On the reinterpretation of non-resonant searches for Higgs boson pairs

Alexandra Carvalho, Florian Goertz, Ken Mimasu, Maxime Gouzevitch and Anamika Aggarwal

Abstract: The detection of production of a pair of Higgs bosons before the end of LHC operation would be a clear evidence of New Physics (NP). As searches for non-resonant production of Higgs pairs are being designed it is of particular importance to be able to conveniently present current experimental results in terms of limits in the most ‘model-independent’ fashion possible. To this end, in this article we provide an analytic parametrization of the differential Higgs-pair production at the LHC in the effective field theory (EFT) extension of the SM. It results from a fit to the theory prediction for the $gg \to hh$ cross section at the 13 TeV LHC. Subsequently the resulting formula is used for a reweighing technique that allows to recast exclusion bounds from ATLAS and CMS $hh \to \gamma\gamma b\bar{b}$ searches to any point of the considered EFT parameter space. We demonstrate with a fast simulation of the LHC detectors that with this approach it is possible to cover the continuous EFT parameter space, taking correctly into account the efficiencies of signal selections, without the necessity of rerunning a large number of full detector simulations. Finally, the resulting exclusion bounds are confronted with several explicit models, such as setups with additional scalars, including 2HDM, vector-like fermions, and minimal composite Higgs models, which are mapped to the EFT.

Keywords: Beyond Standard Model, Effective Field Theories, Higgs Physics

ArXiv ePrint: 1710.08261
1 Introduction

Examining the production of pairs of Higgs bosons $h$ is a crucial task in the long term strategy of the LHC. Analyzing this process offers a unique direct window on the Higgs potential, whose behavior lies at the heart of many unresolved questions on nature, such as the hierarchy problem, the question of vacuum stability [1–3], the feasibility of electroweak baryogenesis [4–6], and the potential dynamical trigger for electroweak symmetry breaking (EWSB) [7–9] — just to name a few. Of particular importance regarding the search for physics beyond the Standard Model (SM) are differential observables, such as the distribution of the invariant mass of the Higgs pairs $m_{hh}$ (see, e.g., [10–12]), since they are especially sensitive to the effects of new physics (NP).

As we eventually want to confront corresponding experimental results with theory, we need to assume a theoretical framework, which in our case would be the effective field theory (EFT) extension of the SM, supplementing it with higher dimensional operators. This captures the effects of any yet undiscovered physics beyond the SM, given that it is separated from the latter by a mass gap and in particular is significantly heavier than the scales probed experimentally. In that sense, the setup is ‘model-independent’, providing the most general parametrization of nature, based on what we observe at low energies. On the other hand, any known UV model (residing at scales well beyond the electroweak scale) can be mapped to such an effective Lagrangian, rendering a formulation of experimental results in terms of bounds on its parameters, an instrumental bridge between data and theory. Having at hand analytic formulae which translate between observables in a realistic collider environment and effective couplings would be of utmost importance for the experimental results to be interpreted appropriately, either by the collaborations themselves or by theorists. In the end, it furnishes the missing link between theoretical models and measurements at colliders.

In this paper, we present on the one hand such a fit of the theoretical prediction for the Higgs-pair production cross-section in gluon fusion at the LHC in terms of effective couplings, considering both the total cross-section as well as the ones differential in $m_{hh}$ and $|\cos \theta^*|$. At Leading Order (LO) the totality of the $hh$ process is described by those two variables.\footnote{The latter results can, in particular, be used to generalize the parameter-space scan of ref. [10] (where the simulations were restricted to sub-regions of the parameter space) to arbitrary combinations of EFT couplings.} Beyond that, we will employ the formula to reinterpret exclusion bounds, derived assuming specific benchmark points in parameter space with given kinematics, for points with modified kinematic distributions. Such a reweighting procedure is frequently employed by experimental collaborations in various contexts as well as explored by theorists, see e.g. [13, 14]. In fact, different distributions in variables, such as $m_{hh}$, would lead to different efficiencies of selection cuts, which can be accounted for, knowing the distributions in terms of effective couplings. Without such an analytic knowledge, a full Monte-Carlo (MC) simulation would be required for each parameter-space point, which is not feasible for a dense scan of the parameters. Thus our analysis fills this gap, employing publicly available information.

We use the interpreted results to confront explicit models, mapped to an effective Lagrangian, efficiently with constraints from the LHC. To perform this exercise, we col-
lect from the literature setups with additional scalars, including two-Higgs-doublet models (2HDM), with vector-like fermions, and minimal composite Higgs models and provide a collection of relations between their explicit parameters and effective couplings after EWSB.

Finally, we show how, in cases where the experimental analysis does not allow for straightforward reinterpretation, experimental limits obtained for shape benchmarks as representatives of clusters of similar kinematic distributions can be used to derive bounds for the full parameter space, applying the efficiencies from the closest benchmarks. This allows to reinterpret also results derived via multivariate analyses, employing for example neural networks or boosted decision trees.

## 2 Setup

The theory, on which our analysis is based, is the EFT extension of the SM with $D = 6$ operators $O_i$ with coefficients $\sim c_i$, suppressed by the NP scale $\Lambda^2$ (assuming for the moment a linearly realized SM gauge symmetry).

The terms relevant for the production of Higgs pairs in gluon fusion to leading order (LO) are contained in the effective Lagrangian[^1] 

$$L = L_{SM} + \frac{c_H}{2\Lambda^2} (\partial^\mu |H|^2)^2 - \frac{c_6}{\Lambda^2} \lambda_{SM} |H|^6 - \left( \frac{c_t}{\Lambda^2} y_t |H|^2 \tilde{Q}L HcR + h.c. \right) + \frac{\alpha_s c_g}{4\pi\Lambda^2} |H|^2 G^{\mu\nu} G_\mu^{\nu}. \tag{2.1}$$

Here, $L_{SM}$ is the SM Lagrangian, $\lambda_{SM} = m_h^2/2v$ is the SM value of the tri-linear Higgs self coupling, with $m_h \approx 125$ GeV the mass of the Higgs boson, $y_t$ is the top-quark Yukawa coupling, and $\alpha_s$ is the strong coupling constant[^2].

After EWSB (in unitary gauge) $H \to 1/\sqrt{2} (0, v + h)^T$, where $v/\sqrt{2} \equiv \langle |H| \rangle \approx 174$ GeV is the vacuum expectation value of the Higgs field, the terms relevant for the $gg \to hh$ process become [10–12]

$$L_h = \frac{1}{2} \partial_\mu h \partial^\mu h - \frac{1}{2} m_h^2 h^2 - \kappa_\lambda \lambda_{SM} v h^3 - \frac{m_t}{v} \left( v + \kappa_t h + \frac{c_2}{v} h h \right) (\tilde{t}L tR + h.c.) + \frac{1}{4\pi v} \left( c_g h - \frac{c_2 g}{2v} h h \right) G^{\mu\nu} G_\mu^{\nu}. \tag{2.2}$$

In fact, the operators in (2.1) modify the trilinear Higgs self coupling and the top yukawa coupling, parametrized via multiplicative factors $\kappa_\lambda$ and $\kappa_t$, where in the SM $\kappa_\lambda = \kappa_t = 1$. Beyond that, they induce additional contact interactions between two Higgs bosons and two fermions and between a gluon pair and one or two Higgs bosons, parametrized by $c_2, c_g,$ and $c_2 g$, which vanish in the SM, i.e., $c_2 = c_g = c_2 g = 0$. The relations between these effective couplings and the coefficients of the operators in eq. (2.1) can be obtained straightforwardly (see, e.g., [12]) and imply the constraint $c_g = -c_2 g$. This correlation is however broken in the non-linear realization of EWSB and is not assumed a priori in this work, where we treat all couplings in eq. (2.2) as free. In the next section, we provide a parametrization

[^1]: The validity of this most general parametrization of heavy NP has been scrutinized in [15].

[^2]: We neglect changes in the bottom Yukawa coupling, as it is already constrained in a way that it plays at most a subleading role in Higgs pair production (see, e.g., [16]).
of the Higgs-pair production cross-sections in terms of the effective couplings entering the effective Lagrangian (2.2).

We now describe our fit to the cross-section for Higgs pair production in gluon fusion at the 13 TeV LHC in the EFT defined above.\(^4\) We consider the total cross-section but as well the ones differential in \(m_{hh}\) and \(|\cos \theta^*|\), with \(\theta^*\) being the polar angle of either of the Higgs bosons with respect to the beam axis (featuring the same \(|\cos \theta^*|\) at parton level).

Examining the different contributions to the full amplitude, it is easy to see that the \(gg \to hh\) cross section \(\sigma_{hh} \equiv \sigma(gg \to hh)\) (normalized to its SM value \(\sigma_{hh}^{SM}\)) can be parametrized to LO in terms of the parameters in (2.2) as \(^{10}\)

\[
R_{hh} \equiv \frac{\sigma_{hh}}{\sigma_{hh}^{SM}} = \text{Poly}(\vec{A}) = A_1 \kappa_t^4 + A_2 c_2^2 + (A_3 \kappa_t^2 + A_4 c_2^2) \kappa_\lambda^2 + A_5 c_2 g + (A_6 c_2 + A_7 \kappa_t \kappa_\lambda) \kappa_t^2
+ (A_8 \kappa_t \kappa_\lambda + A_9 c_g \kappa_\lambda) c_2 + A_{10} c_2 c_2 g + (A_{11} c_g \kappa_\lambda + A_{12} c_2 g) \kappa_t^2
+ (A_{13} \kappa_\lambda c_g + A_{14} c_2 g) \kappa_t \kappa_\lambda + A_{15} c_g c_2 g \kappa_\lambda. \tag{2.3}
\]

In ref. \(^{10}\), the parameters \(A_i\), that form a vector \(\vec{A} \equiv (A_i)\) of coefficients of the polynomial \(\text{Poly}(\vec{A})\), were extracted from simulations to describe the total cross section. In this paper, we push the logic further to describe the differential cross section. We slice the kinematic space into bins with sufficient granularity. In each bin, eq. (2.3) holds, but the coefficients become bin dependent, \(A_i \to A_i^j\), following the matrix-element (ME) integration and the PDF evolution. We define the cross section in a bin \(j\) by \(\sigma_{hh,j} \equiv \sigma_{hh} \times \text{Frac}_j\), where \(\text{Frac}_j\) is the fraction of events contained in bin \(j\). The differential \(hh\) production ratio now becomes

\[
R_{hh}^j \equiv \frac{\sigma_{hh}}{\sigma_{hh}^{SM}} \frac{\text{Frac}_j}{\text{Frac}_j^{SM}} = \text{Poly}(\vec{A}^j). \tag{2.4}
\]

We finally extract the differential coefficients by performing a bin-by-bin fit, extending the procedure of our HXSWG note \(^{17}\), see appendix B for the numerical results.

3 Recast of ATLAS and CMS measurements

In this section, we establish and validate the reweighting method and estimate the present limits on the Higgs couplings in a multi-dimensional space. The constraints are obtained by a recast of preliminary results from the ATLAS and CMS collaborations \(^{18, 19}\) in the \(hh \to \gamma\gamma b\bar{b}\) channel, that refer, respectively, to samples of 3.2 fb\(^{-1}\) and 2.7 fb\(^{-1}\) of 13 TeV data. Results with a slightly better sensitivity were published employing 8 TeV data \(^{20–22}\), but we decided to use the preliminary results since they are obtained at the same center-of-mass energy that will be used for the next generation of LHC results. Between the publications of these and the release of this paper, other searches were performed by both collaborations in several channels \(^{20, 23, 24, 47}\) and with more luminosity, where the CMS analyses also include results for beyond the SM (BSM) scans in \(\kappa_\lambda\) and \(\kappa_t\), and we will in fact use more recent analyses to derive constraints in section 5.\(^5\)

\(^4\)We do not consider other production mechanisms in this work, such as vector-boson fusion, which is sub-dominant in the bulk of the EFT parameter space.

\(^5\)This section was updated for v2 of this paper.
The choice of the $hh \to \gamma \gamma b\bar{b}$ channel results from the fact that this final state is easy to simulate and to reconstruct using a parametric model of the ATLAS and CMS detectors. Moreover, both analyses were performed with a sequential application of selections (in contrast to a multivariate analysis that is a kind of standard now), making them easy to recast. It is known that this final state is less constrained in regions of the parameter space compared to other final states \cite{20}. Nevertheless, since our goal is the proof of principle of the recast technique, we are not hunting for the latest and best limit. In the case of more sophisticated (multivariate) analyses, a reinterpretation of the results is more subtle, however still possible if samples exist for all relevant kinematic configurations, as provided in \cite{10} (see discussion at the end of section 5).

3.1 Signal simulation and reweighing

We employ eq. (2.4) to reweight a set of base events to different points in the parameter space. As base events, we use the 12 benchmarks defined in \cite{10, 17, 25} and for each of them we simulate 100,000 events. Due to the construction of the benchmarks, the resulting ensemble $\mathcal{S}_{\text{BSM}}$ of 1,200,000 events populates well the EFT space we are probing. Also, 100,000 SM-like events are generated for comparison but kept out of $\mathcal{S}_{\text{BSM}}$.

The hard scattering events are generated at LO accuracy using the model described in section 2. Following ref. \cite{17} we normalize the events to a total cross section with a k-factor applied to include corrections up to NNLO (augmented with NNLL re-summation) in QCD ($\sigma_{\text{SM}}^{hh} = 33.5 \pm 1.4 \text{fb}$ \cite{25}). Note that the normalization only becomes relevant to predict the absolute number of events, as used in section 4. In fact, up to NNLO, the k-factors are found to be mostly flat in kinematic variables and as a function of the EFT parameters (in the bulk of the parameter space) \cite{26} and this suggests that the distributions are well described by performing the reweighing with LO information and employing flat k-factors to correct for QCD corrections at the end is a reasonable approximation. Furthermore, we used the NLO set of the PDF4LHC parton densities \cite{27–30} as well as the parton shower and hadronization infrastructure of the PYTHIA 8 package \cite{31}. Finally, we simulate the response of an LHC-like detector by using DELPHES 3 \cite{32}.

To reweight simulated events from $\mathcal{S}_{\text{BSM}}$ to any other point in parameter space (such as the SM point, which we use for validation), we apply an event-by-event weight

$$W_i = \frac{R_{hh}(m_{hh}^i, \cos \theta_*^i)}{N_i} \frac{\sigma_{hh}^{\text{SM}} \text{Frac}^j_{\text{SM}}}{\sigma_{hh} C_{\text{norm}}}$$

(3.1)

where $m_{hh}^i$ and $\cos \theta_*^i$ refer to the ME level variables corresponding to event $i$, and $N_i$ is the number of events in bin $j$ that contains $m_{hh}^i$ and $\cos \theta_*^i$ in the reweighting sample $\mathcal{S}_{\text{BSM}}$. Along the same lines, $R_{hh}(m_{hh}^i, \cos \theta_*^i)$ is given by $R_{hh}^j$ (eq. (2.4)) for this bin. It is our choice to normalize the weights in such a way that the resulting BSM distribution is normalized to one if no cut is applied and to the signal efficiency if any selection is applied. In this manner, we decouple the shape information from the total cross-section information.

The normalization coefficient $C_{\text{norm}}$ is obtained as

$$C_{\text{norm}} = \sum_{j} R_{hh}^j \frac{\sigma_{hh}^{\text{SM}} \text{Frac}^j_{\text{SM}}}{\sigma_{hh} C_{\text{norm}}}$$

(3.2)
and is used to unitarize the reweighed shape. In the ideal case \( C_{\text{norm}} = 1 \). Nevertheless, fluctuations in the samples used to derive the weights can induce a departure from unitarity. For the SM shape, \( C_{\text{norm}} = 1 \), since the bin-by-bin fits are performed fixing the SM point, and statistical errors occur only for BSM points where for some samples there are bins with inevitably low statistics (for example in the high \( m_{hh} \) bins of threshold-like EFT points). After the renormalization procedure, we observe a good agreement within the statistical uncertainties between MC simulation and reweighted samples, as can be seen in the next section. The values of the weights in eq. (3.1) are implemented in the Rosetta function “AnalyticalReweighter.py”, see section 6, which can be used to perform the reweighting in practice.

### 3.2 The analyses

We consider the number of events \( N_{\text{obs}} \) observed by each experiment in a signal region and \( B \) the number of the expected background events. In the ATLAS documents, both values are directly provided in the text, while in the CMS document, they are extracted from figures imposing a \( M_{\gamma\gamma} \) window identical to the ATLAS one. Both experiments found that the observed events are statistically compatibles with the expected SM backgrounds. Therefore, we can use this observation to derive exclusion limits on the maximal allowed number of (expected) events in a NP scenario \( N_{\text{BSM}} \) that could be compatible at 95% CL with the observed data. We build a Poissonian statistical likelihood

\[
L = e^{- (B + S)} \frac{(B + S)^{N_{\text{obs}}}}{N_{\text{obs}}!},
\]

where \( S \) is the number of Signal events. Bayesian hypothesis with a uniform prior on \( S \) allows to invert the likelihood to obtain a posterior as a function of \( S \), normalized to unity between \( S = [0, \infty] \). The likelihoods are shown in figure 1. We integrate then the likelihood from 0 to 95% to obtain the 2 standard deviation confidence level value for \( N_{\text{BSM}} \). The number of observed, expected and excluded events are provided in table 1. In this table we also include the number of events that would be excluded at \( L=100 \text{fb}^{-1} \), projected assuming \( N_{\text{obs}}(100 \text{fb}^{-1}) = B(100 \text{fb}^{-1}) = L' \), where \( L' \) is the data luminosity currently used.

To translate the limit on \( N_{\text{BSM}} \) into the limit on the BSM cross-section \( \sigma_{\text{BSM}} \) we need to know the selection efficiency \( \epsilon \) within the signal box and the luminosity \( L \): \( \sigma_{\text{BSM}} = \)
Table 1. Upper limits on cross section times luminosity in the analysis window, \(N_{BSM}\), from the recast of ATLAS and CMS results.

| \(N_{BSM}/(L_\epsilon)\) | \(N_{BSM}/(100\text{fb}^{-1})\) | \(N_{BSM}/(100\text{fb}^{-1})\) | \(N_{BSM}/(L_\epsilon)\) | \(N_{BSM}/(100\text{fb}^{-1})\) |
|---------------------------|-------------------------------|-------------------------------|---------------------------|-------------------------------|
| 0.8                       | 5.45                          | 12                            | 0.8                       | 5.45                          |

For both analyses, we can extract (or estimate) \(\epsilon_{SM}\), i.e. the efficiency for the SM-like \(hh\) production, from the information provided in the respective papers. What is not known a priori is the efficiency of a given BSM model with the associated choice of parameters. To estimate this, we employ a mock-up of the ATLAS and CMS analyses with DELPHES 3. The resulting efficiency \(\epsilon_D\) is then rescaled (if needed) by a fudge factor \(f_{SM} = \epsilon_{SM}/\epsilon_{SM,D}\) to get an estimate of the real efficiency \(\epsilon_{BSM} \approx f_{SM} \cdot \epsilon_{D}\). This factor takes into account the limited information we benefit from to emulate the analysis and the approximations made in DELPHES 3 to describe the detector response, such as the efficiency of detecting photons or tagging b-jets. Note that this step could be avoided if the experimental collaborations would make public the full information (including the generation level \(m_{hh}\) and \(|\cos \theta^*|\) variables) on the events that pass the signal selection.

In table 2 we summarize the set of selections of both ATLAS and CMS searches. It is essential to notice that the same DELPHES 3 card was used for both experiments. This approximation is valid for photons in the energy range considered in \(hh\) production since both experiments are more than 90% efficient to tag them. The difference in fiducial acceptance is emulated at the level of selections. For jets, both experiments use the anti-\(k_T\) algorithm [33] with the radius parameter 0.4. Their acceptance and calibrations are similar to those energies. The significant difference comes from the b-tagging working point (WP). Since both experiments require both jets to be b-tagged, we just rescale the typical efficiency of the WP used in DELPHES 3 within our \(p_T\) and \(\eta\) range (63% — see ref. [34]) to the efficiency of the WP declared in the public notes, i.e. 85% for ATLAS and 78% for CMS. Finally, for the projections to 100\text{fb}^{-1} given below, we assume the observed number of events to match the expected and no improvements in the analysis, which is a very conservative assumption.

After the reweighting, according to the b-tagging properties and the selections of table 2, we find the efficiencies shown in table 3. It appears that for the ATLAS analysis, our efficiencies have to be rescaled by \(f_{SM} = 1.41\). Besides the usual suspects described above to explain the size of the factor, we suspect that the photon efficiency is slightly better in the real ATLAS analysis than in the DELPHES 3 card, but the exact amount is hard to guess from the details given in the conference note.

In the CMS case, the efficiency of the 2 b-tag and 1 b-tag categories are only provided together, and no explicit information on how it is distributed between them is given. One may notice that a similar categorization was used at 8\text{ TeV} where the events appeared to be evenly split between the two categories, incidentally [22]. We apply the same assumption to the 13\text{ TeV} analysis and split the efficiencies evenly between the two categories (\(\epsilon_{SM,CMS} \approx 10%\) per category).
### Table 2. Selections used in the recast analyses.

| variable                        | ATLAS | CMS |
|---------------------------------|-------|-----|
| $|\eta_{\text{max}}|$            | 2.37  | 2.5 |
| Rejected fiducial region in $|\eta|$ | [1.37, 1.52] | [1.44, 1.57] |
| Leading photon ($p_{T,\text{min}}$) | —     | 30  |
| Subleading photon ($p_{T,\text{min}}$) | —     | 20  |
| Leading photon ($p_{T,\text{min}}$)/$M_{\gamma\gamma}$ | 0.35  | 1/3 |
| Trailing photon ($p_{T,\text{min}}$)/$M_{\gamma\gamma}$ | 1/4   |     |
| $\Delta R$ with any jet         | $>0.4$|     |
| $|\eta_{\text{max},b}|$          | 2.5   | 2.4 |
| Leading b-jet $p_{T,\text{min}}$ (GeV) | 55    | 25  |
| Trailing b-jet $p_{T,\text{min}}$ (GeV) | 35    | 25  |
| $M_{\gamma\gamma}$ window (GeV) | [122, 128] |     |
| $M_{b\bar{b}}$ window (GeV)     | [95,135] | [80,200] |
| $M_{\gamma\gamma b\bar{b}}$ minimum (GeV) | —     | 350 |

### Table 3. Results from the ATLAS and CMS-like recast. See text for details.

|                | ATLAS | CMS |
|----------------|-------|-----|
| $\epsilon_{\text{SM}}$ | 7.1%  | 10% |
| $f_{\text{SM}}$     | 1.41  |     |
| $\epsilon_{\text{SM}}$ | 10.8% | $\approx$ 10% |
| $f_{\text{SM}}$     | 1     |     |

#### 3.3 Validation of the analytical reweighing

In figure 2 we show the comparison of the actual simulation and the reweighting for an SM-like signal for 3 main reconstructed kinematic variables describing the $hh$ system: $m_X \equiv m_{\gamma\gamma b\bar{b}} - m_{b\bar{b}} - m_{\gamma\gamma} + 250$ GeV, $\cos \theta^*_{\gamma\gamma b\bar{b}}$, and $p_{T,\gamma\gamma b\bar{b}}$. The first variable was shown to be the best estimate of $m_{hh}$ in ref. [19]. The statistical errors per bin are shown only for the plain simulation since the shapes derived by reweighting are constructed from 12 times as many bare events and thus expected to have less statistical fluctuations. In general, we find an excellent agreement between the distributions. Comparisons for different benchmarks provide a similar picture and are given in appendix B. Moreover, in figure 3, we show reweighted distributions in $m_X$, after ATLAS-like selections, for BSM points not within the set of (simulated) benchmarks. Beyond points featuring non-vanishing contact interactions ($c_g = -c_2g = 1$ and $c_2 = 1$), we also consider the maximum box-triangle interference ($\kappa_\lambda = 2.4$), thereby exploring three different kinematic regimes. In all the cases, we see smooth shapes.

Finally, the efficiencies of different benchmark points obtained from the MC simulation are given in table 4 ($\epsilon_{\text{MC}}$), together with the reweighted efficiencies as obtained after applying the renormalization via $C_{\text{norm}}$. The agreement of both efficiencies is remarkable, especially in the points where EFT is linear (where the parameter space scan used for the fit is denser, see table 9).
Figure 2. Reconstructed variables after ATLAS-like selection for the di-Higgs system. The histograms display the signal efficiency times 100,000 events.

Figure 3. Reconstructed variables after ATLAS-like selection for the \( m_X \) variable for three BSM points with different kinematic properties. The histograms display the signal efficiency times 100,000 events.

4 Model dependent interpretations

We illustrate the method described in section 3, reinterpreting EFT bounds and applying them to concrete NP setups. We consider the explicit models collected in table 5.

In table 6 we provide the relations between the effective couplings for the various models under consideration, arising after integrating out the NP at tree level, where we express dependent couplings in terms of independent ones, the latter given in the second column. The relation with the fundamental parameters of the models is given in table 7, which contains equivalent information. Note that the fact that we integrate out NP at tree level might appear problematic, given that the effective couplings in table 6 enter Higgs-pair production only at one-loop and, in principle, potential loop induced contact interactions with gluons (\( c_g, c_g^2 \)) could enter the process at tree-level, thus lifting a loop-suppression in integrating out NP. However, since the scalars considered in models 1-7 are assumed to be color neutral, no such effective interactions are induced at the one-loop level. The same is valid for integrating out vector-like leptons (model 9). The vector-like quarks (model 8), on the other hand, do not couple to the Higgs boson without involving further SM fermions, thus also not generating Higgs-gluon contact interactions at one-loop.
| $N$ | $\kappa_\lambda$ | $\kappa_t$ | $c_2$ | $c_g$ | $c_{2g}$ | $C_{\text{norm}}$ | ATLAS-like $\epsilon_{\text{MC}}$ | ATLAS-like $\epsilon_{\text{rew}}$ | CMS-like $\epsilon_{\text{MC}}$ | CMS-like $\epsilon_{\text{rew}}$ |
|-----|-----------------|-----------|------|------|--------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 0   | 1.0             | 1.0       | 0.0  | 0.0  | 0.0    | 1.0             | 10.0            | 10.0            | 11.46           | 11.47           |
| 1   | 7.5             | 1.0       | -1.0 | 0.0  | 0.0    | 0.94            | 10.6            | 10.6            | 10.83           | 10.79           |
| 2   | 1.0             | 1.0       | 0.5  | -0.8 | 0.6    | 0.71            | 10.3            | 10.4            | 11.46           | 11.26           |
| 3   | 1.0             | 1.5       | -1.5 | 0.0  | -0.8   | 0.96            | 9.93            | 10.5            | 10.89           | 11.68           |
| 4   | -3.5            | 1.0       | -3.0 | 0.0  | 0.0    | 0.98            | 9.42            | 9.37            | 9.89            | 9.72            |
| 5   | 1.0             | 1.0       | 0.0  | 0.8  | -1.0   | 0.88            | 11.6            | 11.5            | 12.47           | 12.50           |
| 6   | 2.4             | 1.0       | 0.0  | 0.2  | -0.2   | 0.96            | 8.47            | 8.73            | 7.27            | 7.74            |
| 7   | 5.0             | 1.0       | 0.0  | 0.2  | -0.2   | 1.01            | 6.24            | 6.17            | 3.30            | 3.23            |
| 8   | 15.0            | 1.0       | 0.0  | -1.0 | 1.0    | 0.92            | 10.2            | 10.3            | 10.64           | 10.74           |
| 9   | 1.0             | 1.0       | 1.0  | -0.6 | 0.6    | 0.86            | 11.6            | 11.6            | 12.92           | 12.79           |
| 10  | 10.0            | 1.5       | -1.0 | 0.0  | 0.0    | 1.01            | 6.67            | 6.60            | 3.48            | 3.48            |
| 11  | 2.4             | 1.0       | 0.0  | 1.0  | -1.0   | 0.95            | 8.16            | 8.26            | 6.72            | 6.81            |
| 12  | 15.0            | 1.0       | 1.0  | 0.0  | 0.0    | 1.0             | 7.80            | 7.52            | 6.23            | 6.17            |

**Table 4.** The renormalization factor ($C_{\text{norm}}$) for benchmark point $N$ as well as MC and reweighted sample efficiencies.

| Model | NP integrated out                                      | Ref.   |
|-------|--------------------------------------------------------|--------|
| 1     | real scalar singlet with explicit $Z_2$ breaking       | [35, 36]|
| 2     | real scalar singlet with spontaneous $Z_2$ breaking    | [35]   |
| 3     | real scalar triplet                                    | [35, 36]|
| 4     | complex scalar triplet                                 | [35, 36]|
| 5     | quartet scalar with $Y = 1/2$                          | [35, 36]|
| 6     | quartet scalar with $Y = 3/2$                          | [35, 36]|
| 7     | 2HDM (addtl. scalars heavy + $Z_2$)                     | [37]   |
| 8     | vector-like quark: $T$ (singlet top partner)            | [38]   |
| 9     | vector-like lepton: $E$ (flavor universal singlet)      | [39]   |
| 10    | MCHM$_5$                                               | [7, 8, 40, 41]|
| 11    | MCHM$_4$                                               | [7, 8, 40, 41]|

**Table 5.** Explicit models considered in this work.
| Model | Free Parameters | \( \kappa_\lambda \) | \( \kappa_t \) | \( c_2 \) |
|-------|----------------|------------------|------|------|
| 1     | \( \kappa_t, \kappa_\lambda \) | \( \kappa_\lambda \) | \( \kappa_t \) | \( \kappa_t - 1 \) |
| 2     | \( \kappa_t \) | \( 3\kappa_t - 2 \) | \( \kappa_t \) | \( \kappa_t - 1 \) |
| 3     | \( \kappa_\lambda, c_2 \) | \( \kappa_\lambda \) | 1    | \( c_2 \) |
| 4     | \( \kappa_t, \kappa_\lambda \) | \( \kappa_\lambda \) | \( \kappa_t \) | \( 2\kappa_t - 2 \) |
| 5     | \( \kappa_\lambda \) | \( \kappa_\lambda \) | 1    | 0    |
| 6     | \( \kappa_\lambda \) | \( \kappa_\lambda \) | 1    | 0    |
| 7     | \( \kappa_t, \kappa_\lambda \) | \( \kappa_\lambda \) | \( \kappa_t \) | \( 3/2(\kappa_t - 1) \) |
| 8     | \( \kappa_t \) | 1 | \( \kappa_t \) | \( 3/2(\kappa_t - 1) \) |
| 9     | \( \kappa_t \) | \( \kappa_t \) | \( \kappa_t \) | 0 |
| 10    | \( \kappa_t \) | \( \kappa_t \) | \( \kappa_t \) | \( \kappa_t + \sqrt{\kappa_t^2 + 8}/4 - 1 \) |
| 11    | \( \kappa_t \) | \( \kappa_t \) | \( \kappa_t \) | \( (\kappa_t^2 - 1)/2 \) |

Table 6. Correlations between effective couplings in various explicit models. While the total number of free parameters is invariant, when there was freedom which parameters to treat as free, we always chose \( \kappa_t \), expressing the other couplings in terms of the latter.

The situation becomes more subtle for the composite Higgs setups (models 10 and 11). Here, the potential effect of loop-induced Higgs-gluon contact interactions due to integrating out fermionic resonances appearing in the models (that could become relevant in parts of the parameter space) cancels with additional corrections to the Yukawa couplings generated by the very same fields. Thus, considering only the Higgs-non-linearities that lead to the anomalous couplings as given in table 6 (see also table 7) leads effectively to an accurate description. In summary, the effective couplings as given in table 6 provide an appropriate description of all models at hand to leading approximation, given that the NP is heavy such that the EFT framework is valid (see, e.g., [15]).

Using the map from the model parameters to the anomalous couplings in tables 6 and 7, we determine the sensitivity of the LHC analyses to the independent model parameters (which differs between the various models due to the different correlations). The reweighted differential information is used to determine the expected number of events that populate the signal region in each analysis and a parameter point is excluded if it exceeds the upper limits derived in section 3. Given the low amount of data in the analyses considered, together with the inherent difficulty of the channel, we do not expect to provide competitive limits on the models at the current stage, but rather present these results as an academic exercise to demonstrate the application of our tool in the future.\(^6\)

\(^6\)Note that at the current stage, we do not include the effect of the anomalous couplings on the Higgs decays (\( \kappa_t \) is entering \( h \to \gamma\gamma \) at the one-loop level and has an indirect effect on the branchings by changing \( h \to gg \)) as well as further operators generated in the models at hand that can modify the Higgs boson branching ratios (in particular \( h \to bb \)). As the Higgs boson is a narrow particle, those effects do not change the \( hh \) kinematics, and the effect is secondary with respect to the scope of this paper. However, in principle it should be included and we provide the corresponding expressions in appendix A.
We begin with the recast limits obtained for models with additional scalars in the $\kappa_\lambda - \kappa_t$ plane, given in figure 4. In the left plot, the cases of a real scalar singlet with explicit $Z_2$ breaking, a complex triplet scalar, and the 2HDM are compared with the scenario where only these two effective couplings are allowed to vary, and the others are set to zero.\(^7\) Note that in each of the models considered, only $\kappa_t$ between 0 and 1 are permitted.\(^8\) In these examples, the presence of $c_2$ (for $\kappa_t \neq 1$) leads to mild variations in the final sensitivity in this plane. The impact of $c_2$ on the bounds is further quantified in the right plot, where we show the effect of setting this coefficient to $\pm 1$, which shifts the contours left or right — in agreement with the tendencies observable in the left plot.

Moving to the models which can be described with one parameter (combination), we show in figure 5 the fiducial cross-section as a function of the free (fundamental) parameter for the MCHM (left), real singlet scalar with spontaneous $Z_2$ breaking (center), and singlet

\(^7\) That has been considered by the CMS collaboration.

\(^8\) While a large depletion seems only viable in the presence of cancellations with other NP contributions to single Higgs boson production and decay, here we focus mostly on $hh$ production and only take into account rigorous theoretical constraints at this stage, leaving a combined phenomenological analysis for future work.

| Model | Fund. Parameters | $\kappa_\lambda$ | $\kappa_t$ | $c_2$ |
|-------|-----------------|-----------------|-------------|-------|
| 1     | $\alpha, m_2, \lambda_\alpha$ | $1 - \frac{3\alpha^2}{2} + \frac{\lambda_\alpha}{v^2} (\alpha - \frac{1}{2})$ | $1 - \frac{\lambda_\alpha}{2v^2}$ | $-\frac{\lambda_\alpha}{2v^2}$ |
| 2     | $\alpha, \beta, m_H, m_H$ | $1 - \frac{\lambda_\alpha}{2v^2}$ | $1 - \frac{\lambda_\alpha}{2v^2}$ | $-\frac{\lambda_\alpha}{2v^2}$ |
| 3     | $\beta, m_A, m_H$ | $1 - 8\lambda_\alpha + \frac{\lambda_\alpha}{\lambda_{SM}} m_H^2$ | $1 - 8\lambda_\alpha + \frac{\lambda_\alpha}{\lambda_{SM}} m_H^2$ | $-8\lambda_\alpha + \frac{\lambda_\alpha}{\lambda_{SM}} m_H^2$ |
| 4     | $\lambda_{T\ell}, M_T$ | $1$ | $1 - V_{tb} \frac{\lambda_{T\ell} v^2}{2M_T^2}$ | $-3V_{tb} \frac{\lambda_{T\ell} v^2}{4M_T^2}$ |
| 5     | $\lambda_{E\ell}, M_E$ | $1 + \frac{\lambda_{E\ell} v^2}{4M_E^2}$ | $1 + \frac{\lambda_{E\ell} v^2}{4M_E^2}$ | $0$ |
| 6     | $\xi, Z_6, m_H$ | $\frac{1-2\xi}{\sqrt{1-\xi}}$ | $\frac{1-2\xi}{\sqrt{1-\xi}}$ | $-2\xi$ |
| 7     | $\xi, Z_6, m_H$ | $\frac{1-2\xi}{\sqrt{1-\xi}}$ | $\frac{1-2\xi}{\sqrt{1-\xi}}$ | $-2\xi$ |

Table 7. Effective couplings in terms of physical parameters of various explicit models. Here, $\alpha$ is the mixing angle between the two scalars, while $\beta = \arccos(v_1/v)$ is the arccosine of the ratio of the vev of the (first) doublet and the electroweak vev $v \approx 246$ GeV, and we defined $s_x \equiv \sin x$ ($t_x \equiv \tan x$). A common mass $m_H$ is assumed for the heavy scalars, besides for $H^+$ and $A$ in models 3-6, with masses $m_{H^+}, m_A$. Beyond that, $m_2$ is the coefficient of the triple-singlet coupling and $\lambda_\alpha$ that of the bi-quadratic scalar term, while $Z_6$ multiplies $|H_1|^2 H_1 H_2$ in the 2HDM. Moreover, $M_T$ and $M_E$ are the masses of the heavy vector-like quark and lepton, respectively, and $\lambda_{T\ell}, \lambda_{E\ell}$ are the coefficients of their (Yukawa-type) couplings with the SM fermions, mediated by the Higgs. Finally, $\xi \equiv v^2/f^2$ parametrizes the composite Higgs non-linearity, with $f$ the Pseudo-Goldstone decay constant. See references given in table 5 for more details.
Figure 4. ATLAS (continuous lines) and CMS (dashed lines) 95% exclusion bounds in the $\kappa_\lambda - \kappa_t$ plane, reinterpreted in terms of different models (left) and in the pure EFT framework for different values of $c_2$ (right). See text for details.

vector-like fermion (right) models. For comparison, we also provide current exclusion limits from ATLAS and CMS as well as projections to 100 fb$^{-1}$, that will exclude the part of the parameter space corresponding to a larger fiducial cross-section.

For both MCHM scenarios, the signal efficiency is observed to be rather flat as the kinematics are close to the SM-like case. The sensitivity thus is mostly determined via a simple rescaling of the total rate. The current analyses are sensitive to values of $\xi \sim 0.9$ in the case of the MCHM$_5$, and values of $\xi \sim 0.2$ can be probed with 100 fb$^{-1}$, while no bound is obtained for the MCHM$_4$. We observe a similar pattern in the case of the singlet, where the kinematics are not significantly affected, and no bound is obtained for mixing angles $\cos \alpha \gtrsim 0.5$. We remark that the same is not true for the two-parameter models, for which the efficiencies do vary significantly within the parameter space, as we will see in more detail in section 5.

In the case of vector-like fermions, the free parameter scales with the Yukawa mixing over the mass of the new states and decoupling is exhibited with the latter approaching infinity, as expected. As shown in table 6, the vector-like quark model does not modify $\kappa_\lambda$. Combining this with the sensitivity information of figure 4, it is not surprising that (expected) limits are rather weak, stemming mostly from a non-vanishing $c_2$ and residing in the region of large couplings ($\lambda_T \gtrsim 1$) and/or small masses ($M_T \lesssim 1$ TeV). The vector-like leptons, on the other hand, do modify $\kappa_\lambda$ and also identically affect $\kappa_t$. Since $c_2$ is however unchanged, they correspond exactly to the benchmark scenario considered in the experimental analyses (with the additional constraint $\kappa_t \equiv \kappa_\lambda$), and the limits are similarly weak.

Finally, figure 6 displays the limits obtained for the models with more than one free parameter, this time in terms of the different ‘fundamental’ model parameters (and in-
including the real scalar triplet): the real singlet with explicit $Z_2$ symmetry breaking (left), the real/complex triplet (center) and the 2HDM (right). Where possible, we reduce the parameter space to two degrees of freedom: in the triplet models we set the heavy scalar masses to be equal while for the real singlet we can define an effective trilinear coupling

$$\lambda_{\text{eff}} = \lambda_\alpha - t_\alpha \frac{m_2^2}{v}.$$  \hfill (4.1)

In the 2HDM, it is not trivial to reduce the $(t_\beta, Z_6, m_H)$ parameter space. However, $Z_6$ is related to the usual alignment parameter, $\cos_{\beta - \alpha}$ and the neutral scalar masses by \cite{ref37}

$$v^2 Z_6 = -\cos_{\beta - \alpha} \sin_{\beta - \alpha} (m_H^2 - m_h^2),$$  \hfill (4.2)

which we use to fix a relation between the latter, showing predictions for two fixed values of $Z_6 = \pm 2.5$.\footnote{Fixing a relation between the alignment parameter and the heavy scalar mass (keeping $Z_6$ finite) leads to a proper decoupling behaviour, with $m_H \to \infty$.}

The current constraints on the real singlet lie in regions of large $\lambda_{\text{eff}}$ of order 5–10 and of sizable mixing with the Higgs. The bounds will, however, improve significantly with increasing luminosity. For the triplets, the analyses are sensitive to large mixings for masses of a few hundred GeV while increasing the heavy scalar mass leads to increased sensitivity, including also smaller mixing angles $\beta$. This subtle behavior is due to the fact that $m_H^2$ is proportional to the size of scalar quartic interactions (times the electroweak vev squared), such that the coupling strength increases with the scalar mass. We thus cut off the parameter space at $m_H^2 \sim 4\pi v^2 \sim 1$ TeV, in order to remain at a reliable (perturbative) behavior of the theory. Finally, in the case of the 2HDM, a sensitivity to scalar masses in the TeV region is only possible for small $\tan \beta$, which amplifies the corrections to $\kappa_t$ and $c_2$.

5 Mappping between smooth scan and shape benchmarks

5.1 The mapping approach

Most of the times the experimental searches are not easy to recast as soon as shape analysis or multivariate analysis is used. In this case we may ask ourselves if at some level...
of confidence we can use results obtained for shape benchmarks for a general study of the parameter space. In this section we test this premise on the scans presented in the last section. We verify how the same criteria used to define the shape benchmarks as representatives of clusters of similar shapes from a large parameter space scan can help to predict the closest experimental limit (provided there is a list of results for the shape benchmarks), and how this limit can approximate the “real” limit.

As a quick reminder, in ref. [10] a two sample Test Statistic (TS) was defined to order the degree of similarity between two samples. The log-likelihood ratio function of the hypothetical case in which the two samples under test share the same parent distribution is the product over the bins of the probability to observe $n_{i1}$ and $n_{i2}$ event counts in bin $i$ from the two samples $S_1$ and $S_2$ and can be written as:

$$TS = -2 \sum_{i=1}^{N_{bins}} \left[ \log(n_{i1}) + \log(n_{i2}) - 2\log\left(\frac{n_{i1} + n_{i2}}{2}\right) \right]. \quad (5.1)$$

This quantity is constructed in a manner that it is “$\chi^2$ distributed” [42, 43] and therefore can be directly used as an ordering parameter to decide between pairs of test samples which of them are the most likely to be compatible with the same parental distribution. In other words, the values $TS_{ij}$ and $TS_{kl}$ obtained respectively by testing the compatibility of samples $ij$ and $kl$ are suitable to determine if samples $S_i$ and $S_j$ are more similar to each other than are samples $S_k$ and $S_l$: this is the case if $TS_{ij} > TS_{kl}$.

One solution for constructing analyses optimal in continuous scans is, for example, the use the $TS$ as quantity to decide, given a test sample to which shape benchmark this one is most similar to and check if this prediction corresponds to the closest experimental limit. At this point we should emphasise that we are after the closest experimental limit. The real experimental limit, that would be directly derived for the test point, would be a bit different from the one from the closest shape benchmark.

This approach has multiple applications: if used by the theory side it is an approximation of reality, that can be employed to obtain a first estimate of the constraints on a
specific model without the need of expensive MC simulations or recast. This usage would hold for the case of including further BSM effects such as the inclusion of additional EFT operators (for example, the chromomagnetic term [44]) or the departure from the EFT framework towards explicit inclusion of particles in the loops (for example [45, 46]). As an example, in figure 7 we show the comparison between the constraints estimated from the limits derived via benchmarks and the explicit result of the recast described in the last section for a scan in the $\kappa_\lambda - \kappa_t$ plane for different values of $c_2$. For completeness, in figure 8 we also provide the mapping between shapes and shape benchmarks as computed via eq. (5.1). It is clear that for different values of $c_2$ the population of shape benchmarks changes in the $\kappa_\lambda - \kappa_t$ plane.
Figure 9. Efficiency maps for the signal region of the CMS-like analysis applied to the 2HDM benchmarks we consider. The markers superimposed correspond to the closest shape benchmark. Details are given in the text.

On the experimental side we can imagine a situation where it is advantageous for the discovery potential of a hypothetical BSM signal to construct different selections (or multivariate variables) for the different shape benchmarks. In this case it would be unclear how to extend the interpretation of the search to a smooth scan to decide which set of cuts to use for a point that is not a shape benchmark. The $TS$-test offers a solution for keeping the best search sensitivity also for the case of smooth parameter space scans.\textsuperscript{10} In particular, one can use a map of shape benchmarks that populates regions of interest in the parameter space scan of specific theories to define the analysis optimization strategy.

In figures 9 and 10 we show the benchmark maps for the cases of the two 2HDM scans and the triplet and singlet extension cases. Of course, the quality of color interpolation in the figures depends on the density of points inspected. In the background of those figures we also show the efficiencies that we obtain from the CMS-like analysis.\textsuperscript{11} As this analysis contains a cut in the $m_{\gamma\gamma b\bar{b}}$ variable the signal efficiency is more sensitive to BSM physics and we will only do this mapping exercise for the CMS-like case. The first fact to notice is that as suggested in \cite{10}, the regions belonging to the same benchmark tend to enclose fully connected regions of parameters space, that we will call islands. We also notice that within those fully connected regions the closest benchmark indeed has a signal efficiency that is the closest to the true one. This conclusion is not so precise near the boundaries between islands.

5.2 Application to the CMS results

In this section, we apply the benchmark-mapping technique to provide some constraints on the EFT parameters using the limits that the CMS collaboration set on different benchmarks. We employ the combined constraints from many channels based on a set of 36 fb$^{-1}$ of data collected at $\sqrt{s} = 13$ TeV in 2016 \cite{48}.

\textsuperscript{10}This is applied recently by the CMS collaboration, see e.g. \cite{47}.

\textsuperscript{11}Since we are considering cut-and-count examples the “experimental limits” can be well identified with the signal efficiency in the signal region.
Figure 10. Efficiency maps for the signal region of the CMS-like analysis applied to the triplet and singlet benchmarks we consider. The markers superimposed correspond to the closest shape benchmark. Details are given in the text.

The EFT limits from the 1D benchmark-mapping scans are shown in figure 11 on the right, and the $\kappa_\lambda$ scan produced from the full shape interpolation within CMS on the left. The resulting constraints are provided in table 8. Finally, a set of 2D-scans in the three main EFT parameters explored in this paper is shown in figure 12, providing again limits using the benchmark interpolation technique, together with the numbers of the benchmarks that correspond to the parameter-space points.

To understand the granular effect of the benchmarks, we can explore the constraints on $\kappa_\lambda$. For that we inspect the 2D scan in figure 12 for $c_2 = 0$ and compare the $\kappa_t = 1$ line to the 1D scan from CMS in figure 11, left panel. On the negative side, we observe that the cross-section limit is constant and always given by benchmark 12 up to $\kappa_\lambda = -1$. At the same time the full analysis shows an improvement by $\approx 20\%$ from $\kappa_\lambda = -20$ to $\kappa_\lambda = -1$. This effect can be easily understood since the shape is slowly drifting towards higher values of $m_{hh}$ with $\kappa_\lambda$ increasing. We see a similar effect on the positive side: it is described by benchmark 7 down to $\kappa_\lambda = 4$, while the full analysis features a variation of 10-20% in the limit. This results in a small difference of $\delta \kappa_\lambda \approx 1$ in bounds predicted by the benchmarks and by the full analysis. This difference seems perfectly acceptable, and the benchmarks give in fact a good estimate of the constraints when no full scan is available.

Relying on the robustness seen in the $\kappa_\lambda$ scan, we derived our bounds for the parameter $c_2$, assuming $\kappa_\lambda = \kappa_t = 1$, and for $\kappa_t$, where $\kappa_\lambda = 1, c_2 = 0$. As reported in table 8, they are excluded below $c_2 = -1.35$ and above $c_2 = 1.45$ as well as below $\kappa_t = -2.05$ and above $\kappa_t = 2.3$.\footnote{Note that due to the granularity of the 2D scan, small deviations can arise in the limits inferred from figure 12.}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Parameter & Lower Limit & Upper Limit & Benchmark \hline
\kappa_\lambda & \kappa_t & c_2 & \hline
\end{tabular}
\caption{Summary of limits inferred from the benchmarks.}
\end{table}
Figure 11. Left: scan in $\kappa_\lambda$ based on full event-by-event shape interpolation from the CMS analysis [48]. Right: scans in $\kappa_\lambda/10$, $\kappa_t$, $c_2$, using the CMS limits on benchmarks from the same analysis. Excluded regions at 95% CL correspond to dashed lines. Note that $\kappa_\lambda$ is divided by 10 to appear on the same scale as the other parameters.

Figure 12. 2D scans in the $\kappa_\lambda \times c_2$ phase space for 3 values of $c_2 = -1, 0, 1$ (from left to right), see text for details.

| EFT parameter | Method       | allowed interval at 95% CL |
|---------------|--------------|----------------------------|
|               | observed     | expected                   |
| $\kappa_\lambda$ | Benchmarks  | -11–20                     | -6–12                       |
|                | CMS          | -12–19                     | -7.1–14                     |
| $\kappa_t$    | Benchmarks  | -2.05–2.3                  |                             |
| $c_2$         | Benchmarks  | -1.35–1.45                 |                             |

Table 8. The observed and expected 95% CL intervals on EFT parameters from benchmark interpolation and from CMS public results when available [48].
6 Implementation in Rosetta

Although the numerical values of the coefficients that can be used to construct the signal predictions are given in appendix C, for convenience, the input and algorithm to perform sample reweighing are implemented as a module in the Rosetta package [49]. Rosetta provides a framework for EFT basis translations along with a suite of modules ranging from the calculation of Higgs branching fractions via an interface to eHDECAY [50] to the compatibility of a given parameter space point to electroweak precision data and Higgs signal strength measurements. All modules can be exploited in a basis-independent way once a basis definition has been implemented in the package. The package may be found at http://rosetta.hepforge.org/.

The results of this paper are implemented as an extension of the existing dihiggs module [51]. In the latter we had implemented the mapping of parameter space points onto shape benchmarks, as explained in the last section, to be able to avoid expensive Monte Carlo simulations. From this implementation it is straightforward for the user to extend it to the case where the calculation of event weights is of interest. After downloading the package, one can invoke the dihiggs interface by calling the command line executable

```bash
>> bin/rosetta dihiggs [OPTIONS]
```

The package is designed to receive SLHA formatted parameter cards specifying the coefficients of an existing basis implementation. The required translations are then performed to obtain the anomalous Higgs couplings parameters relevant for pair production and decay. Alternatively, one can directly specify the values of the anomalous couplings via the self-explanatory \( -k_1, -k_t, -c_2, -c_g \) and \( -c_{2g} \) optional arguments. The input parameter card has therefore been made optional, to be specified by the \(-\text{param\_card}\) option.

Along with computing the inclusive Higgs pair production cross section, the module will also compute the correct Higgs branching fractions for that point in parameter space using either eHDECAY or an internal interpolating function. Using the results of this paper, we have also added the functionality to return the closest benchmark according to the test defined in Formula (5.1). Finally, the computation of the higgs-pair production process is promoted to differential level, via an internal function. Inside interfaces/dihiggs/AnalyticalReweigher.py the user can find a function `weight(variables)`, where a matrix of weights is calculated from vectors containing the generation level variables \( m_{hh} \) and \( \cos \theta^* \) according to Formula (3.1). Support for implementing the same algorithm in C++ fashion may be given by directly emailing us.

7 Conclusions

Understanding the properties of the Higgs potential is of utmost importance to eventually answer one of the key questions in particle physics, namely what is the origin of EWSB. Examining the production of Higgs pairs is a crucial experiment to achieve this goal, with a focus both on the total cross section, as well as on distributions, which contain valuable information on the nature of potential NP.
In this article we presented semi-analytic results for the distributions of the Higgs-pair production cross section in a well defined and maximally general parametrization of nature, as it appears at low energies, i.e., employing EFT. This allows us to express potential deviations from the SM in a consistent way as coefficients of effective operators and constraints on these operators will provide us a guidance on how nature could look at shortest distances.

Furthermore, we employed the obtained formula to recast exclusion bounds, derived assuming certain benchmark points in parameter space with a given kinematics, to points with modified kinematic distributions. The presented method is crucial to cover the full EFT parameter space, taking correctly into account the efficiencies of signal selections.

Finally, the results presented are also useful to confront explicit models, mapped to an EFT, efficiently with constraints from the LHC, providing a bridge between (explicit) theories and data. We demonstrated this procedure, using recent ATLAS and CMS results, for various NP setups, like models with additional scalars, including 2HDM, vector-like fermions, and minimal composite Higgs models, delivering also a dictionary between their explicit parameters and effective couplings after electroweak symmetry breaking.

Acknowledgments

FG and AC are grateful for the hospitality of the CERN theory division during the completion of this work. AC is grateful to Andre Tinoco Mendes, Luca Cadamuro, Giacomo Ortona, Olivier Bondu, Konstantin Androsov, Andrey Pozdnyakov, Rafael Teixeira de Lima, Martino Dall’Osso and Tommaso Dorigo for the fruitful discussions and encouragement to pursue this work. KM is supported in part by the Belgian Federal Science Policy Office through the Interuniversity Attraction Pole P7/37 and by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 707983.

A Rescaling of cross section due to modified branching ratios

In this appendix we discuss the inclusion of modifications of the Higgs branching ratios due to the effective couplings, which can have a considerable impact on the results of the analysis. These effects can be incorporated by re-scaling the cross section with

\[
R_{BR} \equiv \frac{\text{BR}(h \rightarrow bb)\text{BR}(h \rightarrow \gamma\gamma)}{\text{BR}(h \rightarrow bb)_{\text{SM}}\text{BR}(h \rightarrow \gamma\gamma)_{\text{SM}}} = \frac{\Gamma(h \rightarrow bb)\Gamma(h \rightarrow \gamma\gamma)}{\Gamma(h \rightarrow bb)_{\text{SM}}\Gamma(h \rightarrow \gamma\gamma)_{\text{SM}}} \frac{\Gamma_{2H,SM}^2}{\Gamma_{H,SM}^2}
\]

(A.1)

where

\[
\kappa_f^2 = \frac{\Gamma(h \rightarrow ff)}{\Gamma(h \rightarrow ff)_{\text{SM}}},
\]

(A.2)

and (see, e.g., [52, 53])

\[
R_{\Gamma} \equiv \frac{\Gamma_H}{\Gamma_{H,SM}} \approx 0.59\kappa_b^2 + 0.23\kappa_V^2 + 0.07\kappa_f^2 + 0.06\kappa_T^2 + 0.03\kappa_Z^2 + 0.02\kappa_c^2.
\]

(A.3)

While explicit results for the rescaling factors \(\kappa_i\) for the models considered in this article will be discussed below, we already note that for all cases, besides \(i = g, \gamma\), considering tree
level effects is sufficient to good approximation. For the latter, loop effects from NP need to be taken into account since they furnish generically the leading corrections.

One class of such effects in \( \kappa_g \) and \( \kappa_g \) arises indirectly from changed couplings, \( \kappa_q, \kappa_W \neq 1 \), in the loop mediated decays. Neglecting small corrections due to light quarks, \( q = b, c, \ldots \), at LO the decay to gluons is just governed by the modified top coupling entering the triangle diagram

\[
\kappa_g = \kappa_t ,
\]

while the result for the decay to photons can be obtained by the weighted ratio of loop functions (see [52, 53])

\[
\kappa_{\gamma} = \left| \frac{\kappa_W F_W(\tau_W) + N_c Q^2 t_{\kappa_q} F_q(\tau_t)}{F_W(\tau_W) + N_c Q^2 t_{\kappa_q} F_q(\tau_t)} \right| \sim 1.283 \kappa_W - 0.283 \kappa_t ,
\]

where we plugged in numerical results for the mass ratios

\[
\tau_X = \frac{4 m^2_X}{m^2_h} ,
\]

and

\[
F_W(\beta) = 2 + 3\beta + 3\beta(2 - \beta) f(\beta) , \quad F_f(\beta) = -2\beta(1 + (1 - \beta) f(\beta)) ,
\]

\[
f(\beta) = \begin{cases} 
\arcsin^2(\beta^{-\frac{1}{2}}) , & \beta \geq 1 \\
-\frac{1}{4} \left[ \log \frac{1+\sqrt{1-\beta}}{1-\sqrt{1-\beta}} - i\pi \right]^2 , & \beta < 1 .
\end{cases}
\]

Another class of effects is due to NP states running directly in the loop diagrams and thereby generating the operators \( |H|^2 V_{\mu\nu} V^{\mu\nu}, V = G, B, W \), on which we will comment for each scenario separately.

We will now provide explicit results for the re-scaling factor (A.1) for all the explicit models at hand.

### A.1 Scalar singlet

For this simple scenario, all \( \kappa_i \) receive a common factor due to the singlet admixture into the Higgs boson,

\[
\kappa_i \equiv \kappa_{\text{singlet}} \quad \forall i ,
\]

dropping out in (A.1). This is also reflected by the fact that, integrating out a singlet, the only operator generated relevant for the above discussion is \( \mathcal{O}_H \), which renormalizes the Higgs kinetic term (see, e.g., [36]). We thus obtain

\[
R_{BR} = 1 .
\]
A.2 Scalar doublet

As shown in [37], integrating out the additional Higgs doublet at tree level generates (besides four-fermion operators) only modified Yukawa and Higgs trilinear interactions,

\[
\frac{c_6 \lambda_{SM}}{\Lambda^2} = -\frac{Z_6^2}{m_H^2}, \quad \frac{c_f}{\Lambda^2} = -\eta_f \frac{Z_6}{t_\beta m_H^2}, \tag{A.12}
\]

where \( f \) denotes any SM fermion and here we will focus on the third generation (although the first two behave similar). The 2HDM type dictates the relationship between \( \eta_t, \eta_b, \) and \( \eta_\tau, \) which reads

- **Type I:** \( \eta_t = \eta_b = \eta_\tau = 1, \)
- **Type II:** \( \eta_t = 1, \quad \eta_b = \eta_\tau = -t_\beta^2, \)
- **Type X:** \( \eta_t = \eta_b = 1, \quad \eta_\tau = -t_\beta^2, \)
- **Type Y:** \( \eta_t = 1, \quad \eta_b = -t_\beta^2, \quad \eta_\tau = 1, \)

and translates to

\[
\kappa_f = 1 + \frac{c_f v^2}{\Lambda^2}, \tag{A.16}
\]

- **Type I:** \( \kappa_b = \kappa_\tau = \kappa_t, \)
- **Type II:** \( \kappa_b = \kappa_\tau = 1 - t_\beta^2 (\kappa_t - 1), \)
- **Type X:** \( \kappa_b = \kappa_t, \quad \kappa_\tau = 1 - t_\beta^2 (\kappa_t - 1), \)
- **Type Y:** \( \kappa_b = 1 - t_\beta^2 (\kappa_t - 1), \quad \kappa_\tau = \kappa_t. \)

The dominant effect of \( \kappa_b \) is to modify the partial width of the Higgs to \( b \bar{b} \) while \( \kappa_t \) affects the Higgs decay to \( \gamma \gamma \) (see eq. (A.5)). Beyond the direct impact on the decays of the Higgs pair, both coefficients (as well as \( \kappa_\tau \)) enter indirectly in the branching ratios by changing the total Higgs width (A.3), where for \( \kappa_t \) the main effect comes not from the \( \gamma \gamma \) partial width but the \( gg \) one.

We note that alongside \( \kappa_t \) and \( m_H^2, \) the third degree of freedom that defines the 2HDM to leading approximation is \( Z_6, \) which characterises the departure from the alignment limit, in which the 125 GeV Higgs boson has SM-like couplings to the gauge bosons. It can be related to the usual alignment angle, \( \cos (\beta - \alpha) \equiv c_{\beta-\alpha} \) by (see, e.g., [54])

\[
v^2 Z_6 = -c_{\beta-\alpha} s_{\beta-\alpha} (m_H^2 - m_h^2), \tag{A.20}
\]

where \( s_{\beta-\alpha} \) denotes the sin of this angle. Fixing a value of \( Z_6 \) and \( m_H^2 \) determines the value of \( c_{\beta-\alpha} \) and therefore the deviations in the gauge bosons couplings

\[
1 - \kappa_V \equiv c_{\beta-\alpha}^2 = \frac{1}{2} \left[ 1 - \sqrt{1 - \left( \frac{2v^2 Z_6}{m_H^2 - m_h^2} \right)^2} \right], \tag{A.21}
\]

\[
\simeq \frac{v^4 Z_6^2}{m_H^4} + \mathcal{O} \left( \frac{1}{m_H^6} \right). \tag{A.22}
\]
This confirms that the corresponding effects arise at order \(1/m_H^4\) and therefore do not constitute a tree-level modification at dimension-6 in the EFT. It is therefore sufficient to consider only the (direct and indirect) impact of \(\kappa_{t,b,\tau}\) for fixed values of \(m_H\) and \(Z_6\) for our purposes.

The final remaining effect that could be taken into account is the contribution of charged Higgs loops to \(\kappa_\gamma\). These are given for example in \([55]\). However the \(hH^+H^-\) coupling in the 2HDM depends on an additional free parameter \(\mu^2\), which controls the soft breaking of the \(Z_2\) symmetry in the scalar potential. One can therefore find points in parameter space where the effect of this interaction is at least minimised. As discussed in, e.g., \([55]\), the effects of these loops are subdominant when the charged scalars are not too light. We choose to neglect them in our analysis as a simplifying assumption.

Employing eqs. (A.3)-(A.5), we finally arrive for the Type I (II) model at

\[
R_{\Gamma}\approx0.26+0.74\kappa_t^2 (1.26 + 0.65(t_\beta^2 - 1)^2 + 0.09\kappa_t^2)
\]

for Type I and

\[
R_{\Gamma}\approx1-1.53(\kappa_t - 1) + \mathcal{O}((\kappa_t - 1)^2)
\]

for Type II.

Analogously, we obtain

\[
R_{\Gamma}\approx1+0.6t_\beta^2 - 0.93(\kappa_t - 1) + \mathcal{O}((\kappa_t - 1)^2)
\]

for Type I and

\[
R_{\Gamma}\approx1 + (0.24t_\beta^2 - 1.29)(\kappa_t - 1) + \mathcal{O}((\kappa_t - 1)^2)
\]

for Type II.

### A.3 Scalar triplet

Since the presence of the new triplet violates custodial symmetry, the scenario is severely constrained by precision measurements and the ensuing effects in double Higgs production are not expected to be important in general.

In consequence, we refrain from providing a full one-loop analysis considering the impact from the new charged scalars in the loops (which could in principle affect \(\kappa_\gamma\) \([56–58]\)), but rather only include the effects of the modified SM (tree-level) couplings in \(R_{BR}\) for consistency. For this task, we are still lacking the modifications of Higgs couplings to leptons and gauge bosons, which we will give below.

#### A.3.1 Real triplet

For the real triplet we obtain, using results from \([25, 36, 59, 60]\),

\[
\kappa_l = \kappa_q = 1,
\]

with \(l = e, \mu, \tau\) and \(q = u, c, t, s, b\), and

\[
\kappa_Z \approx \sqrt{\frac{s_\beta^4 m_{H^+}^4}{m_H^4 + 5 s_\beta^2 m_{H^+}^4}},
\]

\[
\kappa_W \approx \sqrt{\frac{s_\beta^2 m_{H^+}^4}{m_H^4 + 1.73 s_\beta^2 m_{H^+}^4}},
\]

where we neglected quadratic NP contributions.
Plugging these results into eqs. (A.4) and (A.5) (but neglecting momentum dependent operators entering \( \kappa_W \) and thus \( h \to \gamma \gamma \) at loop level) leads to

\[
\kappa_g = 1, \quad \kappa_g = 1, \quad (A.30)
\]

\[
\kappa_\gamma \approx -0.28 + 1.28 \sqrt{1 + \frac{8s_\beta^2 m_W^4}{m_H^2}}, \quad (A.31)
\]

Finally, expressing everything in terms of \( c_2 = -2s_\beta^2 m_W^4/m_H^2 \), we obtain from eq. (A.3) \( R_T \approx 1 - 0.27c_2 \) and, via eq. (A.1),

\[
R_{BR} \approx 1 - 4.58c_2 + \mathcal{O}(c_2^2), \quad (A.32)
\]

**A.3.2 Complex triplet**

We move on to the complex triplet, for which we obtain analogously \([25, 36, 59, 60]\)

\[
\kappa_l = \kappa_q = 1 - \frac{2s_\beta^2 m_A^4}{m_H^2}, \quad (A.33)
\]

and

\[
\kappa_Z \approx \sqrt{1 + \frac{7s_\beta^2 m_A^4}{m_H^4}}, \quad (A.34)
\]

\[
\kappa_W \approx \sqrt{1 + \frac{10.27s_\beta^2 m_A^4}{m_H^4}}, \quad (A.35)
\]

where we neglected again quadratic NP contributions as well as the potential impact of four-fermion operators modifying muon decay \([25, 59]\).

As for the real triplet, we further derive

\[
\kappa_g = 1 - \frac{2s_\beta^2 m_A^4}{m_H^2}, \quad (A.36)
\]

\[
\kappa_\gamma \approx -0.28 + 0.57 \frac{s_\beta^2 m_A^4}{m_H^4} + 1.28 \sqrt{1 + \frac{4s_\beta^2 m_A^4}{m_H^2}}, \quad (A.37)
\]

Trading finally the model parameters for \( \kappa_t \), we arrive at \( R_T \approx 1.55 - 1.29\kappa_t + 0.74\kappa_t^2 \) and

\[
R_{BR} \approx 1 - 1.52(\kappa_t - 1) + \mathcal{O}((\kappa_t - 1)^2). \quad (A.38)
\]

**A.4 Scalar quartet**

Due to their quantum numbers, the scalar quartets do not modify any Higgs decays at the tree-level. In consequence, neglecting potential charged scalar contributions to \( \kappa_\gamma \) at loop level as for the triplets, we obtain

\[
R_{BR} = 1. \quad (A.39)
\]
A.5 Vector-like quark $T$ (top partner)

The new quark $T$, being an $SU(2)_L$ singlet without a doublet component, cannot propagate alone in a loop contributing to $\kappa_g$ or $\kappa_\gamma$ (before mixing with the SM quarks). Thus, to leading approximation, the only relevant correction is the change in the top yukawa coupling,

$$\kappa_t \neq 1,$$

feeding through to the loop coefficients in (A.4) and (A.5) (with $\kappa_W = \kappa_Z = 1$).

As before, a numerical evaluation leads to $R_\Gamma \approx 0.93 + 0.07 \kappa_t^2$ and

$$R_{BR} \approx \frac{(1.283 - 0.283\kappa_t)^2}{(0.93 + 0.07\kappa_t^2)^2} = 1 - 0.845(\kappa_t - 1) + O\left( (\kappa_t - 1)^2 \right).$$

(A.41)

A.6 Vector-like lepton $E$

We move on, considering a vector-like lepton singlet (coupled universally to the SM leptons). For the Higgs couplings to fermions we obtain, using again results from [25, 39, 59, 60],

$$\kappa_l = 1 - \frac{v^2|\lambda_{El}|^2}{4M_E^2},$$

(A.42)

$$\kappa_q = 1 + \frac{v^2|\lambda_{El}|^2}{4M_E^2},$$

(A.43)

where $l = e, \mu, \tau$, and $q = u, c, t, d, s, b$.

For the bosonic tree-level couplings, on the other hand, we arrive at [60] (neglecting once more quadratic NP contributions)

$$\kappa_Z \approx \sqrt{1 + 1.13 \frac{v^2|\lambda_{El}|^2}{M_E^2}},$$

(A.44)

$$\kappa_W \approx \sqrt{1 + 0.72 \frac{v^2|\lambda_{El}|^2}{M_E^2}}. $$

(A.45)

Along the same lines as for the case of the top partners, the new lepton does not contribute in loops mediating $hgg$ and $h\gamma\gamma$ vertices, such that the corresponding effective couplings are governed by the changes of the top and $W$ couplings to the Higgs boson. This results in

$$\kappa_g = \kappa_t,$$

(A.46)

$$\kappa_\gamma = -0.28 - 0.07 \frac{v^2|\lambda_{El}|^2}{M_E^2} + 0.91 \sqrt{2 + \frac{3v^2|\lambda_{El}|^2}{M_E^2}},$$

(A.47)

where we have again neglected momentum dependent operators entering $h \to \gamma\gamma$ at loop level.

Expressing everything in terms of $\kappa_t$ (and expanding to linear order around $\kappa_t = 1$), the final results read now $R_\Gamma \approx -0.29 + 0.55\kappa_t + 0.74\kappa_t^2$ and

$$R_{BR} \approx 1 + 5.06(\kappa_t - 1) + O\left( (\kappa_t - 1)^2 \right).$$

(A.48)
As a final remark, one should note that also the vector-like lepton scenario is strongly constrained from precision tests, such that in the end, the effects in Higgs pair production are not expected to be visible at the LHC (at least for the simple model considered here).

A.7 MCHM$_4$

Due to the pseudo-Goldstone nature of the Higgs, the couplings to fermions and gauge bosons are suppressed, reading [7, 40]

$$\kappa_q = \kappa_l = \kappa_W = \kappa_Z = \sqrt{1 - \xi},$$

(A.49)

with $\xi = v^2/f^2$, while we neglect fermion mixing. In turn (see, e.g., [53, 61])

$$\kappa_\gamma = \kappa_g = \sqrt{1 - \xi},$$

(A.50)

where loop effects from composite resonances, mixing with SM-like fermions, drop out. In consequence, as for the scalar singlet, all modifications in the MCHM$_4$ cancel to leading approximation and we arrive at

$$R_{BR} = 1.$$  

(A.51)

A.8 MCHM$_5$

In the MCHM$_5$, fermion couplings receive a different suppression, resulting in [7, 40]

$$\kappa_q = \kappa_l = \frac{1 - 2\xi}{\sqrt{1 - \xi}}, \quad \kappa_W = \kappa_Z = \sqrt{1 - \xi}.$$  

(A.52)

While simply

$$\kappa_g = \kappa_t = \frac{1 - 2\xi}{\sqrt{1 - \xi}},$$

(A.53)

this results in a slightly more complicated expression for the branching into photons where, employing eq. (A.5), we arrive at

$$\kappa_\gamma \approx 1.28\kappa_W - 0.28\kappa_t \approx \frac{1 - 0.72\xi}{\sqrt{1 - \xi}}.$$  

(A.54)

We finally obtain $R_\Gamma \approx (1 - 3.48\xi + 3.22\xi^2)/(1 - \xi)$ and

$$R_{BR} \approx 1 + 1.53\xi + O(\xi^2).$$  

(A.55)

B Distributions validation for shape benchmarks

We simulate the differential cross section for various points in the five-dimensional parameter space, spanned by $\{\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g}\}$, to determine $\vec{A}$. At LO accuracy the ME information differential in the invariant mass of the Higgs pair ($m_{hh}$) and in $|\cos \theta^*|$ completely defines the $2 \to 2$ process. At any subsequent simulation step — after the Higgs decays — the process evolution is independent of the NP at high energies, up to detector level. Therefore the results based on the ME-level information are applicable to reconstruct NP shapes at LO at detector level, provided that the former is known on an event-by-event basis.
After an optimization procedure, we converged to an optimal list of 59 bins in $m_{hh}$ that allowed for a precise reweighing of an ensemble of generated events to any NP shape. Since the distribution in the $\cos \theta^*$ variable is rather flat, only 4 bins are considered.\footnote{Here, and in the following, we omit the absolute value signs around $\cos \theta^*$ for brevity.} For the binning in $m_{hh}$ we choose 10 GeV intervals up to 700 GeV, where the bulk of the cross section is observed, and use a more coarse binning above. Explicitly, our binning is given by

$$m_{hh} \text{ [GeV]} = [250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1750, 2000, 13000]$$

$$(B.1)$$

$\cos \theta^* = [0,0.4,0.6,0.8,1].$$

Following the procedure established in \cite{10}, the components of $\vec{A}^j$ are extracted by maximizing the likelihood simultaneously for all the coefficients, employing an ensemble of MC simulated samples, scanning the model parameters, i.e., minimizing

$$\log L(\vec{A}^j) = -\sum_i \left( \frac{T^{(i,j)} - \sigma_{hh}^i(\vec{A}^j)}{\delta T^{(i,j)}} \right)^2,$$

where the index $i$ runs over different points in the \{$\kappa_\lambda, \kappa_\ell, c_2, c_3, c_2g$\} parameter space. Here, $T^{(i,j)} \equiv \sigma_{hh,MC}^i \text{Frac}_i^j$, with $\sigma_{hh,MC}^i$ the total cross section calculated via MC simulation and $\text{Frac}_i^j$ the corresponding bin fraction (also taken from MC), while $\sigma_{hh}^i(\vec{A}^j) \equiv R_{hh}^j \sigma_{hh}^{SM} \text{Frac}_{SM}^j$ is the differential cross section parametrization following eq. (2.4). Only the statistical (MC) uncertainty on the cross section in each bin, $\delta T^{(i,j)}$, is considered. For the SM point we assume a non-zero value of $10^{-4}$ to regularize the likelihood.

A careful choice of NP parameters is considered in the simulation to improve the convergence of the procedure. In particular it is important to avoid negative $R_{hh}^j$ values due to statistical fluctuations and, in general, a very poor population of certain bins (due to the limited total statistics). We thus consider simulated datasets such that the minimization of equation (B.2) can be done in parameter-subspaces (S$k$) with lower dimensionality that nevertheless contain the full information and guarantee appropriate statistics for all bins.\footnote{In principle, a system of 15 equations for each bin, relating the components of $\vec{A}^j$ to $R_{hh}^j$, evaluated at 15 particular sets of values of the effective couplings, could be solved recursively. This method however leads again to significantly less stable results compared to a broader fit in the parameter space, see below.} The list of subspaces and the coefficients that they serve to determine are compiled in table 9.

The simulation of events is performed within the MG5\_aMC@NLO framework \cite{62} where the entire event generation process is automated \cite{63}. We make use of UFO model
files [64] extracted from the Lagrangian (2.1), as constructed by the authors of ref. [65]. For each NP point we simulate 50,000 events, while for the SM benchmark we generate a sample of 13,000,000 events.

Our results for the fit coefficients are compiled in tables 10–13. To illustrate the bin statistics, we also show the number of events in each bin, both for our total BSM sample (see section 3) and the SM sample.

In figure 13 we provide some examples of the fit agreement with the dataset points for selected $m_{hh}$ ranges considering the bin central in $\cos \theta^*$. Here, we display three parameter space directions, which are $\kappa_\lambda$ for fixed $\kappa_t = 1$, $\kappa_t$ for $\kappa_\lambda = 10^{-4}$ and $c_2$ for $\kappa_\lambda = 1$, $\kappa_t = 1.5$. The errors on the data points are purely statistical, while the coloured lines correspond to the fit result, eq. (2.4). It becomes evident from the results that, even if it is true that for every bin one needs only a minimum of $N_{\text{fit}} = 15$ different samples to determine all coefficients, due to limited MC statistics the result of such a procedure would have a limited reliability. Concerning the behaviour of the cross section in different bins, the results confirm the expectation that a change in the scalar self coupling $\kappa_\lambda$ affects mostly the threshold region, while the ($D > 4$) $tthh$ contact interaction has a large effect on the high energy tail. Moreover, the pronounced destructive interference present in the cross section in the vicinity of the SM point $\kappa_\lambda = 1$ becomes evident via the peaked minimum, visible in the left-most plot.

In figure 14 we display the value of the coefficients of $R_{hh}^1$ for the central $\cos \theta^*$ bin as a function of $m_{hh}$. As a first observation, we see that at high invariant masses $A_{2,5,10}$ become most important, which coincide with the terms that parametrize genuine higher dimensional effects (containing exclusively $c_{2,9,2g}$), in agreement with expectations. We further note that for the subspace that corresponds solely to a variation of SM parameters (S1, with coefficients $A_{1,3,7}$), the absolute value of the coefficients does not surpass 100, and they are peaked towards low $m_{hh}$. In particular, the value of $A_7$ (that controls the triangle-box interference term $\sim \kappa_\lambda^2 \kappa_t \kappa_t$) is always negative for $m_{hh} < 400$ GeV. For $A_{2,6,8,9,10}$, which feature the contributions from the new $tthh$ contact interaction (including interference with SM-like contributions), the coefficients are in general larger compared to those in S1, resulting in stronger variations with the corresponding parameters close to the $hh$ threshold.

---

Table 9. Definition of the lower dimensional subsets $(S_k)$ to find the coefficients of equation (2.4).

| Subspace | Parameters to be determined | Coefficients to be determined | scan |
|----------|-----------------------------|-------------------------------|------|
| S1       | $\kappa_\lambda, \kappa_t$  | $A_{1,4,5}$                  | $\kappa_\lambda = 10^{-4}, \pm 2.5, \pm 5, \pm 7.5, \pm 10, \pm 12.5, \pm 15$ |
|          |                             |                               | $k_i$ from 0.5 to 2.5 with step of 0.5 |
| S2       | $\kappa_\lambda, c_1, c_2, c_g$ | $A_{5,6,10}$                 | $c_2 = \pm 1, \pm 3, c_g = \pm 0.5, \pm 1.5$ |
|          |                             |                               | $\kappa_\lambda = 1, \pm 5, \pm 10$ and $(c_1, c_g) = \pm 1$ (independently) |
| S3       | $\kappa_\lambda, c_1, c_g$  | $A_{6,10}$                    | $\kappa_\lambda = 1, \pm 5, \pm 10$ |
| S4       | $\kappa_\lambda, c_1, c_2$  | $A_{5,6}$                     | $\kappa_\lambda = 1, \pm 5, \pm 10$ |
| S5       | $\kappa_\lambda, c_1, c_2, c_g$ | $A_{5,6,10}$                 | $\kappa_\lambda = 1, \pm 5, \pm 10$ |
|          |                             |                               | $c_1$ from 0.5 to 2.5 with step of 0.5 |
|          | Additional scans to more precisely determine coefficients (fixing $\kappa_\lambda = \kappa_t = 10^{-4}$) | $c_2 = \pm 1, \pm 3$ and $c_g = \pm 0.5, \pm 1.5$ |
|          | (independently)             | $c_1$ from 0.5 to 2.5 with step of 0.5 |
|          | $c_2 = \pm 0.5, \pm 1.5$ (independently) | $c_1$ from 0.5 to 2.5 with step of 0.5 |
Figure 13. Comparison of $R_{hh}^i$, obtained from the analytical formula (2.4), with the same ratio, as derived directly via MC, for the most central bin in $\cos \theta^*$ (between 0 and 0.4) and for four different bins in $m_{hh}$, capturing threshold, interference, and high energy regions. From left to right we show a scan in $\kappa_\lambda$, $\kappa_t$ and $c_2$, respectively, while keeping the other parameters at fixed values. The error bars have a pure statistical source related with the generated number of events.

Figure 14. Values of the fit coefficients entering $R_{hh}^i$, differential in $m_{hh}$, see table 9. We only display the results for the central $\cos \theta^*$ bin (between 0 and 0.4).

and show more pronounced $m_{hh}$ tails, as discussed before. All the remaining terms contain gluon-Higgs contact interactions and as such also feature pronounced contributions at large energies. The Feynman-diagramatic representation of the different contributions to the cross section can be studied in figure 1 of [10].

Finally, in figures 15 to 18 we compare the kinematic distributions for reconstructed variables after ATLAS like selections (which are $m_X$, $\cos \theta^*$, $p_T^{\gamma\gamma}$ and $p_T^\gamma(\gamma)$) following from our reweighing procedure with fully simulated distributions, finding excellent agreement.
Figure 15. The reconstructed reduced mass after ATLAS-like selection. The histograms are normalized by signal efficiency times 100,000 events.
Figure 16. The reconstructed $\cos \theta^*$ after ATLAS-like selection. The histograms are normalized by signal efficiency times 100,000 events.
Figure 17. The reconstructed $p_T^{\gamma\gamma b\bar{b}}$ after ATLAS-like selection. The histograms are normalized by signal efficiency times 100,000 events.
Figure 18. The reconstructed $p_T(\gamma\gamma)$ after ATLAS-like selection. The histograms are normalized by signal efficiency times 100,000 events.
Figure 19. The bin-by-bin statistic of events in a 13M SM sample and a 1.2M reweighting sample constructed by the plain sum of shape benchmarks.

C Numeric tables with the coefficients

Tables 10 to 13 contain the numerical values entering eq. (2.3). The first columns of the tables stand for the number of the bin, followed by the mean \( m_{hh} \) and \( \cos \theta^* \) of that bin. We also display the number of events found in that bin for a 13,000,000 events SM sample and for the 1,200,000 events of the reweighting sample \( S_{BSM} \) used in this paper. Similar information can be found visually in figure 19 (for the number of events/bin) and 20 (for the values of the coefficients in the not central \( \cos \theta^* \) bins).
Figure 20. Values of the fit coefficients entering $R_{hh}^j$, differential in $m_{hh}$, see table 9.
Table 10. Coefficients of eq. (2.4), where \( j \) is the bin number and \( m_{hh} \) and \( \cos\theta^* \) are the central values in the corresponding bin. Moreover, \( N_{SM}^A \) (\( N_{\text{Sum}} \)) is the number of events in that bin for the SM sample (sum of benchmarks) out of 13M (1.2M) events.

| \( j \) | \( m_{hh} \) (GeV) | \( N_{SM}^A \) | \( N_{\text{Sum}} \) | A1 | A2 | A3 | A4 | A5 | A6 | A7 | A8 | A9 | A10 | A11 | A12 | A13 | A14 | A15 |
|-----|----------------|-----------|-------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 0   | 495.0          | 1.00      | 1.000       | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1   | 255.0          | 0.5      | 0.983      | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 2   | 255.0          | 0.25     | 0.998      | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| 3   | 255.0          | 0.125    | 0.991      | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| 4   | 255.0          | 0.0625   | 0.982      | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |

**Note:** The table continues with the same structure for each bin, providing detailed coefficients for various mass and cosine values, as well as the sum of benchmarks for the SM sample.
| j | m_{hh} | \cos \theta | N_{SM} | \sum N_{SM} |
|---|---|---|---|---|
| 61 | 405 | 0.2 | 58279 | 803 |
| 62 | 405 | 0.2 | 21620 | 405 |
| 63 | 405 | 0.2 | 29535 | 413 |
| 64 | 405 | 0.2 | 30121 | 414 |
| 65 | 405 | 0.2 | 56734 | 307 |
| 66 | 405 | 0.2 | 28448 | 399 |
| 67 | 405 | 0.2 | 53180 | 780 |
| 68 | 405 | 0.2 | 27049 | 395 |
| 69 | 405 | 0.2 | 27261 | 407 |
| 70 | 405 | 0.2 | 29770 | 392 |
| 71 | 405 | 0.2 | 20181 | 784 |
| 72 | 405 | 0.2 | 48676 | 781 |
| 73 | 405 | 0.2 | 25349 | 395 |
| 74 | 405 | 0.2 | 25350 | 381 |
| 75 | 405 | 0.2 | 23758 | 376 |
| 76 | 405 | 0.2 | 24147 | 388 |
| 77 | 405 | 0.2 | 42819 | 751 |
| 78 | 405 | 0.2 | 29866 | 773 |
| 79 | 405 | 0.2 | 22275 | 3770 |
| 80 | 405 | 0.2 | 22753 | 3793 |
| 81 | 405 | 0.2 | 20095 | 3067 |
| 82 | 405 | 0.2 | 20239 | 3716 |
| 83 | 405 | 0.2 | 20076 | 3561 |
| 84 | 405 | 0.2 | 16626 | 7230 |
| 85 | 405 | 0.2 | 18719 | 3471 |
| 86 | 405 | 0.2 | 15780 | 5919 |
| 87 | 405 | 0.2 | 19475 | 3528 |
| 88 | 405 | 0.2 | 15367 | 6891 |
| 89 | 405 | 0.2 | 17444 | 3473 |
| 90 | 405 | 0.2 | 17889 | 3453 |
| 91 | 405 | 0.2 | 10330 | 2196 |
| 92 | 405 | 0.2 | 13291 | 3322 |
| 93 | 405 | 0.2 | 16167 | 4322 |
| 94 | 405 | 0.2 | 15063 | 3332 |
| 95 | 405 | 0.2 | 27997 | 6594 |
| 96 | 405 | 0.2 | 25283 | 6351 |
| 97 | 405 | 0.2 | 14414 | 3214 |
| 98 | 405 | 0.2 | 15322 | 3179 |
| 99 | 405 | 0.2 | 12528 | 6251 |
| 100 | 405 | 0.2 | 12999 | 3121 |
| 101 | 405 | 0.2 | 13375 | 3082 |
| 102 | 405 | 0.2 | 13254 | 3322 |
| 103 | 405 | 0.2 | 23150 | 6260 |
| 104 | 405 | 0.2 | 13867 | 2985 |
| 105 | 405 | 0.2 | 13763 | 2960 |
| 106 | 405 | 0.2 | 12678 | 3024 |
| 107 | 405 | 0.2 | 21003 | 5873 |
| 108 | 405 | 0.2 | 11421 | 2935 |
| 109 | 405 | 0.2 | 11040 | 2912 |
| 110 | 405 | 0.2 | 11742 | 2872 |
| 111 | 405 | 0.2 | 11372 | 2851 |
| 112 | 405 | 0.2 | 9660 | 2819 |
| 113 | 405 | 0.2 | 10096 | 2814 |
| 114 | 405 | 0.2 | 10066 | 2817 |
| 115 | 405 | 0.2 | 10248 | 2820 |

Table 11. Coefficients of eq. (2.4), where j is the bin number and m_{hh} and \cos \theta are the central values in the corresponding bin. Moreover, N_{SM}^j (\sum N_{SM}^j) is the number of events in that bin for the SM sample (sum of benchmarks) out of 13M (12M) events.
Table 12. Coefficients of eq. (2.4), where $j$ is the bin number and $m_{hh}$ and $\cos \theta^*$ are the central values in the corresponding bins. Moreover, $N_{SM}^{N_{Sum}}$ is the number of events in that bin for the SM sample (sum of benchmarks) out of 13M (1.2M) events.

| $j$ | $m_{hh}$ | $\cos \theta^*$ | $N_{SM}^{N_{Sum}}$ |
|-----|----------|----------------|-------------------|
| A1  | A2       | A3             | A4               |
| A5  | A6       | A7             | A8               |
| A9  | A10      | A11            | A12              |
| A13 | A14      | A15            |                  |
| 120 | 555      | 0.2            | 1.473 ± 0.052    |
| 121 | 555      | 0.5            | 2.914 ± 0.052    |
| 122 | 555      | 0.7            | 3.450 ± 0.052    |
| 123 | 555      | 0.9            | 3.986 ± 0.052    |
| 124 | 555      | 1.1           | 4.521 ± 0.052    |
| 125 | 555      | 1.3           | 5.056 ± 0.052    |
| 126 | 555      | 1.5           | 5.591 ± 0.052    |
| 127 | 555      | 1.7           | 6.126 ± 0.052    |
| 128 | 555      | 1.9           | 6.661 ± 0.052    |
| 129 | 555      | 2.1          | 7.196 ± 0.052    |
| 130 | 555      | 2.3          | 7.731 ± 0.052    |
| 131 | 555      | 2.5          | 8.266 ± 0.052    |
| 132 | 555      | 2.7          | 8.801 ± 0.052    |
| 133 | 555      | 2.9          | 9.336 ± 0.052    |
| 134 | 555      | 3.1          | 9.871 ± 0.052    |
| 135 | 555      | 3.3          | 10.406 ± 0.052   |
| 136 | 555      | 3.5          | 10.941 ± 0.052   |
| 137 | 555      | 3.7          | 11.476 ± 0.052   |
| 138 | 555      | 3.9          | 12.011 ± 0.052   |
| 139 | 555      | 4.1          | 12.546 ± 0.052   |
| 140 | 555      | 4.3          | 13.081 ± 0.052   |
| 141 | 555      | 4.5          | 13.616 ± 0.052   |
| 142 | 555      | 4.7          | 14.151 ± 0.052   |
| 143 | 555      | 4.9          | 14.686 ± 0.052   |
| 144 | 555      | 5.1          | 15.221 ± 0.052   |
| 145 | 555      | 5.3          | 15.756 ± 0.052   |
| 146 | 555      | 5.5          | 16.291 ± 0.052   |
| 147 | 555      | 5.7          | 16.826 ± 0.052   |
| 148 | 555      | 5.9          | 17.361 ± 0.052   |
| 149 | 555      | 6.1          | 17.896 ± 0.052   |
| 150 | 555      | 6.3          | 18.431 ± 0.052   |
| 151 | 555      | 6.5          | 18.966 ± 0.052   |
| 152 | 555      | 6.7          | 19.501 ± 0.052   |
| 153 | 555      | 6.9          | 20.036 ± 0.052   |
| 154 | 555      | 7.1          | 20.571 ± 0.052   |
| 155 | 555      | 7.3          | 21.106 ± 0.052   |
| 156 | 555      | 7.5          | 21.641 ± 0.052   |
| 157 | 555      | 7.7          | 22.176 ± 0.052   |
| 158 | 555      | 7.9          | 22.711 ± 0.052   |
| 159 | 555      | 8.1          | 23.246 ± 0.052   |
| 160 | 555      | 8.3          | 23.781 ± 0.052   |
| 161 | 555      | 8.5          | 24.316 ± 0.052   |
| 162 | 555      | 8.7          | 24.851 ± 0.052   |
| 163 | 555      | 8.9          | 25.386 ± 0.052   |
| 164 | 555      | 9.1          | 25.921 ± 0.052   |
| 165 | 555      | 9.3          | 26.456 ± 0.052   |
| 166 | 555      | 9.5          | 26.991 ± 0.052   |
| 167 | 555      | 9.7          | 27.526 ± 0.052   |
| 168 | 555      | 9.9          | 28.061 ± 0.052   |

$N_{SM}^{N_{Sum}}$ is the number of events in that bin for the SM sample (sum of benchmarks) out of 13M (1.2M) events.
| $j$ | $m_{hh}$ | $\cos \theta$ | $N_{SM}$ | $N_{SM}$  |
|----|--------|--------|--------|--------|
| 181 | 725 | 0.2 | 1750 | 1415 |
| 182 | 725 | 0.5 | 9075 | 7065 |
| 183 | 725 | 0.7 | 2551 | 1600 |
| 184 | 725 | 0.9 | 11138 | 6927 |
| 185 | 725 | 1.1 | 11119 | 1181 |
| 186 | 725 | 1.3 | 5994 | 5994 |
| 187 | 725 | 1.5 | 6339 | 5820 |
| 188 | 725 | 0.9 | 7353 | 5854 |
| 189 | 825 | 0.2 | 7380 | 9003 |
| 190 | 825 | 0.5 | 3736 | 4896 |
| 191 | 825 | 0.7 | 4244 | 4495 |
| 192 | 825 | 0.9 | 4654 | 8172 |
| 193 | 825 | 1.1 | 2965 | 4106 |
| 194 | 825 | 1.3 | 3276 | 4253 |
| 195 | 825 | 1.5 | 3129 | 7083 |
| 196 | 825 | 0.9 | 1742 | 3407 |
| 197 | 825 | 1.1 | 1868 | 3521 |
| 198 | 825 | 1.3 | 2250 | 3568 |
| 199 | 825 | 1.5 | 2201 | 6044 |
| 200 | 825 | 0.9 | 1209 | 2953 |
| 201 | 825 | 1.1 | 1356 | 4610 |
| 202 | 825 | 1.3 | 1465 | 4631 |
| 203 | 825 | 1.5 | 1872 | 4625 |
| 204 | 825 | 0.9 | 660 | 5212 |
| 205 | 825 | 1.1 | 514 | 2542 |
| 206 | 825 | 1.3 | 679 | 3433 |
| 207 | 825 | 1.5 | 776 | 3427 |
| 208 | 825 | 0.9 | 1005 | 3429 |
| 209 | 825 | 1.1 | 660 | 5212 |
| 210 | 825 | 1.3 | 348 | 2542 |
| 211 | 825 | 1.5 | 599 | 2484 |
| 212 | 825 | 0.9 | 1507 | 2986 |
| 213 | 825 | 1.1 | 1568 | 2832 |
| 214 | 825 | 1.3 | 1717 | 1977 |
| 215 | 825 | 1.5 | 225 | 1980 |
| 216 | 825 | 0.9 | 294 | 2051 |
| 217 | 825 | 1.1 | 204 | 2096 |
| 218 | 825 | 1.3 | 7 | 256 |
| 219 | 825 | 1.5 | 294 | 2051 |
| 220 | 825 | 0.9 | 54 | 270 |
| 221 | 825 | 1.1 | 175 | 1447 |
| 222 | 825 | 1.3 | 162 | 1652 |
| 223 | 825 | 1.5 | 132 | 2386 |
| 224 | 825 | 0.9 | 156 | 2441 |
| 225 | 825 | 1.1 | 172 | 2775 |
| 226 | 825 | 1.3 | 185 | 2383 |
| 227 | 825 | 1.5 | 34 | 1209 |
| 228 | 825 | 0.9 | 77 | 2176 |
| 229 | 825 | 1.1 | 225 | 172 |
| 230 | 825 | 1.3 | 200 | 1135 |
| 231 | 825 | 1.5 | 246 | 300 |
| 232 | 825 | 0.9 | 73 | 237 |
| 233 | 825 | 1.1 | 170 | 2353 |
| 234 | 825 | 1.3 | 193 | 2162 |
| 235 | 825 | 1.5 | 203 | 2679 |
| 236 | 825 | 0.9 | 94 | 1529 |
| 237 | 825 | 1.1 | 175 | 1447 |
| 238 | 825 | 1.3 | 200 | 2096 |
| 239 | 825 | 1.5 | 246 | 300 |
| 240 | 825 | 0.9 | 73 | 237 |
| 241 | 825 | 1.1 | 170 | 2353 |
| 242 | 825 | 1.3 | 193 | 2162 |
| 243 | 825 | 1.5 | 203 | 2679 |
| 244 | 825 | 0.9 | 94 | 1529 |
| 245 | 825 | 1.1 | 175 | 1447 |
| 246 | 825 | 1.3 | 200 | 2096 |
| 247 | 825 | 1.5 | 246 | 300 |
| 248 | 825 | 0.9 | 73 | 237 |
| 249 | 825 | 1.1 | 170 | 2353 |
| 250 | 825 | 1.3 | 193 | 2162 |
| 251 | 825 | 1.5 | 203 | 2679 |

Table 13. Coefficients of eq. (2.4), where $j$ is the bin number and $m_{hh}$ and $\cos \theta$ are the central values in the corresponding bin. Moreover, $N_{SM}^j$ ($N_{SM}^{jumi}$) is the number of events in batch $j$ for the SM sample (some mass mode) and sample) output of 13M (12M) events.
Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

[1] I.V. Krive and A.D. Linde, *On the vacuum stability problem in gauge theories*, Nucl. Phys. B **117** (1976) 265 [insPIRE].

[2] M. Lindner, M. Sher and H.W. Zaglauer, *Probing vacuum stability bounds at the Fermilab collider*, Phys. Lett. B **228** (1989) 139 [insPIRE].

[3] G. Degrassi et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, JHEP **08** (2012) 098 [arXiv:1205.6497] [insPIRE].

[4] C. Grojean, G. Servant and J.D. Wells, *First-order electroweak phase transition in the Standard Model with a low cutoff*, Phys. Rev. D **71** (2005) 036001 [hep-ph/0407019] [insPIRE].

[5] F. Goertz, *Electroweak symmetry breaking without the $\mu^2$ term*, Phys. Rev. D **94** (2016) 015013 [arXiv:1504.00355] [insPIRE].

[6] P. Huang, A. Joglekar, B. Li and C.E.M. Wagner, *Probing the electroweak phase transition at the LHC*, Phys. Rev. D **93** (2016) 055049 [arXiv:1512.00068] [insPIRE].

[7] G.F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, *The strongly-interacting light Higgs*, JHEP **06** (2007) 045 [hep-ph/0703164] [insPIRE].

[8] R. Grober and M. Mühlleitner, *Composite Higgs boson pair production at the LHC*, JHEP **06** (2011) 020 [arXiv:1012.1562] [insPIRE].

[9] A. Carmona and F. Goertz, *A naturally light Higgs without light top partners*, JHEP **05** (2015) 002 [arXiv:1410.8555] [insPIRE].

[10] A. Carvalho, M. Dall’Osso, T. Dorigo, F. Goertz, C.A. Gottardo and M. Tosi, *Higgs pair production: choosing benchmarks with cluster analysis*, JHEP **04** (2016) 126 [arXiv:1507.02245] [insPIRE].

[11] A. Azatov, R. Contino, G. Panico and M. Son, *Effective field theory analysis of double Higgs boson production via gluon fusion*, Phys. Rev. D **92** (2015) 035001 [arXiv:1502.00539] [insPIRE].

[12] F. Goertz, A. Papaefstathiou, L.L. Yang and J. Zurita, *Higgs boson pair production in the $D = 6$ extension of the SM*, JHEP **04** (2015) 167 [arXiv:1410.3471] [insPIRE].

[13] J.S. Gainer, J. Lykken, K.T. Matchev, S. Mrenna and M. Park, *Exploring theory space with Monte Carlo reweighting*, JHEP **10** (2014) 078 [arXiv:1404.7129] [insPIRE].

[14] O. Mattelaer, *On the maximal use of Monte Carlo samples: re-weighting events at NLO accuracy*, Eur. Phys. J. C **76** (2016) 674 [arXiv:1607.00763] [insPIRE].

[15] R. Contino, A. Falkowski, F. Goertz, C. Grojean and F. Riva, *On the validity of the effective field theory approach to SM precision tests*, JHEP **07** (2016) 144 [arXiv:1604.06444] [insPIRE].

[16] ATLAS collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in $pp$ collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **97** (2018) 072016 [arXiv:1712.08985] [insPIRE].
[17] A. Carvalho et al., *Analytical parametrization and shape classification of anomalous HH production in the EFT approach*, arXiv:1608.06578 [inSPIRE].

[18] ATLAS collaboration, *Search for Higgs boson pair production in the bbγγ final state using pp collision data at √s = 13 TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-004, CERN, Geneva, Switzerland (2016).

[19] CMS collaboration, *Search for H(bb)H(γγ) decays at 13 TeV*, Tech. Rep. CMS-PAS-HIG-16-032, CERN, Geneva, Switzerland (2016).

[20] ATLAS collaboration, *Searches for Higgs boson pair production in the hh → bbττ, γγWW*, γγbb, bbbh channels with the ATLAS detector, *Phys. Rev. D* 92 (2015) 092004 [arXiv:1509.04670] [inSPIRE].

[21] ATLAS collaboration, *Search for Higgs boson pair production in the γγb¯b final state using pp collision data at √s = 8 TeV from the ATLAS detector*, *Phys. Rev. Lett.* 114 (2015) 081802 [arXiv:1406.5053] [inSPIRE].

[22] CMS collaboration, *Search for two Higgs bosons in final states containing two photons and two bottom quarks in proton-proton collisions at 8 TeV*, *Phys. Rev. D* 94 (2016) 052012 [arXiv:1603.06896] [inSPIRE].

[23] LHC Higgs Cross Section Working Group collaboration, *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*, arXiv:1610.07922 [inSPIRE].

[24] J. Butterworth et al., *PDF4LHC recommendations for LHC run II*, *J. Phys. G* 43 (2016) 023001 [arXiv:1510.03865] [inSPIRE].

[25] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* 93 (2016) 033006 [arXiv:1506.07443] [inSPIRE].

[26] L.A. Harland-Lang, A.D. Martin, P. Motylinski and R.S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* 75 (2015) 204 [arXiv:1412.3989] [inSPIRE].

[27] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* 191 (2015) 159 [arXiv:1410.3012] [inSPIRE].

[28] DELPHES 3 collaboration, *DELPHES 3, a modular framework for fast simulation of a generic collider experiment, JHEP 02 (2014) 057* [arXiv:1307.6346] [inSPIRE].

[29] M. Cacciari, G.P. Salam and G. Soyez, *The anti-kt jet clustering algorithm, JHEP 04 (2008) 063* [arXiv:0802.1189] [inSPIRE].

[30] CMS delphes card, https://github.com/delphes/delphes/blob/master/cards/delphes_card_CMS.tcl.
[35] S. Dawson and C.W. Murphy, Standard Model EFT and extended scalar sectors, Phys. Rev. D 96 (2017) 015041 [arXiv:1704.07851] [nSPIRE].

[36] J. de Blas, M. Chala, M. Pérez-Victoria and J. Santiago, Observable effects of general new scalar particles, JHEP 04 (2015) 078 [arXiv:1412.8480] [nSPIRE].

[37] H. Bélusca-Maïto, A. Falkowski, D. Fontes, J.C. Romão and J.P. Silva, Higgs EFT for 2HDM and beyond, Eur. Phys. J. C 77 (2017) 176 [arXiv:1611.01112] [nSPIRE].

[38] F. del Aguila, M. Pérez-Victoria and J. Santiago, Observable contributions of new exotic quarks to quark mixing, JHEP 09 (2000) 011 [hep-ph/0007316] [nSPIRE].

[39] F. del Aguila, J. de Blas and M. Pérez-Victoria, Effects of new leptons in electroweak precision data, Phys. Rev. D 78 (2008) 013010 [arXiv:0803.4008] [nSPIRE].

[40] R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner and M. Spira, Effective Lagrangian for a light Higgs-like scalar, JHEP 07 (2013) 035 [arXiv:1303.3876] [nSPIRE].

[41] R. Contino, The Higgs as a composite Nambu-Goldstone boson, in Theoretical advanced study institute in elementary particle physics: physics of the large and the small, World Scientific, Singapore (2011), pg. 235 [arXiv:1005.4269] [nSPIRE].

[42] S. Baker and R.D. Cousins, Clarification of the use of chi-square and likelihood functions in fits to histograms, Nucl. Instrum. Meth. 221 (1984) 437 [nSPIRE].

[43] S.S. Wilks, The large-sample distribution of the likelihood ratio for testing composite hypotheses, Ann. Math. Statist. 9 (1938) 60.

[44] F. Maltoni, E. Vryonidou and C. Zhang, Higgs production in association with a top-antitop pair in the Standard Model effective field theory at NLO in QCD, JHEP 10 (2016) 123 [arXiv:1607.05330] [nSPIRE].

[45] S. Dawson, A. Ismail and I. Low, What’s in the loop? The anatomy of double Higgs production, Phys. Rev. D 91 (2015) 115008 [arXiv:1504.05596] [nSPIRE].

[46] G. Cacciapaglia et al., Probing vector-like quark models with Higgs-boson pair production, JHEP 07 (2017) 005 [arXiv:1703.10614] [nSPIRE].

[47] CMS collaboration, Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s} = 13$ TeV, arXiv:2011.12373 [nSPIRE].

[48] CMS collaboration, Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. 122 (2019) 121803 [arXiv:1811.09689] [nSPIRE].

[49] A. Falkowski, B. Fuks, K. Mawatari, K. Mimasu, F. Riva and V. Sanz, Rosetta: an operator basis translator for Standard Model effective field theory, Eur. Phys. J. C 75 (2015) 583 [arXiv:1508.05895] [nSPIRE].

[50] R. Contino, M. Ghezzi, C. Grojean, M. Mühlleitner and M. Spira, eHDECAY: An implementation of the Higgs effective Lagrangian into HDECAY, Comput. Phys. Commun. 185 (2014) 3412 [arXiv:1403.3381] [nSPIRE].

[51] G. Brooijmans et al., Les Houches 2015: Physics at TeV colliders — New physics working group report, in 9th Les Houches workshop on physics at TeV colliders, (2016) [arXiv:1605.02684] [nSPIRE].
[52] A. Djouadi, The anatomy of electro-weak symmetry breaking. I: the Higgs boson in the Standard Model, *Phys. Rept.* **457** (2008) 1 [hep-ph/0503172] [inSPIRE].

[53] A. Carmona and F. Goertz, Custodial leptons and Higgs decays, *JHEP* **04** (2013) 163 [arXiv:1301.5856] [inSPIRE].

[54] J. Bernon, J.F. Gunion, H.E. Haber, Y. Jiang and S. Kraml, Scrutinizing the alignment limit in two-Higgs-doublet models: $m_h = 125$ GeV, *Phys. Rev. D* **92** (2015) 075004 [arXiv:1507.00933] [inSPIRE].

[55] P. Posch, Enhancement of $h \to \gamma \gamma$ in the two Higgs doublet model type I, *Phys. Lett. B* **696** (2011) 447 [arXiv:1001.1759] [inSPIRE].

[56] A.G. Akeroyd and S. Moretti, Enhancement of $H \to \gamma \gamma$ from doubly charged scalars in the Higgs triplet model, *Phys. Rev. D* **86** (2012) 035015 [arXiv:1206.0535] [inSPIRE].

[57] A.G. Akeroyd and S. Moretti, Enhancement of $H \to \gamma \gamma$ from charged Higgs bosons in the Higgs triplet model, *PoS*(CHARGED2012)035 (2012) [arXiv:1210.6882] [inSPIRE].

[58] A. Arhrib, R. Benbrik, M. Chabab, G. Moultaka and L. Rahili, $h \gamma \gamma$ coupling in Higgs triplet model, in *International workshop on future linear colliders (LCWS11)*, (2012) [arXiv:1202.6621] [inSPIRE].

[59] A. Falkowski and A. Falkowski, Higgs basis: proposal for an EFT basis choice for LHC HXSWG, (2015).

[60] A. Falkowski, Effective field theory approach to LHC Higgs data, *Pramana* **87** (2016) 39 [arXiv:1505.00046] [inSPIRE].

[61] M. Carena, L. Da Rold and E. Pontón, Minimal composite Higgs models at the LHC, *JHEP* **06** (2014) 159 [arXiv:1402.2987] [inSPIRE].

[62] J. Alwall et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, *JHEP* **07** (2014) 079 [arXiv:1405.0301] [inSPIRE].

[63] N.D. Christensen et al., A comprehensive approach to new physics simulations, *Eur. Phys. J. C* **71** (2011) 1541 [arXiv:0906.2474] [inSPIRE].

[64] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO — the Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201 [arXiv:1108.2040] [inSPIRE].

[65] B. Hespel, D. Lopez-Val and E. Vryonidou, Higgs pair production via gluon fusion in the two-Higgs-doublet model, *JHEP* **09** (2014) 124 [arXiv:1407.0281] [inSPIRE].