

Machine learning the trilinear and light-quark Yukawa couplings from Higgs pair kinematic shapes

Lina Alasfar,^a Ramona Gröber,^b Christophe Grojean,^{a,c} Ayan Paul^{a,c} and Zhuoni Qian^d

^a*Institut für Physik, Humboldt-Universität zu Berlin,
D-12489 Berlin, Germany*

^b*Dipartimento di Fisica e Astronomia ‘G. Galilei’, Università di Padova and INFN,
Sezione di Padova, I-35131 Padova, Italy*

^c*Deutsches Elektronen-Synchrotron DESY,
Notkestr. 85, 22607 Hamburg, Germany*

^d*School of Physics, Hangzhou Normal University,
Hangzhou, Zhejiang 311121, China*

E-mail: lina.alasfar@physik.hu-berlin.de, ramona.groeber@pd.infn.it,
christophe.grojean@desy.de, ayan.paul@desy.de, zhuoniqian@hznu.edu.cn

ABSTRACT: Revealing the Higgs pair production process is the next big challenge in high energy physics. In this work, we explore the use of interpretable machine learning and cooperative game theory for extraction of the trilinear Higgs self-coupling in Higgs pair production. In particular, we show how a topological decomposition of the gluon-gluon fusion Higgs pair production process can be used to simplify the machine learning analysis flow. Furthermore, we extend the analysis to include $q\bar{q} \rightarrow hh$ production, which is strongly suppressed in the Standard Model, to extract the trilinear Higgs coupling and to bound large deviations of the light-quark Yukawa couplings from the Standard Model values. The constraints on the rescaling of the trilinear Higgs self-coupling, κ_λ , and the rescaling of light-quark Yukawa couplings, κ_u and κ_d , at HL-LHC (FCC-hh) from single parameter fits are:

$$\kappa_\lambda = [0.53, 1.7] \quad ([0.97, 1.03])$$

$$\kappa_u = [-470, 430] \quad ([-58, 55])$$

$$\kappa_d = [-360, 360] \quad ([-26, 28])$$

We show that the simultaneous modification of the Yukawa couplings can dilute the constraints on the trilinear coupling significantly. We perform similar analyses for FCC-hh. We discuss some motivated flavourful new physics scenarios where such an analysis prevails.

KEYWORDS: Anomalous Higgs Couplings, SMEFT

ARXIV EPRINT: [2207.04157](https://arxiv.org/abs/2207.04157)

Contents

1	Introduction	1
2	New physics from Higgs pair production	4
3	Events simulation for HL-LHC and FCC-hh	7
4	Exploring higher dimensional kinematic distributions	10
4.1	Interpretable machine learning	10
5	The di-Higgs channel at future hadron colliders	12
5.1	Constraints on light-quark Yukawa couplings at the HL-LHC and FCC-hh	13
5.2	Constraints on Higgs trilinear self-coupling at the HL-LHC and FCC-hh	14
5.3	Two and three parameter constraints on Higgs couplings	15
5.4	Interpretation of Shapley values	18
5.5	Additional constraints on light-quark Yukawa couplings	22
6	Summary	23
A	Discussion of theoretical and systematic uncertainties	25
B	Light-quark Yukawa and self coupling at future lepton colliders	26
C	Shapley values	27

1 Introduction

As the LHC continues to probe the intricacies of the Higgs boson, its couplings to gauge bosons and third-generation fermions have been determined with an accuracy of around 10%–20% [1–3]. Other couplings such as the couplings of the Higgs bosons to the first two generations and the Higgs self-couplings remain more elusive. The high-luminosity LHC (HL-LHC) will contribute significantly to providing limits on the Higgs trilinear self-coupling and couplings of the first generation of quarks to the Higgs. The former can be measured directly at the HL-LHC in Higgs pair production, where it enters via top quark mediated triangle diagrams via gluon fusion [4–7]. This process also contains box diagrams that do not involve the trilinear Higgs self-coupling. Combining several final state channels, with the $b\bar{b}\gamma\gamma$ being the most sensitive, and combining the results of the ATLAS and CMS experiments a sensitivity reach of $0.5 < \kappa_\lambda = g_{hhh}/g_{hhh}^{\text{SM}} < 1.5$ at 68% CL can be estimated [8]. Alternatively, one can use single Higgs production, where the trilinear Higgs self-coupling enters via electroweak loops, leading to comparable limits on the trilinear

Higgs self-coupling [9–12] to the Higgs pair production. However, the sensitivity to the Higgs self-coupling from single Higgs measurements diminishes once other new physics effects modifying also the single Higgs rates are included [13, 14].

Measurements of the second generation quark Yukawa couplings are more challenging than those for the third generation. The couplings of the Higgs boson to the charm quarks may be limited to be smaller than a factor of 8.5 times their SM value [15, 16] from VH production with subsequent decay of the Higgs boson into charm quarks [17].¹ Other proposals for constraining the charm quark Yukawa coupling include exclusive Higgs boson decays to vector mesons [19, 20], which have been searched for by ATLAS and CMS [21, 22], Higgs + charm production [23], Higgs production in association with a jet [24], as well as the $VVcj$ channel [25]. The strange-quark coupling to the Higgs boson is somewhat more challenging to constrain. However, a future e^+e^- machine might allow reaching SM sensitivity by making use of strange tagging [26]. The first-generation Yukawa couplings are significantly more difficult to constrain and the aforementioned Higgs pair production can be used to constrain these couplings at future colliders [27].

For our analysis, we will adopt the framework of Standard Model effective field theory (SMEFT) which parameterizes potential new physics effects with all possible higher-dimensional operators under the assumption that the Higgs field has the same transformation rules under the SM symmetries. The leading effects in the Higgs sector are given by dimension six operators. The operators that modify the light-quark Yukawa couplings at the dimension six level are schematically denoted by

$$\frac{C_{q\phi}}{\Lambda^2} \phi^\dagger \phi \bar{Q}_L \phi q_R, \tag{1.1}$$

where Q_L are left-handed quark doublets, q_R right-handed quark fields and ϕ the Higgs doublet field. This operator also introduces couplings of the light quarks with two or three Higgs bosons. The contribution of the contact interaction of two Higgs bosons with two light quarks in quark annihilation to Higgs pairs is allowed in the SMEFT, and enhances the sensitivity reach on the concerned operator, comparing the Higgs pair production process to the single Higgs production process, as illustrated by figure 1. The advantage of the Higgs pair process can be understood primarily from the parton luminosity change when the hard scale goes from the single Higgs mass scale (125 GeV) to the di-Higgs threshold (300 GeV for $q\bar{q}$ and 400 GeV for gg -fusion process, read from the peak of m_{hh} distribution in ref. [27]). The parton luminosity reduction from single Higgs to di-Higgs process is correspondingly 0.44 (0.37) for the $u\bar{u}(d\bar{d})$ channels and 0.09 for the ggF channel.² This factor of five differences dictates the cross-section difference. As the $q\bar{q}$ process is proportional to the square of the coupling, and the crossing point of $C_{q\phi}$ is the square-root of the cross-section ratio between ggF and $q\bar{q}$, the difference between single Higgs and di-Higgs cross-sections is about a factor of $\sqrt{5} \sim 2$. In short, the $q\bar{q}$ -fusion process gains sensitivity at a higher scale

¹For the second-generation lepton sector, the analysis of [18] constrains the Higgs couplings to muons by a factor of 1.3 with respect to its SM value.

²For 100 TeV, with the parton luminosity estimated at the same typical scale, the corresponding numbers are now 0.97 (0.96) for the $u\bar{u}(d\bar{d})$ channels and 0.90 for the ggF channel.

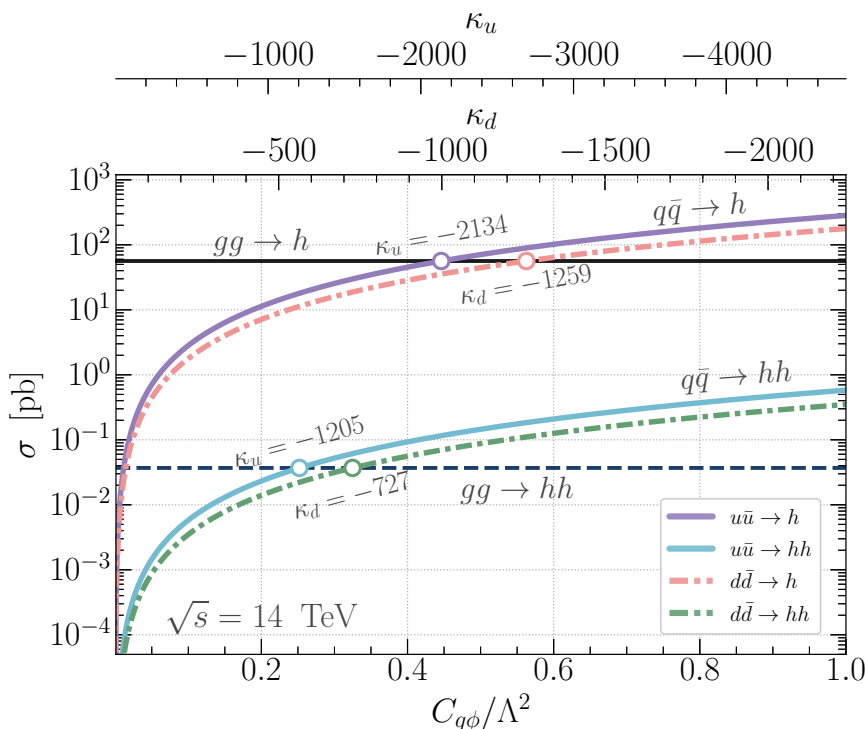


Figure 1. The production cross-section of single Higgs and di-Higgs at 14 TeV from the quark anti-quark annihilation $q\bar{q}hh$ as a function of the Wilson coefficients $C_{u\phi}/\Lambda^2$ and $C_{d\phi}/\Lambda^2$ versus the SM gluon fusion cross-sections (the horizontal solid lines for $gg \rightarrow h$ and $gg \rightarrow hh$). One can observe that for values of $\kappa_u = -1205$ (-2134) and $\kappa_d = -727$ (-1205) the $q\bar{q}hh$ channels become the dominant di-Higgs (single Higgs) production channels.

mostly through the milder drop of parton luminosity compared to the SM ggF processes. For the total cross-section, there is also a “dilution” factor from the modified total width of the Higgs for a final state of a specific (non-“light-jet”) decay channel.

Machine learning techniques are extremely useful in the detection and analyses of the Higgs boson pair production. Boosted decision trees (BDTs) are used both by the ATLAS and CMS collaboration to tag bottom quarks [8]. An increased sensitivity to the trilinear Higgs self-coupling can be achieved by employing neural networks or BDTs [28–32]. The present analysis aims to go a step further. While the interpretation of results obtained by machine learning approaches remains notoriously difficult when machine learning is used as a “black box”, the use of Shapley values, a measure derived from Coalition Game Theory, provides an interpretable analysis framework. We use this framework for the analysis of the $b\bar{b}\gamma\gamma$ final state from Higgs pair production, in a manner similar to what was previously proposed for the $b\bar{b}h$ process [33, 34]. This technique is used to extract the trilinear Higgs self-coupling and to probe the light-quark Yukawa couplings from the measurement of kinematics shapes in Higgs pair production significantly improving on the results from the cut-based analysis [27].

The paper is structured as follows: in section 2 we show the relevant EFT operators for this analysis and briefly discuss some concrete examples of how large light Yukawa coupling

modifications can be realised in models of new physics (NP) making use of the concept of aligned flavour violation (AFV) in order to avoid stringent flavour bounds. In section 3 we introduce the details of the simulation of the $pp \rightarrow hh \rightarrow b\bar{b}\gamma\gamma$ channel and its backgrounds. In section 4 we discuss the multivariate analysis and the interpretable machine learning approach we adopt. We present prospective results in section 5 for the HL-LHC and the Future Circular Collider (FCC-hh) with a center of mass energy of $\sqrt{s} = 100$ TeV and compare them to existing and projected bounds. In section 6 we summarize our main findings. A discussion on the theoretical uncertainties can be found in appendix A and a brief overview of prospective bounds on the light-quark Yukawa couplings and Higgs trilinear coupling at future lepton colliders can be found in appendix B.

2 New physics from Higgs pair production

The potential deformations of the SM in a model-independent manner can be studied through an EFT description parameterizing NP with higher-dimensional operators suppressed by some large energy scale Λ . A complete basis for the higher-dimensional operators has been given in refs. [35, 36]. In this work, we are interested in probing the Higgs trilinear and light-quark Yukawa couplings. Starting with the dimension six operators modifying the Higgs self-couplings, we see that they are given by

$$\mathcal{L} \supset \frac{C_{\phi\Box}}{\Lambda^2} (\phi^\dagger\phi)\Box(\phi^\dagger\phi) + \frac{C_{\phi D}}{\Lambda^2} (\phi^\dagger D_\mu\phi)^* (\phi^\dagger D^\mu\phi) + \frac{C_\phi}{\Lambda^2} |\phi^\dagger\phi|^3, \quad (2.1)$$

where ϕ denotes the Higgs-doublet which, in the unitary gauge, can be written as $\phi = 1/\sqrt{2}(0, v+h)^T$. It is common to quote the constraints on the Higgs couplings in terms of the rescaling with respect to the SM coupling prediction, typically denoted by κ :

$$\kappa = \frac{g_h}{g_h^{\text{SM}}}. \quad (2.2)$$

If the NP contributions do not generate new Lorentz structures, there is a possible translation between the Wilson coefficients in the Warsaw basis of the SMEFT, and the κ formalism used to denote the rescaling of SM couplings. In particular, taking the rescaling of the trilinear coupling, κ_λ , the translation is given by

$$\kappa_\lambda = 1 - \frac{2v^4 C_\phi}{m_h^2 \Lambda^2} + 3C_{\phi,\text{kin}}, \quad (2.3)$$

with

$$C_{\phi,\text{kin}} = \left(C_{\phi\Box} - \frac{1}{4}C_{\phi D} \right) \frac{v^2}{\Lambda^2}, \quad (2.4)$$

and $m_h = 125.1$ GeV. The Wilson coefficients $C_{\phi\Box}$ and $C_{\phi D}$ modify all the Higgs couplings and are in parts strongly constrained by electroweak precision observables (e.g. the T parameter constrains $C_{\phi D}$) [37–39]. Therefore, we set $c_{\phi,\text{kin}} = 0$ in what follows.

In SMEFT, new flavour structures can be introduced through dimension six operators which contain flavour indices. Focusing on the quark coupling to the Higgs boson we have, in particular,

$$\mathcal{L} \supset \frac{\phi^\dagger\phi}{\Lambda^2} \left((C_{u\phi})_{ij} \bar{q}_L^i \tilde{\phi} u_R^j + (C_{d\phi})_{ij} \bar{q}_L^i \phi d_R^j + h.c. \right), \quad (2.5)$$

with $i, j = 1 \dots 3$. Here u and d refer to the up- and down-type sectors respectively, and not the quarks themselves. The mass matrices of the up- and down-type quarks obtained from the Yukawa and the new SMEFT coupling are

$$\begin{aligned}
 M_{ij}^u &= \frac{v}{\sqrt{2}} \left(y_{ij}^u - \frac{1}{2} (C_{u\phi})_{ij} \frac{v^2}{\Lambda^2} \right), \\
 M_{ij}^d &= \frac{v}{\sqrt{2}} \left(y_{ij}^d - \frac{1}{2} (C_{d\phi})_{ij} \frac{v^2}{\Lambda^2} \right).
 \end{aligned}
 \tag{2.6}$$

Due to the modification of the mass matrix, the rotation matrices transforming the quark wavefunction to the mass eigenbasis will be modified with respect to the SM ones. The matrices $C_{q\phi}$ ($q = u, d$) are rotated by a new set of bi-unitary transformations $\mathcal{V}_{L/R}^{u/d}$ that rotate the quark wavefunctions to the mass eigenbasis. We can write $C_{q\phi}$ in terms of $\tilde{C}_{q\phi}$ which are now in the mass eigenbasis:

$$\begin{aligned}
 (C_{u\phi})_{ij} &= (\mathcal{V}_L^u)_{li} \left(\tilde{C}_{u\phi} \right)_{lm} (\mathcal{V}_R^u)_{mj}^\dagger, \\
 (C_{d\phi})_{ij} &= (\mathcal{V}_R^d)_{li}^\dagger \left(\tilde{C}_{d\phi} \right)_{lm} (\mathcal{V}_L^d)_{mj}.
 \end{aligned}
 \tag{2.7}$$

From this, we see that flavour off-diagonal couplings can be generated in the mass basis that drive flavour-changing neutral currents (FCNCs) which we shall discuss later in this section. The couplings of a single and pair of Higgs bosons to fermions in the mass eigenbasis can be defined as

$$\mathcal{L} \supset g_{h\bar{q}_i q_j} \bar{q}_i q_j h + g_{hh\bar{q}_i q_j} \bar{q}_i q_j h^2,
 \tag{2.8}$$

with

$$g_{h\bar{q}_i q_j} = \frac{m_{q_i}}{v} \delta_{ij} - \frac{v^2}{\Lambda^2} \frac{(\tilde{C}_{q\phi})_{ij}}{\sqrt{2}}, \quad g_{hh\bar{q}_i q_j} = -\frac{3}{2\sqrt{2}} \frac{v}{\Lambda^2} (\tilde{C}_{q\phi})_{ij}.
 \tag{2.9}$$

We have assumed that the matrices $\tilde{C}_{q\phi}$ are real, otherwise strong constraints from electric dipole moments apply [40, 41]. A similar relation exists for the rescalings of the quark Yukawa couplings κ_q , for the diagonal elements of $\tilde{C}_{q\phi}$,³

$$\kappa_{q_i} = 1 - \frac{v^3}{\sqrt{2} m_q} \frac{(\tilde{C}_{q\phi})_{ii}}{\Lambda^2},
 \tag{2.10}$$

with $m_u = 2.2 \text{ MeV}$ and $m_d = 4.7 \text{ MeV}$. The $hhq\bar{q}$ coupling, though being linearly related to the quark Yukawa coupling $hq\bar{q}$, is not a rescaling of any SM Higgs coupling. With this in mind, one can remain strictly within an EFT, like SMEFT, where the $SU(2) \times U(1)$ symmetry is linearly realized and links the rescaling of the quark Yukawa, κ_q , to the $hhq\bar{q}$ coupling through

$$g_{hhq\bar{q}}^{\text{linear-EFT}} = \frac{3}{2} \frac{\kappa_q - 1}{v} g_{hhq\bar{q}}^{\text{SM}}.
 \tag{2.11}$$

³For normalisation purposes, we choose to set $\Lambda = 1 \text{ TeV}$ throughout the remainder of this paper. We stress that it does not imply that the scale of new physics is at 1 TeV, but it fixes the normalisation of the Wilson coefficients.

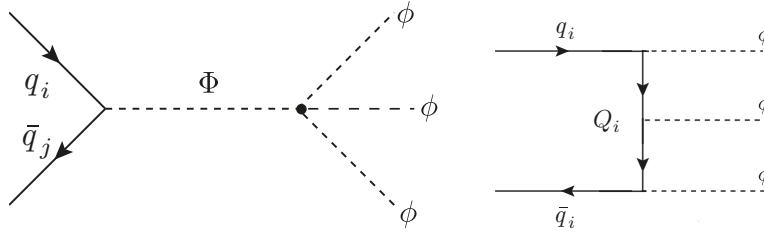


Figure 2. The SMEFT operator $\mathcal{O}_{q\phi}$ can be generated either by a new scalar Φ (left) or a new set of VLQs Q_i (right). When these fields are integrated out, they yield the aforementioned operator in the unbroken phase.

This relation will no longer hold once a non-linear EFT is used. Hence, some caution has to be exercised when the κ -formalism is used in multi-Higgs studies. For clarity we will denote \tilde{C}_i as C_i , i.e., assuming C_i to be in the mass basis, in the rest of the article.

As we have pointed out before, the construction we discussed can lead to contributions from FCNCs which are strongly constrained from low-energy measurements of flavour observables. The bounds are of order $|(C_{u\phi,d\phi})_{12}| \lesssim 10^{-5}\Lambda^2/v^2$ and $|(C_{u\phi,d\phi})_{13}| \lesssim 10^{-4}\Lambda^2/v^2$ and stem from $\Delta F = 2$ transitions [42, 43]. Given that FCNCs need to be suppressed, an effective way of realizing this suppression is by imposing minimal flavour violation (MFV) [44]. The assumption of MFV introduces a strong hierarchy amongst the Higgs couplings to quarks, due to the proportionality of the Wilson coefficients to the Yukawa couplings. Since we want to explore rather large modifications of the light-quark Yukawa couplings, very low values of Λ , the NP scale, or large Wilson coefficients in MFV models or both need to be assumed which render the validity of the EFT questionable. Furthermore, this would potentially generate conflicts with measurements of the third generation couplings to the Higgs boson. To escape the strong hierarchy in the Yukawa couplings imposed by assuming MFV, a less restrictive assumption about the flavour structure can be made which is called flavour alignment [45, 46]. In a more generalized form of this ansatz, it is assumed that the new physics contributions to the Yukawa matrices are aligned in flavour space and can be simultaneously diagonalized to remove the possibility of having flavour non-diagonal couplings that can generate FCNCs at the tree level [47–49].

There exist a handful of potential UV-complete models in which both light-quark Yukawa couplings and the Higgs trilinear couplings are simultaneously enhanced. In general, the operator of eq. (2.5) can be generated at tree-level by vector-like quarks (VLQs) or new scalars figure 2. For example, a model proposed in ref. [50] based on VLQs with the assumption of AFV where the mechanism of suppressing FCNCs is linked to flavour textures imposed by discrete symmetries. The original assumption of this model is excluded as the authors assumed that all the light quark-Higgs couplings to be similar in magnitude to the bottom Yukawa. One could still get a significant enhancement of the light-quark Yukawa couplings from VLQ masses of ~ 2 TeV, which is well above the bounds set by current direct searches excluding VLQ of masses $M_{\text{VLQ}} < 1.6$ TeV [51, 52] for the purely hadronic final state, and $M_{\text{VLQ}} < 1.2$ TeV for final states with leptons and jets [53]. Values of $M_{\text{VLQ}} \sim 2$ TeV would also be in accordance with electroweak precision observables. In

addition, the trilinear Higgs coupling could be modified by the inclusion of an additional scalar singlet as proposed in refs. [54–56].

Another concrete example of a model with enhanced light-quark Yukawa couplings is a two-Higgs-doublet model (2HDM) model proposed in refs. [57, 58]. This model shows a sub-class of AFV, known as spontaneous flavour violation [59]. Enhancements of the light-quark Yukawa couplings stem from the couplings of the second Higgs doublet to the quarks, $K_{u/d}$, which are made diagonal in the flavour space. The model can have either the up-type or the down-type couplings enhanced with respect to the SM values, while the couplings of the other type remain proportional to the SM ones in order to obtain the correct CKM matrix. The addition of the second doublet modifies the Higgs potential, and consequently, the Higgs self-coupling will be modified as well. Like any other 2HDM, the parameter space is rather large. Hence, bounds on the model depend on the part of the parameter space that is under consideration. For a small mass of the “heavy” Higgs H and large Yukawa coupling, K_d , flavour bounds dominate, while for a larger m_H , dijet searches [60–62] dominate due to the decay $H \rightarrow d\bar{d}$. Instead, the decay $H \rightarrow hh$ becomes dominant from smaller values of K_d and larger H mass, rendering in the regime $m_H < 2\text{ TeV}$ resonant di-Higgs searches [63, 64] as the potential discovery channel of such a scenario. Furthermore, in this region of the parameter space constraints can be derived from Zh [65] and ZZ [66, 67] searches. Lastly, for $m_H > 2\text{ TeV}$, the non-resonant Higgs pair production will become the dominant bound on the enhancement of the light-quark Yukawa couplings.

3 Events simulation for HL-LHC and FCC-hh

We consider the final state $b\bar{b}\gamma\gamma$ as this channel has the highest potential for Higgs pair searches [68]. One has the “clean” decay $h \rightarrow \gamma\gamma$ of one Higgs boson, with the other Higgs boson decaying to a $b\bar{b}$ quark pair with a large branching ratio $\sim 58\%$, where b -tagging capabilities for ATLAS and CMS are continuously improving. We consider the $b\bar{b}h$, $t\bar{t}h$, $b\bar{b}\gamma\gamma$ processes as the main sources of background for the hh signal. The details of the simulation for the $b\bar{b}h$ process can be found in ref. [33]. The events are generated at leading order (LO) and then scaled to NLO by K -factors, defined as the ratio of higher order cross section over its LO counterpart. The K -factors are taken from the corresponding literature for $t\bar{t}h$ [69], $b\bar{b}\gamma\gamma$ [70], Zh [71] and the remaining part of the $b\bar{b}h$ processes from [72]. The Higgs particles are further decayed to $\gamma\gamma$ following the Higgs cross-section working group recommendations [73]. The parton-level results are generated using `MadGraph_aMC@NLO` [74], showered using `Pythia 8.3` [75], and a subsequent detector simulation is done using `Delphes 3` [76]. To be inclusive and to explore the capabilities and importance of the full detector coverage, no generator-level cuts are applied on these processes except for the $b\bar{b}\gamma\gamma$ QCD-QED background processes to avoid divergences. These minimal generator-level cuts are

$$\begin{aligned} Xp_T^b &> 20\text{ GeV}, \\ \text{generator level cuts: } \quad \eta_\gamma &< 4.2, \quad \Delta R_{b\gamma} > 0.2, \\ 100\text{ GeV} &< m_{\gamma\gamma} < 150\text{ GeV}. \end{aligned} \tag{3.1}$$

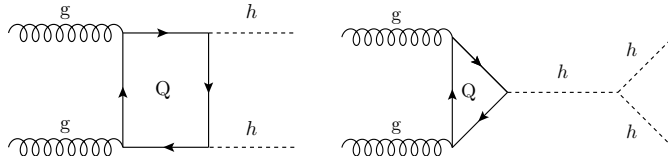


Figure 3. The cross-section of the ggF channel can be decomposed into three subprocesses based on their dependence on the trilinear Higgs self-coupling, λ . The triangle topology depends on λ^2 , the box topology does not depend on λ and the interference amongst the latter two is linear in λ .

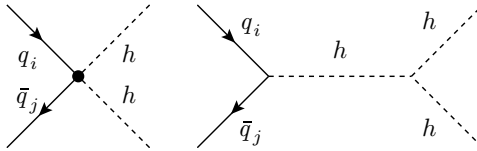


Figure 4. The dominant Feynman diagrams for the quark anti-quark annihilation ($q\bar{q}hh$) production of Higgs pair, via the SMEFT operator $\mathcal{O}_{q\phi}$.

Here Xp_T implies a minimum p_T cut for at least one b -parton. After the showering and detector simulation, further basic selection cuts are applied to select events with

$$\begin{aligned} \text{basic cuts:} \quad & n_{\text{eff}}^{\text{bjet}} \geq 1, & n_{\text{eff}}^{\gamma\text{jet}} \geq 2, \\ & p_T^{y\text{bjet}} > 30 \text{ GeV}, & p_T^{\gamma\text{jet}} > 5 \text{ GeV}, \\ & \eta_{\text{bjet},\gamma\text{jet}} < 4, & 110 \text{ GeV} < m_{\gamma\gamma} < 140 \text{ GeV}, \end{aligned} \quad (3.2)$$

with $n_{\text{eff}}^{b/\gamma\text{jet}}$ representing the number of b/γ -jets that pass the basic selection. The cross-section, K -factors, number of events with 6 ab^{-1} luminosity at 14 TeV are given in table 1.

The simulation of the hh signal is separated into two main channels. The first is the gluon-fusion (ggF) channel which is the dominant channel in the SM and can be further decomposed into three subprocesses based on their dependence on the Higgs trilinear self-interaction, λ , as seen in figure 3. Amongst these subprocesses, the first is the amplitude squared of the contribution from the triangle diagram which is proportional to λ^2 . The second is the squared amplitude of the contribution from the box diagram that does not depend on the trilinear coupling. The third is the contribution from the interference between the triangle and box diagrams, which is proportional to λ . Using this separation allows us to remove the dependence of the total K -factor for hh production on rescaling of the trilinear Higgs coupling [78]. The individual K -factors for each of the subprocesses are independent of the rescaling of the trilinear Higgs coupling making our analysis computationally much simpler. The ggF process is generated using the HH production program implemented in POWHEG [78–80], which has been modified to separate the individual contributions from the three diagrams. The cross-section for these individual contributions and the corresponding K -factors can be found in table 2 as derived from ref. [81].

The other main process, the quark anti-quark annihilation ($q\bar{q}hh$), is strongly suppressed in the SM for first-generation quarks since the SM Yukawa couplings are proportional to the mass of the considered quark flavour. However, since this channel is a tree-level process (cf. figure 4), with sufficiently large enhancement factors of the light-quark Yukawa coupling,

Channel	LO σ [fb]	NLO K -fact	6 ab^{-1} [#evt @ NLO]
y_b^2	0.0648	1.5	583
$y_b y_t$	-0.00829	1.9	-95
y_t^2	0.123	2.5	1,840
Zh	0.0827	1.3	645
$\sum b\bar{b}h$	0.262	-	2,970
$t\bar{t}h$	1.156	1.2	6,938
$b\bar{b}\gamma\gamma$	12.9	1.5	116,000

Table 1. SM cross-section for the main background processes at 14 TeV with 6 ab^{-1} data at the HL-LHC, and the number of events after the basic cuts as defined in eq. (3.2). For $b\bar{b}h$ production, the Higgs boson is decayed to a pair of photons. The total production cross-section of Higgs associated with a $b\bar{b}$ pair is denoted by $\sum b\bar{b}h$ and is the sum of the top four channels.

Channel	LO σ [fb]	K -fact.	6 ab^{-1} [#evt @ order]
hh_{tri}^{ggF}	$7.3 \cdot 10^{-3}$	2.28	96 (NNLO)
hh_{box}^{ggF}	$54 \cdot 10^{-3}$	1.98	680 (NNLO)
hh_{int}^{ggF}	$-36 \cdot 10^{-3}$	2.15	-460 (NNLO)
$u\bar{u}hh$ ($C_{u\phi} = 0.1$)	$5.6 \cdot 10^{-3}$	1.30	43 (NLO)
$d\bar{d}hh$ ($C_{d\phi} = 0.1$)	$3.6 \cdot 10^{-3}$	1.29	28 (NLO)

Table 2. The LO SM cross-section for di-Higgs production at the HL-LHC for 6 ab^{-1} of data multiplied by the $hh \rightarrow b\bar{b}\gamma\gamma$ SM branching ratio, K -factors (taken from [77] for the gluon channels and [27] for the quark channels) and the number of events after the basic cuts for the gluon fusion (ggF) and quark annihilation ($q\bar{q}hh$) at $\sqrt{s} = 14 \text{ TeV}$.

it becomes dominant as shown in figure 1. The $q\bar{q}hh$ cross-section scales like $C_{q\phi}^2/\Lambda^4$ while the ggF production cross-section remains almost unchanged. Therefore, for constraining enhancements of the light-quark Yukawa, we consider this channel as the signal and the ggF channel as part of the background. We note that given the fact that the SM contribution is zero if the light quark masses are neglected, the leading SMEFT contributions to this process are (dimension six)². This does not invalidate the EFT analysis, since the SM \times dimension eight contribution again is vanishing.

The $q\bar{q}hh$ processes are generated with `MadGraph_aMC@NLO` with a UFO model created with `FeynRules` [82]. Samples for both up- and down-quark initiated $q\bar{q}hh$ processes are generated. For all the hh signals, the samples are generated at LO and later scaled by the NLO K -factors given in table 2. The K -factors are adapted from refs. [83–85] as described in [27] for the $q\bar{q}hh$ channel. Moreover, the two Higgs bosons are decayed to $b\bar{b}$ and $\gamma\gamma$ respectively, with `Pythia 8.3` and then showered. The same detector simulation and basic cuts as for the background processes are then performed. In addition, the same sets of parton distribution function sets (`NNPDF31_nlo_as_0118_nf_4`) are used for the signal and the background, implemented via `LHAPDF` [86]. The calculation of the Higgs full width

and branching ratios is done using a modified version of `Hdecay` [87, 88] to include the effects of SMEFT operators $\mathcal{O}_{q\phi}$ and taken into account as a rescaling factor for the total cross-section of the hh signal and the relevant background processes. It should be noted, that in both di-Higgs production and decay calculation, the light quark masses are set to zero. However, when converting between SMEFT and κ -formalism, the $\overline{\text{MS}}$ quark masses at $\mu_R = 2 \text{ GeV}$ are used, in accordance with the PDG.

For FCC-hh, almost everything is done similarly after setting the energy to 100 TeV and the luminosity to 30 ab^{-1} . Since we do not have all K -factors available at a collider energy of 100 TeV we rescaled the LO samples by the same K -factors as for HL-LHC. We note that we explicitly checked that at least within the SM, for Higgs pair production via gluon fusion the difference is of $\mathcal{O}(1\%)$ [89] and hence a good approximation.

4 Exploring higher dimensional kinematic distributions

After detector simulation and jet definition, we have a final state of two photon jets and at least one identified b -jet, where the two photons reconstruct back to a real scalar Higgs mass for all the $b\bar{b}h$ channels. We first define and evaluate a comprehensive set of kinematic observables as follows:

- $p_T^{b_1}, p_T^{b_2}, p_T^{\gamma_1}, p_T^{\gamma\gamma},$
- $\eta_{b_{j1}}, \eta_{b_{j2}}, \eta_{\gamma_1}, \eta_{\gamma\gamma},$
- $n_{b\text{jet}}, n_{\text{jet}}, \Delta R_{\text{min}}^{b\gamma}, \Delta\varphi_{\text{min}}^{bb},$
- $m_{\gamma\gamma}, m_{bb}, m_{b_1h}, m_{b\bar{b}h}, H_T.$

Here $p_T^{b/\gamma_{1,2}}$ and $\eta^{b/\gamma_{1,2}}$ are the transverse momentum and rapidity of the tagged leading and sub-leading b/γ -jets (in our definition the subleading b -jet could be a null four-vector since we require one b -jet inclusive), n_{bj} is the number of tagged and passed b -jets. The variables $\Delta R_{\text{min}}^{b\gamma}$ and $\Delta\varphi_{\text{min}}^{bb}$ are the minimum R -distance and φ -angle between a tagged b -jet and a photon jet. The remaining variables are the invariant masses and H_T is the scalar sum of the transverse mass of the system. We shall show in what follows that it is not necessary to be very selective about the kinematic variables one chooses to use in the analysis. What is necessary is that all possibly useful kinematic variables are included. As can be seen from the list above, some of the variables seem to be interdependent and, probably, highly correlated. The beauty of using interpretable machine learning is that a hierarchy of importance for the variables will be built during the analysis using an over-complete basis of collider observables from which the most important ones can be chosen to further enhance the analysis.

4.1 Interpretable machine learning

Rule-based machine learning algorithms have for long been used as the gold standard for a signal to background discrimination in a wide variety of particle physics analyses. They are known to outperform neural networks in terms of simplicity of implementation,

computational resources required and accuracy in modelling the underlying distributions.⁴ In addition, rule-based algorithms, such as decision trees, are more transparent as far as separating the signal from the background is concerned. The first tree of the BDT always starts off with the most important kinematic cut which can be seen from the structure of the first tree making the process more interpretable. Placing emphasis on interpretability in multivariate analyses [34], we chose to work with Boosted Decision Trees (BDT). However, the interpretability of a machine learning algorithm requires more than just a choice of an interpretable model. The conditions are:

- A variable set that is easily interpretable in terms of the dynamics being studied.
- A machine learning algorithm that is more transparent and not a complete black box.
- A method for interpreting the model and attribute variable importance to understand how the algorithm models the underlying distributions.

Choosing to work with BDTs just satisfies the second condition. For training the BDTs we use XGBoost [91], a publicly available scalable end-to-end boosting system for decision trees. We follow the normal procedures for training and testing the BDT with simulated data. To satisfy the first condition we chose to work with high-level kinematic variables that are representative of the process instead of working with four-vectors. The disadvantage of working with kinematic variables is that a complete set cannot be defined for a particular process unlike the four-vectors associated with the process. So, in principle, a large number of kinematic variables can be formulated and used in a multivariate analysis. While the number is, in general, not too large for any implementation of BDTs, having a large set of variables clouds the understanding of which ones are important for orchestrating the separation of the signal from the background. This is where the third condition listed above is important. Variable importance attribution is a way to “short-list” only those variables that play an important role in the predictive power of the classification (or regression) problem. There are several measures of variable importance used in machine learning like Gini or permutation-based measures [92, 93], local explanations with surrogate models [94] etc., to name a few. However, these suffer from inconsistencies or fail to provide a global explanation of the model [95].

To build a mathematically consistent procedure for variable importance attribution, we use Shapley values [96] from Coalition Game Theory. Formulated by Shapley in the mid-20th century, Shapley values is a formulation of an axiomatic prescription for fairly distributing the payoff of a game amongst the players in a n -player cooperative game. When applied to machine learning, Shapley values tell us how important the presence of a variable is in determining a certain category (like signal or background) when compared to its absence from the multivariate problem being addressed. The process naturally and mathematically lends itself to studying the correlations between different variables since all possible combinations of variables can be taken to check the outcome. A more detailed discussion of the application of Shapley values to signal vs. background classification

⁴Nevertheless, we tested a deep neural network built with Tensorflow [90] and found no improvement in the classification accuracy.

problems for particle physics can be found in refs. [33, 34, 97, 98]. In this work, we follow the same basic procedure as discussed in ref. [33]. The importance of a variable in determining the outcome of a classification will be quantified by the mean of the absolute Shapley value, $|\overline{S_v}|$, larger values signifying greater importance. We will use the SHAP (SHapley Additive exPlanations) [95] package implemented in python based on Shapley values calculated exactly using tree-explainers [99, 100].

To provide an intuition of what Shapley values imply let us look at some edge cases. If there are two variables in the problem that are fully correlated then they are considered “equally good players” and their Shapley values, axiomatically, will be exactly equal. If there is a kinematic variable that does not contribute to the outcome at all, i.e, if one varies the variable but the outcome remains constant, the Shapley value of that variable will remain exactly zero. Hence, Shapley values encapsulate the correlations, even the higher order ones, between the input variables themselves and also the input variables and the outcome.

5 The di-Higgs channel at future hadron colliders

We would like to study the bounds on three specific couplings in this work. The first one being the Higgs trilinear coupling quantified by C_ϕ defined in eq. (2.1) and the other two being the deformation of the first-generation SM Yukawa coupling to the up and down quark defined as $C_{u\phi}$ and $C_{d\phi}$ in eq. (2.9) with $i = j = 1$. We will not consider modifications of the second generation of quarks as their effects in di-Higgs production would be suppressed by the small parton distribution functions.

In the BDT analysis, we combine the $b\bar{b}h$ ($h \rightarrow \gamma\gamma$) and $t\bar{t}h$ ($h \rightarrow \gamma\gamma$) channels into one category calling it $Q\bar{Q}h$ while the other (continuum) background channel, $b\bar{b}\gamma\gamma$, is treated as a separate category. We do not combine all the background channels since the $b\bar{b}\gamma\gamma$ channel is, by far, the dominant background and combining the $Q\bar{Q}h$ channels to it results in an inability of the BDTs to learn the shapes of the $Q\bar{Q}h$ channels which are comparable in size to the signal. Hence, leaving them separate allows for better classification of the signal. Moreover, this also helps with the interpretability of the classifications in terms of the Shapley values of the kinematic variables. For the analysis involving C_ϕ , we simulate three separate categories for the triangle, box and interference terms of the ggF hh production which we refer to as hh_{tri}^{ggF} , hh_{box}^{ggF} and hh_{int}^{ggF} , respectively. The $q\bar{q}hh$ channels include two other categories, one each for probing the Wilson coefficients $C_{u\phi}$ and $C_{d\phi}$, respectively. Note that all the channels with a Higgs boson as an intermediate state are all sensitive to $C_{u\phi}$ and $C_{d\phi}$ through their modification to the total Higgs width and the $h \rightarrow \gamma\gamma$ decay [27], which are all taken into account. In what follows, we refer to the two $q\bar{q}hh$ channels as $u\bar{u}hh$ and $d\bar{d}hh$ explicitly.

As we progress through the analysis we study the modification of one, two and three Wilson coefficients at a time. To find an optimized constraint on C_ϕ from the data we perform a five-channel classification (two signal and three background modes including the hh_{box}^{ggF} contribution that is insensitive to modifications of C_ϕ). To constrain either $C_{u\phi}$ or $C_{d\phi}$ we perform a four-channel classification taking the ggF channel as a single background mode. To constrain C_ϕ and one of $C_{u\phi}$ or $C_{d\phi}$ we perform a six-channel classification. Lastly,

a simultaneous constraint on all three Wilson coefficients is realized with a seven-channel classification. All the codes and data necessary to reproduce the results we got from this interpretable machine learning framework are made available at a `GitHub` repository: <https://github.com/talismanbrandi/IML-diHiggs.git>.

To set the stage, we will define our measure of significance and how we estimate it. We first construct a confusion matrix from the predictions of the trained BDT. This is a $n \times n$ matrix, for n channels. The sum of the elements in the i^{th} row, $\sum_j N_{ij}$, gives the number of events produced in channel i that would be generated in a pseudo-experiment with the projected luminosity corresponding to the actual experiment. The sum of the j^{th} column, $\sum_i N_{ij}$, gives the number of events from channel j predicted (including correct classifications and misclassifications) by the BDT in this pseudo-experiment. Hence, the (i, j) element of the matrix gives the number of events of the i^{th} class that is classified as belonging to the j^{th} class with $i \neq j$ signifying a misclassification. The significance of the j^{th} channel given by $S/\sqrt{S+B}$, S being signal and B being background, can be defined as

$$\mathcal{Z}_j = \frac{|N_{jj}|}{\sqrt{\sum_i N_{ij}}}, \tag{5.1}$$

where i is the row index and j is the column index. For ease of interpretation, we will present our results also in terms of κ_λ , κ_u and κ_d which are simply the rescaling of the SM trilinear coupling and the light-quark Yukawa couplings of the up and down quarks, respectively.

5.1 Constraints on light-quark Yukawa couplings at the HL-LHC and FCC-hh

The fact that machine learning algorithms can far outperform cut-and-count analyses is a bygone conclusion. Preliminary estimates of the HL-LHC reach for SM Higgs pair production can be found in ref. [68] and range from 4σ to 4.5σ signal significance combining several channels and combining the ATLAS and CMS measurements. The $b\bar{b}\gamma\gamma$ final state alone allows for a $\sim 2.7\sigma$ measurement. In ref. [101], a more refined machine learning procedure using Bayesian Optimization has been suggested and it has been shown that, indeed, the measurement of a di-Higgs signal can be further improved over preliminary estimates made by ATLAS and CMS using the $b\bar{b}\gamma\gamma$ final state alone. A sensitivity of about 5σ can be achieved using their procedure with the caveat that they use S/\sqrt{B} as the definition of significance with a very low number of correctly classified signal and background events. As an exercise, we repeated the BDT analysis with our framework and estimated a $\sim 3.4\sigma$ signal significance for SM Higgs pair production, which is similar to the estimate made in ref. [101] without using any optimization.

A better portrayal of the advantages gained by using a multivariate analysis can be made by comparing the constraints set on $C_{u\phi}$, or κ_u , and $C_{d\phi}$, or κ_d , from a cut-and-count (CC) analysis and a multivariate (MV) analysis allowing for the variation of only one Wilson coefficient at a time. The projected 1σ bounds at HL-LHC for 6 ab^{-1} of luminosity for a CC analysis are given in ref. [27] and compared to our results as follows

$$\begin{aligned} C_{u\phi}^{MV}(\kappa_u^{MV}) &= [-0.09, 0.10] \quad ([-466, 454]), & C_{u\phi}^{CC}(\kappa_u^{CC}) &= [-0.18, 0.17] \quad ([-841, 820]), \\ C_{d\phi}^{MV}(\kappa_d^{MV}) &= [-0.16, 0.16] \quad ([-360, 360]), & C_{d\phi}^{CC}(\kappa_d^{CC}) &= [-0.18, 0.18] \quad ([-405, 405]). \end{aligned} \tag{5.2}$$

Predicted no. of events at HL-LHC							
Actual no. of events	Channel	hh_{tri}^{ggF}	hh_{tri}^{ggF}	hh_{box}^{ggF}	$Q\bar{Q}h$	$b\bar{b}\gamma\gamma$	total
	hh_{tri}^{ggF}	28	14	18	38	10	108
	hh_{int}^{ggF}	89	80	129	178	41	517
	hh_{box}^{ggF}	77	105	266	265	50	763
	$Q\bar{Q}h$	177	98	191	5,457	1,835	7,758
	$b\bar{b}\gamma\gamma$	1,743	845	1,074	30,849	287,280	321,791
	\mathcal{Z}_j	0.61	2.37	6.49	28.45	534.1	

Table 3. Trained BDT classification (confusion matrix) of the five channels used to extract constraints on C_ϕ (or κ_λ) at HL-LHC with 6 ab^{-1} luminosity (ATLAS+CMS), assuming SM signal injection. The right-most column gives the total number of events expected in each channel in the SM and the bottom-most row gives the signal significance.

Predicted no. of events at FCC-hh							
Actual no. of events	Channel	hh_{tri}^{ggF}	hh_{tri}^{ggF}	hh_{box}^{ggF}	$Q\bar{Q}h$	$b\bar{b}\gamma\gamma$	total
	hh_{tri}^{ggF}	3,579	1,303	2,372	4,697	337	12,288
	hh_{int}^{ggF}	13,602	7,300	17,075	24,716	1523	64,216
	hh_{box}^{ggF}	14,534	11,416	35,988	415,26	1,996	105,460
	$Q\bar{Q}h$	29,611	12,355	23,279	1,238,266	214,564	1,518,075
	$b\bar{b}\gamma\gamma$	45,574	22,290	26,213	150,935	227,142	24,317,657
	\mathcal{Z}_j	10.95	31.22	111.1	737.7	4,743	

Table 4. Trained BDT classification (confusion matrix) of the five channels used to extract constraints on C_ϕ (or κ_λ) at FCC-hh with 30 ab^{-1} luminosity, assuming SM signal injection. The right-most column gives the total number of events expected in each channel in the SM and the bottom-most row gives the signal significance.

From this, we clearly see a factor of ~ 2 improvement in the bounds on $C_{u\phi}$ and $\mathcal{O}(10\%)$ improvement in the determination of $C_{d\phi}$. The projected bounds on these operators at FCC-hh with 30 ab^{-1} of data using our framework are

$$\begin{aligned}
 C_{u\phi}^{MV} \left(\kappa_u^{MV} \right) &= [-0.012, 0.011] \left([-57.8, 54.7] \right) , \\
 C_{d\phi}^{MV} \left(\kappa_d^{MV} \right) &= [-0.012, 0.012] \left([-26.3, 28.4] \right) .
 \end{aligned}
 \tag{5.3}$$

These projected bounds for FCC-hh are an order of magnitude better than those for HL-LHC. In addition, the bounds on $C_{u\phi}$ and $C_{d\phi}$ are numerically the same displaying a much greater improvement in the bounds on $C_{d\phi}$ than on $C_{u\phi}$ at the higher energy collider.

5.2 Constraints on Higgs trilinear self-coupling at the HL-LHC and FCC-hh

In table 3, we provide the output of the BDT classification for 6 ab^{-1} of data collected at HL-LHC and in table 4, we provide the same for 30 ab^{-1} of data at FCC-hh. It can be seen from these matrices that while the $b\bar{b}\gamma\gamma$ QCD-QED channel is the dominant background, the BDT performs better in separating it from the signal channels than separating $Q\bar{Q}h$. This is due to the kinematic similarities between the signal and the $Q\bar{Q}h$ background.

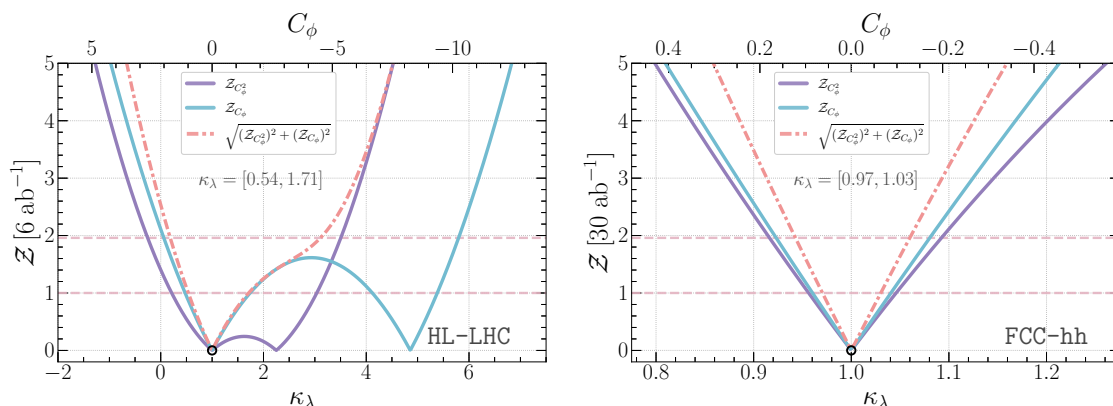


Figure 5. Bounds on κ_λ (or C_ϕ) at the HL-LHC (left panel) and the FCC-hh (right panel). The solid blue lines are the constraints coming from the hh_{int}^{ggF} contribution which scales linearly with the modified coupling and the solid purple line is that from the hh_{tri}^{ggF} contribution that scales quadratically with the modified coupling. The red dot-dashed line is the combination of the quadratic and linear channel. The horizontal light red dashed lines mark the 68% and 95% confidence intervals. The 68% CL bounds on κ_λ are given within square bracket.

In figure 5, we present the constraints on κ_λ (or C_ϕ) that can be set from HL-LHC in the left panel and FCC-hh in the right panel. The hh_{box}^{ggF} topology is not modified by C_ϕ and serves as a background to the measurement of C_ϕ . We separate the constraints from the hh_{tri}^{ggF} , which is quadratic in C_ϕ from the hh_{int}^{ggF} which is linear in C_ϕ . The combination of the two is given by the red dot-dashed line and is asymmetric around the best-fit point, for SM signal injection, $\kappa_\lambda = 1$ ($C_\phi = 0$). The projected 1σ bound on κ_λ is $[0.54, 1.71]$ at HL-LHC. There is a vast improvement projected for the FCC-hh which is not only due to increased luminosity but also due to the measurement being at a higher energy. The projected 1σ bound on κ_λ is about 3%.

5.3 Two and three parameter constraints on Higgs couplings

The primary focus of this work is to move beyond just looking at constraints on C_ϕ from Higgs pair production and to shed light on how simultaneous modifications of the light-quark Yukawa couplings due to non-zero contributions from $C_{u\phi}$ and $C_{d\phi}$ can change the constraints on C_ϕ . The modifications of the light-quark Yukawa couplings manifest themselves in two different ways. Firstly, non-zero $C_{u\phi}$ and $C_{d\phi}$ open up the $q\bar{q} \rightarrow hh$ production mode through a contact interaction (see figure 4) thus changing the production cross-section of the di-Higgs channel. This increase in the production cross-section sets the tightest constraints on $C_{u\phi}$ and $C_{d\phi}$ from Higgs pair production. Secondly, the modifications of the light-quark Yukawa couplings also modify the branching fraction of $h \rightarrow \gamma\gamma$ and the width of the Higgs boson. The latter modifies the channels that are also sensitive to C_ϕ , thus modifying the constraints that can be set on C_ϕ from future measurements. Such constraints from these modifications are the subdominant ones on $C_{u\phi}$ and $C_{d\phi}$ but they need to be taken into account for a holistic picture.

In the two-parameter fits, we consider three possible scenarios. Firstly, one can assume that the trilinear Higgs coupling is not modified and only the light-quark Yukawa couplings

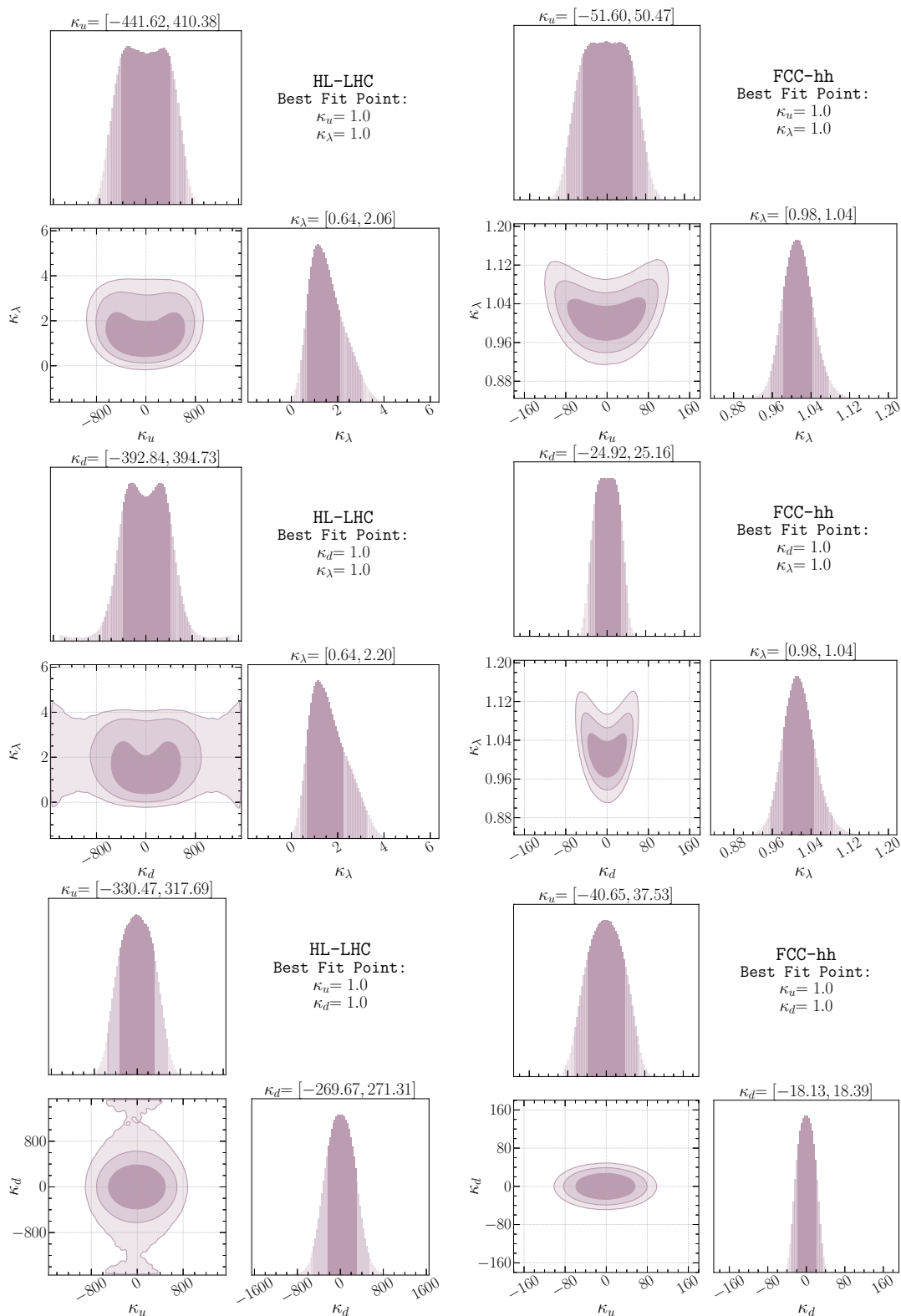


Figure 6. Constraints on pairs of Wilson coefficients for κ_λ , κ_u and κ_d . The panels on the left are for HL-LHC with 6 ab^{-1} of luminosity and the ones on the right are for FCC-hh with 30 ab^{-1} of luminosity. The one-dimensional distributions are derived from the two-dimensional distributions by marginalization using uniformly distributed priors for the independent variables with ranges sufficiently large to accommodate much more than 5σ variation of the variables from their central values.

Operators	$C_{u\phi}$	$C_{d\phi}$	C_ϕ	κ_u	κ_d	κ_λ
HL-LHC 14 TeV 6 ab ⁻¹						
\mathcal{O}_ϕ	–	–	[-1.6, 1.0]	–	–	[0.53, 1.7]
$\mathcal{O}_{u\phi}$	[-0.09, 0.10]	–	–	[-480, 430]	–	–
$\mathcal{O}_{d\phi}$	–	[-0.16, 0.16]	–	–	[-360, 360]	–
$\mathcal{O}_{u\phi}$ & \mathcal{O}_ϕ	[-0.087, 0.091]	–	[-2.4, 0.79]	[-430, 420]	–	[0.63, 2.1]
$\mathcal{O}_{d\phi}$ & \mathcal{O}_ϕ	–	[-0.17, 0.17]	[-2.7, 0.77]	–	[-380, 380]	[0.63, 2.3]
$\mathcal{O}_{u\phi}$ & $\mathcal{O}_{d\phi}$	[-0.066, 0.069]	[-0.12, 0.12]	–	[-330, 310]	[-270, 270]	–
All	[-0.077, 0.084]	[-0.16, 0.16]	[-2.8, 0.43]	[-400, 370]	[-360, 360]	[0.79, 2.3]
FCC-hh 100 TeV 30 ab ⁻¹						
\mathcal{O}_ϕ	–	–	[-0.066, 0.064]	–	–	[0.97, 1.03]
$\mathcal{O}_{u\phi}$	[-0.012, 0.011]	–	–	[-58, 55]	–	–
$\mathcal{O}_{d\phi}$	–	[-0.012, 0.011]	–	–	[-26, 28]	–
$\mathcal{O}_{u\phi}$ & \mathcal{O}_ϕ	[-0.010, 0.011]	–	[-0.091, 0.042]	[-52, 49]	–	[0.98, 1.04]
$\mathcal{O}_{d\phi}$ & \mathcal{O}_ϕ	–	[-0.010, 0.012]	[-0.092, 0.041]	–	[-24, 26]	[0.98, 1.04]
$\mathcal{O}_{u\phi}$ & $\mathcal{O}_{d\phi}$	[-0.008, 0.009]	[-0.008, 0.009]	–	[-42, 39]	[-19, 19]	–
All	[-0.009, 0.010]	[-0.009, 0.010]	[-0.11, 0.023]	[-47, 44]	[-21, 21]	[0.99, 1.05]

Table 5. The 1σ bounds on $C_{u\phi}$, $C_{d\phi}$ and C_ϕ from one-, two- and three-parameter fits for HL-LHC with 6 ab⁻¹ of data and FCC-hh with 30 ab⁻¹ of data. The corresponding bounds on the rescaling of the effective couplings, κ_u , κ_d and κ_λ are presented on the right side of the table.

are. Two other possibilities are the simultaneous modification of the C_ϕ and one of $C_{u\phi}$ and $C_{d\phi}$. These are the three constraints that we show in figure 6 in terms of κ_λ , κ_u and κ_d respectively. As before, the constraints have been obtained by training the BDT to separate the relevant signal channels from the background, the signal used being the one corresponding to the set of Wilson coefficients that we wish to constrain. The confusion matrices for all three cases can be found in the Github repository (<https://github.com/talismanbrandi/IML-diHiggs.git>) for this analysis. The left panels of figure 6 show the projected constraints for HL-LHC and the right panels for the FCC-hh.

Comparing with the constraints on κ_λ given in subsection 5.2 and figure 5, it can be seen from the top and middle left panels of figure 6 that, indeed, the constraints on κ_λ are diluted when the light-quark Yukawa couplings are allowed to vary. This effect is somewhat more prominent for κ_d than for κ_u . This distinction stems from the fact that away from $\kappa_{u,d} = 1$ larger negative values of κ_λ are allowed by the crescent-shaped curves in figure 6. For κ_d vs. κ_λ the 3σ region is unbounded in the domain $|\kappa_d| \gtrsim 1000$. The bounds on κ_u and κ_d from the fits with two parameters including κ_λ remain the same as the bounds on these Wilson coefficients from the single parameter $\kappa_{u,d}$ fits. We summarize the results in table 5.

It should be noted that the two-parameter fit for κ_u and κ_d provide a stronger bound on the two parameters than the fit performed individually. While this might be a bit counter-intuitive considering constraints from fits tend to deteriorate with the increasing number of parameters, we found that is not the case here. The reason is that the two-parameter fit is performed with the predictions made by the BDT trained with simulated events for both $u\bar{u}hh$ and $d\bar{d}hh$. Between these two channels, each forms the background for the other when

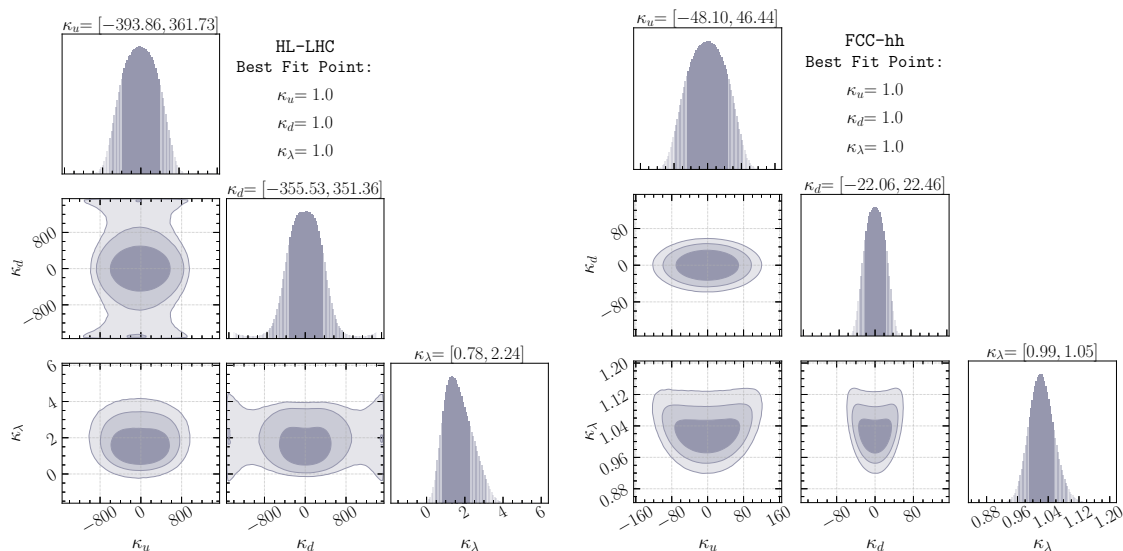


Figure 7. Three parameter fits with $C_{u\phi}$, $C_{d\phi}$ and C_ϕ , 6 ab^{-1} of luminosity at 14 TeV for HL-LHC (left panel) and 30 ab^{-1} of luminosity at 100 TeV for FCC-hh (right panel). The one-dimensional distributions are derived from the two-dimensional distributions by marginalization using uniformly distributed priors for the independent variables with ranges sufficiently large to accommodate much more than 5σ variation of the variables from their central values.

separating them through a confusion matrix. Since the training also gives the proportion of mistagged events, both the signal and the backgrounds are modified by the Wilson coefficients leading to a greater deformation of the likelihood in a favourable direction such that the constraints on the Wilson coefficients in the two-parameter fit is better than for the case in which they were separated from other $b\bar{b}\gamma\gamma$ backgrounds individually.

Finally, we perform a combined three-parameter fit including κ_λ , κ_u and κ_d (C_ϕ , $C_{u\phi}$ and $C_{d\phi}$), with the results shown in figure 7. For the same reason as explained before, the bounds on κ_u and κ_d are somewhat better than the two-parameter fits of these operators individually with κ_λ . The HL-LHC and FCC-hh projected bounds on κ_λ remain nearly the same as those from the corresponding two-parameter fits. In table 5 we also provide the bounds on C_ϕ , $C_{u\phi}$ and $C_{d\phi}$ for comparison.

5.4 Interpretation of Shapley values

Finally, we want to demonstrate the interpretability of the machine learning framework we use and discuss the physics that allows for the separation of the signal channels from the background channels. The advantage of using an interpretable multivariate framework is that one can easily understand which of the kinematic variables are important for orchestrating this separation in a manner that significantly improves upon a cut-and-count analysis. We use a measure derived from Shapley values, $|S_v|$, to understand the importance of each kinematic variable and understand the differences in kinematic shapes that separate the signal from the background.

To give a feeling of what the values of S_v mean, let us examine a single event. Assuming we have trained the BDT with n kinematic variables and each event has $n \times m$ Shapley

values associated with it, m being the number of channels (signal and background channels). For a particular channel, j , and kinematic variable, i , S_v can be positive or negative. A positive value implies that it is more likely that the event belongs to channel j given the value of the kinematic variable i . Conversely, a negative value implies that the event is less likely to belong to channel j given the value of the kinematic variable i . So regardless of whether S_v is positive or negative, it helps in the sorting of events into various channels. Hence, $|S_v|$ for a particular variable represents the strength of the variable to distinguish between channels. When summed over all channels this gives an overall picture of how good a discriminant a kinematic variable is for the processes involved. This is what is shown in figure 8 which we will now elaborate upon.

To begin with, we take a look at the $|S_v|$ computed for the five-channel analysis performed for separating hh_{tri}^{ggF} and hh_{int}^{ggF} channels from hh_{box}^{ggF} , $Q\bar{Q}h$ and $b\bar{b}\gamma\gamma$ QCD-QED background. In figure 8 we see the hierarchy plots for HL-LHC (top left panel) and FCC-hh (top right panel) generated from the predictions made by the BDT for this five-channel analysis. For both the colliders, H_T is the most important variable that is bringing about the separation of the hh_{tri}^{ggF} and hh_{int}^{ggF} channels from the dominating $b\bar{b}\gamma\gamma$ QCD-QED background. The second most important variable is $m_{\gamma\gamma}$. The importance of $m_{\gamma\gamma}$ accentuates the separation of the background by a greater degree at FCC-hh than at HL-LHC.

For the separation between the two $q\bar{q}hh$ channels, the story is very different. From the middle panels of figure 8 we see that the separation of $u\bar{u}hh$ and $d\bar{d}hh$ is truly a multivariate problem. Not surprisingly, the picture is very different for HL-LHC and FCC-hh. The differences between the two channels are driven by the differences in the parton distribution functions (PDF) of the up and down quarks. Since the PDF for the quarks change significantly from 14 TeV to 100 TeV, the variables that affect the separation of the two channels also change. Thus $|S_v|$ give us a true picture of how distributions of several kinematic variables determine the separation of different channels that are mostly similar. When comparing the abscissa of the top two panels with the middle two panels one will also notice that $|S_v|$ assumes much smaller values in the separation of $u\bar{u}hh$ and $d\bar{d}hh$. This clearly shows that the two channels are distributed quite identically and are difficult to separate.

Lastly, in the bottom panels of figure 8 we show the variables that are important in separating the $q\bar{q}hh$ channels from the ggF Higgs pair production channel. The invariant mass of the leading b -jet and h , m_{b_1h} is the most important variable at both HL-LHC and FCC-hh. However, the hierarchy of variables below m_{b_1h} is quite different for HL-LHC and FCC-hh. Both H_T and $p_T^{\gamma\gamma}$ are far less important at FCC-hh than at HL-LHC. This displays the clear advantage that machine learning algorithms have over a cut-and-count analysis where separate cut strategies would have to be built for the two colliders leading to two separate analyses that can, instead, be done with the same framework when using machine learning.

The correlation plots in figure 9 show how the linear correlations amongst the variables evolve when different channels are added. In the top left panel are events sampled from the ggF distribution. One can already see clustering in some of the variables related to

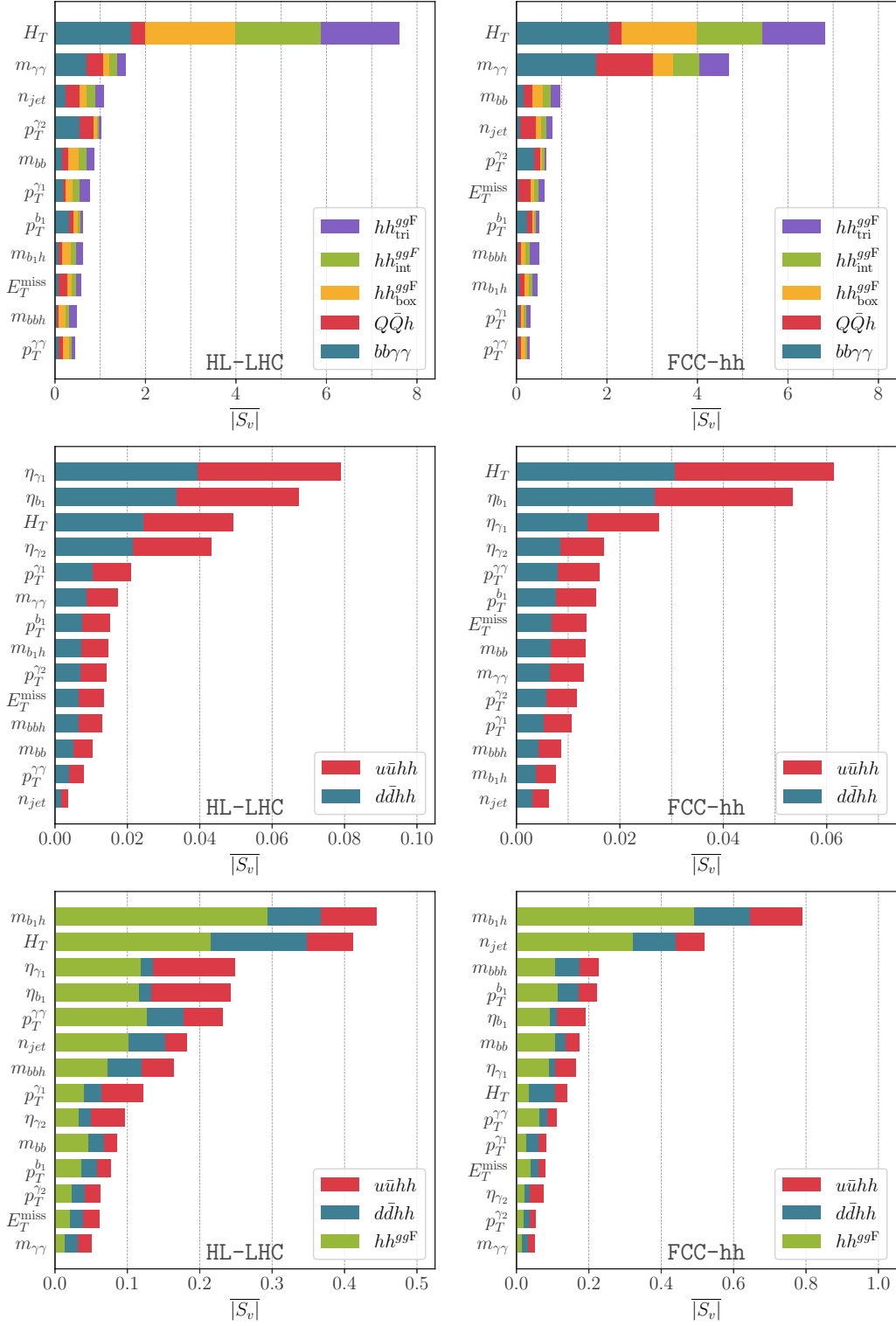


Figure 8. Top panels: the hierarchy of variables important for the separation of hh_{tri}^{ggF} from hh_{int}^{ggF} events from hh_{box}^{ggF} , $Q\bar{Q}h$ and $bb\gamma\gamma$ QCD-QED background at HL-LHC (left panel) and FCC-hh (right panel). Middle panels: the hierarchy of variables important for the separation of $u\bar{u}hh$ from $d\bar{d}hh$ events at HL-LHC (left panel) and FCC-hh (right panel). Lower panels: the hierarchy of variables important for the separation of hh^{ggF} , $u\bar{u}hh$ and $d\bar{d}hh$ events at HL-LHC (left panel) and FCC-hh (right panel). The higher the value of $|\overline{S}_v|$ is, the more important the kinematic variable is in separating the different channels.

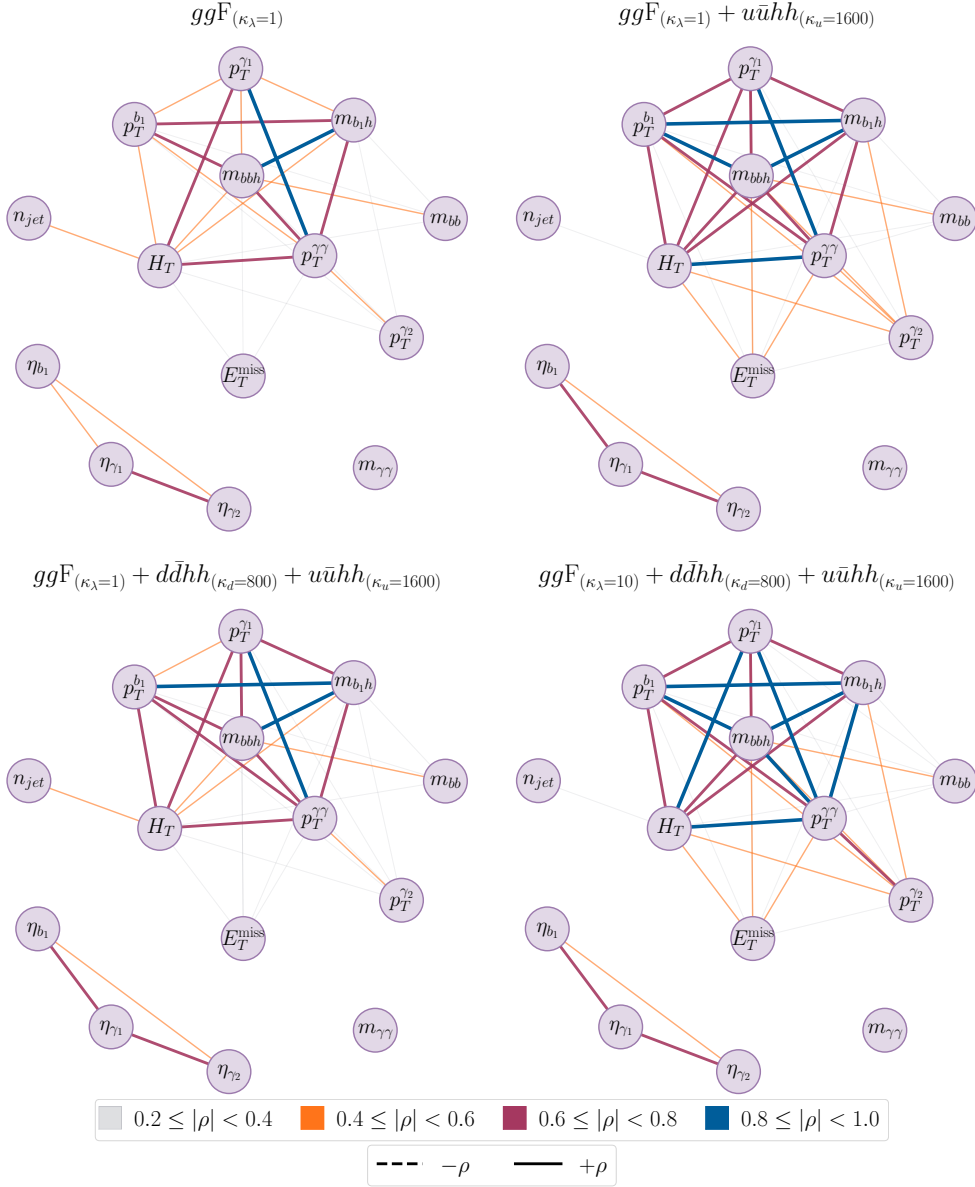


Figure 9. Network diagrams visualization of correlations (ρ) amongst the kinematic variables used in the analysis. Top left: only the gluon-gluon fusion channel. Top right: the gg^F channel along with the $u\bar{u}hh$ channel with $\kappa_u = 1600$. Bottom right: the $d\bar{d}hh$ channel with $\kappa_d = 800$ added to the channels in the top right panel. Bottom left: the same channels as in the bottom right panel but with $\kappa_\lambda = 10$.

momenta and invariant mass. The other cluster is of the pseudorapidity of the particles in the final state. This correlation structure evolves when one adds the $u\bar{u}hh$ channel when E_T^{miss} gets connected to the upper cluster in the top right panel. The correlation is now stronger between η_{γ_1} and η_{b_1} and several correlations in the upper cluster are much stronger too. The change in the correlations continues as one keeps adding channels as can be seen from the bottom right and bottom left panels. It is the capture of this change in the correlations (and higher-order correlations) that enhances the capabilities of the machine learning algorithms to distinguish between the various channels. While $m_{\gamma\gamma}$, by its shape alone, allows for the separation between $b\bar{b}\gamma\gamma$ and the other channels, the correlations between the other kinematic variables aid in the separation of the channels with one or two Higgs in the final state.

5.5 Additional constraints on light-quark Yukawa couplings

There are additional proposed measurements of the light-quark Yukawa couplings that might become relevant at HL-LHC or FCC-hh, a careful study of which is beyond the scope of the current work. We will attempt to include a discussion here so as to provide a comparison with our study and to put it into proper context.

Studies of rare Higgs decays, involving radiative decays to quarkonia have been proposed in [19, 102, 103], as a direct probe for light Yukawa couplings. These studies were followed upon with experimental searches for such decays [104] and set bounds on the branching ratios, $\mathcal{B}(h \rightarrow J/\Psi, \gamma/Z) \sim 10^{-4} - 10^{-6}$ at 95% CL. More recent bounds on κ_c [15, 16] sets it to about $|\kappa_c| < 8.5$ which is hitherto the most stringent *direct* bound on charm quark Yukawa coupling. Another probe for light-quark Yukawa couplings is the associated production of Higgs with a jet. This channel has been shown to be sensitive to $\kappa_c \sim 1$ when using charm-tagged jets [23]. Moreover, by looking at differential distributions for this channel, it is possible to obtain stringent bounds on the first-generation Yukawa couplings [24, 105, 106]. Limits on light-quark Yukawa couplings can also be extracted by studying the untagged branching ratios of the Higgs decay to di-jets, under the assumption that no additional new physics present [107]. However, all the channels mentioned before, suffer from degeneracy amongst up- and down-quark Yukawa couplings. Other channels can be considered as complementary to them to break this degeneracy. It was shown in refs. [108, 109], that the charge asymmetry of the process $pp \rightarrow hW^+$ vs. $pp \rightarrow hW^-$ can be used as a probe for light-quark Yukawa couplings as well as to break the degeneracy amongst quark flavours. Moreover, the rare process $pp \rightarrow h\gamma$ is also a possible way to distinguish between enhancements of the up- and down-quark Yukawa couplings [110]. A cut-based analysis of Higgs pair production using the same final state considered in this work, $hh \rightarrow b\bar{b}\gamma\gamma$, has obtained constraints on light-quark Yukawa couplings [27]. The analysis can be thought of as being complementary to the previously mentioned ones as it mostly probes the coupling between two Higgs bosons to quarks. In addition, we see in our work that using machine learning significantly improves upon the cut-based analysis. Hence this analysis, taken together with other proposals, provides a probe of non-linearities between the Higgs and light quarks parameterized by the electroweak chiral Lagrangian. Constraints for Higgs couplings could also come from processes that do not involve Higgs

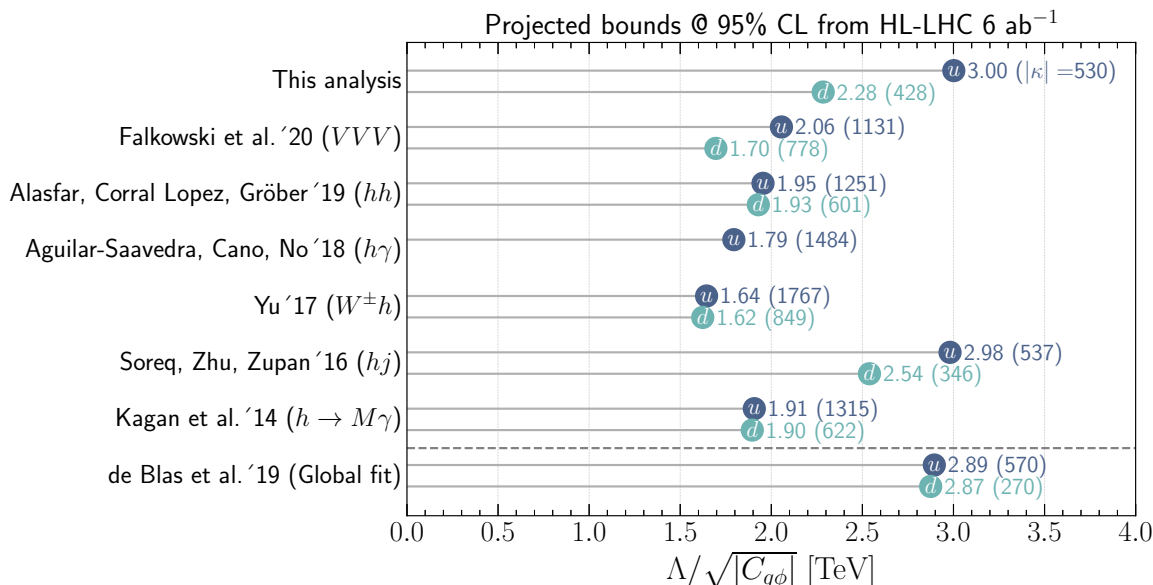


Figure 10. Summary of the 95% CL sensitivity bounds on the SMEFT Wilson coefficients $C_{u\phi}$ (blue), and $C_{d\phi}$ (green). The bounds are interpreted in terms of the NP scale Λ that can be reached through the measurements of the Wilson coefficient at the HL-LHC at 6 ab^{-1} , the corresponding κ_q 's are shown inside the parentheses. The 95% CL bounds from single parameter fits are used from this analysis for comparison with previous studies.

production directly. Three-boson production VVV has been shown to give strong projected bounds for light-quark Yukawa couplings for HL-LHC, with ten-fold improvement expected at FCC-hh [25, 111].

We present a numerical comparison of the strongest bounds from HL-LHC on the first-generation Yukawa couplings from the studies discussed above in figure 10 and compare them to the global fit bounds that have been obtained with no invisible or untagged Higgs decays allowed [112]. For $C_{d\phi}$, the most stringent bound comes from the global fit, and the $h + j$ channel, as a model-independent bound, while our analysis provides the second most stringent model-independent bound. For $C_{u\phi}$, our analysis provides the most stringent constraint while the bound from $h + j$ and the global analysis are comparable. The figure is interpreted in terms of the reach of NP scale Λ that can be achieved by the measurement of these Wilson coefficients. For future colliders, like the FCC-hh at 100 TeV, in addition to Higgs pair production, triple Higgs production might be an interesting channel for constraining the operators with Wilson coefficient $C_{u\phi}$ and $C_{d\phi}$ due to the energy increase of a Feynman diagram coupling the quarks to three Higgs bosons. Finally, we note that there are also signatures from experiments not based on colliders for enhanced light-quark Yukawa couplings manifesting in frequency shifts in atomic clocks from Higgs forces at the atomic level [113].

6 Summary

In this work, we walk through an analysis of how kinematic shapes can be used to glean information about the nuances of various production modes with the same final states but

deformed differentially by the existence of degrees of freedom beyond the Standard Model. We show that this information can be extracted by using an interpretable machine learning framework which is not only very effective in separating these differences in kinematic shapes but also yields itself to interpretations in terms of physics that are known and well understood. The example we chose is Higgs pair production in the $b\bar{b}\gamma\gamma$ final state.

We emphasized that probing Higgs pair production is an important next step for an understanding of the model underlying the fundamental interactions of particles and hence a potential gateway to new physics. We show that even beyond the trilinear Higgs couplings, the light-quark Yukawa couplings can be probed through this production mode. In fact, the $q\bar{q}hh$ channel opens up only in the presence of BSM physics and well-motivated models of new dynamics, bringing about the simultaneous modification of the trilinear Higgs coupling and the light-quark Yukawa couplings. Indeed, we motivated our study by showing that in different frameworks large modifications of the light-quark Yukawa couplings can be obtained. Knowing the difficulty of measuring these couplings, we propose an interpretable machine learning framework that significantly outperforms traditional cut-based analyses.

As opposed to using black-box models, the interpretable framework allows us to gain physics insights into how signal and background separation can be brought into effect, pointing to kinematic variables like H_T and $m_{\gamma\gamma}$ as being important variables that instrument this separation. As a result, we find enhanced sensitivities to C_ϕ or κ_λ that quantify the modification to the Higgs trilinear coupling. Furthermore, we see that the measurement of the light-quark Yukawa couplings is aided by using the methods we advocate bringing about greater sensitivities than would be possible with a cut-based analysis at the HL-LHC and the FCC-hh. The advantage of using an interpretable framework using Shapley values is that it provides added confidence to the robustness of the multivariate analyses that we perform using simulated data.

The salient results of this work are:

- The modification of the Higgs trilinear coupling can be measured at $\mathcal{O}(1)$ precision at the HL-LHC and at $\mathcal{O}(1\%)$ precision at the FCC-hh.
- The rescaling of the light-quark Yukawa couplings, κ_u and κ_d , can be measured to $\mathcal{O}(100)$ at the HL-LHC and $\mathcal{O}(10)$ at FCC-hh.
- The measurement of κ_λ is significantly diluted once the light-quark Yukawa couplings are allowed to vary. Hence, in a joint fit, the bounds on κ_λ are weaker.
- There are theoretical models that motivate the simultaneous modification of the trilinear Higgs coupling and the light-quark Yukawa couplings. Hence, the dilution of the bounds on κ_λ due to the presence of new physics in the light-quark Yukawa sector should be taken into consideration in future phenomenological extraction of κ_λ .
- The bounds obtained with the interpretable machine learning framework that we use not only outperforms cut-based analyses but also allow for physics insights into kinematic distributions of the various channels that help distinguish them in an experiment.

In conclusion, we stress that the interplay between the Yukawa sector and the Higgs trilinear coupling is non-trivial and requires careful consideration. Future experiments at the HL-LHC and FCC-hh will bring significant improvements in the sensitivities to κ_λ , κ_u and κ_d through the Higgs pair production channel. In particular, the bounds on the light-quark Yukawa couplings from Higgs pair production can possibly be the most stringent bounds amongst all other experimental probes of the light-quark Yukawa couplings as a result of the evolution of the parton luminosity functions between the single Higgs threshold and the characteristic energy scale in Higgs pair production.

Acknowledgments

This work benefited from support by the Deutsche Forschungsgemeinschaft under Germany’s Excellence Strategy EXC 2121 “Quantum Universe” – 390833306 as well as under the grant 491245950. L.A.’s research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 417533893/GRK2575 “Rethinking Quantum Field Theory”. The work of A.P. is funded by Volkswagen Foundation within the initiative “Corona Crisis and Beyond – Perspectives for Science, Scholarship and Society”, grant number 99091. Part of the work done by A.P. was funded by the Roux Institute and the Harold Alfond Foundation. R.G. acknowledges support from a departmental research grant under the project “Machine Learning approach to Effective Field Theories in Higgs Physics”. This research was supported in part through the Maxwell computational resources operated at DESY, Hamburg, Germany.

A Discussion of theoretical and systematic uncertainties

In this section, we present an estimate of the systematic uncertainties that can affect the measurements discussed in this work at the HL-LHC. We do not present these estimates for the FCC-hh for lack of sufficient information or the ability to project such uncertainties far into the future. We use two scenarios for systematic uncertainties: the first is a 8.2% uncertainty which corresponds to the current systematic uncertainty that ATLAS has reported for their Run-II search for Higgs pair production [114]. The second scenario is the ATLAS HL-LHC baseline systematic uncertainty of 5.3% reported in [115]. For LHC run-II, statistical uncertainties remain the dominant part of the uncertainty budget for di-Higgs analysis. Regarding the systematic uncertainties, experimental sources remain the dominant part in comparison to the theoretical ones. The story flips for the HL-LHC where the main source of uncertainties is expected to be coming from theoretical uncertainties. The current theoretical uncertainty estimate of the SM gluon fusion process at NNLO is $^{+6\%}_{-23\%}$ for $\sqrt{s} = 14$ TeV and $^{+4\%}_{-21\%}$ for $\sqrt{s} = 100$ TeV [116]. The largest part of the uncertainty stems from the uncertainty due to the renormalization scheme choice of the top quark mass. This uncertainty can, for the moment, only be estimated at NLO since no full mass-dependent results at NNLO are available. Moreover, the top quark mass renormalization scheme uncertainty is not included in the estimated HL-LHC (nor LHC Run II) uncertainties schemes that we have considered.

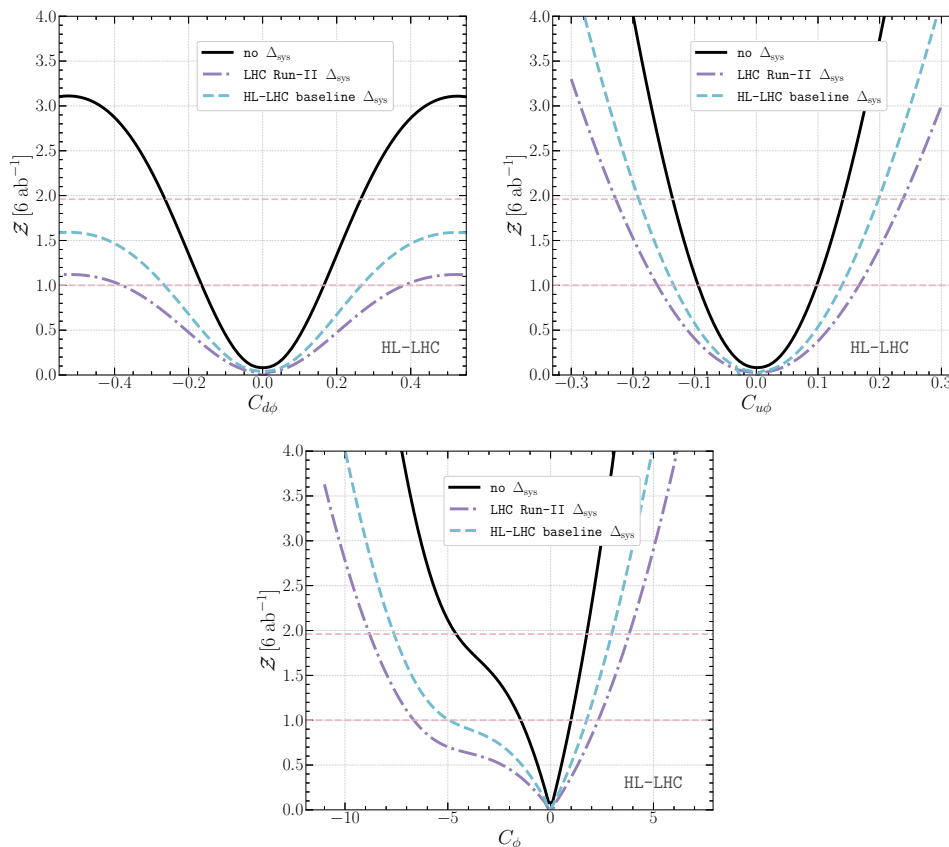


Figure 11. The significance, \mathcal{Z} , from a single parameter fit for $C_{d\phi}$ (upper left panel), $C_{u\phi}$ (upper right panel) and C_ϕ (lower center panel) for the HL-LHC with no systematic uncertainties (black) and two ansatz for systematic uncertainties. The first is the current Run-II 8.2% in violet and the HL-LHC baseline 5.3% estimated by ATLAS in blue, including theoretical uncertainties without the one stemming from the top mass renormalization scheme.

In figure 11 we show the significance \mathcal{Z} for the three Wilson coefficient, C_ϕ , $C_{u\phi}$ and $C_{d\phi}$, at the HL-LHC from single parameter fits with no systematic uncertainties (black), LHC Run-II (violet) and HL-LHC baseline (blue) systematic uncertainties ansatz. We observe that for the current Run-II ansatz, the bounds for all three Wilson coefficients is diluted by 100% or more. As for the HL-LHC baseline, the bounds are diluted by $\sim 70\%$. However, it should be noted, that both systematic uncertainties scenarios are rather conservative. It is likely that the HL-LHC detector upgrade and new theoretical developments in higher-order corrections to the di-Higgs cross-section will reduce the systematic uncertainties from the baseline.

B Light-quark Yukawa and self coupling at future lepton colliders

Future high energy lepton colliders [119–122] offer further alternative and clean signals for measurement of Higgs properties. For example, Higgs decays to “un-tagged” light jets including u, d, s quarks can be further disentangled from $h \rightarrow gg$ using event shape analysis [117] and can reach a sensitivity of $\kappa_d \approx 90$ and $\kappa_u \approx 192$ at 250 GeV with 5 ab^{-1} data compared with a sensitivity of $\kappa_d \approx 470$ and $\kappa_u \approx 900$ at the 6 ab^{-1} HL-LHC [105, 107].

Collider	$ \kappa_u $	$ \kappa_d $	$\delta\kappa_\lambda$ (1σ)
240GeV 5ab ⁻¹ (CECP/FCC)	192 [117]	90 [117]	100% (Indirect [118])
350 GeV 1.5 ab ⁻¹ (FCCee)	310 [112]	140 [112]	40% (Indirect [118])
500 GeV 4 ab ⁻¹ (ILC)	330 [112]	160 [112]	27% [119]
1 TeV 8 ab ⁻¹ (ILC)	–	–	10% [112]
3 TeV 1 ab ⁻¹ (CLIC)	430 [112]	200 [112]	10% [112]
10 TeV 10 ab ⁻¹ (Muon)	–	–	3% [112]

Table 6. Prospective light-quark Yukawa and Higgs self-coupling sensitivities at future lepton colliders. The light-quark Yukawa bounds are 95% CL, while the self-coupling bounds are 1σ or 68% CL sensitivity reach.

The sensitivity to Higgs self-coupling comes indirectly for center of mass energy below 250 GeV from the precision measurement of the Zh production channel (1σ bound on $\delta\kappa_\lambda$ of 0.4 at 250 GeV), and at 500 GeV directly from the Zhh channel (1σ bound on $\delta\kappa_\lambda$ of 0.27 at 500 GeV), and from vector boson fusion like production to $hh\nu\nu$ when 1 TeV or higher energy scales are available (1σ bound on $\delta\kappa_\lambda$ of 10% at 1 TeV). The prospective sensitivity depends on the collider setup, mainly the integrated luminosity and polarization of initial lepton beams. Given the updated prospects of future machine designs [112], we list a short summary in table 6 of the expected sensitivities on the individual parameters in the κ framework. These numbers are all assuming one-parameter fits in κ . No simultaneous fit including both κ_q and κ_λ (or using the corresponding SMEFT operators) has been performed yet.

C Shapley values

Shapley values [96] are defined for a game (v, N) , where $N = \{1, \dots, n\}$ is a set of players in the game and v is the characteristic function that assigns a non-negative real value $v(S)$ to every coalition $S \subseteq N$, and zero to the empty coalition, i.e. $v(\emptyset) = 0$. The function v fully describes the game, as it maps players to payoffs. A subset S of N is referred to as a coalition, and $v(S)$ the value of the coalition. The marginal contribution of a player i to the coalition S is defined as $v(S \cup \{i\}) - v(S)$. The average marginal contribution of player i , over the set S_k of all coalitions containing of k players which exclude i , is

$$\bar{v}_k(i) = \frac{1}{|S_k|} \sum_{S \in S_k} [v(S \cup \{i\}) - v(S)], \quad (\text{C.1})$$

where $|S_k| = \binom{n-1}{k}$. The Shapley value of player i is then

$$\phi_i(v) = \frac{1}{n} \sum_{k=0}^{n-1} \bar{v}_k(i). \quad (\text{C.2})$$

Combining eq. (C.1) and eq. (C.2) we get:

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} (v(S \cup \{i\}) - v(S)), \quad i = 1, \dots, n; \quad (\text{C.3})$$

a weighted mean over all subsets S (including the empty set $S = \emptyset$) not containing i with $|S|$ denoting the cardinality of S . The decomposition into Shapley values is the only solution satisfying a set of four favorable axioms [123]: *Efficiency*, *Symmetry*, *Linearity*, and *Null Player* which are given by:

- *Efficiency*: the sum of the payoff to the individual players equal the payoff of the grand coalition, i.e.

$$\sum_{i \in N} \phi_i(v) = v(N).$$

- *Symmetry*: the contributions of two players j and k should be the same if they contribute equally to all possible coalitions,

$$v(S \cup \{j\}) = v(S \cup \{k\}) \quad \forall S \subseteq \{1, \dots, n\} \setminus \{j, k\} \iff \phi_j = \phi_k.$$

- The *Null Player*: a player j that does not change the payoff — regardless of which coalition of players they are added to — should have a Shapley value of 0,

$$C(v \cup \{i\}) = v(S) \quad \forall S \subseteq \{1, \dots, n\} \iff \phi_i = 0$$

- *Linearity*: the payoffs for the linear sum of two games is the linear sum of the payoffs for each game. For two games (v, N) and (w, N) :

$$\phi_i(\alpha_v v + \alpha_w w) = \alpha_v \phi_i(v) + \alpha_w \phi_i(w).$$

For every coalition that can be formed, value of player i is assessed for when the player is added to the coalition. The averaged payoff that player i gets from all possible coalition is the Shapley value of the player. In a multivariate analysis, the players can be replaced by independent variables and the payoff can be replaced by an outcome or dependent variable(s). Shapley values for the variables can be computed by fitting a machine learning model to data and using the model as the characteristic function. There are distinct challenges to doing this. Firstly, an exact computation of Shapley values scale exponentially with the number of variables and becomes intractable even for a moderate number of variables. Secondly, retraining a model after removing variables is not possible since that would change the model itself and, hence, change the characteristic function.

As a solution to the first problem, SHAP [95] uses either kernel methods or tree-explainers [99, 100] which significantly reduces the computational burden. The tree-explainer addresses the second problem by ignoring the branches of a decision tree that contain the variable not included in a coalition and computing the weighted average of the outcome from the rest of the model. Due to the linearity property of the Shapley value, Shapley values from an ensemble of trees can be added to compute the final Shapley value of a variable. Given that the variable which, on an average, consistently contributed more to the outcome will have a higher Shapley value, an importance ranking can be based on this Shapley value which is known as feature importance in machine learning.

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. SCOAP³ supports the goals of the International Year of Basic Sciences for Sustainable Development.

References

- [1] ATLAS collaboration, *Combined measurements of Higgs boson production and decay using up to 80 fb⁻¹ of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*, *Phys. Rev. D* **101** (2020) 012002 [[arXiv:1909.02845](https://arxiv.org/abs/1909.02845)] [[INSPIRE](#)].
- [2] CMS collaboration, *Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 421 [[arXiv:1809.10733](https://arxiv.org/abs/1809.10733)] [[INSPIRE](#)].
- [3] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01 [[INSPIRE](#)].
- [4] A. Djouadi, W. Kilian, M. Muhlleitner and P.M. Zerwas, *Production of neutral Higgs boson pairs at LHC*, *Eur. Phys. J. C* **10** (1999) 45 [[hep-ph/9904287](https://arxiv.org/abs/hep-ph/9904287)] [[INSPIRE](#)].
- [5] U. Baur, T. Plehn and D.L. Rainwater, *Probing the Higgs selfcoupling at hadron colliders using rare decays*, *Phys. Rev. D* **69** (2004) 053004 [[hep-ph/0310056](https://arxiv.org/abs/hep-ph/0310056)] [[INSPIRE](#)].
- [6] M.J. Dolan, C. Englert and M. Spannowsky, *Higgs self-coupling measurements at the LHC*, *JHEP* **10** (2012) 112 [[arXiv:1206.5001](https://arxiv.org/abs/1206.5001)] [[INSPIRE](#)].
- [7] J. Baglio, A. Djouadi, R. Gröber, M.M. Mühlleitner, J. Quevillon and M. Spira, *The measurement of the Higgs self-coupling at the LHC: theoretical status*, *JHEP* **04** (2013) 151 [[arXiv:1212.5581](https://arxiv.org/abs/1212.5581)] [[INSPIRE](#)].
- [8] J. Alison et al., *Higgs boson potential at colliders: Status and perspectives*, *Rev. Phys.* **5** (2020) 100045 [[arXiv:1910.00012](https://arxiv.org/abs/1910.00012)] [[INSPIRE](#)].
- [9] 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](https://arxiv.org/abs/1607.03773)] [[INSPIRE](#)].
- [10] G. Degrandi, 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](https://arxiv.org/abs/1607.04251)] [[INSPIRE](#)].
- [11] 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](https://arxiv.org/abs/1610.05771)] [[INSPIRE](#)].
- [12] 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](https://arxiv.org/abs/1709.08649)] [[INSPIRE](#)].
- [13] 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](https://arxiv.org/abs/1704.01953)] [[INSPIRE](#)].
- [14] L. Alasfar, J. de Blas and R. Gröber, *Higgs probes of top quark contact interactions and their interplay with the Higgs self-coupling*, *JHEP* **05** (2022) 111 [[arXiv:2202.02333](https://arxiv.org/abs/2202.02333)] [[INSPIRE](#)].
- [15] ATLAS collaboration, *Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector*, *Eur. Phys. J. C* **82** (2022) 717 [[arXiv:2201.11428](https://arxiv.org/abs/2201.11428)] [[INSPIRE](#)].

- [16] CMS collaboration, *A search for the standard model Higgs boson decaying to charm quarks*, *JHEP* **03** (2020) 131 [[arXiv:1912.01662](#)] [[INSPIRE](#)].
- [17] G. Perez, Y. Soreq, E. Stamou and K. Tobioka, *Prospects for measuring the Higgs boson coupling to light quarks*, *Phys. Rev. D* **93** (2016) 013001 [[arXiv:1505.06689](#)] [[INSPIRE](#)].
- [18] ATLAS COLLABORATION collaboration, *A search for the dimuon decay of the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, [ATLAS-CONF-2019-028](#).
- [19] G.T. Bodwin, F. Petriello, S. Stoynev and M. Velasco, *Higgs boson decays to quarkonia and the $H\bar{c}c$ coupling*, *Phys. Rev. D* **88** (2013) 053003 [[arXiv:1306.5770](#)] [[INSPIRE](#)].
- [20] S. Alte, M. König and M. Neubert, *Exclusive Weak Radiative Higgs Decays in the Standard Model and Beyond*, *JHEP* **12** (2016) 037 [[arXiv:1609.06310](#)] [[INSPIRE](#)].
- [21] ATLAS collaboration, *Searches for exclusive Higgs and Z boson decays into $J/\psi\gamma$, $\psi(2S)\gamma$, and $\Upsilon(nS)\gamma$ at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 134 [[arXiv:1807.00802](#)] [[INSPIRE](#)].
- [22] CMS COLLABORATION collaboration, *Search for Higgs boson decays into Z and J/ψ and for Higgs and Z boson decays into J/ψ or Υ pairs at CMS*, [CMS-PAS-HIG-20-008](#).
- [23] I. Brivio, F. Goertz and G. Isidori, *Probing the Charm Quark Yukawa Coupling in Higgs + Charm Production*, *Phys. Rev. Lett.* **115** (2015) 211801 [[arXiv:1507.02916](#)] [[INSPIRE](#)].
- [24] F. Bishara, U. Haisch, P.F. Monni and E. Re, *Constraining Light-Quark Yukawa Couplings from Higgs Distributions*, *Phys. Rev. Lett.* **118** (2017) 121801 [[arXiv:1606.09253](#)] [[INSPIRE](#)].
- [25] N. Vignaroli, *Off-Shell Probes of the Higgs Yukawa Couplings: Light Quarks and Charm*, *Symmetry* **14** (2022) 1183 [[arXiv:2205.09449](#)] [[INSPIRE](#)].
- [26] J. Duarte-Campderros, G. Perez, M. Schlaffer and A. Soffer, *Probing the Higgs-strange-quark coupling at e^+e^- colliders using light-jet flavor tagging*, *Phys. Rev. D* **101** (2020) 115005 [[arXiv:1811.09636](#)] [[INSPIRE](#)].
- [27] L. Alasfar, R. Corral Lopez and R. Gröber, *Probing Higgs couplings to light quarks via Higgs pair production*, *JHEP* **11** (2019) 088 [[arXiv:1909.05279](#)] [[INSPIRE](#)].
- [28] ATLAS collaboration, *Search for non-resonant Higgs boson pair production in the $b\bar{b}l\nu l\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* **801** (2020) 135145 [[arXiv:1908.06765](#)] [[INSPIRE](#)].
- [29] A. Adhikary, R.K. Barman and B. Bhattacharjee, *Prospects of non-resonant di-Higgs searches and Higgs boson self-coupling measurement at the HE-LHC using machine learning techniques*, *JHEP* **12** (2020) 179 [[arXiv:2006.11879](#)] [[INSPIRE](#)].
- [30] J. Amacker et al., *Higgs self-coupling measurements using deep learning in the $b\bar{b}b\bar{b}$ final state*, *JHEP* **12** (2020) 115 [[arXiv:2004.04240](#)] [[INSPIRE](#)].
- [31] B. Tannenwald et al., *Benchmarking Machine Learning Techniques with Di-Higgs Production at the LHC*, [arXiv:2009.06754](#) [[INSPIRE](#)].
- [32] L. Huang, S.-b. Kang, J.H. Kim, K. Kong and J.S. Pi, *Portraying double Higgs at the Large Hadron Collider II*, *JHEP* **08** (2022) 114 [[arXiv:2203.11951](#)] [[INSPIRE](#)].
- [33] C. Grojean, A. Paul and Z. Qian, *Resurrecting $b\bar{b}h$ with kinematic shapes*, *JHEP* **04** (2021) 139 [[arXiv:2011.13945](#)] [[INSPIRE](#)].

- [34] C. Grojean, A. Paul, Z. Qian and I. Strümke, *Lessons on interpretable machine learning from particle physics*, *Nature Rev. Phys.* **4** (2022) 284 [[arXiv:2203.08021](#)] [[INSPIRE](#)].
- [35] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [[arXiv:1008.4884](#)] [[INSPIRE](#)].
- [36] 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](#)] [[INSPIRE](#)].
- [37] SMEFT collaboration, *Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC*, *JHEP* **11** (2021) 089 [[arXiv:2105.00006](#)] [[INSPIRE](#)].
- [38] A. Paul and M. Valli, *Violation of custodial symmetry from W-boson mass measurements*, *Phys. Rev. D* **106** (2022) 013008 [[arXiv:2204.05267](#)] [[INSPIRE](#)].
- [39] L. Di Luzio, R. Gröber and P. Paradisi, *Higgs physics confronts the M_W anomaly*, *Phys. Lett. B* **832** (2022) 137250 [[arXiv:2204.05284](#)] [[INSPIRE](#)].
- [40] Y.T. Chien, V. Cirigliano, W. Dekens, J. de Vries and E. Mereghetti, *Direct and indirect constraints on CP-violating Higgs-quark and Higgs-gluon interactions*, *JHEP* **02** (2016) 011 [[arXiv:1510.00725](#)] [[INSPIRE](#)].
- [41] J. Brod and D. Skodras, *Electric dipole moment constraints on CP-violating light-quark Yukawas*, *JHEP* **01** (2019) 233 [[arXiv:1811.05480](#)] [[INSPIRE](#)].
- [42] G. Blankenburg, J. Ellis and G. Isidori, *Flavour-Changing Decays of a 125 GeV Higgs-like Particle*, *Phys. Lett. B* **712** (2012) 386 [[arXiv:1202.5704](#)] [[INSPIRE](#)].
- [43] R. Harnik, J. Kopp and J. Zupan, *Flavor Violating Higgs Decays*, *JHEP* **03** (2013) 026 [[arXiv:1209.1397](#)] [[INSPIRE](#)].
- [44] G. D'Ambrosio, G.F. Giudice, G. Isidori and A. Strumia, *Minimal flavor violation: An Effective field theory approach*, *Nucl. Phys. B* **645** (2002) 155 [[hep-ph/0207036](#)] [[INSPIRE](#)].
- [45] A. Pich and P. Tuzon, *Yukawa Alignment in the Two-Higgs-Doublet Model*, *Phys. Rev. D* **80** (2009) 091702 [[arXiv:0908.1554](#)] [[INSPIRE](#)].
- [46] A. Pich, *Flavour constraints on multi-Higgs-doublet models: Yukawa alignment*, *Nucl. Phys. B Proc. Suppl.* **209** (2010) 182 [[arXiv:1010.5217](#)] [[INSPIRE](#)].
- [47] P.M. Ferreira, L. Lavoura and J.P. Silva, *Renormalization-group constraints on Yukawa alignment in multi-Higgs-doublet models*, *Phys. Lett. B* **688** (2010) 341 [[arXiv:1001.2561](#)] [[INSPIRE](#)].
- [48] M. Jung, A. Pich and P. Tuzon, *Charged-Higgs phenomenology in the Aligned two-Higgs-doublet model*, *JHEP* **11** (2010) 003 [[arXiv:1006.0470](#)] [[INSPIRE](#)].
- [49] F.J. Botella, G.C. Branco, A.M. Coutinho, M.N. Rebelo and J.I. Silva-Marcos, *Natural Quasi-Alignment with two Higgs Doublets and RGE Stability*, *Eur. Phys. J. C* **75** (2015) 286 [[arXiv:1501.07435](#)] [[INSPIRE](#)].
- [50] S. Bar-Shalom and A. Soni, *Universally enhanced light-quarks Yukawa couplings paradigm*, *Phys. Rev. D* **98** (2018) 055001 [[arXiv:1804.02400](#)] [[INSPIRE](#)].
- [51] ATLAS COLLABORATION collaboration, *Searches for vector-like quarks with the ATLAS Detector*, [ATL-PHYS-SLIDE-2021-363](#).
- [52] CMS collaboration, *Search for pair production of vectorlike quarks in the fully hadronic final state*, *Phys. Rev. D* **100** (2019) 072001 [[arXiv:1906.11903](#)] [[INSPIRE](#)].

- [53] CMS collaboration, *Search for vector-like quarks in events with two oppositely charged leptons and jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 364 [[arXiv:1812.09768](#)] [[INSPIRE](#)].
- [54] 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](#)].
- [55] A. Falkowski and R. Rattazzi, *Which EFT*, *JHEP* **10** (2019) 255 [[arXiv:1902.05936](#)] [[INSPIRE](#)].
- [56] S. Chang and M.A. Luty, *The Higgs Trilinear Coupling and the Scale of New Physics*, *JHEP* **03** (2020) 140 [[arXiv:1902.05556](#)] [[INSPIRE](#)].
- [57] D. Egana-Ugrinovic, S. Homiller and P.R. Meade, *Higgs bosons with large couplings to light quarks*, *Phys. Rev. D* **100** (2019) 115041 [[arXiv:1908.11376](#)] [[INSPIRE](#)].
- [58] D. Egana-Ugrinovic, S. Homiller and P. Meade, *Multi-Higgs Production Probes Higgs Flavor*, *Phys. Rev. D* **103** (2021) 115005 [[arXiv:2101.04119](#)] [[INSPIRE](#)].
- [59] D. Egana-Ugrinovic, S. Homiller and P. Meade, *Aligned and Spontaneous Flavor Violation*, *Phys. Rev. Lett.* **123** (2019) 031802 [[arXiv:1811.00017](#)] [[INSPIRE](#)].
- [60] ATLAS collaboration, *Search for low-mass resonances decaying into two jets and produced in association with a photon using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **795** (2019) 56 [[arXiv:1901.10917](#)] [[INSPIRE](#)].
- [61] ATLAS collaboration, *Search for new resonances in mass distributions of jet pairs using 139fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **03** (2020) 145 [[arXiv:1910.08447](#)] [[INSPIRE](#)].
- [62] CMS collaboration, *Search for high mass dijet resonances with a new background prediction method in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **05** (2020) 033 [[arXiv:1911.03947](#)] [[INSPIRE](#)].
- [63] 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](#)].
- [64] 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](#)] [[INSPIRE](#)].
- [65] ATLAS collaboration, *Search for heavy resonances decaying into a Z boson and a Higgs boson in final states with leptons and b-jets in 139fb^{-1} of pp collisions at $\sqrt{s} = 13\text{TeV}$ with the ATLAS detector*, [ATLAS-CONF-2020-043](#).
- [66] ATLAS collaboration, *Search for heavy resonances decaying into a pair of Z bosons in the $\ell^+\ell^-\ell'^+\ell'^-$ and $\ell^+\ell^-\nu\bar{\nu}$ final states using 139fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 332 [[arXiv:2009.14791](#)] [[INSPIRE](#)].
- [67] CMS collaboration, *Search for a new scalar resonance decaying to a pair of Z bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **06** (2018) 127 [Erratum *ibid.* **03** (2019) 128] [[arXiv:1804.01939](#)] [[INSPIRE](#)].
- [68] M. Cepeda et al., *Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC*, *CERN Yellow Rep. Monogr.* **7** (2019) 221 [[arXiv:1902.00134](#)] [[INSPIRE](#)].

- [69] W. Beenakker, S. Dittmaier, M. Krämer, B. Plumper, M. Spira and P.M. Zerwas, *Higgs radiation off top quarks at the Tevatron and the LHC*, *Phys. Rev. Lett.* **87** (2001) 201805 [[hep-ph/0107081](#)] [[INSPIRE](#)].
- [70] D. Fähr and N. Greiner, *Diphoton production in association with two bottom jets*, *Eur. Phys. J. C* **77** (2017) 750 [[arXiv:1706.08309](#)] [[INSPIRE](#)].
- [71] F. Campanario, R. Roth and D. Zeppenfeld, *QCD radiation in WH and WZ production and anomalous coupling measurements*, *Phys. Rev. D* **91** (2015) 054039 [[arXiv:1410.4840](#)] [[INSPIRE](#)].
- [72] S. Dawson, C.B. Jackson, L. Reina and D. Wackerroth, *Higgs production in association with bottom quarks at hadron colliders*, *Mod. Phys. Lett. A* **21** (2006) 89 [[hep-ph/0508293](#)] [[INSPIRE](#)].
- [73] 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](#)].
- [74] 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](#)].
- [75] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [[arXiv:1410.3012](#)] [[INSPIRE](#)].
- [76] DELPHES 3 collaboration, *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057 [[arXiv:1307.6346](#)] [[INSPIRE](#)].
- [77] D. de Florian, I. Fabre, G. Heinrich, J. Mazzeitelli and L. Scyboz, *Anomalous couplings in Higgs-boson pair production at approximate NNLO QCD*, *JHEP* **09** (2021) 161 [[arXiv:2106.14050](#)] [[INSPIRE](#)].
- [78] G. Heinrich, S.P. Jones, M. Kerner, G. Luisoni and L. Scyboz, *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](#)].
- [79] G. Heinrich, S.P. Jones, M. Kerner, G. Luisoni and E. Vryonidou, *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](#)].
- [80] G. Heinrich, S.P. Jones, M. Kerner and L. Scyboz, *A non-linear EFT description of $gg \rightarrow HH$ at NLO interfaced to POWHEG*, *JHEP* **10** (2020) 021 [[arXiv:2006.16877](#)] [[INSPIRE](#)].
- [81] G. Buchalla, M. Capozzi, A. Celis, G. Heinrich and L. Scyboz, *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](#)].
- [82] A. Alloul, N.D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 — A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250 [[arXiv:1310.1921](#)] [[INSPIRE](#)].
- [83] D. Dicus, T. Stelzer, Z. Sullivan and S. Willenbrock, *Higgs boson production in association with bottom quarks at next-to-leading order*, *Phys. Rev. D* **59** (1999) 094016 [[hep-ph/9811492](#)] [[INSPIRE](#)].
- [84] C. Balázs, H.-J. He and C.P. Yuan, *QCD corrections to scalar production via heavy quark fusion at hadron colliders*, *Phys. Rev. D* **60** (1999) 114001 [[hep-ph/9812263](#)] [[INSPIRE](#)].

- [85] R.V. Harlander and W.B. Kilgore, *Higgs boson production in bottom quark fusion at next-to-next-to leading order*, *Phys. Rev. D* **68** (2003) 013001 [[hep-ph/0304035](#)] [[INSPIRE](#)].
- [86] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132 [[arXiv:1412.7420](#)] [[INSPIRE](#)].
- [87] A. Djouadi, J. Kalinowski and M. Spira, *HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension*, *Comput. Phys. Commun.* **108** (1998) 56 [[hep-ph/9704448](#)] [[INSPIRE](#)].
- [88] A. Djouadi, J. Kalinowski, M. Muehlleitner and M. Spira, *HDECAY: Twenty++ years after*, *Comput. Phys. Commun.* **238** (2019) 214 [[arXiv:1801.09506](#)] [[INSPIRE](#)].
- [89] 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](#)].
- [90] M. Abadi et al., *TensorFlow: Large-scale machine learning on heterogeneous systems*, (2015).
- [91] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, [arXiv:1603.02754](#) [[INSPIRE](#)].
- [92] L. Breiman, *Random forests*, *Mach. Learn.* **45** (2001) 5.
- [93] A. Fisher, C. Rudin and F. Dominici, *All models are wrong, but many are useful: Learning a variable’s importance by studying an entire class of prediction models simultaneously*, *J. Mach. Learn. Res.* **20** (2019) 1 [[arXiv:1801.01489](#)].
- [94] M.T. Ribeiro, S. Singh and C. Guestrin, “*why should i trust you?*”: *Explaining the predictions of any classifier*, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, New York, U.S.A. (2016), pp. 1135–1144 [[DOI](#)].
- [95] S.M. Lundberg and S.-I. Lee, *A unified approach to interpreting model predictions*, in *Advances in Neural Information Processing Systems*, I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan et al. eds., Curran Associates Inc. (2017), pp. 4765–4774 [[arXiv:1705.07874](#)].
- [96] L.S. Shapley, *Notes on the n-Person Game-II: The Value of an n-Person Game*, Rand Corporation (1951).
- [97] D. Alvestad, N. Fomin, J. Kersten, S. Maeland and I. Strümke, *Beyond Cuts in Small Signal Scenarios — Enhanced Sneutrino Detectability Using Machine Learning*, [arXiv:2108.03125](#) [[INSPIRE](#)].
- [98] A.S. Cornell, W. Doorsamy, B. Fuks, G. Harmsen and L. Mason, *Boosted decision trees in the era of new physics: a smuon analysis case study*, *JHEP* **04** (2022) 015 [[arXiv:2109.11815](#)] [[INSPIRE](#)].
- [99] S.M. Lundberg, G.G. Erion and S.-I. Lee, *Consistent Individualized Feature Attribution for Tree Ensembles*, [arXiv:1802.03888](#).
- [100] S.M. Lundberg et al., *From local explanations to global understanding with explainable AI for trees*, *Nature Machine Intelligence* **2** (2020) 56.
- [101] A. Alves, T. Ghosh and K. Sinha, *Can We Discover Double Higgs Production at the LHC?*, *Phys. Rev. D* **96** (2017) 035022 [[arXiv:1704.07395](#)] [[INSPIRE](#)].
- [102] A.L. Kagan, G. Perez, F. Petriello, Y. Soreq, S. Stoynev and J. Zupan, *Exclusive Window onto Higgs Yukawa Couplings*, *Phys. Rev. Lett.* **114** (2015) 101802 [[arXiv:1406.1722](#)] [[INSPIRE](#)].

- [103] M. König and M. Neubert, *Exclusive Radiative Higgs Decays as Probes of Light-Quark Yukawa Couplings*, *JHEP* **08** (2015) 012 [[arXiv:1505.03870](#)] [[INSPIRE](#)].
- [104] CMS collaboration, *Search for rare decays of Z and Higgs bosons to J/ψ and a photon in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 94 [[arXiv:1810.10056](#)] [[INSPIRE](#)].
- [105] Y. Soreq, H.X. Zhu and J. Zupan, *Light quark Yukawa couplings from Higgs kinematics*, *JHEP* **12** (2016) 045 [[arXiv:1606.09621](#)] [[INSPIRE](#)].
- [106] G. Bonner and H.E. Logan, *Constraining the Higgs couplings to up and down quarks using production kinematics at the CERN Large Hadron Collider*, [arXiv:1608.04376](#) [[INSPIRE](#)].
- [107] L.M. Carpenter, T. Han, K. Hendricks, Z. Qian and N. Zhou, *Higgs Boson Decay to Light Jets at the LHC*, *Phys. Rev. D* **95** (2017) 053003 [[arXiv:1611.05463](#)] [[INSPIRE](#)].
- [108] F. Yu, *Phenomenology of Enhanced Light Quark Yukawa Couplings and the $W^\pm h$ Charge Asymmetry*, *JHEP* **02** (2017) 083 [[arXiv:1609.06592](#)] [[INSPIRE](#)].
- [109] F. Yu, *Light Quark Yukawa Couplings and the $W^\pm h$ Charge Asymmetry*, *Nucl. Part. Phys. Proc.* **285-286** (2017) 123 [[INSPIRE](#)].
- [110] J.A. Aguilar-Saavedra, J.M. Cano and J.M. No, *More light on Higgs flavor at the LHC: Higgs boson couplings to light quarks through $h + \gamma$ production*, *Phys. Rev. D* **103** (2021) 095023 [[arXiv:2008.12538](#)] [[INSPIRE](#)].
- [111] A. Falkowski et al., *Light quark Yukawas in triboson final states*, *JHEP* **04** (2021) 023 [[arXiv:2011.09551](#)] [[INSPIRE](#)].
- [112] J. de Blas et al., *Higgs Boson Studies at Future Particle Colliders*, *JHEP* **01** (2020) 139 [[arXiv:1905.03764](#)] [[INSPIRE](#)].
- [113] C. Delaunay, R. Ozeri, G. Perez and Y. Soreq, *Probing Atomic Higgs-like Forces at the Precision Frontier*, *Phys. Rev. D* **96** (2017) 093001 [[arXiv:1601.05087](#)] [[INSPIRE](#)].
- [114] ATLAS COLLABORATION collaboration, *Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [ATLAS-CONF-2021-016](#).
- [115] ATLAS COLLABORATION collaboration, *Measurement prospects of the pair production and self-coupling of the Higgs boson with the ATLAS experiment at the HL-LHC*, [TL-PHYS-PUB-2018-053](#).
- [116] J. Baglio, F. Campanario, S. Glaus, M. Mühlleitner, J. Ronca and M. Spira, *$gg \rightarrow HH$: Combined uncertainties*, *Phys. Rev. D* **103** (2021) 056002 [[arXiv:2008.11626](#)] [[INSPIRE](#)].
- [117] J. Gao, *Probing light-quark Yukawa couplings via hadronic event shapes at lepton colliders*, *JHEP* **01** (2018) 038 [[arXiv:1608.01746](#)] [[INSPIRE](#)].
- [118] S. Di Vita et al., *A global view on the Higgs self-coupling at lepton colliders*, *JHEP* **02** (2018) 178 [[arXiv:1711.03978](#)] [[INSPIRE](#)].
- [119] P. Bambade et al., *The International Linear Collider: A Global Project*, [arXiv:1903.01629](#) [[INSPIRE](#)].
- [120] CLICDP and CLIC collaborations, *The Compact Linear Collider (CLIC) — 2018 Summary Report*, CERN Yellow Reports: Monographs (2018) [[DOI](#)] [[arXiv:1812.06018](#)] [[INSPIRE](#)].
- [121] CEPC STUDY GROUP collaboration, *CEPC Conceptual Design Report: Volume 2 — Physics & Detector*, [arXiv:1811.10545](#) [[INSPIRE](#)].

- [122] A. Vasquez, C. Degrande, A. Tonerio and R. Rosenfeld, *New physics in double Higgs production at future e^+e^- colliders*, *JHEP* **05** (2019) 020 [[arXiv:1901.05979](#)] [[INSPIRE](#)].
- [123] H.P. Young, *Monotonic solutions of cooperative games*, *Int. J. Game Theory* **14** (1985) 65.