

RECEIVED: November 4, 2019

REVISED: February 7, 2020

ACCEPTED: February 29, 2020

PUBLISHED: March 16, 2020

Exploring anomalous couplings in Higgs boson pair production through shape analysis

M. Capozzi and G. Heinrich

*Max Planck Institute for Physics,
Föhringer Ring 6, 80805 München, Germany*

E-mail: mcapozzi@mpp.mpg.de, gudrun@mpp.mpg.de

ABSTRACT: We classify shapes of Higgs boson pair invariant mass distributions m_{hh} , calculated at NLO with full top quark mass dependence, and visualise how distinct classes of shapes relate to the underlying coupling parameter space. Our study is based on a five-dimensional parameter space relevant for Higgs boson pair production in a non-linear Effective Field Theory framework. We use two approaches: an analysis based on predefined shape types and a classification into shape clusters based on unsupervised learning. We find that our method based on unsupervised learning is able to capture shape features very well and therefore allows a more detailed study of the impact of anomalous couplings on the m_{hh} shape compared to more conventional approaches to a shape analysis.

KEYWORDS: QCD Phenomenology, NLO Computations

ARXIV EPRINT: [1908.08923](https://arxiv.org/abs/1908.08923)

Contents

1	Introduction	1
2	Classification through predefined shape types	4
2.1	Parametrisation of anomalous couplings in the Higgs sector	4
2.2	Definition of shape types	5
2.3	Classification of m_{hh} distributions	6
2.4	Classification of $p_{T,h}$ distributions	9
3	Classification and clustering by unsupervised learning	10
3.1	Unsupervised learning procedure	10
3.2	Parameter space underlying the clusters	12
3.3	Identification of benchmark points	17
4	Conclusions	19

1 Introduction

The Higgs sector as we see it today is probably just a glimpse of an underlying more general structure still awaiting to be explored. Manifestations of new physics at higher scales would lead to operators which on one hand introduce new, effective couplings and on the other hand also modify interactions known in the Standard Model (SM). Therefore it is a primary goal for collider physics in the next decades to constrain the couplings, in particular in the Higgs sector, to an unprecedented precision. This is particularly true for the Higgs boson self couplings, in order to find out whether the Higgs potential is indeed of the form assumed by the SM. Deviations from this form could provide strong hints about how to extend the SM.

The trilinear Higgs boson coupling λ can be constrained by measurements of Higgs boson pair production [1, 2], where the gluon fusion channel yields the largest cross section, and the most stringent 95% CL limits on the total $gg \rightarrow HH$ cross section at $\sqrt{s} = 13$ TeV are currently $\sigma_{\max}^{HH} = 6.9 \times \sigma_{\text{SM}}$, constraining trilinear coupling modifications to the range $-5.0 \leq \lambda/\lambda_{\text{SM}} \leq 12.0$ [1].

The trilinear Higgs couplings can also be constrained in an indirect way, through measurements of processes which are sensitive to these couplings via electroweak corrections [3–14]. Such processes offer important complementary information, however they are susceptible to other BSM couplings entering the loop corrections at the same level, and therefore the limits on $c_{hhh} = \lambda/\lambda_{\text{SM}}$ extracted this way may be more model dependent than the ones extracted from the direct production of Higgs boson pairs. A corresponding experimental analysis based on single Higgs boson production processes has

been performed [15], and recently combined constraints from single and double Higgs boson production became available [16]. The idea of indirect constraints through loop corrections also has been employed trying to constrain the quartic Higgs boson self-coupling from (partial) EW corrections to Higgs boson pair production [17, 18].

Theoretical constraints on c_{hhh} are rather loose if derived in a largely model independent way. Recent work based on general concepts like vacuum stability and perturbative unitarity suggests that $|c_{hhh}| \lesssim 4$ for a new physics scale in the few TeV range [19–22]. More specific models can lead to more stringent bounds, see e.g. refs. [23–27]. Recent phenomenological studies about the precision that could be reached for the trilinear coupling at the (HL-)LHC and future hadron colliders are summarised in refs. [28–30].

Higgs boson pair production in gluon fusion in the SM has been calculated at leading order in refs. [31–33], and at NLO in the $m_t \rightarrow \infty$ limit, rescaled with the full Born matrix element, in ref. [34]. ref. [35] contains the full top quark mass dependence in the real radiation, while the virtual part is calculated in the heavy top limit. The NLO QCD corrections with full top quark mass dependence became available more recently [36–38]. Implementations of the full NLO QCD corrections in parton shower Monte Carlo programs are also available [39–41].

NNLO QCD corrections have been computed in the $m_t \rightarrow \infty$ limit in refs. [42–46]. The calculation of ref. [46] has been combined with results including the top quark mass dependence as far as available in ref. [47], and the latter has been supplemented by soft gluon resummation in ref. [48].

The scale uncertainties at NLO are still at the 10% level, while they are decreased to about 5% when including the NNLO corrections. The uncertainties due to the chosen top mass scheme have been assessed in ref. [38].

Analytic approximations for the top quark mass dependence of the two-loop amplitudes in the NLO calculation have been studied in refs. [49–53]. Complete analytic results in the high energy limit have been presented in ref. [54]; the latter have been combined with the full NLO results in the regions where they are more appropriate in ref. [55].

The effects of operators within an Effective Field Theory (EFT) description of Higgs boson pair production have been studied at LO QCD in refs. [22, 56–64] and at NLO in the $m_t \rightarrow \infty$ limit in refs. [65–67], including also CP-violating operators [68]. EFT studies at NNLO in the $m_t \rightarrow \infty$ limit are also available [69]. In ref. [70] for the first time the full NLO QCD corrections have been combined with an EFT approach to study BSM effects.

It is well known that Higgs boson pair production in gluon fusion is a process where delicate cancellations occur between contributions containing the trilinear Higgs coupling and box-type contributions not containing the trilinear coupling. While the destructive interference between these contributions is usually seen as a curse leading to small cross sections, it can be turned into a virtue when analyzing the shapes of distributions, as for example the di-Higgs invariant mass distribution m_{hh} , because even small anomalous couplings can lead to characteristic shape changes. Therefore it is important to investigate in which way the shapes are influenced by a certain configuration in the coupling parameter space.

The idea of a shape analysis has been pursued already in various ways based on LO studies, see e.g. refs. [58–61, 71, 72]. In ref. [61], a cluster analysis is proposed to define 12 benchmark points in a 5-dimensional non-linear EFT parameter space which result from clusters of “similar” shapes. The similarity measure in this case is based on a binned likelihood ratio using LO predictions for the observables m_{hh} , $\cos\theta^*$ and $p_{T,h}$. In ref. [70] it was analyzed how the m_{hh} and $p_{T,h}$ distributions change when going from LO to NLO for the benchmark points defined in ref. [61].

As a function of the 5-dimensional coupling parameter space, the m_{hh} distribution can have a few characterising features, such as an enhanced low- m_{hh} region, a double peak, a single peak or an enhanced tail. Some of these features can be attributed rather easily to a certain anomalous coupling, for example, an enhanced low- m_{hh} region is naturally produced by large values of $|c_{hhh}|$. Other features of the m_{hh} -shape, like a double peak or a SM-like distribution, are harder to attribute to a certain coupling configuration, as there are a multitude of configurations leading to such shapes. This is also reflected in the cluster analysis proposed in ref. [71], where (a) very different coupling configurations can end up in the same cluster, and (b) the same cluster can contain shapes which “by eye” look quite different (for example “double peak” and “single peak”).

Therefore it is desirable to seek for alternative methods to extract information about the underlying parameter space from the shape of distributions in Higgs boson pair production. In this work we first classify the shapes of Higgs boson pair invariant mass distributions, calculated at full NLO, into four characteristic types. We visualise the underlying 5-dimensional EFT parameter space producing these shape types, projecting onto 2-dimensional subspaces. We also comment on the shape of the $p_{T,h}$ distribution. Then we refine the shape analysis, applying an unsupervised learning algorithm based on an autoencoder to identify patterns in the shapes of the m_{hh} distribution. We use the **KMeans** clustering algorithm from **scikit-learn** [73] and ask for a classification of the shapes into a certain number of clusters. One aim of this study is to offer an alternative to the cluster analysis proposed in refs. [61, 71, 72] and earlier work, another aim is to provide an analysis based on full NLO results. The present study allows us to associate certain shapes more globally with distinct regions in the parameter space, in this sense going beyond a benchmark point analysis. Nonetheless, to facilitate a future more quantitative analysis, for example a profile likelihood study, we identify new benchmark points, based on the cluster centers given by our procedure.

The application of machine learning techniques in high energy physics, in particular to constrain the EFT/new physics parameter space, has been brought forward already some time ago [74–77]. There are also successful applications in jet and top quark identification [78–89] and PDF fits [90]. Machine learning in new physics searches is mostly used for anomaly detection [91–94] and to improve the sensitivity to new physics, optimising the signal to background ratio [76, 77, 95–99].

The remainder of this work is structured as follows: in section 2 we explain the framework our data samples are based on. We define four different shape types for the m_{hh} distribution and visualise the parameter space underlying the predefined shape types. In section 3 we describe our cluster analysis based on unsupervised learning and show how

the clusters found by this procedure relate to the underlying parameter space. We also define seven new benchmark points, before we conclude.

2 Classification through predefined shape types

2.1 Parametrisation of anomalous couplings in the Higgs sector

As a starting point we use the effective Lagrangian in a non-linear Effective Field Theory (“Higgs Effective Field Theory, HEFT”) relevant for Higgs boson pair production, assuming CP conservation, up to order 4 in the chiral expansion [70, 100]:

$$\mathcal{L} \supset -m_t \left(c_t \frac{h}{v} + c_{tt} \frac{h^2}{v^2} \right) \bar{t}t - c_{hhh} \frac{m_h^2}{2v} h^3 + \frac{\alpha_s}{8\pi} \left(c_{ggh} \frac{h}{v} + c_{gghh} \frac{h^2}{v^2} \right) G_{\mu\nu}^a G^{a,\mu\nu}. \quad (2.1)$$

In the SM $c_t = c_{hhh} = 1$ and $c_{tt} = c_{ggh} = c_{gghh} = 0$. The chromomagnetic operator is absent in (2.1) because it contributes to $gg \rightarrow hh$ only at higher order in the chiral counting. The coefficients c_{ggh} and c_{gghh} are related in SMEFT (“SM Effective Field Theory”) [30, 101, 102], however in HEFT there is not necessarily a relation between the two parameters. To clarify the relation to the widely used SMEFT operators, we briefly comment on the SMEFT Lagrangian here. The dimension-6 terms relevant for $gg \rightarrow hh$ can be written as [59, 65]

$$\begin{aligned} \Delta\mathcal{L}_6 = & \frac{\bar{c}_H}{2v^2} \partial_\mu(\phi^\dagger\phi)\partial^\mu(\phi^\dagger\phi) + \frac{\bar{c}_u}{v^2} y_t(\phi^\dagger\phi\bar{q}_L\tilde{\phi}t_R + \text{h.c.}) - \frac{\bar{c}_6}{2v^2} \frac{m_h^2}{v^2} (\phi^\dagger\phi)^3 \\ & + \frac{\bar{c}_{ug}}{v^2} g_s(\bar{q}_L\sigma^{\mu\nu}G_{\mu\nu}\tilde{\phi}t_R + \text{h.c.}) + \frac{4\bar{c}_g}{v^2} g_s^2\phi^\dagger\phi G_{\mu\nu}^a G^{a,\mu\nu}. \end{aligned} \quad (2.2)$$

Relating the coefficients \bar{c}_i in eq. (2.2) to the couplings of the physical Higgs field h and comparing with the corresponding parameters of the chiral Lagrangian eq. (2.1), one finds, after a field redefinition of h to eliminate \bar{c}_H from the kinetic term,

$$c_t = 1 - \frac{\bar{c}_H}{2} - \bar{c}_u, \quad c_{tt} = -\frac{\bar{c}_H + 3\bar{c}_u}{2}, \quad c_{hhh} = 1 - \frac{3}{2}\bar{c}_H + \bar{c}_6, \quad (2.3)$$

$$c_{ggh} = 2c_{gghh} = 128\pi^2\bar{c}_g. \quad (2.4)$$

Note that, assuming an underlying weakly coupled gauge theory, dimension-6 operators involving field-strength tensors can only be generated through loop diagrams [103]. Their coefficients then come with an extra factor of $1/16\pi^2$. In this case, the coefficients \bar{c}_{ug} and \bar{c}_g in eq. (2.2) are counted as order $(1/16\pi^2)(v^2/\Lambda^2)$, while \bar{c}_H , \bar{c}_u and \bar{c}_6 are still of order v^2/Λ^2 . For more details about the difference between HEFT and SMEFT we refer to refs. [30, 70].

We produce our data using the differential distributions calculated in ref. [70], parametrised in terms of coefficients A_i for each coupling combination occurring in the

(differential) NLO cross section, which allows for a fast evaluation:

$$\begin{aligned}
 \frac{d\sigma}{dm_{hh}} = & A_1 c_t^4 + A_2 c_{tt}^2 + A_3 c_t^2 c_{hhh}^2 + A_4 c_{ggh}^2 c_{hhh}^2 + A_5 c_{ggh}^2 + A_6 c_{tt} c_t^2 + A_7 c_t^3 c_{hhh} \\
 & + A_8 c_{tt} c_t c_{hhh} + A_9 c_{tt} c_{ggh} c_{hhh} + A_{10} c_{tt} c_{ggh} + A_{11} c_t^2 c_{ggh} c_{hhh} + A_{12} c_t^2 c_{ggh} \\
 & + A_{13} c_t c_{hhh}^2 c_{ggh} + A_{14} c_t c_{hhh} c_{ggh} + A_{15} c_{ggh} c_{hhh} c_{ggh} + A_{16} c_t^3 c_{ggh} + A_{17} c_t c_{tt} c_{ggh} \\
 & + A_{18} c_t c_{ggh}^2 c_{hhh} + A_{19} c_t c_{ggh} c_{ggh} + A_{20} c_t^2 c_{ggh}^2 + A_{21} c_{tt} c_{ggh}^2 + A_{22} c_{ggh}^3 c_{hhh} \\
 & + A_{23} c_{ggh}^2 c_{ggh} .
 \end{aligned} \tag{2.5}$$

The coefficients A_i are evaluated in bins of width 20 GeV from 250 GeV to 1050 GeV, i.e. for 40 bins. They are available with ref. [70] as .csv tables, in units of fb/GeV, for $\sqrt{s} = 13, 14$ and 27 TeV. The median of the statistical uncertainties of the differential coefficients A_i does not exceed 3%, however in the bins beyond $m_{hh} \gtrsim 650$ GeV some A_i coefficients have uncertainties in the 20–30% range.

2.2 Definition of shape types

We distinguish four types of characteristic shapes for the Higgs boson invariant mass distribution m_{hh} :

1. Enhanced low m_{hh} region, constantly falling distribution as m_{hh} increases.
2. Double peak with peaks separated by more than 100 GeV.
3. Single peak near the $t\bar{t}$ production threshold at $m_{hh} \sim 346$ GeV.
4. Double peak with peaks separated by less than 100 GeV.

Examples of the four shape types are shown in figure 1. According to our classification the Standard Model shape is contained in distributions of kind 3. Certainly there is some arbitrariness in the definition of these shapes. For example, shapes of kind 4 would move to kind 1 or 3 for bin widths ≥ 100 GeV. However, the other three shape types are quite robust and would be clearly distinguishable experimentally.

Based on the parametrisation in eq. (2.5), the normalised differential cross section is computed for a 5-dimensional grid in the coupling parameter space and according to its behaviour is classified into one of the four shape types. For this purpose we wrote a function, called “analyzer” in the following, that checks the slopes of the distribution and puts it into the corresponding class. At this stage the shape classes are mutually exclusive. For each point in the coupling parameter space, we also consider the variations of the result in each bin due to inclusion of the statistical uncertainties on the coefficients A_i . If the shape obtained after these variations belongs to a different kind, we exclude that point from the data set. We find that for shape type 4 about 20% of points fall into this category and are therefore excluded, while for shape type 2 it is about 8%, and for types 1 and 3 it is less than 5%. Scale variations have not been included, as they tend to be rather uniform over the whole m_{hh} range [41, 70] and therefore would not significantly modify our shape analysis.

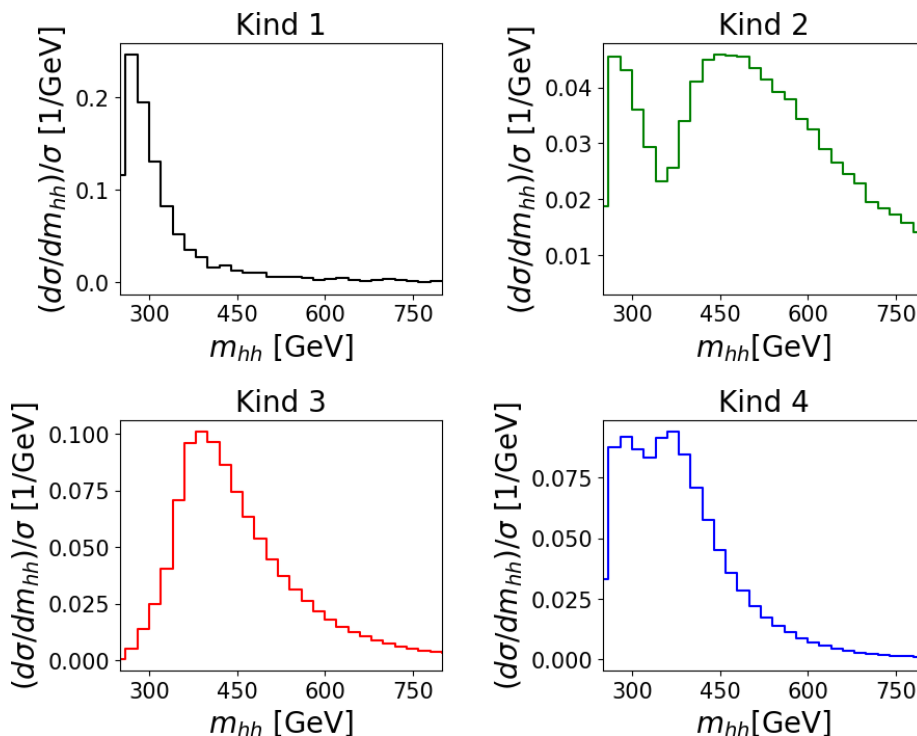


Figure 1. The four kinds of shapes defined in our analyzer to classify the m_{hh} distributions. The colours correspond to the colours shown in figures 2 to 6.

2.3 Classification of m_{hh} distributions

Our results for the $gg \rightarrow hh$ cross sections at NLO are produced for a centre-of-mass energy of $\sqrt{s} = 13$ TeV, using PDF4LHC15_nlo_100_pdfas [104] parton distribution functions interfaced via LHAPDF, along with the corresponding value for α_s . The masses have been set to $m_h = 125$ GeV, $m_t = 173$ GeV and the top quark width has been set to zero.

We study the differential cross section as a function of five anomalous couplings, varying them in the ranges specified below,

$$c_t \in [0.5, 1.5], c_{hhh} \in [-3, 8], c_{tt} \in [-3, 3], c_{ggh}, c_{gggh} \in [-0.5, 0.5]. \quad (2.6)$$

The ranges are motivated by current experimental constraints. For c_{hhh} we use a smaller range than suggested by experiment in order to focus more on the range where interesting shape features are present. In order to visualise the results, we project out 2-dimensional slices of the 5-dimensional parameter space, fixing the other three couplings to their SM values. This leads to a total of ten configurations. For each of these ten projections we generated a set of 10^6 parameter pairs. Feeding them through our analyzer we obtain the shape type produced by the given point in the coupling parameter space. The results are shown in figures 2–6. The white diamonds denote the Standard Model point in the parameter space. Scale variations are not included, as they are rather uniform over the whole m_{hh} range [41, 70] and therefore would not modify our shape analysis significantly.

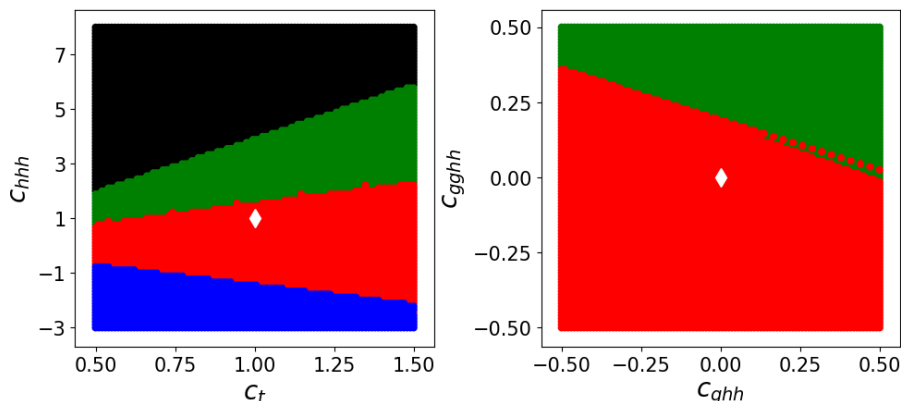


Figure 2. The parameter regions leading to each predefined shape type in the $c_t - c_{hhh}$ (left) and $c_{ggh} - c_{gggh}$ (right) parameter spaces. The black area denotes shapes of kind 1 (enhanced low m_{hh} region; green: kind 2 (well separated double peaks), red: kind 3 (SM-like), blue: kind 4 (close-by double peaks). The white diamonds mark the Standard Model point.

In figure 2 we display variations of the top quark Yukawa coupling c_t versus the trilinear Higgs coupling c_{hhh} (left) and the effective gluon-Higgs couplings, c_{ggh} versus c_{gggh} (right). In all the figures where two couplings are varied, the other three couplings are set to their SM values. It can be clearly seen that the shapes of kind 1, i.e. shapes with an enhanced low m_{hh} region (marked in black), are resulting from larger c_{hhh} values. The total cross section as a function of c_{hhh} is a parabola with a minimum around $c_{hhh} \approx 2.4$, while for $|c_{hhh}| \gtrsim 3$ and $c_t = 1$ the distribution is enhanced in the low m_{hh} region, where the triangle-type contributions dominate. Larger/smaller values of c_t shift this behaviour towards larger/smaller values of c_{hhh} because they enhance/decrease the box-type contributions. For shapes of kind 2, i.e. double peaks with a separation of more than 100 GeV (green), we find that such a shape can be produced for coupling values which are rather close to the SM values. Shapes of kind 3 (red) are SM-like. They only cover about one quarter of the $c_t - c_{hhh}$ plane. Shapes of kind 4 (blue) have a double peak separated by less than 100 GeV. For $c_{tt} = c_{ggh} = c_{gggh} = 0$, such structures only occur for negative values of c_{hhh} , over the whole allowed c_t range.

Considering variations of c_{ggh} versus c_{gggh} , shown in the right-hand panel of figure 2, we find only shapes of kind 2 (green) and SM-like shapes (red). The existence of kind 2 shapes means that a double peak structure could be produced solely by effective Higgs-gluon couplings, while keeping c_{hhh}, c_t and c_{tt} at their SM values. However, for the more likely case that c_{ggh} deviates only slightly from zero [105], and so does c_{gggh} , these couplings do not distort the SM shape significantly.

Variations of c_t versus c_{tt} and c_{hhh} versus c_{tt} are shown in figure 3. Varying only c_t and c_{tt} , the shapes remain mainly SM-like. A small area in the $c_t - c_{tt}$ plane however contains doubly peaked m_{hh} distributions, which thus can originate from anomalous top-Higgs couplings only, while the trilinear Higgs coupling remains fixed at its SM value.

Turning to c_{hhh} versus c_{tt} , displayed in the right-hand panel of figure 3, we find that for kind 1 and kind 4 shapes the parameter regions are split into two disconnected parts.

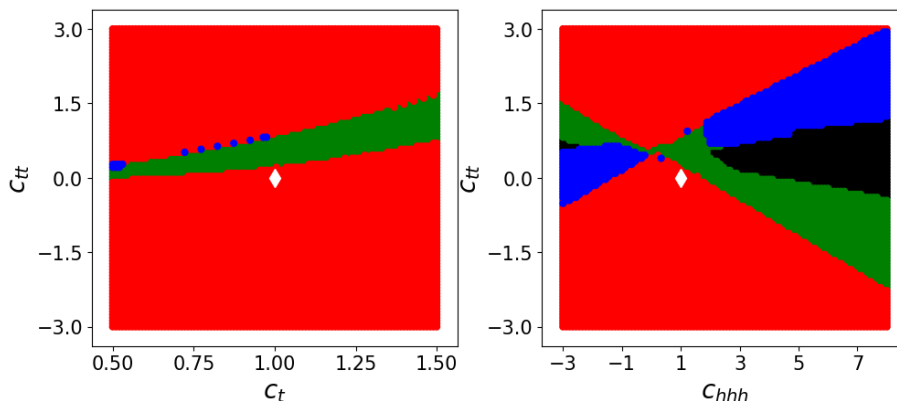


Figure 3. The parameter regions associated to each shape type in the $c_t - c_{tt}$ (left) and $c_{hhh} - c_{tt}$ (right) parameter spaces. For the colour code we refer to figure 1.

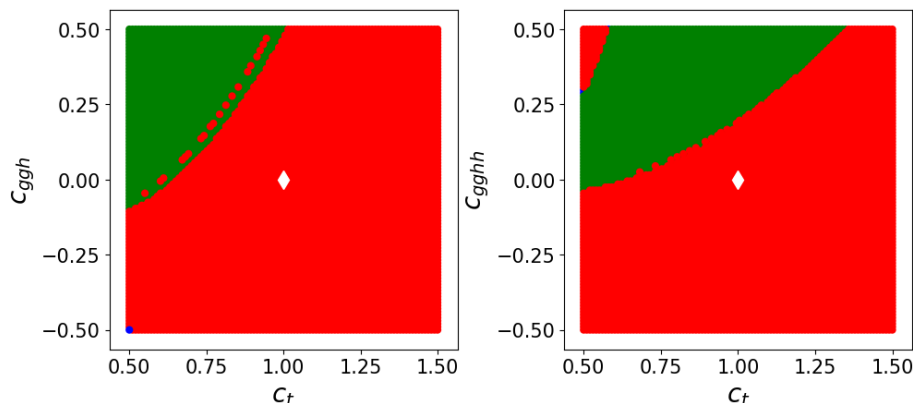


Figure 4. The parameter regions associated to each shape type in the $c_t - c_{ggh}$ (left) and $c_t - c_{gghh}$ (right) planes.

While shapes of kind 1 are favoured by large values of c_{hhh} , it becomes clear that large values of c_{tt} , also related to triangle-type diagrams, can counterbalance this effect, because the top right corner is not a parameter region producing shapes of kind 1. If both c_{hhh} and c_{tt} are large, it is more likely to produce a double peak structure with close-by peaks (kind 4, blue). Further we see that shapes of kind 2 (well separated double peak structure, green) can be produced by values of c_{tt} and c_{hhh} which are rather close to the SM values.

Figure 4 shows variations of c_t versus c_{ggh} (left) and c_t versus c_{gghh} (right). The parameter space is dominated by SM-like shapes (red), however double peaks can occur as well (green). We also see that c_{gghh} acts similarly to c_{ggh} in what concerns the shape.

For variations of c_{hhh} versus c_{ggh} , shown in figure 5 (left), all four shape types can occur. The parameter region related to kind 1 (enhanced low m_{hh} , black) is at high values of c_{hhh} as expected, and the kind 2 shapes (well separated double peak, green) can be seen as a transition from kind 1 to kind 3. Close-by double peaks (kind 4, blue) however are mostly associated to negative c_{hhh} values. Note that a similar pattern can be found in figure 2 (left). Variations of c_{hhh} versus c_{gghh} , shown in figure 5 (right), are similar in the overall behaviour, and again show that c_{gghh} and c_{ggh} have a similar impact on the shape.

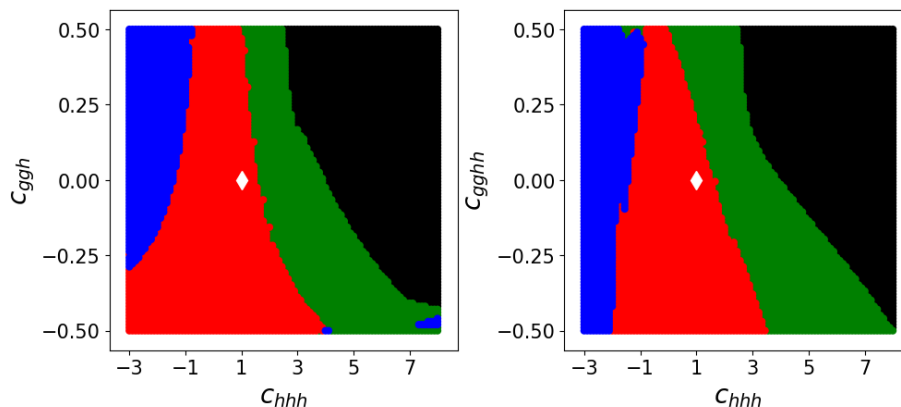


Figure 5. The parameter regions associated to each shape type in the $c_{hhh} - c_{ggh}$ (left) and $c_{hhh} - c_{gggh}$ (right) planes.

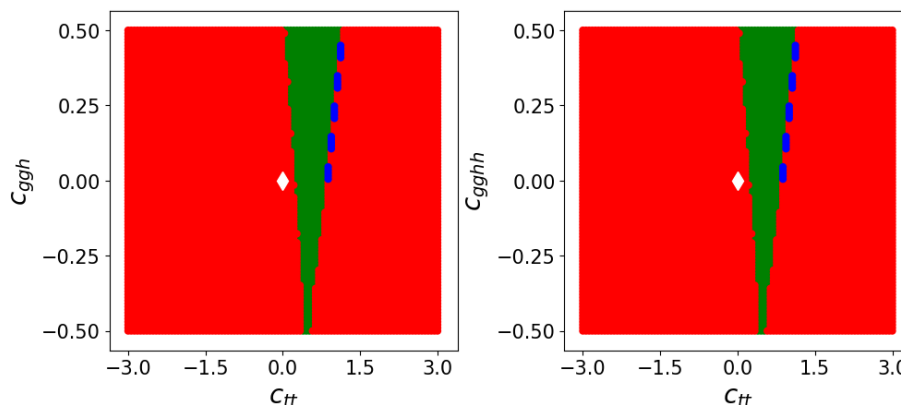


Figure 6. The parameter regions associated to each shape type in the $c_{tt} - c_{ggh}$ (left) and $c_{tt} - c_{gggh}$ (right) planes.

Figure 6 shows variations of c_{tt} versus c_{ggh} (left) and c_{tt} versus c_{gggh} (right). We observe that SM-like shapes (red) are preferred. However, doubly peaked structures are also possible for c_{tt} values not too far from the SM value ($c_{tt} = 0$). We also notice the similarity to figure 3 (left). The behaviour with respect to c_{gggh} is again similar.

Note that in SMEFT, c_{ggh} and c_{gggh} are related, so this behaviour would necessarily be the case. However we will see later that a shape classification algorithm based on unsupervised learning is able to detect shape differences which distinguish effects of c_{ggh} and c_{gggh} . An interesting feature is also that kind 1 (black) and kind 4 (blue) shapes appear only when we modify the value of c_{hhh} : for $c_{hhh} = 1$ shapes of kind 1 never occur, and shapes of kind 4 are very unlikely. Further, the kind 4 shapes tend to point to (moderately) negative values of c_{hhh} as long as c_{tt} is close to zero, as can be seen from figures 2, 3 and 5.

2.4 Classification of $p_{T,h}$ distributions

So far we have studied m_{hh} distributions, assuming that they are very well suited to study the sensitivity to shape changes induced by anomalous couplings. In order to verify that

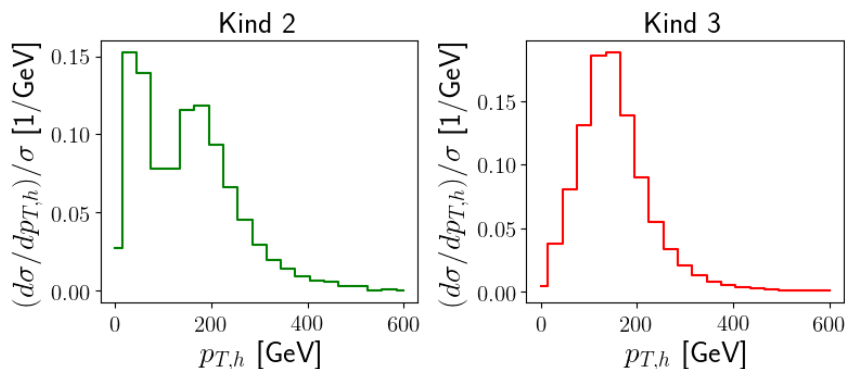


Figure 7. Examples of $p_{T,h}$ distributions with a single or double peak.

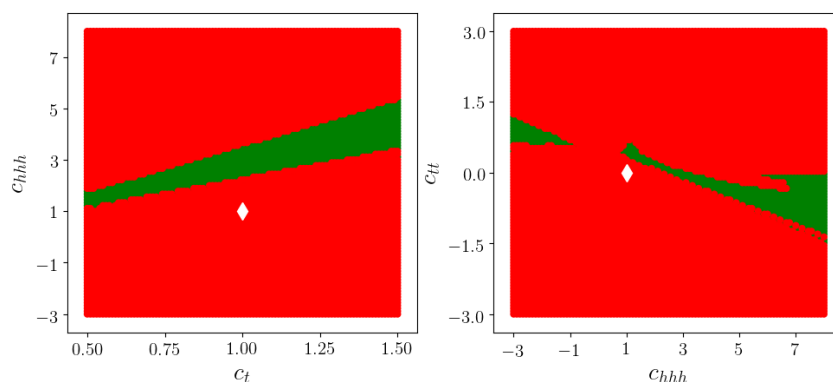


Figure 8. The parameter space associated to each shape type in the $c_t - c_{hh}$ and $c_{hh} - c_{tt}$ planes for the $p_{T,h}$ distribution.

we do not miss out interesting features in the transverse momentum distributions, we also present a study of the $p_{T,h}$ distributions, but only at LO, to assess the salient features. The main difference with respect to the m_{hh} case is that in the $p_{T,h}$ analysis we could identify only two kinds of clearly distinct shapes: single peak (SM-like, which we denote as ‘ $p_{T,h}$ kind 3’) and double peak, with peaks separated by at least 30 GeV (denoted as ‘ $p_{T,h}$ kind 2’). Examples of such $p_{T,h}$ shapes are shown in figure 7.

The parameter spaces leading to singly or doubly peaked shapes are shown in figure 8 for the $c_t - c_{hh}$ and $c_{hh} - c_{tt}$ configurations. The parameter region related to shapes with a well separated double peak (green) is similar to the m_{hh} case, as one can see comparing with figures 2 and 3. This indicates that the underlying parameter space leads to similar characteristics for the distributions differential in $p_{T,h}$ and m_{hh} , however the $p_{T,h}$ distribution is less sensitive than the m_{hh} distribution.

3 Classification and clustering by unsupervised learning

3.1 Unsupervised learning procedure

To assess the bias introduced by the definition of the four shape types, and to find a more flexible classification which can be extended easily to more than four shape types, we

approach the classification problem using unsupervised learning techniques. We construct a classification of the shapes of the m_{hh} distribution into distinct types, where we do not predefine what the types should look like. For this purpose we use an autoencoder to find common patterns in the data and thus achieve a compressed representation. The setup is implemented using KERAS [106] and TENSORFLOW [107]. As input data we use 30 bins of width 20 GeV for the normalised m_{hh} distributions. We train the network based on a set of 10^5 distributions, retaining 10% for the validation. The encoder architecture, i.e. the part compressing the array information, is composed of two dense layers with 20 nodes and a middle layer with 4 nodes, the latter defining the length of the array containing the compressed information. The decoder architecture, which reconstructs the original array from the compressed one, is composed of two dense layers of 20 nodes and an output layer of the length of the input array. We have also tried other encoder architectures, varying the number of nodes in the layers as well as the number of layers, and found that the results deteriorate for less than three layers. Adding more nodes had the tendency to lead to overfitting.

To test how stable our results are against variations of the training data set and the encoding procedure, and to reduce uncertainties, for example due to overfitting, we produced ten different autoencoder models. For each model we picked 10^4 random points from the training set for validation to start from different training and validation sets and a different initialization of the weights. We trained the autoencoder for each model over 10000 epochs using Adam [108] as optimizer and the root mean square error to define the loss function. Based on the trained autoencoder we applied the encoder models to the training and validation data to obtain two sets of compressed arrays for each of the ten models. The ten different encoded training data sets are then fed to a classification algorithm, where we employed the KMeans clustering algorithm from scikit-learn [73], asking for a classification into a given number of clusters. We tested classifications into four to eight clusters.

Asking the KMeans algorithm to find four clusters yielded the shape types shown in figure 9, the result of asking for seven clusters is shown in figure 10. The curves denote the cluster centres determined by the KMeans algorithm, for each of the ten encoder models, with a colour code as defined in table 1.

One can see from figures 9 and 10 that in the case of clustering into four shape types, cluster 2 contains shapes which vary substantially. In contrast, for seven shape types, the cluster centers obtained from the ten different encoder models are quite similar. Asking for 5–8 clusters we found that seven clusters seemed to be the optimal number to capture distinct shape features, while defining eight clusters did not lead to useful additional features but rather to the tendency to focus on local minima in the clustering space, while neglecting more global shape features.

The four clusters shown in figure 9 do only partly coincide with the ones defined in section 2.2. Shapes of kind 1, showing an enhanced m_{hh} region, as well as shapes of kind 3 (SM-like), were clearly identified. Shapes having a double peak were clustered together with shapes showing a shoulder. However, a cluster was formed which was not considered in the predefined types, containing shapes with an enhanced tail.

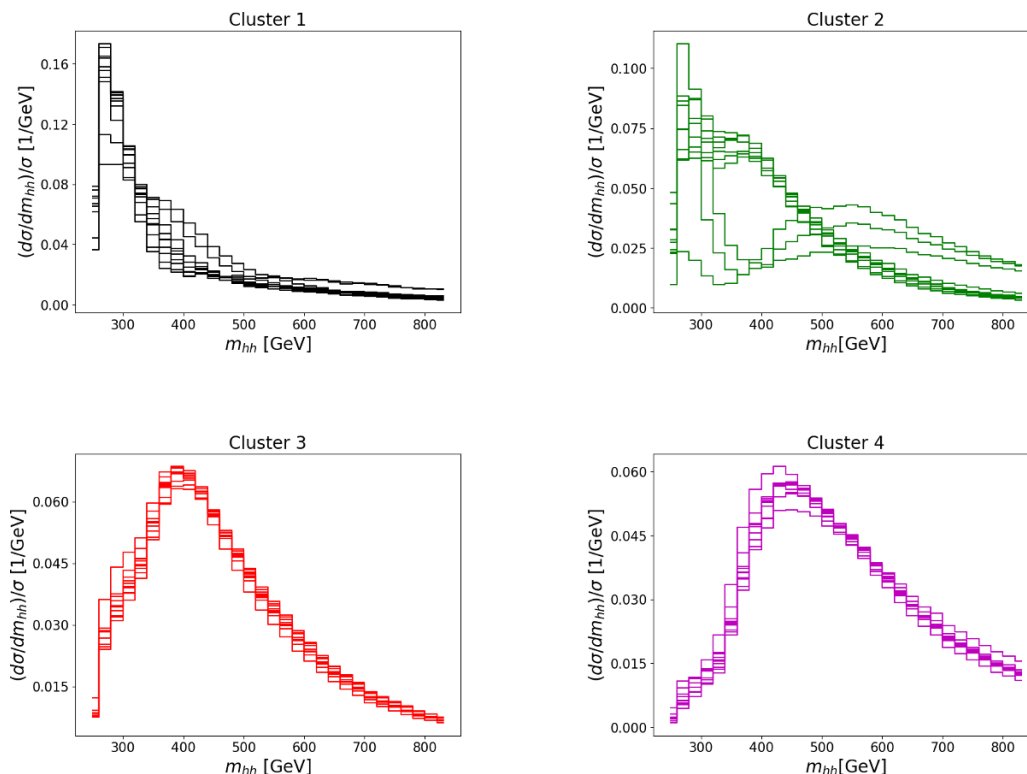


Figure 9. The clusters obtained by asking for a classification into four shape types. We show the cluster centres obtained from 10 different encoder models, in the colour code defined in table 1.

To combine the results from the ten clustering procedures, we adopted the “majority vote” method, i.e. for each of the ten clustering procedures a given point in the coupling parameter space gets a label (“vote”) corresponding to the cluster it belongs to. The final cluster assigned to that point is the one which collected the largest number of votes.

3.2 Parameter space underlying the clusters

In this section we show how the parameter space relates to the clusters if we ask for four or seven clusters. For each parameter configuration of our 5-dimensional grid, we plot the corresponding cluster type in figure 11 and figures 13 to 16. The colour codes are shown in figures 9 and 10, and are also listed in table 1. For clusters which are similar to the shape types defined in section 2.2, we should also find patterns similar to the ones shown in figures 2 to 6.

Comparing figure 11 (top row) with figure 2 (left), both showing variations of c_t versus c_{hhh} , we see that kind 1 shapes (black) are clearly identified. However, for both four and seven clusters the area for SM-like shapes got smaller, as the clustering algorithm also identifies features which were not considered in the predefined shapes. For example, the clustering into four clusters identifies shapes which are almost SM-like but have an enhanced tail (magenta), and the clustering into seven clusters in addition identifies shapes which are almost SM-like but have a shoulder (blue). Certainly we could have defined such features in our analyzer as well, but it is not that easy to define where the tail starts and what

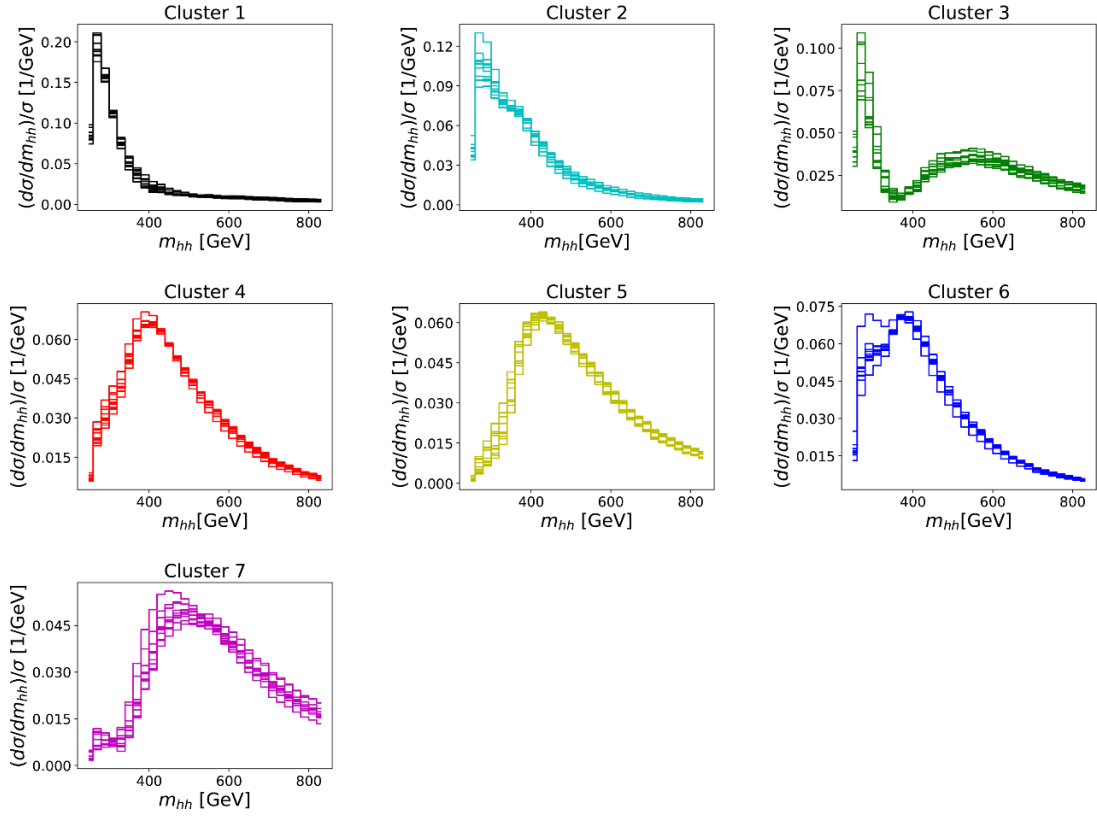


Figure 10. The clusters obtained by asking for a classification into seven shape types. The cluster centres obtained from 10 different encoder models are shown in the colour code defined in table 1.

Cluster	closest predefined type	colour
4 clusters		
1	kind 1 (enhanced low m_{hh})	black
2	double peak/enhanced tail	magenta
3	kind 3 (SM-like)	red
4	kind 4 (close-by double peaks)/shoulder	blue
7 clusters		
1	enhanced low m_{hh}	black
2	enhanced low m_{hh} , slowly falling or shoulder	cyan
3	enhanced low m_{hh} , second local maximum above $m_{hh} \simeq 2m_t$	green
4	SM-like	red
5	SM-like with enhanced tail	yellow
6	close-by double peaks or shoulder left	blue
7	no steep slope at low m_{hh} , enhanced tail	magenta

Table 1. Clusters and shape types with corresponding colour codes for the classification into four and seven clusters.

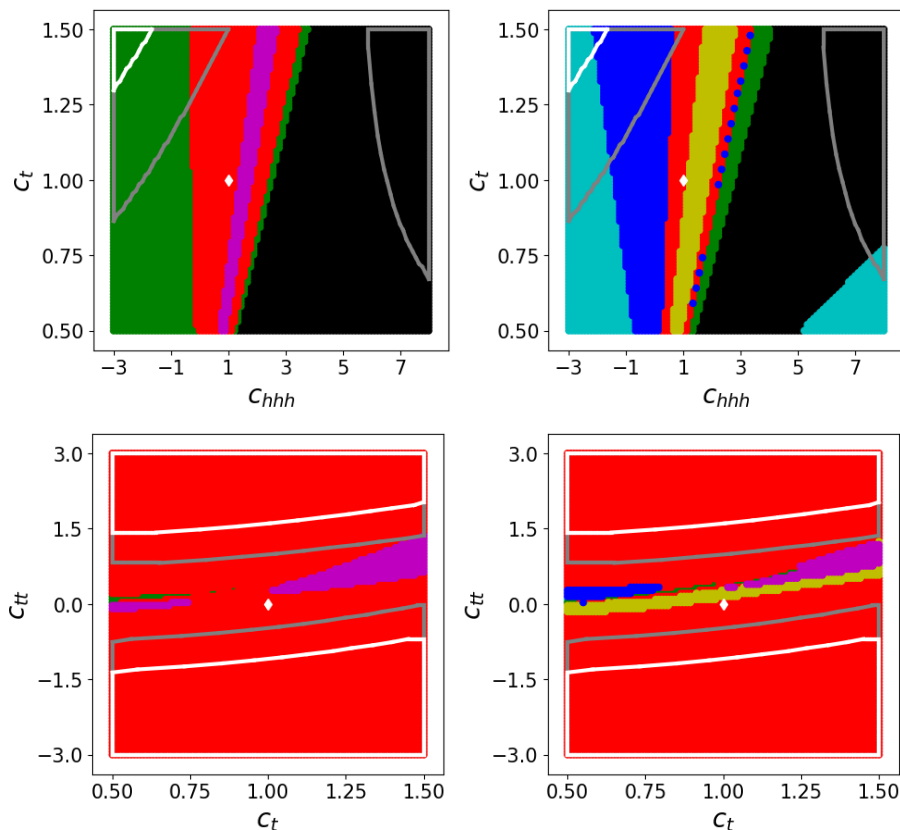


Figure 11. Shape types produced by variations of c_t versus c_{hhh} (top) and c_t versus c_{tt} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [1, 2]. The red areas denote SM-like shapes. The full colour code is given in table 1.

exactly should be considered as “enhanced”. Further, the figure clearly shows that small variations of c_{hhh} can easily distort the SM-like shape, while the shape is more robust against variations of c_t . Figure 11 (bottom row) shows c_t versus c_{tt} . We again see that variations c_t and c_{tt} mostly produce SM-like shapes. Why this is so can be understood from the behaviour of the coefficients A_i in eq. (2.5) which are relevant in these cases. For figure 11 (top row), only the coefficients A_1, A_3 and A_7 are relevant. As A_1 and A_7 have opposite signs and a different peak location, this can generate a rich shape structure. For figure 11 (bottom row), the coefficients A_2, A_6 and A_8 are relevant in addition to A_1, A_3 and A_7 . A_2 being the coefficient of c_{tt}^2 , it is dominant except for very small values of c_{tt} and leads to a SM-like shape. We also observe that c_{tt} has the tendency to enhance the total cross section, such that only a relatively small slice in c_{tt} is left after considering the bounds on the total cross section.

Figure 13 (top row) shows c_{hhh} versus c_{tt} , where we see that the interplay between c_{hhh} and c_{tt} can lead to all shape types. Comparing figure 13 (bottom row) with figure 2 (right), showing variations of c_{ggh} versus c_{gghh} , we observe that the unsupervised learning

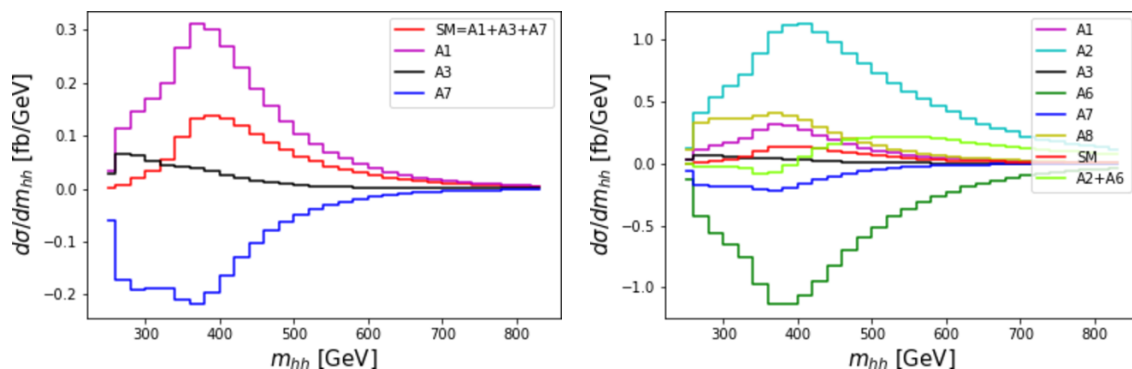


Figure 12. Contributions of the coefficients A_i in eq. (2.5) which are relevant for figure 11.

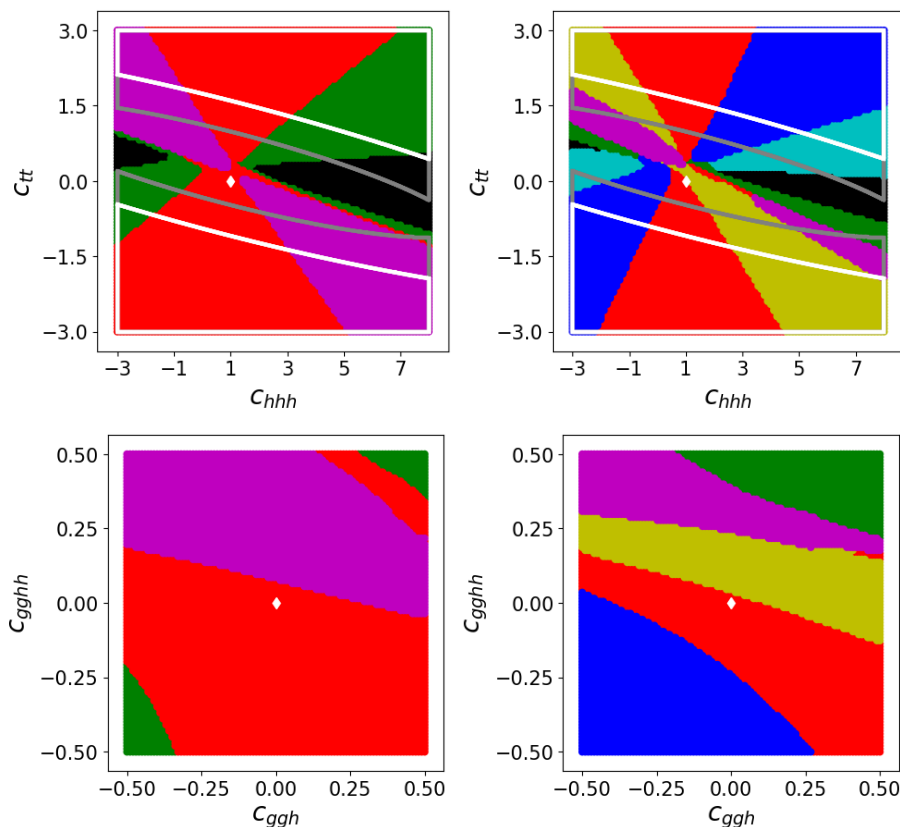


Figure 13. Shape types produced by variations of c_{hhh} versus c_{tt} (top) and c_{gghh} versus c_{ggh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [1, 2]. The red areas denote SM-like shapes. The full colour code is given in table 1.

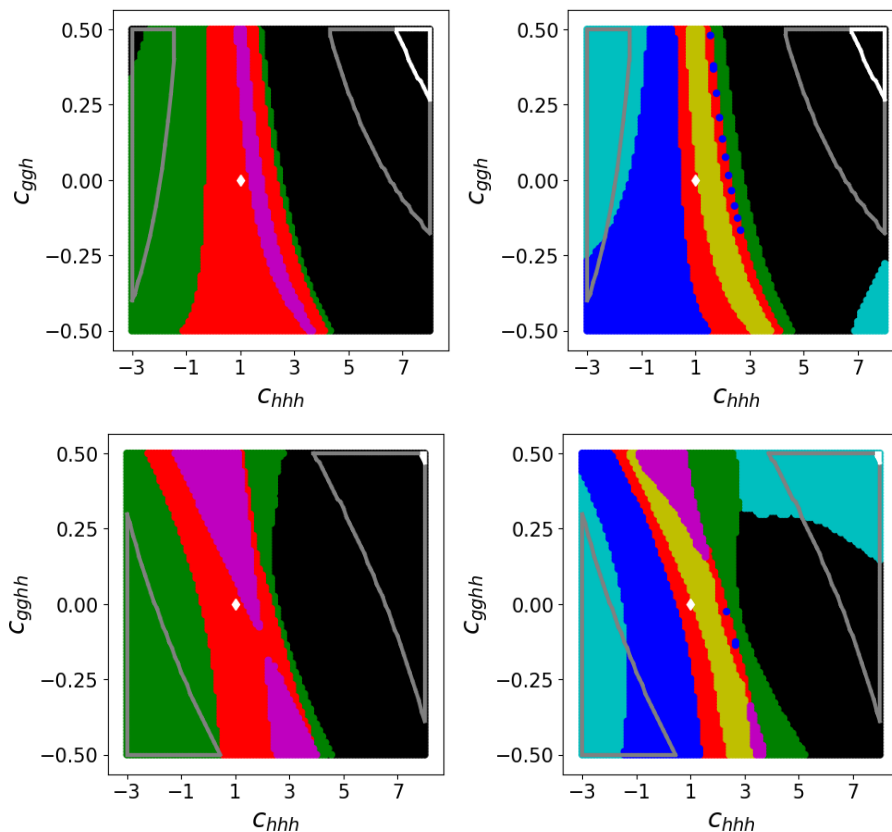


Figure 14. Shape types produced by variations of c_{hhh} versus c_{ggh} (top) and c_{gghh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [1, 2]. The red areas denote SM-like shapes. The full colour code is given in table 1.

algorithm with seven clusters distinguishes four shape types, showing that large values of c_{ggh} and c_{gghh} favour shapes with an enhanced tail (magenta) or/and a double peak (green), while negative values favour a shoulder on the left of the peak (blue). The limits on the total cross section do not exclude any parameter range in this panel.

A behaviour similar to the one in figure 11 can be seen in figure 14: as c_{hhh} varies the distribution goes through various shape types, while variations of c_{ggh} and c_{gghh} affect the shapes to less extent. Figure 14 also shows that a positive c_{gghh} value has the tendency to enhance the tail of the distribution.

Figure 15 shows c_{tt} versus c_{ggh} (top) and c_{tt} versus c_{gghh} (bottom). Compared to figure 6, the clustering into both four and seven clusters shows a better discrimination power between SM-like shapes and small deviations, for example due to an enhanced tail. We again see that c_{tt} has a larger impact on the total cross section than c_{ggh} or c_{gghh} .

Figure 16, showing the $c_t - c_{ggh}$ and $c_t - c_{gghh}$ parameter planes, can be compared to figure 4. Again, both the case of four and of seven clusters indicates that the unsupervised learning algorithm is able to distinguish better subtle influences on the shape than our method based on humanly classified shapes.

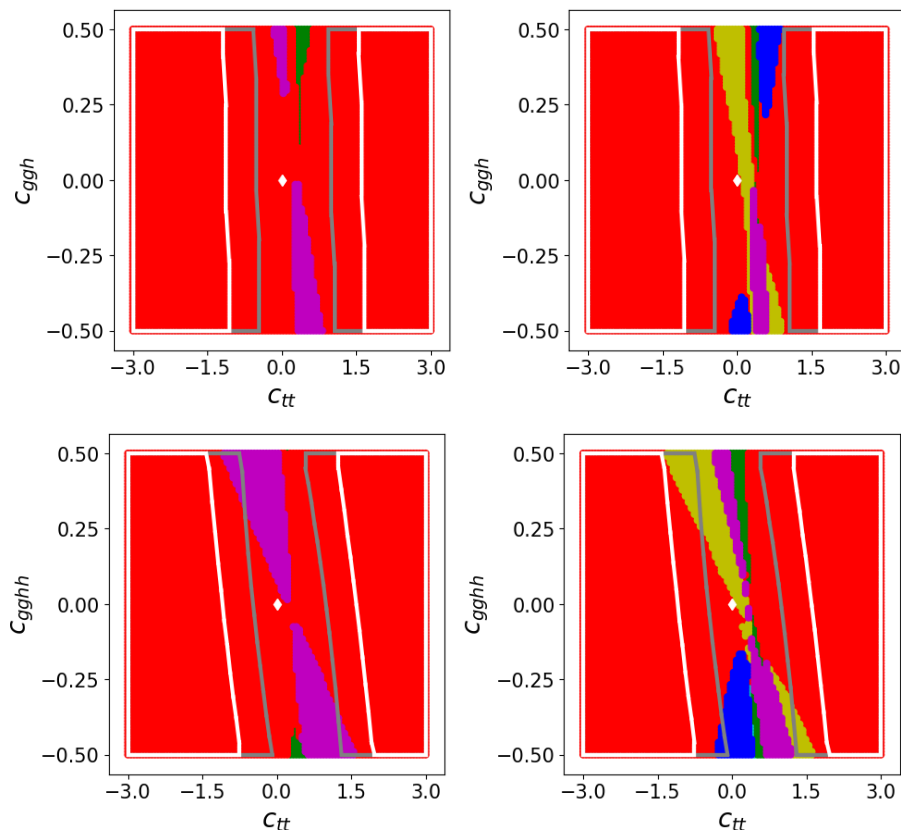


Figure 15. Shape types produced by variations of c_{tt} versus c_{ggh} (top) and c_{gghh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [1, 2]. The red areas denote SM-like shapes.

In figure 17, we compare results of our shape analysis produced with LO and NLO input data. We observe that NLO corrections can change the shape considerably and therefore are important for a shape analysis.

The results above have shown that the parameters c_{hhh} and c_{tt} have the largest influence on the shape. In SMEFT, c_{tt} is suppressed compared to c_t by one order of the large new physics scale [30]. Furthermore, SMEFT imposes the relation (2.4) between c_{ggh} and c_{gghh} . Using this relation and imposing that c_{tt} amounts to 5% of c_t , we obtain a 3-dimensional parameter space simulating the SMEFT situation, which is visualized in figure 18.

3.3 Identification of benchmark points

The determination of the cluster centres allows to identify NLO benchmark points which should be representative for each characteristic shape. As the cluster centres determined by the `KMeans` algorithm do not necessarily correspond to grid points of our input grids in parameter space, and as we work with normalised distributions to find the cluster centres, we determine the benchmark points based on the following procedure

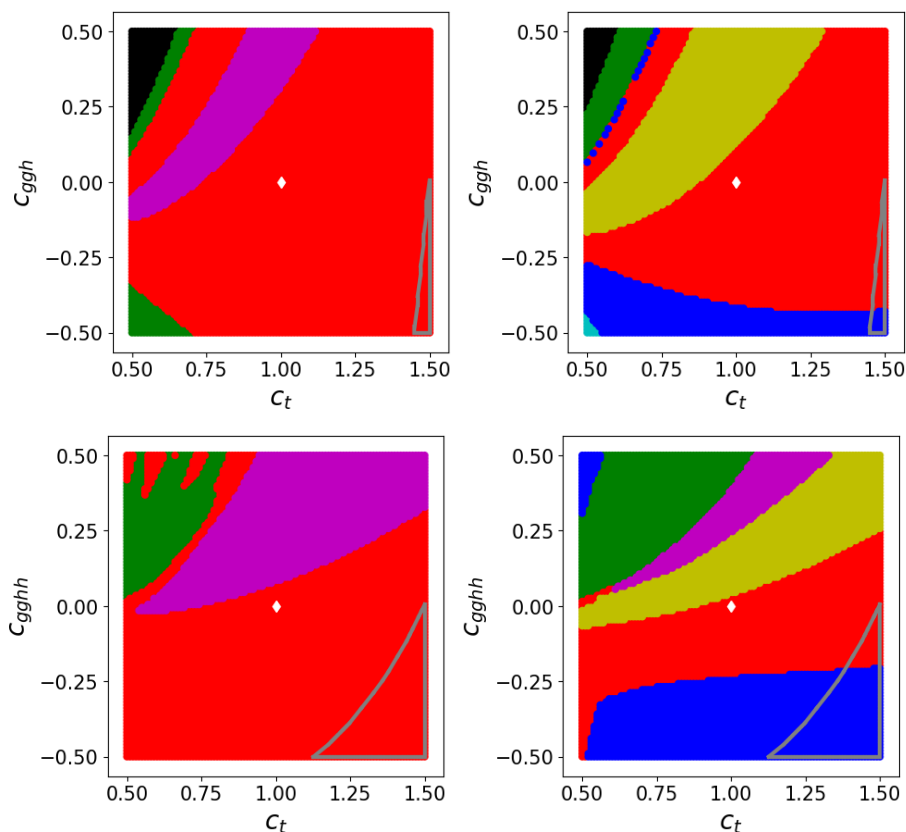


Figure 16. Shape types produced by variations of c_t versus c_{ggh} (top) and c_{gggh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{\text{SM}}$ and $22.2 \times \sigma_{\text{SM}}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [1, 2]. The red areas denote SM-like shapes.

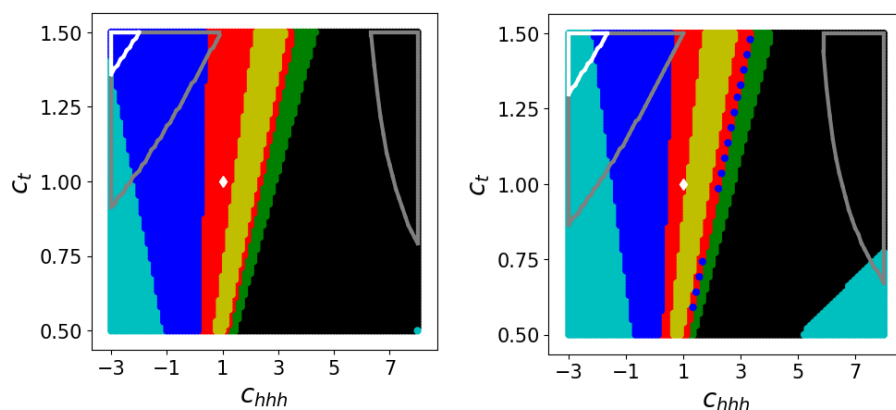


Figure 17. Comparison of LO and NLO results for shape types produced by variations of c_t versus c_{hhh} Left: LO, right: NLO.

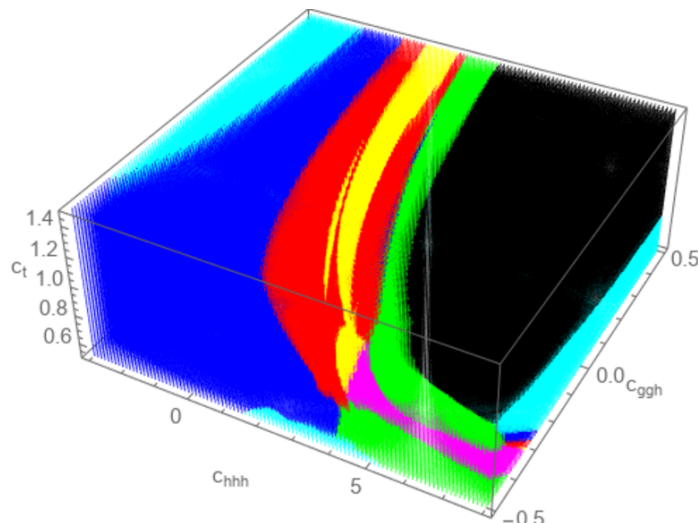


Figure 18. Three-dimensional visualisation of shape types produced by variations of c_t , c_{hhh} and c_{ggh} simulating the SMEFT situation. For c_{ggh} the value given by eq. (2.4) has been used, for c_t we used $c_{tt} = 0.05c_t$.

1. We identify the grid point in parameter space corresponding to a curve which is closest to the cluster centre, where the distance measure is the bin-wise geometric distance to the cluster center of the encoded distribution.
2. If the corresponding total cross section exceeds the limit of $6.9 \times \sigma_{\text{SM}}$ [1], we proceed to the curve which is the next-closest to the cluster center.
3. If several curves determined this way have an identical distance measure, we choose the one where the value of c_t is closer to the SM value (anticipating that the top-Higgs Yukawa coupling will be constrained increasingly well from other processes).
4. If after this procedure there are still several curves satisfying these criteria, from these curves we pick the one closest to the cluster center according to the Kullback–Leibler [109] distance measure, applied to the normalised distributions.

Following this procedure we find the benchmark points listed in table 3.3. In figure 19 we show the m_{hh} distributions corresponding to these benchmark points, at LO as well as at NLO.

4 Conclusions

The aim of this work was to provide more insight how certain configurations of anomalous couplings in the Higgs sector lead to a corresponding characteristic shape of the Higgs boson pair invariant mass distribution. For this purpose we employed the Lagrangian relevant to Higgs boson pair production as given in a non-linear Effective Field Theory framework, which contains five (potentially) anomalous couplings [70]. We produced data for the Higgs boson pair invariant mass distribution, based on a calculation which includes the NLO QCD corrections with full top quark mass dependence, varying all five coupling

benchmark	c_t	c_{hhh}	c_{tt}	c_{ggh}	c_{gggh}	σ_{NLO} [pb]	K-factor	ratio to SM
1	0.94	3.94	$-\frac{1}{3}$	0.5	$\frac{1}{3}$	182.50 ± 5.11	1.93	6.64
2	0.61	6.84	$\frac{1}{3}$	0.0	$-\frac{1}{3}$	135.63 ± 4.27	2.16	4.93
3	1.05	2.21	$-\frac{1}{3}$	0.5	0.5	109.24 ± 2.65	1.86	3.97
4	0.61	2.79	$\frac{1}{3}$	-0.5	$\frac{1}{6}$	50.44 ± 1.53	2.16	1.83
5	1.17	3.95	$-\frac{1}{3}$	$\frac{1}{6}$	-0.5	116.68 ± 6.25	1.63	4.24
6	0.83	5.68	$\frac{1}{3}$	-0.5	$\frac{1}{3}$	145.37 ± 8.25	2.19	5.29
7	0.94	-0.10	1	$\frac{1}{6}$	$-\frac{1}{6}$	96.69 ± 1.45	2.29	3.52

Table 2. NLO benchmark points derived from the cluster centers as described in the text.

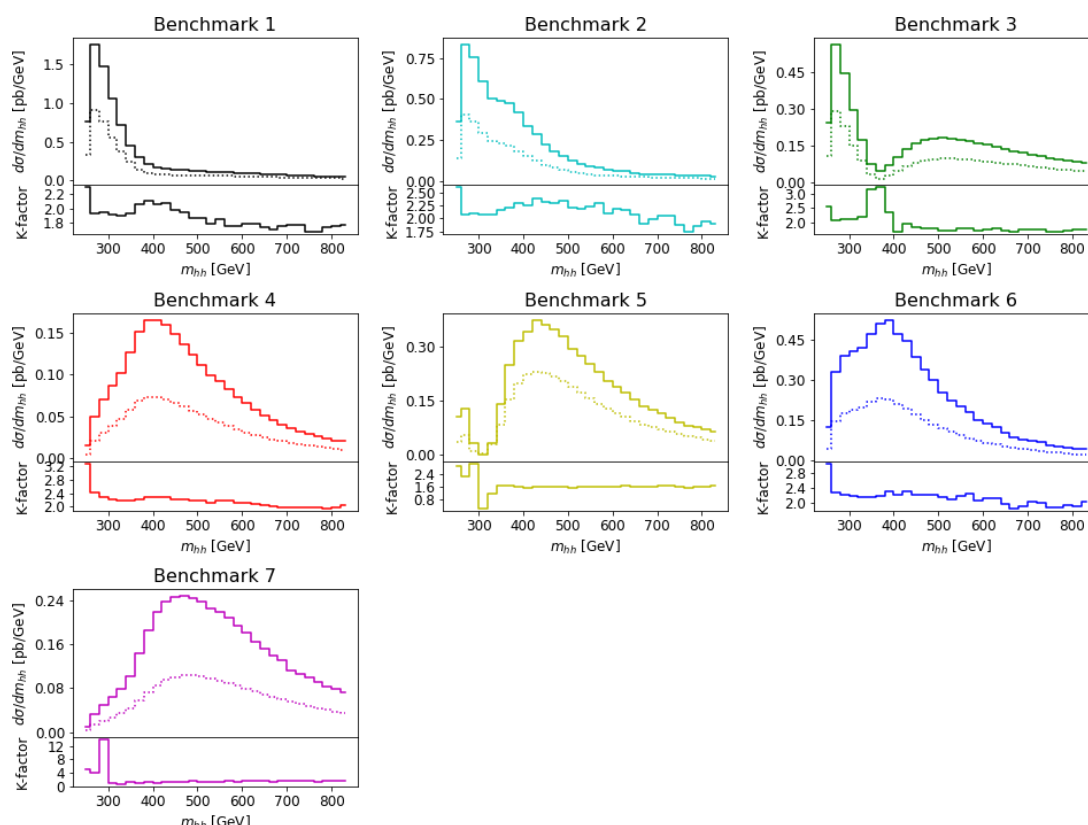


Figure 19. Higgs boson pair invariant mass distributions corresponding to the benchmark points listed in table 3.3. The solid curves denote the NLO result, the dotted curves the LO result. The lower panels show the K-factor, defined as $d\sigma_{\text{NLO}}/d\sigma_{\text{LO}}$.

parameters by finite steps, thus producing a dense grid of data. Then we defined four characteristic shape types for the m_{hh} distribution and visualised the parameter space leading to these shape types. To this aim we projected onto all possible two-dimensional slices of the parameter space, keeping the remaining parameters at their Standard Model values. We also considered $p_{T,h}$ distributions for a shape analysis, however we found that the m_{hh} distribution is more sensitive to shape changes induced by anomalous couplings.

Further, we tested an unsupervised learning approach to classify shapes. We produced 10^5 distributions, trained a neural network based on an autoencoder to extract common shape features and tried to find the number of shape clusters which optimally catches different shape characteristics. Our study demonstrated that some shape features, like an enhanced tail or a shoulder in the m_{hh} distribution, were caught very well by this procedure, and provided more insight about the underlying parameter space leading to such features than the analysis based on predefined shape classes. While machine learning is not essential to define shape clusters, it has the advantage of being easily extendible to a different number of shape types, different binnings or other observables, and of minimising the human bias compared to other shape analysis methods.

The shape analysis revealed that the Standard-Model-like shape is quite stable against variations of c_t , c_{ggh} and c_{gghh} , as long as $c_{hhh} = 1$, while deviations of c_{hhh} from the SM value show a rich shape changing pattern. We also found that small deviations of c_{tt} from zero are very likely to produce a doubly peaked structure in the m_{hh} distribution, while SM-like shapes dominate again as c_{tt} moves further away from zero. However, as c_{tt} leads to a rather fast increase of the total cross section, the shape analysis in combination with the limits on the total cross section allows to put constraints on c_{tt} . This is an interesting feature because, in contrast to c_t and c_{ggh} , c_{tt} cannot be constrained directly from single Higgs boson processes. Further, an enhanced tail or a shoulder of the m_{hh} distribution are likely to be produced by nonzero values of c_{gghh} , however the influence of the effective Higgs-gluon couplings on the shape is milder than the one of c_{hhh} and c_{tt} .

We also provide seven benchmark points, based on the full NLO calculation and taking into account current experimental constraints, which lead to the characteristic shapes as represented by our cluster centers.

The method can also be applied to other processes where anomalous couplings introduce characteristic shape changes to differential cross sections, and it can be extended to consider more than one distribution simultaneously.

Acknowledgments

We would like to thank Gerhard Buchalla, Alejandro Celis, Long Chen, Victor Diaz, Stephan Jahn, Stephen P. Jones, Matthias Kerner, Gionata Luisoni, Ludovic Scyboz and Johannes Schlenk for collaboration on the HH code, for useful discussions and for valuable comments on the manuscript. This research was supported in part by the COST Action CA16201 (‘Particleface’) of the European Union.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] ATLAS collaboration, *Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **800** (2020) 135103 [[arXiv:1906.02025](https://arxiv.org/abs/1906.02025)].

- [2] 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](#)] [[INSPIRE](#)].
- [3] M. McCullough, *An indirect model-dependent probe of the Higgs self-coupling*, *Phys. Rev. D* **90** (2014) 015001 [Erratum *ibid.* **D 92** (2015) 039903] [[arXiv:1312.3322](#)] [[INSPIRE](#)].
- [4] M. Gorbahn and U. Haisch, *Indirect probes of the trilinear Higgs coupling: $gg \rightarrow h$ and $h \rightarrow \gamma\gamma$* , *JHEP* **10** (2016) 094 [[arXiv:1607.03773](#)] [[INSPIRE](#)].
- [5] G. Degrossi, P.P. Giardino, F. Maltoni and D. Pagani, *Probing the Higgs self coupling via single Higgs production at the LHC*, *JHEP* **12** (2016) 080 [[arXiv:1607.04251](#)] [[INSPIRE](#)].
- [6] W. Bizon, M. Gorbahn, U. Haisch and G. Zanderighi, *Constraints on the trilinear Higgs coupling from vector boson fusion and associated Higgs production at the LHC*, *JHEP* **07** (2017) 083 [[arXiv:1610.05771](#)] [[INSPIRE](#)].
- [7] F. Maltoni, D. Pagani, A. Shivaji and X. Zhao, *Trilinear Higgs coupling determination via single-Higgs differential measurements at the LHC*, *Eur. Phys. J. C* **77** (2017) 887 [[arXiv:1709.08649](#)] [[INSPIRE](#)].
- [8] G.D. Kribs, A. Maier, H. Rzehak, M. Spannowsky and P. Waite, *Electroweak oblique parameters as a probe of the trilinear Higgs boson self-interaction*, *Phys. Rev. D* **95** (2017) 093004 [[arXiv:1702.07678](#)] [[INSPIRE](#)].
- [9] G. Degrossi, M. Fedele and P.P. Giardino, *Constraints on the trilinear Higgs self coupling from precision observables*, *JHEP* **04** (2017) 155 [[arXiv:1702.01737](#)] [[INSPIRE](#)].
- [10] J. Nakamura and A. Shivaji, *Direct measurement of the Higgs self-coupling in $e^+e^- \rightarrow ZH$* , *Phys. Lett. B* **797** (2019) 134821 [[arXiv:1812.01576](#)].
- [11] W. Kilian et al., *Multi-Higgs production and unitarity in vector-boson fusion at future hadron colliders*, [arXiv:1808.05534](#) [[INSPIRE](#)].
- [12] F. Maltoni, D. Pagani and X. Zhao, *Constraining the Higgs self-couplings at e^+e^- colliders*, *JHEP* **07** (2018) 087 [[arXiv:1802.07616](#)] [[INSPIRE](#)].
- [13] E. Vryonidou and C. Zhang, *Dimension-six electroweak top-loop effects in Higgs production and decay*, *JHEP* **08** (2018) 036 [[arXiv:1804.09766](#)] [[INSPIRE](#)].
- [14] M. Gorbahn and U. Haisch, *Two-loop amplitudes for Higgs plus jet production involving a modified trilinear Higgs coupling*, *JHEP* **04** (2019) 062 [[arXiv:1902.05480](#)] [[INSPIRE](#)].
- [15] ATLAS collaboration, *Constraint of the Higgs boson self-coupling from Higgs boson differential production and decay measurements*, [ATL-PHYS-PUB-2019-009](#) (2019).
- [16] ATLAS collaboration, *Constraints on the Higgs boson self-coupling from the combination of single-Higgs and double-Higgs production analyses performed with the ATLAS experiment*, [ATL-PHYS-PUB-2019-049](#) (2019).
- [17] W. Bizoń, U. Haisch and L. Rottoli, *Constraints on the quartic Higgs self-coupling from double-Higgs production at future hadron colliders*, *JHEP* **10** (2019) 267 [[arXiv:1810.04665](#)] [[INSPIRE](#)].
- [18] S. Borowka et al., *Probing the scalar potential via double Higgs boson production at hadron colliders*, *JHEP* **04** (2019) 016 [[arXiv:1811.12366](#)] [[INSPIRE](#)].
- [19] A. Falkowski and R. Rattazzi, *Which EFT*, *JHEP* **10** (2019) 255 [[arXiv:1902.05936](#)] [[INSPIRE](#)].
- [20] S. Chang and M.A. Luty, *The Higgs trilinear coupling and the scale of new physics*, [arXiv:1902.05556](#) [[INSPIRE](#)].

- [21] L. Di Luzio, R. Gröber and M. Spannowsky, *Maxi-sizing the trilinear Higgs self-coupling: how large could it be?*, *Eur. Phys. J. C* **77** (2017) 788 [[arXiv:1704.02311](#)] [[INSPIRE](#)].
- [22] S. Di Vita, C. Grojean, G. Panico, M. Riembau and T. Vantalon, *A global view on the Higgs self-coupling*, *JHEP* **09** (2017) 069 [[arXiv:1704.01953](#)] [[INSPIRE](#)].
- [23] J. Braathen and S. Kanemura, *On two-loop corrections to the Higgs trilinear coupling in models with extended scalar sectors*, *Phys. Lett. B* **796** (2019) 38 [[arXiv:1903.05417](#)] [[INSPIRE](#)].
- [24] P. Basler, S. Dawson, C. Englert and M. Mühlleitner, *Showcasing HH production: benchmarks for the LHC and HL-LHC*, *Phys. Rev. D* **99** (2019) 055048 [[arXiv:1812.03542](#)] [[INSPIRE](#)].
- [25] K.S. Babu and S. Jana, *Enhanced di-Higgs production in the two Higgs doublet model*, *JHEP* **02** (2019) 193 [[arXiv:1812.11943](#)] [[INSPIRE](#)].
- [26] A. Adhikary et al., *Revisiting the non-resonant Higgs pair production at the HL-LHC*, *JHEP* **07** (2018) 116 [[arXiv:1712.05346](#)] [[INSPIRE](#)].
- [27] I.M. Lewis and M. Sullivan, *Benchmarks for double Higgs production in the singlet extended standard model at the LHC*, *Phys. Rev. D* **96** (2017) 035037 [[arXiv:1701.08774](#)] [[INSPIRE](#)].
- [28] S. Dawson, C. Englert and T. Plehn, *Higgs Physics: It ain't over till it's over*, *Phys. Rept.* **816** (2019) 1 [[arXiv:1808.01324](#)] [[INSPIRE](#)].
- [29] M. Cepeda et al., *Report from Working Group 2*, *CERN Yellow Rep. Monogr.* **7** (2019) 221 [[arXiv:1902.00134](#)] [[INSPIRE](#)].
- [30] J. Alison et al., *Higgs boson pair production at colliders: status and perspectives*, in the proceedings of *Double Higgs Production at Colliders*, September 2, Batavia, U.S.A. (2019), [FERMILAB-CONF-19-468](#) [[arXiv:1910.00012](#)].
- [31] O.J.P. Eboli, G.C. Marques, S.F. Novaes and A.A. Natale, *Twin Higgs boson production*, *Phys. Lett. B* **197** (1987) 269 [[INSPIRE](#)].
- [32] E.W.N. Glover and J.J. van der Bij, *Higgs boson pair production via gluon fusion*, *Nucl. Phys. B* **309** (1988) 282 [[INSPIRE](#)].
- [33] T. Plehn, M. Spira and P.M. Zerwas, *Pair production of neutral Higgs particles in gluon-gluon collisions*, *Nucl. Phys. B* **479** (1996) 46 [*Erratum* *ibid.* **B 531** (1998) 655] [[hep-ph/9603205](#)] [[INSPIRE](#)].
- [34] S. Dawson, S. Dittmaier and M. Spira, *Neutral Higgs boson pair production at hadron colliders: QCD corrections*, *Phys. Rev. D* **58** (1998) 115012 [[hep-ph/9805244](#)] [[INSPIRE](#)].
- [35] F. Maltoni, E. Vryonidou and M. Zaro, *Top-quark mass effects in double and triple Higgs production in gluon-gluon fusion at NLO*, *JHEP* **11** (2014) 079 [[arXiv:1408.6542](#)] [[INSPIRE](#)].
- [36] S. Borowka et al., *Higgs boson pair production in gluon fusion at next-to-leading order with full top-quark mass dependence*, *Phys. Rev. Lett.* **117** (2016) 012001 [*Erratum* *ibid.* **117** (2016) 079901] [[arXiv:1604.06447](#)] [[INSPIRE](#)].
- [37] S. Borowka et al., *Full top quark mass dependence in Higgs boson pair production at NLO*, *JHEP* **10** (2016) 107 [[arXiv:1608.04798](#)] [[INSPIRE](#)].
- [38] J. Baglio et al., *Gluon fusion into Higgs pairs at NLO QCD and the top mass scheme*, *Eur. Phys. J. C* **79** (2019) 459 [[arXiv:1811.05692](#)] [[INSPIRE](#)].

- [39] G. Heinrich et al., *NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers*, *JHEP* **08** (2017) 088 [[arXiv:1703.09252](#)] [[INSPIRE](#)].
- [40] S. Jones and S. Kuttimalai, *Parton shower and NLO-matching uncertainties in Higgs boson pair production*, *JHEP* **02** (2018) 176 [[arXiv:1711.03319](#)] [[INSPIRE](#)].
- [41] G. Heinrich et al., *Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects*, *JHEP* **06** (2019) 066 [[arXiv:1903.08137](#)] [[INSPIRE](#)].
- [42] D. de Florian and J. Mazzitelli, *Two-loop virtual corrections to Higgs pair production*, *Phys. Lett. B* **724** (2013) 306 [[arXiv:1305.5206](#)] [[INSPIRE](#)].
- [43] D. de Florian and J. Mazzitelli, *Higgs boson pair production at next-to-next-to-leading order in QCD*, *Phys. Rev. Lett.* **111** (2013) 201801 [[arXiv:1309.6594](#)] [[INSPIRE](#)].
- [44] J. Grigo, K. Melnikov and M. Steinhauser, *Virtual corrections to Higgs boson pair production in the large top quark mass limit*, *Nucl. Phys. B* **888** (2014) 17 [[arXiv:1408.2422](#)] [[INSPIRE](#)].
- [45] J. Grigo, J. Hoff and M. Steinhauser, *Higgs boson pair production: top quark mass effects at NLO and NNLO*, *Nucl. Phys. B* **900** (2015) 412 [[arXiv:1508.00909](#)] [[INSPIRE](#)].
- [46] D. de Florian et al., *Differential Higgs boson pair production at next-to-next-to-leading order in QCD*, *JHEP* **09** (2016) 151 [[arXiv:1606.09519](#)] [[INSPIRE](#)].
- [47] M. Grazzini et al., *Higgs boson pair production at NNLO with top quark mass effects*, *JHEP* **05** (2018) 059 [[arXiv:1803.02463](#)] [[INSPIRE](#)].
- [48] D. De Florian and J. Mazzitelli, *Soft gluon resummation for Higgs boson pair production including finite M_t effects*, *JHEP* **08** (2018) 156 [[arXiv:1807.03704](#)] [[INSPIRE](#)].
- [49] R. Gröber, A. Maier and T. Rauh, *Reconstruction of top-quark mass effects in Higgs pair production and other gluon-fusion processes*, *JHEP* **03** (2018) 020 [[arXiv:1709.07799](#)] [[INSPIRE](#)].
- [50] R. Bonciani, G. Degrossi, P.P. Giardino and R. Gröber, *Analytical method for next-to-leading-order QCD corrections to double-Higgs production*, *Phys. Rev. Lett.* **121** (2018) 162003 [[arXiv:1806.11564](#)] [[INSPIRE](#)].
- [51] X. Xu and L.L. Yang, *Towards a new approximation for pair-production and associated-production of the Higgs boson*, *JHEP* **01** (2019) 211 [[arXiv:1810.12002](#)] [[INSPIRE](#)].
- [52] J. Davies, G. Mishima, M. Steinhauser and D. Wellmann, *Double-Higgs boson production in the high-energy limit: planar master integrals*, *JHEP* **03** (2018) 048 [[arXiv:1801.09696](#)] [[INSPIRE](#)].
- [53] J. Davies, F. Herren, G. Mishima and M. Steinhauser, *Real-virtual corrections to Higgs boson pair production at NNLO: three closed top quark loops*, *JHEP* **05** (2019) 157 [[arXiv:1904.11998](#)] [[INSPIRE](#)].
- [54] J. Davies, G. Mishima, M. Steinhauser and D. Wellmann, *Double Higgs boson production at NLO in the high-energy limit: complete analytic results*, *JHEP* **01** (2019) 176 [[arXiv:1811.05489](#)] [[INSPIRE](#)].
- [55] J. Davies et al., *Double Higgs boson production at NLO: combining the exact numerical result and high-energy expansion*, *JHEP* **11** (2019) 024 [[arXiv:1907.06408](#)] [[INSPIRE](#)].
- [56] R. Contino et al., *Anomalous couplings in double Higgs production*, *JHEP* **08** (2012) 154 [[arXiv:1205.5444](#)] [[INSPIRE](#)].

- [57] 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](#)].
- [58] C.-R. Chen and I. Low, *Double take on new physics in double Higgs boson production*, *Phys. Rev. D* **90** (2014) 013018 [[arXiv:1405.7040](#)] [[INSPIRE](#)].
- [59] 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](#)].
- [60] 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](#)] [[INSPIRE](#)].
- [61] A. Carvalho et al., *Higgs pair production: choosing benchmarks with cluster analysis*, *JHEP* **04** (2016) 126 [[arXiv:1507.02245](#)] [[INSPIRE](#)].
- [62] Q.-H. Cao, B. Yan, D.-M. Zhang and H. Zhang, *Resolving the degeneracy in single Higgs production with Higgs pair production*, *Phys. Lett. B* **752** (2016) 285 [[arXiv:1508.06512](#)] [[INSPIRE](#)].
- [63] Q.-H. Cao et al., *Double Higgs production at the 14 TeV LHC and a 100 TeV pp collider*, *Phys. Rev. D* **96** (2017) 095031 [[arXiv:1611.09336](#)] [[INSPIRE](#)].
- [64] J. de Blas, O. Eberhardt and C. Krause, *Current and future constraints on Higgs couplings in the nonlinear effective theory*, *JHEP* **07** (2018) 048 [[arXiv:1803.00939](#)] [[INSPIRE](#)].
- [65] R. Gröber, M. Mühlleitner, M. Spira and J. Streicher, *NLO QCD corrections to Higgs pair production including dimension-6 operators*, *JHEP* **09** (2015) 092 [[arXiv:1504.06577](#)] [[INSPIRE](#)].
- [66] R. Gröber, M. Mühlleitner and M. Spira, *Signs of composite Higgs pair production at next-to-leading order*, *JHEP* **06** (2016) 080 [[arXiv:1602.05851](#)] [[INSPIRE](#)].
- [67] 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](#)] [[INSPIRE](#)].
- [68] R. Gröber, M. Mühlleitner and M. Spira, *Higgs pair production at NLO QCD for CP-violating Higgs Sectors*, *Nucl. Phys. B* **925** (2017) 1 [[arXiv:1705.05314](#)] [[INSPIRE](#)].
- [69] D. de Florian, I. Fabre and J. Mazzitelli, *Higgs boson pair production at NNLO in QCD including dimension 6 operators*, *JHEP* **10** (2017) 215 [[arXiv:1704.05700](#)] [[INSPIRE](#)].
- [70] G. Buchalla et al., *Higgs boson pair production in non-linear effective field theory with full m_t -dependence at NLO QCD*, *JHEP* **09** (2018) 057 [[arXiv:1806.05162](#)] [[INSPIRE](#)].
- [71] A. Carvalho et al., *Analytical parametrization and shape classification of anomalous HH production in the EFT approach*, [arXiv:1608.06578](#) [[INSPIRE](#)].
- [72] A. Carvalho et al., *On the reinterpretation of non-resonant searches for Higgs boson pairs*, [arXiv:1710.08261](#) [[INSPIRE](#)].
- [73] <https://www.scikit-learn.org>
- [74] L. de Oliveira et al., *Jet-images — Deep learning edition*, *JHEP* **07** (2016) 069 [[arXiv:1511.05190](#)] [[INSPIRE](#)].
- [75] J. Brehmer, K. Cranmer, F. Kling and T. Plehn, *Better Higgs boson measurements through information geometry*, *Phys. Rev. D* **95** (2017) 073002 [[arXiv:1612.05261](#)] [[INSPIRE](#)].
- [76] J. Brehmer, K. Cranmer, G. Louppe and J. Pavez, *Constraining effective field theories with machine learning*, *Phys. Rev. Lett.* **121** (2018) 111801 [[arXiv:1805.00013](#)] [[INSPIRE](#)].

- [77] J. Brehmer, K. Cranmer, G. Louppe and J. Pavez, *A guide to constraining effective field theories with machine learning*, *Phys. Rev. D* **98** (2018) 052004 [[arXiv:1805.00020](#)] [[INSPIRE](#)].
- [78] D. Guest et al., *Jet flavor classification in high-energy physics with deep neural networks*, *Phys. Rev. D* **94** (2016) 112002 [[arXiv:1607.08633](#)] [[INSPIRE](#)].
- [79] G. Kasieczka, T. Plehn, M. Russell and T. Schell, *Deep-learning top taggers or the end of QCD?*, *JHEP* **05** (2017) 006 [[arXiv:1701.08784](#)] [[INSPIRE](#)].
- [80] K. Datta and A.J. Larkoski, *Novel jet observables from machine learning*, *JHEP* **03** (2018) 086 [[arXiv:1710.01305](#)] [[INSPIRE](#)].
- [81] G. Louppe, K. Cho, C. Becot and K. Cranmer, *QCD-aware recursive neural networks for jet physics*, *JHEP* **01** (2019) 057 [[arXiv:1702.00748](#)] [[INSPIRE](#)].
- [82] A.J. Larkoski, I. Moult and B. Nachman, *Jet substructure at the Large Hadron Collider: a review of recent advances in theory and machine learning*, *Phys. Rept.* **841** (2020) 1 [[arXiv:1709.04464](#)] [[INSPIRE](#)].
- [83] S. Macaluso and D. Shih, *Pulling out all the tops with computer vision and deep learning*, *JHEP* **10** (2018) 121 [[arXiv:1803.00107](#)] [[INSPIRE](#)].
- [84] S. Bollweg et al., *Deep-learning jets with uncertainties and more*, *SciPost Phys.* **8** (2020) 006 [[arXiv:1904.10004](#)] [[INSPIRE](#)].
- [85] A. Butter et al., *The machine learning landscape of top taggers*, *SciPost Phys.* **7** (2019) 014 [[arXiv:1902.09914](#)] [[INSPIRE](#)].
- [86] A. Butter, T. Plehn and R. Winterhalder, *How to GAN LHC events*, *SciPost Phys.* **7** (2019) 075 [[arXiv:1907.03764](#)] [[INSPIRE](#)].
- [87] E.A. Moreno et al., *JEDI-net: a jet identification algorithm based on interaction networks*, *Eur. Phys. J. C* **80** (2020) 58 [[arXiv:1908.05318](#)] [[INSPIRE](#)].
- [88] Y.-C.J. Chen, C.-W. Chiang, G. Cottin and D. Shih, *Boosted W/Z tagging with jet charge and deep learning*, [arXiv:1908.08256](#) [[INSPIRE](#)].
- [89] J. Chang, K. Cheung, J.S. Lee and J. Park, *Probing the trilinear Higgs boson self-coupling at the high-luminosity LHC via multivariate analysis*, *Phys. Rev. D* **101** (2020) 016004 [[arXiv:1908.00753](#)] [[INSPIRE](#)].
- [90] S. Carrazza and J. Cruz-Martinez, *Towards a new generation of parton densities with deep learning models*, *Eur. Phys. J. C* **79** (2019) 676 [[arXiv:1907.05075](#)] [[INSPIRE](#)].
- [91] J. Hajer, Y.-Y. Li, T. Liu and H. Wang, *Novelty detection meets collider physics*, [arXiv:1807.10261](#) [[INSPIRE](#)].
- [92] A. De Simone and T. Jacques, *Guiding new physics searches with unsupervised learning*, *Eur. Phys. J. C* **79** (2019) 289 [[arXiv:1807.06038](#)] [[INSPIRE](#)].
- [93] A. Andreassen, I. Feige, C. Frye and M.D. Schwartz, *JUNIPR: a framework for unsupervised machine learning in particle physics*, *Eur. Phys. J. C* **79** (2019) 102 [[arXiv:1804.09720](#)] [[INSPIRE](#)].
- [94] R.T. D’Agnolo and A. Wulzer, *Learning new physics from a machine*, *Phys. Rev. D* **99** (2019) 015014 [[arXiv:1806.02350](#)] [[INSPIRE](#)].
- [95] S. Chang, T. Cohen and B. Ostdiek, *What is the machine learning?*, *Phys. Rev. D* **97** (2018) 056009 [[arXiv:1709.10106](#)] [[INSPIRE](#)].
- [96] C. Englert, P. Galler, P. Harris and M. Spannowsky, *Machine learning uncertainties with adversarial neural networks*, *Eur. Phys. J. C* **79** (2019) 4 [[arXiv:1807.08763](#)] [[INSPIRE](#)].

- [97] A. Blance, M. Spannowsky and P. Waite, *Adversarially-trained autoencoders for robust unsupervised new physics searches*, *JHEP* **10** (2019) 047 [[arXiv:1905.10384](#)] [[INSPIRE](#)].
- [98] J. Brehmer et al., *Benchmarking simplified template cross sections in WH production*, *JHEP* **11** (2019) 034 [[arXiv:1908.06980](#)] [[INSPIRE](#)].
- [99] F.F. Freitas, C.K. Khosa and V. Sanz, *Exploring the standard model EFT in VH production with machine learning*, *Phys. Rev. D* **100** (2019) 035040 [[arXiv:1902.05803](#)] [[INSPIRE](#)].
- [100] G. Buchalla, O. Catà, A. Celis and C. Krause, *Note on anomalous Higgs-boson couplings in effective field theory*, *Phys. Lett. B* **750** (2015) 298 [[arXiv:1504.01707](#)] [[INSPIRE](#)].
- [101] L. Berthier and M. Trott, *Towards consistent electroweak precision data constraints in the SMEFT*, *JHEP* **05** (2015) 024 [[arXiv:1502.02570](#)] [[INSPIRE](#)].
- [102] I. Brivio, Y. Jiang and M. Trott, *The SMEFTsim package, theory and tools*, *JHEP* **12** (2017) 070 [[arXiv:1709.06492](#)] [[INSPIRE](#)].
- [103] C. Arzt, M.B. Einhorn and J. Wudka, *Patterns of deviation from the standard model*, *Nucl. Phys. B* **433** (1995) 41 [[hep-ph/9405214](#)] [[INSPIRE](#)].
- [104] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys. G* **43** (2016) 023001 [[arXiv:1510.03865](#)] [[INSPIRE](#)].
- [105] CMS collaboration, *Measurement and interpretation of differential cross sections for Higgs boson production at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* **792** (2019) 369 [[arXiv:1812.06504](#)] [[INSPIRE](#)].
- [106] <https://keras.io>
- [107] <https://www.tensorflow.org>
- [108] D.P. Kingma and J. Ba, *Adam: a method for stochastic optimization*, [arXiv:1412.6980](#) [[INSPIRE](#)].
- [109] S. Kullback, *Information theory and statistics*, John Wiley & Sons, U.S.A. (1978).