Higgs Pair Production at Future Hadron Colliders: From Kinematics to Dynamics

Dorival Gonçalves,1 Tao Han,1 Felix Kling,2 Tilman Plehn,3 and Michihisa Takeuchi4

1PITT PACC, Department of Physics and Astronomy, University of Pittsburgh, USA
2Department of Physics and Astronomy, University of California, Irvine, USA
3Institut für Theoretische Physik, Universität Heidelberg, Germany
4Kavli IPMU (WPI), UTIAS, University of Tokyo, Kashiwa, Japan

The measurement of the triple Higgs coupling is a key benchmark for the LHC and future colliders. It directly probes the Higgs potential and its fundamental properties in connection to new physics beyond the Standard Model. There exist two phase space regions with an enhanced sensitivity to the Higgs self-coupling, the Higgs pair production threshold and an intermediate top pair threshold. We show how the invariant mass distribution of the Higgs pair offers a systematic way to extract the Higgs self-coupling, focusing on the leading channel $pp \to hh + X \to b\bar{b} \gamma \gamma + X$. We utilize new features of the signal events at higher energies and estimate the potential of a high-energy upgrade of the LHC and a future hadron collider with realistic simulations. We find that the high-energy upgrade of the LHC to 27 TeV would reach a 5σ observation with an integrated luminosity of 2.5 ab$^{-1}$. It would have the potential to reach 15% (30%) accuracy at the 68% (95%) confidence level to determine the SM Higgs boson self-coupling. A future 100 TeV collider could improve the self-coupling measurement to better than 5% (10%) at the 68% (95%) confidence level.

CONTENTS

I. Introduction 1
II. Higgs Pair Signature 2
III. The Mother of Distributions 3
IV. Detector-level Analysis 5
V. Summary and Outlook 8
VI. Appendix 9
References 10

I. INTRODUCTION

The discovery of the Higgs boson [1, 2] at the CERN Large Hadron Collider (LHC) is of monumental significance. The completion of the Standard Model (SM) provides us with a consistent theory valid up to high scales. As a perturbative gauge theory, it allows for precision predictions for essentially all LHC observables. In parallel, experimental advances have turned ATLAS and CMS into the first hadron collider precision experiments in history. In combination, these developments open new avenues to tackle fundamental physics questions at the LHC and future high-energy facilities.

On the theory side, we are still lacking an understanding of if and how the Higgs mass, the only dimensionful parameter in the theory, is stabilized against a large new physics scale. The Higgs potential responsible for the electroweak symmetry breaking (EWSB) in the SM is determined by the triple and quartic Higgs self-coupling $\lambda_{SM} \approx 1/8$. It is a true self-interaction in the sense that it is not associated with any conserved charge after EWSB.

With our ignorance for new physics beyond the SM, the shape of the Higgs potential is deeply linked to the fundamental question of electroweak symmetry breaking in the early universe, allowing for a slow second-order phase transition in the SM or a strong first-order phase transition with a modified Higgs potential. It has been argued that a wide range of modified Higgs potentials, which result in a strong first-order EW phase transition, lead to order-one modifications of $\lambda_{SM}$ [3]. All of this points to the Higgs self-coupling $\lambda$ as a benchmark measurement for the coming LHC runs, as well as any kind of planned colliders [4].

Higgs pair production $pp \to hh$ offers a direct path to pin down $\lambda$ at a hadron collider [5, 6]. Previous studies show that promising final states from the $hh$ decays are $b\bar{b}\gamma\gamma$ [7, 8], $b\bar{b}\tau\tau$ [9, 10], $b\bar{b}WW$ [11], $b\bar{b}b\bar{b}$ [12], and $4W$ [13]. Theoretical studies as well as current analyses point to the $b\bar{b}\gamma\gamma$ decay as the most promising signature at the LHC [14]. Combinations with indirect measurements of the self-coupling from quantum effects confirm that Higgs pair production provides the most robust self-coupling measurement [15]. For the high-luminosity LHC (HL-LHC), ATLAS and CMS projections indicate a very modest sensitivity to the Higgs self-coupling [16].

In anticipation to probe new physics beyond the SM, it is customary to parametrize the modification of the self-coupling as

$$\kappa_\lambda = \frac{\lambda}{\lambda_{SM}}.$$  

In the optimistic scenario that we can neglect systematic uncertainties, those studies indicate that the LHC will probe the coupling at 95% confidence level

$$-0.8 < \kappa_\lambda < 7.7.$$  

An issue with those studies is that they are based on the total rate for Higgs production, but neglect a wealth of
available information. Including a full kinematic analysis could lead to an improved measurement [17]

\[-0.2 < \kappa_\lambda < 2.6\,\,,
\]

falling short in precision in comparison to other Higgs property measurements at the LHC, and far from satisfactory in probing the Higgs potential.

In this study, we systematically compare the prospects for measuring the Higgs self-coupling at current and higher energy pp colliders. We focus on the two leading proposals for future hadron colliders:

1. the 27 TeV high-energy LHC (HE-LHC) with an integrated luminosity of 15 ab\(^{-1}\),

2. a 100 TeV hadron collider with 30 ab\(^{-1}\), under consideration at CERN (FCC-hh) [18] and in China (SppC) [19].

We include state of the art signal and background estimates for the \(b\bar{b}\gamma\gamma\) channel, as well as realistic acceptance cuts and efficiencies. While there exist a series of 100 TeV studies of Higgs pair production at different levels of sophistication [20], we include a 100 TeV analysis to be able to compare with the HE-LHC reach on equal footing.

We start with a study of relevant phase space regions using a Neyman-Pearson maximum likelihood approach [17, 21]. This allows us to estimate the impact of using simple kinematic distributions on the measurement of the Higgs self-coupling at the different colliders. Furthermore, we can evaluate the maximum significance of extracting the Higgs pair signal and the significance of detecting a modified self-coupling under idealized conditions.

In the main part of our paper, we perform a state-of-the-art analysis of Higgs pair production including additional jet radiation and a full set of realistic detector efficiencies. Unlike earlier analyses, we include \(b\)-jets from Higgs decays even when they become sub-leading in transverses momentum to the additional jet radiation. Our analysis focuses on the di-Higgs invariant mass distribution, both for the extraction of the Higgs pair signal and for the measurement of the Higgs self-coupling. Using a log-likelihood approach on this single kinematic distribution, we show that the Higgs self-coupling can be properly measured not only at a future 100 TeV collider, but also at the 27 TeV HE-LHC.

II. HIGGS PAIR SIGNATURE

The leading \(hh\) production mechanism in the Standard Model at hadron colliders is depicted by the Feynman diagrams in Fig. 1. Due to the difference of the top quark propagators in the loops, the two diagrams interfere destructively. In Fig. 2 we show the total rate for \(hh\) production as a function of the center of mass energy \(\sqrt{s}\) in TeV, including the next-to-leading order (NLO) corrections [22]. The width of the curve illustrated the theoretical uncertainties around 10% [23]. At the LHC, the signal rate is the limiting factor for Higgs pair studies. At 14 TeV, the cross section including higher-order corrections is in the range of 0.033 pb [23], corresponding to at most 100k events with an integrated luminosity of 3 ab\(^{-1}\) at the HL-LHC. Assuming one Higgs decay to tagged bottom quarks, the available rate is reduced to 60k events in the life time of the HL-LHC. The crucial question is what kind of second Higgs decay allows us to effectively trigger the events and to reduce the QCD backgrounds to a manageable level. The leading candidate is the signature [7]

\[pp \rightarrow hh \rightarrow b\bar{b}\gamma\gamma\,\,,\]

because of the excellent di-photon mass resolution and the guaranteed trigger. The expected number of signal events in the Standard Model at the HL-LHC is 260. Alternatively, the \(b\bar{b}\tau\tau\) signature leads to 7.2k events times the tau tagging probability rate squared, and hampered by a significantly worse signal-to-background ratio.

Because of the rapidly growing gluon luminosity at higher energies, the \(hh\) production cross section increases by about a factor of 4 (40) at 27 (100) TeV. This means that at the HE-LHC with the anticipated integrated lu-
minosity of 15 ab$^{-1}$ the number of events in the $bb\gamma\gamma$ channel increases by a factor $4 \times 5 = 20$ to around 5k events. A 100 TeV hadron collider with a projected integrated luminosity of 30 ab$^{-1}$ features another increase by a factor $10 \times 2 = 20$, to around 100k expected Higgs pair events in the Standard Model.

This estimate shows how the combination of increased energy and increased luminosity slowly turns Higgs pair production into a valid channel for precision measurements. The numbers fundamentally affect our proposed analysis strategy, because the small number of signal and background events suggests a kinematic analysis including as few kinematic distributions as possible. It is possible to improve this situation, for example, using the matrix element technique, as we will discuss below.

We generate the signal with MadGraph5 [24], accounting for a next-to-leading order (NLO) QCD factor $K_{NLO} \sim 1.6$ [22]. In the final state we demand two $b$-tagged jets and two isolated photons with the minimal acceptance and trigger cuts

$$\begin{align*}
|p_T,j| > 30 \text{ GeV}, & \quad |\eta_j| < 2.5 , \\
|p_T,\gamma| > 30 \text{ GeV}, & \quad |\eta_\gamma| < 2.5 , \\
\Delta R_{\gamma,\gamma,j,j} > 0.4 .
\end{align*}$$

(5)

The background to our $bb\gamma\gamma$ signal consists of other Higgs production modes ($tth, Zh$) with $h \to \gamma\gamma$, continuum $bb\gamma\gamma$ production, and of multi-jet events with light-flavor jets faking either photons or $b$-jets ($jj\gamma\gamma, bb\gamma\gamma$) [7]. The different backgrounds are discussed in detail in Sec. IV.

The proper simulation of efficiencies and fake rates are a key ingredient for a realistic background estimate in this analysis. For the HE-LHC and the future 100 TeV collider we follow the ATLAS projections [25]. The efficiency for a tight photon identification can be well parameterized by

$$\epsilon_{\gamma\to\gamma} = 0.863 - 1.07 \cdot e^{-|p_T,\gamma|/34.8 \text{ GeV}} ,$$

(6)

and a jet-to-photon mis-identification rate by

$$\epsilon_{j\to\gamma} = \begin{cases} 5.3 \cdot 10^{-4} \exp \left(-6.5 \left( \frac{248 \text{ GeV}}{|p_T,j|} \right)^2 \right) , \\
9.88 \cdot 10^{-4} \exp \left(- \frac{|p_T,j|}{943 \text{ GeV}} \right) + \frac{248 \text{ GeV}}{|p_T,j|} , \\
\end{cases}$$

where the upper form applied to softer jets with $|p_T,j| < 65$ GeV. This leads to a photon efficiency of about 40% at $|p_T,\gamma| = 30$ GeV, saturating around 85% for $|p_T,\gamma| > 150$ GeV. Note that the Higgs decay products tend to be soft, $p_T,\gamma \sim m_h/2$.

For $b$-tagging, we adopt an efficiency with

$$\epsilon_b = 0.7 ,$$

(7)

associated with mis-tag rates of 15% for charm quarks and 0.3% for light flavors. These flat rates present a conservative estimate from the two dimensional distribution on $(p_T,j,\eta_j)$ shown in the HL-LHC projections [17]. Encouragingly, the small light flavor fake rate projections result in a strong suppression for the initially dominant $jj\gamma\gamma$ background.

Obviously, the final outcome of the analyses would depend on the detector performance for the efficiencies of photon identification and $b$-tagging, as well as the background jet rejection. To have a comprehensive exploration and comparison, we will also examine the other available detector parameters, one from CMS [26] and the other from the CERN Yellow Report [27] for the future collider (FCC), as shown in the Appendix.

### III. THE MOTHER OF DISTRIBUTIONS

As depicted in Fig. 1, Higgs pair production receives contributions from a triangular loop diagram combined with the Higgs self-coupling and from a box or continuum diagram (plus a crossing diagram), where over most of phase space the box contribution completely dominates the total rate. While we can define a number of kinematic observables describing the continuum backgrounds, the measurement of the Higgs self-coupling relies on a simple $2 \to 2$ process with two independent kinematic variables.

Three distinct phase space regions provide valuable information on a modified Higgs self-coupling, both from a large destructive interference between the triangle and box contributions. First, there is the threshold [6, 13] in the partonic center of mass energy

$$m_{hh}^{(th)} \approx 2m_h .$$

(8)

In the absence of hard additional jets, the di-Higgs invariant mass is identical to the partonic collider energy $s = m_{hh}^2$. Note that this threshold is below $2m_t$. Based on the effective Higgs-gluon Lagrangian [28] we can write the corresponding amplitude for Higgs pair production as

$$\frac{\alpha_s}{12 \pi v} \left( \frac{\kappa_\lambda \lambda_{SM}^2}{s - m_t^2} - \frac{1}{v} \right) \to \frac{\alpha_s}{12 \pi v^2} (\kappa_\lambda - 1)^{SM} .$$

(9)

While the heavy-top approximation is known to give a poor description of the signal kinematics as a whole, it does describe the threshold dependence correctly [13]. This indicates that we can search for a deviation of the Higgs self-coupling by looking for an enhancement of the rate at threshold.

Second, an enhanced sensitivity to the self-coupling appears as top mass effect. For large positive values of $\lambda$ absorptive imaginary parts lead to a significant dip in the combined rate at the threshold $p_T,h \approx 100$ GeV [10] or equivalently [17]

$$m_{hh}^{(abs)} \approx 2m_t .$$

(10)

The sharpest interference dip takes place near $\lambda \approx 2$. For negative values of $\lambda$ the interference becomes constructive.
Finally, the triangular and box amplitudes generally have different scaling in the limit [6, 10]

$$m_{hh}^{(\text{high})} \gg m_h, m_t.$$  (11)

While the triangle amplitude features an explicit suppression of either $m_h^2/m_{hh}^2$ or $m_t^2/m_{hh}^2$ at high invariant mass, the box diagrams drop more slowly towards the high-energy regime.

The impact of all three kinematic features can be quantified statistically and is illustrated in detail in Fig. 5 of Ref. [17]. They clearly indicate that essentially the full information on the Higgs self-coupling can be extracted through a shape analysis of the $m_{hh}$ distribution [29].

The practical relevance of the different kinematic regimes has to be estimated including the variation of the signal cross section, the number of expected events at a given collider, and the size of the backgrounds. There exist two similar statistical approaches to answer this problem, the MADMAX approach based on the Neyman-Pearson lemma [21] and the MADFisher approach based on information geometry [30]. While the latter is especially well-suited to estimate the reach for example of precision measurements at the LHC, we employ the former for a simple hypothesis test. The integrated log-likelihood ratio over the full phase space or specific kinematic regimes allows us to estimate the maximum significance with which any multi-variate analysis will be able to extract a signal from backgrounds or distinguish two assumed values of the Higgs self-coupling [17]. Throughout maximum likelihood analysis we limit ourselves to irreducible backgrounds and assume that statistical uncertainties dominate over the relevant phase space regions. Events with soft final states typically contribute little to the search for new particles with weak-scale masses. The exact choice of acceptance cuts in Eq. (5) and the modeling of $b$-tagging or photon identification efficiencies will have a negligible effect on our results.

For our numerical analysis, we account for all backgrounds discussed in Sec. II, except for the $tt$ channel with its significantly different final state. As part of the detailed background analysis in Sec. IV, we will see that this assumption is justified. The setup is essentially identical to Ref. [17], but now using the cuts and fake rates given in Sec. II. In particular, we account for the smearing of the Higgs peak as leading detector effect. The invariant mass distributions are smeared by a Gaussian with width 1.52 GeV for the $\gamma\gamma$ channel [31] and 12.6 GeV for the $bb$ channel [32]. The signal rate is adjusted to account for the loss of signal rate through a poor description of the tails of the distributions [17]. This allows us to restrict ourselves to the two Higgs mass windows

Figure 3. Kinematic distributions (dashed lines with left vertical axes) and significance distribution (solid lines with right vertical axes) assuming a Higgs self-coupling with $\kappa_\lambda = 0, 1, 2$. The significance describes the discrimination of an anomalous self-coupling $\kappa_\lambda \neq 1$ from the SM hypothesis $\kappa_\lambda = 1$. The results are for the HE-LHC (upper row) and for the 100 TeV collider (lower row).
$m_{bb} = 80 \ldots 160\text{ GeV}$ and $m_{\gamma\gamma} = 120 \ldots 130\text{ GeV}$. All other detector effects are left to our actual analysis in Sec. IV.

In Fig. 3 we first show the signal and background distributions for three relevant kinematic variables, $m_{hh}$, $p_{T,h}$, and $\Delta R_{\gamma\gamma}$. The transverse momentum distributions of the two Higgs bosons will be identical, so we can measure them either as $p_{T,\gamma\gamma}$ or as $p_{T,bb}$. Both, for $m_{hh}$ and $p_{T,h}$ the QCD backgrounds reside at small values, with similar signal-to-background ratios at the HE-LHC and the 100 TeV collider. The geometric separation of the two photons from the continuum background has to be large to generate an invariant mass around the Higgs mass.

Also in Fig. 3, we show how the significance of extracting an anomalous self-coupling $\kappa_\lambda \neq 1$ depends on these key observables. The alternative hypothesis in this case is the combination of the backgrounds and the signal with $\kappa_\lambda = 1$. In addition to the signal features, the significance is limited by the rapidly dropping backgrounds, covering both of the above-mentioned regions with an enhanced dependence on the triangle diagram. In the absence of background, the significance indeed peaks between the production threshold and the top-mass threshold [17]. The drop towards large values of $m_{hh}$ is a combination of the dominance of the box diagram in the signal and the limited number of expected signal events. The significance with which we can extract modified self-couplings either smaller ($\kappa_\lambda = 0$) or larger ($\kappa_\lambda = 2$) than in the Standard Model shows a similar phase space dependence. The only difference is a slightly harder significance distributions for $\kappa_\lambda = 2$, an effect of the dip at $m_{hh}$.

Obviously, we can combine the maximum significance distributions into a global maximum significance accumulated over the full phase space. In Fig. 4 we show the idealized, maximum significance with which we can hope for at the HL-LHC, the HE-LHC, and a future 100 TeV collider. The asymmetric behavior for the HL-LHC is a remainder of a degeneracy in the total cross section as a function of the self-coupling, also shown in Fig. 4. A SM-like rate appears when an enhanced triangle diagram overcomes the larger box contribution and flips the sign of the amplitude. Obviously, this degeneracy will be broken by kinematic information, for example the $m_{bb}$ distribution. For the HE-LHC and the 100 TeV collider, the total rate constraint becomes increasingly irrelevant for the measurement of the self-coupling. The expected statistical error bars are narrow and approximately symmetric around on $\kappa_\lambda = 1$. For both future colliders, we can indeed expect a proper measurement of the Higgs self-coupling.

IV. DETECTOR-LEVEL ANALYSIS

Following the analysis path laid out in Sec. III, we now design a detailed analysis strategy to extract the Higgs self-coupling with a focus on the shape of the $m_{hh}$ distribution. Our signal is

$$pp \rightarrow hh + X \rightarrow b\bar{b}\gamma\gamma + X. \quad (12)$$

In anticipation of increasing QCD radiation at higher energies, we inclusively allow extra jets in the events from initial state radiation, along with two tagged $b$-jets and two isolated hard photons, passing the acceptance cuts of Eq. (5). For the detector-level analysis we generate the signal and background samples with MadGraph5+Pythia8 [24, 33], including one extra jet using the MLM scheme [34]. A representative set of Feynman diagrams for the signal is shown in Figs. 1 and 5. Higher-order corrections are included through a next-to-leading order K-factor 1.6 [22, 23, 35], neglecting possible higher-order effects on the $m_{hh}$ distribution. We normalize the $tth$ and $Zh$ to their respective NLO and NNLO rates 2.8 pb and 2.2 pb at 27 TeV (37 pb and 11 pb at 100 TeV) [36]. We also include the full set of detector effects with Delphes3 [37], following the HL-LHC projections [25].

Jets are defined with the anti-$k_T$ algorithm $R = 0.4$.
Table I. Number of signal and background events for the HE-LHC and the 100 TeV collider. We present results for the HE-LHC and the 100 TeV future collider, respectively, with arbitrary units.

| Collider | Process                  | $\kappa_\lambda$ | 1  | 2  | $t\bar{t}h$ | $Zh$ | $b\bar{b}\gamma\gamma$ | $jj\gamma\gamma$ | $bb\gamma_j$ | $BG$ tot. | $S/\sqrt{S + B}_{\text{lab}^{-1}}$ | $S/B$ |
|----------|--------------------------|------------------|----|----|--------------|-----|---------------------|------------------|--------------|----------|-------------------------------|------|
|          | $\sigma$ [fb]            | 0.69             | 0.36 | 0.18 | 6.43         | 0.77 | 1.24 | 36.6 | 506 | 38.3M | 0.07 | 4 \cdot 10^{-5} |
|          | Baseline                 | 2.87K           | 1.57K | 838 | 21.8K | 1.44K | 1.19M | 36M | 1.13M | 38.3M | 0.15 | 1 \cdot 10^{-3} |
|          | $n_j \leq 3$, $n_b = 2$ | 648             | 356 | 190 | 954 | 389 | 200K | 67.4K | 105K | 374K | 0.24 | 0.003 |
|          | $\Delta m_{bb} \leq 25$ GeV | 470 | 260 | 140 | 195 | 66 | 43.7K | 10.6K | 25.8K | 80.4K | 0.24 | 0.003 |
|          | $\Delta m_{\gamma\gamma} \leq 3$ GeV | 459 | 253 | 136 | 197 | 63 | 1.42K | 505 | 758 | 2.94K | 1.2 | 0.09 |
|          | $\Delta m_{\gamma\gamma} \leq 2$ GeV | 459 | 253 | 136 | 197 | 63 | 957 | 342 | 504 | 2.06K | 1.4 | 0.12 |
|          | $\Delta m_{\gamma\gamma} \leq 1$ GeV | 459 | 253 | 136 | 197 | 63 | 485 | 182 | 245 | 1.17K | 0.7 | 0.22 |
|          | $\Delta m_{\gamma\gamma} \leq 3$, $m_{h\ell} > 400$ | 320 | 206 | 120 | 56 | 21 | 324 | 97 | 178 | 676 | 1.8 | 0.30 |
|          | $\Delta m_{\gamma\gamma} \leq 2$, $m_{h\ell} > 400$ | 320 | 206 | 120 | 56 | 21 | 220 | 67 | 122 | 485 | 2.0 | 0.42 |
|          | $\Delta m_{\gamma\gamma} \leq 1$, $m_{h\ell} > 400$ | 320 | 206 | 120 | 56 | 21 | 115 | 41 | 61 | 293 | 2.4 | 0.70 |

|          | $\sigma$ [fb]            | 6.95             | 3.72 | 1.97 | 84.8 | 3.76 | 6.21 | pb | 126 | 5.03 nb | 0.29 | 8 \cdot 10^{-5} |
|          | Baseline                 | 51.8K           | 29.8K | 16.9K | 535K | 13.1K | 13.6M | 330M | 1.86M | 363M | 0.48 | 0.001 |
|          | $n_j \leq 3$, $n_b = 2$ | 9.22K           | 5.28K | 3.02K | 18K | 2.84K | 1.79M | 773K | 1.42M | 4.00M | 0.71 | 0.004 |
|          | $\Delta m_{bb} \leq 25$ GeV | 6.45K | 3.80K | 2.18K | 3.3K | 669 | 361K | 218K | 373K | 956K | 0.71 | 0.004 |
|          | $\Delta m_{\gamma\gamma} \leq 3$ GeV | 6.30K | 3.70K | 2.13K | 3.12K | 653 | 8.34K | 6.06K | 8.99K | 27.2K | 3.9 | 0.14 |
|          | $\Delta m_{\gamma\gamma} \leq 2$ GeV | 6.30K | 3.70K | 2.13K | 3.12K | 653 | 5.66K | 4.13K | 5.99K | 19.5K | 4.4 | 0.19 |
|          | $\Delta m_{\gamma\gamma} \leq 1$ GeV | 6.30K | 3.70K | 2.13K | 3.12K | 653 | 2.82K | 1.91K | 2.99K | 11.4K | 5.5 | 0.32 |
|          | $\Delta m_{\gamma\gamma} \leq 3$, $m_{h\ell} > 400$ | 4.66K | 3.16K | 1.93K | 1.09K | 203 | 1.56K | 1.10K | 1.90K | 5.86K | 6.1 | 0.54 |
|          | $\Delta m_{\gamma\gamma} \leq 2$, $m_{h\ell} > 400$ | 4.66K | 3.16K | 1.93K | 1.09K | 203 | 1.04K | 0.74K | 1.14K | 4.23K | 6.7 | 0.73 |
|          | $\Delta m_{\gamma\gamma} \leq 1$, $m_{h\ell} > 400$ | 4.66K | 3.16K | 1.93K | 1.09K | 203 | 523 | 359 | 617 | 2.79K | 7.5 | 1.13 |

via FastJet [38]. While the $t\bar{t}h$ background is almost irrelevant at the 14 TeV LHC, it becomes increasingly important at higher energies. Obviously, the more complex, high-multiplicity final state offers many handles to tame it. We employ a simple veto on leptons with

$$p_T,\ell > 10 \text{ GeV and } |\eta| < 2.5,$$

combined with a veto of more than three jets passing Eq. (5).

To suppress the initially overwhelming $jj\gamma\gamma$ background, we demand two $b$-tags among the three hardest jets. A crucial observation is that at higher energies, initial state radiation (ISR) often leads to a harder jet than the Higgs decay products, such that either the hardest or second-hardest jet is not a $b$-jet for roughly half of all events. This is illustrated in Fig. 6 as the composition of the second-hardest parton-level jet, requiring that both truth-level $b$-jets pass the selection of Eq. (5). Thus, the $b$-tagging requirement as the two leading jets should be adjusted accordingly.

Based on this observation we account for two patterns of the $p_T$ jets, ($bb, bbj$) and ($jbj, hjb$). This increases our signal efficiency by around 50%. Expanding this scheme to even more jets is not effective because it eventually also increases the continuum backgrounds and the $t\bar{t}h$ contributions. The reliability of our Monte Carlo simula-
tion underlying this procedure is guaranteed by the fact that the hardest three jets are generated using multi-jet merging.

To control the continuum backgrounds, we require two Higgs mass windows,

$$|m_{bb} - m_{h}| < 25 \text{ GeV}, \quad |m_{\gamma\gamma} - m_{h}| < 1 \text{ GeV}. \quad (14)$$

An obvious way to enhance the Higgs pair signal is to improve the resolution on the reconstructed photons and b-jets from the Higgs decays. We adopt the rather conservative resolution for $m_{bb}$ as in Eq. (14). Any improvement on it in experiments would be greatly helpful for the signal identification and background separation. As for the photon resolution, we illustrate this effect by using three representative values where the $m_{\gamma\gamma}$ distribution is smeared by a Gaussian width of 0.75, 1.5, 2.25 GeV, corresponding to Higgs mass windows

$$|m_{\gamma\gamma} - m_{h}| \leq 1, 2, 3 \text{ GeV}. \quad (15)$$

A resolution of 1.5 GeV has already been achieved at the LHC [31].

The results at this stage of the analysis are illustrated in Table I with a full cut flow for the two collider energies and assuming $\kappa_{\lambda} = 0, 1, 2$. We already find a large background suppression $S/B \sim 0.09 \ldots 0.2$ for the HE-LHC and 0.14 ... 0.3 at a future 100 TeV collider. Requiring $m_{hh} > 400$ GeV improves it to $S/B \sim 0.3 \ldots 0.7$ or 0.5 ... 1.1, respectively. This is entirely due to the rapidly falling backgrounds as compared to the $hh$ signal, but will be at the expense of the self-coupling determination. The $m_{hh}$ distribution of the signal and the different backgrounds is shown in Fig. 7.

The signal-to-background ratio can be strongly improved by a better $m_{\gamma\gamma}$ resolution. As long as most of the $h \rightarrow \gamma\gamma$ events are captured by an appropriate $m_{\gamma\gamma}$ window, the contributions from continuum backgrounds can be estimated using the side-band measurements.

Going beyond a cut-based analysis for example on $m_{hh}$, we employ a binned log-likelihood analysis based on the CL$_s$ method, using the full $m_{hh}$ distribution to extract $\kappa_{\lambda}$ [39]. The dominant backgrounds feature powerful control regions or ratio measurements like $t\bar{t}h/ttZ$ [40]. Therefore, we neglect their systematic uncertainties. As a starting point, we show the $5\sigma$ determination on the Higgs pair signal strength in the left panel of Fig. 8, requiring two b-tagged jets among the two or three leading jets. We decompose the latter case in two sub-samples ($bb, bbj$) and ($jbb, jbj$). We see how exploring the extra-jet emission significantly improves the significance as compared to the standard procedure adopted in the literature. The $5\sigma$ measurement for HE-LHC is pushed from 2.8 ab$^{-1}$ to below 2.3 ab$^{-1}$.

In the right panel of Fig. 8 we show the discovery reach for the Higgs pair signal as a function of the luminosity of the HE-LHC and the 100 TeV collider. We assume three di-photon invariant mass resolutions with three Higgs mass windows as in Eq. (15) for a SM self-coupling $\kappa_{\lambda} = 1$. Higgs pair production will be discovered at the HE-LHC with approximately 2.5 ... 5 ab$^{-1}$ and at the 100 TeV collider with 0.2 ... 0.3 ab$^{-1}$ of data, in both cases well below the design luminosity.

As commented in the Introduction, there exist physics scenarios that the Higgs self-coupling could be modified at the level of order one deviation from the SM value. The accurate measurement of the Higgs self-coupling via Higgs pair production at future colliders has the best promise to uncover the new physics associated with the Higgs sector. In Fig. 9, we show the accuracy on this measurement. At the 68% confidence level the triple Higgs coupling can be measured with the precision

$$\kappa_{\lambda} \approx 1 \pm 15\% \quad \text{(HE-LHC, 27 TeV, 15 ab$^{-1}$),}$$

$$\kappa_{\lambda} \approx 1 \pm 5\% \quad \text{(100 TeV, 30 ab$^{-1}$)}. \quad (16)$$

At the 95% confidence level,

$$\kappa_{\lambda} \approx 1 \pm 30\% \quad \text{(HE-LHC, 27 TeV, 15 ab$^{-1}$),}$$

$$\kappa_{\lambda} \approx 1 \pm 10\% \quad \text{(100 TeV, 30 ab$^{-1}$)}. \quad (17)$$

The way to improve these expected limits towards the mathematically-defined best reach shown in Fig. 4 is to

![Figure 7. Higgs pair invariant mass for the signal and backgrounds based on realistic simulations for the HE-LHC (left) and the 100 TeV future collider (right). The $m_{\gamma\gamma}$ distribution is described by a Gaussian with width 0.75 GeV.](image-url)
to illustrate the staged improvements and to give a

exploit more kinematic features and this way also suppress the reducible $tth$ background.

To gain some insight on how robust our results are, we have also examined the other available choices of detector parameters, one from CMS [26] and the other from the CERN Yellow Report (YR) [27] for the future collider (FCC). As shown in Fig. 10 in the Appendix, we find that the results are quite consistent with each other, with the YR performance being slightly better. This indicates possible room for further improvement.

V. SUMMARY AND OUTLOOK

In this paper, we have explored Higgs pair production as a direct way to measure the Higgs self-coupling, the least-known but arguably the most important fundamental parameter of the Standard Model.

We first presented the production cross section for $pp \rightarrow hh$ at future high-energy colliders in Fig. 2, Sec. II. We discussed the signal rate for the process with leading sensitivity $pp \rightarrow hh \rightarrow bb\gamma\gamma$, and laid out the event selection criteria in accordance with the experimental acceptance at the LHC.

In Sec. III, we discussed the kinematic features of the signal and compared with the backgrounds, as shown in Fig. 3. The key variable is the invariant mass distribution of the Higgs pair that presented distinctive behaviors. We first performed a parton-level analysis that combines the maximum significance distributions into a global maximum significance accumulated over the full phase space, for the HL-LHC, the HE-LHC, and a future 100 TeV collider. For both future colliders we found excellent prospects for kinematics-based determinations of the Higgs self-coupling as shown in Fig. 4.

In Sec. IV, we then carried out a search strategy based on a rate combined with kinematic shapes with realistic simulations. The approach is not only more powerful [13, 17] than a purely rate-based measurement but also more stable against systematic and theoretical uncertainties, provided we account for all bin-to-bin correlations. Our method removes all degeneracies which appear in a rate-based measurement and leads to well-defined symmetric error bars on the modified self-coupling.

Higher energy colliders allow for including events with high $m_{hh}$. In such more and more common configurations at high energies, the additional jets from QCD radiation frequently surpass the $b$-jet energy about $m_h/2$, as seen in Fig. 6. To improve the signal efficiency we included at least three observable jets, fully accounting for QCD jet radiation via the MLM merging, with possibly softer $b$-jets from Higgs decays. We showed a cut-flow in Table I to illustrate the staged improvements and to give a

Figure 8. Luminosity required for a 5σ discover of Higgs pair production for the HE-LHC (dashed) and a 100 TeV collider (full). Left: sensitivity in terms of the total rate, demanding two $b$-tags among the two or three leading jets and assuming $|m_{\gamma\gamma} - m_h| < 1$ GeV. Right: sensitivity for three mass windows $|m_{\gamma\gamma} - m_h| < 1, 2, 3$ GeV. We assume the SM hypothesis with $\kappa_\lambda = 1$ and use a binned log-likelihood analysis of the $m_{hh}$ distribution.

Figure 9. Confidence level for separating an anomalous Higgs self-coupling hypothesis from the Standard Model $\kappa_\lambda = 1$. 

1.0

0.1

0.01

0.001

0.0

0.5

1.0

1.5

2.0

$\lambda$

CLs

27 TeV

100 TeV

1σ

2σ

$\kappa_\lambda$
comparison for the two future colliders. We further enhance our measured significances, decomposing the samples into two sub-samples \((bb, bby)\) and \((jbb, bjbj)\).

Finally, we determined the integrated luminosity needed to reach a 5σ significance to observe the SM \(hh\) signal as shown in Fig. 8. We found that the high-energy upgrade of the LHC to 27 TeV would reach a 5σ observation of the Higgs pair production with an integrated luminosity of about 2.5 ab\(^{-1}\). It would have the potential to reach 15% (30%) accuracy at the 68% (95%) confidence level to determine the SM Higgs boson self-coupling. A future 100 TeV collider could improve the self-coupling measurement to better than 5% (10%) at the 68% (95%) confidence level, as shown in Fig. 9. These results roughly agree with the optimal reach shown in Fig. 4. Our conclusions are quite robust against some moderate variations of the detector performances as shown in Fig. 10 in the Appendix. In the hope of searching for effects from physics beyond the SM, our results should provide conclusive information weather or not the Higgs-self-interaction is modified to a level of order one.

While our conclusions on the determination of Higgs-self-interaction at future hadron colliders are robust and important, there is still room for improvement. Although the final state \(bb\gamma\gamma\) is believed to be the most sensitive channel because of the background suppression and signal reconstruction, there exist complementary channels such as \(gg \rightarrow hh \rightarrow bb\tau^+\tau^-\), \(bbW^+W^-\), \(bb\bar{b}\bar{b}\), etc. The kinematics-based measurement and the all features related to QCD radiation at higher energies should be equally applicable to all of them.

VI. APPENDIX

As explained in the text, we optimize our set of selection cuts primarily to reduce the continuum background, which would be accompanied by large systematic uncertainty, and secondarily to reduce the \(tt\bar{t}h\) background, which is the largest background component with a Higgs mass peak structure. To achieve the above optimization, we take the photon identification working point with a reasonably efficient jet-fake rejection [25], and require the additional jet veto \((n_{jj} \leq 3)\).

We believe our selection is almost optimal, but for completeness, we assess the effects of applying different efficiencies taken in the literature and provide the final sensitivities assuming those numbers. For comparison, we have worked on two different efficiency scenarios found for the CMS projections [26] and in the CERN Yellow Report (YR) [27] for the study of Future Circular Colliders (FCC).

We adopt the fitted CMS projections as follows:

\[
\begin{align*}
\epsilon_{b\rightarrow b} &= 0.85 \tanh\left(\frac{p_T}{400 \text{ GeV}}\right) \left[\frac{25.0}{1 + p_T/15.9 \text{ GeV}}\right], \\
\epsilon_{c\rightarrow b} &= 0.25 \tanh\left(\frac{p_T}{55.6 \text{ GeV}}\right) \left[\frac{1}{1 + p_T/769 \text{ GeV}}\right], \\
\epsilon_{j\rightarrow b} &= 0.01.
\end{align*}
\]

The efficiency set used in the YR is the following:

\[
\begin{align*}
\epsilon_{\gamma\rightarrow\gamma} &= 0.9, \\
\epsilon_{j\rightarrow\gamma} &= 0.01 \exp\left(-\frac{p_T}{30 \text{ GeV}}\right), \\
\epsilon_{b\rightarrow b} &= 0.75, \\
\epsilon_{c\rightarrow b} &= 0.1, \\
\epsilon_{j\rightarrow b} &= 0.01.
\end{align*}
\]

Fig. 10 shows the comparison among the final results using the three different sets of the efficiencies for 27 TeV (top) and 100 TeV (bottom). The red lines show the final results assuming our adopted efficiencies (from the ATLAS HL-LHC projection study) [25], while the green and the blue lines show those assuming the YR and the CMS ones, respectively. Our analysis sensitivity is not much improved by taking the working points with a larger photon efficiency used by these two alternative references, due to the corresponding worse light-jet rejection rate, which enhances the continuum background, especially the \(bb\tau\bar{\tau}\) contribution.

![Figure 10](attachment:image.png)
Note that we devise our analysis with a large $S/B$ by targeting to reduce the continuum background $b\bar{b}g_{j}$ and leave the main background contributions from $t\bar{t}h$. In this way we achieve $S/B \sim 0.7$ against the corresponding numbers 0.45 (YR), and 0.4 (CMS), respectively, for the 27 TeV analysis. For the 100 TeV analysis, we achieve $S/B \sim 1.1$ against 0.6 (YR) and 0.5 (CMS). Thus, we can provide a more robust estimate against the systematic uncertainty of the continuum background. Additionally, it allows us to have a larger sensitivity from the lower $m_{ph}$ profile, a regime that is more background contaminated and that displays larger effects on the triple Higgs coupling.

### Acknowledgment

[1] P. W. Higgs, Phys. Lett. 12, 132 (1964); P. W. Higgs, Phys. Rev. Lett. 13, 508 (1964); F. Englert and R. Brout, Phys. Rev. Lett. 13, 321 (1964).

[2] G. Aad et al. [ATLAS Collaboration], Phys. Lett. B 716, 1 (2012) [arXiv:1207.7214 [hep-ex]]; S. Chatrchyan et al. [CMS Collaboration], Phys. Lett. B 716, 30 (2012) [arXiv:1207.7235 [hep-ex]].

[3] X. m. Zhang, Phys. Rev. D 47, 3065 (1993) [arXiv:hep-ph/9301277]; X. Zhang, B. L. Young and S. K. Lee, Phys. Rev. D 51, 5327 (1995) [arXiv:hep-ph/9406322]; C. Grojean, G. Servant and J. D. Wells, Phys. Rev. D 71, 036001 (2005) [arXiv:hep-ph/0407019]; S. Kanemura, Y. Okada and E. Senaha, Phys. Lett. B 606, 361 (2005) [arXiv:hep-ph/0411354]; A. Noble and M. Perelstein, Phys. Rev. D 78, 063518 (2008) [arXiv:0711.3018 [hep-ph]]; K. Fuyuto and E. Senaha, Phys. Rev. D 90, no. 1, 015015 (2014) [arXiv:1406.0433 [hep-ph]]; F. P. Huang, P. H. Gu, P. F. Yin, Z. H. Yu and X. Zhang, Phys. Rev. D 93, no. 10, 103515 (2016) [arXiv:1511.03909 [hep-ph]]; P. Huang, A. Joglekar, B. Li and C. E. M. Wagner, Phys. Rev. D 93, no. 5, 055049 (2016) [arXiv:1512.00068 [hep-ph]]; A. Koba, T. Suzuki, L. Wu and J. Yue, JHEP 1604, 011 (2016) [arXiv:1512.08922 [hep-ph]]; C. Y. Chen, J. Kozaczuk and I. M. Lewis, JHEP 1708, 096 (2017) [arXiv:1704.05844 [hep-ph]]; X. Gan, A. J. Long and L. T. Wang, arXiv:1708.03061 [hep-ph]; Q. H. Cao, F. P. Huang, K. P. Xie and X. Zhang, arXiv:1708.04737 [hep-ph]; B. Jain, S. J. Lee and M. Son, arXiv:1709.03232 [hep-ph]; J. de Vries, M. Postma, J. van de Vis and G. White, arXiv:1710.04661 [hep-ph]; M. Reichert, A. Eichhorn, H. Gies, J. M. Pawlowski, T. Plehn and M. M. Scherer, arXiv:1711.00019 [hep-ph]; M. Carena, Z. Liu and M. Riembau, arXiv:1801.00794 [hep-ph].

[4] N. Arkani-Hamed, T. Han, M. Mangano and L. T. Wang, Phys. Rept. 652, 1 (2016) [arXiv:1511.06495 [hep-ph]].

[5] O. J. P. Eboli, G. C. Marques, S. F. Novaes and A. A. Natale, Phys. Lett. B 197, 269 (1987); D. A. Dicus, C. Kao and S. S. D. Willenbrock, Phys. Lett. B 203, 457 (1988); E. W. N. Glover and J. J. van der Bij, Nucl. Phys. B 309, 282 (1988).

[6] T. Plehn, M. Spira and P. M. Zerwas, Nucl. Phys. B 479, 46 (1996) Erratum: [Nucl. Phys. B 531, 655 (1998)] [arXiv:hep-ph/9603205]; A. Djoudi, W. Kilian, M. Muhlleitner and P. M. Zerwas, Eur. Phys. J. C 10, 45 (1999) [arXiv:hep-ph/9904287]; X. Li and M. B. Voloshin, Phys. Rev. D 89, no. 1, 013012 (2014) [arXiv:1311.5156 [hep-ph]].

[7] U. Baur, T. Plehn and D. L. Rainwater, Phys. Rev. D 69 (2004) 053004 [arXiv:hep-ph/0310056].

[8] V. Barger, L. L. Everett, C. B. Jackson, A. D. Peterson and G. Shaughnessy, Phys. Rev. D 90, no. 9, 095006 (2014) [arXiv:1408.2525 [hep-ph]]; A. Alves, T. Ghosh and K. Sinha, Phys. Rev. D 96, no. 3, 035022 (2017) [arXiv:1704.07395 [hep-ph]]; A. Adhikary, S. Banerjee, R. K. Barman, B. Bhattacharjee and S. Niyogi, arXiv:1712.05346 [hep-ph]; J. H. Kim, Y. Sakaki and M. Son, arXiv:1801.06093 [hep-ph].

[9] U. Baur, T. Plehn and D. L. Rainwater, Phys. Rev. D 68, 033001 (2003) [arXiv:hep-ph/0304015].

[10] M. J. Dolan, C. Englert and M. Spannowsky, JHEP 1210, 112 (2012) [arXiv:1206.5001 [hep-ph]]; A. J. Barr, M. J. Dolan, C. Englert and M. Spannowsky, Phys. Lett. B 728, 308 (2014) [arXiv:1309.6318 [hep-ph]].

[11] A. Papaefstathiou, L. L. Yang and J. Zurita, Phys. Rev. D 87, no. 1, 013001 (2013) [arXiv:1209.1489 [hep-ph]].

[12] D. E. Ferreira de Lima, A. Papaefstathiou and M. Spannowsky, JHEP 1408, 030 (2014) [arXiv:1404.7139 [hep-ph]]; D. Wardrope, E. Jensen, N. Konstantinidis, B. Cooper, R. Falla and N. Norjoharuddeen, Eur. Phys. J. C 75, no. 5, 219 (2015) [arXiv:1410.2794 [hep-ph]]; J. K. Behr, D. Bortoletto, J. A. Frost, N. P. Hartland, C. Issever and J. Rojo, Eur. Phys. J. C 80, no. 5, 348 (2020) [arXiv:1809.02638 [hep-ph]]; A. Papaefstathiou, L. L. Yang and J. Zurita, Phys. Rev. D 87, no. 1, 013001 (2013) [arXiv:1209.1489 [hep-ph]].
