\tilde{d}})$
$C_{Qu}^{(1,8)}$	36	$A_{0000}^{1100}(C_{QQuu})$ $A_{1100}^{0000}(C_{QQ\tilde{u}\tilde{u}})$ $A_{1100}^{1000}(C_{QQ\tilde{u}\tilde{u}})$ $A_{0100}^{1100}(C_{QQ\tilde{u}\tilde{u}})$ $A_{1100}^{1100}(C_{QQ\tilde{u}\tilde{u}})$ $A_{0110}^{1100}(C_{QQ\tilde{u}\tilde{u}})$ $A_{1000}^{1200}(C_{QQ\tilde{u}\tilde{u}})$ $A_{0000}^{2200}(C_{QQuu})$ $A_{2200}^{1100}(C_{QQ\tilde{u}\tilde{u}})$ $A_{0220}^{1100}(C_{QQ\tilde{u}\tilde{u}})$ $A_{0110}^{2200}(C_{QQ\tilde{u}\tilde{u}})$ $A_{1122}^{1100}(C_{QQ\tilde{u}\tilde{u}})$ $A_{1220}^{1200}(C_{QQ\tilde{u}\tilde{u}})$ $A_{1122}^{2200}(C_{QQ\tilde{u}\tilde{u}})$ $B_{0100}^{0000}(C_{QQ\tilde{u}\tilde{u}})$ $B_{1000}^{0000}(C_{QQ\tilde{u}\tilde{u}})$ $B_{0110}^{0000}(C_{QQ\tilde{u}\tilde{u}})$ $B_{0220}^{0000}(C_{QQ\tilde{u}\tilde{u}})$ $B_{1100}^{0000}(C_{QQ\tilde{u}\tilde{u}})$ $B_{0221}^{0000}(C_{QQ\tilde{u}\tilde{u}})$ $B_{1000}^{0100}(C_{QQ\tilde{u}\tilde{u}})$ $B_{1100}^{0100}(C_{QQ\tilde{u}\tilde{u}})$ $B_{2200}^{0100}(C_{QQ\tilde{u}\tilde{u}})$ $B_{2110}^{0100}(C_{QQ\tilde{u}\tilde{u}})$ $B_{2000}^{0200}(C_{QQ\tilde{u}\tilde{u}})$ $B_{2100}^{0200}(C_{QQ\tilde{u}\tilde{u}})$ $B_{2110}^{0200}(C_{QQ\tilde{u}\tilde{u}})$ $B_{0110}^{1000}(C_{QQ\tilde{u}\tilde{u}})$ $B_{0220}^{1000}(C_{QQ\tilde{u}\tilde{u}})$ $B_{0221}^{1000}(C_{QQ\tilde{u}\tilde{u}})$ $B_{1100}^{1100}(C_{QQ\tilde{u}\tilde{u}})$ $B_{2200}^{1100}(C_{QQ\tilde{u}\tilde{u}})$ $B_{2100}^{1200}(C_{QQ\tilde{u}\tilde{u}})$ $B_{2210}^{1200}(C_{QQ\tilde{u}\tilde{u}})$ $B_{1200}^{2100}(C_{QQ\tilde{u}\tilde{u}})$ $B_{0221}^{0110}(C_{QQ\tilde{u}\tilde{u}})$
$C_{Qd}^{(1,8)}$	36	$A_{0000}^{1100}(C_{QQdd})$ $A_{1100}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $A_{1100}^{1000}(C_{QQ\tilde{d}\tilde{d}})$ $A_{1000}^{1100}(C_{QQ\tilde{d}\tilde{d}})$ $A_{0000}^{2200}(C_{QQdd})$ $A_{1100}^{0100}(C_{QQ\tilde{d}\tilde{d}})$ $A_{2200}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $A_{1100}^{1100}(C_{QQ\tilde{d}\tilde{d}})$ $A_{2100}^{1100}(C_{QQ\tilde{d}\tilde{d}})$ $A_{0000}^{1122}(C_{QQdd})$ $A_{1122}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $A_{2200}^{1100}(C_{QQ\tilde{d}\tilde{d}})$ $A_{0220}^{1100}(C_{QQ\tilde{d}\tilde{d}})$ $A_{1122}^{1000}(C_{QQ\tilde{d}\tilde{d}})$ $A_{1100}^{1100}(C_{QQ\tilde{d}\tilde{d}})$ $A_{0122}^{2100}(C_{QQ\tilde{d}\tilde{d}})$ $B_{0100}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{1000}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{0110}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{0220}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{1100}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{0221}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{2200}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{2210}^{0000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{1000}^{0100}(C_{QQ\tilde{d}\tilde{d}})$ $B_{0120}^{0100}(C_{QQ\tilde{d}\tilde{d}})$ $B_{1100}^{0100}(C_{QQ\tilde{d}\tilde{d}})$ $B_{2210}^{0100}(C_{QQ\tilde{d}\tilde{d}})$ $B_{0110}^{1000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{0220}^{1000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{0221}^{1000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{1200}^{1000}(C_{QQ\tilde{d}\tilde{d}})$ $B_{2200}^{1100}(C_{QQ\tilde{d}\tilde{d}})$ $B_{2210}^{1100}(C_{QQ\tilde{d}\tilde{d}})$ $B_{2100}^{1200}(C_{QQ\tilde{d}\tilde{d}})$ $B_{2211}^{2100}(C_{QQ\tilde{d}\tilde{d}})$

Table 12. Continuation of table 11.

Wilson coefficient	Number of phases	Minimal set
$C_{ud}^{(1,8)}$	36	$A_{0000}^{1100}(C_{\tilde u\tilde u dd})$ $A_{1100}^{0000}(C_{uu\tilde d\tilde d})$ $A_{1100}^{1000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{1000}^{1100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{0000}^{2200}(C_{\tilde u\tilde u dd})$ $A_{1100}^{0100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{2200}^{0000}(C_{uu\tilde d\tilde d})$ $A_{1100}^{1100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{0110}^{1100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{2200}^{1000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{2100}^{1100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{0000}^{1122}(C_{\tilde u\tilde u dd})$ $A_{2200}^{0100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{1122}^{0000}(C_{uu\tilde d\tilde d})$ $A_{2200}^{1100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{1122}^{1000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{1122}^{0100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $A_{1122}^{1100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{0100}^{0000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{1000}^{0000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{0110}^{0000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{1100}^{0000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{0221}^{0000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{2200}^{0000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{1000}^{0100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{0110}^{0100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{2110}^{0100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{2000}^{0200}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{2110}^{0200}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{0110}^{1000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{0221}^{1000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{1200}^{1000}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{2200}^{1100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{2211}^{1100}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{2100}^{1200}(C_{\tilde u\tilde u\tilde d\tilde d})$ $B_{1200}^{2100}(C_{\tilde u\tilde u\tilde d\tilde d})$
$C_{QuQd}^{(1,8)}$	81	$A_{0000}^{0000}(C_{Q\tilde u Q\tilde d})$ $A_{1000}^{0000}(C_{Q\tilde u Q\tilde d})$ $A_{0000}^{1000}(C_{Q\tilde u Q\tilde d})$ $A_{1000}^{1000}(C_{Q\tilde u Q\tilde d})$ $A_{0100}^{0000}(C_{Q\tilde u Q\tilde d})$ $A_{0000}^{0100}(C_{Q\tilde u Q\tilde d})$ $A_{1100}^{0000}(C_{Q\tilde u Q\tilde d})$ $A_{0110}^{0000}(C_{Q\tilde u Q\tilde d})$ $A_{1000}^{0100}(C_{Q\tilde u Q\tilde d})$ $A_{0100}^{1000}(C_{Q\tilde u Q\tilde d})$ $A_{0000}^{1100}(C_{Q\tilde u Q\tilde d})$ $A_{0000}^{0110}(C_{Q\tilde u Q\tilde d})$ $A_{1100}^{1000}(C_{Q\tilde u Q\tilde d})$ $A_{0110}^{1000}(C_{Q\tilde u Q\tilde d})$ $A_{1000}^{1100}(C_{Q\tilde u Q\tilde d})$ $A_{0100}^{0100}(C_{Q\tilde u Q\tilde d})$ $A_{1100}^{0100}(C_{Q\tilde u Q\tilde d})$ $A_{0110}^{0100}(C_{Q\tilde u Q\tilde d})$ $A_{0100}^{0110}(C_{Q\tilde u Q\tilde d})$ $A_{2200}^{0000}(C_{Q\tilde u Q\tilde d})$ $A_{0220}^{0000}(C_{Q\tilde u Q\tilde d})$ $A_{2000}^{0200}(C_{Q\tilde u Q\tilde d})$ $A_{1100}^{1100}(C_{Q\tilde u Q\tilde d})$ $A_{0110}^{1100}(C_{Q\tilde u Q\tilde d})$ $A_{0200}^{2000}(C_{Q\tilde u Q\tilde d})$ $A_{0100}^{2100}(C_{Q\tilde u Q\tilde d})$ $A_{1100}^{0110}(C_{Q\tilde u Q\tilde d})$ $A_{0110}^{0110}(C_{Q\tilde u Q\tilde d})$ $A_{1000}^{0210}(C_{Q\tilde u Q\tilde d})$ $A_{1220}^{0000}(C_{Q\tilde u Q\tilde d})$ $A_{2000}^{1200}(C_{Q\tilde u Q\tilde d})$ $A_{0122}^{0000}(C_{Q\tilde u Q\tilde d})$ $A_{1220}^{0100}(C_{Q\tilde u Q\tilde d})$ $A_{0122}^{1000}(C_{Q\tilde u Q\tilde d})$ $A_{2200}^{1100}(C_{Q\tilde u Q\tilde d})$ $A_{0220}^{1100}(C_{Q\tilde u Q\tilde d})$ $A_{2100}^{1200}(C_{Q\tilde u Q\tilde d})$ $A_{1200}^{2100}(C_{Q\tilde u Q\tilde d})$ $A_{0210}^{2100}(C_{Q\tilde u Q\tilde d})$ $A_{0110}^{2200}(C_{Q\tilde u Q\tilde d})$ $A_{2200}^{0110}(C_{Q\tilde u Q\tilde d})$ $A_{0220}^{0110}(C_{Q\tilde u Q\tilde d})$ $A_{2000}^{0112}(C_{Q\tilde u Q\tilde d})$ $A_{1220}^{1100}(C_{Q\tilde u Q\tilde d})$ $A_{0112}^{2100}(C_{Q\tilde u Q\tilde d})$ $A_{1220}^{1200}(C_{Q\tilde u Q\tilde d})$ $A_{2200}^{2200}(C_{Q\tilde u Q\tilde d})$ $A_{1122}^{0110}(C_{Q\tilde u Q\tilde d})$ $A_{2100}^{0122}(C_{Q\tilde u Q\tilde d})$ $A_{0220}^{0220}(C_{Q\tilde u Q\tilde d})$ $B_{0000}^{0000}(C_{Q\tilde u Q\tilde d})$ $B_{0100}^{0000}(C_{Q\tilde u Q\tilde d})$ $B_{1000}^{0000}(C_{Q\tilde u Q\tilde d})$ $B_{1100}^{0000}(C_{Q\tilde u Q\tilde d})$ $B_{2200}^{0000}(C_{Q\tilde u Q\tilde d})$ $B_{0110}^{0000}(C_{Q\tilde u Q\tilde d})$ $B_{0122}^{0000}(C_{Q\tilde u Q\tilde d})$ $B_{0220}^{0000}(C_{Q\tilde u Q\tilde d})$ $B_{0000}^{0100}(C_{Q\tilde u Q\tilde d})$ $B_{1000}^{0100}(C_{Q\tilde u Q\tilde d})$ $B_{1100}^{0100}(C_{Q\tilde u Q\tilde d})$ $B_{2100}^{0100}(C_{Q\tilde u Q\tilde d})$ $B_{0120}^{0100}(C_{Q\tilde u Q\tilde d})$ $B_{1220}^{0100}(C_{Q\tilde u Q\tilde d})$ $B_{1120}^{0200}(C_{Q\tilde u Q\tilde d})$ $B_{0000}^{1000}(C_{Q\tilde u Q\tilde d})$ $B_{0100}^{1000}(C_{Q\tilde u Q\tilde d})$ $B_{1200}^{1000}(C_{Q\tilde u Q\tilde d})$ $B_{0110}^{1000}(C_{Q\tilde u Q\tilde d})$ $B_{0122}^{1000}(C_{Q\tilde u Q\tilde d})$ $B_{0210}^{1000}(C_{Q\tilde u Q\tilde d})$ $B_{0000}^{1100}(C_{Q\tilde u Q\tilde d})$ $B_{1100}^{1100}(C_{Q\tilde u Q\tilde d})$ $B_{2200}^{1100}(C_{Q\tilde u Q\tilde d})$ $B_{0110}^{1100}(C_{Q\tilde u Q\tilde d})$ $B_{0220}^{1100}(C_{Q\tilde u Q\tilde d})$ $B_{1122}^{1100}(C_{Q\tilde u Q\tilde d})$ $B_{2100}^{1200}(C_{Q\tilde u Q\tilde d})$ $B_{0122}^{2100}(C_{Q\tilde u Q\tilde d})$ $B_{0000}^{2200}(C_{Q\tilde u Q\tilde d})$ $A_{1122}^{2200}(C_{Q\tilde u Q\tilde d})$

Table 13. Continuation of tables 11 and 12.

F Invariants featuring θ_{QCD}

In the main text of the paper, we focused on quantities which matter for perturbative computations in SMEFT. However, this left out an important contribution to CPV in the SM, the θ-parameter of QCD, associated to the following topological term,

$$\mathcal{L}_{\text{QCD}} \supset -\theta_{\text{QCD}} \frac{g_s^2}{16\pi^2} \text{Tr}\left(G\tilde{G}\right). \tag{F.1}$$

θ_{QCD} has the following flavor charges (and no lepton-type charge),

	$\text{SU}(3)_{Q_L}$	$\text{U}(1)_{Q_L}$	$\text{SU}(3)_{u_R}$	$\text{U}(1)_{u_R}$	$\text{SU}(3)_{d_R}$	$\text{U}(1)_{d_R}$
$e^{i\theta_{\text{QCD}}}$	**1**	6	**1**	-3	**1**	-3

These charges allow us to build the usual flavor-invariant, physical $\bar{\theta}$-angle, defined as follows,

$$e^{-i\theta_{\text{QCD}}} \det Y_u \det Y_d = |\det (Y_u Y_d)| e^{-i\left[\theta_{\text{QCD}} - \arg \det(Y_u Y_d)\right]} = |\det (Y_u Y_d)| e^{-i\bar{\theta}}. \tag{F.2}$$

As θ_{QCD} provides a dimension-four flavor-charged quantity, one can wonder whether its presence makes new SMEFT coefficients primary, which would mean that new invariants featuring explicitly θ_{QCD} should be included in the minimal sets. The answer is however negative: the secondary sources of CPV from the dimension-six Wilson coefficients are all charged under unbroken vector-like flavor symmetries of the dimension-four Lagrangian, under which θ_{QCD} is neutral. Indeed, as the anomalous angle of a vector-like gauge theory, it only shifts under chiral transformations.

Nevertheless, some SMEFT coefficients can be arranged with θ_{QCD} to form flavor-invariants (albeit redundant in terms of primary parameter counting at dimension-six), which may yield a more natural description of some non-perturbative contributions of the strong interactions to CP-odd observables.[23] Those invariants would not have the single trace structure which we used to build our sets of invariants, since δ_m^n is $\text{U}(3)^5$-invariant, while θ_{QCD} is charged under some abelian parts of the flavor group. Therefore, it will rather offset the abelian charges of determinant-like $\text{SU}(3)^5$-invariants. For instance, for the operator C_{QuQd}, we can form

$$\text{Im}\left(e^{-i\theta_{\text{QCD}}} \epsilon^{ABC} \epsilon^{abc} \epsilon^{DEF} \epsilon^{def} Y_{u,Aa} Y_{u,Bb} C_{QuQd,CcDd} Y_{d,Ee} Y_{d,Ff}\right) =$$
$$\Big|_{\text{up basis}} 4 y_b y_s y_t y_c \text{Im} C_{QuQd,1111} + \dots .$$

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](#)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited. SCOAP³ supports the goals of the International Year of Basic Sciences for Sustainable Development.

---

[23] In the perturbative phase of QCD, the magnitude of such invariants is expected to be suppressed by an additional non-perturbative factor $e^{-8\pi^2/g_s^2}$. For low-energy observables, such as the EDMs of the neutron [75] and of the electron [76, 77], no further suppression would be needed.

References

[1] Particle Data collaboration, *Review of Particle Physics*, *Prog. Theor. Exp. Phys.* **2020** (2020) 083C01 [inSPIRE].

[2] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531 [inSPIRE].

[3] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652 [inSPIRE].

[4] G. Branco, L. Lavoura and J. Silva, *CP Violation*, in *International series of monographs on physics* **103**, Clarendon Press, Oxford, U.K. (1999).

[5] H. Quinn, *Mistakes on the way to a good idea: Peccei-Quinn Symmetry. A tribute to Roberto Peccei*, in *Hertz Lecture*, Deutsches Elektronen-Synchrotron, Hamburg, Germany (2019) and online at https://webcast.desy.de/category/video/Mistakes-on-the-way-to-a-good-idea-Peccei-Quinn-Symmetry-A-tribute-to-Roberto-Peccei/f45c14650f45a617ba41a6b0b1368049/43.

[6] N. Arkani-Hamed, A.G. Cohen and H. Georgi, *Electroweak symmetry breaking from dimensional deconstruction*, *Phys. Lett. B* **513** (2001) 232 [hep-ph/0105239] [inSPIRE].

[7] N. Arkani-Hamed, A.G. Cohen, T. Gregoire and J.G. Wacker, *Phenomenology of electroweak symmetry breaking from theory space*, *JHEP* **08** (2002) 020 [hep-ph/0202089] [inSPIRE].

[8] N. Arkani-Hamed, A.G. Cohen, E. Katz and A.E. Nelson, *The Littlest Higgs*, *JHEP* **07** (2002) 034 [hep-ph/0206021] [inSPIRE].

[9] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation*, *Phys. Rev. Lett.* **55** (1985) 1039 [inSPIRE].

[10] C. Jarlskog, *A Basis Independent Formulation of the Connection Between Quark Mass Matrices, CP-violation and Experiment*, *Z. Phys. C* **29** (1985) 491 [inSPIRE].

[11] J. Bernabeu, G.C. Branco and M. Gronau, *CP Restrictions on Quark Mass Matrices*, *Phys. Lett. B* **169** (1986) 243 [inSPIRE].

[12] A. Mendez and A. Pomarol, *Signals of CP-violation in the Higgs sector*, *Phys. Lett. B* **272** (1991) 313 [inSPIRE].

[13] F.J. Botella and J.P. Silva, *Jarlskog-like invariants for theories with scalars and fermions*, *Phys. Rev. D* **51** (1995) 3870 [hep-ph/9411288] [inSPIRE].

[14] L. Lavoura and J.P. Silva, *Fundamental CP-violating quantities in a* SU(2) × U(1) *model with many Higgs doublets*, *Phys. Rev. D* **50** (1994) 4619 [hep-ph/9404276] [inSPIRE].

[15] J.F. Gunion and H.E. Haber, *Conditions for CP-violation in the general two-Higgs-doublet model*, *Phys. Rev. D* **72** (2005) 095002 [hep-ph/0506227] [inSPIRE].

[16] A. Trautner, *Systematic construction of basis invariants in the* 2*HDM*, *JHEP* **05** (2019) 208 [arXiv:1812.02614] [inSPIRE].

[17] M.P. Bento, *The invariant space of multi-Higgs doublet models*, *JHEP* **05** (2021) 146 [arXiv:2102.13120] [inSPIRE].

[18] G.C. Branco and V.A. Kostelecky, *CP Violation in Supergravity Models*, *Phys. Rev. D* **39** (1989) 2075 [inSPIRE].

[19] O. Lebedev, *CP violating invariants in supersymmetry*, *Phys. Rev. D* **67** (2003) 015013 [hep-ph/0209023] [inSPIRE].

[20] F.J. Botella, M. Nebot and O. Vives, *Invariant approach to flavor-dependent CP-violating phases in the MSSM*, *JHEP* **01** (2006) 106 [hep-ph/0407349] [INSPIRE].

[21] F.J. Botella and L.-L. Chau, *Anticipating the Higher Generations of Quarks from Rephasing Invariance of the Mixing Matrix*, *Phys. Lett. B* **168** (1986) 97 [INSPIRE].

[22] M. Gronau, A. Kfir and R. Loewy, *Basis Independent Tests of CP Violation in Fermion Mass Matrices*, *Phys. Rev. Lett.* **56** (1986) 1538 [INSPIRE].

[23] K. Bering, *Generalized Jarlskog Invariants, Mass Degeneracies and Echelon Crosses*, arXiv:2104.10585 [INSPIRE].

[24] G.C. Branco, L. Lavoura and M.N. Rebelo, *Majorana Neutrinos and CP Violation in the Leptonic Sector*, *Phys. Lett. B* **180** (1986) 264 [INSPIRE].

[25] G.C. Branco, M.N. Rebelo and J.I. Silva-Marcos, *Degenerate and quasidegenerate Majorana neutrinos*, *Phys. Rev. Lett.* **82** (1999) 683 [hep-ph/9810328] [INSPIRE].

[26] H.K. Dreiner, J.S. Kim, O. Lebedev and M. Thormeier, *Supersymmetric Jarlskog invariants: The Neutrino sector*, *Phys. Rev. D* **76** (2007) 015006 [hep-ph/0703074] [INSPIRE].

[27] E.E. Jenkins and A.V. Manohar, *Algebraic Structure of Lepton and Quark Flavor Invariants and CP-violation*, *JHEP* **10** (2009) 094 [arXiv:0907.4763] [INSPIRE].

[28] B. Yu and S. Zhou, *The number of sufficient and necessary conditions for CP conservation with Majorana neutrinos: three or four?*, *Phys. Lett. B* **800** (2020) 135085 [arXiv:1908.09306] [INSPIRE].

[29] B. Yu and S. Zhou, *Sufficient and Necessary Conditions for CP Conservation in the Case of Degenerate Majorana Neutrino Masses*, *Phys. Rev. D* **103** (2021) 035017 [arXiv:2009.12347] [INSPIRE].

[30] J. Lu, *A remark on the critical value of the lightest neutrino mass in the construction of sufficient and necessary conditions for CP conservation with Majorana neutrinos*, *Phys. Lett. B* **821** (2021) 136635 [INSPIRE].

[31] Y. Wang, B. Yu and S. Zhou, *Flavor invariants and renormalization-group equations in the leptonic sector with massive Majorana neutrinos*, *JHEP* **09** (2021) 053 [arXiv:2107.06274] [INSPIRE].

[32] B. Yu and S. Zhou, *Hilbert series for leptonic flavor invariants in the minimal seesaw model*, *JHEP* **10** (2021) 017 [arXiv:2107.11928] [INSPIRE].

[33] J.A. Aguilar-Saavedra, *Measure of the size of CP-violation in extended models*, *J. Phys. G* **24** (1998) L31 [hep-ph/9703461] [INSPIRE].

[34] L. Di Luzio, R. Gröber and P. Paradisi, *Hunting for CP-violating axionlike particle interactions*, *Phys. Rev. D* **104** (2021) 095027 [arXiv:2010.13760] [INSPIRE].

[35] A. Valenti and L. Vecchi, *The CKM phase and $\overline{\theta}$ in Nelson-Barr models*, *JHEP* **07** (2021) 203 [arXiv:2105.09122] [INSPIRE].

[36] W. Buchmüller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621 [INSPIRE].

[37] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [arXiv:1008.4884] [INSPIRE].

[38] 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** (2021) 015026 [arXiv:2005.00008] [INSPIRE].

[39] C.W. Murphy, *Dimension-8 operators in the Standard Model Eective Field Theory*, *JHEP* **10** (2020) 174 [arXiv:2005.00059] [INSPIRE].

[40] R. Alonso, E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators. Part III. Gauge Coupling Dependence and Phenomenology*, *JHEP* **04** (2014) 159 [arXiv:1312.2014] [INSPIRE].

[41] G.F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, *The Strongly-Interacting Light Higgs*, *JHEP* **06** (2007) 045 [hep-ph/0703164] [INSPIRE].

[42] E. Fuchs, M. Losada, Y. Nir and Y. Viernik, *CP violation from τ, t and b dimension-6 Yukawa couplings — interplay of baryogenesis, EDM and Higgs physics*, *JHEP* **05** (2020) 056 [arXiv:2003.00099] [INSPIRE].

[43] I.P. Ivanov and J.P. Silva, *CP-conserving multi-Higgs model with irremovable complex coefficients*, *Phys. Rev. D* **93** (2016) 095014 [arXiv:1512.09276] [INSPIRE].

[44] A. Trautner, *CP and other Symmetries of Symmetries*, Ph.D. Thesis, Technical University of Munich, Munich, Germany (2016) [arXiv:1608.05240] [INSPIRE].

[45] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **53** (1984) 1802 [INSPIRE].

[46] I. Dunietz, O.W. Greenberg and D.-d. Wu, *A Priori Definition of Maximal CP-violation*, *Phys. Rev. Lett.* **55** (1985) 2935 [INSPIRE].

[47] D.-d. Wu, *The Rephasing Invariants and CP*, *Phys. Rev. D* **33** (1986) 860 [INSPIRE].

[48] S.M. Barr and A. Zee, *Electric Dipole Moment of the Electron and of the Neutron*, *Phys. Rev. Lett.* **65** (1990) 21 [*Erratum ibid.* **65** (1990) 2920] [INSPIRE].

[49] J. Brod, U. Haisch and J. Zupan, *Constraints on CP-violating Higgs couplings to the third generation*, *JHEP* **11** (2013) 180 [arXiv:1310.1385] [INSPIRE].

[50] G. Panico, A. Pomarol and M. Riembau, *EFT approach to the electron Electric Dipole Moment at the two-loop level*, *JHEP* **04** (2019) 090 [arXiv:1810.09413] [INSPIRE].

[51] M. Pospelov and A. Ritz, *CKM benchmarks for electron electric dipole moment experiments*, *Phys. Rev. D* **89** (2014) 056006 [arXiv:1311.5537] [INSPIRE].

[52] Y. Yamaguchi and N. Yamanaka, *Quark level and hadronic contributions to the electric dipole moment of charged leptons in the standard model*, *Phys. Rev. D* **103** (2021) 013001 [arXiv:2006.00281] [INSPIRE].

[53] Y. Yamaguchi and N. Yamanaka, *Large long-distance contributions to the electric dipole moments of charged leptons in the standard model*, *Phys. Rev. Lett.* **125** (2020) 241802 [arXiv:2003.08195] [INSPIRE].

[54] Y. Ema, T. Gao and M. Pospelov, *Standard Model prediction for paramagnetic EDMs*, arXiv:2202.10524 [INSPIRE].

[55] O.W. Greenberg, *Rephase Invariant Formulation of CP-violation in the Kobayashi-Maskawa Framework*, *Phys. Rev. D* **32** (1985) 1841 [INSPIRE].

[56] J. Alonso-González, L. Merlo and S. Pokorski, *A new bound on CP-violation in the τ lepton Yukawa coupling and electroweak baryogenesis*, *JHEP* **06** (2021) 166 [arXiv:2103.16569] [INSPIRE].

[57] J. Aebischer, C. Bobeth, A.J. Buras, J.-M. Gérard and D.M. Straub, *Master formula for ε'/ε beyond the Standard Model*, *Phys. Lett. B* **792** (2019) 465 [arXiv:1807.02520] [INSPIRE].

[58] J. Aebischer, C. Bobeth, A.J. Buras and D.M. Straub, *Anatomy of ε'/ε beyond the standard model*, *Eur. Phys. J. C* **79** (2019) 219 [`arXiv:1808.00466`] [ɪɴSPIRE].

[59] J. Aebischer, C. Bobeth and A.J. Buras, *ε'/ε in the Standard Model at the Dawn of the 2020s*, *Eur. Phys. J. C* **80** (2020) 705 [`arXiv:2005.05978`] [ɪɴSPIRE].

[60] C. Degrande and J. Touchèque, *A reduced basis for CP-violation in SMEFT at colliders and its application to diboson production*, *JHEP* **04** (2022) 032 [`arXiv:2110.02993`] [ɪɴSPIRE].

[61] Q. Bonnefoy, E. Gendy, C. Grojean and J.T. Ruderman, in preparation.

[62] I.B. Khriplovich and M.E. Pospelov, *Magnetic quadrupole moment of W boson in Kobayashi-Maskawa model*, *Nucl. Phys. B* **420** (1994) 505 [`hep-ph/9310342`] [ɪɴSPIRE].

[63] M.E. Pospelov, *CP odd effective gluonic Lagrangian in the Kobayashi-Maskawa model*, *Phys. Lett. B* **328** (1994) 441 [`hep-ph/9402317`] [ɪɴSPIRE].

[64] C. Smith and S. Touati, *Electric dipole moments with and beyond flavor invariants*, *Nucl. Phys. B* **924** (2017) 417 [`arXiv:1707.06805`] [ɪɴSPIRE].

[65] J. Kley, T. Theil, E. Venturini and A. Weiler, *Electric dipole moments at one-loop in the dimension-6 SMEFT*, `arXiv:2109.15085` [ɪɴSPIRE].

[66] B. Feng, A. Hanany and Y.-H. He, *Counting gauge invariants: The Plethystic program*, *JHEP* **03** (2007) 090 [`hep-th/0701063`] [ɪɴSPIRE].

[67] S. Benvenuti, B. Feng, A. Hanany and Y.-H. He, *Counting BPS Operators in Gauge Theories: Quivers, Syzygies and Plethystics*, *JHEP* **11** (2007) 050 [`hep-th/0608050`] [ɪɴSPIRE].

[68] C. Romelsberger, *Counting chiral primaries in $N = 1$, $d = 4$ superconformal field theories*, *Nucl. Phys. B* **747** (2006) 329 [`hep-th/0510060`] [ɪɴSPIRE].

[69] P. Pouliot, *Molien function for duality*, *JHEP* **01** (1999) 021 [`hep-th/9812015`] [ɪɴSPIRE].

[70] A. Trautner, *On the systematic construction of basis invariants*, *J. Phys. Conf. Ser.* **1586** (2020) 012005 [`arXiv:2002.12244`] [ɪɴSPIRE].

[71] B. Henning, X. Lu, T. Melia and H. Murayama, *Operator bases, S-matrices, and their partition functions*, *JHEP* **10** (2017) 199 [`arXiv:1706.08520`] [ɪɴSPIRE].

[72] J. Lu, *Comment on "flavor invariants and renormalization-group equations in the leptonic sector with massive Majorana neutrinos"*, *JHEP* **02** (2022) 135 [`arXiv:2110.08210`] [ɪɴSPIRE].

[73] Y. Wang, B. Yu and S. Zhou, *Reply to «Comment on "Flavor invariants and renormalization-group equations in the leptonic sector with massive Majorana neutrinos"»*, `arXiv:2110.13865` [ɪɴSPIRE].

[74] J. Lu, *(Reply to)^2 «Comment on "Flavor invariants and renormalization-group equations in the leptonic sector with massive Majorana neutrinos"»*, `arXiv:2111.02729` [ɪɴSPIRE].

[75] M. Pospelov and A. Ritz, *Theta vacua, QCD sum rules, and the neutron electric dipole moment*, *Nucl. Phys. B* **573** (2000) 177 [`hep-ph/9908508`] [ɪɴSPIRE].

[76] K. Choi and J.-y. Hong, *Electron electric dipole moment and Theta (QCD)*, *Phys. Lett. B* **259** (1991) 340 [ɪɴSPIRE].

[77] D. Ghosh and R. Sato, *Lepton Electric Dipole Moment and Strong CP-violation*, *Phys. Lett. B* **777** (2018) 335 [`arXiv:1709.05866`] [ɪɴSPIRE].