Influence of QCD parton showers in deep learning invisible Higgs bosons through vector boson fusion

Partha Konar[1,*] and Vishal S. Ngairangbam[1,2,†]

[1]*Physical Research Laboratory, Ahmedabad—380009, Gujarat, India*
[2]*Indian Institute of Technology, Gandhinagar—382424, Gujarat, India*

(Received 11 January 2022; accepted 15 May 2022; published 14 June 2022)

Vector boson fusion established itself as a highly reliable channel to probe the Higgs boson and an avenue to uncover new physics at the Large Hadron Collider. This channel provides the most stringent bound on Higgs's invisible decay branching ratio, where the current upper limits are significantly higher than the one expected in the Standard Model. It is remarkable that merely low-level calorimeter data from this characteristically simple process can improve this limit substantially by employing sophisticated deep learning techniques. The construction of such neural networks seems to comprehend the event kinematics and radiation pattern exceptionally well. However, the full potential of this outstanding capability also warrants a precise theoretical projection of QCD parton showering and corresponding radiation pattern. This work demonstrates the relation using different recoil schemes in the parton shower with leading-order and higher-order computation.

DOI: 10.1103/PhysRevD.105.113003

I. INTRODUCTION

The discovery of the last missing piece [1,2] of the Standard Model (SM) of particle physics opened up a plethora of independent searches at the CERN Large Hadron Collider. Apart from looking for new physics signatures, it is paramount to verify whether the scalar is "the Higgs" boson or its properties have some nonstandard nature [3–8] different from those predicted in the SM. With no direct evidence of new physics so far, there is still scope that it can show up in the Higgs boson's physical properties. Hence, the precise determination of Higgs's intrinsic property and coupling with other fundamental particles can provide subtle hints toward new physics.

Although with a relatively lower cross section, the vector boson fusion (VBF) channel [9–13] has long been advocated for as a cleaner alternative in hardon colliders among different Higgs production mechanisms. Eventually, this channel also showed promise in several new physics searches [14–20]. The VBF channel provides the best channel for constraining the Higgs invisible decays [14], providing a strong bound in several dark matter [21–26] scenarios. It is particularly important since the current upper limit of 11% [8] is still significantly larger than the one expected in the SM ($\lesssim$0.1% with the decay chain $h \to ZZ^* \to \nu\bar{\nu}\nu'\bar{\nu}'$). In general, the VBF channel relies on the production of color-neutral heavy states through the collision of electroweak gauge bosons radiated from initial partons and is always associated with two hard forward jets tagged as its characteristic signature. In the absence of colored particle exchange between two parton lines, very little QCD jet activity [10,11] is seen in the detector's central part, in the region between two forward tagged jets. The decay products of the heavy state (in the present case, the Higgs) is expected to be in this region, retaining the color quietness. These characteristic features are well studied and greatly exploited to identify VBF processes over various QCD backgrounds, exhibiting a different jet formation pattern. In this respect, different high-level variables are constructed as proxies for these features. Among them, the rapidity gap, central-jet veto, invariant mass of two forward jets, and Zeppenfeld variable ($z^*_{j_3}$) are some of the crucial ones. The precise theoretical prediction [27–30] of additional jet formation patterns is as important as the experimental measurements. This precise prediction is even more significant in a data-driven paradigm where different characteristics of both signal and background distributions are minutely encapsulated to determine the decision boundary.

The application of machine-learning algorithms has shown immense promise at the LHC [31–36]. They have significantly improved the performance compared to variables motivated by our knowledge of physics. However, this comes at the cost of a reduced understanding, which is

*konar@prl.res.in
†vishalng@prl.res.in

even more profound when we use low-level data to train deep neural networks (DNNs). Moreover, training with simulated data can result in the networks learning features specific to the imperfect simulation, not present in real experimental data. The use of low-level inclusive event information for searches brings a few unique challenges, particularly due to the higher systematic uncertainties associated with the simulations. Given the superior ability of deep learning algorithms to extract features, it is necessary to circumspect the effect of using simulations which better resemble the actual physics.

It is well known that VBF processes receive modest corrections from higher-order QCD corrections [27–30]. Using the full event information in an input image to a convolutional neural network (CNN) for VBF searches has shown a promising avenue [37], where the CNN exploits the lack of central-jet activity in VBF processes. We note that it is essential to scrutinize the differences in leading-order (LO) and next-to-leading-order (NLO) simulations; the presence of a third jet needs the proper introduction of real and virtual corrections to the tree-level process. Another issue of central importance in the simulation of VBF events is the inability of a global-recoil scheme in initial-state radiations (ISRs) of the parton showering algorithm to describe the central-jet activity correctly [38]. This study systematically investigates these issues for the VBF signal search with CNNs, taking the invisible decay as a proxy. Although the preceding arguments apply to the VBF production of weak bosons, we presently ignore its effects as they are much less in proportion (∼5% of the total background for the cuts used here). We also neglect the contribution of the gluon-fusion events in the signal. The global-recoil scheme correctly produces the leading logarithmic behavior, already incorporated in our previous analysis. A precise determination of its various effects demands a very high level of sophistication, requiring much higher perturbative and logarithmic accuracy. Although the cuts used in the analysis have a sizable amount of gluon-fusion contribution, the large amount of data from high-luminosity LHC runs will provide ways to do precision analysis with more stringent cuts, with negligible contribution from gluon-fusion events. These do not impede our final goal, as our intention is not to project experimental sensitivities but to usher pragmatism and careful examination while using inclusive event information as inputs to DNNs.

Although DNNs generally perform better than machine learning (ML) algorithms utilizing high-level variables, their usability in phenomelogical analyses is determined by our ability to simulate subtle aspects of the data accurately. To this end, we show the possibility of CNNs learning inaccurate representations of inclusive events due to a global recoil used in the simulation of VBF events. We find the following:

(i) The training performance is greatly reduced when we use signal simulated with a global recoil scheme on parton-level events generated with leading-order or next-to-leading-order accuracy and improves for a dipole recoil, with events generated at next-to-leading-order accuracy showered with a dipole recoil having the highest training accuracy.

(ii) For each set of signal simulations, the highest validation accuracy is achieved for the network that used the same type during the training process with the same trend as the training accuracies. However, the validation performance of the NLO events showered with dipole recoil (which is the most accurate description of the actual events among the four signals used) is affected mildly by the kind of data used during the training.

Our findings show that CNNs can learn the underlying differences between VBF-type events and the dominant QCD backgrounds, even when trained on suboptimal simulated data.

The rest of this paper is organized as follows. In Sec. II, we outline the CNN-based improvement over the existing study and point out the significance of QCD radiation for better accuracy. Section III discusses the parton shower scheme and NLO effects in the simulation of VBF Higgs signal. In Sec. IV, we examine the impact of the different signal simulations on the trained network output and its performance. We conclude in Sec. V.

II. DEEP LEARNING INVISIBLE HIGGS PRODUCED VIA VBF

In this section, we summarize the VBF search of the invisible Higgs decays proposed in Ref. [37] using deep learning. Since the present study aims to scrutinize the dependence of deep learning algorithms with low-level inputs on the simulation, we focus on the analysis using CNNs with tower images as the input.

A. Data simulation and selection criteria

The background class consists of non-VBF- and VBF-type production of Z and W bosons with at least two hard jets, with the Z boson decaying to neutrinos and the W decaying to a charged lepton and a neutrino (contributes when the lepton fails the identification criteria). The parton-level events are generated using MadGraph5_aMC@NLO [39] (version 2.6.5) at 13 TeV LHC. PYTHIA8 [40] (version 8.243) is used to shower these events in the default global-recoil scheme for the ISR. We match parton-level cross sections of processes where the additional jets arise from QCD vertices via the Michelangelo L. Mangano (MLM), procedure [41,42] (up to four jets for Z and two for W). The showered events were passed through DELPHES3 [43] (version 3.4.1) for simulating a parametrized response of the CMS detector. Jets of radius $R = 0.5$ and $p_T > 20$ are clustered with the anti-k_t [44] algorithm implemented in the FASTJET [45] package.

To have a standard benchmark, we followed the experimental search in Ref. [46] and replicated its shape analysis[1] on our simulated data. Except for the cut on the missing transverse momentum ($\not{E}_T > 200$ GeV instead of $\not{E}_T > 250$ GeV), we implemented the same baseline selection criteria in the shape analysis, which is summarized in the Appendix. Note that the analysis used weaker cuts on the rapidity gap between the two tagging jets $|\Delta\eta_{jj}| > 1$, and the dijet invariant mass $m_{jj} > 200$ GeV, than is normally done in VBF searches, and had no central-jet veto. We therefore expect a significant contribution from gluon-fusion production of the Higgs with two additional jets, along with an increase in selected backgrounds. Thus, the signal class consisted of VBF and gluon-fusion contributions in Ref. [37], decaying to two invisible dark matter particles. The weaker cuts without any central-jet veto retain a larger fraction of the VBF signal, which can be segregated using powerful deep learning techniques. The background and signal classes were formed by combining the different processes according to their expected proportions after the baseline selection criteria. We used 100,000 training and 25,000 validation events for each category. The background class consists of 51.221% and 44.896% of non-VBF Z and $W^{\pm}$ backgrounds, respectively, and 2.295% and 1.587% of VBF-origin Z and $W^{\pm}$ backgrounds, respectively.

B. Input representation and preprocessing

Although the analysis consisted of various high-level and low-level input representations with different preprocessing steps, we concentrate on the low-level representation with better regularized preprocessing steps. We first apply the following operations to the 4-momenta of each particle in each event to regularize their spatial orientations:

(1) Rotate along z axis such that the leading jet's center resides on the z axis ($\phi_{j_1} = 0$).
(2) Reflect along the xy plane, such that the leading jet's η is always positive.

For inputs to CNNs, we form the tower image of the full calorimeter in the (η, ϕ) plane, with the transverse energy E_T as the pixel values, with two resolutions: 0.08×0.08 and 0.17×0.17. It might look like the leftovers due to imperfect divisibility by the pixel resolution from the full range of ϕ can contribute to a boundary effect. However, it is not detrimental to a CNN's performance since the pooling operations would effectively wash away its impact. Moreover, the images formed after the preprocessing sets $\phi_{j_1} = 0$ further reduce the importance of the boundary by concentrating all useful information at the center. Periodicity of the ϕ axis is enforced by padding these images with a fixed number of rows from the opposite side. Therefore, we get preprocessed high- and low-resolution images represented by $\mathcal{P}_J^{HR}$ and $\mathcal{P}_J^{LR}$, respectively.

C. Network architecture and training

The networks for the input images, $\mathcal{P}_J^{HR}$-CNN and $\mathcal{P}_J^{LR}$-CNN, have three blocks of sequential convolution-pooling operations. Each block consists of two convolutions with 64 4×4 filters and an average pooling layer with a 2×2 pool size. The flattened output of the final pooling layer is fed to a dense network with three hidden layers containing 300 nodes each and an output layer with two nodes with a SOFTMAX activation for binary classification. All convolutional layers and the hidden dense layers have ReLu activation. The network was trained for 20 epochs with a NADAM [49] optimizer with a learning rate of 0.001, with cross-entropy loss and a batch size of 300. All trainings were implemented using KERAS (version 2.2.4) [50] with a TensorFlow (version 1.14.1) [51] backend.

The exclusion limits on the invisible branching ratio of the Higgs boson obtained from the network output was compared to the shape analysis of m_{jj} and $|\Delta\eta_{jj}|$. The expected upper limits for an integrated luminosity $L = 300$ fb^{-1} of the four scenarios are shown in Fig. 1. These limits were obtained using the CL$_s$ method [52,53] in the asymptotic approximation [54], with the RooStats [55] package. The statistical model was built using HistFactory [56]. The uncertainty bands are obtained by incorporating the per-bin statistical uncertainty and the normalization uncertainties of the total cross section, Monte Carlo simulation, and the integrated luminosity. The CNN-based approach has better performance, putting stricter limits on the branching ratio. Moreover, it was found that training and validation on events with pileup did not impact the performance and upper bounds on the branching ratio considerable, increasing it mildly within

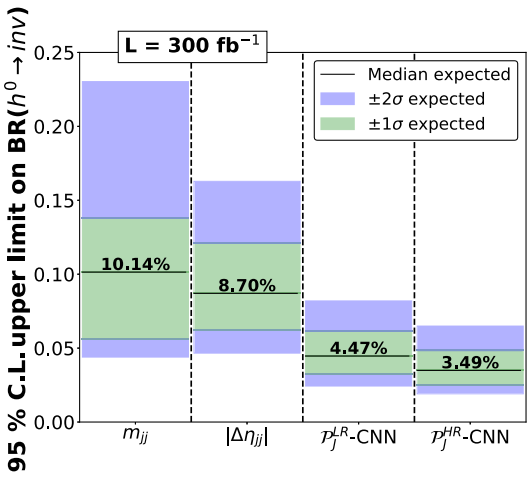


FIG. 1. Expected upper limits on the invisible branching ratio of the Higgs boson, for an integrated luminosity $L = 300$ fb^{-1} from the analysis in Ref. [37].

[1]This study did not use any ML techniques for the final analysis; recent analyses [47,48] on VBF searches have used ML techniques with high-level variables. However, searches using low-level detector information has not been performed for VBF searches to the best of our knowledge.

the one-sigma errors obtained from the one without pileup. Our study [37] did not apply any pileup mitigation technique on the pileup contaminated events and was trained and validated with tower images contaminated with pileup. With extensive research and progress into process independent pileup contamination procedures, we expect this pileup contribution to reduce further. The lower transverse energy in the secondary pileup collisions significantly affects the forward regions. In contrast, for VBF processes, the parton-shower recoil schemes primarily involve the soft radiation in the central rapidity gap between the two (hard) forward jets. Moreover, as will become clear from the results, CNNs primarily look at the amount of hard radiation between the two jets. Therefore, pileup effects would not conceal the impact of a third hard jet in VBF events, thereby demanding an accurate description of such hard jets, which is achieved up to leading-order in a next-to-leading-order simulation of the complete process. Therefore, it is imperative to scrutinize these factors when dealing with deep learning algorithms like CNNs, which enhance the physics reach by setting more stringent bounds on the invisible branching ratio of the Higgs boson by utilizing the inclusive event information efficiently.

III. IMPACT OF NLO CORRECTIONS AND RECOIL SCHEMES

Although VBF processes have relatively lower higher-order corrections, utilizing the hadronic activity between the two tagging jets would use information not captured by a leading-order simulation. This inadequacy is due to the inherent assumption in parton shower generators, primarily focusing on the soft and collinear regions. A next-to-leading-order hard partonic simulation merged with a parton shower algorithm would accurately describe the kinematics of the third leading jet (if present) over the full range of transverse momentum. Additionally, for event topologies with no color flow between the two incoming partons from the colliding protons, a parton shower algorithm with a global-recoil scheme for the ISR is known to have a further inefficient simulation of the wide-angle soft radiation patterns. The cause for this inaccuracy is due to the incorrect assumption of an `II` dipole in the global-recoil scheme [38], while VBF processes have a double deep inelastic scattering topology with an `IF/FI` dipole structure. The above phrasing indicates whether the partons act as a radiator and recoiler are in the final (`F`) or initial (`I`) state. Existing phenomenological studies [57,58] are consistent with this known limitation of the global-recoil scheme, and recent experimental results [59–61] have employed the dipole-recoil scheme [62–65] for the relevant VBF topologies. The effects of both higher-order virtual corrections and the recoil scheme are even more important when using powerful deep learning algorithms with low-level inputs.

A. Signal generation

Since VBF Higgs processes are our primary interest, we do not include the gluon-fusion processes in the present analysis. We, therefore, study the four different possible combinations of the perturbative accuracy and the parton shower's recoil scheme for the VBF channel. These are described as follows:

(1) *Global LO.*—Parton-level events simulated at leading-order perturbative accuracy showered with a global-recoil scheme for the ISR parton shower. This recoil scheme is the default implementation in PYTHIA8 and was used in Ref. [37] for the VBF processes.
(2) *Dipole LO.*—Parton-level events simulated at leading-order perturbative accuracy showered with a dipole-recoil scheme.
(3) *Global NLO.*—Parton-level events simulated at next-to-leading-order accuracy merged with parton shower employing the global-recoil scheme for ISR.
(4) *Dipole NLO.*—Parton-level events simulated at next-to-leading-order accuracy merged with parton shower using the dipole-recoil scheme.

We use the same set of parton-level events for the LO and NLO simulations to shower with the two recoil schemes. The parton-level events at LO were generated with MadGraph5_aMC@NLO, while the NLO events were generated with POWHEG-BOX [66–69]. The renormalization and factorization scales for both orders are set for each event as

$$\mu_0^2 = \frac{m_h}{2}\sqrt{\left(\frac{m_h}{2}\right)^2 + p_{T,h}^2}, \tag{1}$$

where $m_h = 125$ GeV is the mass of the Higgs and $p_{T,h}$ is the transverse momentum of the Higgs boson in the event. For the parton-level generation, we use the PDF4LHC15_NLO_100_PDFAS [70] parton distribution function (PDF) set implemented with LHAPDF6 [71] (version 6.1.6) package. This PDF set is a combination [72] of CT14 [73], MMHT14 [74], and NNPDF3.0 [75] PDF sets using the Hessian reduction method proposed in Ref. [76]. We use MadSpin [77] to decay the Higgs boson at parton level to two scalar dark matter particles for the NLO events, while we simulate the full decay chain for the LO events. All parton showers are performed in PYTHIA8.235. For the NLO events, we perform the POWHEG merging with recommended values from Ref. [78]. The switch to a dipole-recoil scheme is done by setting "SpaceShower:dipoleRecoil = on" for the parton shower. We note that the events generated at NLO and showered with the dipole-recoil scheme should be the most physically accurate simulation of the VBF Higgs process. These four sets of showered events are then passed through the same detector simulation and selection criteria[2]

[2]The details of the baseline selection are given in Appendix.

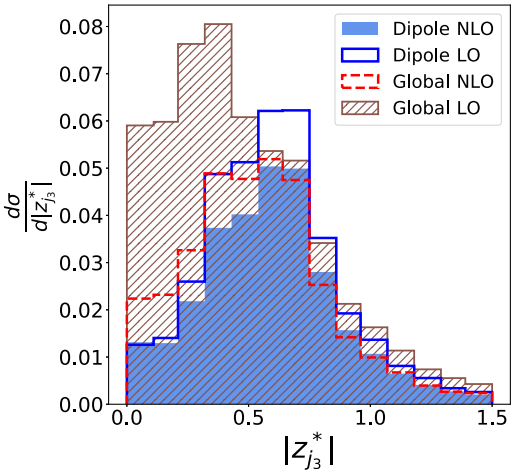


FIG. 2. Distribution of the absolute Zeppenfeld variable $|z^*_{j_3}|$ for the four signal simulations. To capture the relative occurrence of the third jet, we set each event weight so that the total sum with or without an additional jet in each signal simulation sum to unity.

described in Sec. II A. We divide the dataset of each of these simulations into 100,000 training and 25,000 validation samples for the neural network analysis.

B. Characteristics of the third jet

To compare the different signal simulations, we plot distributions of the Zeppenfeld variable $z^*_{j_3}$ in Fig. 2 for events passing the selection criteria and having a third jet with $p_T > 20$ GeV. It is defined as

$$z^*_{j_3} = \frac{\eta_{j_3} - (\eta_{j_1} + \eta_{j_2})/2}{|\Delta\eta_{j_1 j_2}|}, \tag{2}$$

where η_{j_i} is the pseudorapidity of the ith hardest jet and $\Delta\eta_{j_1 j_2}$ is the rapidity gap between the two tagging jets. This variable looks at the position of the third jet relative to the tagging jets and is important when considering the additional information available beyond the two-jet system. We set the normalization such that the cumulative sum of the bins corresponds to the fraction of events that satisfy the requirement on the third jet. The dipole-NLO signal has the least proportion of events passing the additional criteria at 30%, while the global NLO has 35%. The fraction for LO events with dipole and global-recoil schemes are 37% and 55%, respectively. From these values and the shape of the distribution in Fig. 2, we can infer that out of the four global LO should be most similar to the QCD-dominated background and dipole NLO should be the least identical. Consequently, we expect these to be reflected on the performance of any statistical model utilizing radiative information beyond the two jets. Although the proportion of events with a third jet is very close for global NLO and dipole LO, note that the former has more jets in the central regions from the shape of $|z^*_{j_3}|$ distribution. Hence, we would expect better discrimination for dipole LO.

Even though $z^*_{j_3}$ is a good variable, a model like a CNN that uses the inclusive event information will use the third jet's position as well as its transverse momentum implicitly to find the decision boundary. To this end, in Fig. 3, we plot the two-dimensional (2D) histogram plot of the transverse momentum $P_T^{j_3}$ of the third jet and $|z^*_{j_3}|$. Because of the artificial enhancement from the `II` like the global showering scheme in the central regions, we can see that the third jet in global cases are relatively harder than their dipole counterparts for both orders. Moreover, since the third jet results from the parton shower for LO, there is a drastic difference between global LO and dipole LO relative to the same comparison at NLO. From this, we can infer that events that do not have a third reconstructed jet would still follow the same pattern and expect the same effect on the performance of the CNN.

IV. RESULTS

In this section, we examine the performance of CNNs in identifying the different simulations of the same signal from the same background dataset described in Sec. II A. When trained with the same architecture, the relative

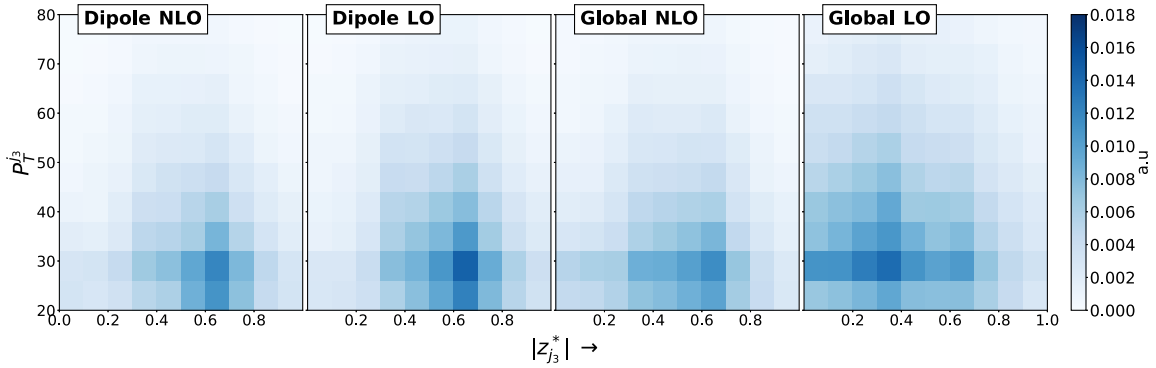
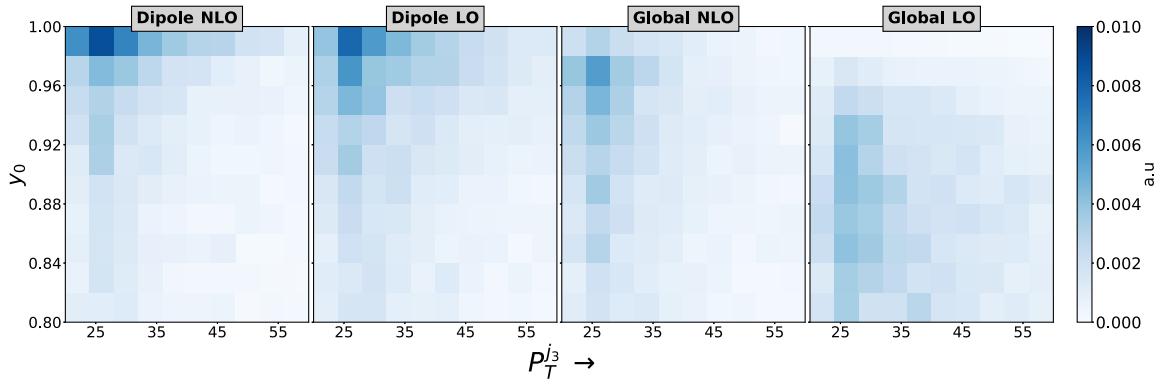
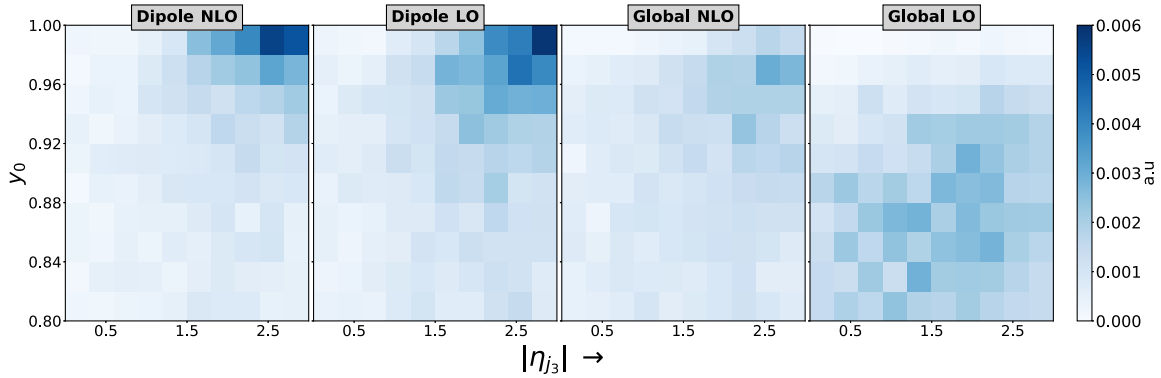


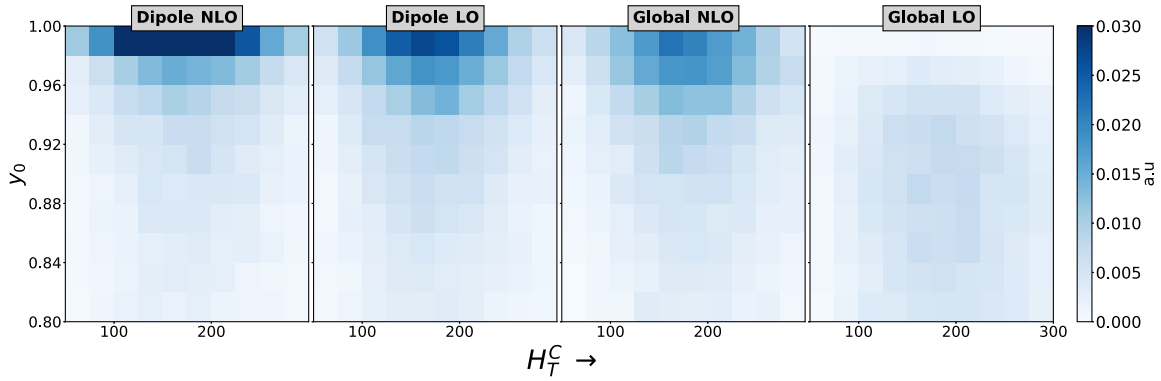
FIG. 3. Two-dimensional histogram of events with the transverse momentum ($P_T^{j_3}$) of the third jet and $|z^*_{j_3}|$ for four different cases of signal simulations, such as dipole NLO, dipole LO, global NLO, and global LO of the VBF Higgs signal.



(a)



(b)



(c)

FIG. 4. Two-dimensional histogram of events of the network output y_0 for each signal simulation with the (a) $P_T^{j_3}$ and (b) $|\eta_{j_3}|$ of the third jet (when present) and (c) the H_T^C between the two tagging jets.

discrimination power should reflect the physical intuition we presented in the preceding section. The four sets of signal events are preprocessed in the same manner as described in Sec. II B for the lower resolution, and the network $\mathcal{P}_J^{LR}$-CNN is trained with the procedure described in Sec. II C. The performance on the higher resolution should follow the same trend, and is hence unwarranted for the aim of the present work.

A. Effects of central radiation on the network output

The two-dimensional histogram of the network output y_0 (the probability of an event being a signal) of the signal validation datasets with various variables quantifying the additional information beyond the two-jet system are shown in Fig. 4. The weight of each event is set such that the total sum of all events with or without the third jet corresponds to 1. Therefore, the total sum of the histogram

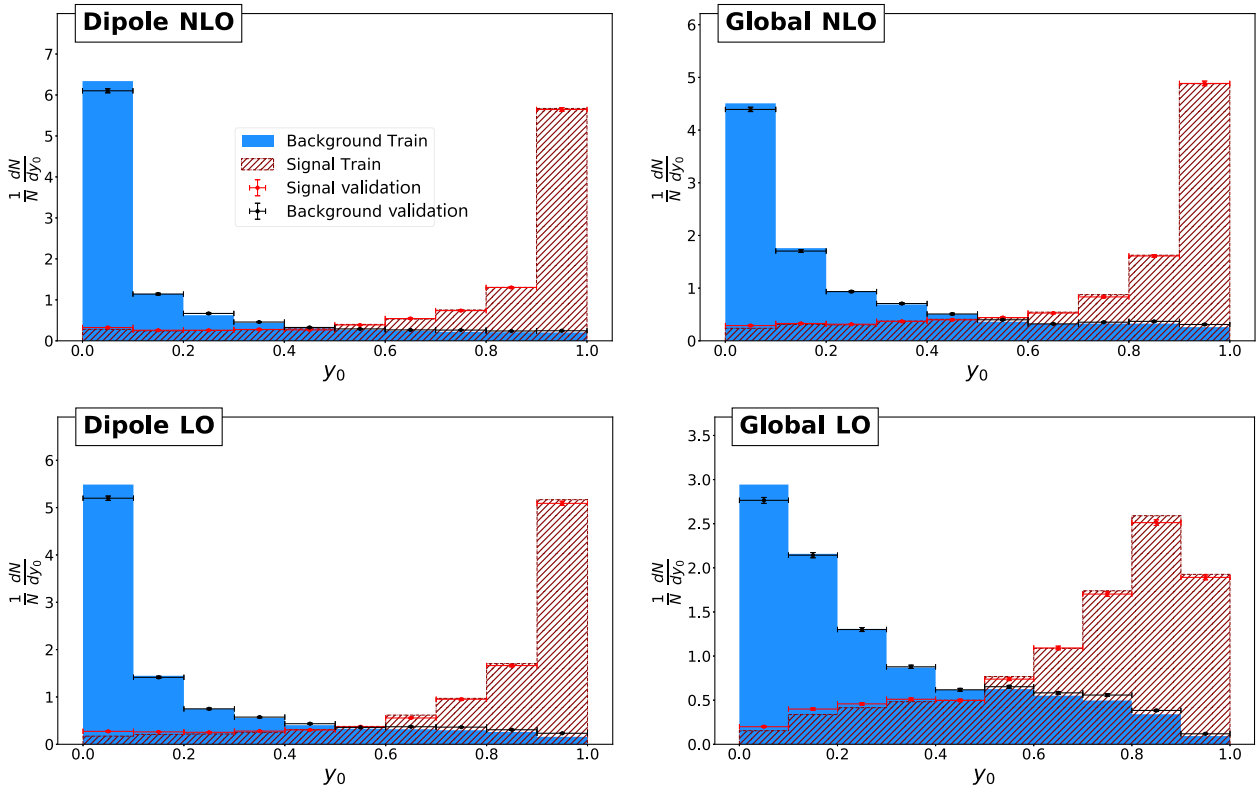


FIG. 5. Normalized binned distributions of the network output discriminating background from signal class for four different instances of signal simulations.

with the physical quantities of the third jet corresponds to the fraction with at least one additional jet.[3] The comparatively lower concentration of events for the global LO simulation is due to the lower performance of the network (presented in Sec. IV B) compared to the other three simulations.

In Fig. 4(a), in which the histogram is with the transverse momentum of the third jet, we see that for the dipole recoil both orders have the maximum concentration of events in the top-left corner. The third jet has the least transverse momentum in this region, and the network identifies the event as most signal-like. For the case of the global recoil, we see that the NLO simulation has a higher concentration near the top-left corner. In contrast, the LO simulation has significantly reduced events near the top left, with the shift toward the bottom in the y axis more prominent. The greater change in the network output can be understood by recalling from Fig. 3 that the relative position of the third jet is much more central for the global LO simulation event if its transverse momentum is in a similar range. This property is further confirmed in Fig. 4(b) where the histogram is on the $|\eta_{j_3}|$ and y_0 plane. The events for the global LO simulation are closer to the left side implying that the third jets are relatively central. At the same time, they populate the histogram on the lower values of the y_0 axis indicating that the network identifies the signal less efficiently. Similarly, the same histogram for the dipole-recoil scheme and different orders shows a concentration of events in the top right corner, where the third jets are more forward, and the network identifies the signal with greater confidence.

To look collectively into the events with or without a third jet, we define the scalar sum of p_T between the two tagging jets as

$$H_T^C = \sum_{\eta_i \in [\eta_l, \eta_u]} p_T^i, \tag{3}$$

where the range is determined by the pseudorapidity of the two jets: η_{j_1} and η_{j_2} mapped such that $\eta_l < \eta_u$. We do not remove the particles within the jets when calculating H_T^C, thus giving a nonzero value for all events. As expected, we see in Fig. 4(c) that the dipole-NLO simulation has the highest proportion of events near the top-left corner, followed by dipole LO and global NLO, with global LO having a larger concentration in the central regions of the (H_T^C, y_0) plane. Therefore, we see that events without the third jet also follow a pattern similar to those with the additional jet.

[3]Because of the range of the variables, the total sum is not equal to the fraction presented in Sec. III B.

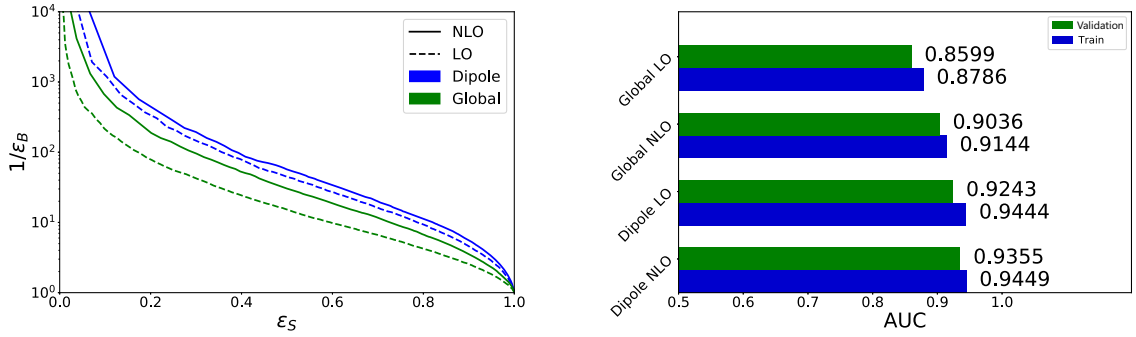


FIG. 6. Comparison of the performances in terms of ROC curves (left plot) on the validation dataset between the signal acceptance ϵ_S and the background rejection $1/\epsilon_B$ and the corresponding AUCs (right plot) for the training and validation data are shown for the four different cases of signal simulations.

B. Dependence of deep learning performance on the signal simulation

The normalized distribution of the network output y_0 of each class are shown in Fig. 5 for the four different signal simulation approaches. One can see that the CNN trained and validated with the dipole-NLO simulation has the highest separation from the background. To better quantify the power, we look at the receiver operating characteristic (ROC) curves between the signal acceptance ϵ_S and the background rejection $1/\epsilon_B$ and the area under the ROC curve (AUC). These are shown in Fig. 6. As expected, the highest discrimination is obtained for dipole NLO with a validation AUC of 0.9355, followed by dipole LO with 0.9243 validation AUC. Inadvertently, the dipole-NLO signal happens to be the most physically accurate simulation. The hierarchy suggests that the recoil scheme is of greater importance than the perturbative accuracy for the CNN analysis with tower images. Looking at the global recoil for each order, we see that global NLO has better performance, with the CNN trained and validated with global LO having the least discriminatory power. To understand this relative power, we note that the third jet in a NLO simulation has a leading-order accuracy. Whereas, for the LO case, the third jet, if present, is a consequence of the parton shower. The global-recoil scheme enhances the radiation in the central regions for both orders; however, it is partially controlled by the NLO simulation of the first real emission, while there is no such control for the LO case.

Although we have seen that the network trained and tested with different signal simulations shows notable differences, it is worth investigating how a CNN trained on a specific simulation fares when tested on other signal simulations. The validation AUC for all signal simulations evaluated on each of the networks trained on the different signal simulations is shown in Table I. For each signal type, the network it was trained on has the maximum discrimination, which is unsurprising given that the purpose of the training is to encode its behavior into the network. Moreover, the trend of increasing performance is the same regardless of the signal dataset used in training, pointing toward all networks learning the underlying difference between the signal and the background. Another feature of interest is the relatively higher range of AUCs for the LO datasets than NLO ones, pointing toward their relatively high uncertainties. Interestingly, regardless of the nature of the simulation used during the training, the most accurate simulation among the four, dipole-NLO events, has a very stable validation AUC with only a 1.6% deviation. This stability shows that CNNs can learn the underlying differences between VBF events and non-VBF events even when the VBF simulation is suboptimal.

To gauge the possible improvement in using a dipole scheme over the global scheme used in our previous work, we train the CNN with the combined gluon-fusion signal and the instance of dipole-NLO simulation of the VBF process in the same proportion as described in Sec. II and extract the bounds on the branching ratio. We find the median upper limit on the invisible branching ratio for an integrated luminosity $L = 300$ fb^{-1} to be 2.22%.

In all preceding analyses, we have used LO samples without any matching, and the third jet originates exclusively from the parton shower, which is inaccurate in describing harder emissions. It is worth examining how a matching procedure between the hard matrix element and the parton shower, which improves the description of the third jet in the harder regions, influences the network

TABLE I. The table shows the test AUC evaluated for all signal simulation for each CNN trained on the different signal simulations.

		Test signal dataset			
Serial number	Train signal dataset	Global LO	Global NLO	Dipole LO	Dipole NLO
1	Global-LO	0.8599	0.8956	0.9027	0.9201
2	Global-NLO	0.8486	0.9036	0.9112	0.9288
3	Dipole-LO	0.8036	0.8878	0.9243	0.9335
4	Dipole-NLO	0.8234	0.8922	0.9200	0.9355

performance. To inspect the possible improvement of such matching procedures, we generate VBF events matched with an additional jet via the MLM procedure [41,42] for both parton-recoil schemes. We found a continuous differential jet rate and transverse momentum distribution of the different jet samples for an XQCUT value of 100. As recommended for VBF processes, the AUTO_PTJ_MJJ flag was set to false. All other aspects of the simulation including the renormalization and factorization scale, PDFs, and baseline selection criteria are the same as described in Sec. III A. We generated about 25,000 events after baseline selection for both recoil schemes. Testing with these samples for the networks trained with the leading-order unmatched samples with the same parton-shower recoil against the validation background dataset, we find an AUC of 0.8651 and 0.9261 for the global and dipole matched LO samples, respectively. Compared to the full NLO simulation values tested on these networks (Table I), these values lie closer to the LO simulation, indicating that the matching procedure does not help alleviate the issues of the global parton shower. In contrast, the matched dipole value is still relatively stable, although closer to the LO value than the NLO value, signifying the relative importance of the virtual corrections of the NLO simulation.

V. CONCLUSION

The Large Hadron Collider, in its previous two runs, has already accumulated enough data to establish the Standard Model on strong footing, pinpointing different properties of the Higgs boson and also setting strong constraints in diverse beyond standard model (BSM) scenarios. The vector boson fusion mechanism of production is unique in many of these Higgs measurements and BSM searches. The strongest upper limit on the Higgs's invisible branching ratio comes from this channel, although a significantly large window remains open above the SM prediction. An improved constraint can enormously squeeze the parameter space on many new physics scenarios, such as Higgs/scalar portal dark matter.

In recent times, advancements in different machine-learning tools have opened the possibility to revisit many of these analyses with sophisticated data-driven methods. Smartly designed neural networks demonstrate a capability to comprehend event kinematics and the radiation pattern to a great extent, even when the event topology is rather simple. The invisible Higgs search in VBF is a process where the phenomenological study relies on two forward jets and a high missing transverse momentum. Moreover, many studies in this direction established some of the fundamental features of the event shape, which has almost been a norm to control the extensive QCD and electroweak backgrounds. Our previous work took the invisible Higgs search in VBF as a case study. We constructed CNN-based deep learning algorithms using just the low-level calorimeter inputs from the entire event topology without exclusive reconstructed objects. This novel approach can indeed provide the most stringent bounds on the invisible branching ratio of the SM-like Higgs boson, significantly outperforming the existing experimental search.

It is evident that deep learning algorithms with multiple nonlinear hidden layers can efficiently characterize complex distribution functions describing the feature space with greater accuracy. This expressivity enhances their capability to distinguish the signal region from the background, exploiting maximal information, even if the event topology is relatively simple. However, to exploit the full potential of this extraordinary capability, a precise theoretical projection of the QCD parton shower and the corresponding radiation pattern is required. The present work demonstrates this interrelation utilizing different showering schemes with leading-order and higher-order computation.

In this work, we carried out a quantitative analysis to investigate the dependence of a CNN's performance on the recoil scheme of the parton shower and the perturbative accuracy of the matrix element simulation for a VBF Higgs signal decaying to invisible particles. The difference between the leading order and next-to-leading order, although present, is not very pronounced for the physically correct dipole-recoil scheme. We found that the training is highly dependent on the recoil scheme, with a better performance coming for the physically accurate dipole recoil. With this fortunate coincidence, a complete analysis with all VBF processes showered with a dipole-recoil scheme will possibly reduce the upper limits on the invisible branching ratio even further than the projection which used a global-recoil scheme.

ACKNOWLEDGMENTS

The work is supported by the Physical Research Laboratory, Department of Space, Government of India. Computations were performed using the HPC resources (Vikram-100 HPC) and TDP project at Physical Research Laboratory. The authors gratefully thank Satyajit Seth and Aruna K. Nayak for stimulating discussion and suggestions throughout this project.

APPENDIX: BASELINE SELECTION CRITERIA

We use the same selection criteria used in the deep learning analysis of Ref. [37]. This cuts except for the one on missing transverse energy were based on the experimental shape analysis of Ref. [46]. Here, we summarize them for completeness:

(i) *Jet p_T.*—We select events with at least two jets, with the hardest (and second hardest) jet having at least 80 (40) GeV transverse momentum.

(ii) *Lepton veto.*—Events should not have any reconstructed electron (muon) with minimum transverse

momentum $p_T > 10$ GeV within the tracker region, i.e., $|\eta| < 2.5(2.4)$.

(iii) *Photon veto.*—Reject events having any photon with $p_T > 15$ GeV within $|\eta| < 2.5$.

(iv) *τ and b veto.*—If an event has any tau-tagged jets in $|\eta| < 2.3$ with $p_T > 18$ GeV or b-tagged jets in $|\eta| < 2.5$ with $p_T > 20$ GeV, they are discarded.

(v) *Minimum missing transverse energy $\not{E}_T$.*—An event must have a minimum transverse energy, $\not{E}_T > 200$ GeV to be selected.

(vi) *Alignment of missing transverse energy/momentum (MET) with respect to jet directions.*—The jets should have an azimuthal separation greater than 0.5 from $\not{E}_T$, i.e., $\min(\Delta\phi(\vec{p}_T^{\not{E}_T}, \vec{p}_T^{\,j})) > 0.5$ for all jets up to the fourth leading jet satisfying $p_T > 30$ GeV and $|\eta| < 4.7$, This requirement rejects QCD multijet backgrounds arising due to severe mismeasurement.

(vii) *Jet rapidity.*—The tagging jets should be well within the calorimeter acceptance ($|\eta_j| < 4.7$), and at least one of them should be within the central regions ($|\eta_{j_i}| < 3$).

(viii) *Jets in opposite hemisphere.*—The tagging jets must reside in opposite hemisphere in η. This is achieved by imposing the condition $\eta_{j_1} \times \eta_{j_2} < 0$.

(ix) *Azimuthal angle separation between jets.*—We require the azimuthal separation between the two tagging jets to satisfy $|\Delta\phi_{j_1 j_2}| < 1.5$.

(x) *Jet rapidity gap.*—The rapidity gap between two leading jets must satisfy $|\Delta\eta_{j_1 j_2}| > 1$.

(xi) *Dijet invariant mass.*—The invariant mass of the two-jet system should satisfy $m_{jj} > 200$ GeV.

After weighting the different background processes by their cross sections and the baseline selection efficiency, we get almost 95% contribution from non-VBF-type production of Z/W bosons, while remaining is of the VBF origin.

[1] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).
[2] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).
[3] CMS Collaboration, Report No. CMS-PAS-HIG-20-003, 2021.
[4] A. M. Sirunyan *et al.* (CMS Collaboration), Phys. Rev. D **104**, 032013 (2021).
[5] CMS Collaboration, Report No. CMS-PAS-HIG-20-007, 2021.
[6] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **125**, 061802 (2020).
[7] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 10 (2021) 013.
[8] ATLAS Collaboration, Report No. ATLAS-CONF-2020-052, 2020.
[9] R. N. Cahn and S. Dawson, Phys. Lett. **136B**, 196 (1984); **138B**, 464(E) (1984).
[10] Yu. L. Dokshitzer, V. Khoze, and S. Troyan, *New perturbative results in hadron jet physics* (World Scientific Pub Co, 1987), p. 365.
[11] J. D. Bjorken, Phys. Rev. D **47**, 101 (1993).
[12] R. S. Fletcher and T. Stelzer, Phys. Rev. D **48**, 5162 (1993).
[13] D. L. Rainwater and D. Zeppenfeld, J. High Energy Phys. 12 (1997) 005.
[14] O. J. Eboli and D. Zeppenfeld, Phys. Lett. B **495**, 147 (2000).
[15] P. Konar and B. Mukhopadhyaya, Phys. Rev. D **70**, 115011 (2004).
[16] P. Konar and D. Zeppenfeld, Phys. Lett. B **647**, 460 (2007).
[17] D. Choudhury, A. Datta, K. Huitu, P. Konar, S. Moretti, and B. Mukhopadhyaya, Phys. Rev. D **68**, 075007 (2003).
[18] A. Datta, P. Konar, and B. Mukhopadhyaya, Phys. Rev. D **65**, 055008 (2002).
[19] A. Datta, P. Konar, and B. Mukhopadhyaya, Phys. Rev. Lett. **88**, 181802 (2002).
[20] A. Datta, P. Konar, and B. Mukhopadhyaya, Phys. Rev. D **63**, 095009 (2001).
[21] H. Han, J. M. Yang, Y. Zhang, and S. Zheng, Phys. Lett. B **756**, 109 (2016).
[22] J. Heisig, M. Krämer, E. Madge, and A. Mück, J. High Energy Phys. 03 (2020) 183.
[23] G. Arcadi, A. Djouadi, and M. Raidal, Phys. Rep. **842**, 1 (2020).
[24] G. Arcadi, A. Djouadi, and M. Kado, Eur. Phys. J. C **81**, 653 (2021).
[25] B. Bhattacherjee, S. Matsumoto, and R. Sengupta, arXiv: 2111.02437.
[26] S. Argyropoulos, O. Brandt, and U. Haisch, Symmetry **2021**, 13 (2021).
[27] T. Han, G. Valencia, and S. Willenbrock, Phys. Rev. Lett. **69**, 3274 (1992).
[28] T. Figy, C. Oleari, and D. Zeppenfeld, Phys. Rev. D **68**, 073005 (2003).
[29] F. A. Dreyer and A. Karlberg, Phys. Rev. Lett. **117**, 072001 (2016).
[30] T. Liu, K. Melnikov, and A. A. Penin, Phys. Rev. Lett. **123**, 122002 (2019).
[31] A. Radovic, M. Williams, D. Rousseau, M. Kagan, D. Bonacorsi, A. Himmel, A. Aurisano, K. Terao, and T. Wongjirad, Nature (London) **560**, 41 (2018).
[32] D. Guest, K. Cranmer, and D. Whiteson, Annu. Rev. Nucl. Part. Sci. **68**, 161 (2018).
[33] D. Bourilkov, Int. J. Mod. Phys. A **34**, 1930019 (2019).

[34] J. H. Kim, M. Kim, K. Kong, K. T. Matchev, and M. Park, J. High Energy Phys. 09 (2019) 047.
[35] J. Amacker *et al.*, J. High Energy Phys. 12 (2020) 115.
[36] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, J. High Energy Phys. 07 (2016) 069.
[37] V. S. Ngairangbam, A. Bhardwaj, P. Konar, and A. K. Nayak, Eur. Phys. J. C **80**, 1055 (2020).
[38] B. Cabouat and T. Sjöstrand, Eur. Phys. J. C **78**, 226 (2018).
[39] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, J. High Energy Phys. 07 (2014) 079.
[40] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, Comput. Phys. Commun. **191**, 159 (2015).
[41] M. Mangano, Fermilab ME/MC Tuning Workshop, October 4, 2002, http://www-cpd.fnal.gov/personal/mrenna/tuning/nov2002/mlm.pdf.gz (2002).
[42] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, J. High Energy Phys. 01 (2007) 013.
[43] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3 Collaboration), J. High Energy Phys. 02 (2014) 057.
[44] M. Cacciari, G. P. Salam, and G. Soyez, J. High Energy Phys. 04 (2008) 063.
[45] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. C **72**, 1896 (2012).
[46] A. M. Sirunyan *et al.* (CMS Collaboration), Phys. Lett. B **793**, 520 (2019).
[47] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **80**, 1165 (2020).
[48] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **81**, 332 (2021).
[49] T. Dozat, in *ICLR 2016 Workshop* (2016), https://openreview.net/forum?id=OM0jvwB8jIp57ZJjtNEZ.
[50] F. Chollet *et al.*, Keras, https://keras.io (2015).
[51] M. Abadi *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems (2015), software available from tensorflow.org.
[52] T. Junk, Nucl. Instrum. Methods Phys. Res., Sect. A **434**, 435 (1999).
[53] A. L. Read, J. Phys. G **28**, 2693 (2002).
[54] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. C **71**, 1554 (2011); **73**, 2501(E) (2013).
[55] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, *Proc. Sci.*, ACAT2010 (2010) 057 [arXiv:1009.1003].
[56] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke (ROOT Collaboration), HistFactory: A tool for creating statistical models for use with RooFit and RooStats, Report No. CERN-OPEN-2012-016 (2012).
[57] A. Ballestrero *et al.*, Eur. Phys. J. C **78**, 671 (2018).
[58] B. Jäger, A. Karlberg, S. Plätzer, J. Scheller, and M. Zaro, Eur. Phys. J. C **80**, 756 (2020).
[59] A. Tumasyan *et al.* (CMS Collaboration), arXiv:2112.05259.
[60] CMS Collaboration, Search for Higgs boson pair production via vector boson fusion with highly Lorentz-boosted Higgs bosons in the four b quark final state at $\sqrt{s} = 13$ TeV, Report No. CMS-PAS-B2G-21-001 (2021).
[61] A. M. Sirunyan *et al.* (CMS Collaboration), J. High Energy Phys. 03 (2021) 257.
[62] G. Gustafson, Phys. Lett. B **175**, 453 (1986).
[63] G. Gustafson and U. Pettersson, Nucl. Phys. **B306**, 746 (1988).
[64] S. Schumann and F. Krauss, J. High Energy Phys. 03 (2008) 038.
[65] S. Platzer and S. Gieseke, J. High Energy Phys. 01 (2011) 024.
[66] P. Nason, J. High Energy Phys. 11 (2004) 040.
[67] S. Frixione, P. Nason, and C. Oleari, J. High Energy Phys. 11 (2007) 070.
[68] P. Nason and C. Oleari, J. High Energy Phys. 02 (2010) 037.
[69] S. Alioli, P. Nason, C. Oleari, and E. Re, J. High Energy Phys. 06 (2010) 043.
[70] J. Butterworth *et al.*, J. Phys. G **43**, 023001 (2016).
[71] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, Eur. Phys. J. C **75**, 132 (2015).
[72] S. Forte, Acta Phys. Pol. B **41**, 2859 (2010), https://inspirehep.net/literature/879362.
[73] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, Phys. Rev. D **93**, 033006 (2016).
[74] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, Eur. Phys. J. C **75**, 204 (2015).
[75] R. D. Ball *et al.* (NNPDF Collaboration), J. High Energy Phys. 04 (2015) 040.
[76] S. Carrazza, S. Forte, Z. Kassabov, J. I. Latorre, and J. Rojo, Eur. Phys. J. C **75**, 369 (2015).
[77] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, J. High Energy Phys. 03 (2013) 015.
[78] https://pythia.org/latest-manual/POWHEGMerging.html.