Sharpening the shape analysis for higher-dimensional operator searches

Sylvain Fichet,[1,*] Alberto Tonero,[1,†] and Patricia Rebello Teles[2,3,‡]

[1]*ICTP-SAIFR & IFT-UNESP R. Dr. Bento Teobaldo Ferraz 271 São Paulo, Brazil*
[2]*CBPF, R. Dr. Xavier Sigaud 150, Rio de Janeiro, Brazil*
[3]*CERN, CH-1211 Geneva 23, Switzerland*
(Received 1 February 2017; published 3 August 2017)

When the Standard Model is interpreted as the renormalizable sector of a low-energy effective theory, the effects of new physics are encoded into a set of higher-dimensional operators. These operators potentially deform the shapes of Standard Model differential distributions of final states observable at colliders. We describe a simple and systematic method to obtain optimal estimations of these deformations when using numerical tools, like Monte Carlo simulations. A crucial aspect of this method is minimization of the estimation uncertainty: We demonstrate how the operator coefficients have to be set in the simulations in order to get optimal results. The uncertainty on the interference term turns out to be the most difficult to control and grows very quickly when the interference is suppressed. We exemplify our method by computing the deformations induced by the $\mathcal{O}_{3W}$ operator in W^+W^- production at the LHC, and by deriving a bound on $\mathcal{O}_{3W}$ using 8 TeV CMS data.

I. INTRODUCTION

After the historical discovery of the Higgs boson [1,2] in July 2012, the Large Hadron Collider (LHC) is currently probing matter and spacetime at unprecedented small distances, looking for a signal of physics beyond the Standard Model (SM). However, new physics remains very elusive so far, as no statistically significant signal has been observed in the collected data up to now.

Although there are good theoretical arguments to expect new particles near the TeV scale, it is also plausible that these states will be somewhat too heavy to be produced on shell at the LHC. In such a scenario, the presence of new physics states is best studied using effective field theory methods, and their effects can be parametrized by higher-dimensional operators made of SM fields. The search for heavy physics beyond the SM then becomes a program of SM precision measurements, aiming at testing the existence of one or several of these higher-dimensional effective operators.

The analysis of the distributions of final-state kinematic variables plays a central role in this scenario because their shapes contain important information about the presence of the effective operators. In principle, a thorough analysis of the shapes of the differential distributions may be the key to

the discovery of new physics. However, such a program is not so straightforward to carry out systematically because of the large number of effective operators and kinematic variables to take into account and because of the computational cost of estimating the differential rates [3]. Much progress is still needed in order to systematically and efficiently search for an arbitrary number of higher-dimensional operators using multidimensional differential rates. An attempt to improve shape analysis techniques using the moments of the differential distributions has been done in Ref. [4].

In this paper we focus on establishing an optimal method for determining the various contributions (SM, interference, and BSM) that constitute the differential rates in the presence of dimension-6 effective operators. The paper is organized as follows. A first inspection of the event rates in the presence of effective operators is done in Sec. II. Based on these considerations, an optimal method to estimate the expected deformations of differential rates induced by an arbitrary number of effective operators is presented in Sec. III and summarized in Sec. III C. As an illustration and check, we obtain the deformations of the differential rate of WW production induced by the $\mathcal{O}_{3W}$ operator at reconstruction level, and we reproduce within two sigma the bound obtained on $\mathcal{O}_{3W}$ from 8 TeV data by CMS. Conclusions are presented in Sec. V.

II. LOW-ENERGY AMPLITUDES AND PHASE-SPACE CONSIDERATIONS

A. Effective theory basics

A field theory can sometimes be approximated by a simpler, so-called "effective" theory, in a given region of phase space [5]. Consider an ultraviolet (UV) theory that

*[*]sylvain@ift.unesp.br
[†]alberto.tonero@gmail.com
[‡]patricia.rebello.teles@cern.ch

lies at a scale Λ which is much larger than the typical energy scale E of a given experiment ($\Lambda \gg E$). From the UV point of view, the low-energy effective theory is obtained by expanding the correlation functions computed within the UV theory with respect to the parameter E/Λ.

It is very plausible that the Standard Model is not the ultimate UV-complete theory of nature, and its Lagrangian corresponds to the relevant and marginal sector of a low-energy effective theory. In this picture, the full SM effective Lagrangian also contains a series of operators of dimension higher than four, built from SM fields and invariant under SM symmetries, which are suppressed by inverse powers of the new physics scale Λ. To the low-energy observer, these higher-dimensional operators parametrize the new physics effects, which appear as new interactions between the known particles, including vertices with extra derivatives. The SM effective Lagrangian takes the general form

$$\mathcal{L}_{\text{eff}} = \mathcal{L}^{\text{SM}} + \sum_{d=5}^{\infty} \mathcal{L}^{(d)}, \tag{1}$$

where

$$\mathcal{L}^{(d)} = \sum_{I} \frac{\alpha_I^{(d)}}{\Lambda^{d-4}} \mathcal{O}_I^{(d)}. \tag{2}$$

The effective operators $\mathcal{O}_I^{(d)}$ have mass dimension $d > 4$, and $\alpha_I^{(d)}$ are dimensionless coefficients. In general, these coefficients can be complex, depending on the operator. However, one can always split the operator into its real and imaginary parts in order to end up with only real coefficients. Because of this, in the following we consider only real coefficients, without loss of generality.

Identifying Λ with the mass scale of a heavy particle, the effective field theory is valid only at energies

$$E < \Lambda, \tag{3}$$

and the contributions of the effective operators to any given observable are arranged as an expansion in E/Λ [6]. The power of the effective theory approach lies in the fact that this series can be consistently truncated in order to match a desired precision, which can be chosen according to the available experimental uncertainties. At any given order of truncation, new physics effects are described by a finite set of universal coefficients. Larger effects are expected to come from operators with lower dimension: For example, the leading contributions to Higgs observables come from the dimension-6 Lagrangian $\mathcal{L}^{(6)}$, while the leading contributions to quartic neutral gauge boson interactions come from the dimension-8 Lagrangian $\mathcal{L}^{(8)}$ [7]. Here and in the following, we truncate the EFT expansion at dimension 6. The complete list of independent dimension-6 SM effective operators has been reported in [8].

Let us consider, for concreteness, the case of a single dimension-6 effective operator. The general case involving an arbitrary number of higher-dimensional operators is very similar. The effective Lagrangian is then given by

$$\mathcal{L}_{\text{eff}} = \mathcal{L}^{\text{SM}} + \frac{\alpha}{\Lambda^2} \mathcal{O}^{(6)}. \tag{4}$$

In the presence of the effective operator of Eq. (4), a generic amplitude $\mathcal{M}$ can be expressed as

$$\mathcal{M} = \mathcal{M}^{\text{SM}} + \frac{\alpha}{\Lambda^2} \mathcal{M}^{\text{BSM}} + O\left(\frac{\alpha^2}{\Lambda^4}\right), \tag{5}$$

where $\mathcal{M}^{\text{SM}}$ is the SM piece and $\mathcal{M}^{\text{BSM}}$ represents the leading BSM contribution obtained by one insertion of the effective operator. The subleading contributions are given by Feynman diagrams with more than one insertion of the effective operator. Notice that, in general, the BSM component $\mathcal{M}^{\text{BSM}}$ is not proportional to $\mathcal{M}^{\text{SM}}$.

Physical observables, like the ones measured at the LHC, are described statistically by event rates σ that are given by the integral of the squared amplitude $|\mathcal{M}|^2$ over some phase space domain $\mathcal{D}$—which can be chosen by the experimentalist to some extent. Typical examples of such observables are total cross sections and decay widths. In the presence of the effective operator of Eq. (4), the observable event rate σ is given by

$$\sigma \propto \int_{\mathcal{D}} d\Phi |\mathcal{M}|^2, \tag{6}$$

where $\mathcal{M}$ is given by Eq. (5) and $\int_{\mathcal{D}} d\Phi$ denotes the integral over the phase space domain. This event rate σ can be further decomposed as a sum of three leading contributions,

$$\sigma = \sigma(\alpha) \equiv \sigma^{\text{SM}} + \frac{\alpha}{\Lambda^2} \sigma^{\text{int}} + \frac{\alpha^2}{\Lambda^4} \sigma^{\text{BSM}} + \dots, \tag{7}$$

where

$$\sigma^{\text{SM}} \propto \int_{\mathcal{D}} d\Phi |\mathcal{M}^{\text{SM}}|^2, \quad \sigma^{\text{int}} \propto \int_{\mathcal{D}} d\Phi \, 2\text{Re}[\mathcal{M}^{\text{BSM}} \mathcal{M}^{\text{SM*}}],$$

$$\sigma^{\text{BSM}} \propto \int_{\mathcal{D}} d\Phi |\mathcal{M}^{\text{BSM}}|^2. \tag{8}$$

The ellipses in Eq. (7) represent both contributions of higher order in $1/\Lambda$ and $O(\Lambda^{-4})$ terms coming from the interference of higher-order diagrams with the SM component. It will be made clear in the next subsection why all the terms in Eq. (7) should be kept in order to describe the leading effects of new physics.

The leading components of a square matrix element in terms of Feynman diagrams are illustrated in Fig. 1. The component σ^{SM} corresponds to the pure SM contribution,
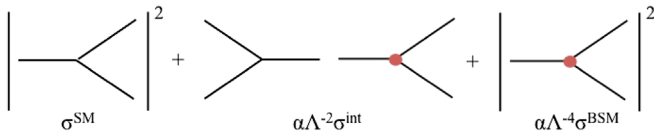
FIG. 1. The leading contributions to the square amplitude in presence of a dimension-6 operator.

which remains in the limit $\alpha/\Lambda^2 \to 0$. The component σ^{int} is obtained from the interference between the SM and BSM amplitudes, while the component σ^{BSM} is the pure BSM contribution. From the point of view of a new physics searcher, these two latter components together constitute the new physics signal $\sigma^{\text{NP}} = \sigma - \sigma^{\text{SM}}$, while σ^{SM} constitutes the irreducible background [9]. Notice that, by definition, σ^{int} and σ^{BSM} do not have the same dimension as σ^{SM}.

Finally, we notice the general fact that the modulus of the interference term has an upper bound

$$|\sigma^{\text{int}}| < 2\sqrt{\sigma^{\text{SM}}}\sqrt{\sigma^{\text{BSM}}}. \tag{9}$$

This is obtained by using

$$
\left| \int d\Phi \text{Re}[\mathcal{M}^{\text{SM}}\mathcal{M}^{\text{BSM}*}] \right|
$$
$$
\leq \left| \int d\Phi \mathcal{M}^{\text{SM}}\mathcal{M}^{\text{BSM}*} \right|
$$
$$
\leq \sqrt{\int d\Phi |\mathcal{M}^{\text{SM}}|^2} \sqrt{\int d\Phi |\mathcal{M}^{\text{BSM}}|^2}, \tag{10}
$$

where in the last step we have used the Cauchy-Schwartz inequality. In the following, we refer to Eq. (9) as the "Cauchy-Schwartz bound" on the interference.

B. Consistent truncation of event rates

We can now consider various limit cases of signal and background over a given phase space domain. Consider first a region of phase space $\mathcal{D}_1$ where the SM contribution is much larger than the BSM contribution,

$$|\mathcal{M}^{\text{SM}}| \gg \left| \frac{\alpha}{\Lambda^2} \mathcal{M}^{\text{BSM}} \right|. \tag{11}$$

The Cauchy-Schwartz bound automatically implies that the interference is much smaller than σ^{SM}, namely, $\sigma^{\text{SM}} \gg |\alpha\Lambda^{-2}\sigma^{\text{int}}|$. The modulus of the interference term can, in principle, take any value between zero and $2|\alpha|\Lambda^{-2}\sqrt{\sigma^{\text{SM}}\sigma^{\text{BSM}}}$. If we assume the case of an unsuppressed interference, then we have

$$\sigma_{\text{SM}} \gg \left| \frac{\alpha}{\Lambda^2}\sigma^{\text{int}} \right| \gg \frac{\alpha^2}{\Lambda^4}\sigma^{\text{BSM}}. \tag{12}$$

In this case, the SM background is large with respect to the signal, and the signal appears predominantly through the interference term σ^{int}. The component σ^{BSM} can thus be neglected at leading order in the effective theory expansion. For practical purposes, one may notice that the σ^{BSM} term can also be kept, as long as it is negligible. Whenever the σ^{BSM} term becomes non-negligible, the truncation of the EFT expansion has to be pushed to next-to-leading order: The higher-order diagrams in Eq. (5) have to be taken into account, as well as the contributions from higher-order operators. In this paper our focus is merely on the leading effects of new physics; thus, aspects of the EFT at next-to-leading order are beyond our scope.

Still, in the case described by Eq. (11), if the interference is vanishing, then the signal comes mainly from σ^{BSM}. In such a case, one should be careful with the next-to-leading order contributions of the EFT, as described in the paragraph above. Since the σ^{BSM} component is $O(1/\Lambda^4)$, other $O(1/\Lambda^4)$ contributions coming from the interference of higher-order diagrams or dimension-8 operators with the SM amplitude can be present. These observations have also been made in [10,11].

Consider now a phase space domain $\mathcal{D}_2$ where the SM contribution is much smaller than the BSM contribution,

$$|\mathcal{M}^{\text{SM}}| \ll \left| \frac{\alpha}{\Lambda^2} \mathcal{M}^{\text{BSM}} \right|. \tag{13}$$

Using again the Cauchy-Schwartz bound, Eq. (13) automatically implies that the interference term is much smaller than the BSM contribution, namely, $\alpha^2\Lambda^{-4}\sigma^{\text{BSM}} \gg |\alpha\Lambda^{-2}\sigma^{\text{int}}|$. Assuming an unsuppressed interference, we have

$$\sigma^{\text{SM}} \ll \left| \frac{\alpha}{\Lambda^2}\sigma^{\text{int}} \right| \ll \frac{\alpha^2}{\Lambda^4}\sigma^{\text{BSM}}. \tag{14}$$

This is the typical situation of a search for "rare events," where the SM background is vanishing and each signal event carries a high statistical significance. In this case, the new physics signal appears predominantly through the pure BSM term σ^{BSM}. This implies that, even at leading order in the EFT expansion, one has to keep the σ^{BSM} term. This fact might seem puzzling at first, as one ends up with a $O(1/\Lambda^4)$ term at leading order. Naively, there are other $O(1/\Lambda^4)$ contributions coming from the interference of higher-order diagrams or dimension-8 operators with the SM amplitude. However, one can easily check that the contribution σ^{BSM} is actually the dominant one because the other $O(\Lambda^{-4})$ contributions would come from an

interference with the SM amplitude, which is small by assumption [see Eq. (13)].

We conclude from the above analysis that, in general, *both* the interference term σ^{int} and the quadratic BSM term σ^{BSM} have to be kept in order to describe the leading effects of new physics for any background configuration. The σ^{int} piece dominates when $|\mathcal{M}^{\text{SM}}| \gg |\alpha\Lambda^{-2}\mathcal{M}^{\text{BSM}}|$—provided that the interference is not suppressed—while the σ^{BSM} piece dominates for $|\mathcal{M}^{\text{SM}}| \ll |\alpha\Lambda^{-2}\mathcal{M}^{\text{BSM}}|$.

Finally, let us comment about specific event rates encountered in collider experiments. In the case of resonant production of an unstable particle, the propagator is resummed, so the amplitude does not initially have the form Eq. (5), but the form

$$|\mathcal{M}|^2_{\text{res}} \propto \frac{1}{(s-m^2)^2 + m^2\Gamma^2_{\text{tot}}}, \qquad (15)$$

where Γ_{tot} is the total decay width. This is instead the total width which has the form of a quadratic function, $\Gamma_{\text{tot}} = \Gamma^{\text{SM}} + \alpha\Lambda^{-2}\Gamma^{\text{int}} + \alpha^2\Lambda^{-4}\Gamma^{\text{BSM}}$. However, if the EFT expansion is valid, one can always expand $|\mathcal{M}|^2_{\text{res}}$ with respect to s/Λ^2 and m^2/Λ^2, to end up with a quadratic form $|\mathcal{M}|^2_{\text{res,SM}} + \alpha\Lambda^{-2}|\mathcal{M}|^2_{\text{res,int}} + \alpha^2\Lambda^{-4}|\mathcal{M}|^2_{\text{res,BSM}} + O(\Lambda^{-6})$. Similarly, when using the narrow-width approximation in a production + decay process, the event rates take the form $\sigma_i \equiv \sigma_{\text{prod}}\Gamma_i/\Gamma_{\text{tot}}$, where σ_{prod} is a production cross section and Γ_i a partial decay width. Again, the three quantities σ_{prod}, Γ_i, Γ_{tot} are, in principle, quadratic functions of α. However, if the EFT expansion is valid, one can always expand σ_i to quadratic order so that $\sigma_i = \sigma^{\text{SM}} + \alpha\Lambda^{-2}\sigma^{\text{int}} + \alpha^2\Lambda^{-4}\sigma^{\text{BSM}} + O(\Lambda^{-6})$.

We can see that the EFT expansion always allows us to express the event rate as a quadratic function in α/Λ^2, even in the case of resonances. The discussion of this section applies similarly to the case of an arbitrary number of effective operators.

A comment on the relevance of rare events.—One may remark that in the case where we have $|\mathcal{M}^{\text{SM}}| \gg |\alpha\Lambda^{-2}\mathcal{M}^{\text{BSM}}|$ and an unsuppressed interference, the signal is proportional to α/Λ^2. On the other hand, in the case $|\mathcal{M}^{\text{SM}}| \ll |\alpha\Lambda^{-2}\mathcal{M}^{\text{BSM}}|$, the signal is proportional to α^2/Λ^4. The absolute magnitude of the signal is thus much larger in the former case than in the latter one, but the magnitude of the *background* is also larger in the former case. We may therefore wonder which of these two configurations is the more advantageous in order to detect the signal. The answer to this seemingly straightforward question is not so easy and needs to involve a test statistic.

Assume a signal discovery test whose significance Z is given by [12]

$$Z = \frac{N - N_{\text{bkg}}}{\sqrt{N_{\text{bkg}}}}, \qquad (16)$$

where N is the total number of events and N_{bkg} is the expected number of background events. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be two domains of phase space satisfying

$$\sigma^{\text{BSM}}_1 \approx \sigma^{\text{BSM}}_2, \qquad \sigma^{\text{SM}}_1 \gg \left|\frac{\alpha}{\Lambda^2}\sigma^{\text{int}}_1\right| \gg \frac{\alpha^2}{\Lambda^4}\sigma^{\text{BSM}}_1,$$

$$\sigma^{\text{SM}}_2 \ll \left|\frac{\alpha}{\Lambda^2}\sigma^{\text{int}}_2\right| \ll \frac{\alpha^2}{\Lambda^4}\sigma^{\text{BSM}}_2, \qquad (17)$$

where $\sigma^{\text{SM}}_{1,2}$, $\sigma^{\text{int}}_{1,2}$, and $\sigma^{\text{BSM}}_{1,2}$ are the components of the event rate σ on $\mathcal{D}_{1,2}$. Denoting Z_1 and Z_2 the discovery significance on $\mathcal{D}_1$ and $\mathcal{D}_2$, respectively, we have

$$\frac{Z_1}{Z_2} \ll 2. \qquad (18)$$

The proof is given in Appendix A. This small theorem makes it clear that the regions of phase space where $|\mathcal{M}^{\text{SM}}| \ll |\alpha\Lambda^{-2}\mathcal{M}^{\text{BSM}}|$, i.e., rare-events regions, can provide a lot of statistical significance, even though the signal is much weaker in that region compared to the $|\mathcal{M}^{\text{SM}}| \gg |\alpha\Lambda^{-2}\mathcal{M}^{\text{BSM}}|$ region. The signal search in such rare-events regions thus deserves special attention. We further emphasize that this rare-events situation naturally happens whenever a higher-dimensional operator carries derivatives, as is the case for many of the SM dimension-6 operators. For such operators, the high-energy tails of the kinematic distributions are typically rare-events zones. Thus, the analysis of high-energy tails also deserves special attention. For example, regarding the presentation of results, the use of overflow bins for the high-energy tail should be prevented as much as possible in order not to lose the precious information from the tail. Instead, all the bins should be kept up to the highest energetic event, no matter how few events are contained in the bins.

III. OPTIMAL ESTIMATION OF DIFFERENTIAL RATES

In the previous section we have considered event rates over a region of phase space $\mathcal{D}$. Let us now assume that the experiment allows us to partition the region $\mathcal{D}$ into subdomains such that $\mathcal{D} = \cup_r \mathcal{D}_r$. We refer to the $\mathcal{D}_r$ subdomains as "bins." In addition to the total rate, the knowledge of the event rate in each bin provides information about the *shape* of the distribution. From the experimental point of view, one talks about binned distribution. In principle, the size of each bin can be made small enough so that each event is seen separately. In this case one talks about unbinned distribution. Notice that, since the experimental precision is finite, the bin size can never shrink to zero; however, it is instructive to keep in mind that the latter case can be seen as a limit of the former.

Let us first consider the case of one single dimension-6 effective operator whose contribution to the amplitudes is

given by Eq. (5). Let X be a variable of phase space with domain $\mathcal{D}_X \subset \mathcal{D}$. For a collider experiment X can be some transverse momentum, invariant mass, etc. One defines the "differential event rate" along X as

$$\sigma_X \equiv \frac{d\sigma}{dX} \propto \int_{\mathcal{D}\backslash\mathcal{D}_X} d\Phi \left| \mathcal{M}^{\text{SM}} + \frac{\alpha}{\Lambda^2} \mathcal{M}^{\text{BSM}} \right|^2, \quad (19)$$

where $\int_{\mathcal{D}\backslash\mathcal{D}_X}$ means that the integral is performed over the complement of $\mathcal{D}_X$. Integrating σ_X over the domain $\mathcal{D}_X$, one recovers the total rate, namely, $\int_{\mathcal{D}_X} dX \sigma_X = \sigma$. The bins $\mathcal{D}_{Xr}$ over the domain $\mathcal{D}_X$ are defined by the partition $\mathcal{D}_X = \cup_r \mathcal{D}_{Xr}$ and the event rate σ_r on a bin $\mathcal{D}_{Xr}$ is then given by

$$\int_{\mathcal{D}_{Xr}} dX \sigma_X = \sigma_r. \quad (20)$$

The set of rates on every bin $\{\sigma_r\}$ forms a histogram that constitutes a discrete estimator of the true differential distribution σ_X.

Let us now proceed to decompose the differential rate σ_X defined in Eq. (19) as a sum of three components, σ_X^{SM}, σ_X^{int}, and σ_X^{BSM}, in complete analogy to the case of the total rate σ. We have

$$\sigma_X = \sigma_X(\alpha) \equiv \sigma_X^{\text{SM}} + \frac{\alpha}{\Lambda^2} \sigma_X^{\text{int}} + \frac{\alpha^2}{\Lambda^4} \sigma_X^{\text{BSM}}, \quad (21)$$

where

$$\sigma_X^{\text{SM}} \propto \int_{\mathcal{D}\backslash\mathcal{D}_X} d\Phi |\mathcal{M}^{\text{SM}}|^2, \quad (22)$$

$$\sigma_X^{\text{int}} \propto \int_{\mathcal{D}\backslash\mathcal{D}_X} d\Phi \, 2\text{Re}[\mathcal{M}^{\text{BSM}}\mathcal{M}^{\text{SM*}}],$$

$$\sigma_X^{\text{BSM}} \propto \int_{\mathcal{D}\backslash\mathcal{D}_X} d\Phi |\mathcal{M}^{\text{BSM}}|^2. \quad (23)$$

The same is also true for the binned rates

$$\sigma_r = \sigma_r(\alpha) \equiv \sigma_r^{\text{SM}} + \frac{\alpha}{\Lambda^2} \sigma_r^{\text{int}} + \frac{\alpha^2}{\Lambda^4} \sigma_r^{\text{BSM}}. \quad (24)$$

The phase space integral is usually difficult or impossible to evaluate analytically, for example, because of the complexity of $\mathcal{D}$. Its evaluation then has to rely on a numerical integration method, for instance, a Monte Carlo (MC) simulation. In the following we assume that such an estimation method is available.

A. Reconstructing the differential rates

Assuming we have a way of evaluating a differential rate in the presence of effective operators with coefficients fixed to given values, we can now figure out how to efficiently determine the deformations induced by the effective operators.

We first consider the case of a unique dimension-6 operator. We have seen in Sec. II that the expansion of the event rate has to be kept up to quadratic order in α, and the same argument also applies to the differential rate σ_X. Since σ_X is a quadratic function of α [see Eq. (21)], in principle, it is sufficient to know σ_X for only *three* different values of α, namely, α_0, α_1, and α_2, in order to reconstruct the exact form of $\sigma_X(\alpha)$. Whenever these three evaluations of σ_X are available, which we denote by $\sigma_X^i = \sigma_X(\alpha_i)$, $i = 0, 1, 2$, then the three components σ_X^{SM}, σ_X^{int}, and σ_X^{BSM} are obtained by simply solving a 3×3 linear system; this simple task needs to be carried out only once.

In the case in which the estimations are made by means of Monte Carlo simulations, one directly deals with the binned rates σ_r which are extracted from the histogram of the σ_X distribution. The three components, σ_r^{SM}, σ_r^{int}, σ_r^{BSM}, are obtained by solving a 3×3 linear system for each bin. The σ_r^{SM} component can be obtained by simply setting $\alpha = \alpha_0 = 0$ in the MC simulation. The components σ_r^{int} and σ_r^{BSM} are instead obtained by running the MC simulation for *two* nonzero values of α, namely, α_1 and α_2. Therefore, we end up with the following solution of our linear system:

$$\sigma_r^{\text{SM}} = \sigma_r^0$$

$$\sigma_r^{\text{int}} = \frac{\Lambda^2}{\alpha_1\alpha_2} \left[\frac{\alpha_2^2\sigma_r^1 - \alpha_1^2\sigma_r^1}{\alpha_2 - \alpha_1} - (\alpha_1 + \alpha_2)\sigma_r^0 \right]$$

$$\sigma_r^{\text{BSM}} = \frac{\Lambda^4}{\alpha_1\alpha_2} \left[-\frac{\alpha_2\sigma_r^1 - \alpha_1\sigma_r^2}{\alpha_2 - \alpha_1} + \sigma_r^0 \right], \quad (25)$$

where $\sigma_r^0 = \sigma_r(0)$, $\sigma_r^1 = \sigma_r(\alpha_1)$, and $\sigma_r^2 = \sigma_r(\alpha_2)$. These components can then be used in Eq. (24), "reconstructing" the formula that gives σ_r for any value of α.

Let us consider the general case of n effective operators. We present only the case of dimension-6 operators for simplicity. A similar analysis applies to the case of operators with arbitrary dimension. The amplitude has, in general, the form

$$\mathcal{M} = \mathcal{M}^{\text{SM}} + \frac{1}{\Lambda^2} \sum_{I=1}^{n} \alpha_I \mathcal{M}_I^{\text{BSM}} + O(\Lambda^{-4}). \quad (26)$$

The n BSM contributions $\mathcal{M}_I^{\text{BSM}}$ are, in general, different from each other. The differential event rate σ_X is proportional to the squared modulus of the amplitude and can be decomposed as

$$\sigma_X = \sigma_X(\alpha_I) = \sigma_X^{\text{SM}} + \frac{1}{\Lambda^2} \sum_{I=1}^{n} \alpha_I \sigma_{X,I}^{\text{int}} + \frac{1}{\Lambda^4} \sum_{I,J=1}^{n} \alpha_I \alpha_J \sigma_{X,IJ}^{\text{BSM}}. \quad (27)$$

It is convenient to rewrite the sum of the BSM quadratic contributions as

$$\sum_{I,J=1}^{n} \alpha_I \alpha_J \sigma_{X,IJ}^{\text{BSM}} = \sum_{I=1}^{n} \alpha_I^2 \sigma_{X,II}^{\text{BSM}} + 2 \sum_{I,J=1,I>J}^{n} \alpha_I \alpha_J \sigma_{X,IJ}^{\text{BSM}} \quad (28)$$

where $\sigma_{X,IJ}^{\text{BSM}} = \text{Re}(\mathcal{M}_I^{\text{BSM}} \mathcal{M}_J^{\text{BSM}*})$. The last piece of Eq. (28) corresponds to the interference term of the effective operators.

In addition to n interference terms $\sigma_{X,I}^{\text{int}}$, we have n terms $\sigma_{X,II}^{\text{BSM}}$ and $n(n-1)/2$ terms $\sigma_{X,IJ}^{\text{BSM}}$. Including the σ_X^{SM} component, the total number of terms to compute is $(n+1)(n+2)/2$. We conclude that $(n+1)(n+2)/2$ simulations are enough to exactly know the event rate σ_X as a function of the operator coefficients. The components are obtained by simply solving a $(n+1)(n+2)/2 \times (n+1)(n+2)/2$ linear system. This operation needs to be done only once. If one uses histograms, the system has to be solved once for each bin, just like in the case of a single operator.

We conclude that the complexity of the simple reconstruction method described above grows quadratically with the number of operators, once simplifications provided by the EFT expansion are taken into account.

B. Minimizing the uncertainties

The numerical evaluations of σ_X that are needed to reconstruct $\sigma_X(\alpha)$ are usually endowed with an intrinsic uncertainty. The correct approach is to think about estimators of σ_X for each value of the chosen α. In the case of a single dimension-6 operator, there are three estimators, and we denote them by $\hat{\sigma}_X^i = \hat{\sigma}_X(\alpha_i)$, $i = 0, 1, 2$. The uncertainties associated with these estimators are naturally propagated to the σ_X^{SM}, σ_X^{int}, σ_X^{BSM} components because these quantities are linear combinations of the σ_X^i. In turn, these uncertainties are propagated to the reconstructed $\sigma_X(\alpha)$ distribution because of the relation in Eq. (21) and, when it comes to confronting $\sigma_X(\alpha)$ with the observed distribution σ_X^{obs}, the uncertainty on $\sigma_X(\alpha)$ should be taken into account. It is therefore important to have a measure of this uncertainty, which is given by the covariance matrix of the estimators of the σ_X^{SM}, σ_X^{int}, σ_X^{BSM} components.

In principle, there is a freedom in choosing the α coefficients that are used to perform the numerical estimations of σ_X which are needed to reconstruct $\sigma_X(\alpha)$. If there were no uncertainties, any set of values would work fine. However, in the presence of uncertainties, it turns out that the choice of the α coefficients has a crucial impact on the reconstruction uncertainties. In the following we will determine the values of α that *minimize* these uncertainties.

Let us work with the binned distributions σ_r and focus on a bin r. The three estimators are denoted by $\hat{\sigma}_r^0 = \hat{\sigma}_r(\alpha_0)$, $\hat{\sigma}_r^1 = \hat{\sigma}_r(\alpha_1)$, $\hat{\sigma}_r^2 = \hat{\sigma}_r(\alpha_2)$. We introduce the relative variance of each estimator $\bar{V}_r^i$, which is given by

$$\bar{V}_r^i = \frac{\text{E}[(\hat{\sigma}_r^i)^2] - \text{E}[\hat{\sigma}_r^i]^2}{\text{E}[\hat{\sigma}_r^i]^2} \quad (i = 0, 1, 2), \quad (29)$$

where $E[\hat{y}]$ represents the expectation value of the random variable $\hat{y}$. No correlation is assumed among estimators related to different values of α_i. Furthermore, we assume that the three relative variances have the same magnitude, namely,

$$\bar{V}_r^0 \sim \bar{V}_r^1 \sim \bar{V}_r^2 \equiv \bar{V}_r. \quad (30)$$

In the case where the estimators are obtained through Monte Carlo simulations with N_{MC} points, we have $\bar{V}_r \sim 1/N_{\text{MC}}$.

The estimators of interest to us are $\hat{\sigma}_r^{\text{SM}}$, $\hat{\sigma}_r^{\text{int}}$, and $\hat{\sigma}_r^{\text{BSM}}$, which are expressed in terms of some linear combinations of the $\hat{\sigma}_r^i$ as shown in Eq. (25). The quantity at the center of our analysis is the relative covariance matrix of these estimators, which is given by

$$\bar{C}_r(\alpha_0, \alpha_1, \alpha_2)$$

$$= \begin{pmatrix} \frac{\text{E}[(\hat{\sigma}_r^{\text{SM}})^2]}{\text{E}[\hat{\sigma}_r^{\text{SM}}]^2} - 1 & \frac{\text{E}[\hat{\sigma}_r^{\text{SM}}\hat{\sigma}_r^{\text{int}}]}{\text{E}[\hat{\sigma}_r^{\text{SM}}]\text{E}[\hat{\sigma}_r^{\text{int}}]} - 1 & \frac{\text{E}[\hat{\sigma}_r^{\text{SM}}\hat{\sigma}_r^{\text{BSM}}]}{\text{E}[\hat{\sigma}_r^{\text{SM}}]\text{E}[\hat{\sigma}_r^{\text{BSM}}]} - 1 \\ & \frac{\text{E}[(\hat{\sigma}_r^{\text{int}})^2]}{\text{E}[\hat{\sigma}_r^{\text{int}}]^2} - 1 & \frac{\text{E}[\hat{\sigma}_r^{\text{int}}\hat{\sigma}_r^{\text{BSM}}]}{\text{E}[\hat{\sigma}_r^{\text{int}}]\text{E}[\hat{\sigma}_r^{\text{BSM}}]} - 1 \\ & & \frac{\text{E}[(\hat{\sigma}_r^{\text{BSM}})^2]}{\text{E}[\hat{\sigma}_r^{\text{BSM}}]^2} - 1 \end{pmatrix}.$$

$$(31)$$

It is important to notice that, for the sake of determining the σ_X^{SM}, σ_X^{int}, and σ_X^{BSM} components, there is *no* need to use values of α and Λ that respect the EFT validity bounds of Eq. (3). Note that to apply this trick, the possible higher-order terms in the expressions of the event rates must be set to zero to avoid any disturbance [13].

After these preliminary steps, we can discuss the uncertainties that affect the $\hat{\sigma}_r^{\text{SM}}$, $\hat{\sigma}_r^{\text{int}}$, and $\hat{\sigma}_r^{\text{BSM}}$ estimators. We aim at simultaneously minimizing the uncertainties for every component. We should therefore consider the trace of the relative covariance matrix of Eq. (31). Let us first notice that choosing $\alpha_0 = 0$ simply gives $\hat{\sigma}_r^{\text{SM}} = \hat{\sigma}_r^0$, and therefore the relative variance for $\hat{\sigma}_r^{\text{SM}}$ is simply $\bar{V}_r^{\text{SM}} = \bar{V}_r$. This choice is arguably optimal, and we are left with finding optimal values for α_1 and α_2.

For a fixed value of α_1, it turns out that $\text{tr}\bar{C}_r(0, \alpha_1, \alpha_2)$ is minimized for α_2 going to infinity [14], and this runaway direction can be seen in Fig. 2. In this limit the $\hat{\sigma}_r^2$ estimator corresponds exactly to $\hat{\sigma}_r^{\text{BSM}}$ with relative variance equal to $\bar{V}_r$ and vanishing correlation with $\hat{\sigma}_r^{\text{SM}}$. Also in this limit the relative covariance matrix takes the form

$$\bar{C}_r(0, \alpha_1, \infty) = \bar{V}_r \begin{pmatrix} 1 & -\frac{\Lambda^2}{\alpha_1}\frac{\sigma_r^{\text{BSM}}}{\sigma_r^{\text{int}}} & 0 \\ -\frac{\Lambda^2}{\alpha_1}\frac{\sigma_r^{\text{SM}}}{\sigma_r^{\text{int}}} & 1 + 2\frac{\Lambda^4\sigma^{\text{SM}}+\alpha_1^2\sigma_r^{\text{BSM}}}{\alpha_1\Lambda^2\sigma_r^{\text{int}}} + 2\frac{\Lambda^8(\sigma^{\text{SM}})^2+\alpha_1^2\Lambda^4\sigma^{\text{SM}}\sigma_r^{\text{BSM}}+\alpha_1^4(\sigma_r^{\text{BSM}})^2}{\alpha_1^2\Lambda^4(\sigma_r^{\text{int}})^2} & -\frac{\alpha_1}{\Lambda^2}\frac{\sigma_r^{\text{BSM}}}{\sigma_r^{\text{int}}} \\ 0 & -\frac{\alpha_1}{\Lambda^2}\frac{\sigma_r^{\text{BSM}}}{\sigma_r^{\text{int}}} & 1 \end{pmatrix}. \tag{32}$$

The trace of the relative variance matrix goes to infinity for both $\alpha_1 \to 0, \infty$ because the linear system becomes degenerate in these limits. Further studying the behavior of $\text{tr}\bar{C}_r(0, \alpha_1, \infty)$, we find that it admits two minima for

$$\alpha_1 = \pm\Lambda^2\sqrt{\frac{\sigma_r^{\text{SM}}}{\sigma_r^{\text{BSM}}}}. \tag{33}$$

It is a nontrivial feature that this expression is independent of σ_r^{int}. This implies that the optimization does not depend on whether the interference is suppressed.

The behavior of the trace of $\bar{C}_r(0, \alpha_1, \alpha_2)$ is shown in Fig. 2, where we have taken Λ to be such that Eq. (C5) is verified. One can observe from the plot that the minimum of the trace of the relative covariance matrix occurs for $\alpha_1 = 1$, $\alpha_2 = \infty$ or vice versa. The relative covariance matrix at the positive optimal α_1 takes the following form:

$$\bar{C}_r^{\min} = \bar{V}_r \begin{pmatrix} 1 & -\frac{\bar{\sigma}_r}{\sigma_r^{\text{int}}} & 0 \\ -\frac{\bar{\sigma}_r}{\sigma_r^{\text{int}}} & 1 + 4\frac{\bar{\sigma}_r}{\sigma_r^{\text{int}}} + 6(\frac{\bar{\sigma}_r}{\sigma_r^{\text{int}}})^2 & -\frac{\bar{\sigma}_r}{\sigma_r^{\text{int}}} \\ 0 & -\frac{\bar{\sigma}_r}{\sigma_r^{\text{int}}} & 1 \end{pmatrix}, \tag{34}$$
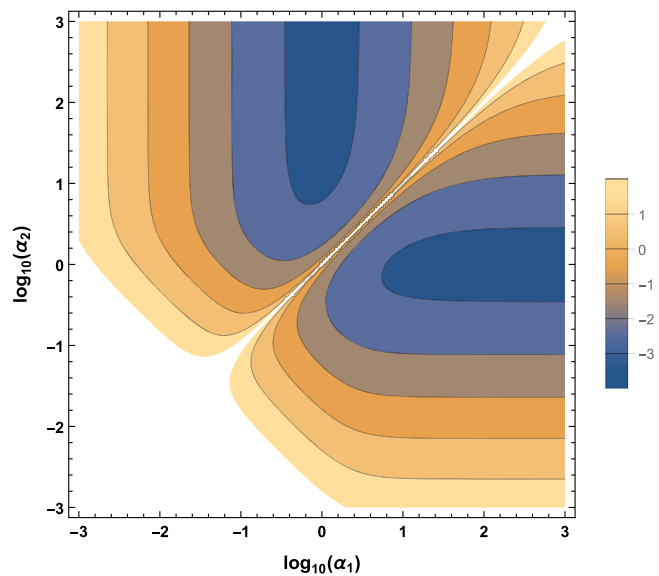

FIG. 2. Trace of the relative covariance of the $\hat{\sigma}^{\text{SM}}$, $\hat{\sigma}^{\text{int}}$, $\hat{\sigma}^{\text{BSM}}$ estimators as a function of α_1, α_2. One assumes $\sigma^{\text{SM}} \sim \sigma^{\text{BSM}}\Lambda^{-4}$, a positive interference saturating the Cauchy-Schwartz bound, and an event number $N_{\text{MC}} = 10^4$ for each simulation.

where

$$\bar{\sigma}_r \equiv \sqrt{\sigma_r^{\text{SM}}\sigma_r^{\text{BSM}}}. \tag{35}$$

Notice that since the minimal relative covariance is, by definition, dimensionless, it can depend on the components of the differential rate only through the combination $\bar{\sigma}_r/\sigma_r^{\text{int}}$. This quantity is also the one appearing in the Cauchy-Schwartz bound of Eq. (10), which can now be rewritten as $|\sigma_r^{\text{int}}| \le 2\bar{\sigma}_r$.

Considering the two values of σ_r^{int} that saturate the Cauchy-Schwartz bound, the relative covariance matrix at the minimum assumes the following form:

$$\bar{V}_r \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{9}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \quad \text{and} \quad \bar{V}_r \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{pmatrix}, \tag{36}$$

for $\sigma_r^{\text{int}} = +2\bar{\sigma}_r$ and $\sigma_r^{\text{int}} = -2\bar{\sigma}_r$, respectively. The relative uncertainty on the interference component is given by the square root of $(\bar{C}_r)_{22}$. From Eq. (36) one sees that the uncertainty on the interference component is larger than the ones on σ_r^{SM}, σ_r^{BSM} by at least a factor $3/\sqrt{2}$ when σ_r^{int} is positive. In the case of a maximal negative interference, the optimal α_1 corresponds, in fact, to a fully destructive interference, $\sigma_r^1 = 0$, and the expected uncertainty associated with this result is vanishing. This can be directly seen in the covariance matrix [Eq. (36), right side], which features a zero eigenvalue. This situation of a large negative interference gives the smallest uncertainties possible. In contrast, whenever the interference is suppressed, $|\sigma_r^{\text{int}}| < 2\bar{\sigma}_r$, the uncertainty on the interference component quickly blows up. For example, taking $\sigma_r^{\text{int}} = \pm0.2\bar{\sigma}_r$, we get $(\bar{C}_r)_{22} \approx 175$ and $(\bar{C}_r)_{22} \approx 130$, respectively. Our calculation provides a quantitative estimate of the difficulties that will be encountered for determining the interference component.

If one evaluates the minimal relative covariance matrix at the negative optimal α_1 of Eq. (33), Eq. (36) and the subsequent expressions and discussions are valid up to a sign flip of σ_r^{int}. Let us notice that if the sign of σ_r^{int} is known in advance, the sign of the optimal α_1 can then be chosen so that the interference term is negative, which gives a better uncertainty on the estimation of σ_r^{int}. But this refinement matters mostly for an interference near the Cauchy-Schwartz bound.

Finally, one may notice that the optimization is, in principle, different for each bin, as made clear by the r dependence of the optimal point given in Eq. (33). However, a simpler version of the optimization can also be obtained by using the α_1 that minimizes the uncertainty on the total rates, namely, $\alpha_1 = \Lambda^2 (\sigma^{\mathrm{SM}}/\sigma^{\mathrm{BSM}})^{1/2}$.

C. Case of n operators

Let us discuss the reconstruction method in the case of n effective operators. We have seen in Sec. III A that $(n+1)(n+2)/2$ simulations have to be performed to fully reconstruct the event rate σ_X as a function of the α_I's. In practice, it turns out that it is sufficient to switch on one or two effective operators at a time in order to reconstruct all the components. Using the results of the previous section, we can provide the complete list of $(n+1)(n+2)/2$ optimized points to be used for the simulations. We have

(1) One point with all $\alpha_I = 0$.
(2) For every I, a single nonzero α_I satisfying $\alpha_I \gg 1$. These are n points.
(3) For every I, a single nonzero α_I satisfying $\alpha_I \sim \Lambda^2 \frac{(\sigma_X^{\mathrm{SM}})^{1/2}}{(\sigma_{X,II}^{\mathrm{BSM}})^{1/2}}$. These are n points.
(4) For every pair (I,J), a single nonzero pair (α_I, α_J) satisfying $\alpha_{I,J} \gg 1$ and $\alpha_I \sigma_{X,II}^{\mathrm{BSM}} \sim \alpha_J \sigma_{X,JJ}^{\mathrm{BSM}}$. These are $n(n-1)/2$ points.

We thus end up with a system of $(n+1)(n+2)/2$ linear equations that has to be solved. A useful side effect of the optimization described in the previous subsection is that the calculation of the various components becomes more transparent. In particular, point (1) provides σ_X^{SM}, and point (2) provides the $\sigma_{X,II}^{\mathrm{BSM}}$ components. Point (3) provides the interference terms $\sigma_{X,I}^{\mathrm{int}}$ which can simply be obtained by subtracting $\sigma_X^{\mathrm{SM}} + \alpha_I^2 \Lambda^{-4} \sigma_{X,II}^{\mathrm{BSM}}$ from the outcome of the simulations in point (3), instead of using the exact formulas of Eqs. (25). This simplification is a consequence of having used arbitrary large α_I's in point (2). Similarly, point (4) provides the $\sigma_{X,IJ}^{\mathrm{BSM}}$ components, which are obtained by subtracting $\alpha_I^2 \sigma_{X,II}^{\mathrm{BSM}} + \alpha_J^2 \sigma_{X,JJ}^{\mathrm{BSM}}$ from the outcome of the simulations in point (4). An analysis similar to the one of Sec. III B shows that the uncertainty on the $\sigma_{X,II}^{\mathrm{BSM}}$ terms is minimized for $\alpha_I \sigma_{X,II}^{\mathrm{BSM}} \sim \alpha_J \sigma_{X,JJ}^{\mathrm{BSM}}$.

Comments on Monte Carlo simulations.—Differential distributions are often estimated using Monte Carlo simulations, which reproduce the experimental setup assuming a fixed number of events N_{MC} or a fixed integrated luminosity $\mathscr{L}_{\mathrm{MC}}$. The expected, relative uncertainties associated with estimation of the event rate are given by $\sqrt{V} \equiv 1/\sqrt{N_{\mathrm{MC}}}(1 + O(N_{\mathrm{MC}}^{-1}))$ in both cases (see Appendix B).

In order to search for a small deviation in a given set of data, one should require that the MC error be small with respect to the statistical error of the data. For binned data, the number of events in each bin $\hat{N}_r$ is Poisson distributed,

and its relative statistical error is given by $1/\sqrt{\hat{N}_r}$. The requirement that the MC error be small with respect to the experimental error in every bin translates into the condition

$$N_{\mathrm{MC},r} \gg \hat{N}_r \quad \text{for any bin } r. \tag{37}$$

Note that this condition depends on the actual data one wants to analyze. It applies for both background-only and signal hypothesis, i.e., for both $\alpha = 0$ and $\alpha \neq 0$. Finally, the binning and range of all the MC histograms have to match the bins chosen for the data and are thus completely fixed.

IV. CONCRETE EXAMPLE: SEARCH FOR ANOMALOUS TRILINEAR GAUGE COUPLING AT THE LHC

To validate our reconstruction method, we apply it to the concrete example of the search for the dimension-6 effective operator $\mathcal{O}_{3W}$ in WW production at the LHC. The operator $\mathcal{O}_{3W}$ is defined as

$$\mathcal{O}_{3W} = \varepsilon_{ijk} W^i_{\mu\nu} W^{j,\nu}{}_\rho W^{k,\rho\mu}, \tag{38}$$

and its coefficient is denoted by $\frac{\alpha_{3W}}{\Lambda^2}$.

After electroweak symmetry breaking, it contributes to anomalous triple gauge couplings [15] that can be parametrized as follows [16]:

$$\begin{aligned} \mathcal{L}_{\mathrm{CGC}}^\partial = {} & \lambda^Z [i g_Z Z_{\mu\nu} (\hat{W}^-_{\nu\rho} \hat{W}^+_{\rho\mu} - \hat{W}^+_{\nu\rho} \hat{W}^-_{\rho\mu})] \\ & + \lambda^\gamma [i e F_{\mu\nu} (\hat{W}^-_{\nu\rho} \hat{W}^+_{\rho\mu} - \hat{W}^+_{\nu\rho} \hat{W}^-_{\rho\mu})], \end{aligned} \tag{39}$$

where $\hat{W}^+_{\mu\nu} = D_\mu W^+_\nu - D_\nu W^+_\mu$ and

$$\lambda^Z = \lambda^\gamma = 3 \frac{\alpha_{3W}}{g\Lambda^2}. \tag{40}$$

The Lagrangian $\mathcal{L}_{\mathrm{CGC}}^\partial$ induces new vertices among the weak gauge bosons which carry extra derivatives with respect to the Standard Model ones. This new interaction will potentially deform the WW differential rates at the LHC, especially in the high-energy range of the distributions.

A search for the $\mathcal{O}_{3W}$ operator has been performed by the CMS Collaboration in [17], where they consider $W^+ W^-$ production in the leptonic decay channel at the LHC, with an integrated luminosity of 19.4 fb^{-1} at 8 TeV center-of-mass energy. In [17] the same operator is defined using a different normalization convention, and the translation to our notation is done by using the following relation:

$$\alpha_{3W} = \frac{g^3}{4} c_{WWW}. \qquad (41)$$

Our aims are as follows:

(i) Determining the deformations of the differential rates in W^+W^- production induced by the $\mathcal{O}_{3W}$ operator using our optimized technique for Monte Carlo simulations.

(ii) Deriving a 95% CL bound on $\frac{\alpha_{3W}}{\Lambda^2}$ using the measured differential distributions of [17].

Therefore, we consider the process $pp \to W^+W^- \to l^+\nu l^-\bar{\nu}$ ($l = e, \mu$) at 8 TeV center-of-mass energy.

The measured unfolded differential distributions for this process are displayed in Fig. 3 of [17], and they can be consistently compared to the outcome of our MC simulations. The dilepton invariant mass (m_{ll}) distribution in the "0-jet category" is the one chosen to put a bound on α_{3W}/Λ^2. The total number of W^+W^- + background events measured in the 0-jet category is ∼4800, and the quoted 95% CL bound [17] translated to our notation is

$$-0.39 < \frac{\alpha_{3W}}{\Lambda^2} < 0.41 \text{ TeV}^{-2}. \qquad (42)$$

We simulate events for W^+W^- production at 8 TeV with MadGraph5 [18] after having implemented the $\mathcal{O}_{3W}$ operator in FeynRules2.0 [19]. These events are showered with Pythia8 [20] and selected using the cuts chosen in the CMS analysis. In particular, the leptons are required to have $p_T > 20$ GeV and $|\eta| < 2.5$. Events with one or more jets with $p_T > 30$ GeV and $|\eta| < 4.7$ are rejected.

Following the notation introduced in the previous sections, the differential rate along the m_{ll} variable will be denoted by $\sigma_{m_{ll}}$. We consider the binned distributions for $\sigma_{m_{ll}}$ in the range [20, 200] GeV, according to the choice made in the CMS analysis.

We evaluate the components of the binned m_{ll} distribution following our optimal method described in Sec. III. We first compute the SM component $\sigma_{m_{ll}}^{SM}$ by setting $\alpha_{3W} = 0$ in

our Monte Carlo simulation. Then we compute the m_{ll} distribution for a very large value of α_{3W}/Λ^2, chosen to be 272 TeV^{-2}, which turns out to be proportional to the BSM component $\sigma_{m_{ll}}^{BSM}$ to a very good approximation. The binned $\sigma_{m_{ll}}^{SM}$ and $\sigma_{m_{ll}}^{BSM}$ components are shown in Fig. 3.

In order to compute the interference component $\sigma_{m_{ll}}^{int}$ following the results of Sec. III B, we used Eq. (33) to determine the third optimal value of α_{3W}/Λ^2. Since the σ^{SM}/σ^{BSM} ratio is roughly 100 TeV^{-4} in most of the bins, we conclude that the third optimal point for the simulation is roughly $\alpha_{3W}/\Lambda^2 \approx 10$ TeV^{-2}.

We still need to set the size of our MC samples. We require, for every bin, the number of MC events to be larger than the observed data analyzed by the CMS Collaboration in the search for the $\mathcal{O}_{3W}$ operator. For example, $N_{MC} = 5 \times 10^5$ events would give a MC uncertainty that is typically 10 times smaller than the statistical ones. Although this number is enough for the sake of analyzing the CMS distribution, it turns out that this amount is not sufficient to resolve the binned interference components because the interference is further suppressed by a m_W^2/E^2 factor with respect to the naive expectation. This can be understood thanks to some helicity selection rules [11,21].

A much larger number of events is needed to properly estimate the interference component. We have indeed used $N_{MC} = 2.4 \times 10^6$ events for each of the three simulation points $\alpha_{3W}/\Lambda^2 = 0$, 8.6, and 272 TeV^{-2}. The relative covariance matrix calculated in Sec. III B readily provides the uncertainty on the $\sigma_{m_{ll}}^{SM}$, $\sigma_{m_{ll}}^{int}$, and $\sigma_{m_{ll}}^{BSM}$ components.

The uncertainties on $\sigma_{m_{ll}}^{SM}$ and $\sigma_{m_{ll}}^{BSM}$ are too small to be visible in Fig. 3. In contrast, the uncertainty on $\sigma_{m_{ll}}^{int}$ is not negligible, even with such a large MC sample. Having fixed a unique optimized value of α_1 for every bin and taking into account that the number of MC events in a given bin depends on this α_1, the uncertainty on σ_r^{int} is obtained from Eq. (C2), where the relative variance associated with a given bin is
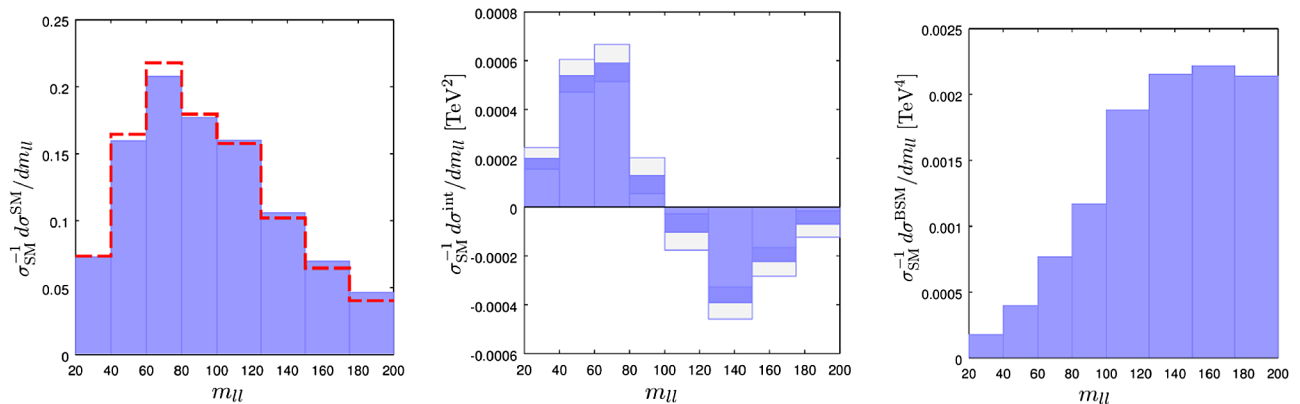

FIG. 3. Differential distributions of dilepton invariant mass. The $\sigma_{m_{ll}}^{SM}, \sigma_{m_{ll}}^{int}, \sigma_{m_{ll}}^{BSM}$ components normalized to σ^{SM} are shown from left to right. The differential distribution from CMS data is shown in red.
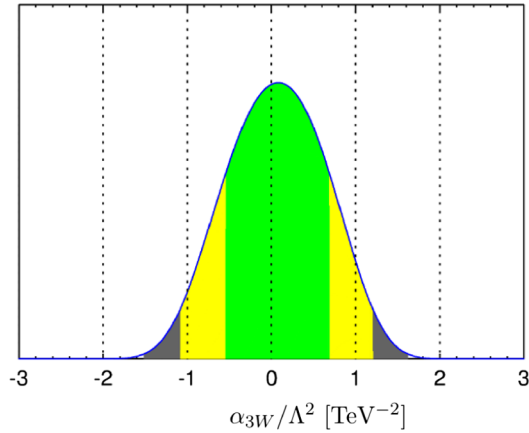
FIG. 4. Posterior density probability for the coefficient of $\mathcal{O}_{3W}$. Green, yellow, and grey areas correspond to Bayesian credible regions with, respectively, 68.27%, 95.45%, 99.73% probability.

$$\bar{V}_r^i = \frac{\sigma_{\text{tot}}^i}{\sigma_r^i} \frac{1}{N_{\text{MC}}}. \tag{43}$$

These uncertainties are shown in Fig. 3.

Finally, we compare our reconstructed differential rate for $pp \to W^+W^- \to l^+\nu l^-\bar{\nu}$ to the measured one shown in Fig. 3 of the CMS study [17]. The uncertainties quoted are a combination of statistical and systematic errors. Because combinations of many sources of uncertainties tend to be governed by the central limit theorem [22], we approximate the likelihood for each bin as a Gaussian. The complete likelihood used in our analysis is

$$L(\alpha_{3W}) = \prod_r \exp\left[-\frac{1}{2\Delta_r^2}\left(\frac{\sigma_r(\alpha_{3W}) - \sigma_r^{\text{obs}}}{\sigma_r^{\text{SM}}}\right)^2\right], \tag{44}$$

where $\sigma_r^{\text{obs}}/\sigma_r^{\text{SM}}$ are the experimental numbers given in [17] and Δ_r are the combined uncertainties for each bin, which are typically $\sim 8\%$. Using Eq. (44), we compute the credible intervals for the α_{3W}/Λ^2 parameter, assuming a flat prior over $[-100, 100]$ TeV. The posterior distribution for $\alpha_{3W}\Lambda^{-2}$ is shown in Fig. 4. We find

$$-1.09 < \frac{\alpha_{3W}}{\Lambda^2} < 1.19 \text{ TeV}^{-2} \quad \text{at } 95\% \text{ CL.} \tag{45}$$

This bound is in agreement within less than two sigma with the one reported by CMS [see Eq. (42)].

V. CONCLUSION

If new particles beyond the Standard Model are too heavy to be produced on shell at the LHC, their presence can still be indirectly detected via the effect of SM higher-dimensional operators. In such a scenario, the LHC precision physics program would play a central role for new physics searches. The key observables for revealing

the existence of higher-dimensional operators may be the distributions of final-state kinematic variables that contain precious information about new physics effects. The analysis of these differential rates thus deserves to be optimized in all its aspects.

We focus on the case where leading effects from new physics arise from dimension-6 operators. We first inspect the event rates and the problem of detecting the deformations induced by the presence of the effective operators. We make it clear that, in general, the pure BSM term should not be neglected in the differential rate analysis—even though it is $O(\alpha^2/\Lambda^4)$—since there can be regions of phase space where its contribution is dominant. We have also found a bound on the interference term which follows from the application of the Cauchy-Schwartz inequality. Using this bound, it can be quantitatively shown that regions of phase space with rare SM events are very important to search for deformations of the differential rates.

Based on this preliminary analysis, we determine an optimal method to obtain the different contributions to the differential rates in the presence of dimension-6 effective operators, assuming that an estimator (e.g., a Monte Carlo tool) of the distributions is available. In the case of n effective operators, the evaluation of the rate at $(n+1)(n+2)/2$ different points is needed. The various contributions to the differential rate are then simply obtained by solving a linear system.

A crucial aspect of the proposed method is the minimization of the uncertainty through an optimal choice of the higher-dimensional operator coefficient to be used in the simulations. In the case of a single dimension-6 operator, we have to estimate the differential rate for three values of the coefficient α. The analysis of the relative covariance matrix of the three estimators reveals that the uncertainty is minimized for values of α equal to zero, infinity, and $\pm\Lambda^2\sqrt{\sigma_X^{\text{SM}}/\sigma_X^{\text{BSM}}}$. Interestingly, this result turns out to be independent of the value of the interference. The covariance matrix provides the uncertainties on the estimated contributions and their correlations, allowing a well-defined control of the estimations. This covariance matrix should be, in principle, implemented in any subsequent statistical analysis. It turns out that the uncertainty on the interference component is larger than the ones on the SM and BSM components by a factor $3/\sqrt{2}$ if the interference saturates the positive Cauchy-Schwartz bound, and it grows very quickly if the interference is smaller than this value.

We illustrate and check our method by determining the deformations induced by the $\mathcal{O}_{3W}$ operator on leptonic final states from WW production, at the 8 TeV LHC. We work at the reconstruction level, aiming to approximately reproduce the analysis made by the CMS Collaboration in Ref. [17]. We ultimately reproduce, within a two-sigma range, the bound on α_{3W}/Λ^2 obtained in this CMS analysis.

ACKNOWLEDGMENTS

Note added.—After completing our work we became aware of similar developments made by members of the LHC Higgs Cross Section Working Group [23] and the ATLAS Collaboration [24]. The basic method presented in these references (called morphing) is essentially the same as our reconstruction method described in Sec. III A, although the parametrization that has been used is slightly different. However, the question of how the input parameters need to be chosen such that the expected uncertainty of the output is minimal has not been addressed in these studies. Our paper fills this important gap by presenting, for the first time, an optimal morphing method and the statistical approach which gives rise to it.

APPENDIX A: PROOF OF EQ. (18)

For our purpose it is enough to assume a p-value based discovery test, which leads to Eq. (16). A similar demonstration can be done using a Bayes factor. Using the inequality in Eq. (17) and keeping only the leading terms in the discovery tests, one gets

$$Z_1 \propto \frac{\alpha}{\Lambda}\frac{\sigma_1^{\mathrm{int}}}{\sqrt{\sigma_1^{\mathrm{SM}}}}, \qquad Z_2 \propto \frac{\alpha^2}{\Lambda^2}\frac{\sigma_2^{\mathrm{BSM}}}{\sqrt{\sigma_2^{\mathrm{SM}}}}. \qquad (A1)$$

Let us consider the ratio

$$\frac{Z_1}{Z_2} = \frac{\Lambda}{\alpha}\frac{\sigma_1^{\mathrm{int}}\sqrt{\sigma_2^{\mathrm{SM}}}}{\sqrt{\sigma_1^{\mathrm{SM}}\sigma_2^{\mathrm{BSM}}}}. \qquad (A2)$$

Using the Cauchy-Schwartz bound on σ_1^{int} [see Eq. (9)], one has

$$\frac{Z_1}{Z_2} \leq \frac{\Lambda}{\alpha}\frac{2\sqrt{\sigma_2^{\mathrm{SM}}\sigma_1^{\mathrm{BSM}}}}{\sigma_2^{\mathrm{BSM}}}. \qquad (A3)$$

Using the equality of BSM components made in the initial assumptions [see Eq. (17)], we obtain

$$\frac{Z_1}{Z_2} \ll 2\left(\frac{\Lambda^2}{\alpha^2}\frac{\sigma_2^{\mathrm{SM}}}{\sigma_2^{\mathrm{BSM}}}\right)^{1/2}. \qquad (A4)$$

The inequality contained in Eq. (17) then proves Eq. (18).

APPENDIX B: UNCERTAINTIES ON MONTE CARLO ESTIMATORS

For a fixed luminosity $\mathscr{L}_{\mathrm{MC}}$, the number of MC events $\hat{N}_{\mathrm{MC}}$ is Poisson distributed. The estimator of the event rate is given by

$$\hat{\sigma} = \frac{\hat{N}_{\mathrm{MC}}}{\mathscr{L}_{\mathrm{MC}}}, \qquad (B1)$$

which satisfies

$$\mathrm{E}[\hat{\sigma}] = \sigma \quad \text{and} \quad \mathrm{V}[\hat{\sigma}] = \frac{\sigma}{\mathcal{L}_{\mathrm{MC}}}, \qquad (B2)$$

where σ is the theoretical rate. The expected relative error associated with the estimation of σ is thus given by $1/\sqrt{\sigma\mathscr{L}_{\mathrm{MC}}}$.

Instead of a fixed luminosity, one can require a fixed number of events N_{MC}. In this case, the random variable is the MC luminosity $\hat{\mathscr{L}}_{\mathrm{MC}}$. The combination $\sigma\hat{\mathscr{L}}_{\mathrm{MC}}$ follows an Erlang distribution

$$\sigma\frac{(\sigma\hat{\mathscr{L}}_{\mathrm{MC}})^{N_{\mathrm{MC}}-1}}{(N_{\mathrm{MC}}-1)!}e^{-\sigma\hat{\mathscr{L}}_{\mathrm{MC}}}, \qquad (B3)$$

for which $\mathrm{E}[\sigma\hat{\mathscr{L}}_{\mathrm{MC}}] = \mathrm{V}[\sigma\hat{\mathscr{L}}_{\mathrm{MC}}] = N_{\mathrm{MC}}$. In this case, the estimator of the event rate is given by

$$\frac{1}{\hat{\sigma}} = \frac{\hat{\mathscr{L}}}{N_{\mathrm{MC}}}, \qquad (B4)$$

which satisfies

$$\mathrm{E}\left[\frac{1}{\hat{\sigma}}\right] = \frac{1}{\sigma} \quad \text{and} \quad \mathrm{V}\left[\frac{1}{\hat{\sigma}}\right] = \frac{1}{\sigma^2}\frac{1}{N_{\mathrm{MC}}}. \qquad (B5)$$

The expected relative error associated with the estimation of $1/\sigma$ is therefore given by $1/\sqrt{N_{\mathrm{MC}}}$. For large N_{MC}, one has that $\mathrm{E}[\hat{\sigma}] = E[1/\hat{\sigma}]^{-1}(1 + O(N_{\mathrm{MC}}^{-1}))$, and the relative uncertainty associated with the estimation of σ is given by $1/\sqrt{N_{\mathrm{MC}}}$ up to $O(1/N_{\mathrm{MC}})$ corrections.

APPENDIX C: GENERAL COVARIANCE MATRIX

The relative covariance matrix for the estimators $(\hat{\sigma}_r^{\mathrm{SM}}, \hat{\sigma}_r^{\mathrm{int}}, \hat{\sigma}_r^{\mathrm{BSM}})$, assuming arbitrary relative covariance for the $\hat{\sigma}_r^i$, namely, $\bar{V}_r^0 \neq \bar{V}_r^1 \neq \bar{V}_r^2$, is given by

$$\bar{C}_r(0, \alpha_1, \infty) = \begin{pmatrix} \bar{V}_r^0 & -\frac{\Lambda^2}{\alpha_1} \frac{\sigma_r^{\mathrm{BSM}}}{\sigma_r^{\mathrm{int}}} \bar{V}_r^0 & 0 \\ -\frac{\Lambda^2}{\alpha_1} \frac{\sigma_r^{\mathrm{SM}}}{\sigma_r^{\mathrm{int}}} \bar{V}_r^0 & (\bar{C}_r)_{22} & -\frac{\alpha_1}{\Lambda^2} \frac{\sigma_r^{\mathrm{BSM}}}{\sigma_r^{\mathrm{int}}} \bar{V}_r^2 \\ 0 & -\frac{\alpha_1}{\Lambda^2} \frac{\sigma_r^{\mathrm{BSM}}}{\sigma_r^{\mathrm{int}}} \bar{V}_r^2 & \bar{V}_r^2 \end{pmatrix}, \tag{C1}$$

where

$$(\bar{C}_r)_{22} = \bar{V}_r^1 \left(1 + 2 \frac{\Lambda^2(\sigma_r^{\mathrm{SM}} + \alpha_1^2 \Lambda^{-4} \sigma_r^{\mathrm{BSM}})}{\alpha_1 \sigma_r^{\mathrm{int}}} + 2 \frac{\sigma_r^{\mathrm{SM}} \sigma_r^{\mathrm{BSM}}}{(\sigma_r^{\mathrm{int}})^2} \right) + (\bar{V}_r^0 + \bar{V}_r^1) \frac{\Lambda^4 (\sigma_r^{\mathrm{SM}})^2}{\alpha_1^2 (\sigma_r^{\mathrm{int}})^2} + (\bar{V}_r^1 + \bar{V}_r^2) \frac{\Lambda^{-4} \alpha_1^2 (\sigma_r^{\mathrm{BSM}})^2}{(\sigma_r^{\mathrm{int}})^2}. \tag{C2}$$

---

[1] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).

[2] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).

[3] One-dimensional differential rates made available by AT-LAS and CMS in run-I Higgs analyses have been included in the global fits performed in Refs. [25,26].

[4] S. Fichet, Int. J. Mod. Phys. A **30**, 1550039 (2015).

[5] One often encounters low-energy effective theories in the literature, but there are other regions of phase space where an effective theory approach can be used. For example, an effective theory for singlet resonances has been described in [27].

[6] Another bound comes from requiring a perturbative expansion of the effective interactions, which implies

$$\frac{|\alpha|}{\Lambda^2} < \frac{4\pi}{E^2}. \tag{C3}$$

Moreover, requiring perturbative couplings in the UV theory implies a bound on α, which depends on the number of fields n in the effective operator (or, equivalently, on the number of legs of the corresponding UV amplitudes). For dimension-6 operators one has

$$|\alpha| < (4\pi)^{n-2}, \tag{C4}$$

with $n \leq 6$. More details about the validity of the SM effective field theory in relation to possible UV completions can be found in [10].

[7] There can sometimes be selection rules suppressing the interference of certain dimension-6 operators with the SM [11]. In that case dimension-6 and 8 operators have a comparable impact. Just in case, we stress that such a feature has nothing to do with a possible breakdown of the E/Λ expansion. The effects of operators with $d > 8$ are still expected to be suppressed by powers of E/Λ with respect to the leading effects. Possible selection rules on the $d > 8$ operators would increase this suppression even more.

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

[9] A necessary condition for the shape analysis of a differential distribution to provide information on new physics is that $\mathcal{M}^{\mathrm{BSM}}$ not be proportional to $\mathcal{M}^{\mathrm{SM}}$. In practice, this condition is realized most of the time.

[10] R. Contino, A. Falkowski, F. Goertz, C. Grojean, and F. Riva, J. High Energy Phys. 07 (2016) 144.

[11] A. Azatov, R. Contino, C. S. Machado, and F. Riva, Phys. Rev. D **95**, 065014 (2017).

[12] This is the significance from the p-value of a likelihood-ratio test, assuming a counting experiment with a large data sample (see, e.g., Ref. [28]).

[13] We notice that, as the EFT validity bounds do not matter for the sake of determining the components of $\sigma_r(\alpha)$, only the α/Λ^2 combination actually appears in the problem. A value for Λ could thus be fixed without loss of generality. In the text we do not make such an assumption. For Fig. 2, it will be convenient to choose Λ so that

$$\sigma_r^{\mathrm{SM}} \sim \frac{1}{\Lambda^4} \sigma_r^{\mathrm{BSM}}. \tag{C5}$$

[14] Because of the symmetry in (α_1, α_2), the same is also true for fixed α_2 and $\alpha_1 \to \infty$.

[15] K. Hagiwara, R. Peccei, D. Zeppenfeld, and K. Hikasa, Nucl. Phys. **B282**, 253 (1987).

[16] We follow the conventions of [29].

[17] V. Khachatryan *et al.* (CMS Collaboration), Eur. Phys. J. C **76**, 401 (2016).

[18] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, J. High Energy Phys. 06 (2011) 128.

[19] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, Comput. Phys. Commun. **185**, 2250 (2014).

[20] T. Sjostrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05 (2006) 026.

[21] A. Falkowski, M. Gonzalez-Alonso, A. Greljo, D. Marzocca, and M. Son, J. High Energy Phys. 02 (2017) 115.

[22] S. Fichet, Nucl. Phys. **B911**, 623 (2016).

[23] L. Brenner, N. Belyaev, V. Bortolotto, C. D. Burgard, M. Duehrssen-Debling, K. M. Ecker, S. Gadatsch, D. S. Gray, A. Kaluza, K. Koeneke, R. Konoplich, S. Kortner, K. Prokofiev, C. Schmitt, and W. Verkerke, Report No. LHCHXSWG-INT-2016-004, 2016, http://cds.cern.ch/record/2143180.

[24] The ATLAS collaboration, Technical Report No. ATL-PHYS-PUB-2015-047, CERN, 2015, http://cds.cern.ch/record/2066980.

[25] T. Corbett, O. J. P. Eboli, D. Goncalves, J. Gonzalez-Fraile, T. Plehn, and M. Rauch, J. High Energy Phys. 08 (2015) 156.

[26] A. Butter, O. J. P. Eboli, J. Gonzalez-Fraile, M. C. Gonzalez-Garcia, T. Plehn, and M. Rauch, J. High Energy Phys. 07 (2016) 152.

[27] S. Fichet and G. von Gersdorff, J. High Energy Phys. 12 (2015) 089.

[28] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. C **71**, 1554 (2011).

[29] S. Fichet and G. von Gersdorff, J. High Energy Phys. 03 (2014) 102.