HWW and HZZ in the fully hadronic final states

Eysermans Jan [1*], Selvaggi Michele [2], Mila Pandurovic [3], Aman Desai [4]

[1] Particle Physics Collaboration, Massachusetts Institute of Technology, 77 Massachusetts Ave, Boston, Cambridge, 02139, MA, USA.
[2] CERN, 1211 Route de Meyrin, Geneva, 1211, Switzerland.
[3] Vinca Institute of Nuclear Sciences, Mike Petrovica Alasa 12-14, Belgrade, 11090, Serbia.
[4]The University of Adelaide, Adelaide, Australia.

*Corresponding author(s). E-mail(s): jan.eysermans@cern.ch;
Contributing authors: michele.selvaggi@cern.ch; milap@vinca.rs; aman.desai@cern.ch;

Abstract

The precise determination of the Higgs boson couplings to electroweak gauge bosons is a key objective of the FCC-ee physics programme. This note presents an analysis of Higgs decays to WW* and ZZ* in the ZH production mode, focusing on the fully hadronic final state, which offers the highest branching fractions and, thus, competitive precision compared to channels involving leptons or neutrinos. The analysis strategy enables the simultaneous extraction of the HWW and HZZ branching ratios by employing dedicated jet clustering and pairing algorithms to maximize the signal significance. The resulting expected uncertainties are 1.48% for HWW and 8.20% for HZZ.

1 Introduction

Measurements of the fundamental properties of the Standard Model (SM) Higgs boson—such as its spin, mass, decay width, and couplings—are a central objective at any future electron-positron collider, including the FCC-ee [1]. Among these, precise measurements of the Higgs boson couplings are particularly important, as the SM provides accurate predictions for the strengths of its interactions with other SM particles. Any deviation from these predictions could indicate the presence of new physics.

At the nominal center-of-mass energy of the FCC-ee, $\sqrt{s} = 240$ GeV, the dominant Higgs production mode is Higgsstrahlung, $\mathrm{e^+e^-} \to \mathrm{HZ}$. This process enables a model-independent determination of the absolute coupling of the Higgs boson to the Z boson, g_{HZZ}, via a measurement of the total production cross section using the recoil mass method [2][3]. This technique relies solely on the reconstruction of the Z boson decay products, allowing inference of the Higgs boson without requiring its direct reconstruction.

In addition, the total Higgs decay width, predicted by the SM to be approximately 4.1 MeV, can also be accessed at the FCC-ee. It is extracted using measurements of the branching fraction for the decay $\mathrm{H} \to \mathrm{ZZ}^*$ in Higgsstrahlung events:

$$\sigma(\mathrm{e^+e^-} \to \mathrm{HZ}) \times \mathcal{B}(\mathrm{H} \to \mathrm{ZZ}^*) \propto \frac{\mathrm{g}^4_{\mathrm{HZZ}}}{\Gamma_{\mathrm{H}}}.$$

Once both g_{HZZ} and Γ_H are determined, all accessible absolute Higgs couplings can be extracted in a model-independent way. Similarly, measuring the decay $\mathrm{H} \to \mathrm{WW}^*$ provides access to the Higgs coupling to W bosons and contributes to the determination of the total width, especially when combined with measurements of $\sigma(\mathrm{HZ})$ and $\sigma \times \mathcal{B}(\mathrm{H} \to \mathrm{b\bar{b}})$.

In this work, we study the sensitivity to the $\mathrm{H} \to \mathrm{WW}^*$ and $\mathrm{H} \to \mathrm{ZZ}^*$ decays in the fully hadronic final state, which yields six jets in the detector. This final state is of particular interest due to its large branching fraction, with $\mathcal{B}(Z \to jj) \approx 0.7$ and $\mathcal{B}(W \to jj) \approx 0.67$, significantly higher than leptonic or invisible channels. However, the analysis is challenging due to the presence of large backgrounds from diboson production and the complexity of the multi-jet environment. Furthermore, the $\mathrm{H} \to \mathrm{WW}^*$ and $\mathrm{H} \to \mathrm{ZZ}^*$ processes constitute irreducible backgrounds to each other, making their separation a central aspect of the analysis strategy.

2 Monte Carlo samples

Monte Carlo samples for various physics processes are generated using multiple event generators. Detector simulation and reconstruction are handled by the fast simulation framework Delphes [4], which applies smearing and efficiency models to generator-level particles to emulate detector response. The full chain—generation, simulation, and reconstruction—is integrated within the Key4Hep framework. This analysis uses centrally produced samples from the Winter2023 campaign.

The Innovative Detector for an Electron-positron Accelerator (IDEA) is used as the default detector model. Samples were generated at center-of-mass energies of $\sqrt{s} = 240$ and 365 GeV using either Pythia [5] or Whizard [6] (interfaced with Pythia8 [7]) for showering, hadronization, and fragmentation), assuming integrated luminosities of 10.8 and 3.0 ab^{-1}, respectively, corresponding to the FCC-ee layout with four interaction points. An overview of the samples used in this analysis and the cross sections at both center-of-mass energies are listed in Table 1.

Table 1 Signal and background samples and the corresponding cross sections in pb at $\sqrt{s} = 240$ and 365 GeV.

Process	$\sqrt{s} = 240$ GeV	$\sqrt{s} = 365$ GeV
`wzp6_ee_(qq,cc,ss,bb)H_HZZ`	0.00357	0.00221
`wzp6_ee_(qq,cc,ss,bb)H_HWW`	0.02914	0.01803
`p8_ee_WW`	16.44	10.72
`p8_ee_ZZ`	1.36	0.64
`wz3p6_ee_(qq,cc,ss,bb)`	55.01	22.80

3 Analysis strategy

A key challenge in this analysis is the separation of the HWW and HZZ decay modes, which is central to the strategy. Due to the relatively poor hadronic resolution and the similar masses of the W and Z bosons, the reconstructed final states from these processes significantly overlap. This overlap especially impacts the HZZ measurement, whose branching ratio is approximately eight times smaller than that of HWW, making it more susceptible to contamination. As a result, the ability to disentangle the two decay modes is crucial for a precise measurement of both processes.

The separation exploits kinematic differences between W and Z bosons and is further enhanced by incorporating flavor information, as discussed in Section 5. More advanced techniques—such as kinematic fitting—will be investigated at a later stage to further improve the hadronic resolution.

After a generic preselection (see Section 4), events are split into HWW-like and HZZ-like categories (see Section 5). Within each category, background suppression is performed using a multivariate analysis that combines all relevant kinematic observables (Section 6 for HZZ and 7 for HWW). The final results are extracted through a simultaneous fit to the output discriminants of both categories, allowing a coherent estimation of the HWW and HZZ signal strengths and their uncertainties (see Section 8).

4 Preselection and event reconstruction

Events are preselected to be pure in hadronic activity by imposing a cut on the missing momentum to be less than 20 GeV and a visible mass above 10 GeV, and an additional cut on the visible energy above 115 GeV (see Fig. 1). Events containing electrons or muons above 10 GeV are vetoed to be orthogonal with other analyses.

The remaining events are clustered into 6 jets using the exclusive Durham k_T algorithm, implemented in FastJet [8] and interfaced to the analysis software. Events that fail to cluster into 6 jets are removed from the event selection. The jet algorithm also returns jet merging variables $\mathrm{d_{ij}}$, which represent the distance parameters when merging $\mathrm{n}+1$ to n jets, and is found useful to reject backgrounds which typically have a lower jet multiplicity. The following cuts are applied: $\sqrt{\mathrm{d}_{23}} > 50$, $\sqrt{\mathrm{d}_{34}} > 25$, $\sqrt{\mathrm{d}_{45}} > 15$, $\sqrt{\mathrm{d}_{56}} > 10$, and are visually shown in Fig. 2.

A Graph Neural Network (GNN) jet flavour tagger [9] is employed to assign each jet a probability of originating from one of seven particle types: gluon (g), up (u), down (d), strange (s), charm (c), bottom (b), or tau (τ). The tagger leverages detailed information from all particle-flow candidates within each jet, incorporating kinematic properties, impact parameters, and particle identification features such as cluster counting and time-of-flight measurements.

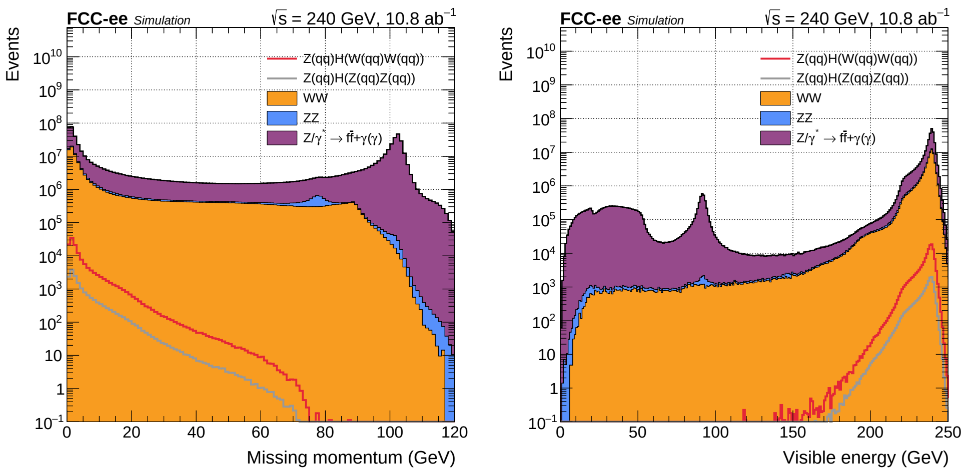

Fig. 1 Missing momentum (left) and visible energy (right).

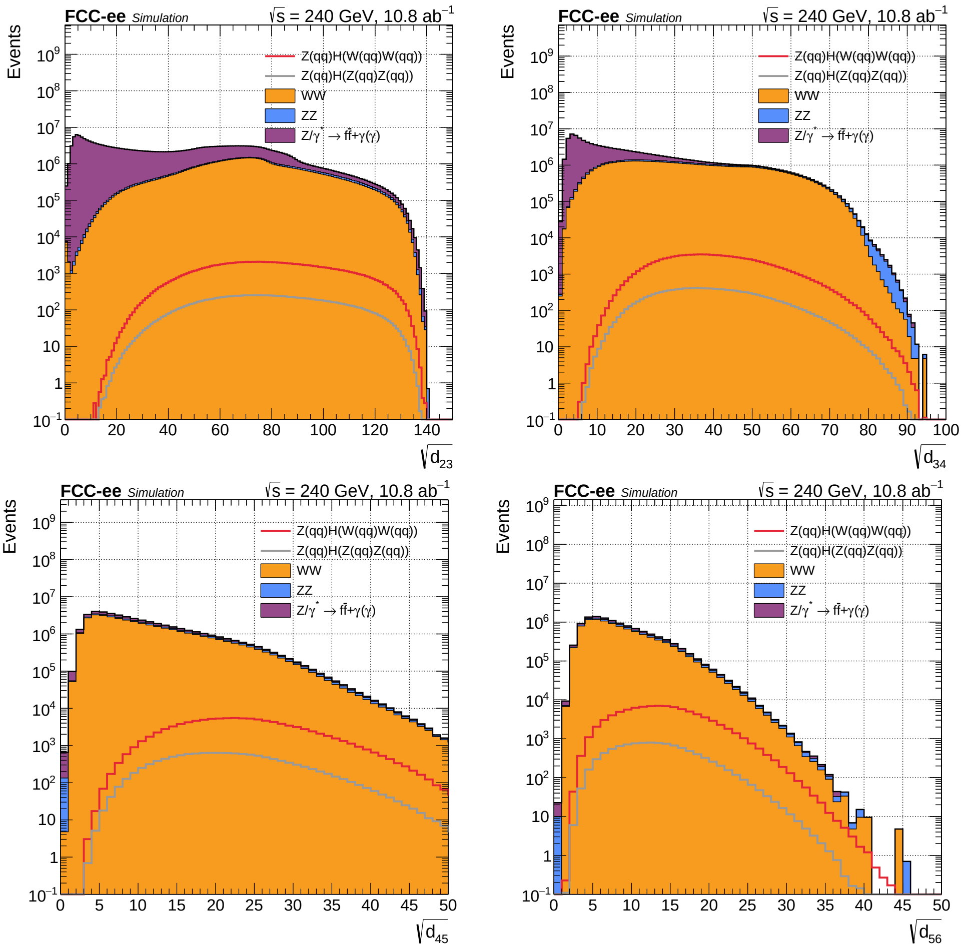

Fig. 2 Jet merging variables $\sqrt{\mathrm{d_{ij}}}$.

The event cutflow and the corresponding yields for signal and background processes are presented in Fig. 3 and Table 2. The final (integrated) significances are found to be 36.6 for the HWW signal and 1.02 for the HZZ signal.

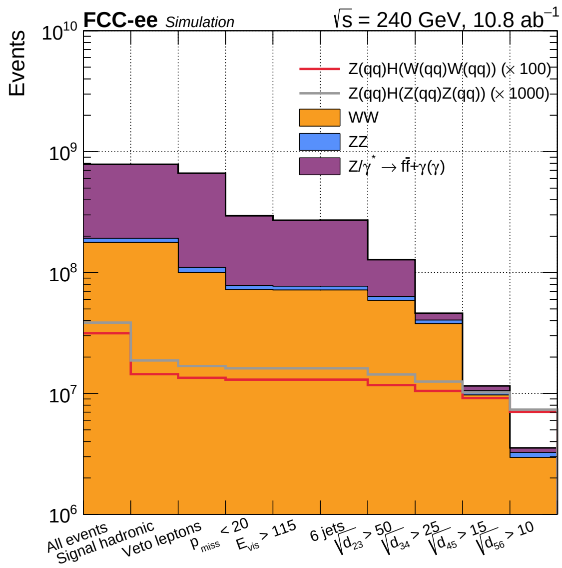

Fig. 3 Cutflow plot with the various selections and cuts.

Table 2 Signal and background yields.

Cut	HWW	HZZ	WW	ZZ	Z/γ
Cut 0	$3.1472 \cdot 10^5$	$5.7484 \cdot 10^4$	$1.7754 \cdot 10^8$	$1.4677 \cdot 10^7$	$5.9411 \cdot 10^8$
Cut 1	$1.4433 \cdot 10^5$	$2.7938 \cdot 10^4$	$1.7754 \cdot 10^8$	$1.4677 \cdot 10^7$	$5.9411 \cdot 10^8$
Cut 2	$1.3479 \cdot 10^5$	$2.3392 \cdot 10^4$	$9.9999 \cdot 10^7$	$1.0502 \cdot 10^7$	$5.5318 \cdot 10^8$
Cut 3	$1.3012 \cdot 10^5$	$1.6398 \cdot 10^4$	$7.1959 \cdot 10^7$	$5.8892 \cdot 10^6$	$2.1719 \cdot 10^8$
Cut 4	$1.3012 \cdot 10^5$	$1.6398 \cdot 10^4$	$7.1782 \cdot 10^7$	$5.4183 \cdot 10^6$	$2.0760 \cdot 10^8$
Cut 5	$1.3012 \cdot 10^5$	$1.6389 \cdot 10^4$	$7.1692 \cdot 10^7$	$5.3945 \cdot 10^6$	$1.9364 \cdot 10^8$
Cut 6	$1.3012 \cdot 10^5$	$1.6389 \cdot 10^4$	$7.1692 \cdot 10^7$	$5.3945 \cdot 10^6$	$1.9364 \cdot 10^8$
Cut 7	$1.1723 \cdot 10^5$	$1.4520 \cdot 10^4$	$5.8936 \cdot 10^7$	$4.3770 \cdot 10^6$	$6.4632 \cdot 10^7$
Cut 8	$1.0495 \cdot 10^5$	$1.2661 \cdot 10^4$	$3.7702 \cdot 10^7$	$2.7232 \cdot 10^6$	$5.6496 \cdot 10^6$
Cut 9	$9.1575 \cdot 10^4$	$1.0287 \cdot 10^4$	$9.6884 \cdot 10^6$	$8.3777 \cdot 10^5$	$1.0195 \cdot 10^6$
Cut 10	$7.0525 \cdot 10^4$	$7.3995 \cdot 10^3$	$2.9556 \cdot 10^6$	$2.9476 \cdot 10^5$	$3.0311 \cdot 10^5$

5 Separation of HWW and HZZ

The six jets in the event are grouped into three pairs to reconstruct the Higgs boson candidate, one on-shell and one off-shell Z or W boson, and the Z boson originating from the associated ZH production. The optimal jet pairing is determined by minimizing the following χ^2, evaluated on an event-by-event basis:

$$\chi^2 = (\mathrm{m_{ij}} - \mathrm{m_V})^2 + (\mathrm{m_{kl}} - \mathrm{m_Z})^2 + (\mathrm{m_{ijmn}} - \mathrm{m_H})^2, \quad \mathrm{i, j, k, l, m, n} = 1...6. \tag{1}$$

For each selected event, the χ^2 is evaluated under both the HWW and HZZ hypotheses by setting the corresponding boson masses in the χ^2 definition, resulting in two sets of reconstructed mass distributions. To assess the performance of the jet pairing, Fig. 4 shows the reconstructed boson masses for both HWW and HZZ signal samples. In the HWW case, the pairing successfully reconstructs all three boson masses. The performance for HZZ is comparatively worse, primarily because the on-shell Z boson tends to be reconstructed with lower invariant masses. Further optimization—potentially incorporating jet flavor information—may improve the jet assignment accuracy for the HZZ hypothesis.

To separate the HWW and HZZ processes, a Boosted Decision Tree (BDT) is trained using HWW as signal and HZZ as background. The input features include kinematic variables and invariant masses from both jet pairings, as well as jet flavor information after pairing. The BDT performance is illustrated in Fig. 5, showing the ROC curve and output discriminator, while the input variable importance is presented in Fig. 6. A separation power of 77% is achieved, with the most discriminating variables being the invariant masses of the boson candidates and the jet flavor scores associated with the HWW hypothesis.

Based on the BDT output, events are classified as HZZ-like (MVA score < 0.5) or HWW-like (MVA score > 0.5). In each of these categories, dedicated analysis optimizations are performed, targeting the corresponding signal, as detailed in Sections 6 and 7, respectively.

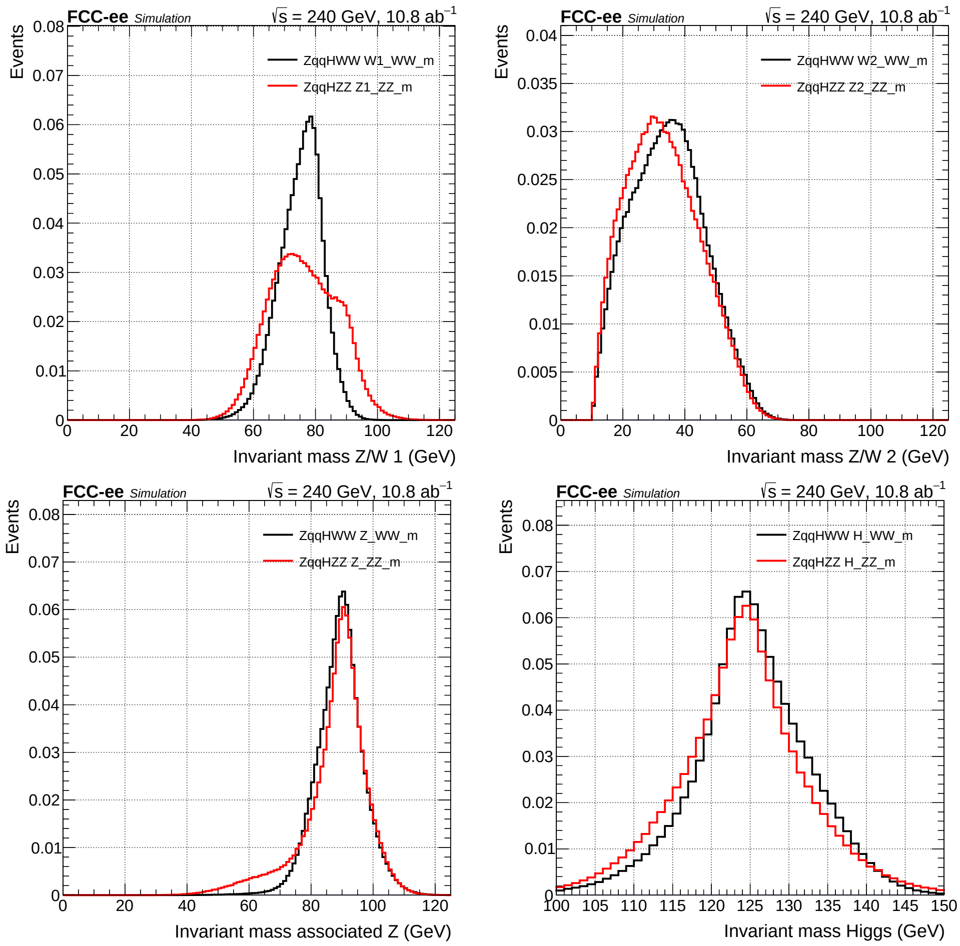

Fig. 4 Outcome of the jet pairing χ^2 for HWW (black curves) and HZZ (red curves): on-shell boson mass (top left), off-shell boson mass (top right), associated production Z mass (bottom left) and reconstructed Higgs mass (bottom right).

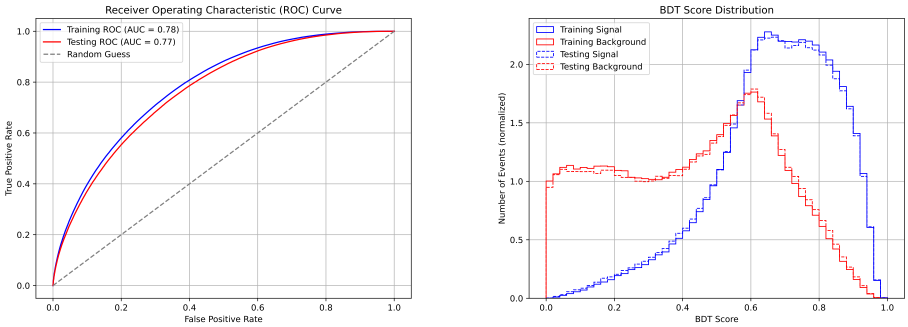

Fig. 5 ROC curve (left) and BDT output discriminator (right). In the discriminator plot, the red curve corresponds to the HZZ process, while the blue curve corresponds to the HWW process.

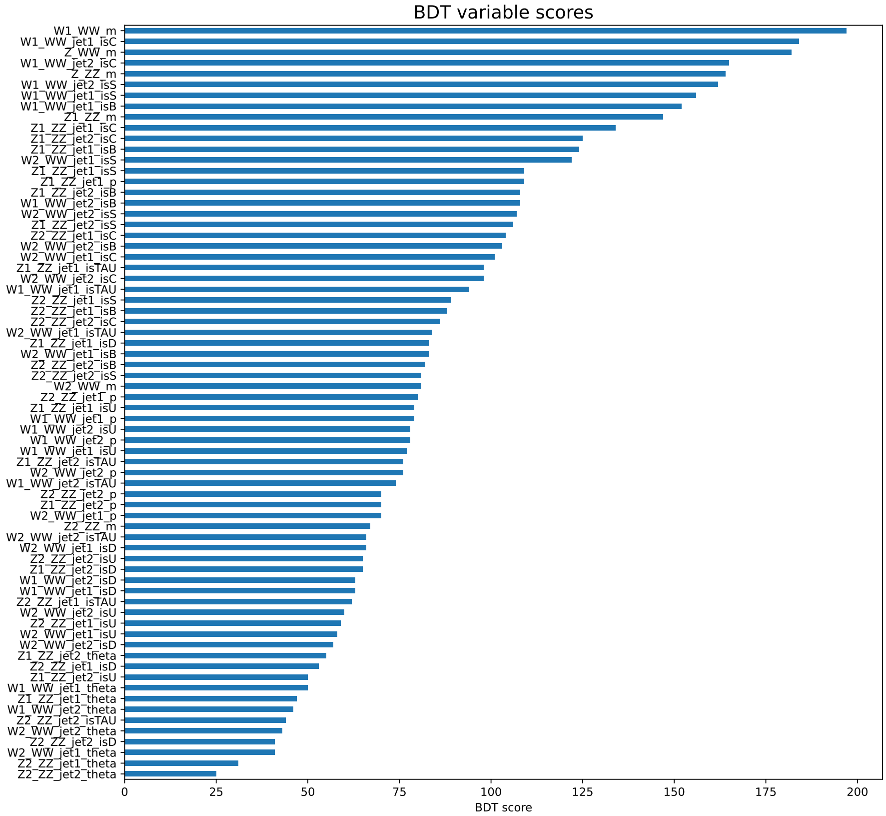

Fig. 6 Ranking of the variables used in the BDT.

6 Analysis optimization for HZZ

The HZZ category is defined by selecting events with a low MVA discriminant score from the HWW versus HZZ separation, as described in the previous section. Further optimization within this category is performed using a dedicated BDT trained (see Figs. 78 to distinguish HZZ signal events from dominant backgrounds, namely WW and ZZ. The input features include all kinematic variables from the HZZ jet pairing procedure, jet kinematics, angular variables, and jet flavor information.

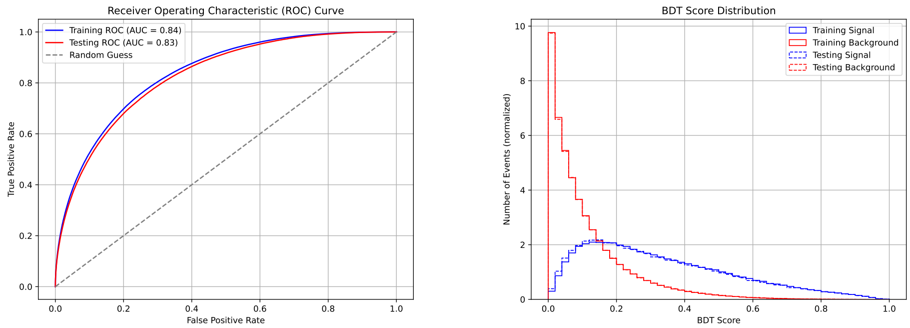

Fig. 7 ROC curve (left) and BDT output discriminator (right) for the HZZ analysis.

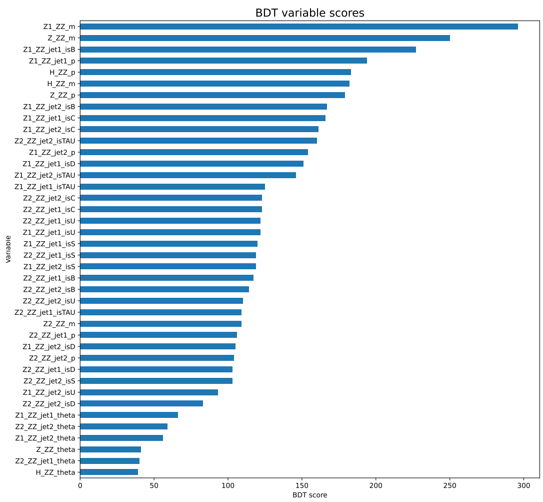

Fig. 8 Ranking of the variables used in the BDT for the HZZ analysis.

7 Analysis optimization for HWW

The HWW analysis, defined by selecting events with a high MVA discriminant score from the HWW versus HZZ separation (as described in the previous section), is optimized similarly to the HZZ analysis. In this case, the input variables are those derived from the HWW jet pairing procedure. The results of the optimization are presented in Figs. 9 and 10.

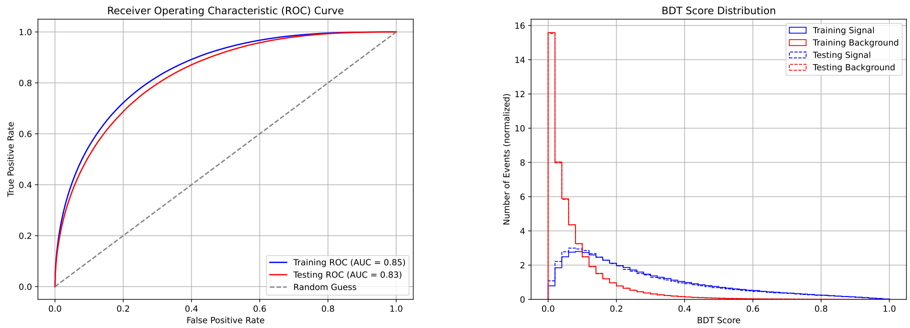

Fig. 9 ROC curve (left) and BDT output discriminator (right) for the HWW analysis.

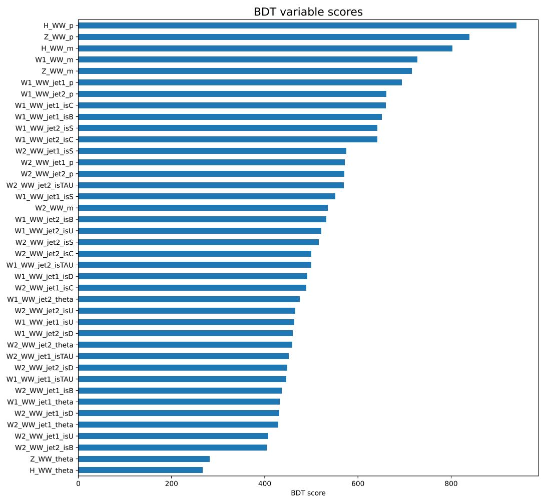

Fig. 10 Ranking of the variables used in the BDT for the HWW analysis.

8 Results

A simultaneous fit is performed using the output discriminants from the HZZ and HWW analyses, with two signal strength parameters extracted to determine their respective uncertainties. Background processes are assigned a 1% normalization uncertainty. The resulting expected uncertainties are 1.48% for $\sigma_{\mathrm{ZH}} \times \mathcal{B}(H \to WW^*)$ and 8.20% for $\sigma_{\mathrm{ZH}} \times \mathcal{B}(H \to ZZ^*)$.

References

[1] Benedikt, M., al.: FCC-ee: The Lepton Collider : Future Circular Collider Conceptual Design Report Volume 2. Eur. Phys. J. Spec. Top. **228**(261), 400 (2019) https://doi.org/10.1140/epjst/e2019-900045-4

[2] Eysermans, J., Selvaggi, M.: FCC note: Total ZH hadronic cross-section at FCC-ee and combination with leptonic final states (2025). https://doi.org/10.17181/9gzts-hsm97

[3] Eysermans, J., Li, A., Bernardi, G.: FCC note: Higgs boson mass and model-independent ZH cross-section at FCC-ee in the di-electron and di-muon final states (2023). https://doi.org/10.17181/jfb44-s0d81

[4] Favereau, J., Delaere, C., Demin, P., Giammanco, A., Lemaître, V., Mertens, A., Selvaggi, M.: DELPHES 3, A modular framework for fast simulation of a generic collider experiment. JHEP **02**, 057 (2014) https://doi.org/10.1007/JHEP02(2014)057 arXiv:1307.6346 [hep-ex]

[5] Sjostrand, T., Mrenna, S., Skands, P.Z.: PYTHIA 6.4 Physics and Manual. JHEP **05**, 026 (2006) https://doi.org/10.1088/1126-6708/2006/05/026 arXiv:hep-ph/0603175

[6] Kilian, W., Ohl, T., Reuter, J.: WHIZARD: Simulating Multi-Particle Processes at LHC and ILC. Eur. Phys. J. C **71**, 1742 (2011) https://doi.org/10.1140/epjc/s10052-011-1742-y arXiv:0708.4233 [hep-ph]

[7] Sjöstrand, T., Ask, S., Christiansen, J.R., Corke, R., Desai, N., Ilten, P., Mrenna, S., Prestel, S., Rasmussen, C.O., Skands, P.Z.: An introduction to PYTHIA 8.2. Comput. Phys. Commun. **191**, 159–177 (2015) https://doi.org/10.1016/j.cpc.2015.01.024 arXiv:1410.3012 [hep-ph]

[8] Cacciari, M., Salam, G.P., Soyez, G.: Fastjet user manual. Eur. Phys. J. C **72**, 1896 (2012) https://doi.org/10.1140/epjc/s10052-012-1896-2

[9] Bedeschi, Franco and Gouskos, Loukas and Selvaggi, Michele: Jet flavour tagging for future colliders with fast simulation. Eur. Phys. J. C **82**, 646 (2022) https://doi.org/10.1140/epjc/s10052-022-10609-1 arXiv:2202.03285 [hep-ex]