Prepared for submission to JHEP IPPP/20/52, CERN-TH-2020-186

arXiv:2011.03555v2 [hep-ph] 11 Mar 2021

# Precision SMEFT bounds from the VBF Higgs at high transverse momentum

**Jack Y. Araz,**[a] **Shankha Banerjee,**[b] **Rick S. Gupta,**[a] **Michael Spannowsky**[a]

[a] *Institute for Particle Physics Phenomenology, Durham University, South Road, Durham, DH1 3LE,*

[b] *CERN, Theoretical Physics Department, CH-1211 Geneva 23, Switzerland*

*E-mail:* jack.araz@durham.ac.uk, shankha.banerjee@cern.ch, sandeepan.gupta@durham.ac.uk, michael.spannowsky@durham.ac.uk

Abstract: We study the production of Higgs bosons at high transverse momenta via vector-boson fusion (VBF) in the Standard Model Effective Field Theory (SMEFT). We find that contributions from four independent operator combinations dominate in this limit. These are the same 'high energy primaries' that control high energy diboson processes, including Higgs-strahlung. We perform detailed collider simulations for the diphoton decay mode of the Higgs boson as well as the three final states arising from the ditau channel. Using the quadratic growth of the SMEFT contributions relative to the Standard Model (SM) contribution, we project very stringent bounds on these operators that far surpass the corresponding bounds from the LEP experiment.

## Contents


**1 Introduction** **1**

**2 VBF Higgs production at high transverse momentum in the D6 SMEFT** **3**

**3 Collider Analyses** **6**
3.1 The $h \to \tau^+\tau^-$ channels 7
3.2 The $h \to \gamma\gamma$ channel 8

**4 Projected sensitivity for EFT couplings** **9**

**5 Conclusions** **12**

**A Rectangular Cut-Based Analyses** **13**
A.1 The $h \to \tau^+\tau^-$ channels 13
A.2 The $h \to \gamma\gamma$ channel 14

**B Neural Network Analysis** **15**

**C Contribution from other anomalous couplings and other processes** **21**


## 1 Introduction

In the absence of any evidence for new physics at the Large Hadron Collider (LHC), the Standard Model Effective Field Theory (SMEFT) is an efficient parametrisation for heavy new physics beyond the reach of the LHC. The effective field theory (EFT) formalism has, in fact, become the standard framework to precision physics at the LHC [1–45]. As we approach higher integrated luminosities, very precise EFT limits will become achievable. This is, in particular true, because, with a higher luminosity, we will gain the ability to probe the high energy tails of various distributions accurately. This can lead to very precise bounds on SMEFT operators whose contributions grow with energy with respect to the SM.

As far as the Higgs and electroweak physics is concerned refs. [36, 37] identified a four-dimensional subspace of the full 59 dimensional space of dimension-6 operators that can be measured very accurately in the diboson processes, $pp \to Vh/VV$, $(V = W^\pm, Z)$ at high energies (see also ref. [46]). That the same set of four operators control both double gauge boson production and Higgs-strahlung is a consequence of the Goldstone

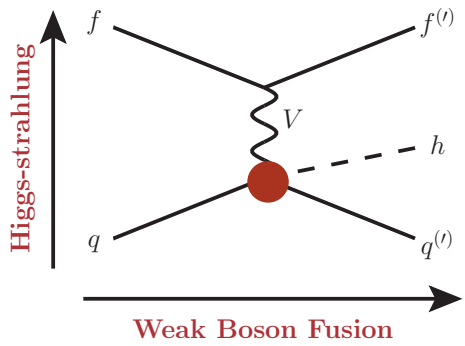


**Figure 1**. Figure shows the crossing symmetry that exists between the Higgs-strahlung and VBF Higgs production processes. The amplitudes for the two processes are the same up to an exchange of the Mandelstam variables, $s \leftrightarrow t$. As a result the same four directions in SMEFT space control VBF Higgs production at high $t$ and Higgs-strahlung at high $s$. The figure has been produced with the help of the JaxoDraw package [47].

Boson Equivalence theorem [1]. These four directions in the EFT space were dubbed the 'high energy primaries' in ref. [36]. It was shown that by utilising the quadratic energy growth of the contributions of these operators with respect to the SM, the LHC sensitivity to probe these operators can far surpass LEP bounds.

In this work, we show that these same high-energy primaries are also sufficient to completely determine the SMEFT amplitude for Higgs production in the Vector Boson Fusion (VBF) channel if the transverse momentum of the Higgs boson is large. The reason for this is a crossing-symmetry that exists between the VBF and Higgs-strahlung diagrams, as shown in Fig. 1. This implies that the two processes have the same amplitude up to an interchange in the Mandelstam variables, $s \leftrightarrow t$. Thus VBF Higgs production probes the same four operators at large $t$ as Higgs-strahlung at large $s$. Furthermore, one can also extend the equivalence theorem argument used in the diboson case in ref. [36] to this case and connect the VBF production of Higgs and gauge bosons.

Thus, the processes, $pp \to VV, Vh$ and VBF production of Higgs or gauge bosons, which are entirely different from each other from a collider physics point of view, actually probe, in a very precise manner, the same set of four operators at high energies. Combining these processes can thus give us the best bounds on the high-energy primaries. Apart from the apparent statistical advantage, it is crucial to combine all these processes because each of them probes a unique linear combination of the four operators; all these processes should, thus, be included to eliminate all flat directions.

As one of the important results of this work, we will present the linear combination of the four operators that are probed by VBF Higgs production. In this work, we carry out a thorough collider analysis of the $h \to \gamma\gamma$ channel and the three final states from the

[1] As a consequence of this theorem, the $Vh/VV$ production amplitudes are equivalent to the amplitude for producing different components of the Higgs doublet in the high energy limit. One can thus connect these amplitudes in the SM as well as the SMEFT using the full $SU(2)_L$ symmetry that is restored at high energies.

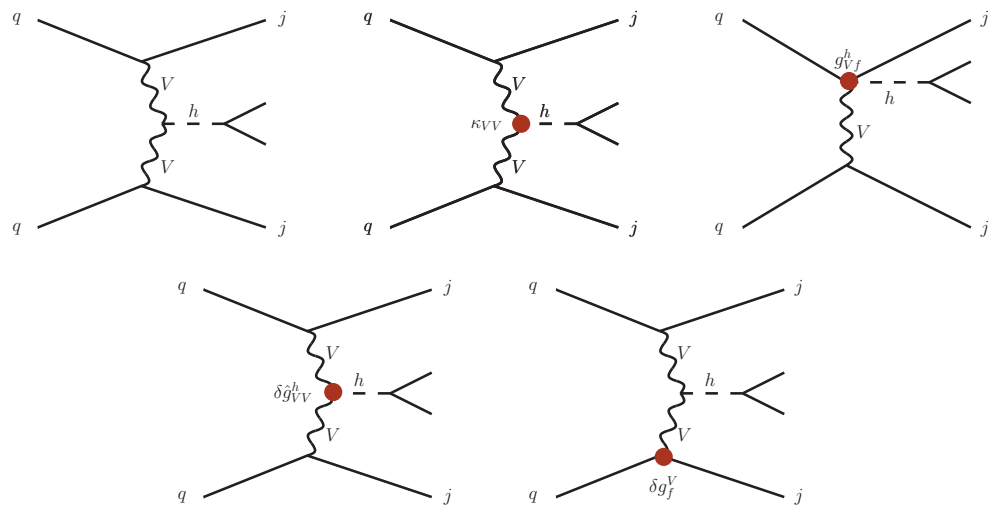


**Figure 2**. Figure shows representative VBF topologies in the SM and in the SMEFT. The red dots signify modified vertices from the EFT couplings in eq. (2.2). The figure has been produced with the help of the JaxoDraw package [47].

$h \to \tau^+\tau^-$ channel, namely, the hadronic, semi-leptonic and fully leptonic final states. We find that including all these channels is important as their sensitivity to the EFT effects is comparable. In the end, we obtain projections, much stronger than LEP bounds, on the high-energy primaries.

## 2 VBF Higgs production at high transverse momentum in the D6 SMEFT

The vertices in the dimension-6 (D6) lagrangian that contribute to the VBF Higgs production are the following,

$$\begin{aligned}
\Delta\mathcal{L}_6 \supset & \sum_f \delta g_f^Z Z_\mu \bar{f}\gamma^\mu f + \delta g_{ud}^W (W_\mu^+ \bar{u}_L \gamma^\mu d_L + h.c.) \\
& + \delta\hat{g}^h_{WW} \frac{2m_W^2}{v} h W^{+\mu} W_\mu^- + \delta\hat{g}^h_{ZZ} \frac{2m_Z^2}{v} h \frac{Z^\mu Z_\mu}{2} \\
& + \sum_f g^h_{Zf} \frac{h}{v} Z_\mu \bar{f}\gamma^\mu f + g^h_{Wud} \frac{h}{v} (W_\mu^+ \bar{u}_L \gamma^\mu d_L + h.c.) \\
& + \kappa_{\gamma\gamma} \frac{h}{v} A^{\mu\nu} A_{\mu\nu} + \frac{h}{v} \kappa_{WW} W^{+\mu\nu} W_{\mu\nu}^- + \kappa_{ZZ} \frac{h}{2v} Z^{\mu\nu} Z_{\mu\nu} + \delta\hat{g}^h_{\tau\tau} \frac{\sqrt{2} m_\tau}{v} h \tau^+ \tau^- \ ,
\end{aligned} \tag{2.1}$$

where we have expanded the D6 SMEFT Lagrangian to obtain lower dimension terms in the broken phase, taking $\alpha_{em}$, $m_Z$ and $m_W$ as the input parameters. Any correction to the SM vector propagators, such as $V_\mu V^\mu, V_{\mu\nu} V^{\mu\nu}$ and $V_{\mu\nu} F^{\mu\nu}$, have been eliminated in favor of the vertex corrections following refs. [14, 48]. The $\delta g_f^Z$ and $g^h_{Zf}$ couplings include only a single generation of fermions such that $f = u_L, d_L, u_R$ and $d_R$. However, we will assume that these couplings are extended to all generations in a flavour universal way which is well justified if we assume Minimal Flavour Violation (MFV) [49].

| SILH Basis | Warsaw Basis |
|---|---|
| $\mathcal{O}_W = \frac{i}{2}\left(H^\dagger \tau^a \overleftrightarrow{D}^\mu H\right) D^\nu W^a_{\mu\nu}$ | $\mathcal{O}_Q^{(3)} = (\bar{Q}\sigma^a\gamma^\mu Q)(iH^\dagger\sigma^a\overleftrightarrow{D}_\mu H)$ |
| $\mathcal{O}_B = \left(H^\dagger \overleftrightarrow{D}^\mu H\right) \partial^\nu B_{\mu\nu}$ | $\mathcal{O}_Q = (\bar{Q}\gamma^\mu Q)(iH^\dagger\overleftrightarrow{D}_\mu H)$ |
| $\mathcal{O}_{HW} = ig(D^\mu H)^\dagger\sigma^a(D^\nu H)W^a_{\mu\nu}$ | $\mathcal{O}_R^u = (\bar{u}_R\gamma^\mu u_R)(iH^\dagger\overleftrightarrow{D}_\mu H)$ |
| $\mathcal{O}_{HB} = ig'(D^\mu H)^\dagger(D^\nu H)B_{\mu\nu}$ | $\mathcal{O}_R^d = (\bar{d}_R\gamma^\mu d_R)(iH^\dagger\overleftrightarrow{D}_\mu H)$ |
| $\mathcal{O}_{2W} = -\frac{1}{2}(D^\mu W^a_{\mu\nu})^2$ | |
| $\mathcal{O}_{2B} = -\frac{1}{2}(\partial^\mu B_{\mu\nu})^2$ | |

**Table 1**. Dimension-six operators contributing to VBF Higgs production at high $p_T^h$.

We show how these vertices give corrections to VBF Higgs production in Fig. 2. It is the subprocess, $qV \to qh$, a common part of all these diagrams, that receives corrections from the D6 lagrangian. As this hard process is a $2 \to 2$ process, its amplitude can be completely specified by two variables, for example, the Mandelstam variable, $t$, and an angle. Up to the leading terms in $t/m_Z^2$ in the EFT correction, we obtain for the $\mathcal{M}(qV_{T,L} \to qh)$ amplitude,

$$
\begin{aligned}
qZ_T \to qh &: g_f^Z \frac{\epsilon^* \cdot J_f}{v}\frac{2m_Z^2}{t}\left[1 + \left(\frac{g^h_{Zf}}{g_f^Z} + \kappa_{ZZ}\right)\frac{t}{2m_Z^2}\right], \\
qZ_L \to qh &: g_f^Z \frac{q \cdot J_f}{v}\frac{2m_Z}{t}\left[1 + \frac{g^h_{Zf}}{g_f^Z}\frac{t}{2m_Z^2}\right], \\
qW_T \to qh &: g_f^W \frac{\epsilon^* \cdot J_f}{v}\frac{2m_W^2}{t}\left[1 + \left(\frac{g^h_{Wff'}}{g_f^W} + \kappa_{WW}\right)\frac{t}{2m_W^2}\right], \\
qW_L \to qh &: g_f^W \frac{q \cdot J_f}{v}\frac{2m_W}{t}\left[1 + \frac{g^h_{Wff'}}{g_f^W}\frac{t}{2m_W^2}\right],
\end{aligned}
\tag{2.2}
$$

where $g_f^Z = g(T_3^f - Q_f s^2_{\theta_W})/c_{\theta_W}$, and $g_f^W = g/\sqrt{2}$; $J_f^\mu = \bar{f}\gamma^\mu f$ is the fermion current, the subscript $L$ ($T$) denotes the longitudinal (transverse) polarisation of the gauge boson, $q$ denotes its four-momentum and $\epsilon$ the associated polarisation vector. The reason we chose to write the EFT corrections to the amplitude as a function of the Mandelstam variable, $t$, and not $s$, can be understood from eq. (2.2). The EFT corrections are functions of only $t$. The additional angular variable required to specify the scattering kinematics does not appear. This is physically important as it means that the EFT corrections grow with the transverse momentum of the Higgs boson as this kinematic variable is highly correlated with $t$.

As we discussed already, the $qV \to qh$ subprocess, is related to the Higgs-strahlung process, $qq \to Vh$, by crossing symmetry, as shown in Fig. 1 such that the expressions in

| | EFT directions probed by high energy $ff \to Vh$ production |
|---|---|
| SILH Lagrangian [2] | $\frac{g}{c_{\theta_W}} \frac{m_W^2}{\Lambda^2}(2\lvert T_3^f\rvert \hat{c}_W - 2t_{\theta_W}^2 Y_f \hat{c}_B)$ |
| BSM Primaries [14] | $\frac{2g}{c_{\theta_W}} Y_f t_{\theta_W}^2 \delta\kappa_\gamma + 2\delta g_f^Z - \frac{2g}{c_{\theta_W}}(T_3^f c_{\theta_W}^2 + Y_f s_{\theta_W}^2)\delta g_1^Z$ |
| Universal observables | $\frac{2g}{c_{\theta_W}} Y_f t_{\theta_W}^2 (\delta\kappa_\gamma - \hat{S} + Y) - \frac{2g}{c_{\theta_W}}(T_3^f c_{\theta_W}^2 + Y_f s_{\theta_W}^2)\delta g_1^Z - \frac{2g}{c_{\theta_W}} T_3^f W$ |
| High Energy Primaries [36] | $-\frac{2m_W^2}{g c_{\theta_W}}(\lvert T_3^f\rvert a_q^{(1)} - T_3^f a_q^{(3)} + (1/2 - \lvert T_3^f\rvert) a_f)$ |

**Table 2**. The contact interaction couplings $g_{Zf}^h$, where $f = u_L, d_L, u_R, d_R$ expressed in different EFT parametrisation. For a given $f$ the expression can be read off from this table by substituting the corresponding value of the $SU(2)_L$ and $U(1)_Y$ quantum numbers $T_3^f$ and $Y_f$. Note that $\hat{c}_W = c_W + c_{HW} - c_{2W}$ and $\hat{c}_B = c_B + c_{HB} - c_{2B}$.

eq. (2.2) are identical to the corresponding ones for the $qq \to Vh$ process if we interchange $t \to s$. This is very significant as it implies that VBF Higgs production at high transverse momentum probes the same set of EFT operators as $qq \to Vh$ at high energies.

For $|t| \gg m_Z^2$, the correction proportional to the contact term coupling, $g_{Vf}^h$ dominates over all other terms[2]. The EFT correction due to $g_{Vf}^h$ grows with $t$ because, unlike the SM diagram, the corresponding diagram in Fig. 2 does not have an intermediate $V$-propagator. The $\kappa_{VV}$ contribution to the transverse amplitude also grows with $t$. This contribution, however, cannot interfere with the dominant longitudinal piece of the SM amplitude and is thus sub-leading with respect to the $g_{Vf}^h$ contribution. The EFT corrections due to the couplings, $\kappa_{\gamma\gamma}, \delta g_f^V, \delta\hat{g}_{\tau\tau}^h$ and $\delta\hat{g}_{VV}^h$, which include corrections to the Higgs decay, do not grow with $t$ at all.

We have checked explicitly that at high $|t|$ only the $g_{Vf}^h$ contributions are important and the effects of the other couplings are negligible provided all the couplings have a similar size, which is a reasonable assumption if a single cut-off is assumed for the different operators. If this assumption is relaxed, however, it is not immediately clear that only the $g_{Vf}^h$ couplings are important at high energies because the different anomalous couplings in Eq. 2.2 will be constrained at different level at the HL-LHC. We discuss this possibility in Appendix C in detail and show that the $g_{Vf}^h$ contribution dominates even if we take into account the different level of expected constraints on $g_{Vf}^h$, $\delta g_f^V, \delta\hat{g}_{\tau\tau}^h$ and $\delta\hat{g}_{VV}^h$. Finally in Appendix C we also show that the process $qq \to hjj$ involving an enhanced $hq\bar{q}$ coupling gives a negligible contribution in our analysis framework once constraints from other processes on $hq\bar{q}$ are taken into account.

Thus, VBF Higgs production at high transverse momentum is controlled by the five contact interaction couplings: $g_{Zf}^h$, with $f = u_L, u_R, d_L$ and $d_R$ and $g_{Wud}^h$. The operators contributing to these five couplings in the Warsaw basis are shown in Table 2. These contribute to these five anomalous couplings as follows,

[2]Note that because of the kinematics, $t$ is always negative in VBF.

$$
\begin{aligned}
g^h_{Zu_L} &= -\frac{g}{c_{\theta_W}}\frac{v^2}{\Lambda^2}(c^1_Q - c^{(3)}_Q) \\
g^h_{Zd_L} &= -\frac{g}{c_{\theta_W}}\frac{v^2}{\Lambda^2}(c^1_Q + c^{(3)}_Q) \\
g^h_{Zu_R} &= -\frac{g}{c_{\theta_W}}\frac{v^2}{\Lambda^2}c_{uR} \\
g^h_{Zd_R} &= -\frac{g}{c_{\theta_W}}\frac{v^2}{\Lambda^2}c_{dR}.
\end{aligned} \tag{2.3}
$$

The coupling $g^h_{Wud}$ is actually not independent of the above four contact interactions at the dimension-6 level, and is given by,

$$
g^h_{Wud} = c_{\theta_W}\frac{g^h_{Zu_L} - g^h_{Zd_L}}{\sqrt{2}}. \tag{2.4}
$$

Thus, only the four $g^h_{Zf}$ couplings are independent and these completely determine the EFT deviations for the VBF Higgs production at high transverse momentum.

In Table 2, we also show the mapping of these four $g^h_{Zf}$ couplings to other EFT parametrisations. In the first row of Table 2, we present the contributions of the universal (bosonic) operators of the SILH Lagrangian. We then show how these four couplings can be predicted/constrained by other independent measurements. The second row provides the mapping to the so-called BSM Primary basis of ref. [14]. In this basis, the correlations between different pseudo-observables are made explicit. For instance, in our case we can see how these 4 Higgs anomalous couplings can be predicted in terms of other measurements, namely, the couplings $\delta g^Z_f$ defined in eq. (2.2) that are strongly constrained by $Z$-pole measurements at LEP, and the anomalous TGCs, $\delta\kappa_\gamma$ and $\delta g^Z_1$ (in the notation of ref. [50]) that were constrained by the $WW$ production during LEP2. In the fourth row of Table 2, we write the 4 couplings in terms of only the "oblique"/universal pseudo-observables, i.e. the TGCs $\delta\kappa_\gamma$ and $\delta g^Z_1$ and the Peskin-Takeuchi $\hat{S}$-parameter [51] in the normalisation of ref. [52]. For a definition of these observables we refer to the Lagrangian presented in ref. [10] (see also ref. [53]). Finally, in the last line of Table 2, we connect these contact terms to the original definition of the high energy primaries in ref. [36].

## 3 Collider Analyses

In this section, we will provide all the details for our collider studies of the three $h \to \tau^+\tau^-$ channels and the $h \to \gamma\gamma$ channel. Utilising the fact that the EFT and SM contributions have the same form apart from a growth in the Mandelstam variable, $t$, we will use a two-step procedure to isolate our EFT signal. First, we will use sophisticated Neural Network (NN) techniques to optimally discriminate between the SM contribution from the other backgrounds in this section. We will then use the $p^h_T$ distribution to isolate the EFT effects from the SM contribution in the next section.

### 3.1 The $h \to \tau^+\tau^-$ channels

The SM Higgs decays 6.27% of the times into a pair of $\tau$-leptons. However, even though this is a significantly large branching ratio, the $\tau$-leptons are not stable, and hence we obtain three distinct final states, depending on the decay modes of the $\tau$s. The cleanest of these final states comprises two light leptons $(e, \mu)$. Thus, we categorise our final states as $\tau_\ell\tau_\ell$, $\tau_\ell\tau_h$ and $\tau_h\tau_h$, where $\tau_h$ is the hadronic remnant of the $\tau$ and is identified as a $\tau$-jet. All of these final states are associated with missing transverse energy, $\not{E}_T$, and at least two hard jets. We consider all three possibilities here. We closely follow the ATLAS analysis [54] and then use multivariate methods to optimally isolate the SM VBF Higgs production from the rest of the backgrounds. Our analysis is done at the centre-of-mass energy of 14 TeV.

The electron (muon) candidates are required to have minimum transverse momentum, $p_T$, of 15 GeV (10 GeV). The electrons (muons) are further required to be in an absolute pseudorapidity region of $|\eta| < 2.47$ (2.50). Furthermore, the electrons are disallowed in the transition region between the barrel and the endcap ($1.37 < |\eta| < 1.52$). Jets are reconstructed using the anti-$k_t$ algorithm [55] with a jet parameter of $R = 0.4$ and with a minimum $p_T$ of 20 GeV. The maximum allowed pseudorapidity range for the jets is required to be 4.5. In order to reconstruct $b$-jets, jets are matched with B-hadrons within $\Delta R(B, j) < 0.2$ and b-jets are required to have $|\eta| < 2.5$. We require a flat $b$-tagging efficiency of 70%. In our setup, a light jet (including $c$-jets) can fake a $b$-jet with a fake-tagging efficiency of 1%. We tag the hadronic $\tau$s with a tagging efficiency of 65%. Light jets can fake $\tau$-jets with a probability of 2.5%.

There are multiple backgrounds to consider for the $\tau\tau$ category. The dominant background comes from $\tau^+\tau^-$ jets excluding the Higgs diagrams. We generate this background keeping in mind that the $\tau$s can also emanate from off-shell photons. We also separately generate $\ell^+\ell^-+$ jets, where $\ell = e, \mu$. The other backgrounds include $t\bar{t}$ which we generate separately for the fully leptonic, semi-leptonic and fully hadronic cases, single top ($tq, tW$ and $t\bar{b}$), $\ell(\tau)\nu+$ jets and $h+$ jets with $h \to WW^*$, where the $W$s decay either leptonically or hadronically. The $h+$ jets samples are generated for the SM scenario as well as with the EFT couplings turned on. The Feynman rules are generated using the FeynRules package [56], through which we obtain the UFO [57] model. All the samples are then generated within the MadGraph version 2.6.5 [58] framework. The fragmentation, showering and hadronisation are done using Pythia version 8.2 [59]. For the full setup, we use the LO set of NNPDF2.3 parton distribution function [60] within the LHAPDF package [61]. For almost all the samples, we use the following cuts at the generation level: $p_{T,j/b(l)} > 20$ (8) GeV, $|\eta_{j(l)}| < 5$ (3), $\Delta R_{jj/bb/bj/ll/jl/bl} > 0.3/0.3/0.3/0/0.2/0.2$, $m_{j_1j_2} > 480$ GeV, where $j_1$ ($j_2$) is the hardest (second-hardest) quark in $p_T$ and can be a $b$-quark as well, and $\Delta\eta_{j_1j_2} > 2.5$. The $m_{j_1j_2}$ and $\Delta\eta_{j_1j_2}$ cuts are not applied to the single top samples at the generation level. All our event generations are at leading order (LO) in perturbation theory and we consider flat $K$-factors to roughly emulate the next-to-leading order (NLO) QCD effects. For the weak-boson fusion samples, the $K$-factor is almost a constant at 1.1 as a function of $p_{T,j_1}$ [62]. For the $l^+l^-+$ jets, $l\nu+$ jets ($l = e, \mu, \tau$), the NLO QCD

$K$-factor is roughly 1 as a function of $p_{T,V}$ [63], with $V$ being the vector boson $W/Z$. For the $t\bar{t}$ samples, we estimated the NNLO $K$-factor be around 1.63 [64]. For the single top channel, there are three sub-processes, *i.e.*, $t$-channel, $s$-channel and associated $Wt$ production. The most dominant of these three sub-processes is the $qb \to q't$ channel followed by the $bg \to tW$ channel. The smallest contribution comes from $q\bar{q}' \to t\bar{b}$. Upon following ref. [65], we consider a conservative $K$-factor of 1.1 [3].

To validate our analysis, we reproduce the rectangular cut-based analysis in the ATLAS paper [54] and find very similar results. The details of our rectangular cut-based analyses are mentioned in Appendix A.1.

To obtain our final results we use a Neural Network (NN) analysis. First, in addition to the $p_T$ and $|\eta|$ requirements, we also impose the following cuts, i.e. $m_{j_{1,2}} > 500$ GeV, $\Delta\eta_{j_{1,2}} > 2.5$ and $m_{\tau\tau}^{\text{col.}} < 300$ GeV. $m_{\tau\tau}^{\text{col.}}$ is the di-tau collinear mass [66]. The variables used in the NN training are shown in Table 6 in Appendix B. In order to prevent the NN from concentrating on the $m_{\tau\tau}$ peak, the sensitivity on the observable has been limited to 5 GeV bins. Table 7 shows the neural network results for the SM $h \to \tau\tau$ events as well as the other backgrounds, divided into three sub-regions, namely hadronic, semileptonic and leptonic. The first row shows the number of events at 0.3 ab$^{-1}$ luminosity after the preprocessing mentioned before and the following row shows the yielding number of events after the classification. The procedure and detailed results regarding the neural network are discussed in Appendix B.

### 3.2 The $h \to \gamma\gamma$ channel

Although the diphoton channel suffers from low branching fractions, due to its clean topology, it is relatively easy to separate from the background. With this in mind, we consider diphoton production with two jets topology to single out the VBF channel to achieve higher sensitivity in the aforementioned EFT operators further. By loosely following ref. [67], we construct two workspaces where first we design a cut-and-count based analysis. Then we studied on a Neural Network (NN) architecture which observed to increase our sensitivity.

Although Higgs-less diphoton with multijet production is the primary background in this channel, it has been shown that in low energy regimes, fake photons can have a significant impact on certain signal regions. The overall fractions of the background sources are presented as 78.7% from Higgsless diphoton channels, 18.6% from single-photon channels and 2.6% from multijet channels [67]. It is important to note that these fractions drastically change depending on the phase-space and the efficiency of the jet vertex tagging algorithm [68], where it has been shown that such techniques can reduce fake photon rates below 0.3% especially at higher energies [68–70]. To test this hypothesis, we generate the SM and other background samples using the aforementioned framework. All samples are generated with a specific set of cuts at the matrix-element level; minimum jet $p_T$ is taken to be 30 GeV, two leading jets' invariant mass is chosen to be greater than 500 GeV, and

[3] The NLO electroweak (EW) corrections have not been considered in this paper. As can be seen from ref. [63], the NLO QCD+EW $K$-factors for the aforementioned backgrounds can be less than 1 for higher values of $p_{T,V}$. Hence, we are overestimating these backgrounds to some extent.

| | Diphoton | | Ditau Hadronic | |
|---|---|---|---|---|
| | Other Background | SM Higgs | Other Background | SM Higgs |
| Preprocessing | 11710 | 4621 | 493756 | 27042 |
| Classifier output | 2251 | 3677 | 69561 | 21897 |

| | Ditau Semileptonic | | Ditau Leptonic | |
|---|---|---|---|---|
| | Other Background | SM Higgs | Other Background | SM Higgs |
| Preprocessing | 4190714 | 32343 | 9191181 | 9401 |
| Classifier output | 91803 | 21469 | 14408 | 3503 |

**Table 3**. Table shows NN results at 3 $\text{ab}^{-1}$ where we present the number of events for SM VBF Higgs production and the rest of the background before and after NN classification.

the pseudorapidity separation between the two leading jets is required to be greater than 1.5. As presented in Appendix A.2, these set of cuts has been chosen with respect to our cut-flow to populate the phase-space that is crucial for this analysis. The generated events are further showered and hadronised via PYTHIA version 8.2 [59].

The analysis of the event samples is performed within MADANALYSIS 5 version 1.8 [71]. The hadronised events are reconstructed using FASTJET version 3.3.2 [72] with the anti-$k_T$ algorithm [55], where the radius parameter has been chosen to be 0.4 with minimum transverse momentum of a reconstructed jet at 30 GeV. In order to simulate a simple detector environment, we apply particular tagging efficiencies on the $b$-jets, $c$-jets, hadronic taus and light jets. The tagging criteria are the same as discussed in the di-tau subsection 3.1.

In order to get definitive objects, detailed preselection requirements are applied. A photon candidate is required to have a minimum 25 GeV transverse momentum and is chosen to be within $|\eta| < 2.37$ and all the photon candidates are required to be separated from each other with $\Delta R > 0.4$. On the other hand, a jet candidate is required to be within $|\eta| < 4.5$. A clear distinction between photon and jet objects is essential in this analysis in order to suppress the background that might arise from misidentified objects. For this reason, we require the two photons which have a maximum of 15% hadronic activity within a cone radius of 0.4. After this point we branched our framework into two where cut-based analysis has been discussed in Appendix A.2 and NN analysis has been discussed in Appendix B. Table 3 shows the NN results presented at 3 $\text{ab}^{-1}$ integrated luminosity. It shows the event yield for the SM Higgs contribution as well as the other backgrounds at preprocessing stage and for the classifier output for this channel.

## 4 Projected sensitivity for EFT couplings

In this section, we present the final sensitivity projections for the EFT couplings. The NN techniques used to optimally isolate the SM VBF Higgs contribution from the other background processes in the previous section also isolate our signal, the EFT interference contribution. This is because as shown in Sec. 2 the dominant EFT contributions have a matrix element that is the same as the SM apart from growth with the magnitude

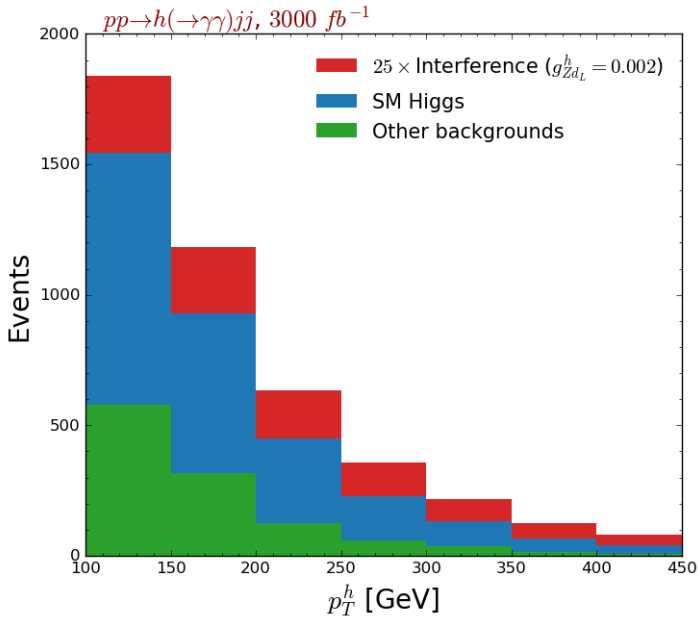


**Figure 3**. The final distribution with respect to the transverse momentum of the Higgs boson, $p_T^h$, for all the events in the diphoton channel that survive NN classification. The red represents the EFT interference contribution; blue represents the SM VBF Higgs contribution and green shows the other background sources which are all presented with 3 ab$^{-1}$ data. The EFT contribution corresponds to interference coming from $g^h_{Zd_L} = 0.002$ amplified 25 folds and with vanishing values for all other anomalous couplings in eq. (2.2).

of the Mandelstam variable $|t|$. We will now use this growth with $|t|$ to distinguish the EFT interference contribution from the SM; we will utilise the distribution of events with respect to $p_T^h$, a variable that is highly correlated to $t$, as the discriminant. We show the $p_T^h$ distribution for the diphoton channel in Fig. 3. The EFT interference contribution can be seen to grow as a fraction of the SM contribution with $p_T^h$. To derive the projected sensitivity for the EFT couplings, we define a $\chi^2$ function as follows.

$$\chi^2 = \sum_i^N \frac{(N_i^{exp} - N_i^{obs})^2}{\sigma_i^2} \ ,$$

where we take the SM as our null hypothesis. $N_i^{exp}$, denotes the expected number of events in the SM for the $i$th bin in the $p_T^h$ distribution. We will then assume that the number of events observed in the $i$-th bin, $N_i^{obs}$ is different from the SM due to the presence of EFT couplings. Finally, $\sigma_i$ includes both the statistical and systematic uncertainties,

$$\sigma_i = \sqrt{N_i^{exp} + (\Delta_{sys} N_i^{exp})^2} \ ,$$

$\Delta_{sys}$ being the percentage of systematic uncertainty.

As can be seen from Fig. 3, although the EFT interference contribution steadily grows with $p_T^h$ as a fraction of the SM, the absolute value for the excess keeps decreasing. As a result of the $\chi^2$ function initially increases with $p_T^h$, peaks at an intermediate value around

$p_T^h \sim 300$ GeV and then decreases again. As discussed in Sec. 2 the four contact couplings $g^h_{Zf}$, with $f = u_L, d_L, u_R, d_R$, give dominant contributions in the high $p_T^h$ region. In a hadron collider, it is impossible to disentangle initial states for the process $qV_{T,L} \to qh$ that gets corrections from these contact terms (see Fig. 2). Thus only a linear combination of these four contact couplings appears in the EFT interference term at a given $p_T^h$. As bins around $p_T^h \sim 300$ GeV yield maximum sensitivity, the direction probed by VBF Higgs production turns out to be the above linear combination at this $p_T^h$ value,

$$\begin{aligned} g_{Zf}^{h(VBF)} &= g^h_{Zu_L} - 0.94\, g^h_{Zd_L} - 0.15\, g^h_{Zu_R} + 0.04\, g^h_{Zd_R} \\ &= (-0.04\, c_Q^1 + 1.4\, c_Q^{(3)} + 0.1\, c_{uR} - 0.03\, c_{dR})\xi \end{aligned} \tag{4.1}$$

where $\xi = v^2/\Lambda^2$. Here the second line expresses the EFT direction in terms of Warsaw basis operators defined in Table 2. If $p_T^h$ is varied, the coefficient of $g^h_{Zd_L}$ varies only by a few per cent whereas the coefficients of $g^h_{Zu_R}$ and $g^h_{Zd_R}$ can decrease by as much as 30 %. The left-handed couplings dominate the above direction as the $W$-boson luminosity is much larger than the $Z$-luminosity in VBF processes and the right-handed couplings cannot contribute to the $qW_L \to qh$ process.

For our final sensitivity estimate we combine all four final states by adding their individual $\chi^2$ functions. We find each final state has a comparable contribution to the final $\chi^2$, value which emphasises the importance of including all these four channels. Including only the bins for which $p_T^h < 400$ GeV we obtain our final bound for an integrated luminosity of 3 ab$^{-1}$ (0.3 ab$^{-1}$) and 10% systematic uncertainty as

$$|g_{Zf}^{h(VBF)}| < 0.003\ (0.005) \tag{4.2}$$

at 68% CL. Assuming the Wilson coefficients in eq. (2.3) are $\mathcal{O}(1)$ the above bound can be translated to the following bound on the cut-off scale,

$$\Lambda > 3.9 \text{ TeV } (\Lambda > 2.9 \text{ TeV}) \ .$$

The $p_T^h < 400$ GeV cut ensures that most events safely respect the EFT validity requirement $\sqrt{t} < \Lambda$. For strongly coupled UV completions, the values of the Wilson coefficients can be much larger than unity giving much larger values for $\Lambda$ and thus higher allowed $p_T^h$ values. This, however, would not lead to much better bounds as the most sensitive bins are around $p_T^h \sim 300$ GeV. Our results, thus, do not depend too much on whether the UV completion is weakly or strongly coupled.

**Combination with diboson channels:** As we discussed in Sec. 1, the diboson channels $pp \to VV/Vh$ at high energies, and the VBF Higgs production process at high $p_T^h$ considered here, probe the same set of four operators. Of these $WZ$ production was studied in ref. [36], $Zh$ production in ref. [37, 43] and $Wh$ production in ref. [43]. Compiling the 68 % CL HL-LHC bounds obtained in these papers with our result in eq. (4.2), we obtain in terms

of Warsaw basis operators,

$$
\begin{aligned}
|(-0.04\ c_Q^1 + 1.4\ c_Q^{(3)} + 0.1\ c_{uR} - 0.03\ c_{dR})\xi| &< 0.003 \qquad [VBF] \\
|(-0.18\ c_Q^1 + 1.3\ c_Q^{(3)} + 0.3\ c_{uR} - 0.1\ c_{dR})\xi| &< 0.0005 \qquad [Zh] \\
|c_Q^{(3)}\xi| &< 0.0004 \qquad [Wh] \\
-0.0004 < c_Q^{(3)}\xi &< 0.0003 \qquad [WZ]
\end{aligned} \tag{4.3}
$$

for 3 ab$^{-1}$ integrated luminosity where $\xi = v^2/\Lambda^2$. It is clear that all these different processes constrain a different direction in four-dimensional space of high energy primaries. As the $WZ$ and $Wh$ process constrain the same direction, the above bounds still leave a flat direction unconstrained. An additional bound from the $WW$ production process will thus close all flat directions and allow us to bound all the four operators simultaneously. The $WW$ process was studied in ref. [38] and it is clear from the results that it puts strong bounds on yet another complementary direction. It is, however, difficult to infer the direction probed by the $pp \to WW$ process from the results of ref. [38] as in this paper the $WW$ and $WZ$ channels have been presented in a combined way including both the interference and EFT squared contributions.

As mentioned in Sec. 1 the VBF production of gauge bosons will also probe the same four-dimensional space. These channels should thus be added to over-constrain the system and maximally constrain the high energy primaries.

**Comparison with LEP bounds:** Using Table 2 we can also write this direction in terms of other pseudo-observables already constrained by LEP,

$$
\begin{aligned}
g_{Zf}^{h(VBF)} &= 2\ \delta g_{Zu_L} - 1.9\ \delta g_{Zd_L} - 0.3\ \delta g_{Zu_R} + 0.08\ \delta g_{Zd_R} - 0.04\ \delta\kappa_\gamma - 1.08\ \delta g_1^Z \\
g_{Zf}^{h(VBF)} &= -0.04\ (\delta\kappa_\gamma - \hat{S} + Y) - 1.08\ \delta g_1^Z - 1.4\ W
\end{aligned} \tag{4.4}
$$

where the first line applies to the general case and the second line to the universal case. The LEP bounds on the above pseudo-observables are given by the second column of Table 4. The LEP bound on the full direction is thus given by the largest term in the right hand sides of the above equations which is $g_{Zf}^{h(VBF)} \lesssim 1.08\ \delta g_1^Z \lesssim 0.03$ which is almost an order of magnitude weaker than the bound in eq. (4.2).

One can also assume that there is no cancellation between the different terms in eq. 4.4. This allows us to require that each term in the right-hand side respects the bound in eq. (4.2). We then obtain the results in the first column of Table 4. We see that with this 'no tuning' assumption, relative to LEP bounds, the results of this work can lead to much stronger bounds on TGCs, comparable bounds on deviations of $Z$ coupling to quarks and weaker bounds for the oblique parameters.

## 5 Conclusions

It is increasingly being recognised that the LHC is a precision machine. This is because various examples are beginning to appear where certain operators can be probed very

| | Our Projection | LEP Bound |
|---|---|---|
| $\delta g^Z_{u_L}$ | $\pm 0.001$ ($\pm 0.002$) | $[-0.0042, -0.001]$ |
| $\delta g^Z_{d_L}$ | $\pm 0.002$ ($\pm 0.003$) | $[0.0013, 0.0033]$ |
| $\delta g^Z_{u_R}$ | $\pm 0.01$ ($\pm 0.02$) | $[-0.0071, -0.0001]$ |
| $\delta g^Z_{d_R}$ | $\pm 0.04$ ($\pm 0.06$) | $[0.0108, 0.0212]$ |
| $\delta g^Z_1$ | $\pm 0.003$ ($\pm 0.005$) | $[-0.03, 0.013]$ |
| $\delta \kappa_\gamma$ | $\pm 0.08$ ($\pm 0.12$) | $[-0.063, 0.026]$ |
| $\hat{S}$ | $\pm 0.07$ ($\pm 0.12$) | $[-0.0003, 0.0011]$ |
| $W$ | $\pm 0.002$ ($\pm 0.004$) | $[-0.0006, 0.0000]]$ |
| $Y$ | $\pm 0.07$ ($\pm 0.12$) | $[-0.0003, 0.0003]$ |

**Table 4**. Table shows the comparison of 68 % CL bounds extracted from the VBF analyses with the existing LEP bounds. The bounds outside the parentheses are projections for 3 ab$^{-1}$ data and those inside are for 0.3 ab$^{-1}$ data. To get our projection, we demand that each term in eq. (4.4) respects the bound in eq. (4.2). The LEP bounds on the $Z$-boson couplings to quarks, $\delta g^Z_f$, are taken from ref. [12], the bounds on the charged TGCs are taken from ref. [73], the bound on $\hat{S}$ from ref. [74], and the bounds on the $W, Y$ observables from ref. [52].

precisely, for instance, by studying the high energy tails of different processes. One of the best examples is that of the high energy primaries, the four operators that dominate the high energy tails in diboson production, including the Higgs-strahlung process. In this work, we highlight how VBF Higgs production probes a linear combination of the same operators given by eq. 4.1. Our results are complementary to those obtained in the Higgsstrahlung, and diboson processes as all these processes probe different directions in this four-dimensional space (see eq. (4.3)). Our final projection for the HL-LHC bounds on the direction corresponding to VBF Higgs production is per mille level (see eq. (4.2)) which translates to a multi-TeV bound on the new physics scale given in eq. (4.3). These bounds far surpass the existing LEP bounds (see discussion below eq. (4.4) and Table 4).

As far as Higgs and electroweak physics is concerned, the highest energy scales the LHC will probe indirectly might well be via a precise measurement of these high energy primaries[4]. These may therefore become part of the legacy measurements of LHC. The VBF Higgs production process studied in this work would be an important and integral part of this program.

## A Rectangular Cut-Based Analyses

### A.1 The $h \to \tau^+\tau^-$ channels

Overall, we follow the relevant cuts listed in tables 3 and 4 of ref. [54]. For the $\tau_\ell\tau_\ell$ case, we demand $m_{j_1j_2} > 500$ GeV instead of 800 GeV as mentioned in the paper. We use the tight VBF category for the $\tau_h\tau_h$ case. For our lepton isolation, we require that the

[4]Another example where LHC can indirectly probe very high scales by studying high energy tails is the Drell-Yan process as shown in ref. [75].

hadronic activity around an isolated lepton $(e, \mu)$ within a cone of $\Delta R = 0.2$, should not exceed 10% of its $p_T$. With the rectangular cut-based analyses, we get the following results. Following ATLAS, for the rectangular cut-based analysis, we dissect the $\tau_\ell \tau_\ell$ scenario into the same flavour and opposite flavour cases. For the same flavour case ($ee$ or $\mu\mu$), the number of events from the SM Higgs signal ($S$) and other background ($B$)[5] at an integrated luminosity ($\mathcal{L}$) of 300 fb$^{-1}$ are 119 and 953, yielding a significance $(S/\sqrt{S+B}) \sim 3.64$. For the different flavour case, $S \sim 149$, $B \sim 1361$ and significance $\sim 3.84$. For the $\tau_\ell \tau_h$ case, $S \sim 436, B \sim 10332$ and significance $\sim 4.20$. Finally, for the $\tau_h \tau_h$ case, we obtain $S \sim 686, B \sim 9647$ and significance $\sim 6.74$.

### A.2 The $h \to \gamma\gamma$ channel

As mentioned in section 3.2, we construct a cut-based analysis by loosely following ref. [67]. After the preselections mentioned above, in order to identify the VBF channel, we use standard VBF cuts where the $b$-jets are vetoed, and we require at least two jets where the leading two are separated into two hemispheres with $|\Delta\eta| > 3$. To identify the boosted VBF topology, the general recipe requires an invariant mass cut between two leading jets at the order of 300-400 GeV as applied in ref. [67]. However, we observe higher sensitivity to the EFT operators achieved when higher $M_{jj}$ requirement is applied. In addition to the isolation requirement presented in section 3.2, we further demand additional angular requirements between the photons and jets to restrict the phase-space for additional emissions. The minimum angular separation between jets and photons, $\Delta R^{min}_{\gamma j}$ is observed to perform as a tremendous discriminatory tool against background rejection. Limiting $\Delta R^{min}_{\gamma j} > 1.5$ is observed to separate the background from the signal events without any loss of the desired phase-space. We also require an azimuthal angle separation between the two-jet and two-photon systems. Although this requirement does not propose relatively active discrimination, it has been shown to be a powerful tool to suppress theoretical uncertainties and veto additional jets in the event sample [67]. These series of requirements cause largely boosted samples, where although our signal does not show any particular azimuthal separation preference, we observe that the background is dominated by highly separated jets in azimuthal angle. For this reason, we require angular separation between the two leading jets to be less than 2. Finally, the most effective cut was expectedly the invariant mass of the two-photon system, which is chosen to be within $125 \pm 3$ GeV. Also, the reconstructed Higgs rapidity is required to lie between the two jets. Table 5 summarises all the cuts and their relative efficiencies for both the signal and the background samples. At the bottom portion of the table, we show various discriminatory variables to asses the quality of the yield events. All results are presented at 0.3 ab$^{-1}$. It is important to note that we also generate single-photon samples to quantify the effect of fake photon contamination in the sample. For this we use the SFS module of MadAnalysis 5 [76] to simulate a light jet mis-tag rate of 0.3%. However, we observe that out of a million events, we do not

[5] We must note that, here and in what follows, we are loosely referring to the SM VBF as the 'signal' ($S$), and the the rest of the samples as 'background' ($B$). In our final analyses, the SM VBF is of course part of the background and the EFT contribution is the true signal.

have any to pass the Higgs mass requirement. Thus, in order to save valuable computation time, we assume that such effects are insignificant in such boosted phase-space regimes.

| | Background | | SM Signal | |
|---|---|---|---|---|
| | Events | $\varepsilon$ [%] | Events | $\varepsilon$ [%] |
| Presel. | 369176.1 | - | 1365.2 | - |
| Njet$\geq 2$ | 286704.2 | 77.66 | 1144.9 | 83.87 |
| Bjet veto | 274869.6 | 95.87 | 1108.2 | 96.79 |
| $\lvert\Delta\eta_{jj}\rvert > 3$ | 164813.0 | 59.96 | 838.0 | 75.62 |
| $\eta_{j_1} \cdot \eta_{j_2} < 0$ | 161844.1 | 98.20 | 827.2 | 98.71 |
| $M_{jj} > 600$ [GeV] | 93105.4 | 57.53 | 658.6 | 79.62 |
| $N_\gamma = 2$ | 20244.1 | 21.74 | 432.8 | 65.72 |
| $I_\gamma^{R=0.4} < 15\%$ | 19876.9 | 98.19 | 431.9 | 99.77 |
| $\Delta R_{\gamma j}^{min} > 1.5$ | 8379.1 | 42.15 | 382.6 | 88.58 |
| $\lvert\Delta\Phi_{\gamma\gamma,jj}\rvert > 1.5$ | 7896.7 | 94.24 | 373.4 | 97.62 |
| $\Delta\Phi_{j_1,j_2} < 2$ | 2393.7 | 30.31 | 227.7 | 60.97 |
| $122 < M_{\gamma\gamma} < 128$ [GeV] | 88.1 | 3.68 | 226.9 | 99.64 |
| $y_{j_{1,2}}^{min} < y_h < y_{j_{1,2}}^{max}$ | 78.0 | 88.63 | 223.2 | 98.40 |
| $S/B$ | 286.06% | | | |
| $S/(B+S)$ | 74.10% | | | |
| $S/\sqrt{B}$ | 25.27 | | | |
| $S/\sqrt{S+B}$ | 12.86 | | | |

**Table 5**. Cut-flow for diphoton channel presented. Both background and signal samples separated in to two columns where on the left of each one yielding number of events are presented at 0.3 ab$^{-1}$ and the right column shows the relative efficiency of each cut. At the bottom part certain ratios are presented. Note that what's referred to as signal ($S$) here is the SM VBF. In the final analysis, the SM VBF is part of the background.

## B Neural Network Analysis

In recent years, the particle physics community has been increasingly adapting to the use of *deep neural networks* (DNN) in challenging signal characterisation problems [77–87]. The `Keras` library [88] offers a python-based, flexible framework using feed-forward networks [89–91] to create mashed layers with connected neurons (nodes). In order to increase our sensitivity to the operators presented above, we design a simple workspace to determine achievable sensitivities with different neural network architectures. We assume to have certain common properties to apply on each architecture. Each architecture is optimised using the `Adam` algorithm [92] and to accommodate multi-class classification, sparse categorical crossentropy loss function is used where the crossentropy is defined as

$$\mathcal{H}(p^{truth}, p^{pred}) = -\frac{1}{N}\sum_{i=1}^{N} p_i^{truth} \log p_i^{pred} \ .$$

Here $p^{truth}$ refers to the vector of the truth values and $p^{pred}$ is the vector of prediction probabilities. We use softmax activation in the output, which is essentially a combination of sigmoid functions for each output class. Furthermore, instead of traditionally used sigmoid activation for each layer, we use rectified linear unit (`ReLU`) [6]. Each model is initialised with a learning rate of $1 \times 10^{-5}$, and the learning rate decayed to its half if the loss value of the validation sample does not improve for 20 epochs. Each training runs for 200 epochs with a requirement of at least 0.01 unit improvement on the loss of the validation sample. The samples that do not satisfy this condition for 50 epochs are terminated before the end of the 200 epochs. The class weights are normalised with respect to their occurrences in the training sample in order to compensate for the difference of the population of each class. We investigate four different signal regions, namely diphoton, and ditau decaying hadronically (hereafter hadronic), ditau decaying semileptonically (hereafter semi-leptonic) and ditau decaying leptonically (hereafter leptonic). Each sample has a separate set of backgrounds, and in order to save computation time, we only use the dominant background samples that have the greatest impact in a given signal region.

In order to prevent over-training for each signal region, we use the dropout and kernel-regularisation methods. Each layer is required to have 25% probability of dropping each node in order to prevent dependency on a given parameter. Additionally, each hidden layer is supported via the $L_2$ kernel regularisation [94] with a penalty strength of $10^{-2}$. This penalty term is directly reflected on the loss as

$$\text{Loss} := \mathcal{H} + \lambda \sum_{i=1}^{N} (||\omega_i||^2 + ||b_i||^2) \ ,$$

where $\omega_i$ is the weight of the node, $b_i$ is the bias and $\lambda$ is the penalty strength. Lower and higher values of the penalty strength are tested as well which lead to the signal regions to over train and, in case of the latter, cumulative accuracy has been observed to drop below 70% respectively.

Each signal region undergoes a specific preprocessing before training. In addition to aforementioned preselection requirements, diphoton sample is required to have at least two jets, and two isolated photons in the data sample and the invariant mass of the two leading jet is required to be at least 550 GeV. In order to prevent the NN from concentrating only on the diphoton invariant mass, we require it to be within $125 \pm 3$ GeV window and remove the $M_{\gamma\gamma}$ from the training parameters. Otherwise, it has been observed that the NN avoids all other variables and concentrate solely on $M_{\gamma\gamma}$ peak, which has been tested up to 5 GeV resolution for the distribution. Although this reduces the accuracy of the test sample significantly, it is necessary to avoid the NN to concentrate only on the sharp invariant mass peak. In order to remain in the realm of VBF, we also veto all $b$-jets and require $|\Delta\eta_{jj}| > 1.5$, without requiring them to be on different hemispheres. Following the same recipe, the ditau samples are preprocessed by requiring $M_{jj} > 500$ GeV, $|\Delta\eta_{jj}| > 2.5$ and $M_{\tau\tau} < 300$ GeV. For the leptonic final state, we demand two isolated leptons, for the semi-leptonic final state, one isolated lepton and one hadronic tau. Finally, for the

[6] See ref. [93] for advantages of the `ReLU` activation over sigmoid function.

hadronic final state, we require two hadronic taus and veto events with isolated leptons. As before, in order to prevent the NN to concentrate on the invariant mass peak of two taus, we require it to have a resolution of 5 GeV. Table 6 summarises all the parameters that have been used for each region. Here, $p_T^{jj}$ refer to the combined vectorial $p_T$ of the two-hardest jets, $\tau_{1h/2h}$ refer to the visible part of the hardest and the second-hardest $\tau$-lepton, which can be $e, \mu$ or $\tau_h$, $m_{T2}$ is the stransverse mass variable [95, 96], $\Delta\phi$s are the azimuthal angle separations, $x_{1/2}$ are the visible momentum fractions for the two $\tau$ leptons and $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. All the other variables are self-explanatory. For the semi-leptonic case, we also have an additional variable, the transverse mass, $m_T$.

| | |
|---|---|
| Ditau Hadronic | $p_T^{j_1/j_2/\tau_{1h}/\tau_{2h}}$, $p_T^{jj}$, $\frac{p_T^{j_1}}{p_T^{j_2}}$, $\Delta R(j_1, j_2)$, $p_T^{vis}(h)$, $p_T^{miss}$, $M_{T2}$, $M_{jj}$, $\lvert\Delta\eta(j_1, j_2)\rvert$, $x_1$, $x_2$, $\Delta\phi(\tau_{1h}/\tau_{2h}, \mathbf{p}_T^{miss})$, $\Delta\phi(j_1, j_2)$, $M_{\tau\tau}^{\text{col.}}$, Jet Multip., $\frac{p_T^{\tau_{1h}}}{p_T^{\tau_{2h}}}$, $\Delta R(\tau_{1h}, \tau_{2h})$, $\Delta R(\tau_{1h}/\tau_{2h}, j_1/j_2)$, $M(j_1/j_2, \tau_{1h}/\tau_{2h})$, $\lvert\Delta\eta(\tau_{1h}, \tau_{2h})\rvert$, $M_{\tau_h\tau_h}$ |
| Ditau Semileptonic | $p_T^{j_1/j_2/\ell/\tau_h}$, $p_T^{jj}$, $\frac{p_T^{j_1}}{p_T^{j_2}}$, $\Delta R(j_1, j_2)$, $p_T^{vis}(h)$, $p_T^{miss}$, $M_{T2}$, $M_{jj}$, $\lvert\Delta\eta(j_1, j_2)\rvert$, $x_1$, $x_2$, $\Delta\phi(\ell/\tau_h, \mathbf{p}_T^{miss})$, $\Delta\phi(j_1, j_2)$, $M_{\tau\tau}^{\text{col.}}$, Jet Multip., $\frac{p_T^{\ell}}{p_T^{\tau_h}}$, $\Delta R(\ell, \tau_h)$, $\Delta R(\ell, j_1/j_2)$, $\Delta R(j_1/j_2, \tau_h)$, $M(j_1, j_2, \ell, \tau_h)$, $\lvert\Delta\eta(\ell, \tau_h)\rvert$, $M_{\ell\tau_h}$, $M_T$ |
| Ditau Leptonic | $p_T^{j_1/j_2/\ell_1/\ell_2}$, $p_T^{jj}$, $\frac{p_T^{j_1}}{p_T^{j_2}}$, $\Delta R(j_1, j_2)$, $p_T^{vis}(h)$, $p_T^{miss}$, $M_{T2}$, $M_{jj}$, $\lvert\Delta\eta(j_1, j_2)\rvert$, $x_1$, $x_2$, $\Delta\phi(\ell_1/\ell_2, \mathbf{p}_T^{miss})$, $\Delta\phi(j_1, j_2)$, $M_{\tau\tau}^{\text{col.}}$, Jet Multip., $\frac{p_T^{\ell_1}}{p_T^{\ell_2}}$, $\Delta R(\ell_1, \ell_2)$, $\Delta R(j_1/j_2, \ell_1/\ell_2)$, $M(j_1, j_2, \ell_1, \ell_2)$, $\lvert\Delta\eta(\ell_1, \ell_2)\rvert$, $M_{\ell\ell}$ |
| Diphoton | $M_{jj}$, $\eta_{j_1/j_2}$, $p_T^{h/\gamma_1/\gamma_2/j_1/j_2}$, $\Delta\phi(j_1, j_2)$, $\Delta\phi(h, j_1/j_2/\gamma_1/\gamma_2)$, $\Delta R(h, j_1/j_2/\gamma_1/\gamma_2)$, $M_{eff}$, $H_T$, $\Delta R(\gamma_1, \gamma_2)$, $y_{j_1/j_2/h}$, $\Delta R_{\gamma j}^{min}$, $\Delta\Phi(h, jj)$, $p_T^{hjj}$, $\lvert\eta_h - 0.5(\eta_{j_1} + \eta_{j_2})\rvert$ |

**Table 6**. Parameters which are used in corresponding NN training. Observables shown with comma, $\mathcal{O}^{i,j}$, represents a system of corresponding $i^{th}$ leading and $j^{th}$ leading particles and the ones shown with slash, $\mathcal{O}^{i/j}$, represents the usage of the same observable for both $i^{th}$ leading and $j^{th}$ leading reconstructed object separately.

Hyper-parameter optimisation is a challenging problem in machine learning. To further understand the phase-space and the effect of the layers, we devise a simple scanning procedure which starts from a linear model and increases number of hidden layers and nodes depending on the performance of the NN. To simplify the process, the number of nodes is chosen to be a certain multiple of the number of input parameters. Table 3 shows the general results of the NN where preselection gives the number of events remaining after the preprocessing and the classifier output is the number of events left after the classification process. All the results are presented at 3 ab$^{-1}$ integrated luminosity. In Table 7, we

| | Diphoton | Ditau Hadronic | Ditau Semileptonic | Ditau Leptonic |
|---|---|---|---|---|
| Test accuracy | $80.04\% \pm 0.01\%$ | $76.04\% \pm 1.19\%$ | $77.23\% \pm 3.31\%$ | $70.31\% \pm 2.43\%$ |
| Signal precision | 98% | 92% | 75% | 94% |
| Signal TPR | 79% | 81% | 37% | 66% |
| Signal F1-Score | 88% | 86% | 50% | 78% |
| Layers $\times$ Nodes | $5 \times 675$ | $3 \times 840$ | $4 \times 1015$ | $5 \times 1400$ |

**Table 7**. Results for NN classification presented. Table shows the statistics of the NN calculated for the test sample. Last row of the table shows the number of hidden layers and node per each layer.

present corresponding classification results for signal in the training sample. As expected, we observe a larger $S/B$ ratio for diphoton channel with respect to the ditau channels. The test accuracy is measured via 10-fold validation in order to see the fluctuations in the results. Although the diphoton channel gives the least amount of uncertainties, none of the uncertainties goes beyond 4% of the mean test accuracy. We also present the signal precision, the true positive rate (TPR) and the F1-score [97] for the test sample. The last row of the Table 7 shows the number of hidden layers and their corresponding number of nodes in each layer. Fig. 4 shows the corresponding receiver operating characteristic (ROC) curve where dark blue, red, green and light blue curves represents hadronic, semileptonic, leptonic and diphoton channels. Area under the ROC (AUC) curve with respect to TPR and false positive rate (FPR) has been attached to each label. Black dashed line presents a reference for random guess.

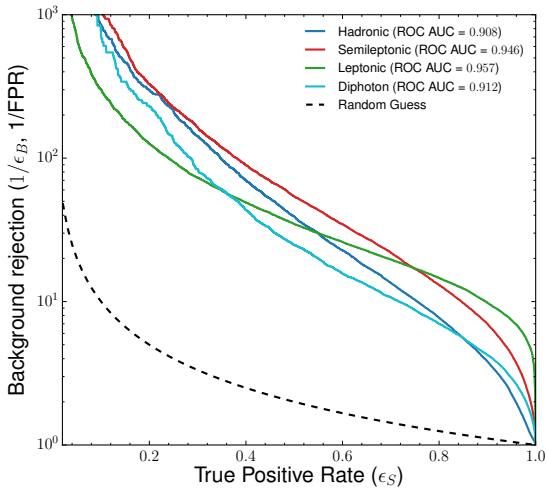


**Figure 4**. ROC curve for diphoton and ditau signal regions. Each label presented with the corresponding area under TPR-FPR curve. The colors blue, red, green and cyan represents ditau hadronic, ditau semileptonic, ditau leptonic and diphoton signal regions.

It is important to understand how the neural network learns and interprets the data,

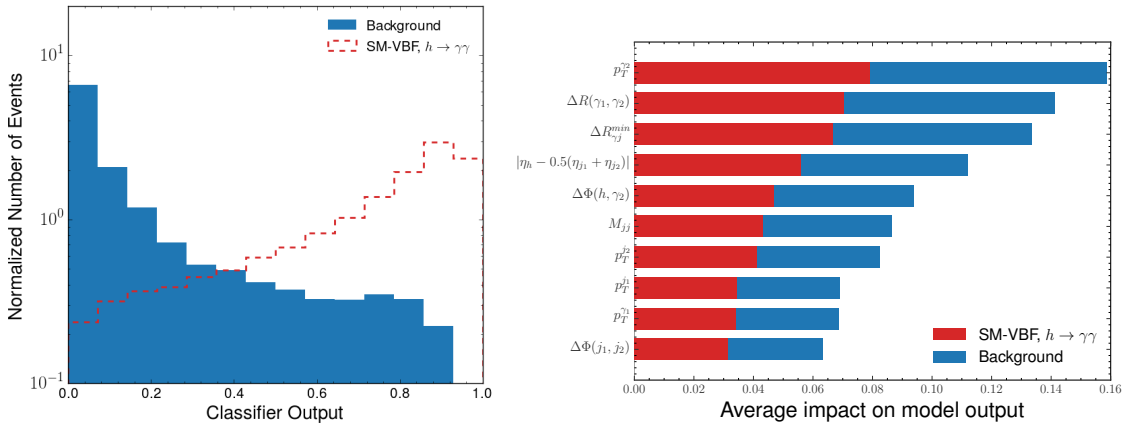


**Figure 5**. Left panel shows the classification output for the diphoton channel and the right panel shows the average SHAP value for ten leading parameter that is used in the training. Red shows the signal and blue represents the background in both figures.

where understanding such features can help in optimising the cut-based analyses as well. For this reason, we adapt the SHapley Additive exPlanations (SHAP) [98] method. The SHAP value shows the average of the marginal contributions of the input parameters to the neural network. In order to measure this value, we use the same training and test samples where the SHAP explainer trained with 2000 events from the training sample and the SHAP values are extracted using 1000 random events from the test sample. For diphoton channel, the most important ten observables with SHAP values are presented in the right panel of Fig. 5 where the signal values are represented with red and the background values are represented with blue bars where the average SHAP value has been divided by the contributions coming from the signal and the background. Although the SHAP values are relatively low, one can immediately see the importance of the angular observables and transverse momenta of the second leading photon in the NN. The left panel of Fig. 5 shows the classifier output for the NN architecture where the red line shows the signal and the blue bars shows the background sample. In figures 6, 7 and 8, we show the classifier output (on the left) and the SHAP values for the ten parameters that have the biggest impact on the classification for hadronic, semi-leptonic and leptonic final states respectively. As seen in the SHAP values, the ditau signal regions mostly rely on angular observables between final state particles. One can see in the classifier outputs, $\tau\tau$+jets is the most dominant background. The loss of sensitivity can also be observed in the classifier outputs where the leptonic signal region can not reach beyond 70%. This outcome renders semi-leptonic and leptonic signal regions as not optimal for sensitivity studies of the particular operator in hand.

In all, we observe much superior results in the diphoton signal region in terms of both statistical significance and accuracy of the NN. As presented in Appendix A, compared to cut-based analysis our $S/B$ ratio is significantly lower in the NN analysis which is by design. We observed that by increasing yielding number of signal events one can populate the high

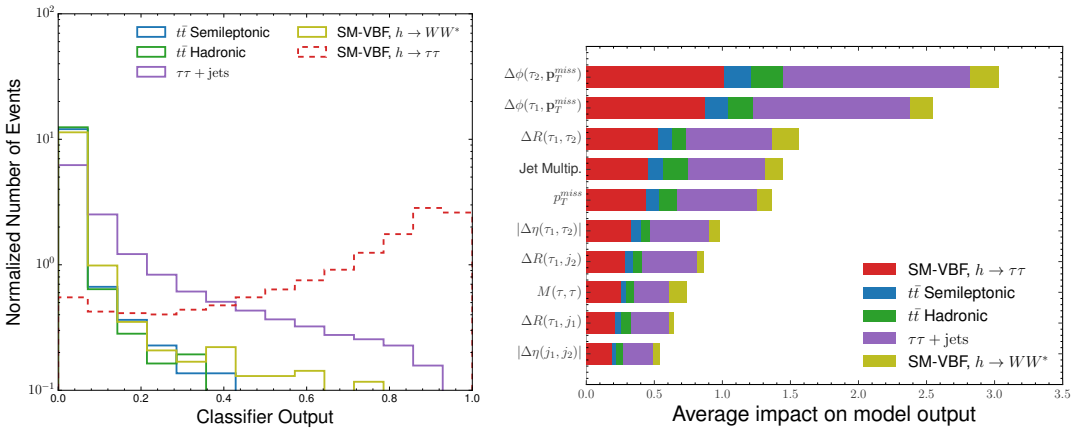


**Figure 6**. Same as Fig. 5 for hadronic channel.

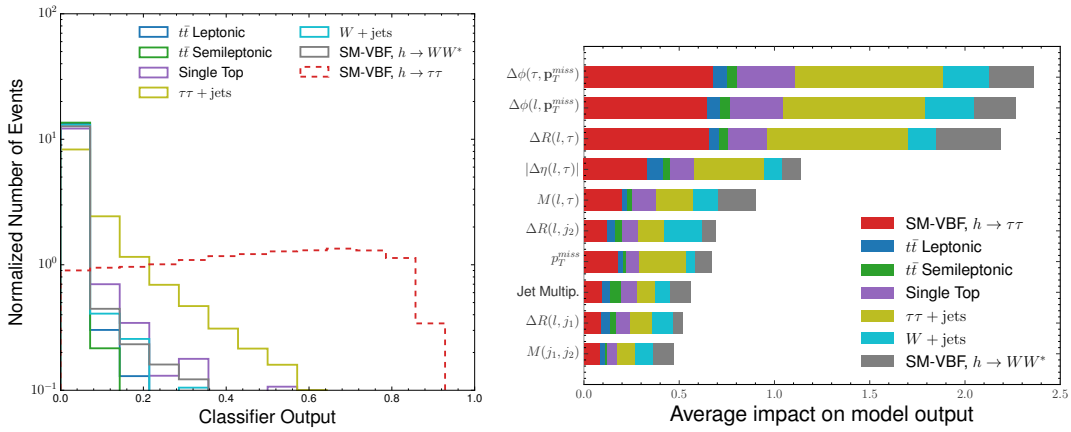


**Figure 7**. Same as Fig. 5 for semileptonic channel.

energetic regions that are crucial for the sensitivity of EFT operators. We observe up to 38% improvement depending on the choice of confidence level, luminosity and systematic uncertainty in the operator sensitivities due to large number of signal events left after the classification. This is because large statistical significance in the rectangular cut-based approach has been achieved with less yielding events which degrades the impact of the events where EFT effects are most prominent. In the ditau signal regions, due to the vast amount of background sources, the classification accuracy is lower than the diphoton channel. Expectedly, the $\tau\tau+$ jets background is the most dominant background source for all ditau subregions. All these results also compared with a boosted decision tree (BDT) algorithm. Although the BDT results were slightly less significant compared to the NN, we observed that both methods were giving priorities to similar observables (as represented by average SHAP values in NN case) to increase signal significance.

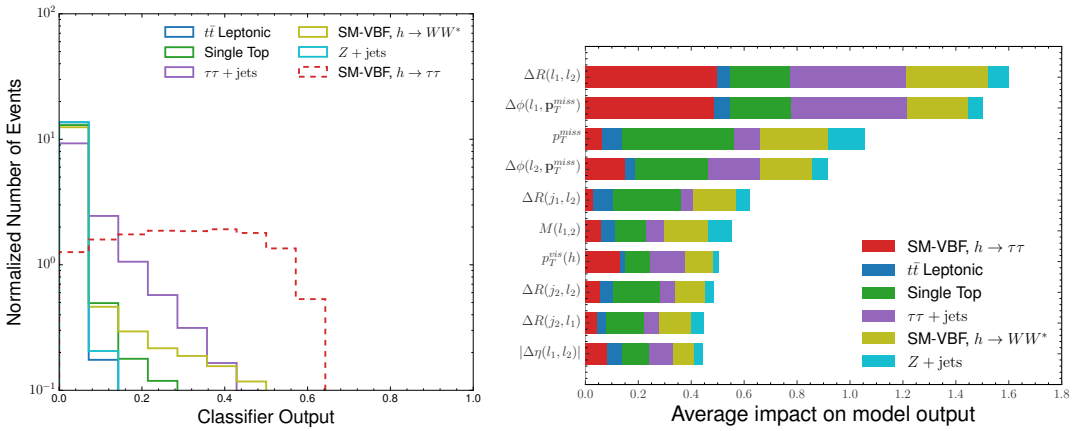


**Figure 8**. Same as Fig. 5 for leptonic channel.

## C Contribution from other anomalous couplings and other processes

In Sec. 2, we argued why the contact terms, $g^h_{Zf}$, dominate at high energies if we assume a similar cut-off for all the couplings so that all the different anomalous couplings have a similar size $\sim v^2/\Lambda^2$. In this appendix we show that the dominance of the linear combination $g^{h(VBF)}_{Zf}$ holds even if we let all the anomalous couplings saturate their bounds; this is not obvious as these bounds are not all of a similar size. Consider first the couplings, $\kappa_{WW}$ and $\kappa_{ZZ}$, which also generate a contribution that grows with $t$, albeit not as rapidly as the $g^h_{Zf}$ contributions. Amongst these two couplings $\kappa_{WW}$ has a much larger contribution to the process because of the greater $W$-luminosity in the VBF process. This coupling can be constrained using the following correlation that holds in the D6 SMEFT,

$$\kappa_{WW} = \kappa_{\gamma\gamma} + \cot\theta_W \kappa_{Z\gamma} + \delta\kappa_\gamma \tag{C.1}$$

where the last two couplings are defined by the lagrangian terms,

$$\Delta\mathcal{L}_6 \supset \kappa_{\gamma\gamma}\frac{h}{2v}A^{\mu\nu}A_{\mu\nu} + \kappa_{Z\gamma}\frac{h}{v}A^{\mu\nu}Z_{\mu\nu} \tag{C.2}$$

The couplings $\kappa_{\gamma\gamma}, \kappa_{Z\gamma}$ can be already constrained at the per-mille level or smaller. On the other hand, even for the less constrained $\delta\kappa_\gamma$, diboson processes could be used to obtain the strong constraint,

$$|\delta\kappa_\gamma| \lesssim 0.005 \tag{C.3}$$

at the HL-LHC [38]. Taking a value of $\delta\kappa_\gamma$ that saturates this we find that, for the most sensitive bins around $p_T^h \sim 300$ GeV, its interference contribution is 8 times smaller than that of $g^{h(VBF)}_{Zf}$ at its maximal value for the HL-LHC in eq. (4.2). Thus the $g^{h(VBF)}_{Zf}$ contribution clearly dominates over the others.

We now discuss the effect of the couplings, $\kappa_{\gamma\gamma}, \delta g^V_f, \delta\hat{g}^h_{\tau\tau}$ and $\delta\hat{g}^h_{VV}$, that rescale the amplitude and thus only modify the total rate. Of these, $\kappa_{\gamma\gamma}$ and $\delta g^V_f$ are highly constrained, at per-mille level or smaller, respectively, from $W, Z$ decays at LEP and the Higgs

diphoton decay mode at the LHC [99]; these couplings can therefore be completely neglected given that their contributions do not grow with energy unlike the contribution of $g_{Zf}^{h(VBF)}$. The other two couplings, $\delta\hat{g}_{\tau\tau}^h$ and $\delta\hat{g}_{VV}^h$ rescale all differential distributions by a constant factor $r = 1 + 2(\delta\hat{g}_{\tau\tau}^h + \delta\hat{g}_{VV}^h)$. While $\delta\hat{g}_{VV}^h$ can be constrained much more stringently in other processes such as, $gg \to h \to 4\ell$ and $pp \to Wh/Zh$, the VBF process considered in this paper is the most sensitive way to directly probe $\delta\hat{g}_{\tau\tau}^h$ .The sensitivity to $\delta\hat{g}_{\tau\tau}^h$ however comes from lower bins with a much higher number of events, unlike the bound on $g_{Zf}^{h(VBF)}$ which arises from the bins around $p_T^h = 300$ GeV. Indeed less than 2-4 % of the events, depending on the $\tau\tau$ decay mode in question, lie in the region $p_T^h > 300$ GeV so that the effect on the bound on $\delta\hat{g}_{\tau\tau}^h$ in our set-up is completely negligible if these events are not considered. Thus $r$ and $g_{Zf}^{h(VBF)}$ can be independently measured by considering separately the bins with $p_T^h$ less than or more than 300 GeV. If this procedure reveals a non-vanishing, $\delta\hat{g}_{\tau\tau}^h$, its effect can be subtracted from the higher bins in a straightforward way. [7] Finally,

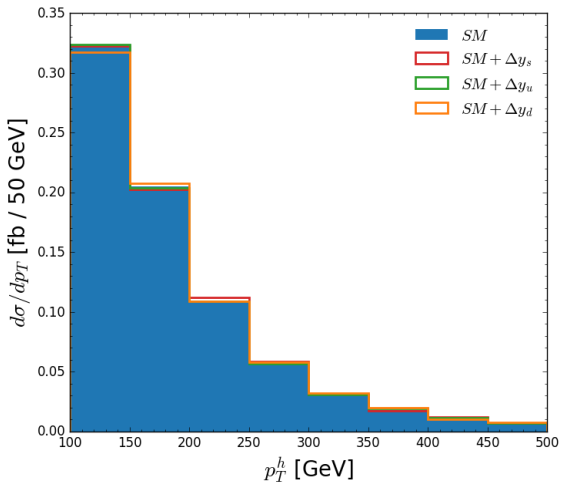


**Figure 9**. Figure shows the $p_T$ spectrum of the Higgs boson reconstructed from the diphoton channel. The blue, red, green and orange histograms show the contributions from the SM, SM + $\Delta y_s$, SM + $\Delta y_u$ and SM + $\Delta y_d$, respectively. $\Delta y_i$ $(i = s, u, d)$ are the interferences with the SM for $\delta y_s \sim 17, \delta y_u \sim 700$ and $\delta y_d \sim 360$.

we discuss how the variation of the Higgs Yukawas to the light quarks affect our results. We consider constraints from the Higgs signal strengths while allowing for enhanced effects from $q\bar{q} \to h$ production [100, 101]. Following the former reference, we consider the modifications to the $u$, $d$ and $s$ Yukawa couplings to be $\delta y_d \lesssim 360, \delta y_u \lesssim 700$ and $\delta y_s \lesssim 17$. These bounds also include a future one on the signal strength to be $\mu = 1.00 \pm 0.03$ at the HL-LHC. Upon using these deformations, and our generation level cuts (as discussed above), we find the cross-sections to be modified as follows. For SM, $\delta y_d, \delta y_u$ and $\delta y_s$, the central values of the cross-sections are respectively 4.550 fb, 4.560 fb, 4.564 fb and 4.552

[7] One should ideally perform a fit to simultaneously ascertain the rescaled SM term proportional to $r$ as well as the the quadratically growing term proportional to $g_{Zf}^{h(VBF)}$, but for our present purposes the above procedure is sufficient.

fb. We show the $p_T$ spectrum of the reconstructed Higgs boson in the diphoton channel in Fig. 9. We find the modifications owing to the change in these Yukawas to be negligible.

## Acknowledgments

We thank Shilpi Jain and Sanmay Ganguly for several helpful discussions concerning the analyses. JYA acknowledges the funding received from the European Union's Horizon 2020 research and innovation programme as part of the Marie Skłodowska-Curie Innovative Training Network MCnetITN3 (grant agreement no. 722104). SB acknowledges the grant received from IPPP, Durham, where the majority of the work was done.

## References

[1] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys.* **B268** (1986) 621–653.

[2] G. F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, *The Strongly-Interacting Light Higgs*, *JHEP* **06** (2007) 045, [hep-ph/0703164].

[3] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085, [1008.4884].

[4] R. S. Gupta, *Probing Quartic Neutral Gauge Boson Couplings using diffractive photon fusion at the LHC*, *Phys. Rev.* **D85** (2012) 014006, [1111.3354].

[5] R. S. Gupta, H. Rzehak and J. D. Wells, *How well do we need to measure Higgs boson couplings?*, *Phys. Rev.* **D86** (2012) 095001, [1206.3560].

[6] S. Banerjee, S. Mukhopadhyay and B. Mukhopadhyaya, *New Higgs interactions and recent data from the LHC and the Tevatron*, *JHEP* **10** (2012) 062, [1207.3588].

[7] R. S. Gupta, M. Montull and F. Riva, *SUSY Faces its Higgs Couplings*, *JHEP* **04** (2013) 132, [1212.5240].

[8] S. Banerjee, S. Mukhopadhyay and B. Mukhopadhyaya, *Higher dimensional operators and the LHC Higgs data: The role of modified kinematics*, *Phys. Rev.* **D89** (2014) 053010, [1308.4860].

[9] R. S. Gupta, H. Rzehak and J. D. Wells, *How well do we need to measure the Higgs boson mass and self-coupling?*, *Phys. Rev.* **D88** (2013) 055024, [1305.6397].

[10] J. Elias-Miró, C. Grojean, R. S. Gupta and D. Marzocca, *Scaling and tuning of EW and Higgs observables*, *JHEP* **05** (2014) 019, [1312.2928].

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

[12] A. Falkowski and F. Riva, *Model-independent precision constraints on dimension-6 operators*, *JHEP* **02** (2015) 039, [1411.0669].

[13] C. Englert and M. Spannowsky, *Effective Theories and Measurements at Colliders*, *Phys. Lett.* **B740** (2015) 8–15, [1408.5147].

[14] R. S. Gupta, A. Pomarol and F. Riva, *BSM Primary Effects*, *Phys. Rev.* **D91** (2015) 035001, [1405.0181].

[15] G. Amar, S. Banerjee, S. von Buddenbrock, A. S. Cornell, T. Mandal, B. Mellado et al., *Exploration of the tensor structure of the Higgs boson coupling to weak bosons in $e^+$ $e^-$ collisions*, *JHEP* **02** (2015) 128, [1405.3957].

[16] M. Buschmann, D. Goncalves, S. Kuttimalai, M. Schonherr, F. Krauss and T. Plehn, *Mass Effects in the Higgs-Gluon Coupling: Boosted vs Off-Shell Production*, *JHEP* **02** (2015) 038, [1410.5806].

[17] N. Craig, M. Farina, M. McCullough and M. Perelstein, *Precision Higgsstrahlung as a Probe of New Physics*, *JHEP* **03** (2015) 146, [1411.0676].

[18] J. Ellis, V. Sanz and T. You, *Complete Higgs Sector Constraints on Dimension-6 Operators*, *JHEP* **07** (2014) 036, [1404.3667].

[19] J. Ellis, V. Sanz and T. You, *The Effective Standard Model after LHC Run I*, *JHEP* **03** (2015) 157, [1410.7703].

[20] S. Banerjee, T. Mandal, B. Mellado and B. Mukhopadhyaya, *Cornering dimension-6 HVV interactions at high luminosity LHC: the role of event ratios*, *JHEP* **09** (2015) 057, [1505.00226].

[21] C. Englert, R. Kogler, H. Schulz and M. Spannowsky, *Higgs coupling measurements at the LHC*, *Eur. Phys. J.* **C76** (2016) 393, [1511.05170].

[22] D. Ghosh, R. S. Gupta and G. Perez, *Is the Higgs Mechanism of Fermion Mass Generation a Fact? A Yukawa-less First-Two-Generation Model*, *Phys. Lett.* **B755** (2016) 504–508, [1508.01501].

[23] C. Degrande, B. Fuks, K. Mawatari, K. Mimasu and V. Sanz, *Electroweak Higgs boson production in the standard model effective field theory beyond leading order in QCD*, *Eur. Phys. J. C* **77** (2017) 262, [1609.04833].

[24] J. Cohen, S. Bar-Shalom and G. Eilam, *Contact Interactions in Higgs-Vector Boson Associated Production at the ILC*, *Phys. Rev.* **D94** (2016) 035030, [1602.01698].

[25] S.-F. Ge, H.-J. He and R.-Q. Xiao, *Probing new physics scales from Higgs and electroweak observables at $e^+$ $e^-$ Higgs factory*, *JHEP* **10** (2016) 007, [1603.03385].

[26] R. Contino, A. Falkowski, F. Goertz, C. Grojean and F. Riva, *On the Validity of the Effective Field Theory Approach to SM Precision Tests*, *JHEP* **07** (2016) 144, [1604.06444].

[27] A. Biekötter, J. Brehmer and T. Plehn, *Extending the limits of Higgs effective theory*, *Phys. Rev.* **D94** (2016) 055032, [1602.05202].

[28] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina et al., *Electroweak precision observables and Higgs-boson signal strengths in the Standard Model and beyond: present and future*, *JHEP* **12** (2016) 135, [1608.01509].

[29] H. Denizli and A. Senol, *Constraints on Higgs effective couplings in $H\nu\bar{\nu}$ production of CLIC at 380 GeV*, *Adv. High Energy Phys.* **2018** (2018) 1627051, [1707.03890].

[30] T. Barklow, K. Fujii, S. Jung, R. Karl, J. List, T. Ogawa et al., *Improved Formalism for Precision Higgs Coupling Fits*, *Phys. Rev.* **D97** (2018) 053003, [1708.08912].

[31] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1–98, [1706.08945].

[32] T. Barklow, K. Fujii, S. Jung, M. E. Peskin and J. Tian, *Model-Independent Determination*

*of the Triple Higgs Coupling at e+e- Colliders*, *Phys. Rev.* **D97** (2018) 053004, [1708.09079].

[33] H. Khanpour and M. Mohammadi Najafabadi, *Constraining Higgs boson effective couplings at electron-positron colliders*, *Phys. Rev.* **D95** (2017) 055026, [1702.00951].

[34] C. Englert, R. Kogler, H. Schulz and M. Spannowsky, *Higgs characterisation in the presence of theoretical uncertainties and invisible decays*, *Eur. Phys. J.* **C77** (2017) 789, [1708.06355].

[35] G. Panico, F. Riva and A. Wulzer, *Diboson Interference Resurrection*, *Phys. Lett.* **B776** (2018) 473–480, [1708.07823].

[36] R. Franceschini, G. Panico, A. Pomarol, F. Riva and A. Wulzer, *Electroweak Precision Tests in High-Energy Diboson Processes*, *JHEP* **02** (2018) 111, [1712.01310].

[37] S. Banerjee, C. Englert, R. S. Gupta and M. Spannowsky, *Probing Electroweak Precision Physics via boosted Higgs-strahlung at the LHC*, *Phys. Rev. D* **98** (2018) 095012, [1807.01796].

[38] C. Grojean, M. Montull and M. Riembau, *Diboson at the lhc vs lep*, *JHEP* **03** (2019) 020, [1810.05149].

[39] A. Biekoetter, T. Corbett and T. Plehn, *The Gauge-Higgs Legacy of the LHC Run II*, *SciPost Phys.* **6** (2019) 064, [1812.07587].

[40] D. Goncalves and J. Nakamura, *Boosting the $H \to$ invisibles searches with $Z$ boson polarization*, *Phys. Rev.* **D99** (2019) 055021, [1809.07327].

[41] R. Gomez-Ambrosio, *Studies of Dimension-Six EFT effects in Vector Boson Scattering*, *Eur. Phys. J. C* **79** (2019) 389, [1809.04189].

[42] F. F. Freitas, C. K. Khosa and V. Sanz, *Exploring SMEFT in VH with Machine Learning*, 1902.05803.

[43] S. Banerjee, R. S. Gupta, J. Y. Reiness and M. Spannowsky, *Resolving the tensor structure of the Higgs coupling to $Z$-bosons via Higgs-strahlung*, *Phys. Rev.* **D100** (2019) 115004, [1905.02728].

[44] S. Banerjee, R. S. Gupta, J. Y. Reiness, S. Seth and M. Spannowsky, *Towards the ultimate differential SMEFT analysis*, *JHEP* **09** (2020) 170, [1912.07628].

[45] A. Biekötter, R. Gomez-Ambrosio, P. Gregg, F. Krauss and M. Schönherr, *Constraining SMEFT operators with associated $h\gamma$ production in Weak Boson Fusion*, 2003.06379.

[46] D. Liu and L.-T. Wang, *Prospects for precision measurement of diboson processes in the semileptonic decay channel in future LHC runs*, *Phys. Rev. D* **99** (2019) 055001, [1804.08688].

[47] D. Binosi, J. Collins, C. Kaufhold and L. Theussl, *JaxoDraw: A Graphical user interface for drawing Feynman diagrams. Version 2.0 release notes*, *Comput. Phys. Commun.* **180** (2009) 1709–1715, [0811.4113].

[48] A. Pomarol, *Higgs Physics*, in *2014 European School of High-Energy Physics*, pp. 59–77, 2016. 1412.4410.

[49] G. D'Ambrosio, G. Giudice, G. Isidori and A. Strumia, *Minimal flavor violation: An Effective field theory approach*, *Nucl. Phys. B* **645** (2002) 155–187, [hep-ph/0207036].

[50] K. Hagiwara, R. D. Peccei, D. Zeppenfeld and K. Hikasa, *Probing the Weak Boson Sector in e+ e- —¿ W+ W-*, *Nucl. Phys.* **B282** (1987) 253–307.

[51] M. E. Peskin and T. Takeuchi, *Estimation of oblique electroweak corrections*, *Phys. Rev.* **D46** (1992) 381–409.

[52] R. Barbieri, A. Pomarol, R. Rattazzi and A. Strumia, *Electroweak symmetry breaking after LEP-1 and LEP-2*, *Nucl. Phys.* **B703** (2004) 127–146, [hep-ph/0405040].

[53] J. D. Wells and Z. Zhang, *Effective theories of universal theories*, *JHEP* **01** (2016) 123, [1510.08462].

[54] ATLAS collaboration, M. Aaboud et al., *Cross-section measurements of the Higgs boson decaying into a pair of $\tau$-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev.* **D99** (2019) 072001, [1811.08856].

[55] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, *JHEP* **04** (2008) 063, [0802.1189].

[56] 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) 2250–2300, [1310.1921].

[57] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *Ufo - the universal feynrules output*, *Comput. Phys. Commun.* **183** (2012) 1201–1214, [1108.2040].

[58] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer 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, [1405.0301].

[59] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177, [1410.3012].

[60] NNPDF collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [1410.8849].

[61] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J.* **C75** (2015) 132, [1412.7420].

[62] A. Greljo, G. Isidori, J. M. Lindert, D. Marzocca and H. Zhang, *Electroweak Higgs production with HiggsPO at NLO QCD*, *Eur. Phys. J. C* **77** (2017) 838, [1710.04143].

[63] S. Kallweit, J. M. Lindert, P. Maierhofer, S. Pozzorini and M. Schönherr, *NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging*, *JHEP* **04** (2016) 021, [1511.08692].

[64] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO.

[65] P. Kant, O. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, *Comput. Phys. Commun.* **191** (2015) 74–89, [1406.4403].

[66] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, *Nucl. Instrum. Meth. A* **654** (2011) 481–489, [1012.4686].

[67] ATLAS collaboration, M. Aaboud et al., *Measurements of Higgs boson properties in the diphoton decay channel with 36 fb$^{-1}$ of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **98** (2018) 052005, [1802.04146].

[68] *Tagging and suppression of pileup jets with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2014-018, CERN, Geneva, May, 2014.

[69] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. A. Khalek, A. A. Abdelalim et al., *Measurement of isolated-photon pair production in pp collisions at $\sqrt{s} = 7$ tev with the atlas detector*, *Journal of High Energy Physics* **2013** (2013) 86.

[70] ATLAS Collaboration collaboration, G. Aad, B. Abbott, J. Abdallah, A. A. Abdelalim, A. Abdesselam, O. Abdinov et al., *Measurement of the isolated diphoton cross section in pp collisions at $\sqrt{s}$=7 TeV with the atlas detector*, *Phys. Rev. D* **85** (Jan, 2012) 012003.

[71] E. Conte and B. Fuks, *Confronting new physics theories to LHC data with MADANALYSIS 5*, *Int. J. Mod. Phys.* **A33** (2018) 1830027, [1808.00480].

[72] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C72** (2012) 1896, [1111.6097].

[73] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, LEP TGC Working Group collaboration, *A Combination of Preliminary Results on Gauge Boson Couplings Measured by the LEP experiments*, Tech. Rep. LEPEWWG-TGC-2003-01. DELPHI-2003-068-PHYS-936. L3-Note-2826. LEPEWWG-2006-01. OPAL-TN-739. ALEPH-2006-016-CONF-2003-012, CERN, Geneva, Jan, 2003.

[74] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, R. Kogler et al., *The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC*, *Eur. Phys. J.* **C72** (2012) 2205, [1209.2716].

[75] M. Farina, G. Panico, D. Pappadopulo, J. T. Ruderman, R. Torre and A. Wulzer, *Energy helps accuracy: electroweak precision tests at hadron colliders*, *Phys. Lett. B* **772** (2017) 210–215, [1609.08157].

[76] J. Y. Araz, B. Fuks and G. Polykratis, *Simplified fast detector simulation in MadAnalysis 5*, 2006.09387.

[77] P. Baldi, P. Sadowski and D. Whiteson, *Enhanced Higgs Boson to $\tau^+\tau^-$ Search with Deep Learning*, *Phys. Rev. Lett.* **114** (2015) 111801, [1410.3469].

[78] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman and A. Schwartzman, *Jet-images — deep learning edition*, *JHEP* **07** (2016) 069, [1511.05190].

[79] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski and D. Whiteson, *Parameterized neural networks for high-energy physics*, *Eur. Phys. J. C* **76** (2016) 235, [1601.07913].

[80] S. Caron, J. S. Kim, K. Rolbiecki, R. Ruiz de Austri and B. Stienen, *The BSM-AI project: SUSY-AI–generalizing LHC limits on supersymmetry with machine learning*, *Eur. Phys. J. C* **77** (2017) 257, [1605.02797].

[81] S. Chang, T. Cohen and B. Ostdiek, *What is the Machine Learning?*, *Phys. Rev. D* **97** (2018) 056009, [1709.10106].

[82] J. Lin, M. Freytsis, I. Moult and B. Nachman, *Boosting $H \to b\bar{b}$ with Machine Learning*, *JHEP* **10** (2018) 101, [1807.10768].

[83] K. Albertsson et al., *Machine Learning in High Energy Physics Community White Paper*, *J. Phys. Conf. Ser.* **1085** (2018) 022008, [1807.02876].

[84] D. Guest, K. Cranmer and D. Whiteson, *Deep Learning and its Application to LHC Physics*, *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 161–181, [1806.11484].

[85] M. Abdughani, J. Ren, L. Wu, J. M. Yang and J. Zhao, *Supervised deep learning in high energy phenomenology: a mini review*, *Commun. Theor. Phys.* **71** (2019) 955, [1905.06047].

[86] P. Windischhofer, M. Zgubiˇ c and D. Bortoletto, *Preserving physically important variables in optimal event selections: A case study in Higgs physics*, 1907.02098.

[87] J. Amacker et al., *Higgs self-coupling measurements using deep learning and jet substructure in the $b\bar{b}b\bar{b}$ final state*, 2004.04240.

[88] F. Chollet et al., "Keras." https://keras.io, 2015.

[89] K. Hornik, *Approximation capabilities of multilayer feedforward networks*, *Neural Networks* **4** (1991) 251 – 257.

[90] R. Lippmann, *An introduction to computing with neural nets*, *IEEE ASSP Magazine* **4** (1987) 4–22.

[91] K. Hornik, M. Stinchcombe and H. White, *Multilayer feedforward networks are universal approximators*, *Neural Networks* **2** (1989) 359 – 366.

[92] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, *CoRR* **abs/1412.6980** (2014) .

[93] ATLAS collaboration, G. Aad et al., *Search for non-resonant Higgs boson pair production in the $bb\ell\nu\ell\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* **801** (2020) 135145, [1908.06765].

[94] C. Cortes, M. Mohri and A. Rostamizadeh, *$l_2$ regularization for learning kernels*, *CoRR* **abs/1205.2653** (2012) , [1205.2653].

[95] P. Baringer, K. Kong, M. McCaskey and D. Noonan, *Revisiting Combinatorial Ambiguities at Hadron Colliders with $M_{T2}$*, *JHEP* **10** (2011) 101, [1109.1563].

[96] A. J. Barr, M. J. Dolan, C. Englert and M. Spannowsky, *Di-Higgs final states augMT2ed – selecting hh events at the high luminosity LHC*, *Phys. Lett. B* **728** (2014) 308–313, [1309.6318].

[97] J. Davis and M. Goadrich, *The relationship between precision-recall and roc curves*, in *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, (New York, NY, USA), pp. 233–240, Association for Computing Machinery, 2006. DOI.

[98] S. Lundberg and S.-I. Lee, *A Unified Approach to Interpreting Model Predictions*, *arXiv e-prints* (May, 2017) arXiv:1705.07874, [1705.07874].

[99] A. Pomarol and F. Riva, *Towards the Ultimate SM Fit to Close in on Higgs Physics*, *JHEP* **01** (2014) 151, [1308.2803].

[100] A. Falkowski, S. Ganguly, P. Gras, J. M. No, K. Tobioka, N. Vignaroli et al., *Light quark Yukawas in triboson final states*, 2011.09551.

[101] C. Delaunay, R. Ozeri, G. Perez and Y. Soreq, *Probing Atomic Higgs-like Forces at the Precision Frontier*, *Phys. Rev. D* **96** (2017) 093001, [1601.05087].