FACHBEREICH C - FACHGRUPPE PHYSIK
BERGISCHE UNIVERSITÄT
WUPPERTAL

Higgs Boson Search in the $H \to WW^{(*)} \to \ell\nu\ell\nu$ Channel using Neural Networks with the ATLAS Detector at 7 TeV

Dissertation

Georg Sartisohn

March 26, 2012

Introduction

Particle physics analyzes the fundamental constituents of matter, the elementary particles and their interactions. Theoretical and experimental efforts lead to the establishment of the Standard Model (SM), a framework based on the concept of quantum field theory. The SM describes the fundamental fermions, quarks and leptons, and their interactions via three fundamental forces. The electromagnetic, the strong and the weak interactions are mediated by gauge bosons: the photon γ, the gluons g, and the Z^0 and $W^{\pm}$ bosons, respectively. Figure I.1 depicts the experimentally verified fermions, grouped in three generations, and the gauge bosons with their properties. More detailed information can be found in References [1, 2].

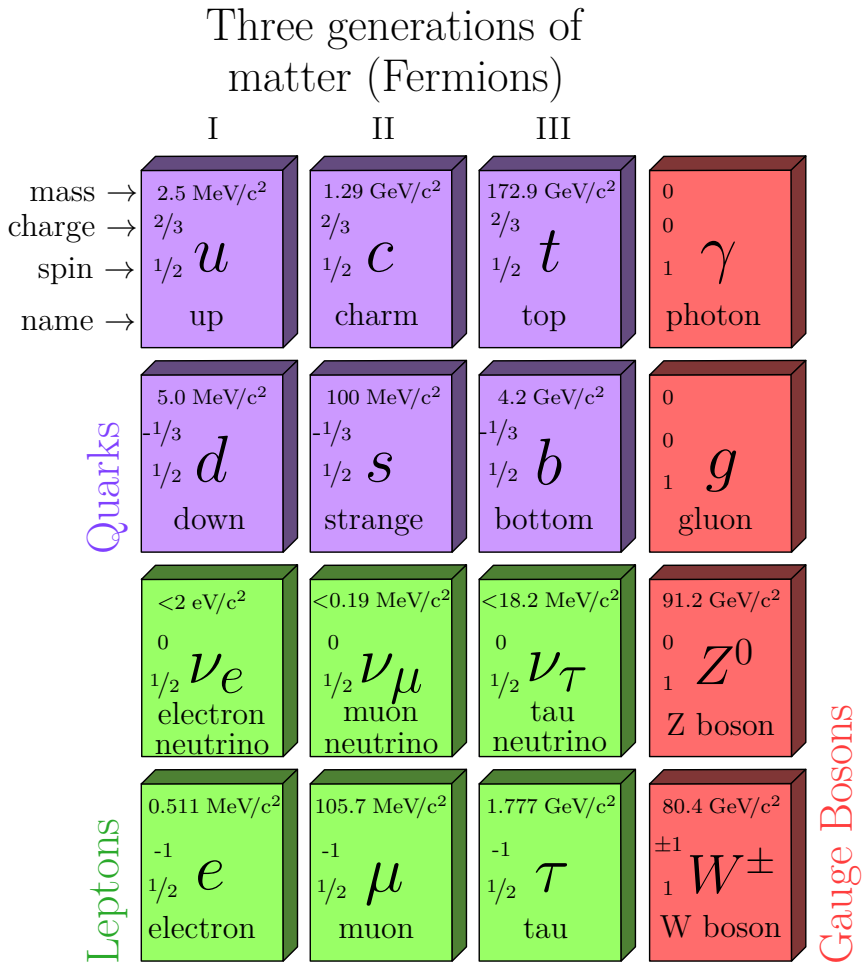


Figure I.1.: *Visualization of the elementary fermions and the gauge bosons of the Standard Model.*

All fermions carry weak isospin, the charge of the weak-interaction, and thus may take part in weak interactions. The charged leptons and the quarks carry also electromagnetic charge and may take part in electromagnetic interactions. The quarks are additionally influenced by the strong interaction, as they carry color, the charge of the strong interaction. Whereas the leptons do not feel the strong interaction and are colorless. Particles carrying color charge, so called colored particles, are not observed directly, i.e. quarks are confined into composite colorless particles, mesons consisting of a quark-antiquark pair or baryons consisting of three quarks.

All ordinary matter is made of first-generation particles. Particles from the other generations decay quickly into the first-generation ones. These decays change flavor and are enabled by the weak interaction. The eigenstates of the weak interaction are not identical to the mass eigenstates. The transformation of mass eigenstates into flavor eigenstates is accomplished by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [3, 4]. Particle decays can be observed in high-energy interactions as they occur in particle experiments, in the cosmos or in the atmosphere.

In modern particle experiments electrons or protons are accelerated and then brought to collision. The kinetic energy of the colliding particles can be transformed into mass and heavier particles may be produced. In collisions of compound particles, proton-proton collisions for example, the interaction happens between the constituents, the so called partons. This means that the kinetic energy of the partons has to be taken into account. The proton is composed of three valence quarks which are bound together by gluons. These gluons can split up into quark-antiquark pairs the sea quarks. All these partons share the momentum of the proton. The parton distribution function (PDF) $f_{i,p}(x_i, \mu^2)$ describes the probability density to find a parton i inside the proton p carrying a momentum fraction x_i, where $p_i = x_i \cdot p_p$. It depends on the factorization scale μ of the considered interaction. Figure I.2 shows the parton distribution function CTEQ5M1 [5] for protons at a factorization scale $\mu^2 = (175\ \text{GeV})^2$.

The masses of composite particles, such as protons and neutrons, are explained by Quantum Chromodynamics (QCD), the gauge theory of strong interactions. The dominant contribution to the nucleon mass is not the masses of the quarks that make up the nucleon, but the energy stored up in confining the quarks in a tiny volume [6]. The masses of the up and down quarks $M_u = 2.5^{+0.6}_{-0.8}\ \text{MeV/c}^2$ and $M_d = 5.0^{+0.7}_{-0.9}\ \text{MeV/c}^2$ [2] contribute about 2% to the mass of the proton (udd) $M_p = 938.272\ \text{MeV/c}^2$ [2].

The masses of the weak gauge bosons are explained by the electroweak theory and the Higgs mechanism. The electroweak theory is like QCD a gauge theory in which interactions follow from symmetries. It unifies the electromagnetic and the weak force in one single theory using weak-isospin and weak-hypercharge symmetries. Four massless gauge bosons emerge from this electroweak symmetry, as local gauge invariance would be violated by explicit mass terms for gauge bosons in the

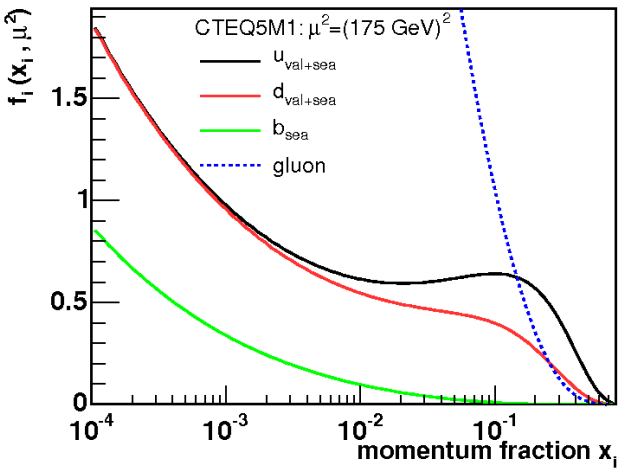


Figure I.2.: *The CTEQ5M1 parton distribution function at a factorization scale* $\mu^2 = (175\ \mathrm{GeV})^2$.

Lagrangian. In order to explain the boson masses an additional field, the Higgs field, is introduced. It breaks the electroweak symmetry and reveals a massive particle, the Higgs boson. The original four generators of the electroweak symmetry are all broken, but a linear combination of two of them is identified as the massless photon, with its coupling to charged particles. The orthogonal combination to the photon is the mediator of the neutral-current weak interaction Z. The Z boson acquires a mass like the two mediators of the charged-current weak interaction $W^{\pm}$. The mass of the weak interaction mediators is predicted as a function of the electroweak mixing angle θ_W.

Yukawa couplings of the Higgs field with the fermions are introduced to explain the masses of the fermions. The strength of the Yukawa couplings and hence the mass of the fermions is not predicted by the SM, but remains a free parameter like the electroweak mixing angle, to be determined by experiments.

The Higgs boson predicted by the SM has neither been discovered nor excluded, yet. It is one of the most urgent and challenging questions of particle physics. This is also one of the main goals of the Large Hadron Collider (LHC) at CERN. The Higgs boson production through gluon-gluon fusion (GGF) ($gg \to H$) via a heavy quark loop is predicted to have the highest production cross-section at the LHC. The production via two weak vector bosons WW or ZZ is referred to as vector boson fusion (VBF) ($qq \to qqH$). Its cross-section is predicted to be an order of magnitude lower than the GGF production cross-section.

The VBF production is not a discovery channel, but has several distinct features. From an experimental point of view, the characteristic two forward jets of the VBF process allows a good suppression of many background processes leading to

a relatively pure signal sample. From the theory point of view the VBF process allow direct access to the coupling of the Higgs boson to the weak vector bosons. Once observed, this process provides the ability to determine the properties of the Higgs boson and to test if they are the same as predicted by the SM. The VBF process does not depend on Yukawa couplings and may characterize the mechanism of electroweak symmetry breaking. It is not guaranteed that the same agent breaks the electroweak symmetry and generates fermion masses. In the fermiophobic model [7, 8] for example the Higgs boson couples only to weak gauge bosons. Summarizing, the VBF process may allow to establish the Higgs boson as the agent of electroweak symmetry breaking [9].

The aim of the analysis presented in this thesis is to analyze the first data of $35.2\,\mathrm{pb}^{-1}$ of the LHC taken by the ATLAS experiment in 2010. Due to low statistics and low production cross-sections of the Higgs boson the analysis does not only concentrate on the VBF process, but also includes the GGF process. Neural networks are employed as multivariate analysis method to fully exploit the kinematic differences between the Higgs signal and backgrounds. Thus, the benefit of this multivariate method compared to a cut-based analysis, the validation of the neural network approach and a robust statistical analysis form the focus of this thesis.

Contents

1. Theory

The first part of this chapter describes the theoretical framework of the Standard Model (SM), which is a gauge theory including the Higgs mechanism. The second part considers the production and decay modes of the Higgs boson as well as theoretical and experimental constraints on the Higgs boson mass.

1.1. The Standard Model Lagrangian

The SM is a non-Abelian gauge theory described by the mathematical framework of a quantum field theory with Lagrangians. It is based on the $\mathrm{SU(3)_C \times SU(2)_L \times U(1)_Y}$ gauge group. The subscripts refer to the physical application, i.e. C refers to color, L to the left-handed nature of the weak coupling and Y to the weak hypercharge quantum number. There are two kinds of fields, matter fields and gauge fields. Matter fields involve three generations i of fermions, quarks and leptons, forming left-handed doublets Q_i, L_i and right-handed singlets u_{R_i}, d_{R_i}, e_{R_i}.

$$
\begin{aligned}
Q_1 &= \begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{L}}, \; u_{R_1} = u_R \,, \; d_{R_1} = d_R \,, \quad L_1 = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_{\mathrm{L}}, \; e_{R_1} = e^-_R \\
Q_2 &= \begin{pmatrix} c \\ s \end{pmatrix}_{\mathrm{L}}, \; u_{R_2} = c_R \,, \; d_{R_2} = s_R \,, \quad L_2 = \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_{\mathrm{L}}, \; e_{R_2} = \mu^-_R \\
Q_3 &= \begin{pmatrix} t \\ b \end{pmatrix}_{\mathrm{L}}, \; u_{R_3} = t_R \,, \; d_{R_3} = b_R \,, \quad L_3 = \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_{\mathrm{L}}, \; e_{R_3} = \tau^-_R
\end{aligned}
$$

Gauge fields correspond to gauge bosons, that mediate the interactions. Quantum Chromodynamics (QCD) is described by the $\mathrm{SU(3)_C}$ group with gauge coupling g_s and eight gauge bosons, the gluons. It acts on the color of the left and right-handed quarks. The electroweak theory is described by the $\mathrm{SU(2)_L \times U(1)_Y}$ group. The $\mathrm{SU(2)_L}$ group has gauge coupling g, three gauge bosons W^i (i = 1,2,3) and acts only on the flavor of the left-handed fermions. The $\mathrm{U(1)_Y}$ group has gauge coupling g' and gauge boson B. It acts on the charge of left and right-handed fermions. The initial $\mathrm{U(2)_L \times U(1)_Y}$ is spontaneously broken to $\mathrm{U(1)_Q}$, the subscript Q refers to the electric charge. $\mathrm{U(1)_Q}$ incorporates Quantum Electrodynamics (QED) with its massless photon γ as linear combination of W^3 and B. The orthogonal combination

to the photon, the Z boson, and the $W^{\pm}$ bosons acquire mass. They mediate the weak force.

The SM Lagrangian is given by

$$\mathcal{L} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_f + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}$$

The gauge term $\mathcal{L}_{\text{gauge}}$ includes the gauge boson kinetic energy and self interactions. The fermion part $\mathcal{L}_f$ consisting of left-handed doublets and right-handed singlets, involves parity violation in SU(2)_L. The Higgs part $\mathcal{L}_{\text{Higgs}}$ incorporates interactions between the weak bosons and the Higgs field. The last term $\mathcal{L}_{\text{Yukawa}}$ represents the Yukawa couplings between the Higgs field and the fermions, which are needed to generate fermion masses. Introductions to gauge theory and the SM can be found in References [1, 9–17].

1.1.1. Local Gauge Invariance

An example of local gauge invariance is shown for QED, leading to a massless gauge boson, the photon. The Dirac Lagrangian

$$\mathcal{L} = i\hbar c\bar{\psi}\gamma^\mu\partial_\mu\psi - mc^2\bar{\psi}\psi$$

with the Dirac spinor ψ and the Dirac matrices γ_μ is invariant under global phase transformation $\psi \to e^{i\theta}\psi$, as the overall phase of the wave function is arbitrary. However, it is not invariant under local phase transformation

$$\psi \to e^{i\theta(x)}\psi$$

where the phase is a function of space-time points x^μ. For convenience this transformation is written as $\psi \to e^{iq\lambda(x)/\hbar c}\psi$ with $\lambda(x) = -\frac{\hbar c}{q}\theta(x)$ where q is the charge of the particle involved.

To achieve local gauge invariance a new vector field A_μ has to be introduced and the derivative ∂_μ has to be replaced by the covariant derivative

$$\mathcal{D}_\mu = \partial_\mu + i\frac{q}{\hbar c}A_\mu.$$

The new field transforms according to the rule

$$A_\mu \to A_\mu + \partial_\mu\lambda.$$

This new gauge field A_μ, introduced to preserve local gauge invariance, implies an additional free term

$$\mathcal{L}_A = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}\left(\frac{m_A c}{\hbar}\right)^2 A^\nu A_\nu$$

to be added to the Lagrangian. The first part with the field strength tensor $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ is locally gauge invariant, while the second part, the mass term of the field, is not. Therefore, the gauge field A_μ has to be massless. It is identified as the photon, the massless gauge boson of QED. The full Lagrangian reads

$$\mathcal{L} = i\hbar c\bar{\psi}\gamma^\mu \mathcal{D}_\mu \psi - mc^2\bar{\psi}\psi - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}$$

and is the Lagrangian of QED describing the interaction between fermions ψ and photons A^μ.

The principle of local gauge invariance is also applied in QCD leading to massless gluons as gauge bosons. In electroweak theory, this principle demands four gauge fields, three fields W_μ^i for the $\mathrm{SU(2)_L}$ group and one field B_μ for the $\mathrm{U(1)_Y}$ group. The covariant derivative is given by:

$$\mathcal{D}_\mu = \partial_\mu + i\frac{g}{2}\tau^i W_\mu^i + i\frac{g'}{2}YB_\mu \tag{1.1}$$

where τ^i denote the Pauli matrices. Using the electroweak mixing angle θ_W defined as $\tan\theta_W = \frac{g'}{g}$, the linear combinations of those fields make up the photon A_μ and the Z_μ and $W_\mu^\pm$ bosons.

$$\begin{aligned} A_\mu &= W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W \\ Z_\mu &= W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W \\ W_\mu^\pm &= \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp W_\mu^2\right) \end{aligned} \tag{1.2}$$

1.1.2. Spontaneous Symmetry Breaking

Gauge theories do not allow mass terms for gauge bosons, as these would break gauge invariance. But the weak interaction is short ranged and requires massive gauge bosons. The concept of spontaneous symmetry breaking is used to give masses to the weak gauge bosons. The symmetry of a Lagrangian, that is symmetric with respect to a symmetry group, is spontaneously broken if the vacuum state is not invariant under a corresponding transformation. The Lagrangian of two real scalar fields

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi_1\,\partial^\mu\phi_1 + \frac{1}{2}\partial_\mu\phi_2\,\partial^\mu\phi_2 - V(\phi)$$

with the potential

$$V(\phi) = -\frac{1}{2}\mu^2\left(\phi_1^2+\phi_2^2\right) + \frac{1}{4}\lambda^2\left(\phi_1^2+\phi_2^2\right)^2$$

is used as an example to present the concept of spontaneous symmetry breaking. The potential $V(\phi)$ depicted in Figure 1.1 is invariant under rotations in the ϕ_1-ϕ_2 plane.

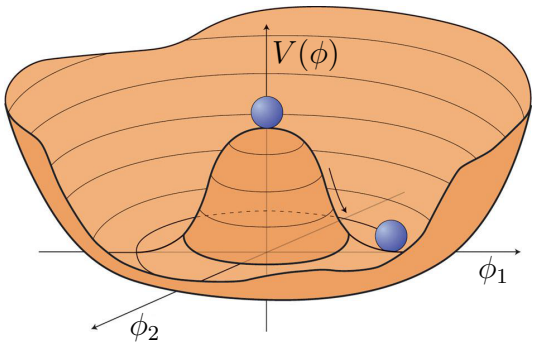


Figure 1.1.: *Illustration of a potential leading to spontaneous symmetry breaking.*

The minima of the potential are located on a circle with radius $\frac{\mu}{\lambda}$ around the origin. The rotational-symmetry is broken by choosing one particular ground state. New variables are introduced to parametrize fluctuations around the chosen ground state:

$$\eta = \phi_1 - \frac{\mu}{\lambda}, \qquad \xi = \phi_2$$

changing the Lagrangian to

$$\begin{aligned}\mathcal{L} =& \frac{1}{2}\partial_\mu\eta\,\partial^\mu\eta - \mu^2\eta^2 + \frac{1}{2}\partial_\mu\xi\,\partial^\mu\xi \\ & - \mu\lambda\left(\eta^3 + \eta\xi^2\right) - \frac{\lambda^2}{4}\left(\eta^4 + \xi^4 + 2\eta^2\xi^2\right) + \frac{\mu^4}{4\lambda^2}.\end{aligned}$$

The field η acquires mass $M = \sqrt{2}\mu\frac{\hbar}{c}$, while the field ξ remains massless. Further terms in the Lagrangian represent different couplings between the two fields, while the final term is just a constant. The massless scalar particle corresponding to the field ξ is referred to as the Goldstone boson.

1.1.3. Higgs Mechanism

The Higgs mechanism is introduced in the SM to generate the masses of the weak gauge bosons and preserve the gauge invariance. The simplest choice to break the symmetry of the $\mathrm{SU(2)_L{\times}U(1)_Y}$ group is a complex doublet of scalar fields ϕ.

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$$

The gauge invariant Lagrangian for the interactions and propagation of these scalars is

$$\mathcal{L}_{Higgs} = (\mathcal{D}^\mu\phi)^\dagger(\mathcal{D}_\mu\phi) - V(\phi^\dagger\phi)$$

with the covariant derivative defined in Equation1.1. The potential is given by

$$V(\phi^\dagger\phi) = \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2$$

with $\mu < 0$ and $\lambda > 0$ to apply spontaneous symmetry breaking. Choosing the minimum of $V(\phi^\dagger\phi)$ as

$$\phi_1 = \phi_2 = \phi_3 = 0 \quad , \quad \phi_4 = v^2 = -\frac{\mu^2}{\lambda}$$

with the vacuum expectation value v, the ground state $\langle\,\phi\,\rangle_0$ is given by

$$\langle\,\phi\,\rangle_0 = \langle\,0\,|\,\phi\,|\,0\,\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}.$$

The fields ϕ are written in terms of new fields θ^i, with $i = 1, 2, 3$ and h to parametrize fluctuations around the ground state:

$$\phi = \begin{pmatrix} \theta^2 + i\theta^1 \\ \frac{1}{\sqrt{2}}(v+h) - i\theta^3 \end{pmatrix} = e^{i\theta^i\tau^i/v}\,\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ (v+h\,) \end{pmatrix}.$$

The factor $e^{i\theta^i\tau^i/v}$ corresponds to three massless Goldstone bosons and is a local SU(2) transformation that vanishes after the appropriate gauge transformation. The only remaining field is the Higgs field h

$$\phi \to \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix}$$

Using this transformed field to expand the Higgs interaction term of the Lagrangian

$$|\mathcal{D}_\mu\phi)|^2 = \left|\left(\partial_\mu + i\frac{g}{2}\tau^i W_\mu^i + i\frac{g'}{2}Y B_\mu\right)\phi\right|^2$$

and replacing W_μ^i and B_μ with the definitions given in equation 1.2, leads to the mass terms of the gauge bosons:

$$M_W = \frac{1}{2}vg \quad , \quad M_Z = \frac{1}{2}v\sqrt{g^2+g'^2} = M_W/\cos\theta_W \quad \text{and} \quad M_A = 0.$$

Finally, inserting the transformed field in the potential $V(\phi^\dagger\phi)$ and using the relation $v^2 = -\frac{\mu^2}{\lambda}$ one obtains

$$V(\phi^\dagger\phi) = \lambda v^2\,h^2 + \lambda v\,h^3 + \frac{\lambda}{4}\,h^4$$

the mass $M_H = \sqrt{2\lambda v^2}$ of the Higgs boson and its self interaction terms. The vacuum expectation value $v = (\sqrt{2}G_F)^{-1/2}$ is fixed in terms of the Fermi constant G_F determined from muon decay. Due to the free parameter λ the mass of the Higgs boson is not predicted by the SM and has to be determined in experiments if the Higgs boson is observed. The Higgs Mechanism is discussed in details in References [18–23].

1.1.4. Yukawa Sector

The fermion masses can be generated using the same scalar field ϕ introduced by the Higgs mechanism, and defining the isodublet $\tilde{\phi} = i\tau_2\phi^*$. The Yukawa term of the SM Lagrangian $\mathcal{L}_{\text{Yukawa}}$, that is invariant under $\text{SU(2)}_\text{L}\times\text{U(1)}_\text{Y}$ transformations, is then given for any fermion generation by

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_e\,\bar{L}\,\phi\,e_R - \lambda_d\,\bar{Q}\,\phi\,d_R - \lambda_u\,\bar{Q}\,\tilde{\phi}\,u_R \;+\; h.c.\;.$$

Applying the procedure presented in the previous section leads in the case of the electron to

$$\mathcal{L}_{\text{Yukawa}} = -\frac{1}{\sqrt{2}}\,\lambda_e\,(\bar{\nu}_e, \bar{e}_L)\begin{pmatrix} 0 \\ v+h \end{pmatrix} e_R \;+\cdots = -\frac{1}{\sqrt{2}}\,\lambda_e\,(v+h)\,\bar{e}_L e_R \;+\cdots\;.$$

Generalizing this, one obtains the fermion mass identified with the term in front of $\bar{f}_L f_R$

$$M_e = \frac{\lambda_e\,v}{\sqrt{2}} \;\;,\;\; M_u = \frac{\lambda_u\,v}{\sqrt{2}} \;\;,\;\; M_d = \frac{\lambda_d\,v}{\sqrt{2}}\;.$$

The masses of neutrinos are not considered in the formalism presented here. In principle one could add additional Yukawa couplings [24], but many physicist prefer a different solution, called see-saw mechanism [25, 26], to explain the large difference between the masses of the quarks and leptons on the one hand side and neutrinos on the other hand side.

The Yukawa term is added to the SM Lagrangian following the rule to use all terms allowed by the symmetry with the chosen fields. The origin of the couplings is unknown, they do not follow from a known symmetry principle. The

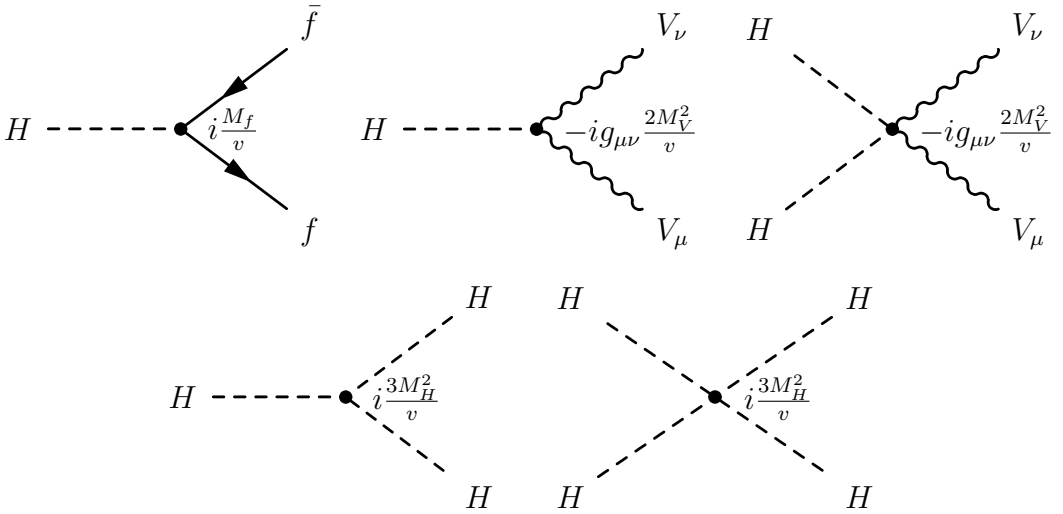


Figure 1.2.: *The Higgs boson couplings to fermions and gauge bosons (top row), and the Higgs self couplings (bottom row) in the SM.*

Yukawa couplings that reproduce the observed quark and lepton masses range over many orders of magnitude, it is unknown what fixes the size of the Yukawa couplings.

Summarized, the Higgs mechanism allows to generate the masses M_V of the weak vector bosons $W^{\pm}$ and Z. It preserves the $\mathrm{SU(2)_L}\times\mathrm{U(1)_Y}$ gauge symmetry which is spontaneously broken, while the $\mathrm{U(3)_C}$ color symmetry and the electromagnetic $\mathrm{U(1)_Q}$ symmetry stay unbroken. The masses of the fermions are generated by interactions with the Higgs field involving Yukawa couplings. The interactions of the Higgs field with fermions, gauge bosons as well as its self interactions and the corresponding couplings are presented in Figure 1.2.

1.2. The Higgs Boson

Many tests of the SM have been performed at the per mill level, measuring the couplings of quarks and leptons to the electroweak gauge bosons, as well as the trilinear couplings among the electroweak gauge bosons. The couplings have been found to be those dictated by the gauge symmetry [27, 28]. With the tests of the strong interactions the SM based on the $\mathrm{SU(3)_C}\times\mathrm{SU(2)_L}\times\mathrm{U(1)_Y}$ gauge symmetry has been established as the theory of the strong and electroweak interactions at present energies. The only missing ingredient is the Higgs boson, which also provides cancellation of high-energy divergences in the production of longitudinally polarized W^+W^- pairs in electron-positron collisions [9, 28], but it has not yet been observed directly.

1.2.1. Higgs Boson Production and Decay

The most important production mechanisms for the Higgs boson at the Large Hadron Collider (LHC) are presented in Figure 1.3:

(a) the gluon-gluon fusion (GGF) ($gg \to H$) process, that proceeds via a virtual top quark loop,

(b) the vector boson fusion (VBF) ($qq \to qqH$) process,

(c) the associated production with vector bosons $W^{\pm}$ or Z denoted as V, and

(d) the associated production with a $t\bar{t}$ pair.

The SM production cross-sections of these processes at a center-of-mass energy of 7 TeV is given in Figure 1.4 as a function of the Higgs boson mass.

The GGF process is predicted to have the highest production cross-section, while the cross-section of the VBF process is an order of magnitude lower. Nevertheless, VBF plays an important role as it features two characteristic forward jets, that

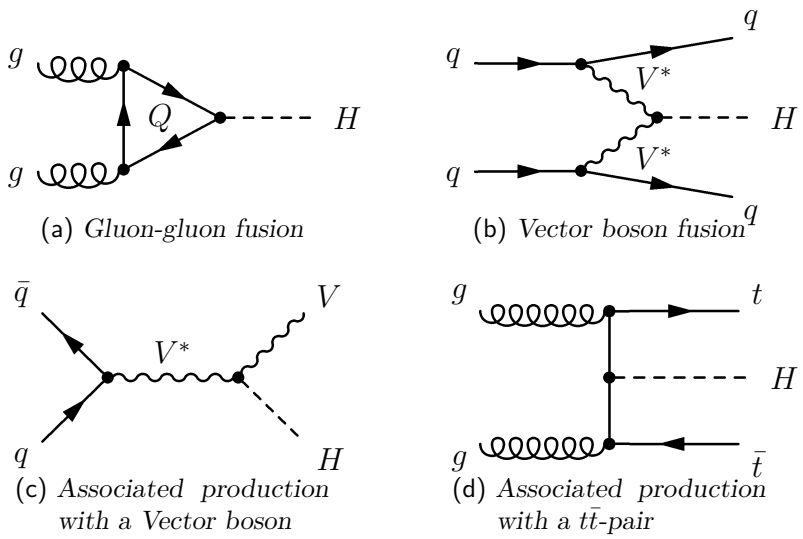


Figure 1.3.: *Most important Higgs boson production diagrams for the LHC, including gluon-gluon fusion (a), vector boson fusion (b), associated production with vector a boson (c) and associated production with a $t\bar{t}$ pair (d).*

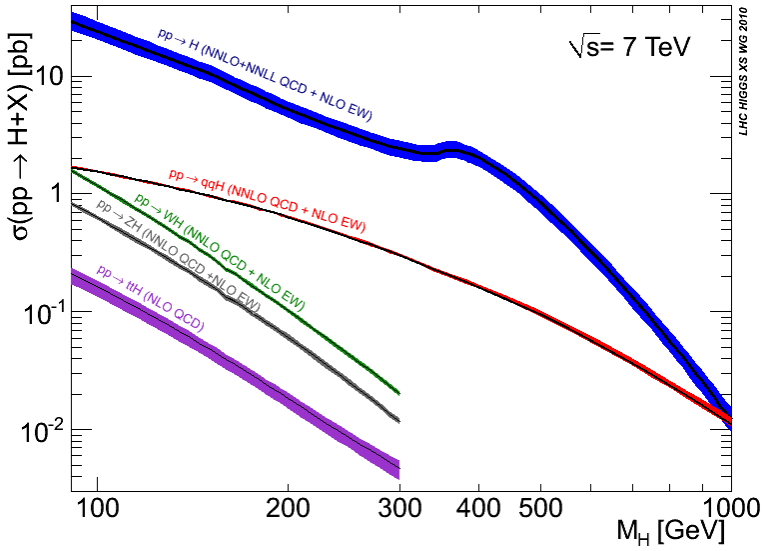


Figure 1.4.: *Standard Model Higgs boson production cross-sections at $\sqrt{s} = 7$ TeV as a function of the Higgs boson mass [29].*

allow for a good suppression of many background processes. Furthermore, it does not depend on Yukawa couplings, but provides direct access to the coupling of the Higgs boson and the weak vector bosons giving the ability to determine if nature implements the properties of electroweak symmetry breaking as predicted by the SM.

As for the production, heavy particles are preferred in the Higgs boson decay. The Higgs boson decay branching ratios and the production cross-section times branching ratios as function of the Higgs boson mass are shown in Figure 1.5. The decays $H \to WW$ and $H \to ZZ$ dominate for large M_H. Below the WW threshold, down to 135 GeV, the decays to WW^* or ZZ^*, with V^* off-shell vector bosons, dominate. For even lower masses the decay $H \to b\bar{b}$ is most important, but challenging due to the huge QCD production of $b\bar{b}$ pairs at the LHC. Thus, rare decay modes as $H \to \tau\bar{\tau}$ or $H \to \gamma\bar{\gamma}$, mediated by a top quark or a W boson loop, become more important in the low mass region.

This analysis considers Higgs boson production via VBF and the decay mode $H \to WW^* \to \ell\nu\ell\nu$, exploiting the specific topological properties: two forward jets, two isolated leptons from the W decays and missing transverse energy caused by the two neutrinos. The properties of this process are discussed in more details in chapter 4. The dataset to be analyzed contains 35 pb^{-1} of data taken with the ATLAS experiment at a center-of-mass energy of 7 TeV. Considering the small amount of data and the small production cross-section of VBF Higgs production the analysis also includes the GGF mode.

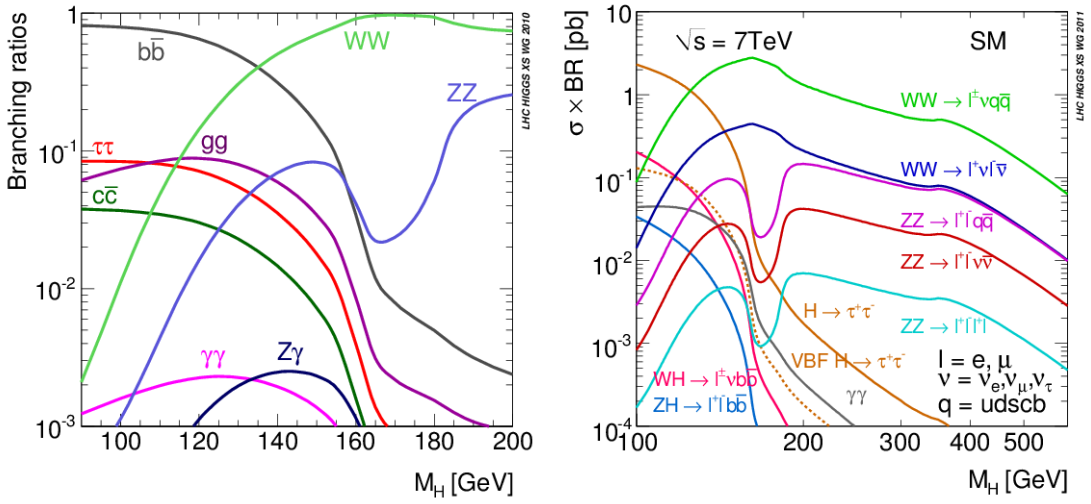


Figure 1.5.: *Standard Model Higgs boson decay branching ratios on the left hand side and production cross-section (at $\sqrt{s} = 7$ TeV) times branching ratios on the right hand side as a function of the Higgs boson mass [29].*

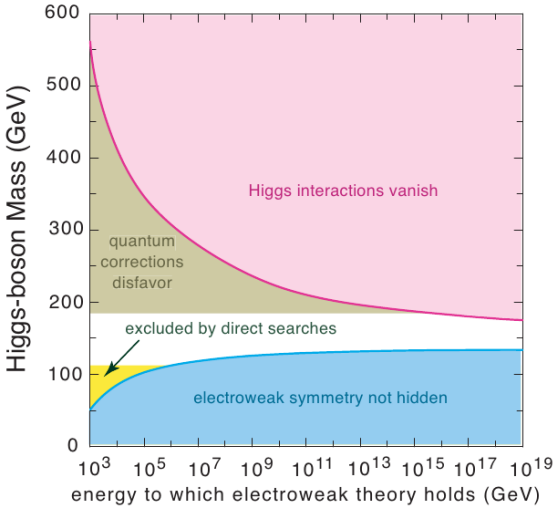
1.2.2. Theoretical and Experimental constraints

The Higgs boson mass obtained by applying spontaneous symmetry breaking is

$$M_H = \sqrt{2\lambda v^2} \qquad \text{with} \qquad \lambda = \frac{gM_H^2}{8M_W^2} = \frac{G_F M_H^2}{\sqrt{2}}\,,$$

the vacuum expectation value $v = 246$ GeV and Fermi's weak interaction coupling constant $G_F = 1.16637 \times 10^{-5}$ GeV^{-2}. The only constraint imposed so far is the vacuum stability $\lambda > 0$, which allows Higgs boson masses from 0 to ∞. However, there are more stringent theoretical constraints that lead to upper and lower bounds. Renormalization constraints [30] impose the so called *"triviality"* upper bound on M_H, where Higgs interactions vanish. The lower bound is obtained by computing quantum corrections [31–36] to the Higgs potential and requiring the ground state $\langle \phi \rangle_0$ to be a global minimum. The upper and lower bounds plotted in Figure 1.6 are results of two-loopcalculations [37, 38]. Figure 1.6 also indicates the upper bound on $M_H < 185$ GeV/c^2 at 95% confidence level derived from precision electroweak measurements [39] and the lower limit $M_H > 114.4$ GeV/c^2 at 95% confidence level from direct searches at LEP [28].

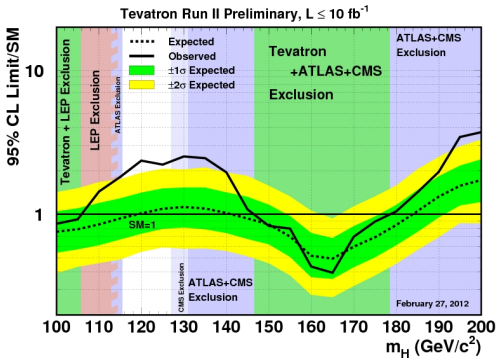
Further experimental limits are presented in Figure 1.7: (a) the combination of results obtained by the CDF and DØ experiments at Tevatron [40]; (b) and (c) results in the full mass range, (d) and (e) results in the low mass range from the LHC experiments ATLAS [41] and CMS [42], respectively. The SM Higgs boson



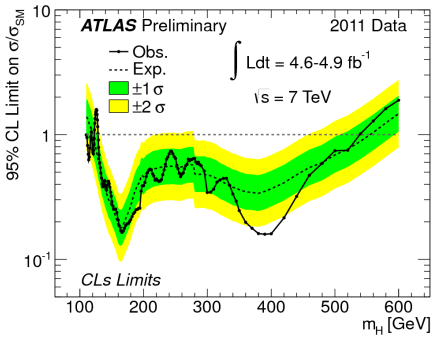
(a) *Theoretical bounds*

Figure 1.6.: *Theoretical bounds [9] on the Higgs boson mass including the lower limit* $M_H > 114.4$ GeV/c^2 *at 95% confidence level [28] from direct searches at LEP.*

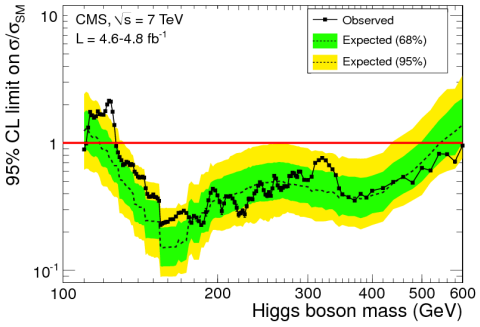
is excluded for all masses up to around 600 GeV/c^2 at the 95% CL, except in the region around 120 GeV/c^2 to 130 GeV/c^2 where a small excess of events is seen by all experiments.



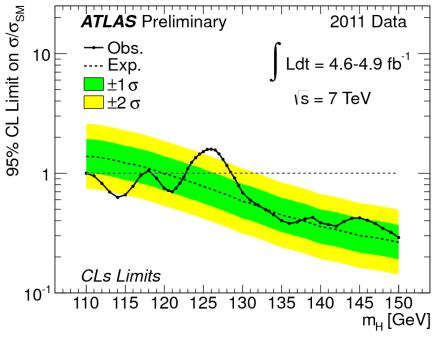
(a) *Tevatron limits*



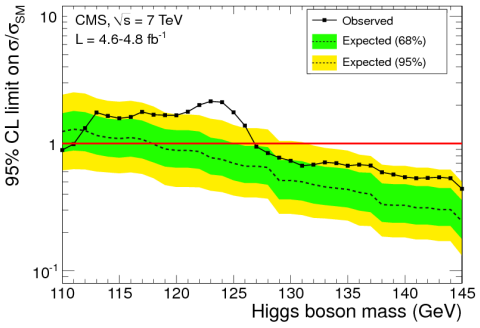
(b) *ATLAS limits full mass range*



(c) *CMS limits full mass range*



(d) *ATLAS limits low mass range*



(e) *CMS limits low mass range*

Figure 1.7.: *Limits on the Higgs boson mass: (a) from Tevatron [43]; (b) and (c) limits in the full mass range, (d) and (e) limits in the low mass range from the LHC experiments ATLAS [41] and CMS [42], respectively.*

2. LHC and Atlas

Particle physics experiments with the currently highest reachable energies are performed at the Large Hadron Collider (LHC) located at CERN. Charged particles, protons or lead-ions, are accelerated by the LHC and brought to collisions at four interaction points that are surrounded by detectors. There are six different experiments recording events that arise in the collisions. A Toroidal LHC Apparatus (ATLAS) [44] and Compact Muon Solenoid (CMS) [45] are two multi-purpose detectors that explore a wide range of physics processes, starting with the established SM processes, the Higgs boson searches, right through to new physics phenomena that may occur. A Large Ion Collider Experiment (ALICE) [46] is a detector dedicated to lead-ion collisions, investigating the properties of a quark-gluon plasma. This plasma presumably existed at the beginning of the universe, where quarks and gluons were not confined in hadrons. The second specialized experiment is LHCb. In contrast to the other experiments, that cover the whole solid angle, the LHCb detector covers only one forward region. Its focus is on b-quark physics, studying boosted b and c hadrons in particular and the small asymmetry between matter and antimatter. The last two experiments LHC forward (LHCf) [47] and Total Elastic and diffractive cross-section Measurement (TOTEM) [48] are smaller and share their interaction points with ATLAS and CMS, respectively. LHCf is designed to measure neutral pions produced close to the beam, in order to estimate the primary energy in high-energy cosmic rays. While TOTEM measures the total proton-proton cross-section studying elastic scattering and diffractive processes.

After a short overview of the LHC and its associated pre-accelerators, the ATLAS detector will be described as this thesis analyzes the data of proton-proton collisions recorded by ATLAS in 2010.

2.1. The Large Hadron Collider

The CERN accelerator complex is sketched in Figure 2.1 indicating the two chains of pre-accelerators needed to fill the LHC with proton or lead-ion beams. Protons are produced in a duoplasmatron source [49], where thermic emission of electrons ionises hydrogen gas. The protons are then injected in the LINear ACcelerator LINAC2 and accelerated to 50 MeV. The resulting pulsed beam is transferred to the Proton Synchrotron BOOSTER [50] and further accelerated to 1.4 GeV. The Protons are

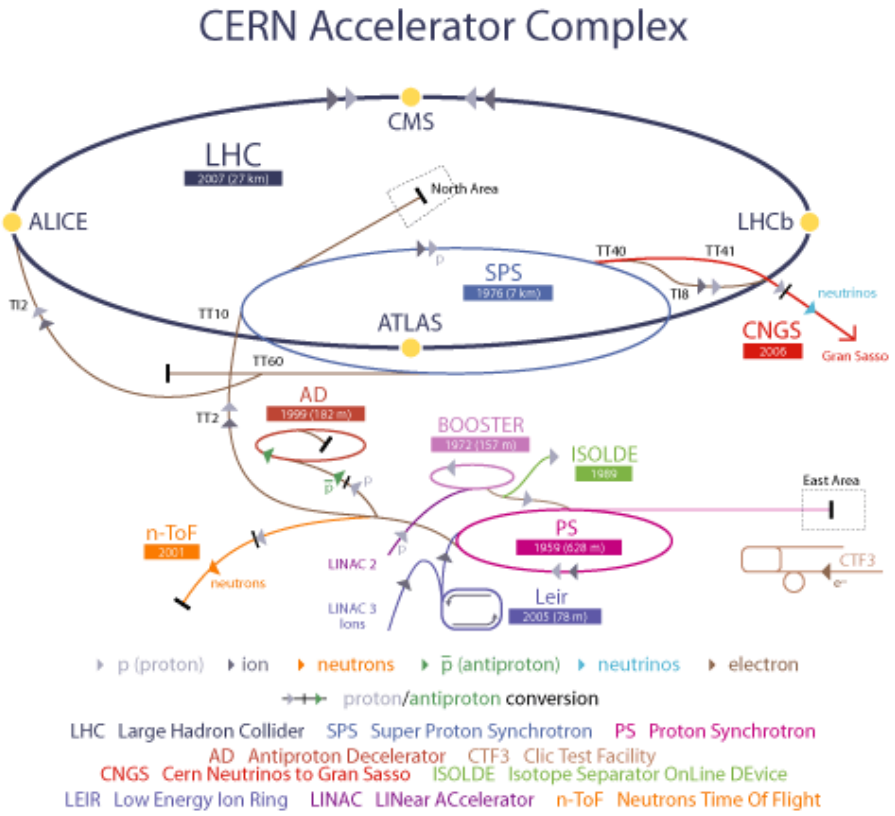


Figure 2.1.: *Overview of CERN accelerator complex indicating the two chains of pre-accelerators needed to fill the LHC with proton or lead-ion beams.*

then injected in the Proton Synchrotron (PS) [51]. After forming the pulsed beam to bunches and accelerating them to 25 GeV the beam is subsequently transferred to the Super Proton Synchrotron (SPS) [52]. The SPS increases the beam energy to 450 GeV. Finally, the beam is injected into the LHC [53] via two transfer lines to obtain a clockwise and a counter-clockwise rotating beam. In 2010 and 2011 the LHC operated with a beam energy of 3.5 TeV, but it is designed to reach a beam energy of 7 TeV and will be prepared for such operations in the future.

The LHC utilizes eight superconducting radiofrequency cavities per beam to accelerate the bunches of hadrons. The cavities are designed to deliver an accelerating field of 5 MV/m with a frequency of 400 MHz while cooled down to 4.5 K. A total of 1232 superconducting dipole magnets bend the beams to their orbit. A magnetic field strength of 8.33 T is necessary to bend the protons at 7 TeV. The dipole magnets are cooled to 1.9 K by superfluid helium which involves a complex cryogenic system. 858 quadrupole magnets and 6208 correcting magnets focus the beams. The total power consumption of the LHC is 180 MW.

Particle colliders are characterized by the instantaneous luminosity $\mathcal{L}$, which

denotes the number of particles going through a square centimeter per second at the interaction point. It depends only on the beam parameters, determined by the accelerator, such as the number of colliding bunches b, the number of particles per bunch N, the revolution frequency f_{revol} and the beam size at the interaction point:

$$\mathcal{L} = \frac{b}{4\pi} \frac{N^2 f_{revol}}{\epsilon^* \beta^*} \gamma F \,,$$

where ϵ^* is the normalized transverse beam emittance and β^* is the beta function that is a measure for the extent of the beam. The * indicates that these parameters are evaluated at the interaction point. The relativistic gamma factor is defined as the fraction of the energy E over the rest energy E_0 of the accelerated particles $\gamma = \frac{E}{E_0}$, while F denotes the geometric luminosity reduction factor due to the crossing angle at the interaction point. The integrated luminosity $\mathcal{L}_{int} = \int \mathcal{L}\, \mathrm{d}t$ characterizes the amount of data, taken in a given time period. The number of events N of a given process with cross-section σ is obtained by $N = \sigma \mathcal{L}_{int}$. The design instantaneous luminosity of the LHC is 10^{34} cm^{-2}s^{-1} and the instantaneous luminosity reached in 2011 was 3.5×10^{33} cm^{-2}s^{-1}.

2.2. The ATLAS Experiment

The ATLAS experiment is a multi-purpose detector designed for precision test of the SM at the TeV scale. It consists of a series of subdetectors built symmetrically around the interaction point. An overview of the ATLAS experiment is shown in Figure 2.2. The performance goals, summarized in Table 2.1, enable efficient tracking provided by the inner detector, electron and photon identification due to good electromagnetic calorimetry, jet and missing energy measurements by a wide-coverage hadronic-calorimetry, and precision muon momentum measurements.

Table 2.1.: *General performance goals of the ATLAS experiment [44]. The units of E and p_{T} are in* GeV.

detector component	required resolution	η coverage measurement	η coverage trigger
tracking	$\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 0.05\% \cdot p_{\mathrm{T}} \oplus 1\%$	± 2.5	-
EM calorimetry	$\sigma_E/E = 10\% \,/\, \sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry			
barrel and end-cap	$\sigma_E/E = 50\% \,/\, \sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\% \,/\, \sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 10\%$ at $p_{\mathrm{T}} = 1$ TeV	± 2.7	± 2.4

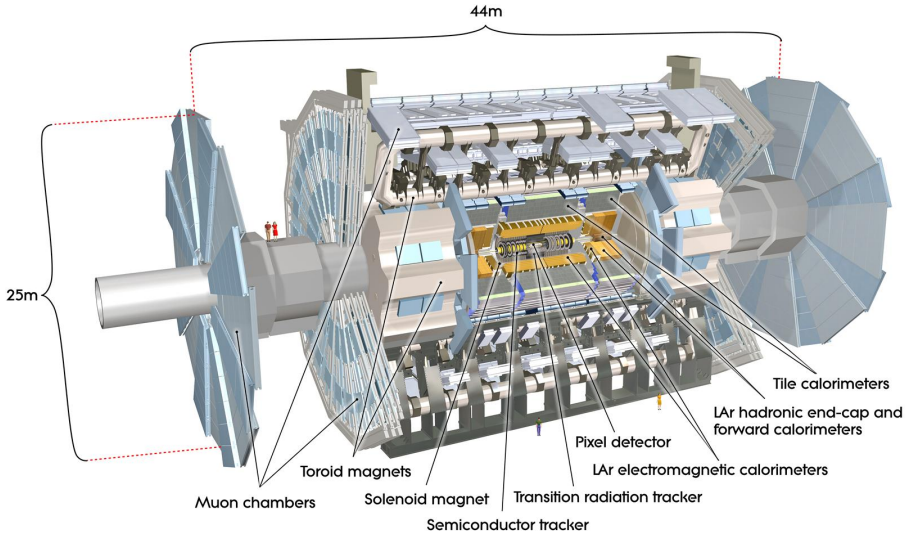


Figure 2.2.: *Overall view of the ATLAS experiment with its subdetectors [44].*

The coordinate system adopted by ATLAS is a right-handed coordinate system. It has the origin centered at the nominal collision point inside the detector, the y-axis pointing vertically upward, the x-axis pointing radially inward toward the center of the LHC, and the z-axis along the beam direction. The azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle θ is measured from the z-axis. Instead of θ it is often more handy to use the rapidity y defined as

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

But in the majority of cases the pseudorapidity defined as

$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$$

is used. For massless particles the rapidity y is equal to the pseudorapidity η. Thus, the angular distance between two point objects, as observed from the origin of the ATLAS detector, is expressed as

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

The momentum measured transverse to the beam direction, denoted by p_{T} is computed from the x and y components. It is the projection of the momentum p onto the x-y plane

$$p_{\mathrm{T}} \equiv \sqrt{p_x^2 + p_y^2}$$

The transverse energy E_{T} is defined as

$$E_{\mathrm{T}} \equiv E \sin\theta$$

and used for objects measured in the calorimeter.

The following sections describe briefly the subdetectors of the ATLAS experiment. A detailed description including important properties and nice illustrations can be found in Reference [44].

2.2.1. The Inner Detector

The Inner Detector (ID) is designed to measure the trajectories of charged particles with high momentum, angle and position resolution. It comprises three subdetectors, from inside out:

- The pixel detector [54] is arranged in three barrel layers and three endcap discs instrumented with silicon pixels with a size of $50 \times 400\ \mu\mathrm{m}^2$. It provides a typical resolution of 10 µm for the $r\phi$ measurements and about 115 µm for the measurement along the z axis. This allows a good reconstruction of vertices and is the basis for b-jet identification.
- The SemiConductor Tracker (SCT) [55] consists of four barrel layers and nine endcap disks, built around the pixel detector. Micro strips with a pitch of 80 µm are used as stereo layers with an angle of 40 mrad, to allow a measurement in z direction, and provide a resolution of 17 µm in $r\phi$ and 580 µm in z.
- The Transition Radiation Tracker (TRT) [56] is the most outer part of the ID. It combines the concept of a straw tracker with transition radiation detection for particle identification. The straws have a diameter of 4 mm and are filled with a xenon based gas mixture. The intrinsic straw resolution is 130 µm.

The ID has a diameter of 2.1 m and is 6.2 m long, providing a coverage of $|\eta| < 2.5$. It is surrounded by a solenoid magnet [57] with a magnetic field strength of 2 T that bends the tracks of charged particles and thus allows the measurement of their momenta. The silicon components of the ID are operated at a temperature of about -10 °C, while the TRT is operated at room temperature.

2.2.2. Electromagnetic Calorimeter

The electromagnetic calorimeter [58] is a sampling detector separating particle absorption and signal readout. It is subdivided in a barrel part with $|\eta| < 1.5$, endcap with $1.5 < |\eta| < 2.5$, and forward calorimeter with $2.5 < |\eta| < 3.2$. All calorimeters are housed in cryostats. A presampler filled with liquid argon is placed in front of the calorimeter to measure the energy loss of the particles caused by

the solenoid. Liquid argon is the active material in the calorimeter. The absorbing part is formed by lead in the barrel and by copper in the forward calorimeters [59]. The liquid argon/lead structure in the barrel is built in accordion shape, ensuring full ϕ coverage. Three layers with increasing segmentation in η and also increasing granularity ($\phi \times \eta$), are used to measure the direction of the particles in the first layer, the main fraction of the energy produced by the electromagnetic shower in the second layer, and the tail of the shower in the third layer, which is also used for trigger purposes. The barrel extends from $r = 1.4$ m to $r = 2$ m and corresponds to a radiation length of $X_0 \approx 24$. The endcaps are 63 cm thick corresponding to $X_0 \approx 26$.

2.2.3. Hadronic Calorimeter

The hadronic calorimeter [60] consists of a barrel with $|\eta| < 1.0$, two extended barrels with $0.8 < |\eta| < 1.7$, two endcaps with $1.5 < |\eta| < 3.2$, and two forward calorimeters with$3.1 < |\eta| < 4.9$. The different parts overlap to ensure an overall coverage in η. The barrel and extended barrel calorimeters use iron plates as absorber and scintillating tiles as active material. The inner and outer radii of the barrel and extended barrel are 2.28 m and 4.25 m. They are divided in three layers with 1.5, 4.1 and 1.8 interaction lengths in the barrel and 1.5, 2.6 and 3.3 interaction lengths in the extended barrel providing good resolution for high energy jets. Both, the endcap and the forward calorimeters use liquid argon as active material. Copper is deployed as the absorber material in the endcaps, while tungsten is used in the forward calorimeter [59], corresponding to 10 interaction lengths. The endcap and forward hadronic calorimeters are contained in the same cryostats as the electromagnetic ones. The total length of the hadronic calorimeter is 11 m.

2.2.4. Muon Spectrometer

The muon spectrometer [61], with its characteristic toroid magnets forms the outermost part of the ATLAS experiment. The toroid consists of eight superconducting coils forming a barrel with a length of 25.3 m and an inner and outer diameter of 9.4 m and 20.1 m. It provides a magnetic field orthogonal to the particle trajectories with a magnetic field strength of 0.5 T in order to deflect the muons and enable momentum measurements. Two smaller toroids are installed in the endcaps providing a field strength of 1 T.

The tracking of muons is accomplished by three layers of monitored drift tubes (MDTs) that cover a pseudorapidity region of $|\eta| < 2.7$. The MDTs, built of aluminum tubes and filled with an argon gas mixture, provide a per chamber resolution of 35 µm. An additional layer of cathode strip chambers (CSCs) is installed in the

endcaps $2.0 < |\eta| < 2.7$ providing a high rate capability and high time resolution to cope with the higher amount of tracks in this region. The spatial resolution of of the CSCs is 40 µm per chamber.

Additional trigger chambers are incorporated in the pseudorapidity range $|\eta| < 2.4$. Resistive plate chambers are used in the barrel and gap chambers in the end-cap regions.

2.2.5. Trigger

A trigger system is necessary to

- reduce the output rate of events, as the high event rate produced by the LHC exceeds the current technical possibilities of storage elements,
- select interesting events with cross-sections that are several magnitudes lower than the total cross-section at the LHC.

The trigger system [62, 63] implemented by ATLAS consists of three levels. The Level-1 (L1) trigger based on custom-build hardware reduces the event rate from 40 MHz to 75 kHz. The decision of the L1 trigger is based on the properties of the following objects reconstructed with reduced granularity: electromagnetic clusters, tau leptons, jets, missing transverse energy, the scalar sum of transverse energies in the calorimeter and muon tracks reconstructed in the muon system. The processing time of the L1 trigger is 2.4 µs.

Further reduction is achieved by the software based Level-2 (L2) trigger. It re-reconstructs the objects defined by the L1 trigger using additional information of the inner detector, and combining the data provided by the different subdetectors. The processing time including data transfers is 40 µs, reducing the output rate to 2 kHz.

The last level is called Event Filter (EF) and is software based. The L2 trigger and the EF are referred to as High Level Trigger. The EF is executed after the event building process. Therefore, it has direct access to the complete properties of an event and uses more complex pattern recognition algorithms. The average processing time of the EF is 4 s and the final output rate is about 200 Hz.

2.2.6. Luminosity and forward detectors

The luminosity is measured by several forward detectors. The first one, the MBTS (Minimum Bias Trigger Scintillators) [64], is located at $z = \pm 3.65$ m from the collision point and consists of segmented scintillator counters. The main purpose of the MBTS is to provide a trigger on minimum collision activity during a pp bunch crossing. The second detector, called LUCID (LUminosity measurement using

Cherenkov Integrating Detector) [65], is placed at a distance of ±17 m from the interaction point. It monitors the beam conditions and the online luminosity, and measures the integrated luminosity by detecting inelastic *pp* scattering in the forward region. An other detector, called ALFA (Absolute Luminosity For ATLAS) [66], is located at a distance of ±240 m. It is designed to measure the elastic scattering at small angles and determines the absolute luminosity. ALFA is not commissioned, yet.

An additional forward detector, the Zero Degree Calorimeter [67], with the purpose of detecting forward neutrons in heavy-ion collisions is placed at a distance of ±140 m from the interaction point. There are proposals to build and install forward proton detectors, called Atlas Forward Protons [68, 69], to enhance the ATLAS baseline physics program, facilitating an improved measurement of the anomalous couplings between γ and W or Z bosons as well as QCD studies.

2.2.7. ATLAS performance and Data Quality

Figure 2.3 presents the luminosity delivered by the LHC in 2010 and 2011 as well as the recorded luminosity by ATLAS, showing a good performance of the ATLAS experiment. More details on luminosity measurements are provided in References [64, 70].

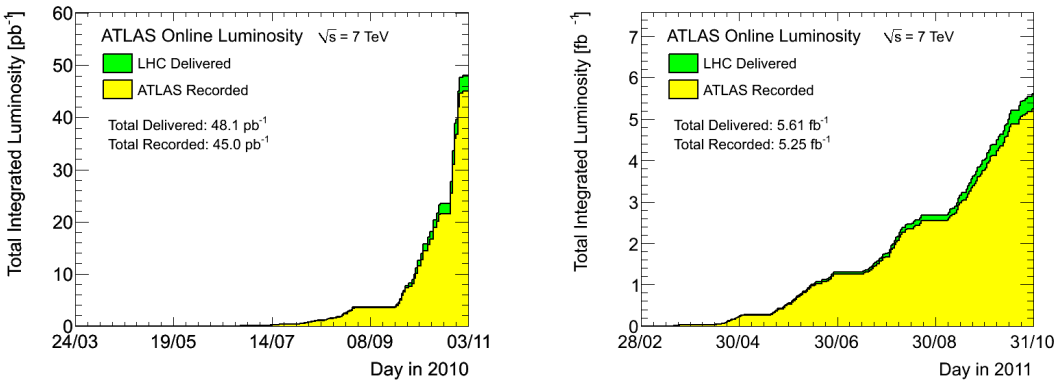


Figure 2.3.: *The delivered and recorded luminosity in 2010 on the left hand side and in 2011 on the right hand side.*

The Data Quality (DQ) selection within ATLAS is based upon inspection of a standard set of distributions that leads to a data quality assessment which is encoded in so-called DQ flags [71]. DQ flags are issued for each detector, usually segmented in subdetectors like barrel, end-caps and forward if such subdetectors exist. DQ flags are also issued for trigger slices and for each physics object reconstruction. Hence, the state of the ATLAS detector from hardware to physics object reconstruction is

expressed through DQ flags, which are saved per luminosity block. A luminosity block is a time interval of typically two minutes.

The DQ information is used in analyses through dedicated lists of good runs/luminosity blocks. Good run lists are formed by DQ selection criteria in addition to other criteria, such as run range, magnetic field configuration and beam energy. A complete list of valid physics runs and luminosity blocks is used in each analysis.

3. Event Generation, Simulation and Reconstruction

The analyses of collision events require deep understanding not only of the physics processes involved, but also of the acceptances and the response of the detectors used, as well as reconstruction effects and selection efficiencies imposed by the analysis techniques. Monte Carlo (MC) simulations help to acquire the needed understanding of these effects. In a first step MC events are generated according to our knowledge of particle physics as described by the SM, simulating kinematic properties of the considered physics processes. In the second step each MC event is subject to the simulation of the detector response, where the interactions of the simulated particles with the detector material are evaluated. Finally, the simulated events are reconstructed with the same algorithms as used for collision data, allowing a direct comparison of simulated events and observed data.

3.1. Monte Carlo Event Generation

MC event generators implement the current knowledge of particle physics to provide simulated events with properties very close to those observed in data. The generation of parton interactions according to known probability distributions and matrix-elements, and the integration of the squared amplitude are performed using MC techniques. The generation is factorized into several components isolating different phases, that may be treated separately because they describe processes at different energy scales. The general structure of a simulated proton-proton collision is sketched in Figure 3.1.

The momentum transfer of the hard process takes place at large scales, thus the cross-section can be computed with perturbation theory. As the partons of the hard process stem from compound particles, protons in this case, the cross-section has to be folded with the PDFs that describe the momentum distribution inside the proton. Initial- and final-state radiation is taken into account by the matrix element method including exact kinematics, possible interferences, and helicity correlations. In the following step, called parton shower, single partons may radiate other partons or split into several partons in a chain of $1 \to 2$ processes. The parton shower uses approximations derived by simplifying the kinematics and the interference structures

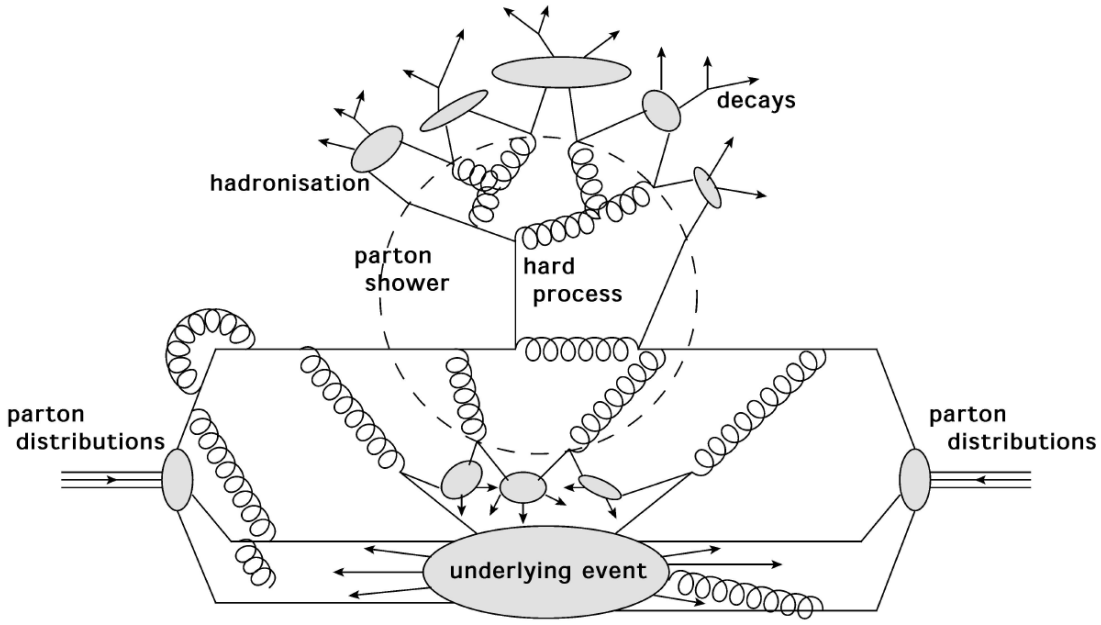


Figure 3.1.: *General structure of a hard proton-proton collision [72].*

of the matrix-elements. After the perturbative shower, the resulting partons that end up close in phase space are grouped into color-singlet clusters. This process simulates the formation of hadrons, the so called hadronization, and is based on phenomenological models that transfer colored partons to colorless hadrons. The decay of unstable hadrons is also simulated.

The remnants of the colliding protons undergo similar shower and hadronization processes forming the underlying event, that depends on the initial hard process as it has some color exchange to guarantee overall neutralization. Finally, additional pile-up or minimum bias events are added to simulate additional interactions that may occur in the same bunch crossing.

3.1.1. MC event generators

There are different types of event generators. Matrix-element generators provide only the simulation of specific hard processes and compute the corresponding cross-section. The output of such generators is fed to multi-purpose generators that also provide showering and hadronization routines. Hard and well separated emission of partons is well described by matrix-element generators, while soft emissions are modeled in a better way by parton showers. The combination of both methods implies an additional requirement. As both methods may lead to the same n-parton

final state configurations, so called matching schemes [73–76] have to be applied to avoid double-counting. There are also software packages dedicated to particle decays. The generators relevant for this analysis are briefly introduced in the following.

Matrix-element Generators

- **Alpgen**
  ALPGEN [77] generates SM processes in hadronic collisions with the emphasis of final states with large parton multiplicities. An exact computation of tree-level matrix elements is performed for a large set of QCD and electroweak processes. The b-quark and top-quark masses are included as well as top-quark and gauge-boson decays with helicity correlations. The full color and flavor structure of each event is provided enabling the usage of shower and hadronisation generators.

- **gg2WW**
  GG2WW [78] calculates the loop-induced gluon-fusion processes $gg \to W^*W^* \to$ leptons. Contribution from massive top- and bottom-quark loops are included as well as intermediate Higgs boson contributions with full spin and decay angle correlations, off-shell and interference effects. Arbitrary invariant masses of W bosons are also allowed.

- **MadGraph/MadEvent**
  MADGRAPH [79] automatically generates the Feynman diagrams and the amplitudes for all relevant subprocesses leading to the same final state. The program produces the mappings for the integration over the phase space, managing processes with up to eight final state partons. The matrix elements and phase-space mappings are passed to MADEVENT to calculate the cross-section and produce the actual events.

- **MC@NLO**
  MC@NLO [80, 81] provides a method for matching the next-to-leading order (NLO) calculation of specific process with the parton shower MC simulation implemented in HERWIG. The method has the following features: generation of fully exclusive events, with hadronization according to the cluster model implemented in HERWIG; total exclusive rates are accurate to NLO; hard emissions are treated as in NLO computations, while soft/collinear emissions are handled by the parton shower, with the same logarithmic accuracy as the MC; and the matching between the hard- and soft/collinear-emission regions is smooth.

Multi-purpose Generators

- **Herwig**
  HERWIG [82] is a general-purpose Monte Carlo event generator, which includes the simulation of hard lepton-lepton, lepton-hadron and hadron-hadron scattering and soft hadron-hadron collisions. It uses the parton-shower approach for initial- and final-state QCD radiation, including color coherence effects and azimuthal correlations both within and between jets. Another feature is its cluster model for jet hadronization based on non-perturbative gluon splitting, and a similar cluster model for soft and underlying hadronic events.

- **Pythia**
  PYTHIA [83] implements the full chain needed to model high-energy physics in pp, ee and ep collisions. It implements the Lund string model [84] to handle the hadronization process. A string represents the color-flux tube, connecting a q and a $\bar{q}$ due to the confinement of color-charged particles. The energy stored in a string increases linearly with the separation of the quarks. If the invariant mass of the string is large enough, the string energy can be converted to further $q\bar{q}$ pairs.

- **Sherpa**
  SHERPA [85] is a Monte Carlo event generator that provides complete hadronic final states in simulations of high-energy particle collisions. It contains a very flexible tree-level matrix-element generator for the calculation of hard scattering processes. The emission of additional QCD partons off the initial and final states is described through a parton-shower model. Multi-parton matrix elements are consistently combined with the QCD parton cascades. A simple model of multiple interactions is used to account for underlying events in hadron–hadron collisions. The fragmentation of partons into primary hadrons is described using a phenomenological cluster-hadronisation model.

Particle Decays

- **TAUOLA**
  TAUOLA [86] supplements events generated by other programs, with τ decays that can be well separated from their production process, because of their narrow width. Effects of spin, including transverse degrees of freedom, genuine weak corrections or effects of new physics may be taken into account at the time when a τ decay is generated and written into an event record.

3.2. Detector Simulation

After the generation of MC events, the simulation of the detector and physics interactions, and the digitization of the energy deposited in the sensitive regions of the detector into voltages and currents is performed. A detailed simulation has been implemented in order to study the detector response for a wide range of physics processes and scenarios. The standard detector simulation [87] of ATLAS relies on the GEANT4 [88] particle simulation toolkit, that provides models for physics and infrastructure for particle transportation through a geometry, including electromagnetic and hadronic interactions of particles with matter within a magnetic field. Additional ATLAS-specific pieces are provided as user-code.

The ATLAS detector geometry used for simulation, digitization, and reconstruction is built from two databases. The first one contains essentially fixed information describing the physical construction, the sensitive parts of the detector as well as the support structures and cabling material. The second one contains time dependent conditions data describing all the information needed to emulate a single data-taking run of the real detector (e.g. status of detector components, temperatures, and so on). The output of the simulation is a format identical to the output of the ATLAS data acquisition system. Thus, both the simulated and real data from the detector can be run through the same ATLAS trigger and reconstruction packages.

3.3. Event Reconstruction

In the reconstruction step the raw data of the detector readout or the detector simulation is converted to reconstructed physics objects by combining the information of the various detector components, to form objects like charged particle tracks and energy clusters. The resulting high-level objects, that fulfill certain quality criteria, are classified as electron, muon or jet candidates and provided for physics analysis.

3.3.1. Tracking

The tracks of charged particles are reconstructed by combining distinct positions measured by the inner detectors. The raw data of the pixel and SCT detectors are converted into space-points. Track seeds are formed by pattern recognition algorithms that combine the space-points of the three pixel layers and the first SCT layer. These seeds are then extended through the SCT applying a Kalman filter [89] formalism that follows the trajectory and includes successive hits on the track candidate in the fit. The found track candidates are refitted to resolve ambiguities where track

candidates share hits, are incomplete or describe fake tracks. Quality cuts are applied and the tracks are extended to the TRT and refitted again with the full information of all three detectors.

To improve the tracking efficiency for secondary tracks, that arise from conversions or decays of long-lived particles, a complementary track fitting strategy, called *"back-tracking"* , searches for unused track segments in the TRT and extends them into the SCT and the pixel detectors. References [90, 91] provide a detailed description of track reconstruction algorithms implemented in the ATLAS software.

3.3.2. Vertex Reconstruction

Vertices are reconstructed by associating reconstructed tracks to particular vertex candidates and performing a fit to determine the actual vertex position and its covariance matrix, that represents the quality of the fit. ATLAS implements two approaches. The *"fitting-after-finding"* approach clusters tracks according to their z-impact parameter position. The track clusters are then fitted and iteratively cleaned from outliers. In this approach the number of vertices is completely determined at the seeding stage and once a track is rejected from a given vertex candidate it is never used in any of the other clusters.

The *"finding-trough-fitting"* approach provides a better treatment of outlying tracks. After a preselection of tracks a single vertex seed is created out of this set by fitting the tracks. Tracks considered to be outliers during the first fit are used to create a new vertex seed. The two vertex candidates are fitted simultaneously in the next iteration. The number of vertices increases in each iteration and the vertices are competing with each other to attain more tracks. The vertex finding and fitting procedures are discussed in References [92, 93].

3.3.3. Reconstruction of Electron Candidates

The reconstruction of electron candidates begins with a seed cluster of $E_{\mathrm{T}} > 2.5$ GeV in the second layer of the electromagnetic calorimeter. A matching track, extrapolated to the second EM calorimeter layer, is searched for amongst all reconstructed tracks with $p_{\mathrm{T}} > 0.5$ GeV/c. The closest-matched track to the center of the cluster is kept as that belonging to the electron candidate. The transverse energy of these electron candidates is obtained from the corresponding calorimeter clusters.

The electron identification selections are based on criteria using calorimeter and tracker information and have been optimized in 10 bins in η and 11 bins in E_{T}. Three reference sets of requirements *"loose"*, *"medium"*, and *"tight"* provide progressively stronger jet rejection at the expense of some identification efficiency loss. Each set adds additional constraints to the previous requirements:

"Loose" is a basic selection that uses electromagnetic shower shape information from the second layer of the electromagnetic calorimeter within the detector acceptance $|\eta| < 2.47$:

- the ratio of the energy in 3×7 cells over the energy in 7×7 cells centered at the electron cluster position,
- lateral width of the shower, and
- energy leakage into the hadronic calorimeters

as discriminant variables. This set of requirements provides high and uniform identification efficiency but a low background rejection.

The *"medium"* selection provides additional rejection against hadrons by evaluating the energy deposit patterns in the first layer of the EM calorimeter:

- the difference between the energy associated with the second largest energy deposit and the energy associated with the minimal value between the first and second maxima,
- second largest energy deposit normalized to the cluster energy,
- total shower width,
- shower width for three strips around the maximum strip,
- fraction of energy outside the core of three central strips but within seven strips.

Strip-based cuts are effective in the rejection of $\pi^0 \to \gamma\gamma$ decays, since the energy-deposit pattern from π^0's is often found to have two maxima. Further cuts on track quality variables are also applied:

- the number of hits in the pixel detector ≥ 1,
- the number of hits in the pixel and the semiconductor tracker ≥ 9,
- transverse impact parameter of the track < 1 mm,
- and a cut on the cluster track matching variable, that is $\Delta\eta$ between the cluster and the track extrapolated to the first layer of the electromagnetic calorimeter.

The *"medium"* cuts increase the jet rejection by a factor of three to four with respect to the loose cuts, while reducing the identification efficiency at a level of ten percent.

The *"tight"* selection further rejects charged hadrons and secondary electrons from conversions by fully exploiting the electron identification potential of the ATLAS detector. It makes requirements on:

- the ratio of cluster energy to track momentum,

- and the number of hits in the transition radiation tracker.
- Electrons from conversions are rejected by requiring at least one hit in the b-layer, the innermost layer of the pixel detector. A conversion-flagging algorithm is also used to further reduce this contribution.
- The impact-parameter requirement applied in the medium selection is further tightened at this level, as well as
- the matching criteria between the track and the cluster.
- The imposed isolation criteria demands a cut on the ratio of transverse energy in a cone of $\Delta R < 0.2$, excluding the cluster associated with the electron, to the total cluster transverse energy.

Further information on the reconstruction of electron candidates and the ATLAS standard electron reconstruction and identification algorithm is provided in References [94, 95]. Detailed measurements of the electron performance of the ATLAS detector using decays of the Z, W and J/ψ particles, as well as the determination of the electron energy scale and resolution are reported in Reference [96].

3.3.4. Reconstruction of Muon Candidates

The ATLAS muon identification and reconstruction algorithms take advantage of multiple subdetector technologies which provide complementary approaches [91]. The *"standalone muon"* reconstruction is based entirely on muon spectrometer information, independently of whether or not the muon spectrometer track is also reconstructed in the inner detector. The muon reconstruction is initiated locally in a muon chamber by a search for straight line track segments. Hits in the precision chambers are used and the segment candidates are required to point to the center of ATLAS. Two or more track segments in different muon stations are combined to form a muon track candidate using three-dimensional tracking in the magnetic field. The track candidates are extrapolated to the interaction point taking into account both multiple scattering and energy loss in the calorimeters. The standalone muon reconstruction algorithms use the least-squares formalism to fit tracks in the muon spectrometer, with most material effects directly integrated into the χ^2 function.

"Standalone muons" have a slightly larger pseudorapidity coverage, up to 2.7 for the muon spectrometer compared to 2.5 for the inner detector. But muons produced in the calorimeter, e.g. from $\pi^{\pm}$ and K decays, are likely to be found in the standalone reconstruction and serve as a background of *"fake"* muons for this physics analysis.

The *"combined muon"* reconstruction associates a standalone muon spectrometer track to an inner detector track. The association is performed using a χ^2-test, defined from the difference between the respective track parameters $\vec{T}$ weighted by their combined covariance matrices C.

$$\chi^2_{\text{match}} = (\vec{T}_{\text{MS}} - \vec{T}_{\text{ID}})^{\text{T}}(C_{\text{ID}} + C_{\text{MS}})^{-1}(\vec{T}_{\text{MS}} - \vec{T}_{\text{ID}})$$

The subscript ID refers to the inner detector and MS to the muon spectrometer. The combined track parameters are derived either from a statistical combination of the two tracks *"staco muons"*:

$$T = (C_{\text{ID}}^{-1} + C_{\text{MS}}^{-1})^{-1}(C_{\text{ID}}^{-1}\vec{T}_{\text{ID}} + C_{\text{MS}}^{-1}\vec{T}_{\text{MS}})$$

or from a refit of the full track adding the inner track vector and covariance matrix to the measurements of the outer track, in the case of *"muid muons"*.

The performance of the ATLAS muon reconstruction and identification, and the measurement of the muon reconstruction efficiency are presented in References [97] and [98], respectively.

3.3.5. Reconstruction of Jet Candidates

The ATLAS jet candidate reconstruction provides several cone and clustering jet algorithms based on signal towers or topological clusters that are constructed from the measurements in the calorimeters. The FastJet package [99] is used for the most jet finder implementations. The anti-k_{T} algorithm [100, 101] based on topological clusters [102] provides the best performance [103].

Clustering algorithms introduce distances d_{ij} between clusters i and j and d_{iB} between cluster i and the beam B:

$$\begin{aligned} d_{ij} &= \min\left(p_{\text{T}i}^{2m}, p_{\text{T}i}^{2m}\right)\frac{(\Delta R)_{ij}^2}{R^2}\,, \\ d_{iB} &= p_{\text{T}i}^{2m}\,, \end{aligned}$$

where m is a parameter defining the specific algorithm. $(\Delta R)_{ij}^2$ denotes the distance of the clusters i and j in the $\eta - \phi$ plane and R defines the resolution at which jets are resolved from each other compared to the beam. The clustering proceeds by identifying the smallest of the distances and if it is one of the d_{ij} recombining the respective clusters i and j and replacing cluster i and j in the list of clusters by the combination. If it is d_{iB} calling i a jet and removing it from the list of clusters. The distances are recalculated and the procedure repeated until no clusters are left.

The anti-k_{T} algorithm defines parameter $m = -1$. The functionality of this algorithm can be understood by considering an event with a few well-separated hard

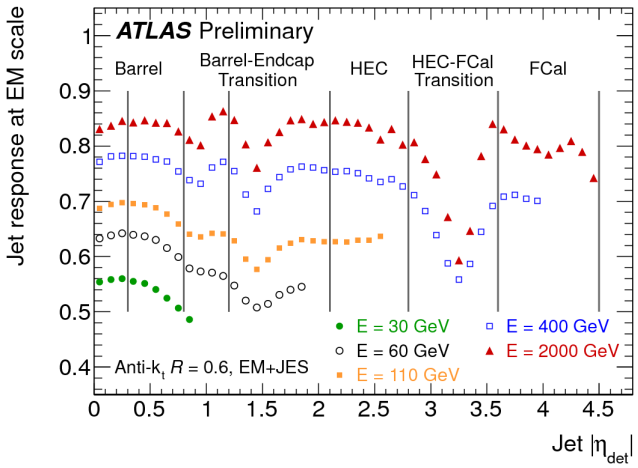


Figure 3.2.: *Simulated jet energy response at the electromagnetic scale as a function of the absolute value of the detector pseudorapidity η for different jet energies. Also shown are the η-intervals used to evaluate the jet energy scale uncertainty. The inverse of the response shown in each bin is equal to the average jet energy scale correction [104].*

particles (with large energy depositions in the clusters) with transverse momenta p_{T1}, p_{T2}, ... and many soft particles (with small energy depositions in the clusters). The $d_{1i} = \min(1/p_{T1}^2, 1/p_{Ti}^2)(\Delta R)_{ij}^2/R^2$ between a hard particle 1 and a soft particle i is exclusively determined by the transverse momentum of the hard particle and the $(\Delta R)_{ij}$ separation. The d_{ij} between similarly separated soft particles will instead be much larger. Therefore, soft particles will tend to cluster with hard ones long before they cluster among themselves.

Jet Energy Calibration

Jets are reconstructed at the electromagnetic scale, which is the basic signal scale for the ATLAS calorimeters. It accounts correctly for the energy deposited in the calorimeter by photons and electrons. The goal of the jet energy scale calibration is to correct the energy and the momentum of the jets measured in the calorimeter, using as a reference the kinematics of the corresponding Monte Carlo truth jets. The jet energy scale calibration is then validated with data using γ+jet and di-jet balancing techniques.

ATLAS has developed several calibration schemes [91, 105] with different levels of complexity and different sensitivity to systematic effects. Results illustrating the performance of these calibration schemes as obtained with a Monte Carlo simulation are shown in Reference [106]. The response of the ATLAS calorimeters to jets is studied by balancing the transverse momenta of dijets [107]. Figure 3.2 shows the

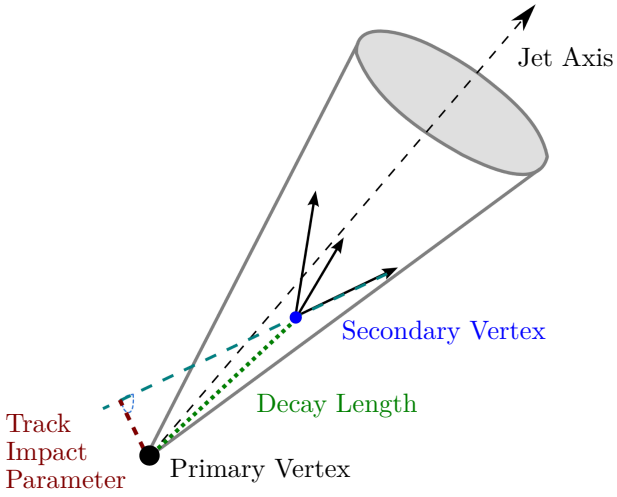


Figure 3.3.: *A secondary vertex with a significant decay length indicates the presence of a long-lived particle in the jet. The secondary vertex is reconstructed from tracks with a large impact parameter*

jet response of the ATLAS calorimeters as a function of the absolute value of the pseudorapidity η for different jet energies. The determination of the jet energy scale and its systematic uncertainty is presented in Reference [104].

This analysis deploys a p_{T}- and η-dependent calibration scheme. This simple calibration scheme corrects for the non-linear correlation between the energy reconstructed in the calorimeter and the energy of the particles forming jets.

b-jet Identification

Jets originating from b quarks are tagged by exploiting the long lifetime of b hadrons, that leads to typical flight paths of a few millimeters observable in the detector, or other properties of the decay products, such as track multiplicity. Several algorithms based on the reconstruction of secondary vertices, likelihood ratios, neural networks or soft leptons are implemented in the ATLAS software.

This analysis vetoes against events containing b-jets and relies on the SV0 algorithm. The SV0 algorithm explicitly reconstructs secondary vertices from tracks, which are associated to a calorimeter jet and fulfill certain quality criteria, e.g. large impact parameter significance, see figure 3.3. Secondary vertices are characterized by the decay length and the decay length significance, defined as the decay length divided by its error. A detailed description of the SV0 algorithm, its performance and calibration can be found in References [108, 109].

3.3.6. Missing Transverse Energy

As the colliding particles have only longitudinal momentum, the momentum perpendicular to the beam axis should compensate due to momentum conservation, unless weakly interacting particles like neutrinos, that leave the detector material without interactions, are present it the event. In these cases the sum of transverse energies measured in the detector is not zero. The missing transverse energy $\not{E}_{\mathrm{T}}$ is defined as:

$$\not{E}_{\mathrm{T}} = \sqrt{(E_x^{\mathrm{miss}})^2 + \left(E_y^{\mathrm{miss}}\right)^2}$$

with

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss,calo}} + E_{x(y)}^{\mathrm{miss},\mu}$$

The calorimeter term is defined as:

$$E_x^{\mathrm{miss,calo}} = -\sum_i^{N_{\mathrm{cell}}} E_i \sin\theta_i \cos\phi_i$$

and

$$E_y^{\mathrm{miss,calo}} = -\sum_i^{N_{\mathrm{cell}}} E_i \sin\theta_i \sin\phi_i$$

where E_i , θ_i and ϕ_i are the energy, the polar angle and the azimuthal angle, respectively, of calorimeter cells i in the pseudorapidity range $|\eta| < 4.5$. Noise contributions are suppressed by limiting the number of cells, N_{cell}, used in the $\not{E}_{\mathrm{T}}$ calculation. This is achieved by using only cells belonging to three-dimensional topological clusters [102] that are seeded by cells with a threshold on the deposited energy. Additional calibration procedures [105, 110] are applied in order to take into account effects from the detector response and the dead material in front of and between the calorimeters.

The muon term is calculated from the momenta of muon tracks reconstructed with $|\eta| < 2.7$:

$$E_{x(y)}^{\mathrm{miss},\mu} = \sum_{\mathrm{muons}} p_{x(y)}$$

Contributions from fake muons, that may arise from high energy jets punching through the calorimeter into the muon system, are reduced by matching the muon track reconstructed in the muon spectrometer with a track reconstructed in the inner detector. The energy deposited by the muon in calorimeters is calculated differently for isolated and non-isolated muons:

- The p_{T} of an isolated muon is determined from the combined measurement of the inner detector and muon spectrometer. In this case the energy lost by the muon in the calorimeters is not added to the calorimeter term to avoid double counting.

- For a non-isolated muon, the energy lost in the calorimeter cannot be separated from the nearby jet energy. The muon spectrometer measurement of the muon momentum after energy loss in the calorimeter is therefore used.

Further contributions due to electron, photon, tau, and jet terms are also included in the calibration of $\not{E}_{\mathrm{T}}$, more details on the calibration can be found in Reference [105, 110].

4. Event Modeling

This analysis relies on Monte Carlo simulations of the considered signal and background processes. Thus, the modeling of these processes is crucial and the simulated samples have to describe the observed collision data correctly in the whole relevant kinematic phase space, including the correlations between the observables, as a multivariate technique is deployed to separate the signal and background processes.

This chapter gives an overview of the considered physics processes, their cross-sections and modeling. The signal processes GGF ($gg \to H \to \ell\nu\ell\nu$) and VBF ($qq \to qqH \to qq\ell\nu\ell\nu$) are discussed in more details. The considered final state of GGF comprises zero or one additional quark that may arise due to initial state radiation. Tests of the modeling in the signal region and in dedicated background control-regions are shown in the following chapters. The properties of the Higgs boson that manifest in the kinematics of its decay products are discussed as well as an approach to reconstruct the Higgs boson using these kinematic features.

4.1. Background Event Modeling

An overview of the background processes, their cross-sections and the used MC generators is given in Table 4.1. The background processes have

- either the same final state as the signal processes, $\ell\nu\ell\nu$+nq, with $n = 0\,..\,2$, observed as $\ell\ell$+$\not{E}_{\mathrm{T}}$+$n$jets,
- or a similar final state and a large production cross-section, which increases their probability to mimic the signal due to mis-reconstruction effects or fake lepton candidates caused by jets.

4.1.1. Vector Boson Production in Association with Quarks

Typical Feynman diagrams for the production of a W or Z boson in association with quarks are shown in Figure 4.1. In the case of W boson production with a subsequent leptonic decay, the resulting final state is $\ell\nu$+quarks, which is observed as a charged lepton, $\not{E}_{\mathrm{T}}$ and jets. This final state has one charged lepton less than the final states of the signal processes. Nevertheless, one of the jets may fake a lepton

Table 4.1.: *Cross-sections at the center-of-mass energy of* 7 TeV *for background processes. The* $W \to \ell\nu$ *and the* $Z/\gamma^* \to \ell\ell$ *cross-sections are single flavor cross-sections.*

Process	Generator	cross-section σ (pb) [× BR]
Inclusive $W \to \ell\nu$	ALPGEN	10.5×10^3 [111, 112]
Inclusive $W \to \tau\nu$	PYTHIA	10.5×10^3 [111, 112]
Inclusive $Z/\gamma^* \to \ell\ell$ ($10 < M_{\ell\ell} < 40$ GeV/c^2)	ALPGEN	3.9×10^3 [113]
Inclusive $Z/\gamma^* \to \ell\ell$ ($M_{\ell\ell} > 40$ GeV/c^2)	ALPGEN	10.7×10^2 [112, 113]
Inclusive $Z/\gamma^* \to \tau\tau$ ($10 < M_{\tau\tau} < 60$ GeV/c^2)	PYTHIA,TAUOLA	4.0×10^3 [113]
Inclusive $Z/\gamma^* \to \tau\tau$ ($M_{\tau\tau} > 60$ GeV/c^2)	PYTHIA,TAUOLA	9.9×10^2 [112, 113]
$t\bar{t}$	MC@NLO	163 [114]
Single top Wt	MC@NLO	15.7 [115]
Single top s-channel	MC@NLO	4.6 [116]
Single top t-channel	MC@NLO	64.6 [117]
WZ	MC@NLO	18.0 [118]
ZZ	MC@NLO	5.6 [118]
$qq/qg \to WW \to \ell\nu\ell\nu(\ell = e, \mu, \tau)$	MC@NLO	4.7 [118]
$gg \to WW \to \ell\nu\ell\nu(\ell = e, \mu, \tau)$	GG2WWW	0.14 [118]
$\gamma W \to \ell\nu(\ell = e, \mu, \tau)$	MADGRAPH, PYTHIA	135.4 [119]

leading to the observed signature of the signal $\ell\ell$+$\not\!\!E_{\mathrm{T}}$+$n$jets. In the case of Z boson production with subsequent leptonic decay, Z+jets $\to \ell\ell$+jets, a not reconstructed jet or mis-measurement in the calorimeters may cause missing transverse energy and thus, mimic the final state of the signal process.

The W+jets and Z+jets processes are modeled using ALPGEN. The parton shower and the hadronization is handled by HERWIG. MC samples with up to 5 final-state parton multiplicities are used and the MLM matching scheme [75] is applied to combine different samples.

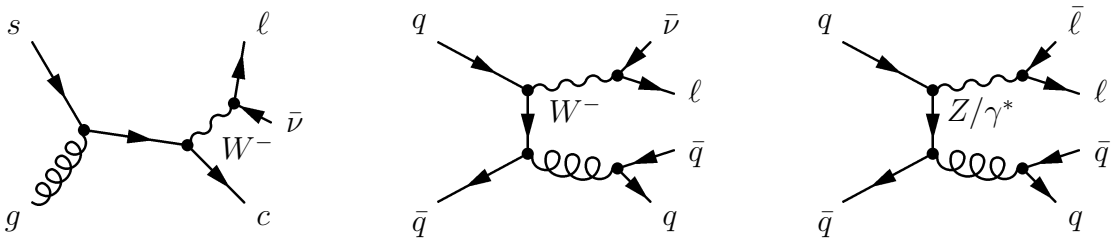


Figure 4.1.: *Feynman diagrams for* W *and* Z *boson production in association with quarks.*

4.1.2. Diboson Production

MC@NLO is used to model events containing two bosons WW, WZ and ZZ. Figure 4.2 displays lowest order Feynman diagrams for diboson production at hadron colliders. WW events where both bosons decay leptonically result in exactly the same observed final state as the GGF signal process. WZ events, with $W \to q\bar{q}'$ and

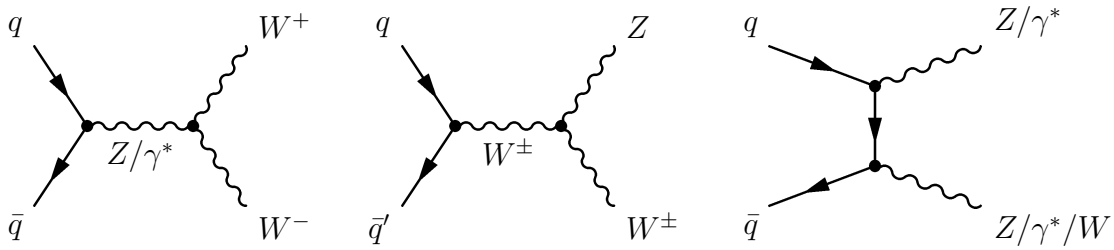


Figure 4.2.: *Lowest order Feynman diagrams for diboson production at hadron colliders.*

$Z \to \ell^+\ell^-$, may mimic the final state of VBF due to a mis-measurement of $\not{E}_{\mathrm{T}}$. This also applies tho $ZZ \to q\bar{q}\ell^+\ell^-$ events.

In $W + \gamma$ events, a hadronic decay of the W boson and $\gamma \to \ell^+\ell^-$ may result in the signal signature. $W + \gamma$ events are modeled with MadGraph calculating the matrix element and Pythia handling the hadronization.

Additional contributions to the continuum WW background is induced by gluons, where a pure QCD initial state couples through a quark loop to a pure electroweak final state, see Figure 4.3(a) and (b). This background is modeled using gg2WW with Herwig hadronization routines. The single-resonant production Figure 4.3(c) with the same final state $\ell\nu\ell\nu$ is also included.

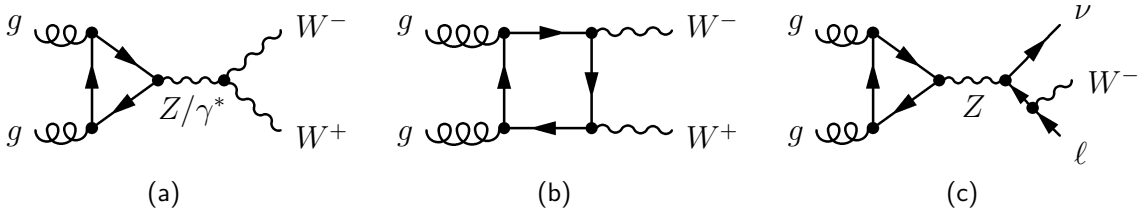


Figure 4.3.: *Gluon-induced W-boson pair production through triangle graph (a) and through a box diagram (b). The single-resonant production (c) is also included. In these production modes a pure QCD initial state couples through a quark loop to a pure electroweak final state.*

4.1.3. Top Quark Production

Top-quark processes including top-quark pair production, single top-quark production, and associated single top-quark production, represented by the Feynman diagrams in Figure 4.4 are modeled using MC@NLO. Due to its huge mass the top quark decays to a W boson and a b quark before it can hadronize. The decay to lighter quarks is suppressed by small CKM matrix elements. Subsequent leptonic decays of the W bosons in $t\bar{t}$ events Figure 4.4(a) leads to the same final state

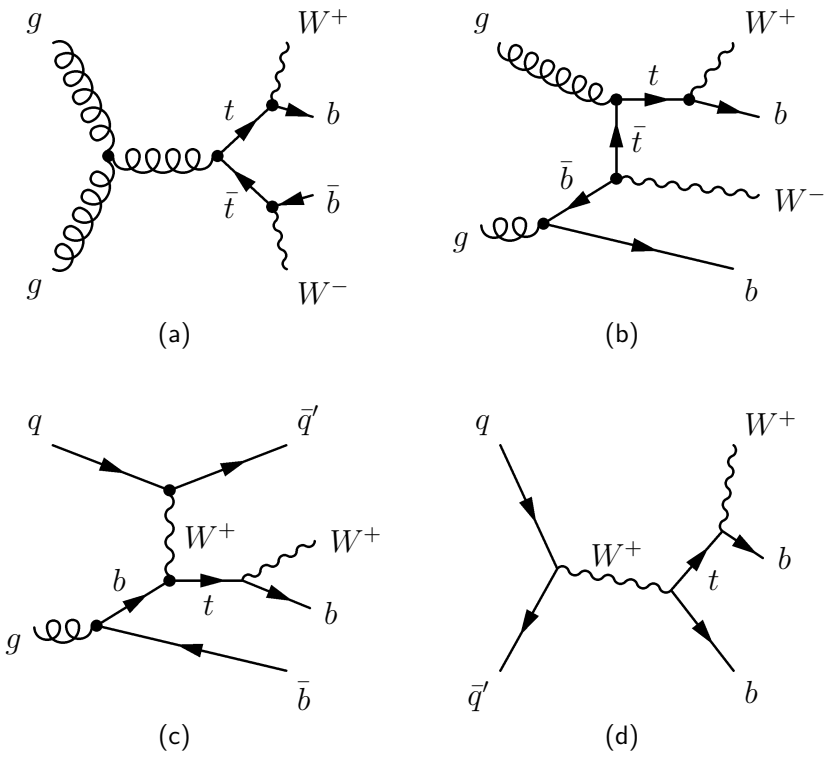


Figure 4.4.: *Top quark Feynman diagrams: top-quark pair production (a), associated single top-quark production (b), and t and s-channel single top-quark production (c) and (d), respectively.*

configuration $\ell\ell$+$\not{E}_{\mathrm{T}}$+2jets as the VBF process, except for the presence of two b quarks. This is also the case for the associated production of a single top quark with a W boson Figure 4.4(b). While the t- and s-channel single top-quark production Figure 4.4(c) and (d) may contribute by jets faking leptons or leptonic b quark decays.

Especially Wt and t-channel may mimic the VBF signature due to the forward b quark from initial-state gluon-splitting, and due to the light quark in the case of the t-channel. As no tracking and thus no b-tagging information is available in the forward region, events with a forward b-quark may provide two jets with a large pseudorapidity gap and thus contribute to the same observed signature as the VBF process.

4.2. Signal Event Modeling

The Higgs boson production cross-section in pp collisions, at a center-of-mass energy of 7 TeV and the Higgs boson branching ratios are studied extensively in Reference [120].

Table 4.2 summarizes the cross-section for the GGF and the VBF processes and the branching ratios for the subsequent decay process $H \to WW$ for all Higgs boson masses considered in this analysis.

4.2.1. Gluon-Gluon Fusion Process

The GGF process presented in Figure 4.5 is the main production mechanism for the SM Higgs boson at hadron colliders and the most important one for searches over the entire mass range, 100 GeV/c^2 $< M_H <$ 1 TeV/c^2, reachable at the LHC. The dynamics of this process is controlled by the strong interaction, producing the Higgs boson via a heavy-quark loop. The main contribution to the production cross-section stems from top quarks caused by the large Yukawa coupling. An overview of the cross-section calculations including QCD radiative corrections at next-to-leading order and next-to-next-to-leading order and two-loop electroweak effects is given in Reference [120]. The GGF process is modeled using MC@NLO and Herwig hadronization routines. This analysis is designed to enhance this process in the zero and one jet bin, while the VBF signal is most important in the two jet bin.

4.2.2. Vector Boson Fusion Process

In the VBF process the Higgs boson is produced in association with two hard jets in the forward and backward regions of the detector. It has the second largest production cross-section at the LHC, and is also a corner stone for the Higgs boson search in the entire mass range. This production mode plays an important role in the determination of the Higgs boson couplings to weak gauge bosons. The production cross-section is proportional to the squared coupling at the HV^*V^* vertices, and thus the observation of this production mode would prove directly the existence of such a tree level coupling and allow to establish the SM Higgs boson as the agent of electroweak symmetry breaking, as Yukawa couplings are not involved.

The genuine VBF production, where the Higgs boson couples to two weak gauge bosons each of them linking to a quark line, is dominated by the t-channel Feynman diagram which is shown in Figure 4.6(a)The scattered quarks have the tendency to produce jets in forward-backward direction in contrast to other jet-production mechanisms. Due to the pure electroweak process there is no color exchange between the quarks and the production of additional jets in the central region of the detector is suppressed. This feature provides a rapidity gap between the two jets in the final state offering a handle to suppress background processes [121, 122].

Table 4.2.: *Higgs boson production cross-sections of the GGF ($gg \to H$) and VBF ($qq \to qqH$) processes along with the $H \to WW^{(*)}$ decay branching ratios from Reference [120].*

m_H [GeV/c^2]	$\sigma(gg \to H)$ [pb]	$\sigma(qq \to qqH)$ [pb]	$Br(H \to WW^{(*)})$
120	16.63	1.270	0.143
125	15.31	1.211	0.216
130	14.12	1.154	0.305
135	13.08	1.099	0.403
140	12.13	1.053	0.504
145	11.27	1.004	0.603
150	10.50	0.9616	0.699
155	9.795	0.9180	0.796
160	9.080	0.8788	0.909
165	8.319	0.8517	0.960
170	7.729	0.8173	0.965
175	7.211	0.7814	0.958
180	6.739	0.7480	0.932
185	6.295	0.7194	0.844
190	5.896	0.6925	0.786
195	5.551	0.6643	0.757
200	5.249	0.6372	0.741

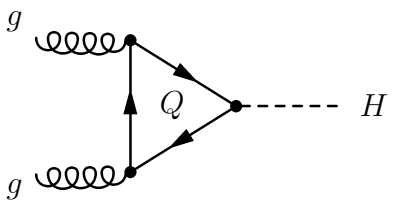


Figure 4.5.: *Feynman diagram of the gluon-gluon fusion process: the Higgs boson is produced via a heavy-quark loop.*

Figure 4.7 depicts the rapidity gap on parton level, as modeled by SHERPA using the following configuration:

$$qq \to qqH \quad \text{with the subsequent decay} \quad H \to WW \to \ell\nu\ell\nu\,,$$

permitting only electroweak couplings. The resulting MC events contain not only the t-channel but also the s-channel as presented in Figure 4.6(b), where the Higgs boson is emitted by a vector boson V^* that subsequently decays into two quarks. The η distribution of the quarks is depicted in Figure 4.7(a), while Figure 4.7(b) presents the $|\Delta\eta|$ distribution between the two quarks, visualizing the characteristic rapidity gap between the scattered quarks. The contributions of the t and s-channel to the distributions in Figure 4.7 are obtained by applying a cut on the invariant mass of the two quarks in the final state. A cut value of 110 GeV/c^2 was chosen for this purpose. The invariant mass distributions M_{qq} below and above this cut value are presented in Figure 4.8(a) and (b), respectively. The W and Z boson mass-peaks are characteristic for the s-channel Figure 4.8(a). The M_{qq} distribution of the t-channels has its maximum around 150 GeV/c^2 and a long tail to larger values Figure 4.8(b).

The most characteristic observables of the VBF process are the $|\Delta\eta|$ and the M_{qq} distributions. They provide a good discrimination and play an important role in separating background processes as will be discussed later.

4.2.3. Higgs Boson Decay

The decay chain $H \to WW^* \to \ell\nu\ell\nu$ where the Higgs boson decays into a W^+W^- pair with the subsequent leptonic decay of both W bosons, as depicted in Figure 4.9, is considered in this analysis. For Higgs boson masses lower than 160 GeV/c^2 only one of the W bosons is produced on-shell.

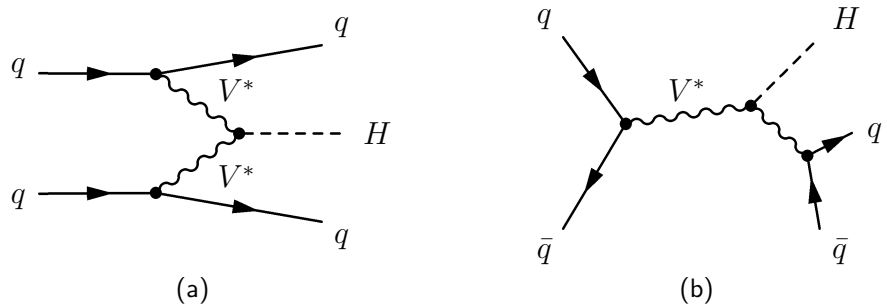


Figure 4.6.: *Higgs boson production diagrams (a) t-channel and (b) s-channel or Higgs strahlung.*

This decay chain exhibits a characteristic topology due to momentum and angular momentum conservation. Figure 4.10 visualizes this feature in the Higgs boson rest frame, where the W bosons are emitted back to back. The W bosons are spin 1 particles, while the Higgs boson is a spin zero particle. Thus, the spins of the W bosons have to point in opposite directions. The produced W bosons are both either left-handed as shown in Figure 4.10, the black arrows represent the momentum and the red arrows the spin of the particles. Or both W bosons are right-handed, if the spin points in the same direction as their momentum vector. The V-A structure of the electroweak theory becomes apparent in the decay of the W bosons, as they only interact with left-handed fermions and right-handed anti fermions. In the decay of the W^- boson the charged lepton has to be left-handed and the corresponding neutrino right-handed, and vice versa for the W^+ boson. As a consequence, the charged leptons are emitted in the same direction in the rest frame of the Higgs boson. A boost of this system, caused by the momentum of the Higgs boson does not change this topology.

The described spin configuration represents a well-defined state only for massless particles. Nevertheless, the lepton masses can be neglected compared to the large masses of the W bosons.

The implication of the decay topology is a small polar angle θ between the charged leptons leading to the characteristic $\cos\theta_{\ell\ell}$ distribution depicted in Figure 4.11(a). A further characteristic of the considered decay chain is the large amount of missing transverse energy caused by the neutrinos in the final state. The transverse energy distribution of the neutrino pair is shown in Figure 4.11(b).

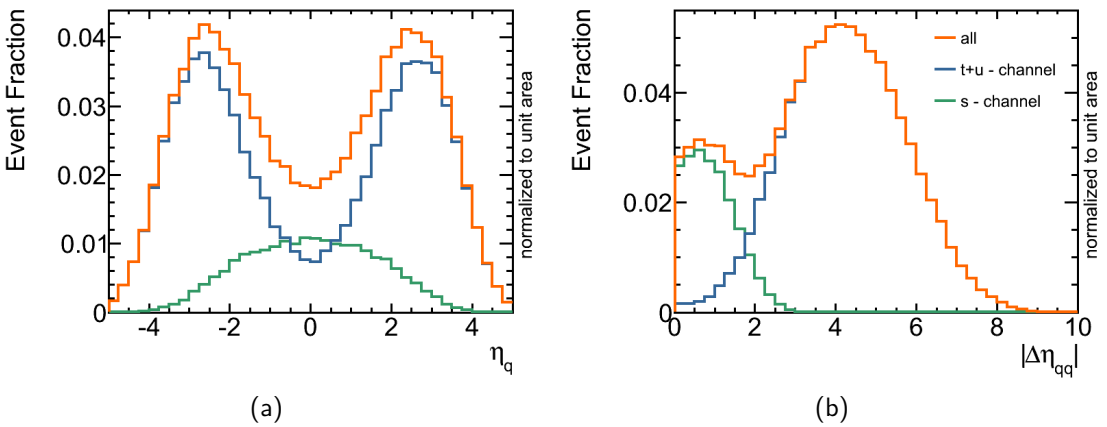


Figure 4.7.: *η (a) and $|\Delta\eta|$ (b) distributions of the scattered quarks in the VBF process.*

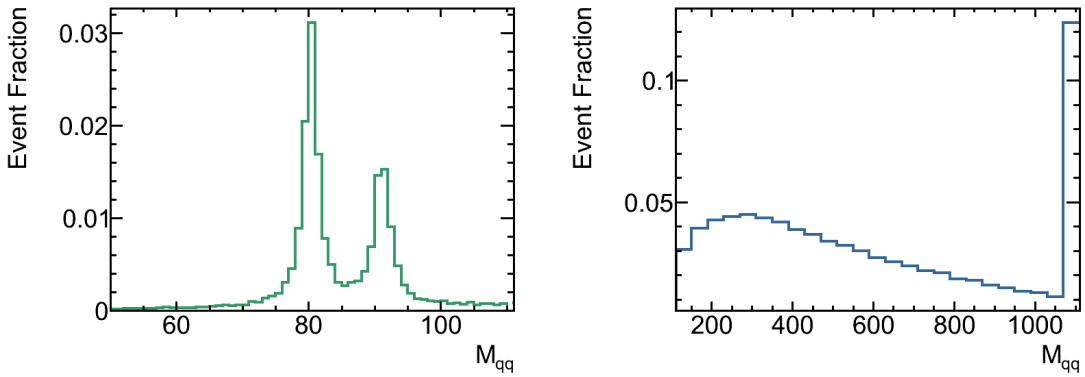


Figure 4.8.: *Invariant mass distribution of VBF: (a) below* 110 GeV/c^2*, depicting the* W *and* Z *boson mass-peaks which are characteristic for the s-channel, and (b) above* 110 GeV/c^2 *depicting the* M_{qq} *distribution characteristic for the t-channel.*

4.2.4. Reconstruction of VBF Higgs Events

The considered final state comprises two neutrinos that escape the detector without being measured. Hence, this final state cannot be fully reconstructed. However there is a special case for Higgs boson masses around 160 GeV/c^2, the W bosons are produced on-shell and their momentum in the Higgs-boson rest-frame is zero. This provides the possibility to completely reconstruct the kinematics of the final state neutrinos and thus, the momentum of the W bosons and of the initial Higgs boson, as proposed in Reference [123].

Neglecting the masses of quarks, charged leptons and neutrinos, the unknown variables are the six momentum components of the neutrinos $\vec{p}^{\,\nu}, \vec{p}^{\,\bar{\nu}}$. Disregarding initial-state radiation, the transverse momentum of the initial state is approximately zero, which means that the sum of transverse momenta in the final state is also zero

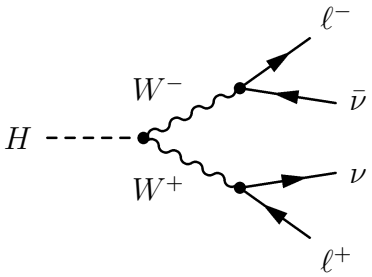


Figure 4.9.: *Higgs boson decay chain* $H \to WW^* \to \ell\nu\ell\nu$ *where the Higgs boson decays into a* W^+W^- *pair with the subsequent leptonic decay of both* W *bosons.*

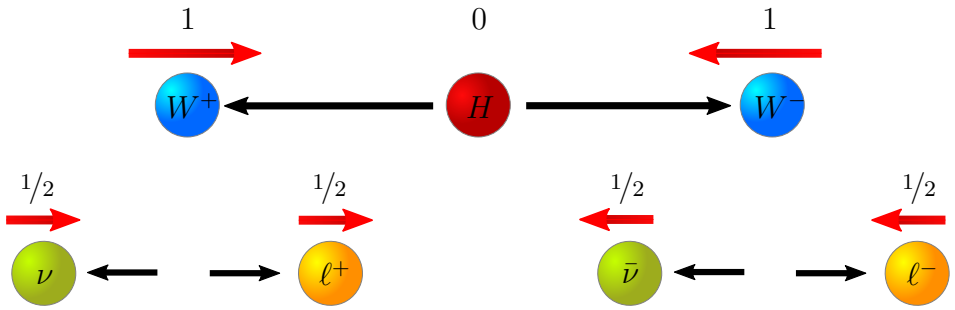


Figure 4.10.: *Visualization of the Higgs boson decay in its rest frame. The black arrows represent the momentum of the particles and the red arrows their spin. The charged leptons are emitted in the same direction due to the V-A structure of the electroweak theory.*

leading to the following equations

$$p_x^{\nu} + p_x^{\bar{\nu}} = -(p_x^{q1} + p_x^{q2} + p_x^{\ell^-} + p_x^{\ell^+})$$
$$p_y^{\nu} + p_y^{\bar{\nu}} = -(p_y^{q1} + p_y^{q2} + p_y^{\ell^-} + p_y^{\ell^+})$$

providing two constraints. Additional three constraints stem from the assumption that the momentum of the W bosons in the Higgs rest frame is zero. Hence, the momenta of the W bosons in the lab frame are the same.

$$\vec{p}^{W^-} = \vec{p}^{W^+}$$
$$\vec{p}^{\ell^-} + \vec{p}^{\bar{\nu}} = \vec{p}^{\ell^+} + \vec{p}^{\nu}$$

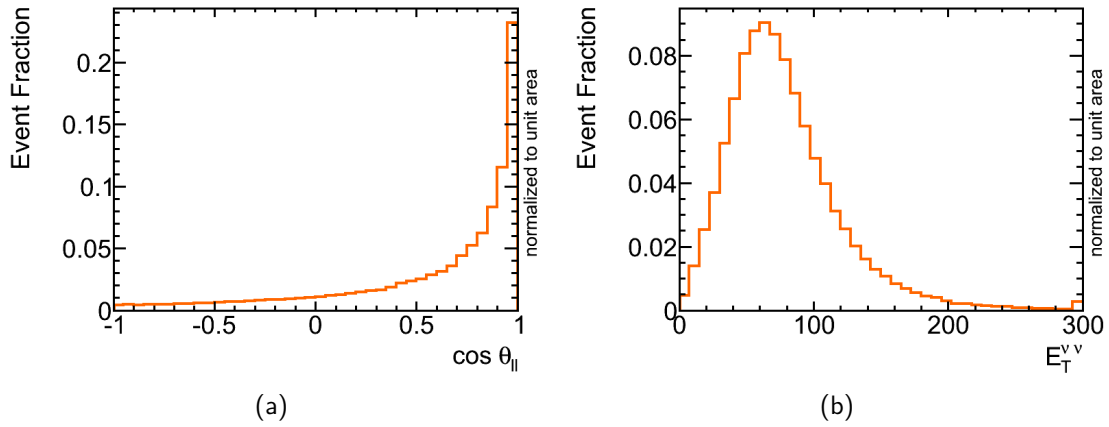


Figure 4.11.: *Characteristic distributions of the Higgs boson decay: the* $\cos\theta_{\ell\ell}$ *distribution of the polar angle* θ *between the charged leptons (a) and the transverse energy of the neutrino pair (b).*

Using the five equations above one obtains the individual transverse momentum components p_x and p_y and a relation between the longitudinal components p_z of the neutrino ant the anti neutrino.

As the charged leptons and the neutrinos are decay products of W bosons, the last constraint arises by requiring one of the two lepton-neutrino pairs to have the W boson mass, leading to a quadratic equation for the p_z component of the corresponding neutrino.

$$(p_z^\nu)^2 - 2 \cdot \frac{\mu \cdot p_z^\ell}{(E^\ell)^2 - (p_z^\ell)^2} \cdot p_z^\nu + \frac{(E^\ell)^2 \cdot (p_\mathrm{T}^\nu)^2 - \mu^2}{(E^\ell)^2 - (p_z^\ell)^2} = 0 \qquad (4.1)$$

$$\text{with} \qquad \mu = \frac{M_W^2}{2} + \cos(\Delta\phi) \cdot p_\mathrm{T}^\ell \cdot p_\mathrm{T}^\nu$$

Here p_z^ℓ and E^ℓ denote the z component of the momentum, and the energy of the charged lepton, respectively. The quantity $\Delta\phi$ is the azimuthal angle difference between the momentum of the charged lepton and the neutrino. In general, a quadratic equation leads to two solutions. The solution with the larger value is chosen, which is the better choice in about 70% of the cases.

In about 41% of all cases there is a complex solution. In these cases the measured transverse mass of the W boson, which is defined as

$$M_{\mathrm{T},W\to\ell\nu_\ell}^2 = (p_\mathrm{T}^\ell + p_\mathrm{T}^\nu)^2 - (p_x^\ell + p_x^\nu)^2 - (p_y^\ell + p_y^\nu)^2 \,,$$

is larger than $M_W = 80.4\ \mathrm{GeV/c^2}$ as demanded using the W mass constraint. Choosing just the real part of the p_z^ν solution, leads to bigger reconstructed W boson mass, since the real part does not obey equation 4.1. To avoid such complex solutions, the transverse mass $M_{\mathrm{T},W}$ is fixed to $80.4\ \mathrm{GeV/c^2}$, and a quadratic dependence $p_y^\nu(p_x^\nu)$ of p_y^ν on p_x^ν is obtained. The difference δ defined as:

$$\delta(p_x^\nu) = \sqrt{(p_x^\nu - \not{E}_{\mathrm{T},x})^2 + (p_y^\nu(p_x^\nu) - \not{E}_{\mathrm{T},y})^2} \,.$$

is minimized with respect to p_x^ν, in a range where p_y^ν does not become complex. Since there are in general two solutions for p_y^ν, two values for δ may be found. In these cases the solution which is closest to $\vec{\not{E}}_\mathrm{T}$, that means the smaller δ value, is used. The W mass constraint is discussed in more details in Reference [124].

The p_z component of the second neutrino follows from the assumption of equal momenta of the W bosons. Thus, no mass constraint is applied on the reconstruction of the second W boson, referred to as W^* in the following. The momentum of the Higgs boson is reconstructed adding up the four vectors of the two W bosons. The reconstruction method is tested using MC events of the t and u-channel.

The following figures present the performance of the Higgs boson reconstruction at the parton level and at the level of physics objects, by comparing the transverse momentum, the pseudorapidity and the invariant mass distributions of the generated and reconstructed particles in the left column. While the relative difference of the generated and the reconstructed values is shown in the right column. The distributions of the generated particles are plotted in black exploiting the MC truth information of the neutrinos. The distributions obtained by the reconstruction at the parton level are plotted in blue, using the four momenta of the quarks and charged leptons provided by the MC generator. The distributions plotted in red show the performance of the Higgs boson reconstruction on the level of physics objects, using the four momenta of jets and charged leptons after full detector simulation and object reconstruction, and applying the event selection described in the next chapter to identify the physics objects.

Figure 4.12, 4.13 and 4.14 present the distributions of the W boson reconstructed using the W-mass constraint, the off-shell W^* boson and the Higgs boson, respectively. In all three cases the reconstruction of the transverse momentum performs better than the reconstruction of the pseudorapidity and the invariant mass, which depend on the p_z component of the neutrinos. As expected, the performance on the level of physics objects is not as well as on parton level, due to the smearing of the momenta of the jets and charged leptons caused by detector simulation and object reconstruction, and due to the event selection.

Except for the p_{T} distributions of the Higgs boson Figure 4.14 all other reconstructed distributions on the level of physics objects are shifted by at least 10% compared to the generated distributions. Thus, only the transverse momentum of the Higgs boson is reconstructed properly, and may provide additional information to the analysis. The resulting resolution of the Higgs-boson mass in about 22%, as can be seen from the last plot in Figure 4.14.

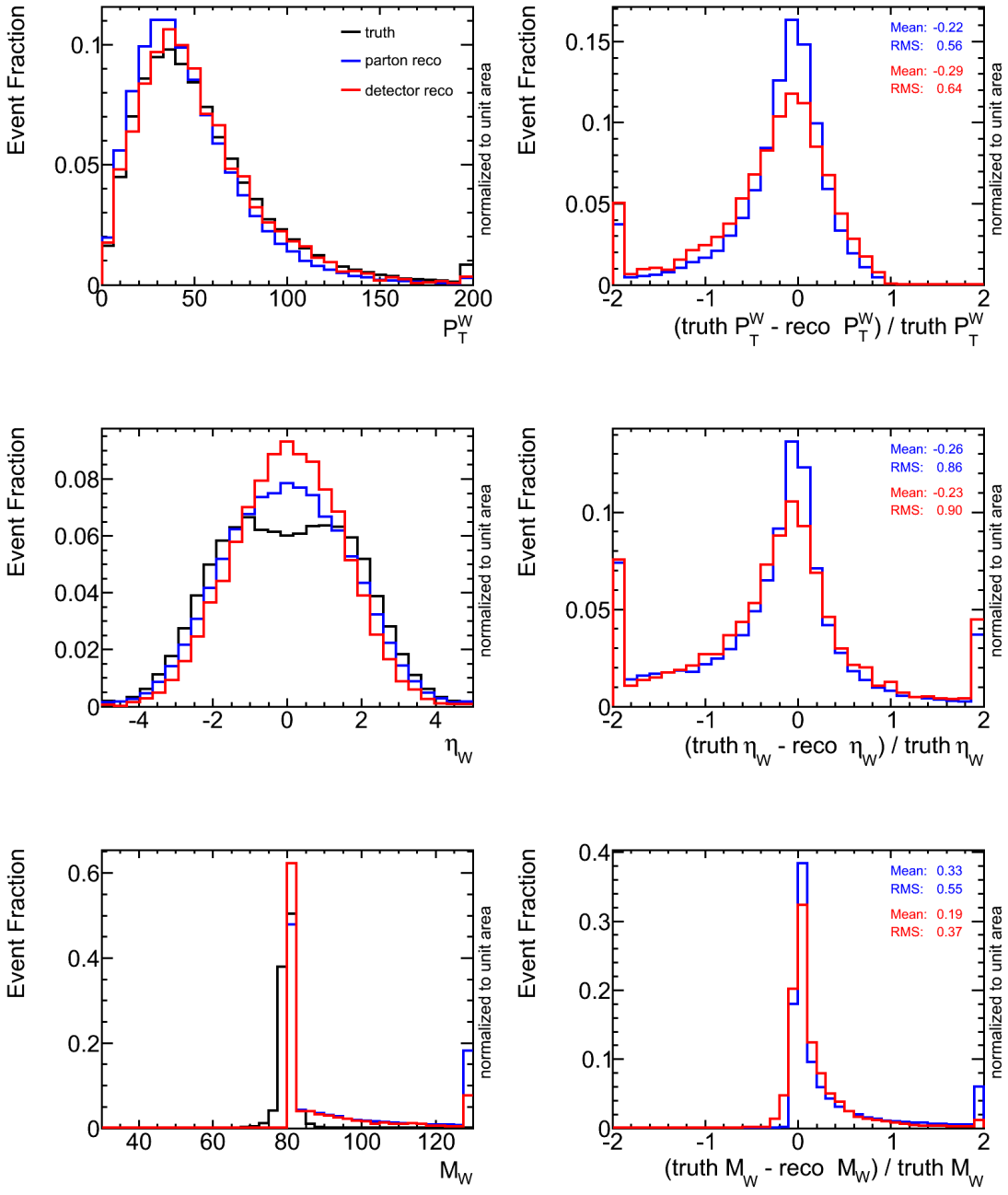


Figure 4.12.: *Properties of the generated and reconstructed W boson: the transverse momentum, the pseudorapidity and the invariant mass distributions shown in the left column and the relative difference of the generated and the reconstructed values in the right column.*

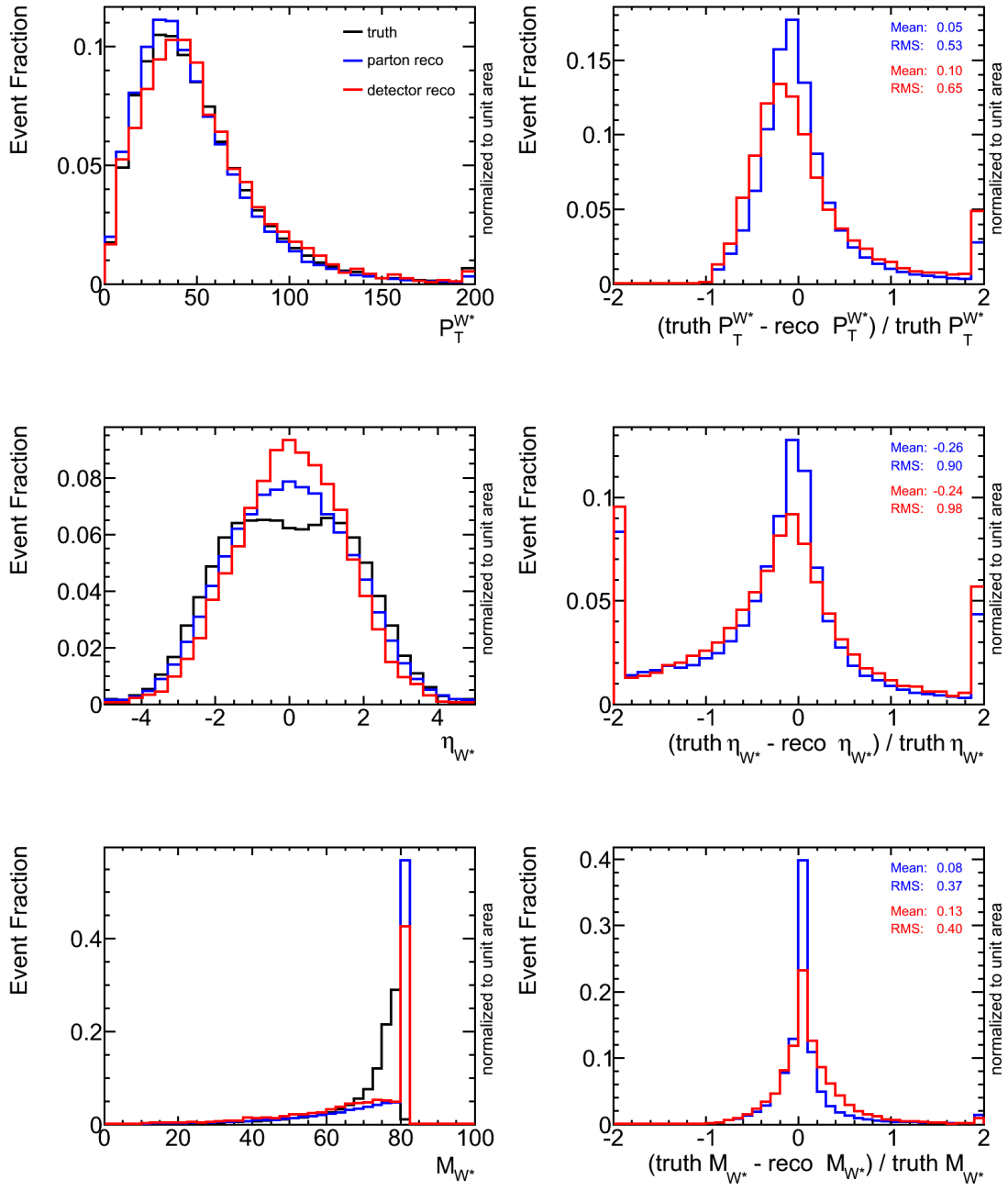


Figure 4.13.: *Properties of the generated and reconstructed W^* boson: the transverse momentum, the pseudorapidity and the invariant mass distributions shown in the left column and the relative difference of the generated and the reconstructed values in the right column.*

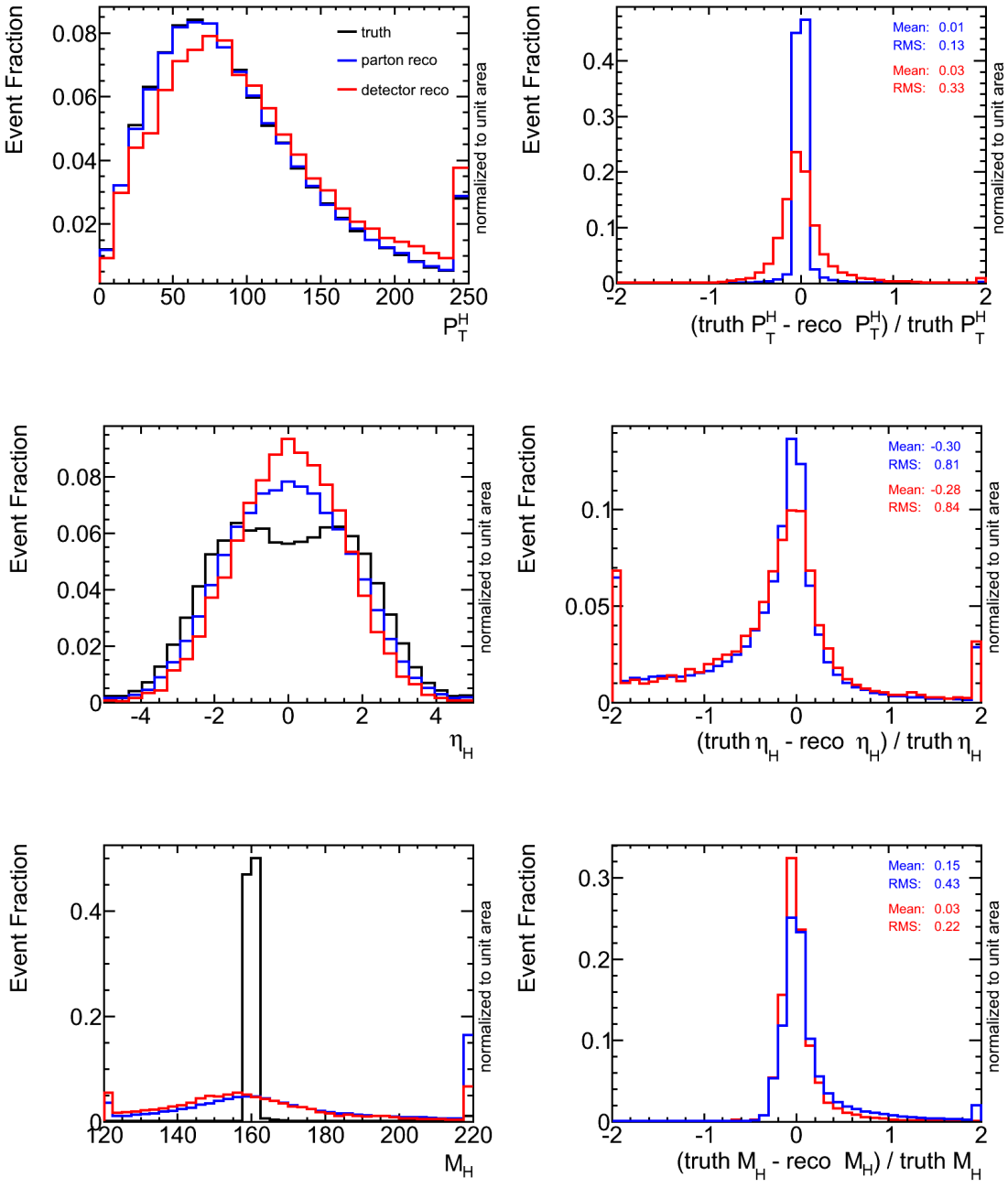


Figure 4.14.: *Properties of the generated and reconstructed Higgs boson: the transverse momentum, the pseudorapidity and the invariant mass distributions shown in the left column and the relative difference of the generated and the reconstructed values in the right column.*

5. Candidate Event Selection

The analysis starts with the candidate event selection. In a first step the objects defined in the reconstruction are identified with physics objects like leptons, jets and $\not{E}_{\mathrm{T}}$. The definitions of these physics objects depend on the purpose of the analysis. In a second step candidate events are selected by applying cuts on the physics objects. These cuts exploit kinematic and topological features of the signal process and thus, suppress events that stem from background processes.

The selection of Higgs boson candidate events is composed of a common preselection and a separation in three analysis channels: Higgs-plus-Zero-Jets ($H + 0j$), Higgs-plus-One-Jet ($H + 1j$) and Higgs-plus-Two-Jets ($H + 2j$). The common preselection requires candidate events being triggered by a single lepton trigger and to have a reconstructed primary vertex, missing transverse energy and two selected lepton candidates. The subsequent separation into jet-based channels ensures an optimized treatment of the relative contributions from signal and background processes in each jet multiplicity bin.

5.1. Object Definitions

Electron criteria

The electron candidates are required to pass the standard ATLAS *"tight"* electron identification quality as described in Section 3.3.3. Candidates are excluded if they are located in regions of the detector that are known not to have been functioning properly during the run when the event was recorded. This information is stored in object quality maps which are two dimensional histograms in η and ϕ. The regions spotted by the maps are caused by:

- isolated dead or high noise channels, masked in the reconstruction
- dead front end boards including regions where the liquid argon calorimeter contains non properly functioning optical transmitters
- dead or non nominal high voltage regions.

Electron candidates are also excluded if they are located in the pseudorapidity regions $1.37 < |\eta| < 1.52$ or $|\eta| > 2.47$, thus outside the coverage of the ATLAS electromagnetic calorimeter. The transverse momentum is required to be larger than

15 GeV/c, it is taken from the calorimeter cluster, while η and ϕ are taken from the track of the electron candidates. In collision data the measured electromagnetic cluster energy of electron candidates is calibrated by applying the energy scales obtained from resonances such as $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ see Reference [125]. This energy scale calibration of electrons depends on the detector region they were measured. They are applied using the following equation:

$$E_{cor} = \frac{E}{1 + f_c}$$

For $|\eta| < 1.4$ the used correction is $f_c = -0.0096$ and $f_c = 0.0189$ for $|\eta| \geq 1.4$.

To suppress electrons which stem from pileup, the track of an electron candidate is required to point back to the primary vertex with transverse impact parameter significance less than 10. The significance is defined by the ratio of the transverse impact parameter to its measured error. The distance of the z-position of the track's origin with respect to the primary vertex should be less than 10 mm. Furthermore electron candidates are required to be isolated by placing cuts on the $\sum p_{\mathrm{T}}^{\mathrm{track}}$, the scalar sum of the transverse momenta of tracks in a cone of $\Delta R = 0.3$ around the candidate and $\sum p_{\mathrm{T}}^{\mathrm{calo}}$, the scaler sum of the transverse momenta of calorimeter energy deposits in a cone of the same size. The track of the electron candidate is excluded in the computation of the sum of transverse momenta. An electron candidate is considered isolated if the relative values $\sum p_{\mathrm{T}}^{\mathrm{track}}/p_{\mathrm{T}}$ and $\sum p_{\mathrm{T}}^{\mathrm{calo}}/p_{\mathrm{T}}$ are lower than 0.1 and 0.15 respectively. The relative isolation provides a good handle to reject QCD background events. Finally, if two electron candidates are reconstructed within a cone of $\Delta R < 0.1$, the candidate with lower p_{T} is discarded.

Muon criteria

Muon candidates are selected from the *"combined muons"* collection. Muon candidates in this collection are reconstructed using a statistical combination of the matched tracks in the Inner Detector and in the Muon Spectrometer and therefor called *"staco muons"*, see Section 3.3.4. The candidates are required to have at least one hit on their track in the pixel detector and at least six hits in the Semi-Conductor Tracker (SCT). Within the acceptance of the Transition Radiation Tracker (TRT) the tracks of muon candidates should meet a successful TRT-extension. The technical details of a TRT-extension are:

- for $|\eta| < 1.9$ require $n > 5$ and $n_{\mathrm{TRT}}^{\mathrm{outliers}} < 0.9 \cdot n$
- for $|\eta| > 1.9$ if $n > 5$ then require $n_{\mathrm{TRT}}^{\mathrm{outliers}} < 0.9 \cdot n$

Where n is given by $n = n_{\mathrm{TRT}}^{\mathrm{hits}} + n_{\mathrm{TRT}}^{\mathrm{outliers}}$. $n_{\mathrm{TRT}}^{\mathrm{hits}}$ denotes the number of TRT hits on the track of the muon candidate and $n_{\mathrm{TRT}}^{\mathrm{outliers}}$ the number of outliers [126].

The momentum p_{ID} as measured using only the Inner Detector (ID) is required to agree with the momentum p_{MS} measured using the Muon Spectrometer (MS) after correcting for the muon energy loss in the calorimeter. If the transverse momentum measured only in the MS is less than 50 GeV/c, the momentum difference $(p_{\mathrm{ID}} - p_{\mathrm{MS}})$ is required to be less than $0.4 \cdot p_{\mathrm{ID}}$ to remove muons from π/K decays. The match between the "ID only" and "MS only" tracks is performed using a χ^2-test as described in Section 3.3.4. If the χ^2 is larger than 150 the muon candidates are also rejected. In addition, the candidates are required to have $p_{\mathrm{T}} > 15$ GeV/c and $|\eta| < 2.4$.

Finally, the same isolation and impact parameter criteria as defined for the electrons must be satisfied. If a selected electron lies within a cone of $\Delta R < 0.1$ around a selected muon the electron candidate is ignored.

Jet criteria

The reconstructed anti-k_{T} jet candidates with $R = 0.4$ are calibrated from the electromagnetic energy scale to the hadronic energy scale using a p_{T} and η dependent correction factor based on Monte Carlo simulation [105]. The candidates are required to have $p_{\mathrm{T}} > 25$ GeV/c at the hadronic energy scale. To meet the signature of vector boson fusion events jet candidates from the full detector coverage up to $|\eta| = 4.5$ are used. An additional cut on the jet-vertex fraction discriminant ε_{jvf} is placed on jets within $|\eta| < 2.1$ where tracking information is available. The quantity ε_{jvf} is the fraction of the summed p_{T} of the tracks, that belong to the jet, pointing back to the primary vertex, to the summed p_{T} of all tracks matched to the jet. Jet candidates with $\varepsilon_{jvf} < 0.75$ are removed thus removing candidates that have a high probability to not stem from the primary vertex. The jet vertex association algorithm is described in [127, 128] and visualized in Figure 5.1.

Selected jets are identified as b-jets if they contain a displaced secondary vertex reconstructed by the SV0 algorithm [109], described in Section 3.3.5, with a weight larger than 5.72. This gives a 50% efficiency on b-jets from $t\bar{t}$-events. If a jet candidate lies within a cone of $\Delta R = 0.3$ around a selected electron the jet candidate is ignored.

$\not{E}_{\mathrm{T}}$ criteria

The missing transverse energy reconstructed from the energy deposits in the ATLAS calorimeters is corrected for the loss in the cryostat and the measured muons, as discussed in Section 3.3.6. More details are provided in References [110, 130].

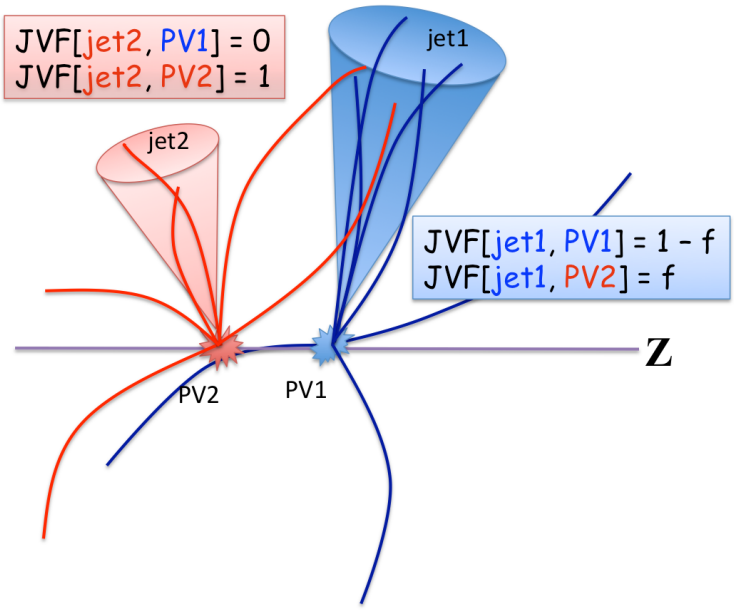


Figure 5.1.: *Visualization of jet-vertex association [129].*

5.2. Common Preselection

A preseletion is applied to enhance the signal-to-background ratio before the analysis proceeds in the three different jet multiplicity bins. The candidate event selection starts with the application of a good run list for data. The good run list used for this $H \to WW \to \ell\nu\ell\nu$ analysis is the same as the one used by the Standard Model $WW \to \ell\nu\ell\nu$ analysis, since both analyses examine events with the same final state. The good run list is compiled upon data quality flags issued for the subdetectors of ATLAS, the used trigger slices and physics object reconstruction. It specifies the conditions of all ATLAS components during data taking and reprocessing.

The next cut applied is the selection of a primary vertex consistent with the beam spot position. It is required to have at least three associated tracks. If more than one primary vertex meets this requirement, the vertex with the highest mass is selected. Monte Carlo events are in addition reweighted depending on the number of primary vertices with more than three tracks. The Monte Carlo reweighting procedure is necessary to adjust Mote Carlo events generated with an average of two in time pileup events to the conditions of the 2010 data. The number of interactions per bunch crossing in this dataset increased due to the increase of the instantaneous luminosity delivered by the LHC. It reached an average of four interactions per bunch crossing in the later data taking periods. The weights for the reweighting procedure are shown in Table 5.1.

Table 5.1.: *Vertex reweighting scheme*

n Vertices	1	2	3	4	5	6+
Weight	1.66781	1.11649	0.846559	0.695447	0.678215	0.672109

Event cleaning criteria are applied on collision candidates to suppress events containing mis-reconstructed jets with subsequently fake $\not{E}_{\mathrm{T}}$. Such jets arise from various sources, ranging from hardware issues like noise in the calorimeter, backgrounds, related to LHC beam conditions, to energy depositions due to cosmic-ray showers. All this indicates non-collision events, hence these events are removed. Details about the event cleaning procedure can be found in [71].

To collect a sample of dilepton events with sensible statistics candidate events are selected using unprescaled single electron and single muon triggers. The electron triggers used for data have a threshold of 14 GeV/c initially, which is then increased to 15 GeV/c as the instanteneous luminosity increases in the later data taking periods. For the muons, a threshold of 10 GeV/c and then 13 GeV/c is used. The usage of triggers with different thresholds is also emulated in the selection of MC events, requiring the appropriate triggers for a fraction of MC events, that corresponds to the integrated luminosity of data where the trigger is used.

Furthermore, the selection requires exactly two leptons as defined in the object definitions with opposite charge. Since the leptons stem from the decay of the neutral Higgs boson. The leading lepton should have a transverse momentum of at least 20 GeV/c and the second one at least 15 GeV/c. The lower p_{T} cut on the second lepton is motivated by an increased acceptance. Requiring the same $p_{\mathrm{T}} > 20$ GeV/c cut on both leptons results in a drop of the event selection efficiency by 40 - 50% on the signal and on the major backgrounds [131].

To adapt the lepton reconstruction efficiencies observed in collision data additional scale factors are applied for MC candidate events. The electron efficiency depends on the transverse energy and the η range. More information on electron performance in the ATLAS experiment can be found in Reference [132]. For muons the efficiency is calculated as a function of the muon transverse momentum, charge, η and ϕ. Ten different regions are used, corresponding to ten different physical regions in the Muon Spectrometer. Figure 5.2 visualizes the locations of these regions in the $\eta - \phi$ plane. Each region represents a particular set of detector layers. A detailed description of the muon reconstruction efficiency is given in References [126, 133].

At this stage, candidate events are separated in three decay channels ee, $e\mu$ and $\mu\mu$ according to the flavor of the two charged leptons. The separation into the decay channels allows to take advantage of different acceptances of the detector and especially the different background contributions. To ensure a robust selection the decay channels where one or both W bosons decay to tau leptons and the correspond-

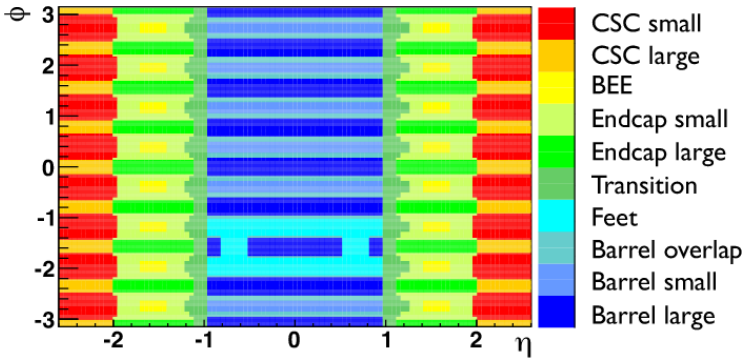


Figure 5.2.: *Location of the ten different physical regions of the Muon Spectrometer in the $\eta - \phi$ plane [126].*

ing neutrino are not included. However, the tau decay channels with a subsequent leptonic tau decay may enter the candidate event selection.

If the two leptons have the same flavor, their invariant mass $M_{\ell\ell}$ is required to be larger than 15 GeV/c^2 in order to suppress Drell-Yan background. Additionally, a Z boson veto is applied to remove candidate events from Z boson production, by rejecting events in which $M_{\ell\ell}$ is in a range consistent with the mass M_Z of the Z boson $|M_Z - M_{\ell\ell}| \leq 10$ GeV/c^2.

As the dilepton decay channel comprises two neutrinos that cannot be measured directly in the detector, a cut on $\not{E}_{\mathrm{T}}^{\mathrm{Rel}} > 25$ GeV is required. This provides a strong suppression of QCD and Z/γ+jets background processes. $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ is calculated as follows:

$$\begin{aligned} \not{E}_{\mathrm{T}}^{\mathrm{Rel}} &= \not{E}_{\mathrm{T}} \sin \Delta\phi && \text{for} \quad \Delta\phi < \frac{\pi}{2} \\ \not{E}_{\mathrm{T}}^{\mathrm{Rel}} &= \not{E}_{\mathrm{T}} && \text{for} \quad \Delta\phi \geq \frac{\pi}{2} \end{aligned} \tag{5.1}$$

where $\Delta\phi$ is the minimum azimuthal separation angle between $\not{E}_{\mathrm{T}}$ and the nearest object in the $\eta - \phi$ plane. The advantage of $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ compared to $\not{E}_{\mathrm{T}}$ is that $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ is less prone to inaccurate measurement of the energy of the nearest object as illustrated in Figure 5.3. The nearest object to $\not{E}_{\mathrm{T}}$ in the scenario illustrated is a lepton ℓ. If the energy of this lepton is mis-reconstructed by ΔE the reconstructed missing energy rec. $\not{E}_{\mathrm{T}}$ is different to the true $\not{E}_{\mathrm{T}}$. In contrast to $\not{E}_{\mathrm{T}}$, $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ does not change if the energy of the nearest lepton or jet is mis-reconstructed with ΔE. $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ is the component of $\not{E}_{\mathrm{T}}$ perpendicular to the direction of ℓ.

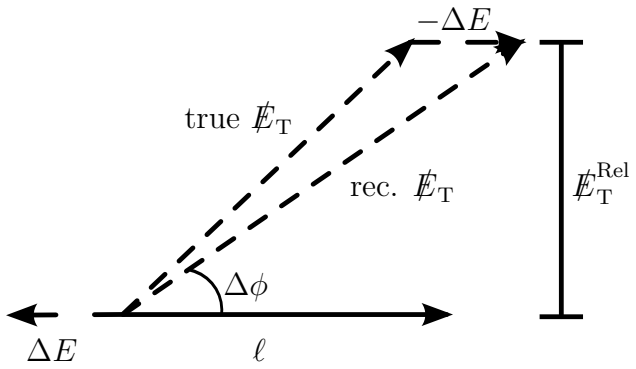


Figure 5.3.: *Visualization of $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$. $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ is the component of $\not{E}_{\mathrm{T}}$ perpendicular to the direction of the lepton ℓ. In contrast to $\not{E}_{\mathrm{T}}$, $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ does not change if the energy of the nearest object, in this illustration assumed to be ℓ, is mis reconstructed with ΔE.*

5.3. Separation into Jet Multiplicity Bins

The candidate events are then separated into zero, one and two jet analyses. Additional cuts are applied to suppress contributions of background processes that are dominant in each jet multiplicity bin.

$H + 0j$ Analysis Channel

Two additional cuts are applied in the Higgs-plus-Zero-Jet analysis channel. One on the upper value of the invariant mass of the two leptons $M_{\ell\ell}$ and one on the upper value of $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ to further suppress the Z+jets background. To cover the entire mass range of the Higgs signal process in an optimal way both cuts depend on the Higgs boson mass.

$$M_{\ell\ell} < 0.5 \cdot M_H$$
$$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$$

These cuts reduce the Z+jets background especially for low Higgs boson masses. They essentially do not reduce the signal acceptance and are also applied in the $H + 1j$ and $H + 2j$ analyses.

$H + 1j$ Analysis Channel

To suppress contributions that arise from the $t\bar{t}$ process, candidate events containing an identified b-jet are vetoed. Jets with a SV0 weight of at least 5.72 are considered as b-jets.

$H + 2j$ Analysis Channel

All events with two and more jets are considered. An additional cut on the pseudo-rapidity gap between the jets $\Delta\eta_{jj}$ is applied. The gap between the jets with the highest transverse momenta is required to be larger than two to separate out the VBF component of the Higgs signal process. The b-jet veto is also applied as the top-quark background is the most important one in this jet multiplicity bin.

The analysis is performed for ten different Higgs boson mass points in the range 120 GeV/c^2 to 200 GeV/c^2. All plots and tables presented here refer to the 170 GeV/c^2 mass point which is used to illustrate different stages of the analysis. Tables 5.2 to 5.4 show the expected number of signal and background events at each cut level of the common preselection and the separation into the three jet multiplicity bins for each decay channel separately. The background yield is calculated with cross-sections predicted by theory and listed in Table 4.1 and 4.2. The expected numbers of signal and background events are normalized to 35.2 pb^{-1} in all tables. The gluon-gluon fusion and the vector-boson-fusion production modes of the Higgs boson are combined as signal.

The dominant background processes in the zero-jet analysis are WW and Z+jets while the $t\bar{t}$ and single top are more important in the one-jet and two-jet analyses.

Table 5.2.: *The expected number of signal (M_H = 170 GeV/c^2) and background events in the ee channel at 35.2 pb^{-1}. The signal and all backgrounds are estimated using MC simulation and theoretical cross-sections.*

Selection	Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$	Total Bkg.	Observed
ee channel	0.86	4.39	1.94	5995	21.71	5.60	6029	6391
$p_T^{\ell 1} > 20$ GeV	0.86	4.37	1.84	5918	21.64	5.52	5952	6252
opposite charge	0.86	4.35	1.22	5879	21.47	4.88	5911	6212
Drell-Yan Veto	0.83	4.31	1.22	5866	21.28	4.85	5897	6174
Z boson Veto	0.77	3.40	1.09	592.3	16.93	0.55	614.3	696
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} > 25$ GeV	0.64	2.09	0.67	2.59	9.68	0.22	15.25	24
0jet	0.37	1.48	0.45	1.45	0.20	0.09	3.68	6
$M_{\ell\ell} < 0.5 \cdot M_H$	0.37	0.83	0.41	1.24	0.10	0.07	2.60	5
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$	0.37	0.78	0.41	1.24	0.01	0.05	2.49	5
1jet	0.18	0.42	0.15	0.37	1.92	0.10	3.02	2
$M_{\ell\ell} < 0.5 \cdot M_H$	0.18	0.23	0	0.37	0.84	0.07	1.51	1
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$	0.18	0.21	0	0.37	0.76	0.07	1.41	1
b-jet Veto	0.18	0.21	0	0.37	0.42	0.07	1.07	1
$\geq$2jet	0.10	0.15	0.07	0.49	6.71	0.07	7.49	16
$M_{\ell\ell} < 0.5 \cdot M_H$	0.07	0.08	0	0.26	3.53	0.02	3.89	6
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$	0.07	0.07	0	0.26	3.16	0.01	3.50	6
b-jet Veto	0.07	0.06	0	0.25	1.05	0.01	1.37	2
$\lvert\Delta\eta_{jj}\rvert > 2$	0.07	0.03	0	0.03	0.27	0	0.33	1

Table 5.3.: *The expected number of signal ($M_H = 170$ GeV/c^2) and background events in the $e\mu$ channel at 35.2 pb^{-1}. The signal and all backgrounds are estimated using MC simulation and theoretical cross-sections.*

Selection	Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$	Total Bkg.	Observed
$e\mu$ channel	2.88	14.1	5.42	74.2	66.14	2.68	162.5	181
$p_T^{\ell 1} > 20$ GeV	2.71	13.24	4.64	59.83	62.5	2.33	142.5	163
opposite charge	2.70	13.18	3.21	59.12	62.18	1.21	138.9	160
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} > 25$ GeV	2.22	8.35	1.79	0.42	35.62	0.54	46.71	52
0jet	1.25	5.83	1.14	0	0.84	0.32	8.12	7
$M_{\ell\ell} < 0.5 \cdot M_H$	1.25	3.08	0.86	0	0.18	0.28	4.49	6
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$	1.25	3.06	0.84	0	0.17	0.27	4.43	5
1jet	0.66	1.87	0.36	0.27	6.76	0.15	9.41	13
$M_{\ell\ell} < 0.5 \cdot M_H$	0.61	0.88	0.36	0.27	3.12	0.09	4.72	8
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$	0.61	0.83	0.36	0.27	2.70	0.08	4.24	8
b-jet Veto	0.61	0.81	0.36	0.27	1.53	0.08	3.05	7
$\geq$2jet	0.33	0.62	0.28	0.17	26.35	0.07	27.49	32
$M_{\ell\ell} < 0.5 \cdot M_H$	0.24	0.29	0.29	0.17	13.03	0.04	13.82	11
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$	0.24	0.25	0.29	0.17	11.55	0.03	12.29	10
b-jet Veto	0.24	0.25	0.15	0.13	3.37	0.03	3.93	2
$\lvert\Delta\eta_{jj}\rvert > 2$	0.17	0.10	0	0.04	0.86	0.01	1.01	1

Table 5.4.: *The expected number of signal ($M_H = 170$ GeV/c^2) and background events in the $\mu\mu$ channel at 35.2 pb^{-1}. The signal and all backgrounds are estimated using MC simulation and theoretical cross-sections.*

Selection	Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$	Total Bkg.	Observed
$\mu\mu$ channel	2.04	10.05	0.82	1.427e+04	47.53	10.01	1.427e+04	1.412e+04
$p_T^{\ell 1} > 20$ GeV	2.02	9.96	0.82	1.377e+04	47.20	1.00	1.384e+04	1.363e+04
opposite charge	2.02	9.96	0.73	1.377e+04	47.00	9.63	1.384e+04	1.362e+04
Drell-Yan Veto	1.93	9.86	0.73	1.372e+04	46.44	9.62	1.379e+04	1.351e+04
Z boson Veto	1.77	7.72	0.53	1735	36.51	0.45	1781	1730
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} > 25$ GeV	1.47	4.93	0.31	8.99	20.86	0.19	35.29	24
0jet	0.83	3.40	0.17	6.26	0.45	0.11	10.39	9
$M_{\ell\ell} < 0.5 \cdot M_H$	0.83	2.07	0.17	5.44	0.28	0.07	8.03	8
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$	0.83	2.06	0.17	5.44	0.27	0.06	8	7
1jet	0.44	1.14	0.14	1.76	4.33	0.05	7.43	5
$M_{\ell\ell} < 0.5 \cdot M_H$	0.43	0.59	0	1.58	2.18	0.04	4.39	3
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$	0.43	0.55	0	1.58	1.85	0.04	4.02	3
b-jet Veto	0.43	0.55	0	1.57	1.01	0.04	3.17	2
$\geq$2jet	0.21	0.37	0	0.80	12.09	0.04	13.03	10
$M_{\ell\ell} < 0.5 \cdot M_H$	0.16	0.19	0	0.70	8.53	0.03	9.45	6
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}} < 0.6 \cdot M_H$	0.16	0.16	0	0.70	7.43	0.02	8.31	5
b-jet Veto	0.16	0.16	0	0.70	2.32	0.02	3.38	2
$\lvert\Delta\eta_{jj}\rvert > 2$	0.11	0.06	0	0.09	0.61	0.01	0.77	0

5.4. Control Regions

In addition to the signal region, defined by the candidate event selection, three separate control regions are defined by reversing particular selection criteria. These control regions are enriched with specific background processes. Their purpose in this analysis is the validation of the used multivariate method, that is introduced in the next chapter. The control regions can also be used to estimate the various background processes.

5.4.1. Z-boson Control Region

The control region enriched with Z+jets background events is defined by reversing the Z boson veto and leaving out the cuts $\not{E}_{\mathrm{T}}^{\mathrm{Rel}} > 25\,\mathrm{GeV}$ and $M_{\ell\ell} < 0.5 \cdot M_H$, while applying all other cuts as described in the candidate event selection. The set of events passing this modified selection is referred to as Z-boson control region in the following sections. It is dominated by Z+jets events as presented in Figure 5.4 visualizing the two main features:

- the invariant mass of the lepton pair $M_{\ell\ell}$, which covers the Z-boson mass peak not included in the default candidate event selection
- and the low $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$.

The agreement between data and Monte Carlo is good in all three jet multiplicity bins and the modeling of the $M_{\ell\ell}$ distributions is very similar to the results of the ATLAS standard model working group [134].

5.4.2. WW Control Region

The dominant background in the $H + 0j$ analysis channel stems from continuum WW decays. The main features that distinguish this background process from the Higgs boson signal are the invariant mass of the lepton pair $M_{\ell\ell}$ and the azimuthal separation angle between the leptons $\Delta\phi_{\ell\ell}$. The WW control region is defined by the same cuts as the signal region except that the cut on $M_{\ell\ell}$ is inversed: $M_{\ell\ell} > 0.5 \cdot M_H$. No cut is applied on $\Delta\phi_{\ell\ell}$ as this observable is used in the multivariate analysis.

Figure 5.5 shows the $M_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ distributions in the zero-jet bin, which is dominated by the WW background process. Both distributions are also shown for the one-jet bin Figure 5.5. In this jet multiplicity bin the top background becomes even more important than WW. As all of the selection cuts are applied the statistics is very low.

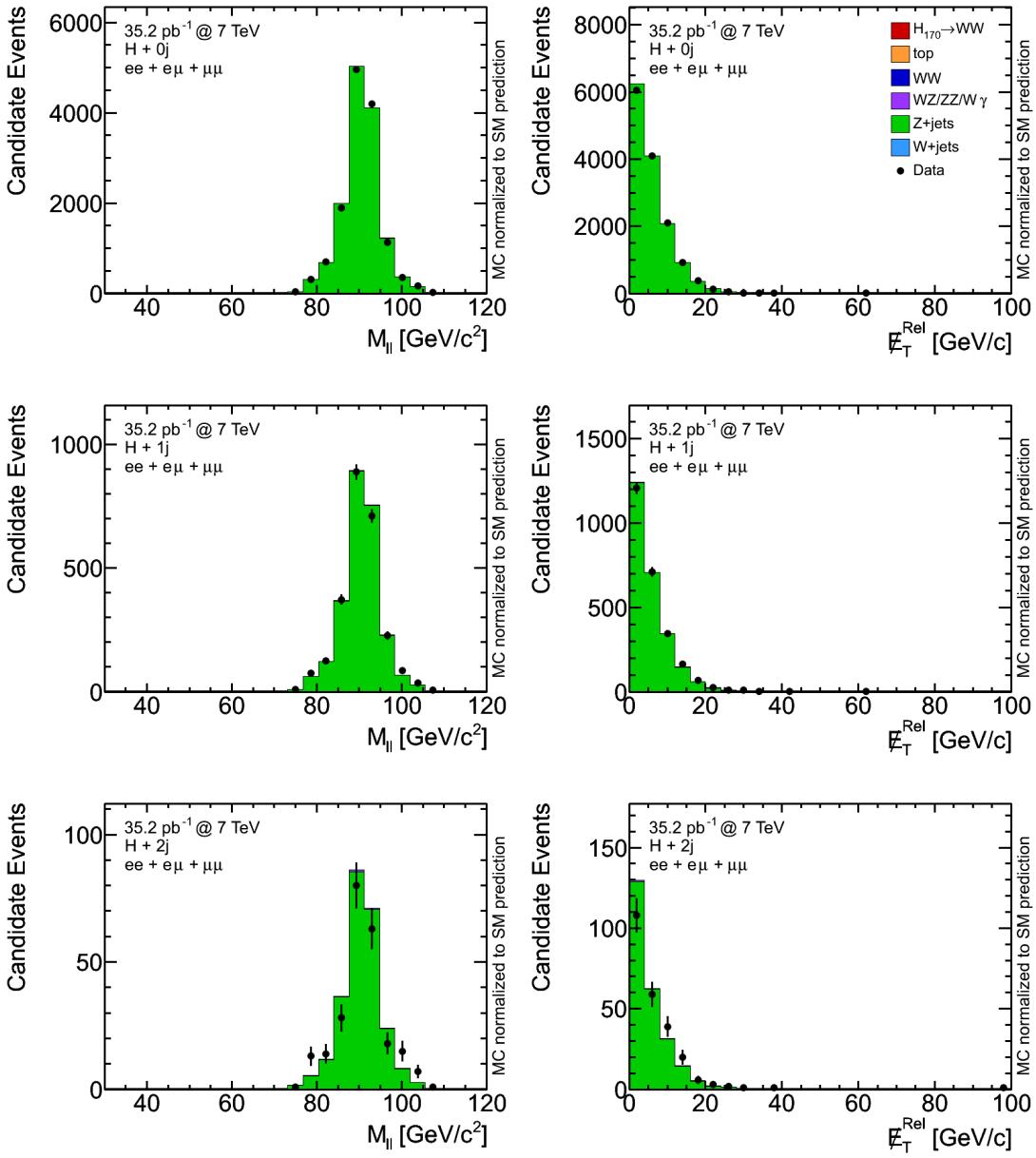


Figure 5.4.: *$M_{\ell\ell}$ and $\not\!\!E_{\mathrm{T}}^{\mathrm{Rel}}$ distributions in the zero-jet, one-jet and two-jet bins of the Z-boson control region for the 170 GeV/c^2 Higgs boson mass point. The agreement between data and Monte Carlo is reasonable in all three jet multiplicity bins.*

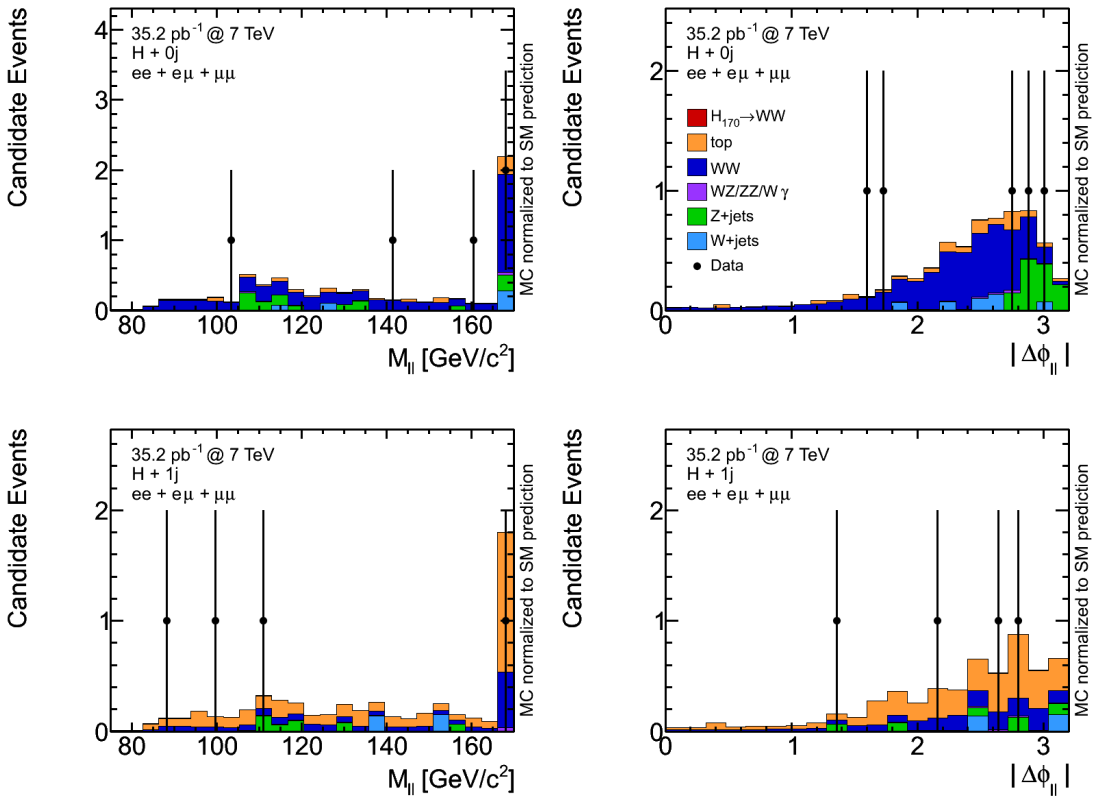


Figure 5.5.: *$M_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ distributions in the zero-jet and one-jet bin of the WW control region for the* 170 GeV/c^2 *Higgs boson mass point.*

5.4.3. Top-Quark Control Region

The dominant background in the $H + 2j$ analysis channel is composed of $t\bar{t}$ and single top processes. The top-quark control region is obtained by inversing the b-jet veto in the one-jet and two-jet multiplicity bins and leaving out the cut on the rapidity gap $\Delta\eta_{jj}$ between the jets. The result for the two-jet bin is shown in Figure 5.6.

This control region is very pure containing almost only top-quark candidate events. A reasonable amount of events is obtained in the two-jet multiplicity bin, with a good agreement between data and Monte Carlo. The top-quark control region in the one-jet bin exhibits a lack of statistics. Therefore, it is not utilized in this analysis.

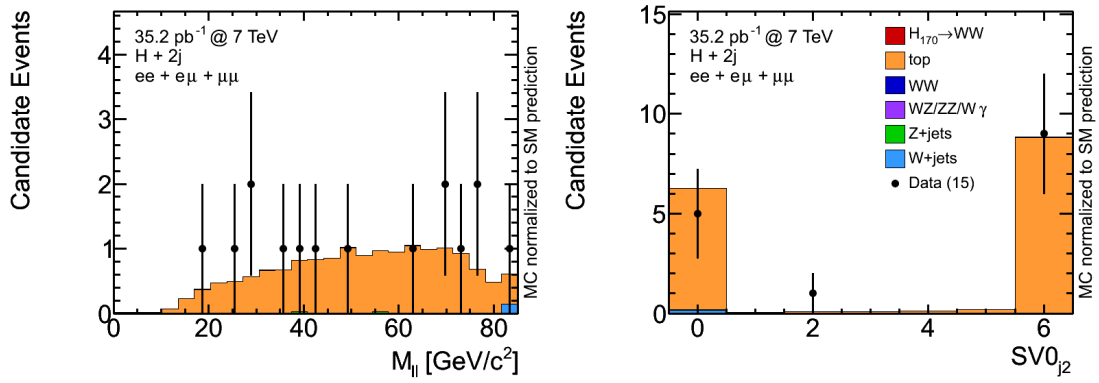


Figure 5.6.: *$M_{\ell\ell}$ and SV0 distributions in the two-jet bin of the top-quark control region for the 170 GeV/c^2 Higgs boson mass point.*

6. Candidate Event Classification

To maximize the sensitivity of the Higgs boson search in the first 35 pb^{-1} of ATLAS data, a multivariate analysis (MVA) technique using artificial neural networks is chosen [135]. Compared to the cut-based analysis [131] the candidate event selection of this analysis applies cuts that are less restrictive, thus gaining in signal acceptance. In particular the requirements on the transverse momentum of the leading lepton and the missing transverse energy are reduced, additionally all topological cuts are omitted.

Table 6.1.: *Number of expected candidate events of signal and background processes for the cut-based and the NN-based analysis performed at the* 170 GeV/c^2 *Higgs boson mass point.*

	Signal		Background	
	cut-based	NN-based	cut-based	NN-based
$H + 0j$	1.26	2.45	1.70	22.19
$H + 1j$	0.60	1.28	1.26	19.89
$H + 2j$	0.06	0.47	0.02	26.44

The number of expected signal and background events for cut-based and the NN-based analysis is presented in Table 6.1. The 170 GeV/c^2 Higgs boson mass point is used again as example to illustrate different stages of the analysis. The signal yield after the candidate event selection of the NN analysis is two to seven times larger. However, it comes at the expense of a dramatically increased background contribution.

The usage of a MVA technique allows to expand the region of phase space, where the signal to background ratio is much smaller, compared to the phase space selected by the cut-based analysis. This is possible as the pattern recognition algorithm of an MVA method does not only use the kinematic and topological differences of signal and background processes provided by the input variables, but takes also additional information from the correlations of the presented variables into account. The output of the MVA method is a discriminant distribution that has different shapes for signal and background processes. These shapes, obtained using Monte Carlo predictions,

are used as templates in a fit to the discriminant distribution observed in data in order to determine the contribution of the signal process.

The multivariate method deployed in this analysis is an artificial neural network provided by the NeuroBayes® package. A brief description of the operating principle of neural networks and the characteristic features of NeuroBayes® is given in the first part of this chapter. The following part describes the training of the used neural networks including the input variables and their validation. Finally, the validation of the trained networks is presented, by evaluating them in the three control regions.

6.1. Neural Networks as Event Classifiers

The NeuroBayes® neural network package [136, 137] provides an algorithm for classification and shape reconstruction. The main feature is the sophisticated and robust preprocessing, which is performed before the input variables are presented to the neural network. NeuroBayes® implements a three-layer feed-forward neural network.

The network infrastructure for classification consists of $n+1$ input nodes or neurons in the first layer. One node for each of the n input variables plus one bias node. The number of nodes in the second or hidden layer is arbitrary and can be chosen by the user. The last layer contains one output node that gives a continuous output distribution between minus one and plus one. Each node in the hidden layer is connected with all nodes in the input layer. The connections are represented by weights w_{ij}. Thus, the input for a hidden node h_j is a weighted sum of the input variables $\vec{x}$ including a bias term β_j, which represents the connection to the bias node

$$h_j(x_1..x_n;\beta_j) = \sum_{i=1}^{n} w_{ij}x_i + \beta_j.$$

The output of each hidden node is determined by passing its input $h_j(x_1..x_n;\beta_j)$ to a transfer function that maps the interval $]-\infty,+\infty[$ to the interval $[-1,+1]$. The transfer function applied is a symmetric sigmoid function S that is sensitive in a small range around zero see Figure 6.1.

$$S(h_j) = \frac{2}{1+e^{-h_j}} - 1$$

The weight β_j of the bias node implements the threshold of the hidden node h_j. It shifts the mean of the input distribution $h_j(x_1..x_n;\beta_j)$, determined by the training sample, to the linear part of the sigmoid function around zero. This procedure avoids saturation effects for large absolute values of $h_j(x_1..x_n;\beta_j)$.

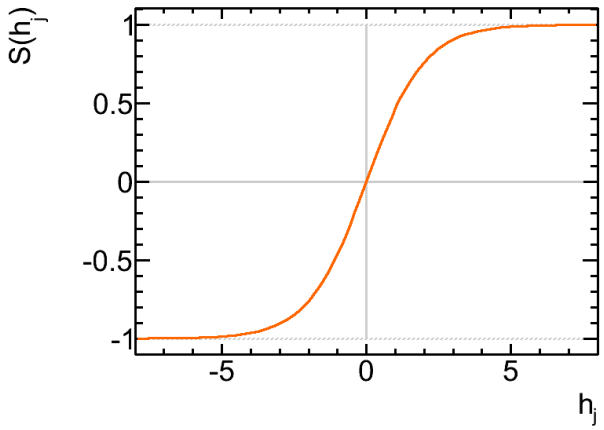


Figure 6.1.: *The symmetric sigmoid function $S(h_j)$ is deployed as transfer function in each hidden node. It maps the interval $]-\infty,+\infty[$ to the interval $[-1,+1]$ and is sensitive in a small range around zero.*

Each hidden node h_j is connected with the output node by a further weight w_j, thus the output o of a neural network with n input variables and m nodes in the hidden layer is given by

$$o = S\left(\sum_{j=1}^{m} w_j \cdot S\left(\sum_{i=1}^{n} w_{ij}x_i + \beta_j\right)\right).$$

Ideally, this output is close to $+1$ for signal and close to -1 for background processes. To achieve this the weights w_{ij} and w_j are adjusted during the training of the neural network, which is performed using the training sample. The training corresponds to finding a minimum in a multidimensional space. The training procedure and the composition of the training sample are described in the next sections.

The preprocessing prior to the network training prepares the input variables in an optimal way to facilitate the training. In the first step the input variable distributions are equalized and scaled, such that the resulting distribution is uniform between minus one and one. This transformation utilizes the cumulative probability density of the input distribution, and inhibits extreme outliers in the input to completely saturate the neurons and thus to dominate the output of these nodes. In a second step the flattened distributions are then converted into Gaussian distributions with a mean value of zero and a standard deviation of one to provide optimal conditions for a fast initial learning.

The procedure described so far is called individual preprocessing as it is applied separately to each input variable. This gives the possibility to control the preprocessing of each variable by assigning individual preprocessing flags. The most important

features that may be set using individual preprocessing flags are:

- fitting a spline curve to the flattened distribution when transforming to a Gaussian distribution to handle statistical fluctuations,
- handling of discrete input variables, which may also represent classes with a certain order,
- using variables that are only given for a subset of events by assigning the missing values to a δ-function.

After the individual preprocessing the variables are ranked according to the significance of their correlation to the target, which is $+1$ for signal events and -1 for background events. In the next step, the correlation matrix of all input variables and their correlation to the target are computed. The significance of a variable is determined by the loss of the total correlation to the target caused by removing variables from the computation one after the other. The variable causing the smallest loss of information is discarded. The same calculation with the remaining $n-1$ variables is repeated removing the second least important variable and so on. The loss of correlation to the target divided by the square root of the training sample size defines the significance of a variable. This leads to a ranking of variables according to their additional significance. Within NeuroBayes®, it is possible to discard variables based on their additional significance to incorporate only those variables that include relevant information that is not already included by other variables.

Finally, the input variables are decorrelated in the global preprocessing by calculating the covariance matrix which is then diagonalized. The rotated input variables are then divided by the square root of the corresponding eigenvalues transforming the covariance matrix into a unit matrix. Preprocessing the multidimensional correlated input variables allows to find the optimal starting point for the subsequent network training.

6.2. Training of Neural Networks

During the training process the weights w_{ij} and w_j are optimized on Monte Carlo predictions, by minimizing the deviation of the neural network output o_k and the target value t_k for each event k of the training sample. With $t_k = +1$ for signal events and $t_k = -1$ for background events as mentioned before. The function minimized during the training is the entropy loss function

$$E = \sum_k \log\left(\frac{1}{2}(1 + t_k - o_k + \epsilon)\right) ,$$

where ϵ is a small regularization constant to avoid numerical problems for untrained networks. It is reduced in each training iteration and is zero just after a few iterations.

The advantage of the entropy loss function is that it leads to infinitely large values for completely wrong classifications.

E is a measure for the deviation of the current network output from the target value. The training process uses the gradient descend method to minimize E iteratively, by changing the weights w_j and w_{ij} proportional to the current gradient of E.

The update of the weights w_j from the hidden nodes to the output node is easily calculated, because the target of the output node and its error are known.

$$w_j \to w_j + \Delta w_j^{\text{new}} \qquad \text{with} \qquad \Delta w_j^{\text{new}} = -\gamma \frac{\partial E}{\partial w_j} + \mu \Delta w_i^{\text{previous}} ,$$

Δw_j^{new} denotes the update of the weight w_j in the current iteration and $\Delta w_i^{\text{previous}}$ the weight update in the previous iteration. γ is a constant of proportionality that is used to control the learning speed. The constant μ controls the step width of gradient descend. Thus $\mu \Delta w_i^{\text{previous}}$ represents a momentum term incorporating information from the previous weight update.

In order to update the weights w_{ij} connecting the input layer with the hidden layer one needs the target value of the corresponding hidden node h_j or its error. However, it is impossible to compute the target value for a hidden node. The solution is provided by the backpropagation algorithm. It derives the error of a hidden node by propagating back the error of the output node. This essentially means applying the chain rule of partial derivatives:

$$w_{ij} \to w_{ij} + \Delta w_{ij}^{\text{new}}$$

with

$$\Delta w_{ij}^{\text{new}} = -\gamma \frac{\partial E}{\partial w_{ij}} + \mu \Delta w_{ij}^{\text{previous}} = -\gamma \frac{\partial E}{\partial S} \frac{\partial S}{\partial h_j} \frac{\partial h_j}{\partial w_{ij}} + \mu \Delta w_{ij}^{\text{previous}} .$$

Both parameters the learning speed γ and the momentum μ are used again to control the weight update. They may be adjusted by the user. A low learning speed makes the algorithm take smaller steps following slowly but accurately the path of steepest descend on the error surface E. The momentum term causes the algorithm to slow down where the gradient changes quickly, while speeding it up in flat areas with small gradient changes. This technique simply adds a fraction of the previous weight update to the current one and helps to get out of local minima.

Furthermore, regularization techniques based on Bayesian statistics deployed in the training avoid the risk of overtraining and enhance the generalization ability of the network. The weights are systematically reduced during the training process. This weight decay regularization may lead to pruning of network connections and even nodes if they become statistically insignificant; thus, reducing the number of free parameters and removing statistical fluctuations. The resulting network represents a minimal topology to reproduce the characteristics of the training data.

6.3. Composition of Training Samples

The sample of events used in the network training is required to emulate the data to be analyzed with the trained network. It should contain the relevant physics processes, each with reasonable statistics, and their relative fraction should correspond to the estimated composition of the observed events. As the predicted production cross-section and hence the number of expected Higgs boson events is very small compared to the background processes, the approach to use the same number of signal events as expected in data is not practicable.

To emphasize the importance of the signal events the training sample is composed of 50% signal and 50% background events. The relative fractions of the considered background processes are given by the respective production cross-sections and acceptance and selection efficiencies as obtained in the candidate event selection. Additionally the composition of the fractions is not based on the number of events, as the process with the smallest generated Monte Carlo statistics or the lowest selection efficiency would also limit the number of events of the other processes. This would imply dropping some of the events of the processes with higher statistics and hence loosing information and restricting the phase space of these processes. Instead the full available Monte Carlo statistics of each process is utilized by weighting each event with the production cross-section and selection efficiency of the corresponding process.

Table 6.2 denotes the composition of the training sample for the neural network training with the aim to separate Higgs signal events with a Higgs boson mass of 170 GeV/c^2. The table presents the available candidate events after the selection, and the computed sum of weights to match the required relative fractions of the considered processes. The Higgs process comprises gluon-gluon fusion (GGF) and vector boson fusion (VBF) production modes and the decay chain $H \to WW \to \ell\nu\ell\nu$. While the content of the background processes are only top-quark pair production $t\bar{t} \to bWbW \to b\ell\nu b\ell\nu$ and continuum $WW \to \ell\nu\ell\nu$ production. Ten different Higgs boson mass points from 120 GeV/c^2 to 200 GeV/c^2 are used in this analysis. A trainings sample is generated for each of the ten mass points, containing signal events with the respective Higgs boson mass and the relative fraction of background processes. The relative fraction of the two background processes changes slightly for different Higgs boson mass points, since there are two selection criteria that depend on the Higgs boson mass.

Backgrounds like other diboson processes $WZ/ZZ/W\gamma$, Z+jets and W+jets exhibit a lack in statistics after the hard selection requirements. Therefore, they are not included in the training process to avoid undesired effects resulting in a bad generalization ability of the trained network.

Table 6.2.: *Signal and background fractions in the training samples used in the three analyses for the Higgs boson mass point of* 170 GeV/c^2*. The events are reweighted to obtain a signal-to-background ratio of 50:50. The relative fractions of background processes is given by the respective production cross-sections and selection efficiencies, to simulate the expectation of the background fractions in data.*

(a) $H + 0j$

Process	Available Events	Sum of Weights	Relative Fraction
Higgs	30261	61064	50.0%
$t\bar{t}$	926	3217	2.6%
WW	55477	57847	47.4%

(b) $H + 1j$

Process	Available Events	Sum of Weights	Relative Fraction
Higgs	13103	26915	50.0%
$t\bar{t}$	5652	15040	27.9%
WW	15877	11874	22.1%

(c) $H + 2j$

Process	Available Events	Sum of Weights	Relative Fraction
Higgs	3935	8292	50.0%
$t\bar{t}$	4683	7422	44.8%
WW	1878	870	5.2%

6.4. Input Variables

To find variables with a good separation power between signal and background events 30 to 60 different input variables were considered, depending on the jet multiplicity bin. Starting with simple distributions like p_{T}, η and ϕ of leptons and jets including also various angular distributions, among these $|\Delta\phi_{\ell\ell}|$ and $|\Delta\eta_{jj}|$ and other combinations. Reconstructed variables as the invariant mass of the two leptons $M_{\ell\ell}$ and the transverse Higgs boson mass M_{T} were checked as well as combined variables like the scalar sum of all transverse energies H_{T} and the momentum balance of the two jets and the Higgs boson decay products in the transverse plane $p_{\mathrm{T}}^{\mathrm{balance}} = p_{\mathrm{T}}^{jj} - p_{\mathrm{T}}^{\ell\ell\not{E}_{\mathrm{T}}}$. Variables that are sensitive to the topology of the events like aplanarity and sphericity were also considered.

A minimal set of most powerful but simple variables is chosen to obtain a robust network for the analysis of the first data and to be comparable with other multivariate analyses. The additional significance computed in the preprocessing step is used for this purpose. A common set of eight input variables, that characterize the properties of the Higgs boson decay products, is used across all three analyses.

$H + 0j$ Input Variables

Eight input variables are used in the $H + 0j$ analysis comprising the properties of the two leptons and $\not{E}_{\mathrm{T}}$.

- $p_{\mathrm{T}}^{\ell 1}$ and $p_{\mathrm{T}}^{\ell 2}$, the transverse momenta of both charged leptons.
- $\eta^{\ell 1}$, the pseudorapidity of the charged lepton with the higher transverse momentum.
- $p_{\mathrm{T}}^{\ell\ell}$, the transverse momentum of the vector sum of the charged leptons.
- $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$, as defined in equation 5.1.
- $|\Delta\phi_{\ell\ell}|$, the azimuthal separation angle between the charged leptons.
- $|\Delta\phi_{\ell 2\not{E}_{\mathrm{T}}}|$, the azimuthal separation angle between the lepton with the lower transverse momentum and the missing transverse energy.
- M_{T}, the transverse Higgs boson mass:

$$M_{\mathrm{T}} = \sqrt{(E_{\mathrm{T}}^{\ell\ell} + \not{E}_{\mathrm{T}})^2 - (p_{\mathrm{T}}^{\ell\ell} + \not{E}_{\mathrm{T}})^2}$$
$$\text{with} \qquad E_{\mathrm{T}}^{\ell\ell} = \sqrt{(p_{\mathrm{T}}^{\ell\ell})^2 + M_{\ell\ell}^2}\,.$$

As neural networks are trained with simulated and reconstructed Monte Carlo events, it is necessary to check the modeling of the input variables. The shape of each input variable obtained by the signal and background models, as described in chapter 4, is compared with the shape observed in data. Figures 6.2 to 6.4 depict all input variables presented above for the $H + 0j$ analysis. The column on the left hand side shows the distributions normalized to unit area, so called shape distributions, of the signal process and the WW and top backgrounds as predicted by the Monte Carlo. The shape distributions of the remaining background processes $WZ/ZZ/W\gamma$, Z+jets and W+jets are not included, because these processes are not used in the training and because the distributions of these processes are very spiky due to the low statistics of the available MC samples. The validation of the modeling is presented in the column on the right hand side containing the observed data and the predicted contributions of all considered processes as stacked histograms. The plots show reasonable agreement between the modeling and the observation.

Figure 6.2 presents the p_{T} and η distributions of the leptons. The $p_{\mathrm{T}}^{\ell 2}$ distribution provides some separation power between the signal and background processes. Figure 6.3 presents the $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ and the angular distributions $|\Delta\phi_{\ell\ell}|$ and $|\Delta\phi_{\ell 2\not{E}_{\mathrm{T}}}|$ that allow a nice separation of the WW process, which is the most dominant background. Finally, the more complex variables $p_{\mathrm{T}}^{\ell\ell}$ and M_{T} are shown in Figure 6.4.

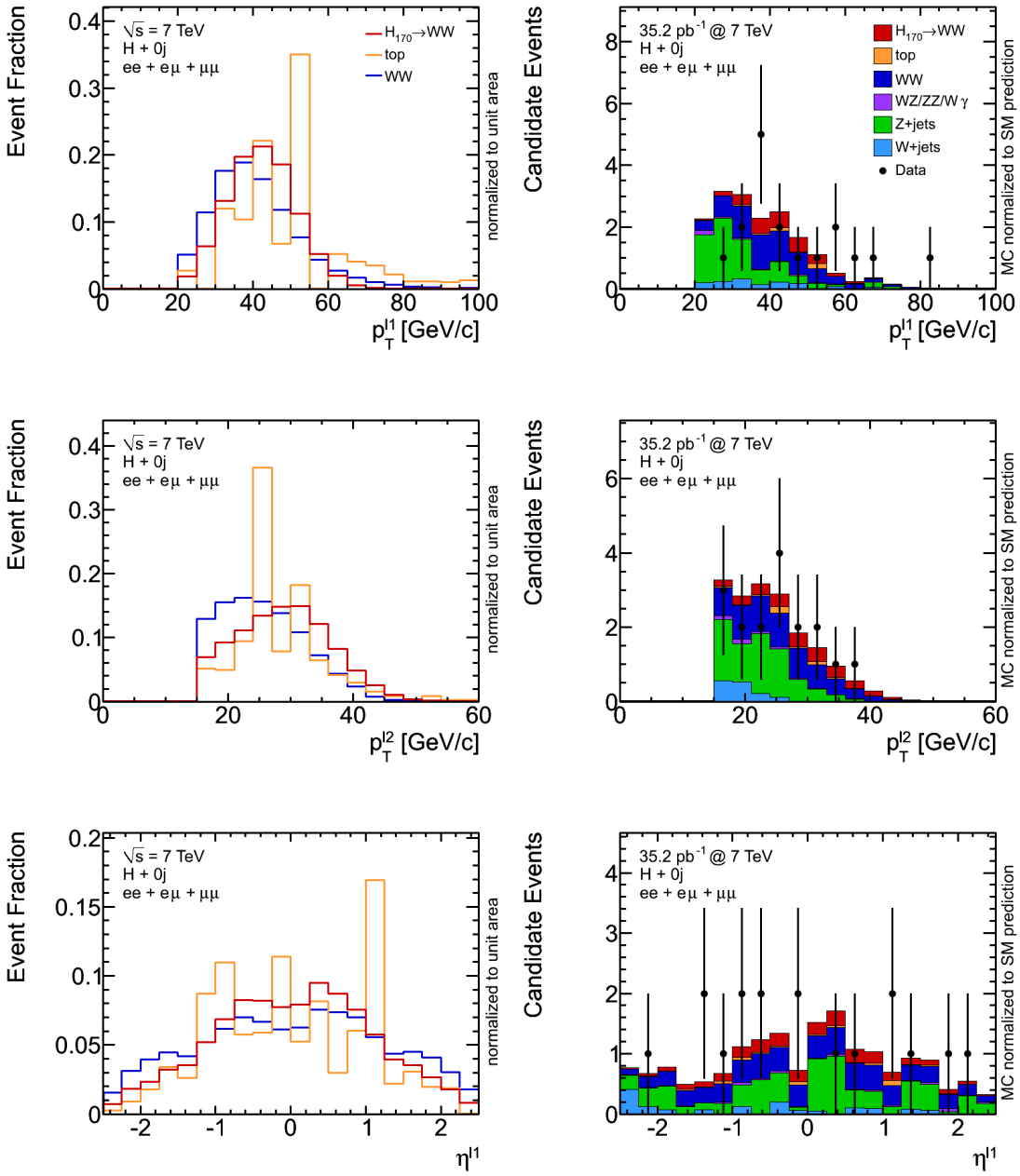


Figure 6.2.: *The p_T and η distributions of the leptons in the $H+0j$ analysis for a Higgs boson mass of 170 GeV/c^2. The column on the left hand side shows the kinematic distributions normalized to unit area, while the validation of the modeling is presented in the column on the right hand side. All three decay channels are combined.*

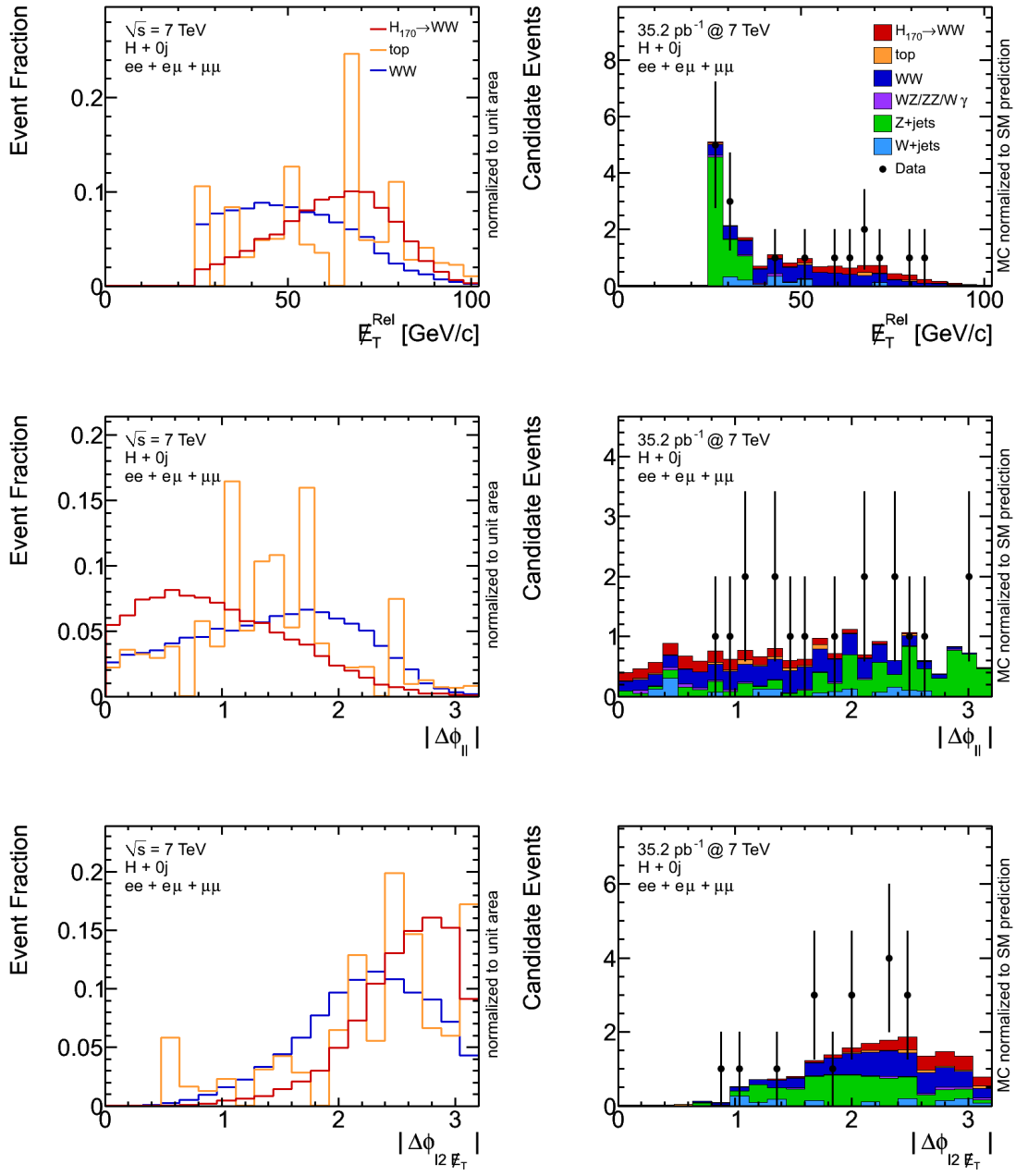


Figure 6.3.: *The $\not\!\!E_T^{\mathrm{Rel}}$ and the angular distributions $|\Delta\phi_{\ell\ell}|$ and $|\Delta\phi_{\ell 2\not\!\!E_T}|$ in the $H+0j$ analysis for a Higgs boson mass of 170 GeV/c^2. The column on the left hand side shows the kinematic distributions normalized to unit area, while the validation of the modeling is presented in the column on the right hand side. All three decay channels are combined.*

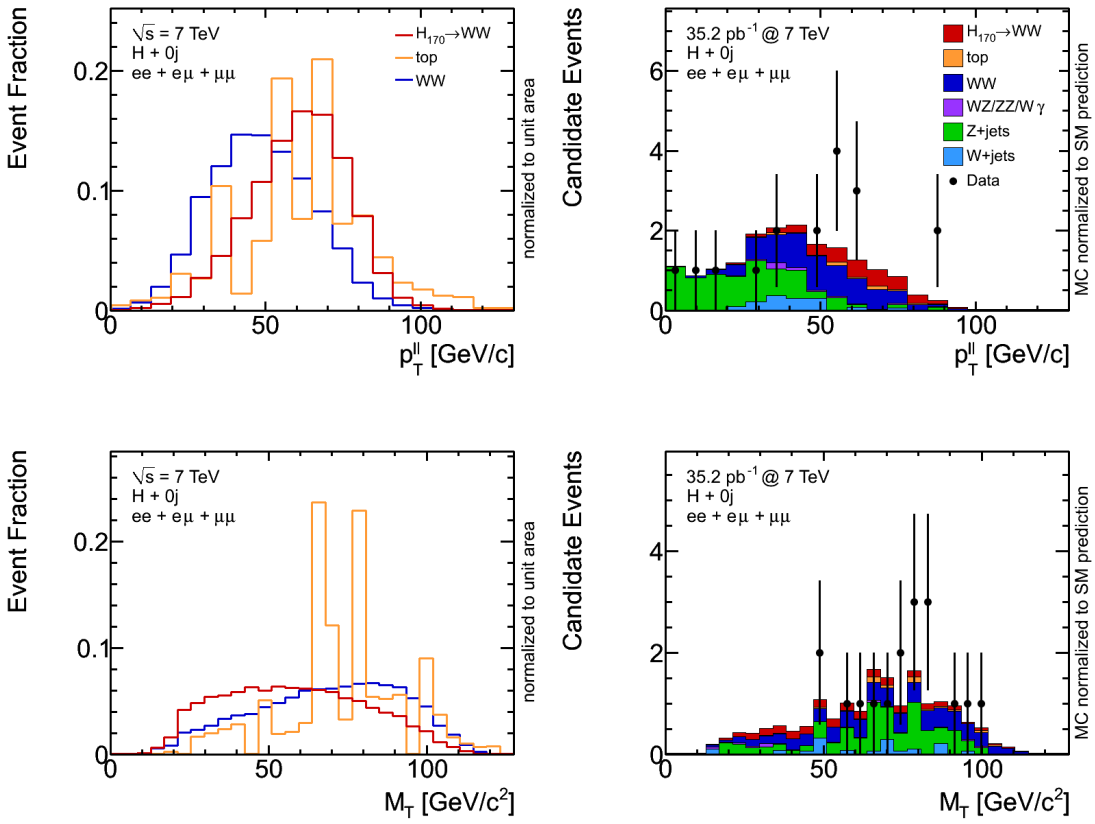


Figure 6.4.: *The $p_{\mathrm{T}}^{\ell\ell}$ and M_{T} distributions in the $H+0j$ analysis for a Higgs boson mass of 170 GeV/c^2. The column on the left hand side shows the kinematic distributions normalized to unit area, while the validation of the modeling is presented in the column on the right hand side. All three decay channels are combined.*

$H+1j$ Input Variables

The following input variables are used in addition to the variables described in the $H+0j$ analysis:

- p_{T}^{j1}, the transverse momentum of the jet,
- η^{j1}, the pseudorapidity of the jet.

The distributions of both variables are depicted in Figure 6.5. The input variables, characterizing the properties of the Higgs boson decay products, are presented in appendix A.1. They do not depend on the jet multiplicity bin. Thus, the shape of these observables and thus their discrimination power compared to the ones presented in the $H+0j$ case does not change. Only the background composition is different in this jet multiplicity bin, which can be seen in the stacked plots of Figure 6.5 and Table 6.2. The WW and the top background processes are equally important in this

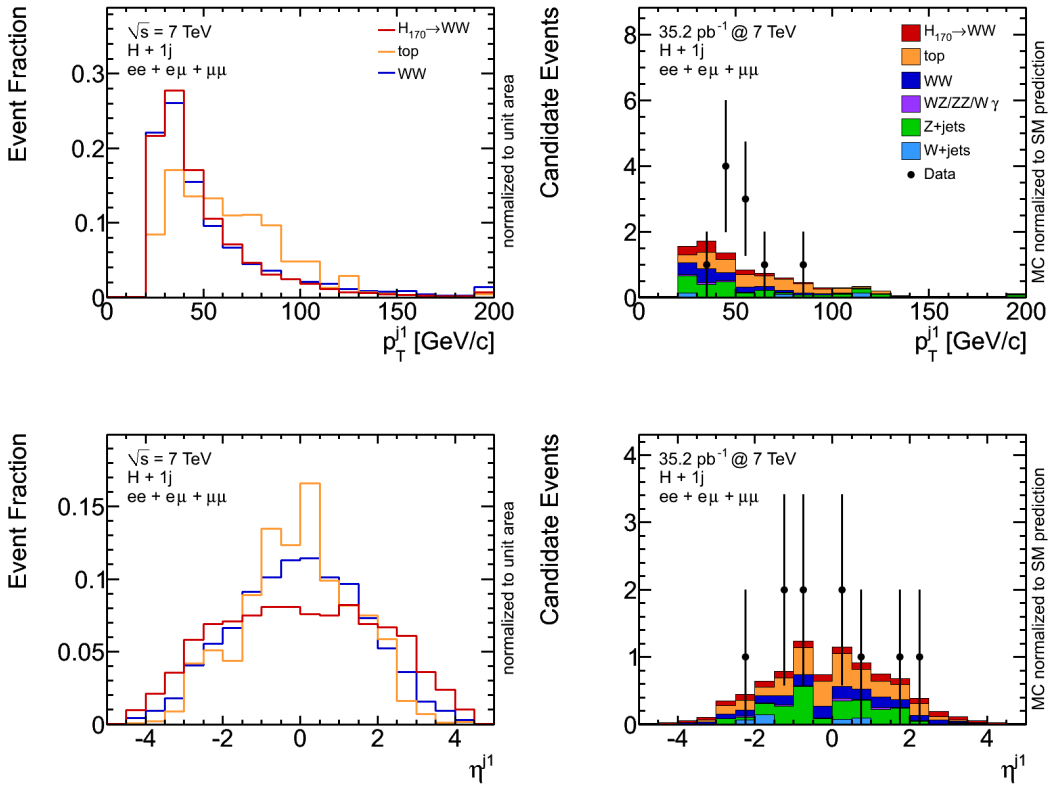


Figure 6.5.: *The p_{T}^{j1} and η^{j1} distributions of the jet in the $H+1j$ analysis for a Higgs boson mass of 170 GeV/c^2. The column on the left hand side shows the kinematic distributions normalized to unit area, while the validation of the modeling is presented in the column on the right hand side. All three decay channels are combined.*

jet bin.

$H + 2j$ Input Variables

The following four variables are used in addition to the variable set of the $H + 1j$ analysis:

- p_{T}^{j2}, the transverse momentum of the second jet.
- $|\Delta\eta_{jj}|$, the pseudorapidity gap between the two jets.
- M_{jj}, the invariant mass of the two jets.
- $p_{\mathrm{T}}^{\mathrm{balance}} = p_{\mathrm{T}}^{jj} - p_{\mathrm{T}}^{\ell\ell\not{E}_{\mathrm{T}}}$, the balance of the two jets and the Higgs boson decay products in the transverse plane.

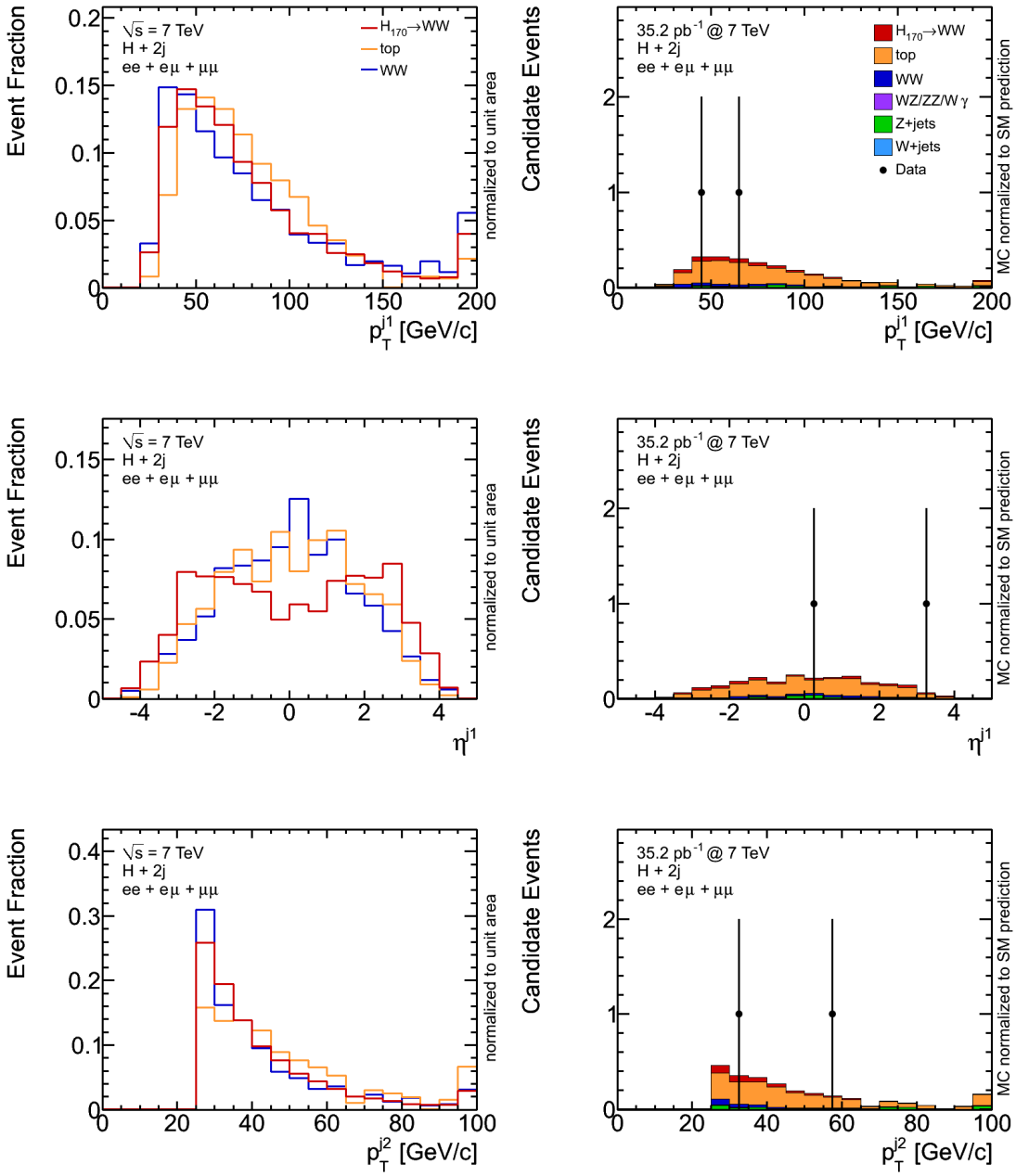


Figure 6.6.: *The p_{T} and η^{j} distributions of the jets in the $H+2j$ analysis for a Higgs boson mass of 170 GeV/c^2. The column on the left hand side shows the kinematic distributions normalized to unit area, while the validation of the modeling is presented in the column on the right hand side. All three decay channels are combined.*

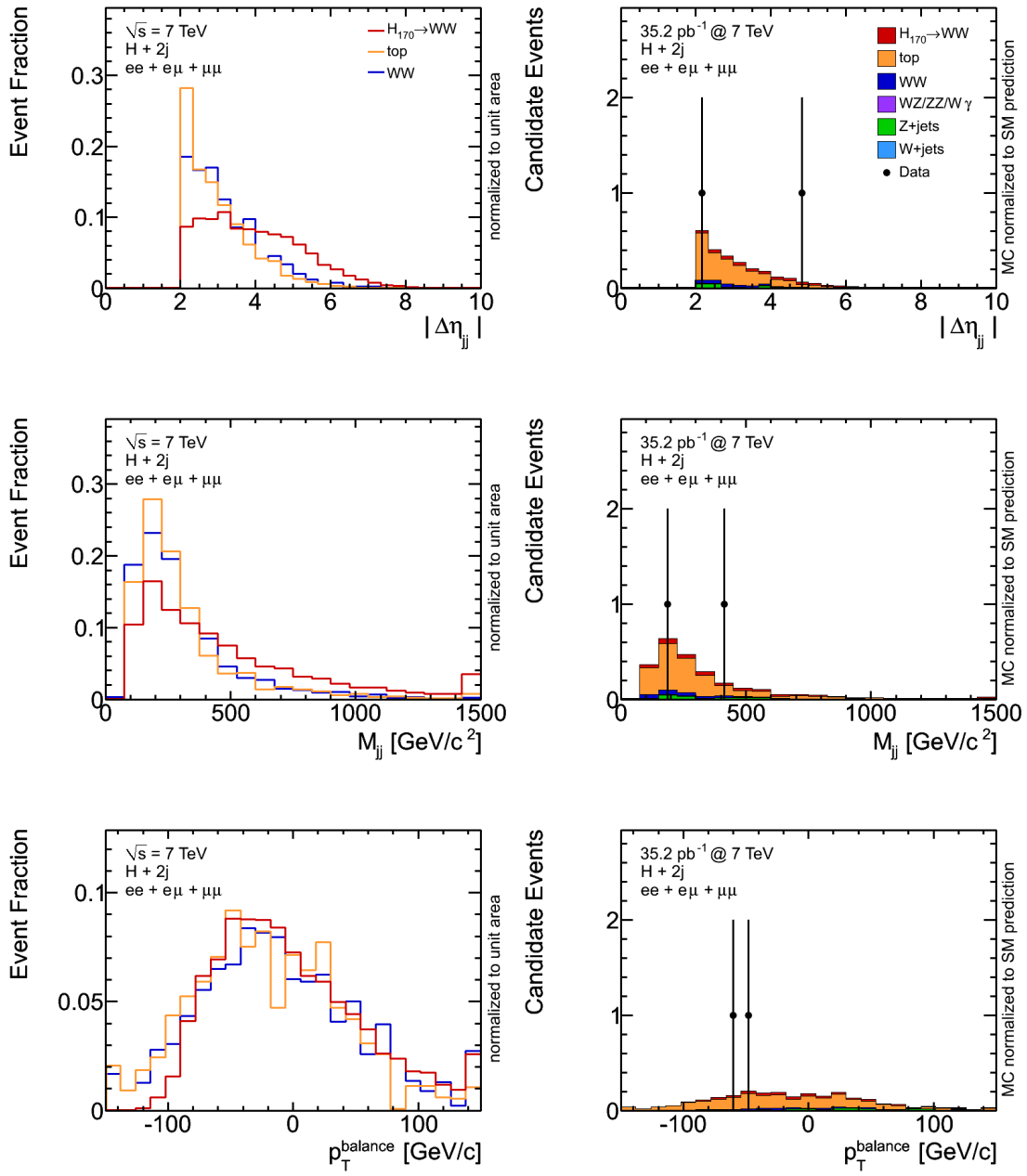


Figure 6.7.: *The $|\Delta\eta_{jj}|$, M_{jj} and $p_{\mathrm{T}}^{\mathrm{balance}} = p_{\mathrm{T}}^{jj} - p_{\mathrm{T}}^{\ell\ell\not{E}_{\mathrm{T}}}$ distributions in the $H + 2j$ analysis for a Higgs boson mass of 170 GeV/c^2. The column on the left hand side shows the kinematic distributions normalized to unit area, while the validation of the modeling is presented in the column on the right hand side. All three decay channels are combined.*

Figure 6.6 depicts the transverse momenta and the pseudorapidity distributions of the jets in the $H+2j$ analysis. The pseudorapidity distribution illustrates that the jets are produced in forward regions of the detector. The most characteristic and also the most important input variable in this jet multiplicity bin is the $|\Delta\eta_{jj}|$ distribution shown in Figure 6.7. It provides a strong separation power between signal and background processes. The invariant mass distribution of the two jets and the balance of transverse momenta of the jets and the Higgs boson decay products are also depicted in that figure.

Validation of Correlations between Input Variables

In addition to the validation of the input variables Figure 6.8 provides validation plots of the correlations between the four most important input variables of the $H + 0j$ analysis. Since correlations are exploited by the neural networks to classify events, their modeling is checked as well. The plots visualize the correlation variable

$$\kappa_{ij} = \frac{x_i - \bar{x_i}}{\sigma_i} \cdot \frac{x_j - \bar{x_j}}{\sigma_j}$$

between two input variables x_i and x_j, where $\bar{x_i}$ denotes the ensemble mean of variable x_i and σ_i its RMS.

The correlation plots show good agreement between data and simulation. The modeling of all correlations is checked before the training of the neural networks.

6.5. Optimization of Neural Network Classifiers

Three neural networks are trained for each of the ten Higgs boson mass points. The global and individual preprocessing flags of NeuroBayes® are chosen carefully to ensure the proper treatment of the input distributions before the training. Table 6.3 presents a summary of all input variables in all three jet multiplicity bins of the analysis at the Higgs boson mass point of 170 GeV/c^2. The variables are sorted in descending order of their additional significance computed in the preprocessing. This ranking does not change much across the analyses of the different Higgs boson mass points.

NeuroBayes® provides several options that may be altered by the user to control the training and to finetune the separation power of the network. In order to find an optimal set of these options, automated trainings are performed scanning through the value ranges of each option. Changing only one option for each training and comparing the training results, the most important options for this analysis are found to be:

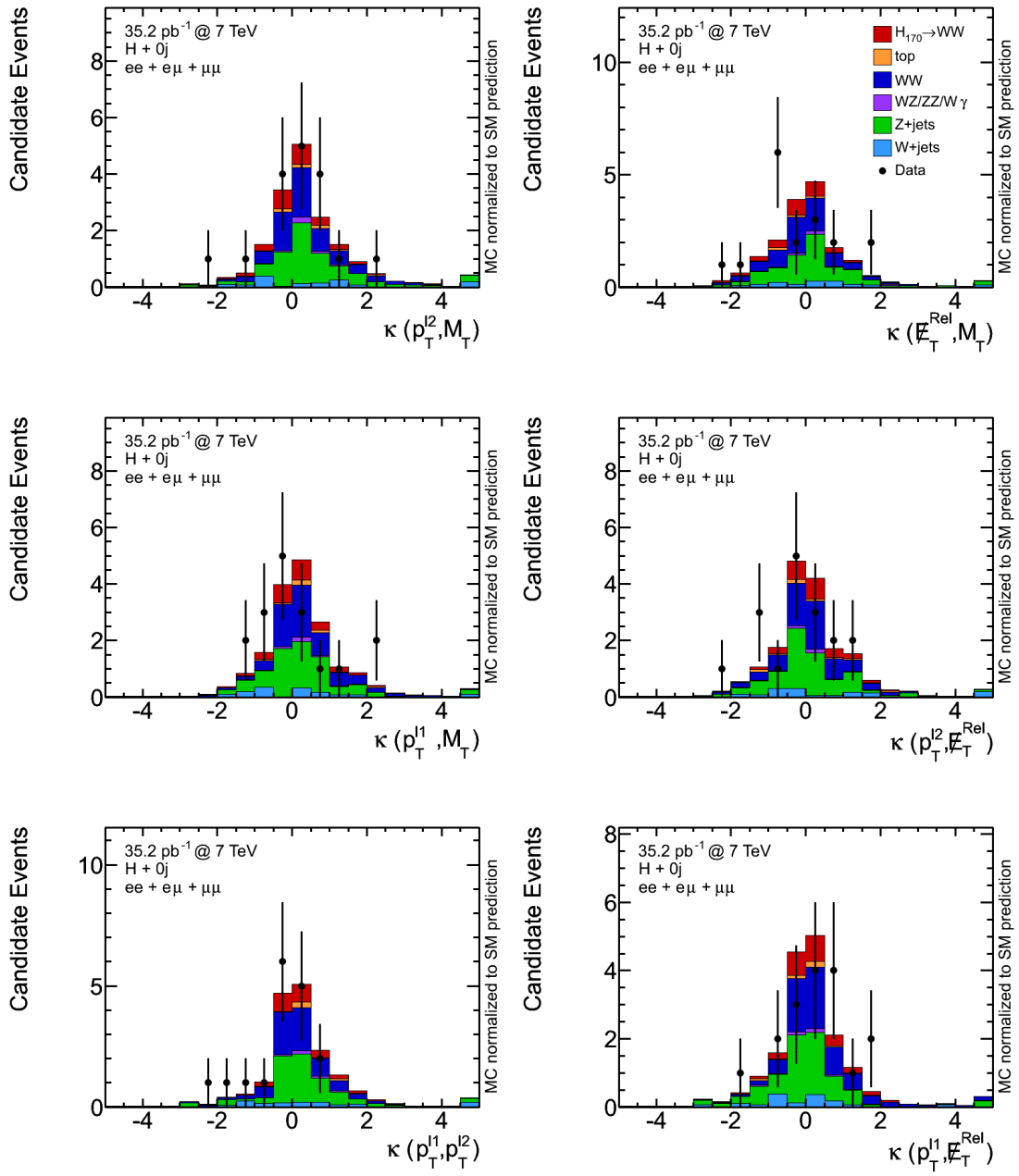


Figure 6.8.: *Validation plots of the correlations between the four most important input variables of the $H + 0j$ analysis at the 170 GeV/c^2 Higgs boson mass point.*

Table 6.3.: *Summary of all of input variables in each jet multiplicity bin for the Higgs boson mass point of* 170 GeV/c^2. *The variables are sorted in descending order of their additional significance as computed in the preprocessing.*

$H+0j$	$H+1j$	$H+2j$
M_T	M_T	$\lvert\Delta\eta_{jj}\rvert$
$p_{\mathrm{T}}^{\ell 2}$	η^{j}	M_T
$\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$	$p_{\mathrm{T}}^{\ell 2}$	$p_{\mathrm{T}}^{\ell 2}$
$p_{\mathrm{T}}^{\ell 1}$	$\lvert\Delta\phi_{\ell 2\not{E}_{\mathrm{T}}}\rvert$	$\lvert\Delta\phi_{\ell 2\not{E}_{\mathrm{T}}}\rvert$
$\eta^{\ell 1}$	$p_{\mathrm{T}}^{\ell 1}$	p_{T}^{j2}
$\lvert\Delta\phi_{\ell 2\not{E}_{\mathrm{T}}}\rvert$	$\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$	M_{jj}
$\lvert\Delta\phi_{\ell\ell}\rvert$	p_{T}^{j}	$p_{\mathrm{T}}^{balance}$
$p_{\mathrm{T}}^{\ell\ell}$	$\eta^{\ell 1}$	$p_{\mathrm{T}}^{\ell\ell}$
	$p_{\mathrm{T}}^{\ell\ell}$	$\eta^{\ell 1}$
	$\lvert\Delta\phi_{\ell\ell}\rvert$	$\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$
		η^{j1}
		p_{T}^{j1}
		$p_{\mathrm{T}}^{\ell 1}$
		$\lvert\Delta\phi_{\ell\ell}\rvert$

- the number of hidden nodes,
- the learning speed and
- the momentum.

The number of hidden nodes changes the layout of the neural network, while the learning speed and the momentum control the weight update during the training process as described in Section 6.2. The dependence of the separation power on the number of hidden nodes is unusual for the NeuroBayes® package and is caused by the low statistics of the available MC samples. Networks with only one hidden node provided the best separation.

To find the network with the best separation power a second scan is performed, now using finer steps in the value ranges of the three most important options and testing all possible combinations. In total, 90 trained networks for each $H+nj$ analysis are obtained. The network with the best performance in terms of separation power between signal and background processes is chosen for each analysis. The final values of the used options are compiled in Table 6.4. This procedure is repeated for each Higgs boson mass point.

Table 6.4.: *Neural network parameters used for the training in the three jet multiplicity bins, for the Higgs boson mass point of* 170 GeV/c^2.

Option	$H+0j$	$H+1j$	$H+0j$
hidden nodes	1	1	1
learning speed	0.5	0.05	0.5
momentum	0.1	0.1	0.9

Another optimization test is performed using the $e\mu$-decay channel only. As mentioned in Section 5.2 the candidate events are separated not only in jet multiplicity bins but also in decay channels ee, $e\mu$ and $\mu\mu$ to exploit the different composition of the background processes in each channel. The training of the neural network classifiers is not performed separately in each decay channel, but incorporates all three decay channels. After the training process the decay channels are considered again, evaluating the trained classifier and constructing the template distributions for the statistical analysis in each decay channel separately.

The advantage of this approach is the availability of the full statistics for the training, which results in an improved generalization ability of the trained classifier. This is scrutinized by comparing the separation power of an event classifier trained with events from all decay channels with the separation power of an event classifier trained with events from the $e\mu$-decay channel only. Both classifiers are evaluated in the $e\mu$-decay channel. Figure 6.9 presents the expected significance of both classifiers as a function of the neural network output (NNoutput). The classifier trained with the full statistics provides a better significance and hence a better separation power. The significance is defined as

$$S_{cL} = \sqrt{2[(S+B)\log(1+\frac{S}{B}) - 1]}$$

where S and B are respectively the expected signal and background events after a cut on the neural network output. S_{cL} is utilized as significance instead of $\frac{S}{\sqrt{B}}$, as S_{cL} provides a better approximation of the discovery significance for low statistics following a poisson distribution [138].

Based on this test only one neural network is trained in each $H+nj$ analysis utilizing events from all decay channels. The separation in the different decay channels is done after the training.

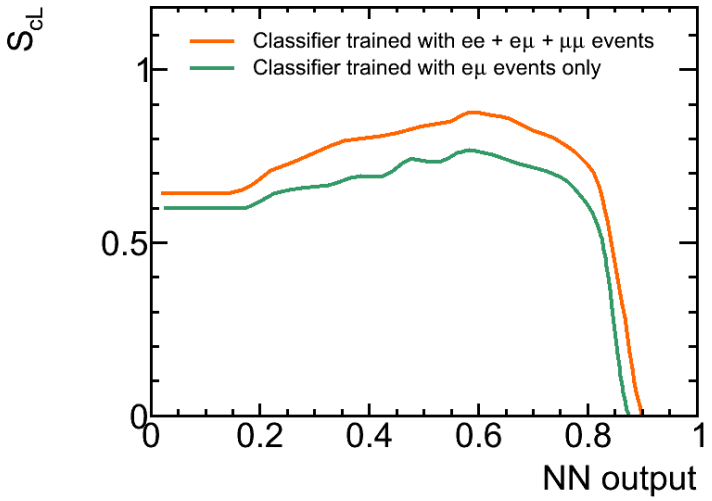


Figure 6.9.: *Expected significance of a neural network classifier trained with $ee + e\mu + \mu\mu$ events, compared to the expected significance of a neural network classifier trained with $e\mu$ events only. Both curves are obtained by evaluating the classifiers in the $e\mu$-decay channel.*

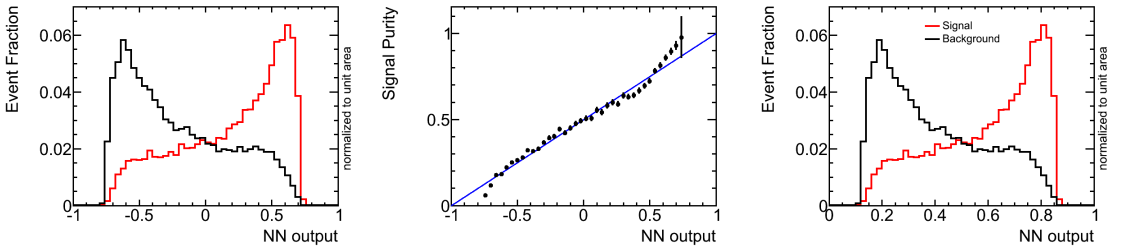
6.6. Training Results and Classifier Validation

One example of the neural network output (NNoutput) distribution for the trained classifier at the Higgs boson mass point of 170 GeV/c^2 is given in Figure 6.10. It is a continuously distributed variable between minus one, which indicates the background like region, and the signal region near plus one Figure 6.10(a). After a good training the signal purity in each bin of the NNoutput distribution is close to the diagonal Figure 6.10(b). This is the precondition to transform the NNoutput distribution to the interval $[0, 1]$ which allows to interpret the output value as the probability for each event to be signal. The transformed NNoutput distribution is depicted in Figure 6.10(c) and is used in this analysis.

The validation of the trained classifiers and the template distributions obtained by evaluating the classifiers on ee, $e\mu$ and $\mu\mu$ events in the different jet multiplicity bins are discussed in the following sections.

6.6.1. Event Classifier Validation

The control regions as described in Section 5.4 are utilized to validate the trained event classifiers. This is a crucial test of the robustness of the event classifiers. The events of these control regions were not available in the training sample and may lie in a phase space region not covered by the training sample.



(a) (b) (c)

Figure 6.10.: *Transformation of the output distribution of NeuroBayes® classifiers from the interval [-1, 1] (a) to [0, 1] (c) which allows to interpret the output value as the probability for each event to be signal. The signal purity in each bin of the NNoutput distribution has to be close to the diagonal (b).*

Z-boson Control Region

The NNoutput distributions for the Z-boson control region are shown for each jet multiplicity bin in Figure 6.11, displaying the shapes of each considered process on the left hand side and the data and Monte Carlo comparison on the right hand side. The main features of the Z-boson control region are the invariant mass of the lepton pair $M_{\ell\ell}$ that covers the Z-boson mass peak and low $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ values, see Figure 5.4.

The classification of the events from this control region as background like events is evident, as can be seen in the plots on the right hand side. These plots depict also a reasonable agreement between data and Monte Carlo. Thus the most important points of the validation are satisfied.

The shape plots on the left hand side provide additional details:

- The top and WW processes that were available in the training sample, but not in this particular phase space, are classified as background like. This is explained by the fact that even $\not{E}_{\mathrm{T}}^{\mathrm{Rel}}$ revealing unusual values, the remaining variables show still distributions similar to the ones in the training sample. Thus, the presented events are classified as background events.
- The Higgs signal process is classified as background, too. Especially in the case of the $H + 0j$ classifier, which demonstrates that the classifier is conservative, because two deviating variables are sufficient to cause this classification also for signal events.
- The main feature of the classifiers is the classification of $WZ/ZZ/W\gamma$, Z+jets and W+jets processes, that were not used in the trainings sample, as background like.

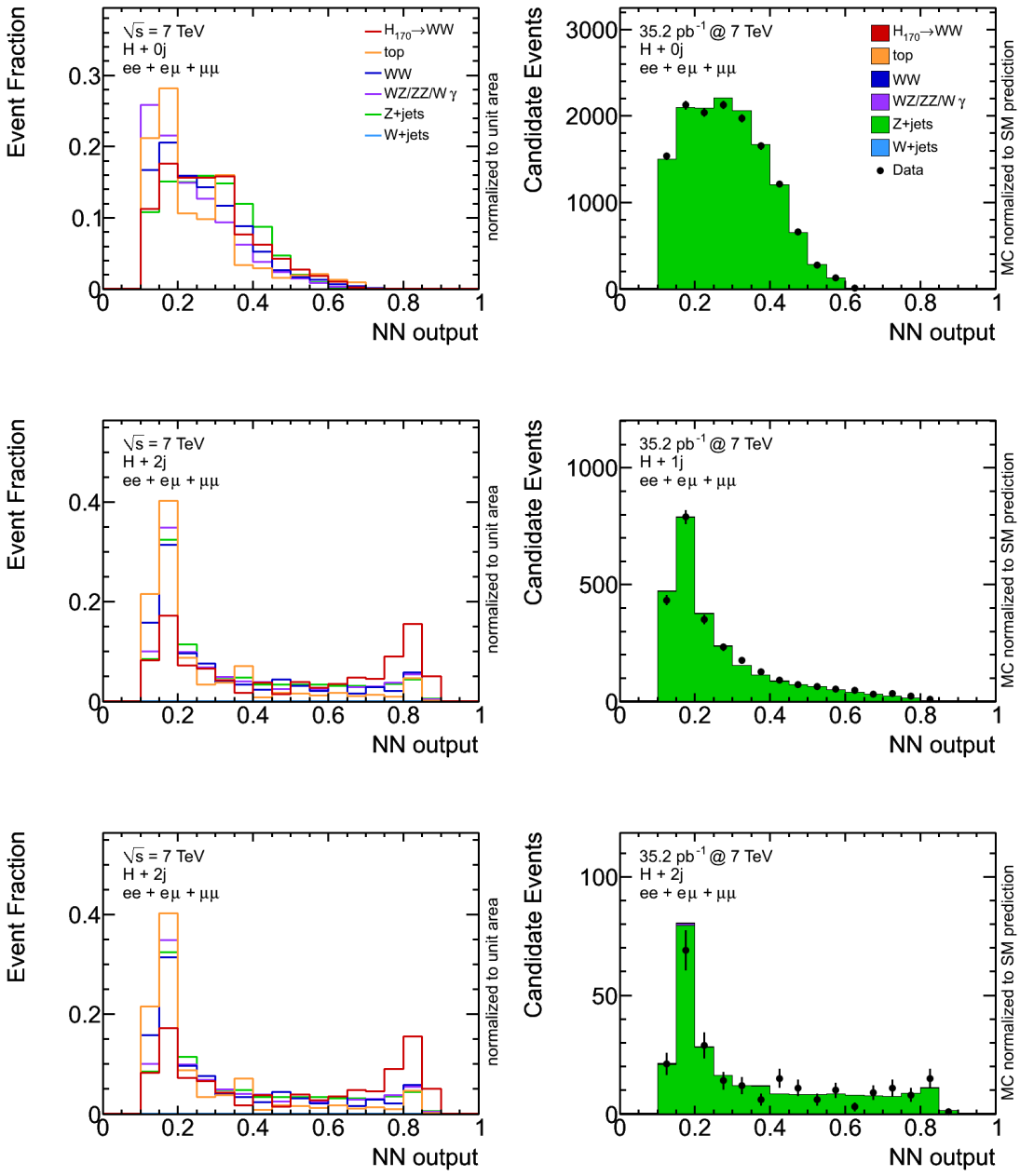


Figure 6.11.: *NNoutput distributions of the three neural network classifiers in the Z-boson control region. The classification of the events from this control region as background-like events is evident.*

The points just mentioned apply to all three $H + nj$ event classifiers. In the zero-jet bin, all events are classified as background like. Some differences due to additional information provided by the jet variables arise in the one-jet multiplicity bin resulting in a tail of the NNoutput distribution, that extends towards the signal region. This is more pronounced in the $H + 2j$ case, where the jet variables are very important, see Table 6.3. As a consequence a large fraction of the Higgs signal process is still classified signal like.

WW Control Region

The evaluation of the classifiers in the WW control region is presented in Figure 6.12. Again, all events in the zero-jet multiplicity bin are classified as background like. While the additional jet variables in the one-jet bin help to identify a fraction of the Higgs boson event as signal. The MC modeling of the data is also quite satisfying.

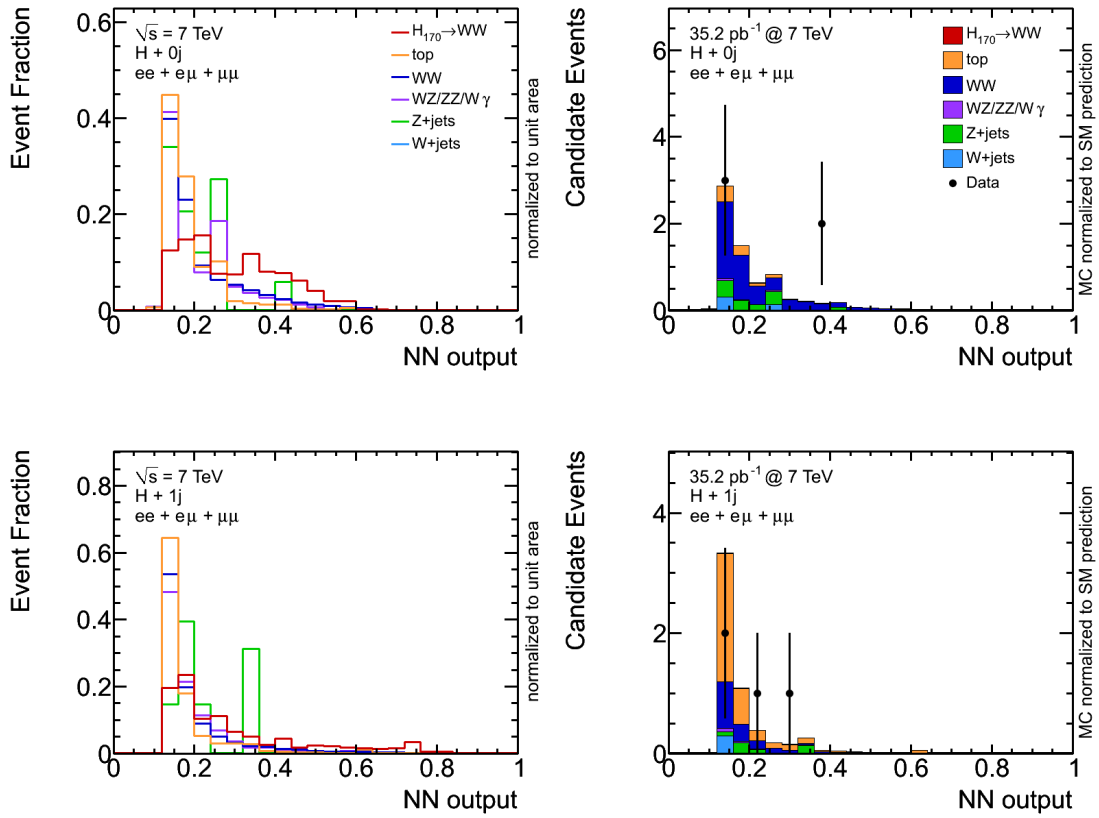


Figure 6.12.: *NNoutput distributions of the neural network classifiers in the WW control region.*

Top-Quark Control Region

Figure 6.13 depicts the neural network output in the top-quark control region, showing a nice agreement between data and MC. The shapes of all background processes peak in the background region, while the shape of the signal processes results in an almost flat distribution.

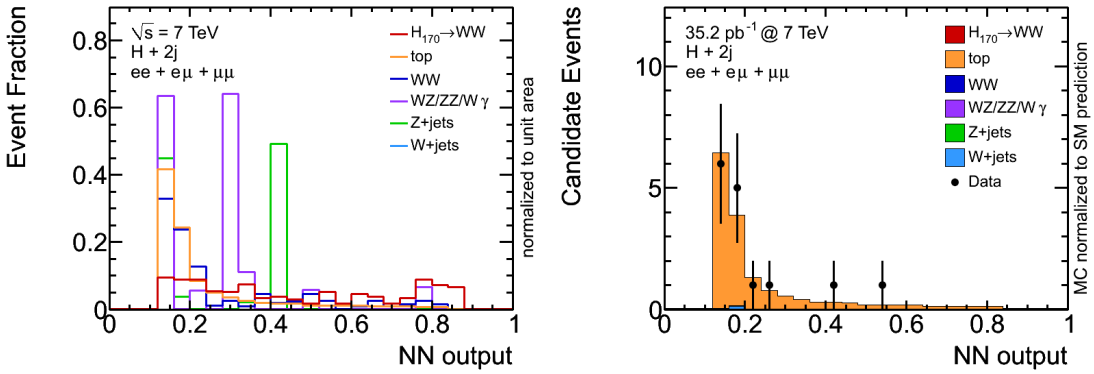


Figure 6.13.: *NNoutput distributions of the neural network classifiers in the top-quark control region.*

6.6.2. Template Distributions

The NNoutput template distributions of the neural network classifiers trained for the Higgs boson mass point of 170 GeV/c^2 can be found in Figures 6.14, 6.15 and 6.16. The three decay channels ee, $e\mu$ and $\mu\mu$ are shown separately for the $H + 0j$ and $H + 1j$ analyses. This separation in decay channels is not used in the $H + 2j$ analysis due to the low statistics expected in the dataset to be analyzed. The sum of all channels is presented for all jet bins. All NNoutput distributions show a nice separation of signal and background events. The $WZ/ZZ/W\gamma$, Z+jets and W+jets processes that were not included in the training sample are classified as background like.

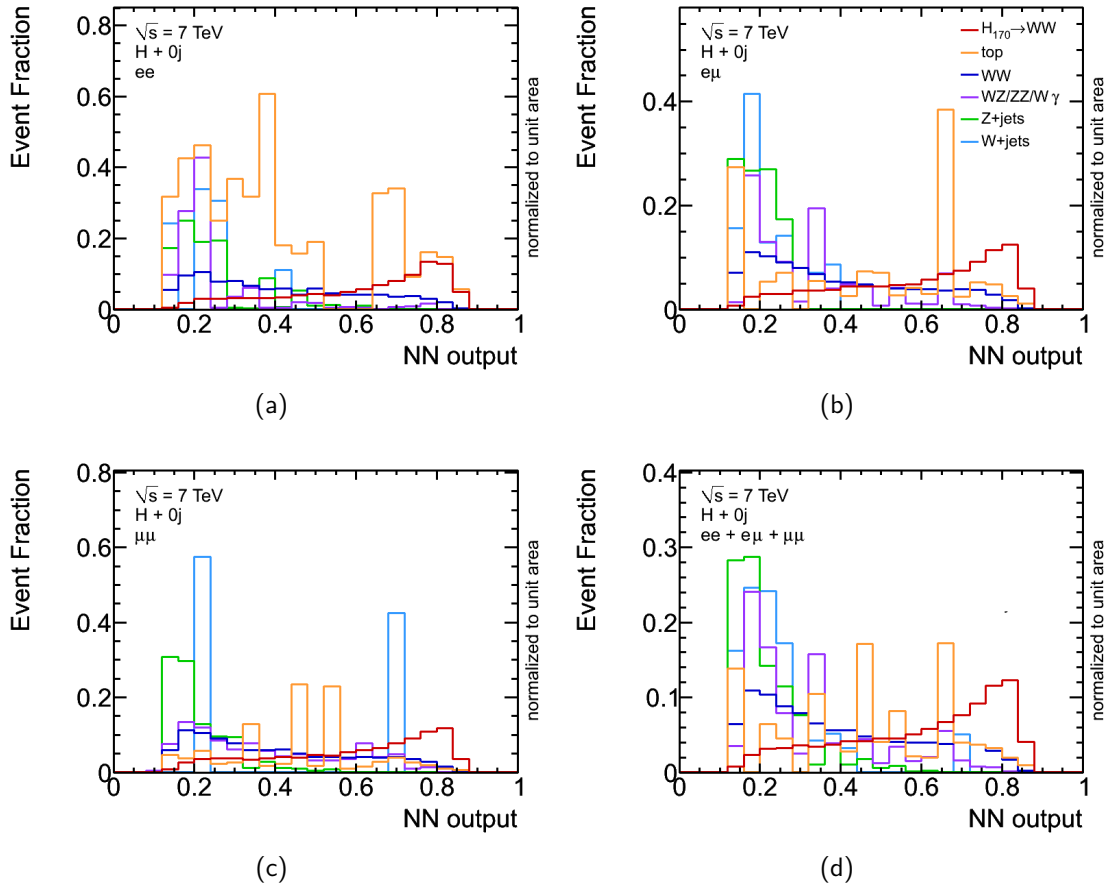


Figure 6.14.: *Template distributions of the NNoutput for the $H + 0j$ analysis and a Higgs boson mass of 170 GeV/c^2. The distributions are shown separately for the three decay channels: ee (a), eμ (b), and μμ (c). The sum of all three decay channels is shown as well (d).*

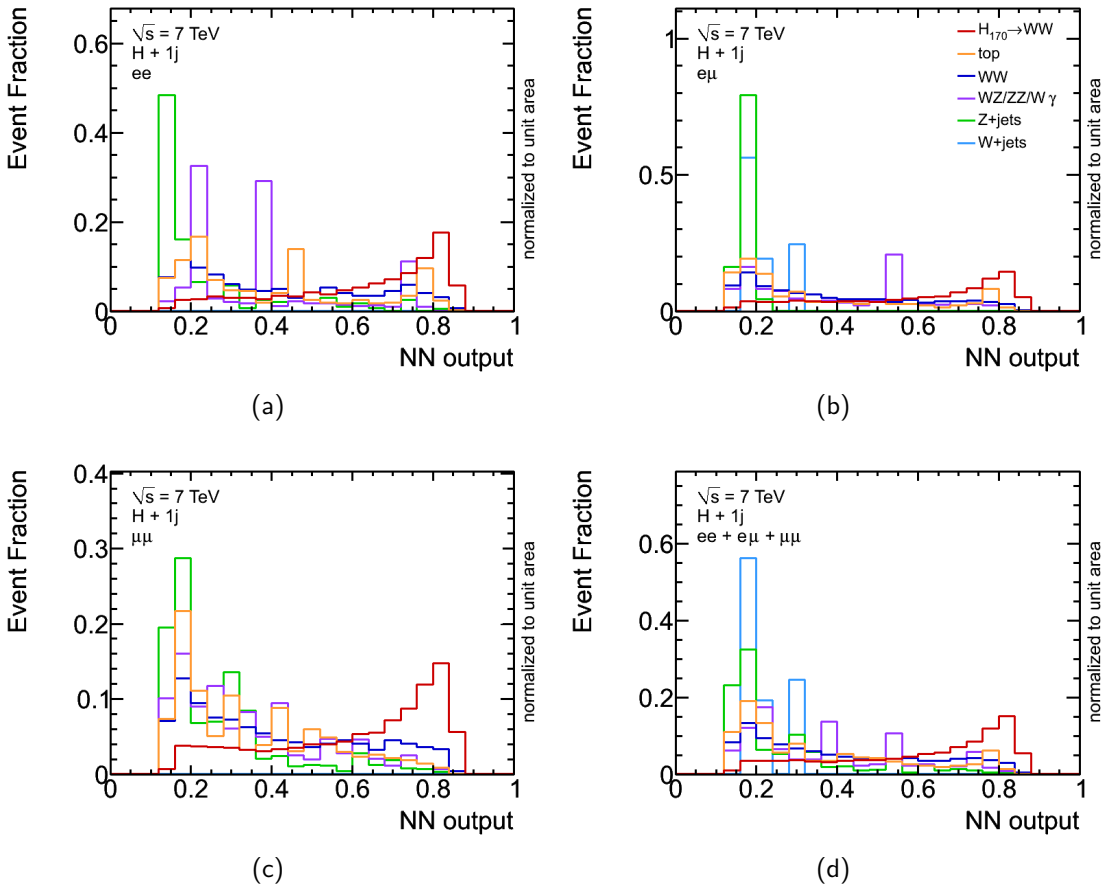


Figure 6.15.: *Template distributions of the NNoutput for the $H + 1j$ analysis and a Higgs boson mass of 170 GeV/c^2. The distributions are shown separately for the three decay channels: ee (a), eμ (b), and μμ (c). The sum of all three decay channels is shown as well (d).*

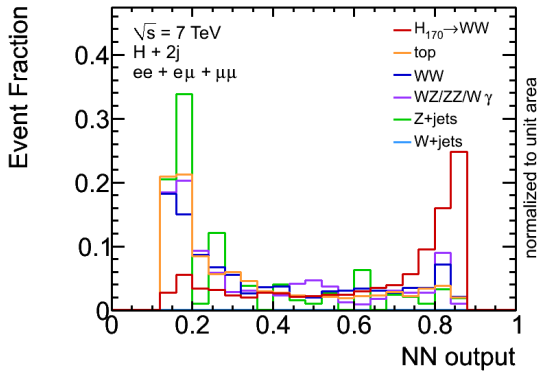


Figure 6.16.: *Template distributions of the NNoutput for the $H + 2j$ analysis and a Higgs boson mass of 170 GeV/c^2, showing the sum of all three decay channels.*

7. Statistical Analysis

In the analysis presented in this chapter the NNoutput template distributions are fitted to the NNoutput distribution observed in collision data. As the template distributions exhibit different shapes for the considered processes, it is possible to determine the fraction of each process in data. Especially the fraction of signal events and thus, the cross-section of the signal process may be determined, because the neural network classifiers were trained to separate the signal from the background processes. Statistical methods are deployed to test whether the distribution observed in data are described best by the signal-plus-background hypothesis or by the background-only hypothesis. After the description of the statistical methods and the systematic uncertainties the expected and observed significances of this analysis are computed. Finally, upper limits on the Higgs boson production cross-section are determined, since the significance is too low to establish the signal process.

7.1. Statistical Methods

A binned likelihood function is utilized to extract the contributions of the different processes, by fitting the NNoutput template distributions to the observed NNoutput distribution. The parametrization of the likelihood function contains the fraction of each process and is described in the first part of this section. The second part illustrates the generation of pseudo experiments that are used by the hypothesis testing procedure, which is introduced subsequently. The hypothesis test is carried out to compute the significance of the measurement extracting the signal contribution. It determines whether the selected collision events prefer the signal-plus-background hypothesis or the background-only hypothesis. The significance is a measure for the probability of statistical fluctuations of the background-only hypothesis to mimic the signal-plus-background hypothesis. A large value of the significance corresponds to a low probability of such a fluctuation. By convention, a significance of at least five standard deviations is required for an observation, while a significance of at least three standard deviations indicates an evidence. Results with a lower significance are usually used to set an upper limit on the production cross-section of the considered signal process. The limit-setting procedure applied in this analysis is described in the last part of this section.

7.1.1. Binned Likelihood Function

The binned likelihood function is built of two products. The first one consists of Poisson likelihoods $\mathcal{P}$ characterizing the content of each bin in the NNoutput distribution, while the second product incorporates Gaussian priors $\mathcal{N}$ as constraints to restrict the acceptance of the background processes.

$$L(\beta^{\mathrm{s}}, \beta_j^{\mathrm{b}}) = \prod_{k=1}^{M} \mathcal{P}(n_k; \mu_k) \cdot \prod_{j=1}^{B} \mathcal{N}(\beta_j^{\mathrm{b}}; 1, \Delta_j)$$

The scale factors β^{s} for the signal and β_j^{b} for the background processes are the parameters of the likelihood function that is maximized to fit the observed data. The index k runs over the number of bins M of the NNoutput distribution, while j is the index of the considered background processes B. The Poisson function $\mathcal{P}(n_k; \mu_k)$ for each bin k is given by

$$\mathcal{P}(n_k; \mu_k) = \frac{e^{-\mu_k} \cdot \mu_k^{n_k}}{n_k!}$$

where n_k is the number of observed events in bin k and μ_k is the mean value of the estimated number of expected events in this bin. Furthermore, the mean value μ_k is the sum of the estimated number of events of each individual physics process.

$$\mu_k = \mu_k^{\mathrm{s}} + \sum_{j=1}^{B} \mu_{jk}^{\mathrm{b}} \qquad \text{with} \quad \mu_k^{s} = \beta^{\mathrm{s}} \cdot \tilde{\nu}^{\mathrm{s}} \cdot \alpha_k^{\mathrm{s}} \quad \text{and} \quad \mu_{jk}^{\mathrm{b}} = \beta_j^{\mathrm{b}} \cdot \tilde{\nu}_j^{\mathrm{b}} \cdot \alpha_{jk}^{\mathrm{b}}$$

Here $\tilde{\nu}^{\mathrm{s}}$ and $\tilde{\nu}_j^{\mathrm{b}}$ denote the predicted number of signal and background events in the selected data set, respectively. The relative fraction of signal events in bin k is given by α_k^{s}, while α_{jk}^{b} represents the relative fraction of events in bin k, that are expected to stem from background process j. The set of α_{jk}^{b} fulfills the normalization condition $\sum_{k=1}^{M} \alpha_{jk} = 1$ as does α_k^{s}. α_k^{s} and α_{jk}^{b} represent the shape of the signal and the shape of the background process j, respectively. They are referred to as template distributions in the following. Thus, β^{s} and β_j^{b} represent the ratio of the measured and the estimated numbers of events of the signal and background processes. As the number of observed events depends linearly on the production cross-section of the considered processes, β^{s} and β_j^{b} can also be interpreted as the ratio of the measured and the predicted cross-section of the respective process.

Finally, the application of Gaussian constraints implements the a priori knowledge on the background processes,

$$\mathcal{N}(\beta_j^{\mathrm{b}}; 1, \Delta_j^2) = \frac{1}{\sqrt{2\pi\Delta_j^2}} \exp\left\{\frac{-(\beta_j^{\mathrm{b}} - 1)^2}{2\Delta_j^2}\right\}$$

with a mean of one and a width of Δ_j, which is the relative uncertainty on the cross-section of each considered background process. This prevents the rates of the background processes from floating freely and hence, improves the robustness of the fit. The cross-section uncertainties of the background processes are either determined by specific studies of control regions in collision data or given by theoretical calculations.

7.1.2. Pseudo Experiments

A frequentist approach based on pseudo experiments is deployed in order to determine the expected significance of the analysis and the impact of its systematic uncertainties [139]. The frequentist interpretation of a measurement requires the specification of an ensemble implementing an independent replication of the same experiment. Such an ensemble can be simulated by running Monte Carlo pseudo experiments. A pseudo experiment represents a possible outcome of a measurement given the expected statistical fluctuations and systematic uncertainties and assuming a certain physics model.

Systematic uncertainties cause variations on the signal acceptance, the background rates and the shape of the NNoutput distributions of each process. These effects are taken into account in the generation of each pseudo experiment. Four different types of uncertainties are considered for each process: the uncertainty on the cross-section, the statistical uncertainty due to the limited amount of simulated Monte Carlo events, and finally acceptance uncertainties and shape uncertainties, which may arise from different systematic sources.

Cross-section Uncertainties

The expectation values of all processes are varied within their cross-section uncertainties. This is done by throwing a random number β^{gen} for each process according to the Log-normal distribution:

$$\ln \mathcal{N}(x; \mu, \sigma^2) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-(\ln x - \mu)^2}{2\sigma^2}\right\} .$$

The Log-normal distribution is favored over the Gaussian distribution as it is not defined for negative values of x which would result in non-physical behavior. The parameters μ and σ are chosen in such a way that the expectation value $E[x]$ and the variance $V[x]$ of the Log-normal distribution correspond to the expectation value and variance of a Gaussian distribution with mean one and width Δ_j. Again, Δ_j denotes the relative cross-section uncertainty of process j. The expectation and

variance are then given by:

$$E[x] = \exp\left\{\mu + \frac{1}{2}\sigma^2\right\} \stackrel{!}{=} 1 \qquad \text{and} \qquad V[x] = \exp\left\{2\mu + \sigma^2\right\}\left(\exp\left\{\sigma^2\right\} - 1\right) \stackrel{!}{=} \Delta_j^2$$

which yields to

$$\mu = -\frac{1}{2}\sigma^2 \qquad \text{and} \qquad \sigma^2 = \ln(\Delta_j^2 + 1)\,.$$

Monte Carlo Statistics Uncertainties

The uncertainty due to the limited amount of simulated events is taken into account by morphing the template distributions α_{jk}. Each bin entry n_k is varied within its statistical uncertainty given by the square root of the bin entry $\sqrt{n_k}$. This is realized by resetting the bin entry with a random number drawn from a Gaussian distribution centered at the original value n_k. The statistical uncertainty $\sqrt{n_k}$ of that bin is used as the standard deviation of the Gaussian distribution.

Acceptance Uncertainties

The signed acceptance uncertainties ϵ_{ij+} and ϵ_{ij-}, of each process j and systematic uncertainty i, are included by varying the expectation values of each process. Technically, this is done by throwing a Gaussian distributed random number δ_i for each systematic uncertainty using a mean of zero and a width of one. The expectation value for each background process in a pseudo experiment is then given by

$$\nu_j^{\mathrm{b,\,gen}} = \tilde{\nu}_j^{\mathrm{b}} \cdot \beta_j^{\mathrm{gen}} \left\{1 + \sum_{i=1}^{S} |\delta_i| \left(\mathcal{H}(\delta_i)\epsilon_{ij+} + \mathcal{H}(-\delta_i)\epsilon_{ij-}\right)\right\}\,,$$

where S denotes the number of considered systematic uncertainties. The Heaviside function $\mathcal{H}$ is used to decide whether to implement an upward or downward fluctuation of systematic i. The absolute value $|\delta_i|$ implements the impact of the systematic uncertainty i. Thus, the number of background events n in a pseudo experiment is given by a sum of j numbers $n_j^{\mathrm{b,\,gen}}$ randomly drawn from Poisson distributions with means $\nu_j^{\mathrm{b,\,gen}}$. The same procedure is used to account for the acceptance uncertainties on the signal process.

Shape Uncertainties

The template distributions α_{jk}, that are obtained by evaluating the event classifiers on the nominal Monte Carlo samples, and systematically altered template distributions α_{ijk}^{+} and α_{ijk}^{-}, obtained by evaluating the event classifiers on each of the i

systematically altered Monte Carlo samples, are utilized to incorporate the uncertainties on the shapes of the NNoutput distributions. A new template distribution α_{jk}^{gen} is generated in each pseudo experiment by interpolating linearly between the nominal template distribution α_{jk} and the systematically altered template distributions α_{ijk}^{+} and α_{ijk}^{-}. Again, the systematic nuisance parameter δ_i is used as a weight, to ensure full correlation between the acceptance variation and the shape variation caused by the systematic uncertainty i. Thus, the generated template distribution $\alpha_{jk}^{\text{b, gen}}$ for each background process j in a pseudo experiment is given by

$$\alpha_{jk}^{\text{b, gen}} = \alpha_{jk}^{\text{b}} + \sum_{i=1}^{S} |\delta_i| \left\{ \mathcal{H}(\delta_i)(\alpha_{ijk}^{+} - \alpha_{jk}^{\text{b}}) + \mathcal{H}(-\delta_i)(\alpha_{ijk}^{-} - \alpha_{jk}^{\text{b}}) \right\} .$$

The template distribution $\alpha_{k}^{\text{s, gen}}$ is obtained likewise, using the corresponding signal templates α_{k}^{s}, $\alpha_{ik}^{\text{s},+}$ and $\alpha_{ik}^{\text{s},-}$.

Pseudo Experiment Generation

Finally, a pseudo experiment is generated applying the recipes just described. First one determines the number of events n^{gen} to be generated for each process. This number incorporates the uncertainty on the production cross-section and acceptance effects due to all considered sources of systematic uncertainties. In the second step the altered template distributions α_{k}^{gen} to be used in this pseudo experiment are generated. This template distributions incorporate the Monte Carlo statistics uncertainties and shape uncertainties induced by the considered systematic sources. The NNoutput distribution $\mathcal{A}_{k}^{\text{gen}}$ for each process is compiled by drawing n^{gen} from the altered template distributions α_{k}^{gen}. This results in a generated NNoutput distribution for the signal process $\mathcal{A}_{k}^{\text{s, gen}}$ and each considered background process $\mathcal{A}_{jk}^{\text{b, gen}}$. The sum of all generated distributions

$$\mathcal{A}_{k}^{\text{gen}} = \mathcal{A}_{k}^{\text{s, gen}} + \sum_{j=1}^{B} \mathcal{A}_{jk}^{\text{b, gen}}$$

represents a pseudo experiment, which simulates a possible outcome of the experiment. The sum $\sum_{k=1}^{M} \mathcal{A}_{k}^{\text{gen}}$ over the bins k gives the number of generated candidate events of this pseudo experiment. Generating many pseudo experiments leads to an ensemble of independent replications of the experiment.

7.1.3. Hypothesis Testing

The hypothesis test is performed by considering the background-only hypothesis H_0 and the signal-plus-background hypothesis H_1. H_0 assumes that there is no signal process $\beta^s = 0$, while H_1 includes the signal process $\beta^s = 1$ with the cross-section

predicted by the Standard Model. The aim of the hypothesis test is to reject one of the considered hypotheses taking the measured data as a basis. The Neyman-Pearson lemma [140] states that the likelihood ratio is the most powerful test to distinguish between the two hypotheses. Thus, the Q-value test statistic defined as:

$$Q = -2 \ln \frac{L(\beta^{\mathrm{s}} = 1, \hat{\beta}_j^{\mathrm{b}})}{L(\beta^{\mathrm{s}} = 0, \hat{\hat{\beta}}_j^{\mathrm{b}})},$$

is deployed to carry out the hypothesis test.

Two ensembles of pseudo experiments are generated. The first ensemble implementing the H_0 hypothesis $\mathcal{A}_k^{\mathrm{gen}} = \sum_{j=1}^{B} \mathcal{A}_{jk}^{\mathrm{b,\,gen}}$, while the second ensemble implements the H_1 hypothesis $\mathcal{A}_k^{\mathrm{gen}} = \mathcal{A}_k^{\mathrm{s,\,gen}} + \sum_{j=1}^{B} \mathcal{A}_{jk}^{\mathrm{b,\,gen}}$. The Q-value is computed for each pseudo experiment of the two ensembles, leading to two Q-value distributions q_0 and q_1, respectively. The $\hat{q}_0$ distribution denotes the q_0 distribution normalized to unit area. It is the probability density distribution given the H_0 hypothesis, while $\hat{q}_1$ is the probability density distribution given the H_1 hypothesis. The observed Q-value is computed applying the test statistic of the observed collision data. The idea of the test is to compare the observed Q-value with the probability densities $\hat{q}_0$ and $\hat{q}_1$ in order to determine whether the observed data prefer the H_0 or the H_1 hypothesis. For this purpose the overlap of the $\hat{q}_0$ and $\hat{q}_1$ distributions is crucial. A small overlap of the probability densities allows a clear distinction between the two hypotheses H_0 and H_1 and thus, indicates a large sensitivity of the analysis. Figure 7.1(a) depicts example distributions for $\hat{q}_0$ and $\hat{q}_1$ and the expected p-value.

The p-value is defined as

$$p(Q) = \int_{-\infty}^{Q} \hat{q}_0(Q')\,\mathrm{d}Q' = \frac{1}{A_{q_0}} \int_{\infty}^{Q} q_0(Q')\,\mathrm{d}Q'$$
$$\text{with} \qquad A_{q_0} = \int_{-\infty}^{\infty} q_0(Q')\,\mathrm{d}Q'.$$

For the expected p-value p^{exp} the upper boundary of the integral is given by the median of the $\hat{q}_1$ distribution. This means, given the hypothesis H_1 one expects a probability of 50% to observe a p-value that is lower than p^{exp}. The evaluation of the likelihood ratio on collision data determines the observed Q-value Q^{obs}, which is used to compute the observed p-value p^{obs}, illustrated in Figure 7.1(b).

Given the background-only hypothesis H_0, the observed p-value denotes the probability to obtain a result as compatible or less with H_0 than the one actually observed. The observed p-value is a measure for the observed significance.

Usually the significance of a measurement is expressed in terms of standard Gaussian deviations σ. Therefore, a one sided integral of a standardized Gaussian distribution is used to convert the p-value. A p-value less than 1.35×10^{-3}, which corresponds to 3σ, indicates an evidence. An observation is claimed for

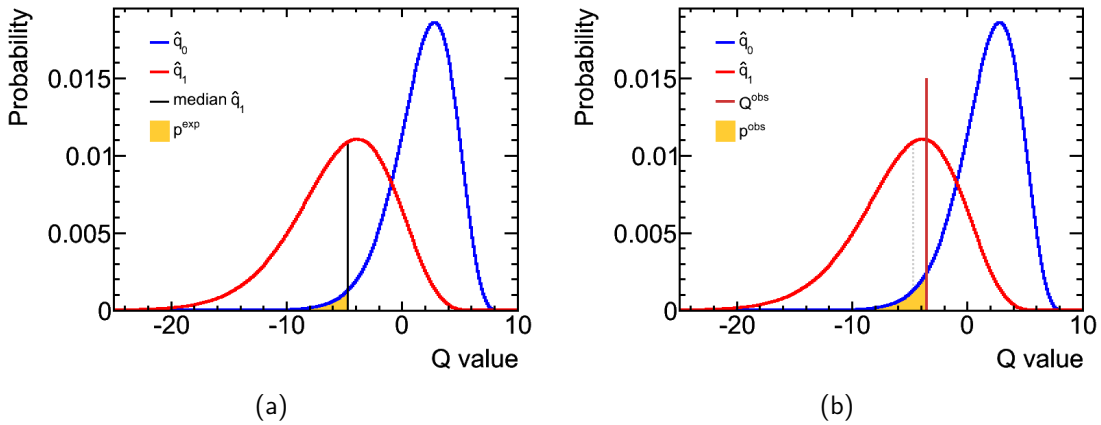


Figure 7.1.: *Example distributions of $\hat{q}_0$, representing the probability to observe a Q-value given the background-only hypothesis H_0, and $\hat{q}_1$, representing the probability to observe a Q-value given the signal-plus-background hypothesis H_1. The expected p-value p^{exp} is visualized in (a) while the observed Q-value Q^{obs} and the observed p-value p^{obs} are shown in (b).*

a p-value lower than 2.87×10^{-7}, that corresponds to a significance of at least 5σ. Hence, assuming the signal-plus-background hypothesis H_1 the observation of that many candidate events corresponds to a background fluctuation of 3σ or 5σ, respectively.

7.1.4. Cross-section Limits

Usually, an upper limit on the production cross-section of the signal process is constructed if the hypothesis test reveals a low significance. Figure 7.2 demonstrates the procedure of setting limits on the basis of the Q-value distributions. In such cases the overlap of the $\hat{q}_1$ distribution, assuming the signal process with the Standard Model production cross-section, and the $\hat{q}_0$ distribution, assuming no signal process at all, is too large. This leads to an expected p-value much larger than 1.35×10^{-3}. The expected p-value is referred to as p_b^{obs} in Figure 7.2(a). It means, that the probability of an upward fluctuation of the background processes is substantially high. Thus, one computes the probability p_{s+b} to observe a Q-value consistent with the background-only hypothesis, given the signal-plus-background hypothesis. The p_{s+b} is computed using the $\hat{q}_1$ distribution, see Figure 7.2(b).

The expected p_{s+b}-value p_{s+b}^{exp} is calculated by deploying the median of the $\hat{q}_0$ distribution as the lower bound of the integral. Given the hypothesis H_0, p_{s+b} denotes the probability of an downward fluctuation of the signal process.

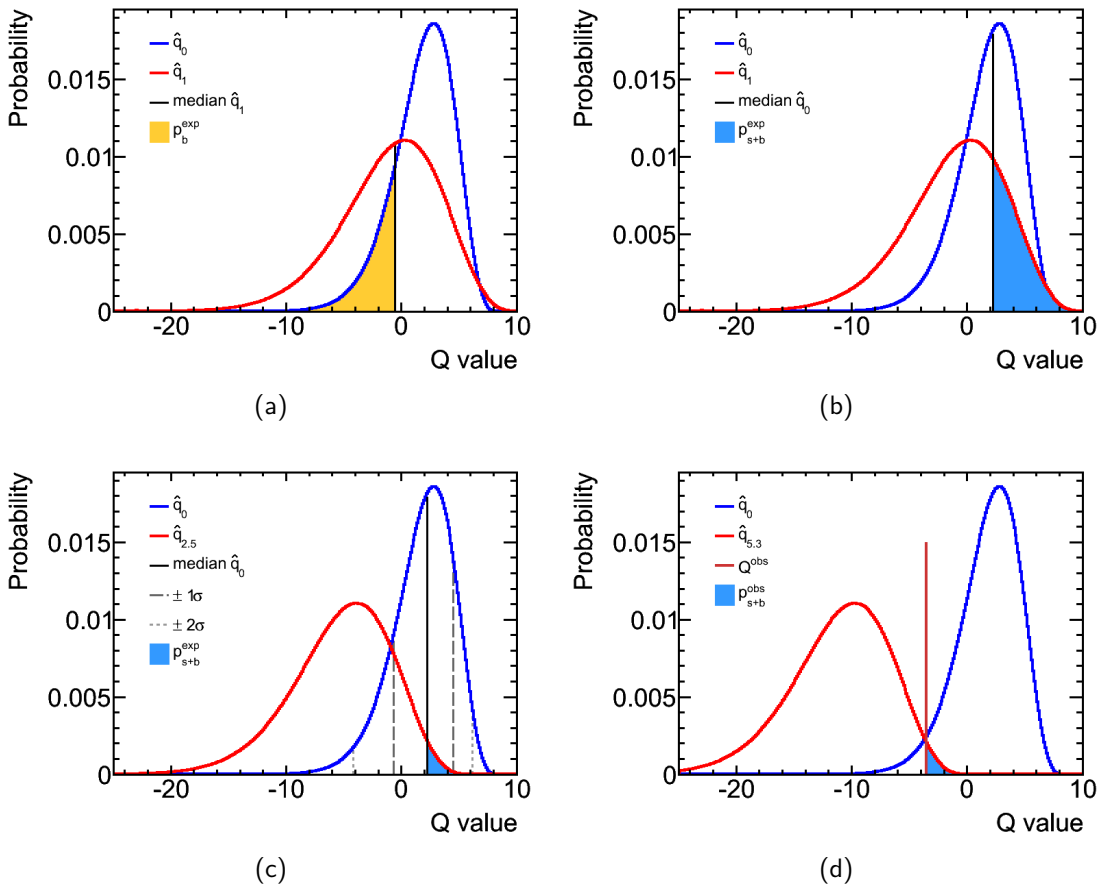


Figure 7.2.: *Demonstration of the limit setting procedure on the basis of the Q-value distributions. A low expected sensitivity leads to a rather large expected p-value (a). Definition of the expected p_{s+b} value (b) that is deployed to compute the expected cross-section limit and its $\pm 1\sigma$ and $\pm 2\sigma$ deviations (c). Therefore, one generates new ensembles of pseudo experiments implementing a larger production cross-section for the signal process until p_{s+b}^{exp} is found to be 0.05. The computation of the observed cross-section limit is visualized in (d) utilizing the observed Q-value as the lower bound for the integral in the p_{s+b}^{obs}-value computation.*

$$p_{s+b}(Q) = \int_Q^\infty \hat{q}_1(Q')\,\mathrm{d}Q' = \frac{1}{A_{q_1}} \int_Q^\infty q_1(Q')\,\mathrm{d}Q'$$
$$\text{with} \qquad A_{q_1} = \int_{-\infty}^\infty q_1(Q')\,\mathrm{d}Q'$$

A confidence level (CL) of 95% is used to set an upper limit on the production cross-section of the signal process. The confidence level is given by $1 - p_{s+b}$, which means that the probability of a downward fluctuation of the signal process is 5%. Figure 7.2(b) depicts a usual case: given the SM production cross-section the p_{s+b}^{exp} value is larger than 5%. Hence, a new ensemble of pseudo experiments, implementing a larger cross-section of the signal process e.g. 2.5 times the SM cross-section, is generated. This new ensemble represents an alternative hypothesis $H_{2.5}$. The Q-value test statistic is applied to obtain the $\hat{q}_{2.5}$ distribution, which is shifted to lower Q-values, see Figure 7.2(c). One computes the new p_{s+b}^{exp} value and the corresponding confidence level using the $\hat{q}_{2.5}$ distribution. This procedure of generating new ensembles of pseudo experiments with larger production cross-sections is repeated, until the expected confidence level reaches 95%. Thus, the production cross-section of the last ensemble denotes the expected limit. The found cross-section and larger ones are excluded at 95% CL.

The same procedure is utilized to generate the one and two sigma errors on the expected limit. The only change is the lower bound of the p_{s+b} integral, using the $\pm 1\sigma$ and $\pm 2\sigma$ standard Gaussian quantiles of the $\hat{q}_0$ distribution instead of its median. The quantiles are indicated as dashed lines in Figure 7.2(c). Finally, the described procedure is also used to compute the observed limit at the 95% CL using the observed Q-value Figure 7.2(d). The obtained limits are called $\mathrm{CL_{s+b}}$. This method of confidence intervals was first developed by Neyman [141]. Further examples can be found in Reference [142].

The $\mathrm{CL_{s+b}}$ method works fine for situations with a large sensitivity level, where the overlap of the considered probability density distributions $\hat{q}_0$ and $\hat{q}_{5.3}$ is rather large, as depicted in Figure 7.2(d). The $H_{5.3}$ hypothesis is rejected or excluded with a confidence level of 95%. Whereas $\mathrm{CL_{s+b}}$ may lead to spurious exclusions in situations where there is only a very small level of sensitivity. Figure 7.3(a) presents such a situation where the probability density distributions $\hat{q}_0$ and $\hat{q}_1$ almost coincide. The H_1 hypothesis is excluded. However, the exclusion is only possible due to a statistical fluctuation of the observed data, which leads to an anomalously strong limit.

To overcome this limitation of $\mathrm{CL_{s+b}}$, a procedure called $\mathrm{CL_s}$ [144, 145] was introduced by LEP experiments.

$$\mathrm{CL_s} = \frac{p_{\mathrm{s+b}}}{1 - p_{\mathrm{b}}}$$

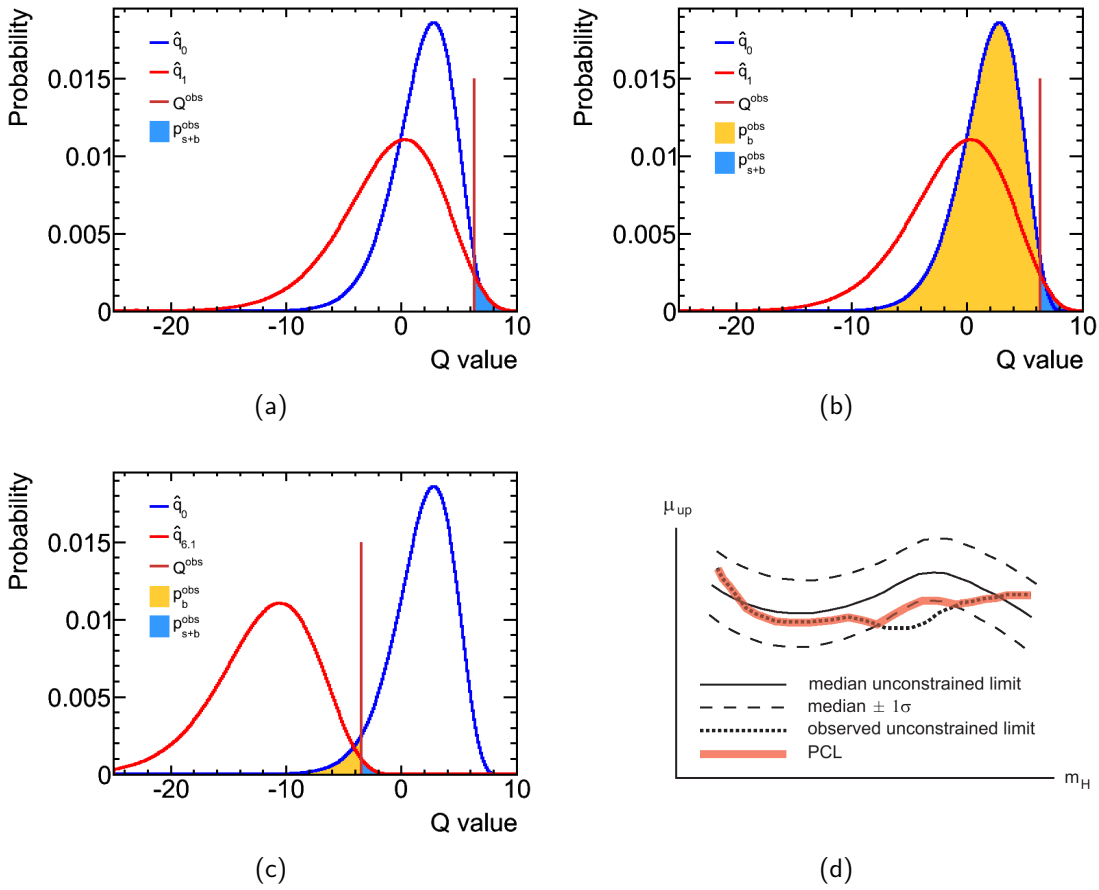


Figure 7.3.: *Demonstration of the limit setting procedure on the basis of the Q-value distributions. Figure (a) depicts a situation where the CL_{s+b} limit procedure leads to a spurious exclusion. The CL_s procedure overcomes the limitation of CL_{s+b} using the fraction $p_{s+b}/1-p_b$ of probabilities, to incorporate the sensitivity of the measurement (b). By definition CL_s leads to more conservative limits (c) compared to CL_{s+b} shown in Figure 7.2(d). The Power-Constrained Limit utilizes the CL_{s+b} procedure to derive the expected limit (d). Additionally, a constraint is applied to the observed limit in situations where the observed limit fluctuates below a predefined lower bound of the expected limit, -1σ it this case [143].*

In the CL_s procedure H_1 is excluded if one finds $CL_s < 0.05$. Because of the ratio the CL_s method avoids spurious exclusions, see Figure 7.3(b). The CL_s value is always larger than p_{s+b}. Thus, the probability to exclude H_1 is less than 0.05%. Upper limits derived with the CL_s procedure are always larger than upper limits derived with CL_{s+b}, compare Figure 7.3(c) and 7.2(d). Therefore, the CL_s procedure is said to be conservative. The upper limits on the Higgs production cross-section reported in the first chapter are computed with the CL_s method. There is a point of criticism on the CL_s procedure: it does not provide a probability because it is by definition a ratio of probabilities.

This analysis implements Power-Constrained Limits (PCL) [138, 143] proposed by the ATLAS statistics committee. PCL is a method for setting limits that avoids excluding parameter values for which the sensitivity falls below a specific predefined threshold. For this example the threshold is the -1σ deviation of the expected limit. In a first step the expected limits are computed according to the CL_{s+b} procedure. In the second step the sensitivity or power of the measurement is tested by comparing the observed Q-value Q^{obs} with the probability density $\hat{q}_0$ of the background-only hypothesis. If Q^{obs} is above the +1σ quantile of the $\hat{q}_0$ distribution, the sensitivity of the measurement is deemed to be to low. Hence, the constraint is applied by quoting the expected CL_{s+b} limit instead. This is equivalent to the Q^{obs} value being found at the +1σ quantile of the $\hat{q}_0$ distribution. An observed Q-value above the +1σ quantile of the $\hat{q}_0$ distribution corresponds to a downward fluctuation of the observed limit below -1σ of the expected limit. If Q^{obs} is below the +1σ quantile one computes the observed limit according to the CL_{s+b} procedure.

Figure 7.3(d) visualizes the application of the PCL. The power constraint is applied if the unconstrained observed limit fluctuates one standard deviation below the expected or median unconstrained limit. The latest recommendations of the ATLAS statistics committee are to apply the power constraint if the unconstrained observed limit fluctuates below the expected unconstrained limit and to report the cases where the power constraint was applied.

7.2. Systematic Uncertainties

Systematic uncertainties may be caused by uncertainties in the modeling of physics processes and detector effects. The various systematic sources may affect the acceptance of the selected signal and background processes and alter the shape of the observables that are used by the candidate event classifiers. Thus, altering the shapes of the NNoutput template distributions utilized in the statistical analysis. Both effects are considered in the analysis as described in Section 7.1.2. This section describes all considered sources of systematic uncertainties and quantifies their

impact on the selection of the different processes and the shape of the template distributions.

For each systematic source under consideration, the affected objects in the Monte Carlo samples are varied up and down according to the uncertainty of that source. The size of the systematic uncertainty is computed for each decay channel. It is given by the difference between the event yield in the object-varied samples and the event yield in the nominal sample, after all candidate event selection criteria have been applied. The neural network classifier trained on the nominal sample is evaluated on the object-varied samples to obtain the altered template distribution for each systematic source.

Jet Energy Scale (JES)

The calibration of jets in the first data recorded by the ATLAS experiment is done by applying jet-by-jet corrections as a function of jet p_{T} and η. This calibration scheme allows a direct evaluation of the systematic uncertainties [104, 106, 107, 146, 147]. The JES uncertainty is estimated using a combination of in-situ techniques and systematic variations of Monte Carlo simulations. For central jets ($|\eta| < 0.8$) it is found to be lower than 6.5% for $p_{\mathrm{T}} < 60$ GeV/c and decreases to 4% for $p_{\mathrm{T}} > 200$ GeV/c. In the endcap ($0.8 < |\eta| < 3.2$) and forward region ($3.2 < |\eta| < 4.5$) the uncertainty for jets with $p_{\mathrm{T}} < 60$ GeV/c amounts to about 9% and 15% respectively. Jets with $p_{\mathrm{T}} > 200$ GeV/c in the endcap region have an uncertainty smaller than 4.5%. Additional uncertainty contributions due to multiple proton-proton interactions are also included as separate contributions to the systematic uncertainty and are evaluated as a function of the number of primary vertices in the event. For events with two primary vertices the contributions are less than 2% in the central region and less than 16% in the endcap and forward region. They become negligible for jets with $p_{\mathrm{T}} > 200$ GeV/c and $|\eta| < 2.8$. The energy and momentum of all jets in the Monte Carlo samples are scaled up or down accordingly to the provided values before applying the jet selection criteria.

Jet Energy Resolution (JER)

The jet resolution is measured utilizing two different methods: the di-jet balance and the bi-sector techniques [148, 149]. The systematic uncertainties associated with the in situ methods are 7% at $p_{\mathrm{T}} > 30$ GeV/c and 4% at $p_{\mathrm{T}} > 500$ GeV/c. The Monte Carlo simulation describes the jet energy resolution measured from data within 10% for jets with p_{T} values between 30 GeV/c and 500 GeV/c in the rapidity range $|\eta| < 2.8$. This systematic uncertainty is considered by varying the jets according to the output of the `JERProvider` tool provided by the ATLAS Jet and Missing Et Group.

Electron Energy Scale (EES)

The electron energy scale in data is calibrated using scale factors of 1% and 3% in the barrel and endcap regions respectively. These scale factors are determined using in-situ calibration of $Z \to ee$ decays. The systematic uncertainty associated with this calibration is considered by studying Monte Carlo samples modified with these values. In this case the "up" and "down" samples are generated by multiplying and dividing the electron energy by the η-dependent electron energy scale factors, respectively [96].

Electron Reconstruction and Identification efficiency (EID)

The efficiencies of the electron reconstruction and identification are different in Monte Carlo and data, such that a correction scale factor for each electron is taken into account as a factor in the event weight for each Monte Carlo event. The p_{T}- and η-dependent weights are obtained using the `EgammaSFClass` tool, which uses the results of the methods described in Reference [95]. This tool also provides the uncertainties associated to the scale factors used to generate the "up" and "down" variation samples by adding and subtracting the uncertainty on the scale factors.

Muon Reconstruction and Identification efficiency (MID)

As for electrons a corrective scale factor is applied to the event weight in Monte Carlo to account for differences between data and Monte Carlo in the reconstruction and identification of muons. A standard procedure using the width of the di-muon mass distribution as described in Reference [150] is used to derive the systematic uncertainty in this case. The "up" and "down" samples are generated similarly as in the EID case.

All these object-based systematic uncertainties affect the acceptance of the selected processes as well as the shape of the template distributions used to fit the observed data. The statistical uncertainty due to the limited amount of Monte Carlo events MCstat is considered in addition, it affects only the shapes of the template distributions. The last two sources of systematic uncertainties have an impact only on the predicted rate of the selected processes:

- the uncertainty on the cross-section as quoted in Tables 4.1 and 4.2, and
- the 11% uncertainty on the luminosity [64, 70].

Tables 7.1 to 7.3 denote the full details of the impact on the acceptance for all considered processes caused by the object-based systematic uncertainties.

Table 7.1.: *Systematic acceptance uncertainties in the zero-jet multiplicity bin for the Higgs boson mass point of* 170 GeV/c^2. *The relative uncertainties are given in %.*

(a) ee channel

Systematic Source		Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$
JES	up	+0.35	+4.84	+1.90	+1.08	+15.23	+0.00
	down	−0.84	+12.81	−2.88	−4.42	−14.46	−0.00
JER	up	−0.22	+0.00	−0.29	+0.92	+24.24	+0.00
	down	−0.07	−8.19	−0.30	+1.01	+24.24	−0.00
EES	up	+0.53	−2.45	−0.41	+2.54	+0.06	+1.36
	down	−0.86	−4.89	+0.52	+0.02	+20.89	−0.00
EID	up	+10.66	−23.31	+13.92	+22.37	+14.03	+19.51
	down	−10.14	+21.88	−13.08	−20.03	−13.18	−18.35

(b) $e\mu$ channel

Systematic Source		Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$
JES	up	+0.13	−0.60	+0.95	+5.28	+28.15	+0.00
	down	−0.11	−1.26	−1.19	−7.45	−13.12	−0.00
JER	up	+0.09	−0.63	−0.14	+4.70	−13.12	+0.00
	down	+0.02	+12.61	+0.01	−6.71	−0.00	−0.00
EES	up	+0.35	+0.74	−0.12	−0.13	−0.03	+0.76
	down	−0.52	+1.39	−0.10	−1.24	−0.00	−22.88
EID	up	+5.26	+6.42	+6.28	+10.23	+8.63	+13.23
	down	−5.26	−6.42	−6.28	−10.23	−8.63	−13.23
MID	up	+0.74	+0.52	+0.74	+0.78	+0.79	+0.70
	down	−0.74	−0.52	−0.74	−0.78	−0.79	−0.70

(c) $\mu\mu$ channel

Systematic Source		Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$
JES	up	−0.19	22.14	−0.07	+0.01	−9.04	+0.00
	down	+0.18	+0.08	−0.04	+0.43	+7.50	−0.00
JER	up	−0.01	−0.34	−0.07	+0.96	−1.24	+0.00
	down	−0.06	−0.26	−0.08	+0.47	−3.27	−0.00
MID	up	+1.49	+1.45	+1.50	+1.46	+1.50	+1.25
	down	−1.48	−1.44	−1.48	−1.45	−1.49	−1.25

Table 7.2.: *Systematic acceptance uncertainties in the one-jet multiplicity bin for the Higgs boson mass point of* 170 GeV/c^2. *The relative uncertainties are given in %.*

(a) ee channel

Systematic Source		Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$
JES	up	−0.56	−0.88	+1.47	+1.47	+16.70	+0.00
	down	−0.75	+1.12	−1.52	−14.36	−2.41	−0.00
JER	up	−0.72	+0.25	+0.95	−10.85	−0.01	+0.00
	down	+0.13	+1.50	−0.04	−0.35	+11.05	−0.00
EES	up	+0.24	−0.92	−0.83	−0.22	+0.57	+0.00
	down	−1.52	+1.39	+0.05	+1.22	+30.32	−0.00
EID	up	+11.04	+14.56	+13.35	+16.45	+25.45	+0.00
	down	−10.48	−13.73	−12.57	−15.37	−22.81	−0.00

(b) $e\mu$ channel

Systematic Source		Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$
JES	up	−1.17	−2.12	+0.49	+1.24	+57.71	+0.00
	down	+0.53	+2.19	−0.66	−0.24	−7.69	−0.00
JER	up	−0.39	−0.64	−0.13	−0.09	+6.30	+0.00
	down	−0.22	−0.46	−0.48	−0.66	+12.24	−0.00
EES	up	+0.44	+0.16	−0.11	−0.17	+14.13	+0.00
	down	−0.41	+0.02	−0.23	+0.02	−0.04	−0.00
EID	up	+5.34	+6.34	+6.39	+6.36	+7.77	+4.29
	down	−5.34	−6.34	−6.39	−6.36	−7.77	−4.29
MID	up	+0.74	+0.78	+0.75	+0.72	+0.80	+0.84
	down	−0.74	−0.78	−0.75	−0.72	−0.80	−0.84

(c) $\mu\mu$ channel

Systematic Source		Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$
JES	up	−1.77	−1.21	−0.38	+1.91	+6.60	+0.00
	down	+1.06	+1.71	+0.02	+0.58	+1.60	−0.00
JER	up	−0.37	+0.18	−0.09	+1.03	+8.36	+0.00
	down	−0.26	+0.74	+0.11	+2.43	+3.62	−0.00
MID	up	+1.48	+1.45	+1.50	+1.50	+1.44	+0.00
	down	−1.47	−1.44	−1.48	−1.49	−1.43	−0.00

Table 7.3.: *Systematic acceptance uncertainties in the two-jet multiplicity bin for the Higgs boson mass point of* 170 GeV/c^2. *All decay channels ee, eμ and μμ are combined. The relative uncertainties are given in %.*

Systematic Source		Signal	WW	W+jets	Z+jets	top	$WZ/ZZ/W\gamma$
JES	up	−0.52	+1.31	+1.09	+4.05	+147.29	+0.00
	down	−0.05	+0.75	−0.39	−6.40	−3.05	−0.00
JER	up	+0.28	−0.86	+0.64	−3.34	+29.73	+0.00
	down	−0.29	+2.06	−0.10	+0.58	+97.01	−0.00
EES	up	+0.05	−0.44	+0.11	+0.11	+0.10	+0.00
	down	−0.23	+0.01	−0.54	−0.62	−5.64	−0.00
EID	up	+4.40	+5.26	+5.10	+6.53	+5.88	+0.00
	down	−4.32	−5.13	−5.00	−6.33	−5.53	−0.00
MID	up	+0.87	+0.89	+0.89	+0.83	+0.99	+0.00
	down	−0.87	−0.89	−0.89	−0.82	−0.98	−0.00

The impact on the shapes of the template distributions is visualized in Figures 7.4 to 7.6 using the JES variation as an example. All tables and figures are ordered by jet multiplicity bins. Additionally, the tables are split up to the level of decay channels. The presented numbers and distributions cover only the Higgs boson mass point of 170 GeV/c^2. Corresponding tables covering the impact of systematic uncertainties for all considered mass points are documented in Reference [135] together with figures presenting the variations of the NNoutput template distributions for all systematic sources and decay channels.

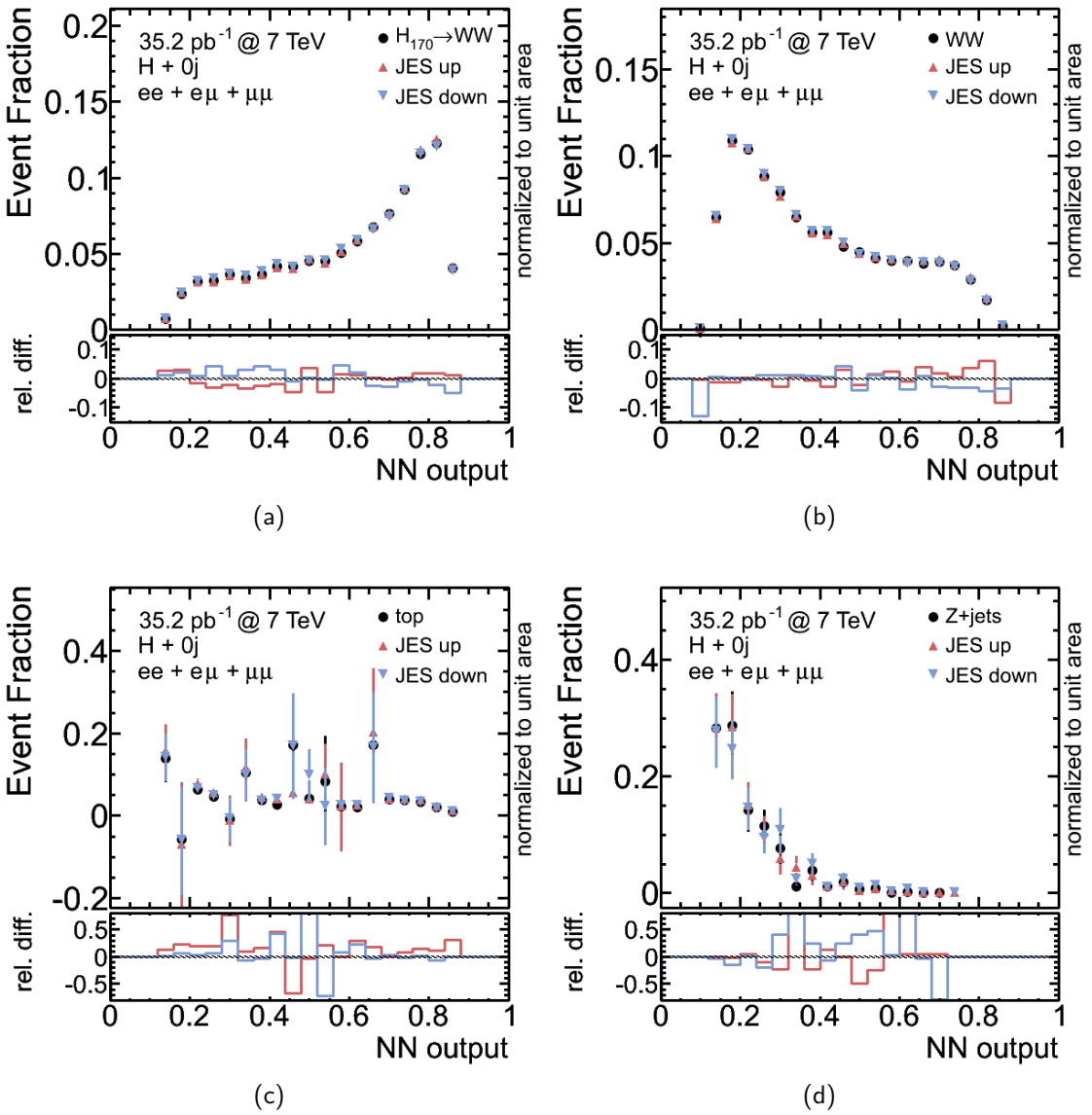


Figure 7.4.: *Systematic shape uncertainties of the NNoutput distributions in the $H + 0j$ analysis caused by the Jet Energy Scale (JES) uncertainty. The black markers represent the nominal NNoutput distribution of the event classifier trained to separate signal events with a Higgs boson mass of* 170 GeV/c^2. *The shape variation, represented by the red and blue markers, are obtained by evaluating the trained event classifiers on the systematically altered samples. Relative differences are visualized bin by bin. They are low for the Higgs signal (a) and the WW background (b). The available top-quark and Z+jets background samples exhibit a lack of statistics after the tight selection criteria. Hence, the bin by bin fluctuations are larger in Figures (c) and (d), respectively. Especially in the case of the top-quark background there is also a negative entry around 0.2 that is caused by a negative event weight of the Monte Carlo. This bin entry is reset to zero in the statistical analysis. Reference [135] provides further NNoutput distributions showing the impact of the systematic uncertainties in the different decay channels ee, eμ and μμ.*

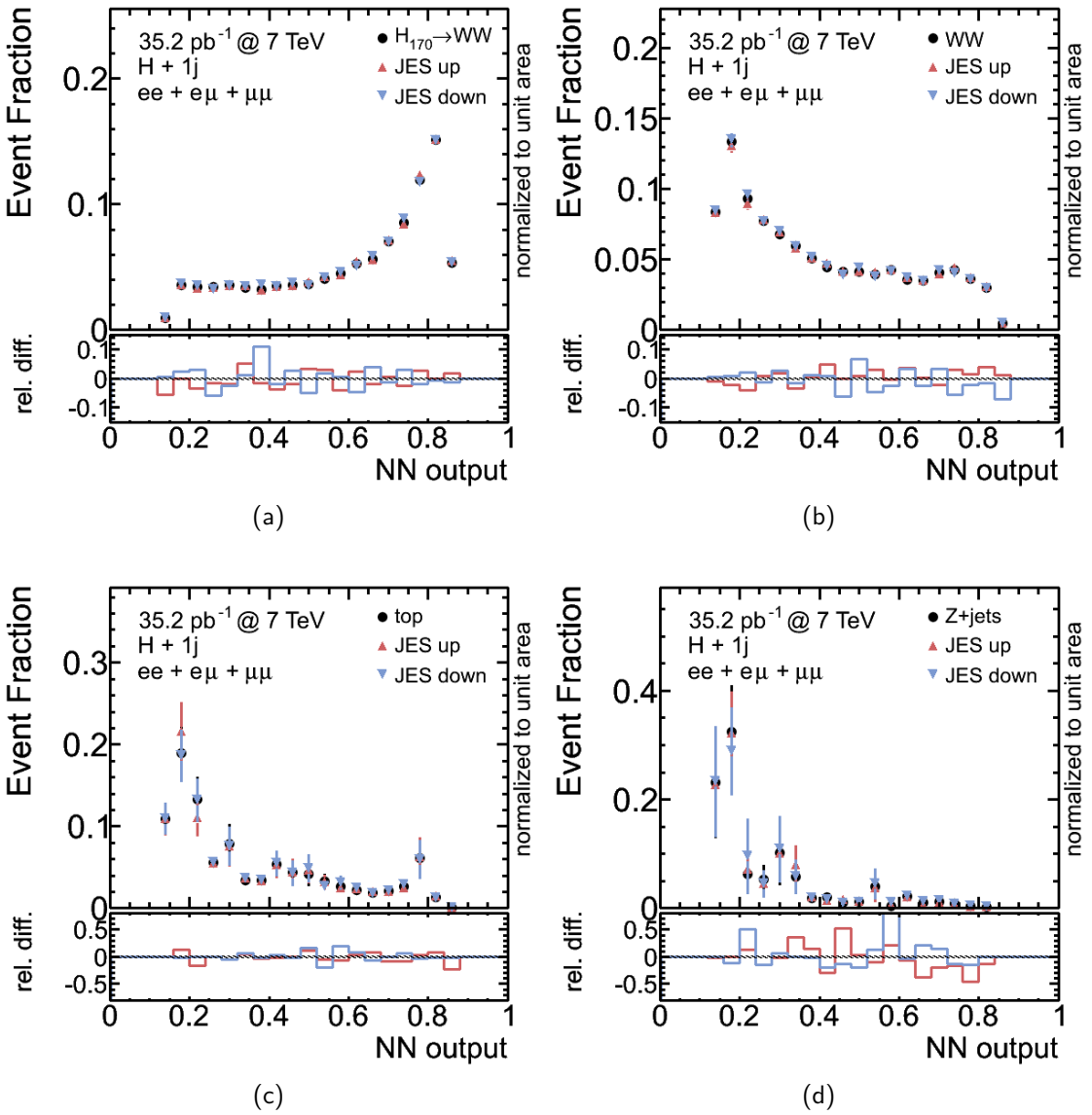


Figure 7.5.: *Systematic shape uncertainties of the NNoutput distributions in the $H + 1j$ analysis caused by the Jet Energy Scale (JES) uncertainty. The black markers represent the nominal NNoutput distribution of the event classifier trained to separate signal events with a Higgs boson mass of 170 GeV/c^2. The shape variation, represented by the red and blue markers, are obtained by evaluating the trained event classifiers on the systematically altered samples. Relative differences are visualized bin by bin. They are low for the Higgs signal (a) and the WW background (b). The available top-quark and Z+jets background samples exhibit a lack of statistics after the tight selection criteria. Hence, the bin by bin fluctuations are larger in Figures (c) and (d), respectively. The top-quark background becomes more important than in the $H + 0j$ analysis. Reference [135] provides further NNoutput distributions showing the impact of the systematic uncertainties in the different decay channels ee, eμ and μμ.*

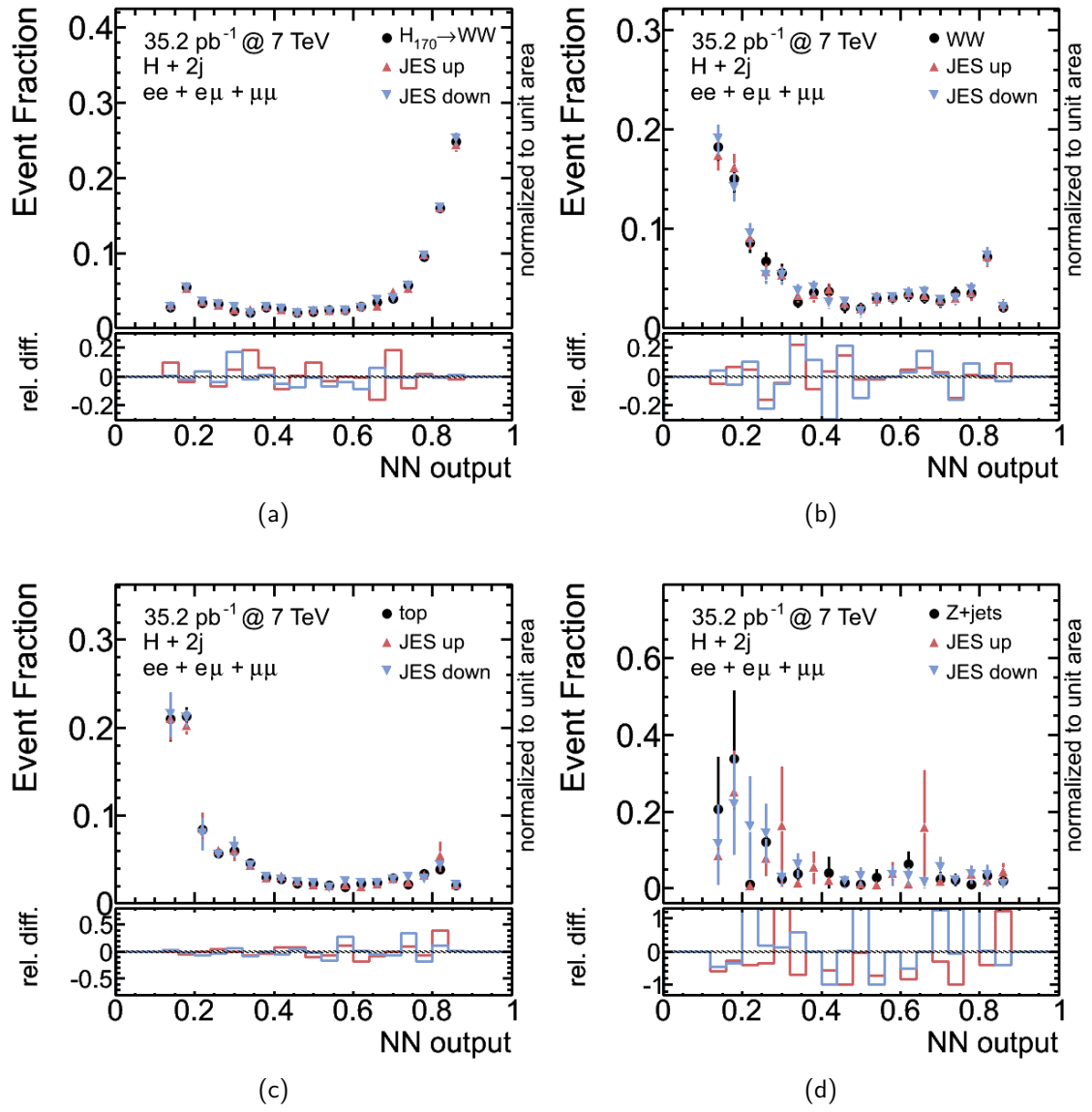


Figure 7.6.: *Systematic shape uncertainties of the NNoutput distributions in the $H + 2j$ analysis caused by the Jet Energy Scale (JES) uncertainty. The black markers represent the nominal NNoutput distribution of the event classifier trained to separate signal events with a Higgs boson mass of* 170 GeV/c^2. *The shape variation, represented by the red and blue markers, are obtained by evaluating the trained event classifiers on the systematically altered samples. Relative differences are visualized bin by bin. They are low for the Higgs signal (a) and the WW background (b). The top-quark background is the most important background in this jet multiplicity bin (c), while the Z+jets background is not (d). Reference [135] provides further NNoutput distributions showing the impact of the systematic uncertainties in the different decay channels ee, eμ and μμ.*

7.3. Results

This section presents the results of the statistical analysis, that is accomplished using the methods described in the first section of this chapter. First of all, pseudo experiments are conducted to evaluate the expected significance of the analysis. Two ensembles of pseudo experiments are generated, one implementing the signal-plus-background hypothesis and the other the background-only hypothesis. All pseudo experiments incorporate cross-section uncertainties, MC statistic uncertainties and acceptance and shape uncertainties due to the considered systematic sources. In each pseudo experiment the NNoutput templates are fitted to the generated pseudo data distributions to determine the cross-section observed in that pseudo experiment. These pseudo observations are used to calculate the expected p-value and the expected significance. The template fit is performed simultaneously in seven different channels, the Higgs boson decay channels ee, $e\mu$, and $\mu\mu$ in the zero-jet and one-jet multiplicity bin and the combined channel $ee + e\mu + \mu\mu$ in the two-jet multiplicity bin. The Q-value test statistic is deployed to carry out the hypothesis test utilizing the fit results.

After the computation of the expected p-values and the corresponding significances, the trained event classifiers are evaluated on 35.2 pb^{-1} of data recorded with the ATLAS experiment in 2010. Now the NNoutput templates are fitted to the observed collision data and the observed p-values are reported. Because no significant signal contribution is observed in data, upper limits on the Higgs boson production cross-section are set. This statistical analysis is performed for ten Higgs boson mass points in the range between 120 GeV/c^2 and 200 GeV/c^2, using the corresponding NNoutput templates of the 30 trained neural network event classifiers.

7.3.1. Significance of the Higgs Boson Search

Figure 7.7 depicts the probability density distributions $\hat{q}_0$ and $\hat{q}_1$, found performing pseudo experiments and deploying the Q-value test statistic for a Higgs boson mass of 170 GeV/c^2. It also depicts the expected p-value of $p^{exp} = 0.07$. This means, given the hypothesis H_1 one expects with a probability of 50% to observe at least that many candidate events, that the measured excess over the background corresponds to an upward fluctuation of the background of 1.48σ or more. Hence, assuming the SM production cross-section for the Higgs boson the discovery potential of the analysis at this mass point is not sufficient to expect an evidence of the signal process with the 2010 dataset.

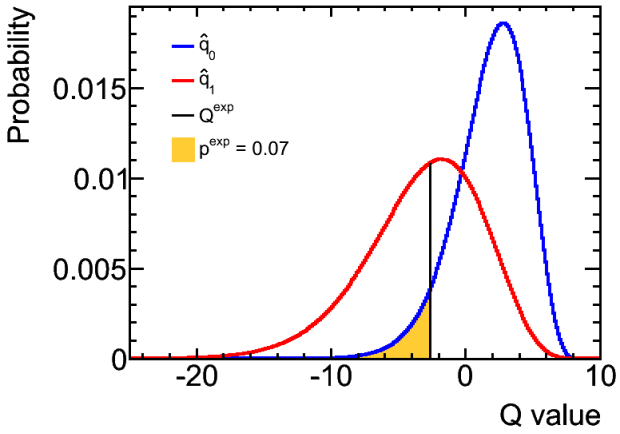


Figure 7.7.: *Probability density distributions found performing pseudo experiments and deploying the Q-value test statistic for the Higgs boson mass point of* 170 GeV/c^2 *and the expected p-value.*

The expected p-values for all considered Higgs boson masses are compiled in Table 7.4, including the corresponding significances in terms of standard Gaussian deviations. The highest expected significance is obtained in the mass range between 160 GeV/c^2 and 170 GeV/c^2. It is not possible to establish an evidence of the signal process with the production cross-section predicted by the SM. Thus, expected and observed upper limits on the Higgs boson cross-section will be computed after the evaluation of the trained event classifiers on observed collision data.

Figures 7.8, 7.9 and 7.10 present the classification of observed events including the expected NNoutput distributions for the $H + 0j$, $H + 1j$ and the $H + 2j$ analysis, respectively. The predicted fractions of the different processes are stacked on top of each other, such that the sum gives the total expectation for the observed amount of 35 pb^{-1} of data. Due to the low production cross-section of the Higgs boson process and the restrictive selection criteria the number of expected events is rather low. The agreement of data and Monte Carlo is reasonable with respect to the statistics.

Table 7.4.: *Expected p-values and the corresponding significances in terms of standard Gaussian deviations for all considered Higgs boson mass points. The highest expected significance is obtained in the mass range between* 160 GeV/c^2 *and* 170 GeV/c^2.

M_H [GeV/c^2]	120	130	140	150	160	165	170	180	190	200
exp. p-value	0.43	0.35	0.25	0.16	0.06	0.05	0.07	0.12	0.22	0.28
significance [σ]	0.18	0.39	0.67	0.99	1.55	1.64	1.48	1.17	0.77	0.58

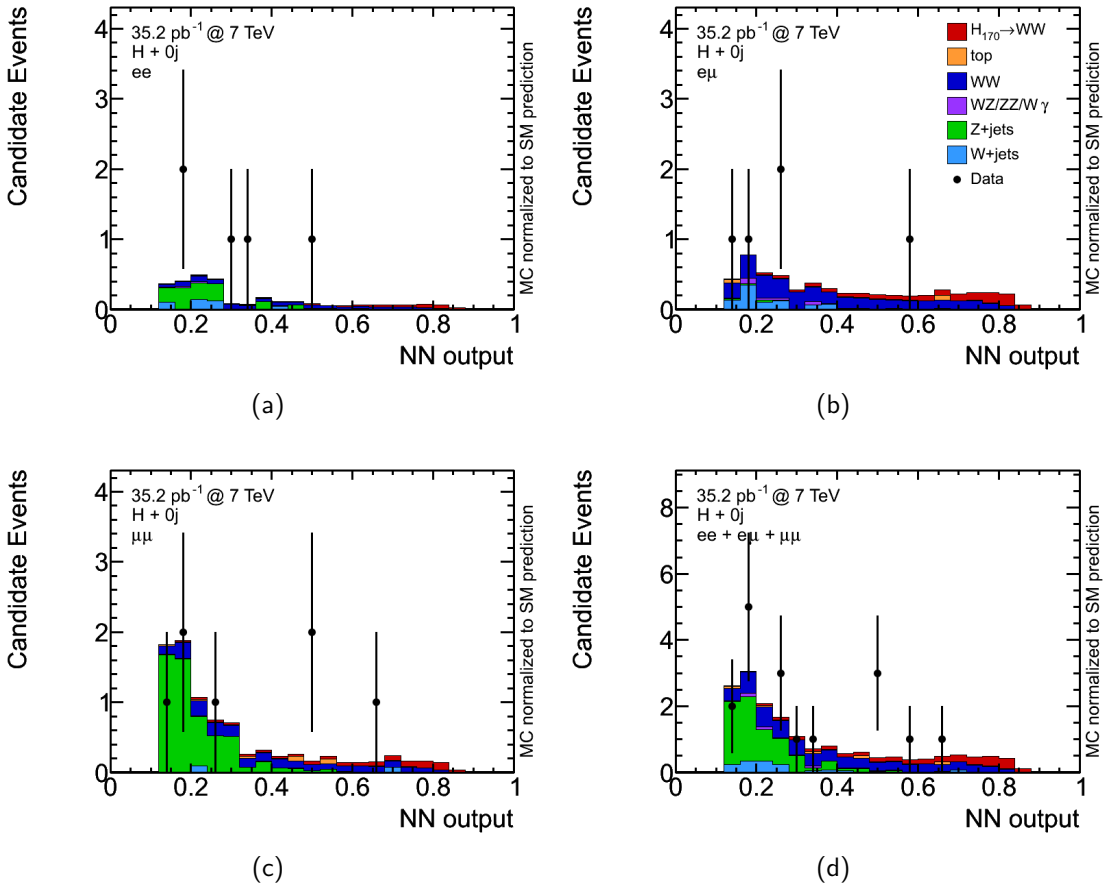


Figure 7.8.: *Observed distributions of the NNoutput for the $H + 0j$ analysis in 35 pb^{-1} of data. The predicted fractions of the different processes are stacked on top of each other, such that the sum gives the total expectation. The distributions are shown separately for the three decay channels: ee (a), eμ (b), and μμ (c). The sum of all three decay channels is shown as well (d).*

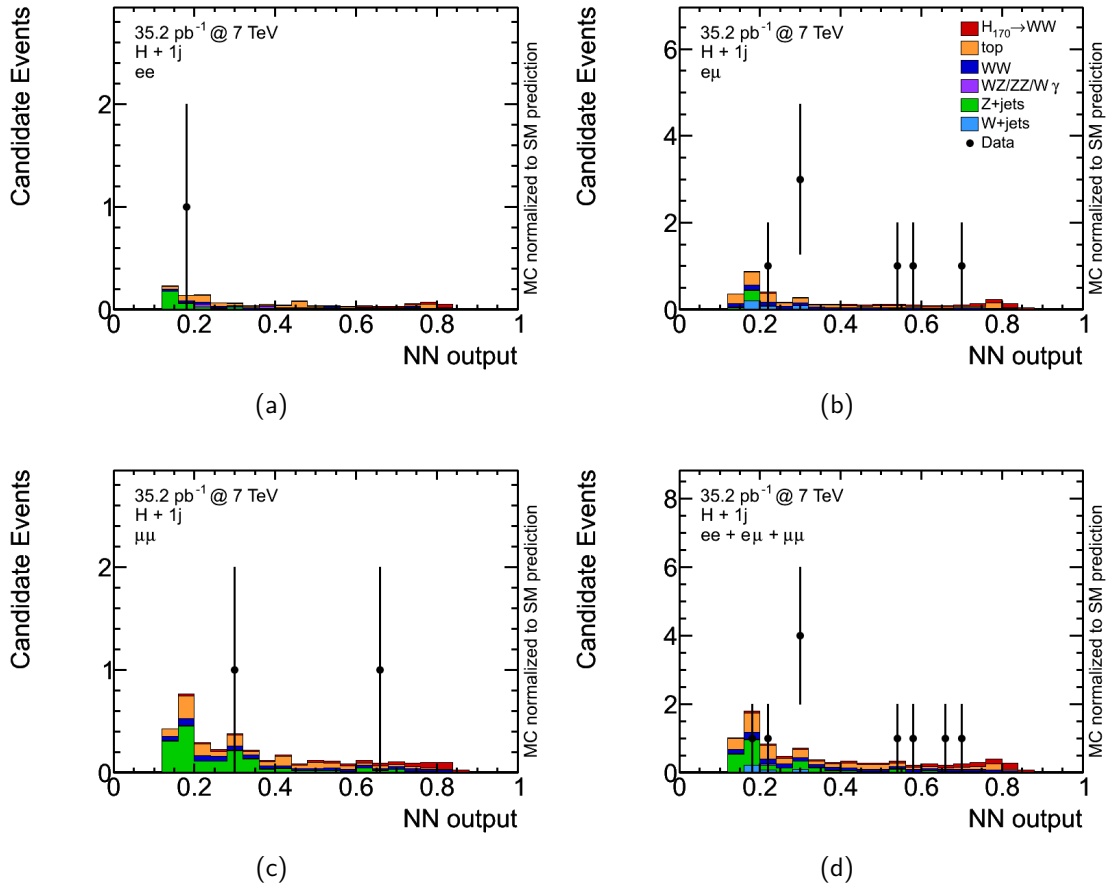


Figure 7.9.: *Observed distributions of the NNoutput for the $H + 1j$ analysis in 35 pb^{-1} of data. The predicted fractions of the different processes are stacked on top of each other, such that the sum gives the total expectation. The distributions are shown separately for the three decay channels: ee (a), eμ (b), and μμ (c). The sum of all three decay channels is shown as well (d).*

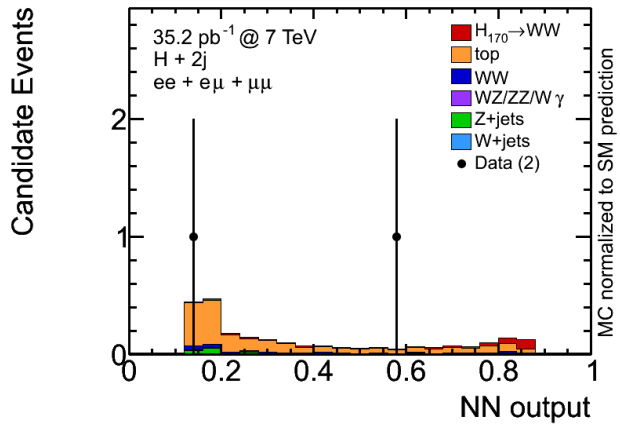


Figure 7.10.: *Observed distribution of the NNoutput for the $H + 2j$ analysis in 35 pb^{-1} of data. The predicted fractions of the different processes are stacked on top of each other, such that the sum gives the total expectation, showing the sum of all three decay channels.*

Table 7.5.: *Observed p-values and the corresponding significances in terms of standard Gaussian deviations for all considered Higgs boson mass points.*

M_H [GeV/c^2]	120	130	140	150	160	165	170	180	190	200
obs. p-value	0.26	0.06	0.09	0.06	0.11	0.18	0.14	0.13	0.08	0.05
significance [σ]	0.64	1.55	1.34	1.55	1.23	0.92	1.08	1.13	1.41	1.64

Table 7.5 provides the observed p-values and the corresponding significances, that are determined applying the statistical analysis on observed candidate event in collision data. The observed significances are similar to the expected ones.

7.3.2. Expected and Observed Limits

Upper limits on the production cross-section of the Standard Model Higgs-boson are determined, as the discovery potential of the analysis is not sufficient to establish the SM Higgs boson. Ensembles of pseudo experiments are generated again. Each ensemble implementing a larger signal cross-section until a exclusion limit of 95% confidence level is reached. The power-constraint method is utilized for limit-setting in the considered mass range. Thus, the expected limits coincide to the ones obtained with the CL_{s+b} procedure, while the observed limits are not allowed to fluctuate below the expected ones.

Figure 7.11 depicts the found upper limits on the Higgs boson cross-section. The observed limits are above the expected limits, which means that it was not necessary

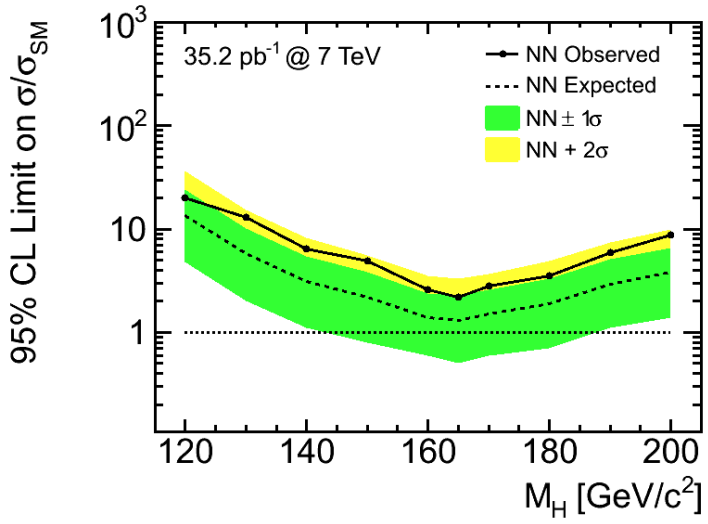


Figure 7.11.: *The Upper limits on the Higgs boson production cross-section obtained by this neural network analysis using* 35 pb^{-1} *of data recorded by the ATLAS experiment in 2010. The plotted values denote the 95% upper limits on the signal rate, as a multiple of the SM rate.*

Table 7.6.: *Observed and expected limits obtained by this neural network analysis and the cut-based analysis [131]. The values denote the 95% upper limits on the signal rate, as a multiple of the SM rate with* 35 pb^{-1} *of integrated luminosity. Both analyses utilize all three jet multiplicity bins. The cut-based analysis implements a counting experiment, while the full shape of the NNoutput distribution is taken as input to the limit-setting procedure in the case of the neural network analysis. As expected the neural network analysis gives better expected limits with smaller error bands than the cut-based analysis at all Higgs boson masses.*

M_H [GeV/c^2]	NN-based					cut-based				
	obs.	exp.	-1σ	$+1\sigma$	$+2\sigma$	obs.	exp.	-1σ	$+1\sigma$	$+2\sigma$
120	19.9	13.4	4.8	24.2	36.4	21.0	18.0	6.0	32.0	56.0
130	12.9	5.8	2.0	10.1	15.3	12.2	7.0	3.0	14.0	24.0
140	6.4	3.1	1.1	5.5	8.2	5.3	4.5	2.0	9.0	16.0
150	4.9	2.2	0.8	3.8	5.6	1.9	3.5	1.3	6.6	11.4
160	2.6	1.4	0.6	2.4	3.5	1.2	2.4	0.9	4.5	8.1
165	2.2	1.3	0.5	2.3	3.3	2.6	2.4	0.9	3.9	6.9
170	2.8	1.4	0.6	2.6	3.7	3.0	2.1	0.9	4.2	7.8
180	3.5	1.9	0.7	3.3	4.9	4.7	2.8	1.2	5.6	10.0
190	5.9	2.9	1.1	5.1	7.5	9.2	5.4	2.4	9.0	15.6
200	8.7	3.8	1.4	6.6	9.9	16.2	6.4	3.2	12.8	21.6

to apply the power-constraint. The most sensitive region is the mass range between 160 GeV/c^2 and 170 GeV/c^2, with an expected limit of 1.4 times the SM cross-section. The observed limits in this region are found to be between 2.2 and 2.8 times the SM cross-section. All determined limits are also summarized in Table 7.6 which additionally includes the upper limits determined by the cut-based analysis [131]. Both analyses analyze the same dataset and make use of all three jet multiplicity bins. The cut-based analysis implements a counting experiment, while the full shape of the NNoutput distribution is taken as input to the limit-setting procedure in the case of the neural network analysis. As expected the neural network analysis gives better expected limits with smaller error bands than the cut-based analysis at all Higgs boson masses. This comparison is also visualized in Figure 7.12(a). The observed upper limits found by the cut-based analysis are lower in the mass range below 165 GeV/c^2 and larger in the mass range above 165 GeV/c^2.

Further upper limits obtained with the same dataset but other multivariate analysis methods [135] are depicted in Figure 7.12(b) and (c). The limits obtained by the boosted decision tree analysis (b) and the matrix element analysis (c) reveal a similar behavior as the limits obtained by this analysis. All three multivariate analyses provide lower expected limits than the cut-based analysis. While the observed limits of the cut-based analysis fluctuate below the expected limits in the mass range between 145 GeV/c^2 and 165 GeV/c^2, there is no such fluctuation of the observed limits in any of the multivariate analyses.

Table 7.7 presents an overview of the upper limits determined by the cut-based analysis and the three multivariate analyses. The results of the neural network analysis described in this thesis are obtained by fitting the entire NNoutput distributions in all seven channels. This is in contrast to the previous neural network results reported in Reference [135] where a counting experiment was performed after a cut on the NNoutput distributions. Fitting all bins of the distribution has the advantage of making maximal use of all signal events remaining after the event selection. In addition, it also allows to constrain the background rates from data.

Hence, comparing the different analysis techniques, it is important to bear in mind that they exploit a different number of jet multiplicity bins and utilize different statistical methods. The boosted decision tree analysis includes six channels, namely the three decay channels ee, $e\mu$ and $\mu\mu$ in the zero-jet and one-jet multiplicity bin. A counting experiment is performed after a cut on the discriminating output variable. The matrix element method exploits only the zero-jet bin and combines the output of the three decay channels into one distribution that is subsequently used as input to the limit setting procedure. The neural network analysis presented in this thesis includes the combined $ee + e\mu + \mu\mu$ channel of the two-jet bin, additionally to the six decay channels of the zero and one-jet bin. All seven distributions are used to determine the upper limits, exploiting the different background contributions in the different decay channels and the shape differences of the considered processes.

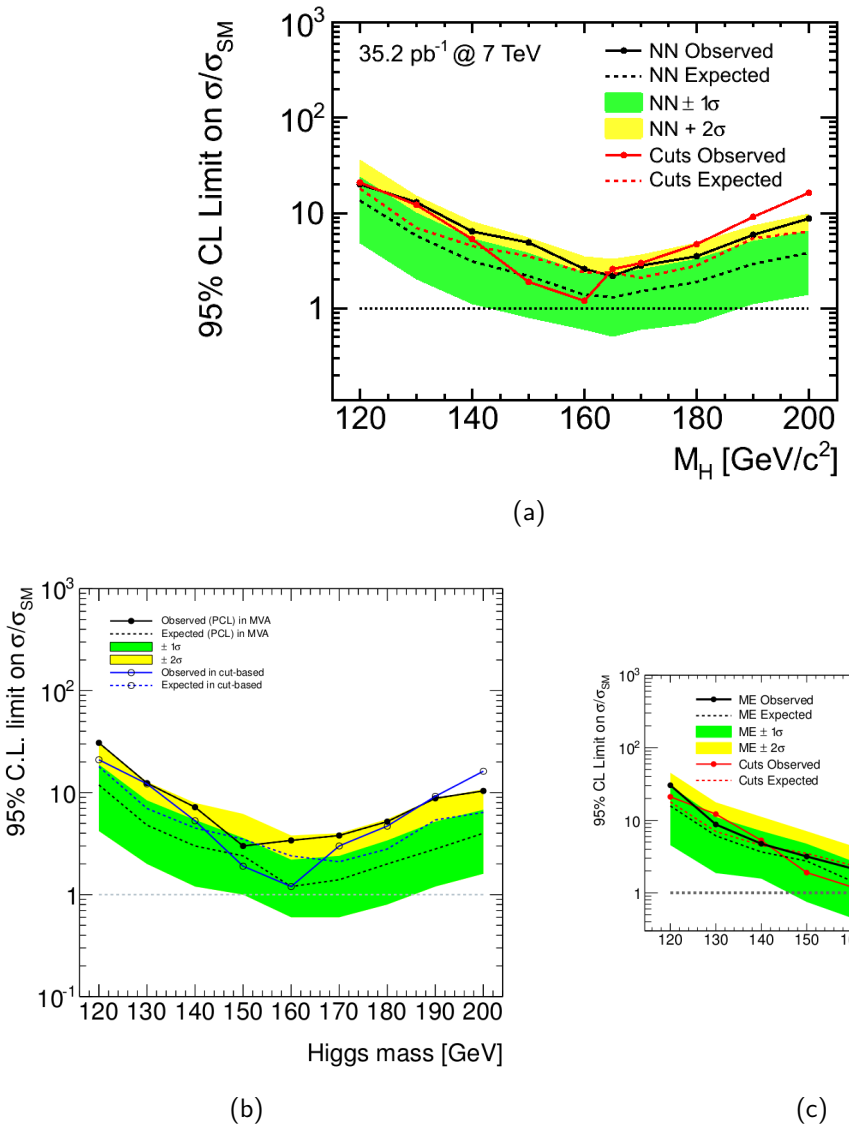


Figure 7.12.: *Upper limits on the Higgs boson production cross-section obtained using* 35 pb^{-1} *of data recorded by the ATLAS experiment in 2010. The plotted values denote the 95% upper limits on the signal rate, as a multiple of the SM rate. (a) presents the comparison of the limits obtained by the neural network analysis presented in this thesis and the cut-based analysis. Additional limits of further multivariate analyses are depicted in (b) and (c) showing the results of the boosted decision tree analysis and the matrix element analysis, respectively.*

Table 7.7.: *Observed and expected limits obtained by the cut-based analysis, the boosted decision tree (BDT) analysis, the neural network (NN) analysis presented in this thesis and the matrix element (ME) analysis. The values denote the 95% upper limits on the signal rate, as a multiple of the SM rate with* 35 pb^{-1} *of integrated luminosity. Comparing the different analysis techniques, it is important to bear in mind that they exploit a different number of jet multiplicity bins and utilize different statistical methods, as explained in the text.*

M_H [GeV/c^2]	cut-based		BDT		NN		ME	
	obs.	exp.	obs.	exp.	obs.	exp.	obs.	exp.
120	21.0	18.0	30.8	11.9	19.9	13.4	30.5	15.8
130	12.2	7.0	12.4	4.8	12.9	5.8	8.8	6.1
140	5.3	4.5	7.2	3.3	6.4	3.1	4.8	3.7
150	1.9	3.5	3.0	2.4	4.9	2.2	3.2	2.8
160	1.2	2.4	3.4	1.2	2.6	1.4	2.2	1.5
170	3.0	2.4	3.8	1.4	2.8	1.4	2.5	1.8
180	4.7	2.1	5.2	2.0	3.5	1.9	2.8	2.5
190	9.2	2.8	8.8	2.8	5.9	2.9	4.3	3.1
200	16.2	5.4	10.4	4.0	8.7	3.8	6.3	5.1

8. Summary

The Standard Model of elementary particle physics has been tested with high accuracy over the last decades [27, 28, 39]. Precision measurements carried out at particle accelerators like LEP and Tevatron have established that the Standard Model provides the correct effective description of strong and electroweak interactions at present energies. However, there is one particle predicted by this theory, the Higgs boson, that has neither been observed nor excluded, yet. The Higgs boson is introduced by the mechanism of spontaneous electroweak symmetry breaking which is a main ingredient of the electroweak sector of the Standard Model. It allows to generate the masses of the weak vector bosons and to formulate a consistent theory. The masses of the fermions can also be explained introducing Yukawa couplings.

Higgs boson production via two electroweak vector bosons WW or ZZ, the so called vector boson fusion (VBF), is the production mode of choice to establish or exclude a Standard Model like Higgs boson. Unlike the gluon-gluon fusion (GGF) mode, where the Higgs boson is produced via a heavy quark loop, the VBF does not involve Yukawa couplings. Measuring the cross-section of the VBF process provides a direct access to the coupling constants HWW and HZZ between the Higgs and the electroweak vector bosons. Such a measurement would proof that nature implements the Higgs boson as the agent of electroweak symmetry breaking.

The predicted Higgs boson production cross-section [120] is very small compared to other processes that form a huge background. The aim of this thesis is to develop and validate an analysis that provides a good separation between the VBF process and the background. Therefore, the Higgs boson decay to a W^+W^- pair is chosen as it has the largest branching ratio for Higgs masses larger than 135 GeV/c^2. To avoid contributions from QCD background events, the subsequent leptonic decay of both W bosons is considered. Together with the two scattered quarks from the vector boson fusion process this leads to a distinct event signature of two jets, two charged isolated leptons and missing transverse energy. Background events with the same final state, $t\bar{t}$ for example, are reduced exploiting the kinematical and topological features of the signal process. For the signal, the jets tend to be produced in the forward region of the detector and exhibit a rapidity gap. The second feature are the angular distributions of the charged leptons, which stem from the decay of the Higgs boson, caused by angular momentum conservation.

The analyzed dataset corresponds to an integrated luminosity of 35 pb^{-1} taken

with the ATLAS experiment at a center-of-mass energy of 7 TeV. Considering the small amount of data and the small production cross-section of VBF, the analysis also includes GGF. After a preselection the candidate events are categorized according the number of jets. Thus, accumulating GGF events in the zero-jet and one-jet multiplicity bin and VBF events in the two-jet multiplicity bin. Additionally, three different control regions are defined to validate the analysis techniques.

A multivariate analysis technique is favored over a cut-based approach to achieve a higher sensitivity. Neural networks provided by the NeuroBayes® package are applied to separate signal and background events. A neural network combines the information contained in the input variables and in the correlations between them into a discriminating output distribution. Multivariate techniques are trained using Monte Carlo predictions. Thus, the modeling of the generated events is crucial. They have to describe the observed collision data correctly.

The analysis is performed for ten Higgs boson mass points in the range between 120 GeV/c^2 and 200 GeV/c^2. After identifying the input variables, the modeling of each variable and the correlations is carefully checked. Three neural networks, each covering one of the three jet multiplicity bins, are optimized for all considered Higgs masses. The 30 networks obtained are validated evaluating them in the control regions and checking the modeling of the output distributions. To exploit the different background compositions in the three different decay channels ee, $e\mu$ and $\mu\mu$, events in the zero-jet and in the one-jet bin are split up according to the lepton flavors. Due to low statistics, this is not applied in the two-jet bin. Finally, the impact of systematic uncertainties is determined by applying the preselection and evaluating the neural networks on systematically altered Monte Carlo samples.

A statistical method based on a frequentist approach is performed to evaluate the expected significance of the analysis. Ensembles of pseudo experiments are generated including systematic uncertainties. Each pseudo experiment simulates a possible outcome of the analysis. The expected significance as determined by the statistical method is too low to find an evidence for the Higgs signal. Consequently, upper limits on the Standard Model Higgs boson cross-section are computed. The most sensitive region is the mass range between 160 GeV/c^2 and 170 GeV/c^2. The observed limits in this region are found to be between 2.2 and 2.8 times the Standard Model cross-section at the 95% confidence level. These results are consistent with other analyses, performed in the same channel [131, 135].

The latest results [41, 42] from ATLAS and CMS exclude the SM Higgs boson for all masses up to around 600 GeV/c^2 at 95% CL, except in the region around 120 GeV/c^2 to 130 GeV/c^2 where a small excess of events is seen. These limits are largely based on the GGF process which depends on Yukawa couplings. Thus, even if the Standard Model Higgs boson is excluded in the whole mass range, the VBF process still remains interesting. It can be used to test a fermiophobic model [7, 8] where the Higgs boson couples only to electroweak vector bosons.

List of Figures

List of Tables

Bibliography

[1] A. Djouadi. "The Anatomy of Electro-Weak Symmetry Breaking." In: (2005). arXiv:hep-ph/0503172v2 (cit. on pp. III, 2).

[2] K. Nakamura et al. (Particle Data Group). "Review of Particle Physics." In: *J. Phys. G* 37 (2010). DOI: 10.1088/0954-3899/37/7A/075021 (cit. on pp. III–IV).

[3] N. Cabbibo. "Unitary Symmetry and Leptonic Decays." In: *Phys. Rev. Lett.* 10 (1963), 531–533. DOI: 10.1103/PhysRevLett.10.531 (cit. on p. IV).

[4] M. Kobayashi and T. Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction." In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652 (cit. on p. IV).

[5] H.L. Lai et al. CTEQ Collaboration. "Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions." In: *Eur. Phys. J. C* 12 (2000), pp. 375–392. DOI: 10.1007/s100529900196. arXiv:hep-ph/9903282v3 (cit. on p. IV).

[6] F Wilczek. "Math without mass. I: Most of matter." In: *Phys.Today* 52 (1999), pp. 11–13 (cit. on p. IV).

[7] A. Barroso, L. Brucher, and R. Santos. "Is there a light fermiophobic Higgs?" In: *Phys. Rev.* D60 (1999), p. 035005. DOI: 10.1103/PhysRevD.60.035005. arXiv:hep-ph/9901293v2 (cit. on pp. VI, 116).

[8] "Experimental Signatures of Fermiophobic Higgs bosons." In: *Eur.Phys.J.* C12 (1999), pp. 87–98. DOI: 10.1007/s100529900252. arXiv:hep-ph/9907434v2 (cit. on pp. VI, 116).

[9] C. Quigg. "Spontaneous Symmetry Breaking as a Basis of Particle Mass." In: (2007). arXiv:0704.2232v2 (cit. on pp. VI, 2, 7, 10).

[10] S.L. Glashow. "Partial-symmetries of weak interactions." In: *Nucl. Phys.* 22 (1961), 579–588. DOI: 10.1016/0029-5582(61)90469-2 (cit. on p. 2).

[11] S. Weinberg. "A Model of Leptons." In: *Phys. Rev. Lett.* 19 (1967), 1264–1266. DOI: 10.1103/PhysRevLett.19.1264 (cit. on p. 2).

[12] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry." In: *Phys. Rev. D* 2 (1970), 1285–1292. DOI: 10.1103/PhysRevD.2.1285 (cit. on p. 2).

[13] C. Quigg. *Gauge Theories in Particle Physics*. Westvew Press, 1997 (cit. on p. 2).

[14] C. Quigg. *Particles and the Standard Model*. Cambridge University Press, 2006 (cit. on p. 2).

[15] D. Griffiths. *Introduction to Elementary Particles*. WILEY-VCH, 2008 (cit. on p. 2).

[16] I.J.R. Aitchinson and A.J.G. Hey. *Introduction to Elementary Particles*. 2nd ed. Hilger, A., 2008 (cit. on p. 2).

[17] P. Langacker. *The Standard Model and Beyond*. Taylor and Francis, 2010 (cit. on p. 2).

[18] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons." In: *Phys. Rev. Lett.* 13 (1964), 321–323. DOI: 10.1103/PhysRevLett.13.321 (cit. on p. 5).

[19] P.W. Higgs. "Broken symmetries, massless particlees and gauge fields." In: *Phys. Lett.* 12 (1964), 132–133. DOI: 10.1016/0031-9163(64)91136-9 (cit. on p. 5).

[20] P.W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons." In: *Phys. Rev. Lett.* 13 (1964), 508–509. DOI: 10.1103/PhysRevLett.13.508 (cit. on p. 5).

[21] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles." In: *Phys. Rev. Lett.* 13 (1964), 585–587. DOI: 10.1103/PhysRevLett.13.585 (cit. on p. 5).

[22] P.W. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons." In: *Phys. Rev.* 145 (1966), 1156–1163. DOI: 10.1103/PhysRev.145.1156 (cit. on p. 5).

[23] T. W. B. Kibble. "Symmetry Breaking in Non-Abelian Gauge Theories." In: *Phys. Rev.* 155 (1967), 1554–1561. DOI: 10.1103/PhysRev.155.1554 (cit. on p. 5).

[24] H. Fritzsch, M. Gell-Mann, and P. Minkowski. "Vectorlike weak currents and new elementary fermions." In: *Phys. Lett. B* 59 (1975), pp. 256–260. DOI: 10.1016/0370-2693(75)90040-4 (cit. on p. 6).

[25] P. Minkowski. "$\mu \to e\gamma$ at a rate of one out of 109 muon decays?" In: *Phys. Lett. B* 67 (1977), 421–428. DOI: 10.1016/0370-2693(77)90435-X (cit. on p. 6).

[26] B. Pontecorvo. In: *Zh. Eksp. Teor. Fiz.* 53 (1967), p. 1717 (cit. on p. 6).

[27] The LEP Electroweak Working Group. *Status of JULY*. 2011. URL: http://lepewwg.web.cern.ch/LEPEWWG/ (cit. on pp. 7, 115).

[28] The LEP Collaborations: ALEPH Collaboration DELPHI Collaboration L3 Collaboration OPAL Collaboration the LEP Electroweak Working Group. "A Combination of Preliminary Electroweak Measurements and Constraints on the Standard Model." In: (2007). arXiv:hep-ex/0612034v2 (cit. on pp. 7, 10, 115).

[29] LHC Higgs Cross Section Working Group. 2011. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections (cit. on pp. 8–9).

[30] N. Cabibbo et al. "Bounds on the fermions and Higgs boson masses in grand unified theories." In: *Nucl. Phys. B* 158 (1979), pp. 295–305. DOI: 10.1016/0550-3213(79)90167-6 (cit. on p. 10).

[31] Linde A.D. In: *Pisma Zh. Eksp. Teor. fiz.* 23 (1976), pp. 73–76 (cit. on p. 10).

[32] S Weinberg. "Mass of the Higgs Boson." In: *Phys. Rev. Lett.* 36 (1976), 294–296. DOI: 10.1103/PhysRevLett.36.294 (cit. on p. 10).

[33] P.Q. Hung. "Vacuum Instability and New Constraints on Fermion Masses." In: *Phys. Rev. Lett.* 42 (1979), 873–876. DOI: 10.1103/PhysRevLett.42.873 (cit. on p. 10).

[34] M. Sher. "Electroweak Higgs potential and vacuum stability." In: *Phys. Rept.* 179 (1989), 273–418. DOI: 10.1016/0370-1573(89)90061-6 (cit. on p. 10).

[35] G. Altarelli and G. Isidori. "Lower limit on the Higgs mass in the standard model: An update." In: *Phys. Lett. B* 337 (1994), 141–144. DOI: 10.1016/0370-2693(94)91458-3. arXiv:hep-ph/9606386v1 (cit. on p. 10).

[36] B. Schrempp and M. Wimmer. "Top Quark and Higgs Boson Masses: Interplay between Infrared and Ultraviolet Physics." In: *Prog.Part.Nucl.Phys* 371-90 (1996), pp. 1–90. DOI: 10.1016/0146-6410(96)00059-2. arXiv:hep-ph/9606386v1 (cit. on p. 10).

[37] J.A. Casas et al. "The Lightest Higgs Boson Mass in the Minimal Supersymmetric Standard Model." In: *Nucl.Phys.B* 436 (1995), pp. 3–29. DOI: 10.1016/0550-3213(94)00508-C. arXiv:hep-ph/9407389v1 (cit. on p. 10).

[38] T. Hambye and K. Riesselmann. "Matching conditions and Higgs mass upper bounds revisited." In: *Phys.Rev. D* 55 (1997), pp. 7255–7262. DOI: 10.1103/PhysRevD.55.7255. arXiv:hep-ph/9610272v1 (cit. on p. 10).

[39] ALEPH Collaboration et al. "Precision Electroweak Measurements and Constraints on the Standard Model." In: (2010). arXiv:1012.2367v2 (cit. on pp. 10, 115).

[40] Tevatron New Phenomena and Higgs Working Group. "Combined CDF and DØ Searches for Standard Model Higgs Boson Production." In: FERMILAB-CONF-12-065-E (2012) (cit. on p. 10).

[41] The ATLAS collaboration. "An update to the combined search for the Standard Model Higgs boson with the ATLAS detector at the LHC using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV." In: (2012). CERN Document Server: ATLAS-CONF-2012-019 (cit. on pp. 10–11, 116).

[42] CMS Collaboration. "Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV." In: (2012). arXiv:1202.1488 (cit. on pp. 10–11, 116).

[43] The CDF D0 Collaborations, the Tevatron New Phenomena, and Higgs Working Group. "Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb^{-1} of Data." In: (2011). arXiv:1107.5518v2 (cit. on p. 11).

[44] The ATLAS Collaboration et al. "The ATLAS Experiment at the CERN Large Hadron Collider." In: *JINST* 3 (2008). DOI: 10.1088/1748-0221/3/08/S08003 (cit. on pp. 12, 14–16).

[45] CMS Collaboration. "CMS Physics Technical Design Report Volume I : Detector Performance and Software." In: (2006). CERN Document Server: CMS-TDR-008-1 (cit. on p. 12).

[46] The ALICE Collaboration et al. "The ALICE experiment at the CERN LHC." In: *JINST* 3 (2008). DOI: 10.1088/1748-0221/3/08/S08002 (cit. on p. 12).

[47] The LHCf Collaboration et al. "The LHCf detector at the CERN Large Hadron Collider." In: *JINST* 3 (2008). DOI: 10.1088/1748-0221/3/08/S08006 (cit. on p. 12).

[48] The TOTEM Collaboration et al. "The TOTEM Experiment at the CERN Large Hadron Collider." In: *JINST* 3 (2008). DOI: 10.1088/1748-0221/3/08/S08007 (cit. on p. 12).

[49] C.E. Hill. "Ion and electron sources." In: (1994). CERN Document Server: CERN-PS-94-36 (cit. on p. 12).

[50] K. Schindl. "The PS Booster as Pre-Injector for LHC." In: *Part. Accel.* 58 (1997), 63–78. CERN Document Server: CERN-PS-97-011-DI (cit. on p. 12).

[51] R. Cappi. "The PS in the LHC injector chain." In: *Part. Accel.* 58 (1997), 79–89 (cit. on p. 13).

[52] T. Linnecar. "Preparing the SPS for LHC." In: *Part. Accel.* 58 (1997), 91–101 (cit. on p. 13).

[53] L Evans and P Bryant. "LHC Machine." In: *JINST* 3 (2008). DOI: 10.1088/1748-0221/3/08/S08001 (cit. on p. 13).

[54] G Aad et al. "ATLAS pixel detector electronics and sensors." In: *JINST* 3 (2008). DOI: 10.1088/1748-0221/3/07/P07007 (cit. on p. 16).

[55] A. Ahmad et al. "The silicon microstrip sensors of the ATLAS semiconductor tracker." In: *Instrum. Meth.* (2008). DOI: 10.1016/j.nima.2007.04.15 (cit. on p. 16).

[56] The ATLAS TRT collaboration et al. "The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance." In: *JINST* 3 (2008). DOI: 10.1088/1748-0221/3/02/P02013 (cit. on p. 16).

[57] R. Ruber et al. "Ultimate Performance of the ATLAS Superconducting Solenoid." In: *IEEE Trans. Appl. Supercond.* (2007) (cit. on p. 16).

[58] B. Aubert et al. "Construction, assembly and tests of the ATLAS electromagnetic barrel calorimeter." In: *Nucl. Instrum. Meth.* A (2006), pp. 558–388. DOI: 10.1016/j.nima.2005.11.212 (cit. on p. 16).

[59] A. Artamonov et al. "The ATLAS forward calorimeters." In: *JINST* 3 (2008). DOI: 10.1088/1748-0221/3/02/P02010 (cit. on p. 17).

[60] P. Adragna et al. "The ATLAS hadronic tile calorimeter: from construction toward physics." In: *IEEE Trans. Nucl. Sci.* (2006). DOI: 10.1109/TNS.2006.872628 (cit. on p. 17).

[61] S. Palestini. "The muon spectrometer of the ATLAS experiment." In: *Nuclear Physics B Proceedings Supplements* 125 (2003), pp. 337–345. DOI: 10.1016/S0920-5632(03)91013-9 (cit. on p. 17).

[62] P. Jenni et al. "ATLAS high-level trigger, data-acquisition and controls : Technical Design Report." In: (2003). CERN Document Server: ATLAS-TDR-016 (cit. on p. 18).

[63] J.C. Vermeulen et al. "The baseline dataflow system of the ATLAS trigger and DAQ." In: (2003). CERN Document Server: ATL-DAQ-2003-032 (cit. on p. 18).

[64] The ATLAS collaboration. "Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC." In: *Eur. Phus. J.* C71 (2011), pp. 333–380. DOI: 10.1140/epjc/s10052-011-1630-5. arXiv:1101.2185v1 (cit. on pp. 18–19, 99).

[65] M. Bruschia. "The ATLAS luminosity monitor." In: *Nuclear Instruments and Methods in Physics Research Section A* 623 (2010), pp. 371–373. DOI: 10.1016/j.nima.2010.02.252 (cit. on p. 19).

[66] S. Franza and P. Barrillon. "ATLAS ALFA—measuring absolute luminosity with scintillating fibres." In: *Nuclear Instruments and Methods in Physics Research Section A* 610 (2009), 35–40. DOI: 10.1016/j.nima.2009.05.148 (cit. on p. 19).

[67] S. White. "The ATLAS zero degree calorimeter." In: *Nuclear Instruments and Methods in Physics Research Section A* 617 (2010), pp. 126–128. DOI: 10.1016/j.nima.2009.09.120 (cit. on p. 19).

[68] M.G. Albrow, R.B. Appleby, and M. Arneodo et al. "The FP420 R&D project: Higgs and New Physics with forward protons at the LHC." In: *JINST* 4 (2009). DOI: 10.1088/1748-0221/4/10/T10001 (cit. on p. 19).

[69] L. Adamczyka et al. *AFP: A Proposal to Install Proton Detectors at 220 m around ATLAS to Complement the ATLAS High Luminosity Physics Program.* 2011. URL: http://atlas-project-lumi-fphys.web.cern.ch/atlas-project-lumi-fphys/AFP_TP.pdf (cit. on p. 19).

[70] The ATLAS collaboration. "Updated Luminosity Determination in pp Collisions at root(s)=7 TeV using the ATLAS Detector." In: (2011). CERN Document Server: ATLAS-CONF-2011-011 (cit. on pp. 19, 99).

[71] The ATLAS collaboration. "Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} =$ 7 TeV." In: (2010). CERN Document Server: ATLAS-CONF-2010-038 (cit. on pp. 19, 53).

[72] M.L. Mangano and T.J. Stelzer. "Tools for the simulation of hard hadronic collisions." In: *Annu. Rev. Nucl. Part. Sci.* 55 (2005), pp. 555–588. DOI: 10.1146/annurev.nucl.55.090704.151505 (cit. on p. 22).

[73] S. Catani et al. "QCD Matrix Elements + Parton Showers." In: *JHEP* 11 (2001), p. 063. DOI: 10.1088/1126-6708/2001/11/063. arXiv:hep-ph/0109231v1 (cit. on p. 23).

[74] F. Krauss. "Matrix Elements and Parton Showers in Hadronic Interactions." In: *JHEP* 08 (2002), p. 015. DOI: 10.1088/1126-6708/2002/08/015. arXiv:hep-ph/0205283v1 (cit. on p. 23).

[75] M.L. Mangano, M. Moretti, and R. Pittau. "Multijet matrix elements and shower evolution in hadronic collisions: W b bbar + n jets as a case study." In: *Nucl.Phys.B* 632 (2002), pp. 343–362. DOI: 10.1016/S0550-3213(02)00249-3. arXiv:hep-ph/0108069v1 (cit. on pp. 23, 35).

[76] S. Hoeche et al. "QCD matrix elements and truncated showers." In: *JHEP* 05 (2009), p. 053. DOI: 10.1088/1126-6708/2009/05/053. arXiv:0903.1219v2 (cit. on p. 23).

[77] M.L. Mangano et al. "ALPGEN, a generator for hard multiparton processes in hadronic collisions." In: *JHEP* 07 (2003), p. 001. DOI: doi:10.1088/1126-6708/2003/07/001 (cit. on p. 23).

[78] T. Binoth et al. "Gluon-induced W-boson pair production at the LHC." In: *JHEP* 12 (2006), p. 046. DOI: doi:10.1088/1126-6708/2006/12/046. arXiv:hep-ph/0611170v1 (cit. on p. 23).

[79] J. Alwall et al. "MadGraph/MadEvent v4: The New Web Generation." In: *JHEP* 09 (2007), p. 028. DOI: 10.1088/1126-6708/2007/09/028. arXiv:0706.2334v1 (cit. on p. 23).

[80] S. Frixione and B.R. Webber. "Matching NLO QCD computations and parton shower simulations." In: *JHEP* 06 (2002), p. 029. DOI: doi:10.1088/1126-6708/2002/06/029. hep-ph/0204244v2: arXiv (cit. on p. 23).

[81] S. Frixione, P. Nason, and B.R. Webber. "Matching NLO QCD and parton showers in heavy flavour production." In: *JHEP* 08 (2003), p. 007. DOI: doi:10.1088/1126-6708/2003/08/007 (cit. on p. 23).

[82] G. Corcella et al. "HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)." In: *JHEP* 01 (2001), p. 001. DOI: doi:10.1088/1126-6708/2001/01/010 (cit. on p. 24).

[83] T. Sjöstrand, S. Mrenna, and P. Skands. "PYTHIA 6.4 physics and manual." In: *JHEP* 05 (2006), p. 026. DOI: doi:10.1088/1126-6708/2006/05/026 (cit. on p. 24).

[84] B. Andersson et al. "Parton fragmentation and string dynamics." In: *Phys. Rept.* 97 (2002), pp. 31–145. DOI: doi:10.1016/0370-1573(83)90080-7 (cit. on p. 24).

[85] T. Gleisberg et al. "Event generation with SHERPA 1.1." In: *JHEP* 02 (2009), p. 007. DOI: doi:10.1088/1126-6708/2009/02/007. arXiv:0811.4622v1 (cit. on p. 24).

[86] M. Jezabek et al. "The tau decay library TAUOLA, update with exact $\mathcal{O}(\alpha)$ QED corrections in $\tau \to \mu(e)\nu\bar{\nu}$ decay modes," in: *Comp. Phys. Commun.* 76 (1993), p. 361 (cit. on p. 24).

[87] The ATLAS Collaboration. "The ATLAS Simulation Infrastructure." In: *Eur.Phys.J.* C70 (2010), pp. 823–874. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv:1005.4568v1 (cit. on p. 25).

[88] S. Agostinelliae et al. "Geant4—a simulation toolkit." In: *NIM A* 506 (2003), 250–303. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on p. 25).

[89] R.E. Kalman. "A new approach to linear filtering and prediction problems." In: *Transactions of the ASME – Journal of Basic Engineering, Series D* 82 (1960), pp. 35–45 (cit. on p. 25).

[90] T. Cornelissen et al. "Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)." In: (2007). CERN Document Server: ATL-SOFT-PUB-2007-007 (cit. on p. 26).

[91] The ATLAS Collaboration. "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics." In: (2008). arXiv:0901.0512v4 (cit. on pp. 26, 28, 30).

[92] G. Piacquadio, K. Prokofiev, and A. Wildauer. "Primary vertex reconstruction in the ATLAS experiment at LHC." In: *J. Phys.: Conf. Ser.* 119 (2008). DOI: 10.1088/1742-6596/119/3/032033 (cit. on p. 26).

[93] E. Bouhova-Thacker et al. "A framework for vertex reconstruction in the ATLAS experiment at LHC." In: *J. Phys.: Conf. Ser.* (2010). DOI: 10.1088/1742-6596/219/3/032019 (cit. on p. 26).

[94] The ATLAS Collaboration. "Measurement of the $W \to \ell\nu$ and $Z/\gamma* \to \ell\ell$ production cross sections in proton-proton collisions at $\sqrt{s} = 7TeV$ with the ATLAS detector." In: *JHEP* 12 (2010), p. 060. DOI: 10.1007/JHEP12(2010)060. arXiv:1010.2130 (cit. on p. 28).

[95] The ATLAS collaboration. "Electron and photon reconstruction and identification in ATLAS: expected performance at high energy and results at 900 GeV." In: (2010). CERN Document Server: ATLAS-CONF-2010-005 (cit. on pp. 28, 99).

[96] ATLAS Collaboration. "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data." In: *Submitted to Eur. Phys. J. C* (2011). arXiv:1110.3174 (cit. on pp. 28, 99).

[97] The ATLAS collaboration. "Muon Reconstruction Performance." In: (2010). CERN Document Server: ATLAS-CONF-2010-064 (cit. on p. 29).

[98] N. Orlando and The ATLAS collaboration. "Muon reconstruction efficiency measurement in the ATLAS experiment." In: (2012). CERN Document Server: ATL-PHYS-PROC-2012-011 (cit. on p. 29).

[99] M. Cacciari, G.P. Salam, and G. Soyez. *FastJet*. URL: http://fastjet.fr/ (cit. on p. 29).

[100] M. Cacciari and G.P. Salam. "Dispelling the N^3 myth for the Kt jet-finder." In: *Phys.Lett. B* 641 (2006), pp. 57–61. DOI: 10.1016/j.physletb.2006.08.037. arXiv:hep-ph/0512210v2 (cit. on p. 29).

[101] M. Cacciari, G.P. Salam, and G. Soyez. "The anti-k_t jet clustering algorithm." In: *JHEP* 0804 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv:0802.1189v2 (cit. on p. 29).

[102] W. Lampl et al. "Calorimeter Clustering Algorithms : Description and Performance." In: (2008). CERN Document Server: ATL-LARG-PUB-2008-002 (cit. on pp. 29, 32).

[103] L. Asquith et al. "Performance of Jet Algorithms in the ATLAS Detector." In: (2010). CERN Document Server: ATL-PHYS-INT-2010-129 (cit. on p. 29).

[104] The ATLAS collaboration. "Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS 2010 data." In: (2011). CERN Document Server: ATLAS-CONF-2011-032 (cit. on pp. 30–31, 98).

[105] The ATLAS collaboration. "Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s} = 7$ TeV." In: (2010). CERN Document Server: ATLAS-CONF-2010-053 (cit. on pp. 30, 32–33, 51).

[106] The ATLAS collaboration. "Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s} = 7$ TeV and measured with the ATLAS detector." In: (2010). CERN Document Server: ATLAS-CONF-2010-056 (cit. on pp. 30, 98).

[107] The ATLAS collaboration. "In-situ pseudo-rapidity inter-calibration to evaluate jet energy scale uncertainty and calorimeter performance in the forward region." In: (2010). CERN Document Server: ATLAS-CONF-2010-055 (cit. on pp. 30, 98).

[108] The ATLAS collaboration. "Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data." In: (2010). CERN Document Server: ATLAS-CONF-2010-042 (cit. on p. 31).

[109] The ATLAS collaboration. "Calibrating the b-Tag and Mistag Efficiencies of the SV0 b-Tagging Algorithm in 3 pb^{-1} of Data with the ATLAS Detector." In: (2010). CERN Document Server: ATLAS-CONF-2010-099 (cit. on pp. 31, 51).

[110] The ATLAS collaboration. "Reconstruction and Calibration of Missing Transverse Energy and Performance in Z and W events in ATLAS Proton-Proton Collisions at 7 TeV." In: (2011). CERN Document Server: ATLAS-CONF-2011-080 (cit. on pp. 32–33, 51).

[111] K. Melnikov and F. Petriello. "The W boson production cross section at the LHC through $O(\alpha_s^2)$." In: *Phys.Rev.Lett.* 96 (2006), p. 231803. DOI: 10.1103/PhysRevLett.96.231803. arXiv:hep-ph/0603182 (cit. on p. 35).

[112] K. Melnikov and F. Petriello. "Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$." In: *Phys.Rev.D* 74 (2006), p. 114017. DOI: 10.1103/PhysRevD.74.114017. arXiv:hep-ph/0609070 (cit. on p. 35).

[113] S. Catani et al. "Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO." In: *Phys.Rev.Lett.* 103 (2009), p. 082001. DOI: 10.1103/PhysRevLett.103.082001. arXiv:0903.2120 (cit. on p. 35).

[114] N. Kidonakis. "Top quark pair and single top production at Tevatron and LHC energies." In: *PoS ICHEP* (2010). arXiv:1008.2460 (cit. on p. 35).

[115] N. Kidonakis. "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-." In: *Phys. Rev. D* 82 (2010), p. 054018. DOI: 10.1103/PhysRevD.82.054018. arXiv:1005.4451 (cit. on p. 35).

[116] N. Kidonakis. "NNLL resummation for s-channel single top quark production." In: *Phys. Rev. D* 81 (2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028. arXiv:1001.5034 (cit. on p. 35).

[117] N. Kidonakis. "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production." In: *Phys. Rev. D* 83 (2011), p. 091503. DOI: 10.1103/PhysRevD.83.091503. arXiv:1103.2792 (cit. on p. 35).

[118] J. Butterworth et al. "Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV." In: (2010). CERN Document Server: ATL-COM-PHYS-2010-695 (cit. on p. 35).

[119] M. Aharrouche et al. "Measurement of the production cross section of $W\gamma$ and $Z\gamma$ at $\sqrt{s} = 7$ TeV with the ATLAS Detector." In: (2011). CERN Document Server: ATL-COM-PHYS-2011-119 (cit. on p. 35).

[120] LHC Higgs Cross Section Working Group et al. "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables." In: (2011). arXiv:1101.0593v3 (cit. on pp. 37–39, 115).

[121] D. Rainwater and D. Zeppenfeld. "Observing $H \to W^*W^* \to e\mu p_{\mathrm{T}}^{(miss)}$ in weak boson fusion with dual forward jet tagging at the LHC." In: *Phys.Rev. D* 60 (1999), p. 113004. DOI: 10.1103/PhysRevD.60.113004. arXiv:hep-ph/9906218v3 (cit. on p. 38).

[122] N.Kauer et al. "$H \to WW$ as the discovery mode for a light Higgs boson." In: *Phys.Lett. B* 503 (2001), pp. 113–120. DOI: 10.1016/S0370-2693(01)00211-8. arXiv:hep-ph/0012351v1 (cit. on p. 38).

[123] D. Green. "Kinematics in Vector Boson Fusion." In: (2006). arXiv:hep-ph/0603022 (cit. on p. 42).

[124] G. Sartisohn. *Studies for the Measurement of the W Boson Helicity in Top Quark Decays with the CMS Experiment*. IEKP-KA/2008-1. 2008. URL: http://www-ekp.physik.uni-karlsruhe.de/pub/web/thesis/iekp-ka2008-01.pdf (cit. on p. 44).

[125] The ATLAS collaboration. "Performance of the Electron and Photon Trigger in p-p Collisions at $\sqrt{s}$ = 7 TeV." In: (2011). CERN Document Server: ATLAS-CONF-2011-114 (cit. on p. 50).

[126] The ATLAS collaboration. "Determination of the muon reconstruction efficiency in ATLAS at the Z resonance in proton-proton collisons at $\sqrt{s}$ = 7 TeV." In: (2011). CERN Document Server: ATLAS-CONF-2011-008 (cit. on pp. 50, 53–54).

[127] D W Miller and D Schwartzman A Su. "Jet-Vertex Association Algorithm." In: (2008). CERN Document Server: ATL-COM-PHYS-2008-008 (cit. on p. 51).

[128] D W Miller and D Schwartzman A Su. "Pile-up jet energy scale corrections using the jet-vertex fraction method." In: (2009). CERN Document Server: ATL-PHYS-INT-2009-090 (cit. on p. 51).

[129] The ATLAS collaboration. *JetVertexFraction: methods to mitigate pile-up effects on jets and the jet-energy scale*. 2010. URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetVertexFraction (cit. on p. 52).

[130] The ATLAS collaboration. "Performance of the Missing Transverse Energy Reconstruction and Calibration in Proton-Proton Collisions at a Center-of-Mass Energy of 7 TeV with the ATLAS Detector." In: (2010). CERN Document Server: ATLAS-CONF-2010-057 (cit. on p. 51).

[131] The ATLAS collaboration. "Higgs Boson Searches using the $H \to WW^* \to \ell\nu\ell\nu$ Decay Mode with the ATLAS Detector at $\sqrt{s}$ = 7 TeV." In: (2011). CERN Document Server: ATLAS-CONF-2011-005 (cit. on pp. 53, 62, 111–112, 116).

[132] The ATLAS collaboration. "Expected electron performance in the ATLAS experiment." In: (2011). CERN Document Server: ATL-PHYS-PUB-2011-006 (cit. on p. 53).

[133] The ATLAS collaboration. "A measurement of the ATLAS muon reconstruction and trigger efficiency using J/psi decays." In: (2011). CERN Document Server: ATLAS-CONF-2011-021 (cit. on p. 53).

[134] The ATLAS collaboration. "Measurement of the $W \to \ell\nu$ and $Z/\gamma^* \to \ell\ell$ production cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector." In: (2010). arXiv:1010.2130v1 (cit. on p. 58).

[135] A Armbruster et al. "Higgs Boson Searches in the $H \to WW^* \to \ell\nu\ell\nu$ Channel using Multivariate Analysis Techniques with the ATLAS Detector at 7 TeV." In: (2011). CERN Document Server: ATL-COM-PHYS-2011-894 (cit. on pp. 62, 102–105, 112, 116).

[136] M Feindt. "A Neural Bayesian Estimator for Conditional Probability Densities." In: (2004). arXiv:physics/0402093v1 (cit. on p. 63).

[137] M Feindt and U Kerzel. "The NeuroBayes neural network package." In: *Nucl. Instrum. Meth.* A559 (2006), 190–194. DOI: 10.1016/j.nima.2005.11.166 (cit. on p. 63).

[138] G Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics." In: *Eur. Phys. J.* C71 (2010), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv:1007.1727v2 (cit. on pp. 79, 97).

[139] CDF Collaboration. "Observation of Single Top Quark Production and Measurement of |Vtb| with CDF." In: *Phys.Rev.* D82 (2010), p. 112005. DOI: 10.1103/PhysRevD.82.112005. arXiv:1004.1181v4 (cit. on p. 89).

[140] J Neyman and E Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses." In: *Phil. Trans. R. Soc. Lond.* A231 (1933), 289–337. DOI: 10.1098/rsta.1933.0009 (cit. on p. 92).

[141] J Neyman. "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability." In: *Phil. Trans. R. Soc. Lond.* A236 (1937), pp. 333–380. DOI: 10.1098/rsta.1937.0005 (cit. on p. 95).

[142] G Cowan. *Statistical Data Analysis*. Clarendon Press, 1998, pp. 118–142 (cit. on p. 95).

[143] G. Cowan et al. "Power-Constrained Limits." In: (2011). arXiv:1105.3166v1 (cit. on pp. 96–97).

[144] T Junk. "Confidence level computation for combining searches with small statistics." In: *Nucl. Instrum. Methods Phys. Res., Sec.* A434 (1999), 435–443. DOI: 10.1016/S0168-9002(99)00498-2 (cit. on p. 95).

[145] A L Read. "Presentation of search results: the CL_s technique." In: *J. Phys. G: Nucl. Part. Phys.* 28 (2002). DOI: 10.1088/0954-3899/28/10/313 (cit. on p. 95).

[146] The ATLAS collaboration. "Update on the jet energy scale systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s} = 7$ TeV measured with the ATLAS detector." In: (2011). CERN Document Server: ATLAS-CONF-2011-007 (cit. on p. 98).

[147] The ATLAS collaboration. "ATLAS jet energy scale uncertainties using tracks in proton proton collisions at $\sqrt{s} = 7$ TeV." In: (2011). CERN Document Server: ATLAS-CONF-2011-067 (cit. on p. 98).

[148] The ATLAS collaboration. "Jet energy resolution and selection efficiency relative to track jets from in-situ techniqu0es with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s} = 7$ TeV." In: (2011). CERN Document Server: ATLAS-CONF-2010-054 (cit. on p. 98).

[149] The ATLAS collaboration. "Jet energy resolution with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010." In: (2011). CERN Document Server: `ATLAS-PERF-2011-04-001` (cit. on p. 98).

[150] The ATLAS collaboration. "Muon Momentum Resolution in First Pass Reconstruction of pp Collision Data Recorded by ATLAS in 2010." In: (2011). CERN Document Server: `ATLAS-CONF-2011-046` (cit. on p. 99).

A. Neural Network Input Variables

A.1. $H + 1j$ analysis

Figures A.1 to A.3 present the input variables, characterizing the properties of the Higgs boson decay products, of the $H + 1j$ analysis for the Higgs boson mass of 170 GeV/c^2.

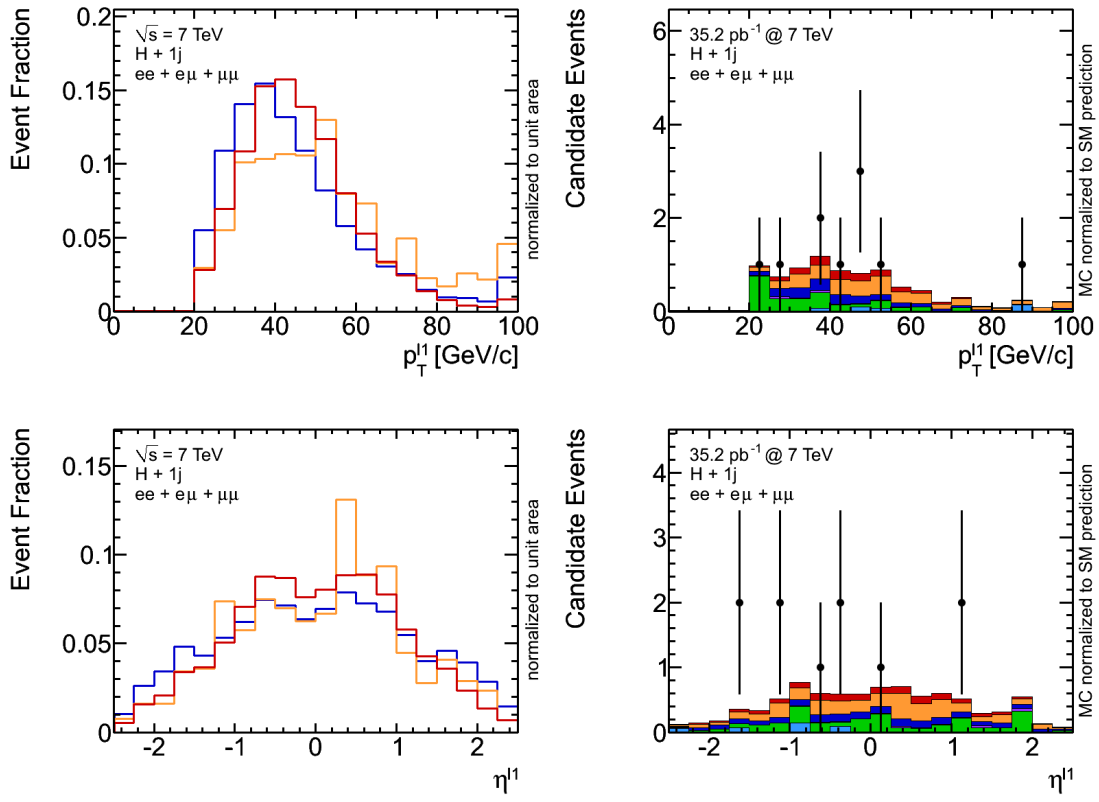


Figure A.1.: *The p_T and η distributions of the leading lepton in the $H + 1j$ analysis for a Higgs boson mass of* 170 GeV/c^2. *The column on the left hand side shows the kinematic distributions normalized to unit area, while the validation of the modeling is presented in the column on the right hand side.*

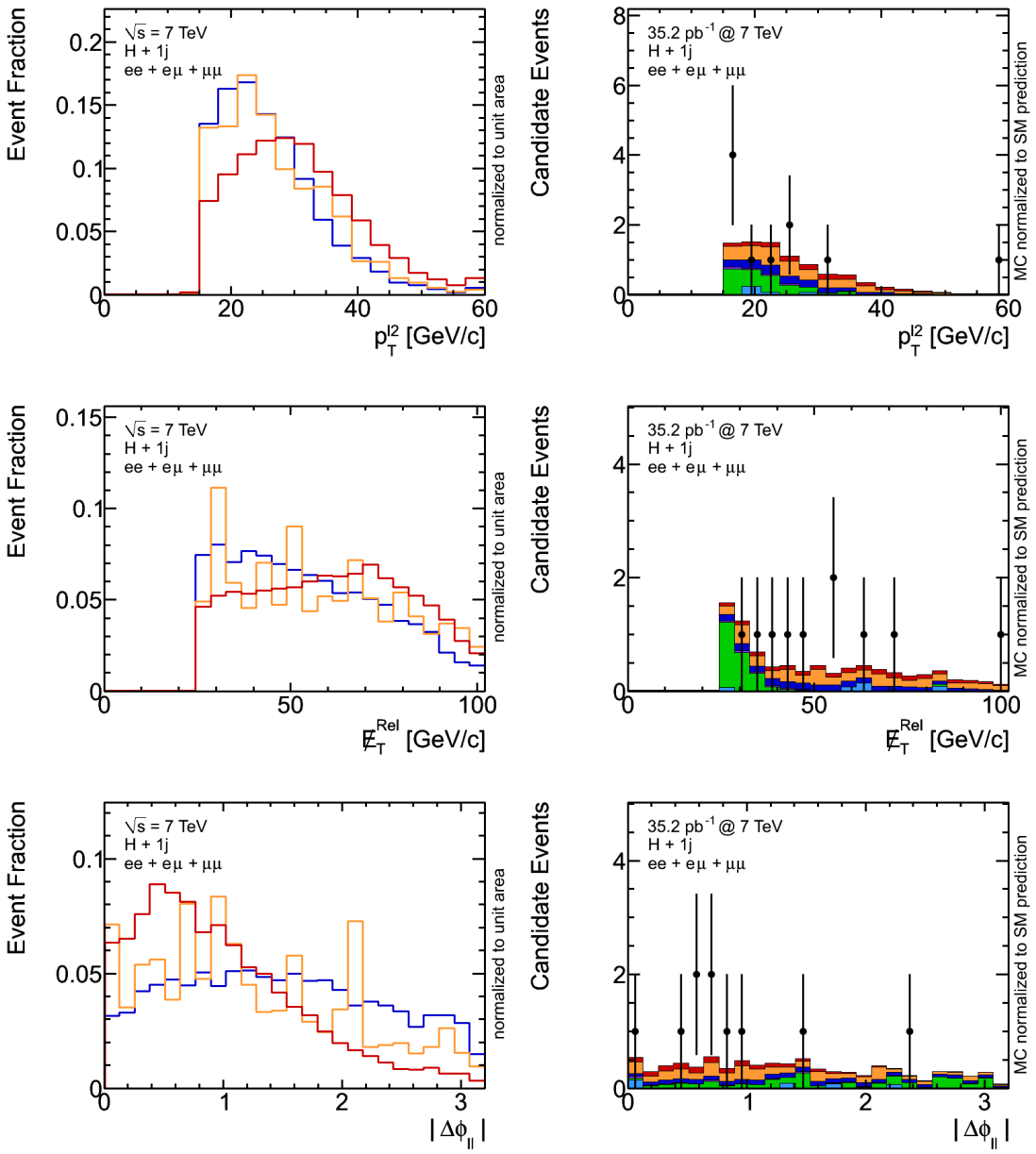


Figure A.2.: *The p_{T} distribution of the second lepton, the $\not\!\!E_{\mathrm{T}}^{\mathrm{Rel}}$ and the angular distribution $|\Delta\phi_{\ell\ell}|$ in the $H + 1j$ analysis for a Higgs boson mass of 170 GeV/c^2. The column on the left hand side shows the kinematic distributions normalized to unit area, while the validation of the modeling is presented in the column on the right hand side.*

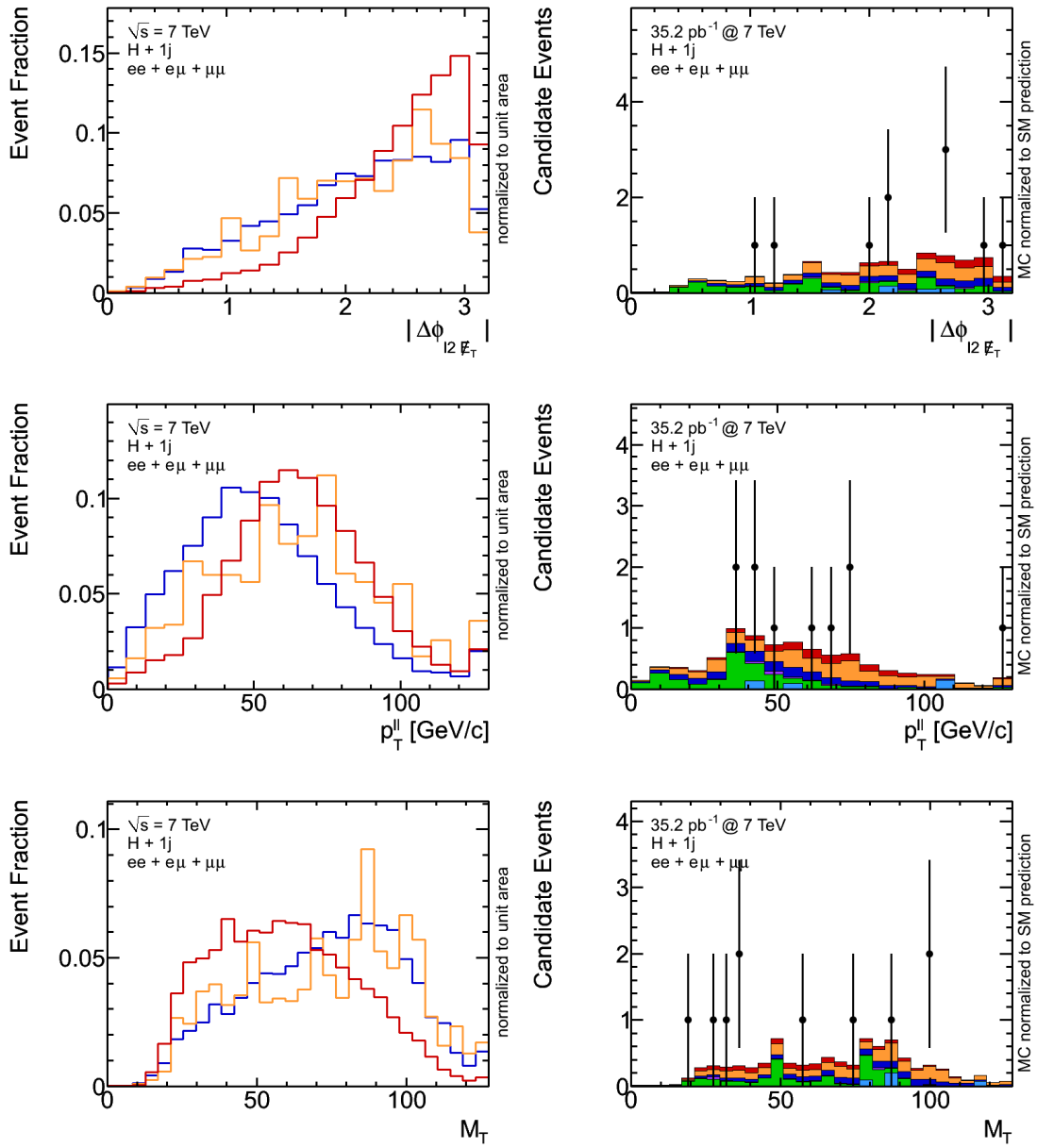


Figure A.3.: *The $|\Delta\phi_{\ell 2 \not\!\!E_{\mathrm{T}}}|$, $p_{\mathrm{T}}^{\ell\ell}$ and M_{T} distributions in the $H+1j$ analysis for a Higgs boson mass of 170 GeV/c^2. The column on the left hand side shows the kinematic distributions normalized to unit area, while the validation of the modeling is presented in the column on the right hand side.*

Danksagung

Mein besonderer Dank geht an Herrn Professor Dr. Wolfgang Wagner für das Aufnehmen in die Arbeitsgruppen in Karlsruhe und Wuppertal, das Ermöglichen des des Aufenthaltes am CERN und seine umfassende Betreuung während der gesamten Zeit. Durch die hervorragende Zusammenarbeit mit ihm und in der von ihm geformten Arbeitsgruppe habe ich sehr viel gelernt.

Ein großes Dankeschön gebührt Dr. Dominic Hirschbühl der mir stets mit Rat und Tat zur Seite stand: von der ersten Programmzeile bis zum Korrekturlesen der letzten Zeile dieser Arbeit. Dr. Philipp Sturm gilt mein Dank für die tolle Zusammenarbeit und das Mitpromovieren.

I would like to acknowledge Professor Jianming Qian for his kind agreement to be the second referee of my thesis. Dr. Klaus Hamacher und Dr. Daniel Wicke möchte ich für das Korrekturlesen der Arbeit danken, sie haben wesentlich zu Verbesserungen beigetragen.

Des Weiteren danke ich allen Mitgliedern der Arbeitsgruppe für das gute Arbeitsklima und die Hilfsbereitschaft, insbesondere Julia Fischer und Simon Köhlmann für ihre vielen Fragen und Astrid Eichler und Phillipp Tepel für die gute Stimmung in unserem Büro. Beim Sekretariat und dem Admin-Team bedanke ich mich für die gute Zusammenarbeit.

Meiner Schwester und meinen Eltern danke ich sehr herzlich für ihre Unterstützung und dafür, dass sie immer für mich da sind.