



CLICdp-Note-2015-001  
11 April 2016

# Measurement of the $H \rightarrow WW^*$ Branching Ratio at 1.4TeV With CLIC ILD Using the $WW^* \rightarrow qq\ell\nu$ Final State

A. Winter<sup>\*</sup>, N. Watson<sup>\*</sup>

On behalf of the CLICdp collaboration

<sup>\*</sup> *University of Birmingham, United Kingdom*

## Abstract

This note summarises a study to evaluate the potential to measure the  $H \rightarrow WW^*$  branching fraction at CLIC, at 1.4TeV centre-of-mass energy, using the  $WW \rightarrow qq\ell\nu$  channel. We will discuss the approach to reconstructing the final state and the methods used to separate signal events from backgrounds. The expected statistical precision on the product branching fraction,  $\delta(\sigma_{H\nu\nu} \times \text{BR}(H \rightarrow WW^*))$ , is found to be 1.3% for an integrated luminosity of  $1.5\text{ab}^{-1}$ .

*This work was carried out in the framework of the CLICdp collaboration*

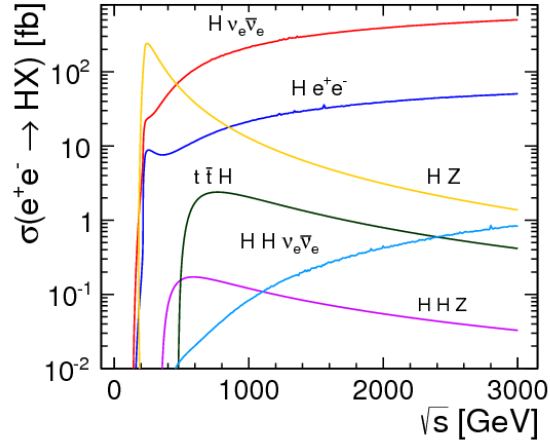


Figure 1: Cross sections for Higgs production mechanisms in an  $e^+e^-$  collider as a function of centre-of-mass energy [3]. For energies above 500GeV, Higgs production is dominated by the WW-fusion process(red.)

## 1 Introduction

One of the key aims for the CLIC experiment will be to perform a model independent measurement of the Higgs total width. Any deviations from the value predicted by the Standard Model for this would be clear evidence that there is new physics beyond the Standard Model involving particles that can interact with the Higgs. The size of any deviations could also give us an indication of the scale at which new physics occurs helping to guide the design of future high energy colliders. No current experiment has the capability of performing this measurement and so it is essential that it is measured at an electron-positron collider. The determination of the Higgs total width is dependent on the measurement of four quantities e.g. as described in Ref.[1]. Here we will consider the measurement of one of these quantities,  $\sigma_{H\nu\nu} \times \text{Br}(H \rightarrow WW^*)$ . As can be seen in Fig.1, this measurement is made possible during the 1.4TeV CLIC run as the cross section for the WW fusion process is the dominant Higgs production mechanism at this energy. This observable has already been studied at 1.4TeV using the  $WW \rightarrow qqqq$  channel [2], yielding an expected statistical precision of 1.4% for the nominal integrated luminosity of  $1.5\text{ab}^{-1}$ . Here we will look at the complementary  $WW \rightarrow qq\nu l$  channel (shown in Fig.2), where  $l=e,\mu$ , with the intention of combining our results with the existing measurement to estimate the overall precision achievable at CLIC.

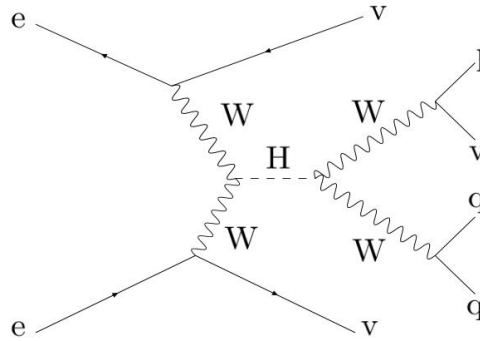


Figure 2: Higgs production via WW fusion with the Higgs decaying to a WW pair, which then decays semileptonically.

Process	Cross Section(fb)	Production ID[5]	Events Used
Signal: $ee \rightarrow H(WW^* \rightarrow qq\ell\nu)\nu\nu$	18.9	2022 <sup>i</sup>	70000
$ee \rightarrow H(WW^* \rightarrow qq\bar{q}q)\nu\nu$	25.6	2022 <sup>i</sup>	140000
$ee \rightarrow H^* \rightarrow \text{Other}$	199.6	2022 <sup>i</sup>	750000
$ee \rightarrow qq$	4009.5	2091	500000
$ee \rightarrow qq\bar{q}q$	1328.1	2163	300000
$e\gamma \rightarrow eqq$ ( $\gamma$ from EPA)	32308	2515 <sup>ii</sup>	500000
$\gamma e \rightarrow eqq$ ( $\gamma$ from BS)	56043	2527 <sup>ii</sup>	500000
$ee \rightarrow qq\nu\nu$	787.7	3243	500000
$ee \rightarrow qq\ell\ell$	2725.8	3246	400000
$ee \rightarrow qq\ell\nu$	4309.7	3249	1000000

Figure 3: Samples used for the analysis

The signal and background processes we have examined in our analysis, along with their cross sections and sample production IDs, are summarised in Table. 3. In all cases the detector model used is CLIC\_ILD\_CDR, a variant of the ILD detector designed for ILC [4]. The main backgrounds of note are: the  $ee \rightarrow qq\ell\nu$  process (dominated by  $e^+e^- \rightarrow W^+W^-$ ) as it has a very similar topology to our signal process and so is expected to be the most difficult to exclude; and the  $ee \rightarrow H(WW^* \rightarrow qq\bar{q}q)\nu\nu$  process as our ability to correctly classify these events will determine how easily our results can be combined with measurements using this channel.

## 2 Event Reconstruction

Reconstruction of the signal events was performed using ILCSoft v01-17-06 and was carried out in two main stages as described below. The first stage was to identify the isolated lepton associated with the leptonic W boson decay. The second stage involved removing this isolated lepton and resolving the remaining particles into two jets that were associated with the two quarks produced by the hadronically decaying W boson. Using the two jets, the W boson could then be reconstructed and combined with the isolated lepton to reconstruct the Higgs boson. The Higgs candidate we reconstruct will not be complete due to the lepton neutrino produced from the W decay. However, the observed properties will still be sufficient for discriminating between signal and background events.

### 2.1 Lepton Identification

Two different methods were used for identifying leptons. Our primary method for particle identification is to assume that the highest energy electron or muon (as identified by PandoraPFA [7]) corresponds to the isolated lepton from the leptonically decaying W boson. This method was found to have an efficiency and purity of 93% and 96% for identifying the isolated lepton.

<sup>i</sup>Sample 2022 is the  $ee \rightarrow H\nu\nu$  inclusive sample. The relevant events were extracted from this main sample

<sup>ii</sup>The cross-section for these events were scaled by a factor of 2 to account for interactions occurring with both the electron and positron. In the case of beamsstrahlung events simulated using GUINEA-PIG [6], a further scaling of 0.75 was applied to account for the lower luminosity of  $e\gamma$  events

The second method used a series of cuts to select the isolated lepton. The first stage of this was to group the particles in the event into four jets. This was done using FastJet [8] to implement the kt-algorithm using the E-scheme for recombination with an R-parameter of 0.4. We then required that the energy of the isolated lepton (electron or muon) constituted more than 35% of the visible energy of the jet it was contained within. For electrons it was then required that at least 90% of the total energy of the particle was deposited in the ECAL, and the ratio of energy to momentum for the particle was between 0.75 and 1.25. For muons it was required that less than 35% of the total energy of the particle was deposited in the ECAL, and the ratio of energy to momentum should be between 0.01 and 0.60. This method yielded an efficiency of 91% and a purity 74%. Although this approach is not as performant as the first method, it allows more than one lepton to be selected. As a result it is useful for discriminating between signal and background processes as requirements can be placed on the number of leptons identified by this selection.

In summary, the first method is used to select a single isolated lepton, which is then used for reconstruction, while the number of lepton candidates selected by the second method is used as a discriminating variable to distinguish between signal and background processes.

## 2.2 Jet Finding

Following the lepton finding, the remaining particles (not including the isolated lepton) are forced into two jets to reconstruct the properties of the two quarks produced from the hadronic W decay. This was carried out using the same jet finding algorithm as was used for the lepton identification. The optimization of the R-parameter was performed by using Monte Carlo information to obtain what mass we would measure for the reconstructed Higgs for various values of R, when using Monte Carlo truth kinematic information of the lepton neutrino in our reconstruction. An acceptably small bias in the reconstructed mass was found for an R value of 0.4, indicating we were successfully reconstructing the quark pair.

## 3 Flavour Tagging

Flavour tagging of events was performed using LCFIPlus v00-05-02 [9]. Three neural nets were trained to identify u/d/s, b and c quarks respectively with 50,000  $ee \rightarrow Z\nu\nu$ ,  $Z \rightarrow qq$  events used for each neural

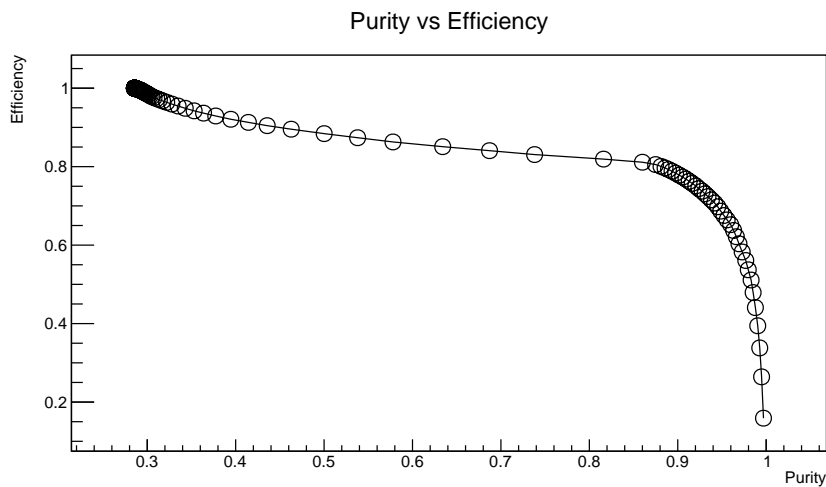


Figure 4: Purity vs efficiency for identifying b-jets, obtained from a sample of  $Z \rightarrow$  light, c and b quark events simulated at  $\sqrt{s} = 1.4\text{TeV}$

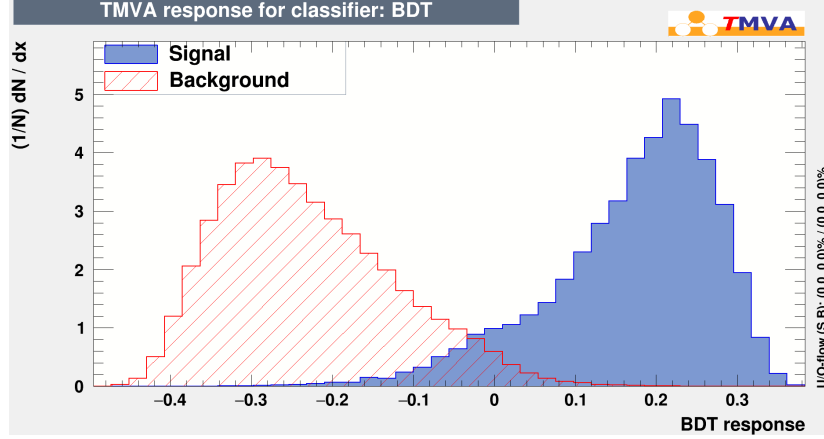


Figure 5: BDT response for signal and background events after TMVA classification

net. Application of these neural nets returned two parameters for jets within the event that quantify the probability of the jet being either a b-jet or c-jet. For this analysis, identifying b-jets is more useful for discriminating against the relevant backgrounds. Performance of the b-tagging was evaluated by applying the neural nets to a sample of 150,000 events containing an equal number of  $Z \rightarrow$  light, c and b quarks. It can be seen from Fig. 4 that a purity of 90% can be achieved while still retaining an efficiency of 80%.

## 4 Event Selection

Event selection was performed using the TMVA package [10] to produce a Boosted Decision Tree (BDT). The BDT used  $7 \times 10^4$  signal events and  $4 \times 10^6$  background events, split evenly between training and testing samples. A collection of 19 variables is used for the training: mass of the reconstructed Higgs and W bosons; energy of the W boson; missing energy and transverse momentum; number of isolated leptons selected; PID of lepton; transverse momentum of lepton; angle of lepton and W boson relative to the beam axis; magnitude of minor thrust value; number of particle flow objects (PFOs) in the two jets; average angle of the two jets relative to the beam axis; kt jet resolution parameter  $y_{12}$ ; number of tightly selected PFOs in the event; angular separation of the isolated lepton and the W boson; minimum angular separation and transverse momentum of the lepton relative to either jet, and the combined b-tag value for both jets. A set of loose pre-selection cuts were also applied before the training to remove events that were clearly background. The cuts used were: energy of the W boson  $< 591 \text{ GeV}$ , Mass of the W boson  $< 231 \text{ GeV}$ , Mass of the reconstructed Higgs  $< 306 \text{ GeV}$  and  $667 \text{ GeV} < \text{total missing energy} < 1400 \text{ GeV}$ . The input signal and background distributions for every input variable after application of these cuts can be seen in the appendix, and the resulting BDT classifier output can be seen in Fig. 5.

Figure 5 shows that there is a high degree of separation achieved between signal and background events. The optimal BDT cut for maximising the signal to background ratio was determined to be at 0.21 and the effect of the pre-selection cuts and applying this BDT cut on the signal and background processes can be seen in Fig. 6. The resulting significance ( $S/\sqrt{S+B}$ ) after these cuts has been calculated to be 77 giving a statistical uncertainty of 1.3% on  $\sigma \cdot \text{Br}$  for an integrated luminosity of  $1.5 \text{ ab}^{-1}$ . This value is similar to that observed for the  $WW \rightarrow q\bar{q}q\bar{q}$  final state, as expected. By neglecting the case where the isolated lepton is a  $\tau$ , we have reduced our statistics to two thirds that of the hadronic channel which inherently limits the precision that can be achieved. However, the backgrounds for the hadronic channel are much larger, making them harder to remove which leads to a reduced precision. Looking in detail at the backgrounds after our selection, we can see that many of the backgrounds have been almost completely removed leaving only  $ee \rightarrow H(\rightarrow \text{other})\nu\nu$  and  $ee \rightarrow q\bar{q}l\nu$  as the dominant backgrounds. This

Process	Cross Section(fb)	Pre-selection Eff(%)	BDT Cut Eff(%)	Events After BDT
Signal	18.9	99.99	42.65	12091
$ee \rightarrow H(WW^* \rightarrow qq\bar{q}\bar{q})\nu\nu$	25.6	99.96	1.79	687
$ee \rightarrow H(\rightarrow \text{Other})\nu\nu$	199.6	99.62	1.26	3767
$ee \rightarrow qq$	4009.5	76.95	<0.01	155
$ee \rightarrow qq\bar{q}\bar{q}$	1328.1	36.03	<0.01	27
$e\gamma \rightarrow eqq$ ( $\gamma$ from EPA)	32308	67.00	<0.01	816
$\gamma e \rightarrow eqq$ ( $\gamma$ from BS)	56043	95.84	<0.01	764
$ee \rightarrow qq\nu\nu$	787.7	96.59	0.07	777
$ee \rightarrow qqll$	2725.8	89.75	<0.01	257
$ee \rightarrow qq\bar{l}\nu$	4309.7	66.44	0.07	4812

Figure 6: Efficiency for all processes following pre-selection and BDT response cuts and the number of events expected to satisfy these requirements, for an integrated luminosity of  $1.5\text{ab}^{-1}$ .

is to be expected as these events most closely mimic our signal, which is mainly distinguished by its large missing energy. In the case of  $H \rightarrow \text{other}$  events it was determined that 26% of the remaining events came from  $H \rightarrow \tau^+\tau^-$  processes with a further 25% from  $H \rightarrow WW^*$  processes with one or more of the Ws decaying to a  $\tau$ . As such, attempts were made to veto  $\tau$  events by rejecting events in which one or more hadronically decaying  $\tau$  was explicitly identified using a Tau Finder [11]. However, the number of  $\tau$ s misidentified in the signal channel was determined to be too high to veto the  $\tau$  events without significantly increasing the overall statistical uncertainty on  $\sigma \cdot \text{Br}$  and so  $\tau$  identification is not used in the final selection. The efficiency for selecting  $WW^* \rightarrow qq\bar{q}\bar{q}$  events in the  $WW^* \rightarrow qq\bar{l}\nu$  channel has been calculated to be 1.8% which should be sufficiently low that a straightforward combination of the uncertainties determined by both channels can be made. However, further investigation must be done to confirm this. The efficiency of the  $H \rightarrow WW \rightarrow qq\bar{l}\nu$  events in the  $H \rightarrow WW \rightarrow qq\bar{q}\bar{q}$  channel has yet to be confirmed, therefore a final combined result has not yet been performed. Where  $H \rightarrow WW$  candidates are identified by both selections, attributing them to a final state on the basis of the predicted purities, would simplify the calculation. Systematic uncertainties have not been described here but are expected to be dominated by the uncertainty on the measured  $WW \rightarrow qq\bar{l}\nu$  branching ratio of 1.1% [12].

## 5 Conclusion

In summary, we have performed a full analysis of the  $ee \rightarrow H(WW^*)\nu\nu$ ,  $WW^* \rightarrow qq\bar{l}\nu$  decay channel using a large set of backgrounds with the aim of measuring the  $H \rightarrow WW^*$  branching ratio. A 19 variable BDT was used to select signal events where the final state charged lepton is either an electron or a muon, and to remove background which was found to be dominated by  $ee \rightarrow H(\rightarrow \text{Other})\nu\nu$  and  $ee \rightarrow qq\bar{l}\nu$  in the final selection. The resulting statistical uncertainty was found to be:

$$\delta\sigma_{H\nu\nu} \times \text{BR}(H \rightarrow WW^*) = 1.3\%$$

The efficiency for incorrectly selecting  $ee \rightarrow H(WW^*)\nu\nu$ , with  $WW^* \rightarrow qq\bar{q}\bar{q}$ , in the  $WW^* \rightarrow qq\bar{l}\nu$  channel, was found to be 1.8%. The correlated overlap in selections developed for the  $WW^* \rightarrow qq\bar{q}\bar{q}$  and  $WW^* \rightarrow qq\bar{l}\nu$  final states would be taken into account when combining the individual results.

## 6 Appendix

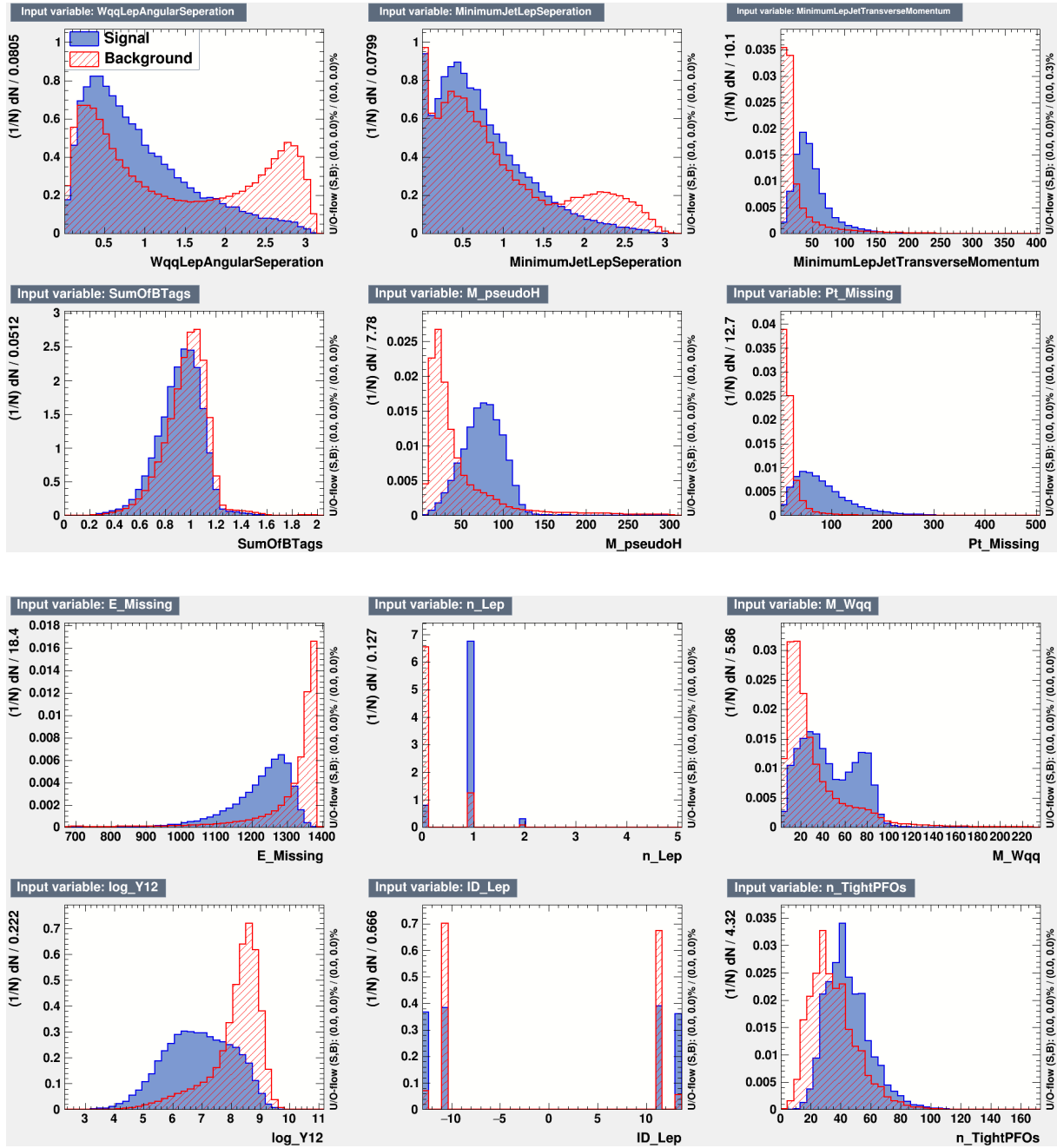


Figure 7: Input variable distributions for training of BDT

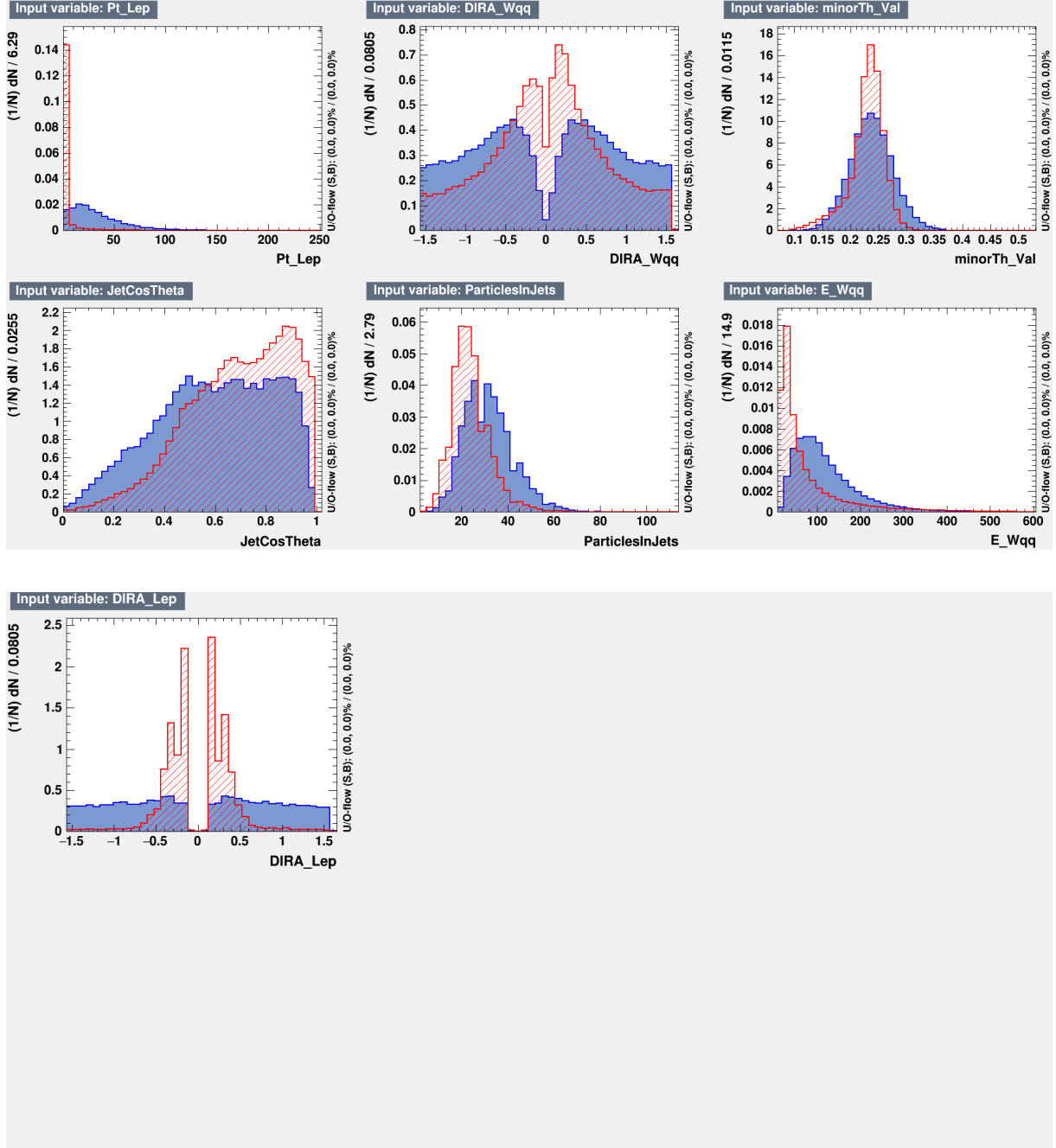


Figure 8: Input variable distributions for training of BDT



## References

- [1] C. Durig et al.,  
*Model Independent Determination of HWW coupling and Higgs total width at ILC*, International Workshop on Future Linear Colliders (LCWS13) Tokyo, Japan, November 11-15, 2013, 2014, arXiv: 1403.7734 [hep-ex], URL: <http://inspirehep.net/record/1287921/files/arXiv:1403.7734.pdf>.
- [2] M. Pandurovic, M. A. Thomson, *Measurement of  $BR(H \rightarrow WW^*)$  in Higgsstrahlung at  $\sqrt{s} = 350$  GeV and WW fusion at  $\sqrt{s} = 1.4$  TeV CLIC*, CLICdp-Draft-2015-009 (in preparation), URL: <https://cds.cern.ch/record/2011508>.
- [3] F. Simon, *Higgs Physics at future Linear Colliders - A Case for precise Vertexing*, PoS **Vertex2013** (2014) 019, arXiv: 1401.6302 [hep-ex].
- [4] P. Lebrun et al., *The CLIC Programme: Towards a Staged  $e+e^-$  Linear Collider Exploring the Terascale : CLIC Conceptual Design Report*, arXiv:1209.2543 [physics.ins-det] (2012), arXiv: 1209.2543 [physics.ins-det].
- [5] *Monte Carlo Samples for CLIC Higgs Studies*, URL: <https://twiki.cern.ch/twiki/bin/view/CLIC/MonteCarloSamplesForTheHiggsPaper>.
- [6] D. Schulte, *Beam-Beam Simulations with GUINEA-PIG* (1999), URL: <https://cds.cern.ch/record/382453>.
- [7] M. Thomson, *Particle flow calorimetry and the PandoraPFA algorithm*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **611** (2009) 25, ISSN: 0168-9002, DOI: <http://dx.doi.org/10.1016/j.nima.2009.09.009>, URL: <http://www.sciencedirect.com/science/article/pii/S0168900209017264>.
- [8] M. Cacciari, G. P. Salam, G. Soyez, *FastJet User Manual*, Eur. Phys. J. **C72** (2012) 1896, DOI: [10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2), arXiv: 1111.6097 [hep-ph].
- [9] T. Suehara, T. Tanabe, *LCFIPlus: A Framework for Jet Analysis in Linear Collider Studies*, Nucl. Instrum. Meth. **A808** (2016) 109, DOI: [10.1016/j.nima.2015.11.054](https://doi.org/10.1016/j.nima.2015.11.054), arXiv: 1506.08371 [physics.ins-det].
- [10] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, ArXiv Physics e-prints (2007), eprint: [physics/0703039](https://arxiv.org/abs/physics/0703039).
- [11] A. Muennich, *TauFinder: A Reconstruction Algorithm for Tau Leptons at Linear Colliders* (2010), URL: <http://cds.cern.ch/record/1443551>.
- [12] K. A. Olive et al., Particle Data Group, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001, DOI: [10.1088/1674-1137/38/9/090001](https://doi.org/10.1088/1674-1137/38/9/090001).