Boosted Higgs shapes

Matthias Schlaffer[1,a], **Michael Spannowsky**[2], **Michihisa Takeuchi**[3], **Andreas Weiler**[1,4], **Chris Wymant**[2,5,6]

[1] DESY, Notkestrasse 85, 22607 Hamburg, Germany
[2] Department of Physics, Institute for Particle Physics Phenomenology, Durham University, Durham DH1 3LE, UK
[3] Theoretical Physics and Cosmology Group, Department of Physics, King's College London, London WC2R 2LS, UK
[4] Theory Division, Physics Department, CERN, 1211 Geneva 23, Switzerland
[5] Laboratoire d'Annecy-le-Vieux de Physique Théorique, 9 Chemin de Bellevue, 74940, Annecy-le-Vieux, France
[6] *Present address:* Department of Infectious Disease Epidemiology, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG, UK

Abstract The inclusive Higgs production rate through gluon fusion has been measured to be in agreement with the Standard Model (SM). We show that even if the inclusive Higgs production rate is very SM-like, a precise determination of the boosted Higgs transverse momentum shape offers the opportunity to see effects of natural new physics. These measurements are generically motivated by effective field theory arguments and specifically in extensions of the SM with a natural weak scale, like composite Higgs models and natural supersymmetry. We show in detail how a measurement at high transverse momentum of $H \rightarrow 2\ell + \not{p}_T$ via $H \rightarrow \tau\tau$ and $H \rightarrow WW^*$ could be performed and demonstrate that it offers a compelling alternative to the $t\bar{t}H$ channel. We discuss the sensitivity to new physics in the most challenging scenario of an exactly SM-like inclusive Higgs cross section.

1 Introduction

After the observation of a Standard Model-like Higgs boson [1,2] the main physics program has shifted to measuring its precise properties. The exotic possibility of the new particle having spin one or two is already disfavored [3,4] by the analysis of its decays into $\gamma\gamma$ [5–8], WW^* [9–11], and ZZ^* [4,12–21]; so is the possibility of it being a pure pseudo-scalar [22,23], though it could still be an admixture of scalar and pseudo-scalar [24–31]. The couplings to gauge bosons and fermions have also been measured for various decay modes, and they are so far consistent with the SM predictions within uncertainties [32,33].

With no apparent deviation from the SM so far, it is important to closely examine the channels where one has a fight-

---

[a] e-mail: Matthias.Schlaffer@desy.de

ing chance to encounter new physics. One such promising process is Higgs production via gluon fusion: In order to avoid unnatural fine-tuning while still obtaining a light Higgs mass, loops of new particles need to soften to the Higgs mass squared UV sensitivity of the top loop. If these particles are charged under the $SU(3)$ color gauge group (which they are in almost all known cases), gluons will couple to the loop. With two gluons coupled to the new physics loop and one Higgs set to its vacuum expectation value, one gets a contribution to gluon fusion, the dominant Higgs production mechanism; see e.g. [34,35].

At the same time, top partners can lead to a modified top-Yukawa coupling. A change in the top-Yukawa affects the Higgs production cross section and can even compensate for new particles in the loop such that a SM-like inclusive cross section is obtained even though new physics is present. The reason for this is that already for the top quark the effective gluon Higgs interaction [36,37] obtained from the low energy theorem is a very good description [38,39] which works even better for heavier particles. Therefore the inclusive amplitude can be expressed as the sum of two identical Feynman diagrams with the effective interaction (one from the top loop and one from the non-SM loop) which differ only by a coefficient, c_t and κ_g, respectively. The cross section is therefore only sensitive to the absolute square of the sum of these coefficients. The effects of this in composite Higgs models were calculated in [40–44] where it is shown that the contributions to the inclusive cross section indeed cancel in minimal models.

The main idea now is to study boosted Higgs shapes above a certain p_T scale. This scale should be high enough to resolve the top loop beyond the effective description, but low enough to keep the effective description of the loop of the new particle valid; see [45] for a discussion in a concrete model.

The simple relation $\sigma \propto |c_t + \kappa_g|^2$ for the inclusive cross sections does not apply but is modified, allowing the two coefficients to be extracted separately, when combined with the inclusive measurement. Early studies looking for New Physics in the Higgs p_T distribution in the gluon-fusion production mode include [46–48] and recent preliminary studies looking at highly boosted Higgs shapes include [45,49–51]. An alternative approach to measure the coupling c_t in boosted $pp \rightarrow HZ$ was presented in [52,53]. Although there are attempts to measure c_t directly by looking into the difficult $t\bar{t}H$ channel [54–60], it is important to explore boosted Higgs production from gluon fusion as a complementary approach.

To simplify the extraction of the small amount of high-p_T signal from the background, we focus on the clean decay of a Higgs to two leptons $\ell = e^\pm, \mu^\pm$ and missing transverse momentum $\not{p}_T$. For a 125 GeV Standard Model-like Higgs boson, this occurs almost entirely via $H \rightarrow WW^*$ and $H \rightarrow \tau\tau$; we will focus on these two channels separately as detailed in Sect. 4.

The organization of the paper is as follows. In Sect. 2 we discuss some examples of beyond the Standard Model physics which motivate this analysis. Section 3 outlines how we generated our signal and background samples. Section 4 contains our signal versus background analyses for boosted $H \rightarrow 2\ell + \not{p}_T$ in the Standard Model. Section 5 contains a discussion of the analysis and we conclude in Sect. 6.

2 New physics models

2.1 Minimal Composite Higgs Model

In the Minimal Composite Higgs Model (MCHM) [61], electroweak symmetry is broken dynamically by a strong interaction based on the coset $SO(5)/SO(4)$. For reviews of MHCM see [62,63]. In this class of models, the Higgs arises as a pseudo-Nambu–Goldstone Boson (pNGB) of the symmetry breaking which naturally explains its small mass. Fermionic resonances of the strong sector, coming in multiplets of $SO(4)$, will contribute to the gluon-fusion loop diagram. These resonances also mix with the SM fermions and thus modify their couplings to the Higgs. Interestingly, the contributions of these resonances to the sum of the coefficients κ_g and c_t cancel exactly in a broad class of MCHM models and lead, up to small corrections which are negligible at the LHC [64], to [40–44]

$$c_t + \kappa_g = f_g(\xi) \tag{1}$$

where f_g is a function satisfying $f_g(\xi \rightarrow 0) = 1$ with $\xi \equiv v^2/f^2$ and f is the decay constant of the non-linear sigma model. The gluon-fusion cross section is therefore independent of the mass spectrum of the fermionic resonances, and for small ξ is even SM-like. This makes it impos-

sible to find traces of the top partner spectrum in the inclusive gluon-fusion process.

While the resonances are needed to cut off the UV-divergences of the Higgs mass and thus must not be too heavy to avoid excessive fine tuning, they should still be heavy enough to allow for an effective description of the boosted Higgs production. In [45] it was shown that as long as the mass of the lightest resonance is at least of the order of the Higgs transverse momentum, the result of the calculation in the heavy top limit lies within $\mathcal{O}(10\,\%)$ of the full calculation. Considering that the masses of the resonances have to be heavier than $600 - 800$ GeV depending on the representation [65–74], the effective description is well justified within the scope of the paper.

2.2 Supersymmetry

In the Minimal Supersymmetric SM (MSSM), an analogous flat direction of the inclusive cross section exists which can be resolved by looking at boosted Higgs shapes. For certain choices of the stop masses $m_{\tilde{t}_1}$, $m_{\tilde{t}_2}$ and A_t, the effects of two contributions cancel, yielding a SM-like inclusive signal strength [75–83]. Assuming the MSSM is in the decoupling limit, and neglecting small D-term contributions, the inclusive signal strength is given by [84]

$$\frac{\Gamma(gg \rightarrow H)}{\Gamma(gg \rightarrow H)_{\mathrm{SM}}} = (1 + \Delta_t)^2 \tag{2}$$

where

$$\Delta_t \approx \frac{m_t^2}{4}\left(\frac{1}{m_{\tilde{t}_1}^2} + \frac{1}{m_{\tilde{t}_2}^2} - \frac{(A_t - \mu/\tan\,\beta)^2}{m_{\tilde{t}_1}^2 m_{\tilde{t}_2}^2}\right) \tag{3}$$

quantifies the deviation from the SM value and can vanish due to the relative minus sign. A 125 GeV Higgs can easily achieved by extending the MSSM by additional D- or F-terms which should, of course, not have a major impact on the couplings of the SM-like lightest Higgs.

Since the A_t-dependent parts of the production cross section are less sensitive to the boost of the Higgs than the A_t-independent ones, the aforementioned degeneracy gets broken in the boosted regime. Therefore the non-SM nature of the Higgs production can be revealed by looking at the boosted production. Moreover, this can make light stops [85–98] accessible which are hidden in the stealth region and challenging to extract given the similarity to the top background [99–103]. An outline showing this sensitivity and taking vacuum stability constraints into account has been presented in [45].

2.3 Effective description

It is useful to parametrize our ignorance of new physics in terms of an effective Lagrangian. Out of the 59 dimension

six operators one can add to the SM [104,105], only four can affect the Higgs production through gluon fusion [106–108]. These four operators as well as the other dimension six operators involving the Higgs are already constrained to some extent by LHC data [108–113]. We will focus on CP-conserving effects and omit the CP-violating operator containing the dual of the QCD gauge field strength. The remaining three important operators are

$$\mathcal{O}_y = \frac{y_t}{v^2}|H|^2 \bar{Q}_L \widetilde{H} t_R, \quad \mathcal{O}_H = \frac{1}{2v^2}\partial_\mu |H|^2 \partial^\mu |H|^2,$$
$$\text{and} \quad \mathcal{O}_g = \frac{\alpha_S}{12\pi v^2}|H|^2 G_{\mu\nu}^a G^{a\,\mu\mu}. \quad (4)$$

After adding them to the SM Lagrangian and extracting the terms relevant for the gluon-fusion process we obtain

$$\mathcal{L}_{\text{eff}} = -c_t \frac{m_t}{v}\bar{t}t H + \kappa_g \frac{\alpha_S}{12\pi}\frac{h}{v}G_{\mu\nu}^a G^{a\,\mu\nu} + \mathcal{L}_{\text{QCD}}, \quad (5)$$

where $c_t = 1 - \text{Re}(C_y) - C_H/2$ scales the top Yukawa coupling which enters the process via the top loop and $\kappa_g = C_g$ controls the direct gluon–Higgs interaction. The C_i are the coefficients of the corresponding operators in (4) and the coefficients are chosen such that for $c_t = 1$ and $\kappa_g = 0$ the SM Lagrangian is obtained.

The full matrix element for boosted Higgs production is then given by[1]

$$\mathcal{M}(c_t, \kappa_g) = c_t \mathcal{M}_{\text{IR}} + \kappa_g \mathcal{M}_{\text{UV}} \quad (6)$$

where $\mathcal{M}_{\text{IR}}$ is the matrix element taking the full top mass dependence into account [115] and $\mathcal{M}_{\text{UV}}$ is the one obtained from $\mathcal{M}_{\text{IR}}$ in the heavy top limit or equivalently from the tree-level diagram generated by $\mathcal{O}_g$. From Eq. (6) we see that the differential cross section, normalized by the SM value, can be described as

$$\frac{\sigma(p_T^{\text{cut}})}{\sigma^{\text{SM}}(p_T^{\text{cut}})} = \frac{\int_{p_T^{\text{cut}}}^\infty dp_T d\Omega |c_t \mathcal{M}_{\text{IR}}(m_t) + \kappa_g \mathcal{M}_{\text{UV}}|^2}{\int_{p_T^{\text{cut}}}^\infty dp_T d\Omega |\mathcal{M}_{\text{IR}}(m_t)|^2}$$
$$= (c_t + \kappa_g)^2 + \delta(p_T^{\text{cut}})c_t \kappa_g + \epsilon(p_T^{\text{cut}})\kappa_g^2, \quad (7)$$

where

$$\delta(p_T^{\text{cut}}) = \frac{2\int_{p_T^{\text{cut}}}^\infty dp_T d\Omega \, \text{Re}(\mathcal{M}_{\text{IR}}(m_t)\mathcal{M}_{\text{UV}}^*)}{\int_{p_T^{\text{cut}}}^\infty dp_T d\Omega |\mathcal{M}_{\text{IR}}(m_t)|^2} - 2, \quad (8)$$

$$\epsilon(p_T^{\text{cut}}) = \frac{\int_{p_T^{\text{cut}}}^\infty dp_T d\Omega |\mathcal{M}_{\text{UV}}|^2}{\int_{p_T^{\text{cut}}}^\infty dp_T d\Omega |\mathcal{M}_{\text{IR}}(m_t)|^2} - 1. \quad (9)$$

For small p_T^{cut}, the coefficients δ, ϵ are very small, modifying the cross section only by a few percent, which is less than the uncertainty expected in the inclusive Higgs cross section measurements [116–118]. This is what is expected

due to the very good description of both the top and the new particle loop by the effective interaction. On the other hand, δ, ϵ grow significantly as p_T^{cut} increases, and they become $\mathcal{O}(1)$ for $p_T^{\text{cut}} > 300$ GeV [45]. It means we can break the degeneracy by measuring the Higgs p_T distribution while we cannot break the degeneracy along $c_t + \kappa_g = \text{const.}$ direction only by determining the inclusive cross section.

3 Event generation

3.1 Signal sample

In this paper we consider H+jet events with subsequent H decays to $WW^* \to \ell^+\ell^-\nu\bar{\nu}$ and $\tau^+\tau^-$ modes as a signal. The signal events are generated with MadGraph5, version 1.5.15 [119] and showered with HERWIG++ [120–122], where only WW^* and $\tau^+\tau^-$ decays are specified.

We have used MadGraph5 to generate H+jet events using the 'HEFT' model with SM couplings which makes use of the low energy theorem. The generated cross section is proportional to $|\mathcal{M}(0, 1)|^2$ and does not take into account finite top mass effects which are crucial to our analysis. To obtain the correct weight of the events we reweighted them by a weight factor

$$w(c_t, \kappa_g) = \frac{|\mathcal{M}(c_t, \kappa_g)|^2}{|\mathcal{M}(0, 1)|^2} \quad (10)$$

making use of our own code, which is based on an implementation of the formulas for the matrix elements given in [115] and also calculated in [123]. At present no finite top mass NLO computation of the SM Higgs p_T spectrum is available. An exact NLO prediction of SM Higgs p_T spectrum would be very desirable and help to exploit the full potential of this observable. Recent progress in the precision prediction of $h + \text{jet}$ can be found in Refs. [124–126]. We will approximate the NNLO (+ NNLL) result of 49.85 pb [127–130] by multiplying the exact LO result with a K factor of 1.71.

We reweight the events for points along the line $c_t + \kappa_g = 1$ for $\kappa_g \in [-0.5, 0.5]$ with steps of 0.1, as shown in the left panel of Fig. 1. This is consistent with the SM inclusive Higgs production cross section. The size of c_t alone is only weakly constrained by the current $t\bar{t}H$ measurement. Although we only consider the most difficult points satisfying $c_t + \kappa_g = 1$ (i.e. an exactly SM-like inclusive cross section), an analysis along different $c_t + \kappa_g = \text{const.}$ lines would be straightforward as a different choice essentially just corresponds to an overall rescaling of the signal.

The right panel of the Fig. 1 shows the $p_{T,H}$ distributions for several model points. In the region with low $p_{T,H}$, the distributions are degenerate but for high $p_{T,H}$ the distributions start to split. For the model points with $\kappa_g > 0$

---

[1] In the SM the effects of the bottom loop are within a few percent if the boost of the Higgs exceeds $\mathcal{O}(50 \text{ GeV})$ [39,48,114] and are therefore neglected.

Fig. 1 *Left panel*: model points generated for this analysis in (c_t, κ_g) plane. The *shaded area* shows parameter space which gives the inclusive cross section consistent to the SM prediction within 20 %. *Right panel*: parton-level $p_{T,H}$ distributions for the SM, and $(c_t, \kappa_g) = (1 - \kappa_g, \kappa_g)$ with $\kappa_g = \pm 0.1, \pm 0.3, \pm 0.5$
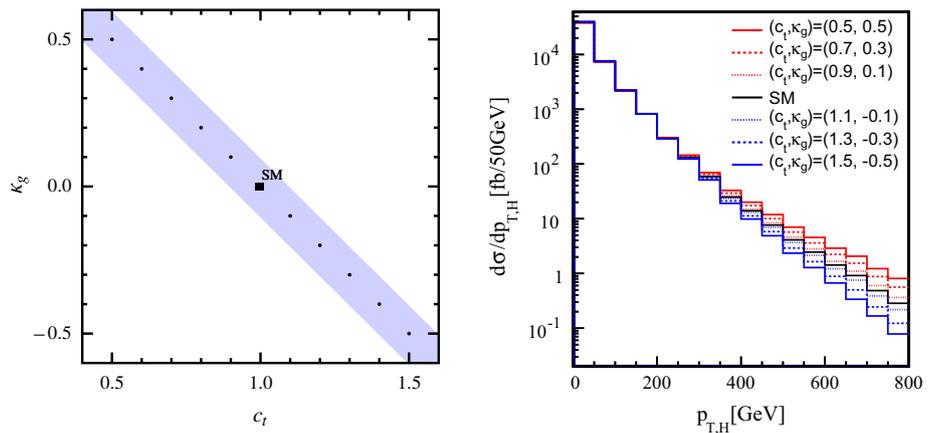
Table 1 Cross-sections normalized by the SM value after applying several $p_{T,H}$ cuts in parton level for several model points (c_t, κ_g)

	(0.5, 0.5)	(0.7, 0.3)	(0.9, 0.1)	SM	(1.1, −0.1)	(1.3, −0.3)	(1.5, −0.5)
$p_{T,H} > 10$ GeV	0.9733	0.9838	0.9945	1.0000	1.0055	1.0167	1.0281
$p_{T,H} > 100$ GeV	1.0044	1.0012	0.9999	1.0000	1.0006	1.0031	1.0076
$p_{T,H} > 200$ GeV	1.1166	1.0646	1.0198	1.0000	0.9820	0.9513	0.9277
$p_{T,H} > 300$ GeV	1.3450	1.1921	1.0591	1.0000	0.9459	0.8526	0.7791
$p_{T,H} > 400$ GeV	1.6531	1.3590	1.1087	1.0000	0.9023	0.7397	0.6210
$p_{T,H} > 500$ GeV	2.0233	1.5520	1.1633	1.0000	0.8573	0.6340	0.4932
$p_{T,H} > 600$ GeV	2.4869	1.7871	1.2274	1.0000	0.8076	0.5279	0.3882
$p_{T,H} > 700$ GeV	3.1213	2.1003	1.3093	1.0000	0.7482	0.4172	0.3161
$p_{T,H} > 800$ GeV	3.7427	2.3989	1.3841	1.0000	0.6981	0.3411	0.3129

we see an enhancement in the high $p_{T,H}$ region while we see the suppression for the model points with $\kappa_g < 0$. Table 1 shows the Higgs production cross sections relative to the SM value for several model points (c_t, κ_g) and $p_{T,H}$ cuts. As one can see, for $p_{T,H} > 10$ GeV the cross sections are essentially the same as the SM value within 3 %, while for increasing p_T^{cut}, significant differences from the SM predictions can be observed. For the model point $(c_t, \kappa_g) = (0.7, 0.3)$, for example, a 6 % difference would be observed for $\sigma(p_{T,H} > 200$ GeV), and a ~ 20 % difference for $\sigma(p_{T,H} > 300$ GeV). We will see that these effects are comparable to the sensitivity of the boosted Higgs shape measurements; see Sect. 4. For very hard cuts, $\mathcal{O}(1)$ differences can be observed, as can be seen from the cross section ratios for $p_{T,H} > 500$ GeV and harder.

3.2 Background sample

We include W+jets, Z+jets and $t\bar{t}$+jets as background processes which we have generated with ALPGEN + PYTHIA [131,132]. Since we consider *boosted* Higgs reconstruction and since we will require the existence of one hard recoil jet, we apply a pre-selection cut in the generation step, where we demand at least one recoil parton of $p_T > 150$ GeV. We merge up to two partons for WW+jets and Z+jets,

and up to one parton for $t\bar{t}$+jets using the MLM matching scheme [133,134]. As we only consider the dilepton mode in this paper we preselect the W decay mode, including W from tops only with leptons, e, μ, and τ. For the Z decay, we consider only $Z \to \tau^+\tau^-$ since for the other leptonic decay modes we can reconstruct the Z-peak and reject them. We rescale the $t\bar{t}$ sample to obtain a NLO inclusive cross section of 918 pb [135–137]. For the Z+jets and WW+jets samples we used LO cross sections.

Our analysis is performed at particle level with a simple detector simulation with the granularity resolution of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. After removing the isolated leptons, the energy of the remaining visible particles falling into each cell are summed up. Cells with transverse energy above 0.5 GeV are used for the further jet reconstruction.

Jet clustering was performed using the FastJet [138] version 3.0.4. We use the Cambridge–Aachen (C/A) algorithm [139,140] with $R = 0.5$ for normal jet and b-tag jet definition. We also define 'fat' jets, as explained later, defined using the C/A algorithm with $R = 1.5$.

In this paper, we only consider the events with isolated leptons for simplicity. There is room for improving the analysis with hadronic tau modes with tau tagging for example [141,142], which is, however, beyond the scope of our current study.

Table 2 Showing the difference in the relative positioning of the neutrinos/$\not{p}_T$ and the dilepton system between $H \to W_\ell W_\ell^*$ and $H \to \tau_\ell \tau_\ell$ decays
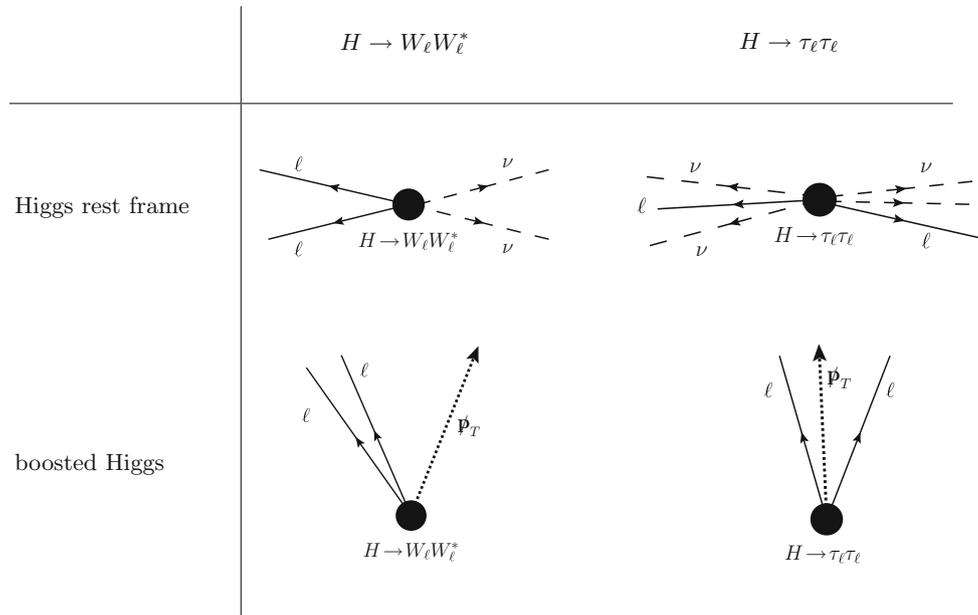
4 Boosted $H \to 2\ell + \not{p}_T$ in the standard model

In our notation a subscript ℓ will denote *leptonically decaying*: τ_ℓ thus represents $\tau \to \ell + 2\nu$, W_ℓ is mostly $W \to l\nu$ with some $W \to \tau_\ell \nu_\tau$, and t_ℓ is $t \to bW_\ell$. The decay $H \to 2\ell + \not{p}_T$ is mostly[2] through $H \to W_\ell W_\ell^*$ and $H \to \tau_\ell \tau_\ell$. As noted in [143,144], in the decay $H \to WW^* \to 2\ell + 2\nu$ spin correlations ensure that the two lepton momenta have similar directions, as do the two neutrino momenta. In $H \to \tau_\ell \tau_\ell$, however, the two τ leptons are back-to-back in the Higgs rest frame, and each of them gives rise to a highly collimated $\ell + 2\nu$ trio. These two facts imply that for a boosted $H \to 2\ell + \not{p}_T$ decay, the $\not{p}_T$ is typically outside the lepton pair for the $H \to W_\ell W_\ell^*$ contribution and inside the lepton pair for $H \to \tau_\ell \tau_\ell$, as shown in Table 2. We use this binary criterion—$\not{p}_T$ inside or outside the leptons—to split our analysis into two sub-analyses, which differ in their background compositions as well as signals.

4.1 Common cuts for $H \to \tau_\ell \tau_\ell$ and $H \to W_\ell W_\ell^*$

In both of our sub-analyses the cuts begin by requiring the following:

- Two opposite-sign isolated leptons each having $p_T > 10$ GeV and $|\eta| < 2.5$. If a third isolated lepton with

$p_T > 5$ GeV and $|\eta| < 2.5$ is present, the event is vetoed. Our isolation criterion is $E_{T,\text{had}}^{R<0.2}/E_{T,\ell} < 0.1$, where $E_{T,\text{had}}^{R<0.2}$ is the sum of transverse energies over all hadronic activity in the cone $\Delta R < 0.2$ around the lepton. (The signal leptons are typically hard, so our p_T threshold could be raised with minimal loss of efficiency.)

- A dilepton mass $m_{\ell\ell}$ exceeding 20 GeV, which is necessary in practice to suppress Drell–Yan dilepton production (not simulated here).

- At least 200 GeV of transverse momentum for the system obtained by vectorially summing the dilepton and missing transverse momenta:

$$|\mathbf{p}_{T,H}| \equiv |\not{p}_T + \mathbf{p}_{T,\ell_1} + \mathbf{p}_{T,\ell_2}| > 200 \text{ GeV}. \tag{11}$$

The system thus defined has a transverse (but not longitudinal) momentum coinciding with the Higgs in the case of the signal: herein lies our restricted focus on highly energetic/boosted Higgs bosons.

- One 'fat' jet, resulting from clustering using the C/A algorithm with a distance parameter $R_{\text{jet}} = 1.5$. This jet should be very hard:

$$p_{T,j} > 200 \text{ GeV}. \tag{12}$$

The presence of a very hard jet coincides with our parton-level picture of the signal process: a boosted Higgs recoiling against a gluon/quark. Defining geometrically large 'fat' jets allows us to capture the radiation emitted by this gluon/quark (which might otherwise be clustered into a

---

[2] The branching ratios for $H \to W_\ell W_\ell^*$ and $H \to \tau_\ell \tau_\ell$ are BR = 1.4 and 0.77 %, respectively; $H \to ZZ^* \to 2\ell + 2\nu$ is negligible with BR = 8×10^{-4}.

Table 3 Cut efficiencies for our analysis aimed at $H \to \tau_\ell \tau_\ell$. The values for each process are cross sections in fb. $S/\sqrt{B}$ has been calculated for 300 fb^{-1}. The W bosons in our $WW/t\bar{t}$+jets backgrounds were forced to decay to e, μ, or τ

Event rate (fb)	$H \to \tau\tau$	$H \to WW^*$	$W_\ell W_\ell$+jets	$Z_{\to \tau\tau}$+jets	$t_\ell \bar{t}_\ell$+jets	S/B	$S/\sqrt{B}$
0. Nominal cross section	3149.779	10719.207	580.000	1.01×10^4	1.02×10^5	–	–
1. $n_\ell = 2$, opposite-sign	118.043	323.531	195.033	347.516	3.72×10^4	–	–
2. $m_{\ell\ell} > 20$ GeV	117.733	264.723	189.522	315.201	3.57×10^4	–	–
3. $p_{T,H}^{\mathrm{rec}} > 200$ GeV	1.987	3.834	91.273	104.434	1.28×10^3	0.004	2.62
4. $n_j^{\mathrm{fat}} = 1$ ($p_{T,j} > 200$ GeV)	0.957	1.858	50.443	58.810	395.602	0.006	2.17
5. $n_b = 0$	0.940	1.825	48.855	57.068	105.851	0.01	3.29
6. $\not{p}_T$ inside the two leptons	0.923	0.533	20.215	55.551	44.050	0.01	2.30
7. $m_{\ell\ell} < 70$ GeV	0.796	0.490	3.860	53.985	8.511	0.02	2.73
8. $\lvert M_{\mathrm{col}} - m_H \rvert < 10$ GeV	0.749	0.046	0.298	1.019	0.758	0.38	9.56
$p_{T,H}^{\mathrm{rec}} > 300$ GeV	0.234	0.012	0.115	0.343	0.166	0.39	5.40
$p_{T,H}^{\mathrm{rec}} > 400$ GeV	0.068	0.006	0.042	0.106	0.049	0.38	2.88
$p_{T,H}^{\mathrm{rec}} > 500$ GeV	0.021	0.001	0.014	0.038	0.010	0.36	1.55
$p_{T,H}^{\mathrm{rec}} > 600$ GeV	0.008	0.001	0.006	0.014	0.005	0.32	0.89

separate jet when clustering with traditional 'skinny' jets). We veto if there is a second fat jet with $p_{T,j} > 100$ GeV. Vetoing on additional hadronic activity beyond the first hard fat jet suppresses higher-multiplicity backgrounds, i.e. $t\bar{t}$+jets. Not vetoing additional fat jets approximately doubles the $t\bar{t}$ background, while the signal increases by roughly 30%. These numbers even hold in case regular jets with a cone size of $R = 0.4$ are vetoed instead. When vetoing jets large logarithms $\sim \ln^2(\sqrt{\hat{s}}/p_{T,\mathrm{veto}})$ can be induced which need to be resummed [145,146]. However, due to the high veto scale we do not expect these contributions to spoil the reliability of our analysis. As an alternative to jet vetos, 2-jet observables can be used to disentangle signal from background in this process [147,148].

- Zero b-tags. This considerably reduces the (until now dominant) $t\bar{t}$+jets background while having negligible effect on the signal. We re-cluster the hadronic activity into jets, again using the C/A algorithm but now with $R_{\mathrm{jet}} = 0.5$, to use for the b-tagging. We assume a flat 70 % (1 %) efficiency for b (light quark or gluon) initiated jets, i.e. a 30 % (99 %) probability for such a jet *not* to provoke the veto. We only consider b-jets of $p_{T,b} > 30$ GeV and $\lvert \eta_b \rvert < 2.5$.

The efficiencies of these cuts for the signal and various backgrounds are shown in the first part of Table 3. At this stage the backgrounds from $WW/Z/t\bar{t}$+jets are seen to contribute at similar levels. The set of cuts described so far are common to both our $H \to \tau_\ell \tau_\ell$ and $H \to W_\ell W_\ell^*$ analyses; from this point onwards they diverge.

4.2 $H \to \tau_\ell \tau_\ell$ analysis

The Higgs mass in the decay $H \to \tau\tau$ can be reconstructed using the *collinear approximation* [123]. The large hierarchy between the Higgs and tau masses ensures a very large boost for the taus, highly collimating their visible and invisible decay products. We can approximate the neutrino momenta by a decomposition of the missing transverse momentum, which assumes that each invisible momentum is parallel to the corresponding visible momentum. (This procedure can be extended to decays of more than one particle; see [149].) As was noted in [123], and further explored in [150], this procedure gains sensitivity with increasing transverse momentum of the Higgs—i.e. when the Higgs recoils against a hard jet. It suffers for a low-p_T Higgs because the two τ daughters are then nearly back-to-back, providing a poor basis for the $\not{p}_T$ decomposition. For our high-p_T Higgs study the mass reconstruction of the signal in this manner is very good and provides a sharp peak.[3]

In more detail, the Higgs mass in $H \to \tau_\ell \tau_\ell$ is reconstructed via the collinear approximation as follows. We require the missing transverse momentum $\not{p}_T$ to be inside the two leptons (more precisely, projecting the two lepton momenta into the transverse plane defines two segments; 'inside' the leptons means inside the smaller segment). We decompose $\not{p}_T$ as a linear combination of the two lepton momenta (defining for it a longitudinal component in the process):

$$\not{p}_T = \mathbf{p}_{T,\nu_1,\mathrm{col}} + \mathbf{p}_{T,\nu_2,\mathrm{col}} : \qquad \mathbf{p}_{\nu_1,\mathrm{col}} = \alpha_1 \mathbf{p}_{\ell_1},$$
$$\mathbf{p}_{\nu_2,\mathrm{col}} = \alpha_2 \mathbf{p}_{\ell_2}. \qquad (13)$$

---

[3] We have also tried a reconstruction using the $m_{T,\mathrm{Bound}}$-if-not-$m_{T,\mathrm{True}}$ prescription of [151] and found very similar results to the collinear approximation; the former is expected to be preferred at lower boosts of the Higgs which we do not consider here.
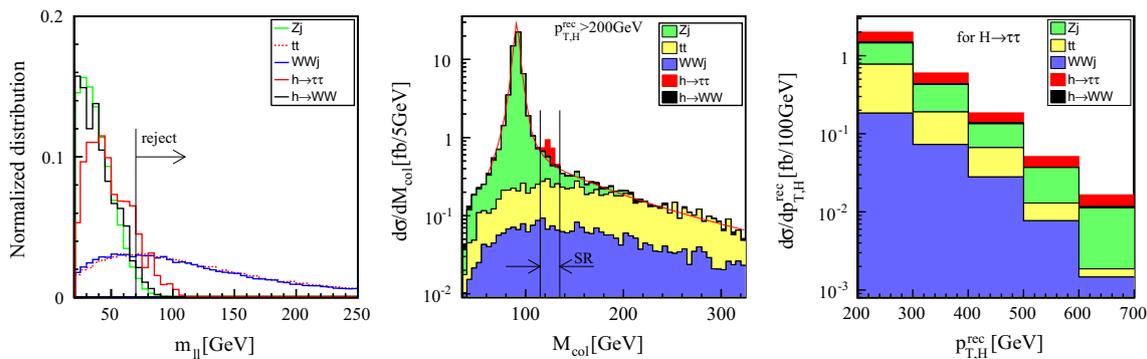
Fig. 2 *Left panel*: the invariant mass of the two leptons, $m_{\ell\ell}$, after cut 6. *Central panel*: the collinear mass M_{col} after cut 7, stacking the different processes. Histograms are normalized to the respective cross sections.

Right panel: stacked distributions of the 'Higgs' transverse momentum $p_{T,H}$ (defined in Eq. 11) after selection cut 8, with a logarithmic scale

The requirement that $\not{p}_T$ be inside the leptons is equivalent to demanding that the decomposition coefficients are both positive:

$$\alpha_1 > 0 \quad \text{and} \quad \alpha_2 > 0. \tag{14}$$

$\mathbf{p}_{\nu_1,col}$ and $\mathbf{p}_{\nu_2,col}$ thus defined approximate the neutrino three-momenta. Promoting them to massless four-momenta and adding them to the lepton four-momenta gives an approximate Higgs four-momentum, the mass of which we refer to as the collinear Higgs mass:

$$p_{col} = p_{\nu_1,col} + p_{\nu_2,col} + p_{\ell_1} + p_{\ell_2}, \quad M^2_{col} = p^2_{col}. \tag{15}$$

We apply one more cut before making use of the collinear mass variable: an upper limit for the dilepton mass, $m_{\ell\ell} < 70$ GeV. This cut reduces the $t\bar{t}$+jets and WW+jets backgrounds very efficiently while leaving most the H+jets signal and $Z \rightarrow \tau\tau$ background (see Fig. 2, left panel). At this stage $Z \rightarrow \tau\tau$ becomes the dominant background for extracting the $H \rightarrow \tau\tau$ signal. The size of the $t\bar{t}$ and WW backgrounds can be estimated in a data-driven way by removing $m_{\ell\ell} < 70$ GeV cuts. We discuss this in detail in the appendix.

The collinear mass is shown in the central panel of Fig. 2. Note that any particle decaying to $\tau_\ell \tau_\ell$ with enough boost that the two τ are not back-to-back will have its mass reconstructed by this process; indeed the most striking feature of the collinear mass distribution is the Z mass peak from the large irreducible $Z \rightarrow \tau_\ell \tau_\ell$ background. A peak due to the signal is visible at $M_{col} \sim m_H = 125$ GeV. By selecting events in the window $|M_{col} - m_H| < 10$ GeV we achieve a $S/B \sim 0.4$ with $S/\sqrt{B} > 9$ for 300 fb^{-1}. The signal is taken to include the $H \rightarrow WW^*$ contribution, which contributes about ~ 10 % the $H \rightarrow \tau\tau$ selection. We estimate the statistical error of the high p_T cross section measurement with $\sqrt{S+B}/S$. We obtain uncertainties of 12 % for $\sigma(p_{T,H} > 200$ GeV), 22 % for $\sigma(p_{T,H} > 300$ GeV), and 41 % for $\sigma(p_{T,H} > 400$ GeV), respectively. Assuming we

can achieve the same efficiencies for high-luminosity run of the LHC (HL-LHC) at 3 ab^{-1}, we obtain $\sim$ 4 % for $\sigma(p_{T,H} > 200$ GeV), $\sim$ 7 % for $\sigma(p_{T,H} > 300$ GeV), and $\sim$ 13 % for $\sigma(p_{T,H} > 400$ GeV).

As seen in the central panel of Fig. 2 the smooth sideband distribution can be used for estimating the background contribution. We show in the appendix that these side-bands are available even after hard $p^{rec}_{T,H}$ cuts. We therefore expect that a data-driven strategy for background estimation will be available, and take the statistical errors as a background uncertainty estimate. There will of course be further systematic uncertainties induced by MC background modeling.

In this analysis we mostly use the recoiling fat jet to remove the $t\bar{t}$+jets background. It could be beneficial to make use of the difference between the jet substructure of gluon and quark jets [152–156] since the dominant background at the last stage is Z+jets, which gives a different fraction of gluon and quark jets than the H+jets signal. We leave this for future work.

4.3 $H \rightarrow W_\ell W^*_\ell$ analysis

Our selection criteria for extracting $H \rightarrow W_\ell W^*_\ell$ from the background begin with those described in Sect. 4.1. In Sect. 4.2 we required that the $\not{p}_T$ vector be inside the two lepton momenta, after which the signal was dominated by $H \rightarrow \tau_\ell \tau_\ell$ and the background by $Z_{\rightarrow \tau_\ell \tau_\ell}$+jets. Here we will remove most of the contribution of these processes by requiring that $\not{p}_T$ be *outside* the two lepton momenta. This is equivalent to demanding that the $m^{\ell\ell}_{T2}$ variable [157] be greater than zero, as $m^{\ell\ell}_{T2} = 0$ when this is not satisfied—the 'trivial zero' [158]. In fact we go further and impose

$$m^{\ell\ell}_{T2} > 10 \text{ GeV}. \tag{16}$$

This rejects essentially all of the contributions from $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$+jets, which have the same end point close to
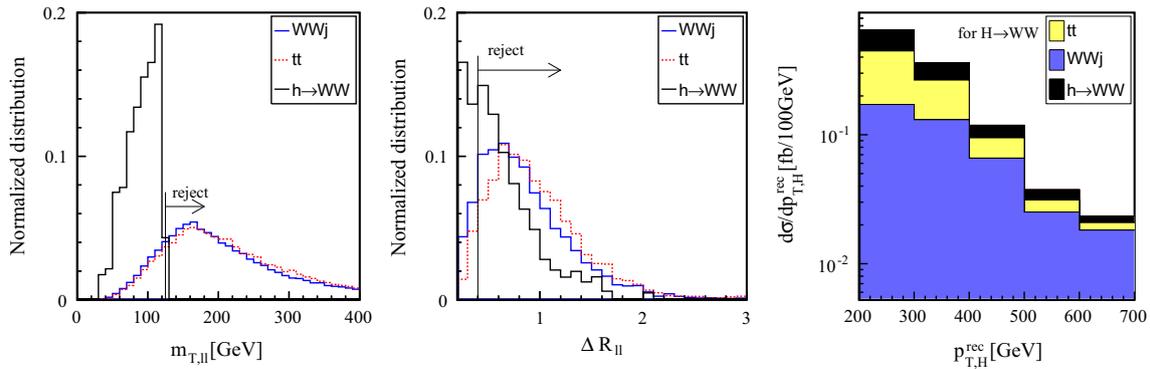
Fig. 3 Distributions of the transverse mass $m_{T,\ell\ell}$ after all selection cuts up to 6′ imposed (*left panel*) and the dilepton separation $\Delta R_{\ell\ell}$ after all selection cuts up to 7′ imposed (*central panel*). *Right*: stacked distribution of the 'Higgs' transverse momentum $p_{T,H}^{rec}$ (defined in Eq. 11) after *all* selection cuts for $H \to W_\ell W_\ell^*$ optimization, with a logarithmic scale

Table 4 Cut efficiencies for our analysis aimed at $H \to W_\ell W_\ell^*$, continued from the first part of Table 3. The values for each process are cross sections in fb. $S/\sqrt{B}$ has been calculated for 300 fb^{-1}

Event rate (fb)	$H \to WW^*$	$H \to \tau\tau$	$W_\ell W_\ell$+jets	$Z_{\to\tau\tau}$+jets	$t_\ell \bar{t}_\ell$+jets	S/B	$S/\sqrt{B}$
5. $n_b = 0$	1.825	0.940	48.855	57.068	105.851	0.01	3.29
6′. $m_{T2}^{\ell\ell} > 10$ GeV	1.096	0.002	25.241	0.028	53.730	0.01	2.14
7′. $m_{T,\ell\ell} < 125$ GeV	1.095	0.002	3.809	0.023	7.235	0.10	5.70
8′. $\Delta R_{\ell\ell} < 0.4$	0.330	0.000	0.426	0.002	0.450	0.38	6.11
$p_{T,H}^{rec} > 300$ GeV	0.128	0.000	0.254	0.002	0.175	0.30	3.38
$p_{T,H}^{rec} > 400$ GeV	0.034	0.000	0.124	0.000	0.040	0.21	1.44
$p_{T,H}^{rec} > 500$ GeV	0.010	0.000	0.058	0.000	0.011	0.14	0.65
$p_{T,H}^{rec} > 600$ GeV	0.004	0.000	0.033	0.000	0.005	0.10	0.33

m_τ. Allowing for endpoint smearing we cut a little harder at 10 GeV instead of m_τ.

We are now left with $H \to W_\ell W_\ell^*$ as our signal process, competing with the $WW/t\bar{t}$+jets backgrounds. Their kinematics unfortunately allow for little discriminating power: all of them contain two leptonic W bosons, with no possibility of mass reconstruction. Luckily, the transverse mass provides some discrimination. As shown in [159], the transverse mass variable satisfying $m_{T,\ell\ell} \leq m_h$ that gives the *greatest* lower bound on the Higgs mass in its decay to $W_\ell W_\ell^*$

$$m_{T,\ell\ell}^2 = m_{\ell\ell}^2 + 2(E_{T,\ell\ell}\not{E}_T - \mathbf{p}_{T,\ell\ell} \cdot \not{\mathbf{p}}_T), \quad (17)$$

where $E_{T,\ell\ell} = (m_{\ell\ell}^2 + p_{T,\ell\ell}^2)^{1/2}$ is the transverse energy of the dilepton system, and $\not{E}_T = |\not{\mathbf{p}}_T|$ is the missing transverse energy. We adopt this definition of $m_{T,\ell\ell}$, also used by the ATLAS Collaboration [160].[4] The end point at m_H for the transverse mass of the signal is shown in the left panel of

Fig. 3, where all the selection cuts up to step 6′ in Table 4 have been applied. We therefore impose

$$m_{T,\ell\ell} < m_H = 125 \text{ GeV}. \quad (18)$$

Finally, backgrounds are further suppressed by requiring that the leptons have similar directions,

$$\Delta R_{\ell\ell} < 0.4, \quad (19)$$

which is typically the case for the signal due to the aforementioned spin correlations.

The efficiencies of the cuts aimed at $H \to W_\ell W_\ell^*$ are shown in Table 4, together with the last common cut—the b veto. We finally find $S/B \sim 0.4$, with $S/\sqrt{B} > 6$ for 300 fb^{-1}. The table also shows the event numbers left after increased p_T cuts on the reconstructed Higgs. The resulting reconstructed Higgs $p_{T,H}^{rec}$ distributions are shown in Fig. 3 (right panel), stacked with the signal and background processes. As $p_{T,H}^{rec}$ increases, the signal over background ratio drops faster for the WW mode selection than for the $\tau\tau$ selection.

---

[4] Setting the dilepton mass $m_{\ell\ell}$ to zero in Eq. (17), despite its non-zero value being measured, gives the transverse mass used by the CMS Collaboration [161]; this results in a less steep end point.

5 Discussion

In this section, we discuss how much of the difference in the $p_{T,H}$ distributions due to the modified couplings can be observed after the realistic reconstruction of the previous section has been performed. The left panel of Fig. 4 shows the signal $M_{\rm col}$ distributions for the model points after applying the analysis described in Sect. 4.2 up to cut 7. We see the peak in the observable for all points. The central and right

panel show the signal $p_{T,H}^{\rm rec}$ distributions after the reconstruction described in Sect. 4 for $H \rightarrow \tau\tau$ and $H \rightarrow W_\ell W_\ell$ optimizations, respectively. As we expect, the difference in shape expected from the parton-level result of Fig. 1 manifests itself also in the reconstructed $p_{T,H}^{\rm rec}$ distributions. A detailed breakdown after successive selection cuts is shown in Table 5 for the $H \rightarrow \tau\tau$ optimization and in Table 6 for the $H \rightarrow W_\ell W_\ell$ optimization, quoting cross sections relative to the corresponding SM value. Compared with the parton level
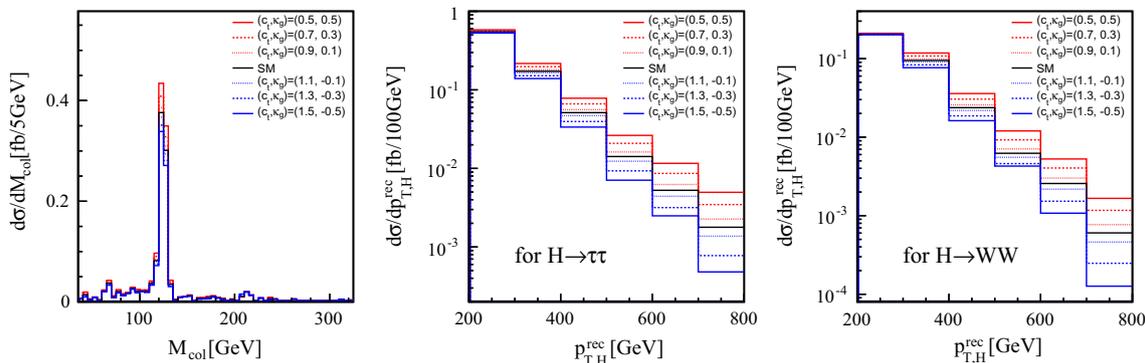
Fig. 4 Signal distributions for the SM and six model points, normalized to the respective cross sections. *Left panel*: the collinear mass $M_{\rm col}$ after cut 7. $p_{T,H}^{\rm rec}$ is shown for $H \rightarrow \tau\tau$ ($H \rightarrow W_\ell W_\ell$) in the *central* (*right*) *panel* after all optimized selection cuts

Table 5 The relative cross section $\sigma/\sigma_{\rm SM}$ for several new physics model points after successive selection cuts for $\tau\tau$ optimization

Model point (κ_g)	0.5	0.4	0.3	0.2	0.1	0 (SM)	−0.1	−0.2	−0.3	−0.4	−0.5
3. $p_{T,H}^{\rm rec} > 200$ GeV	1.109	1.084	1.061	1.039	1.019	1.000	0.983	0.968	0.954	0.942	0.932
4. $n_j^{\rm fat} = 1$	1.143	1.110	1.079	1.050	1.024	1.000	0.978	0.959	0.941	0.926	0.913
5. $n_b = 0$	1.143	1.110	1.079	1.050	1.024	1.000	0.978	0.959	0.941	0.926	0.913
6. $\not{p}_T$ inside two ℓs	1.156	1.120	1.086	1.055	1.026	1.000	0.976	0.954	0.935	0.918	0.903
7. $m_{\ell\ell} < 70$ GeV	1.157	1.121	1.087	1.056	1.027	1.000	0.976	0.954	0.934	0.917	0.902
8. $\|M_{\rm col} - m_H\| < 10$ GeV	1.163	1.125	1.091	1.058	1.028	1.000	0.974	0.951	0.930	0.912	0.896
$\quad p_{T,H}^{\rm rec} > 300$ GeV	1.392	1.303	1.219	1.140	1.067	1.000	0.938	0.882	0.831	0.785	0.745
$\quad p_{T,H}^{\rm rec} > 400$ GeV	1.711	1.544	1.389	1.247	1.117	1.000	0.895	0.802	0.722	0.653	0.597
$\quad p_{T,H}^{\rm rec} > 500$ GeV	2.131	1.857	1.607	1.381	1.179	1.000	0.845	0.715	0.608	0.525	0.465
$\quad p_{T,H}^{\rm rec} > 600$ GeV	2.602	2.201	1.840	1.520	1.240	1.000	0.801	0.642	0.523	0.445	0.407

Table 6 Relative size $\sigma/\sigma_{\rm SM}$ for several new physics model points after successive selection cuts for WW optimization

Model point (κ_g)	0.5	0.4	0.3	0.2	0.1	0 (SM)	−0.1	−0.2	−0.3	−0.4	−0.5
5. $n_b = 0$	1.143	1.110	1.079	1.050	1.024	1.000	0.978	0.959	0.941	0.926	0.913
6′. $m_{T2}^{\ell\ell} > 10$ GeV	1.117	1.089	1.064	1.040	1.019	1.000	0.983	0.968	0.955	0.944	0.936
7′. $m_{T,\ell\ell} < 125$ GeV	1.117	1.089	1.064	1.040	1.019	1.000	0.983	0.968	0.955	0.944	0.936
8′. $\Delta R_{\ell\ell} < 0.4$	1.164	1.125	1.088	1.056	1.026	1.000	0.977	0.958	0.942	0.929	0.920
$\quad p_{T,H}^{\rm rec} > 300$ GeV	1.360	1.278	1.201	1.129	1.062	1.000	0.943	0.892	0.845	0.803	0.767
$\quad p_{T,H}^{\rm rec} > 400$ GeV	1.684	1.520	1.370	1.234	1.110	1.000	0.903	0.819	0.749	0.692	0.648
$\quad p_{T,H}^{\rm rec} > 500$ GeV	2.093	1.822	1.577	1.358	1.166	1.000	0.860	0.747	0.659	0.598	0.564
$\quad p_{T,H}^{\rm rec} > 600$ GeV	2.377	2.043	1.738	1.463	1.217	1.000	0.812	0.654	0.525	0.425	0.354
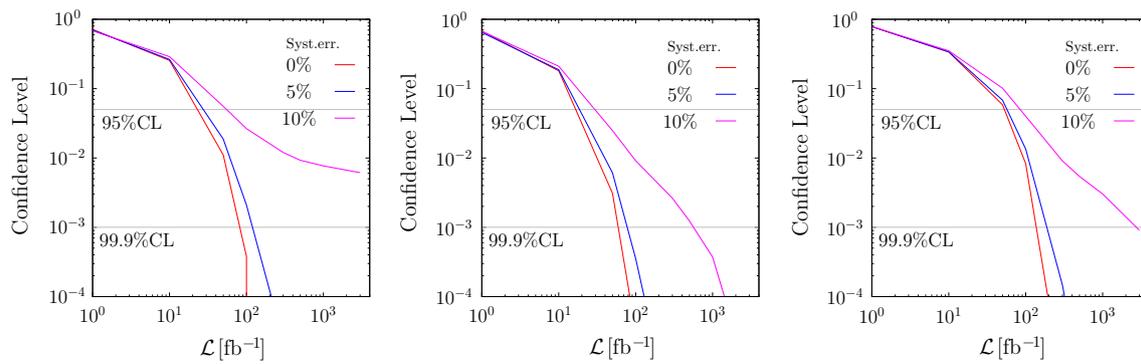
Fig. 5 CL_s vs. the integrated luminosity for the model points $\kappa_g = 0$ (SM, *left*) and $\kappa_g = 0.5$ (*central*) against a background-only hypothesis using the $\tau\tau$ mode. *Right panel*: CL_s plot for the model point of $\kappa_g = 0.5$ against a background-only hypothesis using the WW mode
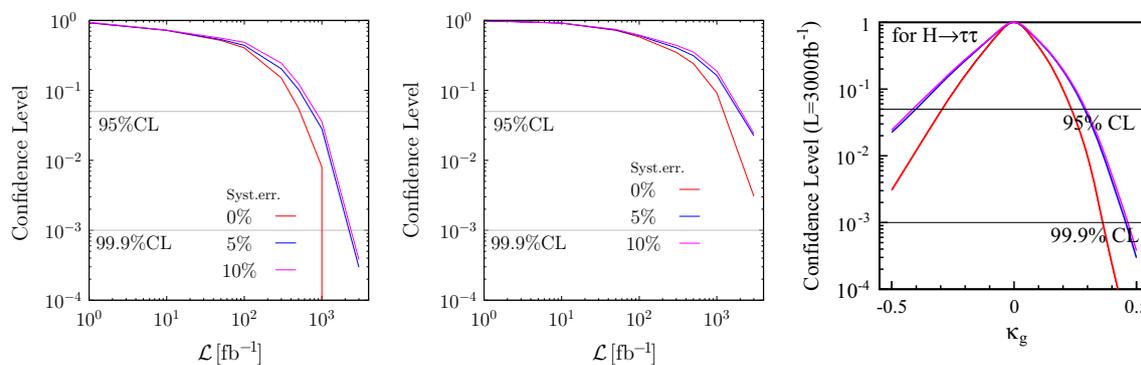
Fig. 6 CL_s vs. the integrated luminosity using $\tau\tau$ mode for the model point of $\kappa_g = 0.5$ (*left*) and $\kappa_g = -0.5$ (*central*). *Right*: CL_s as a function of κ_g for an integrated luminosity of 3000 fb^{-1}

numbers in Table 1, the $p_{T,H}^{rec}$ dependence is more enhanced at the reconstructed level. This is because most of the selection cuts are more efficient for the boosted Higgs event topology.

We will now estimate how much integrated luminosity is needed to find a certain significance for the signal. We perform a binned likelihood analysis of signal and background using the CL_s method, as described in [162]. We include systematic errors on the cross section normalization assuming a Gaussian probability distribution.

Figure 5 shows the expected p values as a function of the integrated luminosity $\mathcal{L}$ in the SM (left panel), the model point of $\kappa_g = 0.5$ (central panel) using the $H \to \tau\tau$ analysis and $\kappa_g = 0.5$ using the $H \to W_\ell W_\ell$ analysis (right panel). The analysis is based on the expected signal-plus-background against a background-only hypothesis. In the analysis, three different systematic errors on the cross section normalization of 0, 5, and 10 % are assumed. While achieving theoretical uncertainties of less than 10 % is challenging, in the separation of signal and background we rely predominantly on the lepton momenta which can be measured very precisely. As one can see from the left panel in Fig. 5, with

$\mathcal{L} = 20 \sim 60$ fb^{-1}, we are able to see the SM signal at 95 % confidence level depending on the assumed systematic uncertainty.

For $\kappa_g > 0$, the signal is enhanced and the required integrated luminosity decreases: it would be $\mathcal{L} = 15 \sim 30$ fb^{-1} for $\kappa_g = 0.5$ to observe the signal at 95 % CL, as shown in the central panel.

The right panel of Fig. 5 shows the p values for $\kappa_g = 0.5$ using the $H \to W_\ell W_\ell$ mode. The sensitivity compared to the $\tau\tau$ mode is slightly reduced. However, it is still possible to exploit the $W_\ell W_\ell$ final state to observe a boosted Higgs boson.

We also perform a binned likelihood analysis to estimate how well we can distinguish these model points from the SM given the presence of backgrounds. The left panel of Fig. 6 shows the expected p values to observe the signal and background against the SM and background hypothesis as a function of the integrated luminosity $\mathcal{L}$ for the model point of $\kappa_g = 0.5$ using the $H \to \tau\tau$ analysis. Again, systematic errors of 0, 5, and 10 % are assumed. We find that we are able to distinguish the model point $\kappa_g = 0.5$ from the SM with $\mathcal{L} = 1000$ fb^{-1} even assuming 10 % systematic uncertainty.

It is more difficult to prove a deviation from the SM for model points with $\kappa_g < 0$, compared to $\kappa_g > 0$ with the same $|\kappa_g|$ value, since this gives a deficit rather than a surplus of signal events. The central panel of Fig. 6 shows the p values for $\kappa_g = -0.5$ using the $H \rightarrow \tau\tau$ analysis. As expected we have less sensitivity, and even smaller values of $|\kappa_g|$ require larger integrated luminosities.

The right panel of Fig. 6 shows the p values as a function of κ_g using the $H \rightarrow \tau\tau$ for an integrated luminosity of 3000 fb^{-1}. If we assume 0 % systematic uncertainty we can exclude $\kappa_g < -0.29$ and $\kappa_g > 0.24$ for $\mathcal{L} = 3000$ fb^{-1} at 95 % CL. For the same integrated luminosity, assuming 10 % systematic uncertainty, we can still exclude $\kappa_g < -0.4$ and $\kappa_g > 0.3$ at 95 % CL.

We have not combined the $\tau\tau$ and $W_\ell W_\ell$ analyses although it could improve our sensitivity by some amount. Combining both channels is a complex task since the systematic uncertainties of both channels have to be evaluated by the experimental collaborations. Furthermore, it is not easy to avoid double-counting of events when combining both decay modes, as the final state reconstructions discussed in Sect. 4 are not able to strictly separate them (see Table 3).

6 Conclusions

The dominant production mode of the Higgs boson at the LHC—gluon fusion—is an important probe of new physics. Even though the inclusive rate has been measured to be in agreement with the SM, the study of a Higgs boosted by recoil against a hard jet constitutes an interesting, albeit challenging, measurement. It is motivated in the context of supersymmetry and composite Higgs models, and indeed generically in natural new physics: the Higgs coupling to a top-quark loop is both central to the question of natural electroweak symmetry breaking, and the chief source of gluon fusion. Due to the low energy theorem, however, the details of this loop-induced process are entirely obscured unless one can access the boosted Higgs regime.

We have shown boosted Higgs signal isolation in the dilepton channel via $H \rightarrow \tau\tau$ and $H \rightarrow WW$. The boost enhances the efficiency of the collinear approximation for mass reconstruction in the $H \rightarrow \tau\tau$ mode, giving a peak at m_H visible above the dominant Z+jets background. Z+jets provides its own peak for this reconstructed mass distribution; using the side-bands around the m_H peak we expect a relatively precise background estimate. In the end we achieve $S/B \sim 0.4$. For $H \rightarrow WW$ mode, we can also achieve $S/B \sim 0.4$ but with fewer events. This is nevertheless a helpful addition to the statistical significance. We expect a 12 % error for the cross section measurement for $p_T > 200$ GeV,

22 % for $p_T > 300$ GeV, and 41 % for $p_T > 400$ GeV with an integrated luminosity of 300 fb^{-1}.

A direct measurement of the top Yukawa coupling in the $t\bar{t}H$ channel is also instrumental for breaking the degeneracy concerning the coupling of the Higgs to gluons and to the top quark, and the H+jets mode provides a complementary determination. We have shown that we can distinguish several new physics models in an effective field theory approach using the reconstructed Higgs p_T distribution. With an integrated luminosity of 3000 fb^{-1} at the 14 TeV LHC, we can exclude $\kappa_g < -0.4$ and $\kappa_g > 0.3$ along the line $c_t + \kappa_g = 1$ at 95 % confidence level assuming the systematic uncertainty of 10 %.

Appendix: M_{col} distributions and data-driven background estimates

We collect distributions of the collinear mass M_{col} for several minimal values of the reconstructed Higgs p_T and discuss how a data-driven background estimate could be performed. In Fig. 7, we show distributions of M_{col} for $p_{T,H}^{\text{rec}} > 200$, 300, and 400 GeV. The upper three plots include the selection cuts up to step 7 in Table 3, while the lower three plots are up to cut 6 (i.e. without the $m_{\ell\ell}$ cut). The red lines show the fitting curves for the background distributions. We take the fitting function as the sum of a Breit–Wigner function and a log-normal function. As one can see, the Z-peak and the tail distributions are well fitted for a wide $p_{T,H}^{\text{rec}}$ range. This means we can estimate the contributions of the background processes using side-bands, which reduces the sensitivity to theoretical uncertainties.

Moreover, the lower plots without the $m_{\ell\ell}$ cut have larger $t\bar{t}$ and WW contributions but are still well fitted with the same fitting function. Thus, we can extract the normalizations of $t\bar{t}$ and WW contributions and control part of the Monte Carlo uncertainties using data. We therefore only consider the statistical uncertainty of the total background contributions in the signal region in the main text.
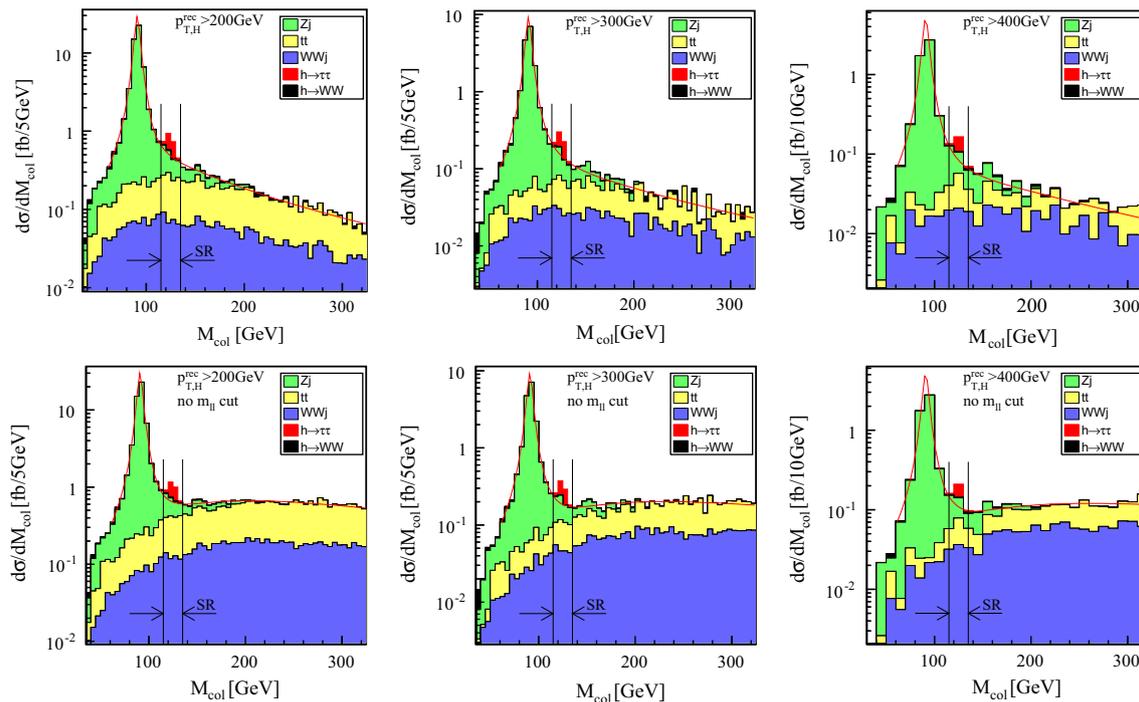
Fig. 7 The collinear mass M_{col} distributions after (*upper*) and before (*lower*) cut 7, stacking the different processes. Histograms are normalized to the respective cross sections. The p_T cut on the reconstructed Higgs are $p_{T,H}^{rec} > 200$ GeV, $p_{T,H}^{rec} > 300$ GeV, and $p_{T,H}^{rec} > 400$ GeV from *left* to *right*

References

1. ATLAS Collaboration, G. Aad et al., Phys. Lett. B **716**, 1 (2012). arXiv:1207.7214
2. CMS Collaboration, S. Chatrchyan et al., Phys. Lett. B **716**, 30 (2012). arXiv:1207.7235
3. ATLAS Collaboration, G. Aad et al., Phys. Lett. B **726**, 120 (2013). arXiv:1307.1432
4. S. Bolognesi et al., Phys. Rev. D **86**, 095031 (2012). arXiv:1208.4018
5. ATLAS Collaboration, CERN Report No. ATLAS-CONF-2013-029 (2013) (unpublished)
6. CMS Collaboration, CERN Report No. CMS-PAS-HIG-13-016 (2013) (unpublished)
7. L. Landau, Dokl. Akad. Nauk Ser. Fiz. **60**, 207 (1948)
8. C.-N. Yang, Phys. Rev. **77**, 242 (1950)
9. ATLAS Collaboration, CERN Report No. ATLAS-CONF-2013-031 (2013) (unpublished)
10. CMS Collaboration, CERN Report No. CMS-PAS-HIG-13-003 (2013) (unpublished)
11. J. Ellis, R. Fok, D. S. Hwang, V. Sanz, T. You, Eur. Phys. J. C **73**, 2488 (2013). arXiv:1210.5229
12. ATLAS Collaboration, CERN Report No. ATLAS-CONF-2013-013 (2013) (unpublished)
13. CMS Collaboration, S. Chatrchyan et al., (2013). arXiv:1312.5353
14. S. Choi, D. Miller, M. Muhlleitner, P. Zerwas, Phys. Lett. B **553**, 61 (2003). hep-ph/0210077
15. C. Buszello, I. Fleck, P. Marquard, J. van der Bij, Eur. Phys. J. **C32**, 209 (2004). hep-ph/0212396
16. A. Bredenstein, A. Denner, S. Dittmaier, M. Weber, Phys. Rev. **D74**, 013004 (2006). hep-ph/0604011
17. Q.-H. Cao, C. Jackson, W.-Y. Keung, I. Low, J. Shu, Phys. Rev. D **81**, 015010 (2010). arXiv:0911.3398
18. Y. Gao et al., Phys. Rev. D **81**, 075022 (2010). arXiv:1001.3396
19. C. Englert, C. Hackstein, M. Spannowsky, Phys. Rev. D **82**, 114024 (2010). arXiv:1010.0676
20. J. Ellis, D. S. Hwang, JHEP **1209**, 071 (2012). arXiv:1202.6660
21. J. Ellis, V. Sanz, T. You, Phys. Lett. B **726**, 244 (2013). arXiv:1211.3068
22. A. Freitas, P. Schwaller, Phys. Rev. D **87**, 055014 (2013). arXiv:1211.1980
23. B. Coleppa, K. Kumar, H. E. Logan, Phys. Rev. D **86**, 075022 (2012). arXiv:1208.2692
24. T. Plehn, D.L. Rainwater, D. Zeppenfeld, Phys. Rev. Lett. **88**, 051801 (2002). hep-ph/0105325
25. G. Klamke, D. Zeppenfeld, JHEP **0704**, 052 (2007). hep-ph/0703202
26. R. M. Godbole, D. Miller, M. M. Muhlleitner, JHEP **0712**, 031 (2007). arXiv:0708.0458
27. S. Berge, W. Bernreuther, B. Niepelt, H. Spiesberger, Phys. Rev. D **84**, 116003 (2011). arXiv:1108.0670
28. C. Englert, M. Spannowsky, M. Takeuchi, JHEP **1206**, 108 (2012). arXiv:1203.5788
29. A. Djouadi, G. Moreau (2013). arXiv:1303.6591
30. R. Harnik, A. Martin, T. Okui, R. Primulando, F. Yu, Phys. Rev. D **88**, 076009 (2013). arXiv:1308.1094
31. J. Ellis, D. S. Hwang, K. Sakurai, M. Takeuchi, JHEP **1404**, 004 (2014). arXiv:1312.5736
32. The ATLAS collaboration, CERN Report No. ATLAS-CONF-2014-009 (2014) (unpublished)
33. CMS Collaboration, S. Chatrchyan et al. (2014). arXiv:1401.6527
34. I. Low, R. Rattazzi, A. Vichi, JHEP **1004**, 126 (2010). arXiv:0907.5413

35. J. R. Espinosa, C. Grojean, V. Sanz, M. Trott, JHEP **1212**, 077 (2012). arXiv:1207.7355
36. M.A. Shifman, A. Vainshtein, M. Voloshin, V.I. Zakharov, Sov. J. Nucl. Phys. **30**, 711 (1979)
37. J.R. Ellis, M.K. Gaillard, D.V. Nanopoulos, Nucl. Phys. **B106**, 292 (1976)
38. B.A. Kniehl, M. Spira, Z. Phys. **C69**, 77 (1995). hep-ph/9505225
39. M. Grazzini, H. Sargsyan, JHEP **1309**, 129 (2013). arXiv:1306.4581
40. A. Falkowski, Phys. Rev. D **77**, 055018 (2008). arXiv:0711.0828
41. I. Low, A. Vichi, Phys. Rev. D **84**, 045019 (2011). arXiv:1010.2753
42. A. Azatov, J. Galloway, Phys. Rev. D **85**, 055013 (2012). arXiv:1110.5646
43. M. Montull, F. Riva, E. Salvioni, R. Torre, Phys. Rev. D **88**, 095006 (2013). arXiv:1308.0559
44. C. Delaunay, C. Grojean, G. Perez, JHEP **1309**, 090 (2013). arXiv:1303.5701
45. C. Grojean, E. Salvioni, M. Schlaffer, A. Weiler, JHEP **1405**, 022 (2014). arXiv:1312.3317
46. U. Langenegger, M. Spira, A. Starodumov, P. Trueb, JHEP **0606**, 035 (2006). hep-ph/0604156
47. C. Arnesen, I. Z. Rothstein, J. Zupan, Phys. Rev. Lett. **103**, 151801 (2009). arXiv:0809.1429
48. E. Bagnaschi, G. Degrassi, P. Slavich, A. Vicini, JHEP **1202**, 088 (2012). arXiv:1111.2854
49. A. Azatov, A. Paul, JHEP **1401**, 014 (2014). arXiv:1309.5273
50. R. V. Harlander, T. Neumann, Phys. Rev. D **88**, 074015 (2013). arXiv:1308.2225
51. A. Banfi, A. Martin, V. Sanz (2013). arXiv:1308.4771
52. C. Englert, M. McCullough, M. Spannowsky, Phys. Rev. D **89**, 013013 (2014). arXiv:1310.4828
53. R. V. Harlander, S. Liebler, T. Zirke, JHEP **1402**, 023 (2014). arXiv:1307.8122
54. T. Plehn, G. P. Salam, M. Spannowsky, Phys. Rev. Lett. **104**, 111801 (2010). arXiv:0910.5472
55. The ATLAS collaboration, CERN Report No. ATLAS-CONF-2014-011 (2014) (unpublished)
56. CMS Collaboration, CERN Report No. CMS-PAS-HIG-13-015 (2013) (unpublished)
57. CMS Collaboration, CERN Report No. CMS-PAS-HIG-13-019 (2013) (unpublished)
58. CMS Collaboration, CERN Report No. CMS-PAS-HIG-13-020 (2013) (unpublished)
59. P. Artoisenet, P. de Aquino, F. Maltoni, O. Mattelaer, Phys. Rev. Lett. **111**, 091802 (2013). arXiv:1304.6414
60. M. R. Buckley, T. Plehn, T. Schell, M. Takeuchi, JHEP **1402**, 130 (2014). arXiv:1310.6034
61. K. Agashe, R. Contino, A. Pomarol, Nucl. Phys. **B719**, 165 (2005). hep-ph/0412089
62. R. Contino (2010). arXiv:1005.4269
63. B. Bellazzini, C. Cski, J. Serra (2014). arXiv:1401.2457
64. M. Gillioz, R. Grober, C. Grojean, M. Muhlleitner, E. Salvioni, JHEP **1210**, 004 (2012). arXiv:1206.7120
65. CMS Collaboration, S. Chatrchyan et al., Phys. Lett. B **729**, 149 (2014). arXiv:1311.7667
66. ATLAS Collaboration, CERN Report No. ATLAS-CONF-2013-018 (2013) (unpublished)
67. ATLAS collaboration, CERN Report No. ATLAS-CONF-2013-056 (2013) (unpublished)
68. ATLAS collaboration, CERN Report No. ATLAS-CONF-2013-051 (2013) (unpublished)
69. A. De Simone, O. Matsedonskyi, R. Rattazzi, A. Wulzer, JHEP **1304**, 004 (2013). arXiv:1211.5663
70. J. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, M. Perez-Victoria, Phys. Rev. D **88**, 094010 (2013). arXiv:1306.0572
71. M. Buchkremer, G. Cacciapaglia, A. Deandrea, L. Panizzi, Nucl. Phys. B **876**, 376 (2013). arXiv:1305.4172
72. A. Azatov, M. Salvarezza, M. Son, M. Spannowsky, Phys. Rev. D **89**, 075001 (2014). arXiv:1308.6601
73. J. Kearney, A. Pierce, J. Thaler, JHEP **1308**, 130 (2013). arXiv:1304.4233
74. C. Delaunay et al., JHEP **1402**, 055 (2014). arXiv:1311.2072
75. Z. Kunszt, F. Zwirner, Nucl. Phys. **B385**, 3 (1992). hep-ph/9203223
76. V.D. Barger, M. Berger, A. Stange, R. Phillips, Phys. Rev. **D45**, 4128 (1992)
77. H. Baer, M. Bisset, C. Kao, X. Tata, Phys. Rev. **D46**, 1067 (1992)
78. J.F. Gunion, L.H. Orr, Phys. Rev. **D46**, 2052 (1992)
79. J.F. Gunion, H.E. Haber, C. Kao, Phys. Rev. **D46**, 2907 (1992)
80. A. Djouadi, M. Spira, P. Zerwas, Phys. Lett. **B264**, 440 (1991)
81. M. Spira, A. Djouadi, D. Graudenz, P. Zerwas, Nucl. Phys. **B453**, 17 (1995). hep-ph/9504378
82. S. Dawson, Nucl. Phys. **B359**, 283 (1991)
83. D. Graudenz, M. Spira, P. Zerwas, Phys. Rev. Lett. **70**, 1372 (1993)
84. H.E. Haber, G.L. Kane, Nucl. Phys. **B250**, 716 (1985)
85. ATLAS, G. Aad et al., JHEP **1310**, 189 (2013). arXiv:1308.2631
86. ATLAS collaboration, CERN Report No. ATLAS-CONF-2013-068 (2013) (unpublished)
87. ATLAS collaboration, CERN Report No. ATLAS-CONF-2013-065 (2013) (unpublished)
88. ATLAS Collaboration, CERN Report No. ATLAS-CONF-2013-061 (2013) (unpublished)
89. ATLAS collaboration, CERN Report No. ATLAS-CONF-2013-048 (2013) (unpublished)
90. ATLAS collaboration, CERN Report No. ATLAS-CONF-2013-037 (2013) (unpublished)
91. ATLAS Collaboration, CERN Report No. ATLAS-CONF-2013-025 (2013) (unpublished)
92. ATLAS Collaboration, CERN Report No. ATLAS-CONF-2013-024 (2013) (unpublished)
93. CMS Collaboration, S. Chatrchyan et al., Eur. Phys. J. C **73**, 2677 (2013). arXiv:1308.1586
94. CMS Collaboration, CERN Report No. CMS-PAS-SUS-13-014 (2013) (unpublished)
95. CMS Collaboration, CERN Report No. CMS-PAS-SUS-13-008 (2013) (unpublished)
96. CMS Collaboration, CERN Report No. CMS-PAS-SUS-12-023 (2012) (unpublished)
97. M. Papucci, K. Sakurai, A. Weiler, L. Zeune (2014). arXiv:1402.0492
98. M. Papucci, J. T. Ruderman, A. Weiler, JHEP **1209**, 035 (2012). arXiv:1110.6926
99. A. Delgado, G. F. Giudice, G. Isidori, M. Pierini, A. Strumia, Eur. Phys. J. C **73**, 2370 (2013). arXiv:1212.6847
100. Y. Kats, D. Shih, JHEP **1108**, 049 (2011). arXiv:1106.0030
101. J. Fan, M. Reece, J. T. Ruderman, JHEP **1207**, 196 (2012). arXiv:1201.4875
102. J. Fan, M. Reece, J. T. Ruderman, JHEP **1111**, 012 (2011). arXiv:1105.5135
103. Y. Bai, A. Katz, B. Tweedie, JHEP **1401**, 040 (2014). arXiv:1309.6631
104. W. Buchmuller, D. Wyler, Nucl. Phys. **B268**, 621 (1986)
105. B. Grzadkowski, M. Iskrzynski, M. Misiak, J. Rosiek, JHEP **1010**, 085 (2010). arXiv:1008.4884
106. G. Giudice, C. Grojean, A. Pomarol, R. Rattazzi, JHEP **0706**, 045 (2007). hep-ph/0703164
107. R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner, M. Spira, JHEP **1307**, 035 (2013). arXiv:1303.3876
108. J. Elias-Miro, J. Espinosa, E. Masso, A. Pomarol, JHEP **1311**, 066 (2013). arXiv:1308.1879
109. A. Pomarol, F. Riva, JHEP **1401**, 151 (2014). arXiv:1308.2803

110. C. Englert et al. (2014). arXiv:1403.7191
111. J. Bramante, A. Delgado, A. Martin, Phys. Rev. D **89**, 093006 (2014). arXiv:1402.5985
112. J. Ellis, V. Sanz, T. You (2014). arXiv:1404.3667
113. B. Dumont, S. Fichet, G. von Gersdorff, JHEP **1307**, 065 (2013). arXiv:1304.3369
114. H. Mantler, M. Wiesemann, Eur. Phys. J. C **73**, 2467 (2013). arXiv:1210.8263
115. U. Baur, E.N. Glover, Nucl. Phys. **B339**, 38 (1990)
116. CERN Report No. ATL-PHYS-PUB-2013-014 (2013) (unpublished)
117. ATLAS Collaboration, A. Liss, J. Nielsen, CERN Report No. ATL-PHYS-PUB-2013-007 (2013) (unpublished)
118. CMS Collaboration, CERN Report No. CMS-NOTE-2012-006. CERN-CMS-NOTE-2012-006 (2012) (unpublished)
119. J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, T. Stelzer, JHEP **1106**, 128 (2011). arXiv:1106.0522
120. M. Bahr et al., Eur. Phys. J. C **58**, 639 (2008). arXiv:0803.0883
121. K. Arnold et al. (2012). arXiv:1205.4902
122. J. Bellm et al. (2013). arXiv:1310.6877
123. R.K. Ellis, I. Hinchliffe, M. Soldate, J. van der Bij, Nucl. Phys. **B297**, 221 (1988)
124. R. V. Harlander, T. Neumann, K. J. Ozeren, M. Wiesemann, JHEP **1208**, 139 (2012). arXiv:1206.0157
125. R. Boughezal, F. Caola, K. Melnikov, F. Petriello, M. Schulze, JHEP **1306**, 072 (2013). arXiv:1302.6216
126. S. Hoeche, F. Krauss, M. Schonherr (2014). arXiv:1401.7971
127. LHC Higgs Cross Section Working Group, S. Dittmaier et al. (2011). arXiv:1101.0593
128. C. Anastasiou, R. Boughezal, F. Petriello, JHEP **0904**, 003 (2009). arXiv:0811.3458
129. D. de Florian, M. Grazzini, Phys. Lett. B **674**, 291 (2009). arXiv:0901.2427
130. M. Bonvini, S. Marzani (2014). arXiv:1405.3654
131. M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A.D. Polosa, JHEP **0307**, 001 (2003). hep-ph/0206293
132. T. Sjostrand, S. Mrenna, P.Z. Skands, JHEP **0605**, 026 (2006). hep-ph/0603175
133. M.L. Mangano, M. Moretti, R. Pittau, Nucl. Phys. B **632**, 343 (2002). hep-ph/0108069
134. S. Hoeche et al. (2006). hep-ph/0602031
135. P. Nason, S. Dawson, R.K. Ellis, Nucl. Phys. **B303**, 607 (1988)
136. W. Beenakker, H. Kuijf, W. van Neerven, J. Smith, Phys. Rev. **D40**, 54 (1989)
137. S. Moch, P. Uwer, Phys. Rev. D **78**, 034003 (2008). arXiv:0804.1476
138. M. Cacciari, G. P. Salam, G. Soyez, Eur. Phys. J. C **72**, 1896 (2012). arXiv:1111.6097
139. Y.L. Dokshitzer, G. Leder, S. Moretti, B. Webber, JHEP **9708**, 001 (1997). hep-ph/9707323
140. M. Wobisch, T. Wengler (1998). hep-ph/9907280
141. B. Gripaios, K. Nagao, M. Nojiri, K. Sakurai, B. Webber, JHEP **1303**, 106 (2013). arXiv:1210.1938
142. A. Katz, M. Son, B. Tweedie, Phys. Rev. D **83**, 114033 (2011). arXiv:1011.4523
143. M. Dittmar, H. Dreiner, Report No. CMS-NOTE-1997-083, CERN-CMS-NOTE-1997-083 (1997) (unpublished)
144. M. Dittmar, H.K. Dreiner, Phys. Rev. **D55**, 167 (1997). hep-ph/9608317
145. C. F. Berger, C. Marcantonini, I. W. Stewart, F. J. Tackmann, W. J. Waalewijn, JHEP **1104**, 092 (2011). arXiv:1012.4480
146. A. Banfi, P. F. Monni, G. P. Salam, G. Zanderighi, Phys. Rev. Lett. **109**, 202001 (2012). arXiv:1206.4998
147. M. Buschmann, C. Englert, D. Goncalves, T. Plehn, M. Spannowsky, Phys. Rev. D **90**, 013010 (2014). arXiv:1405.7651
148. M.J. Dolan, P. Harris, M. Jankowiak, M. Spannowsky, **1406**, 3322 (2014)
149. M. Spannowsky, C. Wymant, Phys. Rev. D **87**, 074004 (2013). arXiv:1301.0345
150. B. Mellado, X. Quayle, S.L. Wu, Phys. Lett. B **611**, 60 (2005)
151. A. J. Barr, S. T. French, J. A. Frost, C. G. Lester, JHEP **1110**, 080 (2011). arXiv:1106.2322
152. J. Gallicchio, M. D. Schwartz, Phys. Rev. Lett. **107**, 172001 (2011). arXiv:1106.3076
153. ATLAS Collaboration, CERN Report No. ATLAS-CONF-2012-138, ATLAS-COM-CONF-2012-108 (2012) (unpublished)
154. CMS Collaboration, CERN Report No. CMS-PAS-JME-13-002 (2013) (unpublished)
155. A. J. Larkoski, G. P. Salam, J. Thaler, JHEP **1306**, 108 (2013). arXiv:1305.0007
156. S. Mukhopadhyay, M. M. Nojiri, T. T. Yanagida (2014). arXiv:1403.6028
157. C. Lester, D. Summers, Phys. Lett. B **463**, 99 (1999). hep-ph/9906349
158. C. G. Lester, JHEP **1105**, 076 (2011). arXiv:1103.5682
159. A. J. Barr, B. Gripaios, C. G. Lester, JHEP **0907**, 072 (2009). arXiv:0902.4864
160. ATLAS Collaboration, CERN Report No. ATLAS-CONF-2013-030 (2013) (unpublished)
161. CMS Collaboration, S. Chatrchyan et al., JHEP **1401**, 096 (2014). arXiv:1312.1129
162. T. Junk, Nucl. Instrum. Meth. A **434**, 435 (1999). hep-ex/9902006