

Learning Physics at Future e^-e^+ Colliders with Machine

Lingfeng Li,^a Ying-Ying Li,^{b,c} Tao Liu,^c Si-Jun Xu^c

^a *Jockey Club Institute for Advanced Study, The Hong Kong University of Science and Technology, Hong Kong S.A.R., P.R.China*

^b *Theoretical Physics Department, Fermi National Accelerator Laboratory, Batavia, IL 60510, U.S.A.*

^c *Department of Physics, The Hong Kong University of Science and Technology, Hong Kong S.A.R., P.R.China*

E-mail: iaslfli@ust.hk, yingying@fnal.gov, taoliu@ust.hk,
sxuaw@connect.ust.hk

ABSTRACT: Information deformation and loss in jet clustering are one of the major limitations for precisely measuring hadronic events at future e^-e^+ colliders. Because of their dominance in data, the measurements of such events are crucial for advancing the precision frontier of Higgs and electroweak physics in the next decades. We show that this difficulty can be well-addressed by synergizing the event-level information into the data analysis, with the techniques of deep neural network. In relation to this, we introduce a CMB-like observable scheme, where the event-level kinematics is encoded as Fox-Wolfram (FW) moments at leading order and multi-spectra at higher orders. Then we develop a series of jet-level (w/ and w/o the FW moments) and event-level classifiers, and analyze their sensitivity performance comparatively with two-jet and four-jet events. As an application, we analyze measuring Higgs decay width at e^-e^+ colliders with the data of 5ab^{-1} @240GeV. The precision obtained is significantly better than the baseline ones presented in documents. We expect this strategy to be applied to many other hadronic-event measurements at future e^-e^+ colliders, and to open a new angle for evaluating their physics capability.

Contents

1	Introduction	1
2	General Strategies	6
2.1	CMB-like Observable Scheme	6
2.2	Machine Learning with Event-level Kinematics	8
3	Benchmark Study	10
3.1	Two-Jet Events	10
3.2	Four-Jet Events	17
4	Application: Higgs Decay Width (Γ_h)	21
4.1	Γ_h Measurement at e^-e^+ Colliders	21
4.2	Measuring $\sigma(Zh)\text{BR}(h \rightarrow WW^*)$	24
4.3	Measuring $\sigma(\nu\nu h)\text{BR}(h \rightarrow \text{hadrons})$	26
4.4	Robustness against Detector Resolution	29
5	Summary and Discussion	32
A	Event Response to Γ_h Classifiers	34

1 Introduction

The precision frontier of next decades in Higgs and electroweak (EW) physics is expected to be defined by next-generation e^-e^+ colliders. The proposed projects so far mainly include [1] circular machines such as CEPC and FCC-ee, and linear machines such as ILC and CLIC. As a Higgs factory, the CEPC and FCC-ee will operate at $\sqrt{s} = 240\text{GeV}$ with an integrated luminosity $\sim 5\text{ab}^{-1}$, in addition to their low- \sqrt{s} runs. About 10^6 clean Higgs events will be produced during this period together with $\sim 6 \times 10^6 ZZ$, $\sim 8 \times 10^7 WW$ and $\sim 3 \times 10^8 qq$ events. After this phase, the FCC-ee operation is expected to be upgraded to a mode beyond Higgs factory, with a collection of $\sim 1.7\text{ab}^{-1}$ data at the $t\bar{t}$ threshold. The ILC project aims for $\sim 2\text{ab}^{-1}$ @ 250 GeV run and $\sim 4\text{ab}^{-1}$ @ 500 GeV run. As for the CLIC, the planned 380 GeV run will collect $\sim 1.5\text{ab}^{-1}$ data. Then the operation will be upgraded to 1.5 and 3.0 TeV, with a collection of $\sim 2.5\text{ab}^{-1}$ and $\sim 5\text{ab}^{-1}$ data, respectively.

So far, a lot of efforts have been made to explore the prospect of measuring Higgs and EW physics at these e^-e^+ colliders [2–16]. The primary Higgs and electroweak processes for the collider low- \sqrt{s} runs and their branching ratios are summarized in Table 1. Clearly the hadronic modes containing (anti-)quarks or/and gluons are dominant and even overwhelmingly dominant over the purely leptonic ones. Because of this, the baseline sensitivities in

$N_{q/g}$	0	2	4	6
$e^-e^+ \rightarrow WW$	11%	44%	45%	0%
$e^-e^+ \rightarrow ZZ$	9%	42%	48%	0%
$e^-e^+ \rightarrow Zh$	3%	32%	55%	11%
$e^-e^+ \rightarrow \nu\nu h$	20%	69%	11%	0%
$e^-e^+ \rightarrow t\bar{t}$	0%	11%	44%	45%

Table 1. Branching ratios of hadronic modes, for the primary Higgs and electroweak processes in low- \sqrt{s} runs of future e^-e^+ colliders. Here $N_{q/g} = 0, 2, 4$ and 6 represents the number of (anti-)quarks and gluons in their final states. τ decays are defined to be leptonic. W/Z decays are assumed to be two-body only.

documents for many benchmark precision measurements are based on such hadronic modes, with jet-level analysis being generally applied. One prominent example is the Higgs decay width (Γ_h) in the Standard Model (SM). At low- \sqrt{s} e^-e^+ colliders, its baseline precision is mainly determined by the measurement of $\sigma(\nu\nu h_b)$ ¹ (the Vector-Boson-Fusion (VBF) Higgs production rate with the Higgs decaying into bb), with the mainstream method (*i.e.*, Method B defined in Subsec. 4.1). At 240GeV, the CEPC and FCC-ee are expected to measure this quantity with a precision of 3.2% and 3.1% (see Table 3), respectively. Combined with the measurements of other intermediate quantities which rely on hadronic data also, this yields a precision of 3.5% for both for the SM Γ_h measurement [17, 18].

Yet, the precision based on the jet-level analysis is limited for several reasons. First of all, due to the imperfection of jet clustering algorithms, some visible particles could be clustered into a wrong jet. This becomes especially significant if the jet ancestral partons are collimated, where their hadronizations might badly overlap with each other in space. This effect will deform the jet kinematics from its truth, and may negatively impact the reconstruction of the intermediate particles or events with jets². In the performance study of the CEPC detector [19], this problem is termed as ‘‘jet confusion’’. Secondly, the jet clustering in essence is an operation of dimensionality reduction in the feature space of the visible particles. This operation aims reconstructing four momentum of the jet ancestral partons. But, it removes the dimensions reflecting jet substructure and superstructure, generically resulting in a loss of kinematic information. The jet substructure manifests color, electric charge, flavor, etc. of the jet ancestral partons and hence is useful in discriminating, e.g., quark/gluon jets [20, 21, 21–25]. The jet superstructure is usually formulated if the jet ancestral partons share the same parent particle, where they tend to be showered in a

¹In this paper we will take a shorthand notation, using the subscript to denote particle decay mode. For example, we will use $h_{q,b,W,Z}$ to denote $h \rightarrow qq, bb, WW^*, ZZ^*$, $h_{W_{lq},qq}$ to denote $h \rightarrow WW^* \rightarrow l\nu qq, qqqq$ and h_h to denote $h \rightarrow qq, bb, gg, \tau\tau$. Additionally, we will use ‘‘ q ’’ to denote the quarks of the first two generations and in a general sense, and ‘‘ q_3 ’’ and ‘‘ q_5 ’’ to denote $\{u, d, s\}$ and $\{u, d, s, c, b\}$, respectively.

²The reconstruction of four momentum of visible particles and their derived quantities such as visible and recoil masses is not influenced by the said information deformation, if no visible particle is missed in jet clustering. But, such high-level observables are often insufficient and even irrelevant for the reconstruction of the whole event. One typical example is the $W_q W_q$ and $Z_q Z_q$ measurement [19], which will be discussed in Subsec 3.2.

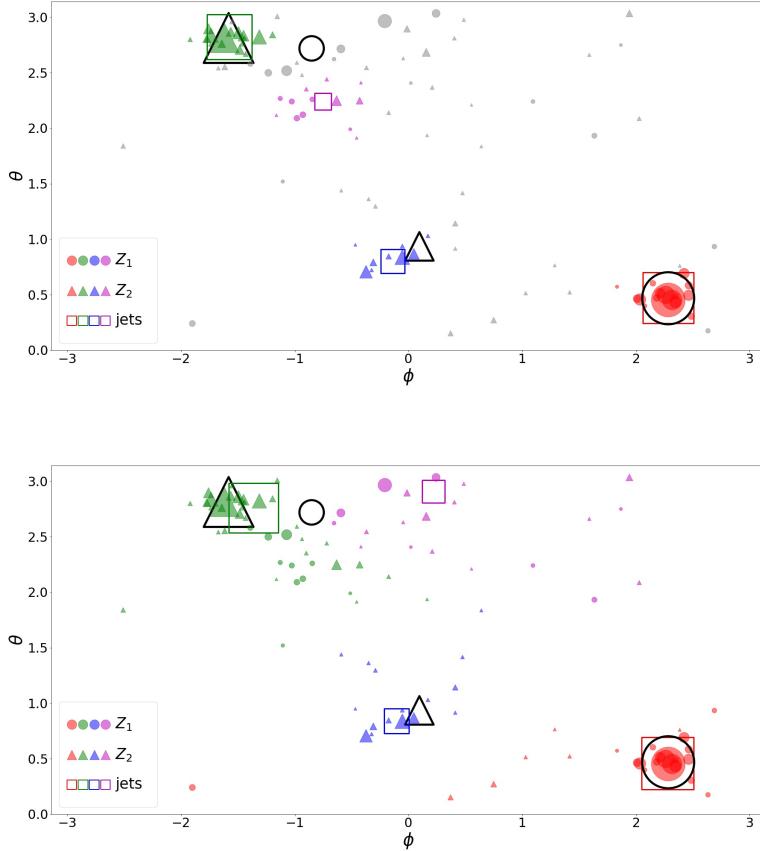


Figure 1. An illustration of information deformation and loss in jet clustering. One representative $e^-e^+ \rightarrow Z_qZ_q$ event is projected to the $\phi - \theta$ plane, without any detector effects. For jet clustering, the anti- k_t algorithm with $\Delta R = 0.5$ is applied in the upper panel and the $ee - k_t$ algorithm in the bottom panel. In both panels, each solid symbol represents a particle visible to the detector, with its size scaling with energy, its color denoting the relevant jet, and its shape (circle and triangle) labeling its parent Z boson. The grey symbols in the upper panel represent the particles not clustered to any jets. We use unfilled symbols, *i.e.*, colored boxes and black circles and triangles, to represent the jets and their ancestral quarks, respectively.

correlated way [26]. This structure encodes quantum numbers of these parent particles such as color, spin and CP-property, and is valuable for their collider search. Thirdly, some particles could be either missed by the detector, due to its limited coverage in space, or not clustered to any jets, because of their large distance to the jets. These effects can contribute to the said information deformation and loss also.

To make these problems more explicit, in Fig. 1 we show jet clustering in an $e^-e^+ \rightarrow Z_qZ_q$ event, using two representative algorithms. The first one is anti- k_t algorithm [27] which has been extensively applied for data analysis at Large Hadron Collider (LHC). The second one is $ee - k_t$ algorithm [28] which has a relatively long history and was originally designed for clean environment at e^-e^+ colliders such as the Large Electron-Positron

Collider. Unlike the anti- k_t , which clusters hard particles first and hence is relatively insensitive to soft radiation and detector noise, the $ee - k_t$ gives a priority to soft particles. This difference allows it to use a priori knowledge on the jet ancestral partons to implement the jet clustering. Explicitly, with the $ee - k_t$ one can request all particles recorded by the detector in each event to be clustered into jets with a given number. For the $e^-e^+ \rightarrow Z_qZ_q$ event shown in Fig. 1, the anti- k_t clustering is implemented with a jet cone of $\Delta R = 0.5$ and the $ee - k_t$ clustering is required to generate four jets. The whole event is projected to the $\phi - \theta$ plane, with no detector simulation being applied. As is shown, the unfilled black circle and triangle on the top of the $\phi - \theta$ plane, which represent jet ancestral partons from different parent Z bosons, are relatively collimated. This results in an overlap between the distributions of solid circles and triangles nearby. Although this unfilled black triangle, together with the unfilled black circle and triangle at the bottom, overlap well with three of the jets at the plane which are denoted by colored boxes, the unfilled black circle on the top is left not close to the fourth jet (magenta box) in both panels. This clearly displays a deformation of kinematic information from the truth for the fourth jet. Indeed, this jet is composed of both solid circles and triangles. Also, as the visible particles are clustered into jets, the information on their correlation and distribution is lost. This information can be partly taken away also by the visible particles which are not clustered to any jets (denoted by grey symbols), if the anti- k_t algorithm is applied.

The limitations of the jet-level analyses naturally raise the question whether the baseline precisions presented in documents fully reflect the physics potential of future e^-e^+ colliders. After all, a significant improvement to many of these baseline precisions would be expected, if the information deformation and loss in jet clustering can be well-addressed. There are two potential methods to solve or partly solve these two problems. The first one is to pursue jet-level analysis by properly incorporating subjet-scale or event-level observables. This method does not solve the problem of information deformation for the jets directly. But it may mitigate its negative impact by incorporating the event-level message. Also, the information lost at jet level can be partly incorporated in this method. The tool of jet-substructure represents such a success which was originally introduced to test QCD [29]. This tool has been extensively applied to searching for boosted heavy objects at LHC. Additionally, a series of event-level observables have been introduced for data analysis at colliders since decades before. These observables fall into two classes roughly. The first class manifest event shape. One prominent example is thrust

$$T \equiv \max_{\mathbf{n}_T} \frac{\sum_{i \in \text{event}} |\mathbf{p}_i \cdot \mathbf{n}_T|}{\sum_{i \in \text{event}} |\mathbf{p}_i|} \quad (1.1)$$

introduced in 1970s [30]. Here \mathbf{n}_T is a unit spacelike vector and defines the thrust axis, and i runs over all visible particles in the event. The thrust was subsequently generalized to many other possibilities [31–35] (for a review, see, e.g., [36, 37]). It is interesting that most of these event-level observables were originally proposed for the e^-e^+ and e^-h events [36] rather than the hh ones [37]. Recently, an observable to measure event isotropy was also proposed [38]. Another class encode the event-level information at different angular scales. The most famous example is probably Fox-Wolfram (FW) moments [39]. The FW moments

were introduced in 1970s also, for analyzing the e^-e^+ collision events. They are defined as

$$H_{AB;l} = \sum_{m=-l}^l H_{AB;l,m} = \frac{4\pi}{2l+1} \sum_{i,j} \frac{A_i B_j}{s} \sum_{m=-l}^l (Y_l^m(\Omega_i)^* Y_l^m(\Omega_j)) = \sum_{i,j} \frac{A_i B_j}{s} P_l(\cos \Omega_{ij}) . \quad (1.2)$$

Here $Y_l^m(\Omega_i)$ is spherical harmonics of degree l and order m , $P_l(\cos \Omega_{ij})$ is Legendre polynomials,

$$\cos \Omega_{ij} = \cos \theta_i \cos \theta_j + \sin \theta_i \sin \theta_j \cos(\phi_i - \phi_j) \quad (1.3)$$

is the cosine of the included angle between the i^{th} and j^{th} visible particles, and A and B are infrared-safe kinematic variables such as p_T , E , etc. In this summation i and j run over all visible particles in the event ³. These two classes of observables are both physically intuitive, but less organized or incomplete in representing the event-level kinematics. Another one is to pursue the analysis in a brute-force way, using the event-level data as input. With this method, the problem of information deformation at jet level becomes irrelevant, while the kinematic information at event level could be exploited to the greatest extent for data analysis. Despite this, both methods are confronted with a challenge, *i.e.*, how to efficiently synergize the event-level information into the data analysis, given the complexity of its structure.

The machine learning (ML) techniques based on deep neural network (DNN) bring a great opportunity to address this challenge, due to their revolutionary capability to mine data. This tool became popularized in last two decades for hardware development and big data availability. This motivates us to pursue the study below. Our primary goal is to

- provide an angle to evaluate the physics capability of future e^-e^+ colliders, which is different from the ones taken in most relevant literatures and documents, by properly synergizing the event-level information into the DNN-based data analysis.

We will develop a set of DNN-based binary classifiers using both methods and apply them to a series of benchmark studies comparatively. Yet, implementing the first method in an organized manner requires the event-level kinematics to be encoded as a complete or approximately complete set of prioritized observables. So, we would also

- construct an observable scheme to systematically represent the event-level information in each event.

By incorporating the observables in such a scheme order by order, we would expect the performance of the jet-level classifiers to approach that of the event-level ones gradually. The information lost at jet level then could be reconstructed based on these observables. Beyond that, such an information-representing scheme is valuable for seeing into the event-level kinematics and revealing the underlying physics, a task generically difficult for the second method.

³To avoid being distracted from the QCD information, one can modify the definition of the FW moments by excluding the isolated leptons or photons from this summation, as we will do in the analysis of measuring $\sigma(Z_\nu h_{W_{lq}})$ in Subsec. 4.2.

The e^-e^+ colliders stand on a better position, compared to hadron colliders, in this regard. They are characterized with negligible pileups, colorless beam, and especially isotropy of event four momentum (*i.e.*, $\vec{p}_{\text{event}} \equiv 0$). This is reminiscent of all-sky Cosmic Microwave Background (CMB) map, and motivates us to introduce a CMB-like observable scheme. In this scheme, the event-level kinematics is encoded as the FW moments at leading order and multi-spectra at higher orders. We will be less ambitious in this paper, and test only to what extent the FW moments of energy can compensate for the information lost at jet level and reduce the performance gap between the jet-level and event-level classifiers.

This paper is organized as follows. In Sec. 2 we introduce the CMB-like observable scheme and the strategies for the DNN-based analyses. Then we develop a series of jet-level (w/ and w/o the FW moments) and event-level classifiers, and analyze their performance comparatively with two-jet and four-jet events in Sec. 3. Similar strategies are subsequently applied to the analysis of measuring Γ_h at 240GeV in Sec. 4. We summarize our results and take an outlook in Sec. 5.

2 General Strategies

2.1 CMB-like Observable Scheme

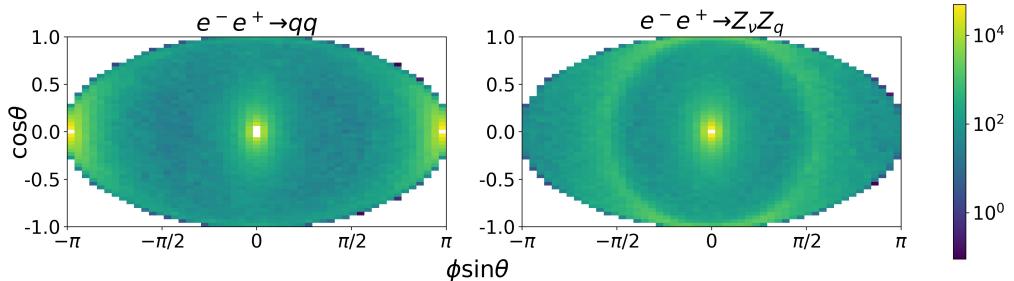


Figure 2. Cumulative Mollweide projections of 10000 events: qq (left) and $Z_\nu Z_q$ (right), with the brightness of each cell scaling with the total energy (GeV) of the particle hits received.

In this subsection, we will introduce a CMB-like observable scheme to encode the kinematic information in each event. Let us consider first the cumulative Mollweide projection of two classes of hadron-level events: $e^-e^+ \rightarrow qq$ and $e^-e^+ \rightarrow Z_\nu Z_q$. Here only the visible particles are relevant. We make the projection in the following way: (1) define a Cartesian coordinate system at the collider, with its z -axis being along the beam line; (2) rotate the momentum of the most energetic visible particle to be along x -axis which points from the paper to outside; and (3) scale the brightness of each cell at the projection sphere with the total energy of the particle hits received. The projections are shown in Fig. 2. Both of them demonstrate anisotropic features. In the $e^-e^+ \rightarrow qq$ projection, there are two bright points centered at $\{\phi \sin \theta = 0, \cos \theta = 0\}$ and $\{\phi \sin \theta = \pm \pi, \cos \theta = 0\}$, respectively. This reflects that before hadronization the two ancestral quarks move oppositely due to momentum conservation. Differently, a pupil-like structure is formed in the $e^-e^+ \rightarrow Z_\nu Z_q$

projection, with its circumference and radius being determined by the relative position of these two ancestral quarks and their included angle, respectively. By energy and momentum conservation, the position contour of the second ancestral quark at the $\phi \sin \theta - \cos \theta$ plane is given by

$$\frac{1}{2} - \frac{1}{2} \sqrt{1 - \cos^2 \theta} \cos \left(\frac{\phi \sin \theta}{\sqrt{1 - \cos^2 \theta}} \right) = \frac{m_Z^2}{E_Z^2}. \quad (2.1)$$

This equation predicts the said included angle at $\cos \theta = 0$, *i.e.*, its characteristic value, to be $\sim \phi \sin \theta \sim \pi/2$. With hadronization, the position points and contours for these ancestral quarks are smeared into halos, with their energy density varying spatially.

Mollweide projection at e^-e^+ colliders	All-sky CMB map
Projection sphere	Celestial sphere
Equatorial plane	Galactic plane
Energy (p_T , timing, charge, d_0 , etc.) projection	Temperature (polarization) map
Event-level kinematics	Anisotropy
Fox-Wolfram moments	Power spectrum (TT, TB, BB , etc.)
Multi-spectra	Bispectrum, trispectrum, etc.
...

Table 2. Dictionary between the Mollweide projection at e^-e^+ colliders and the all-sky CMB map.

These observations are reminiscent of the all-sky CMB map where the message on the early Universe is encoded as its power spectrum and multi-spectra. Quite generally, we can build up a dictionary between the Mollweide projection of each e^-e^+ collision event and the all-sky CMB map, as is summarized in Table 2. Here the projection sphere plays a role of the celestial sphere in the all-sky CMB map, with its equatorial plane mimicking the disc of Milky Way. The spherical projection of collider observables including energy and momentum, timing, tracker parameters such as charge, impact parameter d_0 , etc., can be mapped to the all-sky map of the CMB temperature and polarization. Naturally, the event-level kinematics is manifested as the anisotropy of the projection. The relevant information thus can be encoded as the FW moments at leading order and multi-spectra at higher orders, an analogue to the CMB power spectrum and its bispectrum, trispectrum, etc.⁴ Despite these correspondences, it is noteworthy that there exist some important differences between the Mollweide projection at e^-e^+ colliders and the all-sky CMB map. First, to high accuracy, the CMB temperature fluctuation is a Gaussian random field, with its non-Gaussian effects being expected to be tiny. In contrast, the Mollweide projection of each event at e^-e^+ colliders is physically non-Gaussian. Its multi-spectra thus may contain significant information beyond the FW moments on the event-level kinematics⁵.

⁴To ensure their infrared- and collinear-safety, for some observables in this CMB-like scheme such as the charge FW moments and multi-spectra, one needs to properly weight the contribution of each particle using, e.g., its energy or momentum (as was done for defining jet charge in [40]).

⁵Recall, a Gaussian random field is completely determined by its two-point correlator. All of its higher-order correlators can be expressed in term of the two-point correlator according to Wick's theorem.

Second, the CMB power spectrum is measured in the universe where we live and hence one realization of all the possible CMBs is recorded only. This limits its measurement precision at large angular scale, causing the notorious “cosmic variance” problem. Differently, the collider data sample is typically of large size. The variance of their mean over samples are suppressed for the FW moments at all multipoles, according to the central limit theorem.

2.2 Machine Learning with Event-level Kinematics

In this paper we will apply to our study the two methods to address the information deformation and loss in jet clustering, in a comparative way. Explicitly, we will develop five types of DNN-based binary classifiers in each analysis:

- J1 classifier: jet-level, without FW moments and track information;
- J2 classifier: jet-level, with $H_{EE;l \leq 50}$ (FW moments of energy with $l \leq 50$) and without track information;
- J3 classifier: jet-level, with $H_{EE;l \leq 50}$ and track information;
- E1 classifier: event-level, without track information;
- E2 classifier: event-level, with track information.

Among these, J1 will serve as a reference classifier. J2, J3 and E1, E2 classifiers are based on the first and second methods, respectively. We will test the effectiveness of E1 classifier by comparing its performance with J1’s. By expectation, E1 classifier should perform better than J1. J2 classifier will tell us to what extent $H_{EE;l \leq 50}$, as part of the leading-order observables in the CMB-like observable scheme, can compensate for the information lost at jet level⁶, and reduce the performance gap between J1 and E1 classifiers. The track observables of secondary vertex (SV) will be incorporated at last in J3 and E2 classifiers.

The event-level classifiers are somewhat related to the end-to-end ones proposed in [41, 42]. Yet, instead of using the raw detector response as input for improving particle reconstruction [41, 42], we are more dedicated to addressing the information deformation and loss in jet clustering. Hence we will use the reconstructed particles as input for the analyses. Several difficulties arise in this setup. If the feature space is defined with the momenta of the visible particles in each event, its dimension is not fixed, due to the fluctuation of the particle number. Also, the dimension of the feature space is generically high for the hadronic events and hard to sort. These complexities could be addressed in several ways with the ML techniques. The first one is to image the events and then apply the ML techniques of image recognition, such as Convolutional Neural Network (CNN) [20, 43, 44], for their classification. In this case, the pixel intensity in each image represents the total contribution of the visible particles hitting this pixel to some kinematic variables such as E . The dimension of the input parameters is thus determined not by the particle number, but by the pixel number. Similar techniques of image recognition have been applied to tagging

⁶The information carried by the jets and the FW moments could overlap to some extent. If all observables in this CMB-like scheme are incorporated, we would expect that the jet information become irrelevant.

light jets [23], boosted W boson [45] and top quark [46], selecting events [20, 42, 47–49], mitigating pileups at the LHC [50], etc. For its directviewing and effectiveness, we will take this method below. The second method is based on Recurrent Neural Network [51] or its variants such as Recursive Neural Network (RecNN) [24]. These ML models take inputs from each particle sequentially, yielding a hidden state with fixed dimension. Hence they can deal well with the particle number fluctuation in the events. The third method takes the event-level information as a graph where the graph nodes and edges represent some kind of property of particles and their correlation with each other (e.g, ΔR_{ij} and θ_{ij}). The hidden state of each node gets updated based on its own properties and the properties of its adjacent edges/nodes. The applications of the graph-based models including the Graph Neural Network and its many variants can be found in [52–56] (for a review, see [57]).

Explicitly, we implement all DNNs used in our study in PyTorch [58]. We first define three modules of fully-connected neural network (FCN), using jets, FW moments and track observables as their inputs, respectively. Each of them is comprised of 5 hidden layers, with [16, 128, 128, 128, 16] neurons and activation function of ReLU. These modules are then properly connected to construct J1, J2 and J3 classifiers. For the event-level classifiers, image recognition is based on ResNet-50 CNN [59]. E1 classifier first passes the event images to the convolution part of a ResNet-50 network, and then flattens the convolution output to be the input layer of its FCN part. E2 classifier is defined as a FCN with the output neurons of E1 and the track module being its input. The CNN input is taken from a 50×50 evenly gridded $\theta - \phi$ plane where the energy intensity is defined at each pixel. The assumed image resolution is consistent with the multipole range of $l \leq 50$ for the FW moments incorporated in J2 and J3 classifiers. As a comparison, the proposed CEPC detector template has a granularity ($\phi - \eta$) of 300×360 (150×180) in the central region of ECAL (HCAL) [60], and the IDEA detector design of FCC-ee has a dual readout, with the granularity ($\phi - \eta$) of the ECAL/HCAL being 240×300 [61]. Both of them are finer than the image pixel assumed above. This leaves a space for the simulation setups to absorb the uncertainties arising from the detector granularity which could be achieved. The ResNet-50 network is trained for 50 epochs with a batch size of 512 and a learning rate of 0.0001, using the loss function of binary cross entropy. Adam optimizer [62] is used for gradient descending of the loss function. All FCNs are trained for 300 epochs with a batch size of 512 and learning rate of 0.001.

The size of the samples for training and testing each classifier is set to $10^5 + 10^5$ and 5×10^4 , respectively, in Sec. 3. In Sec. 4, we set it to $3 \times 10^5 + 3 \times 10^5$ and the real event number for 5ab^{-1} data. For the backgrounds, the training samples are defined based on their real budget after preselection. These samples are simulated with Madgraph5 [63] and parton shower with Pythia8 [64], unless otherwise specified. For the jet-level classifiers, the visible particles in each event are clustered into jets with $ee - k_T$ algorithm [28], using FastJet [65]. We assume the detector to be perfect in Sec. 3 and use the built-in CEPC-detector [66] and FCC-ee-IDEA templates [61] in DELPHES3 [67] for the Γ_h analysis in Sec. 4. In the Γ_h analysis, we also simulate the detector effects on the track observables by smearing the displacement of the SV from the primary vertex (PV) (d_{vertex}) by $5\mu\text{m}$, *i.e.*, the typical d_0 resolution for such detectors [68, 69].

3 Benchmark Study

In this section, we will analyze two benchmark scenarios, with each of their events containing two (Subsec. 3.1) and four (Subsec. 3.2) jets, respectively.

3.1 Two-Jet Events

In the two-jet benchmark study, we will develop binary classifiers to distinguish between the four classes of Zh Higgs events at $\sqrt{s} = 240$ GeV, including:

- $e^-e^+ \rightarrow Z_\nu h_b \rightarrow \nu\nu bb$;
- $e^-e^+ \rightarrow Z_\nu h_g \rightarrow \nu\nu gg$;
- $e^-e^+ \rightarrow Z_\nu h_{q_3} \rightarrow \nu\nu q_3 q_3$;
- $e^-e^+ \rightarrow Z_\nu h_{W_{qq}} \rightarrow \nu\nu W_q W_q^* \rightarrow \nu\nu qqqq$.

These four classes of events share the same production mechanism, but are differentiated by the number of jet ancestral partons and their color, electric charge, flavor, etc. At (two) jet level, these events benefit very little from the reconstructed jet kinematics (e.g., four momentum) except b tagging for their mutual distinguishing. A large portion of the information on jet ancestral partons, manifested by their showing, gets lost because of jet clustering. J2, J3, E1, E2 classifiers are expected to be able to utilize such lost information to various extents.

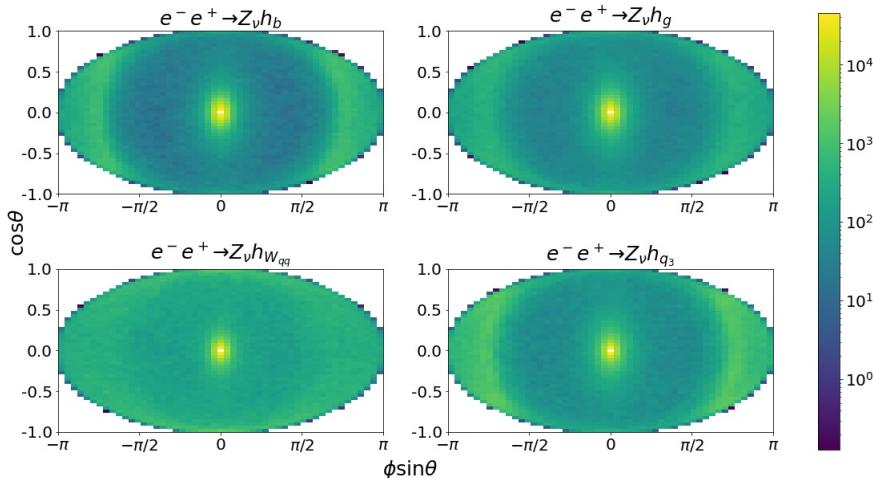


Figure 3. Cumulative Mollweide projections of 10000 events: $Z_\nu h_b$ (upper-left), $Z_\nu h_g$ (upper-right), $Z_\nu h_{W_{qq}}$ (bottom-left) and $Z_\nu h_{q_3}$ (bottom-right), with the brightness of each cell scaling with the total energy (GeV) of the particle hits received.

We present the cumulative Mollweide projections of these four classes of Zh events in Fig. 3. In the $Z_\nu h_{g,b,q_3}$ projections, we see a pupil-like structure again. But, compared

to that of the $Z_\nu Z_q$ projection in Fig. 2, the size of these pupils appears bigger. This is because the parent particle of the two jet ancestral partons for the Zh events (Higgs boson) is heavier than that of the $Z_\nu Z_q$ events (Z boson), which makes its two descendant partons to be less collimated. Analytically, the included angle between these two jet ancestral partons is determined by the formula in Eq.(2.1), but with the factor $\frac{m_Z^2}{E_Z^2}$ being replaced with $\frac{m_h^2}{E_h^2}$. The pupil-like structure becomes vague in the $Z_\nu h_{W_{qq}}$ projection. In this case, there exist multiple ways to define the said included angle, upon the jet ancestral-parton pair to consider, and especially, the one contributing the most to the projection predicts a broad distribution for its included angle (to be discussed below). These effects significantly smear such a structure.

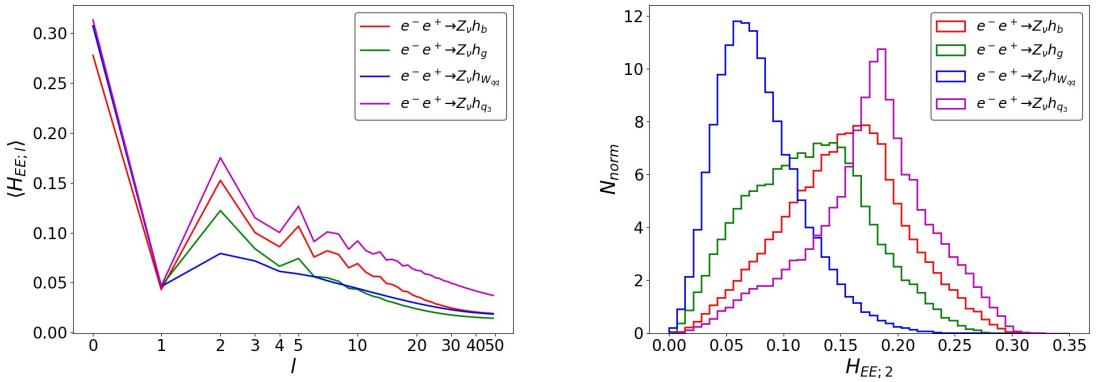


Figure 4. FW spectra of $\langle H_{EE;l} \rangle$ (left) and event distributions of $H_{EE;2}$ (right) for the two-jet samples.

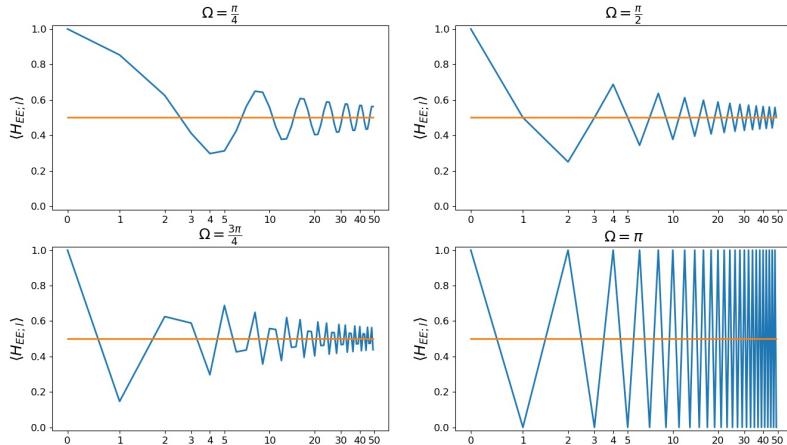


Figure 5. FW spectra of $\langle H_{EE;l} \rangle$ for a system of two partons with equal energy and varied included angle Ω . The orange lines at $\langle H_{EE;l} \rangle = 0.5$ represent the contribution from particle self-correlation to each FW moment.

As a manifestation of the event-level kinematics, the FW spectra of $\langle H_{EE;l} \rangle$ and the event distributions of $H_{EE;2}$ for the two-jet samples are presented in Fig. 4. Here and below, $\langle H_{EE;l} \rangle$ represents an average of $H_{EE;l}$ over each sample. The $\langle H_{EE;l} \rangle$ spectra are reminiscent of the CMB power spectrum. Similar to the latter, they are characterized by a series of “acoustic peaks and valleys” containing rich physical information. To demonstrate this, we plot in Fig. 5 the FW spectra for a system of two jet ancestral partons with equal energy and varied included angle. These spectra are expected to encode the four momenta of these two partons or approximately the jets that they initiate. In comparison, the FW spectra in Fig. 4 encode not only the four momenta of the jet ancestral partons but also their quantum numbers and even their parent particles’. The physical effects resulting from these quantum numbers (e.g., QCD radiation due to color charge) will deform and smear the FW spectra of the jet ancestral partons. Such deformation and smearing represent the information lost at jet level generally. Below are a series of comments on these FW spectra.

- The FW moments with odd l are zero for parity-even events, because of $P_l(-x) = (-1)^l P_l(x)$, as happens to the two-parton system with $\Omega = \pi$ in Fig. 5. In this system, the two-parton correlations contribute to each FW moment

$$\frac{1}{2} \left(1 + P_l(\cos \Omega = -1) \right) = \frac{1}{2} \left(1 + (-1)^l \right), \quad (3.1)$$

yielding a zigzag oscillation in the spectrum along the orange line.

- For $0 < \Omega < \pi$, the FW oscillation becomes less periodic w.r.t. the multipole l . Its amplitude gradually decreases as l increases, due to a suppression caused by Legendre polynomials at high l . As Ω decreases, the enhanced collinearity between the two partons gradually raises the FW spectrum at large angular scales (except $H_{EE;0}$), pushing its first acoustic peak to the high- l end. The FW spectrum becomes a straight line with $\langle H_{EE;l} \rangle \equiv 1$ in the limit of $\Omega = 0$. These effects yield various oscillation patterns for the FW spectrum. Interestingly, the FW moments at low l only are able to determine the nature of Ω qualitatively. For example, the $\langle H_{EE;1 \leq l \leq 4} \rangle$ moments for the $Z_\nu h_{g,b,q_3}$ events in Fig. 4 define a peak at $l = 2$ in their respective spectra. This matches approximately with the pattern of the two-parton system with $\Omega = \frac{3}{4}\pi$ in Fig. 5, and also consists with the indication of the cumulative Mollweide projections in Fig. 3 on the included angle between the two jet ancestral partons in these events. This method will be often used for the discussions below.
- QCD radiation and hadronization will deform and smear the parton-level FW spectrum. To understand this better, one can split the FW moments into self- and inter-correlation parts, *i.e.*,

$$H_{EE;l} = H_{EE;l}^{\text{self}} + H_{EE;l}^{\text{inter}}. \quad (3.2)$$

The two terms in Eq. (3.1) represent such a splitting also, but at parton level instead. The self-correlation of the visible particles makes a universal contribution

$$H_{EE;l}^{\text{self}} = \sum_i \frac{E_i^2}{s} \quad (3.3)$$

to all- l FW moments in each event. Its magnitude is an anti-measure of the democracy of allocating visible energy among these particles, and is irrelevant to their spatial distribution inside the detector. With more particles (e.g., because of stronger QCD radiation) and fairer energy allocation, this contribution will be reduced. In Fig. 5, the orange lines represent such a contribution from two jet ancestral partons with the same energy. If the parton showing is turned on, these orange lines will be shifted downward. This effect results in damped tails for the FW spectra in Fig. 4, and ensures the FW moments to be infrared- and collinear-safe theoretically.

- The inter-correlation of the visible particles makes an l -dependent contribution

$$H_{EE;l}^{\text{inter}} = \sum_{i,j}^{i \neq j} \frac{E_i E_j}{s} P_l(\cos \Omega_{ij}) \quad (3.4)$$

to the FW moments in each event. $H_{EE;l}^{\text{inter}}$ is sensitive to the spatial distribution of these particles inside the detector (except at $l = 0$ since $P_0(\cos \Omega_{ij}) \equiv 1$) and determines the oscillation pattern of the FW spectrum. If these particles are highly collimated, the inter-correlation between any two of them tends to be positive, due to $P_l(\cos(\Omega_{ij} \rightarrow 0)) \rightarrow 1$. The FW moments at large angular scales will gain more from this since $P_l(\cos \Omega_{ij})$ converges to one faster for the low- l modes as Ω_{ij} approaches zero. Similar argument can be applied to explain why in Fig. 5 decreasing the included angle between the two partons will raise the FW spectrum at low l and push its first acoustic peak to the high- l end. The FW oscillation pattern thus can serve as a probe to the collimation of the visible particles and jet ancestral partons in each event.

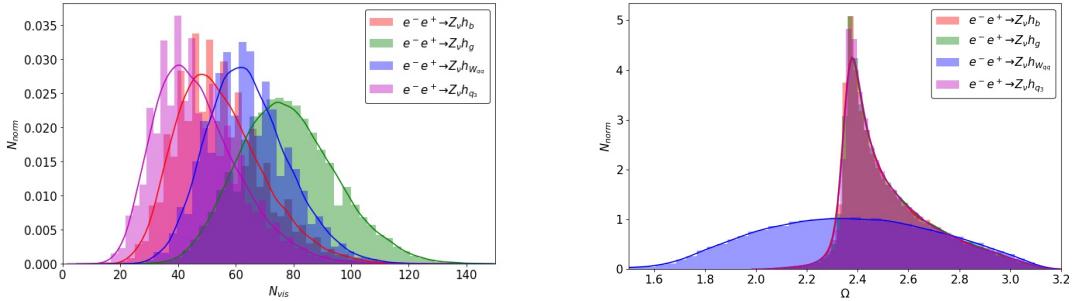


Figure 6. Distributions of the visible-particle number in each event (left) and the included angle between its two jet ancestral partons (right), for the two-jet samples. For the $Z_v h_{W_{qq}}$ events, the two partons are from the W boson with a larger mass. The solid curves are generated by fitting.

- The FW spectrum picturizes the summation of the self- and inter-correlations of the visible particles (or the jet ancestral partons at leading order) at different angular scales. At $l = 0$, the FW moment is given by $H_{EE;0} = \frac{(\sum_i E_i)^2}{s}$. It represents the squared share of the visible energy among the total in each event. As is shown in Fig. 4, the $Z_v h_b$ events tend to have more missing energy, compared to the other ones. This can be explained by leptonic decays of bottom quarks. At $l = 1$, the FW

moment is given by $H_{EE;1} = \frac{|\sum_i \vec{p}_i|^2}{s}$ (assuming all visible particles to be massless), with \vec{p}_i being the particle three-momentum. $H_{EE;1}$ hence can serve as a measure of apparent momentum violation for the events at e^-e^+ colliders. For the two-parton system in Fig. 5, the $H_{EE;1}$ values show that its momentum is equal to zero at $\Omega = \pi$ and maximized at $\Omega = 0$. Interestingly, a combination of $H_{EE;0}$ and $H_{EE;1}$ determines the visible and recoil mass of each event completely⁷. As was discussed above, the democracy effect arising from the self-correlation of the visible particles in each event and the collimation effect caused by their inter-correlation determine the profile of its FW spectrum. The FW spectra in Fig. 4 clearly demonstrate this. For example, the $Z_\nu h_g$ FW spectrum has the most-suppressed damping tail. This is mainly due to the stronger QCD radiation of gluons compared to those of quarks. As is shown in the left panel of Fig. 6, this results in more visible particles in the $Z_\nu h_g$ events than the others. Notably, the flavor of jet ancestral partons can impact the showering also. From Fig. 6 and Fig. 4, we can see that the $Z_\nu h_b$ events contain more visible particles (mainly due to bottom quark decays), compared to the $Z_\nu h_{q_3}$, and hence their FW tail is suppressed more. Another example is the $Z_\nu h_{W_{qq}}$ spectrum. Each $Z_\nu h_{W_{qq}}$ event contains four jet ancestral quarks, in comparison to two of the $Z_\nu h_{g,b,q_3}$ events. The inter-correlation at parton level thus becomes more involved in this case since it represents a collective effect of all possible parton pairings. To make the picture clear, we show the distributions of the included angle between the two (representative) jet ancestral partons in each event, in the right panel of Fig. 6. For the $Z_\nu h_{W_{qq}}$ events, the two jet ancestral partons are selected to be from the W boson with a larger invariant mass. These two partons tend to be harder, compared to the other two, and hence represent a more important contribution to $H_{EE;l}^{\text{inter}}$ at parton level. This plot shows that the $Z_\nu h_{W_{qq}}$ distribution is much broader than the others. This is consistent with the observation in Fig. 3 that there is no clear pupil-like structure in the $Z_\nu h_{W_{qq}}$ cumulative Mollweide projection. More than that, different from the $Z_\nu h_{g,b,q_3}$ parton pairs most of which have an included angle $\gtrsim \frac{3\pi}{4}$, a large portion of the $Z_\nu h_{W_{qq}}$ parton pairs prefer one $\lesssim \frac{3\pi}{4}$. This explains why at $l = 3$ there is a convex in the $Z_\nu h_{W_{qq}}$ spectrum, and a concave instead for the others.

- The two-jet events of each sample define a distribution w.r.t. $H_{EE;l}$, with the ones at $l = 2$ being shown in the right panel of Fig. 4. This reminds us that, unlike the CMB power spectrum, the $\langle H_{EE;l} \rangle$ spectrum is free from big sample variance, because of the large size of collider data. These distributions also manifest the order of the heights of the first acoustic peaks in the $\langle H_{EE;l} \rangle$ spectra, as is shown in the left panel of this figure. Notably, the $\langle H_{EE;l} \rangle$ spectra do not fully reflect the discrimination power of the FW moments. This power also relies on the event-distribution profiles of the data samples at each multipole.

⁷As a comparison, the CMB power spectrum for temperature fluctuation is physically less interesting for $l < 2$. At $l = 0$ the moment is zero by definition. At $l = 1$ the moment is dominated by the Doppler effect caused by the motion of the solar system w.r.t. the last scattering surface, which makes inseparable the cosmological dipole caused by large-scale perturbations.

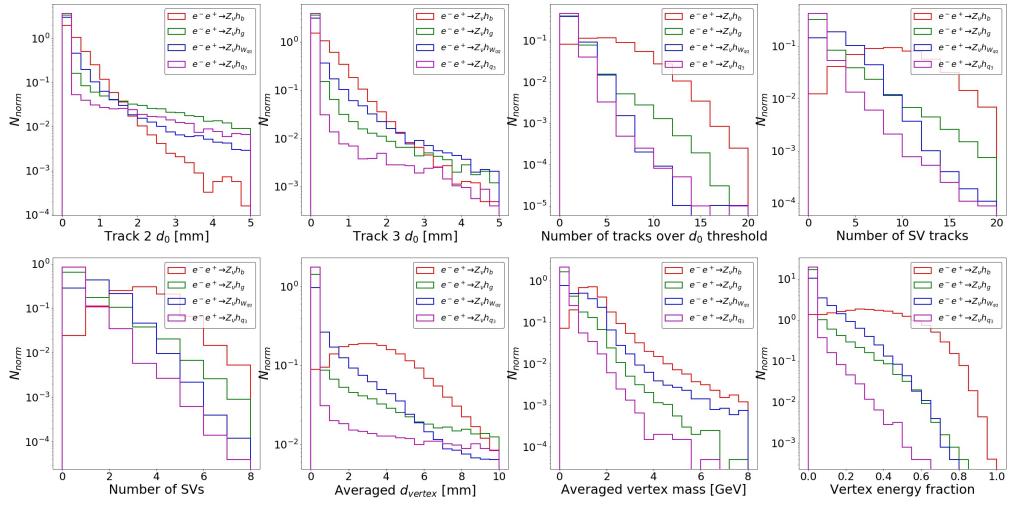


Figure 7. Normalized distributions of the (SV) track observables for the two-jet samples. Here the tracks in each event are sorted ascendingly with the impact parameter d_0 . The d_0 threshold is set to be 0.5 mm. The vertex energy fraction is defined w.r.t. the total (PV + SV) track energy.

Another class of kinematic information arises from the tracks of the SV (defined by $d_{\text{vertex}} \geq 10\mu m$). As is well-known, the heavy-flavor quarks such as bottom tend to decay as a SV. This provides an important sign for recognizing these particles. Fig. 7 displays the normalized distributions of a set of (SV) track observables for the two-jet samples (for simplicity, below we will not stress their “SV” nature). Here the tracks are defined at truth level with no detector smearing. These observables have been used for the DNN-based jet classification in [70]. As is expected, the $Z_\nu h_b$ events demonstrate highly distinguishable track features from the others, especially from the $Z_\nu h_{q_3}$ events. These features are shared to some extent by the $Z_\nu h_g$ and $Z_\nu h_{Wq}$ events. This is largely because some heavy-flavor quarks such as charm quarks can be generated from gluon splitting and hadronic W decay.

The receiver operating characteristic (ROC) curves and their area under the curve (AUC) for the binary classifiers to distinguish the four classes of two-jet events are presented in Fig. 8⁸. In this figure, the blue and purple curves display the performance of J1 and E1 classifiers respectively. With the ML techniques of image recognition, the event-level classifiers are expected to be able to utilize the kinematic information to the greatest extent, upon the detector and image resolutions. Indeed, E1 classifiers yield an AUC universally bigger than that of J1 ones. The light-blue curves in this figure display the performance of J2 classifiers. They indicate that the FW moments of $H_{EE;l \leq 50}$ compensate for a large portion of the information lost at jet level. The AUC of J2 classifier even becomes comparable to that of E1 for the acceptance of the $Z_\nu h_b$ against the $Z_\nu h_{Wq}$, due to this.

⁸In this paper, the ROC curves are drawn as the acceptance of one classes of events against another one. So the “AUC” of each ROC curve is not really “area under the curve”, but the area above the curve.

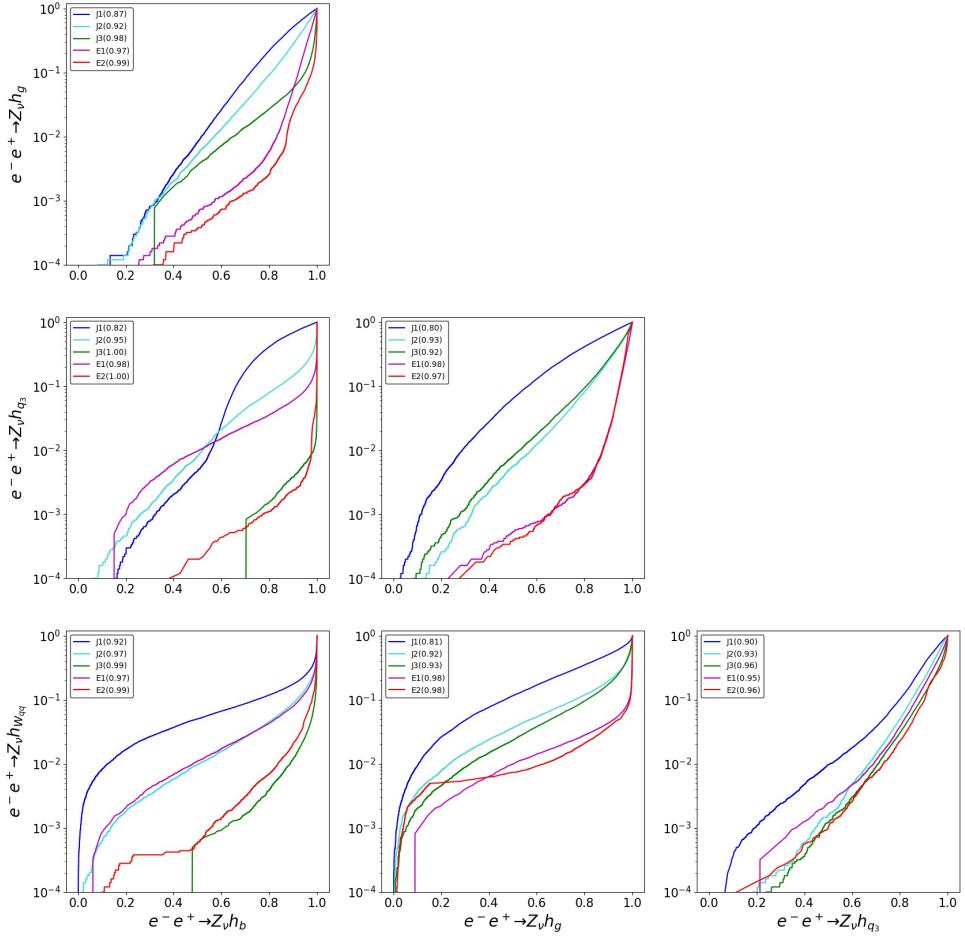


Figure 8. ROC curves and their AUC for the binary classifiers to distinguish between the four classes of two-jet events.

It is also encouraging to see that the $Z_\nu h_b$ events can be tagged at 50% level, with a faking rate of 1% or below for the others, with no track observables being applied yet. For the case against the $Z_\nu h_{q_3}$, this mainly benefits from the flavor-related kinematics at event level such as multiplicity of the visible particles (see Fig. 6 and Fig. 4). Despite these, the AUC gap between J1 and E1 classifiers is not fully addressed by $H_{EE;l \leq 50}$ in most cases. This may imply that the FW moments not included here or/and the multi-spectra carry part of the information lost at jet level. We will leave the relevant exploration to a future work. At last, the green and red curves display the performance of J3 and E2 classifiers. Not surprisingly, the incorporation of track observables yields a remarkable improvement to both jet-level and event-level classifiers in distinguishing the $Z_\nu h_b$ events from the others.

3.2 Four-Jet Events

In the four-jet benchmark study, we will develop binary classifiers to distinguish between the four classes of WW/ZZ events at $\sqrt{s} = 240$ GeV, including:

- $e^-e^+ \rightarrow Z_qZ_q \rightarrow qqqq$;
- $e^-e^+ \rightarrow Z_qZ_b \rightarrow qqbb$;
- $e^-e^+ \rightarrow Z_bZ_b \rightarrow bbbb$;
- $e^-e^+ \rightarrow W_qW_q \rightarrow qqqq$.

Among these, the first three classes of events share the ZZ production, but are differentiated by the flavor of their descendant partons. J2, J3, E1 and E2 classifiers are expected to be able to utilize the flavor-related event-level kinematics for their classification. The last class of events have different intermediate gauge bosons from those of the others. Their distinction may benefit additionally from the event-level kinematics manifesting the four momenta of the jet ancestral partons and even the nature of their parent gauge bosons. This is especially important for distinguishing between the W_qW_q and Z_qZ_q events, or probing anomalous triple-gauge couplings at e^-e^+ colliders.

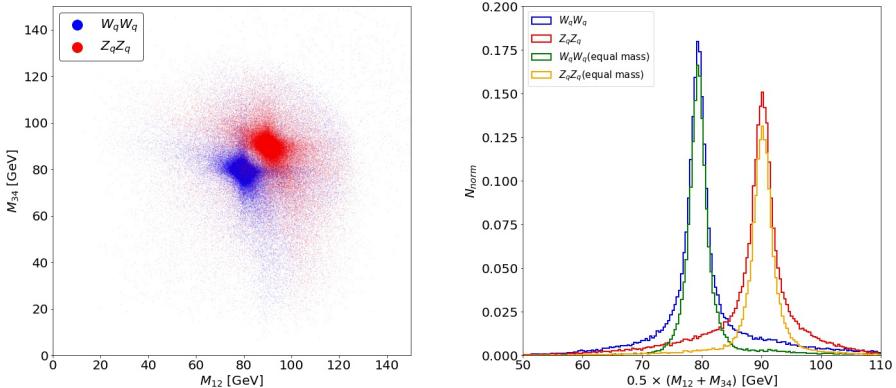


Figure 9. Distributions of the reconstructed W_qW_q and Z_qZ_q events at the $m_{12} - m_{34}$ plane (left) and w.r.t. $(m_{12} + m_{34})/2$ (right), with a perfect detector.

Actually, the W_qW_q and Z_qZ_q events have been applied to illustrate the problem of information deformation in jet clustering at CEPC [19]. In this study, these events were reconstructed by minimizing the measure

$$\chi^2 = \frac{(m_{ab} - m_X)^2 + (m_{cd} - m_X)^2}{\sigma_B^2}. \quad (3.5)$$

Here $\{ab, cd\}$ runs over all possible jet pairings among the four, with $\{12, 34\}$ representing the optimal one, X runs over W and Z bosons, and σ_B denotes the standard deviation of the jet-pair invariant mass. In Fig. 9 we show the distributions of the reconstructed W_qW_q

and Z_qZ_q events at the $m_{12} - m_{34}$ plane and w.r.t. $(m_{12} + m_{34})/2$, with a perfect detector. Mainly due to the information deformation of jets, a good portion of these events are not well-reconstructed. This results in a separation of 50% between these two classes of events. With a condition of mass equality $|m_{12} - m_{34}| < 10\text{GeV}$ [19] being applied, this separation increases to 78%, at the cost of losing 35% W_qW_q and 39% Z_qZ_q events. These results are consistent with the observations made in [19].

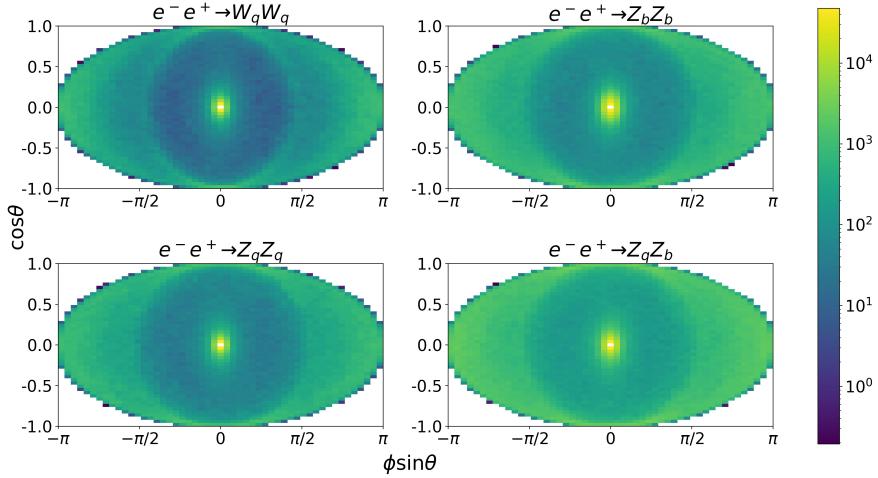


Figure 10. Cumulative Mollweide projections of 10000 events: W_qW_q (upper-left), Z_bZ_b (upper-right), Z_qZ_q (bottom-left) and Z_qZ_b (bottom-right), with the brightness of each cell scaling with the total energy (GeV) of the particle hits received.

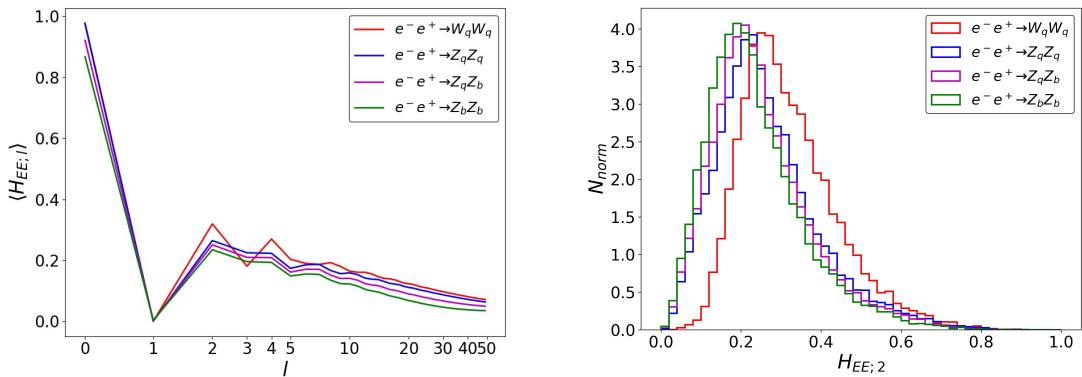


Figure 11. FW spectra of $\langle H_{EE;l} \rangle$ (left) and event distributions of $H_{EE;2}$ (right) for the four-jet samples.

We present the cumulative Mollweide projections of these four classes of four-jet events in Fig. 10. All of them display a pupil-like structure clearly. For these events, each of them contains two intermediate gauge bosons moving oppositely. The Z_qZ_q event is somewhat

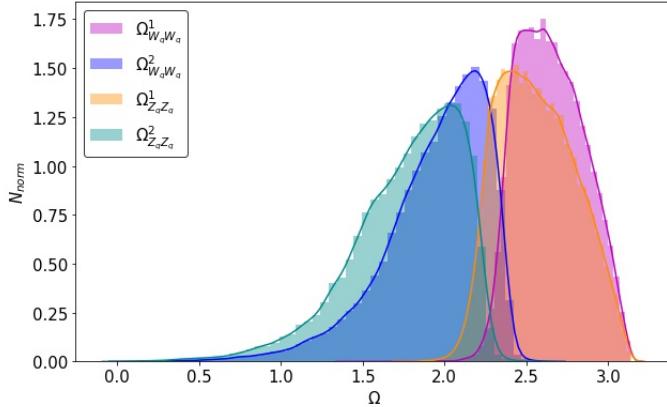


Figure 12. Distributions of the largest (Ω^1) and the second largest (Ω^2) included angles between the paired jet ancestral partons, for the $W_q W_q$ and $Z_q Z_q$ events. Here one of the paired partons is required to be the most energetic among the four, and another one to be from a different parent particle.

like a double copy of Z_q in the $Z_q Z_q$ event, which explains why the pupil size of the $Z_q Z_q$ projection in Fig. 10 is comparable to that of the $Z_q Z_q$ projection in Fig. 2. The $W_q W_q$ event has two more-boosted intermediate gauge bosons, due to $m_W < m_Z$. Its cumulative Mollweide projection thus has a smaller pupil structure, with its radius being smaller than $\frac{\pi}{2}$ at $\cos \theta = 0$. In relation to this, one has a bigger chance to find the jet ancestral-parton pairs with an included angle close to π in the $W_q W_q$ events, compared to the $Z_{q,b} Z_{q,b}$ ones. Most of such pairs are composed of the partons with different parent particles. Mainly due to this, in the $W_q W_q$ cumulative Mollweide projection, the region outside the pupil is approximately split to two: the dark one in the middle and the bright one close to the edge. We make this more explicit with Fig. 12. It is exactly one to the partons comprising the magenta pair of $W_q W_q$ in this figure that results in the said bright region. Fig. 11 displays the FW spectra of $\langle H_{EE;l} \rangle$ and the event distributions of $H_{EE;2}$ for the four-jet samples. As was discussed before, the FW oscillation pattern is a collective manifestation of the inter-correlations among all jet ancestral partons at leading order. Since the $W_q W_q$ events have a bigger chance to be found to have a parton pair with an obtuse included angle, it is not strange that the peaks in its FW spectrum are sharper than those of the $Z_{q,b} Z_{q,b}$ FW spectra.

The ROC curves and their AUC for the binary classifiers to distinguish between the four classes of four-jet events are presented in Fig. 13. Similar to the two-jet case, E1 classifiers perform universally better than J1 ones. But, by incorporating the FW moments $H_{EE;l \leq 50}$, J2 classifiers greatly reduce their AUC gap in most cases. With the track information, J3 and E2 classifiers further improve the AUC values of J2 and E1. The extent is positively correlated with the difference of the bottom-quark number between the two classes of events to classify. Among these, the classifiers of the $Z_b Z_b$ against the others are especially informative. As is expected, its J1 classifiers perform best against the $W_q W_q$ and

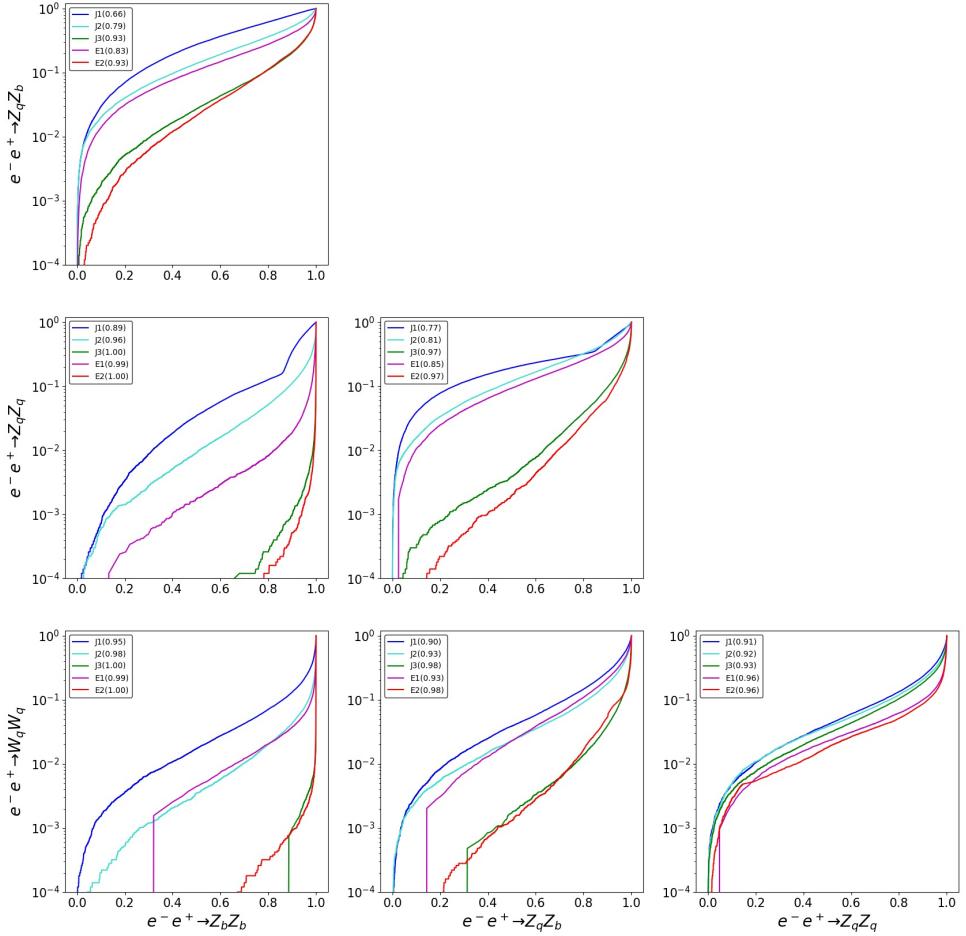


Figure 13. ROC curves and their AUC for the binary classifiers to distinguish between the four classes of four-jet events.

worst against the Z_qZ_b . Thus a space is created for the FW moments (or the event-level kinematics) and the track observables to play a non-trivial role in the latter case. Indeed, the Z_bZ_b classifiers against the Z_qZ_b gain the most from them among all constructions in Fig. 13, resulting in a great AUC improvement from 0.66 to 0.93 for both J3 and E2 ones.

At last, let us take a look at the classification of the W_qW_q and Z_qZ_q events. We show the response of these two classes of events to J1, J2, J3, E1 and E2 classifiers in Fig. 14. These classifiers all improve the event tagging accuracy to some extent, compared to the original analysis discussed above. But, E1 and E2 ones have a better performance than the others. They yield a separation of 70% and 73% respectively (J1: 53%; J2: 56%; J3: 59%), without losing any events, in comparison to the original 50%.

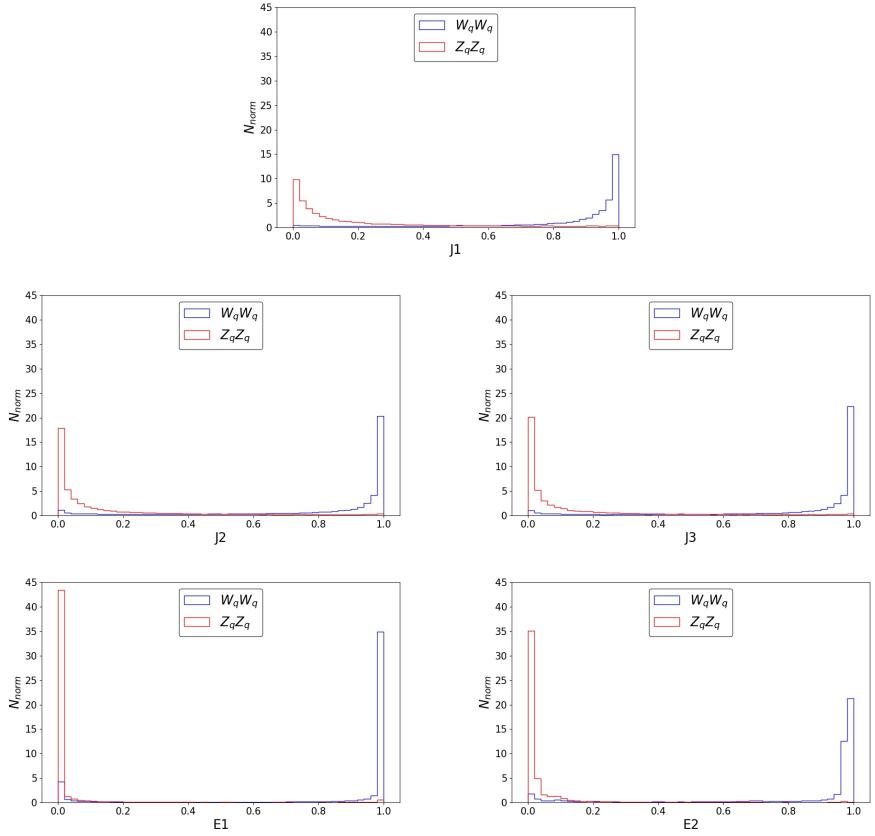


Figure 14. Response of the W_qW_q and Z_qZ_q events to J_1 , J_2 , J_3 , E_1 and E_2 classifiers.

4 Application: Higgs Decay Width (Γ_h)

In this section we will apply the binary classifiers to measuring the SM Γ_h , one of the most important tasks at future e^-e^+ colliders, with the data of 5ab⁻¹@240GeV. Upon the strategy taken, this measurement may involve analyzing the events with two, four, and six jets. In this study, only the first two cases are relevant.

4.1 Γ_h Measurement at e^-e^+ Colliders

There exist multiple methods to measure the SM Γ_h at e^-e^+ colliders. Here are several representative ones.

- Method A. Γ_h is measured using the relation

$$\Gamma_h = \frac{\Gamma(h \rightarrow ZZ^*)}{\text{BR}(h \rightarrow ZZ^*)} \propto \frac{[\sigma(Zh)]}{\text{BR}(h \rightarrow ZZ^*)} = \frac{[\sigma(Zh)]^2}{[\sigma(Zh_Z)]}. \quad (4.1)$$

One needs to measure the quantities in the square brackets first for determining Γ_h . This method requires analyzing the Zh data only, and hence is more straightforward compared to many other strategies. Its major drawback is that the $\sigma(Zh_Z)$ signal rate is small, while its irreducible backgrounds are relatively large.

- Method B. Γ_h is measured using the relation

$$\Gamma_h = \frac{\Gamma(h \rightarrow WW^*)}{\text{BR}(h \rightarrow WW^*)} \propto \frac{\sigma(\nu\nu h)}{\text{BR}(h \rightarrow WW^*)} = \frac{[\sigma(\nu\nu h_b)][\sigma(Zh)]^2}{[\sigma(Zh_b)][\sigma(Zh_W)]} . \quad (4.2)$$

This method utilizes the large signal rates of $\sigma(Zh_b)$ and $\sigma(Zh_W)$, and hence largely avoids the drawback of method A.

- Method C. Γ_h is measured using the relation [71]

$$\Gamma_h = \frac{\Gamma(h \rightarrow WW^*)}{\text{BR}(h \rightarrow WW^*)} \propto \frac{\sigma(\nu\nu h)}{\text{BR}(h \rightarrow WW^*)} = \frac{[\sigma(\nu\nu h_b)]^2[\sigma(Zh)]^2}{[\sigma(\nu\nu h_W)][\sigma(Zh_b)]^2} . \quad (4.3)$$

This method is similar to Method B, except that $\sigma(Zh_W)$, one of the key intermediate quantities to measure, is replaced with $\sigma(\nu\nu h_W)$. This method mainly benefits from the enhancement of the $\sigma(\nu\nu h)$ rate, as \sqrt{s} increases.

These three methods totally involve six intermediate quantities to measure:

$$\sigma(Zh), \sigma(Zh_Z), \sigma(\nu\nu h_b), \sigma(Zh_b), \sigma(Zh_W) \text{ and } \sigma(\nu\nu h_W) . \quad (4.4)$$

With some of them, one can measure Γ_h with a fourth method, *i.e.*,

- Method D. Γ_h is measured using the relation

$$\Gamma_h = \frac{\Gamma(h \rightarrow WW^*)}{\text{BR}(h \rightarrow WW^*)} \propto \frac{\sigma(\nu\nu h)}{\text{BR}(h \rightarrow WW^*)} = \frac{[\sigma(\nu\nu h_W)][\sigma(Zh)]^2}{[\sigma(Zh_W)]^2} . \quad (4.5)$$

This method shares the advantage of Method C, mainly benefitting from the enhancement of the $\sigma(\nu\nu h)$ rate at high \sqrt{s} .

With the relevant intermediate quantities being measured, one can calculate the precision of measuring Γ_h , using the formula of Gaussian statistics

$$\frac{\delta\Gamma_h}{\Gamma_h} = \sqrt{\sum_i \left(\frac{n_i \delta \mathcal{O}_i}{\mathcal{O}_i} \right)^2} . \quad (4.6)$$

Here \mathcal{O}_i represents the intermediate quantities to measure in Eq. (4.1 - 4.3, 4.5), $\delta \mathcal{O}_i$ its absolute precision, and n_i its power.

The expected precisions of measuring \mathcal{O}_i and Γ_h in a variety of low- \sqrt{s} operation scenarios at future e^-e^+ colliders are summarized in Table 3 and Table 4, respectively. Most of these analyses were pursued at jet level with a cut-based strategy. Method B provides a better precision of measuring Γ_h , compared to the other methods. Yet, to reduce the complication in Method B that the Higgs events with different decay modes serve as backgrounds mutually, one can instead measure the hadronic Higgs events in a relatively inclusive way. This idea can be also applied to Method C. We term these inclusive methods as

Measurements (%)	CEPC ₂₄₀₍₂₅₀₎ [17, 68]	FCC ₂₄₀ [18]	FCC ₃₆₅ [18]	CILC ₃₅₀ [72]	ILC ₂₅₀ [71, 73–75]
$\sigma(Zh)$	0.5 (0.5)	0.5	0.9	1.6	2.8
$\sigma(Zh_b)$	0.27 (0.26)	0.3	0.5	0.86	1.2
$\sigma(Zh_c)$	3.3 (3.1)	2.2	3.5	14	8.3
$\sigma(Zh_g)$	1.3 (1.2)	1.9	6.5	6.1	7.0
$\sigma(Zh_W)$	1.0 (0.9)	1.2	2.6	5.1	6.4
$\sigma(Zh_Z)$	5.1 (4.9)	4.4	12	-	19
$\sigma(\nu\nu h_b)$	3.2 (2.9)	3.1	0.9	1.9	10.5
$\sigma(\nu\nu h_c)$	-	-	10	26	-
$\sigma(\nu\nu h_W)$	-	-	3.0	-	-

Table 3. Expected precisions of measuring \mathcal{O}_i at e^-e^+ colliders. The CEPC₂₄₀ precisions are extrapolated from the CEPC₂₅₀ ones (inside the parentheses) [17, 68]. The ILC₂₅₀ results are based on its previous baseline luminosity (250 fb^{-1}).

Γ_h (%)	CEPC ₂₄₀₍₂₅₀₎ [17, 68]	FCC ₂₄₀ [18]	FCC ₂₄₀₊₃₆₅ [18]	CLIC ₃₅₀ [72]	ILC ₂₅₀ [71, 73, 74]
Method A	5.1 (5.0)	4.5*	4.2*	-	20*
Method B	3.5 (3.2)	3.5*	1.7*	6.7	13
Method C	-	-	3.4*	-	-
Combined	2.8 (2.7)	2.7	1.3	6.7	11

Table 4. Expected precisions of measuring the SM Γ_h at e^-e^+ colliders. The CEPC₂₄₀ precisions are extrapolated from the CEPC₂₅₀ ones (inside the parentheses) [17, 68]. The numbers marked with “*” are derived from Table 3, using Eq. (4.6).

- Method B'. Γ_h is measured using the relation

$$\Gamma_h = \frac{\Gamma(h \rightarrow WW^*)}{\text{BR}(h \rightarrow WW^*)} \propto \frac{\sigma(\nu\nu h)}{\text{BR}(h \rightarrow WW^*)} = \frac{[\sigma(\nu\nu h_h)][\sigma(Zh)]^2}{[\sigma(Zh_h)][\sigma(Zh_W)]}. \quad (4.7)$$

Here h_h denotes the inclusive two-body Higgs decays $h \rightarrow bb$, cc , gg and $\tau\tau$. We exclude $h \rightarrow V_q V_q^*$ from h_h , to avoid a correlation between the $\sigma(Zh_h)$ and $\sigma(Zh_W)$ measurements.

- Method C'. Γ_h is measured using the relation

$$\Gamma_h = \frac{\Gamma(h \rightarrow WW^*)}{\text{BR}(h \rightarrow WW^*)} \propto \frac{\sigma(\nu\nu h)}{\text{BR}(h \rightarrow WW^*)} = \frac{[\sigma(\nu\nu h_h)]^2 [\sigma(Zh)]^2}{[\sigma(\nu\nu h_W)][\sigma(Zh_h)]^2}. \quad (4.8)$$

Then Γ_h can be determined using either of Method A, B', C' and D and Eq. (4.6), with the new set of intermediate quantities

$$\sigma(Zh), \sigma(Zh_Z), \sigma(\nu\nu h_h), \sigma(Zh_h), \sigma(Zh_W) \text{ and } \sigma(\nu\nu h_W) \quad (4.9)$$

being measured.

We will take Method B' in this study. Among the four intermediate quantities, $\sigma(Zh)$ can be measured with a precision of sub-percent level. $\sigma(Zh_h)$ is expected to be well-measured also, given the dominance of $\sigma(Zh_b)$ in its signal rate. As is shown in Table 3, the precisions for both the $\sigma(Zh)$ and $\sigma(Zh_b)$ measurements are high. So we would expect the limitations for precisely measuring Γ_h to mainly arise from the $\sigma(Zh_W)$ and $\sigma(\nu\nu h_h)$

Signal	Backgrounds			
$Z_\nu h_{W_{lq}}$	$W_l W_q$	$Z_l Z_{q_5}$	$Z_\nu h_\tau$	
8.57×10^3	2.41×10^5	1.04×10^3	3.22×10^3	
$Z_\nu h_{W_{qq}}$	$Z_\nu Z_{q_5}$	$q_5 q_5(\gamma)$	$\gamma\gamma \rightarrow q_5 q_5$	$W_q W_q/Z_{q_5} Z_{q_5}$
1.65×10^4	5.61×10^4	4.01×10^4	4.41×10^2	1.42×10^4
	$Z_\nu h_b$	$Z_\nu h_c$	$Z_\nu h_g$	$Z_\nu h_{Z_{q_5} q_5}$
	8.78×10^4	4.71×10^3	1.41×10^4	2.10×10^3

Table 5. Numbers of the signal and its main background events after preselection, for measuring $\sigma(Z_\nu h_{W_{lq}})$ and $\sigma(Z_\nu h_{W_{qq}})$. We use Pythia8 to simulate the impact of initial state radiation and beamstrahlung for the $q_5 q_5$ production. The τ decays of the W/Z bosons are also incorporated in the $Z_\nu h_{W_{lq}}$ analysis, with $l = e, \mu$ and τ .

measurements. Below we will focus on these two difficult cases. We will assume that all parameters relevant to their analyses have been precisely measured, and will not consider the impact of systematic errors.

4.2 Measuring $\sigma(Zh)\text{BR}(h \rightarrow WW^*)$

At CEPC, the measurement of $\sigma(Zh_W)$ was simulated with four decay modes of Z boson [17]: e^-e^+ , $\mu^-\mu^+$, $\nu\nu$ and qq . A combination of these yields a precision of 0.9% at 250GeV (see Table 3). The most important contribution arises from the two processes of $Z_\nu h_{W_{lq}}$ and $Z_\nu h_{W_{qq}}$ which give a combined precision of 1.5% [17]. Below we will develop the binary classifiers for their measurements.

To improve the training efficiency of these classifiers, we apply a set of preselection cuts first. For the $\sigma(Z_\nu h_{W_{lq}})$ analysis, we pass the events with one isolated lepton (either e^-/e^+ or μ^-/μ^+ ; $p_T > 10\text{GeV}$), and require the visible particles in each event including this lepton to have an invariant mass $\in [35, 125]\text{GeV}$, a recoil mass $\in [100, 200]\text{GeV}$, and a vector sum of transverse momentum $\in [10, 75]\text{GeV}$. The $W_l W_q$ events are then dominant in the backgrounds⁹. The visible particles are subsequently clustered into two jets for the J1-, J2- and J3-based analyses. For the $\sigma(Z_\nu h_{W_{qq}})$ analysis, we veto the events with any isolated leptons and require the visible particles in each event to have an invariant mass $\in [100, 150]\text{GeV}$, a recoil mass $\in [75, 150]\text{GeV}$, and a vector sum of transverse momentum $\in [20, 80]\text{GeV}$. Different from the $Z_\nu h_{W_{lq}}$ case, both non-Higgs events such as $Z_\nu Z_{q_5}$ and $q_5 q_5(\gamma)$ and Higgs events including $Z_\nu h_b$ contribute to the backgrounds significantly. The visible particles are subsequently clustered into four jets for the jet-level analyses. The numbers of the signal and its main background events after preselection, for measuring $\sigma(Z_\nu h_{W_{lq}})$ and $\sigma(Z_\nu h_{W_{qq}})$, are summarized in Table 5.

The FW spectra of $\langle H_{EE;l} \rangle$ for the $Z_\nu h_{W_{lq}}$ and $Z_\nu h_{W_{qq}}$, and their respective background samples are presented in the upper panels of Fig. 15. In both cases, the signal and backgrounds have close $\langle H_{EE;0} \rangle$ and $\langle H_{EE;1} \rangle$ values. This is not very surprising since these

⁹By requiring one hard isolated lepton, we notice that $\sigma(e\nu_e W_q)$, *i.e.*, the single- W rate, is greatly suppressed compared to $\sigma(Z_\nu h_{W_{lq}})$. So we will not include these events in this analysis.

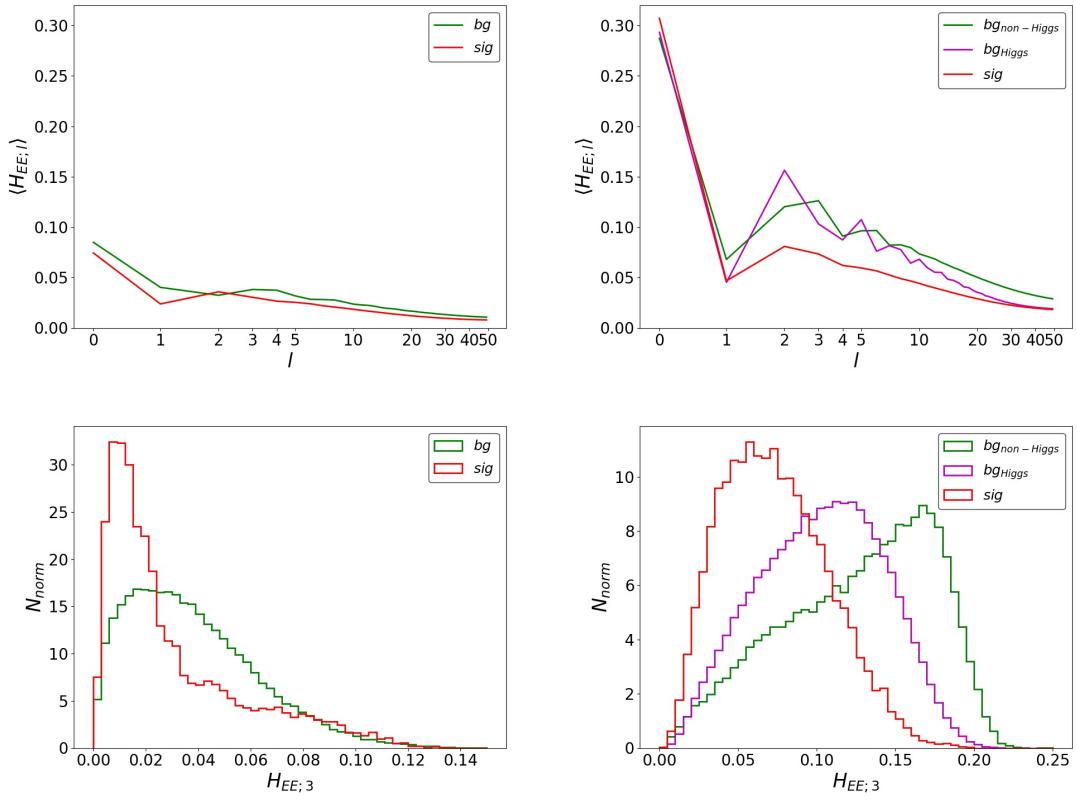


Figure 15. FW spectra of $\langle H_{EE;l} \rangle$ (upper) and event distributions of $H_{EE;3}$ (bottom) for the $Z_\nu h_{W_{lq}}$ (left) and $Z_\nu h_{W_{qq}}$ (right) and their respective background samples after preselection. In the left panels, the contributions from the isolated lepton have been excluded.

events are all preselected from the phase space in favor of the signal. Despite this, these $\langle H_{EE;l} \rangle$ spectra demonstrate a series of characteristic features which may assist distinguishing between the signal and backgrounds. Some of them are related to the discussions in Subsec. 3.1. In the upper-left panel, the signal and background spectra are characterized by a peak and a valley, respectively, at $l = 2$. This indicates that the included angle between the two jet ancestral quarks of $W_l W_q$ is not far from $\frac{\pi}{2}$ (see Fig. 5), while the one of $Z_\nu h_{W_{lq}}$ tends to be wider. In the upper-right panel, the convex-concave structure of the FW spectra at $l = 3$ indicates that the Higgs backgrounds tend to have a wider included angle between their jet ancestral partons, while the non-Higgs backgrounds favor a narrower one. The former case has been discussed before. The latter one can be understood also. For the $Z_\nu Z_{q_5}$ events, without preselection this angle will be reduced to $\sim \frac{\pi}{2}$, while for the $q_5 q_5(\gamma)$ events, most of the jet ancestral quarks are produced at Z pole with the Z boson being recoiled against their initial state radiation.

As was discussed in Subsec. 3.1, the discrimination power of the FW moments also relies on the distribution profiles of the signal and background events at each multipole. For illustration, we present the ones of $H_{EE;3}$ for the $\sigma(Z_\nu h_{W_{lq}})$ and $\sigma(Z_\nu h_{W_{qq}})$ analyses in the bottom panels of Fig. 15. The relevant FW moments demonstrate certain features

to distinguish between the signal and backgrounds in both cases. In the bottom-left panel, the signal distribution has a sharp peak at small $H_{EE;3}$, in comparison to a shape curved down for the background profile. In the bottom-right panel the non-Higgs background events tend to have a bigger $H_{EE;3}$ value for the relatively small included angle between their two jet ancestral partons.

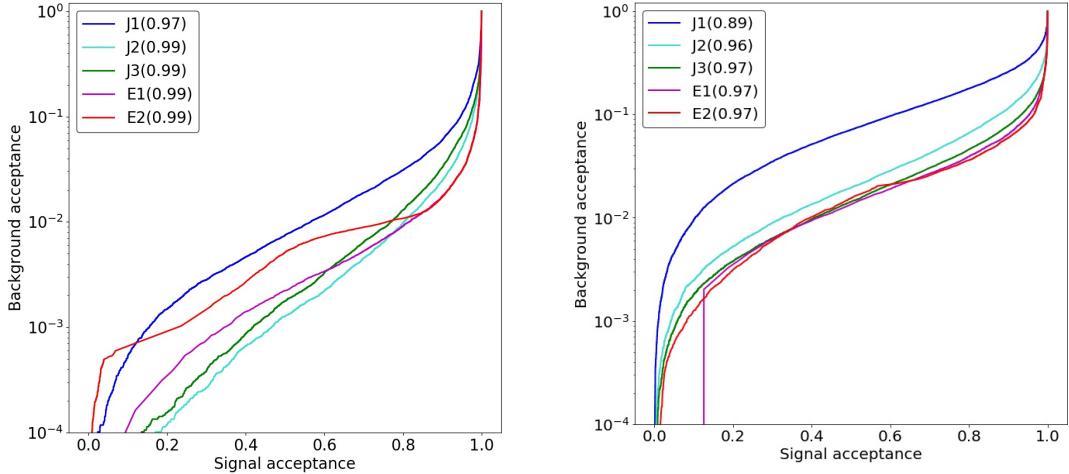


Figure 16. ROC curves and their AUC for the binary classifiers to distinguish the $Z_{\nu}h_{W_{lq}}$ (left) and $Z_{\nu}h_{W_{qq}}$ (right), from their respective background events.

Fig. 16 displays the ROC curves and their AUC for the binary classifiers to distinguish the $Z_{\nu}h_{W_{lq}}$ and $Z_{\nu}h_{W_{qq}}$ from their respective backgrounds (the event responses to these classifiers are shown in Fig. 21 in Appendix A). In both analyses, E1 classifiers yield an AUC bigger than that of J1 ones. Yet, by incorporating the FW moments $H_{EE;l \leq 50}$, J2 classifiers fill up their gaps almost completely. The track observables are then applied to J3 and E2 classifiers, which improve the $\sigma(Z_{\nu}h_{W_{qq}})$ analysis slightly. In this case, the track information is useful in rejecting the $Z_{\nu}h_b$ events, the dominant Higgs background. A combination of these yields the same AUC values (up to $\mathcal{O}(10^{-2})$) for J2, J3, E1 and E2 classifiers of $\sigma(Z_{\nu}h_{W_{lq}})$ and for J3, E1 and E2 classifiers of $\sigma(Z_{\nu}h_{W_{qq}})$.

The precisions of measuring $\sigma(Z_{\nu}h_{W_{lq}})$ and $\sigma(Z_{\nu}h_{W_{qq}})$ with these classifiers are summarized in Table 7. Not surprisingly, J2, J3, E1 and E2 classifiers result in comparable precisions in both analyses. Combining them allows the $\sigma(Zh_W)$ to be measured with a precision of 0.9% at 240GeV. As a comparison, the CEPC baseline precision is 1.5%, which is achieved based on a cut-based analysis of the same channels, with the data of 5.6ab^{-1} @250GeV [17].

4.3 Measuring $\sigma(\nu\nu h)\text{BR}(h \rightarrow \text{hadrons})$

In Method B, $\sigma(\nu\nu h_b)$ is a crucial intermediate quantity to measure. As is shown in Table 3, the precision for its measurement is 3.2 (2.9)% and 3.1%, at CEPC₂₄₀₍₂₅₀₎ and FCC₂₄₀,

respectively. This comprises the main bottleneck for improving the precision of measuring Γ_h . In Method B', we replace $\sigma(\nu\nu h_b)$ with an inclusive quantity $\sigma(\nu\nu h_h)$, for reducing unnecessary complexity in the analysis. This also brings an increasement of $\sim 22\%$ in the signal rate after event preselection. Notably, both methods suffer a subtlety caused by the interference between the signal of $\nu\nu h_{b,h}$ and its irreducible background $Z_\nu h_{b,h}$. To apply Eq. (4.2) and Eq. (4.5), one needs to properly simulate this effect in the analysis. Yet, this was not explicitly implemented in [17, 18, 72, 76]. For the purpose of method comparison, we will tolerate this uncertainty below by simply neglecting it. We do not expect that such a treatment will qualitatively change the conclusions reached in this paper.

Signal	$\nu\nu h_b$	$\nu\nu h_c$	$\nu\nu h_g$	$\nu\nu h_\tau$
1.51×10^4	1.24×10^4	6.43×10^2	1.92×10^3	1.50×10^2
Higgs backgrounds	$Z_\nu h_b$	$Z_\nu h_c$	$Z_\nu h_g$	$Z_\nu h_\tau$
1.39×10^5	9.47×10^4	5.08×10^3	1.52×10^4	1.06×10^3
	$Z_\nu h_{V_{q_5 q_5}}$	$\nu\nu h_{V_{q_5 q_5}}$		
	2.01×10^4	2.51×10^3		
Non-Higgs backgrounds	$q_5 q_5(\gamma)/\gamma\gamma \rightarrow q_5 q_5$	$W_q W_q$	$Z_{q_5} Z_{q_5}$	$Z_\nu Z_{q_5}$
1.40×10^5	$6.79 \times 10^4 / 2.81 \times 10^3$	1.26×10^4	6.61×10^2	5.61×10^4

Table 6. Numbers of the signal and its main background events after preselection, for measuring $\sigma(\nu\nu h_h)$.

For training the classifiers efficiently, we preselect the events by requiring their visible particles to have total energy $\in [105, 155]\text{GeV}$, invariant mass $\in [100, 135]\text{GeV}$, recoil mass $\in [65, 135]\text{GeV}$, a vector sum of $p_T > 10\text{GeV}$ and $p_z < 60\text{GeV}$, and vetoing the events with any isolated leptons. The visible particles in each event are subsequently clustered into two jets for the J1-, J2- and J3-based analyses. The numbers of the signal and its main background events after preselection are summarized in Table 6. Both Higgs events such as $Z_\nu h_h$ and non-Higgs events such as $Z_\nu Z_{q_5}$ and $q_5 q_5(\gamma)$ contribute to the backgrounds significantly.

Fig. 17 displays the $\langle H_{EE;l} \rangle$ spectra and the event distributions of $H_{EE;3}$ for the $\nu\nu h_h$ and its background samples after preselection. Partly for preselection, these $\langle H_{EE;l} \rangle$ spectra are close to each other. But, as was stressed, the discrimination power of the FW moments also relies on the distribution profiles of the signal and background events at each multipole. For illustrating this, we show the event distributions of the invariant and recoil masses of the visible particles in Fig. 18. These two observables are determined by $H_{EE;0}$ and $H_{EE;1}$ completely, as was discussed in Subsec. 3.1, with the formulae given by

$$M_{\text{inv}} = \sqrt{s(H_{EE;0} - H_{EE;1})}; \quad M_{\text{rec}} = \sqrt{s + M_{\text{inv}}^2 - 2s\sqrt{H_{EE;0}}} . \quad (4.10)$$

Both M_{inv} and M_{rec} (especially M_{inv}) demonstrate certain discrimination power in Fig. 18. In the left panel, the non-Higgs backgrounds tend to have a smaller M_{inv} , compared to the others, since $Z_\nu Z_{q_5}$ and $q_5 q_5(\gamma)$ are produced or mainly produced at Z pole. In the right panel, as is expected, the Higgs background events are accumulated near Z pole.

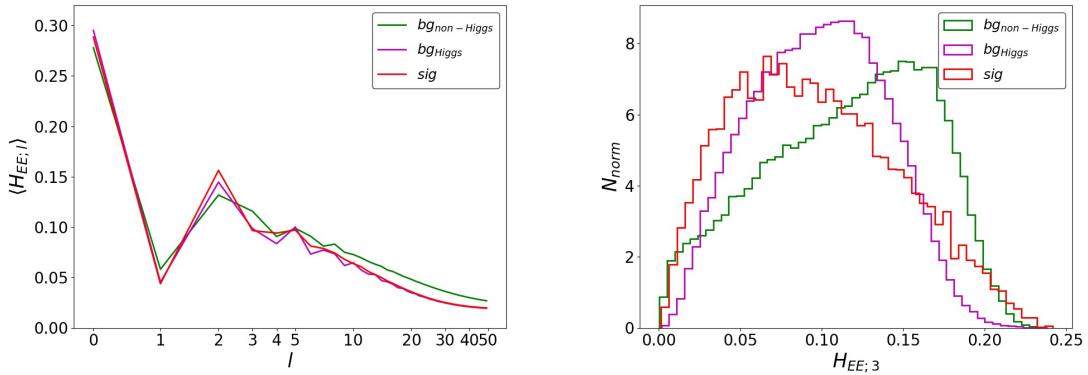


Figure 17. FW spectra of $\langle H_{EE;l} \rangle$ (left) and event distributions of $H_{EE;3}$ (right) for the $\nu\nu h_h$ and its background samples after preselection.

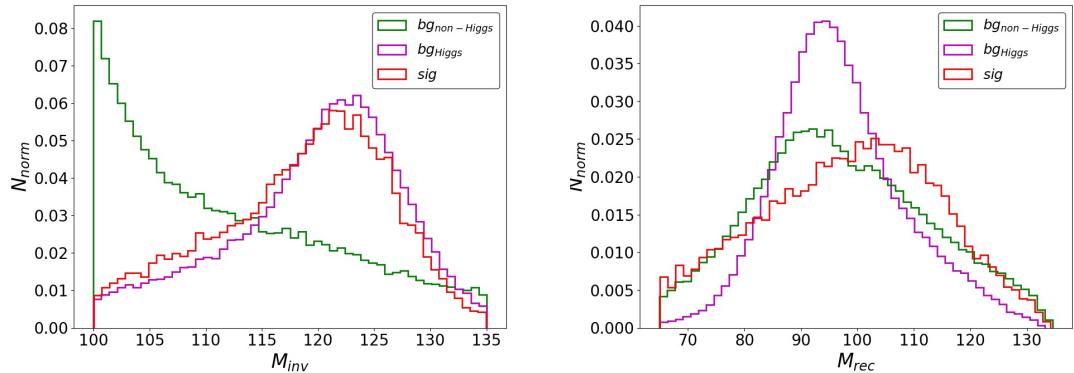


Figure 18. Event distributions of M_{inv} (left) and M_{rec} (right), defined in Eq. (4.10), for the $\nu\nu h_h$ and its background samples after preselection.

Beyond this, the $\langle H_{EE;l} \rangle$ spectra in Fig. 17 demonstrates a series of distinguishable fine structures. One example is related to the FW moments at $l = 2$ and 3 . The $\langle H_{EE;l} \rangle$ peaks at $l = 2$ indicates that the $\nu\nu h_h$ and its main backgrounds favor an included angle bigger than $\frac{\pi}{2}$ for their jet ancestral partons. But, the ordering of the relevant samples w.r.t. $\langle H_{EE;3} \rangle$ (see both panels in Fig. 17) implies that the signal prefers the widest such angle while the non-Higgs backgrounds the smallest one. The kinematic information carried by these fine structures fails to be picked up by the main observables used for the traditional cut-based analysis of measuring $\sigma(\nu\nu h_b)$ (see, e.g., [17]). These observables include M_{inv} , M_{rec} and the polar angle of Higgs boson¹⁰. The J2-, J3- and E1-, E2-based analyses of measuring $\sigma(\nu\nu h_h)$ thus may benefit a lot from such deformed or lost information at jet level and the overall information synergization.

¹⁰By definition the FW moments of $H_{EE;l}$ are not sensitive to event orientation in space, and hence are expected to be independent of the polar angle of Higgs boson. The information carried by the latter could be picked up by $H_{EE;l,m}$, i.e., the FW moments at order m , in this CMB-like observable scheme. The

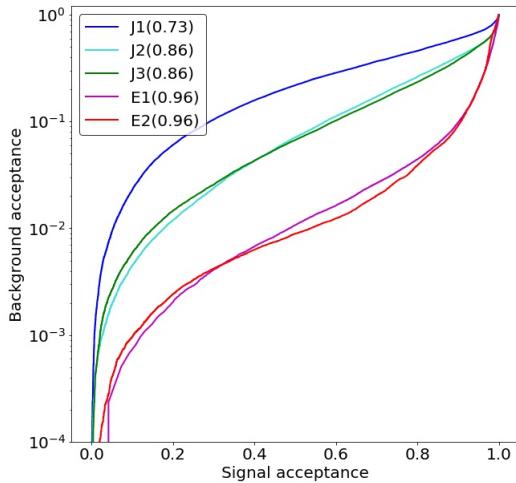


Figure 19. ROC curves and their AUC for the binary classifiers to distinguish the $\nu\nu h_h$ from its background events.

Fig. 19 displays the ROC curves and their AUC for the binary classifiers to distinguish the $\nu\nu h_h$ from its backgrounds (the event responses to these classifiers are shown in Fig. 21 in Appendix A). As is expected, E1 classifier demonstrates a better performance than that of J1. By including the FW moments $H_{EE;l \leq 50}$, J2 classifier yields a significant improvement to the AUC, *i.e.*, from 0.72 to 0.84. But, there is still an AUC gap between J2 and E1 classifiers. This gap could be filled up by the FW moments or/and multi-spectra which are not included in these jet-level analyses.

The precisions of measuring $\sigma(\nu\nu h_h)$ with these classifiers are summarized in Table 7. With the FW moments $H_{EE;l \leq 50}$, the jet-level precision is improved from 2.8% (J1) to 1.8% (J2) and 1.9% (J3). The best precisions of 1.4% and 1.3% are achieved with the two event-level classifiers, *i.e.*, E1 and E2, respectively. These results indicate that, compared to the gain in the signal rate by replacing $\sigma(\nu\nu h_b)$ (Method B) with $\sigma(\nu\nu h_h)$ (Method B'), this measurement benefits more from synergizing the event-level information into the DNN-based analysis. A significant improvement to the precision is thus expected if these classifiers are applied to the exclusive measurement of $\sigma(\nu\nu h_b)$.

4.4 Robustness against Detector Resolution

The precisions of measuring the SM Γ_h with Method B', by applying J1, J2, J3, E1 and E2 classifiers to the data of 5ab^{-1} @ 240GeV , are summarized in Table 7. J1 classifiers yield a precision of 3.2%. It is improved to 2.3% by J2 and 1.9% by E1, with the event-level information being incorporated. The track observables only have a slight impact for the measurements. The best outcome of 1.9% improves the baseline precisions with Method B, *i.e.*, 3.5% at both CEPC₂₄₀ and FCC₂₄₀, by a factor about 1.8.

exploration regarding this is beyond the scope of this paper.

Precision (%)	J1	J2	J3	E1	E2
$\sigma(Z_\nu h_{W_{lq}})$	1.7 (1.6)	1.4 (1.6)	1.5 (1.6)	1.5 (1.4)	1.5 (1.4)
$\sigma(Z_\nu h_{W_{qq}})$	1.6 (1.6)	1.2 (1.2)	1.1 (1.1)	1.1 (1.1)	1.1 (1.1)
$\sigma(\nu\nu h_h)$	2.8 (2.7)	1.8 (1.7)	1.9 (1.8)	1.4 (1.4)	1.3 (1.3)
Γ_h	$3.2^{+0.9}_{-0.3}$ (3.1)	$2.3^{+0.7}_{-0.2}$ (2.2)	$2.3^{+0.7}_{-0.2}$ (2.3)	$1.9^{+0.5}_{-0.1}$ (1.9)	$1.9^{+0.4}_{-0.1}$ (1.9)

Table 7. Expected precisions of measuring Γ_h with Method B', by applying J1, J2, J3, E1 and E2 classifiers to the data of 5ab^{-1} @240GeV. In these analyses, a precision of 0.5% for measuring $\sigma(Zh)$ and 0.3% for measuring $\sigma(Zh_h)$ are assumed. The numbers in the parentheses represent the performance of these classifiers on the data processed with the FCC-ee IDEA template. The superscripts and subscripts for the numbers in last row denote the maximal changes of precision due to the variance of ε from 1.0 to 2.0 and 0.1, respectively.

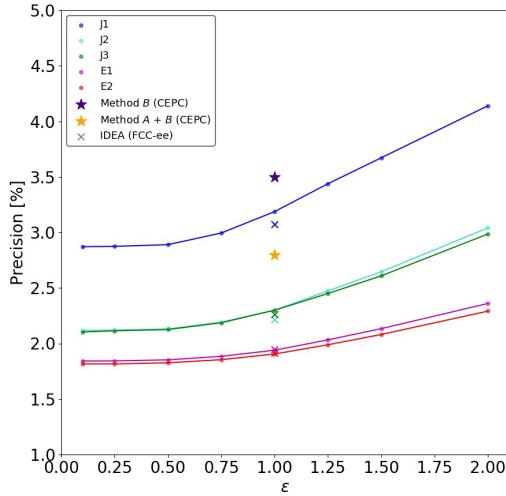


Figure 20. Expected precisions of measuring Γ_h versus detector energy/momentum resolution. $\varepsilon = 1.0$ corresponds to the baseline resolution defined by the CEPC detector template [66]. The purple and orange stars represent the precisions expected to be achieved with 5.6ab^{-1} data at CEPC₂₄₀, using Method B and A+B, respectively [17, 68]. The colored crosses denote the precisions achieved by applying J1, J2, J3, E1 and E2 classifiers to the FCC-ee IDEA data at 240GeV. For the convenience of comparison, we place them at $\varepsilon = 1.0$.

In these analyses, the detector effects are simulated with the built-in CEPC detector template [66] of DELPHES3 [67]. Such a specific choice naturally raises the question whether the classifiers developed are robust against the detector resolutions, including both energy/momentum and angular ones. To get some ideas about this, we take the following test. We first scale the energy/momentum resolution of track, ECAL, HCAL, electrons and muons defined in this template by a factor ε , then reprocess the data by varying its value from 0.1 to 2, and at last apply the classifiers developed at $\varepsilon = 1$ for their testing. The precisions of measuring Γ_h versus ε are presented in Table 7 and Fig. 20. As one can see, the event-level classifiers are slightly more robust against the variation of ε than the

jet-level ones. For J1, J2 and J3 classifiers, the precision is improved by $8.7\% \sim 9.4\%$ as ε decreases from its baseline value to 0.1, and degraded by $28\% \sim 30\%$ as ε increases to 2. As a comparison, the precision is improved by 5.3% and degraded by 21% \sim 26% in these two cases for E1 and E2 classifiers. But, no matter for which classifier, the conclusion reached on its performance has not been qualitatively changed. Additionally, in this study we incorporate the FW moments $H_{EE;l \leq 50}$ for J2 and J3 classifiers and image each event at a 50×50 evenly gridded $\theta - \phi$ plane in the E1- and E2-based analyses. Both setups are not fully synchronized with the baseline resolutions of the CEPC detector. The slightly worsening of the detector granularity or angular resolution thus could be absorbed by this uncertainty-tolerant space. A combination of these justifies the robustness of these classifiers against the detector resolutions. Beyond this, we test the robustness of these classifiers by simulating the testing data with the built-in FCC-ee IDEA template [61] of DELPHES3 [67]. This FCC-ee detector benchmark differentiates itself from the CEPC one mainly by their calorimeter resolutions. The IDEA has an ECAL resolution of $0.11\sqrt{E}$ and an HCAL resolution of $0.30\sqrt{E}$ at leading order, in comparison to $0.20\sqrt{E}$ and $0.60\sqrt{E}$ of the CEPC detector template, respectively. Also, the IDEA calorimeter granularity is higher than that of the CEPC HCAL and slightly lower than that of the CEPC ECAL. The remaining performance of these two detector templates are more or less comparable. The precisions achieved by applying J1, J2, J3, E1 and E2 classifiers to the FCC-ee IDEA data are presented in Table 7 and Fig. 20 also. Compared to the CEPC results, they are better, but by no more than 5%, for the Γ_h measurements.

It is noteworthy that these discussions never mean that better detector resolutions do not help much in improving the precision of measuring Γ_h . Recall, the detector granularity determines the highest multipoles of the FW moments which can be effectively applied to building J2 and J3 classifiers, and the largest pixels which can be legally used for constructing E1 and E2 ones. Also, the classifiers should be trained using the data processed at the new resolution benchmark. But, exploring this is beyond the scope of this study.

Despite these tests, the uncertainty arising from the modeling of parton hadronization could impact the performance of the constructed classifiers (especially E1 and E2). It is known that the shower evolution is not simulated well in some cases such as gluon splitting to heavy flavors, fragmentation functions as $z \rightarrow 1$, etc. But, similar to J1, J2 and J3, which rely on infrared- and collinear-safe observables, E1 and E2 classifiers are based on the images which by definition are infrared- and collinear-safe. The finite resolution and particle-identity irrelevance of these event images grant them some level of immunity to hadronization details. Additionally, the impact of this uncertainty for the image-based classifiers have been studied in the context of jet classification [23, 45], by analyzing the classifier performance with the data simulated by different event generators (Pythia, Herwig, Sherpa). One observation is that the light-quarks jets are less subject to this uncertainty compared with the gluon jets [23]. This might be a sign of small impact of this uncertainty for the Γ_h analysis, where mainly the quark jets are relevant. At last, we would bring it to the attention that the plenty of clean hadronic events produced at future e^-e^+ colliders may allow precisely matching data and simulation, and hence yield a suppression to this uncertainty.

5 Summary and Discussion

The e^-e^+ colliders, because of their clean QCD environment and absence of pileups, play a leading role in advancing the precision frontier in particle physics. One such machine of next generation is expected to push the precisions of measuring Higgs and electroweak physics up to an unprecedented level. Yet, due to the dominance of the hadronic events in data, many of the baseline precisions presented in documents are based on jet-level analysis and hence are limited by the information deformation and loss in jet clustering. We showed that this difficulty can be well-addressed by synergizing the event-level information into the DNN-based data analysis. In relation to this, we introduced a CMB-like observable scheme, where the event-level kinematics is encoded as the FW moments at leading order and multi-spectra at higher orders. Then we developed a series of jet-level (w/ and w/o the FW moments) and event-level binary classifiers, and analyzed their sensitivity performance comparatively with the two-jet and four-jet events. The general conclusion is: the event-level classifiers perform better compared to the jet-level ones; but, incorporating the FW moments into the jet-level classifiers can significantly reduce the performance gap between them. As an application of such classifiers, we analyzed the precision of measuring the SM Γ_h at e^-e^+ colliders with the data of 5ab^{-1} @ 240 GeV. The precisions obtained are significantly better than the baseline ones.

Yet, this is just an initial effort. We can immediately see several directions for future explorations. First of all, we showed that with the classifiers developed we are able to measure Γ_h with a precision of 1.9% (Method B', E1 and E2 classifiers), improving the baseline ones (Method B) by a factor about 1.8. One natural question is if this precision can be pushed to sub percent level in a similar collider operation scenario. After all, the precision of measuring the SM Γ_h is one of the most important indices to measure the physics capability of a future Higgs factory. This could be possible. As was discussed in Sec. 4, $\sigma(Zh_W)$ is one of the main intermediate quantities to determine the precision of measuring Γ_h in both Method B and B'. In the CEPC analysis, the decay modes of $Z \rightarrow qq, l^+l^-, \nu\nu$ are combined for the $\sigma(Zh_W)$ measurement [17]. But, we exclusively focused on the $Z \rightarrow \nu\nu$ mode in this study. Also, we have assumed the precisions obtained from the cut-based analyses for measuring $\sigma(Zh)$ and $\sigma(Zh_h)$. A more complete analysis is thus necessary and important. Beyond that, the SM Γ_h can be determined with four different methods at e^-e^+ colliders, with the set of six intermediate quantities in Eq. (4.9) being measured. One may consider combining these methods, to yield an overall precision. But, as a reminder, one needs to take into account the systematic errors properly, before a firm statement regarding this can be made.

Secondly, this effort opens a new angle to evaluate the physics capability of the future e^-e^+ colliders. We expect the developed strategies to be applied to many other hadronic measurements. Such measurements include Higgs couplings and CP properties, electroweak precision parameters, flavor physics, top physics, QCD parameters, etc. The applications can be even extended to new physics searches via high- \sqrt{s} e^-e^+ collisions. To fully evaluate the collider capability, one needs to pursue a comprehensive study on these aspects.

Thirdly, we expect that with the CMB-like observable scheme the kinematic informa-

tion lost at jet level can be systematically reconstructed. Here we tested only to what extent the FW moments of energy, as part of the leading-order CMB-like observables, can compensate for that. We have observed that the incorporation of these FW moments can greatly reduce the performance gap between the jet-level and event-level classifiers in a general context, but can not eliminate completely. It is thus interesting to explore if the existing gap can be filled by the FW moments not included in this study and the multi-spectra. Taking a step further, we can leave the jet information out, and study comparatively the classifier based on the CMB-like observables only and the one with the techniques of image recognition. This will allow us to test the (approximate) completeness of this CMB-like observable scheme, and dissect the underlying physics of the event-level kinematics.

Last but not least, although the CMB-like observable scheme was introduced for analyzing the data at e^-e^+ colliders, its application can be extended to, e.g., ep colliders, LHC and even future hadron colliders. But, the FW moments and the multi-spectra could be strongly smeared at such machines because of the four-momentum anisotropy of their events, contamination of pileups, etc. If these problems can be well addressed, we would expect the CMB-like observable scheme to be a powerful tool as well in these collider scenarios. We will leave this study and the others to a future work.

Acknowledgements

We would thank Michelangelo Mangano and Manqi Ruan greatly for informative communications on the FCC-ee and CEPC analyses and for valuable comments on this manuscript. We would also thank Spencer Chang, Shirley Ho, Xuhui Jiang, Gang Li and David Shih for useful discussions. This research was supported partly by the General Research Fund (GRF) under Grant No 16302117 and partly by the Area of Excellence under the Grant No AoE/P-404/18-3. Both grants were issued by the Research Grants Council of Hong Kong S.A.R. This manuscript has been authored by Fermi Research Alliance, LLC, under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

A Event Response to Γ_h Classifiers

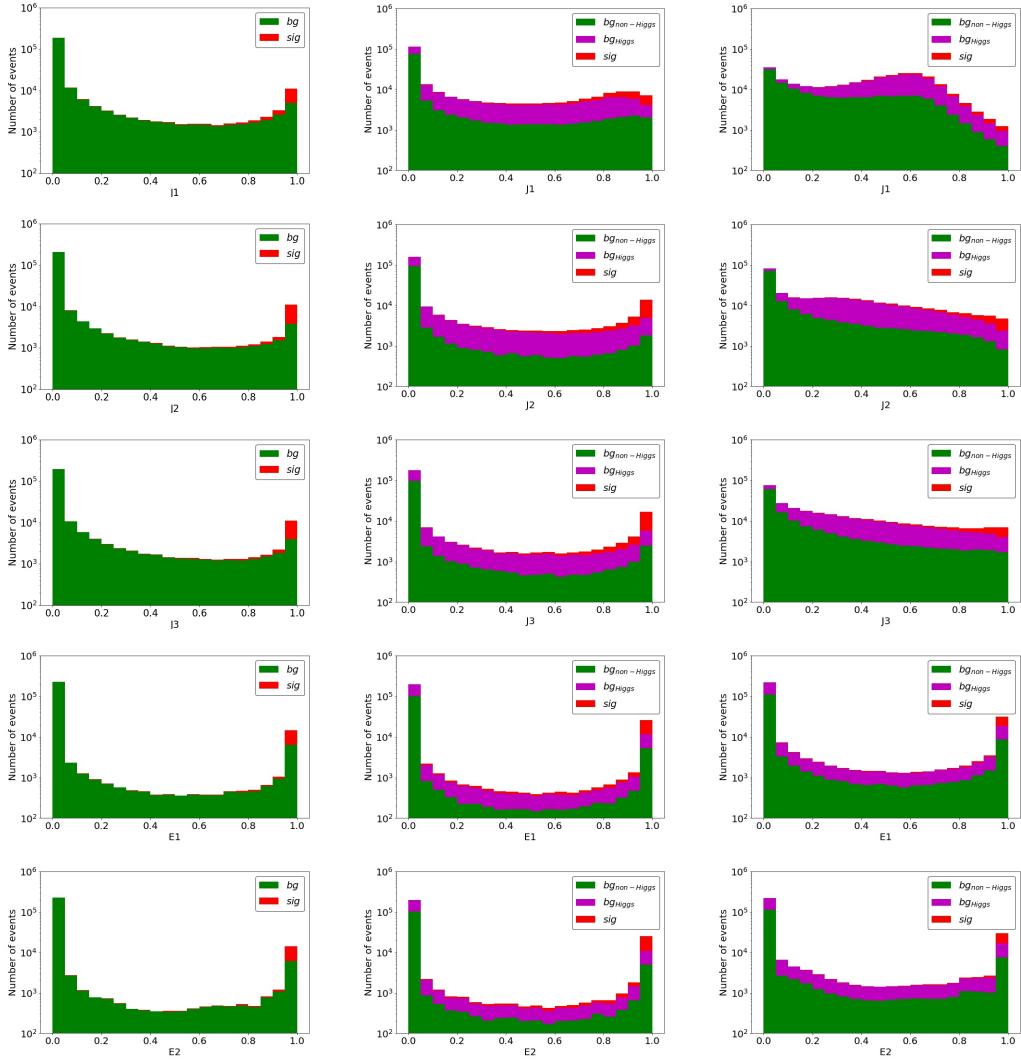


Figure 21. Responses of the $Z_\nu h_{W_{lq}}$ (left), $Z_\nu h_{W_{qq}}$ (middle) and $\nu\nu h_h$ (right) and their respective background events, to the binary classifiers.

References

- [1] G. Taylor, “A Perspective of the Future for HEP,” 2020. http://ias.ust.hk/program/shared_doc/2020/202001hep/conf/20200120_lt_Geoffrey_TAYLOR.pdf.
- [2] J. Fan, M. Reece, and L.-T. Wang, “Possible Futures of Electroweak Precision: ILC, FCC-ee, and CEPC,” *JHEP*, vol. 09, p. 196, 2015, 1411.1054.
- [3] A. Banfi, H. McAslan, P. F. Monni, and G. Zanderighi, “A general method for the resummation of event-shape distributions in e^+e^- annihilation,” *JHEP*, vol. 05, p. 102, 2015, 1412.2126.

- [4] D. d'Enterria, “Physics at the FCC-ee,” in *Proceedings, 17th Lomonosov Conference on Elementary Particle Physics: Moscow, Russia, August 20-26, 2015*, pp. 182–191, 2017, 1602.05043.
- [5] M. A. Fedderke, T. Lin, and L.-T. Wang, “Probing the fermionic Higgs portal at lepton colliders,” *JHEP*, vol. 04, p. 160, 2016, 1506.05465.
- [6] H. Khanpour and M. Mohammadi Najafabadi, “Constraining Higgs boson effective couplings at electron-positron colliders,” *Phys. Rev.*, vol. D95, no. 5, p. 055026, 2017, 1702.00951.
- [7] C. Cai, Z.-H. Yu, and H.-H. Zhang, “CEPC Precision of Electroweak Oblique Parameters and Weakly Interacting Dark Matter: the Fermionic Case,” *Nucl. Phys.*, vol. B921, pp. 181–210, 2017, 1611.02186.
- [8] W. H. Chiu, S. C. Leung, T. Liu, K.-F. Lyu, and L.-T. Wang, “Probing 6D operators at future ee⁺ colliders,” *JHEP*, vol. 05, p. 081, 2018, 1711.04046.
- [9] N. Chen, T. Han, S. Su, W. Su, and Y. Wu, “Type-II 2HDM under the Precision Measurements at the Z-pole and a Higgs Factory,” *JHEP*, vol. 03, p. 023, 2019, 1808.02037.
- [10] G. Durieux, C. Grojean, J. Gu, and K. Wang, “The leptonic future of the Higgs,” *JHEP*, vol. 09, p. 014, 2017, 1704.02333.
- [11] T. Barklow, K. Fujii, S. Jung, R. Karl, J. List, T. Ogawa, M. E. Peskin, and J. Tian, “Improved Formalism for Precision Higgs Coupling Fits,” *Phys. Rev.*, vol. D97, no. 5, p. 053003, 2018, 1708.08912.
- [12] S. Di Vita, G. Durieux, C. Grojean, J. Gu, Z. Liu, G. Panico, M. Riembau, and T. Vantalon, “A global view on the Higgs self-coupling at lepton colliders,” *JHEP*, vol. 02, p. 178, 2018, 1711.03978.
- [13] J. Gu, H. Li, Z. Liu, S. Su, and W. Su, “Learning from Higgs Physics at Future Higgs Factories,” *JHEP*, vol. 12, p. 153, 2017, 1709.06103.
- [14] S.-F. Ge, H.-J. He, and R.-Q. Xiao, “Probing new physics scales from Higgs and electroweak observables at e⁺ e⁻ Higgs factory,” *JHEP*, vol. 10, p. 007, 2016, 1603.03385.
- [15] S.-F. Ge, H.-J. He, and R.-Q. Xiao, “Testing Higgs coupling precision and new physics scales at lepton colliders,” pp. 55–69, 2017, 1612.02718.
- [16] J. Ellis, S.-F. Ge, H.-J. He, and R.-Q. Xiao, “Probing the Scale of New Physics in the ZZ γ Coupling at e⁺e⁻ Colliders,” *Chin. Phys. C*, vol. 44, p. 063106, 2020, 1902.06631.
- [17] F. An *et al.*, “Precision Higgs Physics at CEPC,” *Chin. Phys.*, vol. C43, no. 4, p. 043002, 2019, 1810.09037.
- [18] A. Abada *et al.*, “FCC Physics Opportunities,” *Eur. Phys. J.*, vol. C79, no. 6, p. 474, 2019.
- [19] Y. Zhu and M. Ruan, “Performance study of the separation of the full hadronic WW and ZZ events at the CEPC,” 2018, 1812.09478.
- [20] G. Li, Z. Li, Y. Wang, and Y. Wang, “Improving the measurement of Higgs boson-gluon coupling using Convolutional Neural Networks at e⁺e⁻ colliders,” 2019, 1901.09391.
- [21] M. Andrews, J. Alison, S. An, P. Bryant, B. Burkle, S. Gleyzer, M. Narain, M. Paulini, B. Poczos, and E. Usai, “End-to-End Jet Classification of Quarks and Gluons with the CMS Open Data,” 2019, 1902.08276.
- [22] G. Kasieczka, N. Kiefer, T. Plehn, and J. M. Thompson, “Quark-Gluon Tagging: Machine Learning vs Detector,” *SciPost Phys.*, vol. 6, p. 069, 2019, 1812.09223.

- [23] P. T. Komiske, E. M. Metodiev, and M. D. Schwartz, “Deep learning in color: towards automated quark/gluon jet discrimination,” *JHEP*, vol. 01, p. 110, 2017, 1612.01551.
- [24] T. Cheng, “Recursive Neural Networks in Quark/Gluon Tagging,” *Comput. Softw. Big Sci.*, vol. 2, no. 1, p. 3, 2018, 1711.02633.
- [25] P. T. Komiske, E. M. Metodiev, and J. Thaler, “Energy Flow Networks: Deep Sets for Particle Jets,” 2018, 1810.05165.
- [26] J. Gallicchio and M. D. Schwartz, “Seeing in Color: Jet Superstructure,” *Phys. Rev. Lett.*, vol. 105, p. 022001, 2010, 1001.5027.
- [27] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *JHEP*, vol. 04, p. 063, 2008, 0802.1189.
- [28] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, “New clustering algorithm for multi - jet cross-sections in e+ e- annihilation,” *Phys. Lett.*, vol. B269, pp. 432–438, 1991.
- [29] S. D. Ellis, Z. Kunszt, and D. E. Soper, “Jets at hadron colliders at order $\alpha - s^3$: A Look inside,” *Phys. Rev. Lett.*, vol. 69, pp. 3615–3618, 1992, hep-ph/9208249.
- [30] E. Farhi, “A QCD Test for Jets,” *Phys. Rev. Lett.*, vol. 39, pp. 1587–1588, 1977.
- [31] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, “N-Jettiness: An Inclusive Event Shape to Veto Jets,” *Phys. Rev. Lett.*, vol. 105, p. 092002, 2010, 1004.2489.
- [32] G. Parisi, “Superinclusive cross sections,” *Physics Letters B*, vol. 74, no. 1, pp. 65 – 67, 1978.
- [33] C. F. Berger, T. Kucs, and G. F. Sterman, “Event shape / energy flow correlations,” *Phys. Rev.*, vol. D68, p. 014012, 2003, hep-ph/0303051.
- [34] V. Mateu, I. W. Stewart, and J. Thaler, “Power Corrections to Event Shapes with Mass-Dependent Operators,” *Phys. Rev.*, vol. D87, no. 1, p. 014025, 2013, 1209.3781.
- [35] S. Catani, G. Turnock, and B. R. Webber, “Jet broadening measures in e^+e^- annihilation,” *Phys. Lett.*, vol. B295, pp. 269–276, 1992.
- [36] M. Dasgupta and G. P. Salam, “Event shapes in e+ e- annihilation and deep inelastic scattering,” *J. Phys.*, vol. G30, p. R143, 2004, hep-ph/0312283.
- [37] A. Banfi, G. P. Salam, and G. Zanderighi, “Phenomenology of event shapes at hadron colliders,” *JHEP*, vol. 06, p. 038, 2010, 1001.4082.
- [38] C. Cesarotti and J. Thaler, “A Robust Measure of Event Isotropy at Colliders,” 2020, 2004.06125.
- [39] G. C. Fox and S. Wolfram, “Observables for the Analysis of Event Shapes in e+ e- Annihilation and Other Processes,” *Phys. Rev. Lett.*, vol. 41, p. 1581, 1978.
- [40] D. Krohn, M. D. Schwartz, T. Lin, and W. J. Waalewijn, “Jet Charge at the LHC,” *Phys. Rev. Lett.*, vol. 110, no. 21, p. 212001, 2013, 1209.2421.
- [41] M. Andrews, M. Paulini, S. Gleyzer, and B. Poczos, “End-to-End Event Classification of High-Energy Physics Data,” *J. Phys. Conf. Ser.*, vol. 1085, no. 4, p. 042022, 2018.
- [42] M. Andrews, M. Paulini, S. Gleyzer, and B. Poczos, “End-to-End Physics Event Classification with the CMS Open Data: Applying Image-based Deep Learning on Detector Data to Directly Classify Collision Events at the LHC,” 2018, 1807.11916.

- [43] P. T. Komiske, R. Mastandrea, E. M. Metodiev, P. Naik, and J. Thaler, “Exploring the Space of Jets with CMS Open Data,” 2019, 1908.08542.
- [44] J. W. Monk, “Deep Learning as a Parton Shower,” *JHEP*, vol. 12, p. 021, 2018, 1807.03685.
- [45] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, “Jet-images deep learning edition,” *JHEP*, vol. 07, p. 069, 2016, 1511.05190.
- [46] G. Kasieczka, T. Plehn, M. Russell, and T. Schell, “Deep-learning Top Taggers or The End of QCD?,” *JHEP*, vol. 05, p. 006, 2017, 1701.08784.
- [47] J. Lin, M. Freytsis, I. Moult, and B. Nachman, “Boosting $H \rightarrow b\bar{b}$ with Machine Learning,” *JHEP*, vol. 10, p. 101, 2018, 1807.10768.
- [48] 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.*, vol. 122, no. 9, p. 091801, 2019, 1807.11498.
- [49] J. H. Kim, M. Kim, K. Kong, K. T. Matchev, and M. Park, “Portraying Double Higgs at the Large Hadron Collider,” *JHEP*, vol. 09, p. 047, 2019, 1904.08549.
- [50] P. T. Komiske, E. M. Metodiev, B. Nachman, and M. D. Schwartz, “Pileup Mitigation with Machine Learning (PUMML),” *JHEP*, vol. 12, p. 051, 2017, 1707.08600.
- [51] A. Andreassen, I. Feige, C. Frye, and M. D. Schwartz, “JUNIPR: a Framework for Unsupervised Machine Learning in Particle Physics,” *Eur. Phys. J.*, vol. C79, no. 2, p. 102, 2019, 1804.09720.
- [52] J. Ren, L. Wu, and J. M. Yang, “Unveiling CP property of top-Higgs coupling with graph neural networks at the LHC,” 2019, 1901.05627.
- [53] J. Arjona Martnez, O. Cerri, M. Pierini, M. Spiropulu, and J.-R. Vlimant, “Pileup mitigation at the Large Hadron Collider with graph neural networks,” *Eur. Phys. J. Plus*, vol. 134, no. 7, p. 333, 2019, 1810.07988.
- [54] S. Farrell *et al.*, “Novel deep learning methods for track reconstruction,” in *4th International Workshop Connecting The Dots 2018 (CTD2018) Seattle, Washington, USA, March 20-22, 2018*, 2018, 1810.06111.
- [55] M. Abdughani, J. Ren, L. Wu, and J. M. Yang, “Probing stop pair production at the LHC with graph neural networks,” *JHEP*, vol. 08, p. 055, 2019, 1807.09088.
- [56] H. Qu and L. Gouskos, “ParticleNet: Jet Tagging via Particle Clouds,” 2019, 1902.08570.
- [57] J. Zhou, G. Cui, Z. Zhang, C. Yang, Z. Liu, and M. Sun, “Graph neural networks: A review of methods and applications,” *CoRR*, vol. abs/1812.08434, 2018, 1812.08434.
- [58] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, “Pytorch: An imperative style, high-performance deep learning library,” in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d Alché-Buc, E. Fox, and R. Garnett, eds.), pp. 8024–8035, Curran Associates, Inc., 2019.
- [59] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *CoRR*, vol. abs/1512.03385, 2015, 1512.03385.
- [60] C. Chen, X. Mo, M. Selvaggi, Q. Li, G. Li, M. Ruan, and X. Lou, “Fast simulation of the CEPC detector with Delphes,” 2017, 1712.09517.
- [61] M. A. Elisa Fontanesi, Lorenzo Pezzotti, “FCC-ee IDEA detector model for Delphes,” 2019.

- [62] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings* (Y. Bengio and Y. LeCun, eds.), 2015.
- [63] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, “MadGraph 5 : Going Beyond,” *JHEP*, vol. 06, p. 128, 2011, 1106.0522.
- [64] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1,” *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008, 0710.3820.
- [65] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual,” *Eur. Phys. J.*, vol. C72, p. 1896, 2012, 1111.6097.
- [66] M. Ruan *et al.*, “Reconstruction of physics objects at the Circular Electron Positron Collider with Arbor,” *Eur. Phys. J.*, vol. C78, no. 5, p. 426, 2018, 1806.04879.
- [67] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lematre, A. Mertens, and M. Selvaggi, “DELPHES 3, A modular framework for fast simulation of a generic collider experiment,” *JHEP*, vol. 02, p. 057, 2014, 1307.6346.
- [68] M. Dong and G. Li, “CEPC Conceptual Design Report: Volume 2 - Physics & Detector,” 2018, 1811.10545.
- [69] A. Abada *et al.*, “FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2,” *Eur. Phys. J. ST*, vol. 228, no. 2, pp. 261–623, 2019.
- [70] D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban, and D. Whiteson, “Jet Flavor Classification in High-Energy Physics with Deep Neural Networks,” *Phys. Rev.*, vol. D94, no. 11, p. 112002, 2016, 1607.08633.
- [71] C. Drig, K. Fujii, J. List, and J. Tian, “Model Independent Determination of HWW coupling and Higgs total width at ILC,” in *International Workshop on Future Linear Colliders (LCWS13) Tokyo, Japan, November 11-15, 2013*, 2014, 1403.7734.
- [72] H. Abramowicz *et al.*, “Higgs physics at the CLIC electronpositron linear collider,” *Eur. Phys. J.*, vol. C77, no. 7, p. 475, 2017, 1608.07538.
- [73] H. Ono and A. Miyamoto, “A study of measurement precision of the Higgs boson branching ratios at the International Linear Collider,” *Eur. Phys. J.*, vol. C73, no. 3, p. 2343, 2013, 1207.0300.
- [74] J. Tian and K. Fujii, “Measurement of higgs boson couplings at the international linear collider,” *Nuclear and Particle Physics Proceedings*, vol. 273-275, pp. 826 – 833, 2016. 37th International Conference on High Energy Physics (ICHEP).
- [75] H. Li, K. Ito, R. Poschl, F. Richard, M. Ruan, Y. Takubo, and H. Yamamoto, “HZ Recoil Mass and Cross Section Analysis in ILD,” 2012, 1202.1439.
- [76] K. Fujii *et al.*, “Physics Case for the 250 GeV Stage of the International Linear Collider,” 2017, 1710.07621.