PHYSICAL REVIEW D **95,** 035026 (2017)

Maximizing the significance in Higgs boson pair analyses

Felix Kling,[1,2] Tilman Plehn,[3] and Peter Schichtel[4]
[1]*Department of Physics, University of Arizona, 1118 E. Fourth Street, Tucson, AZ 85721, USA*
[2]*Theoretical Physics Department, Fermilab, Batavia, Illinois 60510, USA*
[3]*Institut für Theoretische Physik, Universität Heidelberg,*
Philosophenweg 16, 69120 Heidelberg, Germany
[4]*Institute for Particle Physics Phenomenology, Durham University, Durham DH1 3LF, United Kingdom*
(Received 28 September 2016; published 22 February 2017)

We study Higgs pair production with a subsequent decay to a pair of photons and a pair of bottoms at the LHC. We use the log-likelihood ratio to identify the kinematic regions which either allow us to separate the di-Higgs signal from backgrounds or to determine the Higgs self-coupling. We find that both regions are separate enough to ensure that details of the background modeling will not affect the determination of the self-coupling. Assuming dominant statistical uncertainties we determine the best precision with which the Higgs self-coupling can be probed in this channel. We finally comment on the same questions at a future 100 TeV collider.

DOI: 10.1103/PhysRevD.95.035026

I. INTRODUCTION

Higgs pair production is one of the key benchmarks for any new collider probing the electroweak scale. This includes the LHC towards large luminosity, but also future electron-positron colliders and future hadron colliders. At hadron colliders the task is clear: we can test Higgs pair production in gluon fusion, with continuum contributions as well as contributions induced by the Higgs self-coupling. The two Feynman diagrams are shown in Fig. 1 [1,2]. In the Standard Model both diagrams rely on a strong interaction between the Higgs and the top quark in the loop to give an observable rate for the LHC. This means that if we want to measure the di-Higgs rate or even the Higgs self-coupling [3–5] from the total rate $\sigma_{gg\to HH}$ we need to make an assumption about the top Yukawa coupling [6]. For a combined Higgs fit we could for example assume that the effective gluon-Higgs coupling is only mediated by the Standard Model quarks, which to date gives a roughly 10% measurement of the top Yukawa [7]. A model independent precision measurement of the top Yukawa coupling at the percent level will only be possible at a 100 TeV collider [8]. Firmly connecting a possible modification of the Higgs self-coupling to a modified top Yukawa in a given model is rather hopeless, as can easily be seen in two-Higgs doublet models. One way to limit the impact of such an assumption would be to include Higgs pair production as a probe of the Higgs self-coupling in a global analysis of the Higgs effective Lagrangian at tree level or at one loop [9,10]. Another way to at least minimize the assumption about the top Yukawa is to test kinematic distributions in Higgs pair production. There are three obvious questions concerning such an analysis of the Higgs pair kinematics which we tackle in this paper:

1. Which kinematic features allow us to extract Higgs pair production from backgrounds?
2. Which kinematic features include information about the Higgs self-coupling?
3. What is the most optimistic LHC sensitivity in the presence of QCD backgrounds?

There are two kinematic regimes which are well known to carry information on the Higgs self-coupling. Both of them exploit the (largely) destructive interference between the two graphs shown in Fig. 1. As we will illustrate later, over most of phase space the continuum contribution dominates. One phase space region where the triangle diagram become comparable is close to threshold [2,3]; if we denote the Higgs effective Higgs-gluon Lagrangian in terms of the gluon field strength $G_{\mu\nu}$ as [11]

$$\begin{aligned}\mathcal{L}_{ggH} &= \frac{\alpha_s}{12\pi} G^{\mu\nu} G_{\mu\nu} \log\left(1+\frac{H}{v}\right) \\ &= G^{\mu\nu} G_{\mu\nu} \frac{\alpha_s}{12\pi v}\left(H - \frac{H^2}{2v} + \cdots\right), \qquad (1)\end{aligned}$$

we can write the amplitude for Higgs pair production as

$$\begin{aligned}&\mathcal{A} \propto \frac{\alpha_s}{12\pi v}\left(\frac{\lambda}{s-m_H^2} - \frac{1}{v}\right) \to \frac{\alpha_s}{12\pi v^2}\left(\frac{3m_H^2}{3m_H^2} - 1\right) = 0 \\ &\quad \text{for } m_{HH} \to 2m_H. \qquad (2)\end{aligned}$$

The exact cancellation is linked to the Standard Model value of the physical Higgs self-coupling $\lambda = 3m_H^2/v$. Note that while the heavy top approximation is well known for giving completely wrong kinematic distributions for Higgs pair production [3], it correctly predicts the threshold region. If we rely on this distribution, the strategy behind an LHC analysis will be to rule out large deviations from the Standard Model

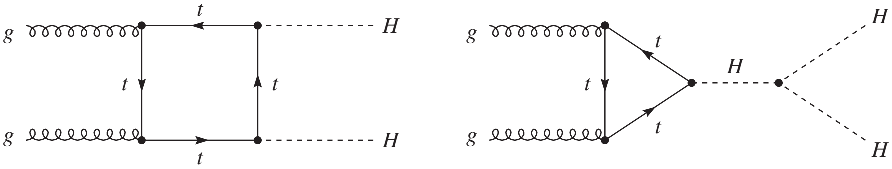

FIG. 1. Feynman diagrams contributing to Higgs pair production at the LHC. Figure from Ref. [3].

Higgs self-coupling based on the fact that any such deviation leads to a strongly enhanced production cross section.

The second sensitive kinematic regime where the two contributions shown in Fig. 1 become comparable is boosted Higgs pair production [12]. Top threshold corrections to the triangle diagram are strongly enhanced around $m_{HH} = 2m_t$. We can translate this into a condition for the transverse momentum of the SM-like Higgs, where we find that around $p_{T,H} \sim 100$ GeV the combined process develops a minimum for large Higgs self-couplings. The fact that there exist two relatively uncorrelated useful kinematic distributions is of course not surprising for an effectively 2-body final state.

At the LHC, we can go through all Higgs decay channels and test their combinations as possible di-Higgs signatures. The most promising channel for a SM-like pair of Higgs bosons is most likely the $b\bar{b}\gamma\gamma$ final state [5,13–15]. Its great advantage is that we can easily reconstruct one of the two Higgs bosons and that the QCD continuum background can be measured in control regions. In addition, it should be possible to use the $b\bar{b}\tau\tau$ final state [4,12], if tau tagging will show a sufficient performance. The combination $b\bar{b}WW$ [16] will only work if we can suppress the $t\bar{t}$ background, while the $4b$ [4,17] and the original $4W$ [3,18] signatures are unlikely to contribute significantly for a pair of SM-like Higgs bosons. Finally, the $b\bar{b}\mu\mu$ shares many beneficial features with the $b\bar{b}\gamma\gamma$ channel [5], but will be further suppressed by the muon branching ratio. The picture changes if we consider either resonant Higgs pair production [5,19] or strongly interacting Higgs pair production [20]. For a pair of SM-like Higgs bosons there exists a large number of theoretical precision calculations [6], including NLO [21] and NNLO [22] predictions for the differential rates. Those will be crucial if we want to study this production process in spite of the top-Yukawa-infected total rate prediction.

In this paper we first generalize MadMax from computing maximum signal significances, globally or distributed over phase space, to comparing two general hypotheses in Sec. I A and I B. In Sec. II we then ask the question which kinematic information should allow us to extract Higgs pair production from the QCD backgrounds. We also recapitulate how some of the key features arise from the combination of the triangle and box diagrams shown in Fig. 1. Next, we test which kinematic distributions allow us to measure the Higgs self-coupling at the LHC in Sec. III. In Appendix A we provide the corresponding information for a future 100 TeV collider.

A. MadMax

The MadMax [23,24] approach of calculating significance distributions for kinematic observables is based on the Neyman–Pearson lemma: the likelihood ratio is the most powerful test statistic for a hypothesis test between a simple null hypothesis—for example background only—and an alternate hypothesis—for example signal plus background [25]. Maximum power is formally defined as the minimum probability for false negative error for a given probability of false positive.

We have established that we can define and compute the maximum significance of a signal-plus-background process as compared to the background-only hypothesis using the standard Monte Carlo tools. As an example we studied Higgs decays to muons in weak boson fusion [23]. Our results can be taken as a benchmark for the performance of multivariate analysis techniques at the LHC, including the matrix element method. In a second step we used the same method to determine which phase space regions contribute to this maximum significance, for example in boosted Higgs production in the ZH or $t\bar{t}H$ channels [24]. Such a study allows us to determine how much of the distinguishing power of a multivariate analysis comes from phase space regions which are systematically and theoretically under control. In this paper we extend this approach to test two signal hypotheses, with the technical complication that over phase space the expected ratio of events can lie on either side of unity.

In general, the likelihood of observing n events assuming a hypothesis H_0 is given by the Poisson distribution $\mathrm{Pois}(n|n_0) = e^{-n_0} n_0^n / n!$. We can generalize this counting experiment by introducing an observable x, where we assume that H_0 is described by the normalized distribution $f_0(x)$. Similarly, the alternative hypothesis H_1 is described by $f_1(x)$. Each likelihood can be factorized into the Poisson likelihood to observe an event and the normalized event likelihood $f_{0,1}(x)$. In combination they give the single-event log-likelihood ratio (LLR)

$$\begin{aligned} q_{n=1}(x) &= \log \frac{L(x|H_1)}{L(x|H_0)} = \log \frac{\mathrm{Pois}(1|n_1) f_1(x)}{\mathrm{Pois}(1|n_0) f_0(x)} \\ &= (n_0 - n_1) + \log \frac{n_1}{n_0} + \log \frac{f_1(x)}{f_0(x)}. \end{aligned} \tag{3}$$

If we know $f_0(x)$ and $f_1(x)$ from Monte Carlo simulations, we can compute the above LLR. A very simple structure containing only the last term appears, when we only rely on kinematic distributions and not on the total rate. We can generalize Eq. (3) to the likelihood of observing n events in a phase space configuration $\vec{x} = \{x_j\}$. The normalized event likelihood is now a product of n likelihoods at the corresponding configurations, so we find

$$\begin{aligned} q_n(\vec{x}) &= \log\frac{L(\vec{x}|H_1)}{L(\vec{x}|H_0)} = \log\frac{\text{Pois}(n|n_1)\prod_{j=1}^n f_1(x_j)}{\text{Pois}(n|n_0)\prod_{j=1}^n f_0(x_j)} \\ &= (n_0 - n_1) + \sum_{j=1}^n \log\frac{n_1 f_1(x_j)}{n_0 f_0(x_j)}. \end{aligned} \tag{4}$$

The combined LLR is additive when we include more than one event. If the argument of the logarithm is allowed to be greater as well as smaller as unity, depending on the position in phase space, their contribution to the overall LLR may cancel.

If we want to use the LLR to distinguish two hypotheses we need to evaluate our events as a function of the LLR, given either H_0 or H_1. Assuming for example the hypothesis H_0 we can integrate over the entire phase space $\vec{x}$ with the normalized event weight $d\sigma_0(\vec{x})/\sigma_{0,\text{tot}}$ and generate a LLR distribution based on the relation $q_n(\vec{x})$ given in Eq. (3),

$$\rho_{0,n=1}(q) = \int dx f_0(x)\delta(q_{n=1}(x) - q). \tag{5}$$

The corresponding likelihood distributions for n events and combined for all possible outcomes n are given by a convolution in q-space

$$\begin{aligned} \rho_{0,n}(q) &= \rho_{0,n=1} \otimes \rho_{0,n=1} \otimes \cdots \otimes \rho_{0,n=1}, \\ \rho_0(q) &= \sum_n \text{Pois}(n|n_0)\rho_{0,n}(q). \end{aligned} \tag{6}$$

The numerical evaluation of such a convolution is best done in Fourier space and will be the topic of the next section.

To compute significance distributions as a function of any phase space variable we use the same procedure as for the maximum significance. However, we restrict the events we use for the construction of $\rho_{0,n=1}(q)$ to be those which populate the phase space of a given bin of an observable of interest. We then iterate for each bin of that observable to fill the differential significance histogram.

B. Computing likelihood distributions

To compute the likelihood distribution we rely on a set of simulated events covering the entire phase space for each of the hypotheses H_0 and H_1. Following Eq. (5) we first construct $\rho_{0,n=1}(q)$ and $\rho_{1,n=1}(q)$. The convolution in LLR space can best be evaluated through a Fourier transform with $q \to \bar{q}$ and $\rho_{0,n} \to \bar{\rho}_{0,n}$. The original convolution in Eq. (6) turns into a product, namely

$$\bar{\rho}_0 = \sum_n \text{Pois}(n|n_0)\bar{\rho}_{0,n} = e^{-n_0}\sum_n \frac{n_0^n}{n!}\bar{\rho}_{0,n=1}^n = e^{n_0(\bar{\rho}_{0,n=1}-1)}, \tag{7}$$

where we have chosen, here and in the following, the H_0 hypothesis as a representative. The logic for H_1 follows in analogy. We then need to transform these distributions back into q space and numerically compute confidence levels by integrating over the relevant q range.

Following the structure of Eq. (6) we know that fast Fourier transforms based on the discrete Fourier transform (DFT) should be a convenient tool to numerically generate the full event likelihood distribution $\rho_0(q)$. DFTs work on an array of numbers a_j where $j \in [0, N-1]$ for fixed N. For a discretized function the Fourier transformation turns into

$$\bar{a}_k = \sum_{j=0}^{N-1} a_j \exp\left[-2\pi\frac{ijk}{N}\right], \tag{8}$$

where $\bar{a}_k$ denotes the discrete Fourier transform on the same size of array $k \in [0, N-1]$ [26]. There are some points which we need to take care of if we are to use this formalism. To use a DFT we need to project our function $\rho_{0,n=1}$ into a binned histogram $\rho_{0,n=1,\text{binned}}$, where in practice we use up to 400 bins for the projection. An issue with the full event likelihood distribution is that it will not have the same support in LLR space as the single event likelihood. We need to predefine an interval in LLR space to perform the complete computation on the same array. Using simple Poisson counting we can estimate the length of this interval to be [27]

$$q_{\text{length}} = \sum_{(q,f_0)\in\rho_{0,n=1,\text{binned}}} |q|(n_0 f_0 + \sigma^{\max}\sqrt{n_0 f_0}). \tag{9}$$

The additional factor in parentheses encodes the feature that a Poisson distribution is most likely localized around the expectation value n_0 with a standard deviation of $\sqrt{n_0}$. For each step or convolution a given q-value will typically move by a factor $n_0 + \sigma^{\max}\sqrt{n_0}$, where $\sigma^{\max}$ counts the number of maximally expected standard deviations. By summing over all bins of $\rho_{0,n=1,\text{binned}}$ and using the absolute value of q, we ensure that the full likelihood distribution will fit into the fixed size array. To be on the conservative side we use $\sigma^{\max} = 8$ and multiply the final result by a factor of two. Because both ρ_0 and ρ_1 have to be mapped to the same array to allow for a meaningful computation of confidence levels we also need to replace $n_0 f_0$ in Eq. (9) by its maximum values for the two hypotheses. This length we divide into 2^{19} bins a_j; see also Ref. [27] for more details.

The main difference between this analysis and the original signal-background study of Refs. [23,24] is that the LLR q can switch signs. As we do not restrict the allowed values of q when we compute the single event likelihood distribution from MC, we also need to know how to map $\rho_{0,n=1,\mathrm{binned}}$ onto the array a_j. By construction, the first bin in the DFT array corresponds to zero on the q axis [26]. The DFT represents a Fourier transform on a finite interval, which means periodic boundary conditions. If we are to use an arbitrary single event likelihood distribution where negative values of q are allowed, we need to respect these boundary conditions by moving all negative values to the very right of the a_j array, while the space in between will be filled with zeros. This way we can make sure to compute the correct full event likelihood distribution. To do so we use the standard fast Fourier transforms as implemented, for example, in the Python package Scipy.

C. Analysis setup

While our analysis closely follows the original proposal in Ref. [5], we ensure that in particular our detector simulation corresponds to the current state of the art. For the signal we use the NLO production cross section $\sigma(pp \to HH) = 34.8$ fb at 14 TeV center-of-mass energy [28]. For the particular $b\bar{b}\gamma\gamma$ final state the SM rate prediction is 0.092 fb. Unlike in the CMS study of Ref. [15] we only consider the high-purity category where both b-jets are tagged and which should dominate the significance at the high-luminosity LHC. Following Ref. [5] and the CMS study [15] main backgrounds are the nonresonant $\gamma\gamma + 2$ jets process—including $\gamma\gamma c\bar{c}$—and $\gamma j + 2$ jets production. Throughout our analysis we will see that these QCD continuum backgrounds have essentially the same shape, while their relative size depends on the photon and bottom identification. This also means that adding mistagged backgrounds can easily be taken into account by scaling the total QCD background rate.

Of the different Higgs processes we include the irreducible, resonant $ZH \to bb\gamma\gamma$ background, because it can be hard to separate from the background. Assuming $\sigma(ZH) = 0.9861$ pb, $\mathrm{BR}(Z \to bb) = 0.1512$, and $\mathrm{BR}(H \to \gamma\gamma) = 0.00227$, its cross section is $\sigma(ZH \to bb\gamma\gamma) = 3.38$ fb. The Z-peak is modeled by a double Gaussian fit to the CMS study [29]. All remaining nonresonant backgrounds as well as resonant backgrounds containing $H \to \gamma\gamma$ we assume to be negligible or comparably easy to control. This also applies to the reducible $t\bar{t}\gamma\gamma$ background, for which we would need to update the MadMax analysis to model the crucial QCD jet activity in the signal and background events.

For the two b-quarks and two photons in the final state we need to simulate the detector performance. We use a p_T-dependent and η-dependent b-tagging rate with $\epsilon_b \approx 0.7$ at high $p_T > 100$ GeV in the barrel, respectively. Towards larger p_T the b-tagging efficiency will be reduced. For our choice of efficiencies the $jj\gamma\gamma$ background ($j = u, d, s, g$) is negligible. If instead we choose a constant b-tagging efficiency and a p_T-dependent misidentification rate the light-flavor backgrounds in the low-p_T regime increase significantly. Similarly, the photon identification efficiency

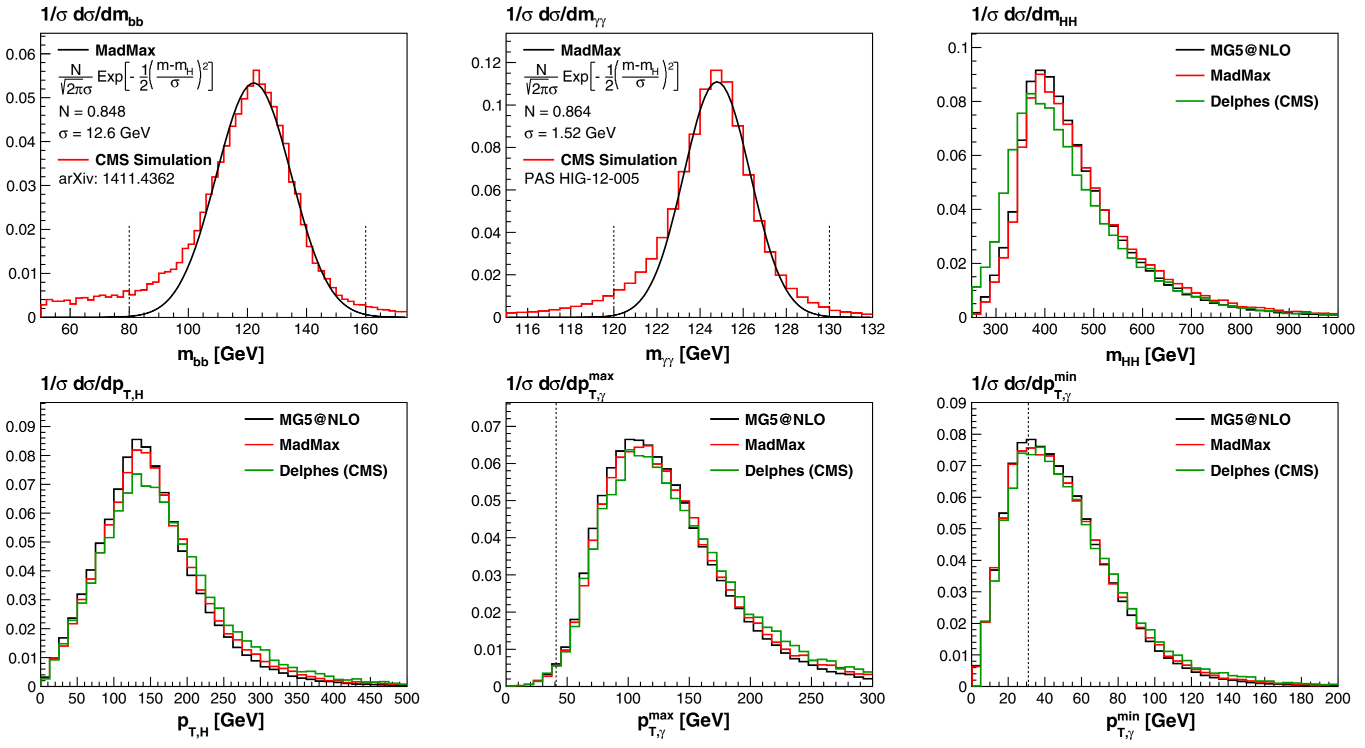

FIG. 2. First two panels: normalized invariant mass distributions of the dijet and diphoton systems as used in MadMax and from CMS [31,32]. Next four panels: di-Higgs invariant mass, transverse momentum spectra used in MadMax, including NLO corrections [28], and including the fast detector simulation Delphes [33].

and fake rate due to jets depend on the barrel vs end-cap position, the transverse momentum, and the parton forming the jet. All tagging efficiencies used in this analysis are described in Appendix B.

To simulate the detector effects in MadMax we modify MadGraph5[30] to smear the Higgs propagator like the square root of a Gaussian distribution. This allows us to reproduce the measured smearing of the CMS detector for $b\bar{b}$ [31] and $\gamma\gamma$ pairs [32]. In the upper panels of Fig. 2 we show that our prescription is adequate for the peak region, but faces limitations for the tails of the invariant mass distributions. However, these tails will hardly contribute to the signal significance. All we need to do is account for the loss of signal rate through these tails, 84.8% for the $b\bar{b}$ peak and 86.4% for the $\gamma\gamma$ peak.

Our choice of trigger cuts is motivated by CMS [15]: $p_{T,\gamma_1} > m_{\gamma\gamma}/3 \approx 41\,\text{GeV}$, $p_{T,\gamma_2} > m_{\gamma\gamma}/4 \approx 31\,\text{GeV}$, $|\eta_\gamma| < 2.5$, $\Delta R_{\gamma\gamma,\gamma j,jj} > 0.4$, $m_{\gamma\gamma} = 100\ldots180$ GeV, $p_{T,j} > 25$ GeV, $|\eta_j| < 2.4$, and $m_{jj} = 60\ldots180$ GeV. Our invariant mass windows $m_{bb} = 80\ldots160$ GeV and $m_{\gamma\gamma} = 120\ldots130$ GeV are designed to fully contain the peaks, as illustrated in the first two panels of Fig. 2. We apply these mass windows for the continuum background simulation throughout our analysis, which means that our background kinematics will not fully reproduce the QCD features.

Also in Fig. 2 we show the m_{HH} distribution and the different transverse momentum spectra entering our analysis. The MadMax model is based on a a loop-improved approach [28] which includes the NLO form factors presented in Ref. [2]. This model is also used by the ATLAS Collaboration [14]. One reference curve shows the full loop calculation using MG5-aMC@NLO [34], another a fast detector simulation with Delphes [33]. Even though our simulation might not correspond to a precision prediction at the percent level, we see that it reproduces the next-to-leading order results well.

II. STANDARD MODEL SIGNAL VS BACKGROUND

The first question we want to address in this study is, *which kinematic features allow us to extract Higgs pair production from backgrounds*? In this section we identify the Higgs self-coupling with its Standard Model value $\lambda_{\text{SM}} = 3m_H^2/v$, ensuring the perfect threshold cancellation in the heavy top limit shown in Eq. (2). In our statistics language the null-hypothesis H_0 is continuum backgrounds only, while the hypothesis H_1 is defined by Standard Model signal plus backgrounds.

In Fig. 3 we show a set of kinematic distributions for the Higgs decay products and for the reconstructed Higgs. At the 2 → 2 level we know that two distributions we want to study are the invariant mass distribution m_{HH} and the transverse momentum of each of the two reconstructed Higgs bosons. The dashed red line simply scales the signal histogram so we can actually see its kinematic distribution.

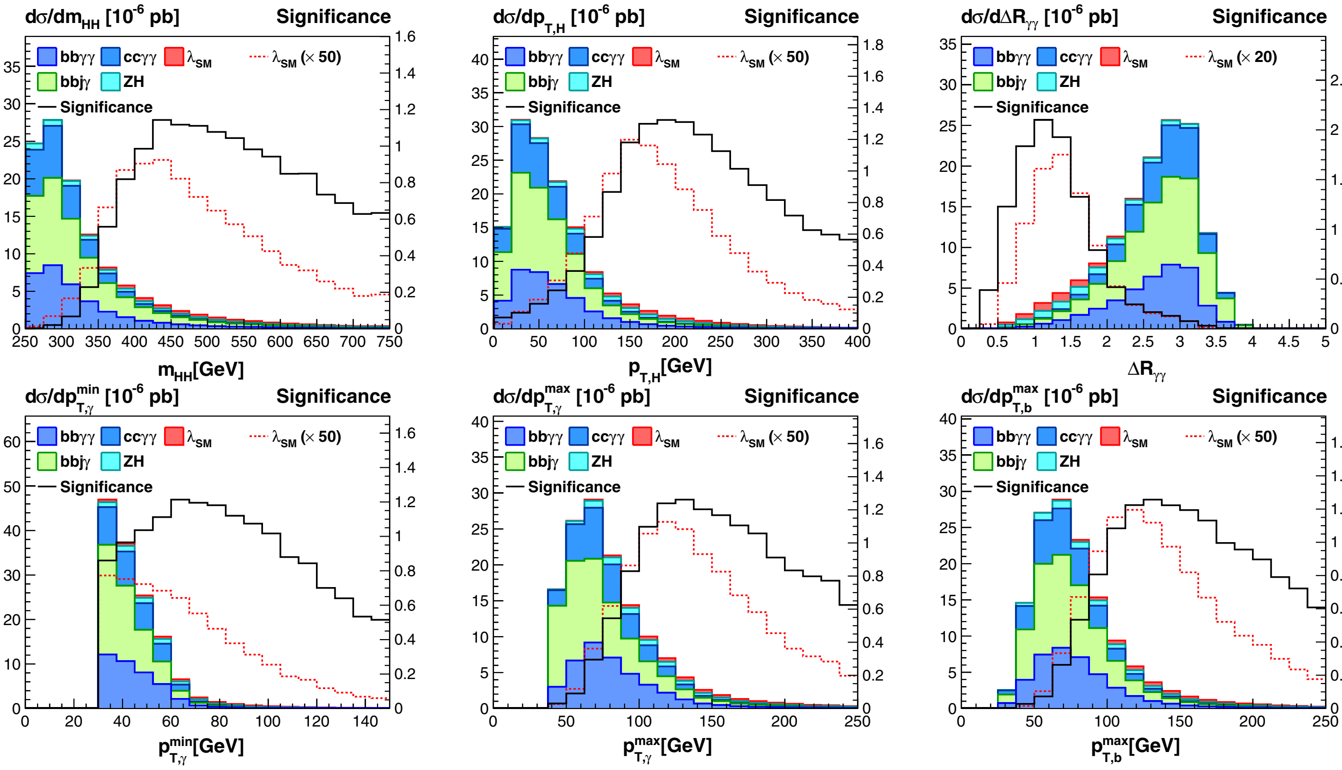

FIG. 3. Kinematic distributions after trigger cuts for the SM signal (red) vs the $bb\gamma\gamma$, $bbj\gamma$, $cc\gamma\gamma$, and ZH backgrounds. The solid black line shows the differential distribution of the significance.

For both Higgs distributions the continuum backgrounds are considerably softer than the Higgs pair signal. The same is true for the transverse momenta of the Higgs decay products. From these distributions it is also clear that an upper bound on $\Delta R_{\gamma\gamma}$ will help us extract the Higgs pair signal [5].

For the maximum significance with which we can extract the Standard Model signal from the continuum background we find a Gaussian equivalent of

$$4.02\sigma \quad \text{for an integrated luminosity of 3 ab}^{-1}. \qquad (10)$$

The black lines in Fig. 3 show how this significance is distributed over phase space. First, we observe that the threshold region $m_{HH} < 350$ GeV hardly contributes to the SM signal extraction. Instead, it seems crucial to include reconstructed Higgs bosons with $p_{T,H} > 150$ GeV, i.e. well in the boosted regime [12,24]. The separation of the significance distribution from the signal distributions occurs because of the rapidly falling background distributions. Due to the Higgs boost, widely separated photons with $\Delta R_{\gamma\gamma} > 1.8$ will not help with the signal extraction, while in particular for the harder of the two photons or bottoms we can completely ignore the soft part of the spectrum.

One key question for any multivariate analysis is if the phase space regions which dominate the signal vs background separation will be safe with respect to systematic and theoretical uncertainties. From the transverse momentum spectra in Fig. 3 we see that soft photons with $p_{T,\gamma} < 50$ GeV play hardly any role in separating the Higgs pair signal from the continuum background. Similarly, for the tagged bottom jets the relevant range is $p_{T,b} = 100\ldots250$ GeV. In this range we do not expect jet radiation and the related combinatorics to have a large effect on our results; the size of the usual perturbative rate corrections for the signal and background process should be theoretically under control. Finally, the relevant diphoton phase space is $\Delta R_{\gamma\gamma} = 0.5\ldots2$, clearly not a challenge for example to photon separation. A multivariate analysis of Higgs pair production in the Standard Model should be straightforward.

From Eq. (2) we know that the two Feynman diagrams contributing to Higgs pair production have distinct kinematic features, and their combination should allow us to understand the signal and significance distributions. In the upper panels of Fig. 4 we illustrate some of the kinematic distributions shown in Fig. 3, but for continuum Higgs pair production only. As expected, the two cases are similar, because the continuum diagram is responsible for almost the entire signal rate in the Standard Model. In the second set of plots we show the kinematic distributions of the triangle diagram only, i.e. the term of the cross section which carries the information on the Higgs self-coupling. As a first piece of information, we see that the significance with which we can extract the triangle diagrams in the

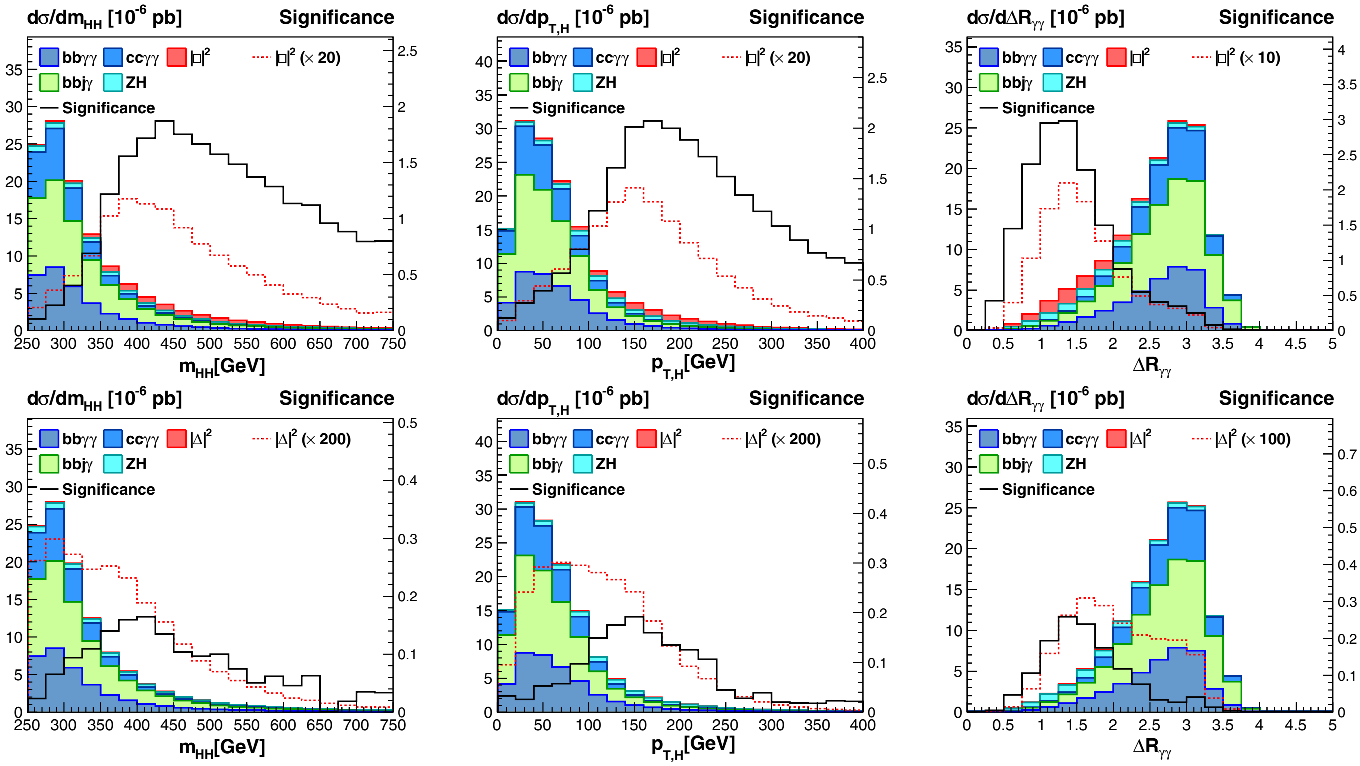

FIG. 4. Kinematic distributions for the SM triangle diagram only (upper) and the SM box diagram only (lower). We require trigger cuts for the signal (red) vs the $bb\gamma\gamma$, $bbj\gamma$, $cc\gamma\gamma$, and ZH backgrounds. The solid black line shows the differential distribution of the significance.

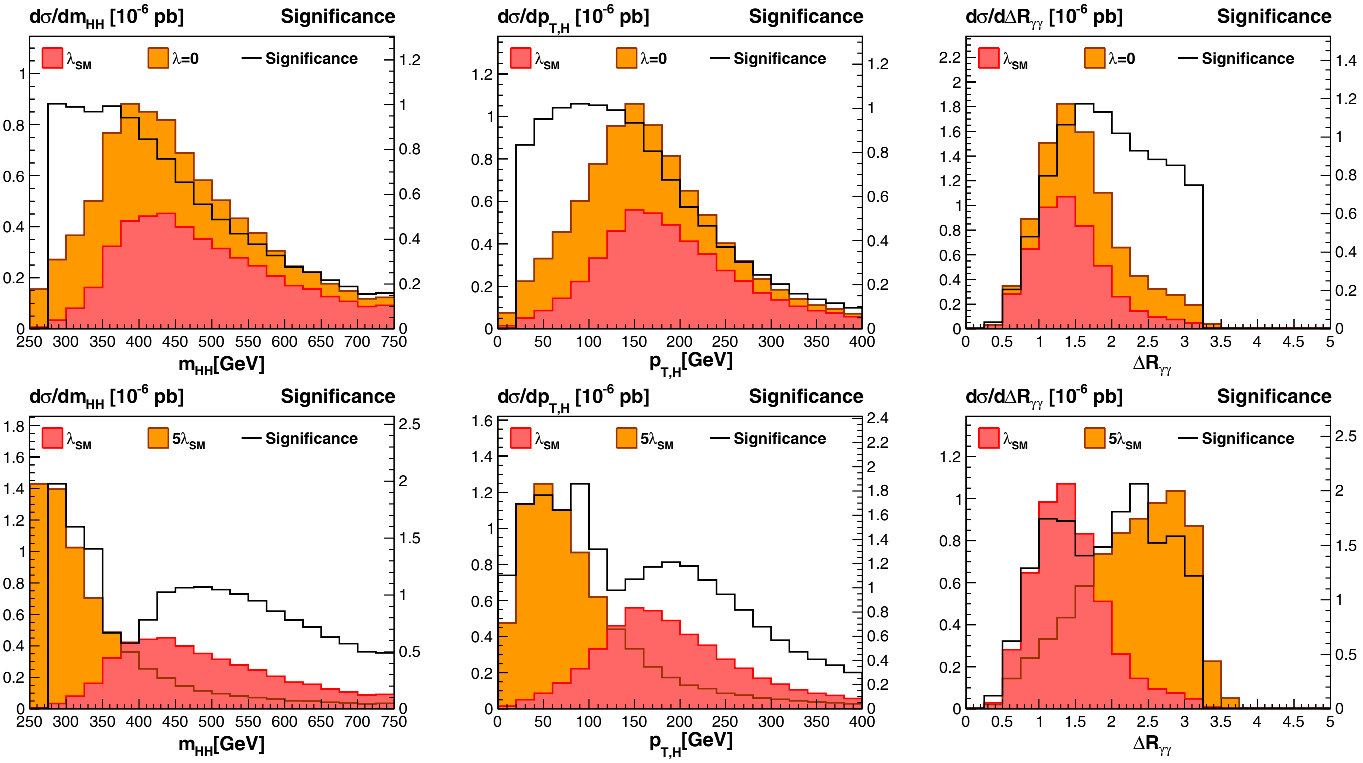

FIG. 5. Differential distributions assuming a modified Higgs self-coupling of $\lambda/\lambda_{\mathrm{SM}} = 0$ (upper panels) and $\lambda/\lambda_{\mathrm{SM}} = 5$ (lower panels). We compare only the two Higgs pair signals and neglect backgrounds for illustration.

absence of an interference with the continuum is extremely small. Kinematically, both in m_{HH} and $p_{T,H}$ this diagram has a considerably softer behavior. The two photons arising from a softer Higgs decay can now be more widely separated. These are the phase space regions where we can expect our signal vs background analysis to gain sensitivity to the value of the Higgs self-coupling, typically through an enhanced cancellation between the two Feynman diagrams shown in Fig. 1.

III. HIGGS SELF-COUPLING

In the second part of the paper we investigate the question, *what are the prospects of measuring the Higgs self-coupling* λ *in the presence of backgrounds*? In Fig. 5 we show some kinematic distributions comparing two signal hypotheses, assuming a variable self-coupling

$$\frac{\lambda}{\lambda_{\mathrm{SM}}} = \frac{\lambda}{\frac{3m_H^2}{v}} = 0, 2, 5. \qquad (11)$$

In the absence of a Higgs self-coupling we see that the Higgs pair production rate is significantly enhanced. While it is not immediately obvious from the two signal curves, the significance distribution over m_{HH} confirms that we can observe the effect of a zero self-coupling mostly close to threshold and for $m_{HH} < 450$ GeV. Similarly, the absence of a self-coupling will modify the transverse Higgs momentum around $p_{T,H} < 200$ GeV. In the photon separation $R_{\gamma\gamma}$ the sensitive range requires widely separated photons. We skip a detailed analysis of a slightly enhanced self-coupling $\lambda = 2\lambda_{\mathrm{SM}}$, because it will be most visible through a significant reduction of the production rate. The dramatic case of $\lambda = 5\lambda_{\mathrm{SM}}$ is shown in the lower panels of Fig. 5. The kinematic distributions are now modified close to threshold and for small transverse momenta of each Higgs

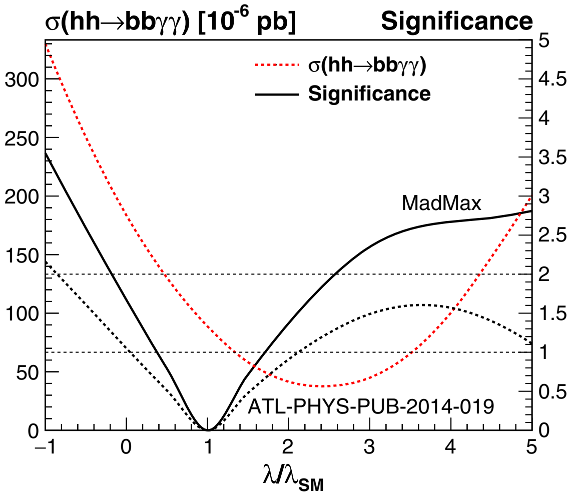

FIG. 6. Signal cross section (red dashed line) and maximum significance (black solid line) for observing an anomalous Higgs self-coupling at the LHC with an integrated luminosity of 3 ab^{-1}. We also show the significance from a cut-based rate measurement using the cuts suggested in Ref. [35] (black dashed line).

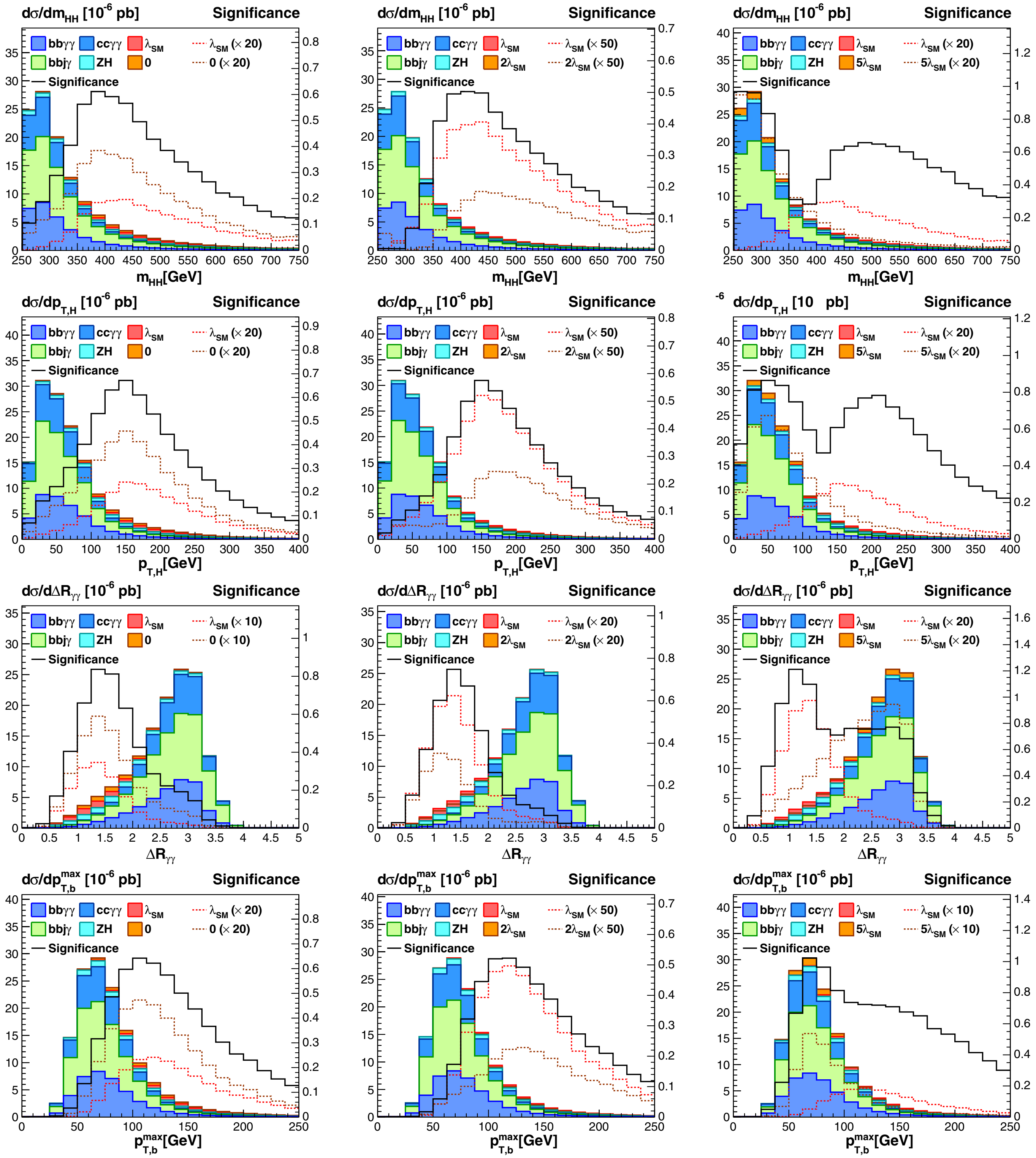

FIG. 7. Differential distributions assuming a modified Higgs self-coupling of $\lambda/\lambda_{SM} = 0, 2, 5$ (left to right). Unlike for the illustration in Fig. 5 we compare the two proper hypotheses including signal and backgrounds.

boson, as expected from an s-channel-mediated process $gg \to H^* \to HH$.

In the next step we add the QCD continuum and ZH backgrounds and test how well we can extract the Higgs self-coupling from the kinematic information. In Fig. 6 we show the maximum significance of a multivariate analysis as well as the significance from a counting experiment based on the total Higgs pair cross section after applying some basic cuts [35]. As mentioned above, the Higgs pair production rate increases towards vanishing self-coupling, has a minimum of roughly a quarter of the Standard Model rate around $\lambda = 2.3\lambda_{\rm SM}$, and increases again for larger self-couplings. The arguably most interesting case of zero self-coupling, related to a Higgs potential without the usual minimum, should be most easily distinguishable from the Standard Model [2,3]. The LHC will at best constrain the Higgs self-coupling range to be

$$\frac{\lambda}{\lambda_{\rm SM}} = 0.4\ldots1.7 \quad \text{at 68\% CL and for 3 ab}^{-1} \qquad (12)$$

and rule out

$$\frac{\lambda}{\lambda_{\rm SM}} = -0.2\ldots2.6 \quad \text{at 95\% CL and for 3 ab}^{-1}. \qquad (13)$$

Obviously, we can try to combine this result with additional signatures to enhance the final reach of the LHC [3–5,12].

From all we have discussed to here, the obvious question is how much of the pure signal interference structure and its sensitivity to the self-coupling survives in the presence of backgrounds, and to what degree background uncertainties can mimic an anomalous self-coupling. In Fig. 7 we compare pairs of hypotheses for signal plus background, i.e. we test how well we can measure the Higgs self-coupling in the presence of the backgrounds. The Standard Model signal is shown in red and the alternative model with its shifted self-coupling $\lambda/\lambda_{\rm SM} = 0, 2, 5$ is shown in orange. In our simulation we always apply mass windows for the $b\bar{b}$ and $\gamma\gamma$ pairs, which already biases our background simulations towards signal-like, hard configurations. Nevertheless, in Fig. 7 we see that even then the QCD backgrounds still populate low-m_{HH} and low-$p_{T,H}$ regions. As already shown in Fig. 5, these are exactly the phase space regions where the sensitivity to an anomalous self-coupling is the largest.

For the measurement of λ in the presence of backgrounds, the sensitive regions of phase space are defined by a combination of background rejection and sensitivity to λ. While for background rejection described in Fig. 3 the region with high $m_{HH} > 400$ GeV and $p_{T,H} > 150$ GeV are most useful, the self-coupling measurement requires lower-m_{HH} and lower-$p_{T,H}$ bins, as seen in Fig. 5. As a result, the significance peak in the SM background rejection around $m_{HH} = 450$ GeV moves to slightly below 400 GeV when we are interested in λ. Similarly, the background-driven significance at $p_{T,H} > 180$ GeV and the self-coupling-sensitive region $p_{T,H} < 150$ GeV together give a distinct peak at 150 GeV for $\lambda/\lambda_{\rm SM} = 0$ or $\lambda/\lambda_{\rm SM} = 2$. In contrast, for a very large self-coupling $\lambda/\lambda_{\rm SM} = 5$ the significance receives contributions from two distinct regions of phase space, $p_{T,H} \approx 50$ GeV and $p_{T,H} \approx 220$ GeV. The geometric separation of the two photons and the transverse momentum of the harder b-jet follow the same pattern. This means that the most significant phase space regions for a measurement of the self-coupling are driven by the background rejection, shifted by the well-known regions of phase space carrying sensitivity to the self-coupling. Large deviations from the phase space regions for background rejection only occur when we test very large self-couplings.

IV. OUTLOOK

Multivariate analyses often challenge our understanding of what limiting factors of important measurements are. To gauge the sensitivity for example of a Higgs self-coupling measurement to different sources of uncertainties we need to understand where the relevant phase space regions for a given measurement are [23,24].

In this paper we have studied the phase space regions which contribute to the extraction of the Higgs pair production signal from the continuum backgrounds, as well as those regions allowing for a measurement of the Higgs self-coupling. We focus on the $HH \to b\bar{b}\gamma\gamma$ signature [5], but expect our results to also hold for other channels with large continuum backgrounds. The two relevant phase space regions for the signal extraction and the coupling measurement are separate and in particular for the signal extraction well understood in terms of systematic and theoretical errors.

The most sensitive phase space region for extracting the self-coupling is close to threshold, where we expect the QCD background to overwhelm the Higgs pair signal. The main question will be how well we understand those backgrounds and how much of this region can still be used for the self-coupling measurement. Assuming SM-like self-coupling, the bulk of the coupling-sensitive region in the presence of QCD backgrounds is only slightly softer than the relevant phase space for the extraction of the Standard Model signal. For large self-couplings this region shifts significantly, forcing us to consider a proper hypothesis test for a variable Higgs self-coupling.

ACKNOWLEDGMENTS

First, we still remember Uli Baur for coming up with this measurement and for developing the first and original analyses. Second, we are grateful to Kyle Cranmer for his continuous support and to Bill Murray for his valuable comments concerning v1. Moreover, we would like to thank Martin Jankowiak for his contributions during an early phase of the project. We thank David Bourne and Anthony Yeates for helpful discussions. T. P. and P. S. would like to thank the MITP for the hospitality and support while this paper was finally written. The work of F. K. was supported by the U.S. Department of Energy under Grant No. DE-FG02-04ER-41298. F. K. also acknowledges support from the Fermilab Graduate

Student Research Program in Theoretical Physics operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy. The work of P. S. was supported in part by the European Union as part of the FP7 Marie Curie Initial Training Network MCnetITN (PITN-GA-2012-315877).

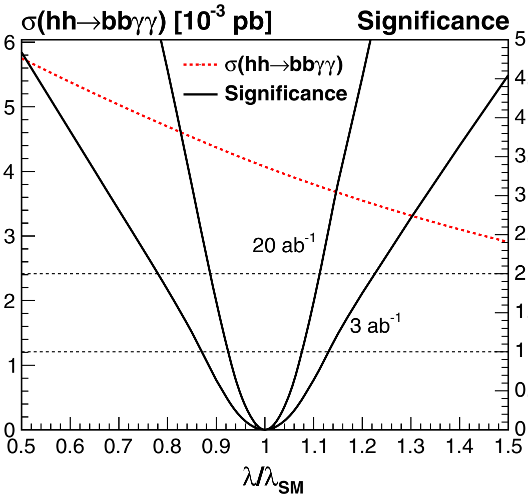

FIG. 8. Signal cross section (red) and maximum significance (black) for observing an anomalous Higgs self-coupling at a 100 TeV hadron collider with an integrated luminosity of 3 ab^{-1} and 20 ab^{-1}. The setup corresponds to Fig. 6 for the high-luminosity LHC.

APPENDIX A: FUTURE 100 TEV COLLIDER

We can use exactly the same setup as in the main body of the paper to estimate the impact of the signal kinematics at a future 100 TeV collider with an integrated luminosity of 20 ab^{-1}. Several studies, largely based on a rate measurement with some background rejection cuts, have shown that Higgs pair production can be probed at such a machine with high precision [36,37].

In Fig. 8 we show the maximum significance for distinguishing a modified Higgs self-coupling from the Standard Model at a future 100 TeV collider. The setup is exactly the same as for our LHC analysis leading to the results shown in Fig. 9. This includes our detector smearing as well as the trigger cuts. The signal cross section is now taken from MadMax without an external normalization. Relying on a multivariate analysis we can at best constrain the Higgs self-coupling range to be

$$\frac{\lambda}{\lambda_{\mathrm{SM}}} = 0.92...1.07 \quad \text{at 68\% CL and for 20 ab}^{-1} \tag{A1}$$

and rule out

$$\frac{\lambda}{\lambda_{\mathrm{SM}}} = 0.89...1.11 \quad \text{at 95\% CL and for 20 ab}^{-1}. \tag{A2}$$

In Fig. 9 we also show the significance distribution over phase space. It is essentially identical to the 14 TeV case,

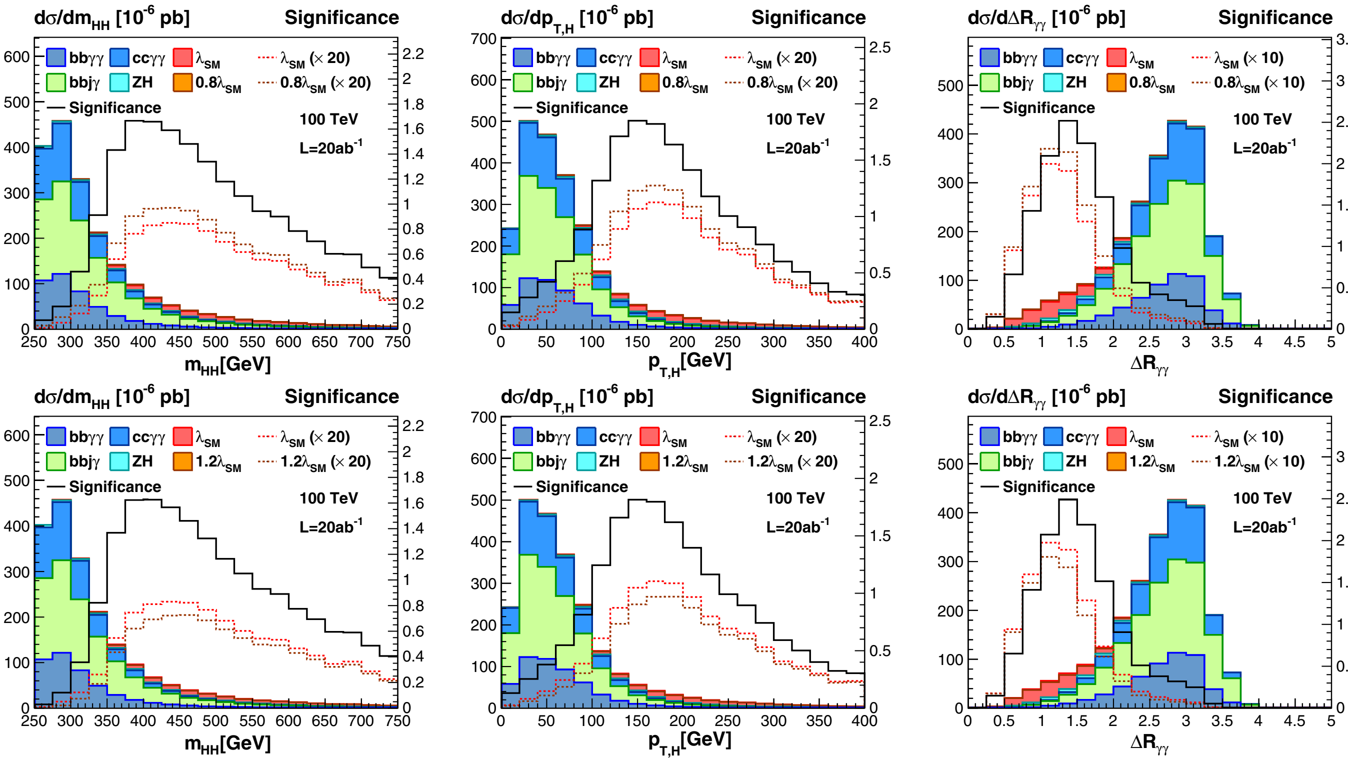

FIG. 9. Differential distributions assuming a modified Higgs self-coupling of $\lambda/\lambda_{\mathrm{SM}} = 0.8, 1.2$ in the presence of backgrounds. The setup corresponds to Fig. 7 for the high-luminosity LHC.

because the relevant energy scale is given by the Standard Model Higgs mass in the two propagators. As for the LHC this implies that systematic and theoretical uncertainties should not pose a major issue for this precision measurement, which is driven by a very large number of signal events.

These 100 TeV results quoted above appear, at first, to disagree with some results shown in the literature. The main reasons for this are diverging assumptions about tagging efficiencies. Since the Higgs decay products are mostly concentrated in the central detector, we do not expect the forward coverage of 100 TeV collider to significantly affect our results. However, we find that assumptions about b-tagging are crucial. The signal's $p_{T,b}$-distribution peaks around $m_H/2$ and therefore in a regime with suppressed tagging efficiency, about 45% according to the parametrization in Appendix B. An improved b-tagging efficiency at low transverse momentum will significantly enhance the signal rate and therefore improve the sensitivity for the triple-Higgs coupling. This explains the better reach quoted in Ref. [37], which assumes a constant 75% tagging efficiency.

APPENDIX B: TAGGING EFFICIENCIES

Because the tagging and identification efficiencies have a crucial effect on our results, we give the necessary details in this appendix. The b-tagging efficiency and the corresponding mistag rate depends on the parton forming the tagged jet, the barrel respectively end-cap position in the detector, and the transverse momentum. The b-related tagging efficiencies we use in our analysis are [38]

$$\begin{aligned}
&b\text{-jets} && \epsilon_b = \begin{cases} 0.7\tanh\left(0.01317 p_T/\text{GeV} - 0.062\right) & |\eta| < 1.2 \\ 0.6\tanh\left(0.01050 p_T/\text{GeV} - 0.101\right) & 1.2 < |\eta| < 2.5 \end{cases} \\
&c\text{-jets} && \epsilon_c = \begin{cases} 0.1873\tanh\left(0.01830 p_T/\text{GeV} - 0.2196\right) & |\eta| < 1.2 \\ 0.1898\tanh\left(0.00997 p_T/\text{GeV} - 0.1430\right) & 1.2 < |\eta| < 2.5 \end{cases} \\
&\text{light-flavor jets} && \epsilon_j = 0.001.
\end{aligned} \tag{B1}$$

For the photons we follow Ref. [39] (Fig. 9.22), which suggests a photon identification efficiency $\epsilon'_\gamma = 0.85$ and a p_T-dependent mistag rate between $\epsilon'_j = 0.002$ and 0.0001 for jets between 50 GeV and 100 GeV:

$$\begin{aligned}
&\text{photon} && \epsilon'_\gamma = 0.85 \\
&\text{light-flavor jets} && \epsilon'_j = \begin{cases} 0.01133\exp\left(-0.038 p_T/\text{GeV}\right) & p_T < 100\ \text{GeV} \\ 0.00025 & p_T > 100\ \text{GeV}. \end{cases}
\end{aligned} \tag{B2}$$

[1] 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. **B309**, 282 (1988).

[2] T. Plehn, M. Spira, and P. M. Zerwas, Nucl. Phys. **B479**, 46 (1996); **B531**, 655 (1998); A. Djouadi, W. Kilian, M. Muhlleitner, and P. M. Zerwas, Eur. Phys. J. C **10**, 45 (1999); X. Li and M. B. Voloshin, Phys. Rev. D **89**, 013012 (2014).

[3] U. Baur, T. Plehn, and D. L. Rainwater, Phys. Rev. Lett. **89**, 151801 (2002); Phys. Rev. D **67**, 033003 (2003).

[4] U. Baur, T. Plehn, and D. L. Rainwater, Phys. Rev. D **68**, 033001 (2003).

[5] U. Baur, T. Plehn, and D. L. Rainwater, Phys. Rev. D **69**, 053004 (2004).

[6] J. Baglio, A. Djouadi, R. Gröber, M. M. Mühlleitner, J. Quevillon, and M. Spira, J. High Energy Phys. 04 (2013) 151.

[7] For a currrent analysis see e.g. T. Corbett, O. J. P. Eboli, D. Goncalves, J. Gonzalez-Fraile, T. Plehn, and M. Rauch, J. High Energy Phys. 08 (2015) 156.

[8] M. L. Mangano, T. Plehn, P. Reimitz, T. Schell, and H. S. Shao, J. Phys. G **43**, 035001 (2016).

[9] For an up-to-date analysis of the Higgs-gauge sector see A. Butter, O. J. P. Eboli, J. Gonzalez-Fraile, M. C. Gonzalez-Garcia, T. Plehn, and M. Rauch, J. High Energy Phys. 07 (2016) 152.

[10] F. Goertz, A. Papaefstathiou, L. L. Yang, and J. Zurita, J. High Energy Phys. 04 (2015) 167; Q. H. Cao, B. Yan, D. M. Zhang, and H. Zhang, Phys. Lett. B **752**, 285 (2016); M. Gorbahn and U. Haisch, J. High Energy Phys. 10 (2016) 094; G. Degrassi, P. P. Giardino, F. Maltoni, and D. Pagani, J. High Energy Phys. 12 (2016) 080.

[11] M. A. Shifman, A. I. Vainshtein, M. B. Voloshin, and V. I. Zakharov, Yad. Fiz. **30**, 1368 (1979) [Sov. J. Nucl. Phys. **30**, 711 (1979)]; B. A. Kniehl and M. Spira, Z. Phys. C

69, 77 (1995); M. Spira, J. High Energy Phys. 10 (2016) 026.

[12] M. J. Dolan, C. Englert, and M. Spannowsky, J. High Energy Phys. 10 (2012) 112; A. J. Barr, M. J. Dolan, C. Englert, and M. Spannowsky, Phys. Lett. B **728**, 308 (2014).

[13] V. Barger, L. L. Everett, C. B. Jackson, A. D. Peterson, and G. Shaughnessy, Phys. Rev. D **90**, 095006 (2014).

[14] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **92**, 092004 (2015).

[15] V. Khachatryan *et al.* (CMS Collaboration), Phys. Rev. D **94**, 052012 (2016).

[16] A. Papaefstathiou, L. L. Yang, and J. Zurita, Phys. Rev. D **87**, 011301 (2013).

[17] D. E. Ferreira de Lima, A. Papaefstathiou, and M. Spannowsky, J. High Energy Phys. 08 (2014) 030; D. Wardrope, E. Jansen, N. Konstantinidis, B. Cooper, R. Falla, and N. Norjoharuddeen, Eur. Phys. J. C **75**, 219 (2015); J. K. Behr, D. Bortoletto, J. A. Frost, N. P. Hartland, C. Issever, and J. Rojo, Eur. Phys. J. C **76**, 386 (2016).

[18] Q. Li, Z. Li, Q. S. Yan, and X. Zhao, Phys. Rev. D **92**, 014015 (2015).

[19] See e.g. J. M. No and M. Ramsey-Musolf, Phys. Rev. D **89**, 095031 (2014); B. Hespel, D. Lopez-Val, and E. Vryonidou, J. High Energy Phys. 09 (2014) 124; C. Y. Chen, S. Dawson, and I. M. Lewis, Phys. Rev. D **91**, 035015 (2015); P. Huang, A. Joglekar, B. Li, and C. E. M. Wagner, Phys. Rev. D **93**, 055049 (2016).

[20] R. Gröber and M. Mühlleitner, J. High Energy Phys. 06 (2011) 020; R. Gröber, M. Mühlleitner, and M. Spira, J. High Energy Phys. 06 (2016) 080.

[21] S. Dawson, S. Dittmaier, and M. Spira, Phys. Rev. D **58**, 115012 (1998); J. Grigo, J. Hoff, K. Melnikov, and M. Steinhauser, Nucl. Phys. **B875**, 1 (2013); F. Maltoni, E. Vryonidou, and M. Zaro, J. High Energy Phys. 11 (2014) 079.

[22] D. de Florian and J. Mazzitelli, J. High Energy Phys. 09 (2015) 053; J. Grigo, J. Hoff, and M. Steinhauser, Nucl. Phys. **B900**, 412 (2015); S. Borowka, N. Greiner, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk, U. Schubert, and T. Zirke, Phys. Rev. Lett. **117**, 012001 (2016); D. de Florian, M. Grazzini, C. Hanga, S. Kallweit, J. M. Lindert, P. Maierhofer, J. Mazzitelli, and D. Rathlev, J. High Energy Phys. 09 (2016) 151.

[23] K. Cranmer and T. Plehn, Eur. Phys. J. C **51**, 415 (2007).

[24] T. Plehn, P. Schichtel, and D. Wiegand, Phys. Rev. D **89**, 054002 (2014).

[25] For a proof and corresponding definitions, see e.g. A. Stuart, K. Ord, and S. Arnold, *Kendall's Advanced Theory of Statistics, Volume 2A, Classical Inference and the Linear Model, 6th Edition* (Oxford University Press, New York, 1994); For a pedagogical introduction in the context of high-energy physics, see F. James, Y. Perrin, and L. Lyons, in CERN Report No. CERN-2000-005, 2000, https://inspirehep.net/record/534129.

[26] L. N. Trefethen, *Spectral Methods in MATLAB*, Software, Environments and Tools (SIAM, Philadelphia, 2000).

[27] K. Cranmer, LEPStats4LHC, http://phystat.org/phystat/packages/0703002.1.html.

[28] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, P. Torrielli, E. Vryonidou, and M. Zaro, Phys. Lett. B **732**, 142 (2014).

[29] P. Eller, arXiv:1408.5044.

[30] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, J. High Energy Phys. 06 (2011) 128.

[31] C. Vernieri (CMS Collaboration), Nucl. Part. Phys. Proc. **273**, 733 (2016).

[32] CMS Collaboration (CMS Collaboration), Report No. CMS-PAS-HIG-15-005.

[33] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3 Collaboration), J. High Energy Phys. 02 (2014) 057.

[34] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, J. High Energy Phys. 07 (2014) 079.

[35] ATLAS Collaboration (ATLAS Collaboration), Report No. ATL-PHYS-PUB-2014-019.

[36] A. J. Barr, M. J. Dolan, C. Englert, D. E. Ferreira de Lima, and M. Spannowsky, J. High Energy Phys. 02 (2015) 016; A. Azatov, R. Contino, G. Panico, and M. Son, Phys. Rev. D **92**, 035001 (2015); A. Papaefstathiou, Phys. Rev. D **91**, 113016 (2015); H. J. He, J. Ren, and W. Yao, Phys. Rev. D **93**, 015003 (2016); A. Papaefstathiou and K. Sakurai, J. High Energy Phys. 02 (2016) 006; C. Y. Chen, Q. S. Yan, X. Zhao, Y. M. Zhong, and Z. Zhao, Phys. Rev. D **93**, 013007 (2016); B. Fuks, J. H. Kim, and S. J. Lee, Phys. Rev. D **93**, 035026 (2016).

[37] R. Contino *et al.*, arXiv:1606.09408.

[38] J. Anderson *et al.*, arXiv:1309.1057.

[39] J. Butler, D. Contardo, M. Klute, J. Mans, and L. Silvestris, Report No. CERN-LHCC-2015-010.