Statistical approach to Higgs boson couplings in the standard model effective field theory

Christopher W. Murphy[*]

Department of Physics, Brookhaven National Laboratory, Upton, New York 11973, USA

We perform a parameter fit in the standard model effective field theory (SMEFT) with an emphasis on using regularized linear regression to tackle the issue of the large number of parameters in the SMEFT. In regularized linear regression, a positive definite function of the parameters of interest is added to the usual cost function. A cross-validation is performed to try to determine the optimal value of the regularization parameter to use, but it selects the standard model (SM) as the best model to explain the measurements. Nevertheless as proof of principle of this technique we apply it to fitting Higgs boson signal strengths in SMEFT, including the latest Run-2 results. Results are presented in terms of the eigensystem of the covariance matrix of the least squares estimators as it has a degree model-independent to it. We find several results in this initial work: the SMEFT predicts the total width of the Higgs boson to be consistent with the SM prediction; the ATLAS and CMS experiments at the LHC are currently sensitive to non-resonant double Higgs boson production. Constraints are derived on the viable parameter space for electroweak baryogenesis in the SMEFT, reinforcing the notion that a first order phase transition requires fairly low-scale beyond the SM physics. Finally, we study which future experimental measurements would give the most improvement on the global constraints on the Higgs sector of the SMEFT.

I. INTRODUCTION

The Higgs boson discovered at the Large Hadron Collider (LHC) very much resembles the one predicted by the standard model (SM) [1]. Unfortunately, to date, no other particles have been discovered at the LHC [2], indicating there is a mass gap between the SM and whatever may lie beyond it.[1] Such a separation of scales lends itself to an effective field theory (EFT) treatment, and the standard model effective field theory (SMEFT) is a well-developed subject [5,6].

An issue when dealing with the SMEFT is the large number of parameters it contains. There are 2,499 baryon number preserving real parameters at dimension-6 [7], and this number grows exponentially with the number of dimensions [8]. Following the pioneering analysis of Ref. [9] many parameter fits in the SMEFT have been performed [10–45]. These more recent analyses have often focused on constraining the Higgs sector of the SMEFT.

In this work, we also perform an SMEFT parameter fit, but with an emphasis on a statistical technique aimed at tackling the issue of the large number of parameters. In particular, the technique we use is a regularized linear regression, where a positive definite function of the parameters of interest is added to the usual cost function. This prevents the fit from falling into an overfit solution, and, in principle, allows information to be obtained about any number of parameters. One application of this in particle physics is unfolding a differential cross section from the detector level to the truth level [46,47] Additionally, it is a commonly used technique in machine learning [48,49], and finds various applications in lattice physics; see e.g. [50–52].

As is typically done, a cross-validation is performed to try to determine the optimal value of the regularization parameter to use. However, it selects the SM as the best model to explain the experimental measurements. Nevertheless we persist in studying the SMEFT, contenting ourselves to performing regularized fits with multiple choices for the regularization parameter, and examining how much regulator dependence various quantities have.

As proof of principle of this technique, we apply it to fitting Higgs boson signal strengths, including the latest Run-2 results. Following Ref. [32], we emphasize presenting results in terms of the eigensystem of the covariance matrix of the least squares estimators as it has a degree

---

[1]Exceptions to this could be a hidden sector with (sub-)GeV particles, possibly related to dark matter [3], or the alignment without decoupling limit of the two-Higgs doublet model [4].

model-independent to it. Despite this being an initial study we obtain several useful physics results. We show the SMEFT predicts the total width of the Higgs boson, which is not yet directly measured, to be consistent with the SM prediction, and that the ATLAS and CMS experiments at the LHC are currently sensitive to nonresonant double Higgs boson production. We derive constraints on the viable parameter space for electroweak (EW) baryogenesis in the SMEFT, and reinforce the notion that a first-order phase transition requires fairly low-scale beyond the SM (BSM) physics. We study which future experimental measurements would improve the global constraints on the Higgs sector of the SMEFT the most. This is quantified using ratios of the global determinant parameter (GDP) of Ref. [53], which has a natural interpretation in terms of the eigensystem of the covariance matrix.

We expect this technique to be of use to practitioners of both bottom-up and top-down approaches to EFTs. In the case of the former, this technique could be applied as described in this work to more sophisticated SMEFT predictions as well as data sets that included differential measurements, or measurements from outside the Higgs sector of the SMEFT such as EW precision data (EWPD), triple gauge couplings, or flavor measurements. For the latter, the regularization matrix provides a convenient way to impose a prior assumptions about possible UV physics. Additionally, another advantage of this approach is that it makes it easy to determine the blind directions in parameter space for a given data set.

The rest of the paper is organized as follows. SMEFT predictions for Higgs boson processes including the Higgs trilinear coupling are given in Sec. II. Next the fitting procedures used in this work, and the statistical approaches they employ, are described in Sec. III. The experimental results used in these fits are compiled in Appendix A. The results of our fits to the Higgs signal strength measurements are then presented in Sec. IV with additional information given in Appendixes B and C. Finally, we summarize our findings in Sec. V.

II. STANDARD MODEL EFT PREDICTIONS

The Lagrangian of the SMEFT is given by

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}_{\mathrm{SM}} + \mathcal{L}^{(5)} + \mathcal{L}^{(6)} + \cdots, \qquad (1)$$

where the superscript n in the non-SM terms indicates the mass dimension of the operators contained in that term.

The Yukawa couplings and the dimension-6 Wilson coefficients implicit in Eq. (1) are in general matrices in flavor space. Additionally, Higgs boson interactions with fermions inherently have a nontrivial flavor structure, and thus it is important that whatever theoretical framework is used to interpret Higgs measurement also have some nontriviality in its flavor structure. With these considerations in mind, the number of parameters can be reduced to a somewhat manageable number of 18 by imposing a $U(2)^5$ symmetry under which the first two generations transform as doublets and the third generation as singlets [54].[2] In the basis of Ref. [59] with an approximate $U(2)^5$ flavor symmetry, these operators are Q_H, $Q_{H\square}$, Q_{HD}, Q_{HG}, Q_{HW}, Q_{HB}, Q_{HWB}, Q_{uH}^{33}, Q_{dH}^{33}, Q_{eH}^{33}, $Q_{H\ell}^{(3)}$, $Q_{H\ell}^{(1)}$, $Q_{Hq}^{(3)}$, $Q_{Hq}^{(1)}$, Q_{He}, Q_{Hu}, Q_{Hd}, and $Q_{\ell\ell}$. Operators without a generation label are $U(2)^5$ symmetric. See Ref. [60] for additional parameter counting along these lines.

This is a proof of principle work regarding the usefulness of the statistical methods in constraining SMEFT coefficients. As such we made an additional simplification with respect to the SMEFT predictions. Specifically, we assume that the production or decay of a Higgs boson involving a pair of W or Z boson does not depend on the type of fermion that produces W or Z, or the type of fermion the W or Z decays into. Clearly both VBF and the associated Vh production mechanisms involve quarks. On the other hand, the best results of Higgs decays to Ws and Zs involve leptonic decays of the vector bosons [1]. Given the aforementioned assumptions, a subset of dimension-6 operators from Eq. (1) that is sufficient for our purposes is

$$\Delta\mathcal{L}^{(6)} = \frac{c_H}{v^2}\partial_\mu(H^\dagger H)\partial^\mu(H^\dagger H) + \frac{c_T}{v^2}\left|H^\dagger\overset{\leftrightarrow}{D}_\mu H\right|^2 + \frac{c_6}{v^2}(H^\dagger H)^3 + \frac{(H^\dagger H)}{v^2}[c_b(\bar{q}_{L3}d_{R3}H) + c_t(\bar{q}_{L3}u_{R3}\tilde{H}) + c_\tau(\bar{\ell}_{L3}e_{R3}H) + \mathrm{H.c.}]$$

$$+ \frac{ic_W}{v^2}(H^\dagger\sigma^i\overset{\leftrightarrow}{D}{}^\mu H)(D^\nu W_{\mu\nu})^i + \frac{ic_B}{v^2}(H^\dagger\overset{\leftrightarrow}{D}{}^\mu H)(D^\nu B_{\mu\nu}) + \frac{ic_{HW}}{v^2}(D^\mu H)^\dagger\sigma^i(D^\nu H)W_{\mu\nu}^i + \frac{ic_{HB}}{v^2}(D^\mu H)^\dagger(D^\nu H)B_{\mu\nu}$$

$$+ \frac{c_\gamma}{v^2}H^\dagger H B_{\mu\nu}B^{\mu\nu} + \frac{c_g}{v^2}H^\dagger H G_{\mu\nu}^a G^{a\mu\nu}, \qquad (2)$$

where $v = (\sqrt{2}G_F)^{-1/2} \approx 246$ GeV, and $H^\dagger\overset{\leftrightarrow}{D}_\mu H \equiv H^\dagger D_\mu H - (D_\mu H^\dagger)H$. Only third generation fermions

appear on the second line of (2), consistent with our assumption of a $U(2)^5$ flavor symmetry. A factor of v^{-2} has been extracted from the Wilson coefficients to make the c_i dimensionless. We will address the effect of different normalizations and UV assumptions later. This set of 12 parameters is collected into a vector for later convenience

---

[2]For applications of this symmetry in semileptonic B physics, see e.g. [55–58].

$$\mathbf{c}^\top = \{c_H, c_T, c_\gamma, c_g, c_{HW}, c_{HB}, c_W c_B, c_t, c_b, c_\tau, c_6\},\tag{3}$$

where $\top$ indicates the transpose.

Numerical results for the Higgs boson decay rates in the SMEFT based on Eq. (2) are given in [5,7,61,62]. The contribution to these decay rates from the Higgs trilinear coupling via electroweak loops is given in Ref. [63]. Combining these results we have

$$\frac{\Gamma(h \to \tau\tau)}{\Gamma_{SM}(h \to \tau\tau)} \simeq 1 - 2c_H - 196c_\tau, \qquad \frac{\Gamma(h \to \mu\mu)}{\Gamma_{SM}(h \to \mu\mu)} \simeq 1 - 2c_H,$$

$$\frac{\Gamma(h \to bb)}{\Gamma_{SM}(h \to bb)} \simeq 1 - 2c_H - 83c_b - 0.0085c_t, \qquad \frac{\Gamma(h \to cc)}{\Gamma_{SM}(h \to cc)} \simeq 1 - 2c_H - 0.015c_t,\tag{4}$$

and

$$\frac{\Gamma(h \to WW^*)}{\Gamma_{SM}(h \to WW^*)} \simeq 1 - 2.02c_H + 0.72c_W + 0.61c_{HW} - 0.057c_6,$$

$$\frac{\Gamma(h \to ZZ^*)}{\Gamma_{SM}(h \to ZZ^*)} \simeq 1 - 2.02c_H - 4c_T + 0.66c_W + 0.34c_B + 0.49c_{HW} + 0.26c_{HB} - 0.24c_\gamma - 0.064c_6,$$

$$\frac{\Gamma(h \to Z\gamma)}{\Gamma_{SM}(h \to Z\gamma)} \simeq 1 - 2c_H + 0.12c_t - 0.12c_b - 0.0088c_\tau + 1.38c_W + 151(0.16c_{HW} - 0.32c_{HB} + 1.58c_\gamma),$$

$$\frac{\Gamma(h \to \gamma\gamma)}{\Gamma_{SM}(h \to \gamma\gamma)} \simeq 1 - 2.01c_H + 0.54c_t - 0.29c_b - 0.69c_\tau + 1.66c_W - 863c_\gamma - 0.038c_6,$$

$$\frac{\Gamma(h \to gg)}{\Gamma_{SM}(h \to gg)} \simeq 1 - 2.02c_H - 2.13c_t + 4.17c_b + 589c_g - 0.051c_6.\tag{5}$$

The width of the Higgs boson in the SMEFT is determined based on Eqs. (4), (5) and the SM branching fractions given in Ref. [64]. We find

$$\frac{\Gamma_h}{\Gamma_{SM,h}} \simeq 1 - 2.007c_H - 0.11c_T - 1.61c_\gamma + 12.3c_g + 0.18c_{HW} - 0.067c_{HB}$$

$$+ 0.18c_W + 0.009c_B - 0.187c_t - 47.4c_b - 12.3c_\tau - 0.018c_6.\tag{6}$$

We take as numerical expression for Higgs boson production in the SMEFT the following:

$$\frac{\sigma(gg \to h)}{\sigma_{SM}(gg \to h)} \simeq \frac{\Gamma(h \to gg)}{\Gamma_{SM}(h \to gg)},$$

$$\frac{\sigma(pp \to jjh)}{\sigma_{SM}(pp \to jjh)} \simeq 1 - 2.02c_H - c_T - 0.06c_\gamma + 0.58c_{HW} + 0.085c_{HB} + 0.71c_W + 0.085c_B - 0.05c_6,$$

$$\frac{\sigma(pp \to Wh)}{\sigma_{SM}(pp \to Wh)} \simeq 1 - 2.03c_H + 0.61c_{HW} + 0.72c_W - 0.081c_6,$$

$$\frac{\sigma(pp \to Zh)}{\sigma_{SM}(pp \to Zh)} \simeq 1 - 2.04c_H - 4c_T - 0.24c_\gamma + 0.49c_{HW} + 0.34c_{HB} + 0.66c_W + 0.34c_B - 0.095c_6,$$

$$\frac{\sigma(pp \to t\bar{t}h)}{\sigma_{SM}(pp \to t\bar{t}h)} \simeq 1 - 2.11c_H - 2.01c_t - 0.29c_6.\tag{7}$$

The relative fractions of WW and ZZ in the vector boson fusion (VBF) production process are approximations based on Ref. [65]. Kinematic differences between production and decays modes, e.g. $\Gamma(h \to WW^*)$ versus $\sigma(pp \to Wh)$, or production cross sections at different center-of-mass energies, are not taken into account. Finally, the prediction for double-Higgs boson production in the SMEFT (at 14 TeV and considering only top quarks in the loop) is [66]

$$\frac{\sigma(gg \to hh)}{\sigma_{SM}(gg \to hh)} \simeq 1 + 4.25c_H - 469c_g + 3.7c_t - 8.8c_6.\tag{8}$$

III. FITTING PROCEDURE

In this section, we discuss the statistics of the two types of fits we perform. The first is the method of least squares that is ubiquitous in high energy physics. We then discuss a variations of this standard approach that can be used to avoid overfitting, regularizing the least squares fit.

The 55 experimental measurements from Run 1 and Run 2 used included in these fits are compiled in Appendix A. The measurements are all Higgs boson signal strengths. We do not consider differential or boosted Higgs measurements in this work. In addition, we do not include EWPD, triple gauge coupling, or flavor results in our fit. Lastly, no attempt is made to take theoretical errors into account in our fit whether they be from the SM prediction or the SMEFT theory error [26,32,36].

We will perform fits to these measurements with and without regularization, and for various choice of which parameters can be nonzero. A cross-validation test is performed to determine the optimal value of the regularization parameter to use in the fit.

A. Least squares review

We closely follow the presentation of the PDG [2] in what follows. The chi-squared function in the case of correlated measurements with covariance matrix V_{ij} is

$$\chi^2(\mathbf{c}) = (\mathbf{y} - \boldsymbol{\mu}(\mathbf{c}))^\top V^{-1} (\mathbf{y} - \boldsymbol{\mu}(\mathbf{c})), \qquad (9)$$

where $\mathbf{y}$ is the vector of measurements, $\boldsymbol{\mu}(\mathbf{c})$ is the vector of predictions and $\mathbf{c}$ is the vector of parameters to be estimated.

We consider the case where the predicted values are linear functions of the parameters

$$\mu(x_i; \mathbf{c}) = \sum_{j=1}^{m} h_j(x_i) c_j. \qquad (10)$$

In the standard case, $h_j(x)$ are m linearly independent functions. In addition, m must be less than the number of measurements, N. Furthermore, at least m of the x_i must be distinct.

It will be useful in what follows to define $H_{i,j} = h_j(x_i)$. Consider as an example the leading-order SMEFT prediction for the $h \to \gamma\gamma$ decay rate from Eq. (5). From this we see that, for instance, $H_{h\to\gamma\gamma,c_W} = 1.66$.

The least squares estimators for the parameters $\mathbf{c}$ are defined through $\boldsymbol{\nabla}\chi^2 = 0$,

$$\hat{\mathbf{c}} = (H^\top V^{-1} H)^{-1} H^\top V^{-1} \mathbf{y}. \qquad (11)$$

The inverse of the covariance matrix for the estimators is given by Hessian of chi-squared function, $\frac{1}{2}\nabla_i \nabla_j \chi^2$, or equivalently

$$U = (H^\top V^{-1} H)^{-1}. \qquad (12)$$

Note that for practical purposes we shift the ones in the SMEFT predictions for the Higgs boson's signal strengths into the measured values.

B. Regularized linear regression

As mentioned in the previous subsection, there are a number of conditions that must be satisfied for the standard least squares approach to be used. This technique is not useful when the covariance matrix of the estimators is ill-defined, which would be the case if, for example, the H_i are not sufficiently unique. These requirements can be bypassed by regularizing the least squares fit. In a regularized linear regression, the cost function is augmented with a positive-definite function of the parameters. In particular, the regularization makes the inverse of the Hessian of the chi-squared function, the Fisher information, well defined.

In this work, we use the following expression for the chi-square function as it admits a closed form solution for the least squares estimators

$$\chi^2(\mathbf{c}) = (\mathbf{y} - \boldsymbol{\mu}(\mathbf{c}))^\top V^{-1} (\mathbf{y} - \boldsymbol{\mu}(\mathbf{c})) + \mathbf{c}^\top \kappa \mathbf{c}, \quad (13)$$

with κ being a positive definite matrix. We primarily use the simple parametrization $\kappa_{ij} = \kappa \delta_{ij}$, which goes by several names: ridge regression, Tikhonov regularization, and ℓ_2 penalization. This choice of κ is the frequentist analog of adding the same Gaussian prior to each parameter of interest. Comments on other choices for κ are made later. Another commonly used regularization term is $\beta \sum_i |c_i|$, $\beta > 0$, which is known as Lasso regression or ℓ_1 penalization. This is a frequentist analog of adding the same Laplacian prior to each parameter of interest. One may also choose to use elastic net regularization, a linear combination of ridge and Lasso regression. We save these methods for future work.

In the case of ridge regression, the least squares estimators are given by

$$\hat{\mathbf{c}} = (H^\top V^{-1} H + \kappa\mathbb{1})^{-1} H^\top V^{-1} \mathbf{y}, \qquad (14)$$

with an obvious generalization for different choices of κ_{ij}. There is a similar modification to the covariance matrix

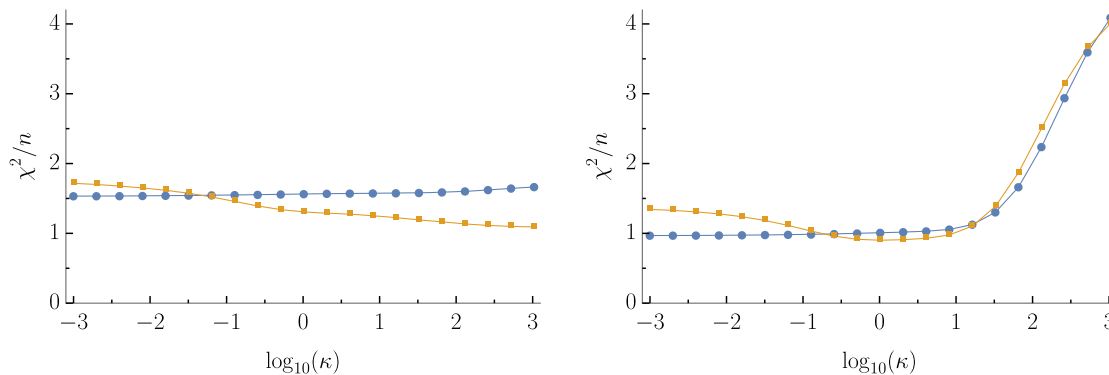$$U = (H^\top V^{-1} H + \kappa\mathbb{1})^{-1}. \qquad (15)$$

FIG. 1.   (Left:) Cross-validation test to determine the optimal value of κ. The average χ^2 per number of measurements is shown as a function of κ. The blue circles and orange squares correspond to the training (χ_t^2/n_t) and validation (χ_v^2/n_v) data sets, respectively. The cross-validation selects the SM ($\kappa \to \infty$) as the best model to explain the measurements. (Right:) The same as on the left, but with an artificial BSM signal injected. Here all tth signal strength central values have been fixed to 3.0 with their uncertainties left unchanged. In this case, the cross-validation selects $\kappa \approx 1$ as the best model as it has the lowest (average) value of χ_v^2/n_v.

C. Cross-validation

In cross-validation, the measurements are randomly split into training and validation groups.[3] The number of measurements assigned to the training group, n_t, is varied between 33 and 39 (60% to 71%) of the total $n = 55$ measurements. The test is performed 300 times for each value of n_t considered, and a chi-squared for the training set, χ_t^2, can be computed. The best-fit parameters are determined from the training set using regularization linear regression with some value of κ. These parameters are then used to compute the chi-squared for just the validation set, χ_v^2, which does not include a regularization term. The optimal choice of κ is given by the value which minimizes χ_v^2/n_v, where $n_v = n - n_t$ is the number of measurements in the validation group. See e.g. Refs. [48,49] for more information about cross-validation.

The average result of the 2100 cross-validation tests for a number of choices of κ between 10^{-3} and 10^3 are presented in the left panel of Fig. 1. Here the cross-validation selects $\kappa \to \infty$, the standard model, as the best model to explain the measurements. This is an uncommon result as typically the validation curve, the orange curve marked with squares in Fig. 1, has a local minimum at a finite value of κ. Another way to think about this is that the SMEFT may give a lower χ^2 than the SM, but the goodness of fit is still better in the SM. A similar conclusion, that the data prefer the SM over the SMEFT, was drawn in a Bayesian analysis of $b \to s\ell\ell$ observables [67–69]

The story would change if there was a (hint of a) signal for BSM physics. This is illustrated in the right panel of Fig. 1. Here we have injected an artificial BSM signal by setting the central values of all the tth signal strengths to

3.0 while leaving their uncertainties unchanged. In this case, the cross-validation selects $\kappa \approx 1$ as the best model as it has the lowest (average) value of χ_v^2/n_v.

Interestingly, in both panels of Fig. 1 the cross-validation suggests the data are being underfit. A hallmark of a model underfitting data is when the validation χ_v^2/n_v is comparable to, or smaller than, the training χ_t^2/n_t. Both panels of Fig. 1 then suggests that to avoid underfitting a value of κ less than one should be chosen. In any case, as no finite value of κ is preferred by the cross-validation, we will typically use two choices for κ and see how much regulator dependence there is in our predicted quantities.

IV. RESULTS

A. Covariance matrix eigensystem

It is instructive to examine the eigensystem of the covariance matrix for the least squares estimators [32]. There are $k = 1\ldots12$ eigenvectors, $W_k = w_{ki}c_i$, normalized such that $|\mathbf{w}_k| = 1$ with $\mathbf{c}$ given in Eq. (3). The square root of an eigenvalue, σ_k, gives the one sigma range on the allowed deviation of eigenvector, W_k, from its central value.

Results for σ_k are reported in Fig. 2. The blue circles and orange squares are the results of 12 parameter, regularized fits with $\kappa = 1$ and 10^{-2}, respectively. From these fits, we see that eigenvectors 1 and 2 are blind directions in parameter space as far as Higgs boson signal strengths are concerned. A direction k is called blind if $\sigma_k = 1/\sqrt{\kappa}$, independent of the choice of κ. An advantage of this approach is that it can quickly pick out these blind directions. Explicit expression for the eigenvectors in the $\kappa = 1$ case are given in Appendix B. Additional plots are presented in Appendix C.

In comparison, the green diamonds correspond to an unregularized, 10-parameter least squares fit. The two

---

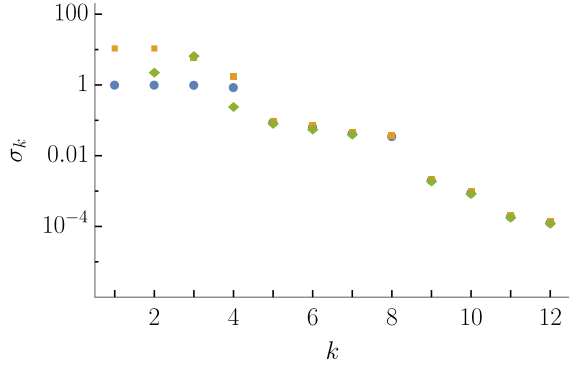[3]We ignore correlations between measurements during cross-validation.

FIG. 2. One sigma limits σ_k on the $k = 1\ldots12$ eigenvectors W_k of the covariance matrix for the least squares estimators. The blue circles and orange squares are the results of 12-parameter, regularized fits with $\kappa = 1$ and 10^{-2}, respectively. The green diamonds correspond to an unregularized, 10-parameter least squares fit.

parameters removed from Eq. (3) are c_T and $g_2 c_W + g_1 c_B$, which are two linear combinations of parameters appearing in eigenvectors 1 and 2 of the regularized fits. This is not a unique choice, but removing these combinations of parameters forces the oblique parameters S and T [70] to be zero at tree level. However, we caution that when there are many operators that can potentially be nonzero, c_T and $g_2 c_W + g_1 c_B$ need not be tiny to be consistent with EWPD [9,13,32,71–73].

When regularization is not important, eigenvectors 5 through 12, there is excellent agreement between the different cases. The bounds on eigenvectors 9 through 12 are at the level of a few permille or stronger. These eigenvectors are composed almost exclusively of c_b, c_τ, c_g, and c_γ, respectively. Weaker, percent level bounds, are found for eigenvectors 5 through 8. The third eigenvector is almost entirely composed of c_6. Not surprisingly the associated bound is weak, especially in the weakly and nonregularized cases, while the regularization with $\kappa = 1$ makes this bound artificially stronger. The correlation between the eigenvectors in regularized and unregularized cases is not exact, so a comparison of eigenvalues 2 and 4 is approximate.

The purpose of the regulation parameter is to the parameters of interest from becoming too large. Specifically, the standard deviations of the least squares estimators are regulated to have a maximum size

$$\Delta c_i \lesssim \frac{1}{\sqrt{\kappa}}. \tag{16}$$

Based on this, and given the normalization of the operators in Eq. (2), a choice of κ can be seen as imposing a prior assumption of the lowest possible scale of BSM physics, $\Lambda_{\min}$, or as imposing an upper limit on a signal strength from an experimental measurement. For example, if κ is taken to be 1 (10) then from (2), (16) we have

$\Lambda_{\min} \sim v (\Lambda_{\min} \sim 800~\text{GeV})$. The choice of $\kappa = 1$ can be seen as minimally enforcing the convergence of the EFT. However, this interpretation depends on the normalization of the operators in Eq. (2), with a different normalization generally leading to a different interpretation. A choice of the regularization parameter less than one could be used to enforce an experimental upper limit on a process that is not yet well measured, such as double Higgs boson production [74,75] or Higgs boson decay to a Z boson and a photon [76–78]. However, we do not pursue this approach, opting to include $h \to Z\gamma$ signal strengths in our fits, and making predictions for $gg \to hh$. The interpretation also depends on the structure of the regularization matrix, κ_{ij}. For instance, if one assumed a UV theory that is strongly-coupled, it might make more sense to relate the entries of κ_{ij} to the size of the coefficients expected from naïve dimensional analysis [79], rather than taking κ_{ij} to be proportional to the identity matrix.

B. Model independence of the eigensystem

Another way to understand the results is to look at two-dimensional profiles of the fits. We focus on c_γ and c_g in what follows. In doing so, we demonstrate that the eigensystem has a certain amount of model-independence to it. It would be interesting to investigate exactly how model-independent the eigensystem is. The key point discussed below is that marginalized allowed regions of parameter space in the parameter basis are sensitive to assumptions about the UV physics, whereas in the eigenbasis this is not the case.

The left panel of Fig. 3 shows the one and two sigma preferred values for c_γ and c_g—these regions are defined by $\Delta\chi^2 = 2.30$ and 6.18, respectively, with all other parameters fixed to their central values—for five scenarios to be defined. The blue and orange regions correspond to regularized fits with all 12 parameters, and $\kappa = 1$ and 10^{-2}, respectively. Darker (lighter) shading indicates the one (two) sigma allowed region. The red region is an unregularized least squares fit where the only two nonzero parameters are c_γ and c_g. Furthermore the purple region is also an unregularized least squares fit, but where the four parameters are nonzero $\{c_\gamma, c_g, c_{HW}, c_{HB}\}$, and where c_{HW}, c_{HB} are marginalized over. There is a noticeable lack of agreement between these different scenarios as to what are the preferred central values of c_γ and c_g are. In fact, for the fifth scenario, the unregularized 10-parameter fit described above, the preferred central values do not show up in the range of parameters plotted in Fig. 3. However, note that the variances of and correlation between c_γ and c_g are the same in all five cases.

To understand what is happening here consider the right panel of Fig. 3. This shows the same five fits, but in the plane of the eigenvectors W_{11} and W_{12}, rather than the parameters c_g, c_γ. All five fits agree perfectly as to what
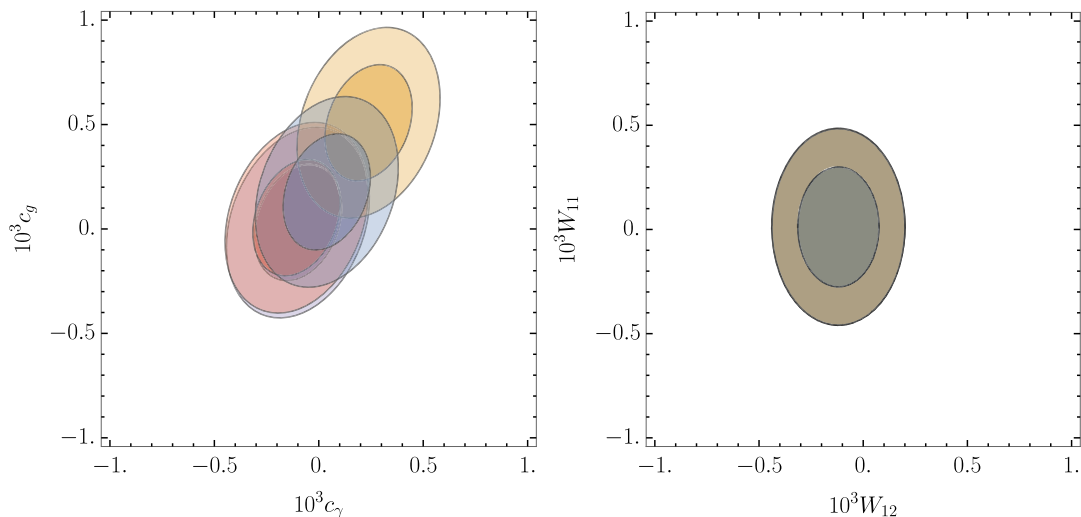
FIG. 3. (Left:) Preferred parameter space in the $c_g - c_\gamma$ plane based on the criteria $\Delta\chi^2 = 2.30$ (darker shading) and 6.18 (lighter shading). The central values of c_γ and c_g depend on the assumption of what additional parameters may be nonzero and how large they can be, while the variances of and correlation between c_γ and c_g are the same in all cases. See the text for details about the five scenarios. (Right:) In contrast, when the same fits are presented in terms of the eigenvectors W_{11} and W_{12}, perfect agreement between the five cases is found. This indicates the eigensystem has a degree of model independence to it.

the preferred region is in this case. That such a difference occurs is interesting because W_{12} and W_{11} are composed almost exclusively c_γ and c_g, respectively; $w_{12,\gamma} \approx w_{11,g} \approx$ 0.93. The difference, or lack thereof, between the central values in a given scenario occurs because when the additional parameters are fixed to their central values, which in turn forces $c_\gamma(c_g)$ away from the central value of $W_{12}(W_{11})$. In the case of the fit with only c_γ and c_g nonzero from the start, there are no additional parameters, and allowed contours in the $c_g - c_\gamma$ and $W_{11} - W_{12}$ plane are identical up to the rotating induced in going from one basis to the other. This shows that eigensystem is a fairly model-independent quantity. It depends only on the SMEFT framework, and in the particular cases of W_{11} or W_{12}, for example, that the parameters c_γ or c_g can be nonzero, but with no additional assumption about which parameters may or may not be nonzero.

C. Predictions

An advantage of having estimates for all of the coefficients under consideration is that predictions can be made for observables that have not been measured yet. For example, a prediction can be made for the total width of the Higgs boson. We find, using only Run-1 results,

$$\frac{\Gamma_{\text{SMEFT},h}}{\Gamma_{\text{SM},h}} \simeq 0.5 \pm 0.4 \quad (\text{Run 1}). \tag{17}$$

The Higgs decay rate to bottom quarks—the largest branching fraction in the SM—was measured to be low during Run 1 of the LHC, which explains why this value is

below the SM prediction. Adding results from Run 2, which are closer to the SM prediction, we instead find

$$\frac{\Gamma_{\text{SMEFT},h}}{\Gamma_{\text{SM},h}} \simeq 0.9 \pm 0.3 \quad (\text{Run 1} + \text{Run 2}). \tag{18}$$

As expected the central value is higher, and now the prediction for the width of the Higgs boson in the SMEFT is consistent with the SM prediction. In addition, predictions can also be made for double Higgs boson production. The CMS Run-2 upper limit for double Higgs production is 19.2 times the SM prediction, at the 95% CL [75]. The upper limit we derive for double Higgs production in the SMEFT in the most general case is not competitive with the experimental upper limit, indicating that the experiments are currently sensitive to nonresonant double Higgs production. Explicit bounds on double Higgs production in the SMEFT in the general case are shown in Appendix C, Fig. 8 specifically. On the other hand, in specific scenarios tight bounds on double Higgs production can be derived. For example, setting c_6 to zero we find $\sigma_{\text{SMEFT}}(gg \to hh)/\sigma_{\text{SM}}(gg \to hh) \simeq 1.4 \pm 0.4$.

D. Electroweak baryogenesis

The trilinear Higgs coupling plays an important role in not only double Higgs production, but also in EW baryogenesis. To investigate the constraints on EW baryogenesis in the SMEFT we switch to a more common notation:

$$c_H = \frac{1}{2}\bar{c}_H, \qquad c_6 = -\frac{m_h^2}{2v^2}\bar{c}_6. \tag{19}$$

Assuming temperature dependence only in the Higgs mass parameter, requiring a first-order phase transition yields the analytic bound [80–83]

$$\frac{2}{3} < \bar{c}_6 < 2. \qquad (20)$$

Reference [84] went further and considered the viable parameter space for a first-order phase transition in the $\bar{c}_H - \bar{c}_6$ plane. This parameter space is bounded by our fits, and is shown in Fig. 4. A first-order phase transition occurs in the wedge bounded by the red lines [84]. The blue and green ellipses give the favored parameter space from the 12 parameter regularized fit with $\kappa = 1$, and the 10 parameter unregularized fit, respectively. Darker and lighter shading again correspond to $\Delta\chi^2 = 2.30$ and 6.18, respectively. The parameter space shaded red is the favored result of standard two parameter fit assuming only $\bar{c}_H$ and $\bar{c}_6$ are nonzero. Although the parameter space is constrained, EW baryogenesis in the SMEFT is still viable provided the cutoff scale is not too large. This is explicitly demonstrated by a regularized fit with $\kappa = 10$, shown by the purple ellipses, which is not consistent with a first-order phase transition at the one sigma level. A regularization parameter of 10 approximately corresponds to an effective scale of 800 GeV, reinforcing the notion that successful EW baryogenesis requires fairly low-scale BSM physics [80,85].
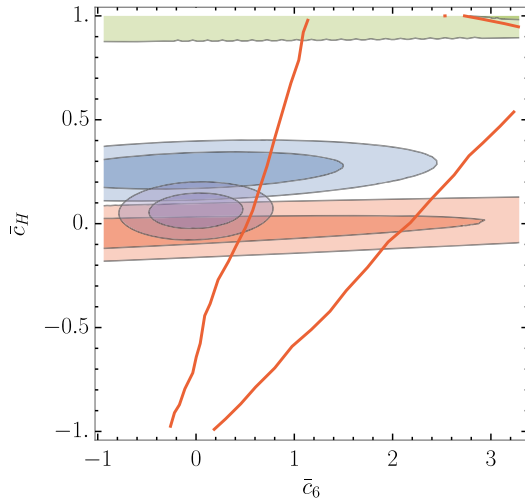

FIG. 4. Constraints on the parameter space relevant for EW baryogenesis in the SMEFT. A first-order phase transition occurs in the wedge bounded by the red lines [84], and viable parameter space still exists. The purple ellipses are the result of a regularized fit with $\kappa = 10$, which is not consistent with a first-order phase transition at the one sigma level. A regularization parameter of 10 approximately corresponds to an effective scale of 800 GeV, reinforcing the notion that successful EW baryogenesis requires fairly low-scale BSM physics [80,85].

TABLE I. Improvement in the global constraints by adding one hypothetical signal strength of 1.0 ± 0.1 to the fit for various Higgs observables. The improvement is quantified using the ratio of the GDPs of the fits with/without the hypothetical measurement.

Observable	GDP ratio	Observable	GDP ratio
$gg \to hh$	0.37	$Wh, h \to ZZ^*$	0.96
$h \to Z\gamma$	0.71	VBF, $h \to b\bar{b}$	0.98
$h \to c\bar{c}$	0.80	Γ_h	0.98
$h \to \mu^+\mu^-$	0.80	$Zh, h \to \tau^+\tau^-$	0.99
$tth, h \to ZZ^*$	0.93	$tth, h \to b\bar{b}$	0.99
$Zh, h \to ZZ^*$	0.94	$ggF, h \to b\bar{b}$	0.99

E. Future measurements

One may wonder which experimental measurements would improve the global constraints the most. To this end, we add to our fit one hypothetical future signal strength of 1.0 ± 0.1 for various Higgs boson observables, and see how this changes the fit. This is quantified using the global determinant parameter of Ref. [53]. The GDP is defined in our notation as

$$\text{GDP} = \left(\prod_{j \subseteq k} \sigma_j^2 \right)^{\frac{1}{m}}, \qquad (21)$$

where m is the total number of eigenvalues considered, which need not be all 12 in general. In particular, we report the ratio of the GDP with this additional hypothetical measurement to the GDP of the 55 measurement fit described previously. We confirm that ratios of GDPs do not depend on the normalization of the operators. The list of all observables that improve the constraints by themselves are given in Tab. I. We use the unregularized, 10 parameter fit in computing these GDPs. Some of the observables, such as double Higgs production, are obvious candidates, but others are less well known to be important for future Higgs coupling constraints.

V. SUMMARY

In this work, we performed an SMEFT parameter fit with an emphasis on a statistical technique aimed at tackling the issue of the large number of parameters. The technique we used is a regularized linear regression, where a positive definite function of the parameters of interest is added to the usual cost function. This prevents the fit from falling into an overfit solution, and, in principle, allows information to be obtained about any number of parameters. A cross-validation was performed to try to determine the optimal value of the regularization parameter to use. The cross-validation instead selected the SM as the best fit, so we contented ourselves to performing regularized fits with multiple choices for the regularization parameter, and examined how much regulator dependence various

quantities had. As proof of principle of this technique we applied it to fitting Higgs boson signal strengths, including the latest Run-2 results. We emphasized presenting results in terms of the eigensystem of the covariance matrix of the least squares estimators as it has a degree model-independent to it. We showed the SMEFT predicts the total width of the Higgs boson, which is not yet directly measured, to be consistent with the SM prediction, and that the ATLAS and CMS experiments at the LHC are currently sensitive to nonresonant double Higgs boson production. We derived constraints on the viable parameter space for EW baryogenesis in the SMEFT, and reinforce the notion that a first-order phase transition requires fairly low-scale BSM physics. We studied which future experimental measurements would improve the global constraints on the Higgs sector of the SMEFT the most. This is quantified using ratios of the GDP, which has a natural interpretation in terms of the eigensystem of the covariance matrix. We expect this technique to be of use to practitioners of both bottom-up and top-down approaches to EFTs.

APPENDIX A: EXPERIMENTAL RESULTS

The experimental results used in this analysis from Run 1 of the LHC are given in Table II. Similarly, the ATLAS and CMS Run-2 results can be found in Tables III and IV, respectively.

TABLE II. Run-1 experimental results used in this work. The $Z\gamma$ result is from ATLAS [77]. CMS does not provide a signal strength for $h \to Z\gamma$ although their 95% CL upper limit is stronger [76] than the ATLAS Run-1 result. All other results are taken from the combined ATLAS + CMS analysis of Ref. [1] with correlations taken into account.

Production	Decay	Signal strength	Production	Decay	Signal strength
ggF	$\gamma\gamma$	$1.10^{+0.23}_{-0.22}$	Wh	bb	1.0 ± 0.5
ggF	ZZ	$1.13^{+0.34}_{-0.31}$	Zh	$\gamma\gamma$	$0.5^{+3.0}_{-2.5}$
ggF	WW	0.84 ± 0.17	Zh	WW	$5.9^{+2.6}_{-2.2}$
ggF	$\tau\tau$	1.0 ± 0.6	Zh	$\tau\tau$	$2.2^{+2.2}_{-1.8}$
VBF	$\gamma\gamma$	1.3 ± 0.5	Zh	bb	0.4 ± 0.4
VBF	ZZ	$0.1^{+1.1}_{-0.6}$	tth	$\gamma\gamma$	$2.2^{+1.6}_{-1.3}$
VBF	WW	1.2 ± 0.4	tth	WW	$5.0^{+1.8}_{-1.7}$
VBF	$\tau\tau$	1.3 ± 0.4	tth	$\tau\tau$	$-1.9^{+3.7}_{-3.3}$
Wh	$\gamma\gamma$	$0.5^{+1.3}_{-1.2}$	tth	bb	1.1 ± 1.0
Wh	WW	$1.6^{+1.2}_{-1.0}$	pp	$\mu\mu$	0.1 ± 2.5
Wh	$\tau\tau$	-1.4 ± 1.4	pp	$Z\gamma$	$2.7^{+4.6}_{-4.5}$

TABLE III. Run-2 ATLAS results used in this work. We estimate the signal strength for $h \to Z\gamma$ from Ref. [78], which states that the upper limit for this process is 6.6 times the SM rate at 95% CL and that the significance of the measurement is 0.5σ.

Production	Decay	Signal strength	Reference	Production	Decay	Signal strength	Reference
pp	$\mu\mu$	-0.1 ± 1.4	[86]	ggF	ZZ	$1.11^{+0.25}_{-0.22}$	[87]
Wh	bb	$1.35^{+0.68}_{-0.59}$	[88]	VBF	ZZ	$4.0^{+1.8}_{-1.5}$	[87]
Zh	bb	$1.12^{+0.50}_{-0.45}$	[88]	VBF	WW	$1.7^{+1.2}_{-0.9}$	[89]
ggF	$\gamma\gamma$	$0.80^{+0.19}_{-0.18}$	[90]	Wh	WW	$3.2^{+4.4}_{-4.2}$	[89]
VBF	$\gamma\gamma$	2.1 ± 0.6	[90]	tth	$2\ell 0\tau_h$	$4.0^{+2.1}_{-1.7}$	[91]
Vh	$\gamma\gamma$	$0.7^{+0.9}_{-0.8}$	[90]	tth	$2\ell 1\tau_h$	$6.2^{+3.6}_{-2.7}$	[91]
tth	$\gamma\gamma$	0.5 ± 0.6	[90]	tth	3ℓ	$0.5^{+1.7}_{-1.6}$	[91]
pp	$Z\gamma$	1.3 ± 2.6	[78]				

TABLE IV. Run-2 CMS results used in this work.

Production	Decay	Signal strength	Reference	Production	Decay	Signal strength	Reference
ggF	ZZ	1.20 ± 0.20	[92]	ggF	$\gamma\gamma$	$1.11^{+0.19}_{-0.18}$	[93]
0-jet	$\tau\tau$	0.84 ± 0.89	[94]	VBF	$\gamma\gamma$	$0.5^{+0.6}_{-0.5}$	[93]
VBF	$\tau\tau$	$1.11^{+0.34}_{-0.35}$	[94]	Vh	$\gamma\gamma$	$2.3^{+1.1}_{-1.0}$	[93]
tth	2ℓ	$1.7^{+0.6}_{-0.5}$	[95]	tth	$\gamma\gamma$	$2.2^{+0.9}_{-0.8}$	[93]
tth	3ℓ	$1.0^{+0.8}_{-0.7}$	[95]	0-jet	WW	$0.9^{+0.4}_{-0.3}$	[96]
tth	4ℓ	$0.9^{+2.3}_{-1.6}$	[95]	VBF	WW	1.4 ± 0.8	[96]
tth	$\tau\tau$	$0.72^{+0.62}_{-0.53}$	[97]	Wh	WW	-1.4 ± 1.5	[96]
Wh	bb	1.7 ± 0.7	[98]	Vh	WW	$2.1^{+2.3}_{-2.2}$	[96]
Zh	bb	0.9 ± 0.5	[98]	tt	bb	$-0.19^{+0.82}_{-0.81}$	[99]

APPENDIX B: EIGENVECTORS

The eigenvectors for the regularized fit with $\kappa = 1$ are

$$W_1 \simeq 0.99c_B + 0.09c_{HW} + 0.09c_T - 0.08c_W + 0.05c_{HB},$$
$$W_2 \simeq 0.67c_{HW} - 0.56c_W - 0.36c_B + 0.33c_{HB} - 0.02c_T,$$
$$W_3 \simeq 0.99c_6 - 0.13c_t + 0.05c_W + 0.03c_{HW} + 0.02c_{HB} - 0.01c_H,$$
$$W_4 \simeq 0.67c_W + 0.45c_{HW} + 0.38c_H - 0.38c_t + 0.24c_{HB} - 0.09c_6,$$
$$W_5 \simeq 0.76c_t + 0.40c_W - 0.32c_H + 0.24c_{HW} + 0.22c_T + 0.20c_{HB} + 0.06c_6 - 0.02c_B,$$
$$W_6 \simeq 0.78c_H + 0.51c_t - 0.32c_T - 0.10c_W + 0.08c_6 - 0.07c_{HB} - 0.05c_{HW} + 0.03c_b + 0.03c_B,$$
$$W_7 \simeq 0.87c_{HB} - 0.48c_{HW} + 0.09c_H + 0.09c_T - 0.06c_W - 0.03c_t + 0.01c_b,$$
$$W_8 \simeq 0.91c_T + 0.34c_H - 0.16c_{HB} - 0.11c_W - 0.08c_B - 0.03c_{HW} + 0.03c_b + 0.01c_6,$$
$$W_9 \simeq 0.97c_b + 0.24c_\tau - 0.07c_g + 0.04c_\gamma - 0.03c_H + 0.02c_W + 0.01c_{HW} - 0.01c_t - 0.01c_T,$$
$$W_{10} \simeq 0.97c_\tau - 0.24c_b + 0.05c_g,$$
$$W_{11} \simeq 0.93c_g + 0.35c_\gamma + 0.06c_b - 0.03c_\tau,$$
$$W_{12} \simeq 0.93c_\gamma - 0.35c_g - 0.07c_b. \tag{B1}$$

All 12 parameters contribute to each eigenvector, but only percent level or higher contributions are shown. The blind directions are (two linear combinations of) eigenvectors 1 and 2.

APPENDIX C: ADDITIONAL PLOTS

Another way to visualize the blind directions in the fit is to use the pseudoinverse without regularization to invert the Fisher information. Given a matrix A, its pseudoinverse, A^p, is defined as

$$AA^pA = A, \tag{C1}$$

as opposed to $A^{-1}A = \mathbb{1}$ for the case of the (genuine) inverse. In this case, the eigenvalues of the blind directions

are zero, and there is no regulator dependence in any of the other eigenvalues. This is shown in Fig. 5.

The combination of Run 1 and Run 2 works exactly as expected, adding more data tightens the resulting bounds; see Fig. 6. Many of the bounds are now driven by the Run-2 results, but the contributions from the Run-1 measurements are still important. Furthermore, Run 2 shows more sensitivity to c_6 as a result of its improved tth measurements.

There is better agreement between different scenarios, involving UV assumption or choice of regularization parameter, in the eigenvectors basis even when the
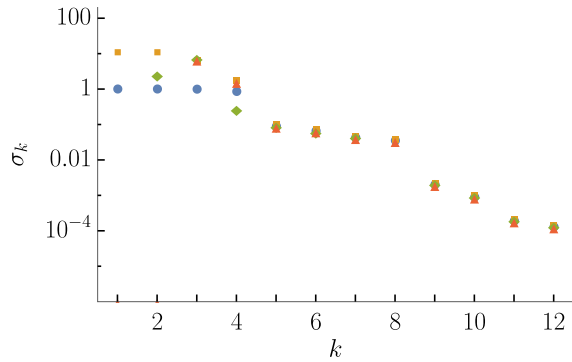
FIG. 5. The same as Fig. 2, but with the results of the pseudoinverse fit included as the red triangles. The eigenvalues of the blind directions are zero when using the pseudoinverse with regularization, and none of the other eigenvalues have regulator dependence.
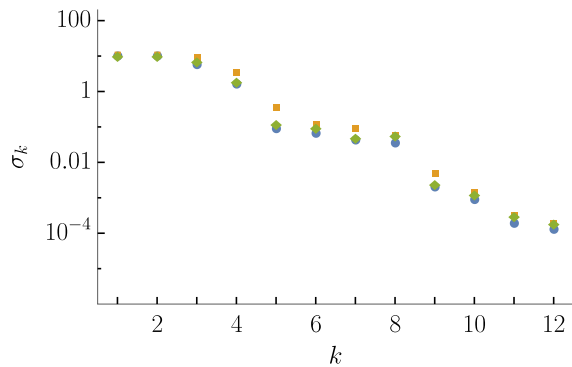

FIG. 6. Results of regularized fits with $\kappa = 10^{-2}$ using only Run-1 results (orange squares), only Run-2 results (green diamonds), and both runs (blue circles). Many of the bounds are now driven by the Run-2 results, but the contributions from the Run-1 measurements are still important.

FIG. 8. Bounds on double Higgs production, $\mu(pp \to hh) = \sigma(pp \to hh)/\sigma_{SM}(pp \to hh)$, when all 12 parameters are allowed to be nonzero as a function of the regularization parameter, κ. The blue curve gives the best fit values, and the darker and lighter shaded regions are allowed at 1 and 2σ, respectively. The dashed line is the experimental upper limit from CMS [75].

correlation between the eigenvectors in the different scenarios is not perfect. This is shown in Fig. 7, which is the same as Fig. 3 but in the $c_g - c_t$ plane (left), and the $W_{11} - W_5$ plane (right).

Bounds on double Higgs production, $\mu(pp \to hh) = \sigma(pp \to hh)/\sigma_{SM}(pp \to hh)$, when all 12 parameters are allowed to be nonzero as a function of the regularization parameter κ are shown in Fig. 8. The blue line gives the best fit values, and the darker and lighter shaded regions are allowed at 1 and 2σ, respectively. The dashed line is the experimental upper limit from CMS [75].
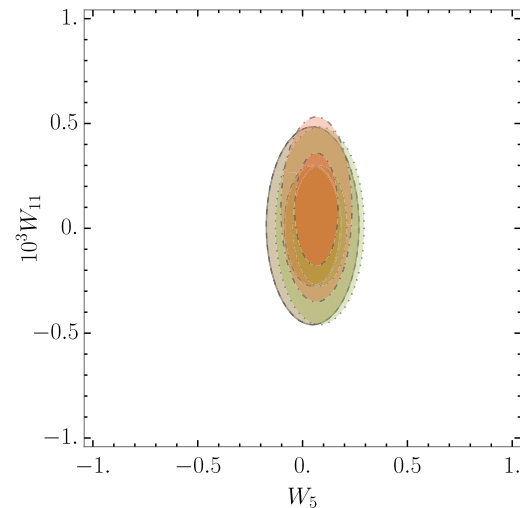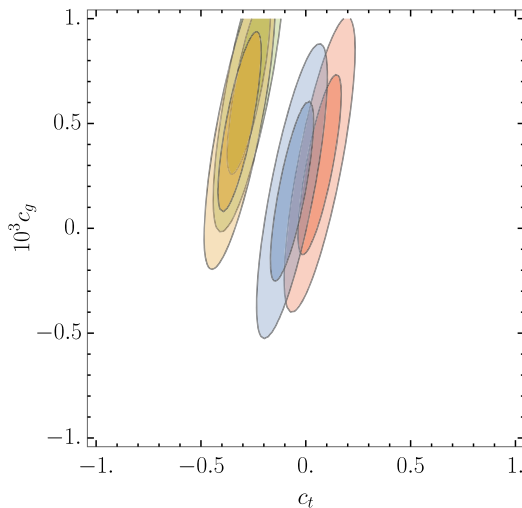

FIG. 7. The same as Fig. 3 but in the $c_g - c_t$ plane (left), and the $W_{11} - W_5$ plane (right). The correlation between the eigenvectors in the different scenarios is not perfect. Nevertheless the agreement between the different scenarios is better in the eigenvector basis.
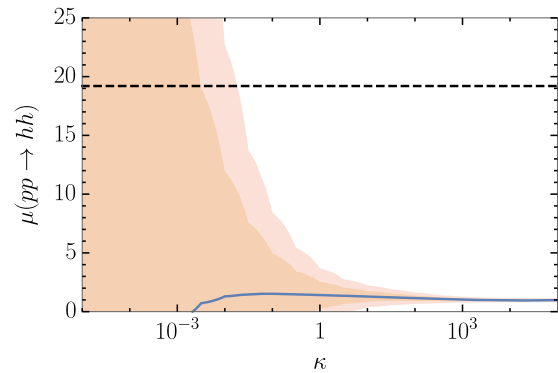
[1] G. Aad *et al.* (ATLAS, CMS Collaborations), Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV, J. High Energy Phys. 08 (2016) 045.

[2] C. Patrignani *et al.* (Particle Data Group Collaboration), Review of particle physics, Chin. Phys. C **40**, 100001 (2016).

[3] M. Battaglieri *et al.*, US Cosmic Visions: New Ideas in Dark Matter 2017: Community Report, arXiv:1707.04591.

[4] J.F. Gunion and H. E. Haber, The *CP* conserving two Higgs doublet model: The approach to the decoupling limit, Phys. Rev. D **67**, 075019 (2003).

[5] I. Brivio and M. Trott, The Standard Model as an Effective Field Theory, arXiv:1706.08945.

[6] A. Dedes, W. Materkowska, M. Paraskevas, J. Rosiek, and K. Suxho, Feynman rules for the standard model effective field theory in R_ξ-gauges, J. High Energy Phys. 06 (2017) 143.

[7] R. Alonso, E. E. Jenkins, A. V. Manohar, and M. Trott, Renormalization group evolution of the standard model dimension six operators III: Gauge coupling dependence and phenomenology, J. High Energy Phys. 04 (2014) 159.

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

[9] Z. Han and W. Skiba, Effective theory analysis of precision electroweak data, Phys. Rev. D **71**, 075009 (2005).

[10] T. Corbett, O. J. P. Eboli, J. Gonzalez-Fraile, and M. C. Gonzalez-Garcia, Robust determination of the Higgs couplings: Power to the data, Phys. Rev. D **87**, 015022 (2013).

[11] T. Corbett, O. J. P. Eboli, J. Gonzalez-Fraile, and M. C. Gonzalez-Garcia, Determining Triple Gauge Boson Couplings from Higgs Data, Phys. Rev. Lett. **111**, 011801 (2013).

[12] B. Dumont, S. Fichet, and G. von Gersdorff, A Bayesian view of the Higgs sector with higher dimensional operators, J. High Energy Phys. 07 (2013) 065.

[13] B. Grinstein, C. W. Murphy, and D. Pirtskhalava, Searching for new physics in the three-body decays of the Higgs-like particle, J. High Energy Phys. 10 (2013) 077.

[14] M. Ciuchini, E. Franco, S. Mishima, and L. Silvestrini, Electroweak precision observables, new physics and the nature of a 126 GeV Higgs boson, J. High Energy Phys. 08 (2013) 106.

[15] A. Pomarol and F. Riva, Towards the ultimate SM fit to close in on Higgs physics, J. High Energy Phys. 01 (2014) 151.

[16] J. Ellis, V. Sanz, and T. You, Complete Higgs sector constraints on dimension-6 operators, J. High Energy Phys. 07 (2014) 036.

[17] J. D. Wells and Z. Zhang, Precision electroweak analysis after the Higgs boson discovery, Phys. Rev. D **90**, 033006 (2014).

[18] M. Baak, J. Cuth, J. Haller, A. Hoecker, R. Kogler, K. Monig, M. Schott, and J. Stelzer (Gfitter Group Collaboration), The global electroweak fit at NNLO and prospects for the LHC and ILC, Eur. Phys. J. C **74**, 3046 (2014).

[19] M. Trott, On the consistent use of constructed observables, J. High Energy Phys. 02 (2015) 046.

[20] M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina, and L. Silvestrini, Update of the electroweak precision fit, interplay with Higgs-boson signal strengths and model-independent constraints on new physics, Nucl. Part. Phys. Proc. **273–275**, 2219 (2016).

[21] J. Ellis, V. Sanz, and T. You, The effective standard model after LHC Run I, J. High Energy Phys. 03 (2015) 157.

[22] A. Falkowski and F. Riva, Model-independent precision constraints on dimension-6 operators, J. High Energy Phys. 02 (2015) 039.

[23] G. Durieux, F. Maltoni, and C. Zhang, Global approach to top-quark flavor-changing interactions, Phys. Rev. D **91**, 074017 (2015).

[24] B. Henning, X. Lu, and H. Murayama, How to use the Standard Model effective field theory, J. High Energy Phys. 01 (2016) 023.

[25] A. A. Petrov, S. Pokorski, J. D. Wells, and Z. Zhang, Role of low-energy observables in precision Higgs boson analyses, Phys. Rev. D **91**, 073001 (2015).

[26] L. Berthier and M. Trott, Towards consistent Electroweak Precision Data constraints in the SMEFT, J. High Energy Phys. 05 (2015) 024.

[27] S. Banerjee, T. Mandal, B. Mellado, and B. Mukhopadhyaya, Cornering dimension-6 *HVV* interactions at high luminosity LHC: the role of event ratios, J. High Energy Phys. 09 (2015) 057.

[28] T. Corbett, O. J. P. Eboli, D. Goncalves, J. Gonzalez-Fraile, T. Plehn, and M. Rauch, The Higgs legacy of the LHC Run I, J. High Energy Phys. 08 (2015) 156.

[29] J. de Blas, M. Chala, and J. Santiago, Renormalization group constraints on new top interactions from electroweak precision data, J. High Energy Phys. 09 (2015) 189.

[30] J. D. Wells and Z. Zhang, Status and prospects of precision analyses with $e^+e^- \to W^+W^-$, Phys. Rev. D **93**, 034001 (2016).

[31] A. Falkowski, M. Gonzalez-Alonso, A. Greljo, and D. Marzocca, Global Constraints on Anomalous Triple Gauge Couplings in Effective Field Theory Approach, Phys. Rev. Lett. **116**, 011801 (2016).

[32] L. Berthier and M. Trott, Consistent constraints on the standard model effective field theory, J. High Energy Phys. 02 (2016) 069.

[33] C. Englert, R. Kogler, H. Schulz, and M. Spannowsky, Higgs coupling measurements at the LHC, Eur. Phys. J. C **76**, 393 (2016).

[34] A. Butter, O. J. P. Eboli, J. Gonzalez-Fraile, M. C. Gonzalez-Garcia, T. Plehn, and M. Rauch, The gauge-Higgs legacy of the LHC Run I, J. High Energy Phys. 07 (2016) 152.

[35] V. Cirigliano, W. Dekens, J. de Vries, and E. Mereghetti, Constraining the top-Higgs sector of the standard model effective field theory, Phys. Rev. D **94**, 034031 (2016).

[36] L. Berthier, M. Bjorn, and M. Trott, Incorporating doubly resonant $W^\pm$ data in a global fit of SMEFT parameters to lift flat directions, J. High Energy Phys. 09 (2016) 157.

[37] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina, and L. Silvestrini, Electroweak precision observables and Higgs-boson signal strengths in the Standard Model and beyond: present and future, J. High Energy Phys. 12 (2016) 135.

[38] Z. Zhang, Time to Go Beyond Triple-Gauge-Boson-Coupling Interpretation of *W* Pair Production, Phys. Rev. Lett. **118,** 011803 (2017).

[39] I. Brivio and M. Trott, Scheming in the SMEFT… and a reparameterization invariance!, J. High Energy Phys. 07 (2017) 148.

[40] H. Khanpour and M. M. Najafabadi, Constraining Higgs boson effective couplings at electron-positron colliders, Phys. Rev. D **95,** 055026 (2017).

[41] H. Khanpour, S. Khatibi, and M. M. Najafabadi, Probing Higgs boson couplings in $H + \gamma$ production at the LHC, Phys. Lett. B **773,** 462 (2017).

[42] S. Alioli, V. Cirigliano, W. Dekens, J. de Vries, and E. Mereghetti, Right-handed charged currents in the era of the Large Hadron Collider, J. High Energy Phys. 05 (2017) 086.

[43] S. Di Vita, C. Grojean, G. Panico, M. Riembau, and T. Vantalon, A global view on the Higgs self-coupling, J. High Energy Phys. 09 (2017) 069.

[44] S. Dawson and C. W. Murphy, Standard model EFT and extended scalar sectors, Phys. Rev. D **96,** 015041 (2017).

[45] S. Jana and S. Nandi, New physics scale from Higgs observables with effective dimension-6 operators, arXiv:1710.00619.

[46] H. B. Prosper and L. Lyons, eds., *Proceedings of PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva, Switzerland 2011* (CERN, 2011), http://inspirehep.net/record/1087459/files/CERN-2011-006.pdf.

[47] *Data Analysis in High Energy Physics*, edited by O. Behnke, K. Kroninger, T. Schorner-Sadenius, and G. Schott. (Wiley, New York, 2013).

[48] A. Ng, Machine learning, https://www.coursera.org/learn/machine-learning.

[49] A. Géron, *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (O'Reilly Media, Sebastopol, CA, 2017).

[50] G. P. Lepage, B. Clark, C. T. H. Davies, K. Hornbostel, P. B. Mackenzie, C. Morningstar, and H. Trottier, Constrained curve fitting, Nucl. Phys. B, Proc. Suppl. **106,** 12 (2002).

[51] A. S. Meyer, R. J. Hill, A. S. Kronfeld, R. Li, and J. N. Simone, Calculation of the Nucleon Axial Form Factor Using Staggered Lattice QCD, in *Proc. Sci.*, LATTICE2016 (2016) 179 [arXiv:1610.04593].

[52] M. T. Hansen, H. B. Meyer, and D. Robaina, From deep inelastic scattering to heavy-flavor semileptonic decays: Total rates into multihadron final states from lattice QCD, Phys. Rev. D **96,** 094513 (2017).

[53] G. Durieux, C. Grojean, J. Gu, and K. Wang, The leptonic future of the Higgs, J. High Energy Phys. 09 (2017) 014.

[54] R. Barbieri, D. Buttazzo, F. Sala, and D. M. Straub, Flavour physics from an approximate $U(2)^3$ symmetry, J. High Energy Phys. 07 (2012) 181.

[55] R. Barbieri, G. Isidori, A. Pattori, and F. Senia, Anomalies in *B*-decays and $U(2)$ flavour symmetry, Eur. Phys. J. C **76,** 67 (2016).

[56] R. Barbieri, C. W. Murphy, and F. Senia, B-decay anomalies in a composite leptoquark model, Eur. Phys. J. C **77,** 8 (2017).

[57] M. Bordone, G. Isidori, and S. Trifinopoulos, Semileptonic *B*-physics anomalies: A general EFT analysis within $U(2)^n$ flavor symmetry, Phys. Rev. D **96,** 015038 (2017).

[58] D. Buttazzo, A. Greljo, G. Isidori, and D. Marzocca, B-physics anomalies: a guide to combined explanations, J. High Energy Phys. 11 (2017) 044.

[59] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, Dimension-six terms in the standard model Lagrangian, J. High Energy Phys. 10 (2010) 085.

[60] I. Brivio, Y. Jiang, and M. Trott, The SMEFTsim package, theory and tools, arXiv:1709.06492.

[61] A. V. Manohar and M. B. Wise, Modifications to the properties of the Higgs boson, Phys. Lett. B **636,** 107 (2006).

[62] R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner, and M. Spira, eHDECAY: an Implementation of the Higgs Effective Lagrangian into HDECAY, Comput. Phys. Commun. **185,** 3412 (2014).

[63] G. Degrassi, P. P. Giardino, F. Maltoni, and D. Pagani, Probing the Higgs self coupling via single Higgs production at the LHC, J. High Energy Phys. 12 (2016) 080.

[64] D. de Florian *et al.* (LHC Higgs Cross Section Working Group Collaboration), Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector, arXiv:1610.07922.

[65] A. Djouadi, The anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model, Phys. Rep. **457,** 1 (2008).

[66] A. Azatov, R. Contino, G. Panico, and M. Son, Effective field theory analysis of double Higgs boson production via gluon fusion, Phys. Rev. D **92,** 035001 (2015).

[67] F. Beaujean, C. Bobeth, D. van Dyk, and C. Wacker, Bayesian fit of exclusive $b \to s\bar{\ell}\ell$ decays: The standard model operator basis, J. High Energy Phys. 08 (2012) 030.

[68] F. Beaujean, C. Bobeth, and D. van Dyk, Comprehensive Bayesian analysis of rare (semi)leptonic and radiative *B* decays, Eur. Phys. J. C **74,** 2897 (2014); Erratum, Eur. Phys. J. C **74,** 3179 (2014).

[69] S. Meinel and D. van Dyk, Using $\Lambda_b \to \Lambda\mu^+\mu^-$ data within a Bayesian analysis of $|\Delta B| = |\Delta S| = 1$ decays, Phys. Rev. D **94,** 013007 (2016).

[70] M. E. Peskin and T. Takeuchi, Estimation of oblique electroweak corrections, Phys. Rev. D **46,** 381 (1992).

[71] G. Sanchez-Colon and J. Wudka, Effective operator contributions to the oblique parameters, Phys. Lett. B **432,** 383 (1998).

[72] W. Kilian and J. Reuter, The low-energy structure of little Higgs models, Phys. Rev. D **70,** 015004 (2004).

[73] C. Grojean, W. Skiba, and J. Terning, Disguising the oblique parameters, Phys. Rev. D **73,** 075008 (2006).

[74] T. A. Collaboration (ATLAS Collaboration)), CERN Technical Report No. ATLAS-CONF-2016-049, 2016.

[75] C. Collaboration (CMS Collaboration), CERN Technical Report No. CMS-PAS-HIG-17-008, 2017.

[76] S. Chatrchyan *et al.* (CMS Collaboration), Search for a Higgs boson decaying into a Z and a photon in pp collisions at sqrt(s) = 7 and 8 TeV, Phys. Lett. B **726,** 587 (2013).

[77] G. Aad *et al.* (ATLAS Collaboration), Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment, Eur. Phys. J. C **76,** 6 (2016).

[78] M. Aaboud *et al.* (ATLAS Collaboration), Searches for the $Z\gamma$ decay mode of the Higgs boson and for new high-mass resonances in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, J. High Energy Phys. 10 (2017) 112.

[79] A. Manohar and H. Georgi, Chiral quarks and the non-relativistic quark model, Nucl. Phys. B **B234**, 189 (1984).

[80] C. Grojean, G. Servant, and J. D. Wells, First-order electroweak phase transition in the standard model with a low cutoff, Phys. Rev. D **71**, 036001 (2005).

[81] F. P. Huang, P.-H. Gu, P.-F. Yin, Z.-H. Yu, and X. Zhang, Testing the electroweak phase transition and electroweak baryogenesis at the LHC and a circular electron-positron collider, Phys. Rev. D **93**, 103515 (2016).

[82] P. Huang, A. Joglekar, B. Li, and C. E. M. Wagner, Probing the electroweak phase transition at the LHC, Phys. Rev. D **93**, 055049 (2016).

[83] B. Jain, S. J. Lee, and M. Son, On the Validity of the Effective Potential and the Precision of Higgs Self Couplings, arXiv:1709.03232.

[84] P. H. Damgaard, A. Haarr, D. O'Connell, and A. Tranberg, Effective field theory and electroweak baryogenesis in the singlet-extended standard model, J. High Energy Phys. 02 (2016) 107.

[85] A. Kobakhidze, L. Wu, and J. Yue, Electroweak baryogenesis with anomalous Higgs couplings, J. High Energy Phys. 04 (2016) 011.

[86] M. Aaboud *et al.* (ATLAS Collaboration), Search for the Dimuon Decay of the Higgs Boson in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector, Phys. Rev. Lett. **119**, 051802 (2017).

[87] T. A. Collaboration (ATLAS Collaboration), CERN Technical Report No. ATLAS-CONF-2017-043, 2017.

[88] M. Aaboud *et al.* (ATLAS Collaboration), Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector, J. High Energy Phys. 12 (2017) 024.

[89] T. A. Collaboration (ATLAS Collaboration), CERN Technical Report No. ATLAS-CONF-2016-112, 2016.

[90] T. A. Collaboration (ATLAS Collaboration), CERN Technical Report No. ATLAS-CONF-2017-045, 2017.

[91] T. A. Collaboration (ATLAS Collaboration), CERN Technical Report No. ATLAS-CONF-2016-058, 2016.

[92] A. M. Sirunyan *et al.* (CMS Collaboration), Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at sqrt(s) = 13 TeV, J. High Energy Phys. 11 (2017) 047.

[93] C. Collaboration (CMS Collaboration), CERN Technical Report No. CMS-PAS-HIG-16-040, 2016.

[94] A. M. Sirunyan *et al.* (CMS Collaboration), Observation of the Higgs boson decay to a pair of tau leptons, arXiv:1708.00373.

[95] C. Collaboration (CMS Collaboration), CERN Technical Report No. CMS-PAS-HIG-17-004, 2017.

[96] C. Collaboration (CMS Collaboration), CERN Technical Report No. CMS-PAS-HIG-16-021, 2016.

[97] C. Collaboration (CMS Collaboration), CERN Technical Report No. CMS-PAS-HIG-17-003, 2017.

[98] C. Collaboration (CMS Collaboration), CERN Technical Report No. CMS-PAS-HIG-16-044, 2016.

[99] C. Collaboration (CMS Collaboration), CERN Technical Report No. CMS-PAS-HIG-16-038, 2016.