Sensitivity of two-Higgs-doublet models on Higgs-pair production via $b\bar{b}b\bar{b}$ final state

Yi-Lun Chung,[1,*] Kingman Cheung,[1,2,†] and Shih-Chieh Hsu[3,‡]

[1]*Department of Physics and Center for Theory and Computation, National Tsing Hua University, Hsinchu 300, Taiwan*

[2]*Division of Quantum Phases and Devices, School of Physics, Konkuk University, Seoul 143-701, Republic of Korea*

[3]*Department of Physics, University of Washington, Seattle, Washington 98195, USA*

(Received 16 August 2022; accepted 24 October 2022; published 14 November 2022)

Higgs boson pair production is a well-known probe of the structure of the electroweak symmetry breaking sector. We illustrate this using the gluon-fusion processes $pp \to H \to hh \to (b\bar{b})(b\bar{b})$ in the framework of two-Higgs-doublet models and show how a machine learning approach (three-stream convolutional neural network) can substantially improve the signal-background discrimination and thus improve the sensitivity coverage of the relevant parameter space. We further show that such $gg \to hh \to b\bar{b}b\bar{b}$ processes can probe the parameter space currently allowed by HIGGSSIGNALS and HIGGSBOUNDS at the HL-LHC. Results are presented for 2HDM types I through IV.

DOI: 10.1103/PhysRevD.106.095015

I. INTRODUCTION

The origin of mass is highly related to the mechanism involved in electroweak symmetry breaking (EWSB), which is believed to give mass to matter and gauge bosons. The simplest implementation of EWSB in the standard model (SM) is to introduce a Higgs doublet field [1–3]. In July 2012, a particle was discovered with properties consistent with the neutral scalar Higgs boson described by the standard model (SM) theory [4,5]. Considering the data accumulated up to 2018, this scalar boson is still described well by the SM Higgs boson [6–9]. However, the SM cannot be a complete theory, because there are a number of observations that it cannot explain, including neutrino oscillations [10], the presence of dark matter and dark energy in the Universe [11], and baryon asymmetry [12]. The Higgs boson also suffers from the so-called gauge hierarchy problem [13].

However, there is no *a priori* reason why the EWSB sector simply contains only one Higgs doublet field. Indeed, many new physics models which resolve one or more of the above problems with the SM involve extensions of the EWSB sector. For example, supersymmetry is a well-known solution to the gauge hierarchy problem and has at least two Higgs doublet fields. Therefore, probing the EWSB sector may give insights into ultraviolet physics at the TeV scale. One of the best ways to probe the structure of the EWSB sector is to probe the Higgs self-couplings. This is because the self-couplings of the Higgs boson are very different among the SM, two Higgs doublet models (2HDM), MSSM, and any composite Higgs models. One of the probes of Higgs self-couplings is Higgs-pair production via gluon fusion at the LHC [14–24]. Higgs-pair production beyond the SM has been very well studied, for example [25] and references therein. Production rates of Higgs boson pairs vary with variations in the Higgs self-coupling, λ_{3H}, or due to the presence of heavier Higgs bosons, providing an experimental signature which can be used to test the various models for physics beyond the SM. In Higgs pair production via gluon fusion, the dominant contributions to the cross section are from triangle and box diagrams, which interfere with each other. In 2HDM, the triangle diagram can involve the Higgs self-trilinear couplings λ_{Hhh} and λ_{hhh}. In particular, the resonance effect of the heavier CP-even Higgs boson can substantially enhance the production rate of hh pairs. The production rate largely depends on the parameters of the 2HDM, such as $\cos(\beta-\alpha)$, $\tan\beta$, and m_{12}^2 in addition to the M_H and Γ_H.

In this work, we study the signal process $pp \to hh \to (b\bar{b})(b\bar{b})$ via gluon fusion against the SM multijet background. It is well known that the signal is overwhelmingly buried under the multijet background. The study using the

*s107022801@m107.nthu.edu.tw
†cheung@phys.nthu.edu.tw
‡schsu@uw.edu

conventional cut-based approach did not give enough significance even at the High-Luminosity LHC (HL-LHC). We make use of the boosted feature of the final-state Higgs boson pair hh, due to the decay of the heavier CP-even Higgs boson. A specific classifier was developed in Ref. [26], which can be employed to significantly enhance the signal-background ratio. We show that a three-stream convolutional neural network (3CNN) can substantially improve the significance of the signal compared to a boosted decision tree (BDT) or conventional cut-based approach. At the end of the analysis, we show the 95% sensitivity coverage of the parameter space of the 2HDM Type I, II, III, and IV. The current study focuses on the boosted regime of Higgs boson pair production, where our classifier is very effective at reducing the SM multijet background. In the literature, there are analyses that probe the Higgs self-coupling in channels such as $hh \to b\bar{b}\gamma\gamma$ [27,28], $hh \to b\bar{b}WW^*$ [29,30], $hh \to b\bar{b}b\bar{b}$ [31], and a resonance search [32].

We summarize our findings that show improvements.

1. Making use of the 3CNN we have successfully cut down the multijet background by more than a factor of 20 compared with the conventional cut-based approach. Thus, we can explore the much larger parameter space of the 2HDMs.
2. Detection of Higgs-pair production via the $4b$ mode is extremely difficult. The current upper limit on the production cross section $pp \to X \to hh$ via the $4b$ mode is about 10 fb for $M_X \simeq 1$ TeV [33]. With the use of 3CNN and running at the HL-LHC, we show that the 95% sensitivity can be down to $\sigma(pp \to X \to hh) \sim O(1)$ fb.

The paper is organized as follows. In the next section, we briefly describe the 2HDM types, as well as the resulting relevant parameters for Higgs-pair production. In Sec. III, we describe the signal and background processes, including the sample generation and event selection. In Sec. IV, we introduce the machine learning approaches, including BDT and 3CNN classifiers. In Sec. V, we scan the parameter space of the 2HDMs to provide a description of the parameter space coverage at the HL-LHC, including the current restrictions on the parameter space due to HIGGSSIGNALS and HIGGSBOUNDS. We conclude in Sec. VI.

II. TWO HIGGS DOUBLET MODELS

The 2HDM extends the SM by adding the second complex Higgs doublet field. The 2HDM Higgs sector, therefore, consists of Φ_1 and Φ_2 [34]:

$$\Phi_i = \begin{pmatrix} w_i^+ \\ \frac{v_i + h_i + i\eta_i}{\sqrt{2}} \end{pmatrix}, \qquad i = 1, 2, \tag{2.1}$$

where v_1 and v_2 are the vacuum expectation values (VEV) of Φ_1 and Φ_2, respectively. The ratio of these two VEVs is defined by $\tan\beta \equiv v_2/v_1$. The dangerous flavor-changing-neutral-currents (FCNC) at tree level are avoided by imposing a discrete Z_2 symmetry, under which $\Phi_1 \to \Phi_1$ and $\Phi_2 \to -\Phi_2$ [35,36]. The scalar potential with softly broken Z_2 and CP invariance is

$$\begin{aligned} V = {} & m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - m_{12}^2 (\Phi_1^\dagger \Phi_2 + \text{H.c.}) \\ & + \frac{1}{2}\lambda_1 (\Phi_1^\dagger \Phi_1)^2 + \frac{1}{2}\lambda_2 (\Phi_2^\dagger \Phi_2)^2 + \lambda_3 (\Phi_1^\dagger \Phi_1)(\Phi_2^\dagger \Phi_2) \\ & + \lambda_4 (\Phi_1^\dagger \Phi_2)(\Phi_2^\dagger \Phi_1) + \frac{1}{2}\lambda_5 [(\Phi_1^\dagger \Phi_2)^2 + \text{H.c.}], \end{aligned} \tag{2.2}$$

where the m_{12}^2 term softly breaks the Z_2 symmetry.

The two Higgs doublets result in five physical Higgs bosons: a pair of CP-even scalar bosons h and H, a CP-odd pseudoscalar A, and a pair of charged Higgs bosons $H^\pm$. The masses of the physical Higgs bosons are related to the values of the couplings λs in the scalar potential, the mixing angle, α, of the CP-even scalar bosons, and β. For the five physical bosons, these couplings are [37]:

$$\begin{aligned} \lambda_1 &= \frac{1}{v^2 c_\beta^2}[c_\alpha^2 M_H^2 + s_\alpha^2 m_h^2 - t_\beta m_{12}^2], \\ \lambda_2 &= \frac{1}{v^2 s_\beta^2}\left[s_\alpha^2 M_H^2 + c_\alpha^2 m_h^2 - \frac{1}{t_\beta} m_{12}^2\right], \\ \lambda_3 &= \frac{1}{v^2}\left[2M_{H^\pm}^2 + \frac{s_{2\alpha}}{s_{2\beta}}(M_H^2 - m_h^2) - \frac{m_{12}^2}{s_\beta c_\beta}\right], \\ \lambda_4 &= \frac{1}{v^2}\left[M_A^2 - 2M_{H^\pm}^2 + \frac{m_{12}^2}{s_\beta c_\beta}\right], \\ \lambda_5 &= \frac{1}{v^2}\left[\frac{m_{12}^2}{s_\beta c_\beta} - M_A^2\right], \end{aligned} \tag{2.3}$$

where $s_\alpha \equiv \sin\alpha$, $c_\alpha \equiv \cos\alpha$, $t_\beta \equiv \tan\beta$, etc. The six free parameters of the 2HDMs are

$$\{m_h, M_H, M_A, M_{H^\pm}, t_\beta, c_{\beta-\alpha}\}. \tag{2.4}$$

Note that we focus on the scenario in which the lighter CP-even scalar Higgs boson h is the SM-like Higgs boson observed, and in this scenario $c_{\beta-\alpha}$ is constrained to be close to zero by the current Higgs boson data.

Conventionally, there are four assignments of the Z_2 parity for the SM fermions, resulting in 2HDM Type I, II, III, and IV, which differ among themselves in the couplings of Higgs bosons to fermions. The Yukawa couplings in 2HDM can be parametrized as

$$\mathcal{L}_Y = -\sum_f \left[\frac{m_f}{v}\xi_f^h \bar{f}fh + \frac{m_f}{v}\xi_f^H \bar{f}fH - i\frac{m_f}{v}\xi_f^A \bar{f}\gamma_5 fA\right] - \left[\frac{\sqrt{2}V_{ud}}{v}H^+\bar{u}(m_u\xi_u^A P_L + m_d\xi_d^A P_R)d + \frac{\sqrt{2}m_\ell}{v}H^+\xi_\ell^A\overline{\nu_L}\ell_R + \text{H.c.}\right], \quad (2.5)$$

where the modifiers $\xi_f^{h,H,A}$ are presented in Table I.

Since the parameter $\cos(\beta-\alpha)$ is constrained to be close to zero by the Higgs boson data, it is instructional to expand the relevant couplings in terms of $\cos(\beta-\alpha)$. We are considering the following gluon-fusion process

$$pp \to H \to hh \to (b\bar{b})(b\bar{b}).$$

The relevant couplings include the trilinear coupling λ_{hhH}, Yukawa couplings $\lambda_{t,b}^{h,H}$ of h and H. The trilinear coupling λ_{hhH} is given by

$$\lambda_{hhH} = \frac{c_{\beta-\alpha}}{s_{2\beta}}\left[s_{2\alpha}(2m_h^2+M_H^2) - \frac{2m_{12}^2}{s_{2\beta}}(3s_{2\alpha}-s_{2\beta})\right]. \quad (2.6)$$

The Yukawa couplings can be expanded similarly. For example, in Type II they are given by

$$\lambda_t^h = \frac{c_\alpha}{s_\beta} = 1 + \frac{c_{\beta-\alpha}}{t_\beta} - \frac{1}{2}c_{\beta-\alpha}^2 + \mathcal{O}(c_{\beta-\alpha}^3) \quad (2.7)$$

$$\lambda_b^h = -\frac{s_\alpha}{c_\beta} = 1 - c_{\beta-\alpha}t_\beta - \frac{1}{2}c_{\beta-\alpha}^2 + \mathcal{O}(c_{\beta-\alpha}^3) \quad (2.8)$$

$$\lambda_t^H = \frac{s_\alpha}{s_\beta} = -\frac{1}{t_\beta} + c_{\beta-\alpha} + \frac{c_{\beta-\alpha}^2}{2t_\beta} + \mathcal{O}(c_{\beta-\alpha}^3) \quad (2.9)$$

$$\lambda_b^H = \frac{c_\alpha}{c_\beta} = t_\beta + c_{\beta-\alpha} - \frac{c_{\beta-\alpha}^2}{2t_\beta} + \mathcal{O}(c_{\beta-\alpha}^3) \quad (2.10)$$

Other types can be expanded similarly. It is straightforward to see that the production cross section via the resonance CP-even H scales as $\cos^2(\beta-\alpha)$.

TABLE I. The Yukawa coupling modifiers in the four types of the 2HDM.

	ξ_u^h	ξ_d^h	ξ_ℓ^h	ξ_u^H	ξ_d^H	ξ_ℓ^H	ξ_u^A	ξ_d^A	ξ_ℓ^A
type-I	$\frac{c_\alpha}{s_\beta}$	$\frac{c_\alpha}{s_\beta}$	$\frac{c_\alpha}{s_\beta}$	$\frac{s_\alpha}{s_\beta}$	$\frac{s_\alpha}{s_\beta}$	$\frac{s_\alpha}{s_\beta}$	$\frac{1}{t_\beta}$	$-\frac{1}{t_\beta}$	$-\frac{1}{t_\beta}$
type-II	$\frac{c_\alpha}{s_\beta}$	$-\frac{s_\alpha}{c_\beta}$	$-\frac{s_\alpha}{c_\beta}$	$\frac{s_\alpha}{s_\beta}$	$\frac{c_\alpha}{c_\beta}$	$\frac{c_\alpha}{c_\beta}$	$\frac{1}{t_\beta}$	t_β	t_β
type-III (lepton-specific)	$\frac{c_\alpha}{s_\beta}$	$\frac{c_\alpha}{s_\beta}$	$-\frac{s_\alpha}{c_\beta}$	$\frac{s_\alpha}{s_\beta}$	$\frac{s_\alpha}{s_\beta}$	$\frac{c_\alpha}{c_\beta}$	$\frac{1}{t_\beta}$	$-\frac{1}{t_\beta}$	t_β
type-IV (flipped)	$\frac{c_\alpha}{s_\beta}$	$-\frac{s_\alpha}{c_\beta}$	$\frac{c_\alpha}{s_\beta}$	$\frac{s_\alpha}{s_\beta}$	$\frac{c_\alpha}{c_\beta}$	$\frac{s_\alpha}{s_\beta}$	$\frac{1}{t_\beta}$	t_β	$-\frac{1}{t_\beta}$

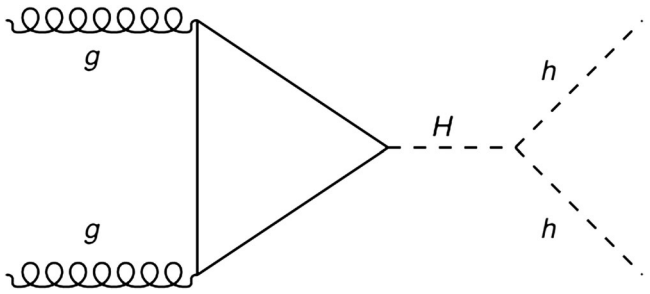

FIG. 1. A Feynman diagram for resonant Higgs boson pair production via gluon fusion in the two-Higgs-doublet model.

2HDM theories are highly constrained by Higgs signal strength data from the LHC [38–45], as well as by direct search bounds on heavy CP-even and CP-odd scalar bosons and charged Higgs bosons. We use HiggsBounds-5.10.2 [46–51] and HiggsSignal-2.6.2 [52–55] to test the validity of the parameter space points considered here.

III. SAMPLE GENERATION AND EVENT SELECTIONS

In this study, the signal is resonant Higgs boson pair production via gluon fusion in 2HDM, as shown in Fig. 1. We consider QCD multijet and top-quark pair production ($t\bar{t}$) processes to be irreducible SM backgrounds. Other background processes, such as the nonresonant (continuum) SM hh[1] and electroweak diboson production, are negligible [33].

We consider processes in which the light Higgs boson h ($m_h = 125$ GeV) pair comes from a heavy CP-even scalar H with mass $M_H = 1000$ GeV. The other physical parameters at this benchmark point are $M_A = M_{H^\pm} = 1000$ GeV, $m_{12}^2 = 400{,}000$ GeV^2, $\tan\beta = 5$, and $\cos(\beta-\alpha) = 0.01$. This benchmark point is still allowed under current limits[2] and close to the alignment limit in Type II 2HDM.

A. Monte Carlo samples

We used MadGraph5_aMC@NLO 2.7.2 [56] to model the signal and background processes in pp collisions at $\sqrt{s} = 14$ TeV. The hard-scattering events are passed to Pythia 8.244 [57] to simulate the parton shower and hadronization, using the default settings. Following Ref. [33], the `NNPDF30_nlo_as_0118` [58] parton distribution function (PDF) is used for the next-to-leading-order calculation of the signal and `NNPDF23_lo_as_0130_qed` [59] for the leading-order calculation of the backgrounds.

For signal event generation, the general 2HDM prepared with FeynRules is used [60]. This model file contains all additional rules needed to promote a 2HDM Lagrangian

[1]We have verified with a parton-level calculation that the resonance peak in the invariant-mass distribution M_{hh} stands tremendously above the continuum hh production.

[2]We verified by HiggsBounds-5.10.2 and HiggsSignal-2.6.2.

to one-loop computations in QCD and electroweak couplings within a completely general 2HDM setup. The loop–induced gluon-fusion production process $gg \to H$ is generated automatically at the leading order within MADGRAPH5_AMC@NLO. The input parameters of the benchmark point for loop propagators and s–channel Higgs boson widths are submitted to MADGRAPH5_AMC@NLO 2.7.2 through parameter cards in the standard setup. The latter are constructed with the public calculator 2HDMC [61] with `HiggsBounds-5.10.2` [47–51] and HIGGSSIGNAL-2.6.2 [52–55] extensions. Moreover, the decay chain $H \to hh \to b\bar{b}b\bar{b}$ is implemented by MADSPIN [62] and the light Higgs boson h is set to decay 100% into $b\bar{b}$ during production to improve selection efficiency. The $H \to b\bar{b}$ branching ratio is then used for the actual event rates.

The $t\bar{t}$ background is simulated at leading order and up to two more jets with the matching scale 20 GeV via MLM prescription [63,64]. The other background, the multijet process, is flavor-inclusive. For the multijet process, we require `ihtmin` (inclusive scalar sum H_t for all partons) to be 850 GeV to enhance simulation efficiency. The cross sections of the background samples are normalized based on the event yield given in the ATLAS analysis.

PYJET [65,66] and the anti-k_t [67] algorithm with cone-radius parameter $R = 1.0$[3] are used to define the boosted jets. After these boosted jets are formed, we require the transverse energy (E_T) of the leading large-R jet to be >420 GeV and invariant mass of the leading large-R jet to be >35 GeV. A trimming procedure [68] is then performed. The constituents in large-R jets are re-clustered into "subjets" using the k_T algorithm [69] with $R = 0.2$. Subjets with less than 5% of the transverse momentum p_T of the large-R jet are then removed.

After the trimming procedure, an event level preselection is applied, following Ref. [33]. Each event is required to contain at least two large-R jets with $p_T(J_1) > 450$ GeV and $p_T(J_2) > 250$ GeV where J_1 and J_2 are the leading and subleading jets, respectively. These two large-R jets are also required to have pseudorapidity $|\eta(J)| < 2$ and jet mass $M(J) > 50$ GeV. After applying the preselection, there remain 600k events from signal and total background for training. In order to generate enough statistics for the analysis, there are 600k events and 2.3M events from the signal and total background for testing, respectively.

As an additional selection, the Higgs jet is required to satisfy double b-tagging. Jets are declared double b-tagged if they have two or more ghost-associated [70,71] B hadrons. This approach is similar to the subjet b-tagging in ATLAS [72,73]. We do not include effects from pileup in this study. This is motivated by [26], which shows that classification performance is relatively unchanged by pileup when neutral particles are not involved.

[3]The angular separation between two jets is defined by $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ where η is the pseudorapidity and ϕ is the azimuth angle.

B. High-level features

In order to distinguish the signal from SM backgrounds via BDTs, the following fifteen commonly-used high-level features are considered:

1. M_{JJ}: invariant mass of the leading and subleading large-R jets;
2. $M(J_1)$ and $M(J_2)$: invariant mass of the leading jet J_1 and the subleading jet J_2, respectively;
3. $|\Delta\eta(JJ)| \equiv |\eta(J_1) - \eta(J_2)|$ where $\eta(J)$ is the pseudorapidity of the jet J;
4. X_{HH}: $\sqrt{\left(\frac{M(J_1)-124\text{ GeV}}{0.1\times M(J_1)}\right)^2 + \left(\frac{M(J_2)-115\text{ GeV}}{0.1\times M(J_2)}\right)^2}$ [33];
5. $\tau_{21} = \tau_2/\tau_1$: ratio of n-subjettiness of the leading jet and the subleading jet [74,75];
6. $D_2^{(\beta)} = e_3^{(\beta)}/(e_2^{(\beta)})^3$ with $\beta = 1, 2$: ratios of energy correlation function of the leading jet and the subleading jet [76];
7. $C_2^{(\beta)} = e_3^{(\beta)}/(e_2^{(\beta)})^2$ with $\beta = 1, 2$: ratios of energy correlation function of the leading jet and the subleading jet [77];

where e_i is the normalized sum over doublets ($i = 2$) or triplets ($i = 3$) of constituents inside jets, weighted by the product of the constituent transverse momenta and pairwise angular distances. In this analysis, β is considered to be 1 and 2.

Distributions of these variables are shown in Figs. 2 and 3. Separation is seen between distributions of signal and background, motivating the choice of these variables for signal and background discrimination. More particularly, the dijet invariant mass distribution peaks near the heavy resonance of 1000 GeV for the signal while it is broad for the background. The resonant signal jets tend to be very central since they are produced through s-channel processes. In this case, the $|\Delta\eta(JJ)|$ provides good discrimination power. X_{HH} represents the distance of an event from the di-Higgs peak in the $M(J_1) - M(J_2)$ plane. In Fig. 2, the peaks of invariant mass of the leading jet and subleading jet are around 124 GeV and 115 for the signal, respectively. It implies that the signal can be distinguished in the small X_{HH} region. The decay of massive objects into two hard QCD partons produces the two-prong structure, which makes the signal jets result in low τ_{21}, $D_2^{(\beta)}$ and $C_2^{(\beta)}$.

C. Low-level features

The low-level inputs to the three-stream convolutional neural networks are full-event images, and images of the leading and subleading jets [78,79]. The resolution is 40×40 pixels for both sets of images and jet images are in 1R × 1R range [26,72]. The images consist of three channels, analogous to the red/green/blue (RGB) channels

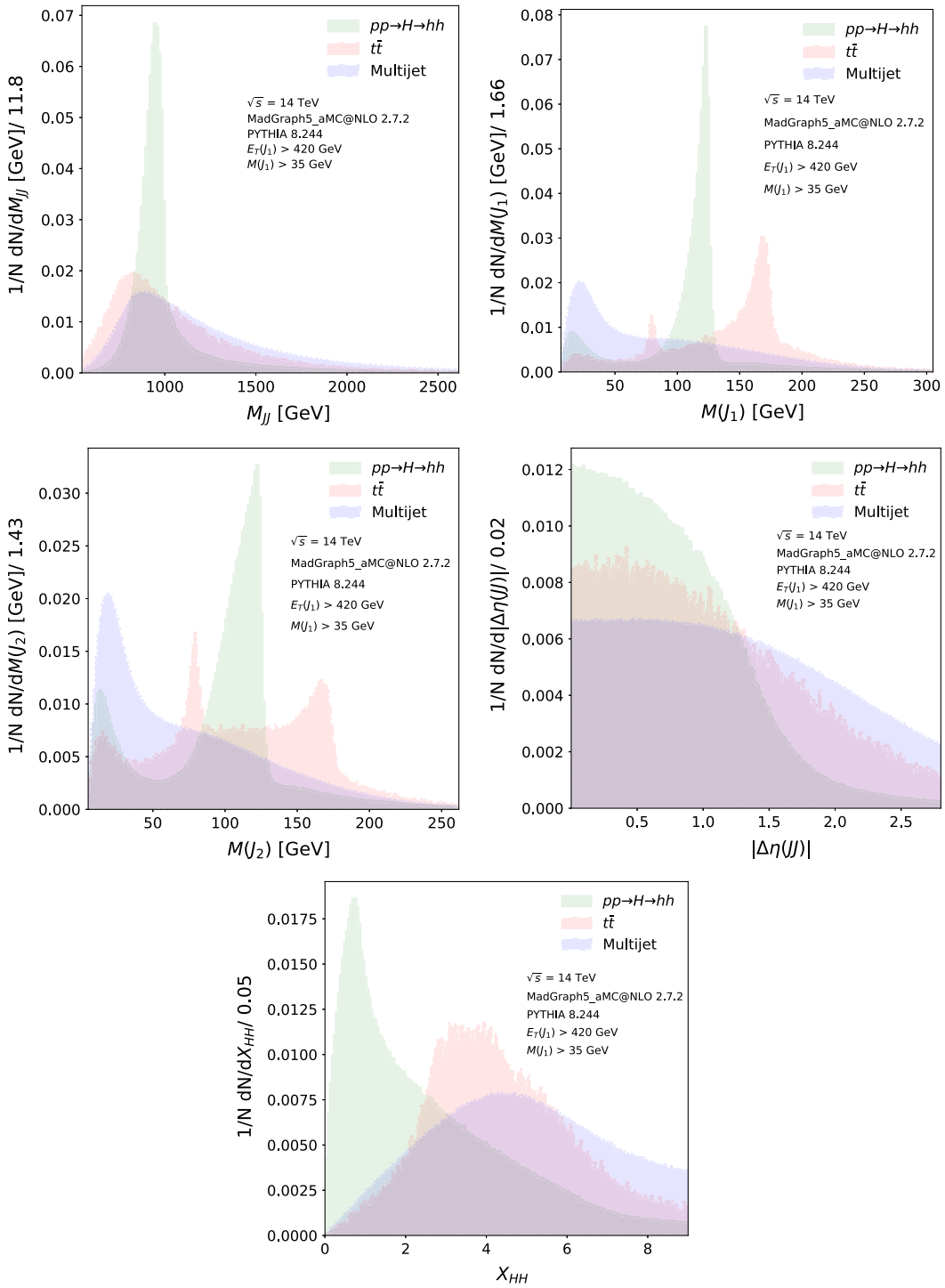

FIG. 2. Distributions of the five kinematic variables used in the BDT. In all figures, $E_T(J_1)$ and $M(J_1)$ represent the transverse energy and invariant mass of the leading jet, respectively.

of a color image. The pixel intensity for the three channels corresponds to the sum of the charged particle p_T, the sum of the neutral particle p_T, and the number of charged particles in a given region of the image. There is no p_T threshold for the contributions to pixel intensity. The full-event image covers effectively the entire $\eta - \phi$ cylinder ($|\eta| < 5$). Moreover, the full-event images are rotated so that the leading jet is always located at $\phi = \pi/2$. Images are then flipped along the axis defined by $\eta = 0$ to put the leading jet centroid in the region with positive η. The jet images are rotated so that the two subjets are aligned along the same axis. The leading subjet is at the origin and the

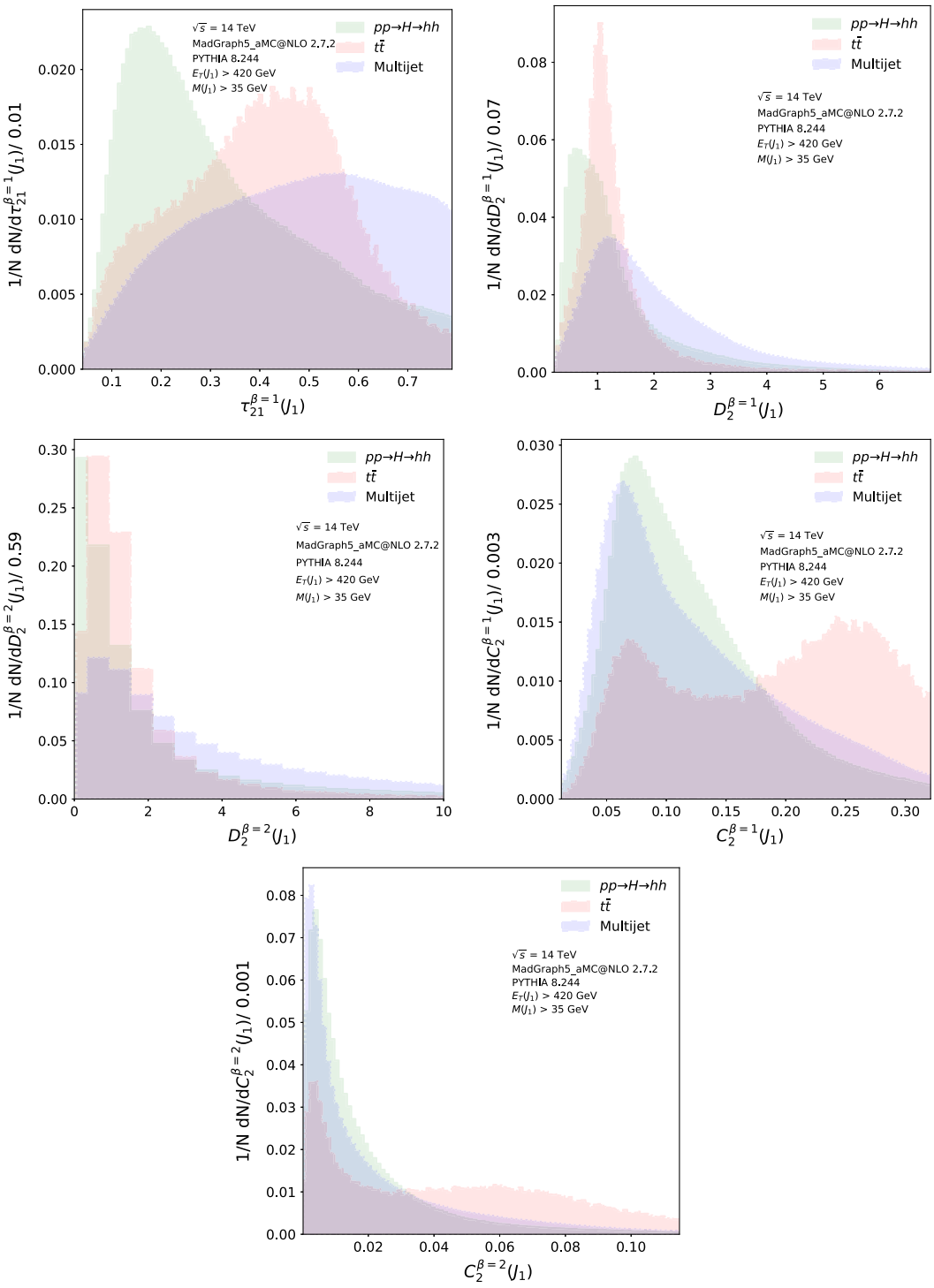

FIG. 3. Distributions of the five substructures used in the BDT. In all figures, $E_T(J_1)$ and $M(J_1)$ represent the transverse energy and invariant mass of the leading jet, respectively. Here, we only show substructures of the leading jet. The distributions of substructures of the subleading jet are similar to those of the leading jet.

subleading subjet is directly below the leading subjet. If there is a third-leading subjet, the image will be reflected. All images are normalized so that the intensities all sum to one. After normalization, the pixel intensities are standardized so that their distribution has mean zero and unit variance. These preprocessing procedures significantly improve the stability of the machine learning training [80]. Figure 4 shows the average images in the charged p_T channel. The patterns in the charged p_T channel are similar to the other two channels.

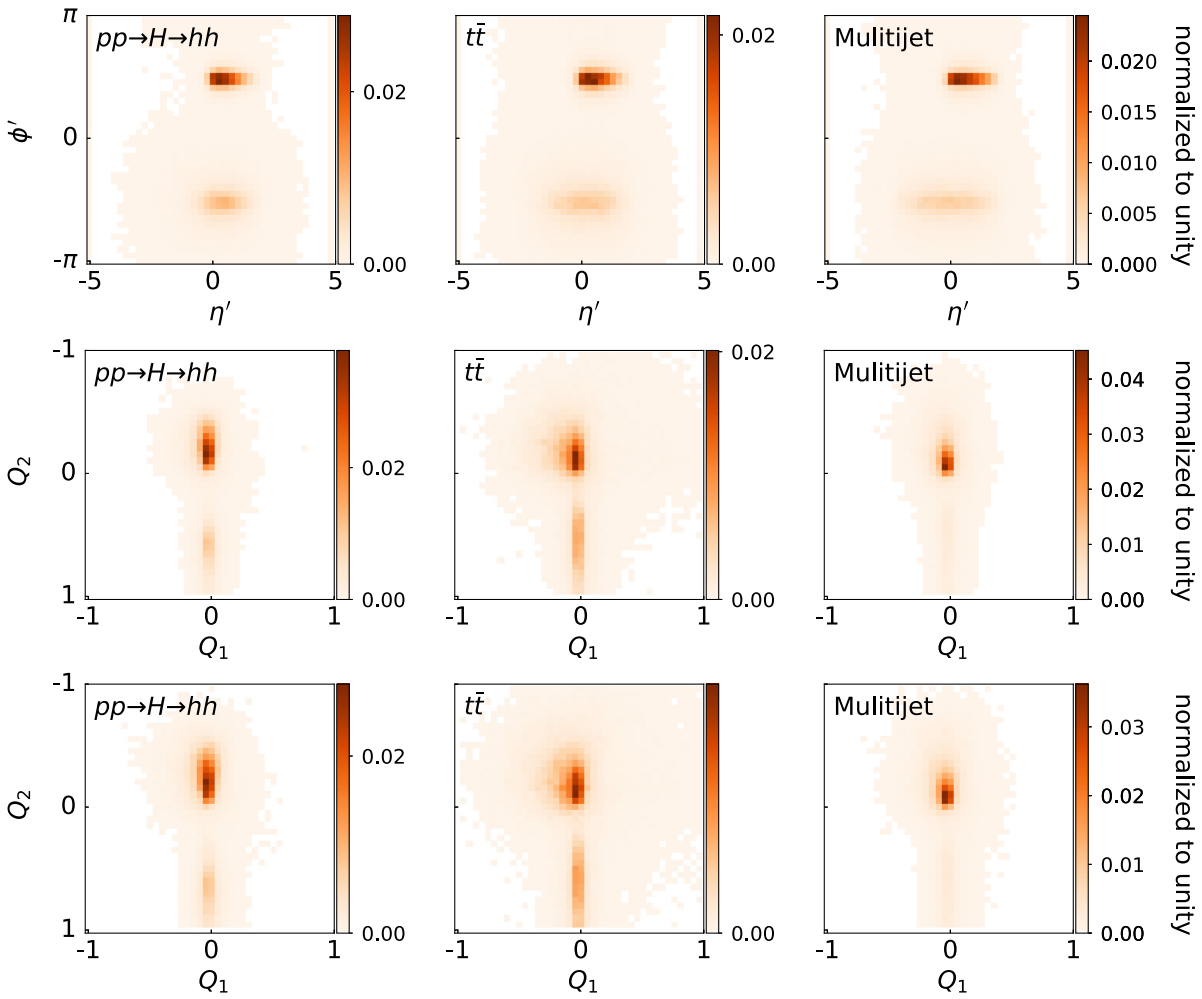

FIG. 4. The average of 10 000 rotated full-event images (top), leading jet images (middle) and subleading jet images (bottom) in the charged p_T channel. The coordinates ϕ' and η' denote the new axis after the full-event images are rotated and flipped. Q_1 and Q_2 denote the new axes after the jet's axis is centralized and rotated. The intensity in each pixel is the sum of the charged particle p_T. The total intensity in each image is normalized to unity. The resolution is 40×40 pixels for each image.

IV. BASELINE AND CLASSIFIERS

In this study, there are three selection methods. The first one is a conventional cut-based selection approach, following Ref. [33]. We refer to this as the "baseline" method in the following. The second approach uses a BDT trained on the high-level features described above. The third approach is based on three-stream convolutional neural networks, as inspired by Refs. [26,72], and is the contribution of this work.

A. The cut-based method

For the cut-based approach, we closely follow the boosted-channel analysis in Ref. [33]. After preselection, which has been described in Sec. III A, we further apply cuts of $|\Delta\eta(JJ)| < 1.3$ and $X_{HH} < 1.6$ to reduce SM backgrounds. We then require that M_{JJ} is in the heavy resonance mass window: $900\,\text{GeV} < M_{JJ} < 1100\,\text{GeV}$. Finally, we require the leading and subleading jets to be the Higgs jets. The cut flow is shown in Table II.

B. The boosted decision tree

The BDT has a fixed number of estimators (2000) with a maximum depth of 5. The minimum number of samples is fixed at 25% as required to split an internal node and 5% as required to be at a leaf node. The deviance of the loss function is set with the learning rate of 0.005. This BDT model is trained on fifteen high-level features of the jet using the `scikit-learn` library [81]. The cut flow for this approach is also shown in Table II.

C. The three-stream convolutional neural networks (3CNN)

The 3CNN in this study is based on Refs. [26,72]. One stream of the 3CNN is dedicated to global full-event

TABLE II. Table showing the cut flow and event yield for the signal process $pp \to H \to hh \to b\bar{b}b\bar{b}$ and the backgrounds at $\sqrt{s} = 14$ TeV with an integrated luminosity $\mathcal{L} = 3000$ fb^{-1}. The signal is Type II of 2HDMs. The B-hadron tagging efficiency $= 0.77$ [33] is applied to calculate the event yield. The preselection is described in the main text.

Selection Flow Table					
		$pp \to H \to hh \to b\bar{b}b\bar{b}$ (Type II)	$t\bar{t}$	Mulitijet	Total Backgrounds
Preselection		8.02×10^1	9.23×10^5	2.76×10^7	2.86×10^7
900 GeV $< M_{JJ} <$ **1100 GeV**		5.29×10^1	2.77×10^5	6.92×10^6	7.20×10^6
2 Higgs jets		4.74×10^1	1.05×10^3	2.34×10^4	2.45×10^4
Baseline	$\lvert\Delta\eta(JJ)\rvert < 1.3$	4.68×10^1	9.99×10^2	2.18×10^4	2.28×10^4
	$X_{HH} < 1.6$	2.82×10^1	2.13×10^1	1.37×10^3	1.39×10^3
BDT score > 0.964		2.56×10^1	5.33	1.37×10^2	1.42×10^2
3CNN score > 0.99		2.56×10^1	2.93×10^1	2.74×10^1	5.67×10^1

information. The other two streams are dedicated to processing local information in the leading jet and subleading jet. In addition, there are two outputs in the last layer, used to help disentangle the signal and SM backgrounds. The three-stream architecture is shown schematically in Fig. 5.

Details of the 3CNN are as follows. The convolution filter is 5×5 in three streams, the maximum pooling layers are 2×2, and the stride length is 1. Rectified linear unit (ReLU) activation functions are used for all intermediate layers of the neural network (NN). The first convolution layer in each stream has 32 filters and the second convolution layer in each stream has 64 filters. There are 300 neurons for the dense layer at the end of each stream. The three dense layers from each stream are fully connected to two output neurons with the softmax activation function $e^{x_i}/\sum_{i=1}^{4} e^{x_i}$, which is the multidimensional generalization of the sigmoid. The AdaDelta optimizer [82] is used to select the network weights. Between the last dense layer and output layer, dropout [83] regularization is added to reduce overfitting with the dropout rate $= 0.1$. The categorical cross-entropy loss function is optimized in the neural network training. For effectively utilizing the full information of the detector in the ϕ direction, a padding method is used to take the information in the bottom four rows of the input images and append them to the top of the image. The `Keras-2.4.0` library is used to train a 3CNN model with the `TTENSORFLOW-2.4.0-rc3` [84] backend, on an `NVIDIA RTX A6000 48 GB`.

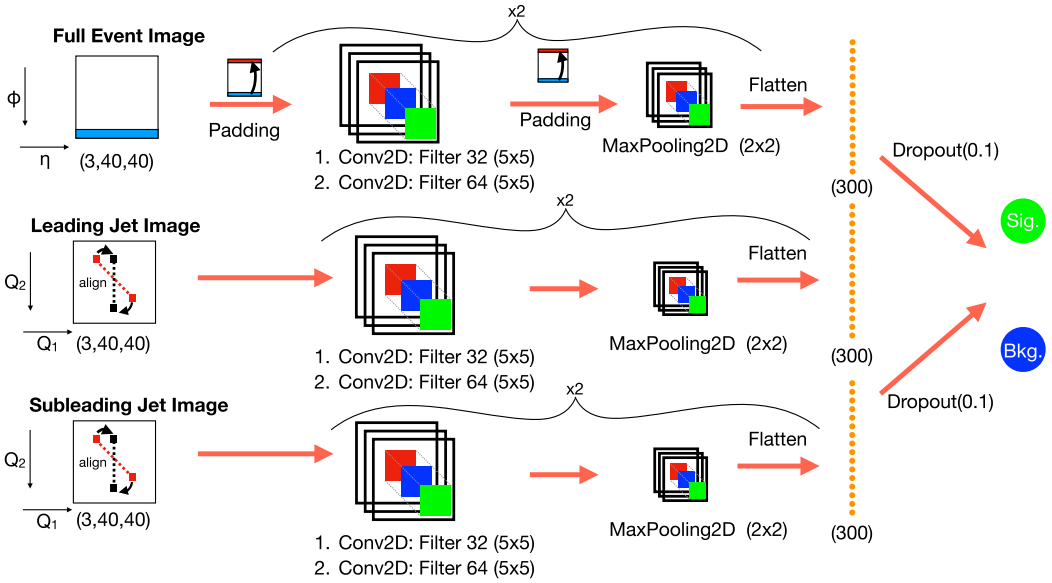

FIG. 5. Architecture of the 3CNN, based on Refs. [26,72]. The first stream (top) is used to process full-event images. The second stream (middle) uses the information from the leading jet. The third stream (bottom) uses the information from the subleading jet.

V. RESULTS AND SENSITIVITY REACH IN 2HDM

The selection for signal and SM backgrounds are shown in Table II. Here we set the light Higgs boson mass $m_h = 125$ GeV, the heavy CP-even scalar mass $M_H = 1000$ GeV, pseudoscalar and charged Higgs mass $M_A = M_{H^\pm} = 1000$ GeV, $m_{12}^2 = 400{,}000$ GeV^2, $\tan\beta = 5$, and $\cos(\beta - \alpha) = 0.01$ in Type II for the benchmark point. We list two major SM backgrounds in this table: $t\bar{t}$ and multijet. The multijet background is the most overwhelming before the selection cuts. The notation **preselection** in Table II is the cut flow that has been described in the Sec. III A. We apply extra B-hadron tagging efficiency = 0.77 [33] for the Higgs jet, requiring double b-tagging via ghosted-associated method, to estimate the event yield. Moreover, to compare the background-discriminating power between the baseline, BDT, and 3CNN, we choose the BDT score cut and 3CNN score cut such that the number of signal events are similar to that in the baseline analysis.

From Table II, we find that the BDT outperforms the baseline. The number of signal events is around 28 while the number of background events is around 1390 events in the baseline analysis. In the BDT analysis, the number of signal events is around 25 and the number of background events is 140; the efficiency of the signal is about the same, while the background rejection power improves by a factor of 10 over the baseline. Finally, in the 3CNN analysis, the number of signal events is 25 and the number of background events is further reduced to 56 events.

Results of these three analyses at $\sqrt{s} = 14$ TeV with an integrated luminosity $\mathcal{L} = 3000$ fb^{-1} are interpreted in the parameter space $(\cos(\beta - \alpha),\ \tan\beta)$ in Fig. 6. We fix

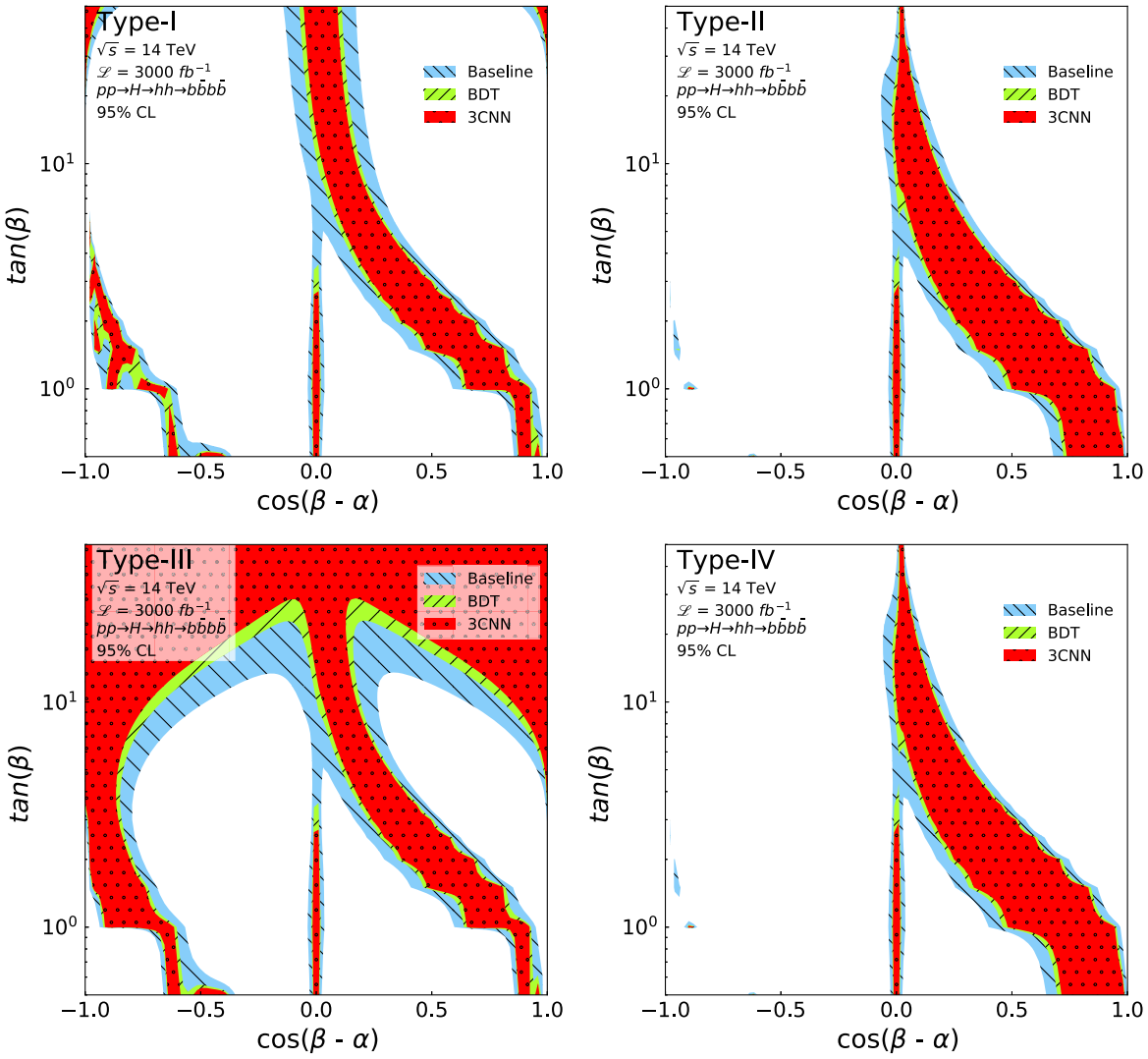

FIG. 6. Allowed regions in all four types of 2HDM for $M_A = M_{H^\pm} = 1000$ GeV, $m_{12}^2 = 400{,}000$ GeV^2 in the $(\cos(\beta - \alpha),\ \tan\beta)$ plane if no excess is seen above the SM background at the HL-LHC. The blue, green, and red regions are the allowed region based on the baseline, BDT, and 3CNN analysis, respectively. Allowed region is the area with significance $\lesssim 2$, significance is $\sqrt{2[(s+b)\ln(1+s/b) - s]}$, where s is the number of signal events and b is the number of background events.

$M_A = M_{H^\pm} = 1000$ GeV, and $m_{12}^2 = 400{,}000$ GeV^2 to find the allowed region at 95% CL in the $(\cos(\beta-\alpha), \tan\beta)$ plane. Note that the colored regions are those with the significance $z \equiv \sqrt{2[(s+b)\ln(1+s/b)-s]} \leq 2$, where s and b stand for the number of signal and background events, respectively. It means that if no excess events are recorded in HL-LHC, the colored regions would be the remaining allowed regions. We see significant gains using the 3CNN analysis for all four types of 2HDMs. The 3CNN analysis has the potential to provide stronger constraints than the baseline method and BDT.

The 3CNN analysis shows stronger background discrimination power and thus providing better coverage of parameter space at the HL-LHC. Therefore, we focus on the 3CNN analysis in the following and combine it with current constraints on Higgs signal strengths obtained at the LHC and direct searches at high-energy colliders. The current constraints are calculated with the public code HIGGSBOUNDS-v5.10.2 and HIGGSSIGNALS-v2.6.2. In the HIGGSBOUNDS-v5.10.2, it includes all processes at LEP, Tevatron, and LHC and determine which is the most sensitive channel and whether the point is still allowed at the 95% CL. In the HIGGSSIGNALS-v2.6.2, it gives the χ^2 output for 111 Higgs observables [44,45,85–90]. Since there are six model parameters, the number of degrees of freedom is 105. We require the p-value to be larger than 0.05, corresponding to 2σ.

In Figs. 7 and 8, we present the sensitivity region (red) with significance $z > 2$ that is still allowed under current constraints and can be covered by the 3CNN at the 14 TeV

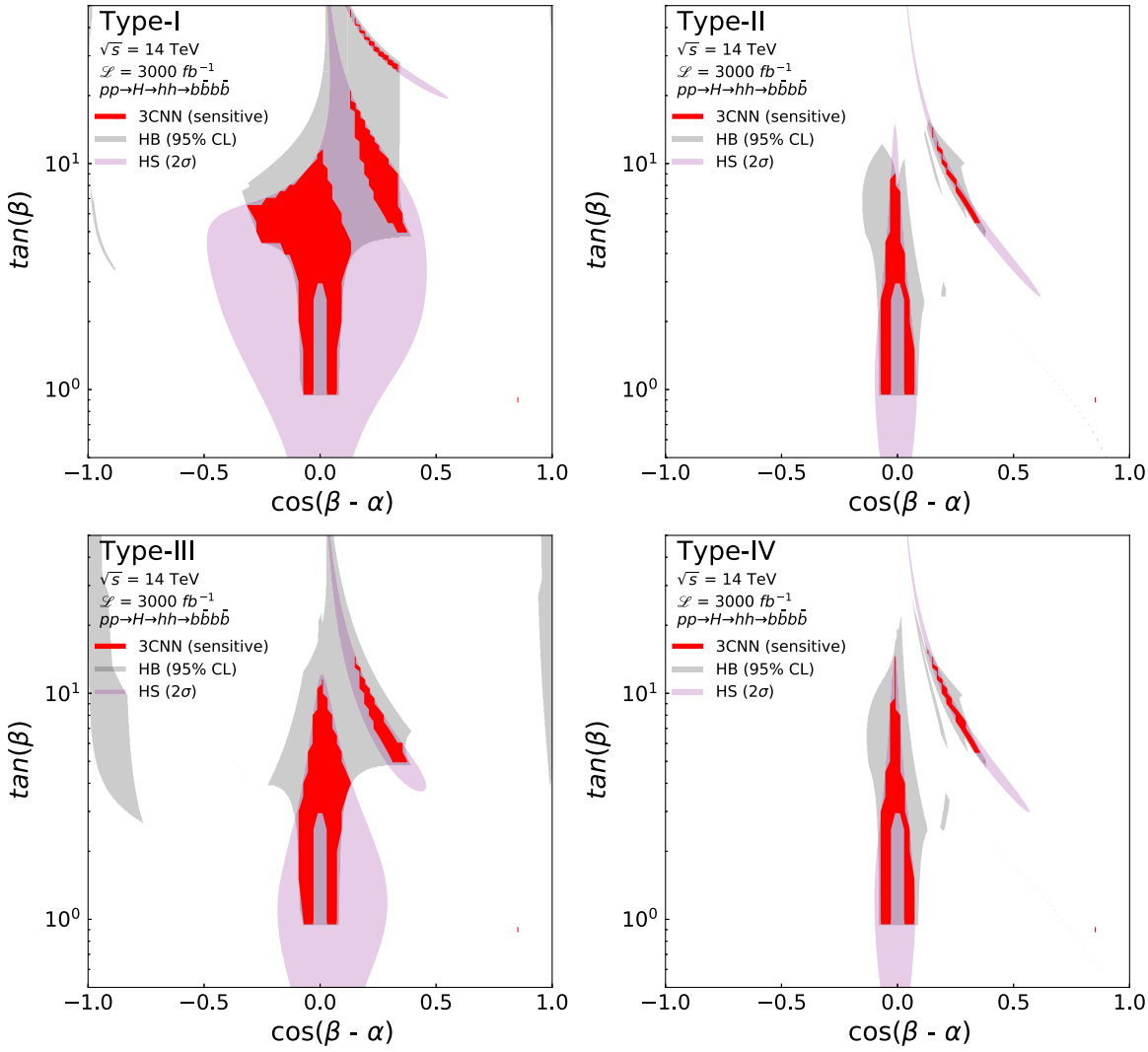

FIG. 7. Sensitive regions in all four types of 2HDM for $M_A = M_{H^\pm} = 1000$ GeV, $m_{12}^2 = 400{,}000$ GeV^2 in the $(\cos(\beta-\alpha), \tan\beta)$ plane. The gray area is the currently allowed area by direct searches at colliders from HIGGSBOUNDS at the 95% C.L. The purple area is due to the constraints from the SM-like Higgs-boson properties from HIGGSSIGNALS at 2σ level. The red region is the sensitive region (significance > 2, where significance is $\sqrt{2[(s+b)\ln(1+s/b)-s]}$, and s is the number of signal events and b is the number of background events) passed the 3CNN analysis, which is still allowed in the current constraints at colliders.

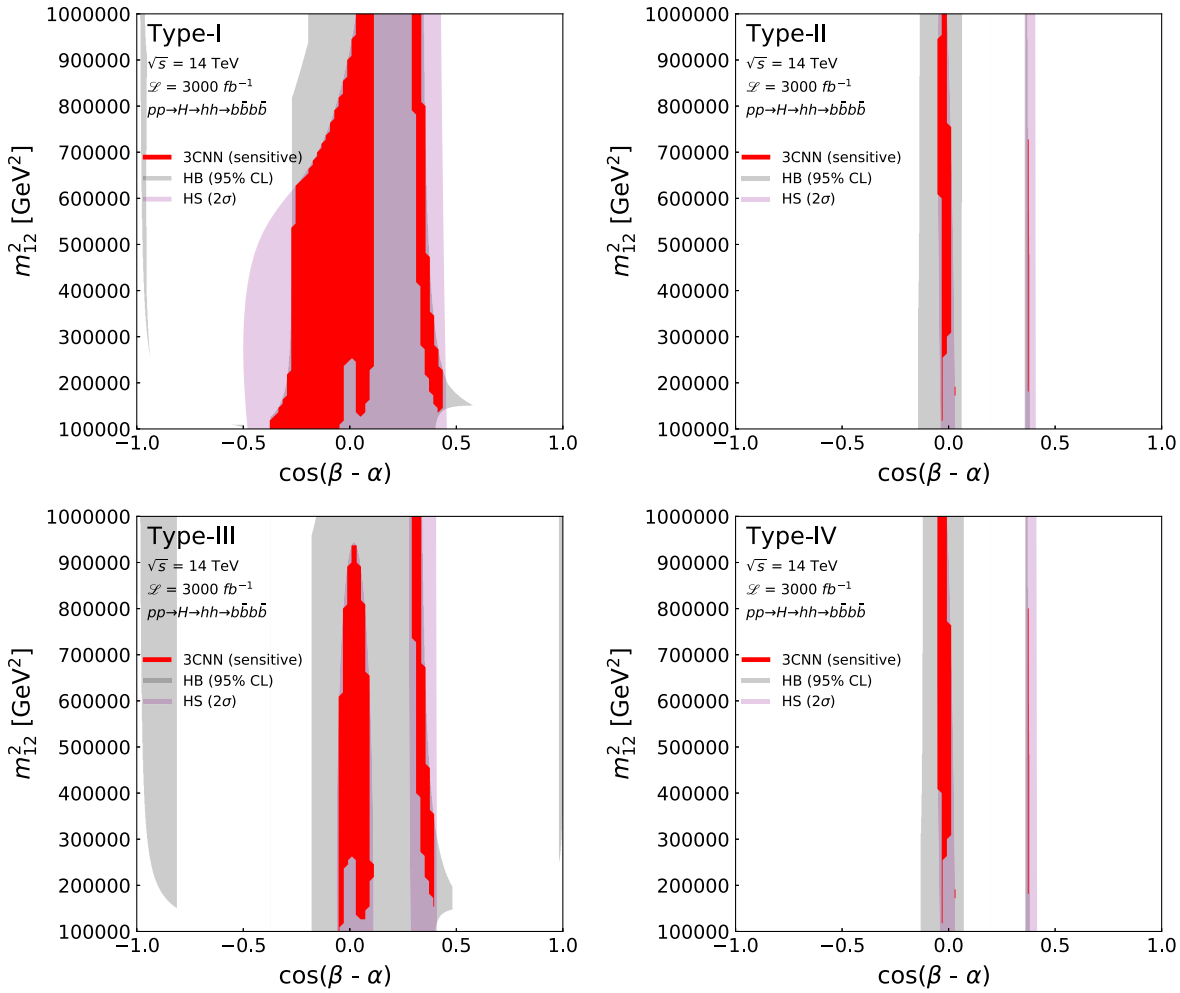

FIG. 8. Sensitive region in all four types of 2HDM for $M_A = M_{H^\pm} = 1000$ GeV, $\tan\beta = 5$ in the $(\cos(\beta-\alpha), m_{12}^2)$ plane. The gray area is the current allowed area by direct searches at colliders from HIGGSBOUNDS at the 95% CL. The purple area is due to the constraints from the SM-like Higgs-boson properties from HIGGSSIGNALS at 2σ level. The red region is the sensitive region (significance >2, significance is $\sqrt{2[(s+b)\ln(1+s/b)-s]}$, where s is the number of signal events and b is the number of background events) passed the 3CNN analysis, where is still allowed in the current constraints at colliders.

HL-LHC in the $(\cos(\beta-\alpha), \tan\beta)$ plane and $(\cos(\beta-\alpha), m_{12}^2)$ plane, respectively. Note that the gray area is the currently allowed region by direct searches at colliders from HIGGSBOUNDS at the 95% CL and the purple area is the allowed region from the SM-like Higgs-boson properties given by HIGGSSIGNALS at 2σ level. We can regard the overlapping regions of the gray and purple areas as the currently allowed parameter space. Note that the overlapping regions can be separated into (i) near the alignment limit and (ii) the wrong-sign Yukawa region. In all 4 types of 2HDM, we see that the 3CNN can cover a large area of the overlapping regions.

In the $(\cos(\beta-\alpha), \tan\beta)$ plane of Fig. 7, we fix $M_A = M_{H^\pm} = 1000$ GeV, and $m_{12}^2 = 400{,}000$ GeV^2. In all 4 types of 2HDM, the red region is the sensitive region where the significance is larger than 2. We can see those large areas (red) in the overlapping region of the gray and purple areas allowed by both HIGGSSIGNALS and HIGGSBOUNDS can be covered by 3CNN, indicating that the process $pp \to H \to hh \to 4b$ can test a large chunk of parameter space at the HL-LHC. We notice that the sensitive regions lie close to the alignment limit, $\cos(\beta-\alpha) = 0$, and in the wrong-sign region in all four types. Around the alignment limit, HIGGSBOUNDS restricts $\tan(\beta)$ to be larger than 1 and the 3CNN analysis shows that the region with $\tan(\beta) \lesssim 10$ is still sensitive. On the other hand, around the wrong sign Yukawa region, the most severe constraint comes from HIGGSSIGNALS. The 3CNN analysis also indicates that it can cover a quite sizable area in the wrong sign Yukawa region.

On the other hand, in the $(\cos(\beta-\alpha), m_{12}^2)$ plane of Fig. 8, we fix $M_A = M_{H^\pm} = 1000$ GeV and $\tan\beta = 5$. The 3CNN analysis indicates the sensitivity in the m_{12}^2 interval around alignment limit and along the wrong sign Yukawa region.

VI. CONCLUSIONS

In this study, we have employed a modern deep-learning approach to improve the search for Higgs boson pair production arising from resonant heavy Higgs enhancement in the $b\bar{b}b\bar{b}$ final state in the framework of two-Higgs-doublet models at the HL-LHC. The resonance production channel plays an important role in probing the structure of the EWSB sector. Using our approach, we have pointed out that the gluon-fusion process $pp \to H \to hh \to 4b$ at the HL-LHC can further probe the currently allowed parameter space of Type I–IV 2HDMs.

The 3CNN architecture in this work is built upon the proposal from Refs. [26,72]. This architecture has 2-class outputs for the signal and background, and contains one stream acting on global event information, and the other two streams acting on information from the leading and subleading jets. This approach is amenable to visualizations that can provide some insights into what the neural network is using for event classification.

We interpret the signal-background discrimination based on our simulations at 14 TeV HL-LHC in the two-Higgs-doublet models' framework. Figures 7 and 8 illustrate our scanning in the parameter space of the 2HDMs for the sensitivity coverage at the HL-LHC, as well as the current restriction on the parameter space due to HIGGSSIGNALS and HIGGSBOUNDS. We find that the 3CNN analysis covers a sizeable region of parameter space.

In summary, we employ the 3CNN architecture to incorporate both local and global information for signal and background identification. Additionally, we have studied the conventional cut-based approach and a boosted decision tree. The conventional cut-based approach does not give enough significance to the signal even at HL-LHC. The BDT is effective but less potent than the neural network. We have shown that the 3CNN can significantly enhance the significance of the signal at HL-LHC and allows us to probe sensitive parameter space in the currently allowed region. This work could be implemented in other Higgs-pair production channels with hadronic or semihadronic final states and may be able to enrich the sensitivity of the signal at the HL-LHC.

ACKNOWLEDGMENTS

We thank SooJin Lee for help with HIGGSSIGNALS and HIGGSBOUNDS, thank Olivier Mattelaer and Kentaro Mawatari for usage on Madgraph. also thank Benjamin Nachman, Chih-Ting Lu and Sean Gasiorowski for their valuable comments on the manuscript, and Alex Schuy for language editing. K. C. and Y. C. were supported by MoST with grant no. MoST-110-2112-M-007-017-MY3. S.-C.H was supported by the National Science Foundation under Grant No. 2110963.

[1] P. W. Higgs, Broken Symmetries and the Masses of Gauge Bosons, Phys. Rev. Lett. **13**, 508 (1964).
[2] F. Englert and R. Brout, Broken Symmetry and the Mass of Gauge Vector Mesons, Phys. Rev. Lett. **13**, 321 (1964).
[3] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Global Conservation Laws and Massless Particles, Phys. Rev. Lett. **13**, 585 (1964).
[4] G. Aad *et al.* (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).
[5] S. Chatrchyan *et al.* (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).
[6] ATLAS Collaboration, A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery, Nature (London) **607**, 52 (2022).
[7] CMS Collaboration, A portrait of the Higgs boson by the CMS experiment ten years after the discovery, Nature (London) **607**, 60 (2022).
[8] G. Aad *et al.* (ATLAS Collaboration), Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment, Phys. Rev. D **101**, 012002 (2020).
[9] A. M. Sirunyan *et al.* (CMS Collaboration), Combined measurements of Higgs boson couplings in proton–proton collisions at $\sqrt{s} = 13$TeV, Eur. Phys. J. C **79**, 421 (2019).
[10] P. F. de Salas, D. V. Forero, C. A. Ternes, M. Tortola, and J. W. F. Valle, Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved *CP* sensitivity, Phys. Lett. B **782**, 633 (2018).
[11] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2015 results. XIII. Cosmological parameters, Astron. Astrophys. **594**, A13 (2016).
[12] G. R. Farrar and M. E. Shaposhnikov, Baryon asymmetry of the universe in the standard electroweak theory, Phys. Rev. D **50**, 774 (1994).
[13] H. Georgi, H. R. Quinn, and S. Weinberg, Hierarchy of Interactions in Unified Gauge Theories, Phys. Rev. Lett. **33**, 451 (1974).
[14] E. W. N. Glover and J. J. van der Bij, Higgs boson pair production via gluon fusion, Nucl. Phys. **B309**, 282 (1988).
[15] D. A. Dicus, C. Kao, and S. S. D. Willenbrock, Higgs boson pair production from gluon fusion, Phys. Lett. B **203**, 457 (1988).

[16] T. Plehn, M. Spira, and P. M. Zerwas, Pair production of neutral Higgs particles in gluon-gluon collisions, Nucl. Phys. **B479**, 46 (1996); **B531**, 655(E) (1998).

[17] A. Djouadi, W. Kilian, M. Muhlleitner, and P. M. Zerwas, Production of neutral Higgs boson pairs at LHC, Eur. Phys. J. C **10**, 45 (1999).

[18] S. Dawson, S. Dittmaier, and M. Spira, Neutral Higgs boson pair production at hadron colliders: QCD corrections, Phys. Rev. D **58**, 115012 (1998).

[19] U. Baur, T. Plehn, and D. L. Rainwater, Determining the Higgs Boson self-coupling at Hadron Colliders, Phys. Rev. D **67**, 033003 (2003).

[20] T. Binoth, S. Karg, N. Kauer, and R. Ruckl, Multi-Higgs boson production in the Standard Model and beyond, Phys. Rev. D **74**, 113008 (2006).

[21] U. Baur, T. Plehn, and D. L. Rainwater, Examining the Higgs boson potential at lepton and hadron colliders: A comparative analysis, Phys. Rev. D **68**, 033001 (2003).

[22] 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.

[23] J. Grigo, J. Hoff, K. Melnikov, and M. Steinhauser, On the Higgs boson pair production at the LHC, Nucl. Phys. **B875**, 1 (2013).

[24] V. Barger, L. L. Everett, C. B. Jackson, and G. Shaughnessy, Higgs-pair production and measurement of the triscalar coupling at LHC (8,14), Phys. Lett. B **728**, 433 (2014).

[25] C.-T. Lu, J. Chang, K. Cheung, and J. S. Lee, An exploratory study of Higgs-boson pair production, J. High Energy Phys. 08 (2015) 133.

[26] Y.-L. Chung, S.-C. Hsu, and B. Nachman, Disentangling boosted Higgs boson production modes with machine learning, J. Instrum. **16**, P07002 (2021).

[27] J. Chang, K. Cheung, J. S. Lee, C.-T. Lu, and J. Park, Higgs-boson-pair production $H(\to b\bar{b})\, H(\to \gamma\gamma)$ from gluon fusion at the HL-LHC and HL-100 TeV hadron collider, Phys. Rev. D **100**, 096001 (2019).

[28] J. Chang, K. Cheung, J. S. Lee, and J. Park, Probing the trilinear Higgs boson self-coupling at the high-luminosity LHC via multivariate analysis, Phys. Rev. D **101**, 016004 (2020).

[29] J. H. Kim, K. Kong, K. T. Matchev, and M. Park, Probing the Triple Higgs Self-Interaction at the Large Hadron Collider, Phys. Rev. Lett. **122**, 091801 (2019).

[30] A. Papaefstathiou, L. L. Yang, and J. Zurita, Higgs boson pair production at the LHC in the $b\bar{b}W^+W^-$ channel, Phys. Rev. D **87**, 011301 (2013).

[31] J. Amacker *et al.*, Higgs self-coupling measurements using deep learning in the $b\bar{b}b\bar{b}$ final state, J. High Energy Phys. 12 (2020) 115.

[32] A. Adhikary, S. Banerjee, R. Kumar Barman, and B. Bhattacherjee, Resonant heavy Higgs searches at the HL-LHC, J. High Energy Phys. 09 (2019) 068.

[33] G. Aad *et al.* (ATLAS Collaboration), Search for resonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Rev. D **105**, 092002 (2022).

[34] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, Theory and phenomenology of two-Higgs-doublet models, Phys. Rep. **516**, 1 (2012).

[35] S. L. Glashow and S. Weinberg, Natural conservation laws for neutral currents, Phys. Rev. D **15**, 1958 (1977).

[36] E. A. Paschos, Diagonal neutral currents, Phys. Rev. D **15**, 1966 (1977).

[37] J. Song and Y. W. Yoon, $W\gamma$ decay of the elusive charged Higgs boson in the two-Higgs-doublet model with vector-like fermions, Phys. Rev. D **100**, 055006 (2019).

[38] M. Aaboud *et al.* (ATLAS Collaboration), Search for Higgs bosons produced via vector-boson fusion and decaying into bottom quark pairs in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector, Phys. Rev. D **98**, 052003 (2018).

[39] M. Aaboud *et al.* (ATLAS Collaboration), Measurements of gluon-gluon fusion and vector-boson fusion Higgs boson production cross-sections in the $H \to WW^* \to e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Lett. B **789**, 508 (2019).

[40] M. Aaboud *et al.* (ATLAS Collaboration), Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Rev. D **99**, 072001 (2019).

[41] G. Aad *et al.* (ATLAS Collaboration), Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector, Eur. Phys. J. C **80**, 957 (2020); **81**, 29 (2021); **81**, 398(E) (2021).

[42] A. M. Sirunyan *et al.* (CMS Collaboration), Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 03 (2019) 026.

[43] A. M. Sirunyan *et al.* (CMS Collaboration), Search for the Higgs Boson Decaying to Two Muons in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **122**, 021801 (2019).

[44] CMS Collaboration, Measurements of properties of the Higgs boson in the four-lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV, CERN, Geneva, Technical Report No. CMS-PAS-HIG-19-001, 2019.

[45] CMS Collaboration, Measurements of differential Higgs boson production cross sections in the leptonic WW decay mode at $\sqrt{s} = 13$ TeV, CERN, Geneva, Technical Report No. CMS-PAS-HIG-19-002.

[46] P. Bechtle, D. Dercks, S. Heinemeyer, T. Klingl, T. Stefaniak, G. Weiglein, and J. Wittbrodt, HiggsBounds-5: Testing Higgs Sectors in the LHC 13 TeV Era, Eur. Phys. J. C **80**, 1211 (2020).

[47] P. Bechtle, O. Brein, S. Heinemeyer, G. Weiglein, and K. E. Williams, HiggsBounds: Confronting arbitrary Higgs sectors with exclusion bounds from LEP and the Tevatron, Comput. Phys. Commun. **181**, 138 (2010).

[48] P. Bechtle, O. Brein, S. Heinemeyer, G. Weiglein, and K. E. Williams, HiggsBounds 2.0.0: Confronting neutral and charged Higgs sector predictions with exclusion bounds from LEP and the Tevatron, Comput. Phys. Commun. **182**, 2605 (2011).

[49] P. Bechtle, O. Brein, S. Heinemeyer, O. Stal, T. Stefaniak, G. Weiglein, and K. Williams, Recent developments in HiggsBounds and a preview of HiggsSignals, *Proc. Sci.*, CHARGED2012 (2012) 024 [arXiv:1301.2345].

[50] P. Bechtle, O. Brein, S. Heinemeyer, O. Stål, T. Stefaniak, G. Weiglein, and K. E. Williams, HiggsBounds – 4: Improved Tests of Extended Higgs Sectors against Exclusion Bounds from LEP, the Tevatron and the LHC, Eur. Phys. J. C **74**, 2693 (2014).

[51] P. Bechtle, S. Heinemeyer, O. Stal, T. Stefaniak, and G. Weiglein, Applying exclusion likelihoods from LHC searches to extended Higgs sectors, Eur. Phys. J. C **75**, 421 (2015).

[52] O. Stål and T. Stefaniak, Constraining extended Higgs sectors with HiggsSignals, *Proc. Sci.*, EPS-HEP2013 (2013) 314 [arXiv:1310.4039].

[53] P. Bechtle, S. Heinemeyer, O. Stål, T. Stefaniak, and G. Weiglein, *HiggsSignals*: Confronting arbitrary Higgs sectors with measurements at the Tevatron and the LHC, Eur. Phys. J. C **74**, 2711 (2014).

[54] P. Bechtle, S. Heinemeyer, O. Stål, T. Stefaniak, and G. Weiglein, Probing the Standard Model with Higgs signal rates from the Tevatron, the LHC and a future ILC, J. High Energy Phys. 11 (2014) 039.

[55] P. Bechtle, S. Heinemeyer, T. Klingl, T. Stefaniak, G. Weiglein, and J. Wittbrodt, HiggsSignals-2: Probing new physics with precision Higgs measurements in the LHC 13 TeV era, Eur. Phys. J. C **81**, 145 (2021).

[56] 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.

[57] T. Sjostrand, S. Mrenna, and P. Z. Skands, A brief introduction to PYTHIA 8.1, Comput. Phys. Commun. **178**, 852 (2008).

[58] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali (NNPDF Collaboration), Parton distributions for the LHC Run II, J. High Energy Phys. 04 (2015) 040.

[59] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali, Parton distributions with LHC data, Nucl. Phys. **B867**, 244 (2013).

[60] C. Degrande, Automatic evaluation of UV and R2 terms for beyond the Standard Model Lagrangians: A proof-of-principle, Comput. Phys. Commun. **197**, 239 (2015).

[61] D. Eriksson, J. Rathsman, and O. Stal, 2HDMC: Two-Higgs-doublet model calculator physics and manual, Comput. Phys. Commun. **181**, 189 (2010).

[62] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations, J. High Energy Phys. 03 (2013) 015.

[63] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, Matching matrix elements and shower evolution for top-pair production in hadronic collisions, J. High Energy Phys. 01 (2007) 013.

[64] J. Alwall *et al.*, Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions, Eur. Phys. J. C **53**, 473 (2008).

[65] N. Dawe, E. Rodrigues, H. Schreiner, B. Ostdiek, D. Kalinkin, R. Marcel, and S. Meehan, aryan26roy, and domen13, scikit-hep/pyjet: Version 1.8.2, 2021.

[66] M. Cacciari, G. P. Salam, and G. Soyez, FastJet user manual, Eur. Phys. J. C **72**, 1896 (2012).

[67] M. Cacciari, G. P. Salam, and G. Soyez, The anti-k_t jet clustering algorithm, J. High Energy Phys. 04 (2008) 063.

[68] D. Krohn, J. Thaler, and L.-T. Wang, Jet trimming, J. High Energy Phys. 02 (2010) 084.

[69] S. D. Ellis and D. E. Soper, Successive combination jet algorithm for hadron collisions, Phys. Rev. D **48**, 3160 (1993).

[70] M. Cacciari, G. P. Salam, and G. Soyez, The catchment area of jets, J. High Energy Phys. 04 (2008) 005.

[71] A. Buckley and C. Pollard, QCD-aware partonic jet clustering for truth-jet flavour labelling, Eur. Phys. J. C **76**, 71 (2016).

[72] J. Lin, M. Freytsis, I. Moult, and B. Nachman, Boosting $H \to b\bar{b}$ with machine learning, J. High Energy Phys. 10 (2018) 101.

[73] M. Aaboud *et al.* (ATLAS Collaboration), Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 08 (2018) 089.

[74] J. Thaler and K. Van Tilburg, Identifying boosted objects with N-subjettiness, J. High Energy Phys. 03 (2011) 015.

[75] J. Thaler and K. Van Tilburg, Maximizing boosted top identification by minimizing N-subjettiness, J. High Energy Phys. 02 (2012) 093.

[76] A. J. Larkoski, I. Moult, and D. Neill, Power counting to better jet observables, J. High Energy Phys. 12 (2014) 009.

[77] A. J. Larkoski, G. P. Salam, and J. Thaler, Energy correlation functions for jet substructure, J. High Energy Phys. 06 (2013) 108.

[78] J. Cogan, M. Kagan, E. Strauss, and A. Schwartzman, Jet-images: Computer vision inspired techniques for jet tagging, J. High Energy Phys. 02 (2015) 118.

[79] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, Jet-images—Deep learning edition, J. High Energy Phys. 07 (**2016**) 069.

[80] L. de Oliveira, M. Paganini, and B. Nachman, Learning particle physics by example: Location-aware generative adversarial networks for physics synthesis, Comput. Software Big Sci. **1**, 4 (2017).

[81] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. **12**, 2825 (2011).

[82] M. D. Zeiler, ADADELTA: An adaptive learning rate method, CoRR (2012), arXiv:1212.5701.

[83] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, Dropout: A simple way to prevent neural

networks from overfitting, J. Mach. Learn. Res. **15**, 1929 (2014).

[84] M. Abadi *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[85] M. Aaboud *et al.* (ATLAS Collaboration), Search for Higgs bosons produced via vector-boson fusion and decaying into bottom quark pairs in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector, Phys. Rev. D **98**, 052003 (2018).

[86] M. Aaboud *et al.* (ATLAS Collaboration), Measurements of gluon-gluon fusion and vector-boson fusion Higgs boson production cross-sections in the $H \to WW^* \to e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Lett. B **789**, 508 (2019).

[87] M. Aaboud *et al.* (ATLAS Collaboration), Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Rev. D **99**, 072001 (2019).

[88] G. Aad *et al.* (ATLAS Collaboration), Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector, Eur. Phys. J. C **80**, 957 (2020); **81**, 29(E) (2021); **81**, 398(E) (2021).

[89] A. M. Sirunyan *et al.* (CMS Collaboration), Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 03 (2019) 026.

[90] A. M. Sirunyan *et al.* (CMS Collaboration), Search for the Higgs Boson Decaying to Two Muons in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **122**, 021801 (2019).