Decoupling theoretical uncertainties from measurements of the Higgs boson

Kyle Cranmer,[1] Sven Kreiss,[1] David López-Val,[2] and Tilman Plehn[3]
[1]*Center for Cosmology and Particle Physics, New York University, New York, New York 10003, USA*
[2]*Centre for Cosmology, Particle Physics and Phenomenology, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium*
[3]*Institut für Theoretische Physik, Universität Heidelberg, 69120 Heidelberg, Germany*
(Received 22 January 2015; published 25 March 2015)

We develop a technique to present Higgs coupling measurements, which decouple the poorly defined theoretical uncertainties associated to inclusive and exclusive cross section predictions. The technique simplifies the combination of multiple measurements and can be used in a more general setting. We illustrate the approach with toy LHC Higgs coupling measurements and a collection of new physics models.

DOI: 10.1103/PhysRevD.91.054032 PACS numbers: 14.80.Bn, 12.60.-i

I. INTRODUCTION

The discovery of the Higgs boson [1–4] has initiated a vigorous program of precision Higgs measurements at the LHC. The primary focus is on the couplings of this new state in terms of an effective Lagrangian based on the Standard Model (SM) gauge structure. The key question is if this Lagrangian is fully renormalizable, as predicted in the Standard Model. Deviations in the properties of the newly found particle from the Standard Model predictions would indicate the first hint for physics beyond the Standard Model.

These precision studies provide one of the key motivations for new accelerators [5,6]. Complicating the interpretation of both the current results and future prospects of the LHC is the presence of uncertainties on the theoretical link between Lagrangian parameters and LHC observables. In this paper, we outline a strategy to decouple the theoretical uncertainties from the experimental results while retaining the ability to incorporate those uncertainties in a subsequent stage we refer to as recoupling.

In reporting results, experimentalists strive to find a representation of the results that is as free of theoretical assumptions as possible while still being convenient for addressing specific theories of interest. For example, the measurement of a cross section in a well-defined fiducial region allows experiments to report results that are not tied to a specific theory at the cost of requiring the reader to estimate a model-dependent acceptance. This approach has advantages but is difficult to generalize. For example, extracting the Higgs couplings requires simultaneous inference of multiple production and decay modes from the combination of multiple searches with correlated experimental systematic uncertainties [7–10]. A convenient approach to combining multiple measurements is the best linear unbiased estimator (BLUE) technique [11]; however, that approach is limited to measurements that are well within the Gaussian regime. The methodology presented here addresses both of these challenges.

The first wave of Higgs coupling measurements [2,12,13] involves signal strength modifiers μ_{pd} that scale the total rate of events for a given combination of production (p) and decay (d) modes,

$$\mu_{pd} = \frac{\sigma_p \times \mathrm{BR}_d}{(\sigma_p \times \mathrm{BR}_d)^{\mathrm{SM}}} = \left(\frac{\sigma_p}{\sigma_p^{\mathrm{SM}}}\right)\left(\frac{\Gamma_d}{\Gamma_d^{\mathrm{SM}}}\right)\left(\frac{\Gamma_H^{\mathrm{SM}}}{\Gamma_H}\right). \quad (1)$$

The ATLAS experiment has made such likelihood functions available [14]. If we focus on measuring the Higgs couplings in the Standard Model Lagrangian, we can link the event rates to the set of shifted Higgs couplings g_x/g_x^{SM} to any Standard Model particle x. Kinematic distributions (and, thus, selection efficiencies and detector acceptances) will not change as long as the Higgs couplings are roughly in their Standard Model ranges. While σ_p and Γ_d can usually be directly linked to a specific coupling g_x, the appearance of the total width forces us to make nontrivial assumptions and induces strong correlations. The LHC Cross Section Working Group has defined a number of benchmark scenarios that specify how the production cross section and branching ratios are modified with respect to the Standard Model values [15], and the ATLAS and CMS collaborations are reporting their results in terms of these benchmark scenarios.

Thus far, results for signal strengths and Higgs couplings include the theoretical uncertainty as part of the Standard Model prediction. Large uncertainties on the production appear because of unknown higher orders in the perturbative QCD expansion. The problem with these theoretical uncertainties is twofold: First, the size of the associated uncertainty in not known. Second, the uncertainties are not statistical in nature; there is no random variable associated to missing higher orders. While there has been effort to assign a Bayesian degree of belief to these uncertainties [16] and an effort to complete the perturbative series [17],

we lack an objective probabilistic interpretation of these uncertainties. While the field may be able to settle on a consensus for both the size and shape of these uncertainties, it is an area of debate and will certainly change with time as theoretical progress is made. Therefore, our view is that we should develop a technique in which the theoretical uncertainties are factorized from the experimental result.

In Sec. II we describe the approach allowing us to decouple and recouple the theoretical uncertainties from the experimental measurement. This approach is independent of our specific application to the Higgs coupling measurements at the LHC. In Sec. III we briefly discuss the issues with theoretical uncertainties at hadron colliders and how they affect the Higgs couplings measurement. Section IV demonstrates the procedure with a toy example based on the ATLAS results presented in Ref. [13]. Next, we consider some specific new physics scenarios that can be tested through Higgs coupling measurements in Sec. V. A simple example is worked in detail in Appendix A and more details on the new physics models in Appendix B.

II. THE APPROACH

A. The statistical model

In this section we outline briefly the statistical modeling approach used by the LHC experiments following Ref. [18]. Once the statistical model has been constructed, the LHC experiments employ the profile likelihood ratio to define confidence intervals [19] on the parameters of interest.

The coupling measurements require defining several disjoint categories of events indexed by c, which satisfy specific selection criteria designed, in part, to be particularly sensitive to a particular production or decay mode. Each category has associated to it an expected (observed) number of events ν_c (n_c). Each category may also have some discriminating variable(s) x, such as an invariant mass and a corresponding probability density function $f(x)$. The data associated to the cth category is denoted $\mathcal{D}_c = \{x_1, \ldots, x_{n_c}\}$. In general, the expected number of events and their distribution will depend on both the signal strength parameters $\boldsymbol{\mu}$ and nuisance parameters $\boldsymbol{\alpha}$. The nuisance parameters $\boldsymbol{\alpha}$ parametrize both theoretical and experimental uncertainties. Typically, the expected number of events is written as

$$\nu_c(\boldsymbol{\mu}, \boldsymbol{\alpha}) = \sum_{p,d} \mu_{pd} s_{cpd}(\boldsymbol{\alpha}) + b_c(\boldsymbol{\alpha}), \tag{2}$$

where $b_c(\boldsymbol{\alpha})$ is the background in this category, and

$$s_{cpd}(\boldsymbol{\alpha}) = L(\boldsymbol{\alpha}) \sigma_p^{\mathrm{SM}}(\boldsymbol{\alpha}) \mathrm{BR}_d^{\mathrm{SM}}(\boldsymbol{\alpha}) \epsilon_{cpd}(\boldsymbol{\alpha}) \tag{3}$$

is the expected Standard Model signal for production mode p and decay mode d predicted by the product of the integrated luminosity, cross section, branching ratio, and selection efficiency, each of which may depend on theoretical uncertainties parametrized by α.

Constraint terms associated with systematic uncertainties are described as $f_i(a_i|\alpha_i)$, where α_i are nuisance parameters and a_i are auxiliary or control measurements designed to estimate those nuisance parameters. In the case of experimental uncertainties, there are often real auxiliary measurements that are summarized by $f_i(a_i|\alpha_i)$. However, in the case of most theoretical uncertainties, the auxiliary measurement does not truly exist and an *ad hoc* $f_i(a_i|\alpha_i)$ is introduced for convenience. The full likelihood function used by the experiments [20] can be written as a product of the main experimental measurement and the constraint terms

$$L_{\mathrm{full}}(\boldsymbol{\mu}, \boldsymbol{\alpha}) = \underbrace{\prod_{c \in \mathrm{category}} \left[\mathrm{Pois}(n_c|\nu_c(\boldsymbol{\mu}, \boldsymbol{\alpha})) \prod_{e=1}^{n_c} f_c(x_e|\boldsymbol{\mu}, \boldsymbol{\alpha}) \right]}_{\equiv L_{\mathrm{main}}(\boldsymbol{\mu}, \boldsymbol{\alpha})} \underbrace{\prod_{i \in \mathrm{syst}} f_i(a_i|\alpha_i)}_{\equiv L_{\mathrm{constr}}(\boldsymbol{\alpha})}. \tag{4}$$

Typically, confidence intervals are then defined by contours of the profile likelihood ratio

$$\lambda(\boldsymbol{\mu}) = \frac{L(\boldsymbol{\mu}, \hat{\hat{\boldsymbol{\alpha}}}(\boldsymbol{\mu}))}{L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\alpha}})}, \tag{5}$$

where $\hat{\hat{\boldsymbol{\alpha}}}(\boldsymbol{\mu})$ is the conditional maximum likelihood estimate, and $\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\alpha}}$ are the unconditional maximum likelihood estimates [19].

B. The effective signal strength

We are interested in inferring the values of the signal strength parameters $\boldsymbol{\mu}$, which scale the signal expectation s_{cpd}; however, the presence of experimental and theoretical uncertainties means that the signal and background expectations are functions of the nuisance parameters as in Eq. (2). Alternatively, we can introduce an effective scale factor with respect to some fixed reference scenario α_0, so the expected number of events can be rewritten,

$$\begin{aligned} \nu_c(\boldsymbol{\mu}, \boldsymbol{\alpha}) &= \sum_{p,d} \mu_{pd} s_{cpd}(\boldsymbol{\alpha}) + b_c(\boldsymbol{\alpha}) \\ &\to \sum_{p,d} \mu_{cpd}^{\mathrm{eff}}(\boldsymbol{\mu}, \boldsymbol{\alpha}) s_{cpd}(\boldsymbol{\alpha}_0) + b_c(\boldsymbol{\alpha}_0). \end{aligned} \tag{6}$$

The key conceptual jump is to realize that we can think of μ_{cpd}^{eff} not as a function but as a well-defined parameter free of theoretical uncertainty that we can infer directly.

While the signal strength parameters μ_{pd} we ultimately want to infer are independent of the details of the individual analysis categories, the effective signal strength $\mu^{\rm eff}_{cpd}$ is specific to the cth category due to the selection efficiency [and, more generally, the distributions $f_c(x|\boldsymbol{\mu},\boldsymbol{\alpha})$]. In particular, selection requirements that leverage exclusive or differential properties of a specific production and decay will introduce category-specific theoretical uncertainties.

State-of-the-art Higgs property measurements can include hundreds of categories of events. The full likelihood $L_{\rm full}$ defined in Eq. (4) encodes a detailed description of the correlated effect of common experimental systematic uncertainties. In practice, we want to find some coarse graining of the many categories into a few groups so that we can ignore the category index within each group. By taking a common category-weighted signal strength factor, we can suppress the category index and make the replacement $\mu^{\rm eff}_{cpd} \to \mu^{\rm eff}_{pd}$. By doing so, we can define

$$L_{\rm eff}(\boldsymbol{\mu}^{\rm eff}) \equiv L_{\rm main}(\boldsymbol{\mu}=\boldsymbol{\mu}^{\rm eff},\boldsymbol{\alpha}=\boldsymbol{\alpha}_0). \tag{7}$$

The goal is to show that by providing $L_{\rm eff}(\boldsymbol{\mu}^{\rm eff})$, the reparametrization $\boldsymbol{\mu}^{\rm eff}(\boldsymbol{\mu},\boldsymbol{\alpha})$, and the constraint terms $f_i(a_i|\alpha_i)$, that we can recouple these ingredients and approximate the full likelihood

$$L_{\rm full}(\boldsymbol{\mu},\boldsymbol{\alpha}) \approx L_{\rm recouple}(\boldsymbol{\mu},\boldsymbol{\alpha}) \equiv L_{\rm eff}(\boldsymbol{\mu}^{\rm eff}(\boldsymbol{\mu},\boldsymbol{\alpha}))\cdot L_{\rm constr}(\boldsymbol{\alpha}). \tag{8}$$

In the case of an inclusive cross section uncertainty, $\mu^{\rm eff}_{cpd}$ is the same for all c—in which case, we say the effect of that uncertainty is category universal. If all uncertainties are category universal, then it is possible for this approach to be exact. More generally, the grouping of categories will lead to $\boldsymbol{\mu}^{\rm eff}(\boldsymbol{\mu},\boldsymbol{\alpha})$ encoding some weighted effect from the individual categories. We will discuss some examples in the next section.

C. Reparametrization templates

The art of this approach lies in choosing a template for the reparametrization in which the coefficients of the template can be effectively deduced from the likelihood. We treat the likelihood as a "black box" since the diversity and complexity of statistical models created by experimentalists and encompassed by Eq. (4) is so diverse. In the case that the reparametrization $\boldsymbol{\mu}^{\rm eff}(\boldsymbol{\mu},\boldsymbol{\alpha})$ is category universal, which can be trivially achieved if the ingredients are explicitly provided for each category, this reformulation of the likelihood can be exact.

For example, a natural way to parametrize the dependence of the expected signal due to uncertainties that modify inclusive production cross sections is

$$s_{cpd}(\boldsymbol{\alpha}) = s_{cpd}(\boldsymbol{\alpha}_0)\left[1+\sum_i \eta_{pi}(\alpha_i-\alpha_{0,i})\right] \quad (\forall\ c,d), \tag{9}$$

which is equivalent to

$$\mu^{\rm eff}_{pd}(\boldsymbol{\mu},\boldsymbol{\alpha}) = \mu_{pd}\left[1+\sum_i \eta_{pi}(\alpha_i-\alpha_{0,i})\right] \quad (\forall\ c,d). \tag{10}$$

In this situation, $\mu^{\rm eff}_{pd}(\boldsymbol{\mu},\boldsymbol{\alpha})$ is not linear in the full set of parameters but is bilinear in $(\mu_{pd},\boldsymbol{\alpha})$. This μ scaling is important for capturing the behavior of the likelihood away from the maximum likelihood estimate and distinguishes this approach from techniques such as BLUE [11].

In the case of uncertainties that only affect the background through

$$b_c(\boldsymbol{\alpha}) = b_c(\boldsymbol{\alpha}_0)\left[1+\sum_i \phi_{ci}(\alpha_i-\alpha_{0,i})\right] \quad (\forall\ p,d), \tag{11}$$

the equivalent form of the effective signal strength is

$$\mu^{\rm eff}_{pd}(\boldsymbol{\mu},\boldsymbol{\alpha}) = \mu_{pd} + \frac{b_c(\boldsymbol{\alpha}_0)}{s_{cpd}(\boldsymbol{\alpha}_0)}\left[\sum_i \phi_{ci}(\alpha_i-\alpha_{0,i})\right] \quad (\forall\ p,d), \tag{12}$$

which is linear in μ_{pd} and α. Because $b_c(\boldsymbol{\alpha}_0)/s_{cpd}(\boldsymbol{\alpha}_0)$ is a constant, this prefactor can be absorbed into ϕ_{ci}, and the category-weighted effect would simply be written ϕ_i.

Another example is motivated by the large uncertainty associated to gluon-fusion (ggF) Higgs production with two additional jets, which populates the categories meant to isolate weak boson fusion (VBF). We would expect the uncertainty to modify the $\mu^{\rm eff}_{c,p={\rm VBF},d}$ signal strength for weak boson fusion but be proportional to $\mu_{\rm ggF}$. Thus, we should anticipate templates of the form

$$\mu^{\rm eff}_{pd}(\boldsymbol{\mu},\boldsymbol{\alpha}) = \mu_{pd} + \sum_{i,p'} \mu_{p'd}\eta^{p'}_{pi}(\alpha_i-\alpha_{0,i}). \tag{13}$$

Combining these three situations, a fairly general template would be

$$\mu^{\rm eff}_{pd}(\boldsymbol{\mu},\boldsymbol{\alpha}) = \mu_{pd} + \sum_{i,p'} \mu_{p'd}\eta^{p'}_{pi}(\alpha_i-\alpha_{0,i}) + \sum_i \phi_i(\alpha_i-\alpha_{0,i}), \tag{14}$$

where we can identify $\eta_{pi}=\eta^{p}_{pi}$ from Eq. (9) and $b_c(\boldsymbol{\alpha}_0)\phi_{ci}/s_{cpd}(\boldsymbol{\alpha}_0)=\phi_i$. This general template involves $n_\alpha(n_p^2+1)$ coefficients (per grouping of categories).

The LHC experiments must cope with additional complications. First is the fact that uncertainties on α_i can be large, and linear extrapolation of the effect of this uncertainty on a signal or background rate can lead to unphysical expectations like $s_{cpd} < 0$ or $b_c < 0$. To cope with this, the LHC experiments typically implement log-normal priors (constraint terms), which are implemented via reparametrization so that $f_i(a_i|\alpha_i)$ is a Gaussian distribution and $s_{cpd}(\alpha_i)$ is an exponential response function. This approach has two advantages: it ensures $s_{cpd}(\alpha_i) > 0$, and it allows for multiple signal expectations with different sensitivities to a common source of uncertainty to have a similar log-normal behavior parametrized by α_i.

Second, experiments often have a few numbers with which to parametrize the signal (and background) expectations. Typically, this is based on a nominal $s_{cpd}(\alpha_i = 0)$ and "$\pm 1\sigma$" variations on $s_{cpd}(\alpha_i = \pm 1)$ (using a conventional scaling of α_i). Often the variation of $s_{cpd}(\alpha_i = \pm 1)$ with respect to $s_{cpd}(\alpha_i = 0)$ is asymmetric, which requires some assumptions about the intermediate behavior and the use of *ad hoc* interpolation algorithms, such as second degree polynomial or higher degree polynomials that match the exponential extrapolation up to the second derivative [21,22].

Finally, to ensure the positivity of $s_{cpd}(\boldsymbol{\alpha})$ under the joint effect of several sources of uncertainty, it is common that the interpolation/extrapolation template is multiplicative over the nuisance parameters. In generic terms, this often leads to signal expectations parametrized as

$$s_{cpd}(\boldsymbol{\alpha}) = s_{cpd}(\boldsymbol{\alpha}_0)\prod_i I(\alpha_i), \tag{15}$$

where $I(\alpha_i)$ is some interpolation/extrapolation function based on s_{cpd} evaluated at several points in α_i. These *ad hoc* choices influence the resulting inference and further strengthen the motivation to decouple theoretical uncertainties from the presentation of experimental results.

D. Determining the coefficients of the reparametrization template

Following Eq. (8), our goal is to show that by providing $L_{\text{eff}}(\boldsymbol{\mu}^{\text{eff}})$, the reparametrization $\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu},\boldsymbol{\alpha})$, and the constraint terms $f_i(a_i,\alpha_i)$, we can "recouple" the ingredients to approximate the original likelihood $L_{\text{full}}(\boldsymbol{\mu},\boldsymbol{\alpha})$. Below we develop two approaches to determine the coefficients of these reparametrization templates.

1. Via the local covariance matrix

A direct path towards this goal is to use a reparametrization template like the ones introduced in Sec. II C and determine the coefficients that reproduce the local covariance structure around the maximum likelihood estimate $(\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\alpha}})$. The maximum likelihood estimate is defined by $-\nabla \ln L = 0$, while the Hessian of $-\ln L$ captures the local covariance structure and is referred to as the observed Fisher information matrix or precision matrix. In particular, the observed Fisher information is an estimate of the inverse of the covariance matrix

$$V^{-1}_{ij} = -\partial_i \partial_j \ln L. \tag{16}$$

It is helpful to factorize the full likelihood in Eq. (4) as a product of the main measurement and the constraint terms for the nuisance parameters $\boldsymbol{\alpha}$. This is equivalent to decomposing the full information matrix into two parts:

$$V^{-1}_{\text{full}} = V^{-1}_{\text{main}} + V^{-1}_{\text{constr}}. \tag{17}$$

By fixing $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$, the experiments can provide the (profile) likelihood for the effective signal strength $L_{\text{eff}}(\boldsymbol{\mu}^{\text{eff}})$ via Eq. (7). We denote the Fisher information for this effective signal strength as V^{-1}_{eff}. Reparametrizing $L_{\text{eff}}(\boldsymbol{\mu}^{\text{eff}})$ via $\boldsymbol{\mu}^{\text{eff}} \to \boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu},\boldsymbol{\alpha})$ implies the following transformation to the Fisher information matrix

$$V^{-1}_{\text{eff}}(\boldsymbol{\mu}^{\text{eff}}) \to V^{-1}_{\text{eff}}(\boldsymbol{\mu},\boldsymbol{\alpha}) = J^T V^{-1}_{\text{eff}} J, \tag{18}$$

where the Jacobian of the transformation is defined by

$$J = \frac{\partial(\boldsymbol{\mu}^{\text{eff}})}{\partial(\boldsymbol{\mu},\boldsymbol{\alpha})}. \tag{19}$$

Even though V^{-1}_{eff} may be a constant matrix, the Jacobian in Eq. (18) depends on $(\boldsymbol{\mu},\boldsymbol{\alpha})$ so that the resulting information matrix need not be constant. This is an important point, as the analysis below is only sensitive to the linear approximation of $\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu},\boldsymbol{\alpha})$ at $(\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\alpha}})$ although the template may encode important nonlinear behavior. Combining Eqs. (17) and (18) gives us

$$V^{-1}_{\text{full}} = J^T V^{-1}_{\text{eff}} J + V^{-1}_{\text{constr}}. \tag{20}$$

The final stage of this procedure is either to check if Eq. (20) holds for a particular reparametrization $\boldsymbol{\mu}^{\text{eff}} \to \boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu},\boldsymbol{\alpha})$ or to determine the coefficients for a template by imposing this equality.

Perhaps the most intuitive way to think of the effect of a given uncertainty is the shift it induces in the best-fit value of the signal strengths $\boldsymbol{\mu}$. This way of thinking requires keeping the $\boldsymbol{\alpha}$ fixed and considering the likelihood is only a function of $\boldsymbol{\mu}$. For example, Fig. 1(b) shows the shift in the $L(\boldsymbol{\mu})$ contour due to a shift in the gluon-fusion inclusive cross section. We denote the best-fit $\boldsymbol{\mu}$ with fixed $\boldsymbol{\alpha}$ as

$$\hat{\boldsymbol{\mu}}^{\text{fix}}(\boldsymbol{\alpha}) \equiv \text{argmax}_{\boldsymbol{\mu}} L_{\text{full}}(\boldsymbol{\mu},\boldsymbol{\alpha}) \tag{21}$$

with $\hat{\boldsymbol{\mu}}^{\text{fix}}(\hat{\boldsymbol{\alpha}}) = \hat{\boldsymbol{\mu}}$. Because of the definition $L_{\text{full}}(\boldsymbol{\mu},\boldsymbol{\alpha}_0) \equiv L_{\text{eff}}(\boldsymbol{\mu})$ in Eq. (7), this implies that $\hat{\boldsymbol{\mu}}^{\text{fix}}(\boldsymbol{\alpha}_0) = \hat{\boldsymbol{\mu}}^{\text{eff}}$. Finally,

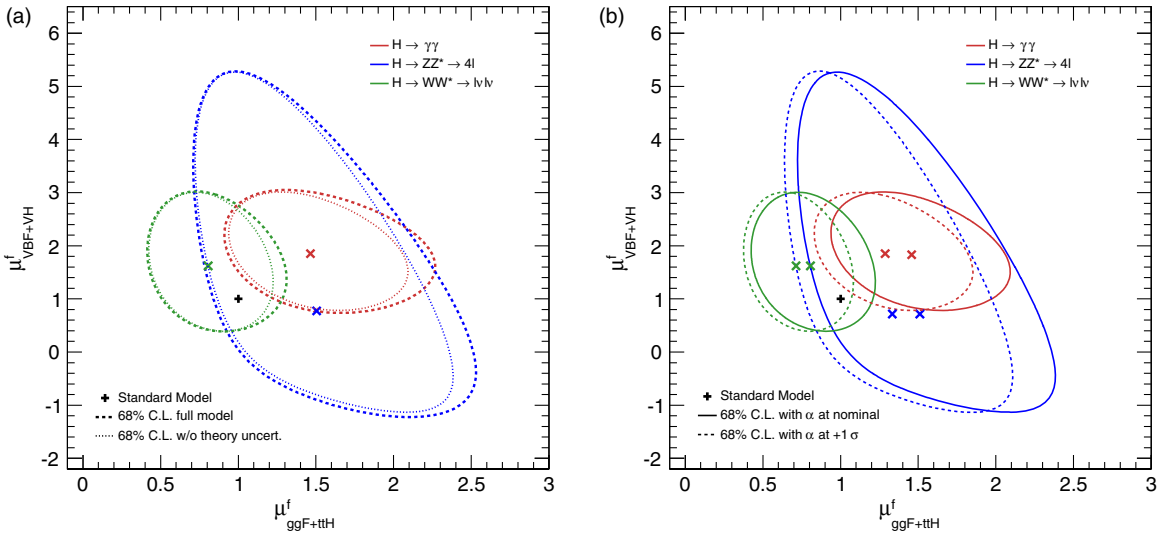

FIG. 1 (color online). Likelihood contours for three different Higgs decays: (a) with (solid) and without (dashed) theoretical uncertainties; (b) without theoretical uncertainties for the nominal gluon-fusion cross section (solid) and a shifted value (dashed) estimated from QCD scale variations.

if the main measurement is not able to measure the nuisance parameters, i.e., there is flat direction in the likelihood or a degeneracy between $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$, then $\boldsymbol{\alpha}_0 = \hat{\boldsymbol{\alpha}}$. By using a specific template—in this case, Eq. (10)—we can equate $\mu_p^{\mathrm{eff}}(\hat{\mu}_p^{\mathrm{fix}}, \alpha_i) = \hat{\mu}_p^{\mathrm{fix}}(1 + \eta_{ip}\alpha_i) = \hat{\mu}_p$ and then explicitly evaluate the partial derivative that quantifies the shift to the best-fit value of $\boldsymbol{\mu}$,

$$\left.\frac{\partial \hat{\mu}_p^{\mathrm{fix}}}{\partial \alpha_i}\right|_{\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\alpha}}} = -\hat{\mu}_p \eta_{ip}. \tag{22}$$

These partial derivatives are visualized as vectors in the signal strength plane in Fig. 2.

Another way to arrive at Eq. (22) is to approximate the likelihood in the neighborhood of the maximum likelihood estimate as a multivariate Gaussian $G(\boldsymbol{\mu}, \boldsymbol{\alpha}|\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\alpha}}, \Sigma)$. The conditional distribution of $\boldsymbol{\mu}$ given $\boldsymbol{\alpha}$ is also a multivariate Gaussian with the mean given by

$$\hat{\boldsymbol{\mu}}^{\mathrm{fix}}(\boldsymbol{\alpha}) = \hat{\boldsymbol{\mu}} + \Sigma_c \Sigma_\alpha^{-1}(\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}), \tag{23}$$

where Σ_c is the upper-right sub-block of the full covariance matrix Σ,

$$(V_{\mathrm{full}}^{-1})^{-1} = \Sigma = \begin{bmatrix} \Sigma_\mu & \Sigma_c \\ \Sigma_c^T & \Sigma_\alpha \end{bmatrix}. \tag{24}$$

In situations that the main measurement does not constrain or pull on the nuisance parameters, Σ_α is just the covariance matrix associated to the constraint term defined in Eq. (4). In general, Σ_c will depend both on the constraint terms and the main measurement. Clearly, the conditional likelihood with α fixed as in Eq. (21) is independent of the constraint term; thus, the product $(\Sigma_c \Sigma_\alpha^{-1})$ can only depend on the details of the main measurement. Through the techniques developed for regression in the general linear model [23], one can show that

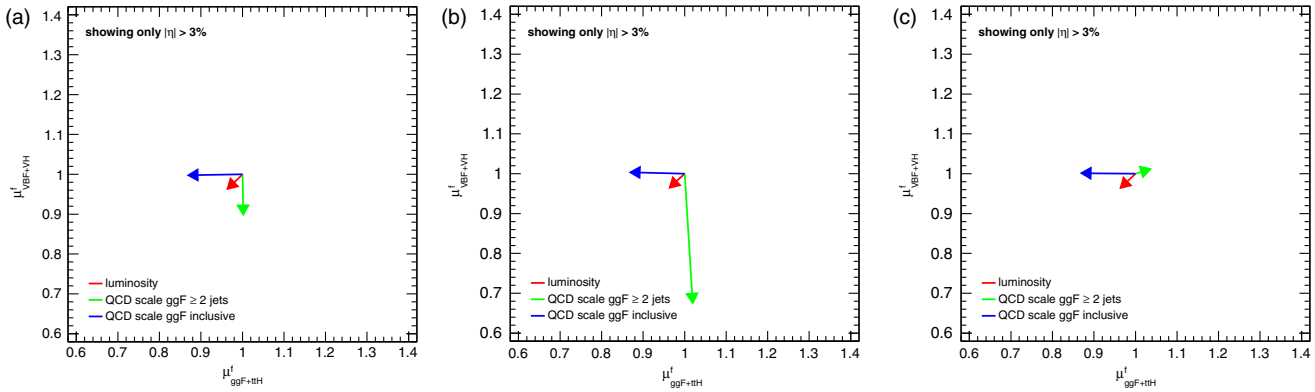

FIG. 2 (color online). Visualization of $\partial\boldsymbol{\mu}^{\mathrm{fix}}/\partial\alpha_i$ for the (a) $H \to \gamma\gamma$, (b) $H \to ZZ \to 4\ell$, and (c) $H \to WW \to \ell\nu\ell\nu$ likelihoods.

$$\left.\frac{\partial \hat{\mu}_p^{\text{fix}}}{\partial \alpha_i}\right|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}} = (\Sigma_c \Sigma_\alpha^{-1})_{ip} = -\hat{\mu}_p \eta_{ip}, \tag{25}$$

where the right-most equivalence is specific to the template of Eq. (10) and explicit computation requires an assumption about the constraint terms even though the result is independent of those assumptions. This is shown explicitly for a simple example in Appendix A. Note, one can eliminate the constraint terms and only consider the main measurement, in which case, Σ in Eq. (25) is a singular matrix and one must use the pseudoinverse. While this approach is mathematically cumbersome, it is equivalent to Eqs. (22) and (26).

A third approach is to impose Eq. (18) with V_{constr}^{-1} subtracted from both sides

$$V_{\text{main}}^{-1} = J^T V_{\text{eff}}^{-1} J. \tag{26}$$

This allows us to proceed without assumptions on the constraint terms, which we are trying to decouple from the procedure. The upper-right sub-bloc of this matrix leads to a system of equations that can be used to determine the coefficients of the template. Again, in the case of template of Eq. (10), these linear equations provide the same solutions for η_{ip} as Eqs. (22) and (25). These equations are solved explicitly for a simple example in Appendix A.

It is worth noting that in the case of an arbitrary template $\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu}, \boldsymbol{\alpha}; \boldsymbol{\eta})$, these equations fix the linear behavior at $\boldsymbol{\alpha}_0$ (i.e., $\partial \boldsymbol{\mu}^{\text{eff}} / \partial \alpha_i|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0}$), which may be nontrivially related to the coefficients $\boldsymbol{\eta}$. For example, this is the case for the most common parametrizations encompassed by Eq. (15) and used by the LHC experiments and to implement asymmetric uncertainties [21,22].

The information matrix from the full likelihood function has $(n^2+n)/2$ independent components, where $n = n_p + n_\alpha$ is the sum of the number of parameters of interest and nuisance parameters. An $n_p \times n_p$ sub-block describes Σ_μ or V_{eff}^{-1} and an $n_\alpha \times n_\alpha$ sub-block describes Σ_α or V_{constr}^{-1}. Thus, there are remaining $n_p \times n_\alpha$ numbers in the local covariance encoded in Σ_c or V_{main}^{-1} that can be used to determine the coefficients of the template. This counting matches precisely for the templates of Eqs. (9) and (11), but leaves ambiguity for the more general template of Eq. (13).

2. The art of choosing the template

As mentioned above, the local information provided by the observed information matrix evaluated at the maximum likelihood estimate is only sensitive to the linear behavior of $\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu}, \boldsymbol{\alpha})$ at $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\alpha}})$. For example, Eqs. (9) and (11) can have the same linear behavior at $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\alpha}})$, though the effect of an uncertainty α_i scales with $\boldsymbol{\mu}$ in the former and not in the latter. This difference is illustrated in Fig. 3. Thus, while both templates will be equivalent locally, if one wants to capture the higher-order corrections encoded in the different templates, some additional information will be required. This information can either be injected by hand or by using information about the likelihood away from $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\alpha}})$, as described below. For instance, one may know that a particular nuisance parameter α_i is associated to the uncertainty in the inclusive cross section for the pth production mode; thus, the coefficients of the general template can be restricted by hand to $\eta_{pi}^{p'} = \delta_{pp'}\eta_{ip}$ and $\phi_{ip} = 0$. In another extreme case, the uncertainty on gluon-fusion production with two jets primarily affects the inference of $\mu_{p=\text{VBF}}^{\text{eff}}$ signal strength for vector boson fusion, but the size of the shift is proportional to μ_{ggF}. Here, the general template would be constrained so that $\eta_{pi}^{p'} = 0$ unless $p' = \text{ggF}$ and $p = \text{VBF}$ or vice versa. We explore a simple model for this situation in scenario C of Appendix A.

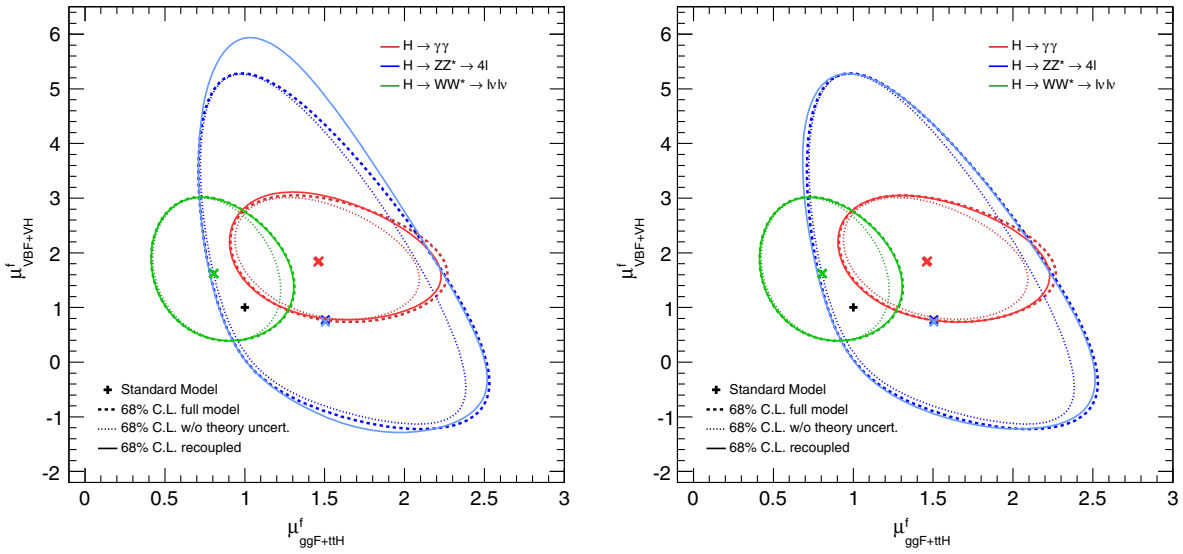

FIG. 3 (color online). Comparison of the effective model using different templates with the same local covariance.

3. Via a global learning approach

Ideally, we would have a formalism that would work with the black box likelihood $L_{\text{full}}(\boldsymbol{\mu},\boldsymbol{\alpha})$ and a general template like Eq. (13) without having to introduce by hand restrictions on that template as described above. In order to do that, we must introduce information about the likelihood away from $(\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\alpha}})$. A flexible approach to that problem is based on the ideas of machine learning and function approximation in which one aims to minimize a loss function with respect to some model parameters (in this case, the template coefficients). The loss function needs to be a scalar evaluated over the $(\boldsymbol{\mu},\boldsymbol{\alpha})$ space, which in frequentist terms has no measure. However, from the point of view of decision theory, one can introduce some weighting over the parameter space (without regarding it as a Bayesian prior) and evaluate

$$\text{Loss}(\boldsymbol{\eta}) = \int d\boldsymbol{\mu} d\boldsymbol{\alpha} \pi(\boldsymbol{\mu},\boldsymbol{\alpha}) |L_{\text{full}}(\boldsymbol{\mu},\boldsymbol{\alpha}) - L_{\text{recouple}}(\boldsymbol{\mu},\boldsymbol{\alpha};\boldsymbol{\eta})|^2. \tag{27}$$

The choice of the weighting function $\pi(\boldsymbol{\mu},\boldsymbol{\alpha})$ is arbitrary, but a reasonable choice is the Bayesian posterior with respect to some baseline constraint terms interpreted as *a prior* on $\boldsymbol{\alpha}$, which leads to

$$\pi(\boldsymbol{\mu},\boldsymbol{\alpha}) \propto L_{\text{main}}(\boldsymbol{\mu},\boldsymbol{\alpha}) L_{\text{constr}}(\boldsymbol{\alpha}). \tag{28}$$

This approach will put the highest weight for L_{recouple} to approximate L_{full} near the best-fit point $(\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\alpha}})$ and lesser weight as one moves away from it. Importantly, this minimum loss solution can be found numerically and is well defined even when the number of parameters in the template is larger than $n_p \times n_\alpha$. Furthermore, this approach may be more robust in the case of very complicated likelihood functions where the numerical accuracy of the covariance matrix, information matrix, and partial derivatives needed in Eqs. (22), (25), and (26) may be poor.

In situations where additional experimental uncertainties α_{exp} have been profiled in providing $L_{\text{eff}}(\boldsymbol{\mu}^{\text{eff}})$—a situation discussed in more detail in Sec. II E—one must take care that the loss function makes the comparison for equivalent values of the profiled nuisance parameters. For example, when creating $L_{\text{eff}}(\boldsymbol{\mu}^{\text{eff}})$, one can keep track of the profiled values $\hat{\hat{\boldsymbol{\alpha}}}_{\text{exp}}(\boldsymbol{\mu}^{\text{eff}})$ and then in Eq. (27) make the replacement

$$L_{\text{full}}(\boldsymbol{\mu},\boldsymbol{\alpha}) \to L_{\text{full}}(\boldsymbol{\mu},\boldsymbol{\alpha},\hat{\hat{\boldsymbol{\alpha}}}_{\text{exp}}(\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu},\boldsymbol{\alpha}))). \tag{29}$$

We demonstrate the effectiveness of this approach in Sec. IV and scenarios B and C of Appendix A.

4. Software

The software implementation of the reparametrization templates described in Sec. II C as well as the three strategies for determining the coefficients of those templates from the local covariance matrix and the learning approach described in Sec. II D are available in Ref. [24]. Experiments can use this software on their full ROOFIT/ROOSTATS [25] models and obtain the effective likelihood $L_{\text{eff}}(\boldsymbol{\mu}^{\text{eff}})$ as well as the reparametrization $\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu},\boldsymbol{\alpha})$ for publication. These ingredients can be supplied in a technology-independent format enabling others to perform the recoupling stage, modify constraint terms associated to theoretical uncertainties, combine multiple results, and create likelihood scans in benchmark models.

E. Grouping of categories and combinations

In Sec. II B we discuss the coarse graining of the many categories into a few groups. By taking a common category-weighted signal strength factor, we are able to suppress the category index and make the replacement $\mu^{\text{eff}}_{cpd} \to \mu^{\text{eff}}_{pd}$ for each group. In practice, this grouping is often based on the Higgs decay mode indexed by d. For example, ATLAS provided the profile likelihood $\lambda(\mu_{\text{ggF}},\mu_{\text{VBF}})$ for the three decay modes $d = \gamma\gamma, WW, ZZ$ [14]. These profile likelihoods are colloquially referred to as "likelihoods"; however, they have eliminated several nuisance parameters via profiling as defined in Eq. (5). In addition to the common theoretical uncertainties, these three likelihoods share common experimental systematic uncertainties. Naively combining these likelihoods will, thus, double count the common constraint terms $f_i(a_i|\alpha_i)$ and lead to an artificial reduction in the uncertainty. This effect can be seen by comparing a naive combination of the three individual profile likelihoods provided by ATLAS with the official ATLAS combined result (which avoids the erroneous double counting). The same effect is demonstrated with a toy example in Fig. 4.

The decouple/recouple procedure described here has focused on theoretical uncertainties for reasons that will be elaborated in Sec. III; however, the same technique can be used for experimental uncertainties. In particular, sources of systematic uncertainties that are anticipated to be common to other analyses should not be profiled, or it will not be possible to avoid double counting of constraint terms. The luminosity uncertainty is an example of an experimental uncertainty that should not be profiled, but instead the effect of the uncertainty should be parametrized in the template $\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu},\boldsymbol{\alpha})$. As a rule of thumb, one can safely profile uncertainties that are specific to the categories in a given analysis but should include common (correlated) sources of uncertainty in the reparametrization.

F. Anticipating higher-dimension operators and other effects

At this point, the Higgs coupling analysis is primarily focused on small deviations in the coefficients of the SM operators. This justifies the scaling of the SM Higgs expectations as $\mu_{pd} s_{cpd}$, since the efficiency and acceptance

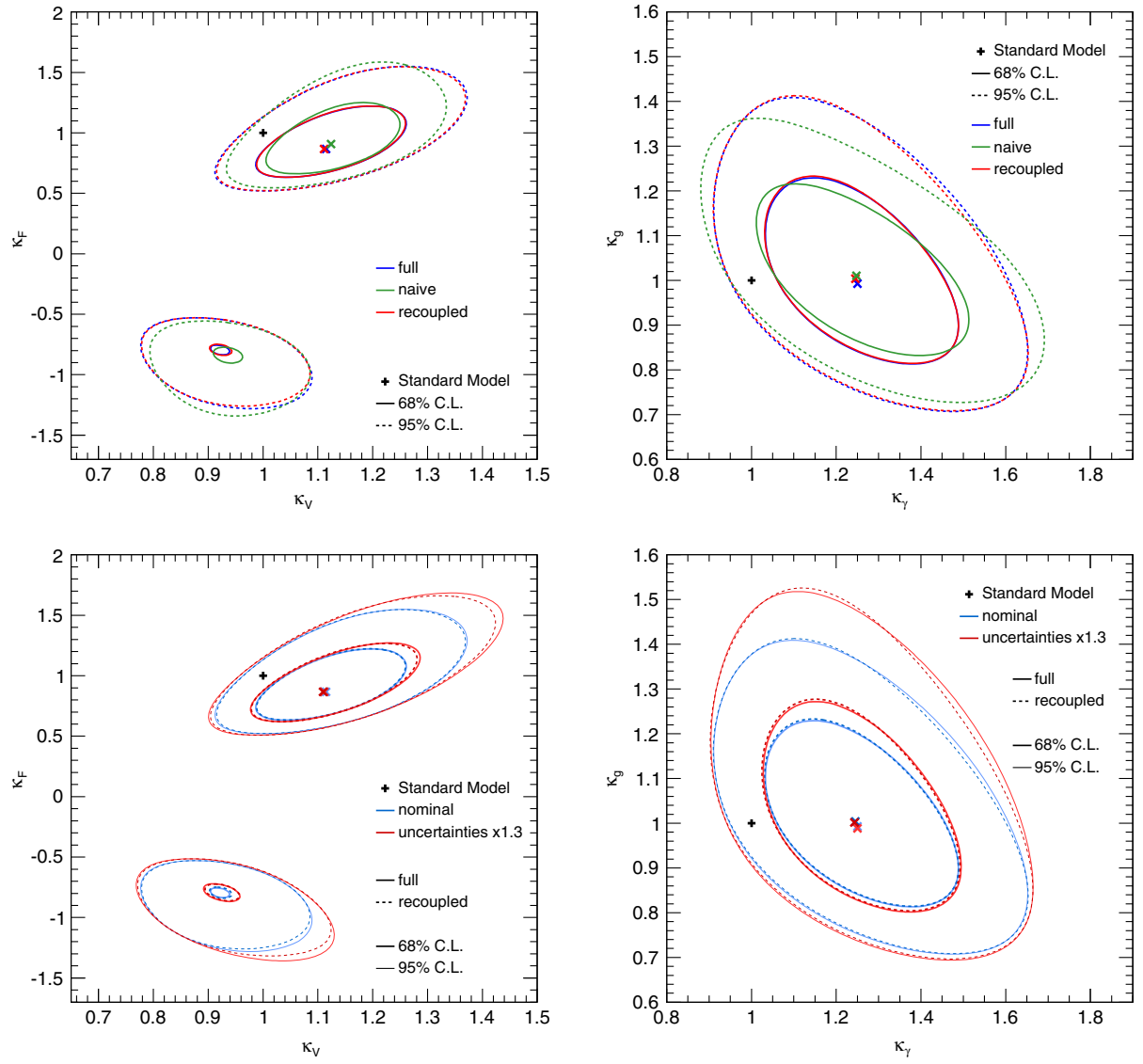

FIG. 4 (color online). Various comparisons of the combined $\gamma\gamma, ZZ, WW$ likelihoods in the (κ_V, κ_F) (left) and $(\kappa_\gamma, \kappa_g)$ (right) planes. Top: comparison of the full combined likelihood, a naive combination with inconsistent profiling and double-counted constraint terms, and the combination of recoupled likelihoods with consistent profiling and without double-counted constraint terms. Bottom: comparison of the full combined likelihood and the combination of recoupled likelihoods using the nominal uncertainties and modified constraint terms with uncertainties inflated by 30%.

of the modified signal in the cth category is not affected. The presence of non-SM operators will generically change the pattern of the signal across the categories in a way specific to the operators under consideration [26]. In particular, operators that modify kinematic distributions will affect the cut efficiencies and acceptances $\epsilon_{cpd}(\boldsymbol{\alpha})$ of Eq. (3) One approach to anticipate these studies is to characterize the effect of a deviation in the number of events in each individual category. In particular, by considering a perturbation of the form

$$\nu_c(\boldsymbol{\mu}, \boldsymbol{\alpha}) = \sum_{p,d} \mu_{pd} s_{cpd}(\boldsymbol{\alpha}) + b_c(\boldsymbol{\alpha}) + \alpha_c, \qquad (30)$$

where α_c is some addition or reduction to the number of events in that specific category. This would add one nuisance parameter for each category, each of which would have the effect of shifting $\hat{\boldsymbol{\mu}}^{\text{eff}}$. These α_c can be seen as a basis for possible new physics effects and can be incorporated into $\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu}, \boldsymbol{\alpha})$ using the same formalism. To use this information in the context of a specific new physics model, one would need to estimate the perturbation to the number of events in each category (α_c) through knowledge of the details of the selection for that category. If that is possible, then one can parametrize each α_c in terms of the coefficients of these new operators. This may involve parametrizing several α_c in terms of just a few operator coefficients.

III. THEORETICAL UNCERTAINTIES

In this section we focus on the dominant theoretical uncertainties plaguing Higgs coupling measurements at the LHC. There are several sources of uncertainties, which can be associated with the theoretical description of LHC processes [4]:

(1) the inclusive rate for a given production and decay; for example, the total cross section $\sigma_{\text{tot}}(gg \to H + X)$,
(2) efficiency and acceptance of kinematic cuts for the signal process; for example, cuts on the transverse momentum of the Higgs boson or a central jet veto,
(3) the parton densities describing the initial state, which have both experimental and theoretical uncertainties,
(4) the parton shower, hadronization, underlying event, and other key parts of the Monte Carlo description of signal and background events.

Linking event rates in a given kinematic region to perturbative theoretical predictions might be considered the critical problem for the LHC's precision Higgs program. According to the above list, it can be separated into two parts: First, the total Higgs production cross section parametrically depends on the Higgs coupling in a given renormalization scheme. This relation is based, for example, on collinear factorization, includes a small number of scales, and can be computed in perturbative QCD. However, this connection cannot be replaced or checked experimentally, which means that Higgs coupling measurements will eventually be limited by theory predictions [5,6].

Kinematic cuts linked to Higgs-plus-jet production [27], tagging jets in weak boson fusion [28,29], or advanced analysis methods lead to additional complications. They do not automatically obey collinear factorization and induce a large number of energy scales and possibly large logarithms [30–34]. Already for the relatively safe inclusive Higgs production cross section, the next-to-next-to-leading-order corrections more than double the predicted number of produced Higgs bosons [35–38]. The source of these problems is the choice of protons as colliding particles, combined with the large value of the strong coupling constant, which leads to a poor convergence of the perturbative QCD description.

Aside from the poor convergence in perturbative QCD, the corresponding theoretical uncertainties face a more fundamental problem: there does not exist any well-defined estimate for an uncertainty on a production cross section computed in perturbative QCD. Traditionally, we derive a range of allowed cross section values using a variation of the factorization and renormalization scales. This is based on the fact that these scales are artifacts of the perturbative expansion, so the scale dependence has to vanish once we include all orders in perturbation theory. This recipe captures some of the effects of the theoretical uncertainties, but we know from the Drell-Yan process or Higgs production in gluon fusion that it does not give a conservative estimate. For Higgs production at the LHC, we know that equal variation of the factorization and renormalization scales leads to a cancellation of the two scale dependencies as one possible reason for the underestimate of the theoretical uncertainties [36]. On the other hand, a separate variation of the factorization and renormalization scales ruins the theoretical description of the scale-dependent parton densities in terms of resummed large logarithms of transverse jet momenta [4]. Independently varying the two scales makes sense phenomenologically but not in terms of a formal QCD description. A recent promising approach might be to combine the scale variation with an extrapolation of the perturbative series [17]. However, no matter what recipe we choose, it is clear that the size of any theoretical uncertainty is poorly defined.

In addition to the overall size of the uncertainty, most statistical analysis techniques require the uncertainty to be quantified in the likelihood, i.e., the constraint terms of Eq. (4). Frequentist interpretation of the likelihood requires one to be able to identify some random observable whose probability is defined with a corresponding ensemble. No traditional ensemble defined by repeated observation exists for theoretical uncertainties. Missing higher-order terms in a perturbative prediction are not random in nature; there is no ensemble, thus, $L_{\text{constr}}(\boldsymbol{\alpha})$ is ill-defined. Attempts to provide a meaningful degree of belief for the perturbative uncertainties [16] require a Bayesian interpretation and attribute a fundamental meaning to the perturbative QCD expansion.

In a practical sense, we must make some choice for the likelihood L_{constr} even if it is fundamentally ill-defined. Unfortunately, the various choices have their own unique pathologies. The often-used Gaussian and log-normal distributions inject information by ascribing a preferred (peak) value and allowing for long tails into what are untenable values for the theoretical prediction. Alternatively, we can try to avoid introducing a peak by adopting a constant likelihood function, as proposed in the RFit scheme as described in Refs. [10,39]. The choice of distributions might not appear numerically relevant, but in combination with a profile likelihood, it can lead to significant differences. For instance, the combined effect of uncertainties described in the RFit scheme add linearly rather than in quadrature [4].

Given that the size as well as the shape of the theoretical uncertainties entering the Higgs couplings measurements are subject to variations in time, geography, and personal taste, the best solution is to decouple them from the experimental result as proposed in Sec. II. This allows for efficient tests of different assumptions on the theoretical uncertainties as well as an efficient implementation of a perceived or actual improvement.

IV. A TOY EXAMPLE

We now consider a toy example that is representative of the current ATLAS results for $H \to \gamma\gamma, WW, ZZ$ [13]. The statistical model here is based purely on a number of events

in various categories without including discriminating variable distributions for $m_{\gamma\gamma}$, $m_{4\ell}$, or m_T [the terms $f_c(x|\boldsymbol{\mu},\boldsymbol{\alpha})$ in Eq. (4)]. Each decay mode groups together several categories of events that together provide sensitivity to the underlying production modes. We model the uncertainty on the signal expectation from luminosity, parton distribution functions, the inclusive gluon-fusion cross section, and the uncertainty on the cross section for gluon fusion in association with two or more jets.

The $H \to \gamma\gamma$ likelihood includes a simplified version of the 14 categories considered by ATLAS including the low- and high-p_T categories, the low- and high-mass two-jet categories, the high-$\not{E}_T$ significance category, and the lepton-tagged category. The $H \to ZZ \to 4\ell$ likelihood includes three ggF-like categories (for 4μ, $2e2\mu$, and $4e$) as well as a VBF-like and a VH-like category. The $H \to WW \to \ell\nu\ell\nu$ likelihood includes zero-, one-, and two-jet categories. The HISTFACTORY script and ROOFIT/ROOSTATS workspace for this toy model can be found in Ref. [40].

Figure 1(a) shows the likelihood contours for the three different decays with and without theoretical uncertainties, which are modeled using Gaussian constraint terms and a linear response as in Eq. (9). Figure 1(b) shows the shift to the contours without theory uncertainty due to fixing the inclusive gluon-fusion cross section to its "$+1\sigma$" value estimated from QCD scale variation. The larger gluon-fusion cross section leads to a smaller inferred value for μ_{ggF}. This can be repeated for each of the nuisance parameters α_i as in Eq. (22). The corresponding partial derivatives are visualized in Fig. 2.

Next we demonstrate the recoupling stage based on the local covariance structure at $(\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\alpha}})$. If we choose the reparametrization template in Eq. (10), we can solve for the η_{pi} coefficients that reproduce the local covariance structure using either Eqs. (22), (25), or (26). Having determined the η_{pi} coefficients, the reparametrization $\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu},\boldsymbol{\alpha})$ is specified, and we can create the decoupled likelihood via Eq. (8). Figure 3(a) compares the full likelihood to the recoupled likelihood using this template. The result is not bad. The effect of the theoretical uncertainties have been recovered; however, there is a significant discrepancy for $H \to ZZ \to 4\ell$ at the top of the contour.

The source of this discrepancy between L_{full} and L_{recouple} contours is the non-category-universal nature of the theoretical uncertainty associated to gluon fusion $+ \geq 2$ jets, which primarily affects the VBF-like categories. Correspondingly, the effect of this uncertainty is a shift in the μ_{VBF} direction, as illustrated in Fig. 2. The reparametrization template in Eq. (10) implies that this uncertainty would scale with μ_{VBF}, though in the full model this source of uncertainty scales with gluon fusion. At the top of the contour, $\mu_{\text{VBF}} \approx 5$, while $\mu_{\text{ggF}} \approx 1$, which leads to a large inflation on the effect of this uncertainty.

As discussed in Sec. II D, the local covariance structure of L_{recouple} is only sensitive to the linear behavior of the template, so this physical insight must either be put in "by hand" or one should instead use more global information about the likelihood as in the learning approach. To illustrate the by-hand approach, we can consider the more general template in Eq. (14). The general template can be constrained so that $\eta_{pi}^{p'} = 0$ unless $p' = \text{ggF}$ and $p = \text{VBF}$ (or vice versa), where i is the index for the nuisance parameter associated to the gluon fusion $+ \geq 2$ jet cross section. In addition, we set $\phi_i = 0$ for all i. The same procedures are followed to determine the coefficients that reproduce the local covariance structure for this template. The resulting contour is shown in Fig. 3(b), where the agreement is improved, particularly near the top of the $H \to ZZ \to 4\ell$ contour. The alternative learning approach works with the general template in Eq. (14) without restricting the terms by hand. The optimized values of the coefficients lead to even better agreement of the contours compared to those shown in Fig. 3(b).

Next, we combine these three likelihoods. As discussed in Sec. II E, the experiments typically present results grouped by decay mode. In this example, the three likelihoods share common sources of systematic uncertainty. Naively combining these likelihoods double counts the common constraint terms $f_i(a_i|\alpha_i)$ and leads to an artificial reduction in the uncertainty. This effect can be seen in the top plots of Fig. 4, which compare a naive combination to the full combined result that avoids the erroneous double counting. There is no meaning to a combined contour in the signal strength plane $\mu_{\text{ggF},d} - \mu_{\text{VBF},d}$ due to the decay index d. Thus, two simple two-parameter benchmark models are used to present the result. Figure 4(a) shows a benchmark model that dictates all $\mu_{p,d}$ based on the scaling of the fermionic couplings $g_f = \kappa_f g_f^{\text{SM}}$ and weak boson couplings $g_V = \kappa_V g_V^{\text{SM}}$. Similarly, Fig. 4(b) considers a simple two-parameter benchmark in which we scale the effective Hgg by a factor κ_g and the effective $H\gamma\gamma$ coupling by a factor κ_γ [15]. The contours from the naive combination are considerably smaller in the κ_V and κ_g directions, leading to poor agreement with the full combination (based on all 22 categories of events without double counting constraint terms).

In contrast, the recoupling approach allows one to avoid double counting constraint terms and for a consistent profiling over the common sources of uncertainty (both theoretical and experimental), which leads to an improved agreement. In addition, the learning approach of Eq. (27) has been used for L_{recouple} in Fig. 4. Note, in this example the sources of experimental uncertainty that are unique to one of the decays were profiled in $L_{\text{eff}}(\boldsymbol{\mu}^{\text{eff}})$, as would be done by the experiments. Thus, there is a small effect that is neglected due to the fact that the profiling of these analysis-specific uncertainties is slightly affected by identifying the common $\boldsymbol{\alpha}$ in the combination.

Perhaps the greatest benefit of this approach is that the theoretical uncertainties have been decoupled from the experimental result encoded in $L_{\rm eff}(\boldsymbol{\mu}^{\rm eff})$. Moreover, the recoupling procedure in Eq. (8) allows one to replace the constraint terms in $L_{\rm constr}(\boldsymbol{\alpha})$ with some other constraint terms $L'_{\rm constr}(\boldsymbol{\alpha})$. The bottom plots of Fig. 4 demonstrate the change by using the same $L_{\rm eff}$ and reparametrization template but with uncertainties inflated by 30%. This recouped approach is compared to the full model where the same modification is made to the constraint term. In addition, an example of replacing Gaussian constraint terms with the RFIT scheme is given in Appendix A.

V. NEW PHYSICS EFFECTS VS THEORETICAL UNCERTAINTY

Any coupling measurement based on a Lagrangian description of underlying field theory assumes a set of operators, specifically those describing the Higgs interactions. In the Standard Model, this includes the renormalizable Higgs Lagrangian with the tree-level interactions to massive gauge bosons and fermions. The effective Higgs couplings to photons and gluons are loop induced, but they avoid the decoupling theorem and are only suppressed by $1/v$. This means that at LHC energies, they have to be included in the description of the Higgs signatures. Additional higher-dimensional operators can be included in the analysis once the corresponding measurements separate them from deviations in the renormalizable Higgs operators [26]. The prefactors of each operator, combined with a renormalization scheme, constitute the measurable couplings. When extracting the couplings of a light Higgs boson with mass 126 GeV, we follow an effective theory approach. This allows us to only consider modifications of the light Higgs boson while keeping the ultraviolet properties of a functioning partly decoupled Higgs sector [41]. We analyze extensions of the SM Higgs sector based on Higgs coupling modifications written as [10,41–43]

$$\begin{aligned} g_x &\equiv g_x^{\rm SM}(1+\Delta_x), \\ g_\gamma &\equiv g_\gamma^{\rm SM}(1+\Delta_\gamma^{\rm SM}+\Delta_\gamma). \end{aligned} \tag{31}$$

For loop-induced Higgs couplings, it is important that we first define the deviation due to a shift in the SM-like loops and then separately treat additional particles in the loop. The Higgs signal strength is given in Eq. (1). The three ratios of the production rate, the decay width, and the total width depend on one or more Higgs couplings. Because the total Higgs width cannot be measured at the LHC, we assume that it is given by the sum of all observable partial widths, where we align the relevant second generation modes with the corresponding third generation decay measurements.

In different new physics models, we describe the leading deviations from the Standard Model in terms of one free parameter ξ. For more details and a complete set of references, we refer to Appendix B and to Ref. [41]. Because new physics effects which violate custodial symmetry are unlikely to be discovered in the Higgs sector, we simplify our example analysis by assuming $\Delta_W = \Delta_Z \equiv \Delta_V$. In all models considered below, we also find $\Delta_V < 0$. For example, two-Higgs-doublet models can violate both of these features at the loop level, so eventually we should release this simplification. We choose ξ such that in the custodial limit the (effective) Higgs coupling to gauge bosons gets modified as

$$\frac{g_V}{g_V^{\rm SM}} = 1 - \frac{\xi^2}{2} + \mathcal{O}(\xi^3) \quad \text{or} \quad \Delta_V = -\frac{\xi^2}{2} + \mathcal{O}(\xi^3). \tag{32}$$

A. Dark singlet

Dark singlet models include an additional scalar S which couples to the Higgs through dimension-four portal interactions [44]. The extra scalar does not form a vacuum expectation value (VEV), and its decays are precluded by a discrete Z_2 symmetry. Because it couples to the Higgs, it leads to an invisible decay width and, hence, impacts the Higgs measurements through an invisible width,

$$\begin{aligned} &\Gamma_{\rm inv} = \xi^2 \Gamma_{\rm SM} \quad \text{with} \\ &\mu_{p,d} = \frac{\Gamma_{\rm SM}}{\Gamma_{\rm SM}+\Gamma_{\rm inv}} = 1 - \xi^2 + \mathcal{O}(\xi^3) < 1. \end{aligned} \tag{33}$$

In this case. the scaling pattern of Eq. (32) does not hold for the actual coupling g_V but for its apparent value from the rate measurement.

B. Additional singlet

In the presence of an additional $SU(2)_L$ singlet with nonzero VEV, its mixing with the Higgs boson is described by an angle $\sin\theta \equiv \xi$ [45]. All Higgs couplings to fermions and gauge bosons are rescaled by the common factor

$$\begin{aligned} 1+\Delta_x = \cos\theta = \sqrt{1-\xi^2} \quad &\text{with} \\ \mu_{p,d} = 1 - \xi^2 + \mathcal{O}(\xi^3) < 1. \end{aligned} \tag{34}$$

The phenomenological equivalence of the dark singlet, the singlet mixing, and the simplest strongly interacting single form factor models can only be broken by an observation of invisible Higgs decays.

C. Composite Higgs

Minimal composite Higgs models (MCHM) describe the Higgs boson as a pseudo-Nambu-Goldstone boson in a new strongly interacting sector with a spontaneously broken global symmetry [46,47]. In the Randall-Sundrum picture,

this global symmetry has to include the local gauge groups of the Standard Model and can, in addition, include a global custodial symmetry of the Standard Model. This way, one of the states in the Higgs sector will be light, while the others reside at $f \gg m_H$. The light Higgs couplings are shifted proportionally to the ratio[1]

$$\xi = \frac{v}{f}. \tag{35}$$

Depending on the symmetry structure, this Goldstone-protected strongly interacting Higgs sector predicts different coupling patterns for fermions and gauge bosons. In the MCHM5 setup, the ratio of production rates scales like

$$\frac{\mu_{\mathrm{VBF},d}}{\mu_{\mathrm{GF},d}} = \left(\frac{1+\Delta_V}{1+\Delta_f}\right)^2 = \frac{(1-\xi^2)^2}{(1-2\xi^2)^2} = 1 + 2\xi^2 + \mathcal{O}(\xi^3). \tag{36}$$

D. Additional doublet

In the most general setup with an additional Higgs doublet [48], the Yukawa-aligned two Higgs doublet model (2HDM), the different Higgs couplings to the heavy fermions vary independently. However, the light Higgs couplings to the massive gauge bosons are universally modified by the mixing angle $\cos(\beta-\alpha) \equiv \xi$,

$$1 + \Delta_V = \sin(\beta-\alpha) = \sqrt{1-\xi^2}. \tag{37}$$

The decoupling parameter ξ parametrizes the distance from the Standard Model limit. In the fermion sector, four setups accommodate the flavor symmetry of the Standard Model [49]:

(i) type I, where all fermions couple to just one Higgs doublet Φ_2,
(ii) type II, where up-type (down-type) fermions couple exclusively to Φ_2 (Φ_1),
(iii) lepton specific, with a type-I quark sector and a type-II lepton sector,
(iv) flipped, with a type-II quark sector and a type-I lepton sector.

These natural flavor-conserving models correspond to particular cases of the aligned 2HDM with specific alignment angles. The modification factors $\Delta_{t,b,\tau}$ depend on the mixing angles α and β. We vary $\sin(\beta-\alpha)$ while fixing $\tan\beta = 1.5$. We also fix the charged Higgs contribution to g_γ via $m_{H^\pm} = 500$ GeV and $\tilde{\lambda} = 1$ for the relevant self-coupling defined in Ref. [41]. This choice leads to a small charged Higgs contribution with the same sign as the top loop. It sharpens the destructive interference with the W loop and yields $\mu_{p,\gamma\gamma} < 1$ in the limit $\xi \to 0$.

[1]Note that for consistency reasons, this definition of ξ differs from the original $\xi = (v/f)^2$.

E. Minimal supersymmetric Standard Model

The Higgs sector of the minimal supersymmetric Standard Model (MSSM) is a particular case of a type-II 2HDM with a supersymmetric Higgs potential [50]. This identifies the Higgs self-couplings with gauge couplings, so neither the additional heavy Higgs masses nor the mixing angle α are free quantities anymore. The decoupling limit can be described in terms of the general 2HDM, as defined in Eq. (37). The only difference is that for the MSSM, the gauge couplings are related to the heavy Higgs masses as

$$\xi^2 = \cos^2(\beta-\alpha) \simeq \frac{m_{h^0}^2(m_Z^2 - m_{h^0}^2)}{m_{A^0}^2(m_{H^0}^2 - m_{h^0}^2)} \simeq \frac{m_Z^4 \sin^2(2\beta)}{m_{A^0}^4} \tag{38}$$

entering the gauge boson couplings just as in Eq. (37). These tree-level Higgs mass and coupling patterns may be strongly modified by the inclusion of quantum effects [51]. In the computation of the decoupling parameter ξ, we, therefore, use α_{eff} defined as the mixing angle of the two scalar Higgs states with the loop-corrected mass matrix. The ratio of the two vacuum expectation values is fixed to $\tan\beta = 10$. The resulting size of ξ is limited by the range $m_{A^0} = 200$–1000 GeV, which will mean that we never reach $\xi = 0.2$.

We use the MSSM Higgs boson cross sections and branching ratios given by FeynHiggs [52]. The Higgs signal strengths are defined after identifying the (lightest) Higgs masses in both models $m_H^{\mathrm{SM}} = m_{h^0}^{\mathrm{MSSM}} = 126$ GeV. We show a benchmark scenario with a maximum light Higgs mass generated through large stop mixing (m_h^{max}). Compared to the general 2HDM, the possible departures from the linear correlations are milder. This can be understood from the more constrained 2HDM potential in the MSSM case, which implies a fast transition to the decoupling regime [41]. The largest deviations arise in the low-m_{A^0} regime or for light stop and stau masses.

F. Signal strengths

For the above-described modifications to a Standard-Model-like Higgs sector, we find simple patterns in the two-dimensional plane of coupling strengths. In Fig. 5, we compare the correlated modifications as functions of the decoupling parameter ξ. In general, the behavior in the $H \to VV$ and $H \to \gamma\gamma$ decay planes should be similar, as long as the loop-induced Higgs-photon coupling is dominated by the W loop. We focus on $\xi < 0.4$ corresponding to a modification of g_V by 8%, which the LHC is sensitive to with sufficient luminosity. In addition, we mark a deviation by $\xi = 0.2$ equivalent to a 2% coupling deviation. The latter could be considered the target of a linear collider analysis.

For the simplest models, the dark singlet and the singlet mixing, all correlations follow a straight diagonal line

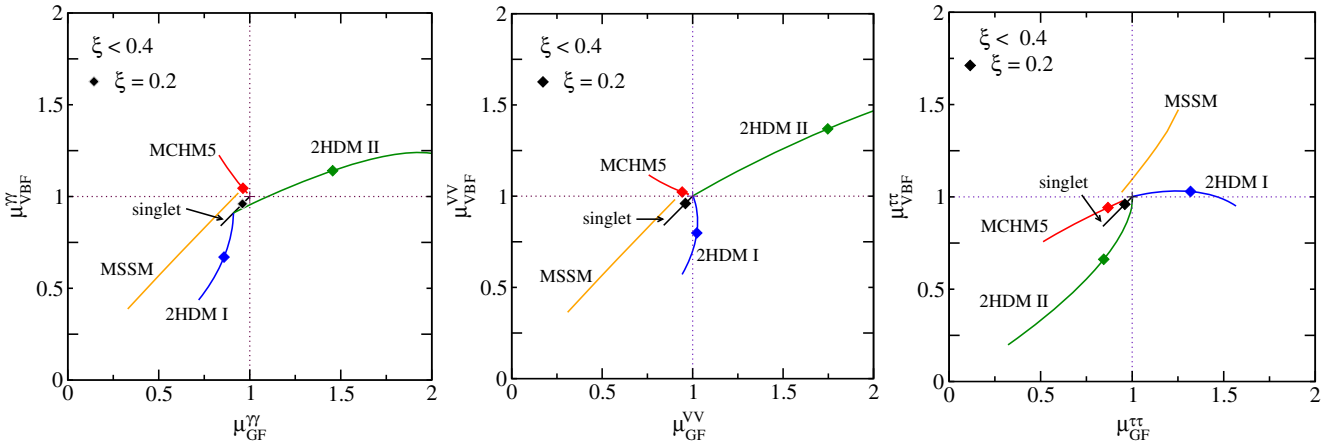

FIG. 5 (color online). Decay-diagonal correlations of signal strengths $\mu_{\mathrm{GF},d}$ vs $\mu_{\mathrm{VBF},d}$ for $d = \gamma\gamma, VV, \tau\tau$ in different models. The coupling variation is limited to $\xi < 0.4$ and the value $\xi = 0.2$ is singled out. The slight deviations from a complete decoupling are discussed in the text.

towards reduced coupling strengths $\mu_{p,d} < 1$. This is due to the simple mixing pattern and the net scaling of the LHC event rates as $\sigma \times \mathrm{BR} \propto g^2$. The same pattern appears for the simplest strongly interacting models with a single Higgs form factor. For the more complex strongly interacting model MCHM5, we find both correlated ($\mu_{\mathrm{GF},\tau\tau} - \mu_{\mathrm{VBF},\tau\tau}$) and anticorrelated ($\mu_{\mathrm{GF},VV} - \mu_{\mathrm{VBF},VV}$) patterns. It predicts an increased number of weak-boson-fusion events whenever the couplings Δ_V in the production process decrease more slowly than the coupling Δ_f in the total width, as shown in Eq. (36).

Larger departures from the Standard Model are possible in the 2HDM, thanks to its more flexible coupling structure. In the type-I setup, there is essentially no mechanism to increase the number of weak-boson-fusion events as compared to the Standard Model because of the increase in the Higgs width combined with the reduced gauge boson coupling. This is described in more detail in Appendix B. For gluon-fusion production combined with a fermionic decay, the suppression by the total width can be compensated by the production and decay couplings. In the type-II setup, both signal strengths can, unexpectedly, be enhanced for bosonic Higgs decays. The reason is a strongly decreased partial Higgs width to bottoms and taus which cannot be generated in the type-I model. For fermionic decays, the direct link between the bottom and tau down-type Yukawas leads to a systematically decreased event rate. The fact that for $\xi \to 0$ the 2HDM rates do not match the Standard Model is linked to the finite contribution of the charged Higgs to the effective photon-Higgs coupling.

Finally, the MSSM as a constrained type-II 2HDM shows limited signal strength variations because of the supersymmetric constraints. Unlike the general 2HDM, even in the type-II setup the MSSM does not allow for a free variation of the two parameters α and β, which affect the Yukawa couplings in a complicated manner. Departures from the decoupling limit in this case lie below $\xi \lesssim 0.2$ for the considered parameter space configurations. In the MSSM, deviations from the Standard Model in the limit $\xi \to 0$ arise through contributions of the sfermions and the charged Higgs to both the effective gluon-Higgs and photon-Higgs couplings. In that sense, the parameter ξ does not fully track down the decoupling limit for the loop-induced Higgs couplings, similar to the 2HDM case. The effect of a shifted bottom Yukawa is not sufficient to overcome the reduction in g_V, which means that unlike for the 2HDM, both signal strength deviations for the VV and $\gamma\gamma$ decays are negative, leading to a (almost linear) correlated suppression. Moreover, we see that the typical deviations in the signal strengths can be achieved for small values $\xi < 0.2$ in the MSSM because quantum effects dominate over the mere tree-level rescaling Δ_V.

G. A heuristic for robustness to theory uncertainty

One of the great challenges of the Higgs coupling program is to understand what type of deviation from the Standard Model prediction would be compelling enough to make a claim for new physics. It is clear that the pattern of deviations in the various production and decay modes carries much more information than considering them individually. Furthermore, the inability to measure the total width of the Higgs necessitates either assumptions on the total width or consideration of various ratios in which the total width cancels. Both approaches lead to strong correlations in the inferred couplings. If the theoretical uncertainties were well defined and statistical in nature, the significance of any given deviation could be readily assessed by standard statistical methods. However, the ill-defined nature of theoretical uncertainties is beyond the scope of rigorous statistical procedures.

For example, if one were to see a 4σ deviation from the Standard Model that could be reduced to a 2σ deviation by

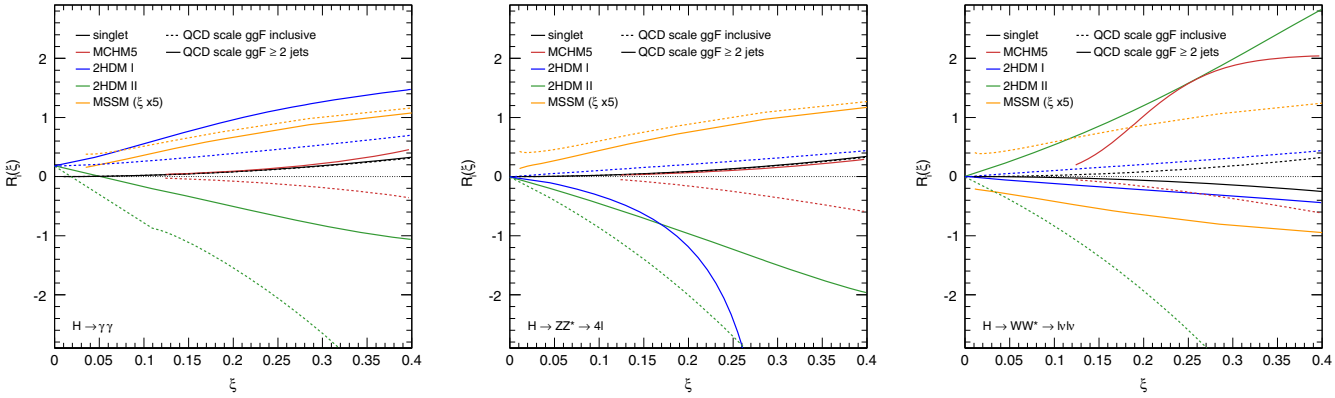

FIG. 6 (color online). The sensitivity heuristic $R_i(\xi)$ evaluated for various new physics models and the theoretical uncertainties i associated to the gluon-fusion cross section for ≥ 0 jets and ≥ 2 jets.

inflating a theoretical uncertainty by a factor of 2 or by changing from a Gaussian constraint to a RFIT constraint, then the deviation would most likely be met with healthy skepticism by the community. However, a deviation that is orthogonal to the effect of a theoretical uncertainty is much more robust.

This motivates a heuristic to evaluate the robustness of an observed deviation $\hat{\boldsymbol{\mu}}$ to a theoretical uncertainty parametrized by α_i, which we denote $R_i(\boldsymbol{\mu})$. We want a larger deviation from the Standard Model to be reflected as a larger value for this robustness heuristic, so we begin with the length of the vector $|\boldsymbol{\mu} - \boldsymbol{1}|$. We also want the robustness to be larger as the deviation from the Standard Model $(\boldsymbol{\mu} - \boldsymbol{1})$ becomes more orthogonal to the shift in the signal strength induced from varying α_i as defined in Eq. (22) and denoted as $\partial_{\alpha_i}\boldsymbol{\mu}^{\text{fix}}$. Since the magnitude of the theoretical uncertainty is poorly defined, it is natural that we only consider the angle between $(\boldsymbol{\mu} - \boldsymbol{1})$ and $\partial_{\alpha_i}\boldsymbol{\mu}^{\text{fix}}$. This leads to the robustness heuristic

$$R_i(\boldsymbol{\mu}) = \frac{|\boldsymbol{\mu} - \boldsymbol{1}|^2 |\partial_{\alpha_i}\boldsymbol{\mu}^{\text{fix}}|}{(\boldsymbol{\mu} - \boldsymbol{1}) \cdot (\partial_{\alpha_i}\boldsymbol{\mu}^{\text{fix}})}. \tag{39}$$

This same heuristic can be used to assess an expected departure from the Standard Model based on a new physics model parametrized by ξ via composition $R_i(\boldsymbol{\mu}(\xi))$. Figure 6 shows the heuristic for the new physics effects in Fig. 5 combined with the theoretical uncertainties in the toy model illustrated in Fig. 2. As expected from the discussion of the different models, the different two-Higgs-doublet models can be distinguished from the Standard Model even for $\xi \sim 0$. While for the type-I model, the VBF topology is more robust with respect to potential QCD effects, deviations due to a type-II model are more robust to the inclusive Higgs production rate. The most reliable signature for the strongly interacting MCHM5 model would be observed in weak-boson-fusion Higgs production with a decay $H \to WW$. For the MSSM, the ξ axis is rescaled because deviations $\xi \gtrsim 0.1$ are hardly generated in our scan over MSSM spectra. The robustness of a supersymmetric Higgs sector is roughly equal for the inclusive and VBF Higgs topologies.

VI. CONCLUSION

Motivated by the fact that ill-defined theoretical uncertainties will eventually be the limiting factor in Higgs coupling measurements at the LHC, we have developed a technique to decouple the theoretical uncertainties from the experimental results while retaining the ability to incorporate those uncertainties in a subsequent stage we refer to as recoupling.

This approach is amenable to simultaneously measuring multiple quantities, such as a vector of signal strength parameters $\boldsymbol{\mu}$ for different Higgs production and decay signatures. Moreover, the technique lends itself well to combinations with several common sources of uncertainty that induce correlations among the contributing measurements. In that respect, it is similar to the BLUE [11] technique but not restricted to Gaussian measurements or linear response to the source of uncertainty. We have considered a toy example modeled after the current ATLAS Higgs coupling measurements where the measurements are not in the Gaussian regime, and these nonlinear effects are important for approximating the full likelihood function.

One of the most powerful features of this approach is that it allows one to change the assumptions on both the magnitude and the shape of the uncertainty in the recoupling stage, which may occur long after the experimental groups have released their results. This includes the ability to introduce *a priori* correlations in the source of the systematics, which might have been neglected originally.

These same capabilities would be possible if the experiments published the full statistical model using ROOFIT/ROOSTATS [25] as suggested in Ref. [53]; however, the approach outlined here is less technology dependent. The

conceptual picture of these uncertainties leading to shifts in the inferred values of the parameters is intuitive and provides convenient visualizations as in Fig. 2.

In order for the experiments to present their results in this way, they would need to

(i) publish the effective likelihood $L_{\text{eff}}(\boldsymbol{\mu}^{\text{eff}})$ profiling only parameters that are not anticipated to be common to other measurements,
(ii) publish the reparametrization template $\boldsymbol{\mu}^{\text{eff}}(\boldsymbol{\mu}, \boldsymbol{\alpha})$, and
(iii) document the conventions that establish meaning to the nuisance parameters $\boldsymbol{\alpha}$.

In return, our approach allows for a flexible treatment of systematic uncertainties and removes the burden of choosing a description of theoretical uncertainties from the experimental groups. It also allows for future improvements in the theoretical description of Higgs processes at the LHC to be easily incorporated. We have discussed multiple strategies to determine the reparametrization template in Sec. II.

In the absence of a reliable measure of theoretical uncertainties, the key question becomes how easily an apparent deviation of experimental measurements from the Standard Model description can be explained by a change in the assumed theoretical uncertainties. If we describe new physics effects in the Higgs sector as one-parameter deviations from the Standard Model decoupling limit in the signal strength planes, the directions of these deviations can be compared to the effects of a change in the theoretical uncertainties in the same planes. The direct comparison of these two possible explanations for an experimental observation leads us to the robustness heuristic presented in Eq. (39).

While this paper has focused on the application of our approach to Higgs coupling measurements and the theoretical uncertainties associated with QCD, the technique is quite general and may find broad applications.

ACKNOWLEDGMENTS

The authors thank M. Dührssen for discussion and encouragement early in the process of this project. We thank M. Rauch for presenting preliminary results on our behalf. We also thank Jamison Galloway for feedback on an early draft. D. L. V. is supported by the F.R.S.-FNRS "Fonds de la Recherche Scientifique" (Belgium). T. P. would like to thank the CCPP at New York University for their hospitality and acknowledges BMBF support under Project No. 05H12VHE. K. C. and S. K. are supported by the U.S. National Science Foundation Grants No. PHY-0854724 and No. PHY-0955626.

APPENDIX A: A WORKED EXAMPLE

In this appendix, we work through a simple example that illustrates explicitly the procedure described in Sec. II. A MATHEMATICA notebook carrying out these calculations can be found in Ref. [40]. We consider two Gaussian measurements (indexed by $c = 1, 2$), which can be thought of as approximating the Poisson distribution for number counting analyses or approximating the maximum likelihood estimator of some more complicated analysis. We consider two signal processes (indexed by $p = 1, 2$) with nominal (e.g., Standard Model) expectations s_{cp} and signal strength modifiers μ_p. Finally, we consider one systematic effect parametrized by α that shifts the signal expectation from the nominal $\alpha_0 = 0$ so that the expectation is

$$\nu_c(\mu_1, \mu_2, \alpha) = \sum_{p=1,2} \mu_p s_{cp}(1 + \eta_{cp}\alpha) + b_c. \tag{A1}$$

Including a Gaussian constraint term for the parameter α leads to the full likelihood

$$L_{\text{full}}(\nu_1, \nu_2, \alpha) = \text{Gaus}(x_1|\nu_1, \sigma_1)\text{Gaus}(x_2|\nu_2, \sigma_2)\text{Gaus}(a = 0|\alpha, \sigma_\alpha). \tag{A2}$$

The maximum likelihood estimators are

$$\begin{aligned}
\hat{\mu}_1 &= \frac{s_{12}(x_2 - b_2) - s_{22}(x_1 - b_1)}{s_{12}s_{21} - s_{11}s_{22}}, \\
\hat{\mu}_2 &= \frac{s_{21}(x_1 - b_1) - s_{11}(x_2 - b_2)}{s_{12}s_{21} - s_{11}s_{22}}, \\
\hat{\alpha} &= 0.
\end{aligned} \tag{A3}$$

The full Fisher information matrix is straightforward to calculate but cumbersome to write explicitly. The results below only require the off-diagonal block elements:

$$V^{-1}_{\text{full }13} = \frac{1}{\sigma_1^2\sigma_2^2}[b_1\eta_1 s_{11}\sigma_2^2 + \eta_2\hat{\mu}_2(s_{11}s_{12}\sigma_2^2 + \sigma_1^2 s_{21}s_{22}) + \eta_1(b_2\sigma_1^2 s_{21} + 2\hat{\mu}_1(s_{11}^2\sigma_2^2 + \sigma_1^2 s_{21}^2) + \hat{\mu}_2(s_{11}s_{12}\sigma_2^2 + \sigma_1^2 s_{21}s_{22}) - s_{11}\sigma_2^2 x_1 - \sigma_1^2 s_{21}x_2)],$$

$$V^{-1}_{\text{full }23} = \frac{1}{\sigma_1^2\sigma_2^2}[b_1\eta_2 s_{12}\sigma_2^2 + \eta_1\hat{\mu}_1(s_{11}s_{12}\sigma_2^2 + \sigma_1^2 s_{21}s_{22}) + \eta_2(b_2\sigma_1^2 s_{22} + \hat{\mu}_1(s_{11}s_{12}\sigma_2^2 + \sigma_1^2 s_{21}s_{22}) + 2\hat{\mu}_2(s_{12}^2\sigma_2^2 + \sigma_1^2 s_{22}^2) - s_{12}\sigma_2^2 x_1 - \sigma_1^2 s_{22}x_2)]. \quad \text{(A4)}$$

In order to decouple the uncertainty, the experiment would provide the likelihood for an effective signal strength with respect to the nominal prediction $\alpha_0 = 0$. The effective likelihood would be based on

$$\nu_c(\mu_1^{\text{eff}}, \mu_2^{\text{eff}}) = \sum_{p=1,2} \mu_p^{\text{eff}} s_{cp} + b_c, \quad \text{(A5)}$$

which has the same maximum likelihood estimates above and the following information matrix:

$$V^{-1}_{\text{eff}} = \begin{bmatrix} \frac{s_{11}^2}{\sigma_1^2} + \frac{s_{21}^2}{\sigma_2^2} & \frac{s_{11}s_{12}}{\sigma_1^2} + \frac{s_{21}s_{22}}{\sigma_2^2} \\ \frac{s_{11}s_{12}}{\sigma_1^2} + \frac{s_{21}s_{22}}{\sigma_2^2} & \frac{s_{22}^2}{\sigma_2^2 + \frac{s_{12}^2}{\sigma_1^2}} \end{bmatrix}. \quad \text{(A6)}$$

The effective likelihood has the form:

$$L_{\text{eff}}(\mu_1^{\text{eff}}, \mu_2^{\text{eff}}) \propto G(\hat{\mu}_1, \hat{\mu}_2 | \mu_1^{\text{eff}}, \mu_2^{\text{eff}}, V^{-1}_{\text{eff}}). \quad \text{(A7)}$$

Now we must choose a reparametrization template to be used to recouple the uncertainty due to α. First, let us choose the template of Eq. (10) so that $\mu_p^{\text{eff}} = \mu_p(1 + \eta_p\alpha)$. Note that the original model in Eq. (A2) had four η_{cp} while the template only has two η_p. If the effect of α is category universal so that $\eta_{c=1,p} = \eta_{c=2,p}$, then recoupling based on this template will reproduce the full model exactly; however, in the more general situation, $\eta_{c=1,p} \neq \eta_{c=2,p}$, it will not.

Now we solve for the coefficients of the template based on the local covariance matrix as described in Sec. II D. We outlined three equivalent approaches based on Eqs. (22), (25), and (26). Let us demonstrate the last of these three approaches. Based on the template of Eq. (10), the reparametrization will lead to the Jacobian

$$J = \frac{\partial(\mu_1^{\text{eff}}, \mu_2^{\text{eff}})}{\partial(\mu_1, \mu_2, \alpha)} = \begin{bmatrix} (1+\eta_1\alpha) & 0 & \mu_1\eta_1 \\ 0 & (1+\eta_2\alpha) & \mu_2\eta_2 \end{bmatrix}. \quad \text{(A8)}$$

As in Eq. (26), we use this Jacobian to relate the information matrix of the effective likelihood and the main measurement

$$V^{-1}_{\text{main}}(\mu_1, \mu_2, \alpha) = J^T V^{-1}_{\text{eff}} J. \quad \text{(A9)}$$

The $(i = \mu_p, j = \alpha)$ sub-block of this matrix leads to the following system of linear equations

$$\hat{\mu}_1\eta_1 V^{-1}_{\text{eff }11} + \hat{\mu}_2\eta_2 V^{-1}_{\text{eff }12} = V^{-1}_{\text{main }13},$$
$$\hat{\mu}_1\eta_1 V^{-1}_{\text{eff }12} + \hat{\mu}_2\eta_2 V^{-1}_{\text{eff }22} = V^{-1}_{\text{main }23}, \quad \text{(A10)}$$

which can easily be inverted to provide solutions for η_p. In practice, one would work with numerical representations of the maximum likelihood estimators and Fisher information matrices, but here we present the result symbolically:

$$\eta_1 = [b_1 s_{22}(-\eta_{12}s_{12}s_{21} - \eta_{21}s_{12}s_{21} + \eta_{22}s_{12}s_{21} + \eta_{11}s_{11}s_{22}) + b_2 s_{12}(\eta_{21}s_{12}s_{21} - (\eta_{11} - \eta_{12} + \eta_{22})s_{11}s_{22}) + \eta_{12}s_{12}s_{21}s_{22}x_1 + \eta_{21}s_{12}s_{21}s_{22}x_1 - \eta_{22}s_{12}s_{21}s_{22}x_1 - \eta_{11}s_{11}s_{22}^2 x_1 - \eta_{21}s_{12}^2 s_{21}x_2 + \eta_{11}s_{11}s_{12}s_{22}x_2 - \eta_{12}s_{11}s_{12}s_{22}x_2 + \eta_{22}s_{11}s_{12}s_{22}x_2]/[(s_{12}s_{21} - s_{11}s_{22})(b_2 s_{12} - b_1 s_{22} + s_{22}x_1 - s_{12}x_2)],$$

$$\eta_2 = [b_2 s_{11}(\eta_{11}s_{12}s_{21} - \eta_{12}s_{12}s_{21} - \eta_{21}s_{12}s_{21} + \eta_{22}s_{11}s_{22}) + b_1 s_{21}(\eta_{12}s_{12}s_{21} - (\eta_{11} - \eta_{21} + \eta_{22})s_{11}s_{22}) - \eta_{12}s_{12}s_{21}^2 x_1 + \eta_{11}s_{11}s_{21}s_{22}x_1 - \eta_{21}s_{11}s_{21}s_{22}x_1 + \eta_{22}s_{11}s_{21}s_{22}x_1 - \eta_{11}s_{11}s_{12}s_{21}x_2 + \eta_{12}s_{11}s_{12}s_{21}x_2 + \eta_{21}s_{11}s_{12}s_{21}x_2 - \eta_{22}s_{11}^2 s_{22}x_2]/[(s_{12}s_{21} - s_{11}s_{22})(-b_2 s_{11} + b_1 s_{21} - s_{21}x_1 + s_{11}x_2)]. \quad \text{(A11)}$$

Note, in the category-universal situation, the solution simplifies to $\eta_p = \eta_{1p} = \eta_{2p}$ as expected. In the non-category-universal situation, it is possible for η_p to lie outside of the range $[\eta_{1p}, \eta_{2p}]$.

TABLE I. The values for the coefficients in Eq. (A1) used to define three example scenarios.

Scenario	σ_α	s_{11}	s_{12}	b_1	x_1	σ_1	s_{21}	s_{22}	b_2	x_2	σ_2	η_{11}	η_{12}	η_{21}	η_{22}
A	1	45	5	50	100	10	10	90	0	100	10	0.2	0.2	0.2	0.2
B	1	45	5	50	100	10	10	90	0	100	10	0.1	0.2	0.3	0.2
C	1	45	5	50	100	10	40	60	0	100	10	0	0	0.2	0

1. Three example scenarios

To make these examples more explicit, Table I specifies three scenarios for the coefficients in Eq. (A1). Scenarios A and B are meant to be representative of the LHC Higgs measurements in which the first category is gluon-fusion-like with significant background and signal dominated by the first production mode ($p = 1 =$ ggF), while the second category is VBF-like with negligible background and signal dominated by the second production mode ($p = 2 =$ VBF). The only difference between scenarios A and B is the η_{cp} that quantify the response to the source of uncertainty. Since scenario A is category universal, we achieve an exact reproduction of the full likelihood, while for scenario B we do not expect exact results. Scenario C is meant to probe the extreme case in which the uncertainty for the first production mode only affects the expectation in the second category (which is representative of the theory uncertainty of gluon fusion $+ \geq 2$ jets).

Figure 7 shows a comparison of the full and recoupled likelihood using the "aligned" template of Eq. (10) and the solutions to the coefficients in Eq. (A11). Scenario A is reproduced exactly; the non-category-universal property of scenario B leads to a slight discrepancy, and the extreme non-category-universal property of scenario C leads to a substantial discrepancy. In particular, while the effect of the uncertainty in scenario C is to change the inferred value of μ_{VBF}, the size of this effect should scale with μ_{ggF}. The agreement can be improved dramatically by moving to the more general template of Eq. (14) that includes $\eta_{pi}^{p'}$. One approach is to fix by hand the coefficients of the template $\eta_{\mathrm{ggF}}^{\mathrm{ggF}} = 0$, $\eta_{\mathrm{VBF}}^{\mathrm{VBF}} = 0$, and $\phi = 0$ and determine the two remaining coefficients using the local covariance matrix

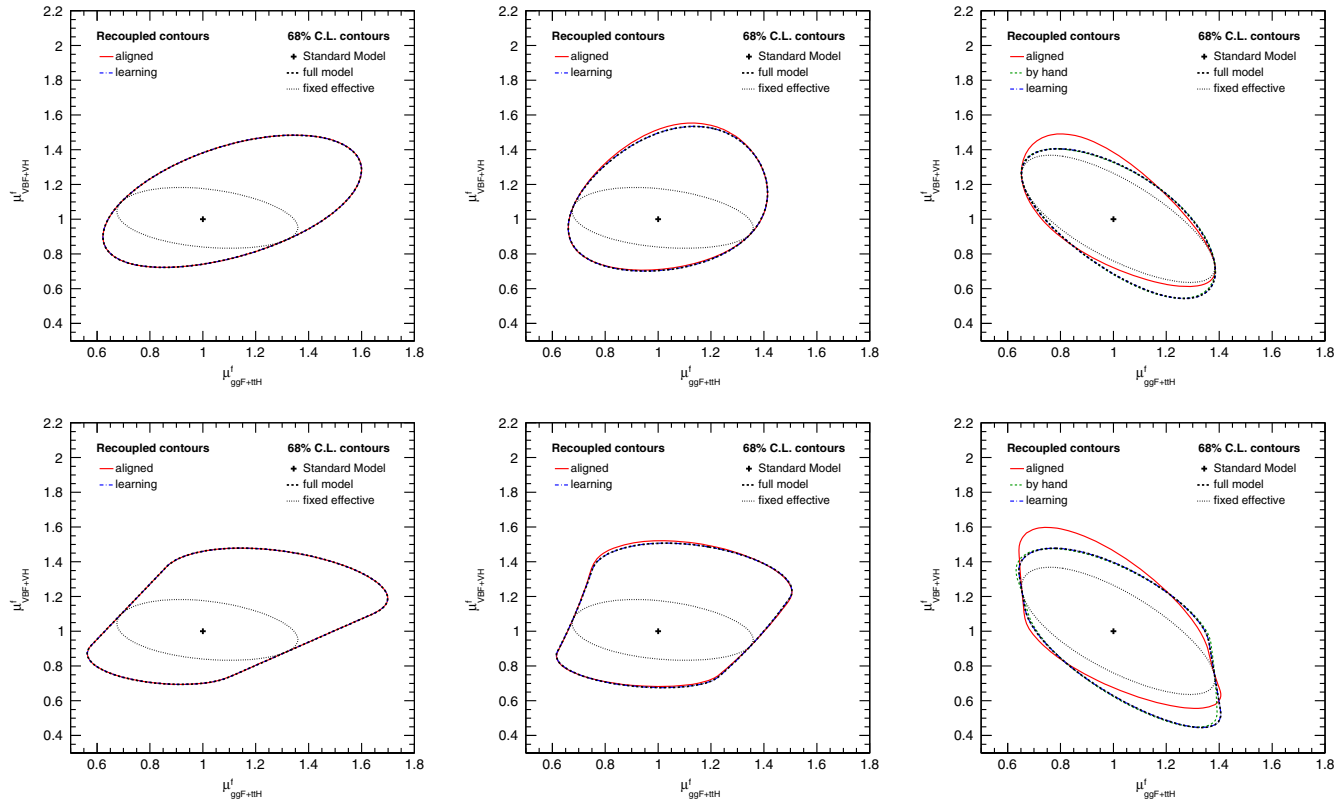

FIG. 7 (color online). Comparison of full likelihood (solid) and recouped (dashed) likelihood for scenarios A, B, and C. Scenario C illustrates the impact of using three templates aligned (red), by hand (green), and learning (blue) as described in the text. The top row is based on the nominal Gaussian constraint, and the bottom row shows the result of replacing it with an alternative RFIT constraint term. The effective likelihood with $\alpha = 0$ is shown as a dotted line.

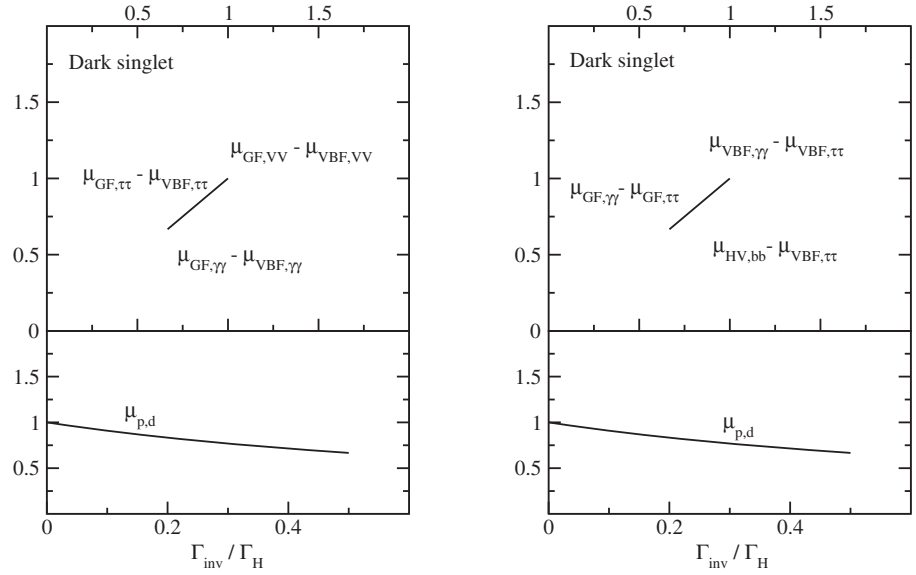

FIG. 8. Dark singlet: correlated Higgs signal strengths for the decay-diagonal channels (left) and nondiagonal channels (right). The first of the two signal strengths in the notation $\mu_{p_1,d_1} - \mu_{p_2,d_2}$ is shown on the vertical axis, the second on the horizontal axis. In the lower panels, we give the different signal strengths as a function of $\Gamma_{\rm inv}/\Gamma_{\rm SM} = \xi^2$.

(which proceeds as above with a different Jacobian transformation). An alternate approach is to use the unrestricted template of Eq. (14) and utilize the "learning" approach of Eq. (27).

Finally, the second row of Fig. 7 shows a comparison of the full and decoupled likelihood with a modified constraint term. In particular, the Gaussian constraint is replaced with a RFIT constraint term: $\text{Gaus}(0|\alpha, 1) \to \text{Uniform}(-1, 1)$. The coefficients for the templates are the same as for both rows of Fig. 7.

APPENDIX B: NEW PHYSICS MODELS

In this appendix, we will give a more detailed picture of the new physics models and their features briefly discussed in Sec. V. In particular, we will motivate and discuss the description of new physics effects by a single parameter ξ defined as the modification of the Higgs couplings to massive gauge bosons, i.e., $\Delta_V \simeq -\xi^2/2$. Note that this unified definition of ξ differs from Ref. [41] for some of the new physics models. All signal strength deviations we compute by rescaling the SM production cross section, branching ratio, and total width [54], while for the MSSM case, we use FEYNHIGGS [52].

1. Dark singlet

A dark singlet is defined as a model with an additional scalar particle S which does not have a vacuum expectation value and, hence, cannot mix with the Higgs boson. In addition, we assume that its only interaction with the Standard Model will be the dimension-four portal interactions in the combined scalar potential [44],

$$V(\Phi, S) = \mu_1^2(\Phi^\dagger\Phi) + \lambda_1|\Phi^\dagger\Phi|^2 + \lambda_3|\Phi^\dagger\Phi|S^2. \quad \text{(B1)}$$

This interaction with strength λ_3 can, if kinematically allowed, lead to an invisible Higgs decay width [55] directly linked to a possible dark matter agent [56]

$$\Gamma_{\rm inv} = \Gamma(h \to ss) = \frac{\lambda_3^2 v^2}{32\pi m_h}\sqrt{1 - \frac{4m_s^2}{m_h^2}} \equiv \xi^2\Gamma_{\rm SM}. \quad \text{(B2)}$$

Such an invisible Higgs width contributes to the total Higgs width and, hence, to the number of predicted Higgs events at the LHC. This leads to an apparent reduction of all couplings, including the Higgs coupling to massive gauge bosons, shown in Eq. (33). Such a universal modification of all predicted event numbers corresponds to a diagonal pattern, for example, in the two-dimensional $\mu_{\text{VBF},d}$ vs $\mu_{\text{GF},d}$ plane. In the upper panels of Fig. 8, we show these identical diagonal correlations for a set of decay-diagonal channels (left panels) and nondiagonal channels (right panels). The correlations are all identical. In the lower panels, we show the expected signal strengths relative to the Standard Model as a function of the invisible width.

2. Additional singlet

If the additional $SU(2)_L$ singlet acquires a finite VEV v_S, the combined Higgs potential [45,57]

$$V(\Phi, S) = \mu_1^2(\Phi^\dagger\Phi) + \lambda_1|\Phi^\dagger\Phi|^2 + \mu_2^2S^2 + \kappa S^3 + \lambda_2 S^4 + \lambda_3|\Phi^\dagger\Phi|S^2 \quad \text{(B3)}$$

with the portal interaction λ_3 leads to singlet-doublet mixing. The rotation to mass eigenstates h and H defines the angle

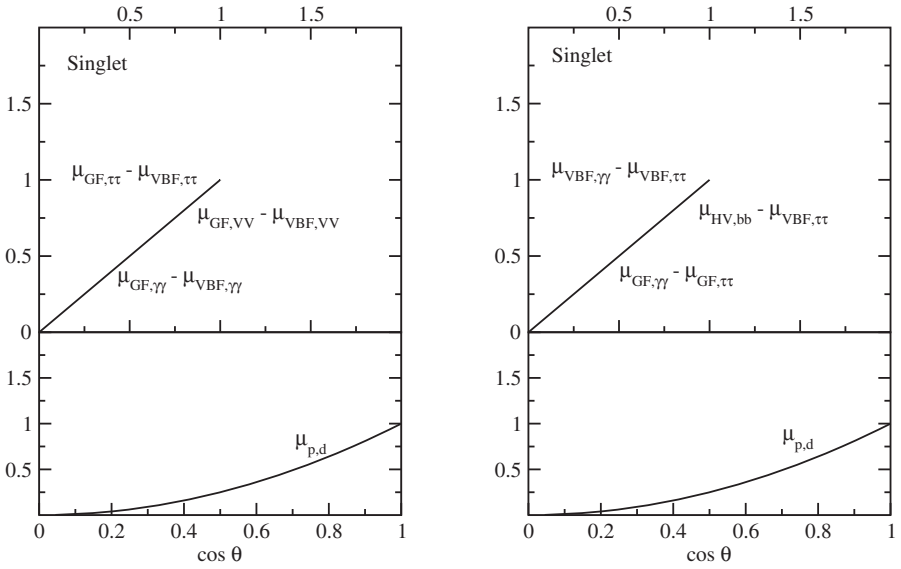

FIG. 9. Singlet mixing: correlated Higgs signal strengths for the decay-diagonal channels (left) and nondiagonal channels (right). The first of the two signal strengths in the notation $\mu_{p_1,d_1} - \mu_{p_2,d_2}$ is shown on the vertical axis, the second on the horizontal axis. In the lower panels, we give the different signal strengths as a function of $\cos\theta = \sqrt{1-\xi^2}$.

$$\tan^2(2\theta) = \frac{\lambda_3^2 v_H^2 v_S^2}{(\lambda_1 v_H^2 - \lambda_2 v_S^2)^2}. \tag{B4}$$

All Higgs couplings to fermions and gauge bosons are rescaled by $\cos\theta < 1$. In addition, the mostly Higgs state can decay into two lighter mostly singlet states [58]

$$\Gamma(H \to hh) = \frac{|\lambda_{Hhh}|^2}{32\pi m_H}\sqrt{1 - \frac{4m_h^2}{m_H^2}}. \tag{B5}$$

The signature for such a decay depends on the lifetime and the decay channels of the lighter state h. If sizeable, this additional decay channel contributes to the total Higgs width, entering the predicted number of LHC events as a second universal modification.

This universal coupling modification again predicts diagonal lines for all $\mu_{p_1,d_1} - \mu_{p_2,d_2}$ correlations, as shown in Fig. 9. If we do not observe the new Higgs decay modes for each of the additional singlet scenarios, the degeneracy in the signal strength deviations makes it impossible to distinguish a dark singlet, an additional singlet, and the simplest strongly interacting form factor models (as discussed in the next paragraph) [59].

3. Composite Higgs

Depending on the symmetry structure, this strongly interacting but Goldstone-protected Higgs sector predicts different coupling patterns for fermions and gauge bosons [47]. The MCHM4 setup with $\Delta_V = \Delta_f = \sqrt{1-\xi^2} - 1$ is experimentally equivalent to a mixing term in a singlet extension. The reduced couplings of the lightest Higgs state reflect the fact that at the energy scale f there exist many strongly interacting Higgs fields which share the unitarization of the usual $2 \to 2$ scattering processes. The MCHM5 setup is more interesting. While Δ_V is identical to the MCHM4 case, the fermions follow a different pattern $1 + \Delta_f = (1-2\xi^2)/\sqrt{1-\xi^2}$. If we ignore the heavy states' contributions to the effective Higgs-photon and Higgs-gluon couplings, we find that the ratio of production rates scales like

$$\begin{aligned}
\frac{\mu_{\mathrm{VBF},d}}{\mu_{\mathrm{GF},d}} &= \left(\frac{1+\Delta_V}{1+\Delta_f}\right)^2 = \frac{1-\xi^2}{\frac{(1-2\xi^2)^2}{1-\xi^2}} = \left(\frac{1-\xi^2}{1-2\xi^2}\right)^2 \\
&= 1 + 2\xi^2 + \mathcal{O}(\xi^3), \\
\frac{\mu_{\mathrm{VBF},VV}}{\mu_{\mathrm{GF},ff}} &= \left(\frac{1+\Delta_V}{1+\Delta_f}\right)^4 = 1 + 4\xi^2 + \mathcal{O}(\xi^3), \\
\frac{\mu_{\mathrm{VBF},ff}}{\mu_{\mathrm{GF},VV}} &= 1.
\end{aligned} \tag{B6}$$

Unlike all previous models, the MCHM5 setup accommodates $\mu_{p,d} > 1$, for example,

$$\begin{aligned}
\mu_{\mathrm{VBF},VV} &\simeq \frac{(1-\frac{\xi^2}{2})^4}{0.3(1-\frac{\xi^2}{2})^2 + 0.7(1-\frac{3\xi^2}{2})^2} \\
&\simeq 1 + 0.4\xi^2 + \mathcal{O}(\xi^3) > 1.
\end{aligned} \tag{B7}$$

The sharper suppression in the fermion couplings still leads to depleted gluon-fusion channels. In Fig. 10, we show a set of correlations between different signal strengths. The different possibilities portrayed in Eq. (B7) explain some

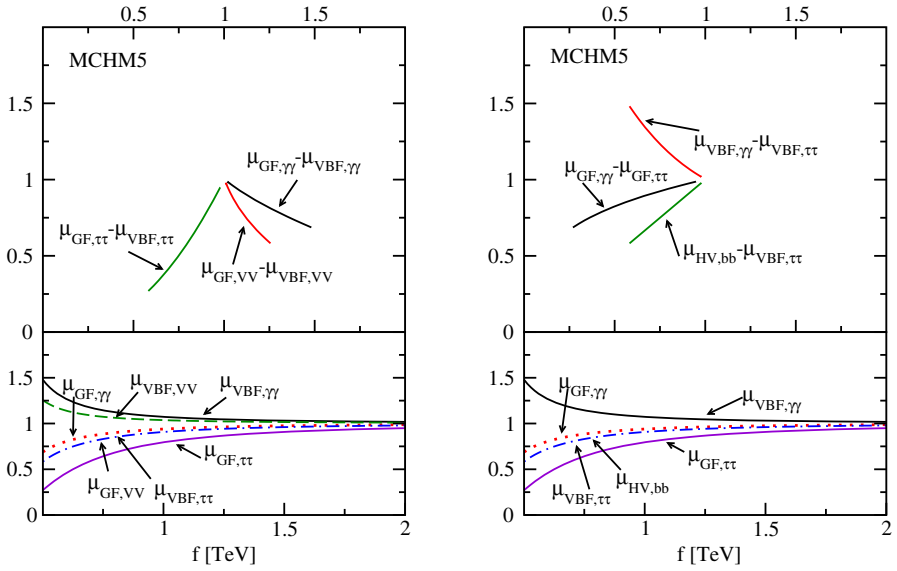

FIG. 10 (color online). Composite Higgs: correlated Higgs signal strengths for the decay-diagonal channels (left) and nondiagonal channels (right). The first of the two signal strengths in the notation $\mu_{p_1,d_1} - \mu_{p_2,d_2}$ is shown on the vertical axis, the second on the horizontal axis. In the lower panels, we give the different signal strengths as a function of $f = v/\xi$.

of the patterns in the signal strength planes. Corrections to the straight linear correlations arise at $\mathcal{O}(\xi^3)$.

4. Additional doublet

An even richer structure, in particular, in separating the gauge and Yukawa couplings of the light Higgs, appears when we add a second Higgs doublet [48,49]. Up to terms of $\mathcal{O}(\xi^3)$, the modification to the LHC signal strengths $(1+\Delta_p)^2(1+\Delta_d)^2 - 1$ is displayed in Table II. These expressions hold for the general Yukawa-aligned model [41,60]. For example, a type-I model corresponds to $\gamma_b = \gamma_\tau = \pi/2$, while a type-II model appears if we set $\gamma_b = \gamma_\tau = 0$. The main feature is that the leading signal strength deviations involving fermions arise at order ξ, while gauge couplings only vary with ξ^2 [61].

The detailed 2HDM signal strength patterns illustrated in Fig. 11 are to a large extend model dependent, as they are tied to the specific Yukawa structures. In the type-I setup, all Yukawas couplings and, hence, the effective Higgs-gluon coupling, are shifted by a common factor $\cos\alpha/\sin\beta$. This way, they are not suppressed for $\sin(\beta-\alpha) \lesssim 1$. The Higgs couplings to gauge bosons scale like $\sin(\beta-\alpha)$. As a result, all gluon-fusion channels show an increased signal strength as compared to their weak-boson-fusion counterparts. Type-II models link uplike and downlike fermions to different Higgs doublets, implying separate Yukawa modifications. The bottom and tau final states are suppressed for a wide range of $\sin(\beta-\alpha)$, eventually leading to the unphysically large signal strength deviations visible in the plots. For an even larger deviation from $\sin(\beta-\alpha) \simeq 1$ models of type-II leptonspecific models, and flipped models feature sign-inverted Yukawas. These sign ambiguities are visible in the ellipsoidal correlated variations in the signal strength plane. The additional charged Higgs-mediated contribution Δ_γ is responsible for the mild offset $\mu_{p,\gamma\gamma} \neq 1$ in the decoupling limit.

5. MSSM

A supersymmetric Lagrangian requires two Higgs doublets to give mass to up-type and down-type fermions. At tree level, the MSSM Higgs sector is a type-II 2HDM but with the different quartic Higgs couplings fixed to gauge couplings. If we assume the lighter of the scalar Higgs bosons to lie at 126 GeV, the heavy Higgs masses are

TABLE II. Higgs coupling modifications from an additional Higgs doublet.

	$b\bar{b}$	$\tau\tau$	VV
GF	$2\xi[\cot\beta - \tan(\beta-\gamma_b)]$ $-\xi^2[2-\tan^2(\beta-\gamma_b) - \cot^2\beta$ $+4\cot\beta\tan(\beta-\gamma_b)]$	$2\xi[\cot\beta + \tan(\beta-\gamma_\tau)]$ $-\xi^2[2-\tan^2(\beta-\gamma_\tau) - \cot^2\beta$ $+4\cos\beta\tan(\beta-\gamma_\tau)]$	$2\xi\cot\beta$ $-\xi^2(2-\cot^2\beta)$
VBF	$-2\xi\tan(\beta-\gamma_b)$ $-\xi^2[2-\tan^2(\beta-\gamma_b)]$	$-2\xi\tan(\beta-\gamma_\tau)$ $-\xi^2[2-\tan^2(\beta-\gamma_\tau)]$	$-2\xi^2$

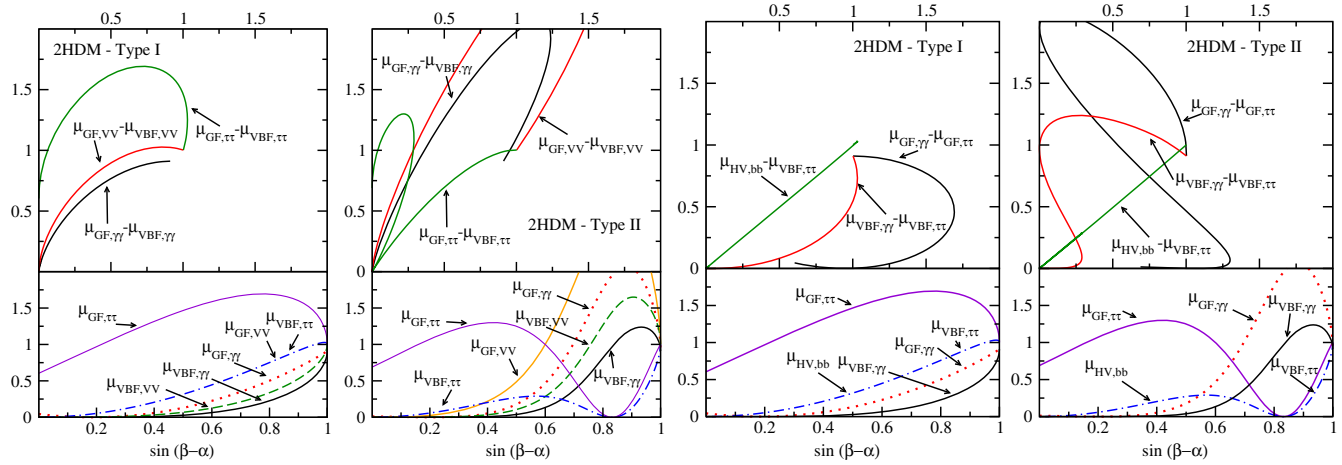

FIG. 11 (color online). 2HDM: correlated Higgs signal strengths for decay-diagonal channels (left two panels) and nondiagonal channels (right two panels), each for type-I and type-II setups. The first of the two signal strengths in the notation $\mu_{p_1,d_1} - \mu_{p_2,d_2}$ is shown on the vertical axis, the second on the horizontal axis. In the lower panels, we give the different signal strengths as a function of $\cos(\beta - \alpha) = \xi$.

almost degenerate at $m_{H^0} \simeq m_{A^0} \simeq m_{H^\pm}$ up to $\mathcal{O}(m_Z^2/m_{A^0}^2)$ corrections. Just like for the general 2HDM, the decoupling is described by $\xi = \cos(\beta - \alpha)$.

At tree level, the entire MSSM Higgs sector is fully described by m_{A^0} and $\tan\beta$. Loop corrections in the top-stop sector lead to significant corrections and yield an additional parameter dependence, for example, on the stop trilinear coupling A_t. The leading quantum corrections to the Higgs couplings can be accounted for through an effective mixing angle $\alpha \to \alpha_{\text{eff}}$ which we use in the definition of ξ. Additional contributions may further modify the effective Higgs coupling pattern. One-loop triangle corrections to the Yukawa couplings are governed by gluino-sbottom loops [62] and shift the bottom quark Yukawas by $\Delta_b \sim \mathcal{O}(m_Z^2 \tan^2\beta/m_{A^0}^2)$. They give rise to a delayed decoupling [50], meaning that the decoupling limit now demands $m_Z^2 \tan^2\beta \ll m_{A^0}^2$ for large $\tan\beta$. These effects depend on the detailed supersymmetric mass spectrum and have to be included in a full MSSM parameter study [63]. In this discussion of the light Higgs couplings, we omit any constraint from the new particles outside the Higgs sector. This also means that we ignore the possibility of light neutralinos $2m_{\chi_1^0} < m_{h^0}$ contributing to the invisible Higgs width. In specific supersymmetric models, the coupling of the lightest neutralino to the light Higgs boson can be important, leading to an accelerated dark matter annihilation in the early Universe. Nevertheless, even the

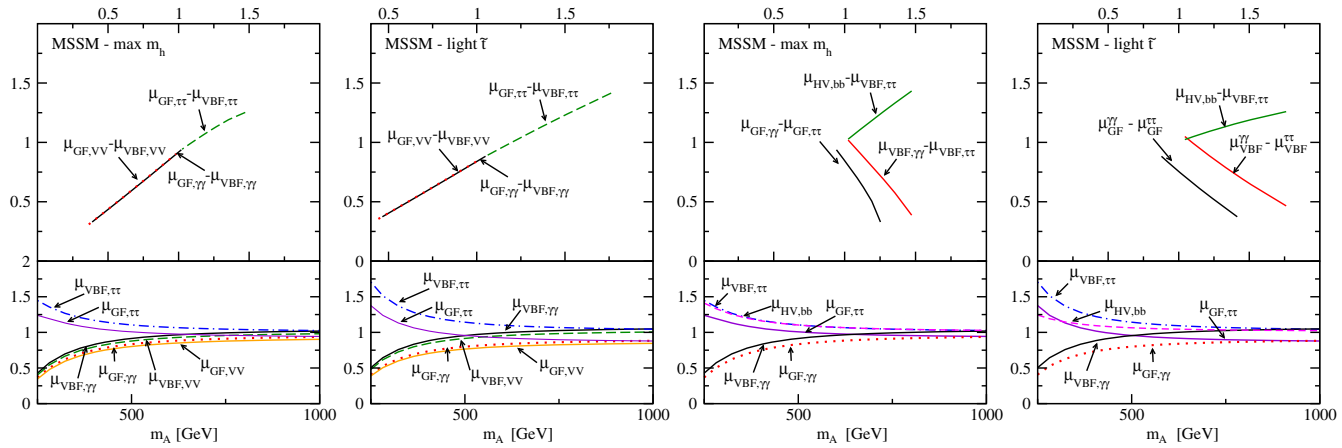

FIG. 12 (color online). MSSM: correlated Higgs signal strengths for decay-diagonal channels (left two panels) and nondiagonal channels (right two panels), each for two MSSM benchmark spectra. The first of the two signal strengths in the notation $\mu_{p_1,d_1} - \mu_{p_2,d_2}$ is shown on the vertical axis, the second on the horizontal axis. In the lower panels, we give the different signal strengths as a function of m_A, linked to ξ via Eq. (38).

couplings required for this light Higgs funnel are unlikely to give a measurable invisible Higgs decay rate at the LHC.

Technically, we use the MSSM Higgs cross sections and branching ratios given by FeynHiggs 2.9.5 [52]. The corresponding Higgs signal strengths are obtained by normalizing the individual production and decay rates to their SM counterparts [54] after identifying the (lightest) Higgs mass in both models $m_H^{\mathrm{SM}} = m_{h^0}^{\mathrm{MSSM}}$. We consider two benchmarks relevant to LHC searches [64], all compatible with the observed ∼126 GeV resonance: the m_h^{max} case with maximum stop mixing, heavy squarks, and $\tan\beta = 10$ and the light stop case with maximum stop mixing, but generally light squarks, and $\tan\beta = 35$. Both of them are numerically similar. The difference is that only the light stop scenario generates sizeable $\mathcal{O}(10\%)$ corrections to Δ_g and, at the same time, gives relatively large negative Δ_b corrections to the bottom Yukawa. In Fig. 12, we see that compared to the general 2HDM, the possible departures from the linear correlations are milder. The two benchmarks predict very similar signal strengths. All this can be understood from the more constrained Higgs potential and the moderate negative corrections to the total width. The largest deviations obviously arise in the low-m_{A^0} regime.

[1] P. W. Higgs, Phys. Lett. **12**, 132 (1964); Phys. Rev. Lett. **13**, 508 (1964); F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).
[2] ATLAS Collaboration, Phys. Lett. B **716**, 1 (2012).
[3] CMS Collaboration, Phys. Lett. B **716**, 30 (2012).
[4] For pedagogical introductions, see, e.g., A. Djouadi, Phys. Rep. **457**, 1 (2008); S. Dittmaier and M. Schumacher, Prog. Part. Nucl. Phys. **70**, 1 (2013); T. Plehn, Lect. Notes Phys. **844**, 1 (2012).
[5] ECFA High Luminosity LHC Experiments Workshop Report.
[6] M. E. Peskin, arXiv:1207.2516; M. Klute, R. Lafaye, T. Plehn, M. Rauch, and D. Zerwas, Europhys. Lett. **101**, 51001 (2013); S. Dawson *et al.*, arXiv:1310.8361.
[7] M. Dührssen, Report No. ATL-PHYS-2003-030, 2003; M. Dührssen, S. Heinemeyer, H. Logan, D. Rainwater, G. Weiglein, and D. Zeppenfeld, Phys. Rev. D **70**, 113009 (2004); M. Dührssen, Czech. J. Phys. **55**, B145 (2005).
[8] ATLAS Collaboration, Report No. ATLAS-CONF-2013-034, 2013.
[9] CMS Collaboration, Report No. CMS-PAS-HIG-13-005, 2013.
[10] R. Lafaye, T. Plehn, M. Rauch, D. Zerwas, and M. Dührssen, J. High Energy Phys. 08 (2009) 009; M. Klute, R. Lafaye, T. Plehn, M. Rauch, and D. Zerwas, Phys. Rev. Lett. **109**, 101801 (2012); T. Plehn and M. Rauch, Europhys. Lett. **100**, 11002 (2012).
[11] L. Lyons, D. Gibaut, and P. Clifford, Nucl. Instrum. Methods Phys. Res., Sect. A **270**, 110 (1988); A. Valassi, Nucl. Instrum. Methods Phys. Res., Sect. A **500**, 391 (2003).
[12] F. Boudjema *et al.*, arXiv:1307.5865.
[13] ATLAS Collaboration, Phys. Lett. B **726**, 88 (2013).
[14] ATLAS Collaboration, data from Fig. 7 from Ref. [13]; http://doi.org/10.7484/INSPIREHEP.DATA.A78C.HK44; http://doi.org/10.7484/INSPIREHEP.DATA.RF5P.6M3K; http://doi.org/10.7484/INSPIREHEP.DATA.26B4.TY5F.
[15] A. David, A. Denner, M. Dührssen, M. Grazzini, C. Grojean, G. Passarino, M. Schumacher *et al.* (LHC Higgs Cross Section Working Group), arXiv:1209.0040.
[16] M. Cacciari and N. Houdeau, J. High Energy Phys. 09 (2011) 039.
[17] A. David and G. Passarino, Phys. Lett. B **726**, 266 (2013).
[18] ATLAS Collaboration, Report No. ATL-PHYS-PUB-2011-011, 2011; Report No. CMS-NOTE-2011-005, 2011.
[19] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. C **71**, 1 (2011).
[20] ATLAS Collaboration, Phys. Rev. D **86**, 032003 (2012).
[21] J. Conway, arXiv:1103.0354.
[22] K. Cranmer *et al.*, Report No. CERN-OPEN-2012-016, 2012.
[23] A. Stuart, J. K. Ord, and S. Arnold, *Kendall's Advanced Theory of Statistics*, 6th ed., Classical Inference and the Linear Model Vol. 2A (Oxford University Press, New York, 1999), Chap. 29.
[24] K. Cranmer, S. Kreiss, D. López-Val, and T. Plehn, arXiv:1401.0080.
[25] W. Verkerke and D. P. Kirkby, arXiv:physics/0306116; L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke *et al.*, *Proc. Sci.*, ACAT2010 (2010) 057 [arXiv:1009.1003].
[26] For different aspects of such analyses, see, e.g., C. Englert, D. Goncalves-Netto, K. Mawatari, and T. Plehn, J. High Energy Phys. 01 (2013) 148; A. Djouadi, R. M. Godbole, B. Mellado, and K. Mohan, Phys. Lett. B **723**, 307 (2013); S. Banerjee, S. Mukhopadhyay, and B. Mukhopadhyaya, Phys. Rev. D **89**, 053010 (2014); I. Brivio, T. Corbett, O. J. P. Eboli, M. B. Gavela, J. Gonzalez-Fraile, M. C. Gonzalez-Garcia, L. Merlo, and S. Rigolin, J. High Energy Phys. 03 (2014) 024.
[27] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, Nucl. Phys. **B737**, 73 (2006); D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, J. High Energy Phys. 11 (2011) 064; .
[28] R. Kleiss and W. J. Stirling, Phys. Lett. B **200**, 193 (1988); U. Baur and E. W. N. Glover, Phys. Lett. B **252**, 683 (1990); V. D. Barger, K. Cheung, T. Han, J. Ohnemus, and D. Zeppenfeld, Phys. Rev. D **44**, 1426 (1991).

[29] V. D. Barger, R. J. N. Phillips, and D. Zeppenfeld, Phys. Lett. B **346**, 106 (1995); D. L. Rainwater, R. Szalapski, and D. Zeppenfeld, Phys. Rev. D **54**, 6680 (1996); V. Del Duca, G. Klämke, M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. D. Polosa, and D. Zeppenfeld, J. High Energy Phys. 10 (2006) 016; C. Bernaciak, B. Mellado, T. Plehn, P. Schichtel, and X. Ruan, Phys. Rev. D **89**, 053006 (2014).

[30] E. Gerwick, T. Plehn, and S. Schumann, Phys. Rev. Lett. **108**, 032003 (2012); E. Gerwick, T. Plehn, S. Schumann, and P. Schichtel, J. High Energy Phys. 10 (2012) 162; E. Gerwick, S. Schumann, B. Gripaios, and B. Webber, J. High Energy Phys. 04 (2013) 089.

[31] J. R. Forshaw and M. Sjodahl, J. High Energy Phys. 09 (2007) 119; J. R. Andersen, V. Del Duca, and C. D. White, J. High Energy Phys. 02 (2009) 015; B. E. Cox, J. R. Forshaw, and A. D. Pilkington, Phys. Lett. B **696**, 87 (2011).

[32] T. Becher and M. Neubert, J. High Energy Phys. 07 (2012) 108; T. Becher, M. Neubert, and L. Rothen, J. High Energy Phys. 10 (2013) 125.

[33] A. Banfi, G. P. Salam, and G. Zanderighi, J. High Energy Phys. 06 (2012) 159; A. Banfi, P. F. Monni, G. P. Salam, and G. Zanderighi, Phys. Rev. Lett. **109**, 202001 (2012).

[34] I. W. Stewart and F. J. Tackmann, Phys. Rev. D **85**, 034011 (2012); I. W. Stewart, F. J. Tackmann, J. R. Walsh, and S. Zuberi, Phys. Rev. D **89**, 054001 (2014).

[35] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas, Nucl. Phys. **B453**, 17 (1995).

[36] R. V. Harlander and W. B. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002); C. Anastasiou and K. Melnikov, Nucl. Phys. **B646**, 220 (2002); V. Ravindran, J. Smith, and W. L. van Neerven, Nucl. Phys. **B665**, 325 (2003).

[37] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, J. High Energy Phys. 12 (2013) 088.

[38] V. Ahrens, T. Becher, M. Neubert, and L. L. Yang, Phys. Rev. D **79**, 033013 (2009); V. Ahrens, T. Becher, M. Neubert, and L. L. Yang, Eur. Phys. J. C **62**, 333 (2009).

[39] A. Höcker, H. Lacker, S. Laplace, and F. Le Diberder, Eur. Phys. J. C **21**, 225 (2001).

[40] K. Cranmer, S. Kreiss, D. López-Val, and T. Plehn, Supplemental Material for A Novel Approach to Higgs Coupling Measurements, http://dx.doi.org/10.6084/m9.figshare.888607.

[41] D. López-Val, T. Plehn, and M. Rauch, J. High Energy Phys. 10 (2013) 134.

[42] For some up-to-date coupling analyses, see T. Corbett, O. J. P. Eboli, J. Gonzalez-Fraile, and M. C. Gonzalez-Garcia, arXiv:1306.0006; A. Azatov, R. Contino, and J. Galloway, J. High Energy Phys. 04 (2012) 127; A. Azatov, R. Contino, D. Del Re, J. Galloway, M. Grassi, and S. Rahatlou, J. High Energy Phys. 06 (2012) 134; J. Ellis and T. You, J. High Energy Phys. 06 (2013) 103; G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, J. High Energy Phys. 06 (2007) 045; J. R. Espinosa, C. Grojean, M. Mühlleitner, and M. Trott, J. High Energy Phys. 12 (2012) 045; S. Banerjee, S. Mukhopadhyay, and B. Mukhopadhyaya, J. High Energy Phys. 10 (2012) 062; N. Craig and S. Thomas, J. High Energy Phys. 11 (2012) 083; F. Bonnet, T. Ota, M. Rauch, and W. Winter, Phys. Rev. D **86**, 093014 (2012); A. Djouadi, Eur. Phys. J. C **73**, 2498 (2013); B. A. Dobrescu and J. D. Lykken, J. High Energy Phys. 02 (2013) 073; E. Massó and V. Sanz, Phys. Rev. D **87**, 033001 (2013); G. Belanger, B. Dumont, U. Ellwanger, J. F. Gunion, and S. Kraml, J. High Energy Phys. 02 (2013) 053; P. P. Giardino, K. Kannike, I. Masina, M. Raidal, and A. Strumia, J. High Energy Phys. 05 (2014) 046; A. Djouadi and G. Moreau, Eur. Phys. J. C **73**, 2512 (2013); D. Carmi, A. Falkowski, E. Kuflik, T. Volansky, and J. Zupan, J. High Energy Phys. 10 (2012) 196; A. Falkowski, F. Riva, and A. Urbano, J. High Energy Phys. 11 (2013) 111; P. P. Giardino, K. Kannike, I. Masina, M. Raidal, and A. Strumia, arXiv:1303.3570.

[43] O. Buchmueller *et al.*, Eur. Phys. J. C **72**, 2243 (2012); P. Bechtle, S. Heinemeyer, O. Stal, T. Stefaniak, and G. Weiglein, Eur. Phys. J. C **74**, 2711 (2014); P. Bechtle *et al.*, *Proc. Sci.*, EPS-HEP2013 (2013) 313 [arXiv:1310.3045].

[44] J. McDonald, Phys. Rev. D **50**, 3637 (1994); C. P. Burgess, M. Pospelov, and T. ter Veldhuis, Nucl. Phys. **B619**, 709 (2001); E. Pontón and L. Randall, J. High Energy Phys. 04 (2009) 080.

[45] B. Patt and F. Wilczek, arXiv:hep-ph/0605188; R. Schabinger and J. D. Wells, Phys. Rev. D **72**, 093007 (2005); R. S. Gupta and J. D. Wells, Phys. Lett. B **710**, 154 (2012).

[46] D. B. Kaplan, H. Georgi, and S. Dimopoulos, Phys. Lett. **136B**, 187 (1984); K. Agashe, R. Contino, and A. Pomarol, Nucl. Phys. **B719**, 165 (2005); G. Burdman and C. E. F. Haluch, J. High Energy Phys. 12 (2011) 038.

[47] R. Contino, Y. Nomura, and A. Pomarol, Nucl. Phys. **B671**, 148 (2003); K. Agashe, R. Contino, and A. Pomarol, Nucl. Phys. **B719**, 165 (2005); R. Contino, arXiv:1005.4269; D. Marzocca, M. Serone, and J. Shu, J. High Energy Phys. 08 (2012) 013.

[48] For useful reviews, see, e.g., J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, *The Higgs Hunter's Guide* (Addison-Wesley, Menlo Park, 1990); P. M. Ferreira, H. E. Haber, M. Maniatis, O. Nachtmann, and J. P. Silva, Int. J. Mod. Phys. A **26**, 769 (2011); G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, Phys. Rep. **516**, 1 (2012).

[49] S. L. Glashow and S. Weinberg, Phys. Rev. D **15**, 1958 (1977); A. Pich and P. Tuzón, Phys. Rev. D **80**, 091702 (2009).

[50] For useful reviews, see, e.g., J. F. Gunion and H. E. Haber, Phys. Rev. D **67**, 075019 (2003); M. S. Carena and H. E. Haber, Prog. Part. Nucl. Phys. **50**, 63 (2003); S. Heinemeyer, Int. J. Mod. Phys. A **21**, 2659 (2006); A. Djouadi, Phys. Rep. **459**, 1 (2008).

[51] For early studies, see, e.g., M. S. Carena, J. R. Espinosa, M. Quiros, and C. E. M. Wagner, Phys. Lett. B **355**, 209 (1995); H. E. Haber, R. Hempfling, and A. H. Hoang, Z. Phys. C **75**, 539 (1997); S. Heinemeyer, W. Hollik, and G. Weiglein, Eur. Phys. J. C **9**, 343 (1999); G. Degrassi, S. Heinemeyer, W. Hollik, P. Slavich, and G. Weiglein, Eur. Phys. J. C **28**, 133 (2003).

[52] S. Heinemeyer, W. Hollik, and G. Weiglein, Comput. Phys. Commun. **124**, 76 (2000); Nucl. Phys. B, Proc. Suppl. **205–206**, 283 (2010).

[53] S. Kraml, B. C. Allanach, M. Mangano, H. B. Prosper, S. Sekmen, C. Balazs, A. Barr, P. Bechtle *et al.*, Eur. Phys. J. C **72**, 1976 (2012); F. Boudjema, G. Cacciapaglia, K. Cranmer,

G. Dissertori, A. Deandrea, G. Drieu la Rochelle, B. Dumont, U. Ellwanger *et al.*, arXiv:1307.5865.

[54] S. Heinemeyer *et al.* (LHC Higgs Cross Section Working Group Collaboration), arXiv:1307.1347.

[55] I. Low, P. Schwaller, G. Shaughnessy, and C. E. M. Wagner, Phys. Rev. D **85**, 015009 (2012); E. Weihs and J. Zurita, J. High Energy Phys. 02 (2012) 041.

[56] S. Andreas, C. Arina, T. Hambye, F.-S. Ling, and M. H. G. Tytgat, Phys. Rev. D **82**, 043522 (2010); A. Djouadi, O. Lebedev, Y. Mambrini, and J. Quevillon, Phys. Lett. B **709**, 65 (2012); A. Drozd, B. Grzadkowski, and J. Wudka, J. High Energy Phys. 04 (2012) 006; J. M. Cline, K. Kainulainen, P. Scott, and C. Weniger, Phys. Rev. D **88**, 055025 (2013).

[57] V. Silveira and A. Zee, Phys. Lett. **161B**, 136 (1985); M. J. G. Veltman and F. J. Ynduráin, Nucl. Phys. **B325**, 1 (1989); H. Davoudiasl, R. Kitano, T. Li, and H. Murayama, Phys. Lett. B **609**, 117 (2005); O. Bahat-Treidel, Y. Grossman, and Y. Rozen, J. High Energy Phys. 05 (2007) 022.

[58] D. O'Connell, M. J. Ramsey-Musolf, and M. B. Wise, Phys. Rev. D **75**, 037701 (2007); V. Barger, P. Langacker, M. McCaskey, M. J. Ramsey-Musolf, and G. Shaughnessy, Phys. Rev. D **77**, 035005 (2008); G. Bélanger, B. Dumont, U. Ellwanger, J. F. Gunion, and S. Kraml, Phys. Lett. B **723**, 340 (2013); G. M. Pruna and T. Robens, Phys. Rev. D **88**, 115012 (2013).

[59] C. Englert, T. Plehn, D. Zerwas, and P. M. Zerwas, Phys. Lett. B **703**, 298 (2011); C. Englert, T. Plehn, M. Rauch, D. Zerwas, and P. M. Zerwas, Phys. Lett. B **707**, 512 (2012); B. Batell, S. Gori, and L.-T. Wang, J. High Energy Phys. 06 (2012) 172.

[60] E. Cerveró and J.-M. Gérard, Phys. Lett. B **712**, 255 (2012); W. Altmannshofer, S. Gori, and G. D. Kribs, Phys. Rev. D **86**, 115009 (2012); Y. Bai, V. Barger, L. L. Everett, and G. Shaughnessy, Phys. Rev. D **87**, 115013 (2013); A. Celis, V. Ilisie, and A. Pich, J. High Energy Phys. 07 (2013) 053; A. Celis, V. Ilisie, and A. Pich, J. High Energy Phys. 12 (2013) 095.

[61] See, e.g., A. W. El Kaffas, P. Osland, and O. M. Ogreid, Phys. Rev. D **76**, 095001 (2007); D. S. M. Alves, P. J. Fox, and N. J. Weiner, arXiv:1207.5499; C.-Y. Chen and S. Dawson, Phys. Rev. D **87**, 055016 (2013); A. Barroso, P. M. Ferreira, R. Santos, M. Sher, and J. P. Silva, arXiv:1304.5225; C.-Y. Chen, S. Dawson, and M. Sher, Phys. Rev. D **88**, 015018 (2013); N. Craig, J. Galloway, and S. Thomas, arXiv:1305.2424; N. Craig, J. A. Evans, R. Gray, C. Kilic, M. Park, S. Somalwar, and S. Thomas, J. High Energy Phys. 02 (2013) 033; A. Arhrib, P. M. Ferreira, and R. Santos, J. High Energy Phys. 03 (2014) 053.

[62] L. J. Hall, R. Rattazzi, and U. Sarid, Phys. Rev. D **50**, 7048 (1994); M. S. Carena, M. Olechowski, S. Pokorski, and C. E. M. Wagner, Nucl. Phys. **B426**, 269 (1994); M. S. Carena, D. Garcia, U. Nierste, and C. E. M. Wagner, Nucl. Phys. **B577**, 88 (2000).

[63] For complete studies of the supersymmetric paarameter space, see, e.g., P. Bechtle, S. Heinemeyer, O. Stal, T. Stefaniak, G. Weiglein, and L. Zeune, Eur. Phys. J. C **73**, 2354 (2013); G. D. Kribs, A. Martin, and A. Menon, Phys. Rev. D **88**, 035025 (2013); A. Djouadi, L. Maiani, G. Moreau, A. Polosa, J. Quevillon, and V. Riquer, Eur. Phys. J. C **73**, 2650 (2013); S. Henrot-Versillé, R. Lafaye, T. Plehn, M. Rauch, D. Zerwas, S. Plaszczynski, B. R. d'Orfeuil, and M. Spinelli, Phys. Rev. D **89**, 055017 (2014).

[64] M. Carena, S. Heinemeyer, O. Stal, C. E. M. Wagner, and G. Weiglein, Eur. Phys. J. C **73**, 2552 (2013).