Methodology for theory uncertainties in the standard model effective field theory

Michael Trott◉

Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen, Denmark

A process specific methodology is defined to systematically assign theoretical uncertainties in the Standard Model effective field theory when performing leading-order global fits. The method outlined also minimizes the computational and theoretical burden to systematically advance such analyses to dimension eight.

I. INTRODUCTION

The Standard Model (SM) effective field theory (SMEFT) is based on the infrared assumptions that physics beyond the SM is present at scales $\Lambda > \bar{v}_T \equiv \sqrt{2\langle H^\dagger H \rangle}$, with no light hidden states, and a $SU(2)_L$ scalar doublet (H) with hypercharge $\mathsf{y}_h = 1/2$ is in the effective field theory (EFT). An expansion in the ratio of scales $q^2/\Lambda^2 < 1$, with q^2 a kinematic invariant associated with experimental measurements in the domain of validity of the EFT, defines the SMEFT Lagrangian

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \mathcal{L}^{(5)} + \mathcal{L}^{(6)} + \mathcal{L}^{(7)} + \cdots,$$

$$\mathcal{L}^{(d)} = \sum_i \frac{C_i^{(d)}}{\Lambda^{d-4}} \mathcal{Q}_i^{(d)} \quad \text{for } d > 4. \tag{1}$$

Operators $(\mathcal{Q}_i^{(d)})$ define SMEFT corrections to the SM predictions, and carry a mass-dimension d superscript. Here $\tilde{C}_i^{(6)} = \bar{v}_T^2 C_i^{(6)}/\Lambda^2$ and $\tilde{C}_i^{(8)} = \bar{v}_T^4 C_i^{(8)}/\Lambda^4$. The operators multiply Wilson coefficients $C_i^{(d)}$, which take on specific values as a result of the Taylor expanded effects of physics beyond the SM. We treat Wilson coefficients as free constraint parameters. The sum over i, runs over the operators in a particular operator basis. We use the Warsaw basis [1,2] for $\mathcal{L}^{(6)}$. For $d > 6$, we use Refs. [3–6], which defines the geoSMEFT formulation of this theory to all orders in $\sqrt{2\langle H^\dagger H \rangle}/\Lambda$ for n-point functions, with $n \leq 3$.

Using SMEFT, ATLAS and CMS have started to perform global analysis of LHC data. As approximations are used in the theory predictions, it is required to define a theoretical error, to avoid misinterpreting experimental results in global SMEFT studies. In this paper we define such a methodology.

II. MISSING HIGHER-ORDER TERMS

When a prediction is made in the SMEFT at leading order (LO), subleading terms are neglected. The expansions present are:
(a) the *loop expansion* in $g_{\text{SM}}^2/16\pi^2$, where $g_{\text{SM}} \subset [g_1, g_2, g_3, \lambda, y_\psi]$ and y_ψ is the Yukawa coupling for the fermion species ψ,
(b) the *vacuum expectation value expansion* in $\sqrt{2\langle H^\dagger H \rangle}/\Lambda$; this expansion is relevant when SM kinematics is present in a process due to a resonant SM state fixing $q^2 \simeq \bar{v}_T^2$, and
(c) the *derivative expansion* in q^2/Λ^2.
Collectively, we refer to the $\bar{v}_T/\Lambda$ and q^2/Λ^2 expansion as the SMEFT operator expansion.[1] LO in the SMEFT means considering $\mathcal{L}^{(6)}$ perturbations to the SM predictions, and roughly ~ 30 new parameters impact Higgs, electroweak, and top-quark processes [7]. Global SMEFT fits constrain these parameters. Uncertainties due to the truncation of the EFT expansion and missing perturbative corrections, should be assigned in this effort, following the usual methodology of EFT studies.[2]

At LO, a dimensionless SM amplitude includes a perturbation due to $\mathcal{L}^{(6)}$. We illustrate the methodology with a pole observable where

---

---

[1]The expansion in $\bar{v}_T/\Lambda$ is relevant if some SM particles go on shell in an observable. We refer to these as pole observables. The q^2/Λ^2 expansion is relevant when considering nonresonant regions of phase space. We refer to these as tail observables.
[2]See for example Refs. [8–16].

$$\mathcal{A} = \mathcal{A}_{\text{SM}} + \tilde{C}_i^{(6)}\, a_i + \cdots, \qquad (2)$$

where a_i is a numerical coefficient that is process dependent. Sums over repeated indices are implied. The expression is not exact [it should be understood to have arbitrary and unfixed corrections of $\mathcal{O}(1/\Lambda^4)$] until the SMEFT corrections for this process are defined to $\mathcal{O}(1/\Lambda^4)$. We return to this point in Sec. V. Quantum corrections cannot be forbidden, as the SMEFT is built of the SM fields. As higher-order terms in the loop and operator expansion are unknown it is important to access the impact of neglecting these terms in LO SMEFT analyses of the data. The next order in the missing terms in both expansions are

$$\mathcal{A} = \mathcal{A}_{\text{SM}} + \tilde{C}_i^{(6)}\, a_i + \tilde{C}_j^{(6)}\tilde{C}_k^{(6)}\, b_{jk} + \tilde{C}_l^{(8)}\, c_l$$
$$+ \frac{1}{16\pi^2}\left[d_m\tilde{C}_m^{(6)} + e_n\tilde{C}_n^{(6)}\, \log\left(\frac{\mu^2}{\Lambda^2}\right)\right] + \cdots. \quad (3)$$

The expression is not exact and $a_i, b_{jk}, c_l, d_m, e_n$, are process-dependent numerical coefficients. In each case, the indicies $i, j, k, \cdots$ run over a subset of the full set of operators in $\mathcal{L}^{(6)}$ and/or $\mathcal{L}^{(8)}$.

Squaring this expression, integrating over phase space with relevant experimental cuts, one finds a cross section that is (schematically)[3]

$$\sigma_{1/\Lambda^4} = \sigma_{\text{SM}} + \sum A_i\tilde{C}_i^{(6)} + \sum B_{jk}\tilde{C}_j^{(6)}\tilde{C}_k^{(6)}$$
$$+ \sum D_l\tilde{C}_l^{(8)} + \cdots. \qquad (4)$$

We assume that a LO simulation, including relevant experimental cuts and acceptance corrections (i.e., A_i) are known for all i in $\mathcal{L}^{(6)}$ using a SMEFTsim [7,17] simulation. A class of terms in B_{jk} with operators residing on different vertices, can also be known (see Ref. [17]). A SMEFTsim simulation does not produce the effects of canonically normalizing the SMEFT to $\mathcal{O}(1/\Lambda^4)$ and $\mathcal{L}^{(8)}$ operators.

To estimate the effect of neglecting higher-order terms for σ, knowing the prediction of all terms to subleading order and varying the unknown parameters numerically is a well defined and straightforward procedure. As such, we focus on how to directly extract subleading terms from the results of a LO simulation.

Higher-order terms in the SMEFT expansion that have common kinematic populations of phase space as operators at $\mathcal{L}^{(6)}$ (already in a LO simulation) receive common numerical corrections due to Monte Carlo event generation, and phase space/acceptance cuts. This means that for classes of $\mathcal{L}^{(8)}$, $(\mathcal{L}^{(6)})^2$ and $\mathcal{L}^{(6)}/16\pi^2$ terms, these terms

---

[3]The number of terms appearing in the sum over SMEFT operators depends on the perturbative order of σ_{SM}.

can be produced with appropriate rescaling from LO simulation results, without the need of redundant/costly event generation. The methodology we lay out defines how to leverage this fact in practice.

III. MISSING PERTURBATIVE CORRECTIONS

Consider a partonic scattering process $X \to Y$ defining a tree-level SMEFT amplitude $\mathcal{A}_{X\to Y}(C_i/\Lambda^2)$ that interferes with the SM amplitude $\mathcal{A}_{X\to Y}^{\text{SM}}$, the latter of which can be of any perturbative order. By definition

$$\sum A_i\tilde{C}_i^{(6)} \propto \mathcal{A}_{X\to Y}^{\text{SM}} \times \mathcal{A}_{X\to Y}(\tilde{C}_i^{(6)}), \qquad (5)$$

up to suppressed phase space integrals with cuts.

We estimate perturbative uncertainties on the EFT parameters $C_i^{(6)}$ using

$$\left[a_i\tilde{C}_i^{(6)} + \frac{d_m\tilde{C}_m^{(6)}}{16\pi^2} + \frac{e_{\text{in}}\tilde{C}_n^{(6)}}{16\pi^2}\log\frac{\mu^2}{\Lambda^2}\right]. \qquad (6)$$

The d_m terms can be determined with a dedicated one-loop calculation and are, in general, unknown. The coefficients e_{in} of the log-enhanced terms, which generally give the largest contribution to the perturbative correction, are known. The subscript has explicit dependence on i, the coefficient appearing in the LO simulation, as these corrections have associated divergences. Such divergences are canceled by the renormalization group evolution (RGE) of the SMEFT was reported in Refs. [18–20] and must feed in via a tree-level operator dependence. SMEFTsim has bare SMEFT Lagrangian $\mathcal{L}^{(6),0}$ coded which is related to the renormalized Lagrangian [denoted with an (r) superscript] where the SMEFT RGE counterterms $Z_{i,j}$ are introduced via

$$\mathcal{L}^{(6),r} = Z_{\text{SM}}Z_{i,n}\frac{C_i^{(6)}}{\Lambda^2}\mathcal{Q}_n^{(6),r}. \qquad (7)$$

This rescaling by $Z_{i,n}$ is *transparent to the simulation chain*. It is UV physics that scales with the leading-order Wilson coefficient dependence, that is known, and being used in the fit. The divergences exactly cancel after renormalization, however the log terms that are associated with the divergences, defining e_{in}, do not cancel, but are predicted.

The log terms have the interpretation of the Wilson coefficients at the scale Λ vs the measurement scale μ. For inclusive on-shell Higgs decay $\mu^2 = m_h^2$. If RGE-improved perturbation theory is used, the log terms can be summed using standard EFT techniques. When doing a LO SMEFT fit in fixed-order perturbation theory the log terms are the (generally largest) part neglected terms in perturbation theory. Z_{SM} refers to the SM wave function/$\bar{v}_T$ renormalization. This causes extra log terms in addition to

those inferred with this procedure [21], and $Z_{\rm SM}$ is also modified with dependence on $C_i^{(6)}$. This dependence [18] is an example of operators mixing down, and is always $\propto \lambda \bar{v}_T^2$. Such corrections can be included into the error estimate by rescaling the SM couplings with these modifications. The RGE is insufficient to characterize the full perturbative corrections for lower values of Λ and are used for determining the central values of parameters [21]. Practically, for such Λ values, the $\mathcal{L}^{(8)}$ corrections that are also neglected dominate the error estimate. Using the log terms as a reasonable proxy for unknown perturbative corrections, is sufficient, well defined, and known at this time.

To preserve the Ward identities of the SMEFT [4] operator by operator, it is necessary to expand out the propagator shifts in the SM masses due to higher-dimensional operators [22]. These shifts should be considered part of the theoretical prediction at LO. The perturbative error algorithm can be applied to such corrections. Width shifts are not required to preserve the Ward ID at one loop, and the pert error algorithm should not be applied to such terms.

The log-enhanced one-loop correction to the SMEFT amplitude modifies each observable according to

$$\nabla \sigma_{1/16\pi^2\Lambda^2} \approx \sum A_i \tilde{C}_j \frac{e_{ij}}{16\pi^2} \log \frac{\mu^2}{\Lambda^2}, \qquad (8)$$

and provides an estimate of the neglected perturbative corrections. We use ∇ for errors in this work.

A. Example: VBF Higgs production

To demonstrate the determination of perturbative uncertainties we consider the inclusive result for the "$\bar{q}q \to h\bar{q}q$ VBF-like" process quoted in Table 9 of Ref. [17] ("direct" contributions)

$$\frac{A_i^{h\bar{q}q \text{ VBF-like}}}{\sigma_{\rm SM}^{h\bar{q}q \text{ VBF-like}}} = \{-6, 0.109, -5.345, -0.323, 0.103, 3\}, \quad (9)$$

for $\tilde{C}_i^{(6)} = \{\tilde{C}_{H\ell}^{(3)}, \tilde{C}_{Hq}^{(1)}, \tilde{C}_{Hq}^{(3)}, \tilde{C}_{Hu}, \tilde{C}_{Hd}, \tilde{C}_{\ell\ell}'\}$. We restrict our attention to perturbative corrections proportional to the top-quark Yukawa in this example, see Eqs. (A.27–30, A.33,A.35) of Ref. [19].

$$\tilde{C}_j^{(6)} e_{ij} = -2 N_c \tilde{C}_{\substack{\ell q \\ pp33}}^{(3)} y_t^2 + 2 N_c y_t^2 \tilde{C}_{\substack{H\ell \\ pp}}^{(3)} + \cdots \quad (10)$$

for $\tilde{C}_i^{(6)} = \tilde{C}_{H\ell}^{(3)}$. The perturbative error that follows is

$$\sim A_i^I \tilde{C}_j \frac{e_{ij}}{16\pi^2} \log \frac{\mu^2}{\Lambda^2},$$

$$\sim -6 \frac{\sigma_{VBF\text{-like}}^{h\bar{q}q}}{16\pi^2} \left[-2N_c \tilde{C}_{\substack{\ell q \\ pp33}}^{(3)} y_t^2 + 2N_c y_t^2 \tilde{C}_{\substack{H\ell \\ pp}}^{(3)} \right] \log \frac{\mu^2}{\Lambda^2}. \quad (11)$$

The results from Ref. [17] have the cuts $m_{jj} > 350$ GeV and $p_T(h) < 200$ GeV. A reasonable kinematic invariant to choose for the μ dependence in the logarithm is $\mu \sim 200$ GeV. p is a flavor label. Consistent with the lepton flavor assumption, it is summed over $p = \{1, 2, 3\}$. This procedure is repeated for all of $\tilde{C}_i^{(6)} = \{\tilde{C}_{H\ell}^{(3)}, \tilde{C}_{Hq}^{(1)}, \tilde{C}_{Hq}^{(3)}, \tilde{C}_{Hu}, \tilde{C}_{Hd}, \tilde{C}_{\ell\ell}'\}$ for each SM coupling dependence that one wishes to retain. For practical purposes retaining the dependence on $y_t, y_b, g_{1,2,3}$, is sufficient. An error estimate results that is a linear sum of unknown (nuisance) parameters, with calculated coefficients dominated by y_t, g_3 corrections. A distribution of the unknown Wilson coefficients is chosen to produce a number, to add in quadrature with other errors. A very weak dependence is present on the distribution chosen, consistent with the [23] central limit theorem, as shown in Refs. [6,24].

IV. MISSING $\mathcal{O}(1/\Lambda^4)$ CORRECTIONS

Many of the $\mathcal{O}(1/\Lambda^4)$ corrections can be determined from LO SMEFT results with rescalings. It is appropriate to organize the theory as specific composite-operator kinematics, with scalar dressings that do not introduce new kinematics, to identify these rescalings. This is the geoSMEFT approach developed in Refs. [3–6,25] where scalar field dependent field-space connections G_i multiply composite operator forms f_i as

$$\mathcal{L}_{\rm SMEFT} = \sum_i G_i(I, A, \phi\ldots) f_i. \qquad (12)$$

Powers of $D^\mu H$ are included in f_i. The kinematic dependence is factorized into the f_i and the rescalings by G_i are the rescalings needed to produce $\mathcal{O}(1/\Lambda^4)$ corrections from LO simulation results.

Example: $h \to \gamma\gamma$—The geoSMEFT is defined up to four-point interactions at this time. The rescaling procedure is best illustrated with a specific example. The three-point function $h - \gamma - \gamma$ in the SMEFT to all orders is given in Ref. [5]. $\langle h\gamma^{\mu\nu}\gamma_{\mu\nu}\rangle$ is a common kinematic factor for the $\mathcal{L}^{(6,8)}$ contributions. Replacements can be made on the $\mathcal{L}^{(6)}$ dependence to directly generate $\mathcal{O}(1/\Lambda^4)$ terms. Using the all-orders definition of the decay width Ref. [5,6] at LO one finds

$$\frac{\Gamma^{\hat{m}_W}(h \to \gamma\gamma)}{\Gamma_{\rm SM}^{\hat{m}_W}(h \to \gamma\gamma)} = 1 - 788 f_1^{\hat{m}_W}, \qquad (13)$$

where $f_i^{\hat{m}_W} \simeq f_i^{\hat{\alpha}_{ew}}$ for $i = 1, 2, 3$ and these functions (and $\delta G_F^{(6)}$) are defined in Ref. [6]. The full result to $\mathcal{O}(1/\Lambda^4)$, in the m_W scheme is [6]

$$\frac{\Gamma^{\hat{m}_W}(h \to \gamma\gamma)}{\Gamma^{\hat{m}_W}_{\rm SM}(h \to \gamma\gamma)} = 1 - 788 f_1^{\hat{m}_W} + 394^2 (f_1^{\hat{m}_W})^2 - 351 (\tilde{C}^{(6)}_{HW} - \tilde{C}^{(6)}_{HB}) f_3^{\hat{m}_W} + 2228 \, \delta G^{(6)}_F \, f_1^{\hat{m}_W}$$

$$+ 979 \tilde{C}^{(6)}_{HD} (\tilde{C}^{(6)}_{HB} + 0.80 \tilde{C}^{(6)}_{HW} - 1.02 \tilde{C}^{(6)}_{HWB}) - 788 \left[\left(\tilde{C}^{(6)}_{H\Box} - \frac{\tilde{C}^{(6)}_{HD}}{4} \right) f_1^{\hat{m}_W} + f_2^{\hat{m}_W} \right]$$

$$+ 2283 \tilde{C}^{(6)}_{HWB} (\tilde{C}^{(6)}_{HB} + 0.66 \tilde{C}^{(6)}_{HW} - 0.88 \tilde{C}^{(6)}_{HWB}) - 1224 (f_1^{\hat{m}_W})^2. \tag{14}$$

A leading-order simulation result using SMEFTsim that gives dependence on the Wilson coefficients in $f_1^{\hat{m}_W}$ or $f_1^{\hat{\alpha}_{ew}}$, with numerical dependence multiplying this function in an observable N_i, can generate the remaining $\mathcal{O}(1/\Lambda^4)$ terms via the replacement

$$N_i f_1^{\hat{m}_W} \to \frac{-N_i}{788} \left[394^2 (f_1^{\hat{m}_W})^2 - 351 (\tilde{C}^{(6)}_{HW} - \tilde{C}^{(6)}_{HB}) f_3^{\hat{m}_W} + 2228 \, \delta G^{(6)}_F \, f_1^{\hat{m}_W} \right.$$

$$+ 979 \tilde{C}^{(6)}_{HD} (\tilde{C}^{(6)}_{HB} + 0.80 \, \tilde{C}^{(6)}_{HW} - 1.02 \tilde{C}^{(6)}_{HWB}) - 788 \left[\left(\tilde{C}^{(6)}_{H\Box} - \frac{\tilde{C}^{(6)}_{HD}}{4} \right) f_1^{\hat{m}_W} + f_2^{\hat{m}_W} \right]$$

$$\left. + 2283 \tilde{C}^{(6)}_{HWB} (\tilde{C}^{(6)}_{HB} + 0.66 \, \tilde{C}^{(6)}_{HW} - 0.88 \tilde{C}^{(6)}_{HWB}) - 1224 (f_1^{\hat{m}_W})^2 \right]. \tag{15}$$

The use of this replacement in the one-loop result for the $f_1^{\hat{m}_W}$ dependence in $\Gamma(h \to \gamma\gamma)$ introduces a relative uncertainty of $(\bar{v}_T^4/\Lambda^4)(1/16\pi^2)$. The replacement generates not only the quadratic terms, but also the full set of $\bar{v}_T^4/\Lambda^4$ corrections contributing to $\nabla\Gamma(h \to \gamma\gamma)$. An error is assigned by choosing a set of distributions for the $C_i^{(6)}, C_i^{(8)}$, and a value for Λ when neglecting this class of terms. The choice of Λ dictates the size of the error induced, and it is appropriate to choose multiple values of Λ when determining errors. A straightforward choice is $\Lambda \sim 1$ TeV and $\Lambda \sim 3$ TeV. $\mathcal{L}^{(8)}$ induced errors dominate for the former choice, while errors due to neglected perturbative corrections dominate for the latter choice.

$h - \gamma - \gamma$ has only one vertex to rescale. Extending this procedure to multiple Feynman diagrams, where each vertex building up the individual Feynman diagrams is generalized into the case of the geoSMEFT, requires the individual dependence on at least one Wilson coefficient present at $\mathcal{L}^{(6)}$ in each type of vertex be identified and isolated, so that a rescaling procedure can be performed. The same $\mathcal{L}^{(6)}$ correction can appear in multiple vertices, so projecting out the dependence of a Wilson coefficient at a particular vertex via linear algebra is required. Consider a process where the same $\mathcal{L}^{(6)}$ Wilson coefficient, $C_1^{(6)}$, appears in two vertices in a Feynman diagram, or sum of diagrams with dependences

$$\delta V_1 \propto a_1 \tilde{C}^{(6)}_1 + a_2 \tilde{C}^{(6)}_2 + a_3 \tilde{C}^{(6)}_3,$$
$$\delta V_2 \propto b_1 \tilde{C}^{(6)}_1 + b_2 \tilde{C}^{(6)}_2 + b_3 \tilde{C}^{(6)}_3. \tag{16}$$

The kinematics associated with the vertices $\delta V_{1,2}$ can differ in what follows.[4] Each of the V_1 and V_2 have a rescaling in the geoSMEFT, but the appearance in the overall result is a convolution of dependence on $\tilde{C}^{(6)}_{1,2,3}$ from both vertices. Isolating $C_1^{(6)}$ in V_1 to generate the $\mathcal{O}(1/\Lambda^4)$ result, one can choose to fix $\delta V_2 = 0$ in the known SMEFT result by choosing $\tilde{C}^{(6)}_2 = (-b_1 \tilde{C}^{(6)}_1 - b_3 \tilde{C}^{(6)}_3)/b_2$ in the LO result. The resulting shift due to δV_1 is then modified to

$$\delta V'_1 \propto \frac{a_1 - a_2 b_1}{b_2} \tilde{C}^{(6)}_1 + \frac{a_3 - a_2 b_3}{b_2} \tilde{C}^{(6)}_3. \tag{17}$$

The geoSMEFT based rescaling uses the known $\delta V_1 = a_1 \tilde{C}^{(6)}_1$ and $\mathcal{L}^{(8)}$ result $\delta V_1 = b_{jk} \tilde{C}^{(6)}_j \tilde{C}^{(6)}_k + c_l \tilde{C}^{(8)}_l$. The net dependence on $\tilde{C}^{(6)}_1$ in Eq. (17) can be rescaled back to a net a_1 dependence using the known dependence on all vertices in the contribution to the observable; i.e., a_1, a_2, b_{1-3} are known analytically in the LO SMEFT results encoded in SMEFTsim. This procedure can be iterated. Performing the full set of rescaling replacements then generalizes the LO SMEFTsim result with a well-defined class of terms at $\mathcal{O}(1/\Lambda^4)$. Distributions of $C_i^{(6)}$ and $C_i^{(8)}$, and a chosen Λ scale, then defines a numerical error.

---

[4]Here $a_{1,2,3}$ and $b_{1,2,3}$ refer to explicit analytic (or numerically evaluated) dependence on a Wilson coefficient in a vertex. For example, when $\delta V_1 = f_1^{\hat{m}_W}$ then $a_{HB} = 1$, $a_{HW} = 0.29$, $a_{HWB} = -0.54$.

V. NEW KINEMATICS BEYOND $\mathcal{L}^{(6)}$

The procedure outlined above relies on the LO simulation for global SMEFT studies having a complete set of kinematic functions to rescale. New kinematic forms appear first at $\mathcal{L}^{(8)}$ in SMEFT global fits of Higgs and EW data, but such anomalous kinematics is remarkably limited when only considering $n \leq 3$ point interactions building up pole observables. The anomalous kinematics are limited to the field space connections of the geoSMEFT of the form[5]

$$k_{IJA}(\phi)(D^\mu\phi)^I(D^\nu\phi)^J \mathcal{W}^A_{\mu\nu}. \tag{18}$$

The operator contributions to this field space connection are equation of motion (EOM) reducible at $\mathcal{L}^{(6)}$, and hence are not present in the Warsaw basis. At $\mathcal{L}^{(8)}$ such terms are no longer EOM reducible. Such contributions modify $VV \rightarrow h$ production, and $h \rightarrow 4\ell$ through the modification of the hVV vertex (here V is a general vector). These corrections must be added in a dedicated extension of SMEFTsim [26].[6] Similarly, modified kinematics is present in VH production which requires a dedicated extension of simulation tools, see Ref. [27]. These contributions can be directly targeted for code extensions and direct simulation to complete the calculation of relevant observables to $\mathcal{O}(1/\Lambda^4)$ using the $\mathcal{L}^{(8)}$ operator basis in Refs. [28,29]. Such dedicated extensions to SMEFTsim by subsets of $\mathcal{L}^{(8)}$ operators, combined with the algorithm above results in a fully well-defined theoretical result to $\mathcal{O}(1/\Lambda^4)$ process by process, and such results can then be used in order to generate a theory error by directly varying the subleading terms. The majority of the $\mathcal{O}(1/\Lambda^4)$ pole observables, can be generated with geoSMEFT based rescaling, avoiding the need for simulation or code modification. When considering tail observables, four-point interactions are generally important and unsuppressed compared to other SMEFT corrections.[7] It is then required to add more operators to simulation tools to characterize observables to $\mathcal{O}(1/\Lambda^4)$.

VI. QUADRATIC TERMS

Here we assess the use of quadratic terms for a theory error estimate, and if the use of quadratic terms for SMEFT fits to define central values in the fit is well defined.

Quadratic terms means the result of squaring Eq. (2); the resulting $\mathcal{O}(1/\Lambda^4)$ term is the quadratic term. In general, retaining only a subset of terms in the power counting of an EFT is an ill-defined procedure, which is not invariant

under the field redefinitions that define the theory.[8] In the case of the SMEFT, retaining quadratic terms is subject to the following field redefinition based ambiguities. Eq. (2) should be understood to have unspecified but existent terms of the form

$$\mathcal{A} = \mathcal{A}_{\text{SM}}\left(1 + n_i\frac{\bar{v}_T^4}{\Lambda^4}\right) + \tilde{C}_k^{(6)}a_k\left(1 + o_j\frac{\bar{v}_T^2}{\Lambda^2}\right) + \cdots. \tag{19}$$

There are also corrections of $\mathcal{O}(1/\Lambda^4)$ with dynamical fields of dimension four and two in each case. Field redefinitions of $\mathcal{O}(1/\Lambda^4)$ is fundamental to defining the SMEFT, and such an ambiguity is not fixed by defining the theory to $\mathcal{O}(1/\Lambda^2)$. Predictions proportional to n_i and o_j are ambiguous until the full set of corrections are defined at $\mathcal{O}(1/\Lambda^4)$ by defining an operator basis for $\mathcal{L}^{(8)}$. Squaring this result gives terms of $\mathcal{O}(1/\Lambda^4)$

$$\left(a_k\tilde{C}_k^{(6)}a_l\tilde{C}_l^{(6)} + 2n_i\frac{\bar{v}_T^4}{\Lambda^4}|\mathcal{A}_{\text{SM}}|^2 + 2a_k\mathcal{A}_{\text{SM}}\tilde{C}_k^{(6)}o_j\frac{\bar{v}_T^2}{\Lambda^2}\right). \tag{20}$$

All terms are of $\mathcal{O}(1/\Lambda^4)$ and the n_i and o_j are arbitrary as it stands. These terms can be chosen to have $\tilde{C}_i^{(6)}$ dependence, modifying the dependence on $\tilde{C}_k^{(6)}\tilde{C}_l^{(6)}$.

Ref. [6] demonstrated that the quadratic term in $h - \gamma - \gamma$ was not correctly predicting dependence on $(f_1^{\hat{m}_W})^2$ due to the redefinition of the electric coupling in the geoSMEFT. This redefinition is a specific example of field redefinition based correction of the form shown in Eq. (19). In addition, the arbitrariness represented by n_i, o_j until $\mathcal{L}^{(8)}$ is defined is required to ensure that SMEFT predictions are well defined at $\mathcal{O}(1/\Lambda^4)$ and not intrinsically dependent on the basis chosen for $\mathcal{L}^{(8)}$. When defining $\mathcal{L}^{(8)}$ the use of the Higgs EOM

$$D^2 H_k - \lambda\bar{v}_T^2 H_k + 2\lambda(H^\dagger H)H_k + \bar{q}^j Y_u^\dagger u\epsilon_{jk}$$
$$+ \bar{d}Y_d q_k + \bar{e}Y_e l_k = 0, \tag{21}$$

leads to modifications of $\mathcal{L}^{(6)}$ terms that are $\propto \lambda\bar{v}_T^2$. These terms cancel in a full matching (see Ref. [6,24]). This correlated matching of $\mathcal{L}^{(8)}$ and $\bar{v}_T^2$ corrections to $\mathcal{L}^{(6)}$ is an example of matching effects descending down in an operator mass dimension, which is similar to the mixing of operators of different mass dimensions in the SMEFT RGE; both effects come about due to $\bar{v}_T^2$. $\tilde{C}_k^{(6)}a_k$ is not well defined in its predictions to $\mathcal{O}(1/\Lambda^4)$ neglecting such effects as the Higgs EOM is not consistently applied when neglecting such terms. The quadratic terms are in general

---

[5]Here ϕ is a four-component real vector defining the components of the Higgs field H.

[6]This modification is available from the author and T. Corbett.

[7]See for example Ref. [30].

[8]One can directly confirm that inconsistent expansions in the power counting expansion among vertex functions violate SMEFT-Ward identities [4].

not well-defined contributions at $\mathcal{O}(1/\Lambda^4)$ for these reasons and they should not be used to fix central values in the global SMEFT fit as a default prediction. Such corrections cannot be translated unambiguously between operator bases until the theory is fixed to $\mathcal{O}(1/\Lambda^4)$ as Eq. (21), and other EOM relations between $\mathcal{L}^{(6)}$ operators, have implicit (usually unspecified) corrections that are further suppressed by $\mathcal{O}(1/\Lambda^2)$.[9]

Nevertheless, the use of quadratic terms to define an error estimate is a reasonable procedure [32–35], in the absence of complete results developed using this methodology for pole observables. As such, the use of quadratic terms to estimate an error for tail observables as advocated in Ref. [33–35] can be appropriate.

VII. CONCLUSIONS

We have defined a methodology to improve predictions for pole observables in the SMEFT to $\mathcal{O}(1/\Lambda^4)$ using geoSMEFT results. When new kinematics are first present at $\mathcal{L}^{(8)}$, modifications to code tools, and new simulation and event generation is required to complete results to $\mathcal{O}(1/\Lambda^4)$. However such corrections are a small subset of the full set of corrections extending predictions

to $\mathcal{O}(1/\Lambda^4)$. This approach to improving LO results to subleading order in the operator expansion relies on simple linear algebra, Taylor expansions of known closed form all-orders expressions in the geoSMEFT, and the known dependence on $\mathcal{L}^{(6)}$ encoded in SMEFTsim. The approach outlined here can be combined with the approach of Refs. [32–36] for tail observables.

Truncation errors result from taking the resulting exact expressions to $\mathcal{O}(1/\Lambda^4)$ and varying the unknown higher-order terms in a range of values. Then $\sqrt{(\nabla\sigma_{1/\Lambda^4})^2 + (\nabla\sigma_{1/16\pi^2\Lambda^2})^2}$ defines a theory error estimate. As the operator expansion in the SMEFT involves many terms at $\mathcal{O}(1/\Lambda^4)$ that are randomly chosen via distributions in linear sums, the central limit theorem applies. In global combinations, common values of Λ should be chosen to define errors for observables. The resulting SMEFT theory error for a LO fit is a Gaussian-distributed numerical value for each observable, with magnitude determined by the chosen Λ.

---

[9]Such corrections to the SMEFT EOM are reported in Ref. [31].

[1] W. Buchmuller and D. Wyler, Nucl. Phys. **B268**, 621 (1986).

[2] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, J. High Energy Phys. 10 (2010) 085.

[3] A. Helset, M. Paraskevas, and M. Trott, Phys. Rev. Lett. **120**, 251801 (2018).

[4] T. Corbett, A. Helset, and M. Trott, Phys. Rev. D **101**, 013005 (2020).

[5] A. Helset, A. Martin, and M. Trott, J. High Energy Phys. 03 (2020) 163.

[6] C. Hays, A. Helset, A. Martin, and M. Trott, J. High Energy Phys. 11 (2020) 087.

[7] I. Brivio, Y. Jiang, and M. Trott, J. High Energy Phys. 12 (2017) 070.

[8] M. J. Savage and B. R. Holstein AIP Conf.Proc. **549**, 120 (2000).

[9] A. Pich, Rep. Prog. Phys. **58**, 563 (1995).

[10] C. W. Bauer, Z. Ligeti, M. Luke, A. V. Manohar, and M. Trott, Phys. Rev. D **70**, 094017 (2004).

[11] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, Phys. Rev. D **83**, 074021 (2011).

[12] G. Passarino, Nucl. Phys. **B868**, 416 (2013).

[13] A. David and G. Passarino, Phys. Lett. B **726**, 266 (2013).

[14] J. Dobaczewski, W. Nazarewicz, and P. G. Reinhard, J. Phys. G **41**, 074001 (2014).

[15] B. D. Carlsson, A. Ekström, C. Forssén, D. F. Strömberg, G. R. Jansen, O. Lilja, M. Lindby, B. A. Mattsson, and K. A. Wendt, Phys. Rev. X **6**, 011019 (2016).

[16] R. J. Hill and G. Paz, Phys. Rev. D **82**, 113005 (2010).

[17] I. Brivio, J. High Energy Phys. 04 (2021) 073.

[18] E. E. Jenkins, A. V. Manohar, and M. Trott, J. High Energy Phys. 10 (2013) 087.

[19] E. E. Jenkins, A. V. Manohar, and M. Trott, J. High Energy Phys. 01 (2014) 035.

[20] R. Alonso, E. E. Jenkins, A. V. Manohar, and M. Trott, J. High Energy Phys. 04 (2014) 159.

[21] C. Hartmann and M. Trott, J. High Energy Phys. 07 (2015) 151.

[22] T. Corbett and M. Trott, SciPost Phys. **10**, 144 (2021).

[23] http://www.medicine.mcgill.ca/epidemiology/hanley/bios601/GaussianModel/HistoryCentralLimitTheorem.pdf.

[24] T. Corbett, A. Helset, A. Martin, and M. Trott, J. High Energy Phys. 06 (2021) 076.

[25] T. Corbett, J. High Energy Phys. 03 (2021) 001.

[26] M. Trott (private communication).

[27] C. Hays, A. Martin, V. Sanz, and J. Setford, J. High Energy Phys. 02 (2019) 123.

[28] C. W. Murphy, J. High Energy Phys. 10 (2020) 174.

[29] H.-L. Li, Z. Ren, J. Shu, M.-L. Xiao, J.-H. Yu, and Y.-H. Zheng, Phys. Rev. D **104,** 015026 (2021).

[30] R. Boughezal, E. Mereghetti, and F. Petriello, arXiv:2106 .05337.

[31] A. Barzinji, M. Trott, and A. Vasudevan, Phys. Rev. D **98,** 116005 (2018).

[32] L. Berthier and M. Trott, J. High Energy Phys. 02 (2016) 069.

[33] S. Alte, M. Knig, and W. Shepherd, J. High Energy Phys. 07 (2019) 144.

[34] E. Keilmann and W. Shepherd, J. High Energy Phys. 09 (2019) 086.

[35] S. Alte, M. König, and W. Shepherd, J. High Energy Phys. 01 (2018) 094.

[36] A. Horne, J. Pittman, M. Snedeker, W. Shepherd, and J. W. Walker, J. High Energy Phys. 03 (2021) 118.