Bachelor's Thesis

Suche nach Quantenverschränkung in Dileptonischen $H \to WW^*$ Endzuständen mit ATLAS bei $\sqrt{s} = 13\,\text{TeV}$

Exploring Quantum Entanglement in Dileptonic $H \to WW^*$ Final States with ATLAS at $\sqrt{s} = 13\,\text{TeV}$

prepared by

Nico Schiel

from Braunschweig

at the II. Physikalischen Institut

Zusammenfassung

Nach dem Standardmodell sind W-Bosonenpaare, die bei Higgs-Zerfällen entstehen, miteinander verschränkt. Diese Eigenschaft wurde noch nicht experimentell getestet. Um die Quantenverschränkung in W-Bosonenpaaren zu messen, ist es sinnvoll, die Hypothese eines separablen Zustands auszuschließen. Die Standardmodell-Hypothese wird in dieser Arbeit als Pseudodaten behandelt. Da Elektronen und Myonen präzise gemessen werden können, wird der dileptonische Zerfallskanal mit einem Elektron und einem Myon im Endzustand verwendet. Für eine Analyse von Quantenverschränkungen müssen sensitive Variablen mit einem deutlichen Unterschied in separablen und verschränkten Zuständen gefunden werden. Die in dieser Arbeit untersuchten Observablen sind m_{ll}, ϕ_{ll} und η_{ll}. Für die Signal- und Untergrundtrennung wird ein dichtes neuronales Netzwerk verwendet. Um die Hypothesen zu vergleichen, wird eine Profil Likelihood Entfaltung und ein χ^2 Hypothesentest durchgeführt.

Die empfindlichste Observable ist m_{ll}. Unter der Verwendung von nur statistischen Unsicherheiten (einschließlich systematischer Unsicherheiten) kann die Hypothese eines separierbaren Zustands mit 6σ $(4,3\sigma)$ ausgeschlossen werden. Die Observable ϕ_{ll} ist weniger empfindlich gegenüber Quantenverschränkung mit Ausschlussgrenzen von $5,7\sigma$ $(3,7\sigma)$ für den separablen Zustand unter der Verwendung von nur statistischen Unsicherheiten (einschließlich systematischer Unsicherheiten). Die am wenigsten empfindliche untersuchte Observable ist η_{ll}. Die Ausschlussgrenze für einen separablen Zustand in dieser Observable beträgt weniger als 1σ.

Abstract

According to the Standard Model, W boson pairs produced in Higgs decays are entangled. This property has not yet been probed experimentally. To probe quantum entanglement in W boson pairs, it is suitable to first exclude the hypothesis of a separable state. The Standard Model hypothesis is treated as pseudo data in this thesis. Since electrons and muons can be measured precisely the dileptonic decay channel with one electron and one muon in the final state is used. To investigate quantum entanglement, sensitive variables with a clear difference between the separable and entangled state must be known. The studied observables are m_{ll}, ϕ_{ll} and η_{ll}. For the signal and background separation, a dense neural network is used. To compare the two hypotheses, a binned profile likelihood unfolding and a χ^2 hypothesis test is performed.

The most sensitive observable is m_{ll}. It provides exclusion limits for the hypothesis of a separable state of 6σ (4.3σ) using only statistical uncertainties (including systematic uncertainties). The observable, ϕ_{ll}, is less sensitive to quantum entanglement with 5.7σ (3.7σ) exclusion limits for the separable state using only statistical uncertainties (including systematic uncertainties). The least sensitive studied observable is η_{ll}. The exclusion limit for a separable state using this observable is less than 1σ.

Contents

Contents

1. Introduction

Entangled particles cannot be described independently of each other. From a classical point of view, this is an unintuitive behaviour. This is why Einstein, Podolsky and Rosen argued in 1935 that quantum mechanics must be incomplete [1]. According to them, entanglement violates the principle of locality. As a solution, deterministic hidden variables were proposed. In 1964 John Bell showed theoretically, through Bell's inequality, that quantum entanglement allows stronger correlations than possible with the theory of hidden variables [2, 3]. Since then, entanglement has been confirmed as a property of quantum mechanical systems in many experiments [4–7] establishing a crucial difference between quantum mechanics and classical physics. Previously, these experiments were focused on low-energy scales. However, in recent years, entanglement at the highest energy scales at the LHC has become a popular research topic. One of the highlights of this research was the observation of entanglement in top quark pairs at the LHC by the ATLAS and by the CMS collaboration [8, 9]. These measurements confirm entanglement as a quantum mechanical property of top quarks at a high energy scale. Despite this progress, further tests of quantum entanglement in different particles and at various energy scales are necessary to test the rules of quantum mechanics and the Standard Model of particle physics.

This work focuses on the entanglement of $W^{\pm}$ boson pairs, produced in decays of on-shell Higgs bosons. The aim is to study sensitivity to quantum entanglement in the dileptonic decay channel in the ATLAS run 2 data. The first part of the thesis introduces the Standard model of particle physics and the experimental setup with a focus on the ATLAS detector. After that, the $H \rightarrow WW^*$ process and especially quantum entanglement the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ final state are explained. Chapter 5 describes and shows the results of the event simulation. In addition, the object definitions used and event selection are defined and evaluated in Chapter 6. The following chapter then presents the analysis strategy and the methods used in this thesis. In Chapter 8, the research results are presented. Finally, the results are summarized and an outlook for further analyses is given.

2. The Standard Model of Particle Physics

The Standard Model of particle physics (SM) is a theory that describes the known elementary particles and their interaction through three of the four fundamental forces. This includes the electromagnetic, the weak and the strong interaction, but not gravity. The SM further contains the Higgs mechanism, which explains the masses of the elementary particles [10–16]. The SM is a renormalisable, relativistic quantum field theory. It has a $U(1)_Y \times SU(2)_L \times SU(3)_C$ gauge symmetry [10–14] and predicts the behaviour of the elementary particles with great precision on various energy scales [17]. But it is not a complete theory. It does not explain several observed phenomena like dark matter and gravitation [18].

2.1. The Elementary Particles

The elementary particles of the SM are shown in Figure 2.1. The spin S divides the particles into fermions and bosons. Fermions have the spin $S = 1/2$ and bosons have an integer spin. In addition to the spin, the electric charge Q in units of the elementary charge e and the third component of the weak isospin I_3 are properties that characterise the elementary particles.

The SM consists of 12 fermions and their corresponding antiparticles, which have the same mass but inverted quantum numbers. For example, the antiparticles have a reversed sign of the charge. The fermions are separated into 6 leptons and 6 quarks and are organised into 3 generations of matter. Each generation consists of two leptons and two quarks. The left-handed fermions of each generation are grouped into weak isospin doublets with $I_3 = \pm 1/2$. Each right-handed fermion forms a weak isospin singlet with $I_3 = 0$. For the quarks, each doublet consists of an up-type quark with $I_3 = 1/2$ and an electric charge of $Q = 2/3$. The second quark in the doublet is the down-type quark with $I_3 = -1/2$ and $Q = -1/3$. The lepton isospin doublet consists of a charged particle with $I_3 = -1/2$ and $Q = -1$ and a neutrino with $I_3 = 1/2$ and $Q = 0$. The charged leptons

Standard Model of Elementary Particles
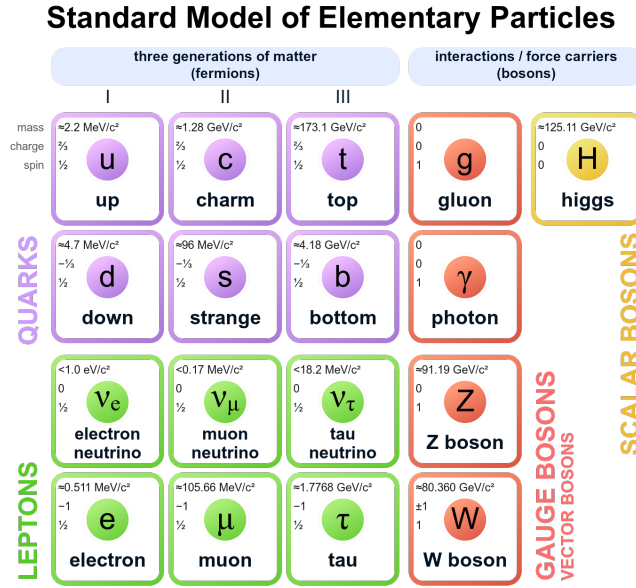
Figure 2.1.: An overview of the elementary particles of the SM.

and the quarks have a mass, while the neutrinos are massless in the SM. Even though it is experimentally known that at least one neutrino has a mass since neutrino oscillation was observed [19, 20]. In each generation, the masses of the particles increases [17].

The five gauge bosons are the mediators of the fundamental forces. They are vector bosons with a spin $S = 1$. The photon mediates the electromagnetic force, the gluons the strong force and the Z^0 and $W^\pm$ bosons the weak force. Furthermore, there is the scalar Higgs boson with $S = 0$ [17].

2.2. The Interactions

The strong interaction is described by the theory of quantum chromodynamics (QCD) which has an $SU(3)_C$ symmetry group. The eight gluons mediate the strong force. The gluons couple to colour-charged particles as referred to by the C in $SU(3)_C$ [11, 12]. All gluons are massless and carry no electric charge and no isospin. There are three different colour charges, blue, green and red. In addition, there are the corresponding anticolours for the antimatter particles. Colour is only carried by the quarks and by the gluons. Each gluon carries a colour and an anticolour. In addition, gluons themselfs interact because $SU(3)_C$ is non-Abelian [11, 12]. A fundamental property of the strong force is confinement, which states that colour-charged particles always appear in colour-neutral bound states. These bound states are mesons consisting of a quark and an antiquark as well as baryons and antibaryons consisting of three quarks or three antiquarks. When

quarks are separated, a new quark-antiquark pair is created, which creates new bound states [21]. The coupling constant α_s depends on the energy. It decreases with increasing energy. This behaviour of the strong force is called asymptotic freedom [22].

The electroweak force is the unification of the electromagnetic and the weak force. This electroweak force is described by a $U(1)_Y \times SU(2)_L$ symmetry. Quantum electrodynamics (QED) describes the electromagnetic force. It has a $U(1)_Y$ gauge symmetry. The Y stands for the weak hypercharge $Y = 2(Q - I_3)$. It has one gauge boson, the massless photon. The photon couples to the electric charge and the photon itself is electrically neutral. Thus, there is no self-interaction. The photon is colourless and has $I_3 = 0$ [17, 23–25].

The weak force is described by the quantum flavour dynamic (QFD). It has a $SU(2)_L$ gauge symmetry. Since $SU(2)_L$ has 3 generators, three gauge bosons mediate the weak force. These are the massive Z^0 and $W^\pm$ bosons [17]. In the following, Z boson and W boson are written for the gauge bosons. The correct charges are implied. The W bosons have a weak isospin of $I_3 = \pm 1$ and an electric charge of $Q = \pm 1$. It mediates the charged weak currents and couples only to left-handed particles and right-handed antiparticles as emphasized by the L in the symmetry group. The vertex factor of the weak interaction for leptons and W bosons [13, 14],

$$i\frac{g_W}{\sqrt{2}}\gamma^\mu\frac{1-\gamma^5}{2},\tag{2.1}$$

describes this. Where γ^μ are the gamma-matrices and g_W is the coupling strength. The operator $\frac{1-\gamma^5}{2}$ projects a particle on the left-handed part of its state. It is called the left-handed chirality projector. The interaction of a W boson with a quark allows flavour changes between the three generations. The Cabibbo-Kobayashi-Maskawa-matrix (CKM-matrix) elements V_{ij} [26] accounts for the transition of an up-type quark of the flavour i to a down-type quark of the flavour j. The vertex factor for quarks interacting with a W boson is the same as the vertex factor of the lepton interaction in Equation (2.1) multiplied with the CKM-matrix element V_{ij} [26]. The Z boson is electrically neutral and has a weak isospin of $I_3 = 0$. It mediates the neutral weak current. In contrast to the W boson, the Z boson cannot mix the flavour of the quarks. Further, it can couple to charged right-handed particles, although it couples to them differently than to the left-handed particles [10].

The masses of the gauge bosons Z and W violate the gauge invariance of the SM. As a solution, the Higgs field and the Higgs boson are introduced through the process of electroweak symmetry breaking [15, 16]. The ATLAS and CMS collaborations at the LHC discovered the Higgs boson in 2012 [27, 28]. It is an electrically neutral, massive

Table 2.1.: Cross-section of the most important production modes of the Higgs boson at pp collisions at the LHC at $\sqrt{s} = 13\,\text{TeV}$ [17].

Process	ggf	VBF	$t\bar{t}H$	WH	ZH
Cross-section [pb]	$48.6^{+5.6\,\%}_{-7.4\,\%}$	$3.78^{+2.1\,\%}_{-2.1\,\%}$	$0.50^{+6.6\,\%}_{-7.4\,\%}$	$1.37^{+2.0\,\%}_{-2.0\,\%}$	$0.88^{+4.1\,\%}_{-3.5\,\%}$

and colourless scalar boson with $I_3 = -\frac{1}{2}$, which couples to the masses of the elementary particles [17]. Since the Higgs boson is massive, it also couples to itself. The Yukawa coupling of the Higgs field to the fermions explains the masses of the fermions [10, 29].

2.3. Higgs Production at the LHC Run 2

The most dominating production mode of the Higgs boson at a pp collider at $\sqrt{s} = 13\,\text{TeV}$ is the gluon-gluon-fusion (ggf) [17]. In this process, two gluons initiate a virtual loop of quarks in which the Higgs boson is produced. Since the top quark is the heaviest quark and couples the strongest to the Higgs boson, this production mechanism is dominated by loops with top quarks. Other production modes at the LHC are the vector-boson-fusion (VBF). In this mode, two Z bosons or two W bosons with opposite charges are emitted by two quarks and fuse together to form a Higgs boson. A Higgs boson can also be produced in association with a top anti-top pair ($t\bar{t}H$). A W or Z boson, which emerges in a quark anti-quark annihilation, can radiate a Higgs boson. This process is called Higgs Strahlung. The cross-sections are listed in Table 2.1.

2.4. Quantum Entanglement of a two-Particle State

A state of a quantum system of multiple particles is described in the tensor product of the one-particle Hilbert spaces. A state on this product Hilbert space is called entangled if it is not separable. In this case, it is not possible to describe the particles independently of each other. The density matrix is often used to describe a many-particle state and entanglement. A density matrix ρ on a two-particle Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ describes a separable state if it can be written as,

$$\rho = \sum_i m_i \rho_{A,i} \otimes \rho_{B,i}.$$

Where the m_i are coefficients and the $\rho_{A,i}$ and $\rho_{B,i}$ are density matrices on the one-particle Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$.

3. Experimental Setup

High energetic particle collisions are required to measure processes involving a Higgs boson. These particle collisions are provided by the large hadron collider and recorded by the ATLAS detector. Both experiments are located at CERN near Geneva.

3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) is a synchrotron with a circumference of approximately 27 km. The accelerator is primarily used for symmetric proton-proton collisions, but it is also used for lead-lead collisions.

The particles are first accelerated in several smaller linear or circular accelerators before they are injected into the LHC. The accelerated proton beams are divided into bunches of protons. These bunches collide in one of the four detectors at the LHC. A collision of the proton beam happens every 25 ns [30]. The second run of the LHC had a centre of mass energy of $\sqrt{s} = 13$ TeV [30].
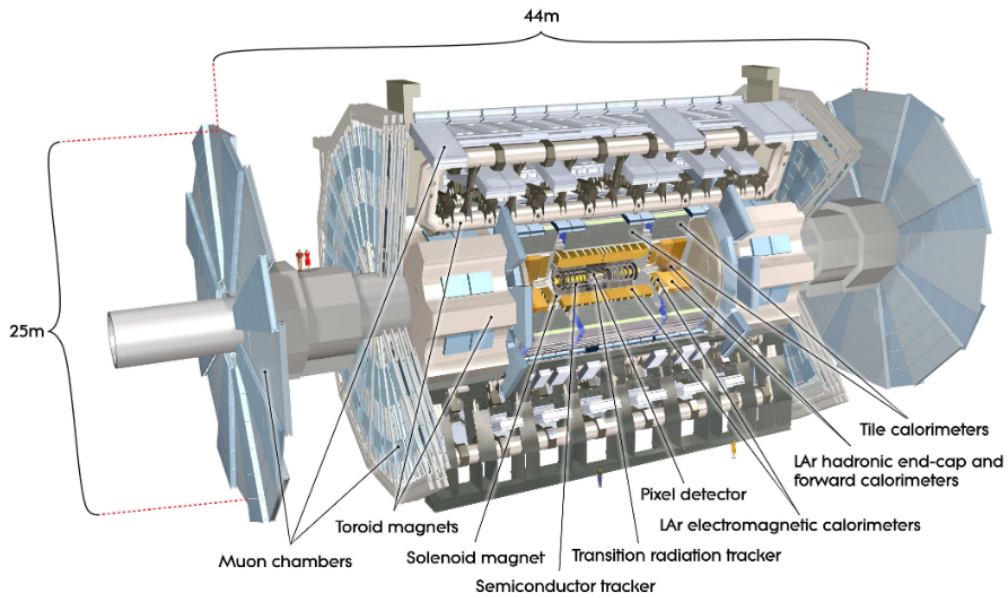
3.2. The ATLAS Detector

The ATLAS detector is a multipurpose detector at the LHC. It has a cylindrical shape and an end cap on each detector side to close the cylinder. The height of the detector is 25 m and the width is 44 m. The detector can roughly be divided into the tracking detector, the calorimeters, the magnets and the muon spectrometers [31]. At the second run of the LHC the ATLAS detector had an integrated Luminosity of $L_{\text{int}} = 140.1(12) \, \text{fb}^{-1}$ [30].

The origin of the coordinate system of ATLAS is the interaction point of the collision. The z-axis is oriented along the beamline, the x-axis is directed to the centre of the LHC and the y-axis points upwards.

The xy-plane is orthogonal to the beam pipe. The transverse momentum is therefore defined as,

$$p_{\text{T}} = \sqrt{p_x^2 + p_y^2},$$

Figure 3.1.: A computer generated picture of the ATLAS detector. The different layers are labelled. ©CERN

the momentum orthogonal to the beam pipe. Here p_x and p_y are the momentum in the x and y direction. The colliding protons at the LHC do not have any transverse momentum. The total transversal momentum after the collision is zero since momentum is conserved. The measured transversal momentum can be non-zero for an event. This is the case if particles are not measured by the detector. The negative of the total measured transverse momentum is defined as the missing transverse momentum. At high energies, the missing transverse momentum is also called the missing transverse energy.

Instead of Euclidean coordinates, polar coordinates can be used. In these coordinates ϕ is the azimuthal angle and θ is the polar angle in the xy-plane. Instead of the polar angle, often the pseudorapidity,

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right),$$

is used. The pseudorapidity is the relativistic limit of the rapidity,

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right),$$

when the mass is negligible compared to the energy $E \gg m$. Here p_z is the momentum in the z direction. Differences in the pseudorapidity $\Delta\eta$ and rapidity Δy are Lorentz

invariant under Lorentz transformations along the *z* axis.

The ATLAS detector consists of individual subdetectors. These subdetectors are arranged in cylindrical layers, shown in Figure 3.1. Each subdetector fulfils special tasks.

The inner part of the ATLAS detector, the vertex and tracking detector, is positioned directly around the beam pipe. It consists of a pixel detector, a semiconductor detector and a transition radiation tracker [31]. The tracking detector measures the trajectories of charged particles, such as electrons and muons. The Inner Detector is embedded in a solenoid magnet [31], which produces a magnetic field inside the tracking detector. This magnetic field bends the trajectories of the electrically charged particles inside the detector for momentum measurements.

The next layer are the calorimeters, which measure the energy of the particles. The calorimeter is divided into an electromagnetic and a hadronic calorimeter. The electromagnetic calorimeter measures the energy of electrons and photons through electromagnetic interactions such as bremsstrahlung and pair production. ATLAS uses a liquid Argon calorimeter with lead as the passive medium [31]. Since the masses of muons and charged hadrons are large compared to the electron mass, they lose less energy through bremsstrahlung. Therefore, they do not shower in the electromagnetic calorimeter. Electrically neutral particles, except the photon, do not interact with the electromagnetic calorimeter. Using a trace in the tracking detector and a particle shower in the electromagnetic calorimeter electrons can be identified.

The hadronic calorimeter further out measures the energy and position of hadrons through the strong interaction. It consists of a Tile Calorimeter around the beam pipe with plastic scintillators as the active medium and steel as the passive medium [31]. A liquid Argon calorimeter is used at the end caps.

The outermost part of the detector is the muon spectrometer. There are additional toroid magnets between the muon chamber and the hadronic calorimeter [31]. The magnetic field is used in the muon spectrometers to determine the momentum of muons. The magnetic field turns the muons in a different direction as in the Inner Detector. With the information about the tracks in the Inner Detector and the muon spectrometer, the momentum and energy of the muon are precisely reconstructed. The muon identification through the spectrometer is accurate because most other particles stop in the calorimeters [31].

The neutrinos do not interact with the detector at all. Therefore, they cannot be identified directly. They are reconstructed, via the missing transverse momentum.

The collision rate at the *L*HC is 40 MHz [30]. The high data rate that is processed is too large for all events to be stored. To filter the data for interesting events which are

eventually saved a two-level trigger system is used in Atlas. The rate of events which pass the trigger is approximately 1 kHZ. The Level-1 trigger is hardware based and it makes its decision in less than 2.5 µs. It takes information from the calorimeters and the muon chamber and retains high energetic events. The software-based Level-2 trigger makes more complex decisions. In addition to the information from the calorimeters and the muon chambers, information from the Inner Detector is used. The decision time is under 200 µs [31].

4. The $H \to WW^* \to \ell\nu\ell\nu$ Process

The Higgs boson is an unstable particle with a lifetime of approximately $\tau_H = 2 \times 10^{-22}\,\text{s}$ [17] and a measured mass of $m_H = (125.25 \pm 0.17)\,\text{GeV}$ [17]. One of the decay modes of the Higgs bosons is the decay in two W bosons. The W has a measured mass of $m_W = (80.0377 \pm 0.00012)\,\text{GeV}$ [17]. Therefore, the mass of two W bosons is larger than the mass of one Higgs boson. As a result, one of the W bosons has to be an off-shell virtual particle. The W boson has a lifetime of approximately $\tau_W = 3 \times 10^{-25}\,\text{s}$ [17]. Thus, the W bosons decay before they are detected directly. The decay products are therefore used to analyse the WW^* state. It can decay leptonically into a charged lepton and the corresponding neutrino from the same isospin doublet. The τ leptons decay further, for example into electrons or muons and two neutrinos, before they are measured in the detector. τ leptons are therefore experimentally challenging. Hence, only electrons and muons are treated as signal in this analysis. Due to their leptonic decay, τ decays are a background to the electrons and muons. The resulting neutrino cannot be measured directly in the detector. The kinematic properties of the charged leptons $\mu^\pm$ and $e^\pm$ as energy and momentum can, however, be precisely measured in the detector as described in Section 3.2. The decay of the W bosons into quarks is not used since the hadronic background is large in hadron colliders.

4.1. Entanglement in $H \to WW^* \to \ell\nu\ell\nu$

The following section is based on references [32, 33]. The Higgs boson is a spin-0 particle and it decays into two spin-1 particles. Due to angular momentum conservation, both W bosons have the same polarisation state. Consequently, both W bosons are either left-handed, right-handed or longitudinally polarised. Hereinafter W_1 is the on-shell and W_2 is the off-shell W boson. Let (x_H, y_H, z_H) be the coordinate system of the Higgs bosons rest frame. Choose the z_H axis such that it points in the direction of the momentum of the W_2 boson. The y_H-axis is chosen such that the momentum of the incoming protons and the momentum of the W bosons lies in the $y_H z_H$-plane. The x_H-axis is defined as $x_H = y_H \times z_H$. Furthermore let (x_i, y_i, z_i) be the coordinate system of the rest frame

4. The $H \to WW^* \to \ell\nu\ell\nu$ Process
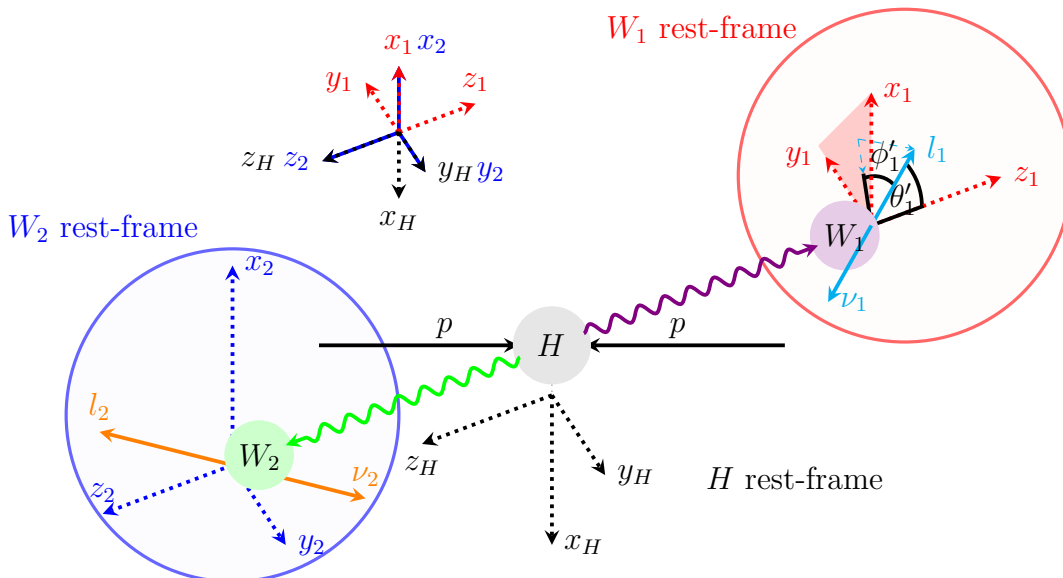
Figure 4.1.: Coordinate System used in the analysis. The parts in the blue and red circles are the rest-frames of the W bosons. Outside the circle is the rest-frame of the Higgs boson. The angles are only shown for the lepton l_1.

of the W_i boson. Denote with $\hat{x}$, $\hat{y}$ and $\hat{z}$ the unit vector in the direction of the x-axis, y-axis and z-axis. Then the axes of the rest frame of the W_1 are defined by $\hat{x}_1 = -\hat{x}_H$, $\hat{y}_1 = -\hat{y}_H$ and $\hat{z}_1 = -\hat{z}_H$. The axes of the rest frame of the W_2 are defined by $\hat{x}_2 = -\hat{x}_H$, $\hat{y}_2 = \hat{y}_H$ and $\hat{z}_2 = \hat{z}_H$. The coordinate systems are shown in Figure 4.1.

The spin density matrix of the two W bosons can be determined as [32, 33]

$$
\rho = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{6}(\sqrt{2}A^1_{2,0}+2) & 0 & \frac{1}{3}C_{2,1,2,-1} & 0 & \frac{1}{3}C_{2,2,2,-2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{3}C_{2,1,2,-1} & 0 & \frac{1}{3}(1-\sqrt{2}A^1_{2,0}) & 0 & \frac{1}{3}C_{2,1,2,-1} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{3}C_{2,2,2,-2} & 0 & \frac{1}{3}C_{2,1,2,-1} & 0 & \frac{1}{6}(\sqrt{2}A^1_{2,0}+2) & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.
$$

The C_{L_1,M_1,L_2,M_2} and $A^i_{L_i,M_i}$ are constant depending on the polarisation state of the W_i bosons with $L_i = 1, 2$ and $-L_i \leq M_i \leq L_i$. The helicity state of the W bosons can be written in terms of the coefficients a_{-1-1}, a_{11} and a_{00}. Here a_{-1-1} decribes the left-handed helicity state, a_{11} the right-handed helicity state and a_{00} the longitudinal helicity state. Assuming CP is conserved and performing a calculation at leading order, only the

following coefficients [32, 33],

$$A_{2,0}^1 = A_{2,0}^2 = \frac{1}{\sqrt{2}N}(|a_{11}|^2 + |a_{-1-1}|^2 - 2|a_{00}|^2),$$

$$C_{2,2,2,-2} = C_{2,-2,2,2}^* = \frac{3}{N}a_{11}a_{-1-1}^*,$$

$$C_{2,1,2,-1} = -C_{1,1,1,-1} = -C_{1,-1,1,1}^* = C_{2,-1,2,1}^* = \frac{3}{2N}(a_{11}a_{00}^* + a_{00}a_{-1-1}^*),$$

$$C_{1,0,1,0} = \frac{-3}{2N}(|a_{11}|^2 + |a_{-1-1}|^2),$$

$$C_{2,0,2,0} = \frac{1}{2N}(|a_{11}|^2 + |a_{-1-1}|^2 + 4|a_{00}|^2),$$

are non-zero. Here $N = (|a_{11}|^2 + |a_{-1-1}|^2 + |a_{00}|^2)$ is a normalisation factor. Additionally, $B_1 = \pm\sqrt{2\pi}$ and $B_2 = \sqrt{\frac{2\pi}{5}}$ are defined. The upper sign is for the case that $W_1 = W^+$ and the lower for the case that $W_1 = W^-$. With the spherical harmonics Y_L^M the angular distribution for the density operator ρ is determined as [32, 33]

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_1\mathrm{d}\Omega_2} = \frac{1}{4\pi}\Big[1 + A_{L_1M_1}^1 B_{L_1}Y_{L_1}^{M_1}(\theta_1', \phi_1') + A_{L_2M_2}^2 B_{L_2}Y_{L_2}^{M_2}(\theta_2', \phi_2')$$
$$+ C_{L_1,M_1,L_2,M_2}B_{L_1}B_{L_2}Y_{L_1}^{M_1}(\theta_1', \phi_1')Y_{L_2}^{M_2}(\theta_2', \phi_2')\Big],$$

where summation over indices is implied. The angles θ_i' and ϕ_i' are the polar and azimuthal angles of the charged lepton in the rest frame of the W_i boson.

In the separable state, all non-diagonal elements vanish in the density matrix. Hence, $C_{2,1,2,-1} = 0$ and $C_{2,2,2-2} = 0$. Consequently are $a_{11} = a_{-1-1} = 0$. The only non-vanishing coefficients in this case are $A_{2,0}^1 = A_{2,0}^2 = -\sqrt{2}$ and $C_{2,0,2,0} = 2$. The W bosons are therefore exclusively longitudinally polarized. In the separable case, the angular distribution simplifies to,

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_1\mathrm{d}\Omega_2} = \frac{1}{16}\Big[4 - (3\cos^2(\theta_1') - 1) - (3\cos^2(\theta_2') - 1)$$
$$+ (3\cos^2(\theta_1') - 1)(3\cos^2(\theta_2') - 1)\Big]. \tag{4.1}$$

Since the angular distribution for the separable state is independent of ϕ_i, each value of ϕ_i is assumed to be equally likely. The separable hypothesis, that the state only contains longitudinally polarised W bosons can be compared, to the SM hypothesis that the state also contains different polarizations.

Since the neutrinos cannot be measured precisely as described in Section 3.1, observables
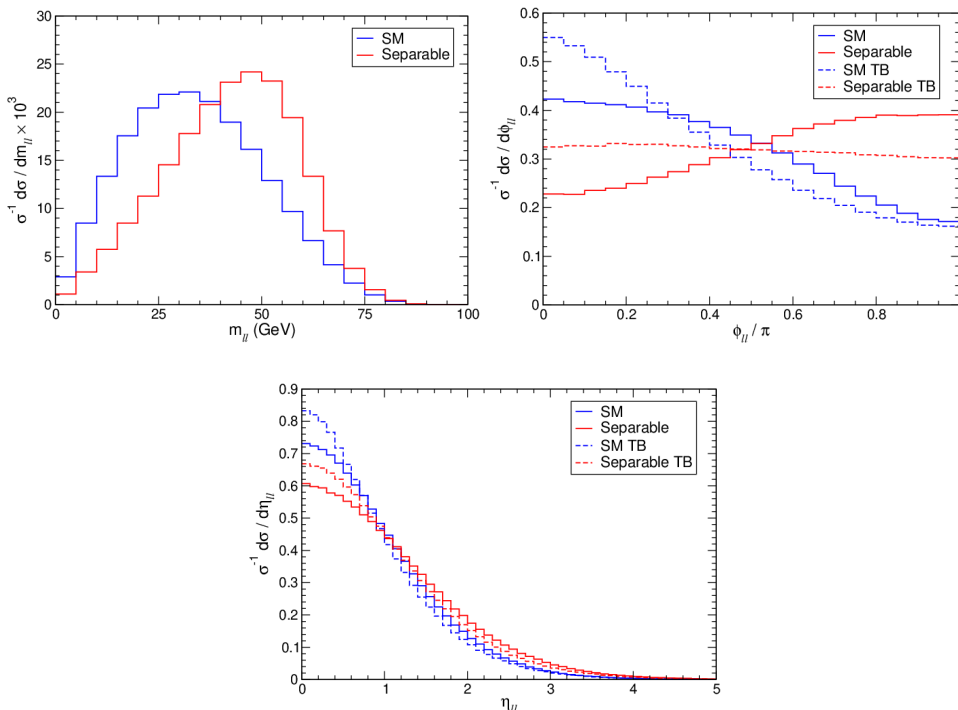
Figure 4.2.: Simulated observables for an entangled and a separable state. The variables marked with TB are for a Higgs boson boosted with $p_t^H = 20\,\text{GeV}$. The solid lines are for the Higgs boson at rest [32]. The top left shows m_{ll}, the top right ϕ_{ll} and the bottom η_{ll}.

which depend solely on the charged leptons are instead favourable for the analysis. Especially, the invariant mass m_{ll} of the charged leptons, the absolute value of the difference in azimuthal angle ϕ_{ll} between the momentum of the charged leptons in the laboratory frame and the modulus of the difference in the pseudorapidity η_{ll} in the laboratory frame of the two charged leptons are relevant observables. These observables can be determined as shown in Figure 4.2. The dashed lines marked with TB are for a Higgs boson with a transversal momentum of $p_T^H = 20\,\text{GeV}$ in the laboratory frame. The solid lines are for the Higgs boson at rest. Especially the invariant mass m_{ll} and the angle between the leptons ϕ_{ll} show a clear difference between the separable and the SM case. Therefore, these variables are well suited for studying entanglement. The difference in η_{ll} is less clear. This observable is less likely sensitive to the difference between the non-separable and separable state.

In this thesis, the experimental sensitivity to quantum entanglement of the 3 observables m_{ll}, ϕ_{ll} and η_{ll} in the ATLAS run 2 data will be studied.

5. Event Generation

To compare measured data to theory predictions in particle physics simulations of the studied processes are used. One collision in the detector is called an event. Many events are simulated to have larger statistics. An event can be simulated at different levels. These are explained in the following chapter.

5.1. Monte Carlo Event Generators

At first, the parton level is simulated. Parton level events are simulated with *Monte Carlo* (MC) generators. MC generators simulate the hard particle scattering in the accelerator. The parton level is also called truth information because all information of every particle involved in the hard interaction process is known. The probability for a certain process is given by SM calculations. For these calculations, a given order of perturbation theory is used. This level considers no further showering processes and decays after the hard interaction.

Next is the particle level. Here the effects of the parton shower and other decays are considered. The parton shower describes the emission of gluons and quarks from the decay products and the initial proton-proton state. The emerging quarks hadronise and form bound states such as mesons and baryons. The hadronisation can not be calculated with perturbation theory. Instead, heuristic models like the Lund-String [34] or clustering models [35, 36] are used. The hadronisation is simulated in independent shower generators. The decay products from one initial parton are clustered into jets. Consequently, a particle level event does not include the initial products of the hard scattering process.

The MC generators used for the different signal and background samples are shown in Table A.1 in the Appendix.

5.2. Detector Simulations

The detectors used in particle physics are limited in their detection and resolution capabilities. These are not included in the MC simulations. In the last step of the event

generation, the effects of the detector are simulated, according to the detector used in the experiment. This level is called the reconstruction (reco) level. The MC data at particle level is turned into a simulated detector signal. To simulate the ATLAS detector Geant4 is used [37].

5.3. Custom Angle Replacement

In the MC sample of the HWW_{ggf} process, the polarisations are given by the SM distribution and not by the separable state. With a technique called *custom angle replacement* (CAR) the desired polarisation that describes a separable state can be produced [38], by modifying the angles according to the angular distribution of the separable state in Equation (4.1). The used coordinate systems are described in Section 4.1.

At first, the four-momenta of the leptons in the final state are calculated in the laboratory frame. After that, a 2-step Lorentz transformation into the rest frame of the respective W_i boson, via the Higgs rest frame, is performed. In this, the four-momenta of the leptons are determined again. Then the new $\cos\theta_i'$ and ϕ_i' variables are randomly drawn from the angular distribution of the separable state according to Equation (4.1). Since the distribution is independent of ϕ_i', the angle is randomly drawn from a uniform distribution. These are the new polar coordinates of the charged lepton in the W_i boson rest frame. The direction of the momentum of the charged lepton gets rotated, such that it points in the new direction (θ_i', ϕ_i'). The new direction of the momentum of the neutrino is in the opposite direction of the charged lepton in the W boson rest frame. This does not change the total energy and the total momentum. In the last step, the leptons are transformed back into the laboratory frame to determine the new four-momenta. [38]

The CAR algorithm is applied to each event of the truth level SM sample to obtain the separable state at truth level. The separable events at truth level are therefore the SM events, with a changed angular distribution of the leptons in the W bosons rest frame. The events of the separable state are also called CAR events.

5.4. Pseudo-Reconstruction CAR Sample

Currently, there are no CAR samples at reconstruction level because no dedicated theory samples are available for simulations. However, later in this analysis, a reconstruction CAR sample is needed. For this, a *pseudo-reconstruction* CAR sample is produced. This pseudo-reconstruction CAR sample uses no complete detector simulation since this would be outside the scope of this bachelor thesis. Despite that, it is still usable in the signal

background classification to reduce the bias towards the SM or the CAR sample. Apart from the signal background classification, the pseudo-reconstruction CAR sample is not used in the analysis, due to its incompleteness with respect to a proper detector simulation. If the CAR sample is shown in a figure it is denoted as such. In graphics HWW_{ggf} always means the SM sample.

The pseudo-reconstruction CAR sample is produced with the truth information of the CAR and SM sample and the reconstruction SM simulation. The idea is that the momentum changes from truth to reconstruction level are the same for SM and CAR sample. At first, the leptons at reconstruction and truth level in the SM simulation are matched by lepton type. If that is not possible they are matched by minimizing in

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$$

where $\Delta\phi$ and $\Delta\eta$ are the difference in the azimuthal angle and the pseudorapidity between the lepton at truth and at reconstruction level. The 3 momenta of the pseudo-reconstruction CAR sample $\mathbf{p}_{\text{reco}}^{\text{CAR}}(l_i)$ of the ith lepton l_i is then calculated with the truth three-momentum of the CAR and SM sample $\mathbf{p}_{\text{truth}}^{\text{CAR}}(l_i)$ and $\mathbf{p}_{\text{truth}}^{\text{SM}}(l_i)$ and the reconstruction level three-momentum of the SM sample $\mathbf{p}_{\text{reco}}^{\text{SM}}(l_i)$ by

$$\mathbf{p}_{\text{reco}}^{\text{CAR}}(l_i) = \mathbf{p}_{\text{reco}}^{\text{SM}}(l_i) + \mathbf{p}_{\text{truth}}^{\text{CAR}}(l_i) - \mathbf{p}_{\text{truth}}^{\text{SM}}(l_i).$$

The energy $E_{\text{reco}}^{\text{CAR}}(l_i)$ is then calculated by the four-momentum mass relation with the mass at the reconstruction level of the ith lepton $m_{\text{reco}}^{\text{SM}}(l_i)$ by

$$E_{\text{reco}}^{\text{CAR}}(l_i) = \sqrt{(m_{\text{reco}}^{\text{SM}}(l_i))^2 + (\mathbf{p}_{\text{reco}}^{\text{CAR}}(l_i))^2}.$$

Additionally the cuts in lepton acceptance in the pseudorapidity and transverse momentum in the object definition in Section 6.1 of $|\eta| < 2.5$ for muons and $|\eta| < 1.32$ and $1.52 < |\eta| < 2.47$ for electrons are implemented. The additional jets in the CAR sample are the same as in the SM simulations. The vector of the missing transverse energy $\mathbf{E}_{\text{T, reco}}^{\text{Miss, CAR}}$ is then calculated with the transverse momentum $\mathbf{p}_{\text{T}}$ of the leptons l_i and jets j_i as

$$\mathbf{E}_{\text{T, reco}}^{\text{Miss, CAR}} = -\left(\sum_{i=1}^{2} \mathbf{p}_{\text{T reco}}^{\text{CAR}}(l_i) + \sum_{i=1}^{N_{\text{jets}}} \mathbf{p}_{\text{T}}(j_i)\right).$$

The notation with superscript and subscript variables is as before. Here N_{jets} is the number of jets in the event. If $N_{\text{jets}} = 0$ the sum for the jets is dropped. The modulus of the vector is the missing transverse energy $E_{\text{T}}^{\text{Miss}}$ and the angle in the transverse plane is
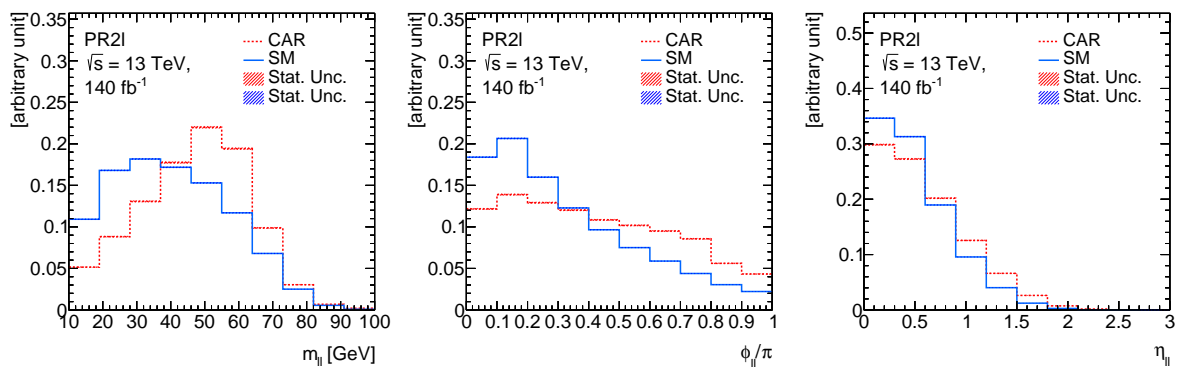
Figure 5.1.: Separation of the SM reconstruction sample and the pseudo-reconstruction CAR sample. On the left for m_{ll}, in the middle for ϕ_{ll} and on the right for η_{ll}.

the angle of the missing transverse energy E_ϕ^{Miss}.

The difference in the distribution of the kinematic variables of the leptons and the missing transverse energy between the SM events and CAR events are shown in Appendix B. The distribution of the pseudorapidity of the leptons is centred around zero for both hypothesises. The CAR sample is narrower than the distribution of the SM events. The difference in ϕ of the leptons is negligible. This is expected since ϕ is uniformly distributed in the SM hypothesis and the CAR algorithm draws ϕ from a uniform distribution. The difference in the transverse momentum of the leading lepton between the two hypotheses is negligible. However, the energy of the leading lepton in the CAR sample is on average smaller than the SM sample. Furthermore, the transverse momentum and the energy of the sub-leading lepton are smaller in the CAR sample than in the SM sample. The leptons in the CAR sample are therefore softer. The missing transverse energy $E_{\mathrm{T}}^{\mathrm{Miss}}$ is on average broader distributed for the CAR events. The shapes of both samples are similar, but the SM distribution has a higher maximum. The observable E_ϕ^{Miss} is uniformly distributed as expected. The separation for the observable m_{ll}, ϕ_{ll} and η_{ll} are shown in Figure 5.1. Especially, m_{ll} shows a significant difference between the SM and CAR events. The maximum of the pseudo-reconstruction CAR distribution is at a higher value compared to the SM events. The difference in the shape of the distribution in ϕ_{ll} is not as large as in m_{ll}. However, the pseudo-reconstruction CAR distribution is more flat than the SM distribution. They are still significantly different. The difference in η_{ll} is minimal in the width of the two distributions. It is broader distributed in the CAR sample than in the reconstruction SM sample.

6. Object Definition & Event Selection

Particles are reconstructed from the signal in the detector. For the same particle, different definitions can be used. The definitions of leptons and jets used in this analysis are explained in the object definition. After that, in the preselection, cuts in the kinematics of the events are explained. The sample is also divided into the preselection regions. A region is a part of the total sample with specific characteristics in the event and object kinematics. Finally, the categorization into signal and control regions is explained in the event classification. Both event selection and classification are important to minimize the number of background events in the signal regions.

6.1. Object Definition

Electrons are reconstructed from their path through the Inner Detector and their energy decomposition inside the electromagnetic calorimeter. The detector limits the pseudorapidity of electrons to $|\eta| < 2.47$ [31]. Furthermore, electrons are not detected in the area between the barrel and the endcap calorimeter, with a pseudorapidity of $1.32 < |\eta| < 1.52$ [31]. Muons are reconstructed via a track in the inner detector where they do not shower. Therefore, they pass through the calorimeters and are measured in the muon chambers. Their pseudorapidity needs to be inside the detector range of $|\eta| < 2.5$. The electrons and muons also require a transverse momentum of at least $p_\mathrm{T} > 7\,\mathrm{GeV}$. For the electron and muon identification a tight likelihood working point is chosen. A tight working point on the lepton isolation reduces the effect of non-prompt leptons. This also works with a likelihood-based identification algorithm [39, 40].

Jets are reconstructed from the energy decomposition in the hadronic calorimeter. For the clustering of the energy decomposition the anti-k_t algorithm is used with a cone width of $R = 0.4$ [41]. The jets are required to have transverse momentum $p_T > 25\,\mathrm{GeV}$ and a pseudorapidity $|\eta| < 2.5$. The `"DL1dv01:Continuous"` algorithm is used for b-tagging. `DL1dv01` is an algorithm that uses track and secondary vertex information to identify b
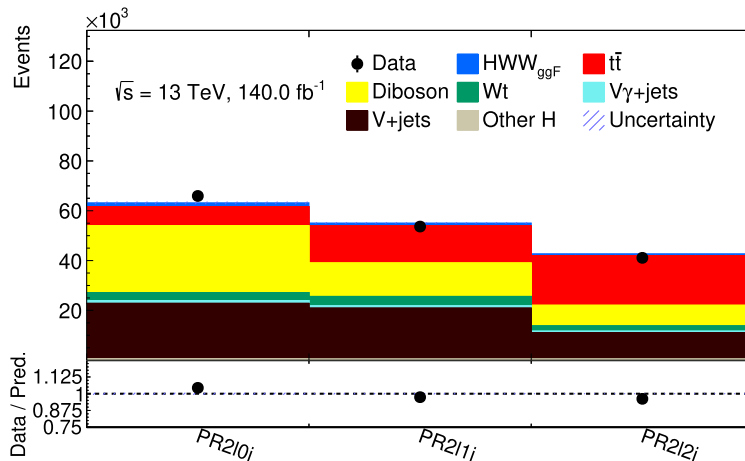
Figure 6.1.: Summary plot of the different signal regions with the cuts according to the preselection. Only statistical uncertainties are shown.

quarks [42]. The continuous refers to the division of events in semi-continuous bins. For example, 85% of all true *b*-jets pass the selection to the 85% efficiency bin.

6.2. Preselection

The preselection aims at reducing the backgrounds. The sample is also divided into regions in the preselection. The backgrounds covered in this analysis are the $t\bar{t}$ background, the diboson background which includes WW and WZ events which do not emerge from Higgs decays. In addition, there are the Vjets and $V\gamma$jets backgrounds. Here V is a W or a Z boson. The Wt process and other decays originating from Higgs bosons are also included. Other relevant backgrounds like fake leptons are not considered in the analysis. The HWW signal only includes gluon-gluon-fusion (ggf) and not the other production channels. The SM hypothesis is used for the HWW_{ggf} sample.

The preselection is based on previous $H \to WW^*$ analyses as in reference [43]. Top quarks almost always decay in a b quark and into a W boson [17]. To account for background with top quarks, zero b-jets with an efficiency of 85% are required. Exactly one electron and one muon are needed to reduce background events with Z boson decays where two electrons or two muons are created. Since the W bosons appear in the Higgs decay they have opposite electrical charges. The leptons must therefore also have opposite charges. Additionally, the leading lepton is required to have a transverse momentum of $p_T(l_1) > 22\,\text{GeV}$ and the sub leading lepton of $p_T(l_2) > 15\,\text{GeV}$. The missing transverse energy is suspected to be non-zero because neutrinos are in the final state. Therefore,

Table 6.1.: Selection for the preselection regions at reconstruction level.

Variable description	reconstruction level cuts		
Preselected regions	PR2l0j	PR2l1j	PR2l2j
Jet multiplicity	$N_{\text{jets}} = 0$	$N_{\text{jets}} = 1$	$N_{\text{jets}} \geq 2$
b-jet multiplicity	$N_{b\text{-jets85\%}} = 0$		
Lepton multiplicity	$N_e = 1$ & $N_\mu = 1$		
Lepton charge	$q_e = -q_\mu$		
Lepton transverse momentum	$p_T(l_1) > 22\,\text{GeV}$ & $p_T(l_2) > 15\,\text{GeV}$		
Missing transverse energy	$E_\text{T}^\text{Miss} > 20\,\text{GeV}$		
Invariant mass leptons	$10\,\text{GeV} < m_{ll} < 100\,\text{GeV}$		

Table 6.2.: Selection for the fiducial region at truth level used in the analysis.

Variable description	truth level cuts
Lepton multiplicity	$N_e = 1$ & $N_\mu = 1$
Lepton charge	$q_e = -q_\mu$
Lepton transverse momentum	$p_T(l_1) > 22\,\text{GeV}$ & $p_T(l_2) > 15\,\text{GeV}$
Missing transverse energy	$E_\text{T}^\text{Miss} > 20\,\text{GeV}$
Invariant mass leptons	$10\,\text{GeV} < m_{ll} < 100\,\text{GeV}$

$E_\text{T}^\text{Miss} > 20\,\text{GeV}$ is used as a cut. Also, $m_{ll} > 10\,\text{GeV}$ is required to remove meson resonances and Drell-Yan background. Additionally, $m_{ll} < 100\,\text{GeV}$ is required to reduce the $t\bar{t}$ background. Since the background varies for different numbers of jets the analysis is divided into three preselection regions. One region with no jets (PR2l0j), one with exactly one jet (PR2l1j) and one with two or more jets (PR2l2j). This has the disadvantage that the size of each sample in each region is smaller. The advantage, however, is that the individual backgrounds in each signal region can be examined in more detail. This is especially helpful since each region has different backgrounds as shown in the summary plot Figure 6.1. The strongest background in PR2l0j are diboson and V jets and in PR2l2j $t\bar{t}$. The three backgrounds are also the strongest in PR2l1j. Table 6.1 summarizes the preselection.

The definition of the fiducial region at truth level matches the one at reconstruction level in the cuts in missing transverse energy and the lepton selections. Definitions including the jet multiplicity are not possible because they are not accessible at parton level for the HWW_{ggf} sample. The selections at parton level are summarised in Table 6.2.

6.3. Event Classification

The selected events in the different preselected regions are classified into signal and control regions. The event classification is divided into two steps. At first, HWW^* events and the strongest backgrounds are classified into signal and control regions. The second step filters out HWW^* events in the signal regions that contain τ-decays. Dense neural networks are used for the event classification. For this reason, the foundations of dense neural networks are explained below.

Basics of Dense Neural Networks

Dense neural networks (DNN) are machine learning tools that train a model on an input sample to make predictions without additional user input. A DNN consists of different layers of nodes. The nodes are individual units in a layer. For each layer, the activation function σ is defined. The input function receives an input $\vec{x}$ from every node in the previous layer and produces an output for every node in the next layer x'. The output of a node is

$$x' = \sigma(\vec{x} \cdot \vec{w} + b)$$

where $\vec{w}$ are the weights of the previous layer and b is the bias of the node. The first layer is called the input layer, and the last is the output layer. Between these two layers are hidden layers, where the input sample is processed. The activation function of the output layer is called the output activation function.

The target is compared with the output of the DNN. The loss function L, that takes the target and the output of the DNN as input, is defined. This loss function returns a positive value, the loss. The training of a DNN refers to optimising the model to classify the input more precisely. To accomplish this goal, the loss is minimised. This is achieved by recalculating the weights and biases in each training iteration. The initial weights and biases are chosen randomly. The vector $\vec{\omega}_l$ stores all weights and biases of one layer l. In the training, the gradient of the loss function $\nabla_{\vec{\omega}_l} L$ for the weights and biases in one layer is calculated, starting with the output layer. The initial weights and biases are then updated into the opposite direction of $\nabla_{\vec{\omega}_l} L$. The step size is given by the learning rate λ. The updated weights and biases of the layers $\vec{\omega}_l{}'$ are then given by

$$\vec{\omega}_l{}' = \vec{\omega}_l - \lambda \nabla_{\vec{\omega}_l} L$$

The other layers are recursively updated similarly from back to front after the output layer. This technique of calculating the gradient of the individual layers from back to

Table 6.3.: Input variables for the DNN. The input variables for the τ classifier are marked with τ and for the signal background classifier with S/B.

variable	description	Classifier
E_T^{Miss}	missing transverse energy	S/B, τ
E_ϕ^{Miss}	azimuthal angle of themissing transverse energy	S/B, τ
$p_T(l_i)$	transverse momentum of the ith lepton	S/B, τ
$\eta(l_i)$	pseudorapidity of the ith lepton	S/B, τ
$\phi(l_i)$	azimuthal angle of the ith lepton	S/B, τ
$E(l_i)$	energy of the ith lepton	S/B, τ
$p_T(j_i)$	transverse momentum of the ith jet	S/B
$\eta(j_i)$	pseudorapidity of the ith jet	S/B
$\phi(j_i)$	azimuthal angle of the ith lepton	S/B
$E(j_i)$	energy of the ith jet	S/B

front recursively is called *backpropagation* [44, 45]. The training is divided into epochs, in which the same steps are repeated. After each iteration through the network, the loss is calculated again. The DNN learns in an epoch when the difference between the loss after two iterations is larger than the positive real number Δ_{min}. If the loss is smaller than Δ_{min} the machine learning model no longer learns significantly. Training repeats until the classifier does not learn something after a certain number of iterations. This number of iterations is called *patience.*

The machine learning model must be validated on a second statistically independent sample. This sample is called the validation sample. The machine learning model is overtrained if it classifies the validation sample worse than the training sample. This is the case, if the loss of the validation sample is greater than that of the training sample. On the other hand, the model is underfitted if the loss of the training sample is larger than the loss of the validation sample.

HWW^* **Signal Classification**

A multiclass classifier based on a DNN is used to separate the sample into control and signal regions. It is trained to distinguish between the signal and the dominating backgrounds diboson, $t\bar{t}$ and V jets. This multiclass classification allows greater control over these backgrounds. A classifier is trained for each preselection region. Therefore, control regions for each dominating background and each jet multiplicity are used. The signal in the training includes both the pseudo-reconstruction CAR and the reconstruction SM sample, to reduce bias towards one of these samples. The results of the pseudo-reconstruction CAR sample are reweighted during the training, such that the yield of CAR and SM

Table 6.4.: Selection for each signal region at reconstruction level used in the analysis. The preselection is as in Table 6.1. The additional selection criteria are written in bold.

Variable description	reconstruction level cuts		
Signal regions	SR2l0j	SR2l1j	SR2l2j
Jet multiplicity	$N_{\text{jets}} = 0$	$N_{\text{jets}} = 1$	$N_{\text{jets}} \geq 2$
b-jet multiplicity	$N_{b\text{-jets}85\%} = 0$		
Lepton multiplicity	$N_e = 1$ & $N_\mu = 1$		
Lepton charge	$q_e = -q_\mu$		
Lepton transverse momentum	$p_T(l_1) > 22\,\text{GeV}$ & $p_T(l_2) > 15\,\text{GeV}$		
Missing transverse energy	$E_{\text{T}}^{\text{Miss}} > 20\,\text{GeV}$		
Invariant mass leptons	$10\,\text{GeV} < m_{ll} < 100\,\text{GeV}$		
HWW^* classifier	$P(HWW)$ is maximal probability		
τ-tagging	$P(e\mu) > 0.35$		

sample are the same in every region. The classifier for each region of 0, 1 and more than 2 jets is trained on the missing transverse energy, the lepton kinematics and jet kinematics. PR2l2j is only trained on the two leading jets. All functions that are used in the training of the classifiers are defined in Appendix C. The used neural networks consist of three layers with 50 nodes each and a learning rate of 0.001. To reduce the effect of overtraining, the classifiers are regularised. In the first two layers the batch normalisation [46] is added, and the last two layers are dropout layers [47] with a dropout probability of 10%. The used optimizer is `Nadam` [48] and the output activation function is `softmax`. `ReLU` and `tanh` are used as activation functions. The patience is 30 and $\Delta_{\text{min}} = 0.001$. The `categorical_crossentropy` is used as the loss function. The multiclass classifier has one output variable for the signal and each background type it is trained on. These variables describe the probability that an event is of the corresponding type according to the model. Thus, the sum of all output variables is one. This is because the `softmax` is used as the output activation function. The training is divided into two statistically independent folds. This means that the sample is divided into two samples. For each sample, a model is trained independently. An event is assigned to the signal HWW_{ggf} or one of the backgrounds B if the corresponding probability $P(HWW)$ or $P(B)$ is the maximal probability for the event. The signal and control regions are denoted similarly to the preselection regions. For example, the signal region for 0 jets is denoted as SR2l0j and the $t\bar{t}$ control regions as CR2l0j($t\bar{t}$).
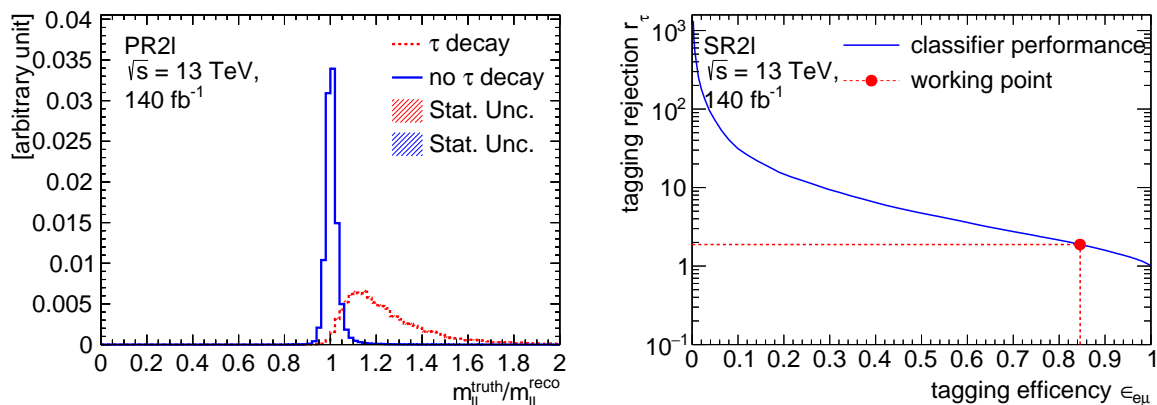
Figure 6.2.: On the left is the ratio of the truth and reconstruction level m_{ll}. On the right is the rejection plotted over the efficiency. The rejection axis is scaled logarithmically. The working point with $\epsilon_{e\mu} = 0.85$ and $r_\tau = 1.88$ is marked with a red dot.

τ Decay Classification

Another background are τ leptons from the W decay, which decay further to the lighter leptons. The m_{ll} of the truth events of the τ-decays is mostly larger than the m_{ll} of the reco events. This can be seen from the ratio between the truth and the reco m_{ll} in Figure 6.2. The distribution of the τ decays has many events where the ratio is larger than 1. The τ-decays influence therefore the studied observable m_{ll}.

Only about $20,000$ events with τ decays are included in these samples. Due to the low statistics of the τ decays in the HWW_{ggf} sample, it is not useful to divide them into signal regions. Therefore, one binary classifier is trained on all signal regions to distinguish between the τ decays and the $e\mu$ events without τ-decay. The binary classifier is only trained on the missing transverse energy and the kinematics of the leptons. The structure of the DNN is the same as for the other classifiers. Only the output activation is changed to `sigmoid` and the loss function to `binary crossentropy`. The classifier has therefore only one output. The probability, that an event does not contain a τ-decay, is $P(e\mu)$.

The tagging efficiency $\epsilon_{e\mu}$ is defined as the probability that an $e\mu$ event is identified as an $e\mu$ event by the classifier. The inverse of the probability, that the classifier identifies an event with a τ-decay as such is the tagging rejection r_τ. For a suitable working point of the classifier, the tagging rejection and the tagging efficiency should be as large as possible. In Figure 6.2, the connection between the tagging rejection and the tagging efficiency is shown. This shows that high tagging rejection and high tagging efficiency can not be reached simultaneously. A working point of $\epsilon_{e\mu} = 0.85$ which corresponds to $r_\tau = 1.88$ is used. This keeps most of the $e\mu$ events while approximately 47% of τ decays

are rejected. An event is therefore classified as an $e\mu$ without τ-decay if $P(e\mu) > 0.35$. There is no additional control region for τ decay. This control region would not receive enough events to improve the understanding of the backgrounds. The selection for the signal regions is summarised in Table 6.2 and for the control regions in Table D.1 in the Appendix.

6.4. Evaluation of the Event Classification

The confusion matrix compares the classifier's prediction to the true label. Each row shows the true label and the column is the label the classifier assigned. A matrix element of the confusion matrix is the probability that the classifier assigns an event one type to the corresponding column type.

The *receiver operator characteristics* (ROC) is the true positive rate plotted against the false positive rate. It describes the separation qualities of the classifier for one output variable. Furthermore, *the area under the ROC curve* (AUC) measures the classifier's overall performance. The higher the AUC, the better the separation quality of a classifier. The AUC_{Nom} is the AUC of the original input and the AUC is the AUC when one input variable is permuted with random noise. With this, the permutation importance is calculated as,

$$\frac{\text{AUC}_{\text{Nom}} - \text{AUC}}{\text{AUC}_{\text{Nom}}}.$$

The permutation importance measures the impact of an input variable on the output.

This thesis considers only the training sample in the permutation importance and confusion matrix. The other parts consider the training and the validation sample.

Evaluation of the HWW^* Signal Classification

In the first step, an event is classified as HWW_{ggf} if the probability, of an event being HWW_{ggf}, namely $P(HWW)$, is larger than any of the background probabilities. To study potential biases of the classifier with respect to SM events or CAR events the two distributions of $P(HWW)$ are compared as shown in Figure 6.3. The classification in all regions slightly favours SM events over CAR events. This manifests itself in general higher $P(HWW)$ values for SM events compared to CAR events. This behaviour is expected since the SM sample differs more from the background than the CAR sample. This is shown in the separation of CAR events and SM events with respect to the background in m_{ll} in Figure 6.4. It is noticeable, that the CAR distribution in PR2l0j has an increasing number of events for small values of $P(HWW)$ and PR2l1j for high values of $P(HWW)$.
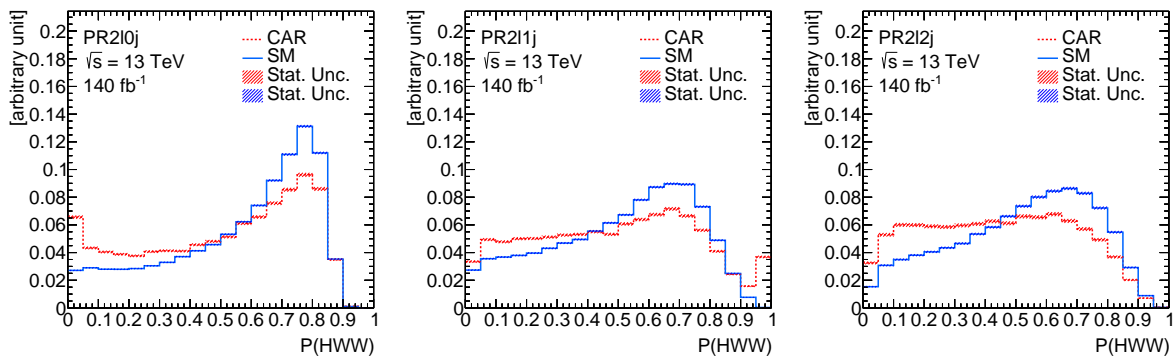
Figure 6.3.: Separation of the pseudo-reconstruction CAR and the reconstruction SM sample in the classifiers output $P(HWW)$. On the left for PR2l0j, in the middle for PR2l1j and on the right for PR2l2j.
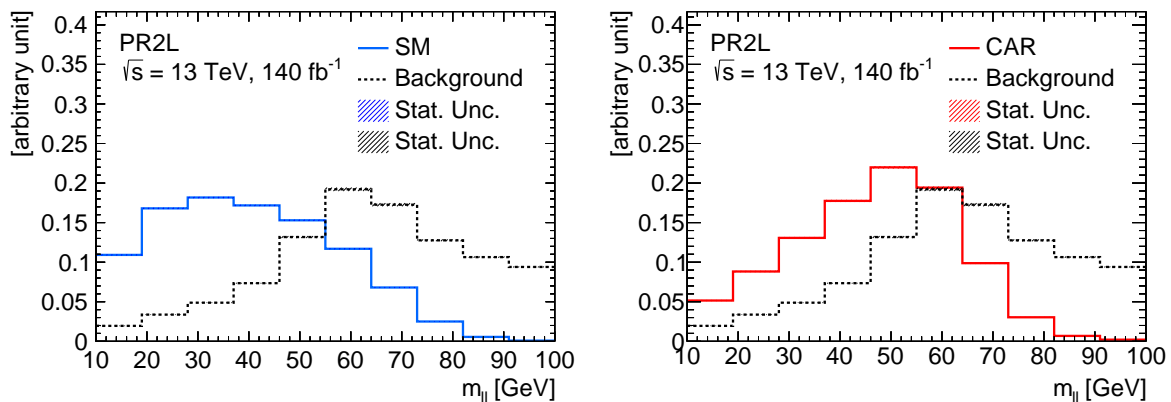
Figure 6.4.: Separation of m_{ll} for the SM and the background on the left and pseudo-reconstruction CAR and background on the right.

This is not the case in the other regions. This could indicate some outliers in these two regions. Figure D.3 in the Appendix D shows the loss for the different folds of the classifiers. The loss of the validation and training sample in PR2l0j shows no difference in the first fold. In the second fold, the validation sample has a negligible higher loss. The loss converges in both cases. The effect of overtraining is therefore negligible for this classifier. In the other two regions, the difference between validation and training samples is also negligible and the loss converges except in the first fold in PR2l2j. The validation sample has a smaller loss in this fold. Thus the model is minimally underfitted. The loss of the validation sample has spikes in the folds in PR2l1j and in the second fold in PR2l2j. This is likely due to the small size of the signal sample in these two regions. The number of events in each region is shown in Figure D.1 in Appendix D. The SM signal in SR2l0j has a sample size of about $200,000$ events compared to about $100,000$ events in PR2l2j. The validation phase space in PR2l2j is therefore only half as large as in PR2l0j. The

27
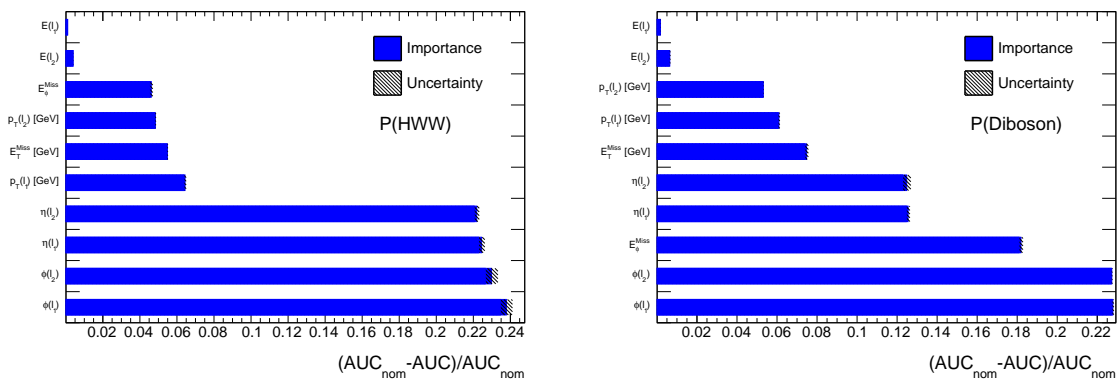
Figure 6.5.: The perturbation importance of the multiclass classifier in PR2l0j of Diboson on the left and HWW_{ggf} on the right.

validation size can be increased by including additional folds to the training. With three folds the sample size of 50% would be increased to 75%.

The permutation importance of the input variables for the training is shown in Figure 6.5 and in Figure D.4, Figure D.5 and Figure D.6 in the Appendix. Both the energy of the leptons and jets are negligible for the training. The other kinematic variables of the jet are also mostly negligible. They are however relevant for the classification of the $t\bar{t}$ and the diboson background events. The importance in the classification of $t\bar{t}$ events is expected since the jets originate most likely from top decays. The angular variables of the leptons $\phi(l_i)$ and $\eta(l_i)$ are important for all classifications and especially for the signal HWW_{ggf} classsification. This is interesting since these influence the behaviour of the observables m_{ll}, η_{ll} and ϕ_{ll} strongly. The other kinematics of the leptons are less important but have a non-negligible contribution. The missing transverse energy, on the other hand, is important. Since many of the processes like HWW_{ggf} and $t\bar{t}$ contain neutrinos, this is to be expected.

The confusion matrix for PR2j0j is shown in Figure 6.6 and in Figure D.2 in the Appendix for PR2l1j and PR2l2j. Additionally, the ROC curves are shown in Figure D.4 for PR2l0j, Figure D.5 for PR2l1j and Figure D.6 for PR2l2j in Appendix D. The classification of the signal HWW_{ggf} is correct for more than 70% of the events in SR2l0j and SR2l1j. The classification in SR2l2j is worse with only 66%. This could be because the jets originate from initial state radiation that influences the signal. The diboson background is the background that is most frequently identified as HWW_{ggf}. More than 20% of the diboson events are classified as signal. This is anticipated since diboson events comprise WW events. The Vjets event classification is the best. Over 80% are assigned correctly and the AUC is larger than 0.9. The diboson and $t\bar{t}$ background have the best classification in CR2l0j with 66% and 45% correctly assigned events. It is also noticeable
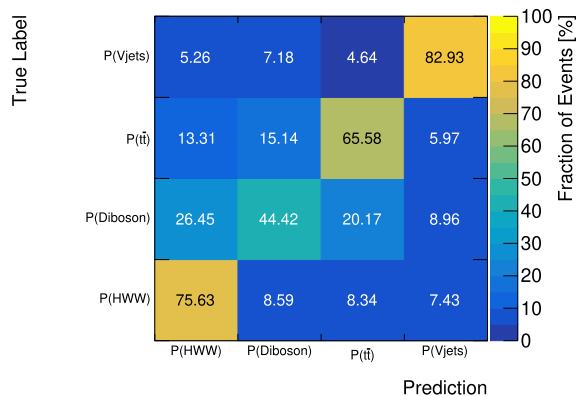
Figure 6.6.: The confusion matrix for the classifier in PR2l0j.

that diboson is more often misidentified as $t\bar{t}$ in the regions with jets. This could again be due to the influence of next-to-leading order processes like initial state radiation.

Evaluation of the τ Decay Classification

The network output is the probability that an event includes no τ-decay $P(e\mu)$. The separation is shown in Figure 6.7 and the working point is marked with a vertical line. If $P(e\mu) > 0.35$ the events are classified as $e\mu$ without τ-decays. The CAR distribution has a larger amount of events in the rejected domain $P(HWW) \leq 0.35$. The SM sample is therefore favoured over the CAR sample. The difference in rejected CAR and SM events is however negligible. As described in Section 6.3 a dense neural network is used to tag and filter the τ decays. The validation loss of both folds is around 0.3 larger than the loss of the training sample. The model is therefore overtrained. The cause is the small number of τ decays in the sample. Only about 20.000 events with τ decays are in the sample. The AUC is small with values around 0.75 as shown in Figure D.7 in the Appendix. The separation quality of the classifier is therefore limited. The reason for this could also be the small sample size. For better training results, a larger sample size is required. The most important variable for the training is E_ϕ^{Miss} as seen in the figures in Figure D.7. The variables of the missing transverse energy are important for the training since additional neutrinos are involved in the τ decays. As before for the other classifiers, the angular variables of the leptons are also important for the training.

Summary of the Event Classification

Figure 6.8 shows the signal and background distribution in the regions. In particular the background of the *V* jets dominates its control regions. The diboson control regions CR2l0j
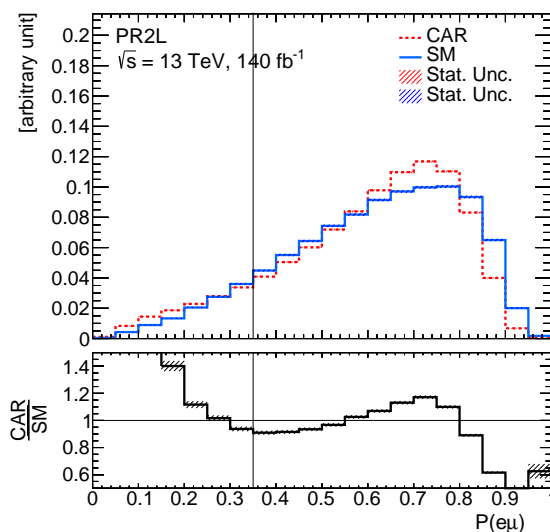
Figure 6.7.: Separation for the probability $P(e\mu)$ for the pseudo-reconstruction CAR and the reconstruction SM sample. The working point is marked with a vertical line.
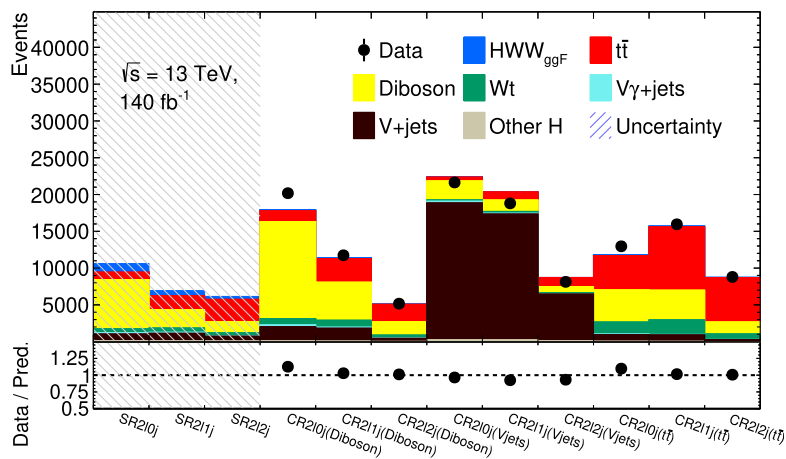
Figure 6.8.: Summary of the different signal and control regions. Only the statistical uncertainties are shown. The signal regions are shaded gray and the measured data points are not shown in those regions.

and CR2l1j are dominated by diboson. However, the CR2l2j for dibsoson is dominated by $t\bar{t}$. In the $t\bar{t}$ control regions CR2l1j and CR2l2j has $t\bar{t}$ the largest yield. In CR2l0j the diboson background is as strong as the $t\bar{t}$ background. The signal is mostly in the signal regions and the amount of signal in the control regions is negligible. SR2l1j and especially SR2l0j have large contributions from diboson events while SR2l2j has a large contribution of $t\bar{t}$ events.

7. Analysis Strategy and Methods

As mentioned before, the analysis aims at investigating if the observables m_{ll}, ϕ_{ll} and η_{ll} are sensitive towards quantum entanglement. For this investigation, two hypotheses are considered. The first is the SM hypothesis that the state is entangled. This hypothesis is assumed to be true and is treated as pseudo data. The second hypothesis is the null hypothesis of a separable state, the CAR hypothesis. To examine the sensitivity of the observables to quantum entanglement, exclusion limits for the CAR hypothesis of a separable state are calculated under the assumption that the SM pseudo data is true.

In the analysis, the sample of the measured observables is at reconstruction level, while the simulated sample for the CAR hypothesis of a separable state is only available at parton level. To compare the input samples, either the measured data needs to be transformed from reconstruction to parton level or the simulations have to be transformed to reconstruction level. The first technique, to transform reconstruction to parton level, is called *unfolding*. In this thesis, a binned unfolding is performed. The truth and the reconstruction information are stored in one-dimensional histograms with N and M bins. Then the parton yield t_n in bin n is transformed to the reconstruction level data r_m in bin m by folding [49]

$$r_m = \sum_n^N R_{mn} t_n \tag{7.1}$$

with the response matrix R. The response matrix element R_{mn} contains the probability that a parton in bin n produces an event at reconstruction level in bin m. The efficiency ϵ_n of the nth bin is defined as the probability of passing the reconstruction level selection for an event in the nth truth bin. The acceptance A_m is the probability of passing the truth level selection for an event in the mth reconstruction bin. Further, the migration matrix is the number of events that are reconstructed in the mth reconstruction bin and generated in the nth truth bin. The normalised migration matrix M_{mn} is normalised in every column to one. With these definitions, the response matrix can be calculated as [49]

$$R_{mn} = M_{mn} \frac{\epsilon_n}{A_m}.$$

The unfolding technique used in this analysis is a binned profile likelihood unfolding

(PLU). Each truth bin of the signal distribution is folded with the response matrix to get a reconstruction distribution according to Equation (7.1). Thus, there is a distribution of M reconstruction bins for each truth bin. A normalisation factor μ_i is assigned to each truth bin. These normalisation factors are determined by fitting the calculated reconstruction distribution to given data. The unfolded distribution is then reconstructed by applying the normalisation factors on the data [50]. To account for systematic uncertainties of the simulated sample nuisance parameters θ_j are added to the fit. The systematic uncertainties are assumed to be Gaussian. The measured data is divided into regions. The measured data in each region r and in n^{th} bin $n_{r,i}$ is assumed to be Poisson distributed. The expected signal strength and the expected background strength in the region r in the ith bin are $S_{r,i}$ and $B_{r,i}$. The normalisation factors μ_i and the nuisance parameters θ_j are contained in the vectors $\vec{\mu}$ and $\vec{\theta}$. With these, the likelihood of the model is [50]

$$\mathcal{L}(\vec{n}|\vec{\mu},\vec{\theta}) = \prod_{r\in\text{regions}} \prod_{i\in\text{bin}} \text{Pois}(n_{r,i}|S_{r,i}(\vec{\theta},\vec{\mu}) + B_{r,i}(\vec{\theta},\vec{\mu})) \prod_{j\in\text{syst}} \text{Gauss}(\theta_j).$$

The profile likelihood can then be calculated as,

$$\mathcal{L}_{\text{profile}}(\vec{n}|\vec{\mu}) = \max_{\vec{\mu}=\text{const.}} \left[\mathcal{L}(\vec{n}|\vec{\mu},\vec{\theta}) \right],$$

and a regular profile likelihood fit can be performed to determine the normalisation factors.

In this thesis, no real data is used in the PLU. Instead, fits with the simulated SM distributions as pseudo data are performed. The unfolded SM distribution is compared with the truth CAR distribution and an exclusion limit for the separable CAR hypothesis is calculated.

The exclusion limits of the null hypothesis are calculated using a χ^2-test [51]. From the test results, the σ exclusion limits are calculated. This thesis aims to find sensitivity with exclusion limits larger than 5σ for the separable CAR hypothesis.

8. Analysis Results

In the following chapter, the results are presented and discussed. First, the binning of the observables, the migration matrix, the efficiency and the acceptance are described. The implemented uncertainties are then named. Afterwards, the unfolding results and exclusion limits are displayed. Finally, the uncertainties of the fit are discussed.

8.1. Preparation of the Observables for the Unfolding

The binning must be chosen such that the statistical error is small. Furthermore, a sufficient number of degrees of freedom (NDF) are required for the unfolding. Four bins are the best regarding the unfolding results. The binning of the observables further meets two conditions in all signal regions. The relative statistical MC error of the diagonal elements in the migrating matrices has to be smaller than 0.8% and no diagonal bins of the migration matrix are allowed to be empty. This choice of binning prevents uncertainties in one bin from being much larger than in the other bins. To achieve that binning, the bins of the finely binned migration matrix are merged from left to right recursively until both conditions are satisfied. The normalised migration matrices, the efficiencies, the acceptance and the distribution of the events in the signal regions are shown in Figure 8.1 for SR2l0j and in Figure E.1 and Figure E.2 in Appendix E for SR2l1j and SR2l2j. The migration matrices should be as diagonal as possible for the unfolding. That reduces the statistical fluctuations and correlations between the different bins. The portion of diagonal elements together in all signal regions are $94.6^{+0.6}_{-0.6}$ % for m_{ll}, $99.6^{+0.3}_{-0.6}$ % for ϕ_{ll} and $99.5^{+0.3}_{-0.6}$ % for η_{ll}. The three observables are therefore mostly measured within the bin in which they are generated.

The efficiency is smaller than 10%. Thus, most events in the fiducial region at truth level are not measured in the signal regions at reconstruction level. The efficiency is the smallest in the fourth bin. In contrast, the acceptance is larger than 85% and therefore significantly larger than the efficiency. The acceptance is distributed approximately uniformly in SR2l0j and SR2l1j. In SR2l2j, the acceptance is not as uniformly distributed. The acceptance is the smallest in the first bin.
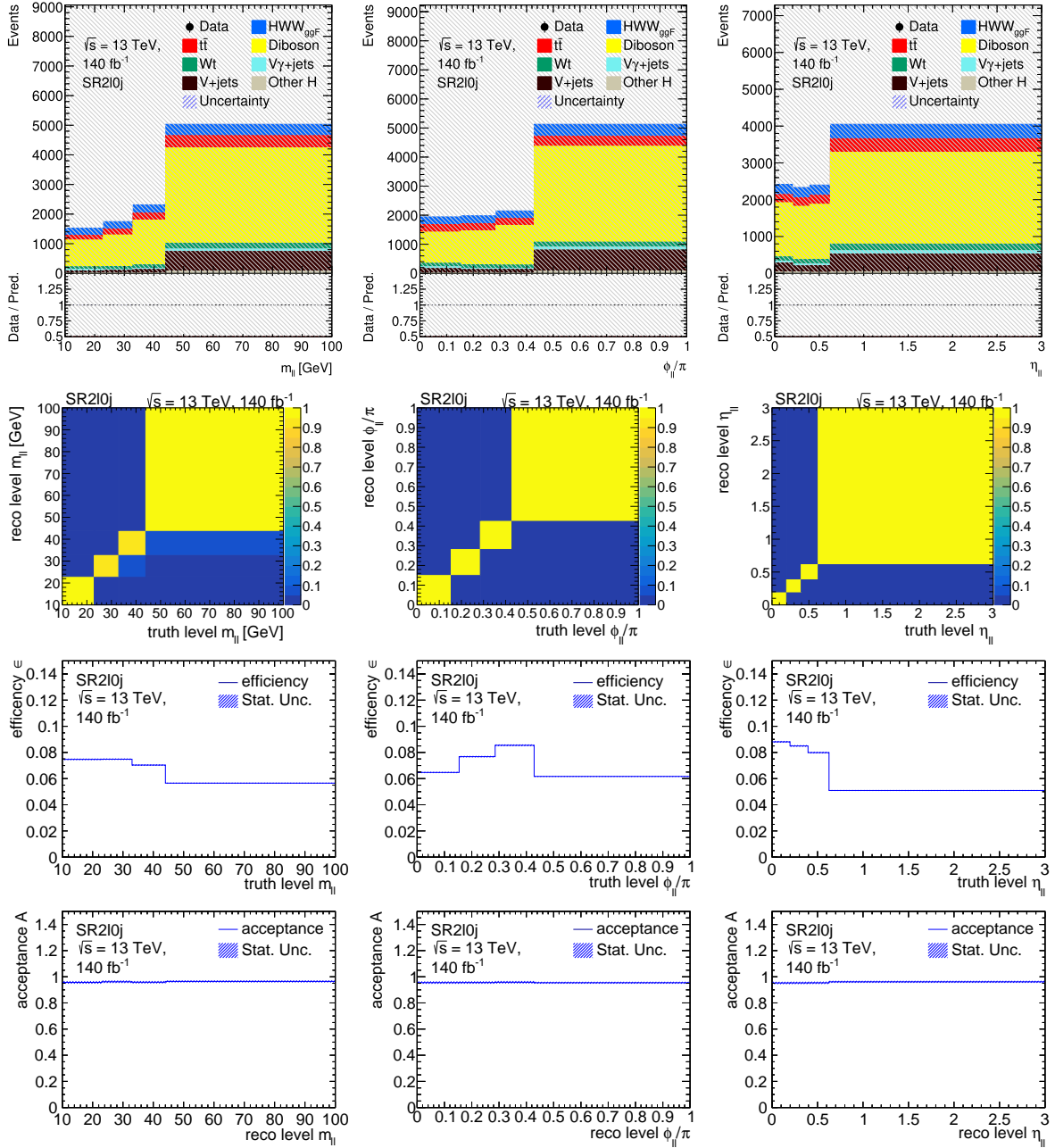
Figure 8.1.: Different figures for the studied observables in SR2l0j. On the left is m_{ll}, in the middle is ϕ_{ll} and on the right is η_{ll}. On the top is the distribution in the signal and the backgrounds. The signal regions are shaded gray and the measured data points are not shown. The second row shows the normalised migration matrices, the third row shows the efficiency and the fourth row shows the acceptance. Only statistical uncertainties are shown in all figures.

Table 8.1.: Experimental systematic uncertainties included in the PLU and their description.

experimental systematic	description
- pileup	Uncertainties in the modelling of jets due to pileup
- variations of b-tagging scale factor in light, c and b quarks	Uncertainties of the b-tagging in the DL1dv01 neural network in identifying b, c- and light quarks
- electron scale factor - muon scale factor identification - muon scale factor identification track to vertex association (TTVA)	Uncertainties in the reconstruction of electrons and muons

8.2. Systematic Uncertainties

The simulated sample is affected by systematic uncertainties. These can be divided into different groups. The first group is the experimental uncertainties. The studied experimental uncertainties are flavour tagging uncertainties, lepton identification uncertainties and uncertainties in the pileup estimations. The included experimental uncertainties are listed in Table 8.1. Other experimental uncertainties are not included in the analysis because the simulated samples used in this analysis are incomplete. To account for the systematic uncertainties the event weights are shifted up or down by one standard deviation. Using the shifted weights the migration matrix, the efficiency and the acceptance are calculated again and the unfolding is performed. The uncertainties on the fit are estimated with these results. The systematic uncertainties of the cross-section of the backgrounds are included. The backgrounds without dedicated control regions, i.e. Wt, $V\gamma$jets and other H are estimated conservatively to have normalisation uncertainties of overall 50% [17]. The other dominant backgrounds, i.e. $t\bar{t}$, diboson and Vjets are estimated in the unfolding procedure using free-floating normalisation factors.

8.3. Unfolding Results

As mentioned before, fits with the simulated SM distributions as pseudo data are performed for the PLU. All regions are unfolded simultaneously for each observable. The Unfolded distributions of the three observables m_{ll}, ϕ_{ll} and η_{ll} including τ-tagging are shown in Figure 8.2. The unfolding is also performed without τ-tagging to profit from larger statistics as a comparison. The unfolded distributions without τ-tagging are shown in Figure E.3 in the Appendix. In both cases, the exclusion limits are calculated for only
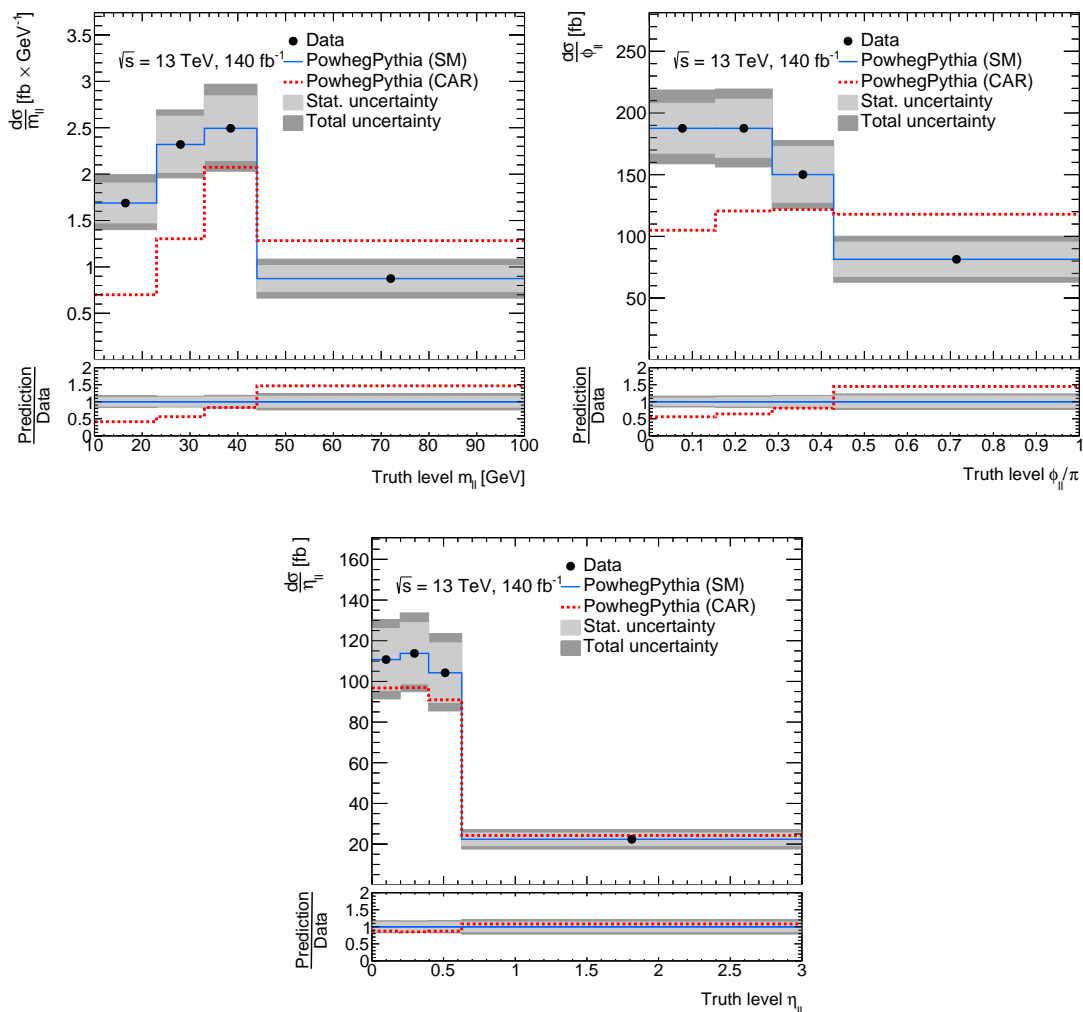
Figure 8.2.: The unfolded distributions of the three observables. The bin entries are divided by the bin width. The results include τ-tagging at 85%. On the top left is m_{ll}, on the top right is ϕ_{ll} and at the bottom is η_{ll}

the statistical uncertainty and the total uncertainty. The total uncertainty contains the statistical and systematic uncertainties. The normalisation of the histograms is a free parameter in the fit. The number of degrees of freedom is therefore the number of bins of the histogram, NDF = 4. The calculated exclusion limits are summarised in Table 8.2. The CAR distribution using η_{ll} is within the uncertainty of the unfolded SM distribution with and without using τ-tagging. All the calculated exclusion limits provided by η_{ll} are smaller than 1σ. It is noticeable that the statistical uncertainty is better without using τ-tagging, due to the larger number of events. However, all the statistical-only exclusion limits provided by m_{ll} and ϕ_{ll} are larger than 5σ. They are therefore larger than the desired exclusion limit of 5σ. The τ-tagging increases the control over the systematic

Table 8.2.: Exclusion limits for the CAR (separable) hypothesis for statistical and to-
tal uncertainty provided by the three observables. The total uncertainty
includes statistical and systematic uncertainties. Rounding to the first digit
is applied.

Observable	τ-tagging	σ exclusion limits	
		total	stat. only
m_{ll}	85% efficiency	4.3	6.0
m_{ll}	-	4.1	6.8
ϕ_{ll}	85% efficiency	3.7	5.2
ϕ_{ll}	-	3.7	5.7
η_{ll}	85% efficiency	0.4	0.7
η_{ll}	-	0.4	0.9

Table 8.3.: Normalisation factors for the different bins and observables in the PLU in-
cluding τ-tagging at 85% efficiency.

Observable	Bin 1	Bin 2	Bin 3	Bin 4
m_{ll}	$1.00^{+0.24}_{-0.24}$	$1.00^{+0.19}_{-0.19}$	$1.00^{+0.16}_{-0.16}$	$1.00^{+0.18}_{-0.19}$
ϕ_{ll}	$1.00^{+0.23}_{-0.23}$	$1.00^{+0.18}_{-0.18}$	$1.00^{+0.17}_{-0.17}$	$1.00^{+0.16}_{-0.15}$
η_{ll}	$1.00^{+0.22}_{-0.21}$	$1.00^{+0.18}_{-0.17}$	$1.00^{+0.17}_{-0.16}$	$1.00^{+0.18}_{-0.18}$

uncertainties using the variable m_{ll}. The exclusion limit with the total uncertainty using
m_{ll} including τ-tagging is 4.3σ. Without τ-tagging the exclusion limit is significantly
lower with 4.1σ. In all cases m_{ll} provides the best sensitivity. The observable ϕ_{ll} provides
worse sensitivity to quantum entanglement. Nevertheless, the exclusion limits provided
for the separable state by ϕ_{ll} are still larger than 3σ and 5σ for total and statistics-only
uncertainty, respectively. The τ-tagging does not influence the exclusion limits if the total
uncertainty is considered for ϕ_{ll}, both are 3.7σ.

The τ-tagging increases therefore the control over the systematic uncertainties, while
it decreases statistics. Below, the results using tau tagging are discussed in more detail
since the results for the total uncertainty are more sensitive.

All normalisation factors are one since a fit with the SM prediction as pseudo data is
performed. Their uncertainties are listed in Table 8.3.

8.4. Evaluation of the Unfolding Results

A closure test is performed to examine if the PLU is statistically independent. The sample
is divided into two statistically independent sets with even and odd event numbers. The
set with even event numbers is treated as pseudo data and is called the test sample. The

8. Analysis Results
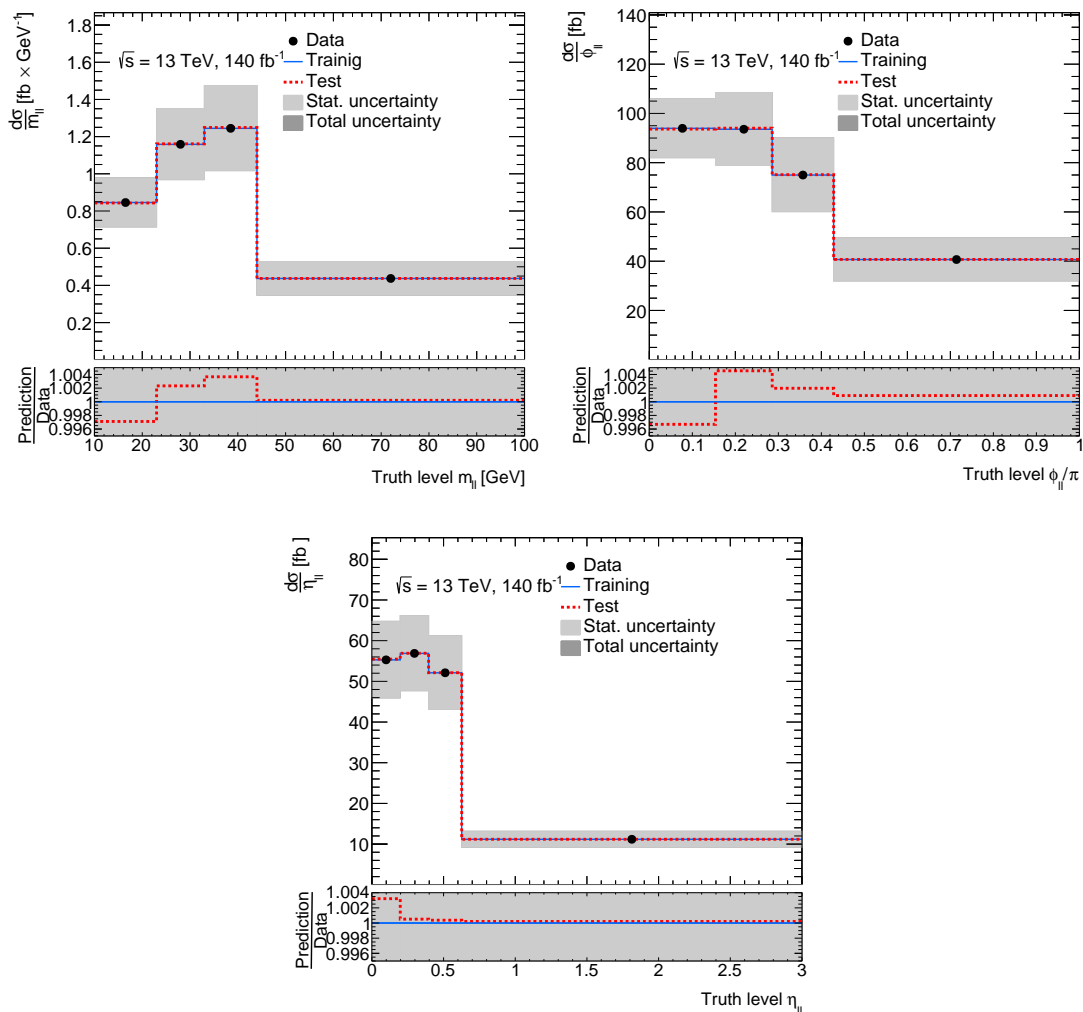
Figure 8.3.: The closure of the different observables. Only statistical uncertainties are shown. The bin entries are divided by the bin width. On the top left is m_{ll}, on the top right is ϕ_{ll} and on the bottom is η_{ll}.

sample with odd event numbers is used to derive the response matrix for the unfolding. It is called the training sample. Then the PLU is performed with the pseudo-data and the response matrix derived from the training sample. The unfolded distribution with the test and the training sample are shown in Figure 8.3. The pseudo data agrees with the unfolded distribution despite negligible deviations within the uncertainties. These deviations are smaller than 0.5%. Therefore, the two samples' statistical differences do not significantly affect the fit.

To examine the systematic uncertainties let $\hat{\theta}$ be the profile likelihood estimator of the uncertainties and θ_0 and $\Delta\theta$ the pre-fit value and uncertainty. Then the pull distributions of the nuisance parameters are defined as $(\hat{\theta} - \theta_0)/(\Delta\theta)$. The pull distributions are
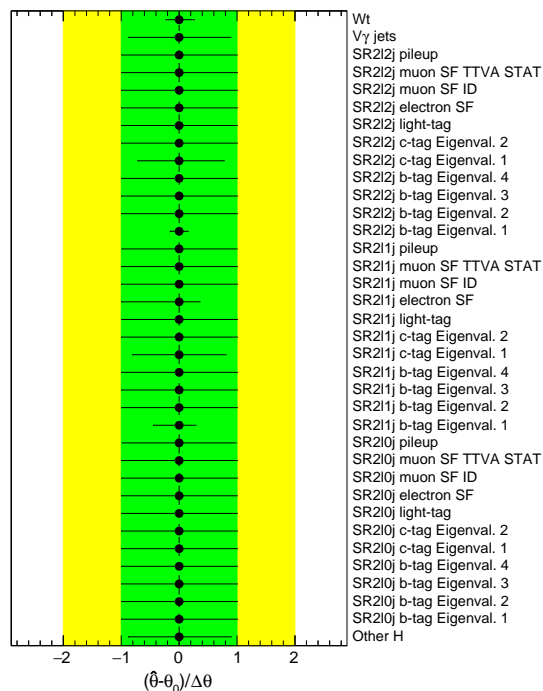
Figure 8.4.: Pulls of the nuisance parameters of the PLU of the observable m_{ll}.

predicted to be a standard Gaussian distribution with an expectation value of 0 and a standard deviation of 1. The pulls for m_{ll} are shown in fig. 8.4 and for ϕ_{ll} and η_{ll} in Figure E.4 in Appendix E. The desired expectation value of 0 is always met in the fits since a fit with the simulated pseudo-data is performed. However, the standard deviation is smaller than 1 for some systematics. These systematics are constrained by the data. That means that the expected uncertainty is larger than the uncertainty obtained from the fit. The background normalisations of Wt, $V\gamma$ jets and other H are constrained. These constraints are expected since these normalisation uncertainties are estimated conservatively. However, some of the experimental uncertainties are also constrained by the fit. The constrained experimental uncertainties are the first eigenvalue of the b-tagging and c tagging uncertainties in SR2l1j and SR2l2j and the scale factors (SF) of the electron in SR2l1j. There should not be large constraints because the experimental uncertainties are estimated more precisely. An issue in the modelling of the named experimental uncertainties is suspected.

The ranked impact of the systematic uncertainties on the fit is shown in Figure 8.5 for m_{ll} and in Figure E.5 and Figure E.6 in the Appendix for ϕ_{ll} and η_{ll}. The dominating uncertainties for m_{ll} in the first bin and ϕ_{ll} in the first two bins is the $V\gamma$ jets normalisation. This could be due to the assumed normalisation uncertainty of 50% and that it cannot

Figure 8.5.: Ranking of the impact of the different systematic uncertainties of the PLU for m_{ll}. The normalisation factors are marked with a red dot. The other uncertainties are marked with a black dot. On the top left is the first bin, on the top right is the second bin, on the bottom left is the third bin, and on the bottom right is the fourth bin.

be controlled well in the control regions. Anyway, it is large compared to the other backgrounds that are modelled in the same way. The impact of the pileup uncertainty in SR2l0j is large in every bin. The uncertainties of the normalisation factors of the diboson and V jets background are also relatively large. Since these two are some of the largest backgrounds this is expected. Further, it is surprising that the impact of the normalisation factor of $t\bar{t}$ is much smaller compared to the other two dominating backgrounds. The impact of the systematics on the fourth bin is the largest in all the observables. It is also noticeable that some of the pre-fit impacts of the uncertainties are asymmetrical. This is especially the case for the constrained b-tagging uncertainties. However, the uncertainties are expected to be more symmetrical. This supports the assumption of an issue in the calculations of these uncertainties.

9. Conclusion and Outlook

The thesis investigates the experimental sensitivity towards quantum entanglement in the $H \to WW^* \to e\nu\mu\nu$ final state. Two hypotheses are considered: One of an entangled state (the SM hypothesis) and one of a separable state (the CAR hypothesis). The SM model hypothesis is used as pseudo data, while the CAR hypothesis is treated as a null hypothesis to exclude. The aim is to exclude the CAR hypothesis at 5σ. The studied observables are m_{ll}, ϕ_{ll} and η_{ll}. The SM hypothesis is produced with MC simulations and detector simulations with Geant4. The truth information of the separable state is produced with the custom angle replacement algorithm from the already created SM distribution. No full detector simulation is performed to produce a CAR sample at reconstruction level. Nevertheless, a full detector simulation would be more precise and should be done in the future.

The signal and background are separated into multiple signal and control regions by multiclass classifiers. The classifiers are a dense neural network. The classification into different regions is successful, but they favour the entangled state over the separable state. A dense neural network is also used as a binary classifier to distinguish between τ decays and $e\mu$ events. This classifier is overtrained. To improve the training, a larger sample size is required. The τ-tagging has no significant influence on the observables ϕ_{ll} and η_{ll} but it improves the exclusion limits for the separable state in m_{ll}. Therefore, τ-tagging is a promising approach. Even though all classifiers show promising results they need to be revised and improved. The multiclass classifier needs to be improved regarding the bias towards the entangled state and the binary classifier regarding the overtraining. Additionally, a hyperparameter optimisation should be performed. An improved classifier could also offer better control over the statistical uncertainties. Furthermore, the fake lepton background and the Higgs production modes vector-boson-fusion, Higgs Strahlung and associated with a top pair need to be included into the analysis.

A binned profile likelihood unfolding of the pseudo data is performed. A performed closure test shows that the fit is statistically independent. However, additional closure tests of the fit should be done. Both statistical and systematic uncertainties influence the fit. The included experimental systematics are the flavour tagging uncertainties, pileup

uncertainties and lepton scale factor uncertainties. The pseudo data constrains some experimental systematics. These constraints need to be investigated further, to improve the understanding of the systematics. To investigate the sensitivity towards quantum entanglement a χ^2 test of the unfolded distribution and the separable hypothesis is done. The observable ϕ_{ll}, however, is promising for studying entanglement. This observable excludes the separable state at exclusion limits of 3.7σ. From the three studied observables, m_{ll} provides the best exclusion limit with 4.3σ. Even though m_{ll} is the most promising observable of the three examined it can not be used to exclude a separable state at 5σ thus far. The exclusion limits provided by ϕ_{ll} and m_{ll} based on only statistical uncertainties are beyond 5σ. Again m_{ll} provides the highest sensitivity here. The calculated exclusion limit for the CAR hypothesises using η_{ll} is smaller than one sigma. Therefore, η_{ll} is experimentally less sensitive to quantum entanglement than m_{ll} and η_{ll}.

Through further studies, for example with the previously mentioned approaches, an improvement to a 5σ exclusion limit, for the total uncertainty using m_{ll} and ϕ_{ll}, is likely achievable. As a final result, the approach of custom angle replacement is promising for investigating quantum entanglement in $H \to WW^* \to e\nu\mu\nu$ final states. Especially the observables ϕ_{ll} and m_{ll} are promising.

A. Monte Carlo Generators

Table A.1.: The different MC generators [52–54] used in this analysis for the different signal and background samples.

Sample	Generator	Parton Shower
HWW$_{\text{ggf}}$ (Signal)	Powheg [52]	Pythia 8 [53]
Diboson	Sherpa 2.2.2 [54]	
$t\bar{t}$	Powheg	Pythia 8
Wt	Powheg	Pythia 8
Vjets	Sherpa 2.2.11	
$V\gamma$jets	Sherpa 2.2.11	
Other Higgs	Powheg	Pythia 8

B. Additional Evaluation of the Pseudo-Reco CAR Sample
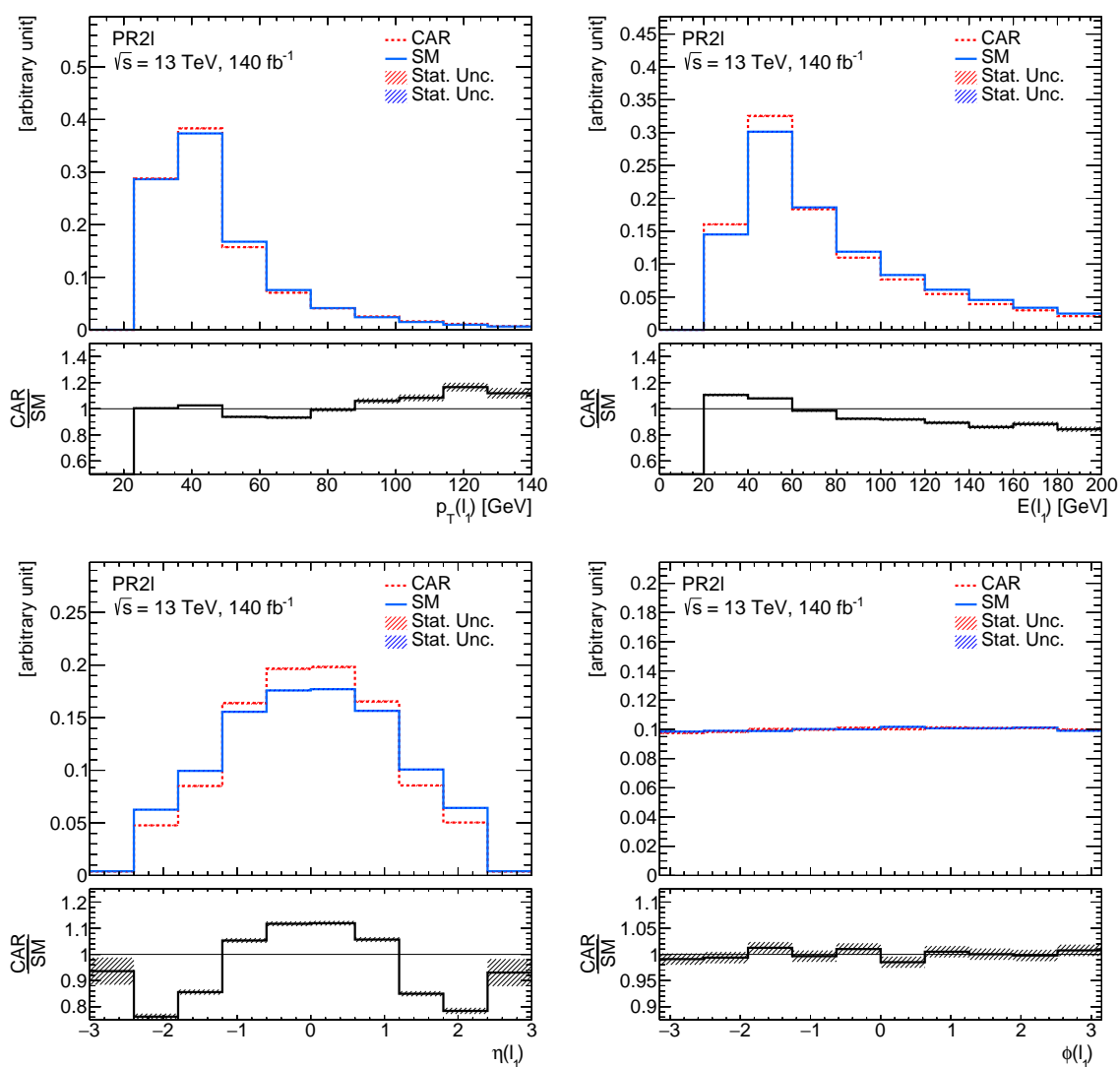
Figure B.1.: Separation plots for the kinematic variables of the leading lepton for the SM sample on reconstruction level and the pseudo-reconstruction CAR sample.
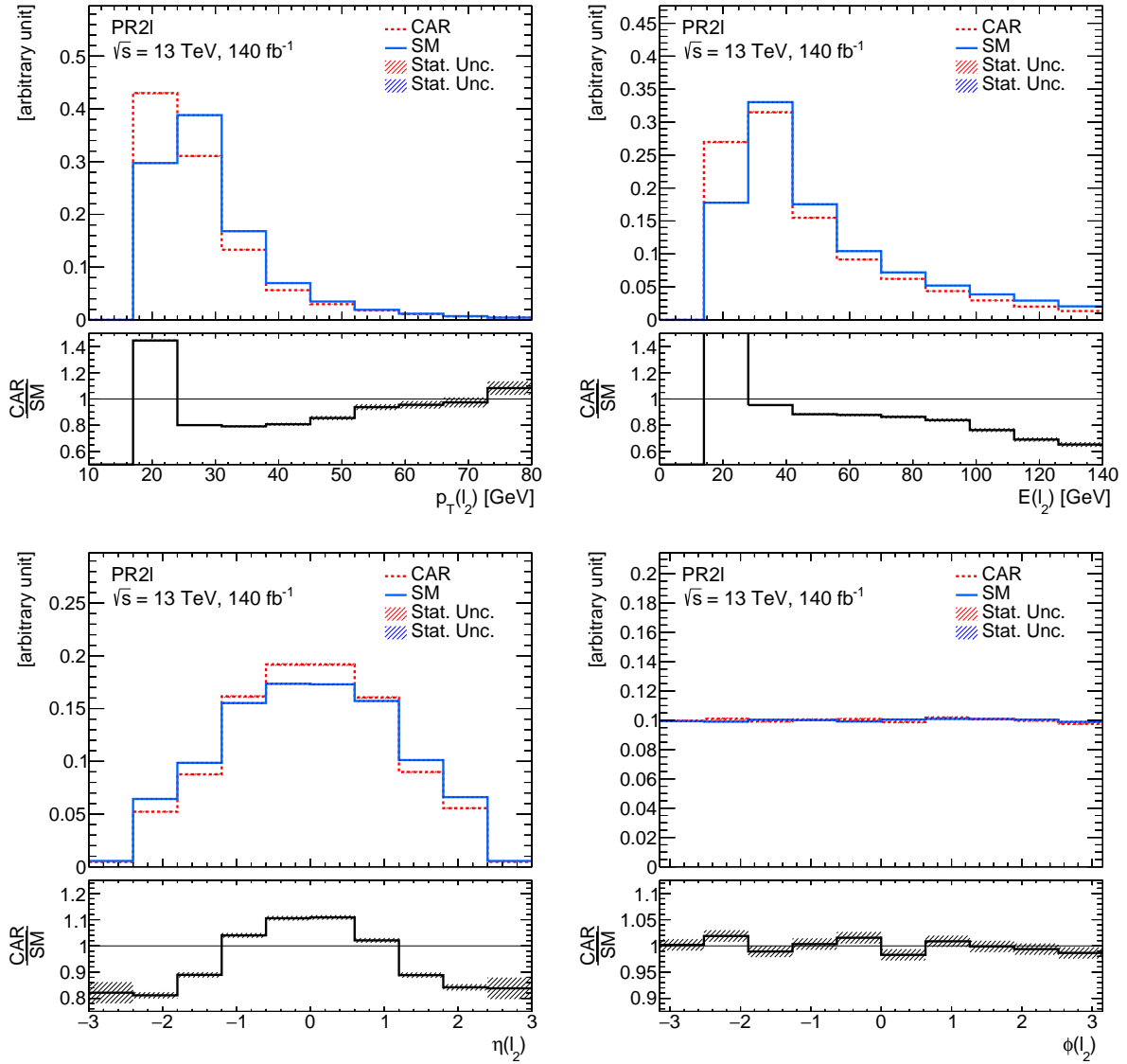
Figure B.2.: Separation plots for the kinematic variables of the sub-leading lepton for the SM sample on reconstruction level and the pseudo-reconstruction CAR sample.
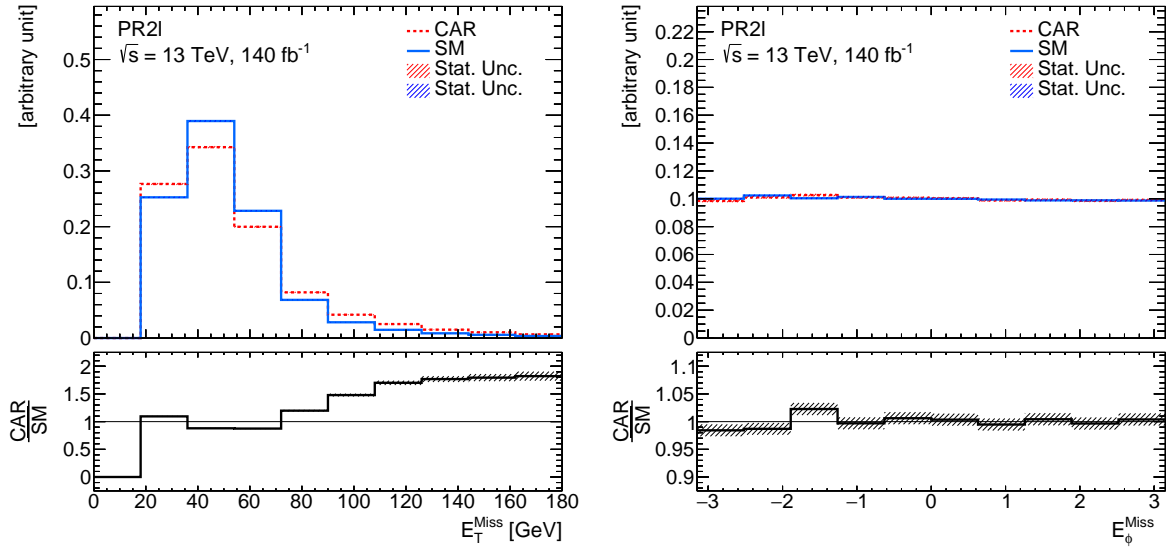
Figure B.3.: Separation plots for the missing transverse Energy for the SM sample on reconstruction level and the pseudo-reconstruction CAR sample.

C. Machine Learning Functions

Functions that are used in the machine learning classifiers in Section 6.3 are defined here.

Activation Functions

ReLU

The `ReLU` function for an input x is defined as

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

tanh

The `tanh` function for an input x is defined as

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Output Activation Functions

Softmax

The `softmax` function transforms a vector of real-valued $\vec{x}$ into a set of probabilities $\vec{x'}$. It is defined as

$$x'_i = \text{softmax}(\vec{x}, i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

Sigmoid

The `sigmoid` function transforms a real-valued x into a probability x'. It is defined as

$$x' = \text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$$

Loss Functions

Categorical Cross-Entropy

Let $\vec{t}$ be the target and $\vec{p}(\vec{x})$ be the prediction of the classifier for the input $\vec{x}$. The `categorical cross-entropy` loss function for the classification into four classes is defined as

$$\text{loss}(\vec{t}, \vec{p}(\vec{x})) = \text{CE}(\vec{t}, \vec{p}(\vec{x})) = -\sum_{i=1}^{4} t_i \log\left(p_i(\vec{x})\right)$$

Binary Cross-Entropy

Let y be the target and $p(\vec{x})$ be the prediction of the classifier for the input $\vec{x}$. The `binary cross-entropy` loss function for the classification into two classes is defined as

$$\text{loss}(t, p(\vec{x})) = \text{BC}(t, p(\vec{x})) = -[t\log\left(p(\vec{x})\right) + (1-t)\log\left(1 - p(\vec{x})\right)]$$

D. Additional Information for the Event Classification
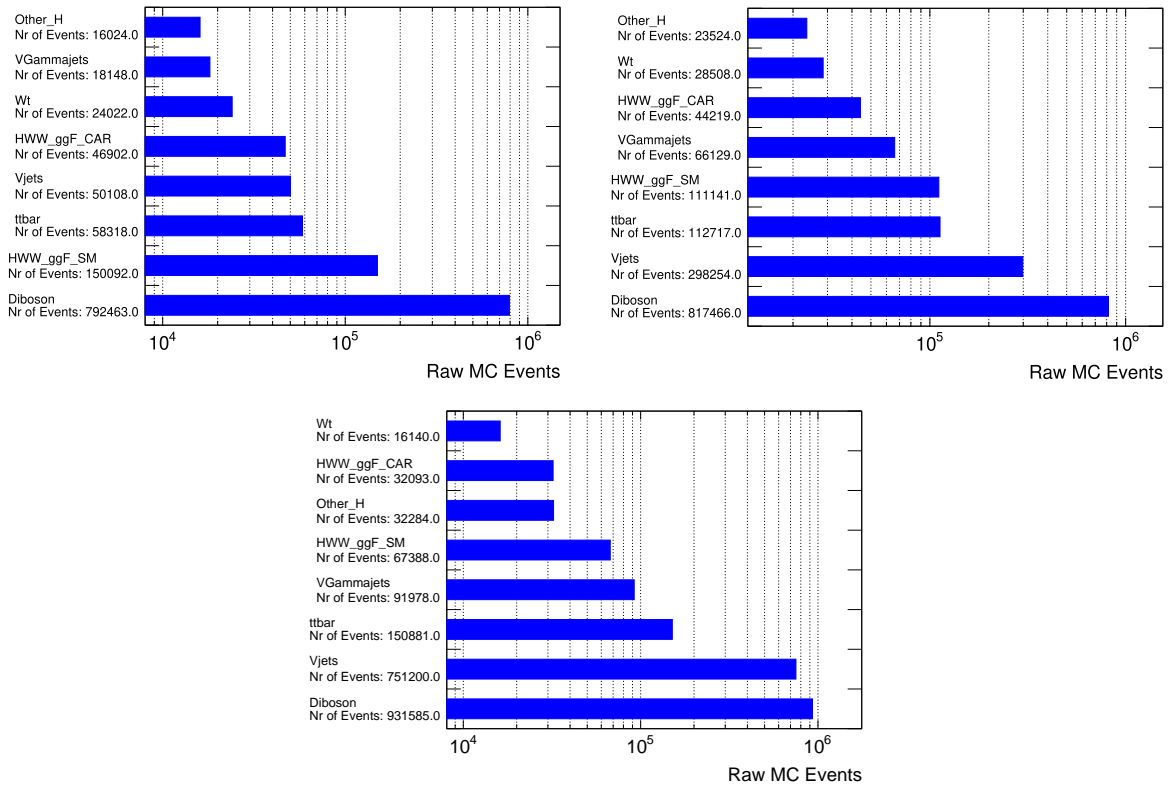
Figure D.1.: Number of MC events for the signal and each background type in each preselection region. On the top left is PR2l0j, on the top right is PR2l1j and on the bottom is PR2l2j.
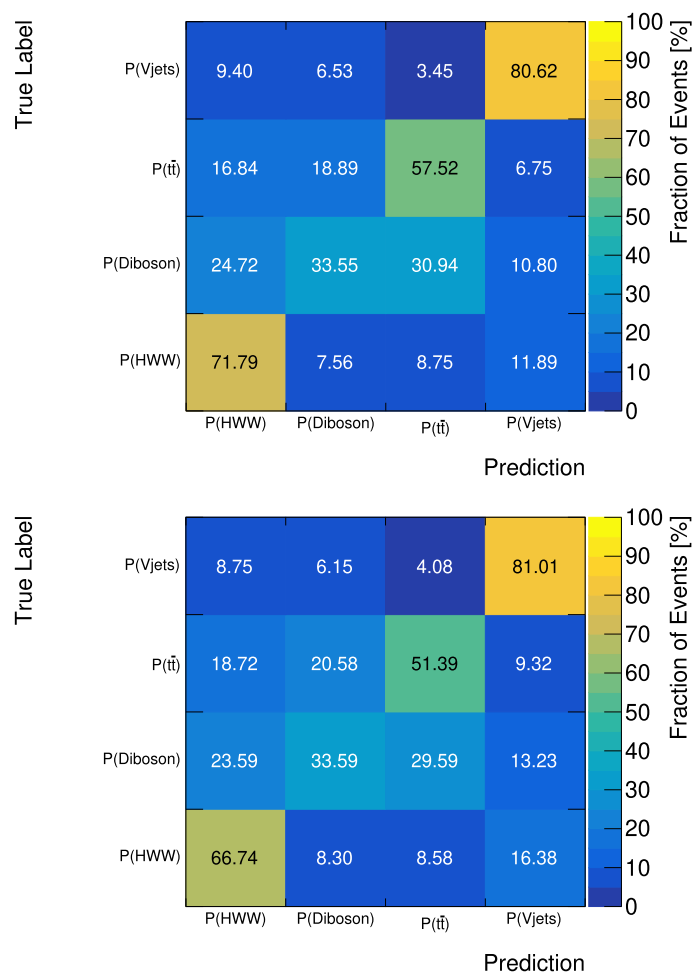
Figure D.2.: The confusion matrices for SR2l1j on the top for SR2l2j on the bottom.
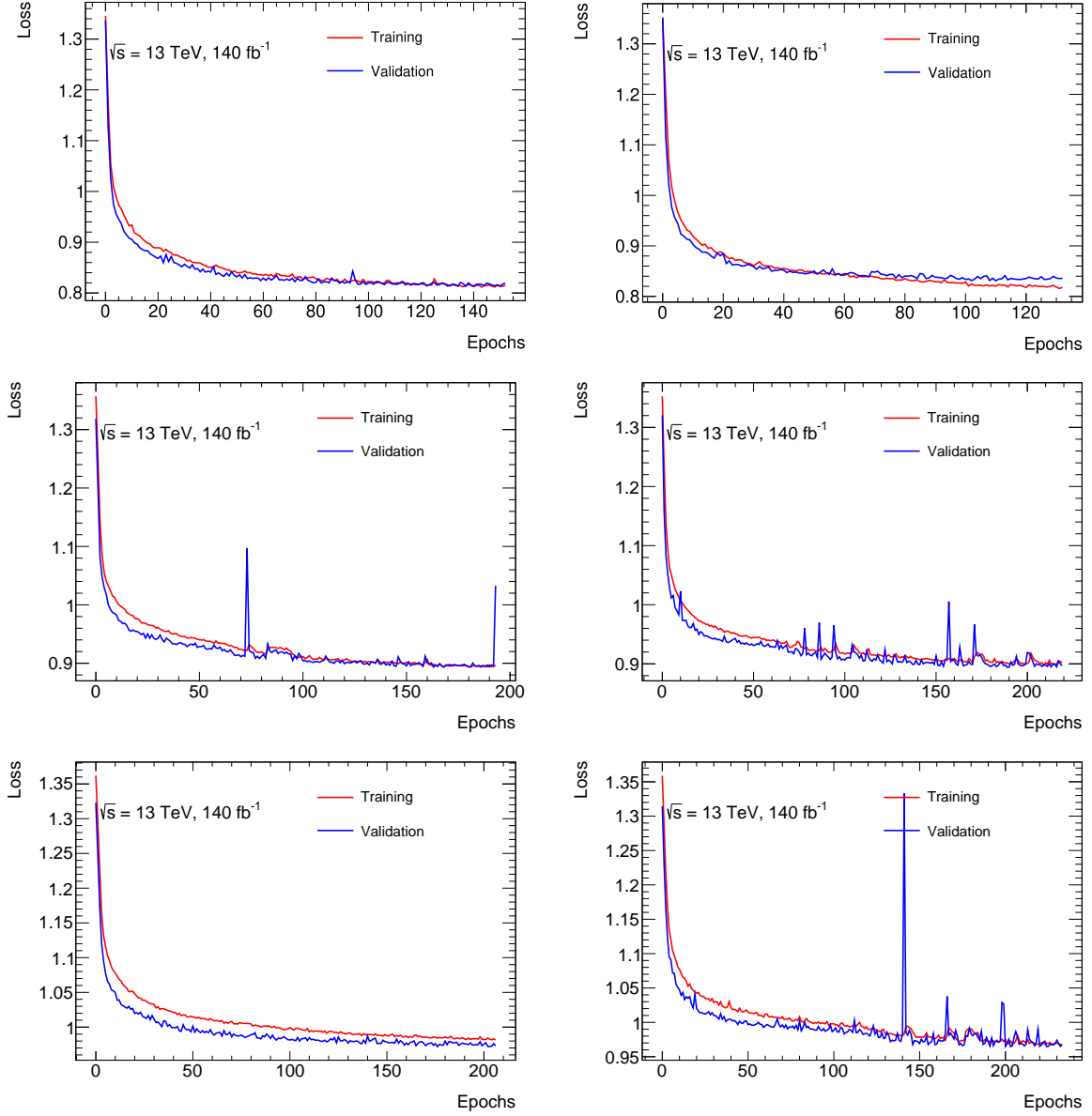
Figure D.3.: The loss for each fold for each preselection region for the multiclass classifier. On the top left is PR2l0j, in the middle is PR2l1j and on the bottom is PR2l2j. The first fold is on the left-hand side and the second fold is on the right-hand side.

Figure D.4.: The perturbation importance and the ROC curve for the output variables of multiclass classifier in PR2l0j. On the top left is *HWW*, on the top right is diboson, on the bottom left is $t\bar{t}$ and on the top right is *V*jets.
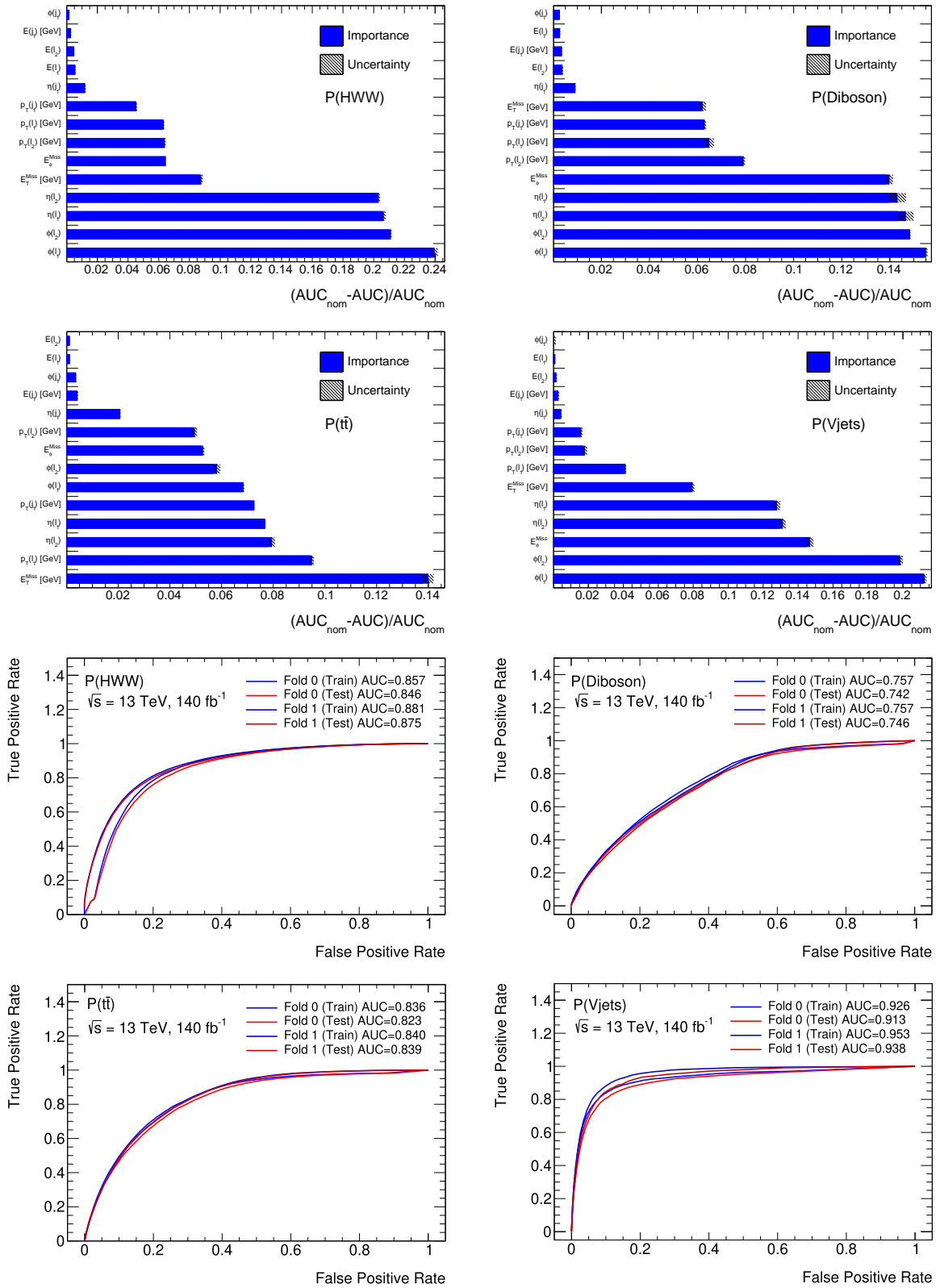
Figure D.5.: The perturbation importance and the ROC curve for the output variables of multiclass classifier in PR2l1j. On the top left is *HWW*, on the top right is diboson, on the bottom left is $t\bar{t}$ and on the top right is *V*jets.

Figure D.6.: The perturbation importance and the ROC curve for the output variables of multiclass classifier in PR2l2j. On the top left is HWW, on the top right is diboson, on the bottom left is $t\bar{t}$ and on the top right is Vjets.
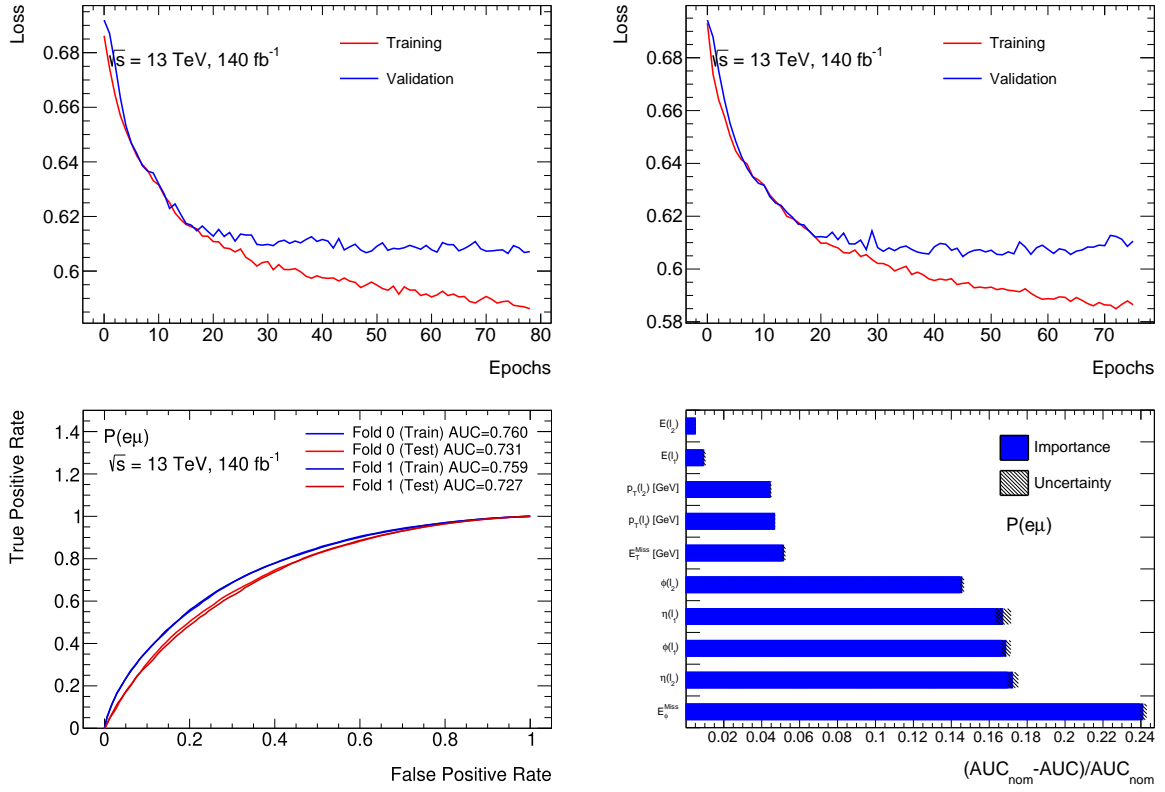
Figure D.7.: Evaluation of the DNN model for the τ classifier. The two figures on the top show the loss of the two folds, the bottom left shows the ROC curve and the bottom right shows the permutation importance.

Table D.1.: Summary of the event selection and classification for the signal and the control regions. The table is divided into a bottom and a top part by a triple horizontal line. The top part and bottom part show each the definition of 6 regions.

Variable description	SR2l0j	SR2l1j	SR2l2j	CR2l0j(Diboson)	CR2l1j(Diboson)	CR2l2j(Diboson)
Signal regions			reconstruction level cuts			
Jet multiplicity	$N_\text{jets} = 0$	$N_\text{jets} = 1$	$N_\text{jets} \geq 2$	$N_\text{jets} = 0$	$N_\text{jets} = 1$	$N_\text{jets} \geq 2$
b-jet multiplicity	$N_{b\text{-jets85\%}} = 0$					
Lepton multiplicity	$N_e = 1 \ \& \ N_\mu = 1$					
Lepton charge	$q_e = -q_\mu$					
Lepton transverse momentum	$p_T(l_1) > 22 \ \& \ p_T(l_2) > 15$					
Missing transverse energy	$E_T^\text{Miss} > 20\,\text{GeV}$					
Invariant mass leptons	$10\,\text{GeV} < m_{ll} < 100\,\text{GeV}$					
HWW^* classifier	$P(HWW)$ is maximal probability			$P(\text{Diboson})$ is maximal probability		
τ-tagging	$P(e\mu) > 0.35$			–		

Variable description	CR2l0j($t\bar{t}$)	CR2l1j($t\bar{t}$)	CR2l2j($t\bar{t}$)	CR2l0j(Vjets)	CR2l1j(Vjets)	CR2l2j(Vjets)
Signal regions			reconstruction level cuts			
Jet multiplicity	$N_\text{jets} = 0$	$N_\text{jets} = 1$	$N_\text{jets} \geq 2$	$N_\text{jets} = 0$	$N_\text{jets} = 1$	$N_\text{jets} \geq 2$
b-jet multiplicity	$N_{b\text{-jets85\%}} = 0$					
Lepton multiplicity	$N_e = 1 \ \& \ N_\mu = 1$					
Lepton charge	$q_e = -q_\mu$					
Lepton transverse momentum	$p_T(l_1) > 22 \ \& \ p_T(l_2) > 15$					
Missing transverse energy	$E_T^\text{Miss} > 20\,\text{GeV}$					
Invariant mass leptons	$10\,\text{GeV} < m_{ll} < 100\,\text{GeV}$					
HWW^* classifier	$P(t\bar{t})$ is maximal probability			$P(\text{Vjets})$ is maximal probability		
τ-tagging	–					

E. Additional Unfolding Results
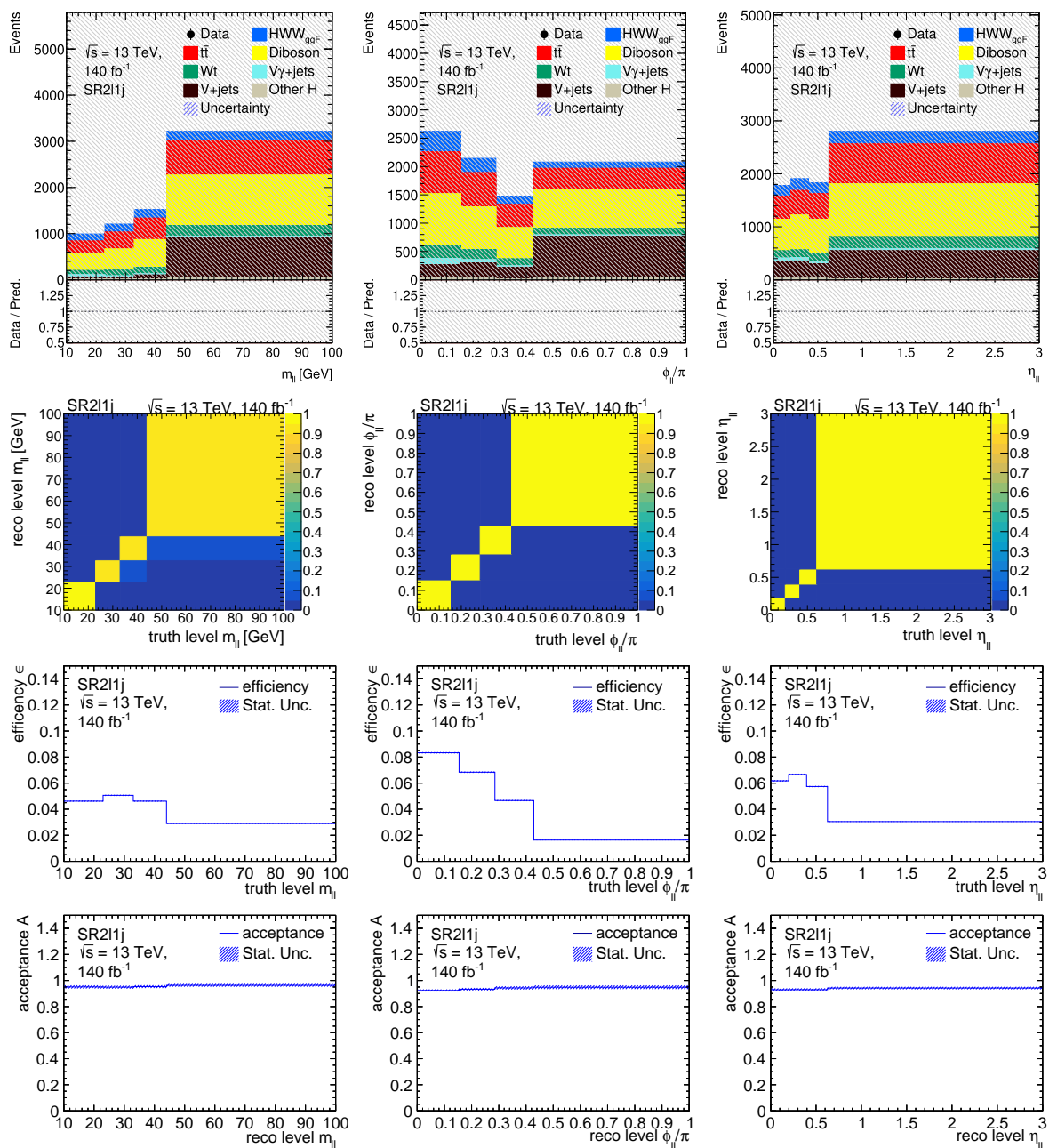
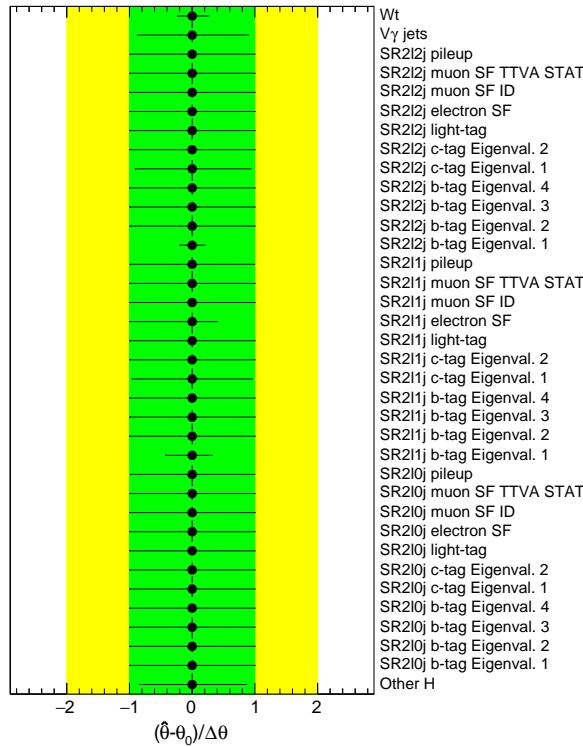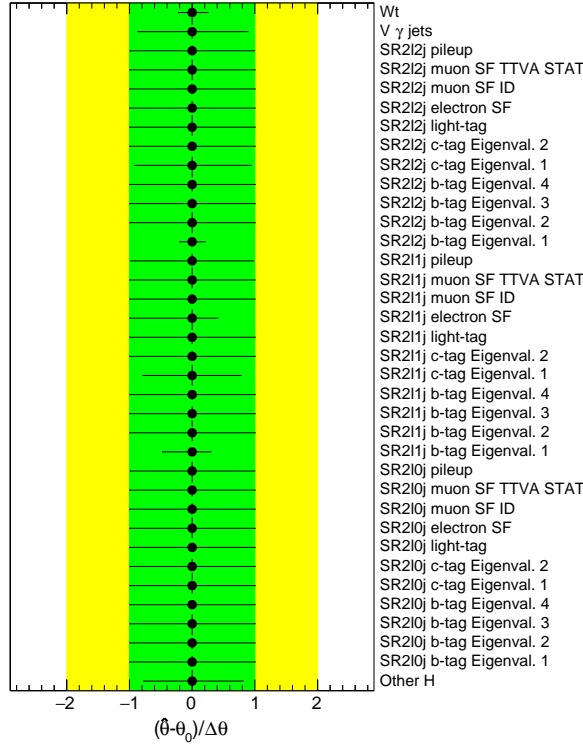***Figure E.1.:*** The same results as in Figure 8.1 for SR2l1j.
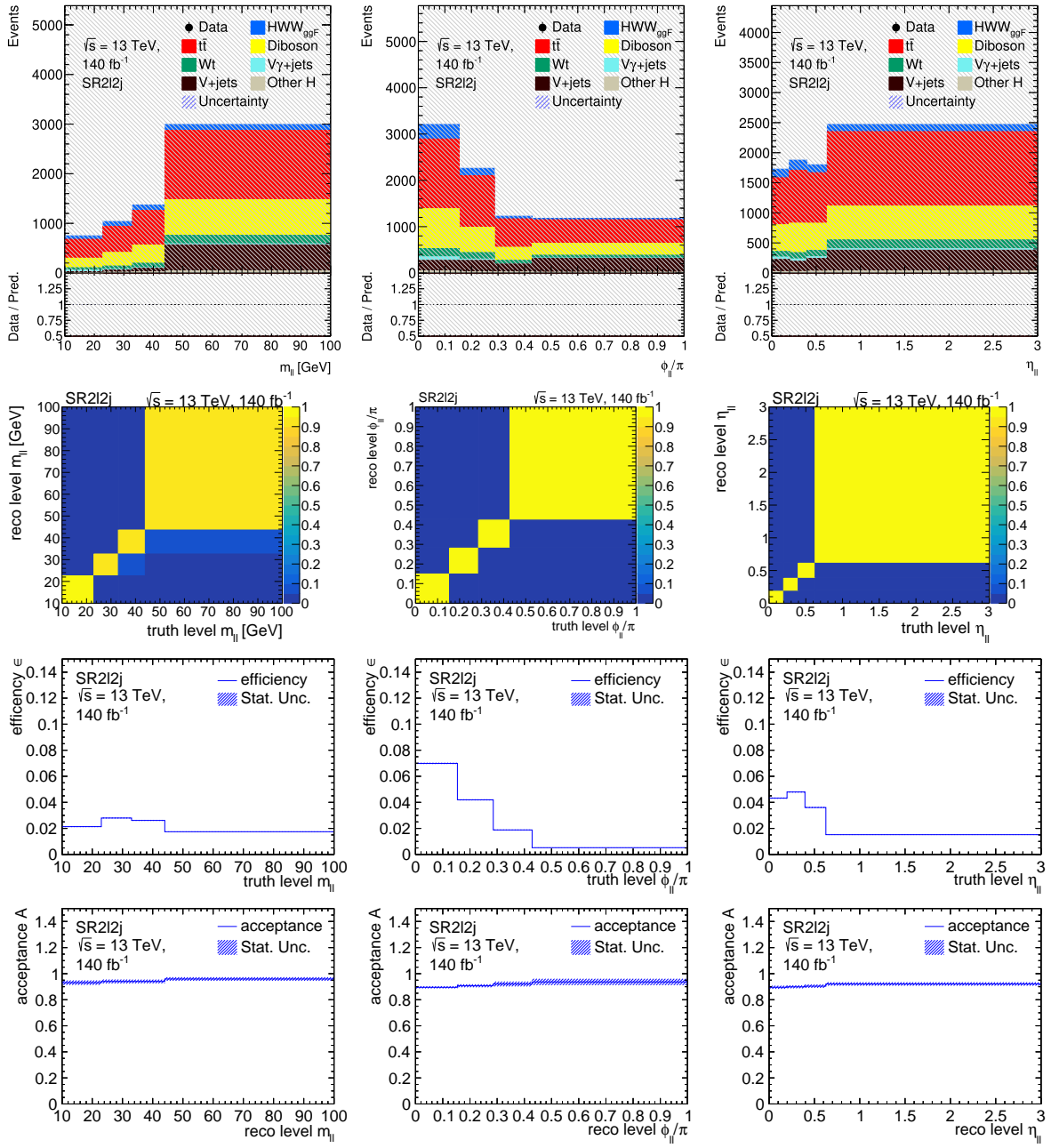
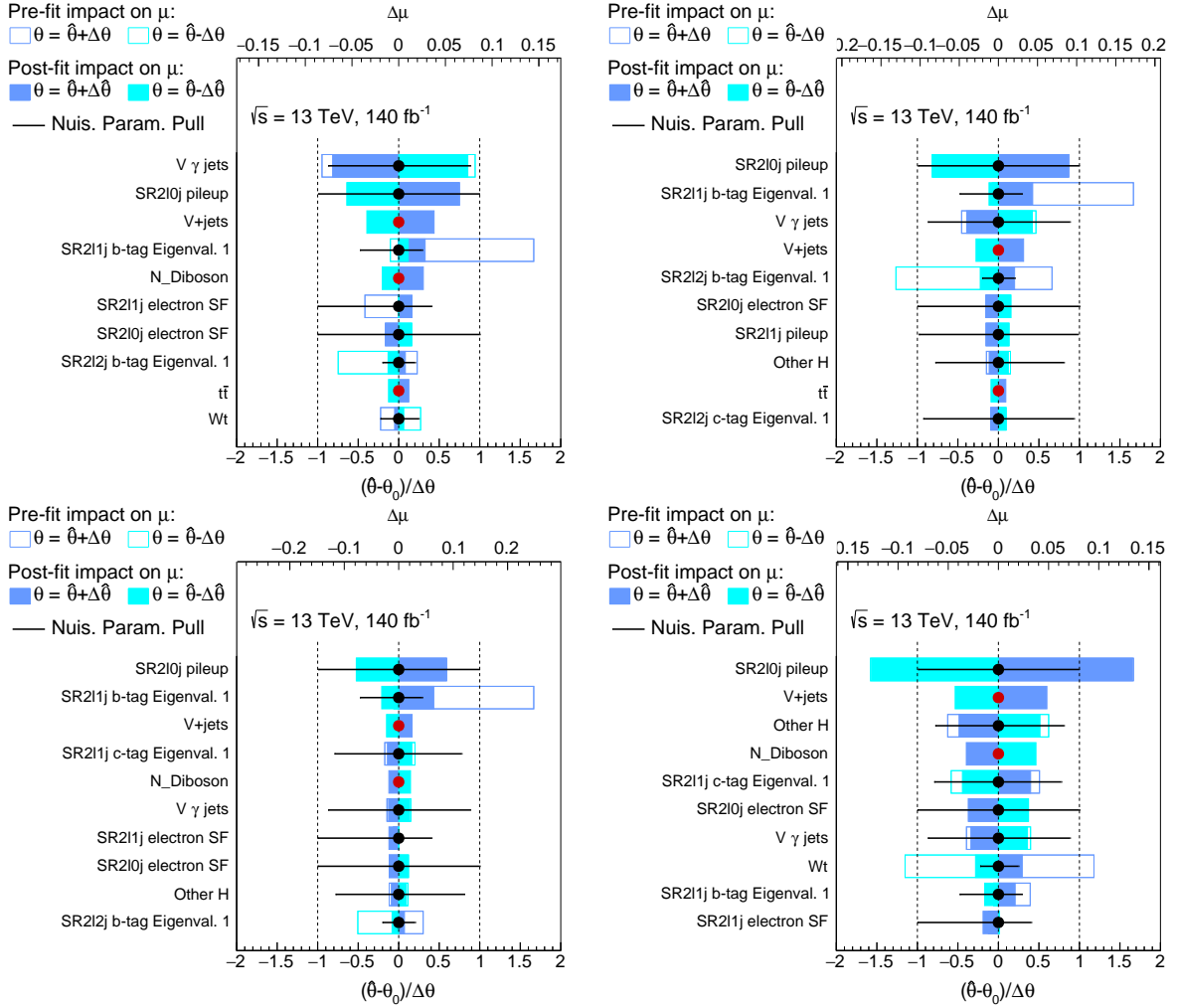***Figure E.2.:*** The same results as in Figure 8.1 for SR2l2j.
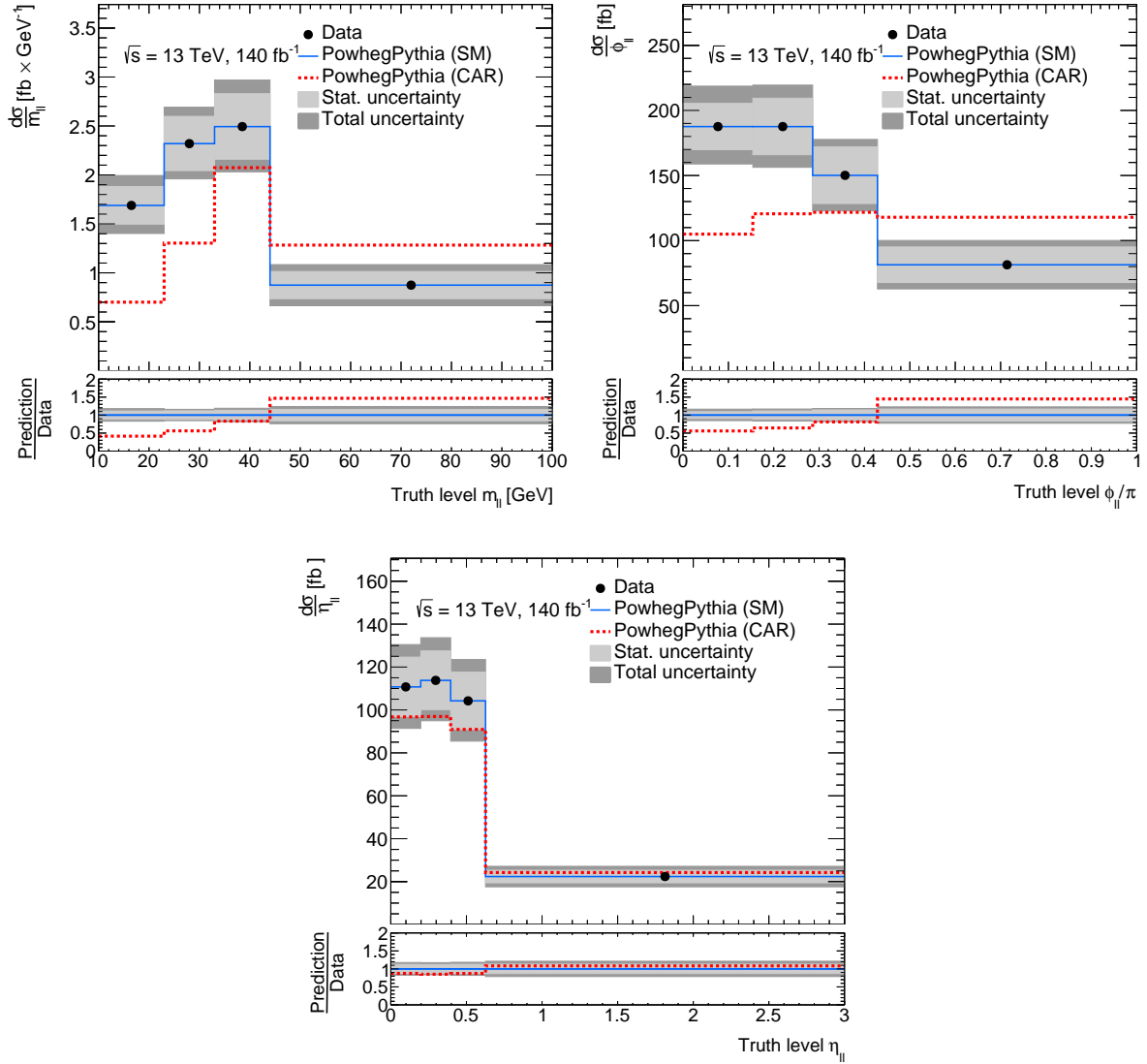
Figure E.3.: The unfolded distributions of the three observables. The bin entries are divided by the bin width. The results do not include τ-tagging. On the top left is m_{ll}, on the top right is ϕ_{ll} and at the bottom is η_{ll}.
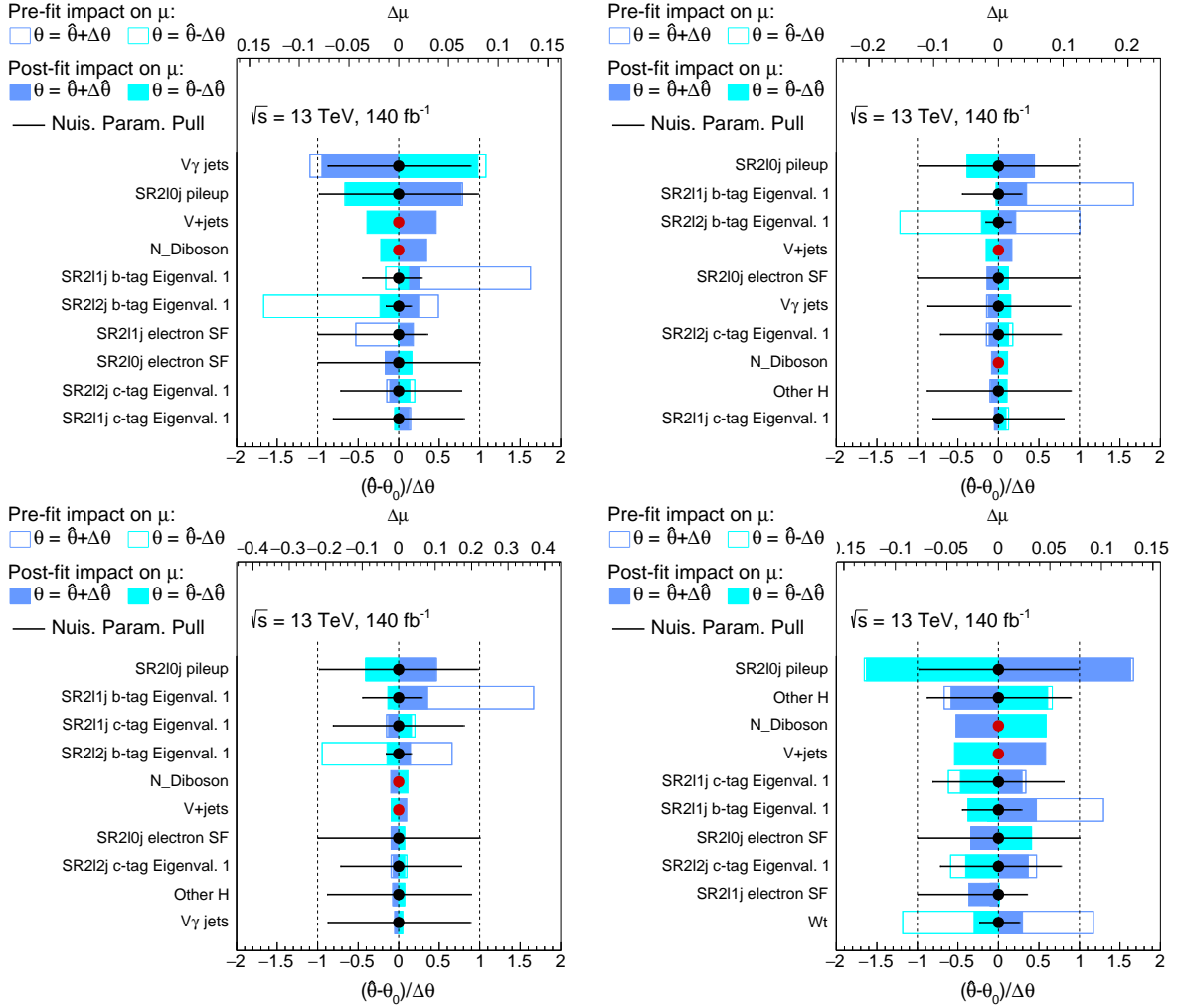
Figure E.4.: Pulls of the nuisance parameters in the PLU. On the top is ϕ_{ll} and on the bottom is η_{ll}.

Figure E.5.: Ranking of the impact of the different systematic uncertainties of the PLU for ϕ_{ll}. The norm factors are marked with a red dot. The other uncertainties are marked with a black dot. On the top left is the first bin, on the top right the second bin, on the bottom left the third bin and on the bottom right the fourth bin.

Figure E.6.: Ranking of the impact of the different systematic uncertainties of the PLU for η_{ll}. The norm factors are marked with a red dot. The other uncertainties are marked with a black dot. On the top left is the first bin, on the top right the second bin, on the bottom left the third bin and on the bottom right the fourth bin.

Bibliography

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

[2] J. S. Bell, *On the Einstein-Podolsky-Rosen paradox*, Phys. Phys. Fiz. **1** (1964)

[3] J. S. Bell, *On the Problem of Hidden Variables in Quantum Mechanics*, Rev. Mod. Phys. **38** (1966)

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

[5] A. Aspect, P. Grangier, G. Roger, *Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities*, Phys. Rev. Lett. **49** (1982)

[6] J. Yin, et al., *Lower Bound on the Speed of Nonlocal Correlations without Locality and Measurement Choice Loopholes*, Phys. Rev. Lett. **110** (2013)

[7] M. Giustina, et al., *Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons*, Phys. Rev. Lett. **115** (2015)

[8] ATLAS Collaboration, *Observation of quantum entanglement in top-quark pairs using the ATLAS detector*, CERN-EP-2023-230 (2023)

[9] CMS Collaboration, *Observation of quantum entanglement in top quark pair production in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, CMS-TOP-23-001, CERN-EP-2024-137 (2024)

[10] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967)

[11] H. Fritzsch, M. Gell-Mann, H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. B **47** (1973)

[12] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964)

Bibliography

[13] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968)

[14] H. Georgi, S. L. Glashow, *Unified weak and electromagnetic interactions without neutral currents*, Phys. Rev. Lett. **28** (1972)

[15] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964)

[16] F. Englert, R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964)

[17] Particle-Data-Group-Collaboration, *Review of Particle Physics*, PTEP **2022** (2022)

[18] Planck Collaboration, *Planck 2018 results. VI. Cosmological parameters*, Astron. Astrophys. **641** (2020)

[19] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, Zh. Eksp. Teor. Fiz. **53** (1967)

[20] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998)

[21] K. G. Wilson, *Confinement of Quarks*, Phys. Rev. D **10** (1974)

[22] D. J. Gross, F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30** (1973)

[23] Z. Koba, T. Tati, S. Tomonaga, *On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields. II: Case of Interacting Electromagnetic and Electron Fields*, Prog. Theor. Phys. **2(3)** (1947)

[24] R. P. Feynman, *Space - time approach to quantum electrodynamics*, Phys. Rev. **76** (1949)

[25] R. P. Feynman, *Mathematical formulation of the quantum theory of electromagnetic interaction*, Phys. Rev. **80** (1950)

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

[27] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012)

[28] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012)

[29] E. E. Salpeter, H. A. Bethe, *A Relativistic equation for bound state problems*, Phys. Rev. **84** (1951)

[30] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector at the LHC*, European Physical Journal C **83** (2023)

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

[32] J. A. Aguilar-Saavedra, *Laboratory-frame tests of quantum entanglement in H→WW*, Phys. Rev. D **107(7)** (2023)

[33] J. A. Aguilar-Saavedra, et al., *Testing entanglement and Bell inequalities in H→ZZ*, Phys. Rev. D **107(1)** (2023)

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

[35] T. Sjostrand, *Jet Fragmentation of Nearby Partons*, Nucl. Phys. B **248** (1984)

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

[37] GEANT4-Collaboration, *GEANT4−a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003)

[38] J. A. Aguilar-Saavedra, *Crafting polarizations for top, W, and Z*, Phys. Rev. D **106(11)** (2022)

[39] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80(1)** (2020)

[40] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. C **76(5)** (2016)

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

Bibliography

[42] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79(11)** (2019)

[43] ATLAS Collaboration, *Measurements of Higgs boson production by gluon-gluon fusion and vector-boson fusion using $H \to WW^* \to e\nu\mu\nu$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **108** (2023)

[44] D. E. Rumelhart, G. E. Hinton, R. J. Williams, *Learning representations by back-propagating errors*, Nature **323** (1986)

[45] S. Amari, *Backpropagation and stochastic gradient descent method*, Neurocomputing **5(4)** (1993)

[46] S. Ioffe, C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, in *Proceedings of Machine Learning Research*, volume 37, Lille, France (2015)

[47] N. Srivastava, et al., *Dropout: a simple way to prevent neural networks from overfitting*, J. Mach. Learn. Res. **15(1)** (2014)

[48] T. Dozat, *Incorporating Nesterov Momentum into Adam*, in *Proceedings of the 4th International Conference on Learning Representations* (2016)

[49] P. Baroň, J. Kvita, *Extending the Fully Bayesian Unfolding with Regularization Using a Combined Sampling Method*, Symmetry **12(12)** (2020)

[50] G. Cowan, et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011)

[51] K. Pearson, *X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling*, London Edinburgh Philos. Mag. & J. Sci. **50(302)** (1900)

[52] S. Alioli, et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010)

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

[54] E. Bothmann, et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7(3)** (2019)

Acknowledgements

Erklärung nach 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 17. August 2024

(Nico Schiel)