New method for fitting coefficients in standard model effective theory

Geoffrey T. Bodwin\textsuperscript{1,*} and Hee Sok Chung\textsuperscript{2,3,†}

\textsuperscript{1}High Energy Physics Division, Argonne National Laboratory, Argonne, Illinois 60439, USA
\textsuperscript{2}Physik-Department, Technische Universität München, James-Franck-Str. 1, 85748 Garching, Germany
\textsuperscript{3}Excellence Cluster ORIGINS, Boltzmannstrasse 2, D-85748 Garching, Germany

(Dated: December 23, 2019)

Abstract

We address the issue that certain directions in the space of standard-model-effective-theory operator coefficients may be poorly constrained by experimental observables. We argue that this issue is best dealt with by making the use of a principal-component analysis, and we present an efficient method for carrying out that analysis that is based on singular-value decomposition. We demonstrate the method by applying it to the process of top-quark decays to a $b$ quark and a $W$ boson.
I. INTRODUCTION

In recent years, standard model effective theory (SMEFT) [1–4] has been a focus of activity in both the theoretical and experimental particle-physics communities. SMEFT has been advocated as a means to quantify systematically deviations of the global set of experimental measurements from the predictions of the standard model. To this end, a number of efforts have been undertaken to perform global fits of the Wilson coefficients of SMEFT operators [5–10].

Several difficulties can arise in using data to constrain coefficients in SMEFT. First, there are many operators in SMEFT (59 in dimension-6) and potentially many data points that could be used in fitting the coefficients of these operators. Second, theoretical expressions for the SMEFT contributions to a given set of experimental observables can contain “flat directions” (or nearly flat directions) in the space of SMEFT coefficients, that is, directions for which the observables are insensitive to the values of SMEFT coefficients. Third, for analyses involving a limited sector of observables, there may be fewer observables than SMEFT coefficients. This situation will necessarily result in the existence of exactly flat directions.

All of these difficulties can pose computational problems in fitting SMEFT coefficients to data. A global fit may be computationally challenging because standard methods for carrying out fits may bog down when the number of observables and coefficients is large. When the number of coefficients is greater than the number of observables and/or there are flat directions, methods of fitting that minimize $\chi^2$ numerically may not converge reliably. In addition to these technical issues, there is also an issue of principle: When there are more coefficients than observables and/or flat directions, the uncertainties in the coefficients can be highly correlated, and presentation of bounds in terms of ranges on the individual SMEFT coefficients may be very misleading.

In this paper, we advocate addressing this latter issue of principle by carrying out principal-component analyses of the SMEFT coefficients. We address the technical issues that arise in carrying out fits of the SMEFT coefficients in a reliable and computationally efficient manner by making use of the method of singular-value decomposition (SVD). As we will show, the SVD method is also a very convenient one for carrying out the principal-component analysis.
The method that we present is based on a Gaussian uncertainty analysis. While this, of course, is not completely general, it should prove to be adequate at least for initial exploratory studies of the bounds on SMEFT coefficients. The SVD method also requires that the observables depend linearly on the SMEFT coefficients. This is the case in a computation at leading order in the effective-field-theory expansion. As we will describe later, the method can also be used iteratively to carry out a principal-component analysis of the SMEFT coefficients in the case in which higher-order contributions in the effective-field-theory expansion are considered—provided that the expansion itself converges.

We demonstrate the principal-component/SVD method by applying it to a restricted class of observables that appear in decays of the top quark to a $b$ quark and a $W$ boson.\footnote{We note that global fits of SMEFT coefficients in the top-quark sector have been carried out in Refs.\cite{11–14}.} This example allows us to demonstrate the use of the SVD method to deal with the difficulties of flat directions. While we have not specifically demonstrated the utility of the the principal-component/SVD method for a situation in which there is a large number of observables and a large number of SMEFT coefficients, we are confident that it would work reliably and efficiently in such a situation because of experience with an application of SVD, in a different context, that involved the fitting of thousands of data points with hundreds of coefficients\cite{15}.

The remainder of this paper is organized as follows. In Sec.\ref{sec:methods}, we outline the basics of SVD and present a method for using SVD to carry out principal-component analyses of the SMEFT coefficients. Section\ref{sec:results} contains an illustration of the use of the principal-component/SVD analysis in top-quark decays to a $b$ quark and a $W$ boson. Here, we present examples of fits involving flat directions and various numbers of SMEFT coefficients, and we contrast the results from the principal-component/SVD approach with those from traditional fitting approaches. In Sec.\ref{sec:discussion}, we discuss the extension of the principal-component/SVD approach to situations in which the theoretical expressions for the observables depend nonlinearly on the SMEFT coefficients. Finally, in Sec.\ref{sec:conclusions} we summarize our results.
II. SINGULAR-VALUE DECOMPOSITION METHOD

A. Singular-value decomposition

The singular-value-decomposition theorem states that an \( m \times n \) matrix \( M \) that contains either real or complex entries can always be decomposed as \[ M = U W V^\dagger, \] (1)

where \( U \) is an \( m \times m \) unitary matrix, \( V \) is an \( n \times n \) unitary matrix, \( W \) is a \( m \times n \) diagonal matrix with nonnegative real numbers on the diagonals, and \( \dagger \) denotes the hermitian conjugate (complex conjugate transpose). The diagonal entries of \( W \) are called the singular values. The matrix \( W \) is unique, but the matrices \( U \) and \( V \) are not. If \( M \) is a square matrix, then \( U \) and \( V \) are unique, up to phases that multiply each row of \( V^\dagger \) and corresponding inverse phases that multiply each column of \( U \). If \( M \) is not a square matrix, additional ambiguities in \( U \) and \( V \) can arise. Efficient computer codes exist for carrying out the SVD decomposition of large matrices numerically. See, for example, Refs. [17–20].

SVD has the important property that it can be used to solve the linear least-squares problem, as we will now explain. Suppose that \( M \) is an \( m \times n \) matrix, \( C \) is an \( n \)-dimensional column vector, and \( O \) is an \( m \)-dimensional column vector. Further suppose that we wish to minimize

\[ (MC - O)^2 \] (2)

with respect to the elements of \( C \) (coefficients). Here, the square denotes the inner product of \( MC - O \) with itself. The solution of this problem is given by \[ \overline{C} = VW^{-1}U^\dagger O, \] (3)

where \( U, V, \) and \( W \) are the matrices that appear in the SVD decomposition of \( M \) [Eq. (1)] and \( W^{-1} \) is the Moore-Penrose pseudoinverse of \( W \) [21–24], which is obtained by replacing the nonzero elements of \( W \) with their inverses, while leaving the zero elements unchanged.\(^2\)

Note that the solution in Eq. (3) exists even if the matrix \( M \) is non-invertible. The solution exists, for example, if \( M \) is not a square matrix or if \( M \) has a vanishing determinant.

\(^2\) Throughout this paper, when we refer to the pseudoinverse of a matrix, we mean the Moore-Penrose pseudoinverse, and we use the \(-1\) power of a noninvertible matrix to denote its Moore-Penrose pseudoinverse.
The matrix $V^\dagger$ takes the elements of $C$ from the original basis of coefficients to the principal-component basis. Each element of $P = V^\dagger C$ is one of the principal components. Owing to the phase ambiguity in each row in $V^\dagger$ in the SVD, each principal component is defined only up to an overall phase. By virtue of the unitarity of $V$, the principal components are orthogonal to each other. The best-fit values of the principal components are given by the elements of

$$\overline{P} = V^\dagger C = W^{-1}U^\dagger \mathcal{O},$$

where we have used Eq. (3). The fluctuation in $MC$ that is produced by a unit fluctuation in a principal component is given by the corresponding singular value in $W$. Hence, the uncertainties in the best-fit values of the principal components, $\Delta P$, are given by the inverses of the diagonal values of $W$. The principal components that are ill-constrained because of the existence of flat directions correspond to near-vanishing diagonal values of $W$.

A key property of a principal-component analysis is that, because $W$ is diagonal, the fluctuation in $MC$ that is produced by a fluctuation in a given principal component is independent of the fluctuations in the other principal components. That is, the uncertainties in the principal components are uncorrelated, which is equivalent to the statement that the covariance matrix of the coefficients is diagonal in the basis of the principal components. As we will see, this property will prove to be very useful in constraining the SMEFT coefficients. A related property of an SVD/principal-component analysis is that it sequesters linear combinations of coefficients that are ill-constrained because of the presence of flat directions in the coefficient space.

### B. Application of singular-value decomposition to the fitting of SMEFT coefficients

The fitting of the SMEFT coefficients is carried out by minimizing the $\chi^2$, which is defined by

$$\chi^2 = (O^{\text{SMEFT}} - O^{\text{exp}})^T (\sigma^2)^{-1} (O^{\text{SMEFT}} - O^{\text{exp}}),$$

where $O^{\text{exp}}$ is the $N_{\text{obs}}$-dimensional column vector of experimental observables, $O^{\text{SMEFT}}$ is the $N_{\text{obs}}$-dimensional column vector of theoretical predictions for the observables in the SMEFT, and $\sigma^2$ is the $N_{\text{obs}} \times N_{\text{obs}}$ covariance matrix of experimental and theoretical uncertainties. We decompose the theoretical predictions in the SMEFT into standard-model
(SM) contributions and beyond-the-standard-model (BSM) contributions as
\[
O^{\text{SMEFT}} = O^{\text{SM}} + O^{\text{BSM}},
\]
and re-write \( \chi^2 \) as
\[
\chi^2 = (O^{\text{diff}} - O^{\text{BSM}})^T (\sigma^2)^{-1} (O^{\text{diff}} - O^{\text{BSM}}),
\]
where
\[
O^{\text{diff}} = O^{\exp} - O^{\text{SM}}.
\]

Now we wish to put \( \chi^2 \) in Eq. (7) into the linear-least-squares form. First, since the covariance matrix is symmetric, we can diagonalize it:
\[
U_{\text{cov}}^{-1} \sigma^2 U_{\text{cov}} = \hat{\sigma}^2.
\]
We note that the diagonal matrix \( \hat{\sigma}^2 \) and the unitary matrix \( U_{\exp} \) can be found conveniently from the SVD decomposition of \( \sigma^2 \), although other diagonalization methods could also be used. Then, we can write \( \chi^2 \) as
\[
\chi^2 = (O^{\text{diff}} - O^{\text{BSM}})^T U_{\text{cov}} (\sigma^2)^{-1} U_{\text{cov}}^{-1} (O^{\text{diff}} - O^{\text{BSM}}).
\]

Since the diagonal matrix \( \hat{\sigma}^2 \) is positive definite, the quantity \( (\hat{\sigma}^2)^{-\frac{1}{2}} \) is well defined. Therefore, we can normalize the observables in the new basis to unit error by writing
\[
\hat{O}^{\text{diff}} = (\hat{\sigma}^2)^{-\frac{1}{2}} U_{\text{cov}}^{-1} O^{\text{diff}},
\]
\[
\hat{O}^{\text{BSM}} = (\hat{\sigma}^2)^{-\frac{1}{2}} U_{\text{cov}}^{-1} O^{\text{BSM}}.
\]
Now \( \chi^2 \) has the form
\[
\chi^2 = (\hat{O}^{\text{BSM}} - \hat{O}^{\text{diff}})^2.
\]
Since \( \hat{O}^{\text{BSM}} \) is linear in the SMEFT coefficients, we can write it in the form
\[
\hat{O}^{\text{BSM}} = M^{\text{SMEFT}} C,
\]
where \( C \) is an \( N_{\text{coeff}} \)-dimensional column vector of SMEFT coefficients and \( M^{\text{SMEFT}} \) is an \( N_{\text{obs}} \times N_{\text{coeff}} \) matrix. Hence, in order to constrain the SMEFT coefficients, we minimize
\[
\chi^2 = (M^{\text{SMEFT}} C - \hat{O}^{\text{diff}})^2.
\]
which is a linear-least-squares problem. As was described in Sec. II A, the solution of this minimization problem can be obtained from the SVD decomposition $M_{\text{SMEFT}} = UWV^\dagger$:

$$C = VW^{-1}U^\dagger\hat{O}^{\text{diff}}.$$  \hspace{1cm} (15)

As we have mentioned, each element of $V^\dagger C$ gives one of the principal components. The best-fit values of the principal components are given by the elements of $\hat{P}$ in Eq. (4), and the one-standard-deviation uncertainty on each principal component is given by the inverse of corresponding singular value in $W$.

The covariance matrix of the SMEFT coefficients is obtained by using $V$ to rotate $W$ back to the original basis of SMEFT coefficients:

$$\sigma_{\text{SMEFT}}^2 = VW^{-1}(W^{-1}V^\dagger)^\dagger.$$  \hspace{1cm} (16)

As is standard, the covariance matrix for the situation in which one has marginalized over some of the coefficients is obtained by striking from the full covariance matrix the rows and columns that correspond to the marginalized coefficients. Although the covariance matrix contains the same information as the uncertainties in the principal components, we will see that the presentation of uncertainties in principal-component form leads to a clearer picture of the constraints on the SMEFT coefficients.

III. APPLICATION TO TOP-QUARK DECAYS

In this section, we illustrate the principal-component/SVD method for fitting the SMEFT coefficients by applying it to the case of top-quark decay to a $b$ quark and a $W$ boson.

A. SMEFT operators

We work in the Warsaw-basis of SMEFT operators, and our notation is similar to that of Ref. [26]. Following Ref. [27], we fit the coefficients $C_{tW}$, $C_{bW}$, $C_{\phi tb}$, $C_{tg}$, and $C_{bg}$, which
correspond to the operators

\[ O_{tW} = O_{uW}, \]
\[ O_{bW} = O_{dW}, \]
\[ O_{\phi tb} = O_{\phi ud}, \]
\[ O_{tg} = O_{ug}, \]
\[ O_{bg} = O_{dg}. \]  

(17)

with

\[ O_{uW}^{pr} = \bar{q}_p \sigma^{\mu \nu} u_r \tau^I \tilde{\phi} W_I^{\mu \nu}, \]
\[ O_{dW}^{pr} = \bar{q}_p \sigma^{\mu \nu} d_r \tau^I \phi W_I^{\mu \nu}, \]
\[ O_{\phi ud}^{pr} = i (\bar{\tilde{\phi}} D_\mu \phi) (\bar{u}_r \gamma^{\mu} d_r), \]
\[ O_{uq}^{pr} = \bar{q}_p \sigma^{\mu \nu} T^A u_r \bar{\phi} G^A_{\mu \nu}, \]
\[ O_{dg}^{pr} = \bar{q}_p \sigma^{\mu \nu} T^A d_r \bar{\phi} G^A_{\mu \nu}. \]  

(18)

We also consider the coefficients \( C^{(1)}_{qq}, C^{(3)}_{qq}, C^{(1)}_{qu}, C^{(8)}_{qu} \) and \( C^{(3)}_{lq} \), which correspond to the four-fermion operators

\[ O^{(1)}_{qq} = (\bar{q}_p \gamma^\mu q_r) (\bar{q}_s \gamma^\mu q_t), \]
\[ O^{(3)}_{qq} = (\bar{q}_p \gamma^\mu \tau^a q_r) (\bar{q}_s \gamma^\mu \tau^a q_t), \]
\[ O^{(1)}_{qu} = (\bar{q}_p \gamma^\mu q_r) (\bar{u}_s \gamma^\mu u_t), \]
\[ O^{(8)}_{qu} = (\bar{q}_p \gamma^\mu T^A q_r) (\bar{u}_s \gamma^\mu T^A u_t), \]
\[ O^{(3)}_{lq} = (\bar{q}_p \gamma^\mu \tau^a q_r) (\bar{q}_s \gamma^\mu \tau^a q_t). \]  

(19)

Here, \( q_p \) (\( l_p \)) is a left-handed quark (lepton) isospin doublet with generation index \( p \), \( u_r \) and \( d_r \) are the up and down right-handed isospin singlets with generation index \( r \), \( l_r \) is the lepton right-handed isospin singlet with generation index \( r \), \( \phi \) is the Higgs isospin doublet, \( \tilde{\phi} = i \tau_2 \phi^* \) is the hypercharge-conjugate Higgs doublet, \( \tau \) is a Pauli matrix, \( W_{\mu \nu}^I \) is the field-strength tensor for the \( SU(2)_I \) gauge bosons with isospin index \( I \), the \( \gamma \)'s are Dirac matrices, \( \sigma_{\mu \nu} = i[\gamma_\mu, \gamma_\nu] \), \( G_{\mu \nu}^A \) is the gluon field-strength tensor with color index \( A \), and \( T^A \) is a color matrix in the fundamental representation with tensor with color index \( A \).
B. Experimental inputs

We take experimental values of the total top-quark decay rate and the helicity fractions from the Particle Data Group compilation [28]:

\[
\begin{align*}
\Gamma_{\text{exp}}^\text{tot} &= 1.41^{+0.19}_{-0.15} \text{ GeV}, \\
F_L^{\text{exp}} &= 0.687 \pm 0.018, \\
F_-^{\text{exp}} &= 0.320 \pm 0.013.
\end{align*}
\] (20)

In our analysis, we symmetrize the uncertainties in \(\Gamma_{\text{exp}}^\text{tot}\) by shifting the central value. That is, we take \(\Gamma_{\text{exp}}^\text{tot} = 1.43 \pm 0.17 \text{ GeV}\). The correlation matrix of the experimental uncertainties is [29]

\[
\rho = \begin{pmatrix}
1.0 & 0 & 0 \\
0 & 1.0 & -0.87 \\
0 & -0.89 & 1.0
\end{pmatrix}.
\] (21)

Then, the experimental inputs for our SVD analysis are \(O^{\text{exp}} = (\Gamma_{\text{exp}}^\text{tot}, F_L^{\text{exp}}, F_-^{\text{exp}})^T\), and \(\sigma_{ij}^2 = \sigma_i \rho_{ij} \sigma_j\), where \(i, j = 1, 2, 3\).

C. Theoretical inputs

We make use of the expressions in Ref. [27] for SMEFT contributions to \(\Gamma_{\text{tot}}\), the total decay width to \(bW\), \(F_L\), the fractional decay rate for a longitudinally polarized \(W\) boson, and \(F_-\), the fractional decay rate for a \(W\) boson with negative helicity.\(^3\) We include the QCD corrections that are given in Ref. [27]. We note that the standard-model QCD corrections are also given in Ref. [31] and that an analysis of the SMEFT contributions to \(t\) quark decay has also been given in Ref. [32].

For purposes of this demonstration, we do not include uncertainties in the theoretical predictions. They could be incorporated into the analysis by adding the theoretical covariance

\[\text{\footnotesize \textsuperscript{\cite{footnote}}\footnotetext{The published forms of these expressions in Ref. 27 have been corrected 30. In }\Gamma_{\text{tot}}\text{ in Eq. (13), the denominator factor }x_W\text{ has been replaced with }x_W^2.\text{ In }F_L^{\text{LO}}\text{ in Eq. (13), the second numerator parenthesis has been moved to the start of the numerator. In }\Delta F_L^{\text{QCD}}\text{ in Eq. (A2), the standard-model contribution has been changed to agree with the expression in Eq. (15) of Ref. 31.}}]}\]
matrix to the experimental covariance matrix.\textsuperscript{4}

The input parameters for the theoretical calculation are given in Table \ref{tab:input} and are identical to those in Ref. \cite{27}, except that we evaluate $\alpha_s$ at the scale $m_t$, rather than the scale $M_Z$. We compute the electroweak coupling $\bar{g}$ from \cite{33}

\begin{equation}
\bar{g}^2 = 2 \sqrt{2} G_F M_Z^2 \left( 1 + \sqrt{1 - \frac{4 \pi \alpha_{em}}{\sqrt{2} G_F M_Z^2}} \right). \tag{22}
\end{equation}

We set the SMEFT cutoff to be $\Lambda = 500$ GeV.

\textbf{D. Fit with one SMEFT coefficient}

In Table \ref{tab:bestfit}, we show the best-fit values of the SMEFT coefficients and their two-standard-deviation uncertainties that are obtained by setting all of the coefficients to zero, except for one. This is a widely used approach to constraining the SMEFT coefficients. However, as we will see, it can be quite misleading. Throughout the remainder of this paper, when we

\begin{table}[h]
\centering
\begin{tabular}{ccc}
$M_Z$ & 91.1876 GeV & $M_W$ & 80.379 GeV \\
v & 246 GeV & $m_t$ & 173.0 GeV \\
$m_b$ & 4.78 GeV & $G_F$ & $1.1664 \times 10^{-5}$ GeV$^{-2}$ \\
$\alpha_{em}^{-1}$ & 137.036 & $\alpha_s(m_t)$ & 0.1081 \\
\end{tabular}
\caption{Input parameters for the theoretical calculation.}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{cc}
$C_{tW}$ & 0.0644 $\pm$ 0.100 \\
$C_{tg}$ & $-3.65 \pm 8.22$ \\
$C_{bW}$ & 0.679 $\pm$ 1.03 \\
$C_{bg}$ & $-13.5 \pm 17.9$ \\
$C_{qtb}$ & 2.94 $\pm$ 5.56 \\
$C_{qq}$ & $-4.81 \pm 37.1$ \\
$C_{qq}^{(1)}$ & $0.656 \pm 5.07$ \\
$C_{qq}^{(3)}$ & $14.1 \pm 56.4$ \\
$C_{qq}^{(8)}$ & 10.6 $\pm$ 42.3 \\
$C_{qg}^{(3)}$ & $-4.35 \pm 33.6$ \\
\end{tabular}
\caption{Best-fit values of the SMEFT coefficients and their two-standard-deviation uncertainties computed by setting all of the coefficients to zero, except for one.}
\end{table}

\textsuperscript{4} If the theoretical and experimental uncertainties are correlated, say, through the use of a common parameter in the theoretical and experimental analyses, then one would need to construct a complete covariance matrix of theoretical and experimental uncertainties, including an entries for that parameter, and then marginalize over that parameter.
TABLE III: Best-fit values of the SMEFT coefficients obtained by keeping only three coefficients nonzero. The uncertainties are the two-standard-deviation uncertainties that are obtained by marginalizing over two of the coefficients.

| C_{tW} | 0.0209 ± 0.333 |
|--------|-----------------|
| C_{tg} | 4.73 ± 62.8     |
| C_{bW} | 1.13 ± 8.72     |

E. Fit with three SMEFT coefficients

We now contrast the results in Table II with those that can be obtained through a principal-component analysis. We begin by considering the case in which only the first three coefficients in Table II are nonzero. Then, we can compute the best-fit values of those coefficients and their uncertainties, marginalized over the other two coefficients. As we explained earlier, the latter can be obtained from the diagonal values of the covariance matrix. The results of this computation are shown in Table III. As can be seen, the central values have shifted substantially relative to those in Table II and the uncertainties have increased, in some cases by almost an order of magnitude. Clearly, the single-coefficient values and uncertainties in Table II are not indicative of the true constraints on the SMEFT coefficients in the presence of three non-zero coefficients. However, the large uncertainties in Table III paint an unduly pessimistic picture of the constraints that can be achieved.

In order to see this, let us consider the results of the SVD analysis with three nonzero coefficients. We obtain

$$V^\dagger = \begin{pmatrix} -0.999 & 0.000697 & -0.0414 \\ 0.0411 & 0.136 & -0.990 \\ -0.00496 & 0.991 & 0.136 \end{pmatrix}$$ (23)
and
\[ W = \begin{pmatrix}
19.9 & 0 & 0 \\
0 & 1.77 & 0 \\
0 & 0 & 0.0315
\end{pmatrix}. \tag{24} \]

From the expression for \( V^\dagger \), we see that the principal components are
\[
\begin{align*}
P_1 &= 0.999 C_{tW} - 0.000697 C_{tg} + 0.0414 C_{bW}, \\
P_2 &= -0.0411 C_{tW} - 0.136 C_{tg} + 0.990 C_{bW}, \\
P_3 &= -0.00496 C_{tW} + 0.991 C_{tg} + 0.136 C_{bW}. \tag{25}
\end{align*}
\]

Each row of \( V^\dagger \) gives the coefficients of the SMEFT coefficients for one of the principal components. The best-fit values of the principal components are given by the elements of \( P \) in Eq. (4), and their two-standard-deviation uncertainties are given by twice the inverse of the corresponding diagonal element of \( W \). The results are
\[
\begin{align*}
P_1 &= -0.0645 \pm 0.100, \\
P_2 &= -0.477 \pm 1.13, \\
P_3 &= 4.84 \pm 63.4. \tag{26}
\end{align*}
\]

We see that \( P_1 \) and \( P_2 \) are much better constrained than any of the individual coefficients and that only \( P_3 \) is poorly constrained. The principal-component analysis clearly allows one to access a much more powerful set of constraints than do the analyses of individual SMEFT coefficients.

**F. Fit with five SMEFT coefficients**

Now suppose that we keep only the first five SMEFT coefficients in Table II nonzero. In this case, we have more SMEFT coefficients than observables, and so the individual coefficients cannot be fit unambiguously. Furthermore, because there are necessarily flat directions, the marginalization over some sets of SMEFT coefficients is ill-defined. Never-
theless, the SVD approach allows us to find meaningful constraints. We obtain

$$V^\dagger = \begin{pmatrix}
-0.999 & 0.000698 & -0.0414 & 0.00390 & -0.00316 \\
0.0411 & 0.134 & -0.969 & 0.0443 & -0.196 \\
0.00282 & -0.931 & -0.0541 & 0.0222 & -0.361 \\
0.00199 & 0.0163 & 0.0453 & 0.999 & 0.0128 \\
0.00647 & -0.340 & -0.231 & 0.00431 & 0.912 
\end{pmatrix}$$

(27)

and

$$W = \begin{pmatrix}
19.9 & 0 & 0 & 0 & 0 \\
0 & 1.80 & 0 & 0 & 0 \\
0 & 0 & 0.0339 & 0 & 0 
\end{pmatrix}.$$  

(28)

The coefficients of the SMEFT coefficients of a given principal component are given in the entries in a the corresponding row in $V^\dagger$. The best-fit values for the first three principal components and their two-standard-deviation errors are

$$\overline{P}_1 = -0.0645 \pm 0.100,$$

$$\overline{P}_2 = -0.467 \pm 1.11,$$

$$\overline{P}_3 = -4.46 \pm 59.0.$$  

(29)

The linear combinations of SMEFT coefficients that are unconstrained have been sequestered by the SVD procedure. Comparing Eq. (29) with Eq. (26), we see that the best-fit values and uncertainties in the first two principal components are remarkably stable as new SMEFT contributions are introduced. This reflects the fact that the observables are relatively insensitive to the SMEFT contributions that are proportional to $C_{bg}$ and $C_{\phi tb}$, as can be seen from the small coefficients of $C_{bg}$ and $C_{\phi tb}$ the first two principal components.
G. Fit with ten SMEFT coefficients

Next we apply the SVD method to the complete set of ten SMEFT coefficients in Table II. We list only the three principal components that are constrained. They are

\[ P_1 = -0.999C_{tW} + 0.000698C_{tg} - 0.0414C_{bW} + 0.00390C_{bg} - 0.00316C_{\phi tb} \]
\[ + 0.000737C_{qq}^{(1)} - 0.00540C_{qq}^{(3)} - 0.000570C_{qu}^{(1)} - 0.000760C_{qu}^{(8)} + 0.000815C_{lq}^{(3)}, \]
\[ P_2 = 0.0406C_{tW} + 0.134C_{tg} - 0.967C_{bW} + 0.0442C_{bg} - 0.196C_{\phi tb} \]
\[ - 0.00856C_{qq}^{(1)} + 0.0628C_{qq}^{(3)} - 0.00550C_{qu}^{(1)} - 0.00733C_{qu}^{(8)} - 0.00947C_{lq}^{(3)}, \]
\[ P_3 = 0.00819C_{tW} - 0.0753C_{tg} - 0.0672C_{bW} + 0.00484C_{bg} - 0.0452C_{\phi tb} \]
\[ + 0.131C_{qq}^{(1)} - 0.964C_{qq}^{(3)} + 0.0854C_{qu}^{(1)} + 0.114C_{qu}^{(8)} + 0.145C_{lq}^{(3)}. \] (30)

The best-fit values and two-standard-deviation uncertainties for these principal components are

\[ \overline{P}_1 = -0.0645 \pm 0.100, \]
\[ \overline{P}_2 = -0.465 \pm 1.11, \]
\[ \overline{P}_3 = -0.432 \pm 5.32. \] (31)

As can be seen, the first two principal components remain quite stable in best-fit value and uncertainty as new SMEFT coefficients are introduced, reflecting the relative insensitivity of the observables to the new SMEFT coefficients.

H. Fit with a flat direction in coefficient space

Finally, we examine the case in which the number of SMEFT coefficients and the number of observables are equal, but there is a hidden flat direction. In order to construct an example of this situation, we keep three SMEFT coefficients, \( C_{tW}, C_{tg}, \) and \( C_{bW}, \) nonzero and set the remaining SMEFT coefficients to zero. Let \( a_{tW}, a_{tg}, \) and \( a_{bW} \) be the coefficients of \( C_{tW}, C_{tg}, \) and \( C_{bW} \) in \( O^{\text{SMEFT}}. \) Then, the following replacement creates an approximate artificial flat direction in the space of \( C_{tg} \) and \( C_{bW}: \)

\[ a_{tg} \to \epsilon a_{tg} + (1 - \epsilon)ra_{bW}. \] (32)

In the limit \( \epsilon \to 0, \) there is an exact flat direction in the space of SMEFT coefficients.
As a numerical example, we take scale factor \( r \) in Eq. (32) to be \(-3.2\). As \( \epsilon \) approaches zero, conventional fitting procedures that use gradients of \( \chi^2 \) to find a minimum in \( \chi^2 \) have numerical difficulties. Let us consider, for example, the situation for \( \epsilon = 10^{-6} \). The Mathematica routine FindMinimum can be used to minimize the \( \chi^2 \). This routine comes with a number of options for the method to be used in finding the minimum. Using Mathematica version 11.3 [17], we find that the conjugate-gradient method algorithm yields

\[
C_{tw} = 0.0445, \\
C_{tg} = 0.236, \\
C_{bw} = 1.24, 
\]

(33)

Newton’s method yields

\[
C_{tw} = 0.0410, \\
C_{tg} = 7.00 \times 10^5, \\
C_{bw} = 2.24 \times 10^6, 
\]

(34)

and the principal-axis method yields

\[
C_{tw} = 0.0445, \\
C_{tg} = 0.161, \\
C_{bw} = 1.00. 
\]

(35)

Owing to the existence of a nearly flat direction, the results vary wildly, depending on the algorithm that is used in minimizing \( \chi^2 \). While this behavior is to be expected in under-determined problems, it is difficult to draw any conclusion from such results. In contrast, meaningful constraints can be found by using principal-component analysis. From SVD, we obtain

\[
V^\dagger = \begin{pmatrix}
-0.988845 & 0.142167 & -0.0444274 \\
-0.148947 & -0.943833 & 0.294948 \\
-1.49286 \times 10^{-9} & 0.298275 & 0.95448
\end{pmatrix}, 
\]

(36)

and

\[
W = \begin{pmatrix}
20.1307 & 0 & 0 \\
0 & 5.80765 & 0 \\
0 & 0 & 9.49318 \times 10^{-9}
\end{pmatrix}, 
\]

(37)
which implies that, among the principal components,

\[ P_1 = -0.989C_{tW} + 0.142C_{tg} - 0.0444C_{bW}, \]
\[ P_2 = -0.149C_{tW} - 0.944C_{tg} + 0.295C_{bW}, \]
\[ P_3 = -1.49 \times 10^{-9}C_{tW} + 0.298C_{tg} + 0.954C_{bW}, \] (38)

only \( P_1 \) and \( P_2 \) are well constrained, while \( P_3 \) is not, as is evident from the near-vanishing of the corresponding diagonal value of \( W \).\(^5\) We see that the principal-component/SVD method has constrained the principal components that contain a contribution that is proportional to \(-3.2C_{tg} + C_{bW}\) and has identified as unconstrained the principal component that contains a contribution that is proportional to \(C_{tg} + 3.2C_{bW}\), which corresponds to the flat direction. The best-fit values of the principal components and their two-standard deviation uncertainties are

\[ \bar{P}_1 = -0.0655 \pm 0.0994, \]
\[ \bar{P}_2 = 0.136 \pm 0.344, \]
\[ \bar{P}_3 = 1.58 \times 10^7 \pm 2.11 \times 10^8. \] (39)

We can invert the relations in Eq. (38), using the rows of \( V \) to obtain the coefficients. The result is

\[ C_{tW} = -0.989P_1 - 0.149P_2 - 1.49 \times 10^{-9}P_3, \]
\[ C_{tg} = 0.142P_1 - 0.944P_2 + 0.298P_3, \]
\[ C_{bW} = -0.0444P_1 + 0.295P_2 + 0.954P_3. \] (40)

From Eq. (40), it is easily seen that, to good approximation, the differences between the three results from FindMinimum correspond to differences in the value of \( P_3 \).

When \( \epsilon = 0 \) and there is an exact flat direction, numerical minimization of \( \chi^2 \) with respect to the SMEFT coefficients would fail to converge to a result. However, the SVD method still yields meaningful constraints. Specifically, we have

\[ V^\dagger = \begin{pmatrix} -0.988845 & 0.142168 & -0.0444274 \\ -0.148948 & -0.943833 & 0.294948 \\ 1.29707 \times 10^{-17} & 0.298275 & 0.95448 \end{pmatrix} \] (41)

\(^5\) We keep more significant digits than usual in the results for \( V^\dagger \) and \( W \) in Eqs. (30) and (37) for purposes of later comparison with the results for \( V^\dagger \) and \( W \) when \( \epsilon = 0 \).
and

\[ W = \begin{pmatrix} 20.1307 & 0 & 0 \\ 0 & 5.80765 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (42) \]

which are very close to the results for \( \epsilon = 10^{-6} \).

**IV. EXTENSION TO HIGHER ORDERS IN THE SMEFT EXPANSION**

The principal-component/SVD method that we have presented is limited to fitting problems in which the observables depend linearly on SMEFT coefficients. This is the case for fits at the leading nontrivial order in the effective-field-theory expansion, in which one considers only the contributions of the interference of the dimension-6 SMEFT-operator amplitudes with the dimension-4 standard-model-operator amplitudes. At the next order in the SMEFT expansion, one would need to consider not only the contribution from the square of the dimension-6 SMEFT-operator amplitudes, but also the interference of the dimension-4 standard-model-operator amplitudes with the thousands of dimension-8 SMEFT-operator amplitudes—a task that is not likely to be undertaken soon.

Nevertheless, beyond leading order, one might still apply the SVD method by making use of an iterative procedure. One could first carry out a fit that retains only the interference of the dimension-6 amplitude with the dimension-4 amplitude. Then, one could compute the contribution of the square of the dimension-6 amplitude, subtract it from the experimental values of the observables, and carry out a new fit, including the dimension-8 operators. This last step could be iterated to produce fits of the desired accuracy. The iteration process should converge if the effective-field-theory expansion is valid, that is, if the square of the dimension-6 amplitude is less than the interference of the dimension-6 amplitude with the dimension-4 amplitude.

This method would yield best-fit values of the coefficients, but would not give accurate results for the principal components. Instead, one could compute the principal components as follows. First one could obtain the inverse covariance matrix by computing analytically two derivatives of \( \chi^2 \) with respect to the SMEFT coefficients and evaluating the result at the best-fit values of the coefficients from the iterative procedure. The inverse covariance matrix could be diagonalized by standard methods, and the principal components could
then be obtained from the elements of the unitary transformation that effects the diagonal-
ization. The uncertainties would be given by the inverses of the square roots of the diagonal
components of the inverse covariance matrix.

V. SUMMARY

In this paper, we have addressed a difficulty that appears in using experimental mea-
surements of observables to constrain the Wilson coefficients in standard model effective
theory (SMEFT). The difficulty is that the observables may be insensitive to certain linear
combinations of SMEFT coefficients. That is, there may be “flat directions” in the space of
SMEFT coefficients. This difficulty can arise because, in a partial analysis that is restricted
to a particular set of physical processes, the number of experimental observables may be less
than the number of SMEFT coefficients. In this case, it is clear that some linear combina-
tions of SMEFT coefficients would are not constrained. However, it can happen that some
linear combinations of SMEFT coefficients are poorly constrained even when the number of
observables is equal to or greater than the number of SMEFT coefficients to be fit.

We have advocated for the use of a principal-component analysis to address the issue of
flat directions, and we have presented an efficient method for carrying out the principal-
component analysis that is based on singular-value decomposition (SVD). The principal-
component/SVD analysis isolates the principal components that correspond to poorly con-
strained linear combinations of SMEFT coefficients. Furthermore, the reduction to princi-
pal components allows one to ascribe independent uncertainties to linear combinations of
SMEFT coefficients, thereby leading to more stringent bounds than one could obtain sim-
ply by marginalizing over some of the SMEFT coefficients. In particular, one can obtain
meaningful bounds on SMEFT coefficients, even in a partial analysis in which the number
of observables is less than the number of coefficients. We have demonstrated the application
of this method to the process of top-quark decays to a $W$ boson and a $b$ quark.

A further difficulty in fitting SMEFT coefficients to data that we did not address explicitly
in this paper is that the sheer numbers of SMEFT coefficients and observables may lead to a
formidable computational problem. In this regard, we can mention that the SVD method has
been used in another context to fit hundreds of coefficients to thousands of observables [15]
and can do so in a few seconds of CPU time on a typical present-day desktop computer.
Although the method that we have presented is limited to the case in which the observables depend linearly on the SMEFT coefficients, we have outlined in Sec. IV an iterative extension of the method that can be applied to the non-linear situation, provided that the contributions of the non-linear terms to the observables are small in comparison with the contributions of the linear terms. This is the case if the SMEFT expansion converges.

Finally, the method that we have presented relies on a Gaussian probability analysis. While one might ultimately want to improve on a Gaussian approach, it should certainly be adequate for the purpose of carrying out exploratory studies in SMEFT.

Acknowledgments

We thank Radja Boughezal, Chien-Yi Chen, Frank Petriello, and Daniel Wiegand for providing us with a version of Ref. [27] that contains corrected expressions for the total width and helicity fractions in top-quark decay to a $b$ quark and a $W$ boson. We also thank Xiang-peng Wang for checking some of our Mathematica expressions. The work of G.T.B. is supported by the U.S. Department of Energy, Division of High Energy Physics, under Contract No. DE-AC02-06CH11357. The work of H.S.C is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2094 – 390783311. The submitted manuscript has been created in part by UChicago Argonne, LLC, Operator of Argonne National Laboratory. Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

[1] W. Buchmuller and D. Wyler, Nucl. Phys. B 268, 621 (1986). doi:10.1016/0550-3213(86)90262-2

[2] K. Hagiwara, S. Ishihara, R. Szalapski and D. Zeppenfeld, Phys. Rev. D 48, 2182 (1993). doi:10.1103/PhysRevD.48.2182
[3] G. F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, JHEP 0706, 045 (2007) doi:10.1088/1126-6708/2007/06/045 [hep-ph/0703164].

[4] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, JHEP 1010, 085 (2010) doi:10.1007/JHEP10(2010)085 [arXiv:1008.4884 [hep-ph]].

[5] Z. Han and W. Skiba, Phys. Rev. D 71, 075009 (2005) doi:10.1103/PhysRevD.71.075009 [hep-ph/0412166].

[6] A. Pomarol and F. Riva, JHEP 1401, 151 (2014) doi:10.1007/JHEP01(2014)151 [arXiv:1308.2803 [hep-ph]].

[7] C. Y. Chen, S. Dawson and C. Zhang, Phys. Rev. D 89, no. 1, 015016 (2014) doi:10.1103/PhysRevD.89.015016 [arXiv:1311.3107 [hep-ph]].

[8] J. Ellis, V. Sanz and T. You, JHEP 1407, 036 (2014) doi:10.1007/JHEP07(2014)036 [arXiv:1404.3667 [hep-ph]].

[9] J. D. Wells and Z. Zhang, Phys. Rev. D 90, no. 3, 033006 (2014) doi:10.1103/PhysRevD.90.033006 [arXiv:1406.6070 [hep-ph]].

[10] A. Falkowski and F. Riva, JHEP 1502, 039 (2015) doi:10.1007/JHEP02(2015)039 [arXiv:1411.0669 [hep-ph]].

[11] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell and C. D. White, JHEP 1604, 015 (2016) doi:10.1007/JHEP04(2016)015 [arXiv:1512.03360 [hep-ph]].

[12] V. Cirigliano, W. Dekens, J. de Vries and E. Mereghetti, Phys. Rev. D 94, no. 3, 034031 (2016) doi:10.1103/PhysRevD.94.034031 [arXiv:1605.04311 [hep-ph]].

[13] J. A. Aguilar-Saavedra et al., arXiv:1802.07237 [hep-ph].

[14] N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou and C. Zhang, arXiv:1901.05965 [hep-ph].

[15] G. T. Bodwin, H. S. Chung and J. Repond, JINST 14, no. 07, P07002 (2019) doi:10.1088/1748-0221/14/07/P07002 [arXiv:1805.12163 [physics.ins-det]].

[16] See, for example, C. L. Lawson and R. J. Hanson, “Solving Least Squares Problems,” Society for Industrial and Applied Mathematics, Philadelphia (1995).

[17] Wolfram Research, Inc., “Mathematica,” version 11.3, Champaign, Illinois, 2018.

[18] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, “Numerical Recipes in C,” Cambridge University Press, Cambridge, 2002.

[19] S. van der Walt, S. C. Colbert, and G. Varoquaux, “The NumPy Array: A Structure for
Efficient Numerical Computation,” Computing in Science Engineering, 13 no. 2, 22 (2011).

[20] M. Galassi et al., GNU Scientific Library Reference Manual, third edition (January 2009), http://www.gnu.org/software/gsl/

[21] A. Bjerhammar, Trans. Roy. Inst. Tech. Stockholm 49, 1 (1951).

[22] E. H. Moore, Bulletin of the American Mathematical Society 26, 394 (1920).

[23] R. Penrose, Proceedings of the Cambridge Philosophical Society 51, 406 (1955).

[24] R. Penrose, Proceedings of the Cambridge Philosophical Society 52, 17 (1956).

[25] T. B. Schön and F. Lindsten, Manipulating the multivariate Gaussian Density, Technical Report, Division of Automatic Control, Linköping University, Sweden (January, 2011).

[26] A. Dedes, W. Materkowska, M. Paraskevas, J. Rosiek and K. Suxho, JHEP 1706, 143 (2017) doi:10.1007/JHEP06(2017)143 [arXiv:1704.03888 [hep-ph]].

[27] R. Boughezal, C. Y. Chen, F. Petriello and D. Wiegand, Phys. Rev. D 100, no. 5, 056023 (2019) doi:10.1103/PhysRevD.100.056023 [arXiv:1907.00997 [hep-ph]].

[28] M. Tanabashi et al. [Particle Data Group], Phys. Rev. D 98, no. 3, 030001 (2018) doi:10.1103/PhysRevD.98.030001

[29] V. Khachatryan et al. [CMS Collaboration], Phys. Lett. B 762, 512 (2016) doi:10.1016/j.physletb.2016.10.007 [arXiv:1605.09047 [hep-ex]].

[30] D. Wiegand, private communication.

[31] M. Fischer, S. Groote, J. G. Korner and M. C. Mauser, Phys. Rev. D 63, 031501 (2001) doi:10.1103/PhysRevD.63.031501 [hep-ph/0011075].

[32] C. Zhang, Phys. Rev. D 90, no. 1, 014008 (2014) doi:10.1103/PhysRevD.90.014008 [arXiv:1404.1264 [hep-ph]].

[33] S. Dawson and P. P. Giardino, [arXiv:1909.02000] [hep-ph].

21