GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Bachelor's Thesis

Untersuchung quantenmechanischer Signaturen in $H \to WW^* \to l\nu l\nu$ Endzuständen mit ATLAS bei $\sqrt{s} = 13$ TeV

Probing Quantum Signatures in $H \to WW^* \to l\nu l\nu$ Final States with ATLAS at $\sqrt{s} = 13$ TeV

prepared by

Neele Bäumann

from Neumünster

at the II. Physikalischen Institut

Thesis number:	II.Physik-UniGö-BSc-2025/05
Thesis period:	31st March 2025 until 4th July 2025
First referee:	Prof. Dr. Arnulf Quadt
Second referee:	Prof. Dr. Jörn Große-Knetter

Zusammenfassung

Das Standardmodell der Teilchenphysik sagt voraus, dass die W Bosonen, die aus einem Higgs Zerfall stammen, quantenverschränkt sind. Für die Untersuchung der Verschränkung ist eine präzise Rekonstruktion quantensensitiver Observablen erforderlich, um die Anwendung der Quantentomographie zu ermöglichen. Es werden dileptonische Endzustände mit einem Elektron und einem Muon betrachtet, um den Zerfall mit hoher Präzision zu analysieren. Die in diesem Endzustand auftretenden Neutrinos werden nicht vom Detektor gemessen und müssen rekonstruiert werden.

Diese Arbeit konzentriert sich auf die Rekonstruktion der Neutrinos in diesem Prozess. Drei Rekonstruktionsalgorithmen werden eingesetzt und hinsichtlich ihrer Genauigkeit verglichen. Unter Berücksichtigung der für die Rekonstruktion erforderlichen Randbedingungen und Annahmen zeigen alle Algorithmen Potenzial. Besonders vielversprechend ist dabei der Ansatz, über einen Wertebereich von η und $M_{\nu\nu}$ zu iterieren. Zur Trennung von Signal- und Untergrundprozessen wird ein mehrlagiges neuronales Netzwerk für Mehrklassenklassifikation eingesetzt.

Abstract

The Standard Model of Particle Physics predicts that the W bosons originating from a Higgs decay are entangled. To study the entanglement, a precise reconstruction of quantum-sensitive observables towards an application of quantum tomography must be enabled. A dileptonic final state including one electron and one muon is considered, to analyse the decay with high precision. The neutrinos present in this final state are not measured by the detector and must be reconstructed.

This work focuses on the reconstruction of the neutrinos in this process. Three reconstruction algorithms are employed and compared regarding their accuracy. Taking into account the number of constraints and assumptions necessary for the reconstruction, all algorithms show potential. Especially the approach of scanning over a range of values for η and $M_{\nu\nu}$ is promising. For the separation of signal and background processes, a deep neural network designed for multiclass classification is employed.

Contents

1. Introduction **1**

2. The Standard Model of Particle Physics **3**

2.1. The Decay $H \to WW^* \to l\nu l\nu$ 5

2.1.1. Higgs Production and Decay 5

2.1.2. W boson Decays 5

2.1.3. Helicity Configurations of the Decay 6

2.2. Quantum Entanglement 6

3. Experimental Setup **9**

3.1. Large Hadron Collider 9

3.2. ATLAS Experiment 9

4. Event Generation **13**

4.1. Monte Carlo Event Generation 13

4.2. Detector Simulation 14

5. Object Definition and Preselection **15**

5.1. Object Definition 15

5.2. Preselection 16

6. Event Classification **19**

6.1. Deep Neural Networks 19

6.2. Multiclass Classification 20

6.3. Square Cut Method 22

7. Event Reconstruction **27**

7.1. Reconstruction Algorithms 27

7.2. Evaluation of the Reconstruction Algorithms 31

7.2.1. Evaluation of the Reconstructed Neutrino Kinematics 34

8. Conclusion and Outlook **41**

A. MC Event Generation **43**

B. DNN Output Distributions of the Individual Folds **45**

C. Additional Square Cut Results **47**

D. Selection Criteria **49**

E. Additional Derivations for the Reco Algorithm **51**

F. Evaluation of the Reconstruction Algorithms **53**

1. Introduction

One hundred years ago, Göttingen marked the birthplace of quantum mechanics as Born, Heisenberg and Jordan published their paper "Zur Quantenmechanik. II." [1]. Quantum mechanics describes many phenomena that appear to be in conflict with our everyday experiences. One of them is quantum entanglement. This was first introduced in the EPR paper in 1935, where Einstein, Podolsky, and Rosen argue that quantum mechanics must be incomplete, as such a distant influence contradicts the principle of locality [2]. In 1964, in his paper "On the Einstein-Podolsky-Rosen paradox" John S. Bell introduced his famous inequality, providing a criterion to test whether quantum mechanical phenomena are consistent with a local and realistic theory [3]. Since then, numerous Bell tests have been successfully conducted to verify whether quantum mechanical phenomena indeed violate the Bell inequalities [4–6].

Recently, efforts have begun to investigate quantum entanglement in high-energy physics [7]. This opens the possibility to test quantum mechanics at the highest energy scales and with fermions and bosons. It has been proposed that quantum entanglement could be measured in the $H \to WW^*$ decay [8]. This thesis explores various analysis strategies aimed at studying quantum signatures for this process, using the experimental collisions provided by the LHC during its Run 2 and taken with the ATLAS experiment. To study the process with high precision, leptonic final states are considered. The primary challenge lies in accurately reconstructing the dineutrino system, as neutrinos are not directly measured by the ATLAS detector. To suppress background contributions, machine learning algorithms are employed. The goal is to establish a foundation for the reconstruction of observables towards an application of quantum tomography and, eventually, the study of Bell inequalities in the process $H \to WW^* \to l\nu l\nu$.

This thesis first presents the theoretical and experimental foundations of the topic. The event generation as well as object definition and preselection are described in Chapters 4 and 5. The following chapter explains and analyses the event classification using machine learning techniques. Chapter 7 introduces three algorithms to reconstruct the neutrino system and compares their performance. Finally, the main results are summarised and an outlook for further analyses towards probing quantum signatures in this process is given.

2. The Standard Model of Particle Physics

All known particles are described in the Standard Model of Particle Physics [9]. An overview of these particles is shown in Figure 2.1. The Standard Model also describes three of the known fundamental forces, which are the strong, the weak and the electromagnetic force. Gravity is excluded. The particles are divided into two groups: fermions having half-integer spin and bosons having integer spin. There are twelve fermions: six leptons and six quarks.

The three charged leptons, electron (e), muon (μ), and tau (τ), have an electromagnetic charge of -1. They interact electromagnetically, the corresponding exchange particle is the photon which carries no electric charge and is massless. The leptons are divided into three generations, each containing one electrically charged and one electrically neutral lepton. The mass of the charged leptons rises with the generation. For each charged lepton there is a corresponding neutrino with which it forms a weak isospin doublet. The neutrinos carry no charge and the upper bound of their mass is very small.

The six quarks are also divided into three generation. Each generation contains an up-type quark (up, charm, top) and a down-type quark (down, strange, bottom). For each generation they form a weak isospin doublet. The up-type quarks carry a charge of $+\frac{2}{3}e$, the down-type quarks carry a charge of $-\frac{1}{3}e$. Similar to the leptons, the mass of the quarks rises with the generation. The top quark with a mass of $m_t \approx 172.6$ GeV is the heaviest known elementary particle [9]. Quarks also carry a colour charge which can be blue, red or green. Antiquarks carry the corresponding anticolour, antiblue, antired or antigreen. Since quarks carry colour charge they engage in the strong interaction. The force carrier, the gluon, carries colour and anticolour. Since quarks carry an electric charge they also can interact electromagnetically.

The third force described by the Standard Model is the weak force. Its force carriers are the massive, electrically neutral Z boson, the electrically positively charged W^+ and the electrically negatively charged W^- boson. The W bosons are the exchange particles for the charged-current interaction. Only left-handed particles and right-handed antiparticles

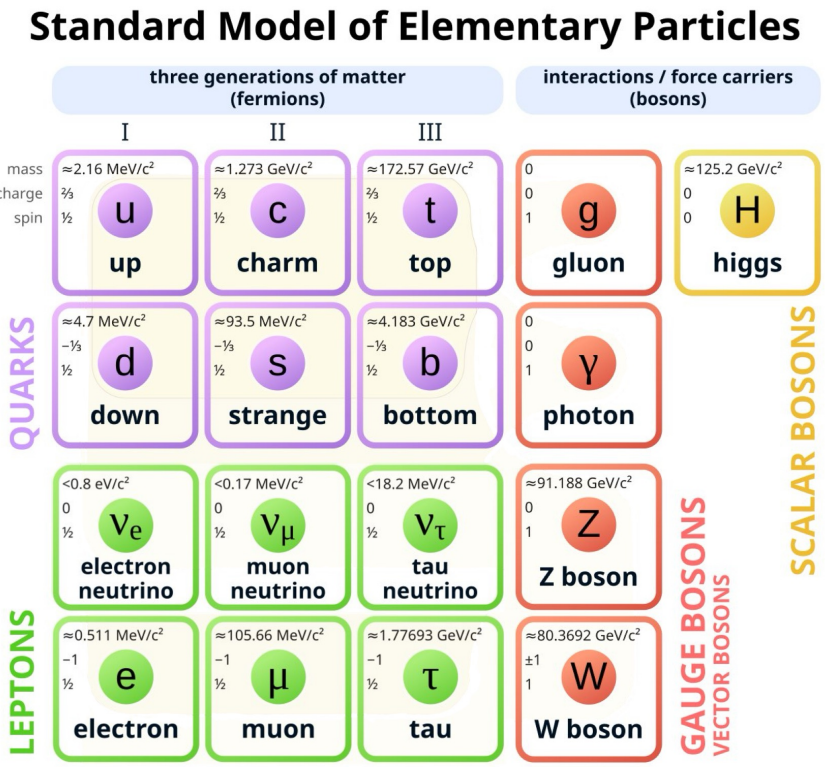

Figure 2.1.: The figure shows the elementary particles described by the Standard Model. The particles' masses, spin and charge are shown [9].

take part in this interaction. An up-type lepton interacts with its corresponding down-type lepton and vice versa. For quarks, the interaction partners can be any up- and down-type quark. The probability for the quarks of their respective generations to interact with each other are determined in the Cabibbo-Kobayashi-Maskawa matrix (CKM matrix) [10].

At very high energies the electromagnetic and the weak force merge into a single force. Therefore they are unified in the electroweak theory. This leads to a symmetry breaking because photons are massless whereas the weak exchange particles are massive. This is caused by the Higgs mechanism [11]. The Standard Model includes the Higgs field which is a scalar field and via the Higgs mechanism gives massive fermions and bosons its mass. The Higgs boson is a quantum excitation of this field. It is a scalar boson, has spin 0 and no charge.

Table 2.1.: The cross sections for several Higgs boson production channels at a centre-of-mass energy of $\sqrt{s} = 13$ TeV are listed [9].

process	ggF	VBF	WH	ZH	$t\bar{t}$H	total
cross section (in pb)	$48.6^{+5.6\%}_{-7.4\%}$	$3.78^{+2.1\%}_{-2.1\%}$	$1.37^{+2.0\%}_{-2.0\%}$	$0.88^{+4.1\%}_{-3.5\%}$	$0.50^{+6.8\%}_{-9.9\%}$	$55.1^{+5\%}_{-7\%}$

2.1. The Decay $H \to WW^* \to l\nu l\nu$

For this thesis, the decay of interest is $H \to WW^* \to l\nu l\nu$. In the following subsections, Higgs boson production at a centre-of-mass energy of $\sqrt{s} = 13$ TeV and the considered decay into two W bosons are discussed. Furthermore, the considered W boson decays are explained.

2.1.1. Higgs Production and Decay

There are several ways a Higgs boson can be produced. The most relevant production channels for a centre-of-mass energy of $\sqrt{s} = 13$ TeV are discussed in the following. The most common one is gluon fusion (ggF). Two gluons exchange a virtual top quark. A top loop is formed and eventually a Higgs boson is produced. The process with the second largest cross section is vector boson fusion. Two (anti-)quarks exchange a W or Z boson which radiates a Higgs boson. Less relevant production channels are WH, ZH and $t\bar{t}$H associated mechanisms. The respective production cross sections are listed in Table 2.1.

The Higgs decay under consideration is the $H \to WW^*$ decay. Its branching ratio is 25.7(2.5)% [9]. Given that the Higgs has a mass of $m_H \approx 125$ GeV and the W boson has a mass of $m_W \approx 80$ GeV, one of the W bosons is off-shell [9]. In the following, the off-shell W boson will be referred to as W^*.

2.1.2. W boson Decays

W bosons can decay hadronically and leptonically. The considered decay is the leptonic channel because leptons can be detected with a higher sensitivity. The W boson decays into a lepton ℓ and its corresponding neutrino ν_l. The branching ration for each lepton neutrino pair is $\sim 10\%$ [9]. Tau leptons have a very short mean lifetime of $\sim 290 \cdot 10^{-15}$ s, and therefore do not reach the detector [9]. Instead, its decay products, primarily other leptons or π-mesons, are detected.

2.1.3. Helicity Configurations of the Decay

The Higgs boson is a scalar particle with spin zero [9]. In its rest frame, it possesses no orbital angular momentum. Consequently, due to the conservation of total angular momentum, the spins Z-component of the resulting W bosons, projected along a common axis, must sum to zero. In this frame, the W bosons are emitted back-to-back as a result of momentum conservation. Therefore, their helicities, that is the projection of their spins onto their respective directions of motion, has o be equal and can be explored. As spin 1 particles, W bosons can have helicity values of +1, 0, or -1. The helicity combinations consistent with angular momentum conservation are (+1, +1) and (-1, -1), corresponding to transverse polarisation and (0, 0), which corresponds to longitudinal polarisation. These combinations are allowed because the W bosons move in opposite directions in the Higgs rest frame.

In the weak decay $W \to \ell\nu$, the W boson couples exclusively to left-handed neutrinos and right-handed antineutrinos. Since leptons are spin $\frac{1}{2}$ particles, their spins must align appropriately with the W boson's polarisation to conserve angular momentum. In the case of transverse polarisation, the lepton and neutrino are preferentially emitted along the W boson's direction of motion, ensuring angular momentum conservation. This results in the two neutrinos in the decay $H \to WW^* \to l\nu l\nu$ being emitted collinearly. In contrast, when the W boson is longitudinally polarised, no particular emission direction is favoured, as the helicity projection is zero and does not impose directional constraints.

2.2. Quantum Entanglement

The Standard Model describes elementary particles as quantum mechanical entities. A phenomenon that occurs for these particles is quantum entanglement. A quantum state $\psi_A \in \mathcal{H}_\mathcal{A}$ of a system A can be written as

$$\psi_A = \sum_i a_i i_A. \tag{2.1}$$

The i_A form a basis of the Hilbert space $\mathcal{H}_\mathcal{A}$, the a_i are the corresponding coefficients. The Hilbert space of a combination of two systems A and B is the tensor product $\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$. A state ψ_{AB} in the combined Hilbert space can be written as

$$\psi_{AB} = \sum_{i,j} c_{ij} i_A j_B, \tag{2.2}$$

where j_B is a basis of $\mathcal{H}_\mathcal{B}$. If the coefficients c_{ij} factorize to coefficients for system A and B as $c_{ij} = a_i b_j$ the state is separable. If such a factorisation does not exist the states ψ_A and ψ_B are entangled. This implies that their behaviour is intrinsically correlated, despite their spatial separation, and they can no longer be described independently. The considered signature of correlation is the polarisation of two particles.

The entanglement of a combined system can also be inferred from its spin density matrix. The process of determining the spin density matrix from quantum-sensitive observables is called quantum tomography. It requires the reconstruction of the system's kinematic or state variables from the measured data, in order to compute the observables needed to infer the density matrix.

3. Experimental Setup

To study the elementary particles high energies are needed. Particle accelerators are used to collide particles at high centre-of-mass energies. The decay products of these collisions are detected in particle detectors.

3.1. Large Hadron Collider

The Large Hadron Collider (LHC) [12] at CERN near Geneva is a synchrotron that accelerates protons and ions. Electromagnetic cavities and superconducting magnets are arranged in a circular shape with a circumference of approximately 27 km. In Run 2 the two colliding proton beams had a centre-of-mass energy of $\sqrt{s} = 13$ TeV and bunch spacing of 25 ns [12]. The integrated luminosity for Run 2 is $\mathcal{L} = 140.1(1.2)$ fb^{-1} [13].

3.2. ATLAS Experiment

ATLAS is a multi-purpose particle detector at the LHC. It is barrel-shaped and consists of multiple detector layers with different purposes. An overview of the detector and its sub-detectors is shown in Figure 3.1.

The ATLAS Detector uses a right-handed cartesian coordinate system. Its origin lies in the interaction point. The x-axis points to the centre of the LHC, the y-axis points upwards and the z-axis is tangential to the beam. In the plane transversal to the beam polar coordinates r and φ are used. Furthermore the polar angle θ describes the angle between a decay particle's track and the z-axis. The pseudo-rapidity is defined as

$$\eta = -\ln \tan \frac{\theta}{2}. \tag{3.1}$$

The angular separation of two particles is defined as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}. \tag{3.2}$$

The detector and its sub-detector layers are symmetrically arranged around the beam.

The Inner Detector (ID) consists of three sub-detectors. The Pixel Detector consists of silicon pixel layers and is located closest to the interaction point. It is followed by the Semi-Conductor Tracker (SCT) which contains of many silicon microstrip layers. The outermost sub-detector is the Transition Radiation Tracker (TRT) built of drift tubes. The ID is located in a solenoidal magnetic field of 2 T which causes charged particles to curve. It measures the particle's momentum and primary and secondary vertices [14].

The ID is surrounded by the electromagnetic calorimeter that measures the energy of electrons, positrons and photons. It is a sampling calorimeter and consists of layers of lead (passive absorber) and liquid argon (active detector) [14]. In the absorber the particles form electromagnetic showers. The main processes are Bremsstrahlung and pair production. Pair production refers to the process in which a single photon generates an electron-positron pair, while Bremsstrahlung occurs when an electron or positron is deflected by another charged particle, emitting a photon. Those secondary particles produce even more particles by the same mechanisms until the particles' energy is too low. The energy of the shower particles is measured in the active layer of the detector. Even though muons and tau leptons interact electromagnetically they are not detected in the electromagnetic calorimeter. Tau leptons have a mean lifetime of $290.3(0.5) \cdot 10^{-15}$ s and, compared to the electron, a much larger mass of approximately $m_\tau \approx 1.8$ GeV [9]. Therefore they do not reach the detector and Bremsstrahlung does not occur. Muons have a much higher lifetime but also a mass that is 200 times heavier than electrons. Hence their energy loss through Bremsstrahlung is negligible.

The next layer is the hadronic calorimeter. It also is a sampling calorimeter which uses steal as absorber and plastic scintillators as active layer [14]. The energy of hadronic decay products is measured.

The only particles beside neutrinos that have not been stopped yet are muons. They are detected in the outermost detector, the Muon Spectrometer. It is located in a system of three superconducting air-core toroids which bend the muon tracks. Several drift and multiwire proportional chambers then measure the momentum of the muons [14].

The only particles that do not interact with the detector are neutrinos. Since the colliding particles have no transversal momentum neutrinos can be reconstructed with the help of missing transverse momentum or energy $E_{\mathrm{t}}^{\mathrm{miss}}$.

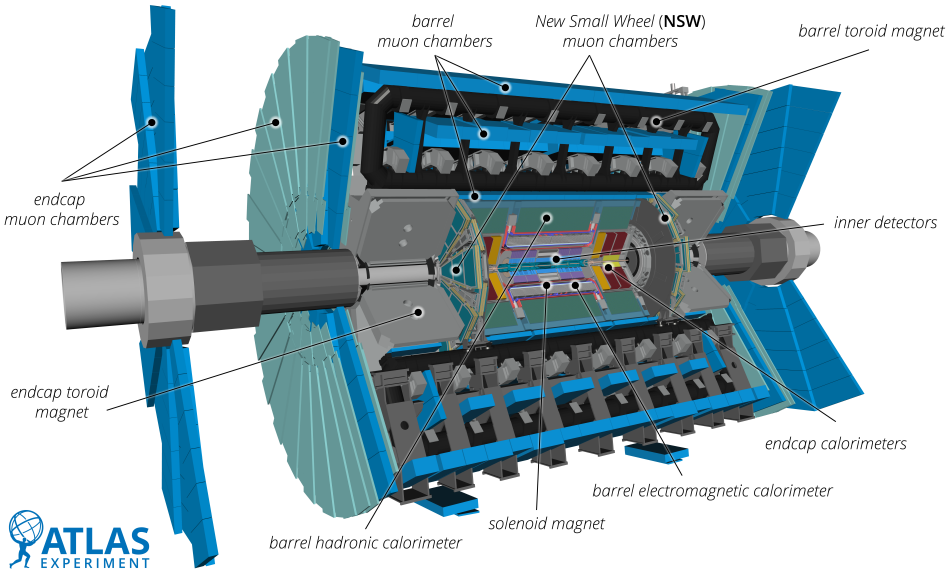

Figure 3.1.: The picture shows a schematic overview of the ATLAS detector and its sub-detectors, © Cern.

4. Event Generation

Particle Physics aims to describe and understand the fundamental properties and interactions of elementary particles. To develop strategies and algorithms to precisely reconstruct the measured data, which is then compared with theoretical predictions, simulations are employed. They enable a direct comparison between reconstructed observables and their corresponding values at truth-level. To achieve high statistical precision, a large number of collisions is simulated. In particle physics, such a collision is referred to as an event. First, the event itself is generated using Monte Carlo event generators. Then, the resulting measurement of the detector is simulated.

4.1. Monte Carlo Event Generation

First, the event is simulated at the parton level. This stage describes the hard scattering of the collision, excluding showering processes and further decays. The simulated events follow the probability distributions predicted by the Standard Model. Monte Carlo event generators are used to randomly generate individual events according to these distributions. The particles which reach the detector are not necessarily the same as those produced in the initial hard scattering process. Therefore, the event is simulated at the particle level, which includes parton showering and further particle decays. Parton showering describes the emission of gluons and quarks from the initial partons as well as from their decay products. These partons then undergo hadronisation, a process by which they form colour-neutral hadrons. Since hadronisation cannot be calculated using perturbative methods, phenomenological models such as the Lund-String model [15] or cluster models [16] are employed to describe this process.

The Monte Carlo generators used in this analysis are Powheg [17], Sherpa [18] and Pythia [19]. The considered samples and their respective event generators are listed in Table A.1 in the appendix.

4.2. Detector Simulation

To access the information at the particle level experimentally, particle detectors are used. Their measurement accuracy is limited by factors such as resolution, acceptance and trigger efficiency. To account for these detector effects, the measurement process is simulated to obtain the reconstructed level of an event. The ATLAS detector response is simulated using Geant4 [20], which models the interaction of particle-level information with the detector and produces the corresponding detector signals.

5. Object Definition and Preselection

In order to reconstruct an event, physical objects have to be derived from the signals measured by the detector. The object definition defines, which measured signals correspond to which physical object, and is explained in Section 5.1. To study the process of interest, the events are separated into signal and background events. These preselection criteria are defined in Section 5.2.

5.1. Object Definition

For this analysis, the objects of interest are the decay products of the process $H \rightarrow WW^* \rightarrow \ell^+\nu\ell^-\overline{\nu}$. The measurable decay products of this process are electrons and muons; decays into tau leptons are not considered as signal processes. The neutrinos produced in the W boson decays are not directly detected, instead their transverse energy is inferred from the missing transverse energy $E_{\mathrm{t}}^{\mathrm{miss}}$. Furthermore, jets are taken into account, since they may be emitted prior to Higgs boson production and the Higgs recoils against them, resulting in different kinematics.

Electrons are reconstructed from their signals in the electromagnetic calorimeter and the associated tracks in the Inner Detector. Due to detector acceptance, they are required to have a transverse momentum of $p_{\mathrm{t}} \geq 15$ GeV and a pseudo-rapidity within $|\eta| \leq 2.47$. To reduce the contribution from non-prompt electrons, electrons passing a tight working point are used for this analysis. A tight likelihood-based identification criterion and a tight isolation requirement using a variable isolation radius are applied [21, 22].

Muons are identified from their signals in the Muon Spectrometer. Their trajectories also are determined from their tracks in the Inner Detector. The minimum transverse momentum of an object to be considered as a muon is $p_{\mathrm{t}} \geq 15$ GeV, their pseudo-rapidity is required to be within $|\eta| \leq 2.5$. Muons are required to pass a medium working point, which is defined by a tight isolation criterion using a variable isolation radius [23].

Jets are reconstructed using the anti-k_t algorithm [24] with a radius parameter $R = 0.4$.

They have to have a transverse momentum of $p_{\mathrm{t}} \geq 25$ GeV and a pseudo-rapidity of $|\eta| \leq 2.5$. Jets must satisfy the 'jet vertex tagger' requirements [25]. To identify bottom jets, the GN2v01 algorithm [26] is used. It is based on a graph neural network and tags b-jets at different efficiencies. For example, at the 90% working point, 90% of all true b-jets are tagged as b-jets.

Missing transverse energy is calculated considering all the objects mentioned above.

5.2. Preselection

The preselection aims to reduce background events and separate them from the signal events. It is based on reference [27]. The signal process is HWW^*, where the Higgs boson is produced either via gluon fusion (HWW_{ggf}) or vector boson fusion (HWW_{vbf}). The main background processes in this analysis are top-quark pair production ($t\bar{t}$) and diboson production. Dedicated control regions are defined for these processes. For the diboson background, production via electroweak processes and via the strong force are included. Furthermore, Wt, V+jets, $Z\tau\tau$, $V\gamma$+jets and other Higgs processes (Other H) are considered. V denotes a W or Z boson. Fake-lepton backgrounds originating from heavy-flavour ($\mathrm{F_{HF}}$), light-flavour ($\mathrm{F_{LF}}$), and other sources ($\mathrm{F_{Other}}$) are also included.

Cuts are imposed on the properties of each event to assign it to a region. For the signal region, the leading and subleading lepton must have a transverse momentum of $p_{\mathrm{t}} \geq 25$ GeV each. The final state of the signal process contains two neutrinos, therefore a missing transverse energy of $E_{\mathrm{t}}^{\mathrm{miss}} > 20$ GeV is required. To suppress $t\bar{t}$ background, events containing any b-tagged jets using a working point with 90% efficiency are rejected, and the dilepton mass must be less than 100 GeV. Furthermore, it must be greater than 10 GeV, to reduce low-mass meson resonances and Drell-Yan background. To exclude Z boson processes, which emit two charged leptons of the same flavour, exactly one electron and one muon are required. As mentioned before, the kinematic properties of the signal process differ depending on whether the Higgs boson is at rest or boosted by jets that were emitted before its production. Therefore, three signal regions are defined, accounting for the jet multiplicity. The regions are SR0j, SR1j and SR2j, having no jets, exactly one and two or more jets. The cuts defining the signal regions at the preselection level are listed in Table D.1.

The diboson background includes WZ, ZZ and WW events, which do not originate from a Higgs decay. Since the diboson background leads to final states very similar to those of the signal, the same preselection is applied to define the control region (CRDiboson). To differentiate between this background and the signal, machine learning methods are

Table 5.1.: The table lists the preselection criteria used for the signal regions.

Preselection Criteria	SR0j	SR1j	SR2j
Jet multiplicity	$N_{\text{jets}} = 0$	$N_{\text{jets}} = 1$	$N_{\text{jets}} \geq 2$
Missing transverse energy	$E_{\text{t}}^{\text{miss}} > 20$ GeV		
b-jet multiplicity	$N_{b\text{-jet, }90\%} = 0$		
Leading lepton p_T	$p_{\text{t}} \geq 25$ GeV		
Subleading lepton p_T	$p_{\text{t}} \geq 25$ GeV		
Dilepton mass	10 GeV $< m_{\ell\ell} <$ 100 GeV		
Number of electrons	$N_e = 1$		
Number of muons	$N_\mu = 1$		

Table 5.2.: The table shows the preselection criteria used for the control regions.

Preselection Criteria	CRDiboson	CR$t\bar{t}$
Missing transverse energy	$E_{\text{t}}^{\text{miss}} > 20$ GeV	
b-jet multiplicity	$N_{b\text{-jet, }90\%} = 0$	$N_{b\text{-jet, }65\%} \geq 2$
Leading lepton p_T	$p_{\text{t}} \geq 25$ GeV	
Subleading lepton p_T	$p_{\text{t}} \geq 25$ GeV	
Dilepton mass	10 GeV $< m_{\ell\ell} <$ 100 GeV	
Number of electrons	$N_e = 1$	
Number of muons	$N_\mu = 1$	
Jet multiplicity		$N_{\text{jets}} \geq 2$

employed to define an additional selection. This is described in the following section. To define the $t\bar{t}$ control region (CR$t\bar{t}$), two or more b-jets with an efficiency of 65% are required, as top quarks always decay into a bottom quark and a W boson. For the leptons, the same preselection criteria as those used in the signal region are applied. In addition to the preselection, machine learning methods are also used to improve the discrimination of the $t\bar{t}$ background. The preselection criteria of the control regions are shown in Table D.2.

6. Event Classification

To separate signal events from background events, machine learning techniques are employed. A fully connected deep neural network (DNN) is used for multiclass classification. Based on the classification scores, cuts are applied to define signal and control regions. The basics of DNNs and the used structure are explained below.

6.1. Deep Neural Networks

A DNN consists of an input layer, an output layer and several hidden layers inbetween. Each hidden layer consists of nodes that determine the capacity of the network. The number of nodes in a hidden layer is referred to as the hidden dimension, denoted by d_{hid}. The transformation from one layer of dimension d_i to the next of dimension d_{i+1} is performed by applying a linear mapping followed by a non-linear activation function. For the linear transformation, a weight matrix W of dimension $d_{i+1} \times d_i$ and a bias vector $\vec{b}$ of dimension d_{i+1} are randomly initialised. The output of the linear transformation is then passed through a non-linear activation function σ. If $\vec{x}$ is the output of the previous layer, the input for the next layer is calculated as

$$\vec{y} = \sigma(W\vec{x} + \vec{b}). \tag{6.1}$$

After the input has propagated through all layers of the network once, the output is compared to the true labels and the loss is computed via a loss function. The loss quantifies how well the DNN is able to correctly classify the input. In order to train a reliable and accurate DNN model, the goal is to minimise the loss. To achieve this, the weight matrices and biases are optimised after every epoch. Let θ denote the vector of all trainable parameters (weights and biases) of the model. To determine the influence of every model parameter on the loss, backpropagation is used to calculate the gradient of the loss with respect to all parameters. A gradient descent algorithm is used to compute the optimised model parameters, which give a minimum in the loss. They are stored in

θ'. A basic algorithm to update the parameters is

$$\theta' = \theta - \eta \nabla_\theta L, \tag{6.2}$$

where η is the learning rate. If a model is trained for too long, it specialises on the training data and is no longer able to generalise to unseen data. This phenomenon is called overfitting. To prevent this, the dataset is often split into a training set and a validation set. The model is trained on the training set, and after every epoch, the validation loss is computed. If training improves only the training loss but not the validation loss, early stopping is triggered. The patience parameter defines how many epochs without improvement in the validation loss are tolerated before early stopping is activated. The explained procedure is repeated until the maximum number of epochs is reached or early stopping is triggered.

6.2. Multiclass Classification

The DNN trained for this analysis is designed to perform a multiclass classification. Each input sample belongs to one of three classes and is assigned a probability that the sample belongs to the respective class. The classes consist of the signal process HWW_{ggf} and the two background processes $t\bar{t}$ and $\text{Diboson}_{\text{QCD}}$, as these represent the most relevant contributions. The DNN consists of three hidden layers, that have 64 nodes each. The activation function is ReLU, which is defined as

$$\text{ReLU}(x) = \max(0, x). \tag{6.3}$$

To convert the model output to probabilities, the softmax function is used. The loss is computed with the CrossEntropyLoss function, the optimiser is Adam [28]. To achieve a stable training process, the training rate is $\eta = 0.0005$. In order to prevent overfitting, the patience is 50, the number of maximal epochs is 20000. The input features include the energy, azimuthal angle, pseudo-rapidity, and transverse momentum of the leading electron and muon, as well as those of the leading and sub-leading jets. In addition, the missing transverse energy, its azimuthal angle, and its significance, defined as $S = \frac{E_{\text{t}}^{\text{miss}}}{\Sigma E_{\text{t}}}$, are taken into account. Lastly, the scalar sum of transverse energy is included. Therefore, the input layer has dimension $d_{in} = 20$. To achieve stable and fast training and prevent features with high values from dominating over others, all features are scaled to lie in the interval $[0, 1]$. For a set A containing all values of one feature, every value $a \in A$ is scaled

via

$$a' = \frac{a - a_{\min}}{a_{\max} - a_{\min}}, \tag{6.4}$$

where a' is the scaled value and $a_{\min}$ and $a_{\max}$ are the minimum and maximum of A. To evaluate the model's ability to generalise to unseen data, k-fold cross-validation is employed. This technique divides the dataset into k equally sized folds. Each of the k models is trained on $k-1$ folds, ensuring that for every fold there is a model that was not trained on it. Therefore, the entire dataset can be evaluated by models that have not seen the corresponding data during training, allowing for a reliable assessment of the models' performance. For this analysis, k is set to two. All events with an even event number are assigned to one fold, while events with an odd event number are assigned to the other. Within each fold, 20% of the data are used for validation and 80% for training. Events entering the training process must satisfy the selection criteria for one of the signal or background regions, as described in the Tables D.1 and D.2. Furthermore, they are required to contain no b-jets at an 85% efficiency working point, in order to include only background events that closely resemble the signal.

The training process is evaluated using the loss curves, which are shown in Figure 6.1. They display the training and validation loss after every epoch for each fold. The fluctuations in the validation loss are due to the smaller sample size of the validation dataset. Both training and testing loss values converge for both folds and early stopping was triggered after 63 and 161 epochs, indicating that overfitting was prevented. The models' ability to correctly classify events of different classes is assessed by analysing their output score distributions. The average output distributions across both folds are presented in Figure 6.2. For each of the three possible output scores of the models the distributions of the true classes are shown. The bottom plot, which displays the HWW output score, indicates that signal and background are well separated, both background processes are mostly assigned very low probabilities $P(HWW)$, whereas the signal process accumulates at high probability scores. The upper plots, showing the distributions for both background processes, confirm that HWW processes mostly are assigned low probabilities to be background. However, these plots also suggest that the models struggle to distinguish between the two background processes themselves. Despite this, the training behaviour observed in these plots still can be considered successful, since the main goal is to separate the signal from the background processes. The output distributions for the individual folds are shown in Appendix B.

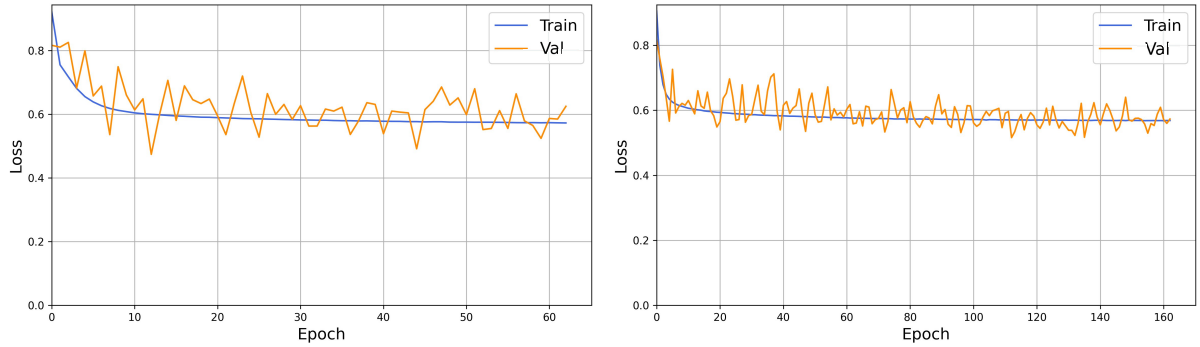

Figure 6.1.: The plots show the training and validation loss curves for both folds.

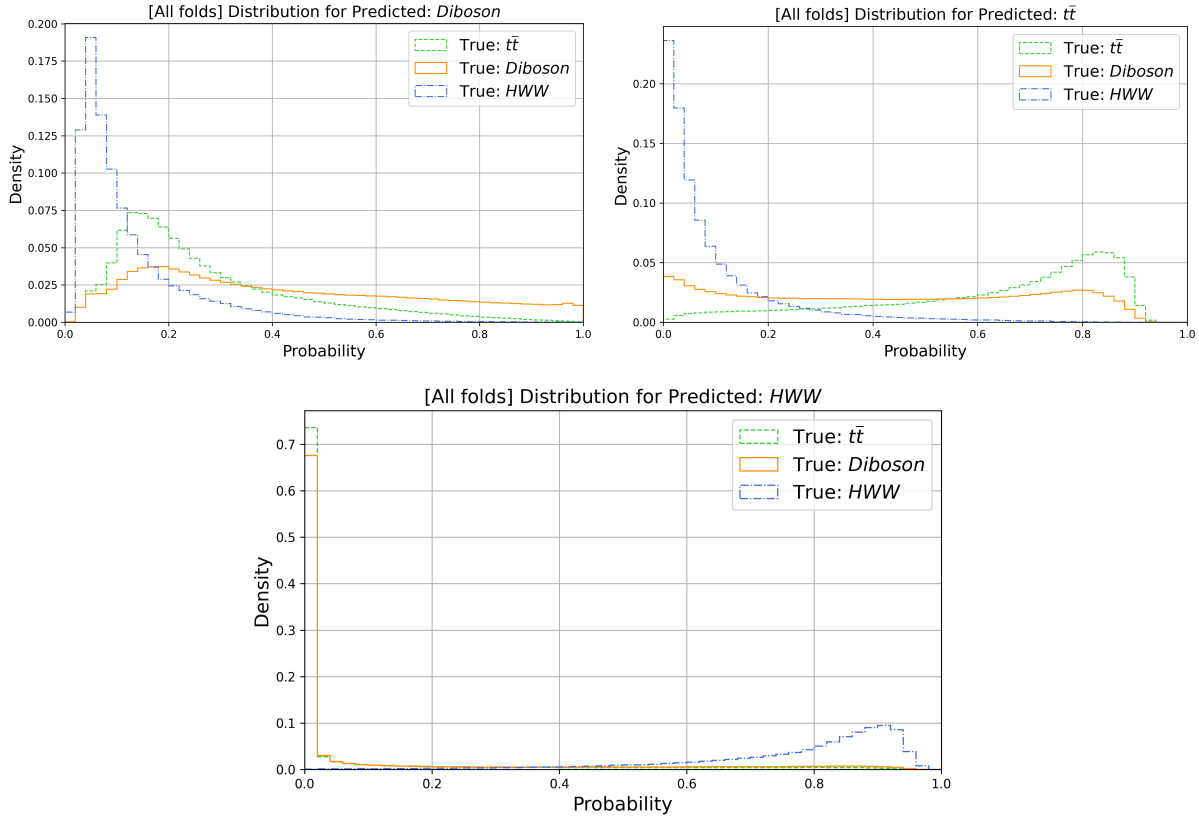

Figure 6.2.: The output score distributions for the signal score (bottom), the $t\bar{t}$ score (upper right) and the *Diboson* score (upper left) are displayed. The distributions are normalised and represent the average of both folds.

6.3. Square Cut Method

The probability scores assigned to each event are used as selection criteria to better distinguish between signal and background processes. Due to the softmax function, for every event the three scores sum up to one. Therefore, they can be represented in a two-

dimensional plane, where x- and y-axes are the $t\bar{t}$ and *Diboson* score and the HWW score is shown implicitly. Since probabilities cannot exceed one, triangular shapes are expected. For each of the three classes, these plots are created, with the results for the 0-jet region shown in Figure 6.3. A clear separation can be observed: HWW events accumulate near the origin, indicating high HWW scores, while both background processes mostly appear near the hypotenuse of the triangle. This supports the earlier statement, that the model poorly differentiates between the two background processes. Otherwise, the events would be expected to concentrate in the upper and lower left corners of the triangle. However, for all classes, especially the background, the events spread almost across the whole triangle, showing that the model sometimes confuses them with signal events and vice versa. The plots for the other jet regions can be found in the appendix in Figure C.1 and show similar distributions.

For a successful analysis, both the signal purity and its statistical significance should be maximal. To achieve this, the signal-to-background ratio, $\frac{S}{B}$, and the statistical measure $\frac{S}{\sqrt{B}}$ are maximised. To find the cuts on the probability scores, that maximise these fractions, the square cut method is employed. For every possible combination of cuts on the $t\bar{t}$ and *Diboson* scores, using a step size of 0.02, the cumulative sums of $\frac{S}{B}$ and $\frac{S}{\sqrt{B}}$ are computed over all bins within the rectangle from the origin, representing events with zero probability to be background, to the given cut point. The resulting $\frac{S}{B}$ and $\frac{S}{\sqrt{B}}$ distributions for the 0-jet region are shown in Figure 6.4. The signal-to-background ratio $\frac{S}{B}$ reaches its maximum value of approximately 0.8 near the origin, and then rapidly decreases to around 0.4. Therefore the highest value would be obtained by applying very tight cuts on both scores. However, inspection of the $\frac{S}{\sqrt{B}}$ distribution shows that such cuts would result in very low statistics. The $\frac{S}{\sqrt{B}}$ metric reaches high values slightly farther from the origin and remains high over a broader region. In this area, the $\frac{S}{B}$ ratio shows a relatively small gradient. Therefore, the cut combination that maximises $\frac{S}{\sqrt{B}}$ is chosen as the final selection criterion, as the signal purity remains reasonably high. The corresponding distributions for the other jet regions are presented in the appendix in Figure C.2 and show similar behaviour. The cut combinations maximising the respective fractions are listed in Table 6.1. For all signal regions, the cut combinations leading to the highest $\frac{S}{\sqrt{B}}$ are chosen as final selection criteria. All events having smaller or equal scores pass the selection. To enrich the background processes in the respective control regions for $t\bar{t}$ and *Diboson*, the cuts are also applied on these regions. Since events in the $t\bar{t}$ region are required to have two or more jets, the cut combination for the 2-jet region is chosen. The selected $t\bar{t}$ score is extended to divide the triangular distribution into three regions representing the 2-jet signal region near the origin, the *Diboson* control region at

Table 6.1.: The table lists the cuts on the DNN scores which maximise the signal-to-background ratio and statistical significance for different jet regions. The cuts calculated for $\frac{S}{\sqrt{B}}$ are chosen as selection criteria for the signal regions. All events with an output score less than or equal to the cut pass the selection.

Jet region	**S/B**		**S/$\sqrt{\mathbf{B}}$**	
	$t\bar{t}$ cut	*Diboson* cut	$t\bar{t}$ cut	*Diboson* cut
0 jets	0.01	0.03	0.17	0.11
1 jet	0.01	0.01	0.13	0.11
$\geq$2 jets	0.01	0.03	0.13	0.13

the top and the $t\bar{t}$ control region in the right corner. The final selection criteria for the tight regions, where the DNN classification cuts have been applied, are listed in Appendix D.

The summary plot of all signal and control regions, shown in Figure 6.5, illustrates the contributions of all considered processes to the total number of events in a region, both at the preselection level and after applying the cuts on the DNN scores. The regions satisfying tighter selection criteria are indicated by (t). A significant reduction of the respective background contributions is observed, demonstrating the effectiveness of the applied selection.

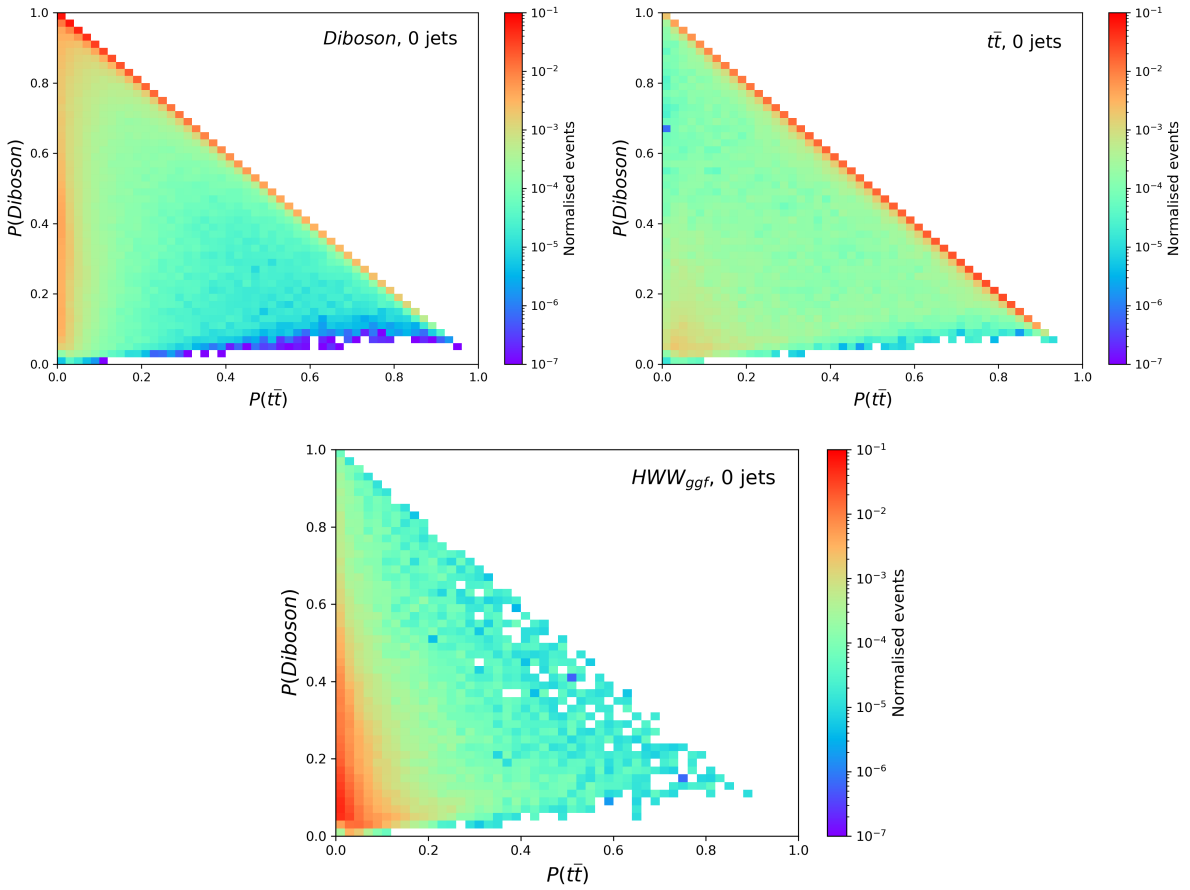

Figure 6.3.: The normalised event distributions of all considered classes are shown for the 0-jet region.

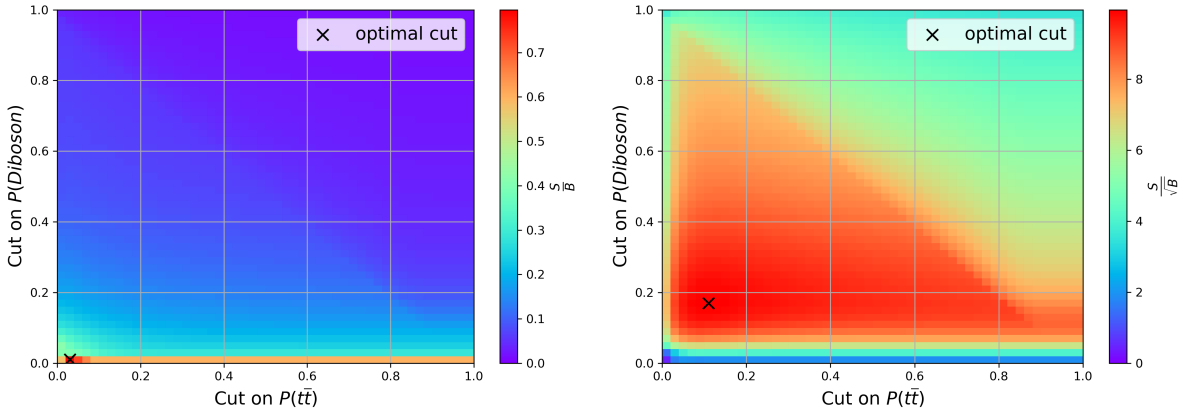

Figure 6.4.: The signal-to-background ratio distribution (left) and the $\frac{S}{\sqrt{B}}$ metric distribution (right) depending on the background scores are displayed for the 0-jet region.

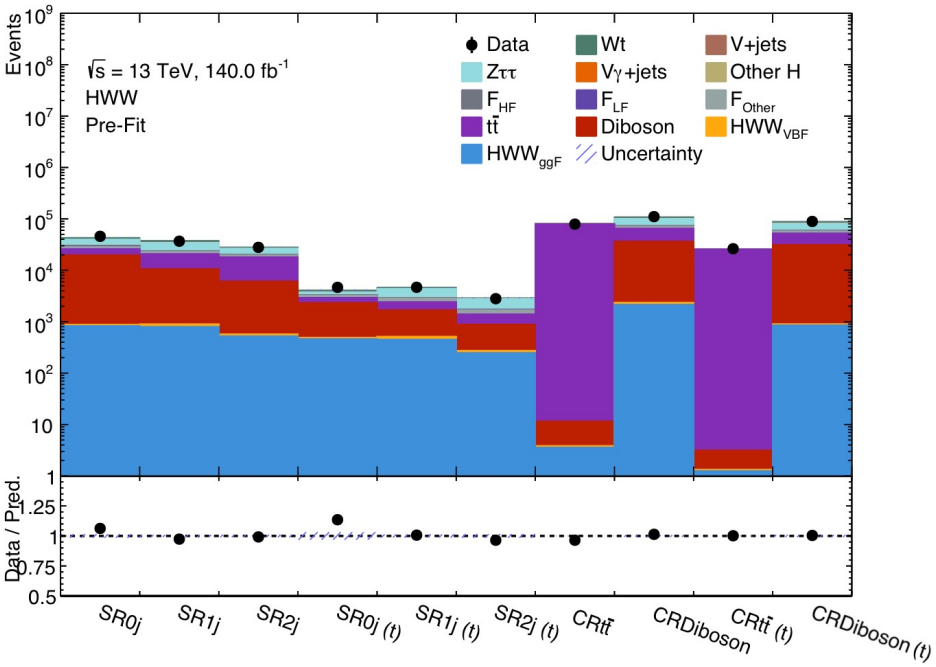

Figure 6.5.: The figure shows the summary plot for all signal and control regions. The (t) indicates that the cuts on the output scores have been applied.

7. Event Reconstruction

To study the entanglement of the W bosons, each event has to be fully reconstructed. To determine the kinematic properties of the decay, the four-momenta of the two charged leptons and neutrinos originating from the W boson decays are required. The charged leptons are measured in the detector, therefore their four-momenta are precisely known. Neutrinos are not directly detected; the only accessible kinematic properties of the dineutrino system are its transverse energy and its azimuthal angle in the transverse plane, both of which are indirectly inferred from the missing transverse energy, $E_{\mathrm{t}}^{\mathrm{miss}}$, and its direction, $\varphi_{E_{\mathrm{t}}^{\mathrm{miss}}}$. To fully reconstruct the neutrinos, reconstruction algorithms must be employed.

7.1. Reconstruction Algorithms

The reconstruction is performed with three reconstruction algorithms, which are tested and compared on their ability to precisely reconstruct the four-momenta of the neutrinos and other properties of the event. The algorithms are referred to as *Reco algorithm*, *Weighting algorithm* and *Extended reco algorithm*. The procedure for each algorithm is explained in the following.

Reco Algorithm

The Reco algorithm was originally proposed in [29]. First, the four-momentum of the dineutrino system is reconstructed. Due to conservation of energy and momentum, the squared four-momentum of the Higgs boson p_H, which equals its squared invariant mass M_H, can be expressed with the four-momenta of the dilepton system $p_{\ell\ell}$ and the dineutrino system $p_{\nu\nu}$, which leads to the expression $p_H^2 = M_H^2 = (E_{\ell\ell} + E_{\nu\nu})^2 - (\vec{p}_{\ell\ell} + \vec{p}_{\nu\nu})^2$. E_{ii} and $\vec{p}_{ii}$ refer to the energy and three-momentum of the charged lepton or neutrino system. With the missing transverse energy, $E_{\mathrm{t}}^{\mathrm{miss}}$, and its angle, $\varphi_{E_{\mathrm{t}}^{\mathrm{miss}}}$, the x- and y-component of the dineutrino system are determined to be $p_{x,y}^{\nu\nu} = E_{x,y}^{\mathrm{miss}}$. Finally, the Higgs mass is

expressed as

$$M_H^2 = 2E_{\ell\ell}\sqrt{M_{\nu\nu}^2 + (E_{\mathrm{t}}^{\mathrm{miss}})^2 + (p_z^{\nu\nu})^2} - 2p_x^{\ell\ell}E_x^{\mathrm{miss}} - 2p_y^{\ell\ell}E_y^{\mathrm{miss}} - 2p_z^{\ell\ell}p_z^{\nu\nu} + M_{\ell\ell}^2 + M_{\nu\nu}^2, \tag{7.1}$$

where the energy of the neutrinos is written as $E_{\nu\nu} = \sqrt{M_{\nu\nu}^2 + (E_{\mathrm{t}}^{\mathrm{miss}})^2 + (p_z^{\nu\nu})^2}$. The only parameters in this equation, which are not measured by the detector, are the masses of the Higgs boson and the dineutrino system, and the z-component of the dineutrino momentum. The dilepton mass is not directly measured, but is easily derived from its four-momentum. To solve the equation with three unknown parameters, two constraints are imposed on the system. The mass of the Higgs boson is fixed to its rounded world-average value of $M_H = 125$ GeV [9]. The mass of the dineutrino system is set to its average value in this decay, which is $M_{\nu\nu} = 30$ GeV according to reference [29]. These constraints leave $p_z^{\nu\nu}$ as the only unknown parameter. It can be determined by solving the obtained quadratic equation

$$a(p_z^{\nu\nu})^2 + bp_z^{\nu\nu} + c = 0, \tag{7.2}$$

where $a = (p_z^{\ell\ell})^2 - E_{\ell\ell}2$, $b = M_{fix}^2 p_z^{\ell\ell}$, $c = \frac{1}{4}M_{fix}^4 - E_{\ell\ell}^2((E_{\mathrm{t}}^{\mathrm{miss}})^2 + M_{\nu\nu}^2)$ and $M_{fix}^2 = M_H^2 - M_{\ell\ell}^2 - M_{\nu\nu}^2 + 2p_x^{\ell\ell}E_x^{\mathrm{miss}} + 2p_y^{\ell\ell}E_y^{\mathrm{miss}}$. The quadratic equation is solved for $p_z^{\nu\nu}$, yielding the solutions $p_z^{\nu\nu} = \frac{-b\pm\sqrt{\Delta}}{2a}$, where the discriminant is defined as $\Delta = b^2 - 4ac$. Depending on the value of the discriminant, there are three ways to proceed.

$\Delta = 0$: The quadratic equation returns one solution for $p_z^{\nu\nu}$, which is taken as the result.

$\Delta < 0$: Equation (7.2) yields two solutions, which are imaginary. For these cases, the mean value of $M_{\nu\nu}$ is 17 GeV, which is lower than 30 GeV. Therefore, the initial approach to obtain a real solution is to relax the constraint on the dilepton mass by setting $M_{\nu\nu} = 0$, with the aim of reconstructing cases where $0 < M_{\nu\nu} < 30$ GeV. If the discriminant remains negative, $p_z^{\nu\nu}$ is chosen such that $\frac{\partial M_H^2}{\partial p_z^{\nu\nu}} = 0$, which yields the determining equation $p_z^{\nu\nu} = p_z^{\ell\ell}\frac{\sqrt{(E_{\mathrm{t}}^{\mathrm{miss}})^2 + M_{\nu\nu}^2}}{\sqrt{E_{\ell\ell}^2 - (p_z^{\ell\ell})^2}}$. If the discriminant becomes positive, the event is treated as explained in the following case.

$\Delta > 0$: In this case, equation (7.2) yields two real solutions for $p_z^{\nu\nu}$. The solution, which pushes the dilepton system more towards the transverse plane, is chosen. This is done by minimising $|\cos\psi_{\ell\ell}^*|$, where $\psi_{\ell\ell}^*$ is the opening angle of the dilepton system with the z-axis in the Higgs restframe.

Once the four-momentum of the dineutrino system has been reconstructed, the four-momenta of the individual neutrinos need to be extracted. As discussed in Subsection

2.1.3, the neutrinos are often collinear in the Higgs restframe. Therefore, they are assumed to be collinear. This leads to the ansatz $\vec{p}_{\nu,\text{on}} = \alpha^2 \vec{p}_{\nu\nu}$ and $\vec{p}_{\nu,\text{off}} = (1-\alpha^2)\vec{p}_{\nu\nu}$. Here, $\vec{p}_{\nu,\text{on/off}}$ refers to the three-momentum of the neutrino, that will be matched to the on/off-shell W boson. The corresponding neutrinos are denoted as ν_{on} and ν_{off}. The positive scaling factor is α^2. An overview of calculating α^2 and fully determining the separate neutrino momenta is given in the following. The complete derivation is performed in Appendix E. To determine α^2, one W boson is constrained to be on-shell and consequently having a mass of $M_W = 80.379$ GeV [9]. This leads to an expression for α^2, which is

$$\alpha^2 = \frac{M_{W,\text{on}}}{2(E_\ell E_{\nu\nu} - \vec{p}_\ell \vec{p}_{\nu\nu})}. \tag{7.3}$$

To match the charged leptons to the on- and off-shell W bosons, the invariant mass is calculated for the sum of each charged lepton's four-momentum and that of the dineutrino system. The lepton that yields the higher invariant mass with the dineutrino system is associated with the on-shell W boson, whereas the other is assigned to the off-shell W boson. If (7.3) returns a solution $\alpha^2 > 1$, it is fixed to be $\alpha^2 = 1$.

To resolve the contradiction between the assumption that the neutrinos are collinear, which implies a massless dineutrino system, and the imposed constraint of $M_{\nu\nu} = 30$ GeV, a fictive mass is assigned to the neutrino associated with the off-shell W boson. The fictive mass is calculated as

$$M_{\nu\text{f}} = \sqrt{(E_{\nu\nu} - \alpha^2 |\vec{p}_{\nu\nu}|^2)^2 - (1-\alpha^2)^2 |\vec{p}_{\nu\nu}|^2}. \tag{7.4}$$

The final four-momenta of the neutrinos are defined over their three-momentum-components and respective masses according to

$$p_{\nu,\text{on}} = (\alpha^2 E_x^{\text{miss}},\ \alpha^2 E_y^{\text{miss}},\ \alpha^2 p_z^{\nu\nu},\ 0) \tag{7.5}$$

and

$$p_{\nu,\text{off}} = ((1-\alpha^2) E_x^{\text{miss}},\ (1-\alpha^2) E_y^{\text{miss}},\ (1-\alpha^2) p_z^{\nu\nu},\ M_{\nu\text{f}}). \tag{7.6}$$

Weighting Algorithm

The approach for the Weighting algorithm is to once again assume that the neutrinos are collinear and scan over different values of the dineutrino pseudo-rapidity $\eta_{\nu\nu}$ and the scaling factor α^2, which is defined as before. Every solution receives a weight, and the solution corresponding to the highest weight is selected. First, the dineutrino four-

momentum is defined through $E_{\mathrm{t}}^{\mathrm{miss}}$, $\varphi_{E_{\mathrm{t}}^{\mathrm{miss}}}$ and $M_{\nu\nu} = 30$ GeV, which corresponds to the constraint from the Reco algorithm. The remaining component needed to fully define the four-momentum is $\eta_{\nu\nu}$. To determine which values to test, a Gaussian curve is fitted to the distribution of $\eta_{\nu\nu}$ at the truth level, shown in Figure 7.1. The fitted parameters are the mean value $\mu = 0.00 \pm 0.07$, the standard deviation $\sigma = 2.07 \pm 0.07$ and the amplitude $A = 15300 \pm 500$, which is not relevant for further calculations. The inverse cumulative distribution function is employed to determine 100 quantiles, corresponding to the values of $\eta_{\nu\nu}$ that are tested. To determine the separate neutrino four-momenta, 100 equidistant values between 0.1 and 1.0 are used as scaling factor α^2. Given two four-momenta for the charged leptons and two for the neutrinos, four possible candidates for the on-shell W boson four-vector arise. Each candidate is constructed by summing the four-momenta of one charged lepton and one neutrino. A candidate is considered valid, if the mass difference $\Delta_{W_{\mathrm{on}}}$ of the reconstructed on-shell W boson and its theoretical value of $M_W = 80.379$ GeV is less than 25 GeV. Furthermore, the mass difference Δ_H of the reconstructed Higgs boson, which is obtained by adding the four-momenta of both charged leptons and both neutrinos, has to be less than 30 GeV. If a solution is considered as valid, it is assigned a weight,

$$w = \exp\left(-\frac{(\Delta_{W_{\mathrm{on}}})^2}{2\sigma_W^2} - \frac{(\Delta_H)^2}{2\sigma_H^2}\right). \tag{7.7}$$

The widths are defined as $\sigma_W = 2$ GeV and $\sigma_{Higgs} = 2$ GeV. The solution with the highest weight is chosen.

Extended Reco Algorithm

The Extended reco algorithm is based on the Reco algorithm, but the restriction requiring the dineutrino mass to be $M_{\nu\nu} = 30$ GeV is loosened. The distribution of $M_{\nu\nu}$ is plotted on truth level and a Gaussian curve is fitted. The resulting distribution is shown in Figure 7.2. The fitting parameters are the mean value $\mu = 32.01 \pm 0.4$, the standard deviation $\sigma = 18.1 \pm 0.4$ and the amplitude $A = 1770 \pm 30$, which again is not relevant for further calculations. Similar to the Weighting algorithm, the inverse cumulative distribution function is used to determine 100 quantiles, which represent the values of $M_{\nu\nu}$, that are tested. For each value, defining the constrained mass, the complete Reco algorithm is performed. Every solution is assigned with a weight

$$w = \exp\left(-\frac{(\Delta_{W_{\mathrm{on}}})^2}{2\sigma_W^2} - \frac{(\Delta_H)^2}{2\sigma_H^2}\right). \tag{7.8}$$

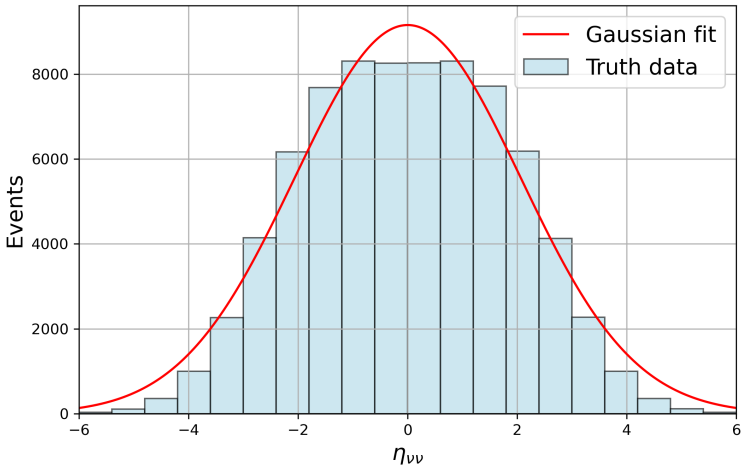

Figure 7.1.: The figure shows the truth-level distribution of $\eta_{\nu\nu}$ and a Gaussian function fitted to the data. The fitted parameters are the mean value $\mu = 0.00 \pm 0.07$, the standard deviation $\sigma = 2.07 \pm 0.07$ and the amplitude $A = 15300 \pm 500$.

The widths are defined as before in Section 7.1. The solution associated with the highest weight is selected.

Reconstruction on Truth Level

To evaluate and compare the performance of the algorithms, the reconstructed values are compared to their corresponding truth-level values. The truth information is included in the signal samples used. In the reconstruction, the particles are referred to as on-shell and off-shell, depending on whether they are matched to the W boson treated as on-shell or to the other W boson, which has a lower mass. In order to compare the reconstructed particles with their correct truth-level counterparts, the truth particles are also labelled as on-shell and off-shell accordingly. To achieve this, the sums of the four-momenta for all possible combinations of charged leptons and neutrinos are computed. The combination that yields the invariant mass closest to $M_W = 80.379$ GeV is assigned to be the on-shell lepton and on-shell neutrino, while the remaining pair is labelled as off-shell.

7.2. Evaluation of the Reconstruction Algorithms

For a successful quantum tomography in order to study the entanglement of the W boson in the process $H \to WW^* \to \ell^+\nu\ell^-\overline{\nu}$, it is essential to reconstruct the neutrino four-momenta precisely. This section evaluates the performance of three reconstruction algorithms with respect to this task and compares them against each other. Initially, gen-

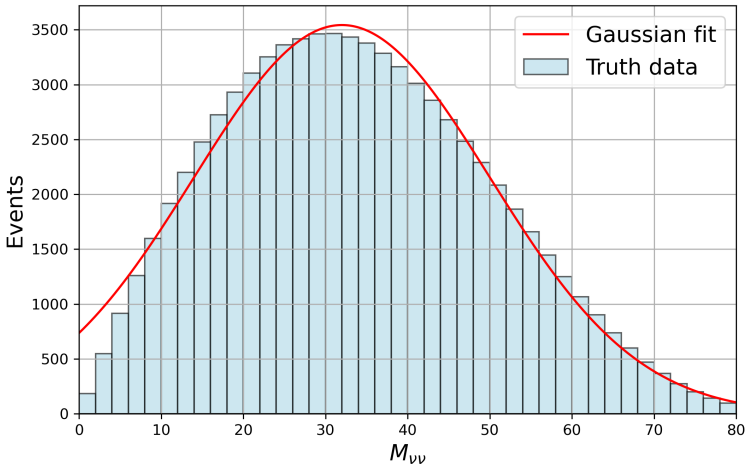

Figure 7.2.: The figure shows the truth-level distribution of $M_{\nu\nu}$ and a Gaussian function fitted to the data. The fitted parameters are the mean value $\mu = 32.01 \pm 0.4$, the standard deviation $\sigma = 18.1 \pm 0.4$ and the amplitude $A = 1770 \pm 30$.

eral characteristics of the reconstruction are discussed. Subsequently, the reconstructed neutrino quantities are compared to their respective truth-level values for the HWW_{ggf} sample.

Evaluation of the Scaling Factor α^2

For all algorithms, a central component in dividing the dineutrino momentum between the individual momenta is the scaling factor α^2. Its distributions for SR0j are shown in Figure 7.3 for all algorithms. In the case of the Weighting algorithm, a large dip at around 0.5 is observed, with the distribution exhibiting symmetry about this point. This suggests that ν_{on} and ν_{off} are treated equivalently when splitting the dineutrino momentum. In contrast, the distributions of the other two algorithms steadily increase towards a value near 0.8 and display higher bin counts at larger values of α^2. The neutrinos share identical initial conditions, except for the mass of their mother particles. Consequently, the neutrino originating from the heavier W boson, ν_{on}, is expected to possess greater momentum than ν_{off}. Both the Reco algorithm and the Extended reco algorithm demonstrate significantly better agreement with this expectation. Furthermore, the distribution of the Reco algorithm shows a lot of entries at $\alpha^2 = 1$, indicating that originally α^2 was often calculated to exceed one. This leads to ν_{off} being at rest and contradicts the collinear hypothesis. This issue no longer arises for the Extended reco algorithm. Allowing a wider range of values for the dineutrino mass appears to result in a

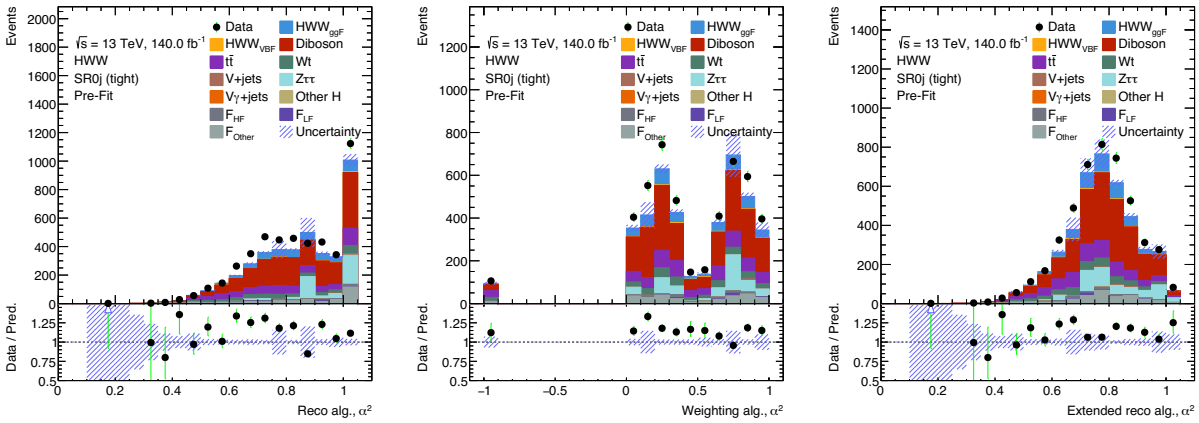

Figure 7.3.: The figure shows the computed scaling factor α^2 for SR0j for all algorithms.

more realistic reconstruction. Consequently, this algorithm produces the most physically realistic distribution of α^2. The α^2 distributions for the other jet regions can be found in the appendix in Figuren F.1 and exhibit similar behaviour, although the dip at 0.5 for the Weighting algorithm is significantly less pronounced in SR2j. The distributions obtained with the Weighting algorithm show a small number of events with $\alpha^2 = -1$. This value does not correspond to a physically relevant solution, but rather serves as a default for events that fail the mass difference requirements with respect to the constrained Higgs and on-shell W boson masses, which are necessary for the algorithm to perform a successful reconstruction. The distributions for the other jet regions exhibit similar behaviour. However, this affects only a few hundred events per jet region. For a further analysis of the quantum-mechanical behaviour of the W bosons, these events could be excluded without significantly reducing the statistics.

Discussion of the Reconstructed Masses M_H and $M_{W_{\text{on}}}$

Furthermore, the reconstructed Higgs boson and on-shell W boson masses are examined. The corresponding distributions for SR0j are shown in Figure 7.4, while the distributions for the other regions are provided in the appendix in Figure F.2. For the Reco algorithm as well as the Extended reco algorithm, these masses are fixed to be $M_H = 125$ GeV and $M_W = 80.379$ GeV. Therefore, the distributions exhibit very sharp peaks at these values across all jet regions. However, the peak in the Reco algorithm contains around 1000 fewer events. These events accumulate at lower mass values. This difference corresponds to the number of events assigned $\alpha^2 = 1$. This is expected from Equation (E.2), because values $\alpha^2 > 1$ imply that a momentum greater than the dineutrino momentum must be assigned to ν_{on} to satisfy the on-shell W mass constraint. Since this is not possible, as α^2

cannot exceed one, the reconstructed mass of the on-shell W boson is smaller than the imposed constraint. This is also observed for the other jet regions.

Examining the Weighting algorithm, the mass distribution of the on-shell W boson exhibits a pronounced peak at the constrained value. In contrast, the peak corresponding to the Higgs boson mass is less sharp, and several events with higher masses are observed. This effect diminishes as the number of considered jets increases. This behaviour is unexpected, since the algorithm assumes the neutrinos to be collinear, which best approximates the true physical behaviour in the absence of jets and when the Higgs boson is at rest. Nevertheless, although the assumed masses are not directly imposed on the system, the algorithm still, in most cases, finds values for α^2 and η, that satisfy the expected mass constraints.

However, it should be noted that the sharpness of the mass peaks is much higher than expected for this process. In particular, although the W boson is commonly assumed to be on-shell, this assumption is not universally valid, as there exist scenarios in which both W bosons are off-shell. Therefore, these tight constraints on the respective masses should be revisited to improve the reconstruction algorithms.

7.2.1. Evaluation of the Reconstructed Neutrino Kinematics

To evaluate the algorithms' ability to accurately reconstruct the neutrino kinematics of the signal process, their reconstructed kinematic quantities are compared to their corresponding truth-level values. Only the HWW_{ggf} sample is evaluated, as the HWW_{vbf} contribution to the signal is comparatively small. For each event, the reconstructed value is plotted against its truth value. An ideal reconstruction is indicated by a diagonal distribution, representing a high degree of accuracy. The diagonality d, defined as the normalised sum of all bins along the diagonal, is used as a quantitative measure of this behaviour. Its values are provided at the corresponding plots. Additionally, the difference between the truth and reconstructed values is analysed. The analysed kinematic quantities are the pseudo-rapidity η, the transverse momentum p_{t} as well as its longitudinal component p_z. It should be noted that the unreconstructed events for the Weighting algorithm have not been excluded from the following plots, because this effect is inherent to the algorithm and should be part of the evaluation. Since the default values of the considered quantities are zero, these events accumulate along a horizontal line where the reconstructed value is zero. As discussed before, this affects only a few hundred events per region. Nevertheless, due to this effect, the Weighting algorithm can be considered slightly more accurate than the plots might suggest.

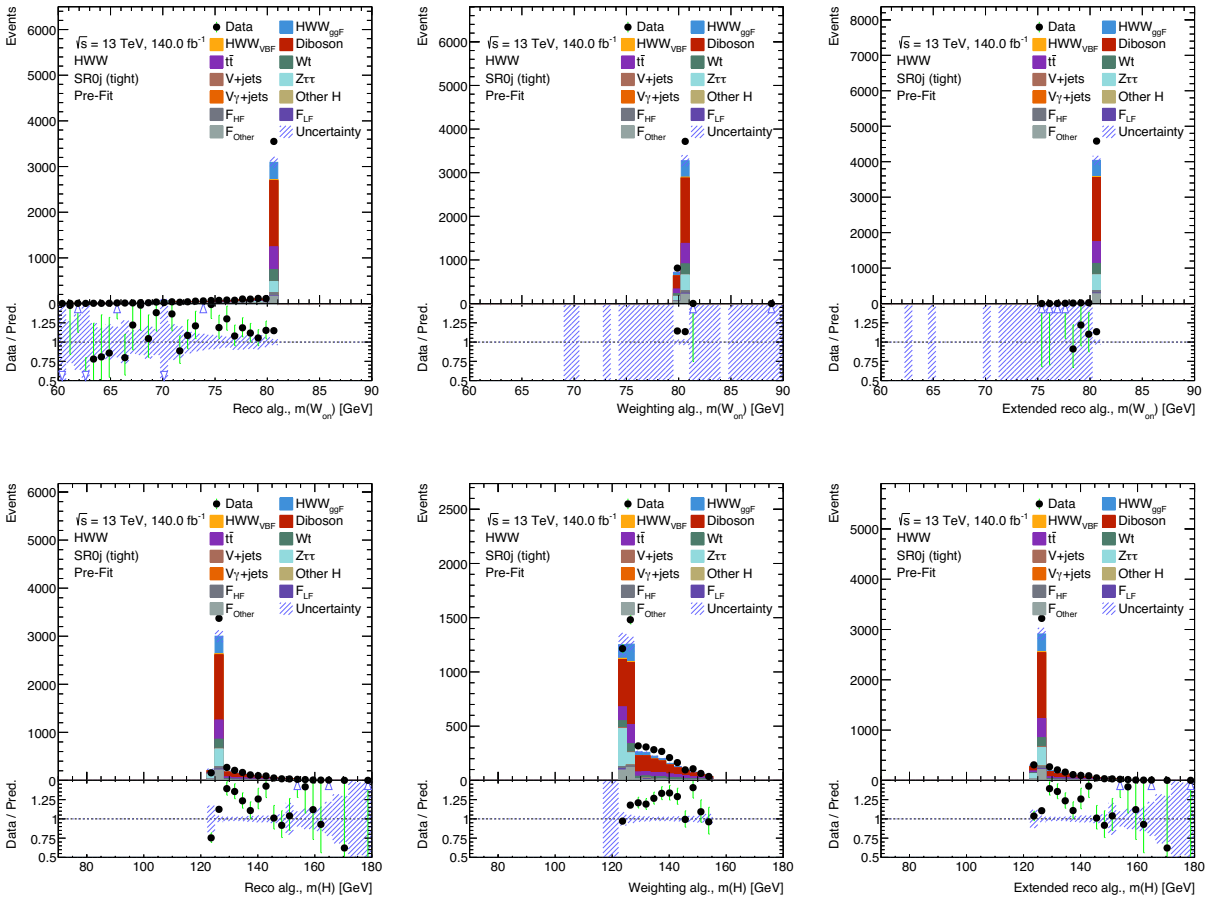

Figure 7.4.: The distributions of the reconstructed Higgs and on-shell W boson masses are shown for SR0j.

Evaluation of the Pseudo-Rapidity η

The two-dimensional distributions as well as the residual distributions of the pseudo-rapidity are shown in Figure F.3 for ν_{on} and in Figure F.4 for ν_{off}. Across all jet regions and for both neutrinos, the Weighting algorithm exhibits the best reconstruction performance at higher values of η, which results in the distribution appearing more diagonal. Due to the scanning over a range of η values, larger values are more frequently selected as the solution. Furthermore, for the off-shell neutrino, the Reco algorithm produces a high event density at a reconstructed value of zero. This corresponds to the large number of events assigned with $\alpha^2 = 1$, which leads to the off-shell neutrino's three-momentum becoming a zero vector, for which $\eta = 0$ holds. Nevertheless, the calculated diagonality values, as well as the residual plots, indicate that overall, the Reco algorithm provides the most accurate reconstruction of η. However, all two-dimensional distributions appear reasonably diagonal, and the calculated diagonality values, which indicate that 5.9 to even

10.7 percent of the events lie on the diagonal, are of comparable magnitude. The residual plots show distinct peaks at zero for all algorithms, and the distributions appear broadly similar, despite minor variations in the peak heights.

Evaluation of the Transverse Momentum p_t

The two-dimensional distributions and the residual plots for the transverse momentum are shown in Figure F.5 for ν_{on} and in Figure F.6 for ν_{off}. Across all algorithms, the two-dimensional distributions appear highly dispersed. While for the on-shell neutrino the event densities tend to accumulate along the diagonal, the distributions for the off-shell neutrino are broadly spread across the entire plane and lack a clear diagonal structure. This is reflected in very low calculated diagonality values. For the on-shell neutrino, the Weighting algorithm in particular yields a lower diagonality and the peak in the difference distributions is not as high as for the other algorithms. This effect cannot be solely attributed to the few hundred unreconstructed events, which only slightly reduce the diagonality, but rather indicates inferior reconstruction performance. The other algorithms, however, perform comparably for the on-shell neutrino. Only the Reco algorithm exhibits a tendency to reconstruct $p_{\text{t,on}}$ as slightly too large, as indicated by a shift towards negative values in the difference distributions. This is due to the large number of events assigned with $\alpha^2 = 1$, since it is not physically realistic for a single neutrino to account for the entire amount of missing transverse energy measured by the detector. The effect of these events is also observed for the off-shell neutrino. The Reco algorithm again exhibits a high event density at $p_{\text{t}}^{\text{reco}} = 0$, representing the events with $\alpha^2 = 1$. A significant shift of the difference distribution towards positive values can be observed, indicating that the reconstructed values are too low, and not actually zero. The other two algorithms show a more accurate reconstruction of $p_{\text{t,off}}$, and the difference distributions are more strongly centred around zero. Nevertheless, all algorithms struggle to precisely reconstruct $p_{\text{t,off}}$, as indicated by the low diagonality values, with only 2.7 to 5.5 percent of events reconstructed on the diagonal, and the absence of a clear diagonal structure in the two-dimensional distributions. Furthermore, for all algorithms, the reconstruction accuracy decreases with an increasing number of jets in the respective region. Since $p_{\text{t,on}}$ and $p_{\text{t,off}}$ are directly derived from the measured missing transverse energy and the scaling factor α^2, the reduced accuracy of α^2 in events with multiple jets is responsible for this behaviour. This is expected, as the use of a scaling factor is only meaningful under the assumption that the neutrinos are collinear, a condition that holds more accurately when fewer jets are present.

Evaluation of the Longitudinal Momentum p_z

The two-dimensional distributions, as well as the residual plots for the longitudinal component of the momentum, p_z, are depicted in Figure F.7 for ν_{on} and in Figure F.8 for ν_{off}. The two-dimensional distributions of ν_{on} are spread across the entire plane, but exhibit a somewhat diagonal shape, indicated by moderate diagonality values compared to the other distributions, with 4.4 to 7.0 percent of all events lying on the diagonal. The Extended reco algorithm provides a slightly more accurate reconstruction than the other two algorithms, as evidenced by the highest diagonality values and the most pronounced peak at zero at the difference distributions. The two-dimensional distributions of the off-shell neutrino's longitudinal momentum are also spread and do not exhibit a clear diagonal structure. Nevertheless, the computed diagonality values are relatively high compared to those of $p_{z,\text{on}}$, indicating that 6.6 to 8.4 percent of all events lie along the diagonal. This can be attributed to a relatively accurate reconstruction of lower values of $p_{z,\text{off}}$. In contrast, for high values of $p_{z,\text{off}}^{\text{truth}}$, all algorithms appear to encounter difficulties. The Higgs boson has spin zero, and therefore the distribution of the W bosons is isotropic in the Higgs restframe, leading to generally lower values of their longitudinal momenta. Since the leptons tend to be emitted collinearly or anti-collinearly with their parent particles in this process, the longitudinal component of the neutrino momenta is often small. This results in a more accurate reconstruction of $p_{z,\text{off}}$ by the Reco algorithm, as in the common case where α^2 is computed to be one, the reconstructed value is identically zero. However, this cannot be considered a genuine indication of good algorithmic performance, as it does not result from an accurate reconstruction; the same outcome would be obtained by simply always setting the value to zero. Moreover, this behaviour frequently occurs even for higher values of $p_{z,\text{off}}^{\text{truth}}$, indicated by a high density of events near the bottom of the plot. Therefore, the Extended reco algorithm yields the most reliable results, closely followed by the Weighting algorithm. For both neutrino types and across all algorithms and jet regions, a shift in the difference distributions towards positive values is observed. This suggests that, for all algorithms, $p_z^{\nu\nu}$ tends to be systematically underestimated. This effect is less pronounced for the Weighting algorithm, likely because the scanning over multiple values of η also increases the range of considered values for $p_z^{\nu\nu}$. Nevertheless, the Extended weighting algorithm can be considered the most reliable, although it struggles to accurately reconstruct especially higher values of the longitudinal momentum.

Conclusion

In conclusion, all algorithms exhibit their individual strengths and weaknesses in the reconstruction of the neutrinos. For some variables, such as $p_{\mathrm{t,off}}$, $p_{z,\mathrm{on}}$ and $p_{z,\mathrm{off}}$, an improvement is observed when moving from the Reco algorithm to the Extended reco algorithm, indicating that scanning over multiple values of $M_{\nu\nu}$ yields a more accurate reconstruction by preventing the scaling factor α^2 from frequently being equal to one. The approach of scanning over a range of values for the pseudo-rapidity also appears promising, as the Weighting algorithm provides a more reliable reconstruction of higher values of η and p_z of both neutrino types. An overview of the advantages and disadvantages is shown in Table 7.1.

However, to achieve a more precise reconstruction, the collinear approach for the neutrinos needs to be revisited, since it only applies when the Higgs boson is assumed to be at rest. The limited accuracy in the computation of α^2, particularly evident in the reconstruction of $p_{\mathrm{t,on}}$ and $p_{\mathrm{t,off}}$, is not only due to the general assumptions underlying the algorithms, but primarily to the fact that an exact value of α^2 cannot be defined, as the neutrinos are not perfectly collinear in reality.

Furthermore, the strategy of selecting solutions that best match the assumed masses of the Higgs boson and the W boson, which is treated as on-shell, should be re-evaluated. As shown in Figures 7.4 and F.2, this condition is well satisfied by all algorithms and across all jet regions, as it is explicitly imposed by the reconstruction strategies. Nevertheless, only a small percentage of events are reconstructed with high precision, as discussed above, suggesting that additional selection criteria should be applied when identifying the most accurate reconstruction of an event.

However, given the number of assumptions required for neutrino reconstruction, the achieved accuracy is promising. The strengths and weaknesses of the individual algorithms have been clearly identified, providing a basis for further improvements in their performance.

Table 7.1.: The table shows an overview of advantages and disadvantages of the evaluated algorithms.

Algorithm	**Advantages**	**Disadvantages**
Reco Algorithm	• M_H and $M_{W_{\text{on}}}$ are reconstructed accurately with respect to their assumed values • accurate reconstruction of η	• α^2 is often equal to one • many assumptions • low accuracy for $p_{\text{t,off}}$ and p_z
Weighting Algorithm	• more accurate for larger values of η	• less precise reconstruction of M_H • reconstruction of p_{t} and p_z less accurate
Extended Reco Algorithm	• α^2 is less often equal to one • more accurate for $p_{\text{t,off}}$ and p_z	• accuracy of $p_{\text{t,off}}$ and $p_{z,\text{off}}$ still only moderate

8. Conclusion and Outlook

This thesis investigates analysis strategies aimed at studying quantum signatures in the process $H \to WW^* \to \ell\nu\ell\nu$. The two key steps required to access quantum-sensitive observables in the considered process are the separation of signal and background events, and a precise reconstruction of the final-state leptons and neutrinos.

To separate signal events from background events, a deep neural network is used. A multiclass classification task is performed, including the signal process HWW as well as the background processes *Diboson* and $t\bar{t}$. K-fold cross validation is used to assess the model's ability to generalise to unseen data. The DNN outputs are used to define signal and control regions via a square cut method. This technique leads to a significant improvement in the signal purity and statistical significance within the signal regions. However, the background contributions remain large. To further improve this, the training should be extended to include the $Z\tau\tau$ sample, and other background processes, such as fake leptons, should be considered. A multi-dimensional square cut optimisation could be conducted to define selection criteria that take all background processes into account.

To study the entanglement of the W bosons, quantum tomography techniques are a common approach. To enable such analyses, each event must be fully reconstructed. The greatest challenge in obtaining the final state for the process $H \to WW^* \to \ell\nu\ell\nu$ is the neutrino reconstruction, as neutrinos are not directly detected by the ATLAS detector. In this thesis, three algorithms are tested for their ability to precisely reconstruct the neutrinos in the considered process. Constraints are imposed on the Higgs mass and the mass of the on-shell W boson, and the neutrinos are assumed to be collinear. The reconstructed values for the pseudo-rapidity, the longitudinal and transverse momentum of both neutrinos are evaluated against their truth-level counterparts, and the algorithms are compared based on their ability to precisely reconstruct these values. The reconstruction accuracy of the pseudo-rapidity and the momentum components of the on-shell neutrino is relatively high across all algorithms. In contrast, all methods struggle to accurately reconstruct $p_{\mathrm{t,off}}$ and $p_{z,\mathrm{off}}$. Nevertheless, considering the number of constraints and assumptions involved in neutrino reconstruction, the overall performance appears promising, and the specific strengths of each algorithm have been identified. In particular, the approach of scan-

ning over a range of values for the pseudo-rapidity, the dineutrino mass and the scaling factor shows potential. This strategy should be extended to include additional variables such as the longitudinal momentum of the dineutrino system, enabling multi-dimensional scanning over possible configurations across the entire phase space. Furthermore, the collinear hypothesis assumed for the neutrinos should be revisited, as it does not hold for the majority of events, particularly those with higher jet multiplicity. Since the angular distribution of the final state particles is an important observable for quantum tomography, greater efforts should be directed towards allowing reconstruction configurations that do not require the neutrinos to be exactly collinear. Besides the mass requirements on the Higgs and the on-shell W boson, additional criteria should be considered when selecting the optimal reconstruction.

In conclusion, this thesis has investigated various analysis strategies to probe quantum signatures in the process $H \to WW^* \to \ell\nu\ell\nu$, revealing key aspects with potential for further improvement. Consequently, a solid foundation has been established to pave the way for the successful application of quantum tomography and the detailed study of quantum entanglement in this process.

A. MC Event Generation

The Monte Carlo event generators used in this analysis are Powheg [17], Pythia [19] and Sherpa [18]. The samples and their corresponding generators are listed below.

Table A.1.: The table lists MC Generator and Parton Shower configurations for the considered samples.

Sample	**Generator**	**Parton Shower**
HWW_{ggf}	Powheg	Pythia 8
HWW_{vbf}	Powheg	Pythia 8
$\text{Diboson}_{\text{EWK}}$	Sherpa 2.2.12	
$\text{Diboson}_{\text{QCD}}$	Sherpa 2.2.14	
$t\bar{t}$	Powheg	Pythia 8
Wt	Powheg	Pythia 8
Wjets	Sherpa 2.2.11	
Zjets	Sherpa 2.2.11	
$Z\tau\tau$	Sherpa 2.2.14	
$V\gamma$jets	Sherpa 2.2.14	
Other Higgs	Powheg	Pythia 8

B. DNN Output Distributions of the Individual Folds

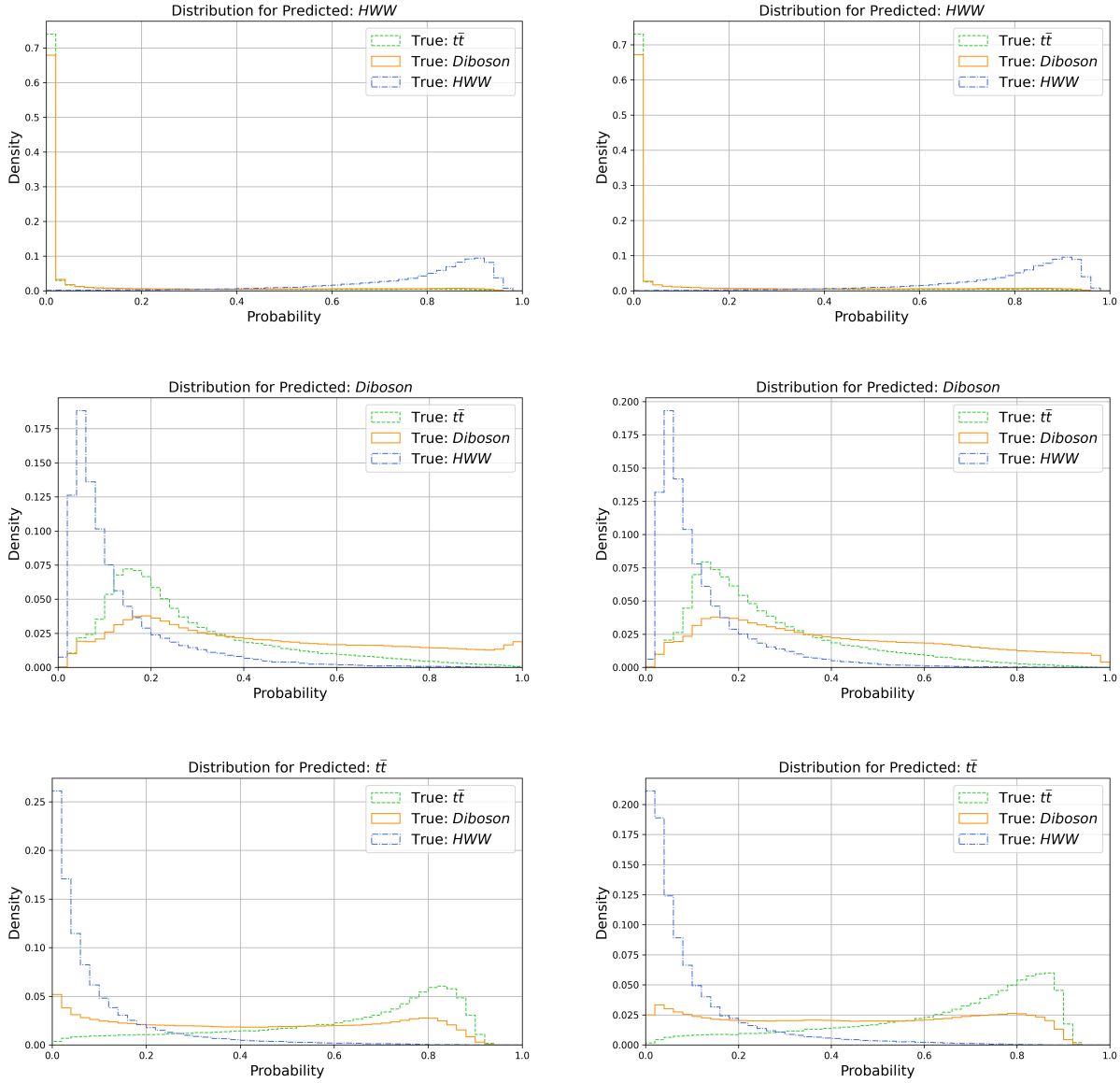

Figure B.1.: The figure shows the output distributions for all classes. Each column represents one fold.

C. Additional Square Cut Results

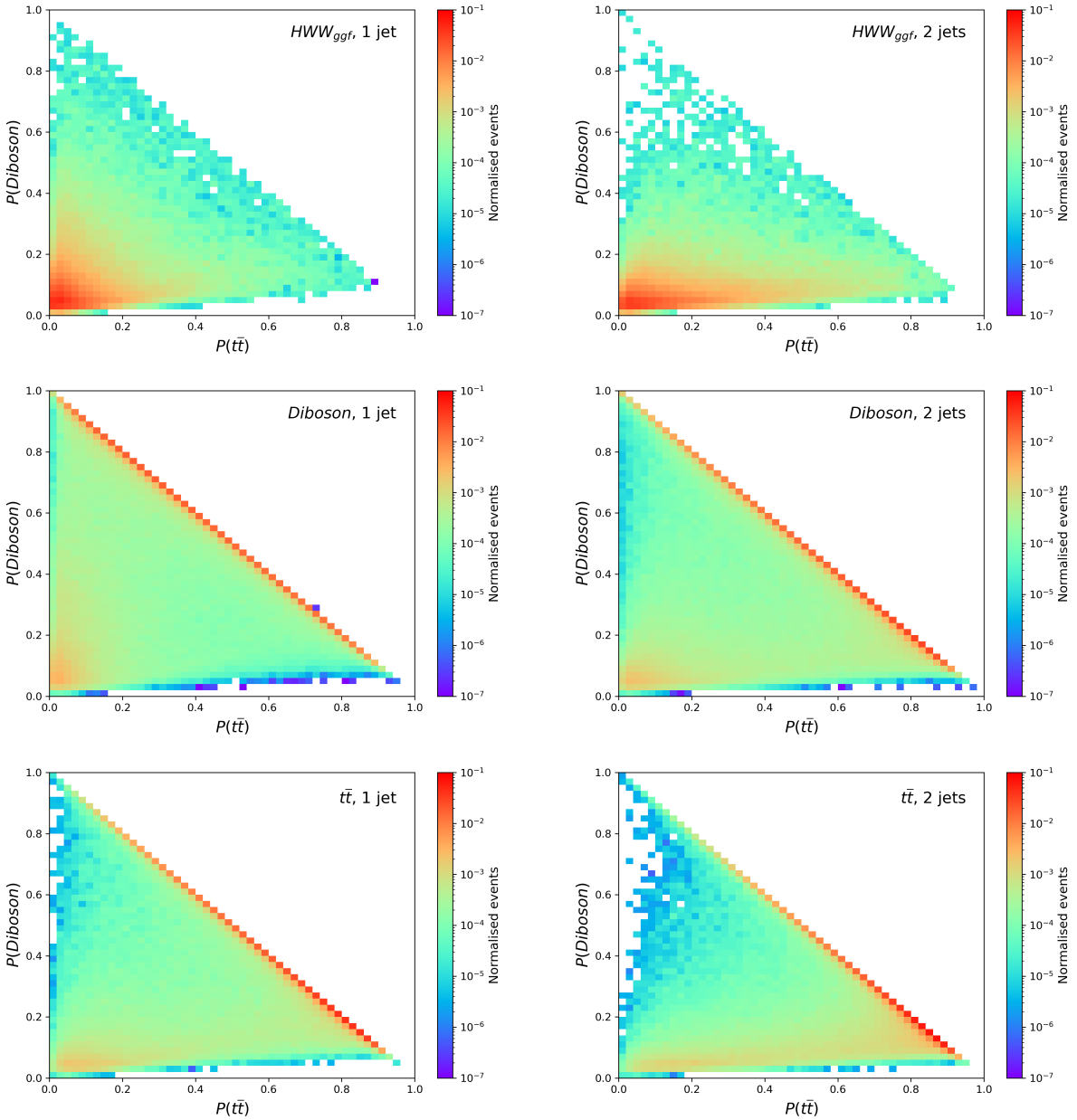

Figure C.1.: The figure shows the normalised event distributions for all classes for SR1j (left) and SR2j (right).

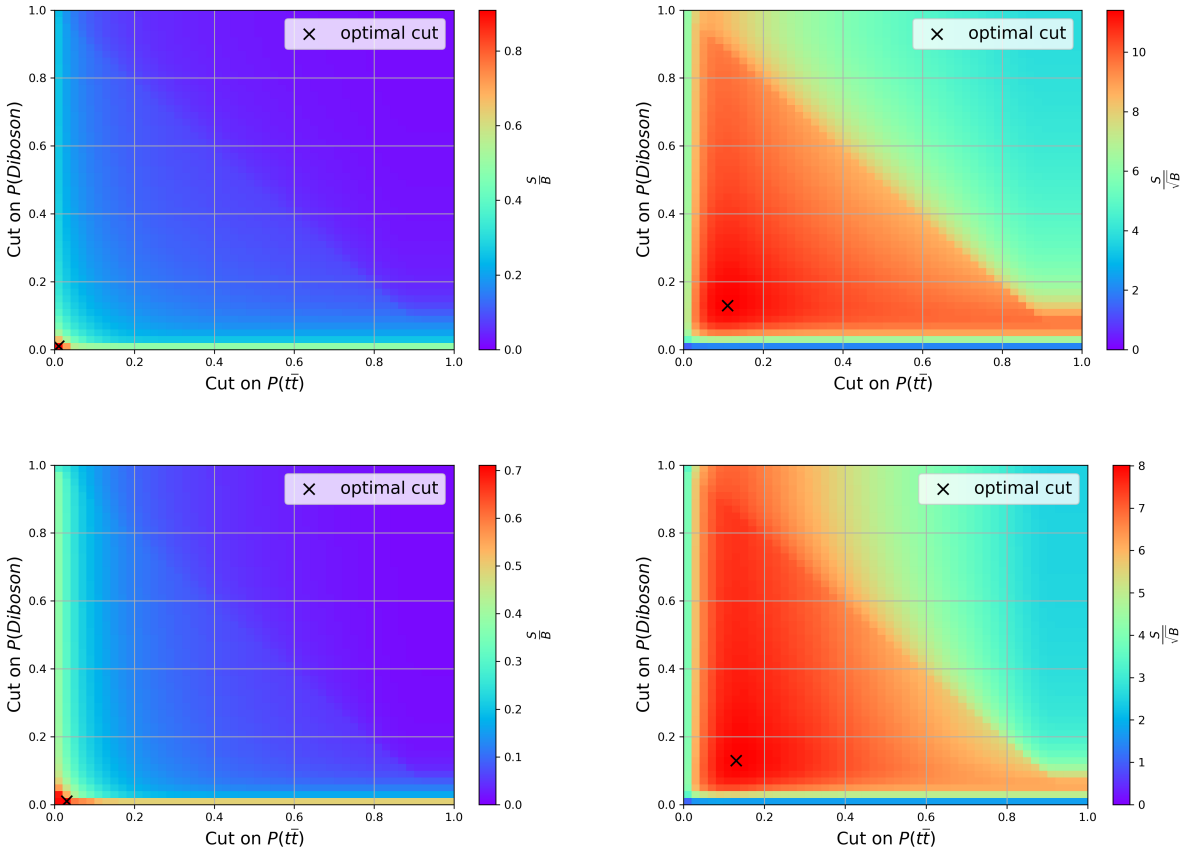

Figure C.2.: The figure shows the distributions of the cumulative sums of $\frac{S}{B}$ and $\frac{S}{\sqrt{B}}$ for SR1j (top) and SR2j (bottom).

D. Selection Criteria

Table D.1.: The table lists the Selection criteria used for the signal regions. The (t) denotes that DNN classifier cuts have been applied.

Selection Criteria	SR0j (t)	SR1j (t)	SR2j (t)
Jet multiplicity	$N_{\text{jets}} = 0$	$N_{\text{jets}} = 1$	$N_{\text{jets}} \geq 2$
Missing transverse energy	$E_{\text{t}}^{\text{miss}} > 20$ GeV		
b-jet multiplicity	$N_{b\text{-jet, }90\%} = 0$		
Leading lepton p_T	$p_{\text{t}} \geq 25$ GeV		
Subleading lepton p_T	$p_{\text{t}} \geq 25$ GeV		
Dilepton mass	10 GeV $< m_{\ell\ell} <$ 100 GeV		
Number of electrons	$N_e = 1$		
Number of muons	$N_\mu = 1$		
Diboson classification score	≤ 0.11	≤ 0.11	≤ 0.13
$t\bar{t}$ classification score	≤ 0.17	≤ 0.13	≤ 0.13

Table D.2.: The table shows the selection criteria used for the control regions. The (t) denotes that DNN classifier cuts have been applied.

Selection Criteria	CRDiboson (t)	CR$t\bar{t}$ (t)
Missing transverse energy	$E_{\text{t}}^{\text{miss}} > 20$ GeV	
b-jet multiplicity	$N_{b\text{-jet, }90\%} = 0$	$N_{b\text{-jet, }65\%} \geq 2$
Leading lepton p_T	$p_{\text{t}} \geq 25$ GeV	
Subleading lepton p_T	$p_{\text{t}} \geq 25$ GeV	
Dilepton mass	10 GeV $< m_{\ell\ell} <$ 100 GeV	
Number of electrons	$N_e = 1$	
Number of muons	$N_\mu = 1$	
Jet multiplicity		$N_{\text{jets}} \geq 2$
Diboson classification score	> 0.13	≤ 0.13
$t\bar{t}$ classification score		> 0.13

E. Additional Derivations for the Reco Algorithm

This section describes the calculation of the scaling factor α^2, the matching of the leptons to the on- and off-shell W bosons and the derivation of the fictive mass $M_{\nu\mathrm{f}}$.

To determine α^2, another constraint is imposed on the system. Since the mass of the Higgs boson is $M_H = 125$ GeV, its mass is smaller than the mass of two W bosons, which have a mass of $M_W \approx 80$ GeV each. Therefore, one of the W bosons has to be off-shell. The other W boson is now constrained to be on-shell and consequently having a mass of $M_W = 80.379$ GeV [9]. Since the W boson decays into a charged lepton and a neutrino, the conservation of energy and momentum allows the constraint mass to be expressed in terms of the reconstructed four-vectors of the decay products as

$$80 \text{ GeV} = M_{W,\mathrm{on}} = (\alpha^2 E_{\nu\nu} + E_\ell)^2 - (\alpha^2 \vec{p}_{\nu\nu} + \vec{p}_\ell)^2. \tag{E.1}$$

This leads to an expression for α^2, which is

$$\alpha^2 = \frac{M_{W,\mathrm{on}}}{2(E_\ell E_{\nu\nu} - \vec{p}_\ell \vec{p}_{\nu\nu})}. \tag{E.2}$$

The masses of the leptons are neglected. To calculate this expression, the charged leptons have to be matched to the on- and off-shell W bosons. Given that the on-shell W boson is expected to have a higher mass than the off-shell W boson, the invariant mass is calculated for the sum of each charged lepton's four-momentum and that of the dineutrino system. The lepton that yields the higher invariant mass with the dineutrino system is associated with the on-shell W boson, whereas the other is assigned to the off-shell W boson. If (E.2) returns a solution $\alpha^2 > 1$, it is fixed to be $\alpha^2 = 1$. The assumption of the neutrinos being collinear leads to the dineutrino system being massless. This is in contradiction with the mean value of $M_{\nu\nu} = 30$ GeV, that was mentioned before. To account for that, the neutrino associated with the off-shell W boson is assigned a fictive mass $M_{\nu\mathrm{f}}$. To obtain it, the energy of the dineutrino system is expressed as the sum of the energies of

the separate neutrinos as

$$E_{\nu\nu} = E_{\nu,\text{on}} + E_{\nu,\text{off}} = \alpha^2 |\vec{p}_{\nu\nu}| + \sqrt{(1-\alpha^2)^2 |\vec{p}_{\nu\nu}|^2 + M_{\nu\text{f}}^2}. \tag{E.3}$$

The energy is $E_{\nu\nu} = \sqrt{M_{\nu\nu}^2 + |\vec{p}_{\nu\nu}|^2}$, such that the mass constraint from before is conserved. Therefore, only events with $M_{\nu\nu} = 30$ GeV have a fictive mass that is non zero. This leads to the expression

$$M_{\nu\text{f}} = \sqrt{(E_{\nu\nu} - \alpha^2 |\vec{p}_{\nu\nu}|^2)^2 - (1-\alpha^2)^2 |\vec{p}_{\nu\nu}|^2}. \tag{E.4}$$

F. Evaluation of the Reconstruction Algorithms

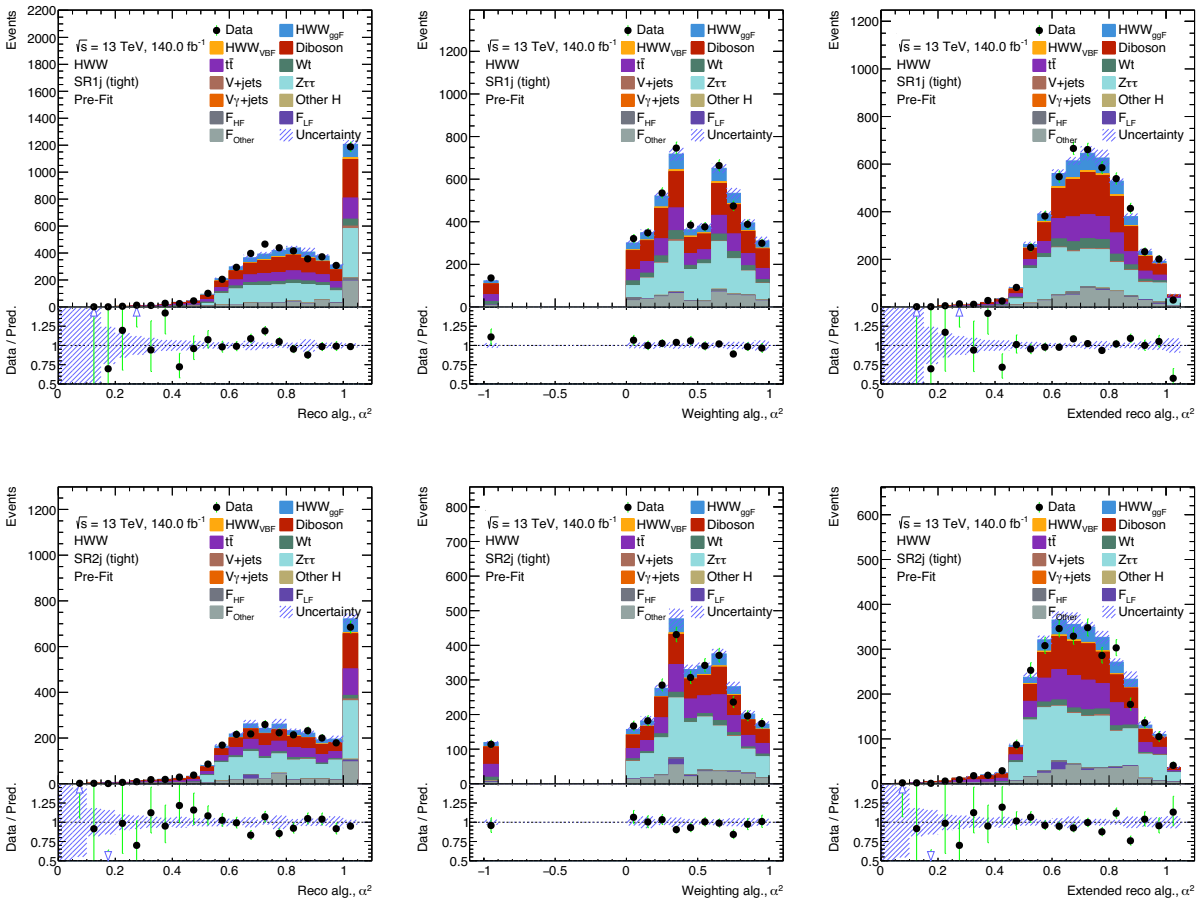

Figure F.1.: The figure shows the scaling factor α^2 for SR1j (top) and SR2j (bottom) for all algorithms.

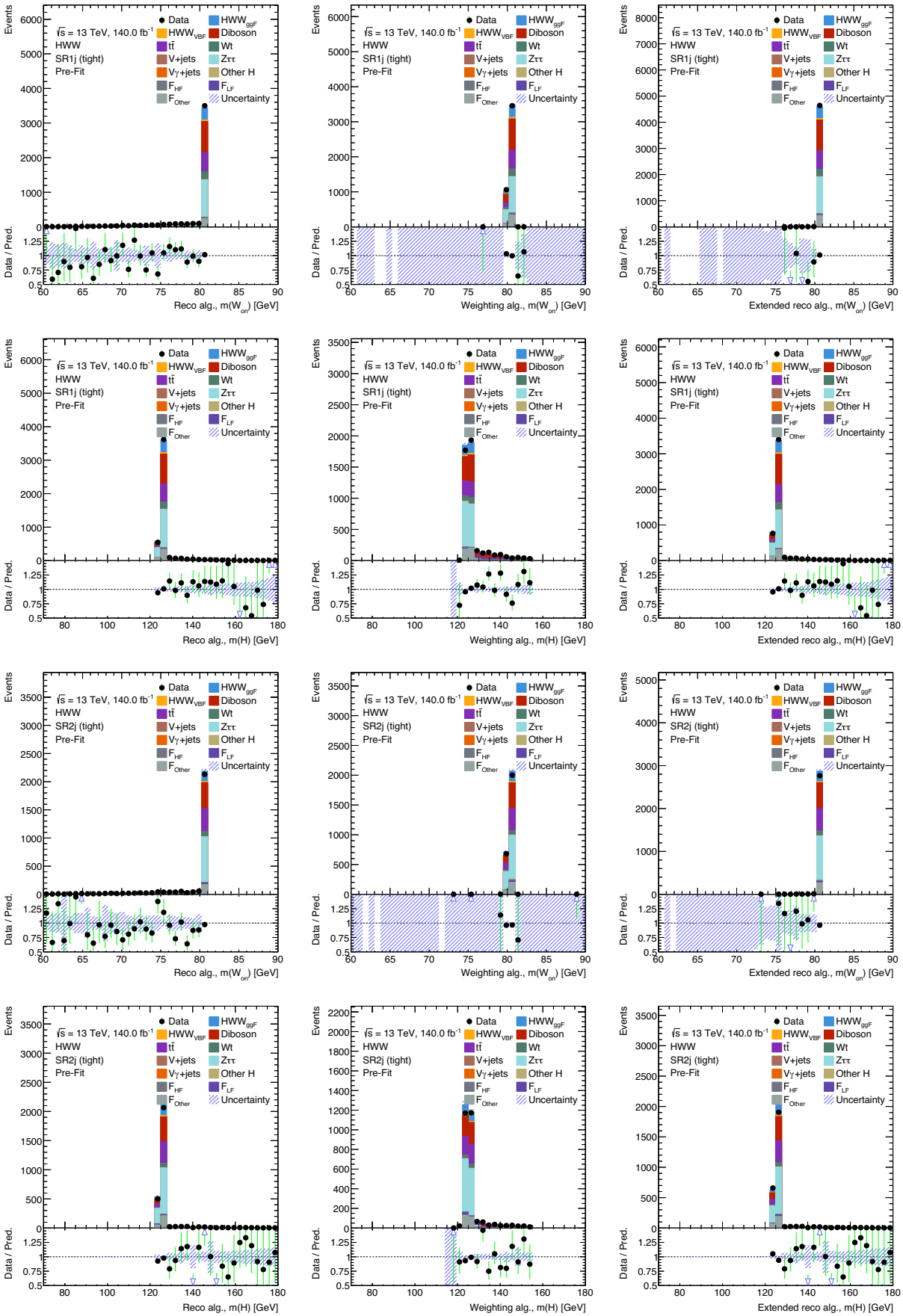

Figure F.2.: The reconstructed masses of the on-shell W boson and the Higgs boson are shown for SR1j and SR2j for all algorithms.

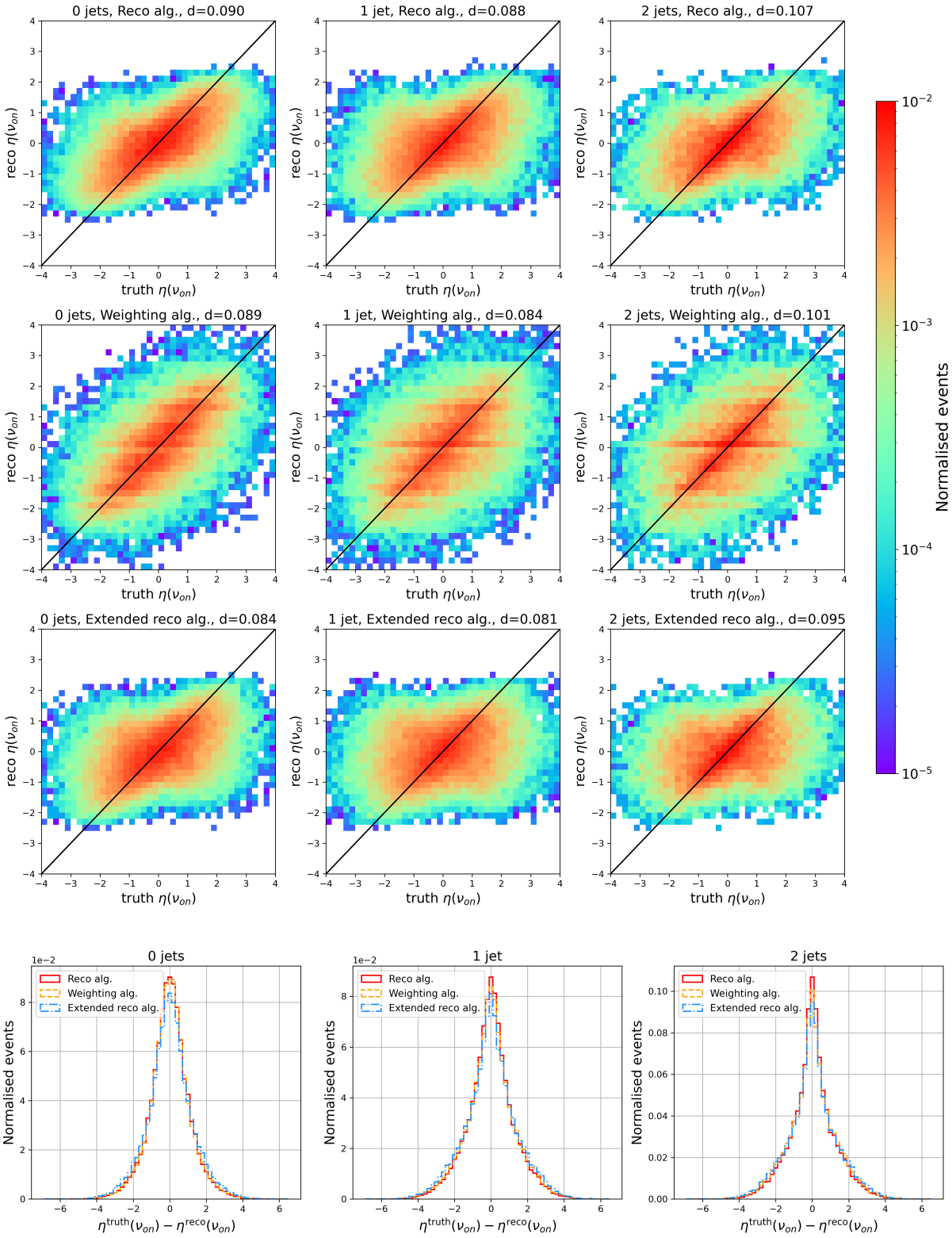

Figure F.3.: This figure shows the distribution of the difference between the reconstructed pseudo-rapidity, η, of ν_{on} and its corresponding truth-level value (bottom). At the top, the reconstructed values are plotted against the corresponding truth values. All algorithms and jet regions are shown, and the calculated diagonality values are listed.

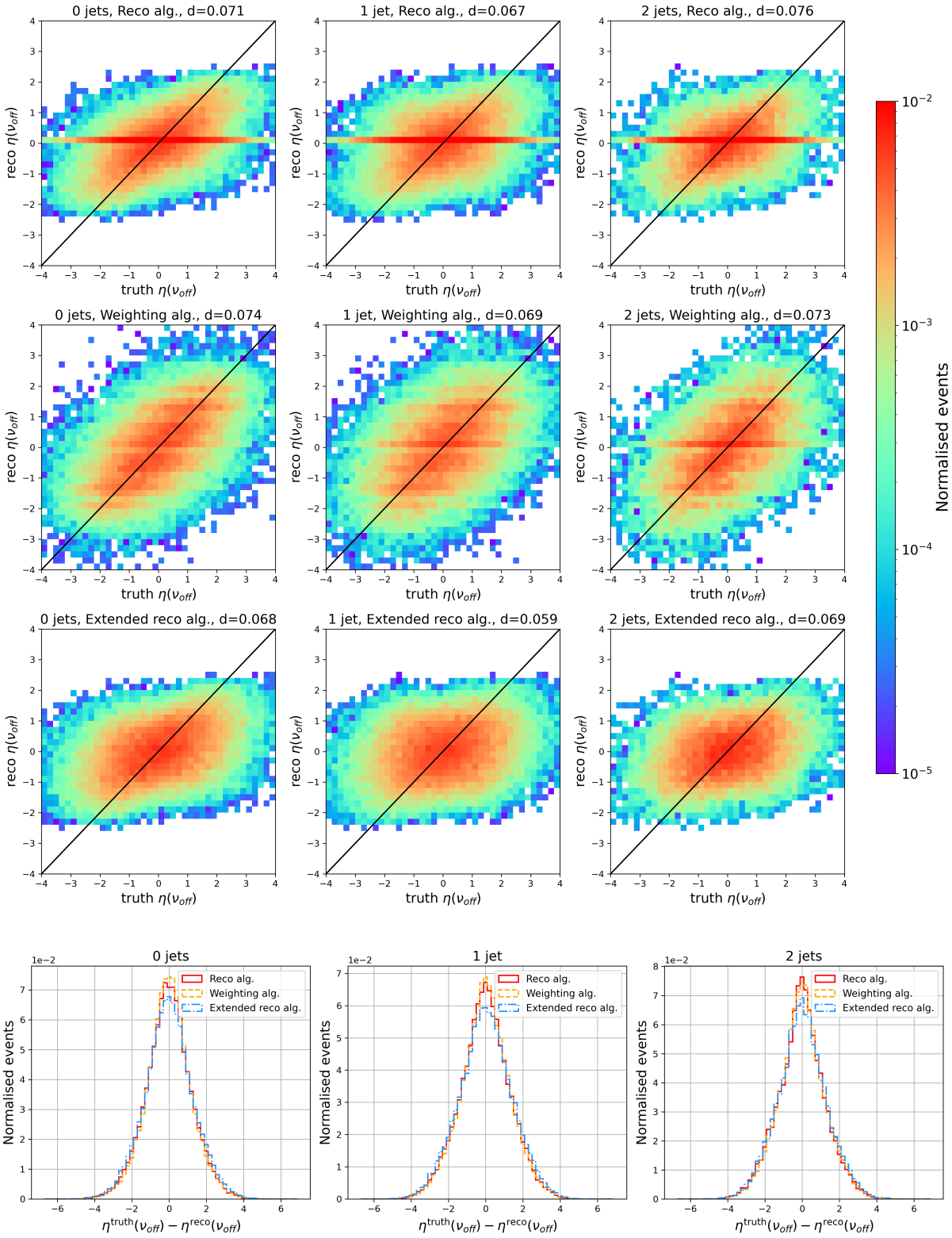

Figure F.4.: The figure shows the distribution of the difference between the reconstructed pseudo-rapidity, η, of ν_{off} and its corresponding truth-level value (bottom). At the top, reconstructed values are plotted against the corresponding truth values. All algorithms and jet regions are shown, and the calculated diagonality values are listed.

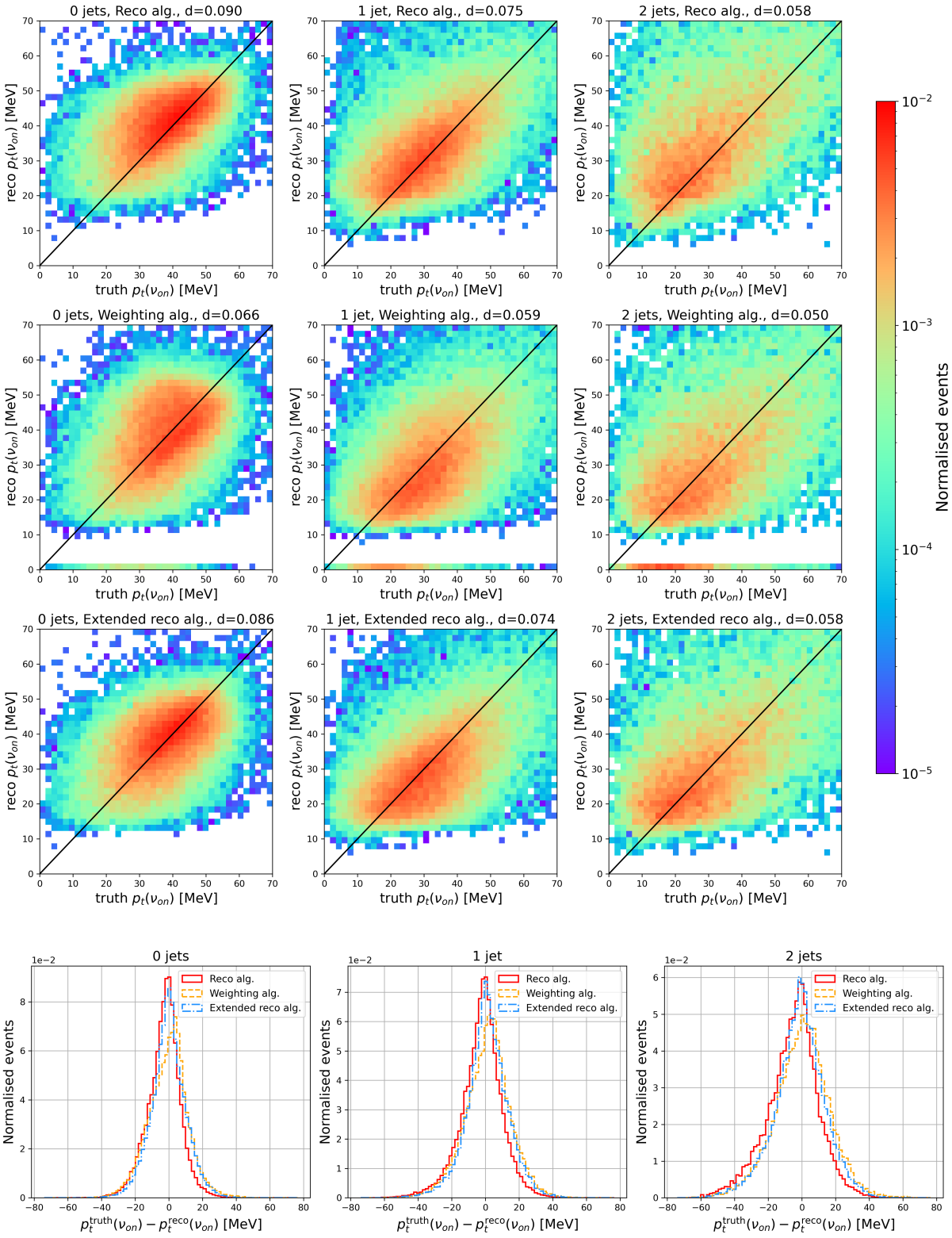

Figure F.5.: The distribution of the difference between the reconstructed transverse momentum, p_t, of ν_{on} and its corresponding truth-level value are shown (bottom). At the top, reconstructed values are plotted against the corresponding truth values. All algorithms and jet regions are shown, and the calculated diagonality values are listed.

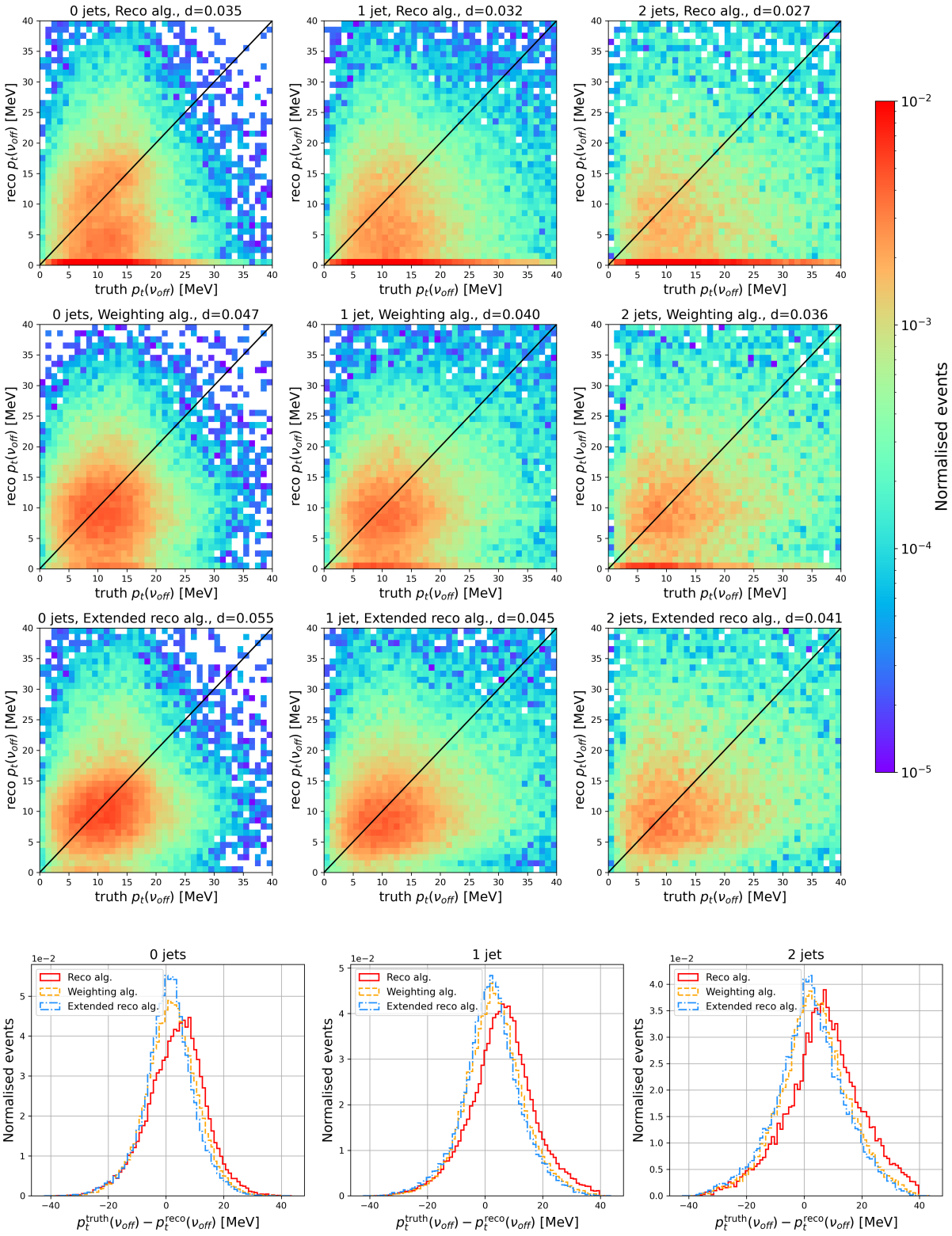

Figure F.6.: The figure displays the distribution of the difference between the reconstructed transverse momentum, p_t, of ν_{off} and its corresponding truth-level value at the bottom. At the top, reconstructed values are plotted against the corresponding truth values. All algorithms and jet regions are shown, and the calculated diagonality values are shown.

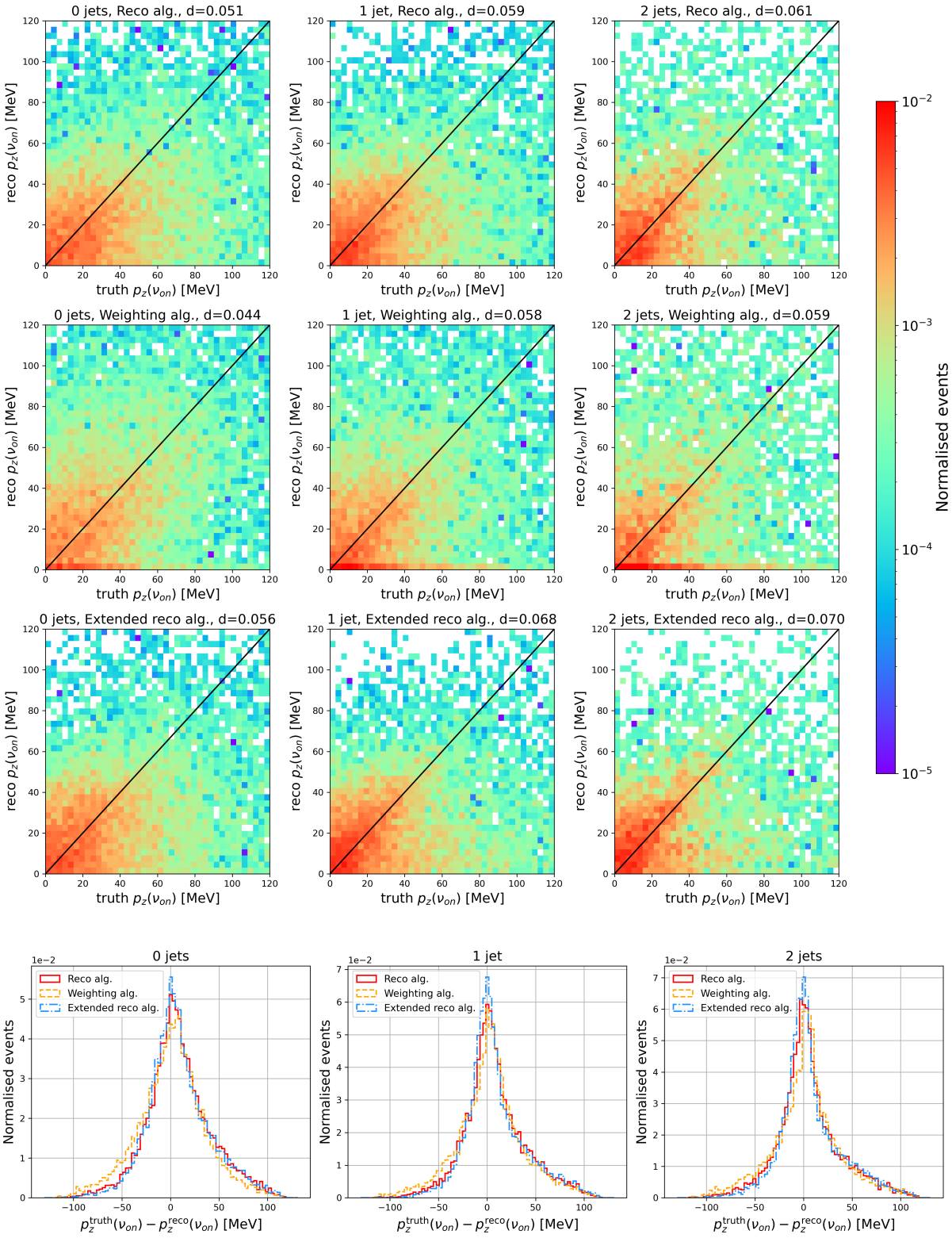

Figure F.7.: The distribution of the difference between the reconstructed longitudinal momentum, p_z, of ν_{on} and its corresponding truth-level value are shown (bottom). At the top, reconstructed values are plotted against the corresponding truth values. All algorithms and jet regions are shown, and the calculated diagonality values are listed.

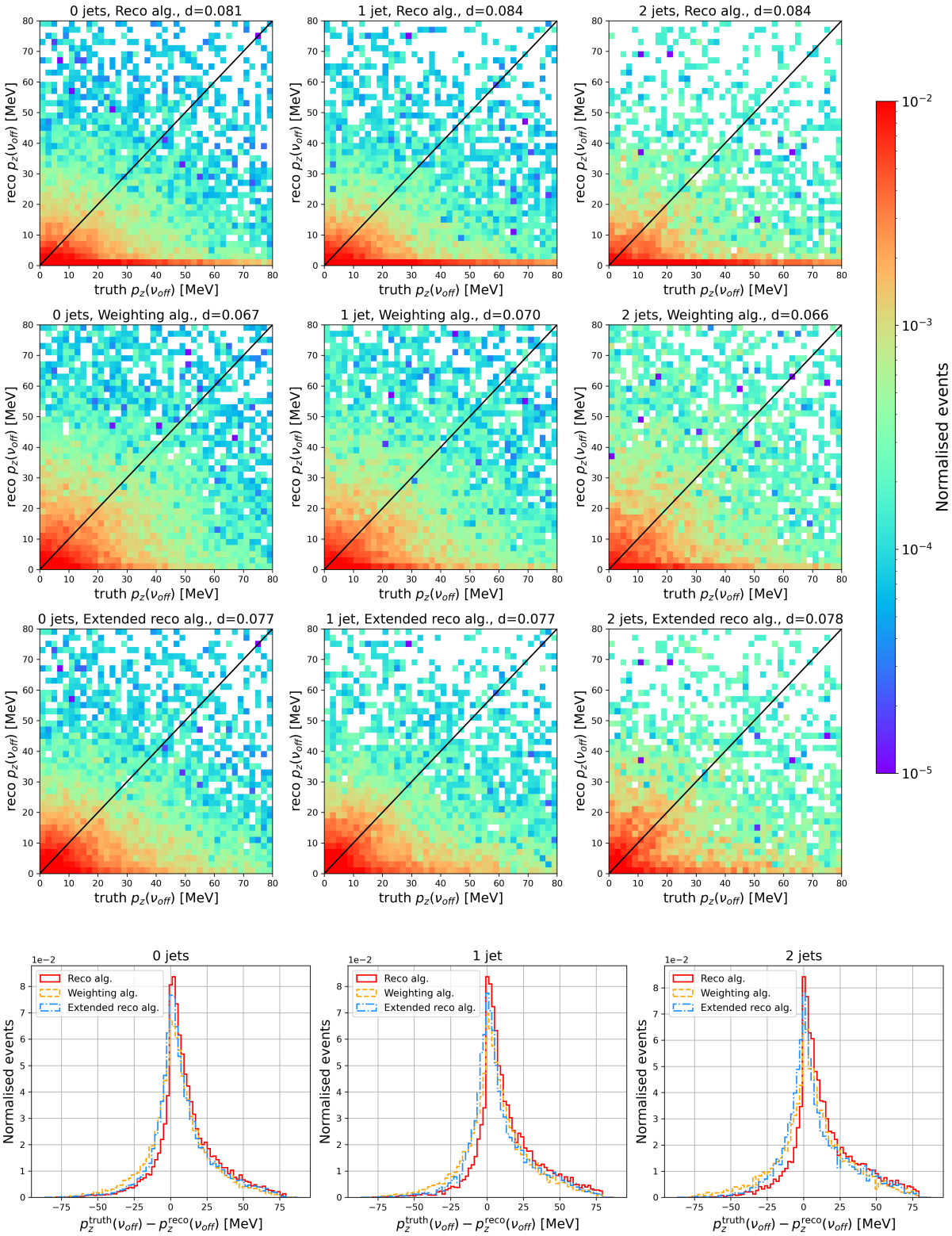

Figure F.8.: The distribution of the difference between the reconstructed longitudinal momentum, p_z, of ν_{off} and its corresponding truth-level value are shown (bottom). At the top, reconstructed values are plotted against the corresponding truth values. All algorithms and jet regions are shown, and the calculated diagonality values are listed.

Bibliography

[1] M. Born, W. Heisenberg, P. Jordan, *Zur Quantenmechanik. II.*, Z. Phys. **35(8-9)**, 557 (1926)

[2] A. Einstein, B. Podolsky, N. Rosen, *Can quantum mechanical description of physical reality be considered complete?*, Phys. Rev. **47**, 777 (1935)

[3] J. S. Bell, *On the Einstein-Podolsky-Rosen paradox*, Physics **1**, 195 (1964)

[4] S. J. Freedman, J. F. Clauser, *Experimental Test of Local Hidden-Variable Theories*, Phys. Rev. Lett. **28**, 938 (1972)

[5] A. Aspect, J. Dalibard, G. Roger, *Experimental Test of Bell's Inequalities Using Time-Varying Analyzers*, Phys. Rev. Lett. **49**, 1804 (1982)

[6] A. Zeilinger, et al., *Violation of Bell's Inequality under Strict Einstein Locality Conditions*, Phys. Rev. Lett. **81**, 5039 (1998)

[7] ATLAS Collaboration, *Observation of quantum entanglement with top quarks at the ATLAS detector*, Nature **633(8030)**, 542 (2024)

[8] A. J. Barr, *Testing Bell inequalities in Higgs boson decays*, Phys. Lett. B **825**, 136866 (2022)

[9] Particle Data Group, *Review of particle physics*, Phys. Rev. D **110(3)**, 030001 (2024)

[10] M. Kobayashi, T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49**, 652 (1973)

[11] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13**, 508 (1964)

[12] E. Lyndon, P. Bryant, *LHC Machine*, JINST **3**, S08001 (2008)

[13] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **83(10)**, 982 (2023)

[14] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3**, S08003 (2008)

[15] B. Andersson, et al., *Parton Fragmentation and String Dynamics*, Phys. Rept. **97**, 31 (1983)

[16] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl. Phys. B **238**, 492 (1984)

[17] S. Frixione, P. Nason, C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11**, 070 (2007)

[18] Sherpa Collaboration, *Event Generation with Sherpa 2.2*, SciPost Phys. **7(3)**, 034 (2019)

[19] T. Sjöstrand, et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**, 159 (2015)

[20] GEANT4 Collaboration, *GEANT4 - A Simulation Toolkit*, Nucl. Instrum. Meth. A **506**, 250 (2003)

[21] ATLAS Collaboration, *Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment*, JHEP **05**, 162 (2024)

[22] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at* $\sqrt{s}$ *= 13 TeV*, Eur. Phys. J. C **79(8)**, 639 (2019)

[23] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **81(7)**, 578 (2021)

[24] M. Cacciari, G. P. Salam, G. Soyez, *The anti-*k_t *jet clustering algorithm*, JHEP **04**, 063 (2008)

[25] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, Eur. Phys. J. C **76(11)**, 581 (2016)

[26] ATLAS Collaboration, *Transforming jet flavour tagging at ATLAS* (2025)

[27] ATLAS Collaboration, *Measurements of differential cross sections of Higgs boson production through gluon fusion in the $H \to WW^* \to e\nu\mu\nu$ final state at $\sqrt{s} =$ 13 TeV with the ATLAS detector*, Eur. Phys. J. C **83(9)**, 774 (2023)

[28] A. Barakat, P. Bianchi, *Convergence and Dynamical Behavior of the ADAM Algorithm for Nonconvex Stochastic Optimization*, SIAM Journal on Optimization **31(1)**, 244 (2021)

[29] R. Z. Aben, *Spinning the Higgs: Spin and parity measurement of the discovered Higgs-like boson in the $H \to WW \to \ell\nu\ell\nu$ decay mode*, Ph.D. thesis, Amsterdam U. (2015)

Acknowledgements

Zuerst möchte ich mich bei Arnulf Quadt für die Möglichkeit bedanken, meine Bachelorarbeit in seiner Arbeitsgruppe schreiben zu dürfen. Außerdem bedanke ich mich bei Jörn Groß-Knetter, dass er die Zweitkorrektur meiner Bachelorarbeit übernimmt.

Furthermore I would like to thank the whole AG Quadt for their warm welcome and especially all participants of the top meetings for your time and constructive comments. Ganz besonders möchte ich mich bei Steffen Korn bedanken, ohne den diese Bachelorarbeit nicht möglich gewesen wäre. Danke, dass du den Rahmen meiner Arbeit gesetzt, mich angeleitet und dir immer Zeit für all meine Fragen genommen hast. Danke auch für das Korrekturlesen meiner Bachelorarbeit.

Außerdem möchte ich meiner Familie und meinen Freunden danken, die mich immer unterstützen. Danke Mama, dass du immer für mich da bist. Danke Vivi, dass du mich schon so lange begleitest und für mich da bist. Danke Miri, dass du die letzten drei Jahre an meiner Seite warst und mich bedingungslos angefeuert, aufgefangen und unterstützt hast.

Erklärung nach §13(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen:
Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.
Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 21. September 2025

(Neele Bäumann)