Search for Higgs boson pair production in the $b\bar{b}\mu^+\mu^-$ final state at the LHC

Botao Guo, Xiaohu Sun,* Licheng Zhang, Zhe Li, and Yong Ban

Department of Physics and State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing 100871, China

(Received 29 August 2022; revised 4 December 2022; accepted 19 January 2023; published 22 February 2023)

The Higgs boson pair production via gluon-gluon fusion and vector boson fusion in the $b\bar{b}\mu^+\mu^-$ final state at the LHC is studied to probe the Higgs self-coupling κ_λ and the four-boson HHVV coupling κ_{2V} for the first time. A cut-based analysis and a machine-learning analysis using boosted decision trees are performed with categorizations and optimizations depending on the variations of these couplings. The expected sensitivities are extracted with different integrated luminosities assumed up to the full high-luminosity LHC runs. The expected upper limit at 95% confidence level on the Higgs boson pair production is calculated as 47 (28) times the Standard Model cross section using the cut-based method (boosted decision trees) for the gluon-gluon fusion production and 928 for the vector boson fusion production, assuming an integrated luminosity of 3000 fb^{-1}. The expected constraints on the couplings at 95% confidence level are calculated to be $-13.8 < \kappa_\lambda < 19.1$ ($-10.0 < \kappa_\lambda < 15.5$) and $-3.4 < \kappa_{2V} < 5.5$ using the cut-based method (boosted decision trees), respectively, assuming an integrated luminosity of 3000 fb^{-1}.

DOI: 10.1103/PhysRevD.107.034014

I. INTRODUCTION

Since the discovery of the Higgs boson by the ATLAS and CMS collaborations [1,2], the measurements of its properties are of high priorities in order to understand the Brout-Englert-Higgs mechanism [3–8] in the Standard Model (SM). The self-interaction of the Higgs boson scaled by the relative self-coupling $\kappa_\lambda = \lambda/\lambda_{\mathrm{SM}}$ is fundamental in the determination of the shape of the Higgs potential. The Higgs boson pair production (HH) is the only accessible mode in the direct probe of the Higgs self-coupling at the LHC and the high-luminosity LHC (HL-LHC). Given its low rates at the LHC, any enhancement of the HH production can also indicate physics beyond the Standard Model (BSM) [9–44].

In proton-proton collisions, the largest HH production mode is gluon-gluon fusion (ggF) with a cross section of $31.05^{+6\%}_{-23\%}(\mathrm{scale} + m_t) \pm 3.0\%(\mathrm{PDF} + \alpha_S)$ fb calculated at next-to-next-to-leading order in QCD with top quark mass effects ($\mathrm{FT}_{\mathrm{approx}}$) [45] at center-of-mass energy ($\sqrt{s}$) of 13 TeV. This is followed by vector boson fusion (VBF) with a cross section of $1.726^{+0.03\%}_{-0.04\%}(\mathrm{scale}) \pm 2.1\%(\mathrm{PDF} + \alpha_S)$ fb calculated at next-to-next-to-next-to-leading order in QCD [46]. The ggF mode provides a direct probe of κ_λ at the leading order (LO), while the VBF mode can additionally open a window to the four-boson HHVV coupling κ_{2V}, as shown in Fig. 1. Other production modes such as VHH, ttHH, and tjHH with much smaller cross sections [47,48] are not discussed here.

Given the two Higgs bosons simultaneously produced in the process, there is a variety in its final states to explore. Recent experimental results performed by ATLAS and

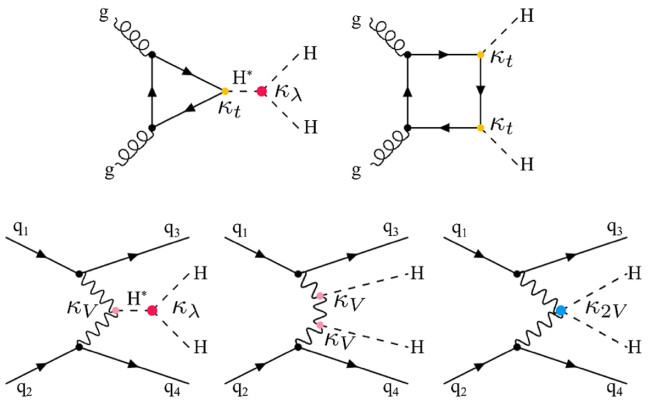


FIG. 1. Leading-order diagrams contributing to the Higgs pair production: ggF and VBF. The ggF mode contains the trilinear Higg self-coupling κ_λ and the top quark Yukawa coupling κ_t. The VBF mode contains the four-boson HHVV coupling κ_{2V}, the HVV coupling κ_V, and the trilinear Higgs self-coupling κ_λ.

*Corresponding author.
xiaohu.Sun@pku.edu.cn

CMS cover the decay channels of $b\bar{b}b\bar{b}$ [49–52], $b\bar{b}\gamma\gamma$ [53–56], $b\bar{b}\tau^+\tau^-$ [57–59], $b\bar{b}ZZ^*$ [60,61], $b\bar{b}WW^*$ [62–64], WW^*WW^* [65,66],$WW^*\tau^+\tau^-$ [66], $WW^*\gamma\gamma$ [67], and $\tau^+\tau^-\tau^+\tau^-$ [66]. In terms of the expected sensitivity, the upper limits from the leading decay channels reach around four times the SM prediction on the ggF HH cross section, while the combined results start to get close to two times [68–72].

By now, the searches of the HH production have covered the Higgs decays to bosons and third-generation fermions, mostly benefiting from the large branching ratios, while they have not fully explored the decays involving the second-generation fermions, such as $HH \to b\bar{b}\mu^+\mu^-$. This is a rare decay channel but can benefit from the excellent mass resolution in the dimuon invariant mass $m_{\mu\mu}$ at a level of 1–2 GeV [73], similar to $m_{\gamma\gamma}$, at the LHC. Moreover, CMS [73] has claimed the evidence of $H \to \mu\mu$, and ATLAS [74] also found large excess in this final state. Experimentally, this decay channel becomes feasible.

We present a comprehensive study of the HH searches in $HH \to b\bar{b}\mu^+\mu^-$, focusing on not only ggF but also VBF production modes, with a dependence on the Higgs self-coupling κ_λ and the HHVV coupling κ_{2V}. In measuring κ_λ and κ_{2V}, other couplings, such as κ_V and κ_t that directly enter the leading HH diagrams, are all set to their SM values, following many of the recent ATLAS and CMS HH analyses. The measurements of κ_V and κ_t mostly rely on single Higgs processes and thus are not discussed in the scope of this paper. Both the cut-based and boosted decision trees (BDT) methods are applied in the optimization of event selections. The final fits are performed on the kinematic shapes including the dimuon invariant mass $m_{\mu\mu}$ together with the dibjet invariant mass m_{bb}. For a comparison to the existing literature, Ref. [75] briefly discusses $HH \to b\bar{b}\mu^+\mu^-$ with a cut-based method using counting experiments, only targeting at the SM ggF HH production. Reference [76] considered $HH \to b\bar{b}\mu^+\mu^-$ with a multivariate analysis using the BDT algorithm at the high-energy LHC.

This paper is structured as follows. Section II introduces the simulated signal and background samples. Section III describes the analysis strategies, event categorizations, and selections. Section IV focuses on the BDT analysis. Finally, Sec. V reports the results. and Sec. VI summarizes the conclusions.

II. SIGNAL AND BACKGROUND SAMPLES

For the signal Monte Carlo (MC) samples, the ggF HH processes are generated at next-to-leading order (NLO) in QCD using POWHEG-BOX-V2 [77,78], while the VBF HH processes are generated at LO in QCD using MG5_aMC@NLO2.6.5 [79]. In total, seven ggF HH MC samples are generated with different κ_λ values, $\kappa_\lambda = -5, 0, 1, 2.4, 5, 10$, and 20, and seven VBF HH MC samples are generated with different κ_{2V} values,

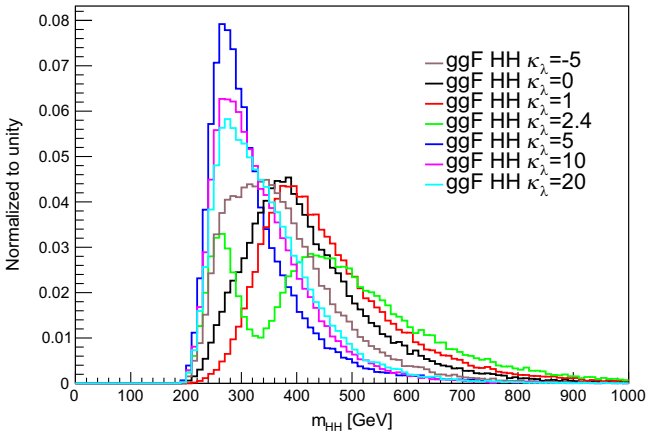


FIG. 2. The m_{HH} distribution of ggF HH signals with different κ_λ values.

$\kappa_{2V} = -10, -5, 0, 1, 2, 5$, and 10. Other coupling values used in the scan later are from the combination of the generated samples, since the differential cross section is scaled by a second-order polynomial of κ_λ and κ_{2V} in the LO electroweak precision. This approach of linearly combining samples from different couplings follows the treatment in Ref. [49].

The relevant couplings modify not only the total rates but also the kinematics significantly, as shown in Figs. 2–5, in which m_{HH} is the reconstructed invariant mass of $b\bar{b}\mu^+\mu^-$ and $p_T^{\mu\mu}$ is the transverse momentum of the $\mu\mu$ system. In these figures, all distributions are normalized to unity in order to have a direct comparison on the shapes. For ggF HH process, in Fig. 2, distributions of the invariant mass m_{HH} of the Higgs boson pair system are displayed for different values of κ_λ. They exhibit a characteristic dip at $m_{HH} \sim 350$ GeV for $\kappa_\lambda \sim 2.4$ [80]. This value of the trilinear Higgs self-coupling corresponds to a maximal destructive interference between the triangle and box diagrams in Fig. 1. For $\kappa_\lambda = 1$, the maximal destructive

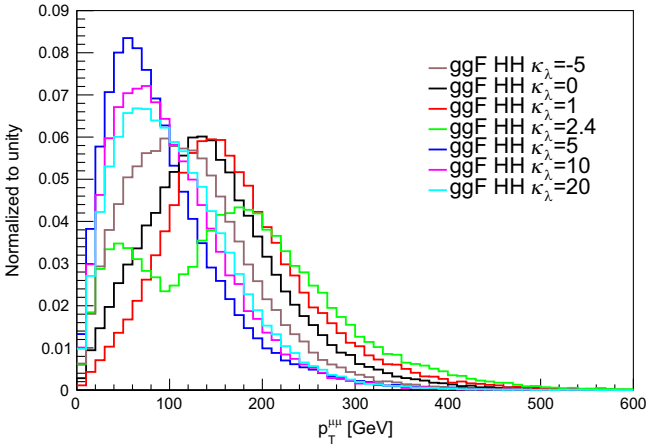


FIG. 3. The $p_T^{\mu\mu}$ distribution of ggF HH signals with different κ_λ values.

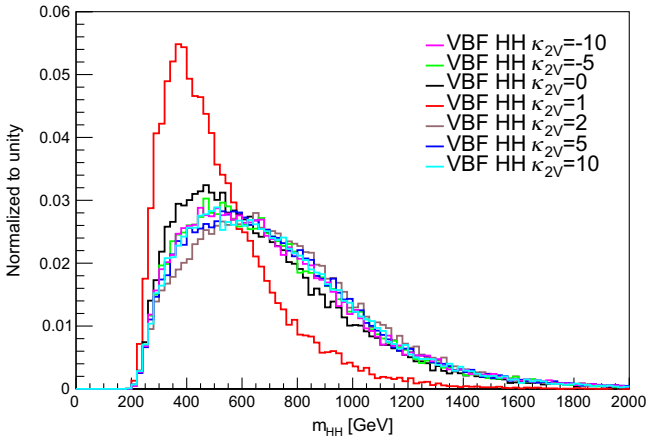


FIG. 4. The m_{HH} distribution of VBF HH signals with different $\kappa_{2\mathrm{V}}$ values.

interference happens at the HH production threshold and therefore does not introduce a visible dip in the distribution. For larger κ_λ values, the triangle diagram starts to dominate, resulting in a softer spectrum. The similar interference structure is manifested in the distribution of $p_{\mathrm{T}}^{\mu\mu}$ in Fig. 3. In the VBF process, the distributions tend to be harder for the BSM cases with $\kappa_{2\mathrm{V}}$ deviating from 1, as shown in Figs. 4 and 5.

Among the background processes, the Drell-Yan (DY) process is dominant in general. The DY samples are generated with at least one b-quark and up to four quarks associated, using MG5_aMC@NLO2.6.5. The generated DY processes include llb, $llbj$, $llbjj$, and $llbjjj$, where l stands for muon, b stands for bottom quark, and j stands for all quarks except the top quark. To enhance the MC statistics in the signal enriched phase space, the samples are generated with m_{ll} slices of [100, 150], [150, 200], and $[200, +\infty]$ GeV. A k factor of 1.23 [81] is used for the NLO QCD correction in the DY process. No k factors are applied for other background processes, given the fact that

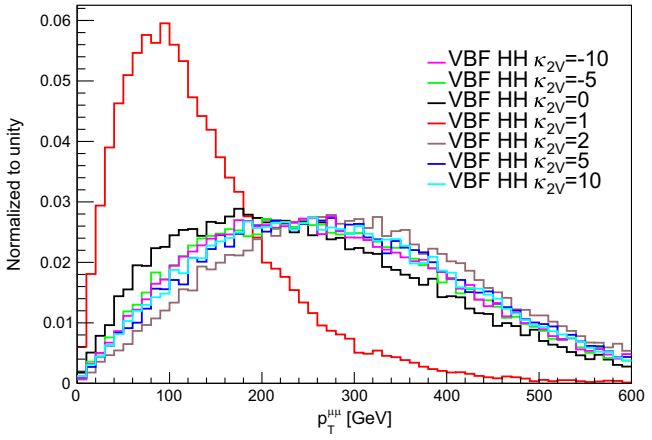


FIG. 5. The $p_{\mathrm{T}}^{\mu\mu}$ distribution of VBF HH signals with different $\kappa_{2\mathrm{V}}$ values.

DY totally dominates the background contribution. The second leading background is top quark pair production $t\bar{t}$. The samples are generated at LO in QCD using MG5_aMC@NLO2.6.5. Single Higgs processes that have a Higgs boson decaying to a pair of muons can also enter the signal regions. These samples are generated in the production modes of gluon-gluon fusion (ggH), vector boson fusion (VBFH), ZH, $t\bar{t}$H, and $b\bar{b}$H, using MG5_aMC@NLO2.6.5.

For all signal and background samples, the decay, parton shower, hadronization, and underlying events are modeled by PYTHIA8.306 [82]. No pileup is considered. To emulate the detector effects, DELPHES3.5.0 [83] is used, and the default CMS configuration card is applied. The jets are reconstructed using the anti-k_t clustering algorithm [84,85] with a radius parameter $R = 0.5$ following the CMS defaults. All the MC samples are summarized in Table I, in which the DY and $t\bar{t}$ cross section contains the branching ratios down to the final state with dimuon, while the single and double Higgs processes do not contain any branching ratio.

TABLE I. Summary of Monte Carlo samples. The first column gives the names of the processes. The second column provides the slicing on $m_{l^+l^-}$ whenever it is applicable. The third column lists the cross section. The last column presents the number of events in the generated MC samples. The DY and $t\bar{t}$ cross section contains the branching ratios down to the final state with dimuon, while the single and double Higgs processes do not contain any branching ratio.

Process	$m_{l^+l^-}$(GeV)	σ (fb)	$N_{\mathrm{events}}^{\mathrm{gen}}(\times 10^6)$
Drell-Yan	[100, 150]	5481	9.98
Drell-Yan	[150, 200]	384	10.0
Drell-Yan	$[200, +\infty]$	201	1.0
$t\bar{t}$	⋯	4864	2.0
ggH	⋯	48580	1.0
VBFH	⋯	3782	1.0
ZH	⋯	883.9	1.0
ttH	⋯	507.1	1.0
bbH	⋯	488	1.0
ggF signal			
$\kappa_\lambda = -5$	⋯	599	0.55
$\kappa_\lambda = 0$	⋯	70	0.55
$\kappa_\lambda = 1$	⋯	31	0.55
$\kappa_\lambda = 2.4$	⋯	13	0.55
$\kappa_\lambda = 5$	⋯	95	0.55
$\kappa_\lambda = 10$	⋯	672	0.55
$\kappa_\lambda = 20$	⋯	3486	0.55
VBF signal			
$\kappa_{2\mathrm{V}} = -10$	⋯	2365	0.50
$\kappa_{2\mathrm{V}} = -5$	⋯	722	0.50
$\kappa_{2\mathrm{V}} = 0$	⋯	27	0.50
$\kappa_{2\mathrm{V}} = 1$	⋯	1.73	0.50
$\kappa_{2\mathrm{V}} = 2$	⋯	14.2	0.50
$\kappa_{2\mathrm{V}} = 5$	⋯	279	0.50
$\kappa_{2\mathrm{V}} = 10$	⋯	1479	0.50

III. CUT-BASED ANALYSIS

A. Analysis strategy

Two analysis strategies are adopted. One using sequential cuts inspired by recent ATLAS and CMS analyses [53,54] serves as a baseline strategy to understand the basic kinematics and get a conservative estimation of the sensitivity, while the other applies the BDTs to seek further improvements in the sensitivity. The latter one is only studied for the ggF HH, as there is insufficient statistics in the VBF enriched regions.

B. Object and basic event selection

Basic acceptance requirements following a CMS-like detector are applied in the physics object selections. For muon candidates, they must satisfy the requirements of $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$ following Ref. [73]. Jets are reconstructed using the anti-k_t clustering algorithm with a radius parameter $R = 0.5$. All jet candidates have to pass $p_{\mathrm{T}} > 20$ GeV within $|\eta| < 4.7$, while the bjets have to be within $|\eta| < 2.4$ limited by the tracker geometry. The bjets are tagged with an parametrized efficiency depending on p_{T} and η mimicking the CMS scenario.

Each event is required to contain at least two opposite charged muons. When more than two muon candidates are found, the muon pair with the highest transverse momentum $p_{\mathrm{T}}^{\mu\mu}$ is chosen to reconstruct the Higgs boson candidate. Moreover, all the events have to sit in $100 < m_{\mu\mu} < 180$ GeV. Next, events are required to contain at least two bjets. In case of more than two bjets, the Higgs boson candidate is reconstructed from the two jets with the highest transverse momentum p_{T}^{bb}. In the end, the invariant mass of two bjets is required to be $70 < m_{bb} < 190$ GeV. Both the choices of the dimuon and dibjet mass range are inspired by Ref. [54]. In the low $m_{\mu\mu}$ or m_{bb} end, the statistics is much higher than in the high end. Thus, the intervals are skewed toward higher mass ranges.

Besides the two bjets, the signal events could have more jets in the VBF HH production. It is featured by the presence of two additional energetic jets (VBF jets), corresponding to two quarks from each of the colliding protons scattered away from the beam line. These VBF jets are expected to have a large spacial separation, which results in a large dijet invariant mass, m_{jj}^{VBF}. The jet pair with the highest m_{jj}^{VBF} in an event is selected as the two VBF jets.

C. Event categorization and background rejection

The events are first grouped by the signal production modes ggF and VBF. In each group, events are then categorized according to the couplings variations. After categorization, the sequential cuts are optimized in the cut-based analysis, and the training of BDT is performed in the machine-learning analysis, both for suppressing the corresponding background in each category. Eventually, the fits are performed on the combined $m_{\mu\mu}$ and m_{bb} distributions.

In the separation of ggF and VBF modes, events that do not have two or more additional jets enter the ggF categories directly. With two or more additional jets, events can be categorized by the invariant mass of the two VBF

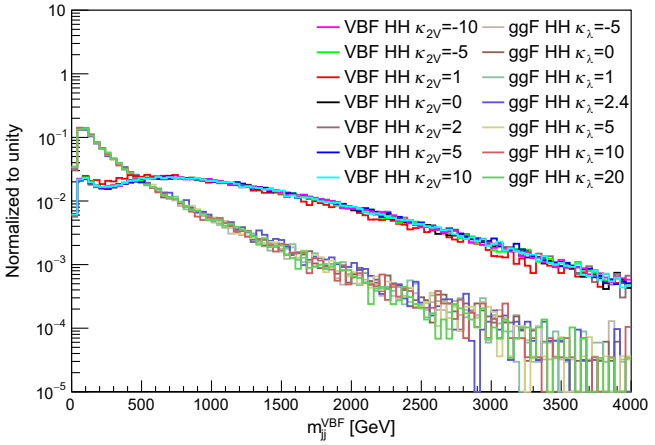


FIG. 6. The m_{jj}^{VBF} distribution of all the signals.

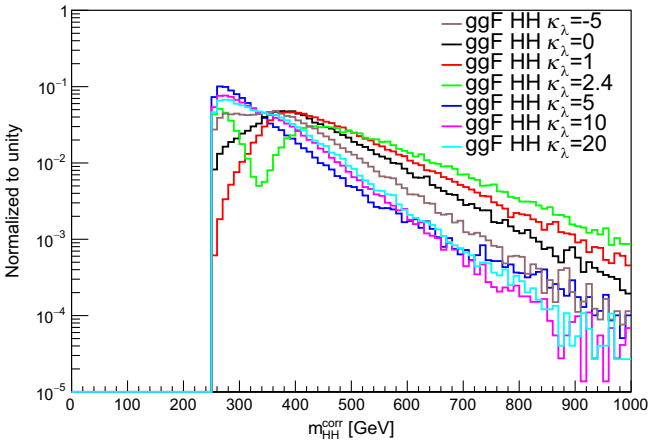


FIG. 7. The $m_{\mathrm{HH}}^{\mathrm{corr}}$ distribution of ggF signals with different κ_λ values.

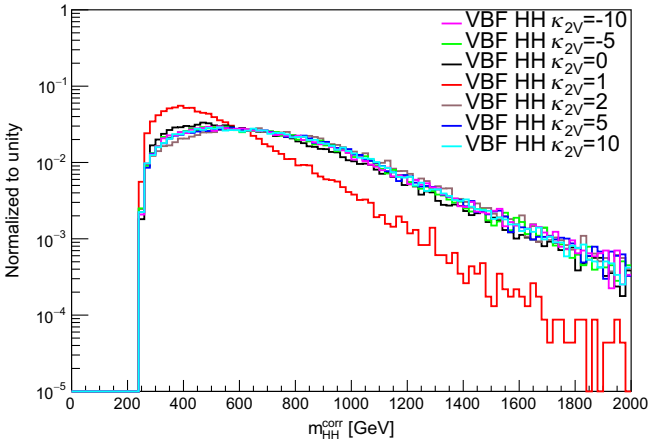
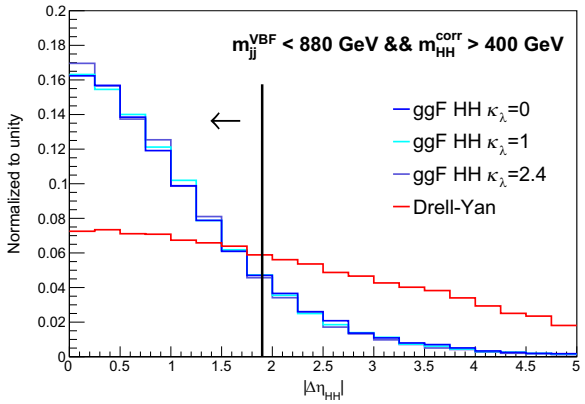


FIG. 8. The $m_{\mathrm{HH}}^{\mathrm{corr}}$ distribution of VBF signals with different κ_{2V} values.

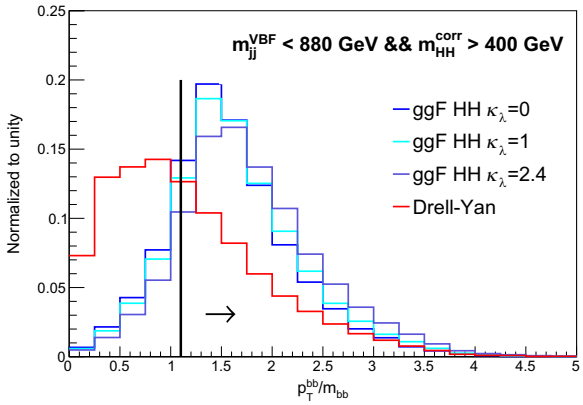
TABLE II. Summary of the event categorization.

Category	m_{jj}^{VBF}(GeV)	$m_{\mathrm{HH}}^{\mathrm{corr}}$(GeV)
ggF SM	<880	>400
ggF BSM	<880	<400
VBF SM	>880	<680
VBF BSM	>880	>680

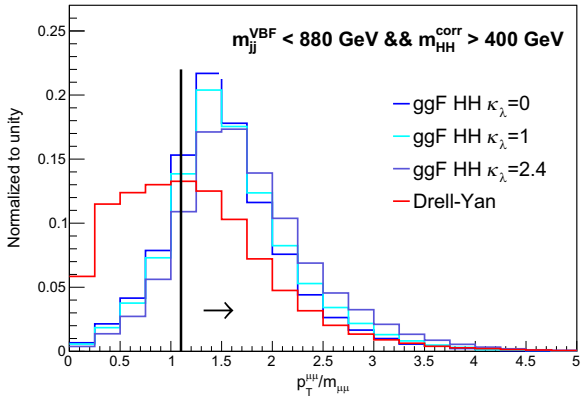
jets m_{jj}^{VBF}, as shown in Fig. 6. The VBF events clearly have a much harder spectrum of m_{jj}^{VBF} with respect to the ggF ones due to the VBF jets that fly out in a large spacial separation, and this feature does not depend either of the couplings κ_λ and $\kappa_{2\mathrm{V}}$. A scan on m_{jj}^{VBF} is performed to maximize the separation of the two production modes. The scan that maximizes the separation determines the threshold $m_{jj}^{\mathrm{VBF}} = 880$ GeV which the ggF category is



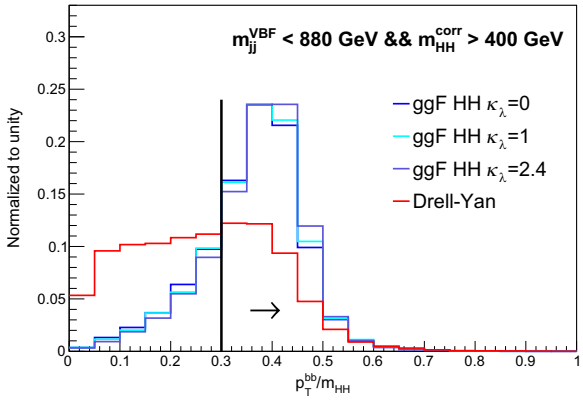
(a) The $|\Delta\eta_{\mathrm{HH}}|$ distribution in the ggF SM category.



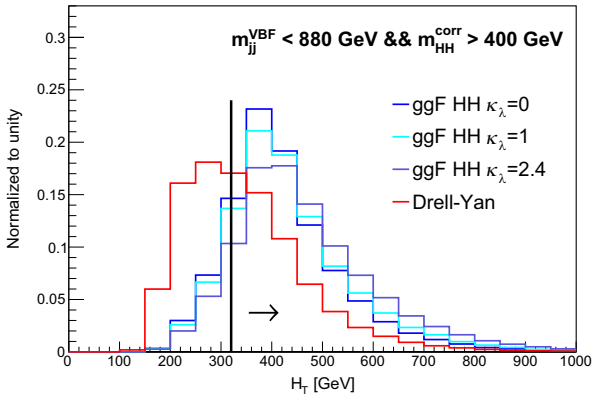
(b) The $p_{\mathrm{T}}^{bb}/m_{bb}$ distribution in the ggF SM category.



(c) The $p_{\mathrm{T}}^{\mu\mu}/m_{\mu\mu}$ distribution in the ggF SM category.



(d) The $p_{\mathrm{T}}^{bb}/m_{\mathrm{HH}}$ distribution in the ggF SM category.



(e) The H_{T} distribution in the ggF SM category.

FIG. 9. The discriminating variables chosen in the ggF SM category.

defined below and VBF is defined above. Other variables, such as the difference in pseudorapidity between the two VBF jets $|\Delta\eta_{jj}^{\mathrm{VBF}}|$, are also tried but do not show any significant improvement once m_{jj}^{VBF} is used.

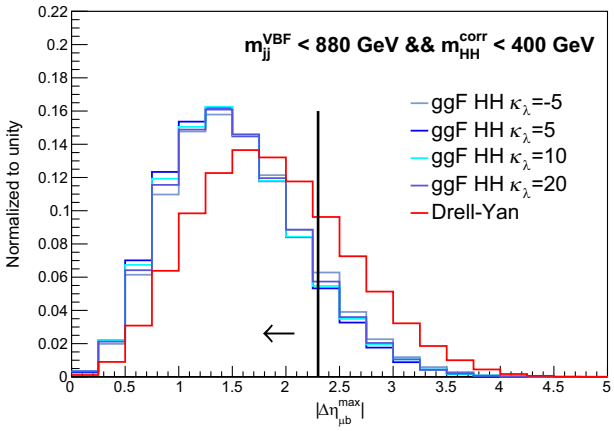
In maximizing the sensitivity on the couplings, two categories are defined in each of the ggF and VBF categories. The variable $m_{\mathrm{HH}}^{\mathrm{corr}}$ as defined in Eq. (1),

$$m_{\mathrm{HH}}^{\mathrm{corr}} = m_{\mathrm{HH}} - (m_{\mu\mu} - 125) - (m_{bb} - 125), \qquad (1)$$

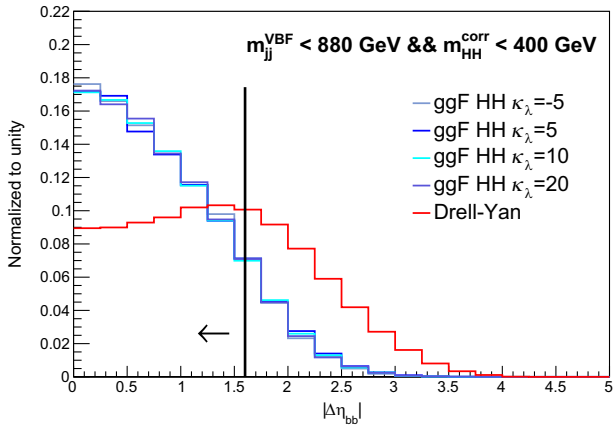
following Ref. [56] is often used in experiments, and it serves as a good proxy of m_{HH} reflecting the κ_λ variation, comparing Figs. 2 and 7. It reduces the systematic uncertainties from the energy scale and resolution than directly using the bare m_{HH}. Maximizing the separation between SM and BSM couplings, a threshold on $m_{\mathrm{HH}}^{\mathrm{corr}}$ is determined to divide the ggF events into the ggF SM category with $m_{\mathrm{HH}}^{\mathrm{corr}} > 400$ GeV for a harder $m_{\mathrm{HH}}^{\mathrm{corr}}$ spectrum largely contributed by the box diagram and the ggF BSM category with $m_{\mathrm{HH}}^{\mathrm{corr}} < 400$ GeV for the less energetic events contributed mostly from the triangle diagram that contains the self-coupling κ_λ. In the VBF category, a further split is optimized for $\kappa_{2\mathrm{V}}$. The $m_{\mathrm{HH}}^{\mathrm{corr}}$ varies rapidly when $\kappa_{2\mathrm{V}}$ starts to deviate from its SM value, making a clear difference between the SM and the BSM cases, as shown in Fig. 8. Maximizing the separation, the threshold $m_{\mathrm{HH}}^{\mathrm{corr}} < 680$ GeV selects events into the VBF SM category, and the opposite defines the VBF BSM category. The thresholds used in the categorization are summarized in Table II.

The categorization based on signal kinematic variations due to the couplings are shared in the cut-based and the BDT analyses. After that, the background amounts and kinematics differ per category. Thus, dedicated cut-based and BDT approaches are optimized or trained separately. In the rest of this section, we focus on the cut-based analysis. The dominant background DY is the main target to suppress. The event selections are optimized by maximizing the expected significance $S/\sqrt{B}$ for each category.

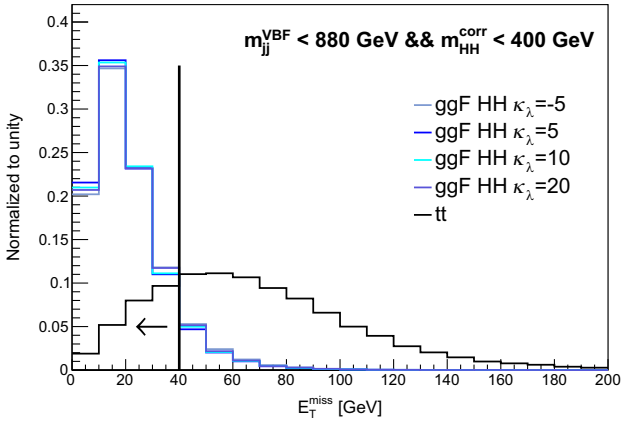
In the ggF SM category, the following discriminating variables are chosen. The $|\Delta\eta_{\mathrm{HH}}|$ variable stands for the absolute value of the η separation between two Higgs bosons. The signal events tend to be more transverse, resulting in smaller $|\Delta\eta_{\mathrm{HH}}|$, while the DY background events are less transverse, leading to larger $|\Delta\eta_{\mathrm{HH}}|$, as



(a) The $|\Delta\eta_{\mu b}^{max}|$ distribution in the ggF BSM category.

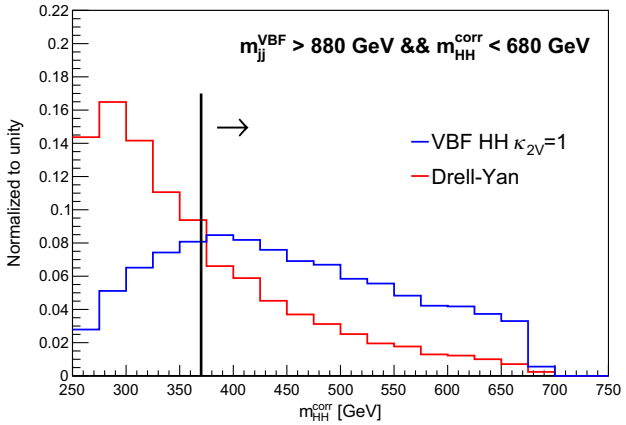


(b) The $|\Delta\eta_{bb}|$ distribution in the ggF BSM category.

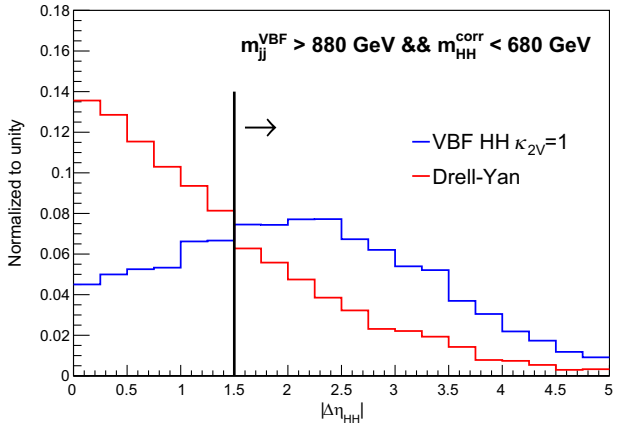


(c) The E_{T}^{miss} distribution in the ggF BSM category.

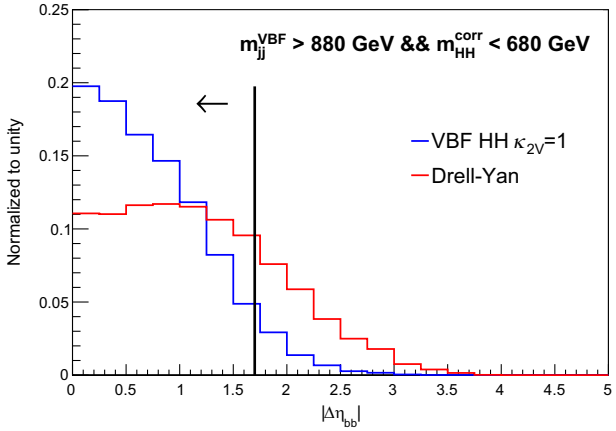
FIG. 10. The discriminating variables chosen in the ggF BSM category.



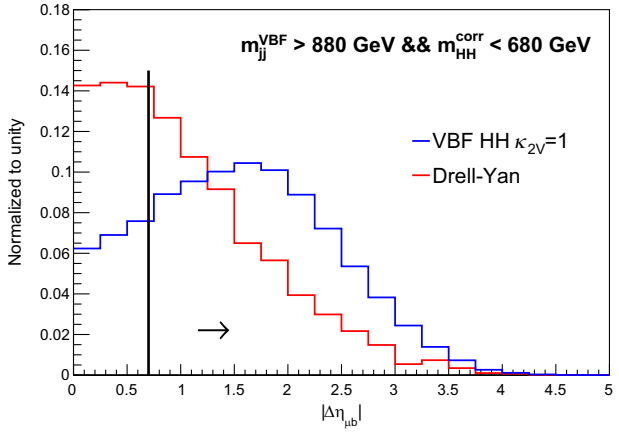
(a) The $m_{\mathrm{HH}}^{\mathrm{corr}}$ distribution in the VBF SM category.



(b) The $|\Delta\eta_{\mathrm{HH}}|$ distribution in the VBF SM category.



(c) The $|\Delta\eta_{bb}|$ distribution in the VBF SM category.



(d) The $p_{\mathrm{T}}^{bb}/m_{\mathrm{HH}}$ distribution in the ggF SM category.

FIG. 11. The discriminating variables chosen in the VBF SM category.

shown in Fig. 9(a). The relative p_{T} variables, such as $p_{\mathrm{T}}^{bb}/m_{bb}$, $p_{\mathrm{T}}^{\mu\mu}/m_{\mu\mu}$, and $p_{\mathrm{T}}^{bb}/m_{\mathrm{HH}}$, where p_{T}^{bb} and $p_{\mathrm{T}}^{\mu\mu}$ are the transverse momenta of the dibjet and dimuon candidates, can also effectively separate the signal and the DY background events. As shown in Figs. 9(b)–9(d), the signal events tend to have harder spectrum. Lastly, H_{T}, which represents the scalar sum of the transverse momentum of bjets and muons, has a strong separation power given more energetic signal events, as shown in Fig. 9(e).

In the ggF BSM category, less energetic events are enriched, resulting in very similar behavior in signal and background events. In this challenging phase space, we find two outstanding variables that can improve the signal significance by calculating $\mathrm{S}/\sqrt{\mathrm{B}}$: $|\Delta\eta_{\mu b}^{\mathrm{max}}|$, which is the absolute value of the maximal η separation between the muons and the bjets, and $|\Delta\eta_{bb}|$, which stands for the absolute value of the η separation between two bjets. As shown in Fig. 10(a), the signal events are more centrally produced, leading to slightly smaller $|\Delta\eta_{\mu b}^{\mathrm{max}}|$ than the background. In Fig. 10(b), the signal events have two bjets originating from the Higgs boson, resulting in smaller $|\Delta\eta_{bb}|$ than the background. Different than other categories, there emerge non-negligible $t\bar{t}$ background events. The missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$ as a proxy of the neutrino p_{T}, effectively rejects the $t\bar{t}$ events, as shown in Fig. 10(c).

In the VBF SM category, the statistics is in general low. We find four variables that improve the significance in a

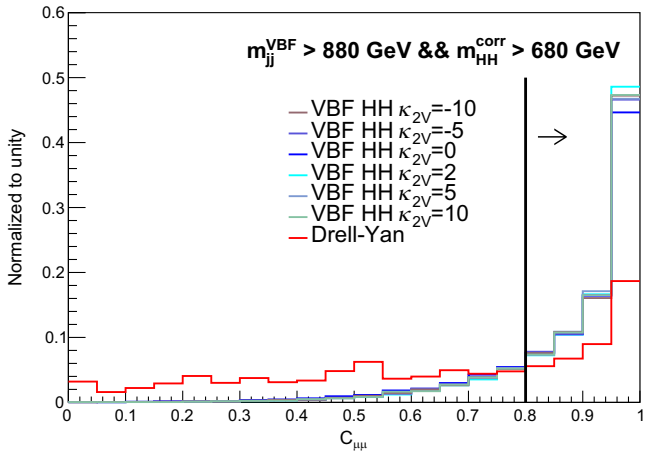


FIG. 12. The $C_{(H\to\mu\mu)}$ distribution in the VBF BSM category.

TABLE III. Summary of the optimized cuts for background suppression and the corresponding efficiencies ϵ in all the four categories. The efficiencies are calculated with the number of events passing the full set of cuts over the number of events that enter the category. For the signal efficiencies, only the relevant signals in the corresponding categories are listed.

Category	Variable cut	ϵ_{Signal}	$\epsilon_{\text{DY+tt}}$
ggF SM	$\lvert\Delta\eta_{\text{HH}}\rvert < 1.9$ $p_{\text{T}}^{bb}/m_{bb} > 1.1$ $p_{\text{T}}^{\mu\mu}/m_{\mu\mu} > 1.1$ $p_{\text{T}}^{bb}/m_{\text{HH}} > 0.3$ $H_{\text{T}} > 320$ GeV	53% ($\kappa_\lambda = 1$)	11%
ggF BSM	$\lvert\Delta\eta_{\mu b}^{\max}\rvert < 2.3$ $\lvert\Delta\eta_{bb}\rvert < 1.6$ $E_{\text{T}}^{\text{miss}} < 40$ GeV	55% ($\kappa_\lambda = 5$)	13%
VBF SM	$m_{\text{HH}}^{\text{corr}} > 370$ GeV $\lvert\Delta\eta_{\text{HH}}\rvert > 1.5$ $\lvert\Delta\eta_{bb}\rvert < 1.7$ $\lvert\Delta\eta_{\mu b}\rvert > 0.7$	39% ($\kappa_{2\text{V}} = 1$)	3%
VBF BSM	$C_{\mu\mu} > 0.8$	69% ($\kappa_{2\text{V}} = 10$)	15%

sizeable way by calculating $\text{S}/\sqrt{\text{B}}$: $m_{\text{HH}}^{\text{corr}}$, $|\Delta\eta_{\text{HH}}|$, $|\Delta\eta_{bb}|$, and $|\Delta\eta_{\mu b}|$. Their distributions are shown in Fig. 11(a)–11(d). The VBF BSM category has the lowest statistics among all due to $m_{\text{HH}}^{\text{corr}} > 680$ GeV, leaving little room for optimization. Only one variable, the centrality variable $C_{\mu\mu}$, defined in Eq. (2) following Ref. [54], is chosen, as shown in Fig. 12. The signal events tend to have larger η separation of the two VBF jets, resulting in $C_{\mu\mu}$ more close to 1 than the background events,

TABLE IV. The event yields of signal and background processes in the four categories after background suppression, assuming an integrated luminosity of 3000 fb^{-1}.

	Category			
Process	ggF SM	ggF BSM	VBF BSM	VBF SM
ggF HH signal				
$\kappa_\lambda = 1$	1.39	0.65	0.01	0.02
$\kappa_\lambda = 5$	0.71	4.27	0.004	0.01
VBF HH signal				
$\kappa_{2\text{V}} = 1$	0.01	0.02	0.004	0.02
$\kappa_{2\text{V}} = 10$	52.5	9.24	27.9	5.61
Background				
Drell-Yan	7976	54369	94.3	713
$t\bar{t}$	832	24619	0	175
ggH	1.25	1.32	0.07	0.07
VBFH	0.08	0.21	0	0.04
ZH	0.74	1.32	0	0.01
ttH	1.32	2.93	0.03	0.08
bbH	0.02	0.18	0	0
Total background	8811.41	78993.96	94.40	888.20

$$C_{\mu\mu} = e^{\left[-\frac{4}{(\eta_1^{\text{VBF}}-\eta_2^{\text{VBF}})^2}(\eta^{\mu\mu}-\frac{\eta_1^{\text{VBF}}+\eta_2^{\text{VBF}}}{2})^2\right]}, \quad (2)$$

where η_1^{VBF} and η_2^{VBF} are the pseudorapidities of the two VBF jets.

Table III summarizes the optimized cuts for background suppression and the corresponding efficiencies in all the four categories, while Table IV lists the event yields of the signal and background processes, assuming an integrated luminosity of 3000 fb^{-1}. We also consider other HH processes, HH $\to b\bar{b}$WW and $b\bar{b}\tau^+\tau^-$, where both can lead to $b\bar{b}\mu^+\mu^- + E_{\text{T}}^{\text{miss}}$ final states. However, the aforementioned processes' contributions to the background are negligible due to relatively smaller dimuon invariant mass than that for the signal.

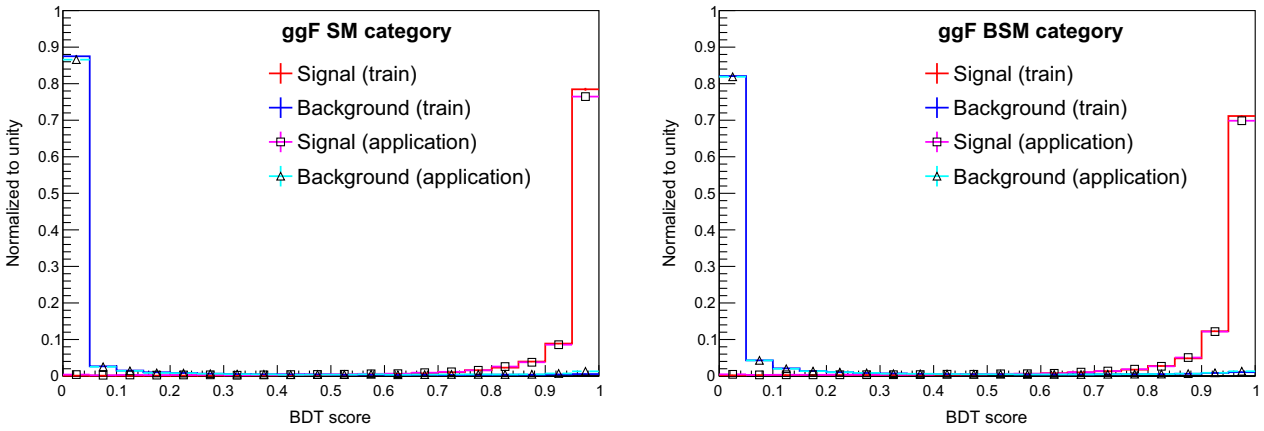
In the boosted regime where the two b quarks cannot be resolved in two separate small-R jets, merged jets are reconstructed using $R = 0.8$ instead. In this case, the final state includes a large-R jet and two muons. However, this region has too low statistics and ends up with negligible contribution with respect to the analysis using resolved b quarks.

IV. BDT ANALYSIS

In the ggF SM and BSM categories, the sufficient statistics allows to apply machine-learning algorithms for

TABLE V. Summary of input variables for the BDT training in the two ggF categories. Besides the variables that are already explained in the texts, $|\Delta\eta_{\mu b}^{\max}|$ is the maximal $|\Delta\eta|$ between muons and bjets, while $|\Delta\eta_{\mu b}^{\text{other}}|$ is for the other muon and bjet. $|\Delta R_{\mu b}^{\min}|$ and $|\Delta R_{\mu b}^{\text{other}}|$ are defined accordingly.

	Category	
Input variable	ggF SM	ggF BSM
$p_{\text{T}}^{\mu 1}$, $p_{\text{T}}^{\mu 2}$, p_{T}^{b1}, p_{T}^{b2}	✓	✓
$E_{\mu 1}$, $E_{\mu 2}$, E_{b1}, E_{b2}	✓	
$\eta^{\mu 1}$, $\eta^{\mu 2}$	✓	
η^{b1}, η^{b2}	✓	✓
η_{j1}^{VBF}		✓
$E_{\mu\mu}$, E_{bb}, $\eta_{\mu\mu}$, η_{bb}, $\cos\theta_{\mu\mu}$, $\cos\theta_{bb}$	✓	
$p_{\text{T}}^{\mu\mu}$, p_{T}^{bb}, $m_{\mu\mu}$, m_{bb}	✓	✓
m_{HH}, $m_{\text{HH}}^{\text{corr}}$	✓	
p_{T}^{b1}/m_{bb}, p_{T}^{b2}/m_{bb}, p_{T}^{bb}/m_{bb}, $p_{\text{T}}^{\mu 1}/m_{\mu\mu}$, $p_{\text{T}}^{\mu 2}/m_{\mu\mu}$	✓	
p_{T}^{bb}/m_{bb}, $p_{\text{T}}^{bb}/m_{\text{HH}}$, $p_{\text{T}}^{\mu\mu}/m_{\mu\mu}$, $p_{\text{T}}^{\mu\mu}/m_{\text{HH}}$	✓	✓
H_{T}, p_{T}^{HH}, $p_{\text{T}}^{\mu\mu}/p_{\text{T}}^{bb}$	✓	✓
$E_{\text{T}}^{\text{miss}}$, η^{miss}	✓	✓
$\lvert\Delta\eta_{\text{HH}}\rvert$, $\lvert\Delta\eta_{\mu b}\rvert$, $\lvert\Delta\eta_{\mu b}^{\max}\rvert$, $\lvert\Delta\eta_{\mu b}^{\text{other}}\rvert$	✓	✓
$\lvert\Delta\eta_{bb}\rvert$, $\lvert\Delta\eta_{\mu\mu}\rvert$	✓	✓
$\lvert\Delta R_{\text{HH}}\rvert$, $\lvert\Delta R_{\mu b}\rvert$, $\lvert\Delta R_{bb}\rvert$, $\lvert\Delta R_{\mu\mu}\rvert$	✓	✓
$\lvert\Delta R_{\mu b}^{\min}\rvert$, $\lvert\Delta R_{\mu b}^{\text{other}}\rvert$, $\lvert\Delta R_{jj}^{\text{VBF}}\rvert$	✓	
$\lvert\Delta\phi_{\text{HH}}\rvert$, $\lvert\Delta\phi_{\mu b}\rvert$, $\lvert\Delta\phi_{bb}\rvert$, $\lvert\Delta\phi_{jj}^{\text{VBF}}\rvert$	✓	✓
$\lvert\Delta\phi_{\mu\mu}\rvert$	✓	



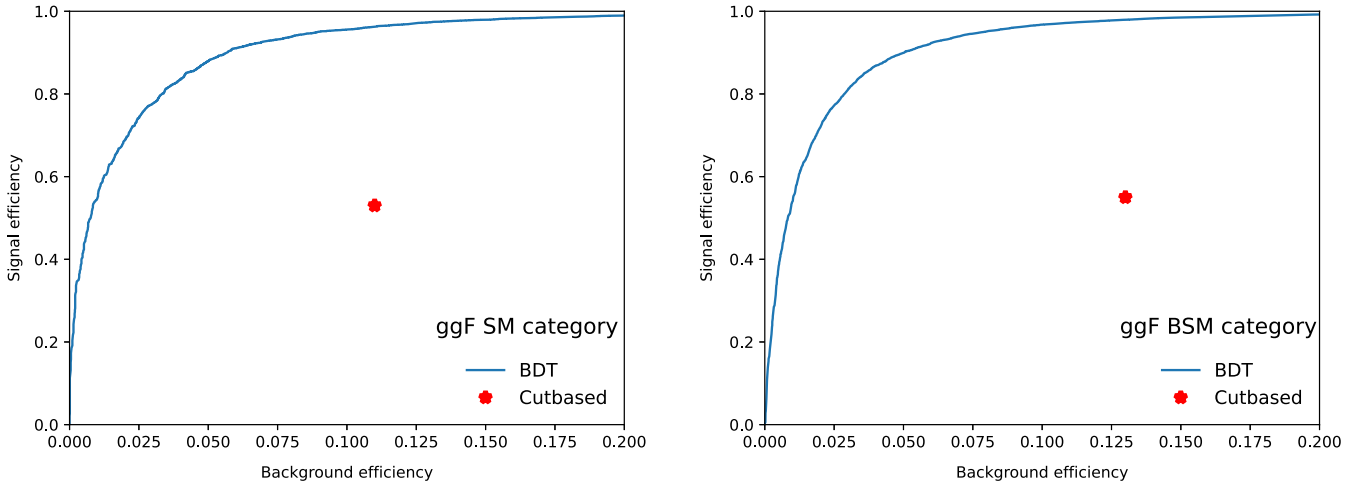
(a) The ggF SM category (b) The ggF BSM category

FIG. 13. The BDT score distributions in the ggF SM and BSM categories.

further improvement in the sensitivity. The BDT algorithm is adopted using the package of XGBoost [86].

The general strategy is to train a huge amount of shallow decision trees and extract a strong separation power from the tree ensemble. The training setup includes 2500 trees, the tree depth of 3 and a learning rate of 0.08 (0.1) for ggF SM (BSM) category. The shallow trees with only a depth of 3 are proven to be an effective way of avoiding overtraining. The MC samples are split into 64%, 16%, and 20% for training, testing, and application and correspond to the numbers of events of 101,000 (335,000), 25,000 (84,000), and 32,000 (104,000) for the ggF SM (BSM) category. The ratio of signal, DY and $t\bar{t}$ events in all samples is 40:60:1 (4:7:1) for the ggF SM (BSM) category. The proportion of $t\bar{t}$ in the training is low in the ggF SM category given its small contribution. The input variables in the training are listed in Table V. The training is performed separately in the ggF SM and BSM categories. The signal sample generated with $\kappa_\lambda = 1$ is used in the training for the ggF SM category, while the signal samples generated with $\kappa_\lambda = 5$, 10, and 20 are used in the training for the ggF BSM category. Both the DY and $t\bar{t}$ processes are used as the background in the training in the two categories.

The BDT score distributions of the training and testing samples are compared, and good agreement is found, which suggests no overtraining issue. The BDT score distributions are then shown with the training and the application samples, the latter of which are used for the inference in the analysis, in Figs. 13(a) and 13(b) for the ggF SM and



(a) The ggF SM category (b) The ggF BSM category

FIG. 14. The ROC curve in the training of the ggF SM and BSM categories. The red dots present the performance of the cut-based approach.

BSM categories, respectively. Still, there is no overtraining issue found.

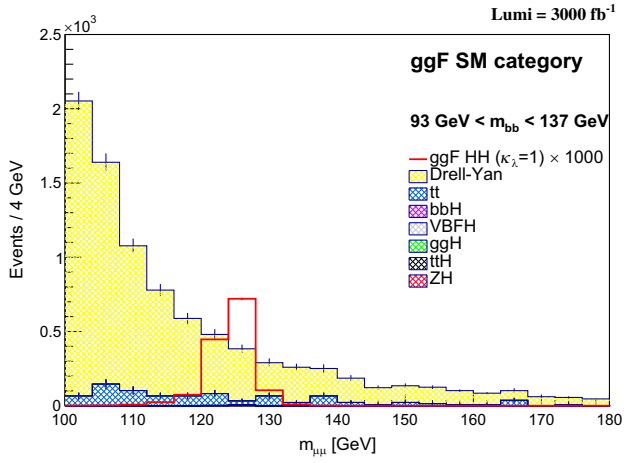
A cut on the BDT score is applied to purify the signal. The threshold is chosen to keep the signal efficiency roughly equal to the one in the cut-based analysis, which is around 50% as listed in Table III. With this threshold, the background efficiency is very suppressed down to 0.85% and 0.94% for the ggF SM and BSM categories, respectively, 1 order of magnitude smaller than the ones from the cut-based analysis as listed in Table III. The improvement by BDT is also visualized in the ROC curves (receiver operating characteristic curve) in Figs. 14(a) and 14(b). The cut-based performance is shown as the red stars for comparisons. The events with the BDT score above the threshold are used in the final fits described in the next section.

V. RESULTS

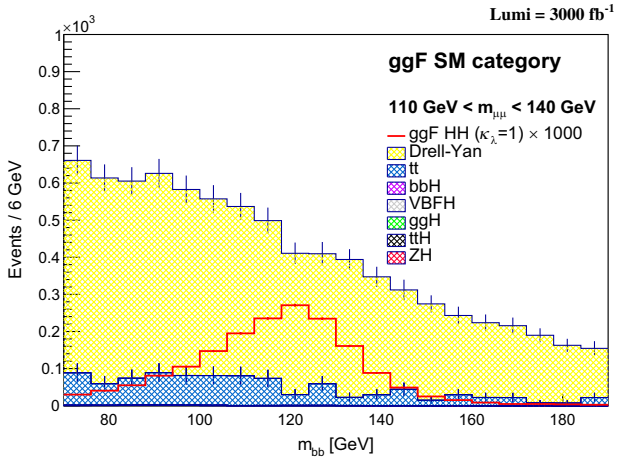
All events are used in searching for the signal combining the four categories after the background suppression, using either the cut-based or the BDT approach. The fits are performed in the ranges of $100 < m_{\mu\mu} < 180$ GeV and $70 < m_{bb} < 190$ GeV in the four categories. The fitting templates are the combined dimuon mass $m_{\mu\mu}$ and dibjet mass m_{bb} distributions as shown in Figs. 15(a)–18(b) for all the four categories with the cut-based method and in Figs. 19(a)–20(b) for the two ggF categories with the BDT approach.

With a integrated luminosity of 3000 fb^{-1}, the search with HH $\to b\bar{b}\mu^+\mu^-$ does not reach the threshold for a discovery. The upper limits on the cross sections at 95% confidence level (CL) are extracted using the modified frequentist CL$_s$ approach [87,88] in which the asymptotic approximation [89] is applied. For the SM ggF and VBF HH cross section, their upper limits at 95% CL are presented in the unit of their SM cross section as shown in Table VI for the integrated luminosities of 300 fb^{-1} from the LHC Run3 only, 450 fb^{-1} from the combined Run2 and Run3, and 3000 fb^{-1} from the full runs in the LHC and the HL-LHC. The expected upper limit at 95% CL on the ggF HH prediction is 47 (28) times that of its SM prediction using the cut-based (BDT) approach with the full integrated luminosity. The expected upper limit at 95% CL on the VBF HH prediction is 928 times that of its SM prediction using the cut-based approach with the full integrated luminosity. The expected upper limits are also set for the combined ggF and VBF HH processes assuming their SM cross sections. Given the small VBF rate, its contribution to the upper limits of the combined ggF and VBF HH is marginal. All results are extracted assuming $m_{\mathrm{H}} = 125$ GeV that is close to the most precise measurement of the Higgs boson mass to date $m_{\mathrm{H}} = 125.38 \pm 0.14$ GeV [90].

Benefiting from the four categories that are defined to maximize the sensitivity not only at the SM coupling but also to the BSM, the upper limits at 95% CL are scanned along κ_λ and $\kappa_{2\mathrm{V}}$, assuming the top-quark Yukawa coupling and the HVV coupling SM-like ($\kappa_{\mathrm{t}} = 1$ and $\kappa_{\mathrm{V}} = 1$). The scan on κ_λ is shown in Fig. 21(a) with the individual contributions from each category and the combination and in Fig. 21(b) with the combined results, using the cut-based approach. The green and yellow bands surrounding the upper limit median represent its 68% and 95% CL uncertainty. The red solid curve with its band represents the theoretical prediction and the corresponding uncertainty [91]. At 95% CL, the expected constraint on κ_λ is $-13.8 < \kappa_\lambda < 19.1$ with the full integrated luminosity. The same scan is performed using the BDT approach as shown in Fig. 23(a) for the breakdown and Fig. 23(b) for the combination. At 95% CL, the expected constraint on κ_λ using the BDT approach is $-10.0 < \kappa_\lambda < 15.5$ with the full integrated luminosity.

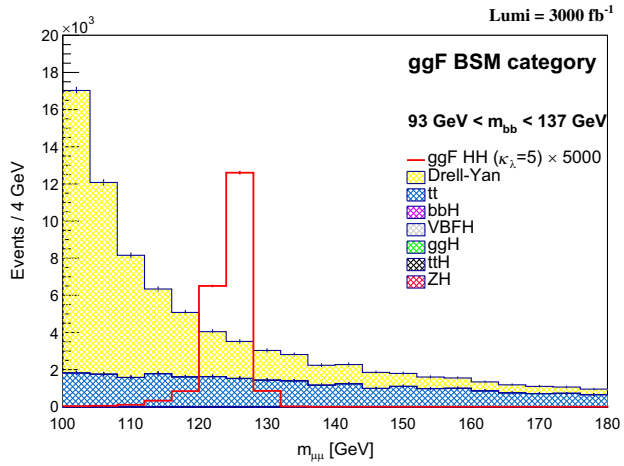


(a) The distribution of $m_{\mu\mu}$ in the ggF SM category.

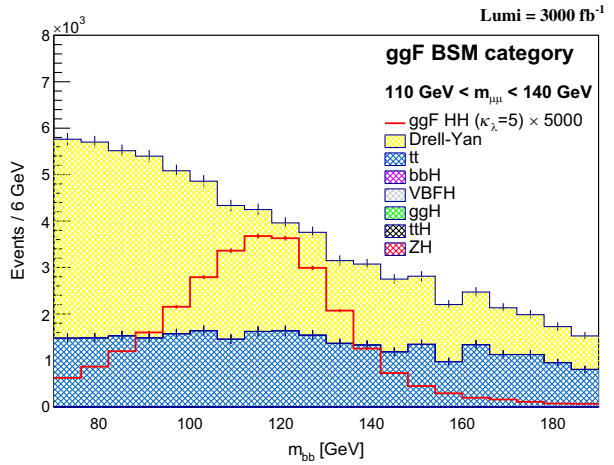


(b) The distribution of m_{bb} in the ggF SM category.

FIG. 15. The distributions of $m_{\mu\mu}$ and m_{bb} in the ggF SM category from the cut-based analysis.

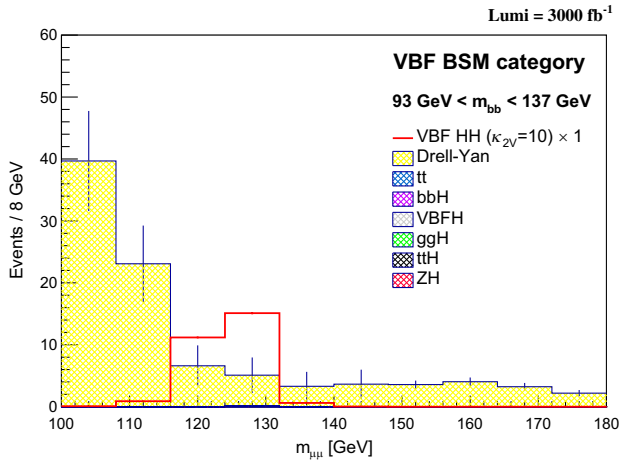


(a) The distribution of $m_{\mu\mu}$ in the ggF BSM category.

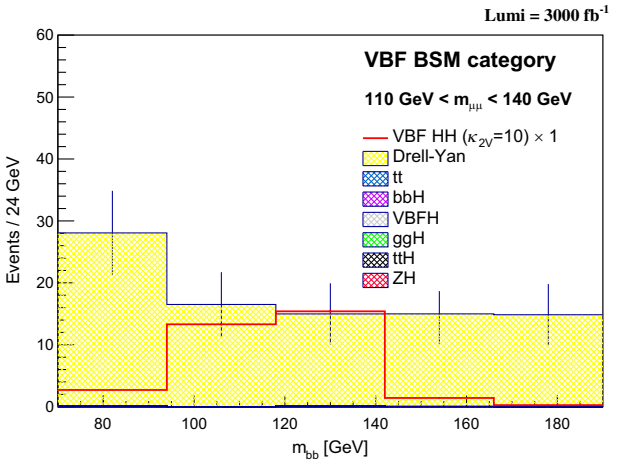


(b) The distribution of m_{bb} in the ggF BSM category.

FIG. 16. The distributions of $m_{\mu\mu}$ and m_{bb} in the ggF BSM category from the cut-based analysis.

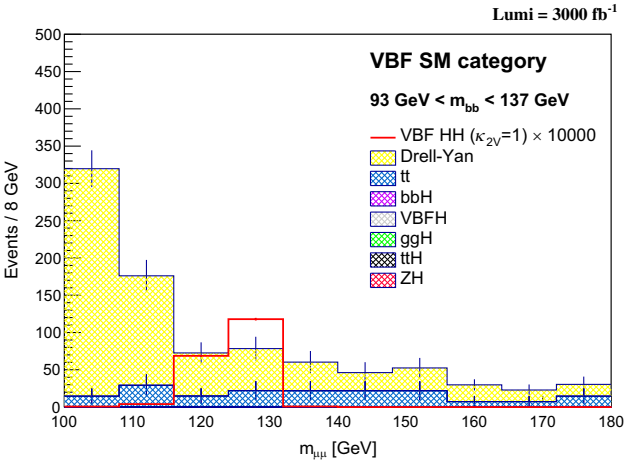


(a) The distribution of $m_{\mu\mu}$ in the VBF BSM category.

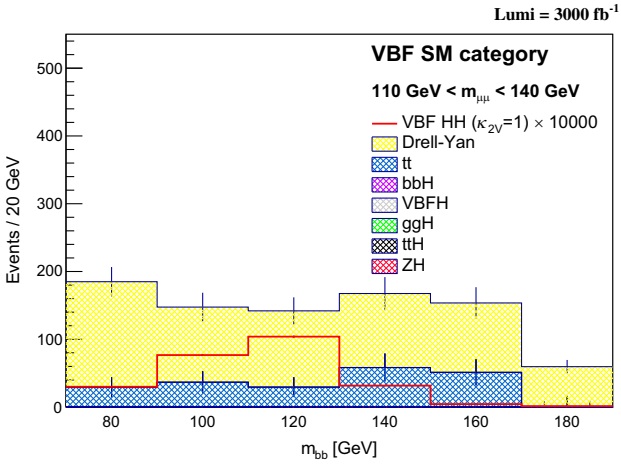


(b) The distribution of m_{bb} in the VBF BSM category.

FIG. 17. The distributions of $m_{\mu\mu}$ and m_{bb} in the VBF BSM category from the cut-based analysis.

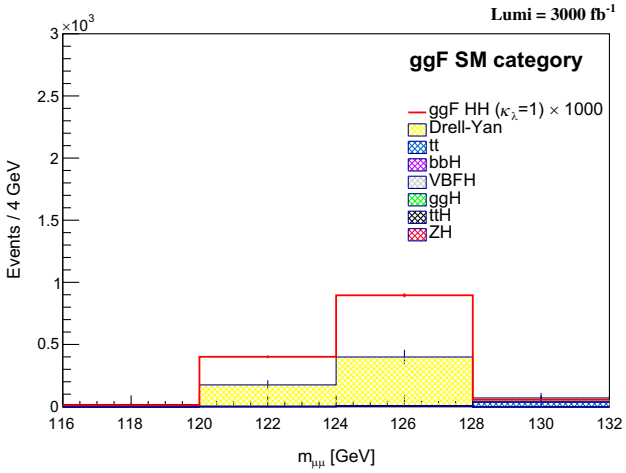


(a) The distribution of $m_{\mu\mu}$ in the VBF SM category.

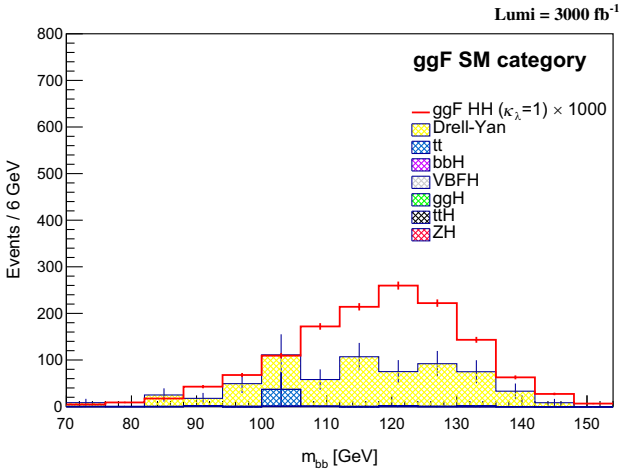


(b) The distribution of m_{bb} in the VBF SM category.

FIG. 18. The distributions of $m_{\mu\mu}$ and m_{bb} in the VBF SM category from the cut-based analysis.

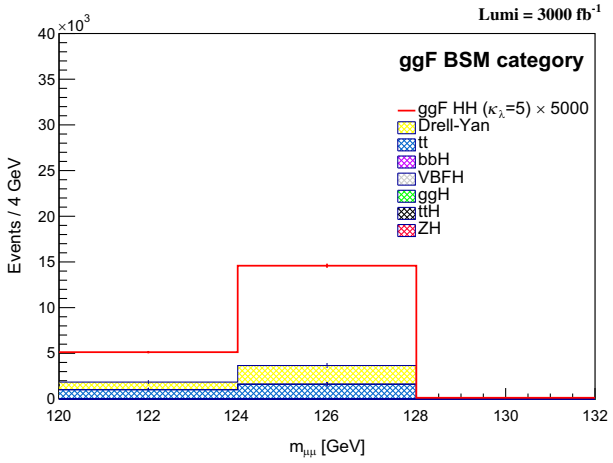


(a) The distribution of $m_{\mu\mu}$ in the ggF SM category.

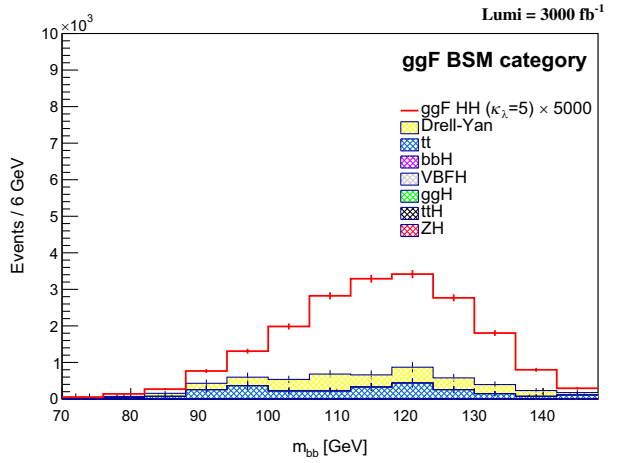


(b) The distribution of m_{bb} in the ggF SM category.

FIG. 19. The distributions of $m_{\mu\mu}$ and m_{bb} in the ggF SM category from the BDT analysis.



(a) The distribution of $m_{\mu\mu}$ in the ggF BSM category.

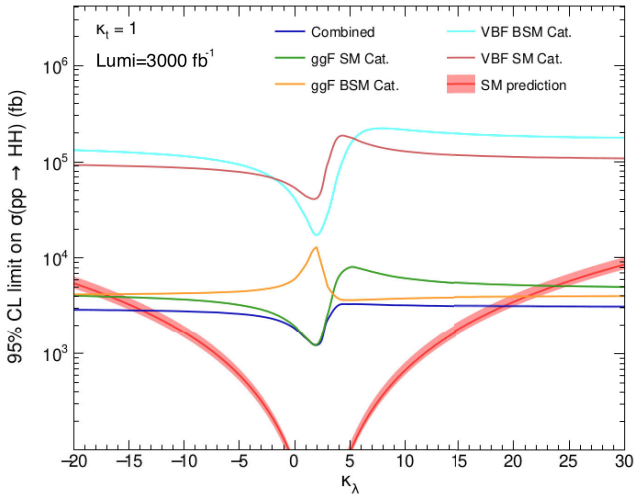


(b) The distribution of m_{bb} in the ggF BSM category.

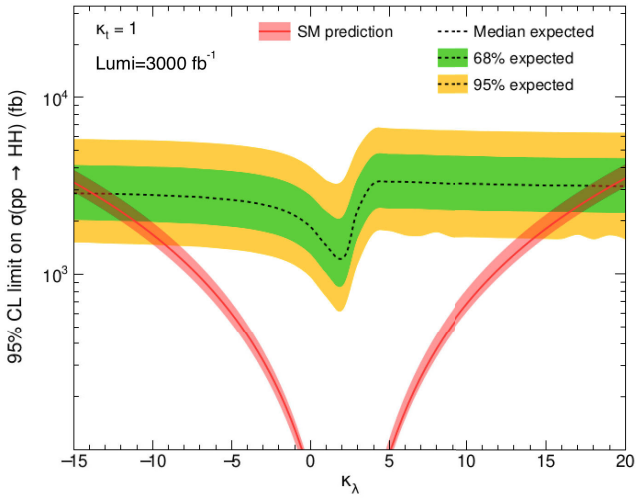
FIG. 20. The distributions of $m_{\mu\mu}$ and m_{bb} in the ggF BSM category from the BDT analysis.

TABLE VI. The upper limits at 95% CL of the ggF and VBF HH cross section in the cut-based and BDT analyses.

Analysis type	300 fb^{-1}	450 fb^{-1}	3000 fb^{-1}
		ggF HH ($\sigma/\sigma_{\mathrm{SM}}$)	
Cut based	152^{+87}_{-46}	123^{+70}_{-37}	$47^{+26.1}_{-14.1}$
BDT	$96^{+56}_{-29.8}$	$77^{+45}_{-23.9}$	$28^{+16.3}_{-8.8}$
		ggF + VBF HH ($\sigma/\sigma_{\mathrm{SM(ggF+VBF)}}$)	
Cut based	152^{+86}_{-46}	122^{+70}_{-37}	$46^{+26.1}_{-13.9}$
BDT	$96^{+56}_{-29.7}$	$77^{+45}_{-23.9}$	$28^{+16.2}_{-8.8}$
		VBF HH ($\sigma/\sigma_{\mathrm{SM}}$)	
Cut based	3195^{+1440}_{-960}	2555^{+1130}_{-760}	928^{+380}_{-265}

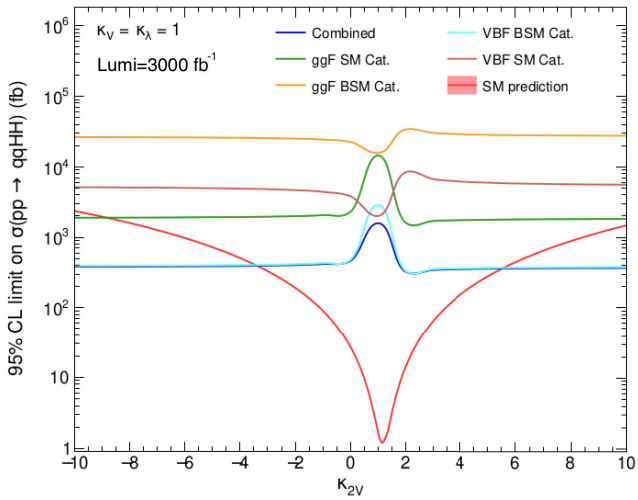


(a) Expected 95% CL upper limits of the ggF HH cross-section as a function of κ_λ in the individual categories (Cat.) and the combination.

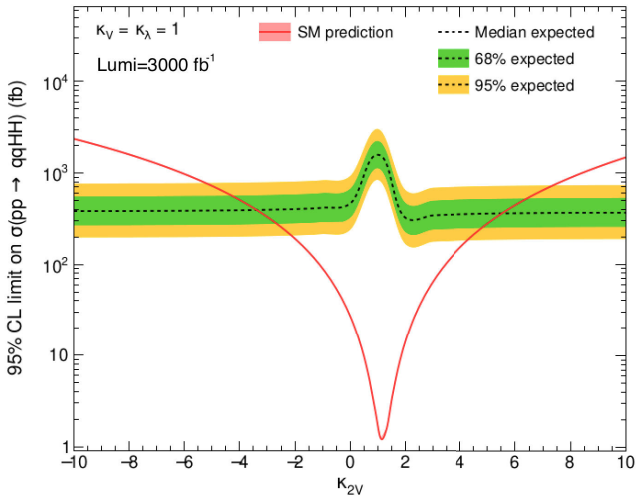


(b) Expected 95% CL upper limits of the ggF HH cross-section as a function of κ_λ.

FIG. 21. Expected 95% CL upper limits of the ggF HH cross section for the breakdown and the combination as a function of κ_λ, using the cut-based approach. The green and yellow bands surrounding the upper limit median represent its 68% and 95% CL uncertainty. The red solid curve with its band represents the theoretical prediction and the corresponding uncertainty [91].



(a) Expected 95% CL upper limits of the VBF HH cross-section as a function of κ_{2V} in the individual categories (Cat.) and the combination.

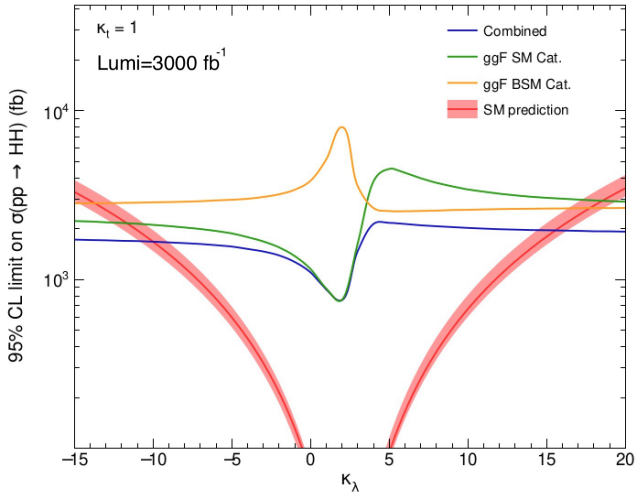


(b) Expected 95% CL upper limits of the VBF HH cross-section as a function of κ_{2V}.

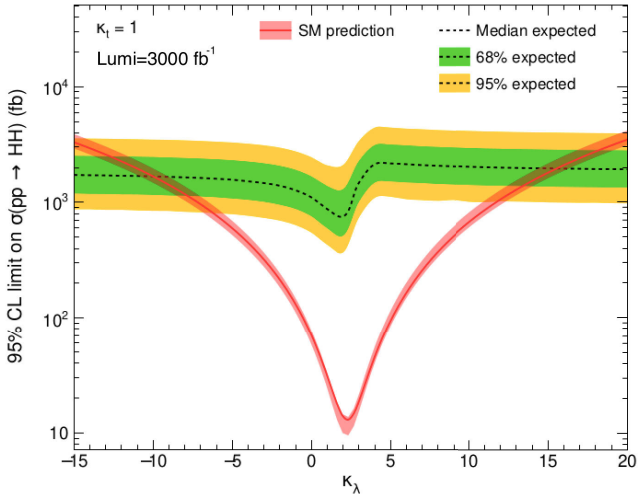
FIG. 22. Expected 95% CL upper limits of the VBF HH cross section for the breakdown and the combination as a function of κ_{2V} using the cut-based approach. The green and yellow bands surrounding the upper limit median represent its 68% and 95% CL uncertainty. The red solid curve with its band represents the theoretical prediction and the corresponding uncertainty [91].

Similarly, the cross section upper limits at 95% CL are also scanned on κ_{2V} as shown in Fig. 22(a) for the individual contributions per category and Fig. 22(b) for the combined results with the uncertainty bands. The green and yellow bands surrounding the upper limit median represent its 68% and 95% CL uncertainty. The red solid curve with its band represents the theoretical prediction and the corresponding uncertainty [91]. At 95% CL, the expected constraint on κ_{2V} is $-3.4 < \kappa_{2V} < 5.5$ with the full integrated luminosity.

In the end, the constraints of κ_λ and κ_{2V} are also extracted for lower integrated luminosities, e.g., 300 fb^{-1} from the LHC Run3 only and 450 fb^{-1} from the combined Run2 and Run3 in Table VII.



(a) Expected 95% CL upper limits of the ggF HH cross-section as a function of κ_λ in the individual categories (Cat.) and the combination.



(b) Expected 95% CL upper limits of the ggF HH cross-section as a function of κ_λ.

FIG. 23. Expected 95% CL upper limits of the ggF HH cross section for the breakdown and the combination as a function of κ_λ using the BDT approach. The green and yellow bands surrounding the upper limit median represent its 68% and 95% CL uncertainty. The red solid curve with its band represents the theoretical prediction and the corresponding uncertainty [91].

VI. CONCLUSIONS

This paper presents a comprehensive study of the Higgs boson pair production in the rare decay of HH $\to b\bar{b}\mu^+\mu^-$ with both the ggF and VBF production modes included to probe the Higgs self-coupling κ_λ and the four-boson HHVV coupling κ_{2V} for the first time. As both of the production rates and the kinematics strongly depend on the two couplings of interest, the analysis is performed with four different event categories, each of which focuses on one of the following cases: SM-like κ_λ, BSM κ_λ, SM-like κ_{2V}, and BSM κ_{2V}. In each of the categories, the corresponding background contributions are suppressed by the dedicated cut-based or BDT approach that is optimized individually. The final result is extracted with fits to the combined spectrum of the dimuon invariant mass $m_{\mu\mu}$ and the dibjet invariant mass m_{bb}. With a integrated luminosity up to 3000 fb^{-1}, the channel HH $\to b\bar{b}\mu^+\mu^-$ cannot lead to the observation of HH with the cut-based or the BDT approach discussed in this paper. The upper limits at 95% confidence level on the cross sections are then extracted using the full HL-LHC integrated luminosity of 3000 fb^{-1}. They are 47 (28) for the ggF HH and 928 for the VBF HH using the cut-based (BDT) approach, both in the unit of their SM predicted cross section. The cross section limits are also used to constrain the couplings. The constraints at 95% confidence level are $-13.8 < \kappa_\lambda < 19.1$ ($-10.0 < \kappa_\lambda < 15.5$) and $-3.4 < \kappa_{2V} < 5.5$ using the cut-based (BDT) approach.

TABLE VII. The 95% CL constraints on κ_λ and κ_{2V} in the cut-based and BDT analyses.

Analysis type	300 fb^{-1}	450 fb^{-1}	3000 fb^{-1}
		κ_λ constraints	
Cut based	$(-26.9, 32.2)$	$(-24.0, 29.3)$	$(-13.8, 19.1)$
BDT	$(-20.7, 26.2)$	$(-18.3, 23.8)$	$(-10.0, 15.5)$
		κ_{2V} constraints	
Cut based	$(-7.6, 9.8)$	$(-6.6, 8.8)$	$(-3.4, 5.5)$

The recent experimental results of HH $\to b\bar{b}\gamma\gamma$ and HH $\to b\bar{b}\tau^+\tau^-$ that lead the HH sensitivity are projected to the HL-LHC with an integrated luminosity of 3000 fb^{-1} under $\sqrt{s} = 14$ TeV [92]. To have a fair comparison with our results, the projections with only statistical uncertainties are quoted in the following. The expected constraints on κ_λ at 95% confidence level are $1.2 < \kappa_\lambda < 4.2$ from HH $\to b\bar{b}\gamma\gamma$ and $2.4 < \kappa_\lambda < 4.5$ from HH $\to b\bar{b}\tau^+\tau^-$. The expected upper limits on the cross section at 95% confidence level are 0.86 and 0.49 times the SM prediction for HH $\to b\bar{b}\gamma\gamma$ and HH $\to b\bar{b}\tau^+\tau^-$, respectively.

In conclusion, the HH $\to b\bar{b}\mu^+\mu^-$ decay channel could not lead to the observation alone up to the HL-LHC using the method discussed in this paper, but it is still able to contribute in a sizeable way to the HH search combination and can be sensitive to BSM enhancement given its small rate and excellent dimuon mass peak.

ACKNOWLEDGMENTS

We thank Qiang Li, Congqiao Li, and Sitian Qian for the helpful discussions. The work is supported in part by the National Science Foundation of China under Grants No. 12175006, No. 12188102, and No. 12061141002 and by Peking University under startup Grant No. 7100603613.

[1] ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys. Lett. B **716**, 1 (2012).

[2] CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Phys. Lett. B **716**, 30 (2012).

[3] F. Englert and R. Brout, Broken Symmetry and the Mass of Gauge Vector Mesons, Phys. Rev. Lett. **13**, 321 (1964).

[4] P. W. Higgs, Broken symmetries, massless particles and gauge fields, Phys. Lett. **12**, 132 (1964).

[5] P. W. Higgs, Broken Symmetries and the Masses of Gauge Bosons, Phys. Rev. Lett. **13**, 508 (1964).

[6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Global Conservation Laws and Massless Particles, Phys. Rev. Lett. **13**, 585 (1964).

[7] P. W. Higgs, Spontaneous symmetry breakdown without massless bosons, Phys. Rev. **145**, 1156 (1966).

[8] T. W. B. Kibble, Symmetry breaking in non-Abelian gauge theories, Phys. Rev. **155**, 1554 (1967).

[9] C. Grojean, G. Servant, and J. D. Wells, First-order electroweak phase transition in the standard model with a low cutoff, Phys. Rev. D **71**, 036001 (2005).

[10] J. Cao, Z. Heng, L. Shang, P. Wan, and J. M. Yang, Pair production of a 125 GeV Higgs boson in MSSM and NMSSM at the LHC, J. High Energy Phys. 04 (2013) 134.

[11] M. Gouzevitch, A. Oliveira, J. Rojo, R. Rosenfeld, G. P. Salam, and V. Sanz, Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production, J. High Energy Phys. 07 (2013) 148.

[12] R. S. Gupta, H. Rzehak, and J. D. Wells, How well do we need to measure the Higgs boson mass and self-coupling?, Phys. Rev. D **88**, 055024 (2013).

[13] C. Han, X. Ji, L. Wu, P. Wu, and J. M. Yang, Higgs pair production with SUSY QCD correction: Revisited under current experimental constraints, J. High Energy Phys. 04 (2014) 003.

[14] K. Nishiwaki, S. Niyogi, and A. Shivaji, ttH anomalous coupling in double Higgs production, J. High Energy Phys. 04 (2014) 011.

[15] F. Goertz, A. Papaefstathiou, L. L. Yang, and J. Zurita, Higgs boson pair production in the $D = 6$ extension of the SM, J. High Energy Phys. 04 (2015) 167.

[16] B. Hespel, D. Lopez-Val, and E. Vryonidou, Higgs pair production via gluon fusion in the two-Higgs-doublet model, J. High Energy Phys. 09 (2014) 124.

[17] J. Cao, D. Li, L. Shang, P. Wu, and Y. Zhang, Exploring the Higgs sector of a most natural NMSSM and its prediction on Higgs pair production at the LHC, J. High Energy Phys. 12 (2014) 026.

[18] A. Azatov, R. Contino, G. Panico, and M. Son, Effective field theory analysis of double Higgs boson production via gluon fusion, Phys. Rev. D **92**, 035001 (2015).

[19] M. Carena, H. E. Haber, I. Low, N. R. Shah, and C. E. M. Wagner, Alignment limit of the NMSSM Higgs sector, Phys. Rev. D **93**, 035013 (2016).

[20] R. Grober, M. Muhlleitner, M. Spira, and J. Streicher, NLO QCD corrections to Higgs pair production including dimension-6 operators, J. High Energy Phys. 09 (2015) 092.

[21] L. Wu, J. M. Yang, C.-P. Yuan, and M. Zhang, Higgs self-coupling in the MSSM and NMSSM after the LHC Run 1, Phys. Lett. B **747**, 378 (2015).

[22] H.-J. He, J. Ren, and W. Yao, Probing new physics of cubic Higgs boson interaction via Higgs pair production at hadron colliders, Phys. Rev. D **93**, 015003 (2016).

[23] A. Carvalho, M. Dall'Osso, T. Dorigo, F. Goertz, C. A. Gottardo, and M. Tosi, Higgs pair production: Choosing benchmarks with cluster analysis, J. High Energy Phys. 04 (2016) 126.

[24] W.-J. Zhang, W.-G. Ma, R.-Y. Zhang, X.-Z. Li, L. Guo, and C. Chen, Double Higgs boson production and decay in Randall-Sundrum model at hadron colliders, Phys. Rev. D **92**, 116005 (2015).

[25] P. Huang, A. Joglekar, B. Li, and C. E. M. Wagner, Probing the electroweak phase transition at the LHC, Phys. Rev. D **93**, 055049 (2016).

[26] K. Nakamura, K. Nishiwaki, K.-y. Oda, S. C. Park, and Y. Yamamoto, Di-Higgs enhancement by neutral scalar as probe of new colored sector, Eur. Phys. J. C **77**, 273 (2017).

[27] L. Di Luzio, R. Gröber, and M. Spannowsky, Maxi-sizing the trilinear Higgs self-coupling: How large could it be?, Eur. Phys. J. C **77**, 788 (2017).

[28] P. Huang, A. Joglekar, M. Li, and C. E. M. Wagner, Corrections to di-Higgs boson production with light stops and modified Higgs couplings, Phys. Rev. D **97**, 075001 (2018).

[29] G. Buchalla, M. Capozi, A. Celis, G. Heinrich, and L. Scyboz, Higgs boson pair production in non-linear effective field theory with full m_t-dependence at NLO QCD, J. High Energy Phys. 09 (2018) 057.

[30] S. Borowka, C. Duhr, F. Maltoni, D. Pagani, A. Shivaji, and X. Zhao, Probing the scalar potential via double Higgs boson production at hadron colliders, J. High Energy Phys. 04 (2019) 016.

[31] S. Chang and M. A. Luty, The Higgs trilinear coupling and the scale of new physics, J. High Energy Phys. 03 (2020) 140.

[32] M. Blanke, S. Kast, J. M. Thompson, S. Westhoff, and J. Zurita, Spotting hidden sectors with Higgs binoculars, J. High Energy Phys. 04 (2019) 160.

[33] H.-L. Li, M. J. Ramsey-Musolf, and S. Willocq, Probing a scalar singlet-catalyzed electroweak phase transition with resonant di-Higgs boson production in the $4b$ channel, Phys. Rev. D **100**, 075035 (2019).

[34] M. Capozi and G. Heinrich, Exploring anomalous couplings in Higgs boson pair production through shape analysis, J. High Energy Phys. 03 (2020) 091.

[35] A. Alves, D. Gonçalves, T. Ghosh, H.-K. Guo, and K. Sinha, Di-Higgs production in the $4b$ channel and gravitational wave complementarity, J. High Energy Phys. 03 (2020) 053.

[36] J. Kozaczuk, M. J. Ramsey-Musolf, and J. Shelton, Exotic Higgs boson decays and the electroweak phase transition, Phys. Rev. D **101**, 115035 (2020).

[37] D. Barducci, K. Mimasu, J. M. No, C. Vernieri, and J. Zurita, Enlarging the scope of resonant di-Higgs searches: Hunting for Higgs-to-Higgs cascades in $4b$ final states at the LHC and future colliders, J. High Energy Phys. 02 (2020) 002.

[38] P. Huang and Y. H. Ng, Di-Higgs production in SUSY models at the LHC, Eur. Phys. J. Plus **135**, 660 (2020).
[39] K. Cheung, A. Jueid, C.-T. Lu, J. Song, and Y. W. Yoon, Disentangling new physics effects on nonresonant Higgs boson pair production from gluon fusion, Phys. Rev. D **103**, 015019 (2021).
[40] Q.-H. Cao, B. Yan, D.-M. Zhang, and H. Zhang, Resolving the degeneracy in single Higgs production with Higgs pair production, Phys. Lett. B **752**, 285 (2016).
[41] Q.-H. Cao, G. Li, B. Yan, D.-M. Zhang, and H. Zhang, Double Higgs production at the 14 TeV LHC and a 100 TeV pp collider, Phys. Rev. D **96**, 095031 (2017).
[42] G. Li, L.-X. Xu, B. Yan, and C. P. Yuan, Resolving the degeneracy in top quark Yukawa coupling with Higgs pair production, Phys. Lett. B **800**, 135070 (2020).
[43] L.-C. Lü, C. Du, Y. Fang, H.-J. He, and H. Zhang, Searching heavier Higgs boson via di-Higgs production at LHC Run-2, Phys. Lett. B **755**, 509 (2016).
[44] J. Ren, R.-Q. Xiao, M. Zhou, Y. Fang, H.-J. He, and W. Yao, LHC search of new Higgs boson via resonant di-Higgs production with decays into 4W, J. High Energy Phys. 06 (2018) 090.
[45] M. Grazzini, G. Heinrich, S. Jones, S. Kallweit, M. Kerner, J. M. Lindert, and J. Mazzitelli, Higgs boson pair production at NNLO with top quark mass effects, J. High Energy Phys. 05 (2018) 059.
[46] F. A. Dreyer and A. Karlberg, Vector-boson fusion Higgs pair production at N^{3}LO, Phys. Rev. D **98**, 114016 (2018).
[47] J. Baglio, A. Djouadi, R. Gröber, M. M. Mühlleitner, J. Quevillon, and M. Spira, The measurement of the Higgs self-coupling at the LHC: Theoretical status, J. High Energy Phys. 04 (2013) 151.
[48] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, P. Torrielli, E. Vryonidou, and M. Zaro, Higgs pair production at the LHC with NLO and parton-shower effects, Phys. Lett. B **732**, 142 (2014).
[49] CMS Collaboration, Search for Higgs Boson Pair Production in the Four b Quark Final State in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **129**, 081802 (2022).
[50] ATLAS Collaboration, Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, J. High Energy Phys. 01 (2019) 030.
[51] CMS Collaboration, Search for nonresonant Higgs boson pair production in the $\mathrm{b\bar{b}b\bar{b}}$ final state at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 04 (2019) 112.
[52] CMS Collaboration, Search for production of Higgs boson pairs in the four b quark final state using large-area jets in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 01 (2019) 040.
[53] ATLAS Collaboration, Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Rev. D **106**, 052001 (2022).
[54] CMS Collaboration, Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 03 (2021) 257.
[55] ATLAS Collaboration, Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV pp collision data collected by the ATLAS experiment, J. High Energy Phys. 11 (2018) 040.
[56] CMS Collaboration, Search for Higgs boson pair production in the $\gamma\gamma\mathrm{b\bar{b}}$ final state in pp collisions at $\sqrt{s} = 13$ TeV, Phys. Lett. B **788**, 7 (2019).
[57] CMS Collaboration, Search for nonresonant Higgs boson pair production in final state with two bottom quarks and two tau leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV, arXiv:2206.09401.
[58] CMS Collaboration, Search for a heavy Higgs boson decaying into two lighter Higgs bosons in the $\tau\tau$bb final state at 13 TeV, J. High Energy Phys. 11 (2021) 057.
[59] ATLAS Collaboration, Search for Resonant and Non-Resonant Higgs Boson Pair Production in the $b\bar{b}\tau^+\tau^-$ Decay Channel in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector, Phys. Rev. Lett. **121**, 191801 (2018); **122**, 089901(E) (2019).
[60] CMS Collaboration, Search for resonant pair production of Higgs bosons in the $bbZZ$ channel in proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. D **102**, 032003 (2020).
[61] CMS Collaboration, Search for nonresonant Higgs boson pair production in the four leptons plus two bjets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV, arXiv:2206.10657.
[62] CMS Collaboration, Search for resonant and nonresonant Higgs boson pair production in the $\mathrm{b\bar{b}}\ell\nu\ell\nu$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 01 (2018) 054.
[63] CMS Collaboration, Search for resonances decaying to a pair of Higgs bosons in the $\mathrm{b\bar{b}q\bar{q}}\prime\ell\nu$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 10 (2019) 125.
[64] ATLAS Collaboration, Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s} = 13$ TeV with the ATLAS detector, J. High Energy Phys. 04 (2019) 092.
[65] ATLAS Collaboration, Search for Higgs boson pair production in the $WW^{(*)}WW^{(*)}$ decay channel using ATLAS data recorded at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 05 (2019) 124.
[66] CMS Collaboration, Search for Higgs boson pairs decaying to WWWW, WW$\tau\tau$, and $\tau\tau\tau\tau$ in proton-proton collisions at $\sqrt{s} = 13$ TeV, arXiv:2206.10268.
[67] ATLAS Collaboration, Search for Higgs boson pair production in the $\gamma\gamma WW^*$ channel using pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS detector, Eur. Phys. J. C **78**, 1007 (2018).
[68] CMS Collaboration, A portrait of the Higgs boson by the CMS experiment ten years after the discovery, Nature (London) **607**, 60 (2022).
[69] ATLAS Collaboration, Constraining the Higgs boson self-coupling from single-and double-Higgs production with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV, Report No. ATLAS-CONF-2022-050, https://cds.cern.ch/record/2816332.
[70] ATLAS Collaboration, Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Lett. B **800**, 135103 (2020).

[71] ATLAS Collaboration, Searches for Higgs boson pair production in the $hh \to bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector, Phys. Rev. D **92**, 092004 (2015).
[72] CMS Collaboration, Combination of Searches for Higgs Boson Pair Production in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **122**, 121803 (2019).
[73] CMS Collaboration, Evidence for Higgs boson decay to a pair of muons, J. High Energy Phys. 01 (2021) 148.
[74] ATLAS Collaboration, A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector, Phys. Lett. B **812**, 135980 (2021).
[75] U. Baur, T. Plehn, and D. L. Rainwater, Probing the Higgs self-coupling at hadron colliders using rare decays, Phys. Rev. D **69**, 053004 (2004).
[76] A. Adhikary, R. K. Barman, and B. Bhattacherjee, Prospects of non-resonant di-Higgs searches and Higgs boson self-coupling measurement at the HE-LHC using machine learning techniques, J. High Energy Phys. 12 (2020) 179.
[77] G. Heinrich, S. P. Jones, M. Kerner, G. Luisoni, and E. Vryonidou, NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers, J. High Energy Phys. 08 (2017) 088.
[78] G. Heinrich, S. P. Jones, M. Kerner, G. Luisoni, and L. Scyboz, Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects, J. High Energy Phys. 06 (2019) 066.
[79] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, J. High Energy Phys. 07 (2014) 079.
[80] G. Heinrich, S. Jones, M. Kerner, G. Luisoni, and L. Scyboz, Trilinear Higgs boson coupling variations for di-Higgs production with full NLO QCD predictions in Powheg, J. Phys. Conf. Ser. **1525**, 012009 (2020).
[81] HiggsWG, Summary table of samples produced for the 1 billion campaign, with 25 ns bunch-crossing, https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns#Summary_table_of_samples_produce.
[82] C. Bierlich *et al.*, A comprehensive guide to the physics and usage of PYTHIA 8.3, arXiv:2203.11601.
[83] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3 Collaboration), DELPHES 3, A modular framework for fast simulation of a generic collider experiment, J. High Energy Phys. 02 (2014) 057.
[84] M. Cacciari, G. P. Salam, and G. Soyez, The anti-k_t jet clustering algorithm, J. High Energy Phys. 04 (2008) 063.
[85] M. Cacciari, G. P. Salam, and G. Soyez, FastJet user manual, Eur. Phys. J. C **72**, 1896 (2012).
[86] T. Chen and C. Guestrin, XGBoost: A scalable tree boosting system, 10.1145/2939672.2939785 (2016).
[87] T. Junk, Confidence level computation for combining searches with small statistics, Nucl. Instrum. Methods Phys. Res., Sect. A **434**, 435 (1999).
[88] A. L. Read, Presentation of search results: The CL(s) technique, J. Phys. G **28**, 2693 (2002).
[89] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71**, 1554 (2011); **73**, 2501(E) (2013).
[90] CMS Collaboration, A measurement of the Higgs boson mass in the diphoton decay channel, Phys. Lett. B **805**, 135425 (2020).
[91] LHCHWGHH, LHC Higgs Cross Section, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWGHH?redirectedfrom=LHCPhysics.LHCHXSWGHH.
[92] ATLAS Collaboration, Projected sensitivity of Higgs boson pair production combining the $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ final states with the ATLAS detector at the HL-LHC, CERN Report No. ATL-PHYS-PUB-2022-005, 2022.