

Feasibility study for $hh \rightarrow b\bar{b}\tau^+\tau^-$ at the FCC-hh using graph neural networks

Jordy Degens^{1*}, Monica D’Onofrio^{1*}, Carl Gwilliam^{1*},
Lennox Wood^{1*}, Samuel Valentine^{1*} and Christiano Sebastiani^{2*}

¹University of Liverpool, Liverpool, United Kingdom.

²CERN, Geneva, Switzerland.

*Corresponding author(s). E-mail(s): jordy.degens@liverpool.ac.uk;
onofrio@liverpool.ac.uk; gwilliam@hep.ph.liv.ac.uk;
sglwoo10@liverpool.ac.uk; S.Valentine@liverpool.ac.uk;
cristiano.sebastiani@cern.ch;

Abstract

A feasibility study of the di-Higgs production channel at the Future Circular Collider proton-proton collider, FCC-hh, has been conducted focusing on the $hh \rightarrow b\bar{b}\tau^+\tau^-$ final state. Both the leptonic-hadronic (lep-had) and fully hadronic (had-had) decays of the taus are considered, leveraging advanced Graph Neural Network (GNN) techniques for event classification capabilities. The analysis employs Graph Attention Networks (GATs), with each event represented as a fully connected graph, where nodes correspond to reconstructed physics objects, and edges capture their relational properties. Additional complex features – such as invariant masses of b -jet and tau leptons, angular separations, and transverse kinematic variables – are integrated to enhance event classification. This innovative application of GNNs significantly improves the ability to extract signal events, offering a powerful technique for precision Higgs self-coupling measurements at the FCC-hh. Considering a centre-of-mass energy of 84 TeV, a statistical significance of $Z \sim 26$ ($Z \sim 92$) is expected for the lep-had (had-had) channel, under the no-systematic assumption and for a specific value of the signal score threshold. Considering systematic uncertainties at the 2% on the SM background, the expected significances is around 30 for the had-had channel alone, demonstrating how the Higgs self-coupling could be constrained at the percentage level at the FCC-hh.

1 Introduction

After the discovery of the Higgs boson (H) in 2012 [1, 2], the ATLAS [3] and CMS [4] Collaborations at the Large Hadron Collider [5] (LHC) have been focusing on an extensive measurement programme of its properties. In this, the measurement of HH production, sensitive to modifications to the Brout–Englert–Higgs (BEH) potential and therefore crucial for our understanding of electroweak symmetry breaking (EWSB), remains one of the main targets of the LHC and its upgrade, the High-Luminosity (HL) LHC. With the HL-LHC, ATLAS and CMS will observe the SM di-Higgs-boson production with a significance exceeding 7σ , and will directly probe the trilinear self-coupling from the Standard Model (SM), λ_3 , with a precision better than 30% in absence of beyond SM (BSM) effects.

With its extraordinary energy and integrated luminosity, the Future Circular hadron-hadron Collider, FCC-hh [6, 7], will provide an unprecedented potential to explore in detail the Higgs boson properties and its self-coupling. The FCC-hh is now expected to produce collisions at the centre of mass (c.o.m) energy of $\sqrt{s}=84$ TeV and to deliver an integrated luminosity of 30 ab^{-1} . Previous studies performed assuming $\sqrt{s}=100$ TeV have shown that Higgs-self coupling will be measured with a 5% precision via double Higgs production [8, 9]. Despite the lower c.o.m. energy, such level of precision can be achieved and possibly surpassed, thanks to improvements in analysis techniques that have already increased the sensitivity of ATLAS and CMS current analyses, i.e. using graph-based architectures for deep-learning based jet tagging for $H \rightarrow b\bar{b}$ and $H \rightarrow \tau^+\tau^-$.

In this note, we report a feasibility study of the di-Higgs production channel at the FCC-hh focusing on the $HH \rightarrow b\bar{b}\tau^+\tau^-$ final state. Figure 1a shows the feynman diagram of the dominant production mode, gluon-fusion (ggF), followed by the decay of the two Higgs boson in $b\bar{b}$ and $\tau^+\tau^-$, respectively. A total cross section $\sigma_{HH} = 0.95 \text{ pb}$ is calculated at next-to-next-to-leading order (NNLO) for a Higgs mass $m_H = 125 \text{ GeV}$ at $\sqrt{s} = 84 \text{ TeV}$ [10]. This does not include contributions from sub-dominant production modes like vector-boson-fusion (VBF) production and $t\bar{t}HH$ production, both expected to be order 10% of the total. The $b\bar{b}\tau^+\tau^-$ final state captures 7.3% of all HH final states (see Figure 1b) and provides a compromise between expected signal event yield and background contamination.

We study the various decay modes of the taus and consider two exclusive final states, the mixed hadronic-leptonic decays $H \rightarrow \tau_\ell\tau_h$ (ℓ -had), and the pure hadronic decays $H \rightarrow \tau_h\tau_h$ (had-had), leveraging advanced Graph Neural Network (GNN) techniques for event classification. The analysis employs Graph Attention Networks (GATs), with each event represented as a fully connected graph, where nodes correspond to reconstructed physics objects, and edges capture their relational properties. Additional complex features – such as invariant masses of b -jet and tau leptons, angular separations, and transverse kinematic variables – are integrated to enhance event classification. With its large cross section ($\sigma_{t\bar{t}}^{NNLO} \simeq 32000 \text{ pb}$ at 84 TeV), top-pair production is the main SM background process considered in the classification procedure. Appropriate corrections (rescaling) are applied, based on the background composition found in current LHC di-Higgs analyses in the same channels [11], to take into account the contributions of other background sources. In particular, the

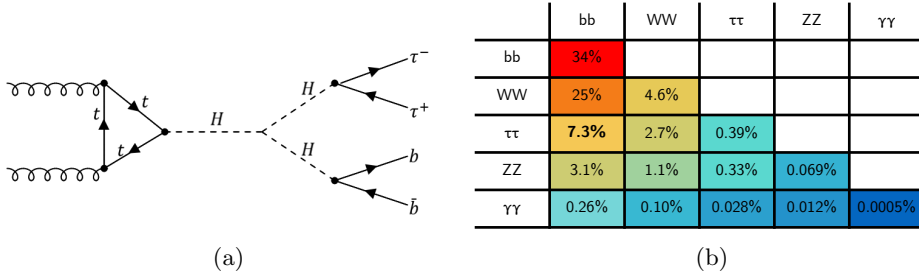


Fig. 1: Left: Di-Higgs $bb\tau\tau$ decay channel. Right: Production rates of different decays of di-higgs.

probability that a light jet is misidentified as tau ($j \rightarrow \tau_h$ fake rate) leads to large contributions of the multi-jet (QCD) and W, Z boson plus jets backgrounds in the had-had channel that must be factored in.

The note is structured as follows: Section 2 presents an overview of the Monte Carlo (MC) simulated samples used, the analysis framework and the definition of the event selection. Section 3 illustrates the GNN structure and inputs, while Section 4 shows the results for both lep-had and had-had analysis channels in terms of statistical significance, with considerations on the impact of systematic uncertainties. Work is in progress to estimate the sensitivity in terms of signal strength and λ_3 coupling precision, and to evaluate the potential of performing differential cross section measurements as a function of the invariant mass of the di-higgs system (m_{HH}). We emphasize that the analysis presented here remain relatively simplistic compared to LHC standards. However, it is valuable in establishing targets for the ultimate achievable precision and overall detector performance.

2 Simulated samples and analysis framework

The software infrastructure utilised in this study has been developed for the previous EU Strategy studies and extensively documented, see for example Ref [12]. While the framework has been improved following recent MC generators updates and further studies and fixes, its structure remains largely unchanged. The prototype of a baseline FCC-hh detector that could fulfill the requirements to perform the presented analyses has been designed for the FCC CD [13].

Hard scattering events are produced in the form of Les Houches Events (LHE) [14] (via MADGRAPH5_aMC@NLO [15, 16] for the $t\bar{t}$ background, with POWHEG [17–19] for the HH signal). Shower evolution and hadronization are performed with PYTHIA 8. The FCC-hh detector response is simulated with DELPHES [20]. In the latest stages of the MC generation, PYTHIA 8 [21] and DELPHES are interfaced directly to the FCC software (FCCSW) [22], allowing to produce final stage events in the FCC Event Data Model (EDM) format [23]. All the generated signal and background events can be found in a database ¹. A detailed description of the FCC-hh

¹<https://fcc-physics-events.web.cern.ch/fcc-hh/index.php>

detector parametrization in DELPHES is available in Ref. [24]. It is noted that the overall contribution of pile-up has been neglected in the parametrised simulation. Although DELPHES allows for such possibility, including pile-up interactions would result in an overly conservative object reconstruction performance since the current DELPHES FCC-hh setup does not possess well-calibrated pile-up rejection tools that allow to recover the nominal detector performance. Studies at the HL-LHC suggest that improvements expected in the reconstruction algorithms and data processing offer significantly enhanced trigger, tracking, and pileup suppression capabilities, which are expected to more than compensate for the challenges posed by the large pile-up by the time the FCC-hh will be in operation.

2.1 Signal and background samples

Signal samples consist of simulated events from non-resonant ggF production of Higgs boson pairs at c.o.m energy 84 TeV and 100 TeV, with one Higgs boson decaying into $b\bar{b}$ and the other one to $\tau^+\tau^-$. A total of 1.2M (5M) events have been generated at $\sqrt{s} = 84(100)$ TeV, assuming their scale factors with respect to the SM predictions, $\kappa_\lambda = \lambda_3/\lambda_3^{SM}$, equal to 1. Additional samples have been generated with value of the self-coupling modifier $\kappa_\lambda = 2.4$ and 3, to quantify the constraints on this parameter.

Top-pair production samples are also generated at $\sqrt{s} = 84$ TeV and 100 TeV, considering inclusive decays of the top-quark (47M and 80M for 84 and 100 TeV, respectively), and with the requirement that both top quarks decay into a leptonic final state (di-leptonic samples, 72M events at 84 TeV). The di-leptonic sample is used for the final training of the GNN for the nominal 84 TeV results as $t\bar{t} \rightarrow \ell\nu b\ell\nu\bar{b}$ constitute the largest background component. Inclusive samples are used to evaluate the contributions from fake taus in case of fully-hadronic or semi-leptonic decays of the top-quark pairs.

Other SM background sources include single-top (Wt), W or Z boson production associated to jets and, relevant for the had-had channel only, QCD multijet backgrounds. Due to the difficulties to produce sufficiently large MC samples for these processes, their contributions is evaluated assuming the background composition found by current LHC analyses of lep-had and had-had HH channels [11], backed-up by studies performed at pre-selection level using 100 TeV samples and reported in the following.

2.2 Objects reconstruction and event selection

Delphes utilizes a parameterized detector response, incorporating resolution functions and efficiency factors. Its simulation framework includes a track propagation system within a magnetic field, along with electromagnetic and hadronic calorimeters and a muon identification system. Delphes generates physics objects such as tracks, calorimeter signals, and higher-level objects like isolated leptons, jets, and missing transverse momentum $\mathbf{p}_{T,\text{miss}}$ (with its modulus referred to as missing transverse energy, $E_{T,\text{miss}}$, in the following). Additionally, it features a particle-flow reconstruction method that integrates tracking and calorimeter data to produce particle-flow candidates, including

charged hadrons, neutral hadrons, and photons. These reconstructed particles serve as inputs for jet clustering, missing energy calculations, and isolation determinations.

Jets are clustered by the anti- k_T algorithm [25] with a parameter $R=0.4$. For electrons and muons, the relative isolation is computed by summing the p_T of all particle-flow candidates in a cone of radius $R=0.3$ around the particle of interest and dividing by the particle's p_T . The reconstruction and identification (ID) efficiencies for leptons are parameterised as function of p_T and pseudo-rapidity, η . Delphes also provides heavy flavour tagging, in particular τ_h and b -jet identification. Both hadronic taus and b -jets are reconstructed using the total visible 4-momentum of the jet. The tagging efficiencies rely on a parameterisation of the (mis-)identification probability as a function of (p_T, η) . The so-called “Medium” working point is used for the identification of b -jets. The b -tagging efficiency ϵ_b and the light (charm) mis-tag rates $\epsilon_{l(c) \rightarrow b}$ are assumed to be $\epsilon_b = 75\text{--}85\%$ and $\epsilon_{l(c) \rightarrow b} = 0.6\text{--}1.5$ (10-20)%. Details of these efficiencies are presented in Table 1.

Table 1: b -jet tagging efficiency, light-jet mistagging rate, and c -jet mistagging rate as parameterized in Delphes for different ranges in p_T and η .

| | $ \eta < 4.0$ | $ \eta > 4.0$ |
|----------------------------------|----------------|----------------|
| b-jet Tagging Efficiency | | |
| $p_T \leq 20$ GeV | 0% | 0% |
| $20 < p_T \leq 50$ GeV | 75% | 0% |
| $50 < p_T \leq 100$ GeV | 84% | 0% |
| $100 < p_T \leq 200$ GeV | 85% | 0% |
| $200 < p_T \leq 500$ GeV | 83% | 0% |
| $500 < p_T \leq 1000$ GeV | 80% | 0% |
| $p_T > 1000$ GeV | 80% | 0% |
| Light-jet Mistagging Rate | | |
| $p_T \leq 20$ GeV | 0% | 0% |
| $20 < p_T \leq 50$ GeV | 1.5% | 0% |
| $50 < p_T \leq 100$ GeV | 0.8% | 0% |
| $100 < p_T \leq 200$ GeV | 0.6% | 0% |
| $200 < p_T \leq 500$ GeV | 0.8% | 0% |
| $500 < p_T \leq 1000$ GeV | 1.0% | 0% |
| $p_T > 1000$ GeV | 1.2% | 0% |
| c-jet Mistagging Rate | | |
| $p_T \leq 20$ GeV | 0% | 0% |
| $20 < p_T \leq 50$ GeV | 20% | 0% |
| $50 < p_T \leq 100$ GeV | 15% | 0% |
| $100 < p_T \leq 200$ GeV | 13% | 0% |
| $200 < p_T \leq 500$ GeV | 12% | 0% |
| $500 < p_T \leq 1000$ GeV | 11% | 0% |
| $p_T > 1000$ GeV | 10% | 0% |

The so-called “Medium” working point is used for τ ID, in order to operate at an optimal signal-to-background rejection in each decay channel, with an efficiency of 80%(65%) in the barrel(forward) region and a $j \rightarrow \tau_h$ misidentification rate of 2% in

the barrel and 0.75% for the forward region. Details on the p_T and η dependencies of tau efficiency and fake rate are presented in Table 2.

Table 2: Tau identification rates as parameterised in Delphes for different range in p_T and η .

| | $ \eta < 4.0$ | $ \eta > 4.0$ |
|---|----------------|----------------|
| τ_h Efficiency | | |
| $p_T \leq 20$ GeV | 0% | 0% |
| $20 < p_T \leq 100$ GeV | 80% | 0% |
| $p_T > 100$ GeV | 88% | 0% |
| $j \rightarrow \tau_h$ misidentification Rate | | |
| $p_T \leq 20$ GeV | 0% | 0% |
| $20 < p_T \leq 100$ GeV | 2% | 0% |
| $p_T > 100$ GeV | 1% | 0% |

Events are required to contain exactly two b -tagged jets with $p_T > 30$ and $|\eta| < 4$. Up to one additional extra jet with the same requirements in p_T and η is allowed. Exactly one or two hadronic taus are required for the lep-had and had-had channels, respectively, also with $p_T > 30$ and $|\eta| < 4$. No additional electrons or muons ($ell = e, \mu$) are required for the had-had channel, while exactly one electron or muon with p_T threshold at 10 GeV within the same η range is required for the lep-had channel, where the lepton ℓ is the proxy for τ_ℓ . The $\ell - \tau_H$ and $\tau_H \tau_H$ pairs are required to be opposite-sign (OS) are originating from the same Higgs boson. Details of the event selection are presented in Table 3. Overlap removal of tau objects and b -jets is also applied to avoid double-counting. The transverse momenta of b -jets and τ_h , as well as the missing transverse energy at this stage of selection is reported as example of kinematic distributions in Figure 2 for the had-had channel, for signal and $t\bar{t}$ background events.

Table 3: Event selection for the lep-had and had-had channels.

| $\tau_{lep}\tau_{had}$ | $\tau_{had}\tau_{had}$ |
|--|--|
| $N(b-jet) = 2, p_T > 30, \eta < 4$ | |
| $N(light-jet) = 0 \text{ or } 1, p_T > 30, \eta < 4$ | |
| $N(\tau) = 1, p_T > 30, \eta < 4$ | $N(\tau) = 2, p_T > 30, \eta < 4$ |
| $N(\ell) = 1, p_T > 10$ | $N(\ell) = 0$ |
| $N(\tau_{had}) = 1$ | $N(\tau_{had}) = 2$ |
| $\ell\tau_{had}$ OS | $\tau_{had}\tau_{had}$ OS |

A set of variables based on the kinematic properties of the b -jets, tau objects, missing transverse energy and the possible extra-jet in the event are also used in the analysis, and are defined as follows:

- The invariant mass of the two b -jets, m_{bb} , as in the signal they originate from the Higgs boson decays.

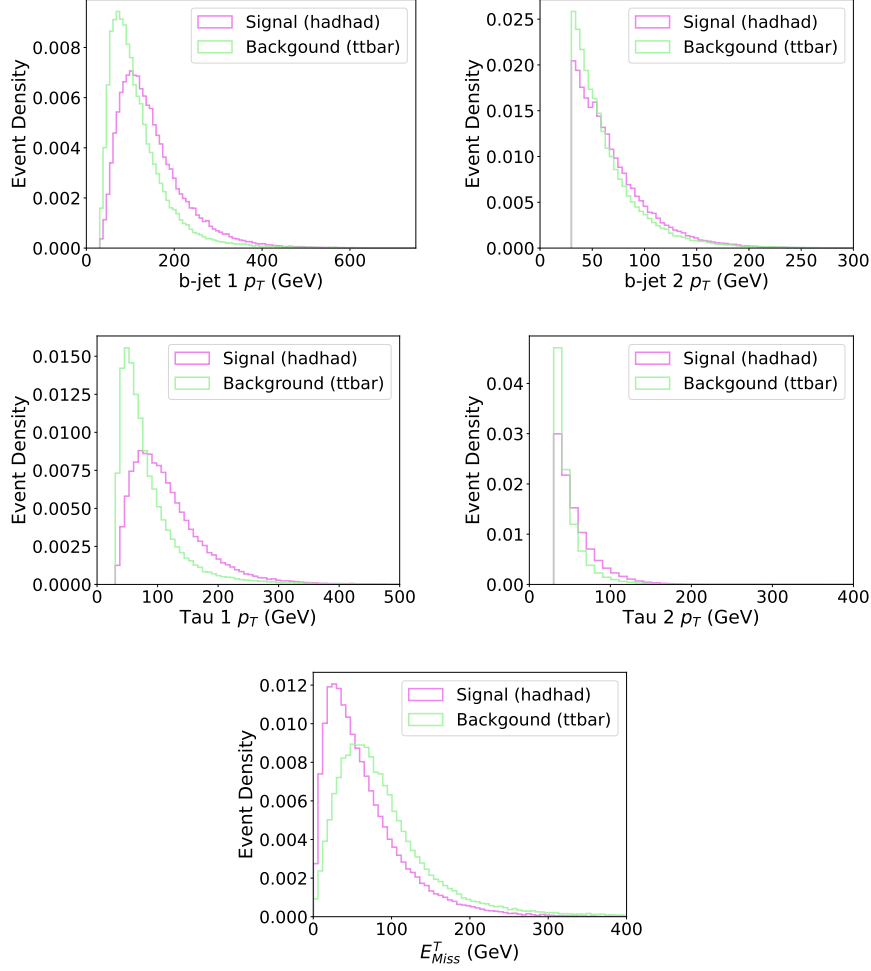


Fig. 2: From top-left: Normalised distributions for the transverse momentum of the b -jets and τ_h , as well as the missing transverse energy for the had-had channel, for signal and $t\bar{t}$ background events.

- The radial distance of the two b -jets, ΔR_{bb} , defined in terms of η and the azimuthal angle, ϕ .
- The invariant mass of the two taus, $m_{\tau_h \ell}$ and $m_{\tau_h \tau_h}$, for the lep-had and had-had channels, respectively, as proxy of the invariant mass of the Higgs decaying into taus.
- The radial distance of the two tau objects, $\Delta R_{\tau_h \ell}$ and $\Delta R_{\tau_h \tau_h}$, depending on the channel.
- The $E_{T,miss}$ centrality, a variable that quantifies the angular position of E_{miss}^T relative to the τ and ℓ products in the transverse plane [26].

- The difference in p_T of the two tau objects, $\Delta p_T(\tau_h, \ell)$ and $\Delta p_T(\tau_h, \tau_h)$, depending on the channel.
- The transverse mass $M_T(\ell)$, defined from the lepton ℓ transverse momentum and the E_{miss}^T , and used only in the lep-had channel. For one on-shell W boson decaying leptonically, as present in $t\bar{t}$ events, the observable has an upper endpoint at the W boson mass.

In addition, the invariant mass and the azimuthal distance of the two higgs candidates reconstructed as b -jets and tau object pairs, m_{hh} and $\Delta\phi_{hh}$, are also used. Figure 3 shows the shape comparison of kinematic variables for the lep-had channel as example, for signal and $t\bar{t}$ background events.

These variables, effective to discriminate between signal and SM backgrounds as amply demonstrated in current LHC analyses [27], are utilised in the GNN-based analysis to aid in the classification of the events, as reported in the following.

3 Graph Neural Network analysis

Graph Neural Networks have emerged as promising tools for analysing complex datasets like those produced by hadron colliders. They are a specialized class of neural networks designed to process and analyse graph-structured data, with their key strength lying in their ability to learn representations that capture complex relationships among nodes in a graph. By representing particles and their interactions as nodes and edges in a graph, GNNs can model complex relationships of particle objects and achieve excellent discrimination between signal and backgrounds [28].

A Graph Attention Network (GAT) ² is used as architecture for the NN, as it is found to perform better than classical Graph Convolutional Networks (GCNs). We represent each collision event as a fully connected graph G with 5-6 nodes (N), where each node represents a particle and contains up to 3 kinematic features. Mathematically, each event is encoded as a feature matrix $X \in \mathbb{R}^{N \times 3}$, where rows correspond to particles and columns to their features. The scheme for particle collisions and their corresponding features are depicted in the graph in Figure 4 and, in detail, in Table 4. With reference to the Table, each row represents a distinct particle. The first two rows correspond to the tau objects, either $\tau 1$ and $\ell 1$ in the lep-had channel, or $\tau 1$ and $\tau 2$ in the had-had channel. The third and fourth rows correspond to the b -jets, $b1$ and $b2$. The fifth row, named *energy*, is the missing transverse energy. Events with no extra jets have 5 nodes, while a 6th node is introduced if it exists ($j1$). The Table columns represent the features (F) which refers to the p_T , η and ϕ of jets and taus for the corresponding objects. For the *energy*, only the azimuthal angle of E_T^{Miss} can be defined due to the unknown longitudinal momentum, so that one of the feature is depicted as 'nan'. The graphs are created so that they were fully connected, meaning all nodes had edges between them. They are also undirected, hence the information can be passed between all nodes of the graph in any direction.

The GAT architecture uses 3 convolutional layers with 50 hidden channels, succeeded by a rectified linear unit (ReLU) activation function after each layer. It employs global mean pooling for aggregation and concludes with a linear layer to output the

²https://pytorch-geometric.readthedocs.io/en/latest/generated/torch_geometric.nn.models.GAT.html

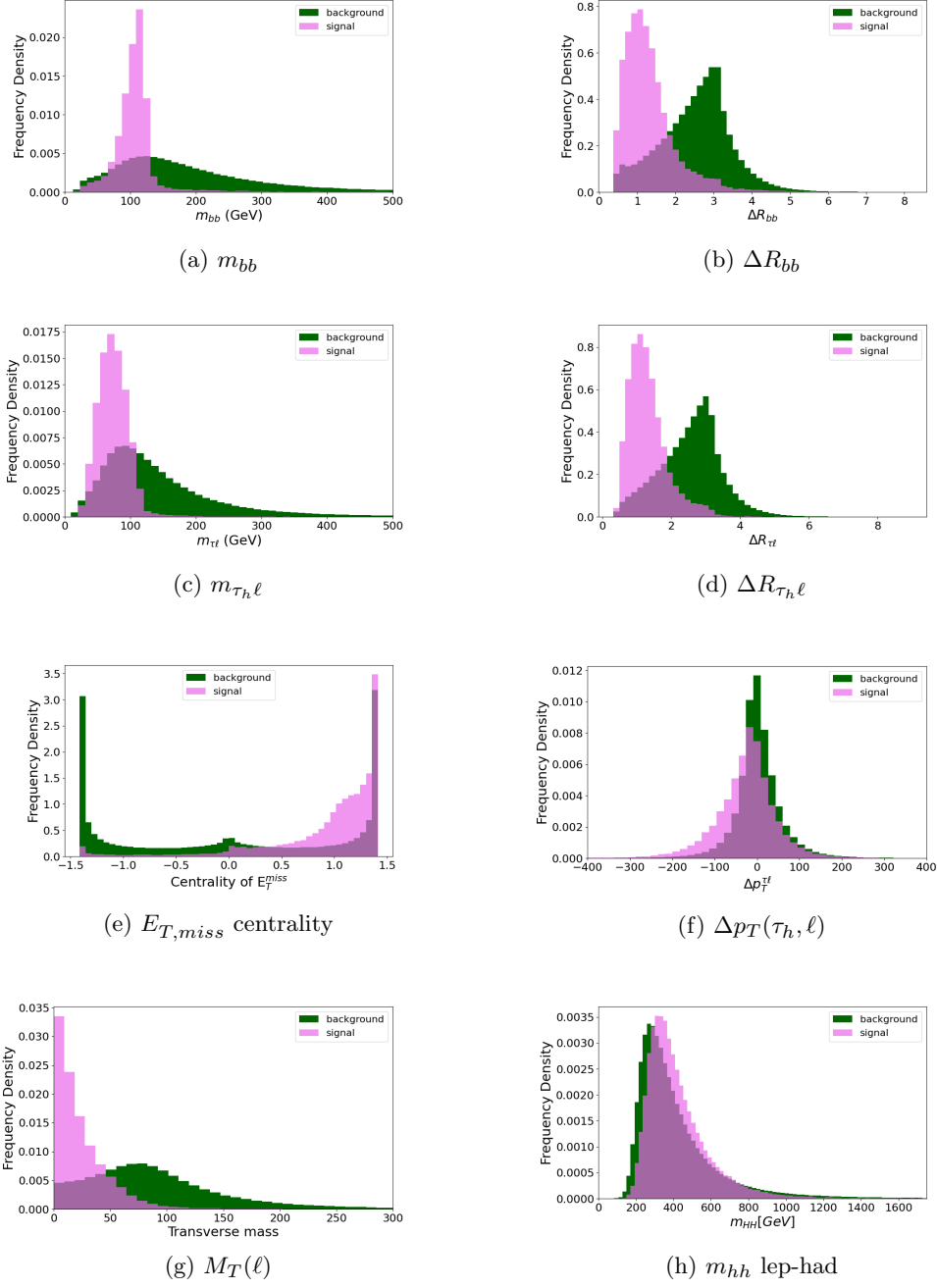


Fig. 3: Normalised distributions for signal and background events for various kinematic variables in the lep-had channel. From (a)-(h): the invariant mass of the b-jets, m_{bb} ; the radial distance between the b-jets, ΔR_{bb} ; the invariant mass of the lepton and the hadronic tau, $m_{\tau_h \ell}$; the radial distance between the tau objects, $\Delta R_{\tau_h \ell}$; the $E_{T,miss}$ centrality; the difference in p_T of the tau objects, $\Delta p_T(\tau_h, \ell)$; the transverse mass of the tau-lepton, $M_T(\ell)$; and the invariant mass of the two reconstructed Higgs candidates, m_{hh} .

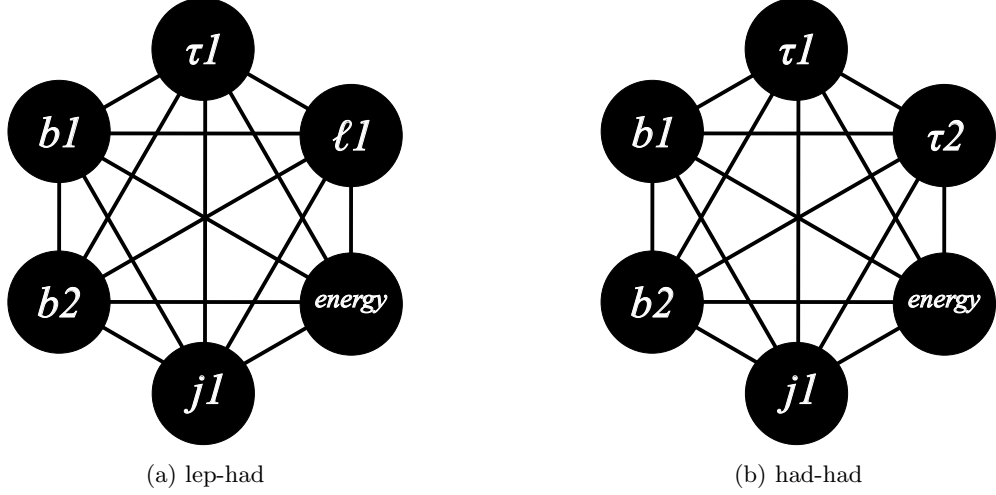


Fig. 4: Schematic representation of events modelled as a fully connected graph, highlighting the reconstructed particles as described in the text.

Table 4: Layout of each event in the dataset, 'nan' means not a number.

| Particle | F1 | F2 | F3 |
|-----------------|-----------------------|------------------------|------------------------|
| $\tau 1$ | $p_T^{\tau 1}$ | $\eta_{\tau 1}$ | $\phi_{\tau 1}$ |
| $\ell 1/\tau 2$ | $p_T^{\ell 1/\tau 2}$ | $\eta_{\ell 1/\tau 2}$ | $\phi_{\ell 1/\tau 2}$ |
| $b 1$ | $p_T^{b 1}$ | $\eta_{b 1}$ | $\phi_{b 1}$ |
| $b 2$ | $p_T^{b 2}$ | $\eta_{b 2}$ | $\phi_{b 2}$ |
| “energy” | E_T^{miss} | nan | ϕ_{MET} |
| (if there) “j1” | p_T^{jet} | η_{jet} | ϕ_{jet} |

class predictions. This process was repeated for 100 epochs for train and test sets to ensure that the GNN’s learning had fully converged. Training, validation and test samples were defined through a k-fold procedure to maximise the statistics of the training samples. In this process, the MC samples are randomly split in 5 distinct sets, grouped such that 60% of the events are used for training to maximise the statistics, with 20% for validation and 20% for testing. The GNN output is in terms of ‘logits’ for each event, and a sigmoid function is used to convert it to a prediction between 0 and 1. At this point, finally, the GNN outputs a prediction for a probability of how likely all the graphs are to be signal, referred to as signal score. The Receiver Operating Characteristic (ROC) curve was also used to visualize the binary classification model’s performance across various thresholds, plotting the True Positive Rate (TPR) against the False Positive Rate (FPR), with the AUC (Area Under the Curve) summarizing the model’s ability to distinguish between positive and negative classes.

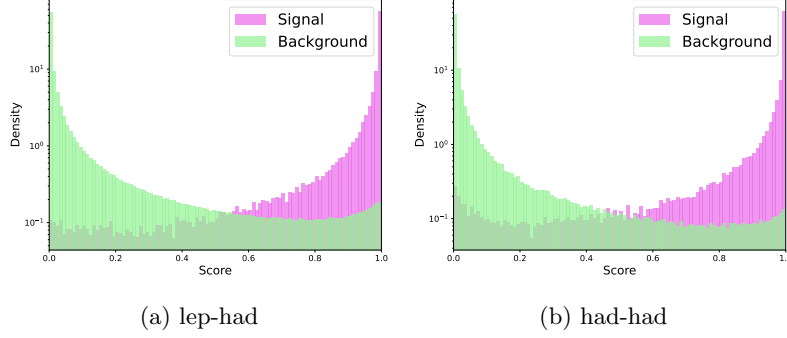


Fig. 5: Density distribution of the GNN output scores for the lep-had (left) and had-had (right) channel.

In addition to the nodes depicted in Table 4, each kinematic “complex” variable introduced in Section 2 – namely m_{bb} , $m_{\tau_h \ell}$ or $m_{\tau_h \tau_h}$, $E_{T,miss}$ centrality, $\Delta p_T(\tau_h, \ell)$ or $\Delta p_T(\tau_h, \tau_h)$, ΔR_{bb} , $M_T(\ell)$, m_{hh} and $\Delta \phi_{hh}$ – is used as additional node to represent the relational features between different objects. The values for each of these variables was calculated and added to the dataset. It was found that adding the kinematic relations between each object (i.e. the invariant mass of the two b -jets, or that of the tau candidates), substantially better distinction between signal and background was achieved, with AUC from 0.6 to about 0.99 when all variables were included. The dataset was then normalised so that all individual and complex features were transformed into values between 0 and 1, with a value of 0 corresponding to the lowest value for a given feature and 1 corresponding to the highest value, to guarantee equal consideration of the GNN in the training.

The GNN output score density distribution for the lep-had (left) and had-had channel is shown in 5 training on large statistics $t\bar{t}$ di-leptonic samples generated at 84 TeV c.o.m. energy. Excellent separation is observed in both channels. Equivalent performance and score distributions are obtained when training signal against inclusive $t\bar{t}$ events, at 84 and 100 TeV. Figure 6 shows the ROC curves for the lep-had and had-had, respectively, for the nominal 84 TeV c.o.m. energy.

4 Results

To measure the sensitivity of the HH signal strengths, the signal and background expected yields are calculated considering a luminosity of 30 ab^{-1} at a c.o.m. energy of 84 TeV. The significance on the SM is calculated using the following formula:

$$Z = \frac{N_s}{\sqrt{N_b + (N_b \sigma)^2}}. \quad (1)$$

Here N_s (N_b) are the number of signal (background) events and σ incorporates the uncertainty on the background prediction. The significance is estimated for events with a GNN score higher than a certain value. The GNN is trained using 84 TeV

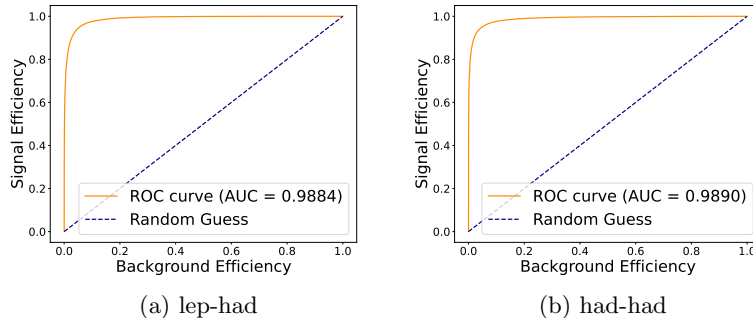


Fig. 6: Performance measurement in terms of ROC curve for classification models in case of lep-had (left) and had-had (right) channel.

signal and $t\bar{t}$ di-leptonic MC samples. Few score threshold values are chosen to show different signal purities. An even more stringent selection would be possible but it is not performed due to limited sizes of the MC samples.

While $t\bar{t}$ background contributions are expected to account for the largest fraction of SM background [8], some assumptions are made for a conservative estimate of N_b based on the SM background composition found by the Run-2 ATLAS analysis [11]. The ATLAS analysis uses BDT approaches based on kinematic variables as used in our study, hence SM processes contributing to the signal regions are expected to be distributed similarly to what expected at the FCC-hh (Figure 7).

For the lep-had channel, $t\bar{t}$ di-leptonic background events are the dominant component, followed by contributions from top-quark events (top-quark pair and single production) where a jet is misidentified as τ_h . Predictions from $t\bar{t} \rightarrow b\ell\nu b\tau_h\nu$ estimates by this analysis are rescaled by a factor ~ 2 . For the had-had channel, contributions from Z +jets and multijet backgrounds must also be taken into account, so that the predictions from $t\bar{t}$ di-leptonic samples are rescaled by a factor of ~ 8 . We note that Z +jets and multijet processes might have different events kinematic with respect to $t\bar{t}$, however in a real analysis, dedicated control regions can be easily set up and the GNN trained against more than one background category to ensure good separation with respect to the signal. Hence at this stage, a conservative scaling such as the one applied is a good proxy to keep into account the contributions.

The significances are summarized for the lep-had channel in Table 5 and in Table 6 for the had-had channel. They are calculated for a very conservative 5% uncertainty³ and more realistic uncertainties of 1 and 2%. Under the no-systematic assumption, significances $Z \sim 26$ and $Z \sim 92$ are expected for signal score thresholds set at 0.999 for the lep-had and had-had channel, respectively. Even for the most conservative uncertainty of 5%, the expected significances is above 10 for the had-had channel

³SM background yields are expected to be estimated in control regions through data-driven approaches, leading to negligible residual experimental systematic uncertainties, and reduced dependencies on theoretical modelling of SM processes. For the latter, further progress in high-precision theoretical calculations are expected once the FCC-hh will be operational.

if a tighter score threshold is considered. The lep-had significances are lower, but a significance above 3 is still obtained for an uncertainty of 2% ($s > 0.9996$).

Table 5: Significances in the lep-had channel for different cut values and different systematic uncertainties. The Cut column shows the lower bound for the GNN output score. The next two columns show the estimated number of signal and background events. The last four columns show the significances for no systematic uncertainty and a systematic uncertainty of 1, 2 and 5%.

| Cut | Signal | Background | Z | Z (1%) | Z (2%) | Z (5%) |
|------------|--------|------------|------|----------|----------|----------|
| > 0.99 | 57 900 | 5 800 000 | 24.0 | 0.996 | 0.498 | 0.199 |
| > 0.995 | 46 800 | 2 690 000 | 28.5 | 1.74 | 0.870 | 0.348 |
| > 0.999 | 14 200 | 287 000 | 26.5 | 4.86 | 2.46 | 0.987 |
| > 0.9996 | 2700 | 43 400 | 13.0 | 5.61 | 3.03 | 1.24 |

For illustration purposes, Figures 8 and 9 show some examples of kinematic distributions for signal and background, with GNN score threshold set at 0.8 to retain good statistics, for the lep-had and had-had channels, respectively. Signal is shown for $s > 0.8$, while background distributions for higher and lower score events are overlaid. It is clear that the GNN “learns” from mass-related correlations (e.g. m_{bb} , $m_{\tau\tau}$), while additional discriminating power exists from radial distances variables (e.g. ΔR_{bb} and $\Delta R_{\tau\tau}$) for both lep-had and had-had.

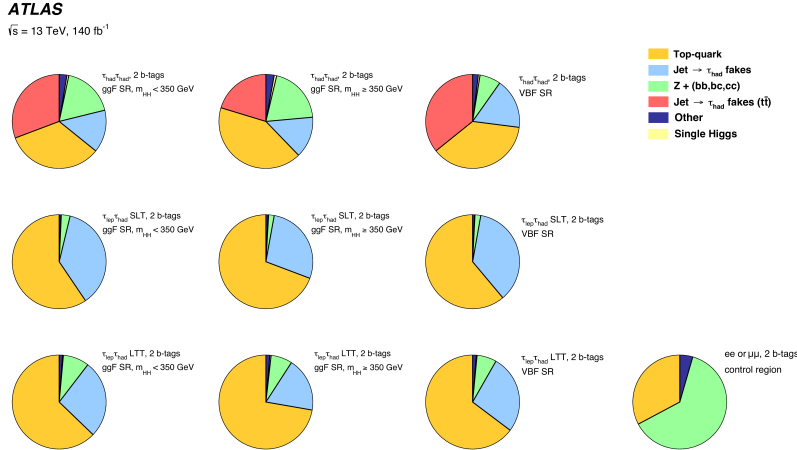


Fig. 7: Composition of SM background for the ATLAS Run-2 analysis used as proxy to rescale MC predictions obtained with $t\bar{t}$ di-leptonic samples only. For details on the analysis, see Reference [11].

Table 6: Significances in the had-had channel for different cut values and different systematic uncertainties. The Cut column shows the lower bound for the GNN output score. The next two columns show the estimated number of signal and background events. The last four columns show the significances for no systematic uncertainty and a systematic uncertainty of 1,2 and 5%.

| Cut | Signal | background | Z | Z (1%) | Z (2%) | Z (5%) |
|----------|--------|------------|------|--------|--------|--------|
| > 0.99 | 72 400 | 1 720 000 | 55.2 | 4.19 | 2.10 | 0.840 |
| > 0.995 | 63 400 | 876 000 | 67.7 | 7.19 | 3.61 | 1.45 |
| > 0.999 | 40 300 | 192 000 | 92.0 | 20.5 | 10.4 | 4.20 |
| > 0.9997 | 22 100 | 36 000 | 117 | 54.3 | 29.7 | 12.2 |

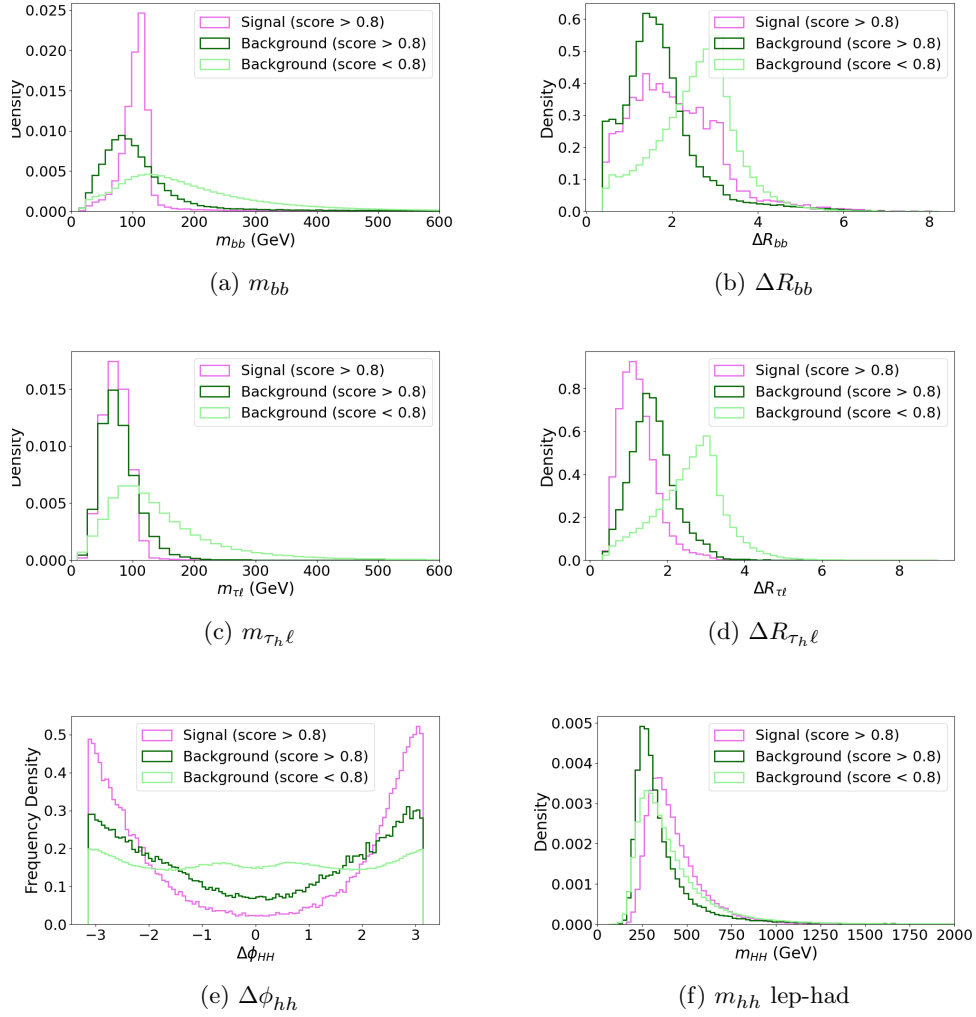


Fig. 8: Normalised distributions for signal and background events for various kinematic variables in the lep-had channel for signal score above 0.8 (signal and background), and below 0.8 for background. From (a)-(f): the invariant mass of the b-jets, m_{bb} ; the radial distance between the b-jets, ΔR_{bb} ; the invariant mass of the lepton and the hadronic tau, $m_{\tau_h \ell}$; the radial distance between the tau objects, $\Delta R_{\tau_h \ell}$; the $\Delta \phi_{hh}$ of the two reconstructed Higgs candidates; and their invariant mass, m_{hh} .

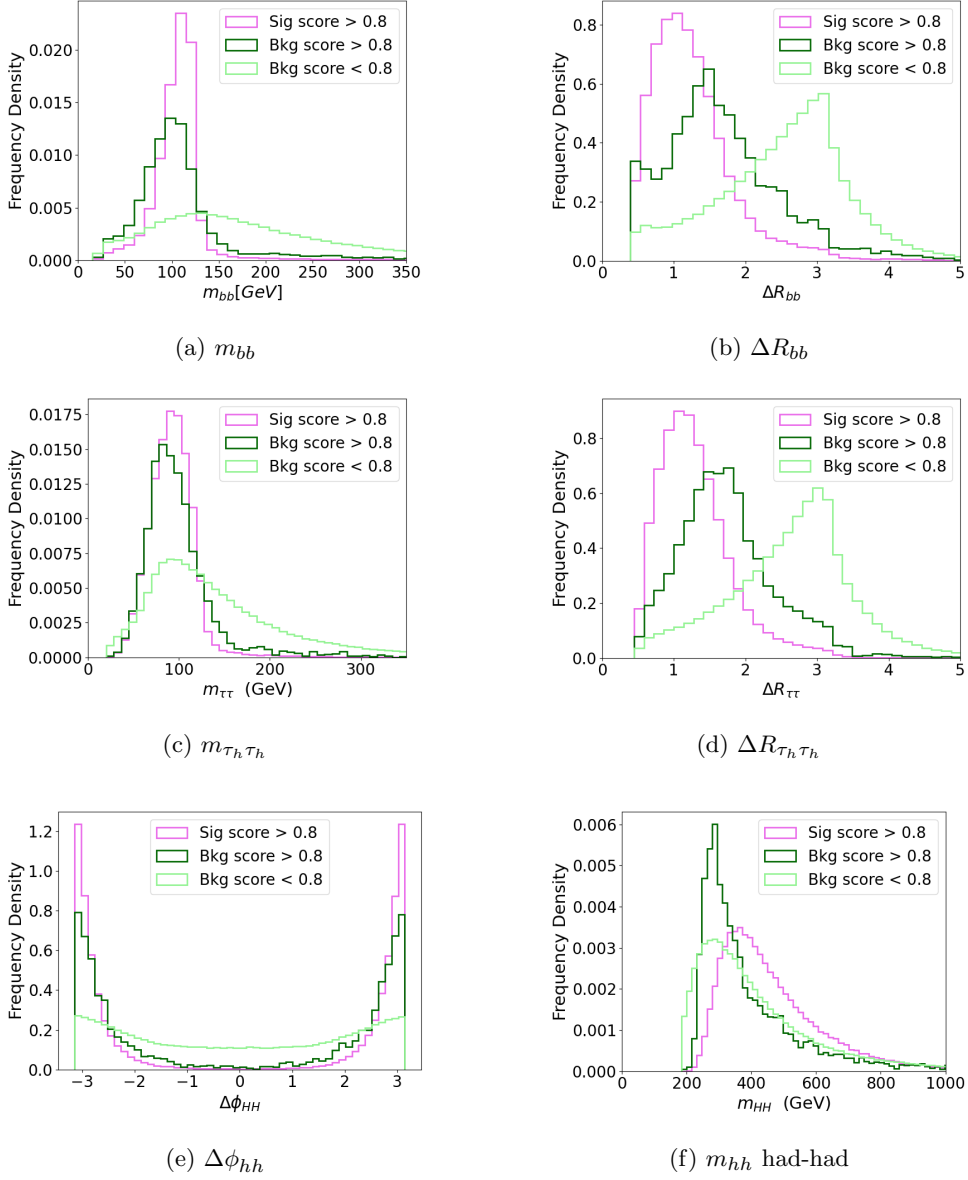


Fig. 9: Normalised distributions for signal and background events for various kinematic variables in the had-had channel for signal score above 0.8 (signal and background), and below 0.8 for background. From (a)-(f): the invariant mass of the b-jets, m_{bb} ; the radial distance between the b-jets, ΔR_{bb} ; the invariant mass of the lepton and the hadronic tau, $m_{\tau_h \tau_h}$; the radial distance between the tau objects, $\Delta R_{\tau_h \tau_h}$; the $\Delta \phi_{hh}$ of the two reconstructed Higgs candidates; and their invariant mass, m_{hh} .

In the had-had channel, events are further categorized into two regions based on the invariant mass of the di-Higgs system to enhance sensitivity to the Higgs self-interaction coupling strength, κ_λ . This categorization uses the value of m_{HH}^* , defined as:

$$m_{HH}^* = m_{HH} - m_{bb} - m_{\tau\tau} + 250 \text{ GeV}.$$

This formula corrects for the missing neutrinos originating from the τ decay in the calculation of m_{HH} , bringing it closer to the true value of the generated di-Higgs mass. Events with m_{HH}^* smaller (larger) than 350 GeV are categorized into the low-mhh (high-mhh) region. This approach mirrors the strategy adopted in recent LHC (ATLAS) di-Higgs measurements [11] and corresponds to the value of maximal interference. The expected number of events and significances, similar as for the inclusive results, are shown in Tables 7 and 8 for the low-mHH and high-mHH region respectively. Most of the sensitivity for the SM di-higgs sample is from high-mHH region, resulting in a significance of 5 sigma for the most conservative scenario of a systematic uncertainty of 5% and much larger significances for the more realistic scenario's of 1 or 2% uncertainty on the background. The significance for the SM is much lower in the low-mHH region due to the kinematics of the di-Higgs mass spectrum resulting in only very few signal events. Different values of κ_λ and many new physics scenarios can however predict an enhancement of the number of events in the low-mHH region. Work is in progress to evaluate the precision on the κ_λ parameter.

Finally, we note that the same studies have been performed using 100 TeV c.o.m. energy samples. The results are consistent with those obtained for $\sqrt{s} = 84$ TeV, so that reducing the proton-proton collisions energy does not impact the expectations to be able to measure precisely higgs-self coupling at the FCC-hh.

Table 7: Significances for the low-mHH region in the had-had channel for different cut values and uncertainties. The Cut column shows the lower bound for the GNN output score. The next two columns show the estimated number of signal and background events. The last four columns show the significances for no systematic uncertainty and a systematic uncertainty of 1,2 and 5%.

| Cut | Signal | background | Z | Z (1%) | Z (2%) | Z (5%) |
|--------|--------|------------|------|--------|--------|--------|
| > 0.9 | 3940 | 2 410 000 | 2.54 | 0.627 | 0.0819 | 0.0328 |
| > 0.95 | 2200 | 816 000 | 2.44 | 1.02 | 0.135 | 0.0540 |
| > 0.98 | 646 | 162 000 | 1.60 | 1.38 | 0.198 | 0.0796 |

5 Conclusion

The precise determination of the Higgs self-coupling will be a key objective for future high-energy collider experiments. A feasibility study of the di-Higgs production channel at the FCC-hh has been performed, focusing on the $hh \rightarrow b\bar{b}\tau^+\tau^-$ final state. Both

Table 8: Significances for the high-mHH region in the had-had channel for different cut values and uncertainties. The Cut column shows the lower bound for the GNN output score. The next two columns show the estimated number of signal and background events. The last four columns show the significances for no systematic uncertainty and a systematic uncertainty of 1,2 and 5%.

| Cut | Signal | background | Z | Z (1%) | Z (2%) | Z (5%) |
|------------|--------|------------|------|----------|----------|----------|
| > 0.9 | 93 100 | 10 500 000 | 28.7 | 3.40 | 0.442 | 0.177 |
| > 0.99 | 72 200 | 1 690 000 | 55.6 | 16.3 | 2.14 | 0.856 |
| > 0.997 | 56 300 | 558 000 | 75.3 | 37.6 | 5.03 | 2.02 |
| > 0.9991 | 35 000 | 138 000 | 94.3 | 86.6 | 12.6 | 5.07 |

the leptonic-hadronic (lep-had) and fully hadronic (had-had) decays of the taus are considered, leveraging advanced GNN techniques for event classification. Considering a centre-of-mass energy of 84 TeV, a statistical significance of $Z \sim 26$ ($Z \sim 92$) is expected for the lep-had (had-had) channel, under the no-systematic assumption for a signal score threshold set at 0.999. Even considering an over-conservative uncertainty of 5%, the expected significances is above 10 for the had-had channel if a tighter score threshold is considered. The lep-had significances are lower, but a significance above 3 is still obtained for an uncertainty of 2% ($s > 0.9996$). With this sensitivity, the Higgs self-coupling could be constrained at the percentage level using the had-had channel alone. Work is in progress to evaluate the precision on the κ_λ parameter considering events above and below a 350 GeV threshold on the invariant mass of the two Higgs candidates corrected for the missing neutrinos originating from the τ decay. Insights into interpretability and performance of the GNN indicates that improvement can be expected i.e. from further exploitation of angular correlations of b -jets and τ objects.

Acknowledgement

This work has been partly supported by the Future Circular Collider Innovation Study (FCCIS) project, that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant No 951754. M.D’O., S.V. and L.W. were partly funded by the European Union’s CHIST-ERA programme under grant agreement CHIST-ERA19-XAI-009 (MUCCA). The authors would like to thank Alessio Devoto, Donatella Genovese and Simone Scardapane, collaborators within the MUCCA project, for their invaluable contributions in the setting of the GNN approach.

References

- [1] G. Aad, et al., Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B* **716**, 1–29 (2012). <https://doi.org/10.1016/j.physletb.2012.08.020>. [arXiv:1207.7214](https://arxiv.org/abs/1207.7214) [hep-ex]
- [2] S. Chatrchyan, et al., Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B* **716**, 30–61 (2012). <https://arxiv.org/abs/1207.3238>

- [//doi.org/10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021). arXiv:1207.7235 [hep-ex]
- [3] G. Aad, et al., The ATLAS Experiment at the CERN Large Hadron Collider. JINST **3**, S08003 (2008). <https://doi.org/10.1088/1748-0221/3/08/S08003>
 - [4] S. Chatrchyan, et al., The CMS Experiment at the CERN LHC. JINST **3**, S08004 (2008). <https://doi.org/10.1088/1748-0221/3/08/S08004>
 - [5] LHC Machine. JINST **3**, S08001 (2008). <https://doi.org/10.1088/1748-0221/3/08/S08001>
 - [6] A. Abada, et al., FCC-hh: The Hadron Collider: Future Circular Collider Conceptual Design Report Volume 3. Eur. Phys. J. ST **228**(4), 755–1107 (2019). <https://doi.org/10.1140/epjst/e2019-900087-0>
 - [7] A. Abada, et al., FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1. Eur. Phys. J. C **79**, 474 (2019). <https://doi.org/10.1140/epjc/s10052-019-6904-3>
 - [8] M.L. Mangano, G. Ortona, M. Selvaggi, Measuring the Higgs self-coupling via Higgs-pair production at a 100 TeV p-p collider. Eur. Phys. J. C **80**(11), 1030 (2020). <https://doi.org/10.1140/epjc/s10052-020-08595-3>. arXiv:2004.03505 [hep-ph]
 - [9] M. Selvaggi, Higgs measurements at the FCC-hh. PoS **ICHEP2018**, 684 (2019). <https://doi.org/10.22323/1.340.0684>
 - [10] M. Grazzini, G. Heinrich, S. Jones, S. Kallweit, M. Kerner, J.M. Lindert, J. Mazzitelli, Higgs boson pair production at NNLO with top quark mass effects. JHEP **05**, 059 (2018). [https://doi.org/10.1007/JHEP05\(2018\)059](https://doi.org/10.1007/JHEP05(2018)059). arXiv:1803.02463 [hep-ph]
 - [11] G. Aad, et al., Search for the nonresonant production of Higgs boson pairs via gluon fusion and vector-boson fusion in the $b\bar{b}\tau^+\tau^-$ final state in proton-proton collisions at $\sqrt{s}=13$ TeV with the ATLAS detector. Phys. Rev. D **110**(3), 032012 (2024). <https://doi.org/10.1103/PhysRevD.110.032012>. arXiv:2404.12660 [hep-ex]
 - [12] *2020 Update of the European Strategy for Particle Physics* (CERN Council, Geneva, 2020). <https://doi.org/10.17181/ESU2020>
 - [13] M. Aleksa, et al., Conceptual design of an experiment at the FCC-hh, a future 100 TeV hadron collider **2/2022** (2022). <https://doi.org/10.23731/CYRM-2022-002>
 - [14] J. Alwall, et al., A Standard format for Les Houches event files. Comput. Phys. Commun. **176**, 300–304 (2007). <https://doi.org/10.1016/j.cpc.2006.11.010>. arXiv:hep-ph/0609017

- [15] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.S. Shao, T. Stelzer, P. Torrielli, M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP* **07**, 079 (2014). [https://doi.org/10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079). [arXiv:1405.0301](https://arxiv.org/abs/1405.0301) [hep-ph]
- [16] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.S. Shao, M. Zaro, The automation of next-to-leading order electroweak calculations. *JHEP* **07**, 185 (2018). [https://doi.org/10.1007/JHEP11\(2021\)085](https://doi.org/10.1007/JHEP11(2021)085). [Erratum: *JHEP* **11**, 085 (2021)]. [arXiv:1804.10017](https://arxiv.org/abs/1804.10017) [hep-ph]
- [17] P. Nason, A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP* **11**, 040 (2004). <https://doi.org/10.1088/1126-6708/2004/11/040>. [arXiv:hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146)
- [18] S. Frixione, P. Nason, C. Oleari, Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP* **11**, 070 (2007). <https://doi.org/10.1088/1126-6708/2007/11/070>. [arXiv:0709.2092](https://arxiv.org/abs/0709.2092) [hep-ph]
- [19] S. Alioli, P. Nason, C. Oleari, E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP* **06**, 043 (2010). [https://doi.org/10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043). [arXiv:1002.2581](https://arxiv.org/abs/1002.2581) [hep-ph]
- [20] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, M. Selvaggi, 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](https://doi.org/10.1007/JHEP02(2014)057). [arXiv:1307.6346](https://arxiv.org/abs/1307.6346) [hep-ex]
- [21] T. Sjostrand, S. Mrenna, P.Z. Skands, A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.* **178**, 852–867 (2008). <https://doi.org/10.1016/j.cpc.2008.01.036>. [arXiv:0710.3820](https://arxiv.org/abs/0710.3820) [hep-ph]
- [22] B. Francois, G. Ganis, *The FCC software for PED studies* (2024). <https://doi.org/10.17181/8k0c4-nkr70>
- [23] F. Gaede, T. Madlener, P. Declara Fernandez, G. Ganis, B. Hegner, C. Helsens, A. Sailer, G. A. Stewart, V. Völkl, EDM4hep - a common event data model for HEP experiments. *PoS ICHEP2022*, 1237 (2022). <https://doi.org/10.22323/1.414.1237>
- [24] M. Selvaggi, A Delphes parameterisation of the FCC-hh detector. Tech. rep., CERN, Geneva (2020). URL <https://cds.cern.ch/record/2717698>
- [25] M. Cacciari, G.P. Salam, G. Soyez, The anti- k_t jet clustering algorithm. *JHEP* **04**, 063 (2008). <https://doi.org/10.1088/1126-6708/2008/04/063>. [arXiv:0802.1189](https://arxiv.org/abs/0802.1189) [hep-ph]

- [26] A. Collaboration, Evidence for the higgs-boson yukawa coupling to tau leptons with the atlas detector. *Journal of High Energy Physics* **2015**(4). [https://doi.org/10.1007/jhep04\(2015\)117](https://doi.org/10.1007/jhep04(2015)117). URL [http://dx.doi.org/10.1007/JHEP04\(2015\)117](http://dx.doi.org/10.1007/JHEP04(2015)117)
- [27] G. Aad, et al., Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector. *JHEP* **07**, 040 (2023). [https://doi.org/10.1007/JHEP07\(2023\)040](https://doi.org/10.1007/JHEP07(2023)040). [arXiv:2209.10910](https://arxiv.org/abs/2209.10910) [hep-ex]
- [28] D. Genovese, A. Sgroi, A. Devoto, S. Valentine, L. Wood, C. Sebastiani, S. Giagu, M. D’Onofrio, S. Scardapane, Mixture-of-Experts Graph Transformers for Interpretable Particle Collision Detection (2025). [arXiv:2501.03432](https://arxiv.org/abs/2501.03432) [cs.LG]