

DIPLOMARBEIT

**Search for the Higgs Boson in the Channel
 $WH \rightarrow e\nu b\bar{b}$ using Boosted Decision Trees
in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV
with the ATLAS Experiment**

Daniel Büscher



Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität Freiburg

**Search for the Higgs Boson in the Channel
 $WH \rightarrow e\nu b\bar{b}$ using Boosted Decision Trees
in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV
with the ATLAS Experiment**

DIPLOMARBEIT

vorgelegt von

Daniel Büscher

9. November 2012

Prof. Dr. Karl Jakobs
Fakultät für Mathematik und Physik der
ALBERT-LUDWIGS-UNIVERSITÄT
Freiburg im Breisgau

Contents

1	Introduction	1
2	Theoretical background	3
2.1	The Standard Model	3
2.1.1	Elementary particles	3
2.1.2	Fundamental interactions	4
2.1.3	Spontaneous symmetry breaking	7
2.2	Phenomenology of proton-proton collisions	9
2.3	Phenomenology of the Higgs Boson	12
2.4	Current knowledge of the Higgs boson	14
3	The ATLAS experiment	17
3.1	The Large Hadron Collider	17
3.2	The ATLAS detector	17
3.2.1	Inner detector	18
3.2.2	Calorimeters	19
3.2.3	Muon spectrometer	20
3.2.4	Trigger system	21
4	Reconstruction and identification of physical objects	23
4.1	Tracks and vertices	23
4.2	Electrons	23
4.3	Muons	24
4.4	Jets	25
4.5	Flavour tagging	26
4.6	Energy calibration	27
4.7	Lepton isolation	28
4.8	Missing transverse energy	28
5	Event selection	31
5.1	Data samples	32
5.2	Monte Carlo samples	32
5.2.1	Pile-up correction	34
5.2.2	Vertex correction	35
5.2.3	Trigger weighting	35
5.2.4	Flavour tagging efficiencies	35
5.2.5	Electron scale factors	36
5.2.6	Final event weights	36
5.3	Event pre-selection	36
5.4	Electron selection	36
5.5	Muon selection	37

5.6	Jet selection	37
5.7	Overlap removal	37
5.8	Missing transverse energy	37
5.9	Signal selection	37
5.10	Event yields	38
6	The $t\bar{t}$ background	43
6.1	Truth composition	43
6.2	Discriminating variables	47
7	Boosted Decision Trees	53
7.1	BDT construction and configuration	53
7.2	Signal BDTs	55
7.3	Background BDTs	56
8	Background estimation	59
8.1	The multijet background	59
8.2	Background Fits	62
9	Limit calculation	67
9.1	Sources of systematic uncertainties	67
9.2	Normalisation systematics	68
9.3	Shape systematics	72
9.4	Limits	76
9.5	Comparison to the reference analysis	77
10	Summary	81
A	Appendix	83
	Bibliography	89

The Standard Model of particle physics, elaborated in the 1960s, describes all known elementary particles and their interactions with great precision. It states the existence of fermions, spin $1/2$ particles, which form the known matter. The interactions among the fermions are mediated by bosons, spin one particles. The Standard Model predicted the existence of several particles and all of them were experimentally observed until the turn of the century, except for one: the Higgs boson.

The Standard Model symmetry, depicted in Quantum Field Theory (QFT), does not allow for mass terms for both bosons and fermions. The masses are constructed in the Standard Model by a spontaneous symmetry breaking via a scalar field: the *Higgs field*. This *Higgs mechanism*¹ predicts aspects of the Standard Model, e.g. the relation between the masses of the W and the Z bosons. In addition, the Higgs mechanism postulates the existence of a new boson, the *Higgs boson* with spin zero. The direct search for the Higgs boson, however, one of the largest scientific ventures of the last decades, is not yet completed.

In 2009 a new particle accelerator started its operation: the Large Hadron Collider (LHC). The LHC is a circular proton-proton collider with a design centre-of-mass energy of $\sqrt{s} = 14$ TeV. In 2011 and 2012 it was operated at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV, respectively. The protons are collided at several points of the ring where four detectors are located. Two of them, the ATLAS and the CMS detectors, are multi-purpose detectors that are designed for the search for new particles.

The operation of the LHC in 2011 and 2012, about 50 years after the postulation of the Higgs mechanism, made the discovery of a new boson with a mass of about 126 GeV possible. This boson is consistent with the Higgs boson predicted in the Standard Model. However, this observation was made mainly in the decay channels $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^* \rightarrow llll$ and $H \rightarrow WW^* \rightarrow l\nu l\nu$. These channels provide a very clean signature in the detector, due to photons or leptons in the final state. Therefore, they can be well separated from the backgrounds. Other channels, such as $H \rightarrow b\bar{b}$ and $H \rightarrow \tau\tau$, mainly produce jets in the detector, which cannot be distinguished very well from the overwhelming hadronic background. Nonetheless, these two channels are very important since they probe the decay of the Higgs boson to fermions, while the three former channels only provide evidence for the decay to other bosons

The $H \rightarrow b\bar{b}$ decay mode has the largest branching fraction for a Higgs boson with a mass of 126 GeV. Therefore, it is particularly important to constrain the coupling of the Higgs boson to Standard Model particles.

The search for the process $H \rightarrow b\bar{b}$ is possible in a specific production mode of the Higgs boson: the associated production with a vector boson. The decay products of the vector boson allow for a better background rejection and for triggering of the events, especially if they are leptons. Therefore, this process delivers a higher sensitivity than other production modes, although it occurs less frequently. This thesis concentrates on the associated production with a W boson that decays into an electron and the corresponding neutrino: $WH \rightarrow e\nu b\bar{b}$.

¹The mechanism of spontaneous symmetry breaking in field theories was independently published by Anderson, Brout, Englert, Guralnik, Hagen, Higgs, Kibble and 't Hooft.

However, even with the presence of the electron the background is still very large compared to the expected signal.

The standard approach to reduce the backgrounds is a simple selection using specific variables for the extraction of the signal, which is e.g. used in a published result of the ATLAS collaboration in Ref. [1]. However, such a *cut-based* analysis is barely capable of reaching the sensitivity for a discovery of a Standard Model Higgs boson in $H \rightarrow b\bar{b}$ based on the data collected in 2011 and 2012 [2]. A *multivariate analysis* (MVA) can improve the sensitivity. MVAs aim to optimise the selection by taking multiple variables and their correlations into account. Several kinds of MVAs exist with different capabilities. *Boosted Decision Trees* (BDTs) are used in this thesis due to their easy configuration and robust separation power.

BDTs provide a one dimensional output classifier, which is calculated from the input variables. In the output distribution the signal is separated from the backgrounds in a presumably optimal way. For the estimation of the background normalisations and the systematic uncertainties novel BDT-based techniques are developed. The output distributions of various BDTs are used to calculate limits on the Higgs boson cross section by fitting the signal and backgrounds to data. The fit uses various systematic uncertainties as nuisance parameters. This is done for Higgs-boson masses in the mass range of 110 to 135 GeV.

The analysis presented here is based on the data collected by ATLAS in 2011 at $\sqrt{s} = 7$ GeV, using the electron channel, $WH \rightarrow e\nu b\bar{b}$. It is not expected to exclude the existence of a Standard Model Higgs boson in the considered mass range or even find evidence for its existence. However, this study can be generalized to the full dataset of 2012 and to include the muon channel. Including also the associated production with Z bosons the BDT-based approach is expected to improve the cut-based results and might be able to deliver evidence for $H \rightarrow b\bar{b}$.

In Chapter 2 a brief introduction to the Standard Model (SM) is given, the phenomenologies of proton-proton collisions and Higgs boson decays are discussed and recent results in the searches for the Higgs boson are presented. The ATLAS experiment is described in Chapter 3 and the reconstruction of physical objects in Chapter 4. A basic event selection is presented in Chapter 5. The most important reducible background, the $t\bar{t}$ background, is investigated in Chapter 6. The Boosted Decision Trees are presented in Chapter 7. In Chapter 8 the background normalisations are estimated. And finally, in Chapter 9 systematic uncertainties are discussed, limits on the Higgs boson cross sections are calculated and compared to cut-based results.

2.1 The Standard Model

The *Standard Model* (SM) of particle physics is a description of the elementary particles and their interactions. It was developed in the 1960s and 1970s and proved very successful in explaining the results of existing experiments and predicting the outcome of later ones. The most important predictions that were experimentally confirmed are the existence of the W^\pm and Z bosons, later found at CERN, as well as the top-quark, later found at the Tevatron¹.

The Standard Model postulates the existence of elementary particles that form the known matter, the *fermions*, which can be classified into *leptons* and *quarks*. Their interactions are described by gauge fields mediated by the exchange of *bosons*. With these ingredients the model is able to depict three of the four known forces: the *electromagnetic*, the *weak* and the *strong* force. Only the *gravitational* force is not included.

The theoretical description is given by quantum field theory, where the key assumption is an invariance under local phase transformations, also called “gauge” invariance. An underlying symmetry of $SU(3) \times SU(2) \times U(1)$ is found and will be motivated in the following. The description follows to a large extent the textbooks of Refs. [3–5].

2.1.1 Elementary particles

Fermions Fermions are particles with spin $1/2$. Experiments show, that they exist in three generations (*flavour*). The generations differ only by the particle masses, increasing steeply from the first to third generation. A summary of all known fermions is given in Table 2.1. Listed is the classification in the three generations, in quarks and leptons and their electromagnetic charge.

All ordinary matter consists only of the first generation particles: up- and down-quarks, forming protons and neutrons, and the electron. The other fermions can occur in high energy processes, e.g. in cosmic rays or in collider experiments, such as the LHC.

¹Proton-antiproton collider at Fermilab, operating 1987-2011

	1 st generation		2 nd generation		3 rd generation		charge [e]
quarks	up	u	charm	c	top	t	$\frac{2}{3}$
	down	d	strange	s	bottom	b	$-\frac{1}{3}$
leptons	e-neutrino	ν_e	μ -neutrino	ν_μ	τ -neutrino	ν_τ	0
	electron	e	muon	μ	tau	τ	-1

Table 2.1: *The fermions divided into quarks and leptons and the three generations with their electric charge. For each particle exists an anti-particle with opposite charge.*

interaction	boson	charge [e]	mass [GeV]
electromagnetic	photon γ	0	0
strong	gluon g	0	0
weak	W^\pm boson	± 1	80.4
	Z boson	0	91.2

Table 2.2: *The bosons with their electric charge and mass.*

Bosons The interactions of fermions are mediated by bosons, which are integer spin particles. The vector bosons with spin 1 are shown in Table 2.2. They consist of

- the well-known *photon* (γ), mediator of the electromagnetic force. The photon is massless and stable, therefore has an infinite range. It interacts with all particles carrying electric charge.
- the *gluons* (g), carrier of the strong force. Gluons interact with coloured particles, the quarks and the gluons themselves, as described in the next Section. They exist in 8 different coloured states. Also gluons are massless, but since QCD forbids the existence of unbound coloured particles the strong interaction is short-ranged.
- the W^\pm and Z bosons, mediators of the weak force, which will be introduced in the next Section. They are massive particles, therefore the weak force has only a small range. They interact with all particles carrying weak charge.

The masses of the particles are generated by the Higgs mechanism. It predicts the only elementary spin 0 particle, the Higgs boson. It is the only particle of the Standard Model that has not been experimentally detected, but current observations at the LHC yield strong indications for its existence, see Section 2.4. All quarks and charged leptons are massive, but neutrinos are assumed to be massless in the following. This agrees with direct mass measurements to a very good level, but it should be mentioned, that neutrino oscillations are observed. These are indications for neutrinos having small, but non-zero masses [6].

2.1.2 Fundamental interactions

The Standard Model consists of several theories describing the dynamics of the elementary particles. Historically, the first part is the Quantum Electro Dynamics (QED). The strong force is described by the Quantum Chromo Dynamics (QCD) and the weak force is unified with QED in the Electroweak (EW) model [7–9], also called GSW model named after Glashow, Salam and Weinberg.

Quantum Electro Dynamics QED is mathematically described by a Lagrangian density \mathcal{L} (simply *Lagrangian* in the following), derived using symmetry arguments, especially the invariance under local phase transformations. A free fermion (spin $\frac{1}{2}$) is described by the Lagrangian

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \quad (2.1)$$

where ψ is the Dirac spinor of the fermion field, $\bar{\psi} = \psi^\dagger \gamma^0$ is the adjoint, γ^μ are the Dirac gamma matrices and $\partial_\mu = \frac{\partial}{\partial x_\mu}$ are the partial derivatives. The Dirac equation,

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \quad (2.2)$$

is the corresponding equation of motion. The Lagrangian is requested to be invariant under local phase transformations of the form

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha(x)}\psi(x), \quad (2.3)$$

where the local phase is given by $\alpha(x)$, depending on space and time. These transformations form the abelian unitary group $U(1)$, since $e^{i\alpha(x)}$ can be written as a 1×1 -matrix U with $U^\dagger U = 1$. One finds that the first term of the Lagrangian (2.1) is not invariant under this transformation, since

$$\partial_\mu \psi \rightarrow \partial_\mu \psi' = e^{i\alpha(x)} \partial_\mu \psi + ie^{i\alpha(x)} \psi \partial_\mu \alpha(x) \quad (2.4)$$

The invariance is established by introducing an additional field A_μ , transforming as

$$A_\mu \rightarrow A'_\mu = A_\mu + \frac{1}{e} \partial_\mu \alpha(x), \quad (2.5)$$

and replacing ∂_μ with the covariant derivative D_μ :

$$D_\mu = \partial_\mu - ieA_\mu. \quad (2.6)$$

One can show that

$$D_\mu \psi \rightarrow D'_\mu \psi' = e^{i\alpha(x)} D_\mu \psi \quad (2.7)$$

holds and the Lagrangian becomes

$$\begin{aligned} \mathcal{L} &= i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi \\ &= \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu \psi A_\mu. \end{aligned} \quad (2.8)$$

Thus it was necessary to introduce the *gauge field* A_μ to preserve local phase invariance. This field couples to the fermion field with the coupling constant e . By introducing a kinetic term, using the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, the gauge field A_μ is identified as the physical photon field and the Lagrangian of QED is complete:

$$\mathcal{L}_{\text{QED}} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi + e\bar{\psi}\gamma^\mu \psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (2.9)$$

The individual terms are the kinetic energy and mass of the fermion, the coupling between fermion and photon fields and the kinetic energy of the photon. The local phase invariance forbids the introduction of a mass term in the form $\frac{1}{2}m^2 A_\mu A^\mu$ for the photon field. Hence this formalism requests the photon to be massless, which is in agreement with all experiments.

Quantum-Chromo Dynamics QCD describes the interaction of quarks and gluons. Quarks are fermions too, but an additional degree of freedom is introduced: the *colour charge*, which exists in the states red (r), green (g) and blue (b). Therefore the simple Dirac spinors are replaced by vectors of three spinors denoting the quarks:

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix} \quad (2.10)$$

	generation			quantum numbers			
	1 st	2 nd	3 rd	I	I_3	Y	$Q [e]$
quarks	$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{2}{3}$
	u_R	c_R	t_R	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{3}$
	d_R	s_R	b_R	0	0	$\frac{4}{3}$	$\frac{2}{3}$
				0	0	$-\frac{2}{3}$	$-\frac{1}{3}$
leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	$\frac{1}{2}$	$\frac{1}{2}$	-1	0
				$\frac{1}{2}$	$-\frac{1}{2}$	-1	-1
	e_R^-	μ_R^-	τ_R^-	0	0	-2	-1

Table 2.3: The fermions in the electroweak model with their quantum numbers.

Together with the eight gauge fields G_μ^a , $a = 1, \dots, 8$, the gluons, the Lagrangian of QCD is derived:

$$\mathcal{L}_{\text{QCD}} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - g_s\left(\bar{\psi}\gamma^\mu\frac{\lambda_a}{2}\psi\right)G_\mu^a - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}. \quad (2.11)$$

It is invariant under transformations of the non-abelian unitary group $SU(3)$. Here g_s denotes the strong coupling constant, λ_a are the eight Gell-Mann-matrices and $G_{\mu\nu}^a$ is the field strength tensor, written as

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - gf_{abc}G_\mu^b G_\nu^c. \quad (2.12)$$

The structure constants f_{abc} fulfill $\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = if_{abc}\frac{\lambda_c}{2}$. The last term of the field strength tensor is due to the $SU(3)$ being non-abelian and does not have an equivalent in QED. It enables the gluons to interact with themselves. As for QED the local phase invariance requests the force carriers, here gluons, to be massless.

Electroweak model All fermions are subject to the weak force, manifested e.g. in nuclear β -decays. The interaction is carried out by exchange of W^\pm bosons, what is called *charged current* interaction, and the Z boson, called *neutral current* interaction. While the exchange of the W^\pm bosons changes the flavour of quarks, there are no flavour-changing neutral currents (FCNC) observed at tree level.

Experiments show, that the charged weak current only couples to left-handed fermions. Therefore these are assigned to $SU(2)_L$ doublets with isospin $I_3 = \pm\frac{1}{2}$, while the right-handed fermions are described by $U(1)_Y$ singlets with $I = 0$. The hypercharge Y is introduced, which is related to the electric charge via $Q = I_3 + \frac{Y}{2}$ (see below). An overview of the fermions and their quantum numbers is given in Table 2.3. The weak eigenstates of the quarks, denoted by d' , s' and b' , are a mixture of the mass eigenstates. The mixing is described by the CKM mechanism [10].

The left-handed isospin doublets χ_L and right-handed singlets ψ_R are transformed as

$$\chi_L(x) \rightarrow \chi'_L(x) = e^{i\alpha_a(x)\tau_a} e^{i\beta(x)Y} \chi_L, \quad (2.13)$$

$$\psi_R(x) \rightarrow \psi'_R(x) = e^{i\beta(x)Y} \psi_R \quad (2.14)$$

and form a $SU(2)_L \times U(1)_Y$ symmetry. Here $\alpha(x)$ and $\beta(x)$ are the local phases, τ_a with $a = 1, 2, 3$ are the generators of $SU(2)_L$ and Y is the weak hypercharge operator, generating $U(1)_Y$. The covariant derivative is given by

$$D_\mu = \partial_\mu + igW_\mu^a \frac{\tau_a}{2} + ig' B_\mu \frac{Y}{2} \quad (2.15)$$

where g is the coupling constant of the $SU(2)_L$ gauge fields W_μ^a and g' is the coupling constant of the $U(1)_Y$ gauge field B_μ . The resulting electroweak Lagrangian is

$$\mathcal{L}_{\text{EW}} = i\overline{\chi}_L^i \gamma^\mu D_\mu \chi_L^i + i\overline{\psi}_R^i \gamma^\mu D_\mu \psi_R^i - \frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \quad (2.16)$$

where a summation over the three isospin doublets and six singlets is done (index i). The field tensors are given by

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g\varepsilon_{abc} W_\mu^b W_\nu^c \quad (2.17)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad (2.18)$$

where the completely antisymmetric tensor ε_{abc} denotes the structure constants of $SU(2)$, while the structure constant of $U(1)$ is zero. This enables the W_μ^a fields to interact with themselves, while the B_μ field only couples to fermions.

The physical fields are given by linear combinations of W_μ^a and B_μ ,

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (2.19)$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \quad (2.20)$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \quad (2.21)$$

where the weak mixing angle θ_W is introduced. It relates the coupling constants via $\cos\theta_W = g/\sqrt{g^2 + g'^2}$ and $\sin\theta_W = g'/\sqrt{g^2 + g'^2}$. By rewriting the Lagrangian in terms of the physical fields and comparing the A_μ components to the photon field of QED (Eq. 2.9) one obtains the relations

$$e = g \sin\theta_W = g' \cos\theta_W \quad (2.22)$$

$$\text{and} \quad Q = I_3 + \frac{Y}{2}. \quad (2.23)$$

As for QED and QCD the local phase invariance forbids the introduction of mass terms for the bosons. In addition mass terms of the form $-m\overline{\psi}\psi$ for the fermions are forbidden in the EW sector. This is in conflict with experiments, where the Z and W^\pm bosons are found to be massive, see Table 2.2.

2.1.3 Spontaneous symmetry breaking

In the electroweak model the local phase invariance requests the fermions and bosons to be massless particles. However, experiments show that the vector bosons W^\pm and Z^0 are massive with large masses, as shown in Table 2.2. This contradiction can be solved by a spontaneous symmetry breaking, introduced by the *Higgs mechanism* [11–13]. The scalar Higgs field Φ , a weak isospin doublet containing complex scalar fields with hypercharge $Y = 1$ and its potential $V(\Phi)$ are postulated as

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (2.24)$$

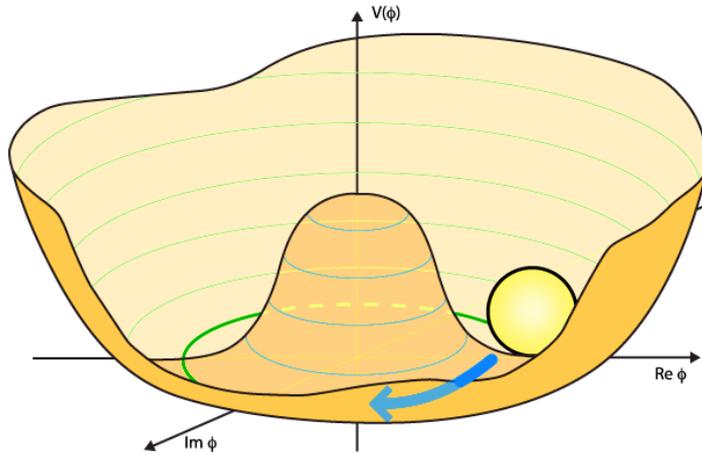


Figure 2.1: Illustration of the Higgs potential in the complex plane [14]. The potential has a minimum at $\phi \neq 0$, causing the spontaneous symmetry breaking as described in the text.

The Lagrangian

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 \quad (2.25)$$

is invariant under $SU(2)_L \times U(1)_Y$ phase transformations. The potential, parametrised as

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad (2.26)$$

has for $\mu^2 < 0$ and $\lambda > 0$ a local minimum at $\Phi \neq 0$. This is illustrated for a single complex scalar field ϕ in Fig. 2.1. One point of the local minimum,

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (2.27)$$

is chosen as ground state, where $v = \sqrt{-\mu^2/\lambda}$ depicts the vacuum expectation value. Since $v > 0$, the symmetry of the group is spontaneously broken. The field is now parametrised as

$$\Phi(x) = \frac{e^{i\tau_a \theta_a(x)/v}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (2.28)$$

where $\theta_a(x)$ ($a = 1, 2, 3$) and $h(x)$ are real fields. The corresponding particle to the $h(x)$ field is called Higgs boson and has spin 0. The exponential containing the θ_a fields (*Goldstone bosons*) is eliminated in the Lagrangian due to the local phase invariance and does not have a physical meaning.

Substituting this parametrization in the Lagrangian (2.25) the term

$$\begin{aligned} & \left| \left(ig \frac{\tau_a}{2} W_\mu^a + ig' \frac{Y}{2} B_\mu \right) \Phi_0 \right|^2 \\ &= \left(\frac{1}{2} vg \right)^2 W_\mu^+ W^{\mu-} + \frac{1}{2} \frac{1}{\cos^2 \theta_W} \left(\frac{1}{2} vg \right)^2 Z_\mu Z^\mu + 0 A_\mu A^\mu \end{aligned} \quad (2.29)$$

is found, where the notation $|\dots|^2 = (\dots)^\dagger (\dots)$ was used and the physical bosons occur in the final expression. The obtained terms can be identified as mass terms of the physical vector bosons and the relations

$$m_W = \frac{1}{2} vg, \quad m_Z = \frac{m_W}{\cos \theta_W} \quad \text{and} \quad m_\gamma = 0 \quad (2.30)$$

for the boson masses are found. Thus, a strong prediction for the relation between M_W and M_Z was found, which was experimentally verified by the experiments at the LEP collider [15]. Further, a mass term for the Higgs boson can be found, which leads to $M_H = \sqrt{-2\mu^2}$. This contains the free parameter μ and does not have predictive power.

In a similar way, one can generate fermion masses, using their coupling to the Higgs boson (*Yukawa coupling*). This is described by an additional $SU(2)_L \times U(1)_Y$ invariant component to the Lagrangian,

$$\mathcal{L}_{\text{Yukawa}} = -G_l^{ij} \bar{L}_L^i \Phi_C l_R^j - G_d^{ij} \bar{Q}_L^i \Phi_C d_R^j - G_u^{ij} \bar{L}_L^i \Phi_C u_R^j + \text{h.c.}, \quad (2.31)$$

where \bar{L}_L^i (\bar{Q}_L^i) are the lepton (quark) isospin doublets, l_R^j (d_R^j , u_R^j) are the lepton (down-/up-type quark) singlets and the Higgs field is parametrised as

$$\Phi_C(x) = \sqrt{\frac{1}{2}} \begin{pmatrix} v + h(x) \\ 0 \end{pmatrix}. \quad (2.32)$$

In summary the Lagrangian of the Standard Model is composed as:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \quad (2.33)$$

This Lagrangian is invariant under local phase transformations of the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group. The Higgs boson interacts with all massive weakly-interacting particles including itself. As can be seen from (2.30) and derived from (2.31) the coupling of the bosons and fermions to the Higgs is proportional to their mass.

2.2 Phenomenology of proton-proton collisions

The predictions of the Standard Model (and models beyond that) can be tested with scattering experiments such as the ATLAS experiment, located at the Large Hadron Collider (LHC), which is a proton-proton collider. The advantage of using protons instead of electrons, the most common elementary particle used in collider experiments, is the higher mass of the proton. This enables the collider to reach higher centre of mass energies. However, protons have a substructure which has to be modeled, since the Standard Model only describes the interaction of elementary particles.

To describe a scatter process its *cross section*, σ , has to be defined. It is a measure for the frequency of occurrence of such a process, also called *event*. From the experimental point of view the cross section is defined as

$$\sigma = \frac{R}{\mathcal{L}}, \quad (2.34)$$

where $R = dN/dt$ is the rate of events and \mathcal{L} is the *instantaneous luminosity*. The commonly used unit of σ is 1 barn = 1 b = 10^{-28} m². Using the *integrated luminosity*, $L_{\text{int}} = \int \mathcal{L} dt$, and the total number of events, N , the cross section can be rewritten as

$$\sigma = \frac{N}{\int \mathcal{L} dt} = \frac{N}{L_{\text{int}}}. \quad (2.35)$$

The luminosity depends on the parameters of the particle accelerator. For a circular accelerator with a rotation frequency f_r and two colliding beams with n_b bunches of particles each, the luminosity can be written as

$$\mathcal{L} = f_r \frac{n_b N_1 N_2}{A}, \quad (2.36)$$

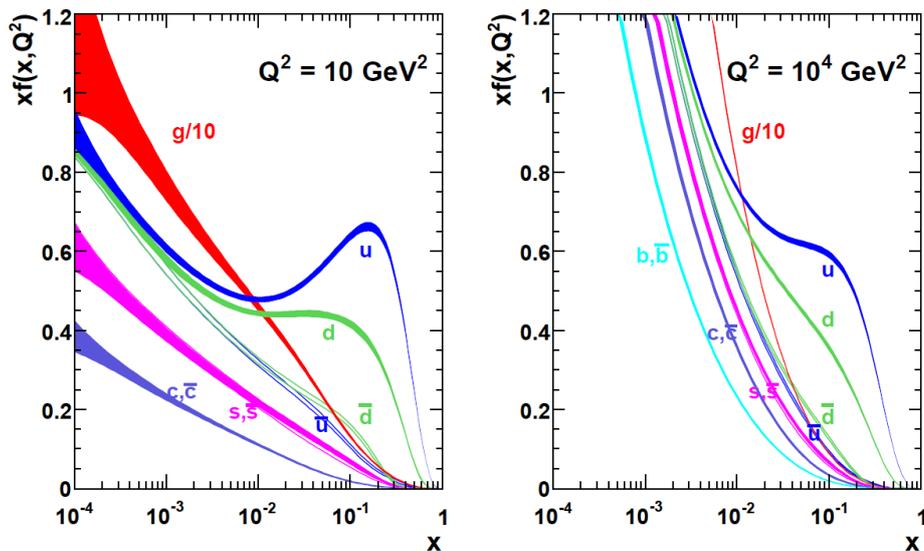


Figure 2.2: Parton distribution functions for $Q^2 = 10 \text{ GeV}$ (left) and $Q^2 = 10^4 \text{ GeV}$ (right) calculated at NNLO [16]. Shown are the function for the valence quarks, sea quarks and gluons inside the proton. The widths of the bands represent the uncertainty.

where A is the transverse area of the beam and $N_{1,2}$ are the numbers of particles inside each bunch of beam 1 and 2, respectively. Assuming $N_1 = N_2 = N_p$ and a Gaussian shape of the beam with widths σ_x and σ_y in the transverse plane the luminosity becomes

$$\mathcal{L} = f_r \frac{n_b N_p^2}{4\pi\sigma_x\sigma_y}. \quad (2.37)$$

If the total cross section for inelastic processes, σ_{inel} , is known, the luminosity can be determined as

$$\mathcal{L} = \frac{R_{\text{inel}}}{\sigma_{\text{inel}}} = \frac{\mu n_b f_r}{\sigma_{\text{inel}}}, \quad (2.38)$$

where μ is the average number of interactions per bunch crossing.

From Eq. 2.35 the number of events for a defined process can be predicted:

$$N = \sigma \int \mathcal{L} dt. \quad (2.39)$$

Therefore a theoretical computation of the cross section is needed. This is provided by the Standard Model of elementary particles. Protons are composite particles, consisting of valence quarks (uud), sea quarks of any flavour and gluons. This structure is described by Parton Distribution Functions (PDFs). These experimentally determined PDFs, $f_{q_i}(x_i, Q^2)$, give the probability to find a parton q_i with a momentum fraction x_i of the total proton momentum in a collision with a momentum transfer Q^2 .

The PDFs measured at a specific Q_0^2 can be transferred to a different Q^2 by using the DGLAP equations [17–19]. In Fig. 2.2 the PDFs for $Q^2 = 10 \text{ GeV}$ and $Q^2 = 10^4 \text{ GeV}$ are shown. One can see that for high x the up and down valence quarks of the proton dominate, while for low x the gluons dominate and the sea quarks have similar contributions as the

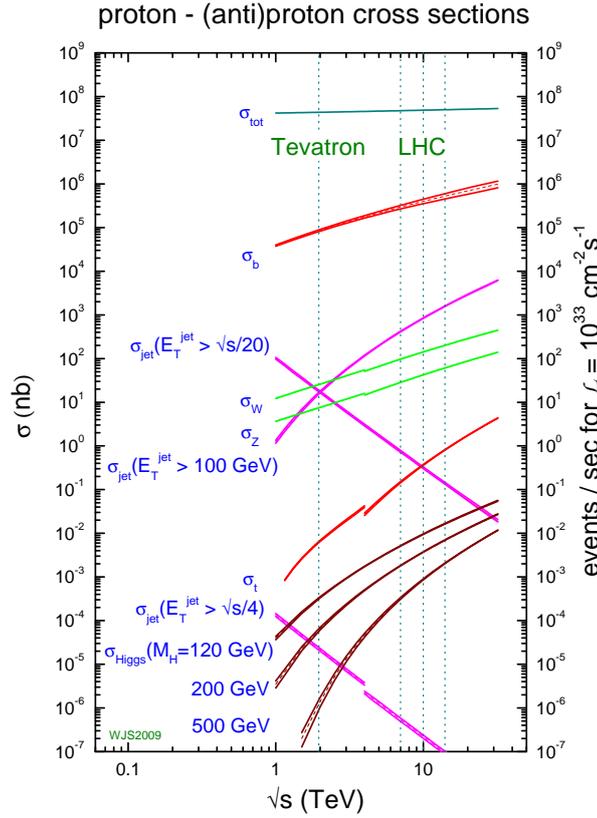


Figure 2.3: Predicted cross sections (left axis) and expected number of events for $\mathcal{L} = 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ (right axis) for various processes occurring in proton-(anti)proton collisions as function of \sqrt{s} [20]. The dotted lines represent the \sqrt{s} of the Tevatron (1.96 GeV) and the LHC (7 GeV for 2010/2011 and 14 GeV design energy) and a dotted line for 10 GeV is shown for comparison.

valence quarks. With rising Q^2 the fraction of sea quarks and gluons rises in comparison to the valence quarks.

The total cross section for the process $pp \rightarrow Y$ is calculated by integrating the PDFs multiplied with the partonic cross section, $\hat{\sigma}$, and summing over all parton combinations leading to the final state Y :

$$\sigma(pp \rightarrow Y) = \sum_{i,j} \int dx_i dx_j f_{q_i}(x_i, Q^2) f_{q_j}(x_j, Q^2) \hat{\sigma}(q_i q_j \rightarrow Y), \quad (2.40)$$

where the partons are denoted by q . The partonic cross section can be written as

$$\hat{\sigma} = \int \frac{|\mathcal{M}|^2}{F} dQ, \quad (2.41)$$

where \mathcal{M} is the matrix element for the transition from initial to final state, F is the particle flow and dQ is the phase space factor of the given kinematics. The probability amplitude is given by $|\mathcal{M}|^2$. The matrix element \mathcal{M} can be calculated using the Lagrangian densities and perturbation theory. For the strong force the hadronic cross section can be written as a power expansion series of the strong coupling constant α_s :

$$\hat{\sigma} = \hat{\sigma}_0 \left(1 + c_1 \alpha_s + c_2 \alpha_s^2 + \dots + c_n \alpha_s^n \right), \quad (2.42)$$

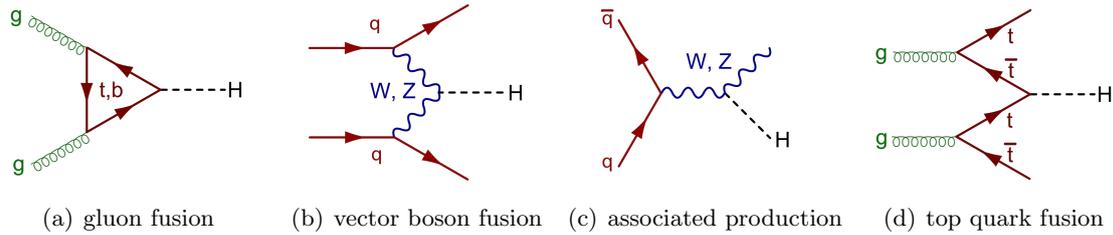


Figure 2.4: Feynman diagrams of the most important production modes of the Higgs boson at the LHC [24]. The process with the highest cross section is the gluon fusion (a), followed by the vector boson fusion (b), the associated production with a vector boson (c) and the top quark fusion (d).

where $\hat{\sigma}_0$ denotes the leading order. The cross sections for electroweak processes can be expressed the same way in orders of the electroweak coupling constant. The numerical calculation of such a series can only take a limited number of orders into account, which is denoted by LO (leading order), NLO (next-to leading order), NNLO (next-to next-to leading order) and so on.

In Fig. 2.3 various predicted cross sections are shown. One can see that processes like W and Z boson production have cross sections that are several orders of magnitude below the total inelastic cross section of $\sigma_{\text{tot}} \approx 70$ mb. The process of interest in this thesis, the Higgs boson production, is again several order below that with $\sigma_{\text{Higgs}} \approx 15$ pb for $m_H = 120$ GeV and $\sqrt{s} = 7$ TeV. Therefore less than one event out of one billion is expected to contain a Higgs boson, making its detection a great challenge.

The remaining partons of the protons that did not take part in the hard scattering are not in a stable compound anymore. They hadronise and cause additional objects in the detector and are called *underlying event*. The fragmentation of the partons is described by *fragmentation functions*. These functions cannot be calculated by perturbation theory, but have to be modeled based on experimental results. Common models are the *cluster fragmentation* [21] and the *string fragmentation* [22].

An additional effect, making the measurements at the LHC even harder, is called *pile-up*. Pile-up is caused by the fact, that each bunch contains a large number of protons and that the total cross section for hard scattering is very high. Therefore up to 20 proton-proton interactions occurred in a single bunch crossing in 2011 [23].

In conclusion the phenomenology of proton-proton collisions is far more complex than for elementary particles. All the discussed effects have to be modeled by the Monte Carlo simulation, which is used to compare the expectation to the data. Since this modeling is not perfect, several corrections have to be applied, which will be discussed in Sections 4 and 5.

2.3 Phenomenology of the Higgs Boson

Production modes The production of the Higgs boson at the LHC can be classified into several modes. The four most important, depicted in Fig. 2.4, are discussed here. The modes differ in their cross sections, which are shown in Fig. 2.5, and their phenomenology.

The dominant mode is the gluon fusion. Its cross section is over 10 times higher than the next important mode. This is due to the large fraction of gluons inside the protons [16]. Two gluons fuse through a quark loop to a Higgs boson. Since the coupling of the Higgs boson is

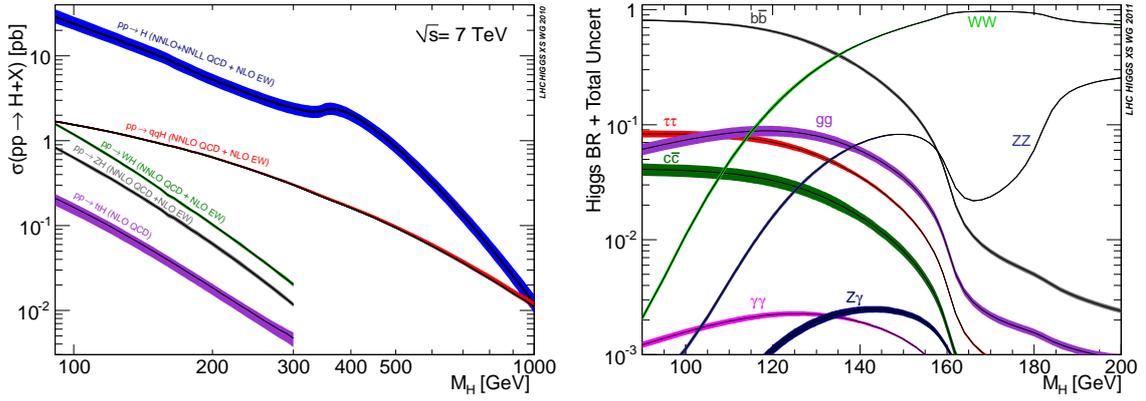


Figure 2.5: Left plot: cross sections for the various Higgs boson production modes at $\sqrt{s} = 7$ GeV as function of m_H [25]. $pp \rightarrow H$ denotes the gluon fusion, $pp \rightarrow qqH$ the vector boson fusion, $pp \rightarrow WH$ and $pp \rightarrow ZH$ the associated production and $pp \rightarrow ttH$ the top quark fusion. Right plot: branching ratios of the Higgs boson decay as function of m_H [25].

proportional to the mass, a top-quark loop is preferred.

The vector boson fusion is the next important production mode. Here, two W or Z bosons are radiated by two quarks inside the protons and fuse to a Higgs boson. The two quarks produce jets inside the detector in opposite directions along the beam. These can be reconstructed and used to reject backgrounds.

The production in association with a vector boson has an even smaller cross section. Here, two quarks fuse to a W or Z boson, which radiates the Higgs boson. The vector boson can decay into leptons. These constitute additional objects in the detector, which can be reconstructed with high efficiencies. This allows for a very good background suppression.

The final production mode is the top quark fusion. The decays of the top quarks produce various objects in the detector, therefore this production mode is the most complex one regarding the reconstruction.

Branching-ratios The branching ratios of the Higgs boson decay are dictated by the masses of the decay products, since the coupling of the Higgs boson is proportional to these masses. The branching ratios are shown in Fig. 2.5 as function of the Higgs boson mass, m_H .

For $m_H < 135$ GeV the decay $H \rightarrow b\bar{b}$ is preferred, since the b-quarks are the heaviest elementary particles that are kinematically accessible. Pairs of heavier decay products, namely W and Z boson and top-quark pairs, are suppressed, since they are in summary heavier than the Higgs boson itself.

For $m_H > 135$ GeV the process $H \rightarrow WW^*$ has the largest branching ratio. The production of virtual vector bosons is allowed due to their intrinsic width.

Even though the decay to two b-quarks is preferred in the low mass region, this channel is not the most sensitive one. The decays into vector bosons produce cleaner signatures in the detector and therefore can be better extracted from the backgrounds.

The mass resolution in the various decay channels differ greatly. The best resolution is reached with a few GeV in $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow lll$, since the energies of the photons and leptons can be measured well and the invariant mass can be fully reconstructed. The processes $H \rightarrow b\bar{b}$ has a worse resolution due to larger uncertainties on the reconstructed jets caused by the b-quarks. Decay channels that contain neutrinos in the final state, e.g.

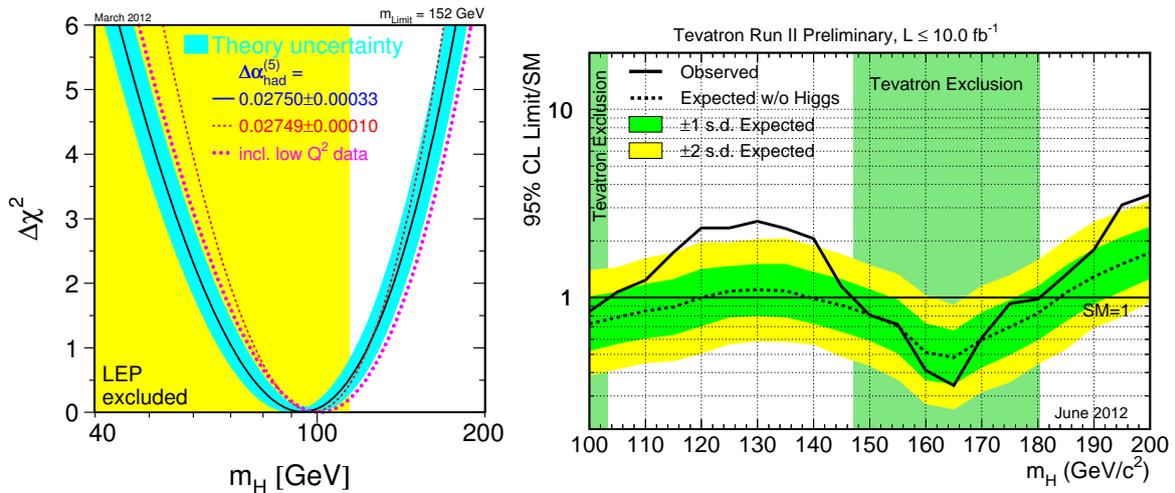


Figure 2.6: Exclusion limits on the Standard Model Higgs boson from LEP, Tevatron. In the left plot the exclusion from LEP up to 114.4 GeV is shown in yellow and the curve denotes the $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ from the fit using loop corrections [27]. The right plot shows the 95 % CL exclusion limit on the Higgs boson cross section from the Tevatron. The ranges below 113 GeV and from 147 GeV to 180 GeV are excluded. An excess is visible around 130 GeV [28].

$H \rightarrow WW^* \rightarrow l\nu l\nu$, do not allow for a complete reconstruction of the mass and therefore have even less resolution.

2.4 Current knowledge of the Higgs boson

In the Standard Model the Higgs boson mass is a free parameter and has to be measured experimentally. Only weak limits on the mass exist from perturbative unification theories, which state a lower limit of about 100 GeV and an upper limit of a few hundred GeV [26]. In this range the LHC is expected to either exclude the existence of the Higgs boson or find evidence for it.

Before the operation of the LHC the mass of the Higgs boson was already constrained by the experiments that were located at the LEP collider. The Higgs mass was excluded for $m_H < 114.4$ GeV at a 95 % confidence level by direct searches [15]. Additionally, it was possible to restrain the mass by the results of electroweak precision measurements. The theoretical predictions for these results involve loop corrections, which are sensitive to the Higgs boson mass. A fit was performed and the resulting χ^2 as function of m_H is shown in Fig. 2.6. One can see, that low masses just above the excluded range of the LEP experiments are preferred.

Also the Tevatron was able to restrain the Higgs boson mass by direct searches. It excludes the ranges below 113 GeV and from 147 GeV to 180 GeV, as shown in Fig. 2.6. Also a small excess was found around 130 GeV, which is mainly driven by $H \rightarrow b\bar{b}$ [28].

In 2011 and 2012 the LHC delivered a large amount of data, which made it possible to constrain the mass of the Higgs boson to a very narrow region around 126 GeV, as shown in Fig. 2.7. The mass was not only constrained, but also an excess in the limits on the Standard Model Higgs boson cross section of over 5σ is observed [29, 30].

This excess is driven by three channels, $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^* \rightarrow lll$ and $H \rightarrow WW^* \rightarrow l\nu l\nu$, which are the most sensitive ones. The $H \rightarrow b\bar{b}$ channel, which is investigated in this thesis,

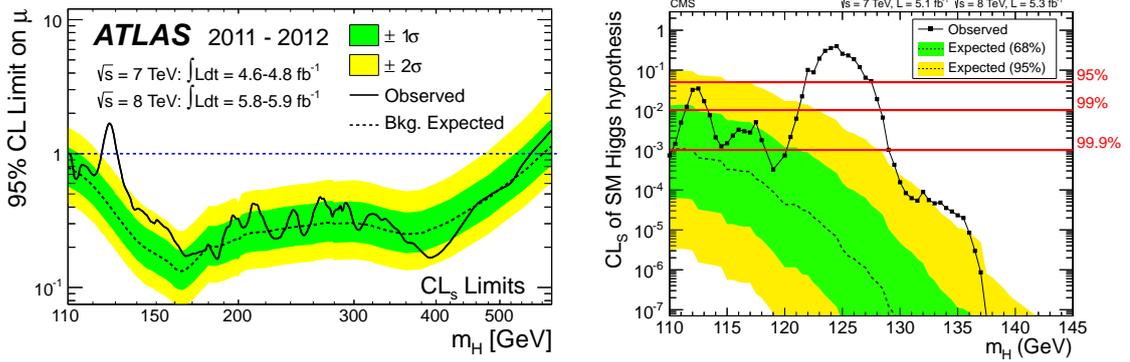


Figure 2.7: Most recent exclusion limits on the Standard Model Higgs boson cross section from the ATLAS and CMS experiments. The left plots shows the limits in the mass range from 100 to 600 GeV from ATLAS [29] and the right plot shows the limits in the low-mass region from CMS [30]. Both experiments are excluding the range from 110 to several hundred GeV with exception of a narrow region around 126 GeV. There, a large excess is visible, which corresponds to a significance of over 5σ [29, 30].

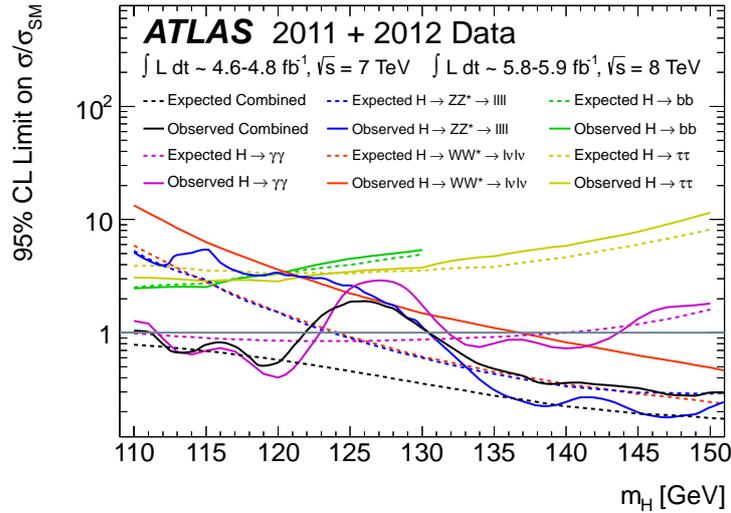


Figure 2.8: Most recent 95% CL exclusion limits on the Standard Model Higgs boson cross section from the ATLAS experiment in various channels [29]. The channels that contribute to the observed excess are the ones with the highest sensitivity: $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^* \rightarrow llll$ and $H \rightarrow WW^* \rightarrow l\nu l\nu$. The channel $H \rightarrow b\bar{b}$, which is investigated in this thesis, is shown in green and has a lower sensitivity.

does not contribute to these limits. It has a lower sensitivity and no excess is visible, as shown in Fig. 2.8: observed limits for $H \rightarrow b\bar{b}$ agree with the expected limits of the background-only hypothesis. The expected limit for the Higgs boson cross section with $m_H = 126$ GeV is about four times the Standard Model expectation.

The observed excess in the combined channels is assessed as the observation of a new boson. But it is not yet clear, if it is a Standard Model Higgs boson. Since the decay of the boson to

two photons is observed, spin one is excluded. But for the Higgs boson it still has to be shown that the spin is zero. Additionally, as long as the discovery is not confirmed in $H \rightarrow b\bar{b}$ and other channels it is not possible to conclude that the observed boson couples to all massive particles, which is required for the Higgs boson.

In case the new boson is in fact a Standard Model Higgs boson the determination of its total decay width will rely essentially on the $H \rightarrow b\bar{b}$ channel, since the branching ratio of $H \rightarrow b\bar{b}$ for $m_H = 126 \text{ GeV}$ is the largest one with about 55%, as shown in Fig. 2.5.

This thesis presents a multivariate analysis of the $H \rightarrow b\bar{b}$ channel, which increases the sensitivity with respect to the cut-based approach, and may help to find evidence in this channel.

3

The ATLAS experiment

The ATLAS¹ Experiment is a multi-purpose detector operating at the Large Hadron Collider (LHC), located at CERN near Geneva, Switzerland. Its main goal is the detection of yet undiscovered particles, such as the Higgs boson and particles predicted in physics models beyond the Standard Model. The following description of the experiment is based on Refs. [31, 32].

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a circular collider, mainly used for proton-proton collisions. It was built inside a tunnel with 27 km circumference and about 100 m below ground-level, formerly housing the Large Electron-Positron Collider (LEP). It was planned in the 1990s and installed after the shutdown of the LEP in the year 2000. The LHC is able to accelerate Protons and heavy ions in two beams running in opposite directions. The beams are bent by dipole magnets, generating fields up to 8.3 T. These magnets are super-conducting and cooled down to 1.9 K by liquid helium. A maximum beam energy of 3.5 TeV for Protons was reached in 2011, thus a maximum centre-of-mass energy of $\sqrt{s} = 7$ TeV. In 2012 the centre-of-mass energy was increased to $\sqrt{s} = 8$ TeV and the design energy of $\sqrt{s} = 14$ TeV is planned to be reached in 2014.

Protons are fed into the LHC with an energy of 450 GeV, pre-accelerated by a chain of linear and circular accelerators. The beams are brought to collision at linear sections of the tunnel, where the experiments are located. These are the LHCb², ALICE³ and the two multi-purpose detectors CMS⁴ and ATLAS. The latter is described in more detail in the following.

3.2 The ATLAS detector

The ATLAS detector is a multi-purpose detector, aiming at the discovery of new particles, such as the Higgs boson and physics beyond the Standard Model in proton-proton collisions. Its cylindrical 4π -design with high density of detection material up to low scattering angles allows for near-complete reconstruction of hard scattering processes.

The common coordinate system used for measurements is centered in the beam crossing, with the z -axis along the beam, the y -axis pointing upwards and the x -axis pointing towards the center of the LHC ring. The azimuthal angle ϕ is measured in the x - y -plane and the polar angle θ is the angle to the z -axis. A commonly used variable, the pseudo-rapidity η , is defined as $\eta = -\ln(\tan(\theta/2))$.

The detector consists of several sub-detectors: the inner detector with a coverage of $|\eta| < 2.5$, the calorimeters with $|\eta| < 4.9$ and the muon system with $|\eta| < 2.7$, see Fig. 3.1. The

¹A Toroidal LHC Apparatus

²Large Hadron Collider beauty

³A Large Ion Collider Experiment

⁴Compact Muon Solenoid

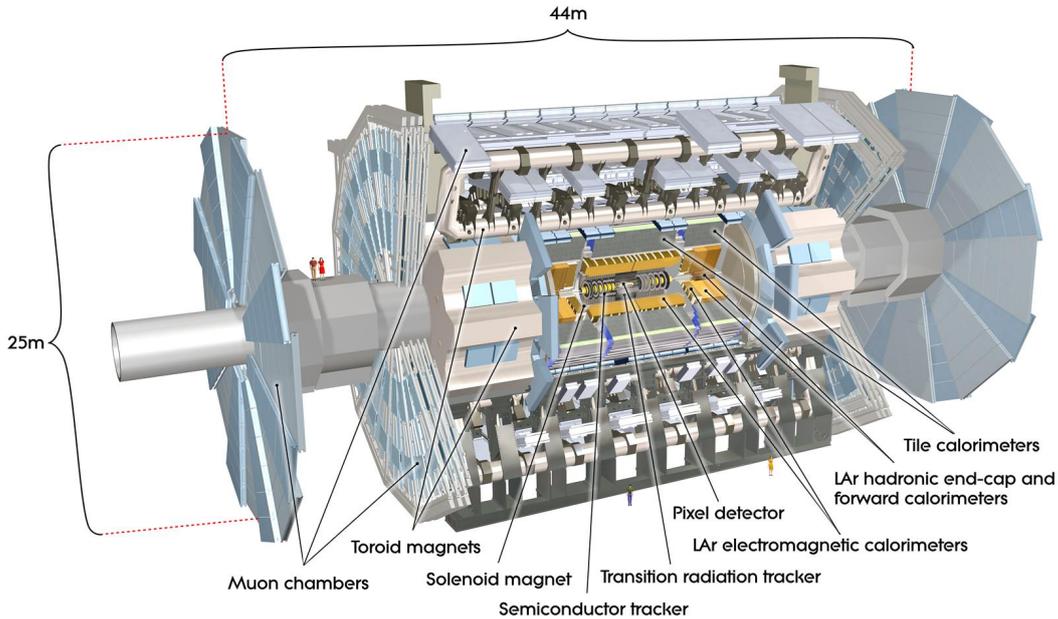


Figure 3.1: Schematic view of the ATLAS detector [32]. The visible components are the muon system, the toroidal and solenoid magnets, the calorimeters and the inner detector.

ATLAS detector has a multi-stage triggering system to cope with the high interaction rates occurring at the LHC. The sub-systems are described in more detail in the following.

3.2.1 Inner detector

The inner detector consists of pixel detectors, semiconductor trackers (SCTs) and transition radiation trackers (TRTs), see Fig. 3.2. The pixel detector allows precise track measurements of charged particles close to the interaction point. This enables reliable reconstruction of primary interaction and secondary vertices. Together with the surrounding tracking detector and a solenoid, generating a magnetic field of 2 T, the track momenta can be determined.

Pixel detector The pixel detector is the closest detector to the interaction point, the innermost layer being located at $R = 50.5$ mm. It has layers cylindrically ordered around the beam in the central range (barrel) and others radially in the end-caps. Particles typically pass three of these layers. The pixels have a minimum size of $R-\phi \times z = 50 \times 400 \mu\text{m}^2$ and a total of 80 million read-out channels is reached. The track resolution reached is about $10 \mu\text{m}$ in the $R-\phi$ -plane and about $115 \mu\text{m}$ along z (barrel) or along R (end-caps) [32].

Semiconductor tracker The SCT surrounds the pixel detector. It consist of four layers of silicon strip detector modules in the barrel and nine layers in the end-caps. Particles typically pass four of these layers. The modules in the barrel have two layers of silicon that are slightly rotated against each other to allow for the determination of the position along the strips. The resolution is about $17 \mu\text{m}$ in the $R-\phi$ -plane and about $580 \mu\text{m}$ along z (barrel) and along R (end-caps) [32].

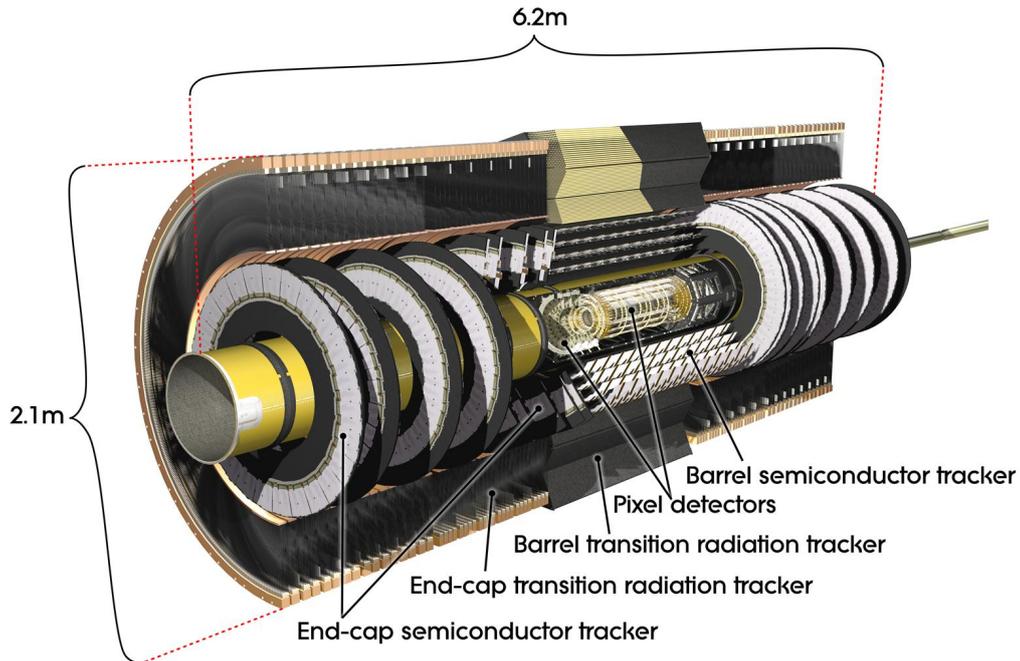


Figure 3.2: Sketch of the inner detector with its pixel detectors, SCT and TRT [32].

Transition radiation tracker The TRT surrounds the SCT and is the outermost part of the inner detector. It consists of gas-filled drift-tubes along the z -axis (144 cm long) and radial in the end-caps (37 cm long). In the barrel it provides only a position measurement in the R - ϕ -plane with a resolution of about $130 \mu\text{m}$. The TRT provides on average 36 coordinate measurements over the radial distance from 55 cm to 108 cm [32]. By having a larger radius than the other trackers it improves significantly the resolution of the momentum measurement. In addition it provides potential for particle identification, since the transition radiation is proportional to $\gamma = E/mc^2$ [33]. This is especially important for electrons, since they are by far the lightest stable charged particles and therefore emit the most transition radiation.

Since the components of the inner detector are very close to the interaction point, they have to cope with high radiation, which damages the detector material. To reduce this damage the pixel detector and the SCT are cooled down to about -7°C , while the TRT operates at room temperature. At the design luminosity of the LHC a total of about 85 kW of heat has to be removed from the inner detector, which is done with an elaborate cooling system [32].

3.2.2 Calorimeters

The calorimeter system of the ATLAS detector consists of an electromagnetic ($|\eta| < 3.2$) and a hadronic calorimeter ($|\eta| < 4.9$), covering the full ϕ -range, see Fig. 3.3. Both are sampling calorimeters, which means they use layers of active detector material and absorber material. Their main purpose is the determination of particle energies. The calorimeters are *non-compensating*, meaning they have a different response to electromagnetic showers than to hadronic showers.

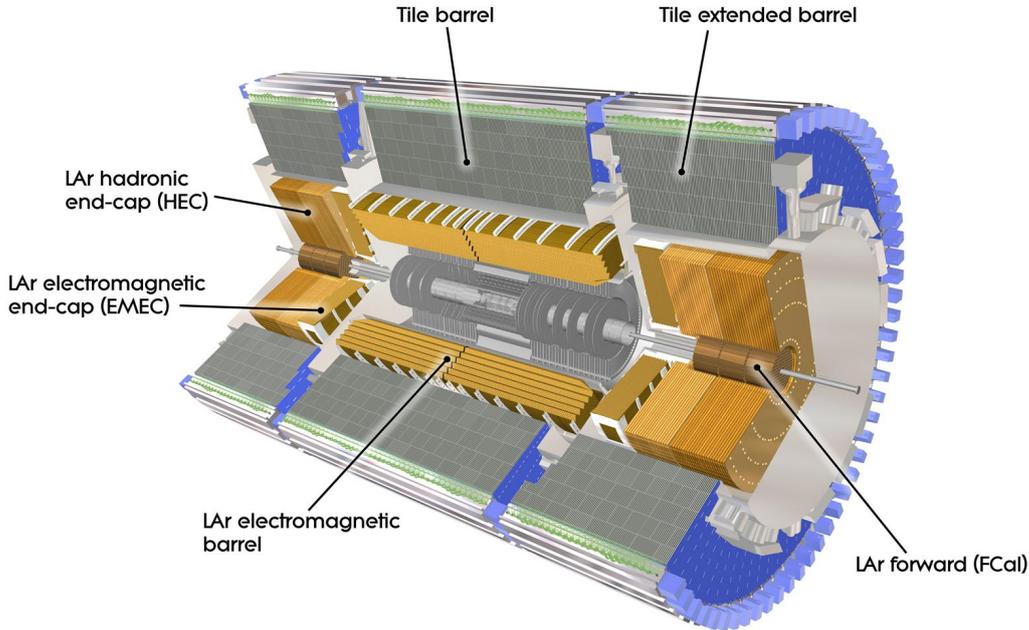


Figure 3.3: Sketch of the ATLAS calorimeters with their electromagnetic and hadronic subcomponents [32].

Electromagnetic calorimeter The electromagnetic (EM) calorimeter is a liquid argon (LAr) sampling calorimeter with lead absorbers. It is divided into a barrel region ($|\eta| < 1.475$) and end-caps (EMEC, $1.375 < |\eta| < 3.2$). The end-caps are divided into a central region ($1.375 < |\eta| < 2.5$) with finer resolution and a forward region ($2.5 < |\eta| < 3.2$) with coarser resolution. The calorimeter has kapton electrodes in accordion geometry, enabling a full ϕ -coverage without gaps. It has a design energy resolution of $\sigma_E/E = 10\%/\sqrt{E(\text{GeV})} \oplus 7\%$ [32], using the notation $a \oplus b = \sqrt{a^2 + b^2}$.

Hadronic calorimeter Also the hadronic calorimeter has a sampling structure. In the central region ($|\eta| < 1.7$) it is called tile calorimeter and uses scintillating tiles and steel as absorber. The Hadronic End-Caps (HEC) at the outer region ($1.5 < |\eta| < 3.2$) use LAr detector material and copper as absorber. The design resolution of the hadronic barrel and end-caps is $\sigma_E/E = 50\%/\sqrt{E(\text{GeV})} \oplus 3\%$ [32].

Forward calorimeter The forward calorimeter (FCal) is dedicated to the very forward region ($3.1 < |\eta| < 4.9$) and has only three layers of absorber material, the first one is copper, optimised for electromagnetic measurements and the other two are made of tungsten for hadronic measurements. The design resolution of the FCal is $\sigma_E/E = 100\%/\sqrt{E(\text{GeV})} \oplus 10\%$ [32].

3.2.3 Muon spectrometer

The muon spectrometer is the outermost subdetector system of the ATLAS experiment. It is the largest part of the detector. Its toroidal magnet system does not only give ATLAS its

name, but also is responsible for the distinct look of the detector, see Fig. 3.1.

It is divided into three regions: barrel ($|\eta| < 1.4$), end-caps ($1.6 < |\eta| < 2.7$) and transition region ($1.4 < |\eta| < 1.6$). It provides triggering capabilities for muons up to $|\eta| < 2.4$. Both, barrel and end-caps have their own toroid magnet system providing the bending fields. The magnets have air-cores and a toroidal geometry. This provides a magnetic field, which bends charged particles orthogonal to the bending direction in the inner solenoid. The eight coils produce a field integral of $\int B_{\perp} dl = 1.5$ to 5.5 Tm in the barrel, 1.0 to 7.5 Tm in the end-caps. For track detection *monitored drift tubes* are installed in the full η -range up to 2.7 . For $2.0 < |\eta| < 2.7$ *cathode strip chambers* provide better spatial resolution and can cope with higher signal rates. The trigger system consists of *resistive plate chambers* in the barrel and *thin gap chambers* in the end-caps. The muon spectrometer allows for precise measurement of muon momenta: the nominal resolution is $\sigma_{p_T}/p_T = 10\%$ for $p_T = 10$ GeV [32].

3.2.4 Trigger system

The design event rate provided by the LHC is about 40 MHz. This rate has to be reduced drastically to about 200 Hz to be able to record the most important events. The decision, which events to record, is taken by a multi-level triggering system [32].

The level-1-trigger (L1) uses the trigger chambers of the muon system and the full calorimeter system with reduced granularity to search for objects with high transverse energy, e.g. electrons, photons, muons, jets or missing transverse energy. The L1-trigger is hardware-based and a decision is taken within $2.5 \mu\text{s}$. It reduces the event rate to $75 - 100$ kHz. The remaining events are passed to the level-2-trigger system (L2) together with η and ϕ information about Regions of Interest (RoIs).

The L2-trigger is software based and uses the full detector information inside the RoIs and brings the event rate down to about 3.5 kHz. The event filter (EF) is the final triggering stage, reducing the rate to about 200 Hz. It uses the full detector information and needs about 4 s for a decision. All events passing the EF are recorded for further offline analysis. A typical event has a size of 1 to 2 MB of raw data.

4 Reconstruction and identification of physical objects

The reconstruction and identification of physical objects in ATLAS is carried out by the software framework ATHENA [34]. It analyses the raw detector-data of all events that passed the trigger system. Dedicated algorithms for several object types are implemented, which will be discussed in the following.

4.1 Tracks and vertices

To identify charged particles and assign them to certain decays their tracks and vertices have to be reconstructed in the inner detector [35]. The first step is to reconstruct three-dimensional space points corresponding to the energy deposits of charged particles in pixel, SCT and TRT detectors.

Track seeds are generated using three or more hits in the pixel detector and the first layer of the SCT. These are extended to the full SCT and a first track fit is applied. Some quality cuts are done to reject fake tracks. Then these tracks are associated to drift circles in the TRT. At last the fit is redone using the full information of the three detector systems.

A different algorithm starts the track reconstruction from the TRT. This approach does not need hits in the innermost detector layers and therefore has sensitivity to particles stemming from secondary vertices distant to the primary interaction point.

After building tracks vertices can be reconstructed. For this, the tracks are extrapolated and intersections are looked for. Fits are carried out, trying to combine the tracks into vertices, applying certain quality criteria. The vertices that are compatible with the beam spot are called primary vertices (PVs). Since the total interaction cross section at the LHC is very large, one expects multiple PVs for each event, which are called pile-up. One of these vertices is identified as *signal vertex*, which should correspond to the hardest interaction. This vertex is required to have the highest sum of squared transverse momenta,

$$\sum_{i=1}^{N_{\text{track}}} p_{T,i}^2, \quad (4.1)$$

where the index i runs over all outgoing tracks of the vertex.

4.2 Electrons

Electrons are light charged particles and therefore leave a track in the inner detector and deposit most of their energy in the EM calorimeter. There are three independent algorithms for reconstruction. Two of them use the information of the inner detector and the calorimeter, one is dedicated to high- p_T and the other to low- p_T electrons. They are limited to the central region $|\eta| < 2.5$. A third one is dedicated to forward electrons and uses only the calorimeter information up to $|\eta| < 4.9$. In this thesis the high- p_T central electrons are used as signal and veto electrons and additionally the forward electrons are used for vetoing.

Central electrons are reconstructed by starting with clusters in the second layer of the EM calorimeter. The clusters are built by a sliding window algorithm. The window consists of 3×5 calorimeter cells, corresponding to about 0.025×0.025 in the $\eta \times \phi$ -space. For electrons with $E_T > 15$ GeV the cluster reconstruction efficiency is expected to be close to 100 % [36].

Tracks from the inner detector are matched to these clusters. The criteria for the matching are $\Delta\eta < 0.1$ and $\Delta\phi < 0.1$ in the bending direction of the solenoid magnet and $\Delta\phi < 0.05$ in the other direction. This accounts for bremsstrahlung, which reduces the bending radius due to energy loss. In case multiple tracks are matched to one cluster, those with pixel and SCT hits are preferred and the one with the closest distance to the cluster is chosen. After this matching the clusters are rebuilt using 3×7 cells in the barrel and 5×5 cells in the end-caps.

After the electrons are reconstructed the identification (ID) is applied to distinct them from other physics objects, mainly jets. This is done with several quality criteria, categorised into *loose++*¹, *medium++* and *tight++*. All categories use information from the inner detector and the calorimeters to calculate certain variables and apply cuts to them. The first category, *loose++*, has the least variables and the loosest cuts, therefore has the highest efficiency, but also has the highest contamination of jets in the sample. It uses mainly shower shape information in the middle layer of the EM calorimeter, energy deposits in the hadronic calorimeter, as well as track information from the inner detector. The *medium++* category tightens the previous requirements and additionally asks for hits in the b-layer² and uses particle ID information from the TRT. The last category, *tight++*, has the most and tightest cuts, the least identified electrons, but the cleanest sample.

The electron ID does not look at the electron isolation. This can be done in later analyses, using certain variables, such as the amount of energy deposits close to the electron.

4.3 Muons

Muons show a special behaviour in the detector. They are long-lived minimum ionizing particles. This means they lose only little energy passing the detector material, especially in the calorimeters. To achieve a reliable muon identification and precise momentum measurement a dedicated muon system (MS) was installed as the outermost part of the ATLAS detector.

There are four independent reconstruction methods implemented in ATHENA. First, for *standalone* muons only hits in the MS are requested together with energy deposits in the calorimeter. This method has the largest coverage in η , but has drawbacks. It cannot reconstruct low energy muons that do not reach the MS and is vulnerable to muons coming from secondary vertices, e.g. from Pion or Kaon decays. Finally, the momentum resolution of this approach is not optimal, since the information from the inner detector is not used.

The *segment tagged* muon algorithm is complementary. It reconstructs the muons by extrapolating tracks from the inner detector to the MS. If a track matches to a track segment in the MS it is identified as muon. The momentum measurement is done by using inner detector information only, therefore this approach is especially suited for low momentum muons.

The *combined* muons are reconstructed using the full inner detector and MS information. First, tracks from both systems are extrapolated to the beam axis. The matching parameter χ_{match}^2 is calculated from the track parameters \mathbf{T}_{ID} , \mathbf{T}_{MS} and covariance matrices \mathbf{C}_{ID} and \mathbf{C}_{MS} as:

¹The appendix “++” refers to improvements made on the 2010 electron ID criteria.

²The innermost layer of the pixel detector.

$$\chi_{\text{match}}^2 = (\mathbf{T}_{\text{MS}} - \mathbf{T}_{\text{ID}})^{\text{T}} (\mathbf{C}_{\text{MS}} + \mathbf{C}_{\text{ID}})^{-1} (\mathbf{T}_{\text{MS}} - \mathbf{T}_{\text{ID}}) \quad (4.2)$$

The track pairs with the lowest matching parameter below a certain threshold are searched for and identified as muon. Finally, the selected track pairs are combined in one fit, that uses the information of both detector systems.

A last algorithm identifies *calorimeter tagged* muons. These are tracks from the inner detector that match to energy deposits in the calorimeter. All above algorithm are implemented in two different frameworks, called STACO and MuID. In this thesis the STACO *combined* muons are used.

4.4 Jets

Quarks and gluons cannot exist in unbound states, as described in Section 2.2. They create collimated hadronic showers in the detector, called Jets. Two algorithm classes for their reconstruction are widely used: the *cone*- and the *cluster*-algorithms [37].

The *cone*-algorithms maximise the total energy of all objects inside a cone in η - ϕ -space. They only look at the distance of objects to group them together. These objects can be energy deposits in the calorimeter, reconstructed tracks, truth particles and so on. The distance of the objects are defined for each algorithm in a specific way, see below. The algorithm used in this thesis, the *anti- k_T* -algorithm, is of the *cluster* type. These use an additional stopping criterion for grouping objects together.

The *anti- k_T* -algorithm defines a measure of distance d_{ij} between objects in the calorimeter [38],

$$d_{ij} = \min \left\{ \frac{1}{k_{T_i}^2}, \frac{1}{k_{T_j}^2} \right\} \times \frac{(\Delta R_{ij})^2}{R^2}, \quad (4.3)$$

where k_{T_i} is the transverse momentum of object i and $\Delta R_{ij} = \sqrt{(\Delta y_{ij})^2 + (\Delta \phi_{ij})^2}$ is the geometrical distance between objects i and j with y being the rapidity. Typical values for the parameter R are 0.4 or 0.6. The stopping criterion d_{iB} is defined as the distance of object i to the beam axis B :

$$d_{iB} = \frac{1}{k_{T_i}^2} \quad (4.4)$$

The algorithm calculates a minimal distance d_{\min} in the event from the list of all objects i and j :

$$d_{\min} = \min\{d_{ij}, d_{iB}\} \quad (4.5)$$

If $d_{\min} = d_{iB}$ the object i is regarded as jet and is removed from the list. Otherwise ($d_{\min} = d_{ij}$) the objects i and j are grouped together. Then d_{\min} is recalculated. This procedure is repeated until all objects are defined as jets.

The *anti- k_T* -algorithm used in this thesis has $R = 0.4$ and uses *topological* clusters as input objects [39]. The corresponding jet collection is called *AntiKt4TopoEM*, where *EM* denotes the calibration to the electromagnetic scale, as described in section 4.6.

ε	$MV1_{\min}$
60%	0.905363
70%	0.601713
75%	0.404219
85%	0.071423

Table 4.1: *MV1 operating points for jets with $p_T > 15$ GeV and $\eta < 2.5$. The tagging efficiency ε for b-jets is given for the cut $MV1(\text{jet}) > MV1_{\min}$ [41].*

4.5 Flavour tagging

Several algorithms exist to identify the flavour of the hard-scattered quarks causing jets in data and Monte Carlo simulation. They are called *tagging algorithms*. Especially jets caused by b-quarks (*b-jets*) are of interest. They are distinct from c-jets or light jets (caused by gluons, u-, d- or s-quarks) by the fact that b-quarks form hadrons with relatively long lifetimes of typically 1.5 ps. This is long enough for them to travel several mm from the primary interaction before they decay and create *secondary vertices*. These can be reconstructed using the tracks from the inner detector. The pixel detector with its innermost *b-layer* delivers the highest precision for tracks close to the interaction point. Also c-quarks form hadrons with comparable, but slightly shorter lifetimes. The algorithms rely either on the *impact parameters* or on reconstruction of *secondary vertices*.

The impact parameter d_0 is the minimal distance of a track to the primary vertex in the transverse plane and z_0 in the longitudinal direction. A track coming from a secondary vertex will have large impact parameters. Commonly used is the impact parameter significance, $S_i = \{d_0/\sigma_{d_0}, z_0/\sigma_{z_0}\}$, where the impact parameters are divided by their uncertainties. The **IP3D** algorithm uses a two dimensional log-likelihood ratio,

$$W_{\text{IP3D}} = \sum_{i=1}^{N_{tr}} \ln \frac{b(S_i)}{u(S_i)}, \quad (4.6)$$

with $u(S_i)$ and $b(S_i)$ being the probability density functions (PDFs) for light and b-jets, respectively.

The algorithm **SV1** tries to reconstruct one secondary vertex for each jet, for which it tries to combine all tracks with impact parameters above a certain threshold. Several of the vertex parameters are combined in a likelihood ratio, such as the invariant mass and energy sum of all outgoing tracks. A more advanced algorithm, **JetFitter**, is trying to fit tracks into secondary vertices using the decay topologies of b- and c-hadrons in the jet. It is not relying on a single geometrical vertex and even secondary vertices with one track are can be reconstructed [40].

Several tagging algorithms can be combined to achieve better performance. **JetFitter-CombNN** uses a neural network to combine IP3D and JetFitter. **MV1** finally combines IP3D, SV1 and jetFitterCombNN into one neural network [42]. It is the most advanced tagger and is used in this thesis. Calibrated working points exist with defined b-tagging efficiencies at certain jet η and p_T , as shown in Table 4.1. In addition JetFitterCombNNc is used, which is a variant of jetFitterCombNN. Its neural network is trained against c-jets instead of light jets to get a higher c-rejection. The performance of the described algorithms is compared in Fig. 4.1. One can see, that MV1 has the highest rejection against light jets, while JetFitterCombNNc has the highest rejection against c-jets.

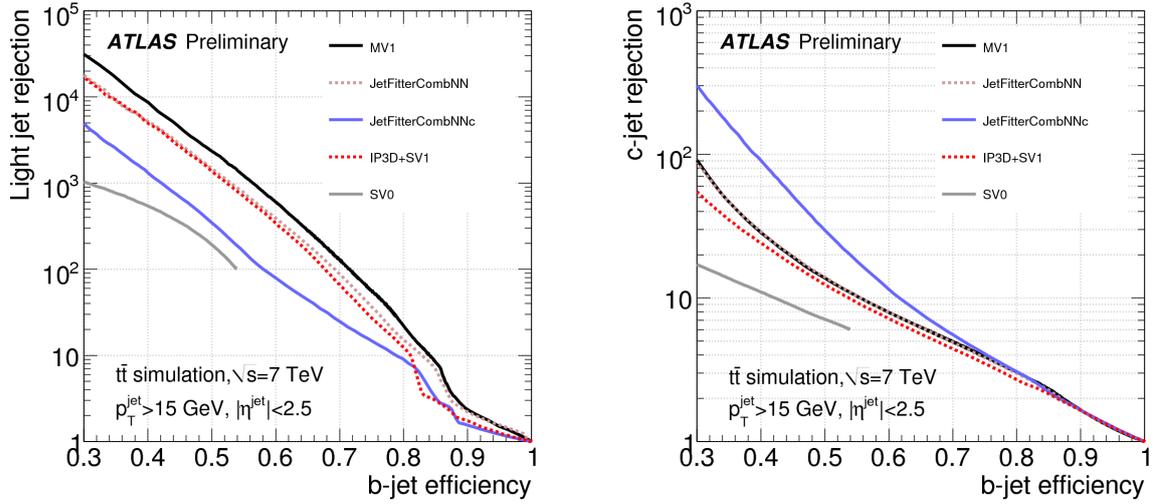


Figure 4.1: Light and c -jet rejection as function of the b -tagging efficiency for several flavour tagging algorithms. The cuts on the jets are $p_T > 15$ GeV and $|\eta| < 2.5$. [41]

4.6 Energy calibration

To achieve precise energy measurements of physical objects, the detector has to be calibrated. This is done for the electromagnetic calorimeter with electron test-beams and for the hadronic calorimeter with muon test-beams, as well as muons from cosmic rays [39]. Monte Carlo based corrections are applied to account for energy losses in the passive material of the detector. Additionally, collision data from 2010 is used to calibrate the energies further using known processes for the following physical objects:

Electrons Using the well-known invariant masses of the processes $Z \rightarrow e^+e^-$ and $J/\Psi \rightarrow e^+e^-$ one can show [43], that the dependence of the measured on the real energy is approximately linear: $E_{\text{meas}} = (1 + \alpha_i)E_{\text{real}}$ where i denotes bins of pseudo-rapidity. The parameters α_i are determined from the two processes and used to correct the measured energies.

Another method, that does not need the invariant mass of a decaying particle, is the comparison of the electron energy with its momentum, independently measured by the inner detector. Since the electron is expected to deposit all its energy in the calorimeter. Therefore the ratio E/p should be close to one, which was confirmed to the level of $\pm 2\%$ with an uncertainty of 0.3 to 1.6% [43].

The energy resolution can be determined using the invariant di-electron mass in $Z \rightarrow ee$. It was found that the resolution in Monte Carlo simulations is better than in data. This is accounted for by smearing the simulated energies with Gaussians of appropriate width.

Muons The measured muon energy can be validated by using the invariant mass of $Z \rightarrow \mu\mu$ decays [44]. It turns out that energy scale is already very precise. But as for electrons the simulated energy resolution is better than the measured. This is confirmed by comparing the momenta measured in the inner detector and the muon system of *combined* muons in $W \rightarrow \mu\nu$ events. This is again corrected for by smearing the simulated muon energies.

Jets Since jets are more complex objects than leptons, their measurement is somewhat more complicated. There are several processes that prevent the deposition of the full energy in the calorimeter.

- Neutrinos and muons can be produced in the jet and leave the detector losing only a small fraction of their energy.
- Electromagnetic showers are produced by numerous pions in the hadronic shower, e.g. via $\pi^0 \rightarrow \gamma\gamma$. These cause a different response than hadronic showers, since the calorimeter is non-compensating.
- High energy particles from the jets can leave the calorimeter.

Due to the complexity of these effects the necessary corrections cannot be simplified to scale factors. The calibration is done with respect to the electromagnetic scale and scale factors are applied to obtain jet energies at the hadronic scale. The energy resolution is already described well by Monte Carlo simulation and does not have to be corrected [45].

4.7 Lepton isolation

The isolation of reconstructed leptons is an important quantity to discriminate between leptons contained in jets and those coming for example from Z - or W -decays.

Two variables can be defined, one based on the isolation in the calorimeter,

$$E_T^{\text{cone}}(R_0) = -E_T^{\text{lep}} + \sum_{R < R_0} E_T^{\text{cell}}, \quad (4.7)$$

and another one based on the track-isolation,

$$p_T^{\text{cone}}(R_0) = -p_T^{\text{lep}} + \sum_{R < R_0} p_T^{\text{track}}, \quad (4.8)$$

where E_T^{cell} and p_T^{track} are the energy deposits and track momenta in a cone with $R \leq R_0$ around the lepton. For isolated leptons these quantities are assumed to be lower than for unisolated ones and cuts can be applied on them for separation.

While E_T^{cone} has the advantage of taking also neutral particles into account, the p_T^{cone} can distinguish between tracks coming from the signal vertex and those from other vertices, especially those from pile-up. In this thesis both isolation variables are used.

4.8 Missing transverse energy

Neutrinos cannot be detected in ATLAS, since they interact only weakly and the cross sections are extremely small. They can be reconstructed however using the transverse energy imbalance of the total event. In an ideal event the total transverse energy vectors of all particles should sum up to zero, but if the undetectable neutrinos are involved this is generally not the case. The absolute value of this energy is called missing transverse energy E_T^{miss} or MET . Unfortunately the z -component cannot be reconstructed, since the longitudinal momenta of the colliding partons are a-priori unknown.

The missing transverse energy can be written in its x - and y -components:

$$E_T^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2} \quad (4.9)$$

The components are composed of the energy deposits in the calorimeter and the energy of the reconstructed muons:

$$E_{x(y)}^{\text{miss}} = -E_{x(y)}^{\text{calo}} - E_{x(y)}^{\mu} \quad (4.10)$$

The energy deposits in the calorimeter,

$$E_{x(y)}^{\text{calo}} = \sum_{\text{obj}} E_{x(y)}^{\text{obj}} + E_{x(y)}^{\text{CellOut}}, \quad (4.11)$$

are composed of all the reconstructed and calibrated physical objects $obj = e, \gamma, \tau, \text{jet}, \mu$ and CellOut are calorimeter-cells not assigned to any physical object, also called *soft terms*. The muon energy deposits in the calorimeter are only taken into account, if they are not isolated ($\Delta R(\text{jet}, \mu) < 0.3$). The energy components of each object can be calculated from the energy deposit E_i in cell i and its angular position:

$$E_x^{\text{obj}} = \sum_i E_i \sin \theta_i \cos \phi_i \quad (4.12)$$

$$E_y^{\text{obj}} = \sum_i E_i \sin \theta_i \sin \phi_i \quad (4.13)$$

For jets the energy contributions are divided further into $p_{\text{T}}^{\text{jet}} > 20 \text{ GeV}$ and $7 \text{ GeV} < p_{\text{T}}^{\text{jet}} < 20 \text{ GeV}$ (*soft jets*), each with its own calibration.

In this chapter the event selection will be discussed. This selection aims to reduce the backgrounds, while keeping as much signal events as possible. The signal process is the Higgs boson production associated with a W boson, where the Higgs boson decays into two b-quarks and the W boson into an electron and an electron neutrino: $WH \rightarrow e\nu b\bar{b}$. This process is sketched in Fig. 5.1. Its signature in the detector consists of the electron, missing transverse energy, caused by the neutrino, and two b-jets, caused by the two b-quarks. The events where the W bosons decay to muons or taus are not considered within this thesis.

Important backgrounds arise from other processes containing W bosons. First, there is the irreducible background $W + b\bar{b}$, a W boson with two additional b-quarks. It has exactly the same final state signature as the signal, but different kinematics, since the two b-quarks stem from a massless gluon. Furthermore, there are W bosons with additional light jets ($W + \text{light}$) or c-jets ($W + c$ and $W + c\bar{c}$) and diboson processes (WW , WZ and ZZ). Other processes contain one or two top-quarks, which decay to W bosons and b-quarks. These are the single-top (Wt and single-top s- and t-channel) and top/anti-top pair ($t\bar{t}$) backgrounds.

The only important background containing no W boson is the multijet background. Here, the electron is caused by a jet that is falsely identified as electron and missing energy is caused by measurement uncertainties. This is not very likely to happen, but since the total cross section for multijet interactions is very high, this background has to be taken into account. Unfortunately this process is hard to model. In addition, since the selection efficiency is very low, huge samples had to be simulated. This is circumvented by estimating the multijet background from data, which will be discussed in Section 8.1.

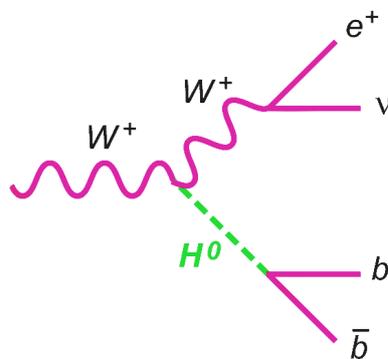


Figure 5.1: *The signal process: a Higgs boson production in association with a W boson. The W boson decays into an electron and an electron neutrino and the Higgs boson decays into two b-quarks.*

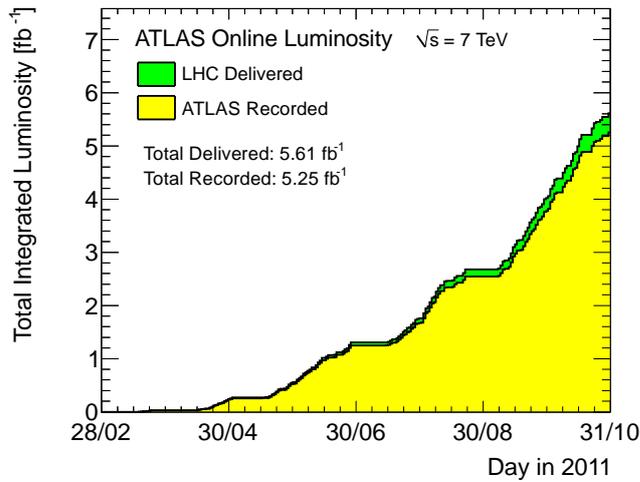


Figure 5.2: *Integrated luminosity of 2011: delivered by the LHC (green) and recorded by the ATLAS experiment (yellow) [23].*

data period	data trigger	MC period	luminosity
D	EF_e20_medium	1	3 %
E, F, G, H	EF_e20_medium	2	17 %
I, J	EF_e20_medium	3	13 %
K	EF_e22_medium	3	13 %
L, M	EF_e22vh_medium1	4	54 %

Table 5.1: *Data-taking periods of 2011 with the corresponding electron trigger and the Monte Carlo (MC) periods. The luminosity is given as fraction of the total integrated luminosity of 2011.*

5.1 Data samples

The analysis performed in this thesis is carried out using the data recorded by the ATLAS experiment in 2011. The data was taken at a centre-of-mass energy of $\sqrt{s} = 7$ TeV and correspond to an integrated luminosity of 5.61 fb^{-1} delivered by the LHC. A fraction of 93.6 % was recorded by ATLAS, corresponding to 5.25 fb^{-1} , as shown in Fig. 5.2.

The events in data that passed certain quality criteria for the detector conditions amount to 4.643 fb^{-1} of integrated luminosity and are used for the analysis. The data-taking divides into the periods D to M, which have different detector conditions and electron trigger configurations, which are shown in Table 5.1. The trigger thresholds were increased between the periods to cope with the increasing instantaneous luminosity.

5.2 Monte Carlo samples

All background processes except the multijet background are simulated by Monte Carlo. The simulation aims to create datasets that can be handled just like data from the detector. This allows for a technically easy comparison of the data to the prediction from the Standard Model. Such a simulation consists of several steps, which are discussed in the following.

process	generator	$\sigma \times BR$ [nb]	k -factor	$\varepsilon_{\text{filter}}$	N_{events}	$corr$
$WH \rightarrow l\nu b\bar{b}$						
$m_H = 110$ GeV	PYTHIA	0.0002109	1	1	$3.000 \cdot 10^4$	1
$m_H = 115$ GeV	PYTHIA	0.0001719	1	1	$3.299 \cdot 10^5$	1
$m_H = 120$ GeV	PYTHIA	0.0001377	1	1	$3.298 \cdot 10^5$	1
$m_H = 125$ GeV	PYTHIA	0.0001071	1	1	$3.298 \cdot 10^5$	1
$m_H = 130$ GeV	PYTHIA	0.00008011	1	1	$3.300 \cdot 10^5$	1
$m_H = 135$ GeV	PYTHIA	0.00005749	1	1	$2.990 \cdot 10^4$	1
$m_H = 140$ GeV	PYTHIA	0.00003934	1	1	$3.000 \cdot 10^4$	1
$t\bar{t}$	MC@NLO+ HERWIG	0.1455	1.146	0.5426	$1.498 \cdot 10^7$	0.773
Wt	ACERMC + PYTHIA	0.01534	1.026	1	$9.949 \cdot 10^5$	1
single- t s-chan.	ACERMC + PYTHIA	0.0003475	1.439	1	$1.999 \cdot 10^5$	1
single- t t-chan.	ACERMC + PYTHIA	0.007838	0.890	1	$9.993 \cdot 10^5$	0.844
$W + b\bar{b}$	ALPGEN + HERWIG	0.1081	1.195	1	$9.245 \cdot 10^5$	1
$W + c\bar{c}$	ALPGEN + HERWIG	0.3051	1.195	1	$3.020 \cdot 10^6$	1
$W + c$	ALPGEN + HERWIG	0.9142	1.195	1	$9.232 \cdot 10^6$	1
$(W \rightarrow e\nu) + \text{light}$	ALPGEN + HERWIG	8.751	1.195	1	$1.705 \cdot 10^7$	1
WW	HERWIG	0.03110	1.437	0.3895	$2.489 \cdot 10^6$	1
WZ	HERWIG	0.01148	1.600	0.3099	$9.999 \cdot 10^5$	1
ZZ	HERWIG	0.004572	1.300	0.2132	$2.500 \cdot 10^5$	1

Table 5.2: *The Monte Carlo samples used in this thesis and the Monte Carlo generators used for their simulation. The cross sections σ , branching ratios BR and k -factors are taken from the references listed in Table 5.3. The filter efficiencies and number of events are from [55]. The correction factor, $corr$, is calculated from the Monte Carlo event weight, as described in the text.*

First, the hard scattering process of the partons is simulated using various Monte Carlo generators: e.g. MC@NLO with *CT10* parton distribution functions (PDFs) [46, 47], ALPGEN with *CTEQ6L1 LO* PDFs [48, 49] and ACERMC with *MRSTMC* *al* PDFs [50]. The samples that are used in this analysis and the corresponding generators are listed in Table 5.2.

After the generation of the hard scattering process the showering and hadronisation has to be modeled. This is done by PYTHIA [51] or HERWIG [52]. While PYTHIA also generates the underlying event, for HERWIG this is implemented by a dedicated generator, named JIMMY [53].

The next step of the simulation is the propagation of the generated particles through the detector material and the simulation of the detector responses. This detector simulation is implemented using GEANT4 [54]. Finally, the read-out electronics of the detector are simulated in the *digitisation*, implemented in ATHENA [34].

Since the number of events in the simulated samples does not necessarily agree with the expected number of events in data, the simulated events have to be given *event weights*. These weights, w_{cs} , are calculated from the cross section σ , branching ratio BR and number of events N_{events} for a given process and the integrated luminosity:

$$w_{\text{cs}} = \frac{\sigma \times BR \times k \times \varepsilon_{\text{filter}}}{N_{\text{events}} \times corr} \times \int \mathcal{L} dt. \quad (5.1)$$

Additional effects are taken into account: the k -factor extrapolates the cross section, which was calculated at a specific order, to higher orders. The filter efficiency, $\varepsilon_{\text{filter}}$, accounts for

process	pert. order	σ unc.	source
WH	NLO	15 %	[25]
$t\bar{t}$	\sim NNLO	11 %	[56]
single-top	NNLO	8 %	[57–59]
$W+b\bar{b}$	\sim NNLO	5 %	[55, 60]
$W+c\bar{c}$	\sim NNLO	5 %	[55, 60]
$W+c$	\sim NNLO	5 %	[55, 60]
W +light	NNLO	5 %	[60]
diboson	NLO	7 %	[60]

Table 5.3: *The perturbative orders and the relative uncertainties of the cross sections for the various simulated processes. The perturbative orders denoted by \sim NNLO are approximately NNLO [56].*

lepton filters on the $t\bar{t}$ and diboson processes.

Finally, the correction factor, $corr$, takes into account negative events weights for the $t\bar{t}$ and single-top t-channel processes. It is calculated from the Monte Carlo event weights, w_i^{MC} , by summing over all events i :

$$corr = \frac{1}{N_{\text{events}}} \sum_{i=1}^{N_{\text{events}}} w_i^{\text{MC}}, \quad (5.2)$$

where the w_i^{MC} are given by the Monte Carlo generators. All numbers are listed in Table 5.2. Additionally, in Table 5.3 the perturbative orders and the relative uncertainties on the cross sections of the various processes are listed.

5.2.1 Pile-up correction

At the LHC bunches of up to $1.3 \cdot 10^{11}$ protons collide and the total cross section for a scattering process is very high with $\sigma_{\text{inel}} = 71.5 \text{ mb}$ [23]. Therefore multiple interactions occur in a single bunch crossing and cause additional energy deposits in the detector. This is called *in-time pile-up* or *minimum bias events*. Another form of pile-up is caused by the short time of 50 ns between the bunch crossings, which causes some of the detector responses to be assigned to the wrong bunch crossing. This is called *out-of-time pile-up*.

Both types of pile-up are modeled by Monte Carlo simulation by adding minimum bias events to the simulated processes. This is done for each MC period with different numbers of interaction per bunch crossing, since the luminosity of the LHC changed during this time.

The number of interactions for a single bunch crossing cannot be measured reliably. Instead, the distribution for multiple bunch crossings is estimated using a Poissonian distribution with mean μ . The mean is calculated using the instantaneous luminosity \mathcal{L} from Eq. 2.38, the inelastic cross section, σ_{inel} , the number of bunches, n_b , and the rotation frequency, f_r , of the beams in LHC:

$$\mu = \frac{\mathcal{L} \cdot \sigma_{\text{inel}}}{n_b \cdot f_r} \quad (5.3)$$

The μ is averaged over *luminosity blocks* [61] with run-times of about 1 min as $\langle \mu \rangle$. For 2011 data $\langle \mu \rangle$ ranges from zero to about 24 with an average of about nine interaction per bunch crossing [23].

The simulated pile-up from Monte Carlo does not agree perfectly with data. By applying scale factors to each Monte Carlo period the description of the $\langle\mu\rangle$ distribution observed in data is improved. An additional pile-up correction is applied by calculating weights, σ_{pu}^i , for each bin i in the $\langle\mu\rangle$ distribution:

$$\sigma_{\text{pu}}^i = \frac{n_{\text{data}}^i}{n_{\text{MC}}^i}, \quad (5.4)$$

where n_{data}^i and n_{MC}^i are the number of events of data and Monte Carlo, respectively. The distributions from data and Monte Carlo are normalised to the same area before this calculation. By applying the derived weights to the Monte Carlo events they are reweighted to describe the shape of the $\langle\mu\rangle$ distribution observed in data.

5.2.2 Vertex correction

The positions of the interactions in the detector in the z -direction follow a gaussian distribution. A width of $\sigma_z = 58$ mm and a mean of $\langle z \rangle = -6.3$ mm of this distribution is observed in data [62]. This distribution is modeled in the Monte Carlo simulation, but with a different mean of $\langle z \rangle = -6.1$ mm and widths of $\sigma_z = 75$ mm and $\sigma_z = 90$ mm, depending on the sample [63]. This effect is corrected for by calculating weights, w_z , to reweight the z -distribution from the simulation to that one measured in data. The weights are calculated in a similar way to those in the pile-up correction.

5.2.3 Trigger weighting

As shown in Table 5.1 the trigger for data-taking changed during Monte Carlo simulation period 3. Thus, for this MC period one cannot just choose one trigger, but one has to take two triggers into account. This is done by a *trigger weighting*, which takes into account the integrated luminosities, L , of the corresponding data periods (I, J and K). The weight, w_{trig} , is calculated for each event as

$$w_{\text{trig}} = \frac{L_{\text{I,J}}}{L_{\text{I,J,K}}} \text{EF_e20_medium} + \frac{L_{\text{K}}}{L_{\text{I,J,K}}} \text{EF_e22_medium}, \quad (5.5)$$

where the values of the triggers are one or zero, depending if the event was triggered or not. The periods I and J provide about 49% of the integrated luminosity and period K the remaining 51% of the MC period 3. Therefore, the resulting weights are one, zero or about 0.5.

5.2.4 Flavour tagging efficiencies

The values of the variables that are used for the flavour tagging differ in data and Monte Carlo simulation, which causes the tagging efficiencies to be different.

This is corrected for by calculating weights, w_{jet} , from data for each selected jet. These weights depend on the Monte Carlo truth flavour, f , the transverse momentum, p_{T} and the pseudo rapidity, η , of the jet. The weights of the two jets that are requested in the selection (see below) are multiplied to give a total flavour tagging weight:

$$w_{\text{tag}} = w_{\text{jet1}}(f, p_{\text{T}}, \eta) \cdot w_{\text{jet2}}(f, p_{\text{T}}, \eta), \quad (5.6)$$

where the f can take the values “light”, “c” or “b”. The calculation of the jet weights is done using the `BTaggingCalibrationDataInterface` [64].

5.2.5 Electron scale factors

As the flavour tagging, also the electron selection has different efficiencies for the Monte Carlo simulation than for data. This is corrected by various scale factors (SFs) that behave as weights and are calculated as

$$w_e^i = \frac{\varepsilon_{\text{data}}^i}{\varepsilon_{\text{MC}}^i}, \quad (5.7)$$

where $\varepsilon_{\text{MC}}^i$ and $\varepsilon_{\text{data}}^i$ are the selection efficiencies in Monte Carlo and data, respectively. The index i denotes various effects, that arise from the triggers (trig), reconstruction (reco), identification (ID) and isolation (iso) of the electrons.

The various effects are combined in a weighting factor for the selected electrons (see below) as

$$w_e = w_e^{\text{trig}} \cdot w_e^{\text{reco}} \cdot w_e^{\text{ID}} \cdot w_e^{\text{iso}}. \quad (5.8)$$

The weights of the various effects are calibrated using the well known processes $Z \rightarrow ee$ and $W \rightarrow e\nu$ [43].

5.2.6 Final event weights

All event weight discussed above are combined in a final event weight, w , as

$$w = w_{\text{CS}} \cdot w_{\text{pu}} \cdot w_{\text{trig}} \cdot w_{\text{Z}} \cdot w_{\text{tag}} \cdot w_e. \quad (5.9)$$

The application of these weights to the simulated events without any further estimation of the normalisations of the processes will be referred to as “normalised using Monte Carlo simulations”.

5.3 Event pre-selection

In the pre-selection each event is requested to contain at least one vertex with three or more outgoing tracks. Furthermore, events with a jet in the *lAr hole* are rejected. This is an inefficient region in the liquid argon (lAr) calorimeter. It is located in $-0.1 < \eta < 1.5$ and $-0.9 < \phi < -0.5$ [65]. This cut is applied only for the data periods E to H and the corresponding Monte Carlo period 3.

5.4 Electron selection

From the reconstructed electrons, described in Section 4.2, so-called *loose* electrons are selected first. They are classified into *central* and *forward* electrons. The central electrons are selected for $|\eta| < 2.47$ and $p_{\text{T}} > 10 \text{ GeV}$ and the forward electrons for $|\eta| < 4.5$ and $p_{\text{T}} > 20 \text{ GeV}$. The identification criterion in both regions is *loose++*.

The central electrons are requested to be isolated in the calorimeter by the cut $E_{\text{T}}^{\text{cone}}(R_0 = 0.3)/E_{\text{T}}^{\text{e}} < 0.14$ and, regarding their tracks, by $p_{\text{T}}^{\text{cone}}(R_0 = 0.2)/p_{\text{T}}^{\text{e}} < 0.1$. Additionally, the impact parameter of the electron track with respect to the signal vertex, d_0 , is required to be smaller than 0.1 mm.

On top of the requirements of the loose electrons the *signal* electrons have tighter cuts. They are requested to pass the *tight++* identification criterion and are used only in the central region for $|\eta| < 2.47$ and $p_{\text{T}} > 25 \text{ GeV}$. The loose electrons that are not selected as signal electrons are used as *veto* electrons.

5.5 Muon selection

Muons are selected only for vetoing and have weak quality criteria. The muons reconstructed by STACO with *loose* requirements, as discussed in Section 4.3, are selected for $|\eta| < 2.7$ and $p_T > 10$ GeV. Cuts are applied on the impact parameters with respect to the signal vertex: $|d_0| < 0.1$ mm and $Z_0 < 10$ mm. The isolation criteria are the same as for central electrons: $E_T^{\text{cone}}(R_0 = 0.3)/E_T^\mu < 0.14$ and $p_T^{\text{cone}}(R_0 = 0.2)/p_T^\mu < 0.1$.

5.6 Jet selection

Jets are selected from the *AntiKt4TopoEM* collection after the correction of their four-vectors, as described in Section 4.4. A cut on the *jet vertex fraction* (JVF) of all jets is performed to reject jets stemming from pile-up:

$$\text{JVF}(\text{jet}, \text{SV}) = \frac{\sum_i p_T^i(\text{SV})}{\sum_j p_T^j} > 0.75, \quad (5.10)$$

where index j runs over all tracks in the jet and i only over tracks that are compatible with the signal vertex (SV). The JVF has small values for a jet, whose tracks are not stemming from the signal vertex, but from pile-up. The cuts $|\eta| < 4.5$ and $p_T > 20$ GeV are applied for *loose* jets and $|\eta| < 2.5$ and $p_T > 25$ GeV for *signal* jets. The loose jets that are not selected as signal jets are used as *veto* jets.

5.7 Overlap removal

The identified physical objects are subject to overlap, meaning the same object in the calorimeter can be identified as several physical objects, i.e. as an electron and as a jet at the same time. To remove this overlap the selected loose jets and electrons are deleted from the list of selected objects if they pass matching conditions. First, jets are deleted if they are within $dR(\text{jet}, e) < 0.4$ to an electron with $p_T^e \geq 20$ GeV. Next, electrons are removed if they are within $dR(\text{jet}, e) < 0.4$ to a jet. Obviously, this affects only electrons with $p_T^e < 20$ GeV.

5.8 Missing transverse energy

The missing transverse energy, E_T^{miss} , is calculated from all reconstructed and corrected jets and electrons, as described in Section 4.8. If the jet collection contains a *bad jet* the event is rejected. Bad jets rise from energy deposits in the calorimeter that do not stem from an interaction in the beam spot. Various sources of bad jets exist, e.g. hardware problems, LHC beam conditions or cosmic rays [65]. No cut on the E_T^{miss} is applied, but it is used in the training of the Boosted Decision Trees, as described in Section 7.

5.9 Signal selection

After the object selection a basic event-based signal selection is applied. Each event is requested to have exactly one signal electron and no veto electron nor muon. Furthermore, two b-tagged signal jets are requested and no additional jet is allowed. The b-tagging is done with MV1 at the 70% working point, as discussed in Section 4.5.

sample	from MC		from fits	
$W+bb$	480.6	± 18.7	759.3	± 29.5
$W+c\bar{c}$	132.1	± 10.0	132.1	± 10.0
$W+c$	133.1	± 9.8	133.1	± 9.8
W +light	105.6	± 23.5	105.6	± 23.5
diboson	52.7	± 1.3	52.7	± 1.3
single-top	280.2	± 3.1	280.2	± 3.1
$t\bar{t}$	976.5	± 7.4	1103.5	± 8.4
multijet	—		1321.8	± 40.0
WH	10.8	± 0.1	10.8	± 0.1
total bkg.	2160.8	± 34.1	3888.3	± 57.4
data	3934.0	± 62.7	3934.0	± 62.7

Table 5.4: Yields of the backgrounds, signal and data after the event selection. The WH signal is shown for $m_H = 125$ GeV. The normalisations in the middle column are taken from Monte Carlo simulation using the cross section from theory. In the right column the normalisations are taken from the fit discussed in Section 8.2. The multijet background is estimated in Section 8.1. The statistical uncertainties are calculated from the number of events.

To this point the event selection follows exactly the cut-based analysis in Ref. [1], which is used as reference analysis. A good agreement to the reference analysis is reached regarding the number of events for various stages of selection (*cutflow agreement*) to the level of a few per mil. In contrast to the reference analysis, no additional cuts are applied in the analysis presented here. Instead, Boosted Decision Trees (BDTs) are used to further separate a possible signal from the backgrounds.

5.10 Event yields

The event yields obtained after the selection are shown in Table 5.4. One can see that the largest background is the multijet background, followed by $t\bar{t}$ and Wbb . However, the multijet background can be easily suppressed by applying cuts on the missing transverse energy or the transverse mass of the reconstructed W boson. This leaves $t\bar{t}$ as most important reducible background.

Various variables, listed in Table 5.5, are calculated from the reconstructed and selected objects. These objects are the electron, the missing transverse energy and the two jets. Additionally, the four-vector of the W boson is defined as the vectorial sum of the missing transverse energy and the electron four-momentum in the transverse plane. The four-vector of the Higgs boson is defined as the vectorial sum of the two jet four-momenta.

The variable distributions after the event selection are shown in Fig. 5.3 and 5.4. Each plot consists of two histograms: The upper histogram shows the measured number of events in data (black points with error bars) and the expected number for all backgrounds stacked on each other. The expected signal for a Standard Model Higgs boson with $m_H = 125$ GeV is shown as black line, scaled up by a factor of 20. The lower histogram shows the ratio data/MC and the statistical errors from MC as orange band.

From the distributions one can guess a slight missmodeling of the multijet background, which is estimated from data in Section 8.1. This is apparent in the distributions of p_T^e and

p_T^e	transverse momentum of the electron
$p_T^{\text{jet1,2}}$	transverse momentum of the leading jet (jet1) and the sub-leading jet (jet2)
E_T^{miss}	missing transverse energy
p_T^W	transverse momentum of the reconstructed W boson
m_T^W	transverse mass of the reconstructed W boson
m_H	invariant mass of the reconstructed Higgs boson
$\Delta R(\text{jet1, jet2})$	distance in R between the two jets
$\Delta\phi(e, E_T^{\text{miss}})$	angle between the missing transverse energy and the electron in the transverse plane
$\Delta\phi_{\text{min}}(e, \text{jet})$	angle between the electron and the closer of the two jets in the transverse plane
H_T	scalar sum of the transverse momenta: $H_T = E_T^{\text{miss}} + p_T^e + p_T^{\text{jet1}} + p_T^{\text{jet2}}$
p_T^{imb}	imbalance of the scalar transverse momenta: $p_T^{\text{imb}} = p_T^H - p_T^W / (p_T^H + p_T^W)$
MV1(jet1,2)	neural network output of the flavour tagging algorithm MV1 for the two jets
JFcombNNc(jet1,2)	neural network output of the flavour tagging algorithm JetFitterCombNNc for the two jets

Table 5.5: Kinematic and flavour tagging variables, which are calculated from the reconstructed and selected physical objects.

p_T^W , where the estimated multijet background has a shift to higher values compared to data.

The variable E_T^{miss} and m_T^W yield good separation between the signal and the multijet background. This is not exploited in the event selection to increase the number of events for the training of Boosted Decision Trees (BDTs), as discussed in Section 7. The BDTs are expected to use the separation in these variables in an optimal way.

The b-tagging variables MV1 and JFcombNNc show separation between the signal and backgrounds that contain only light or c-jets. All listed variables, except p_T^e , p_T^{jet1} and p_T^{jet2} , are used in the BDT training in Section 7.1 and will be motivated further in Section 6.

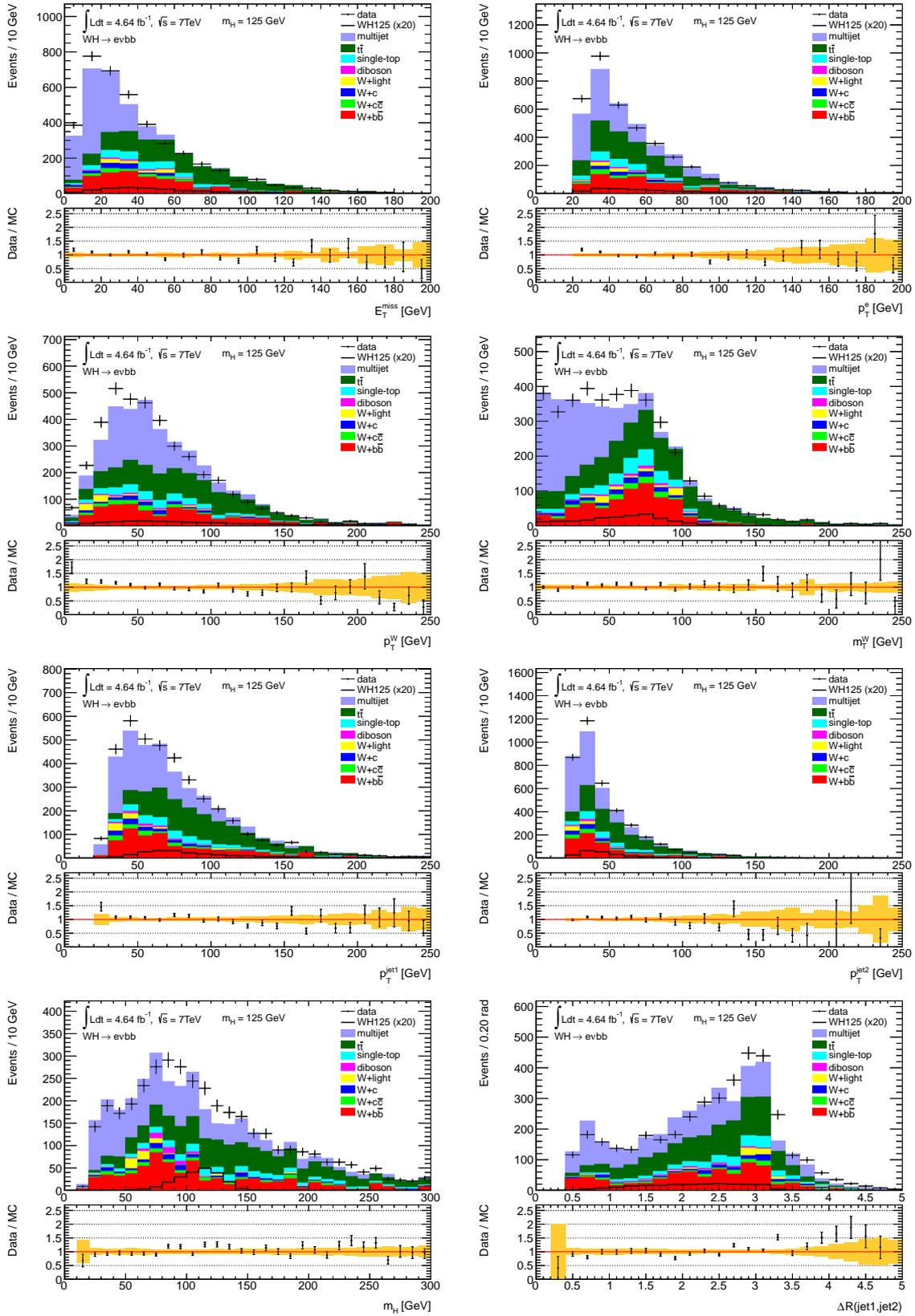


Figure 5.3: Distributions of E_T^{miss} , p_T^e , p_T^W , m_T^W , p_T^{jet1} , p_T^{jet2} , m_H and $\Delta R(\text{jet1}, \text{jet2})$ after the event selection. The definitions of the variables are given in Table 5.5. The backgrounds are normalised using the fit described in Section 8.2. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

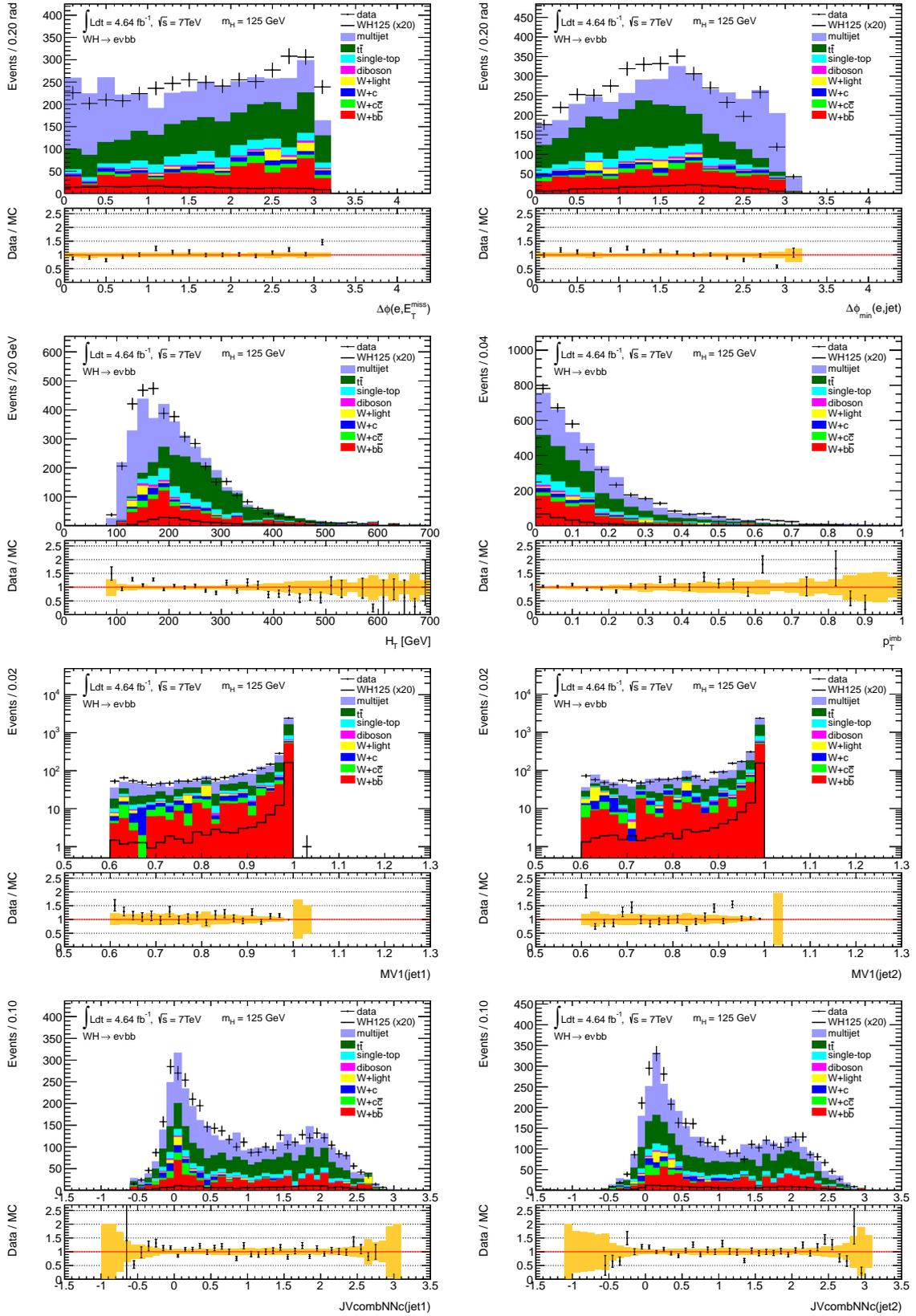


Figure 5.4: Distributions of $\Delta\phi(e, E_T^{\text{miss}})$, $\Delta\phi_{\text{min}}(e, \text{jet})$, H_T , p_T^{imb} , $MV1(\text{jet}1)$, $MV1(\text{jet}2)$, $JF\text{combNNc}(\text{jet}1)$ and $JF\text{combNNc}(\text{jet}2)$ after the event selection. The definitions of the variables are given in Table 5.5. The backgrounds are normalised using the fit described in Section 8.2. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

The $t\bar{t}$ background is one of the most important backgrounds for the $WH \rightarrow e\nu b\bar{b}$ signal. It has a high cross section and similarities in the final state topology. As for the signal, these can contain one electron, missing transverse energy and two b-quarks. This makes $t\bar{t}$ the second largest background after the event selection, as shown in Table 5.4. But in contrast to the signal it has additional objects, thus making this background reducible. The only larger background is the multijet background, which can be easily suppressed by introducing an E_T^{miss} or m_T^W cut.

In almost 100 % of all $t\bar{t}$ decays both top quarks decay into a W boson and a b-quark [66]. The W bosons decay leptonically into a lepton and a corresponding neutrino or hadronically in a light quark and a light or c-quark. Also hadronic decays with b-quarks are possible, but are suppressed by the CKM mechanism [10]. Since in the event selection an electron with high quality criteria is requested, at least one of the W bosons needs to decay into an electron and a neutrino. Otherwise the event is very likely to be rejected.

In the following the $t\bar{t}$ events are classified into three categories by the decay products of two W bosons:

- The *leptonic* decay, $t\bar{t} \rightarrow WbW\bar{b} \rightarrow l\nu l\nu b\bar{b}$
- The *semi-leptonic* decay, $t\bar{t} \rightarrow WbW\bar{b} \rightarrow qql\nu b\bar{b}$
- The *hadronic* decay, $t\bar{t} \rightarrow WbW\bar{b} \rightarrow qqqb\bar{b}$,

The first two processes are shown in Fig. 6.1. The hadronic $t\bar{t}$ is neglected in the following, since it has no lepton and is therefore very unlikely to pass the event selection. In comparison to the signal the leptonic $t\bar{t}$ has an additional lepton and higher E_T^{miss} due to the second neutrino in the final state. The semi-leptonic $t\bar{t}$ on the other hand has two additional quarks, causing jets. One of these quarks is a light quark and the other is a light or c-quark.

The contributions of the various final state topologies of the $t\bar{t}$ background are further investigated in the next Section and their differences to the decay topology of the signal motivate several discriminating variables, as discussed in Section 6.2.

6.1 Truth composition

A study of the Monte Carlo truth is performed to have a closer look at the $t\bar{t}$ background. The truth contains the simulated events at parton level, without any detector simulation. Therefore it does not suffer from reconstruction efficiencies and all simulated particles are preserved.

This study uses the events after the selection, presented in Section 5. This includes b-tagging for two selected jets, which is important when interpreting the results presented here. Since signal-like events are of interest, an additional cut on the reconstructed Higgs boson mass is applied, $80 \text{ GeV} < m_H < 150 \text{ GeV}$. The goal of this study is to see why the $t\bar{t}$ background passes the event selection, despite of having additional decay products. These additional objects should be selected as veto objects and reject the event.

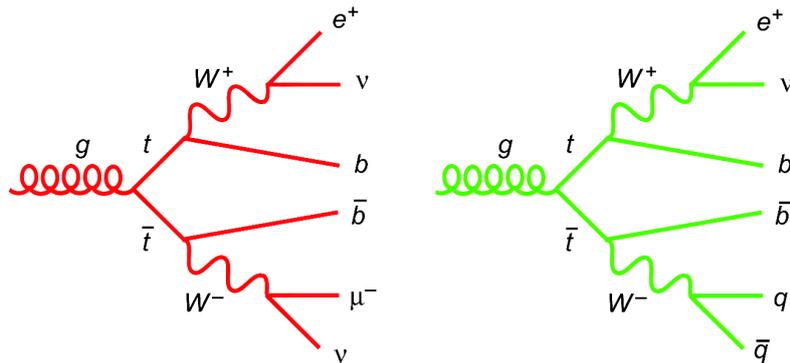


Figure 6.1: The $t\bar{t}$ background processes: leptonic decay (left) and semi-leptonic decay (right). In both cases the top-quarks decay each into a b -quark and a W boson with a branching ratio close to 100% [66]. One of the W bosons decays leptonically into an electron and a neutrino, the other W boson decays leptonically (left) or hadronically (right).

This study unveils what kind of additional objects are missed in the selection and suggests ways to recover them. From the leptonically decaying $t\bar{t}$ an additional lepton and from the semi-leptonic $t\bar{t}$ two additional jets are expected.

The Monte Carlo truth is searched for the $t\bar{t}$ decay products and based on this information the decays are classified as leptonic or semi-leptonic. Further, a separation into categories of the various leptons and quarks from the W boson decays is performed. The combinations of one W boson decaying into an electron and the other into an e , μ , τ_{had} or c -quark¹ in the final state are taken into account. Leptonically decaying taus are included in the e and μ categories. The remaining combinations are summed up in *other* categories. The total number of events of the $t\bar{t}$ background are listed in Table 6.1(a) and the fractions of the various categories described above are listed in Table 6.1(b) and 6.1(c).

The b -quarks from the top-quark decays and the c -quarks from the W boson decays are matched to the two b -tagged signal jets. The matching condition is $\Delta R(\text{jet}, \text{quark}) < 0.4$ after final state radiation for the quarks. The results are shown in Table 6.1(d) in bins of the transverse momentum of the W boson.

The leptonically decaying $t\bar{t}$ background has a large contribution to the overall $t\bar{t}$ background of about 66%. This is rather unexpected, since one might think that a lepton from a W -decay is easy to veto. Instead, the second lepton often fails the loose lepton selection, especially for the lower p_{T}^W region.

Missed electrons cause a contribution of about 18%. The p_{T} and η spectra of these electrons are shown in Fig. 6.2(a). One can see that the bulk of these electrons is lost in the low- p_{T} region and is hard to recover. Depending on the reason for losing these electrons one might think about looser selection criteria or using tracks from the inner detector for their reconstruction. However, these options are not considered in this thesis, but might be subject to a follow-up study. Outside the selected η range of $|\eta| < 4.5$ only a small number of electrons is missed.

The contribution from missed muons is even higher with about 21%. Their p_{T} and η spectra are shown in Fig. 6.2(b). Here the p_{T} distribution shows a peak just below 10 GeV,

¹ C -quarks are of special interest, since they produce jets in the detector that can be misidentified as b -jets with higher probability than jets caused by light quarks. Therefore, the c -quarks from the W boson decays can fake the b -quarks of the signal process, $H \rightarrow b\bar{b}$.

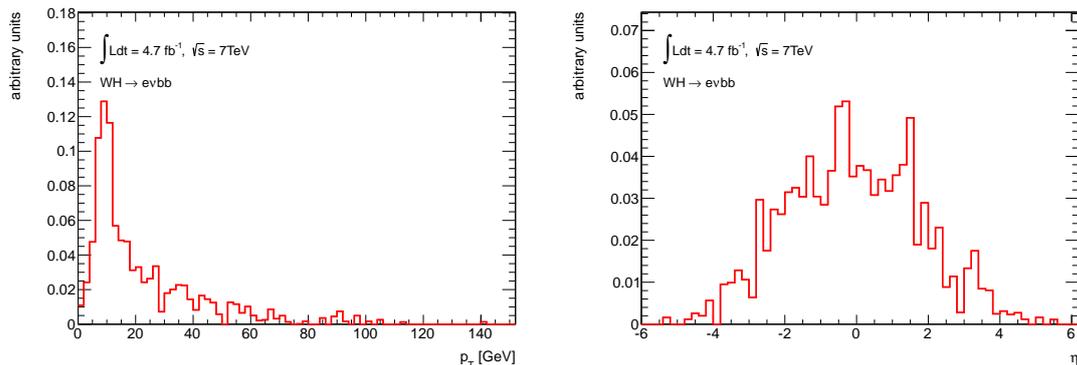
(a) $t\bar{t}$ background, number of events						
p_T^W [GeV]	0 – 50	50 – 100	100–150	150–200	> 200	all
$N_{\text{events}}^{\text{raw}}$	3771	6322	3238	561	168	14060
$N_{\text{events}}^{\text{weighted}}$	107.2	185.3	108.0	19.8	6.2	426.6

(b) leptonic contributions						
p_T^W [GeV]	0 – 50	50 – 100	100–150	150–200	> 200	all
e, e	19.0%	18.2%	19.3%	14.5%	9.1%	18.4%
e, μ	22.1%	21.1%	20.6%	10.6%	7.2%	20.6%
e, τ_{had}	27.4%	24.9%	23.4%	14.9%	5.5%	24.4%
other	3.5%	3.0%	3.0%	1.5%	2.3%	3.0%
sum	72.0%	67.2%	66.3%	41.5%	24.1%	66.4%

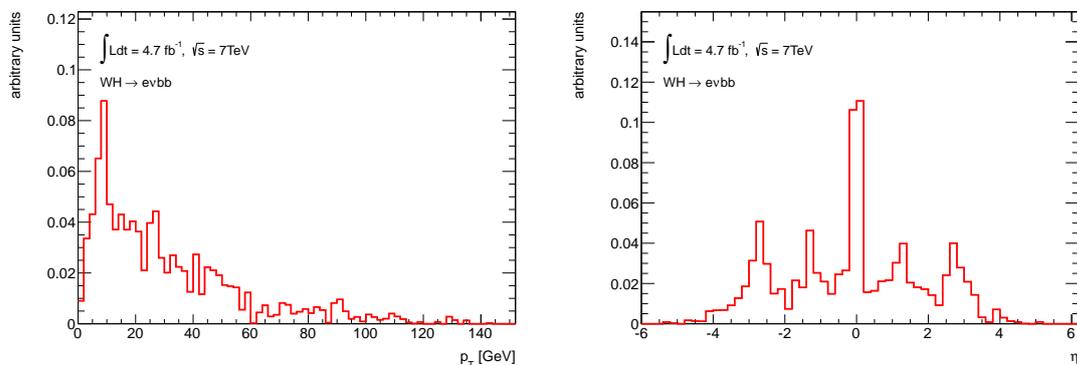
(c) semi-leptonic contributions						
p_T^W [GeV]	0 – 50	50 – 100	100–150	150–200	> 200	all
e, c	17.6%	21.0%	24.0%	46.6%	67.0%	22.7%
other	10.4%	11.8%	9.7%	11.9%	8.9%	10.9%
sum	28.0%	32.8%	33.7%	58.5%	75.9%	33.6%

(d) quark-jet-matching						
p_T^W [GeV]	0 – 50	50 – 100	100–150	150–200	> 200	all
b, b matched	88.2%	86.8%	84.4%	64.4%	37.5%	84.8%
b, c matched	8.1%	9.6%	12.2%	31.5%	53.2%	11.5%
b, bc matched	0.6%	1.2%	2.9%	2.5%	2.5%	1.5%

Table 6.1: Truth composition of the $t\bar{t}$ background after the event selection including b -tagging for the two jets and the additional m_H cut. Shown are the contributions of the various $t\bar{t}$ decay topologies in bins of p_T^W . In (a) the raw number of events, $N_{\text{events}}^{\text{raw}}$, are shown to give an estimate of the statistical uncertainty for each bin. For $p_T^W > 200$ GeV the total uncertainties on the fractions are of the order of a few percent. For the lower p_T^W bins they are about 1% or less. All percentages below are relative to the number of weighted events, $N_{\text{events}}^{\text{weighted}}$, in each p_T^W -bin. The weighting is done using the results of the background estimation, presented in Section 8.2. The leptonic decays are listed in (b) and the semi-leptonic decays in (c). In (d) the quark-matching results are listed. In the events tagged “b, b matched” both b -quarks are successfully matched to the selected jets, in “b, c matched” only one b -quark is matched to a jet and a c -quark to the other. In “b, bc matched” both b -quarks are matched to the two jets and a c -quark is matched to one of the same jets. This can occur if a b -quark and the c -quark are close in ΔR and their jets are merged.



(a) spectra of the electrons failing the selection



(b) spectra of the muons failing the selection

Figure 6.2: p_T and η distributions of the leptons that fail the loose electron selection and therefore cannot be vetoed. The distributions are shown for leptonic $t\bar{t}$ events that passed the event selection and the additional m_H cut. These contain two W bosons that decay each into a lepton and a neutrino. One of the leptons is selected as signal lepton.

the minimal p_T in the loose muon selection. This peak can be reduced by lowering the p_T threshold for muons, since the reconstruction efficiency for muons is high down to a few GeV [32]. A significant contribution of muons is lost in the very central region of the detector around $\eta = 0$. This is due to inefficiencies in the muon spectrometer. Also in the transition regions around $|\eta| = 1.5$ muons are lost, as well as in the forward regions outside the selected range of $|\eta| < 2.7$. This might be improved by using segment- or calorimeter-tagged muons. Again, this might be reconsidered in a follow-up study.

The largest contribution of the leptonic $t\bar{t}$ comes from hadronically decaying taus with about 24%. These are even harder to recover than electrons or muons and are not considered here. One might think about reconstructing them by applying a tau identification with loose selection criteria.

The semi-leptonically decaying $t\bar{t}$ has a contribution to the total $t\bar{t}$ background of about 34%, but for $p_T^W > 200$ GeV this fraction rises to about 76%. This is partly due to fact, that the efficiency for vetoing the second lepton from the leptonic $t\bar{t}$ rises with p_T^W . Therefore, the leptonic $t\bar{t}$ fractions drops and the semi-leptonic fraction rises. An additional effect is, that c-jets with high p_T are hard to distinguish from b-jets: The dominant part of the semi-leptonic

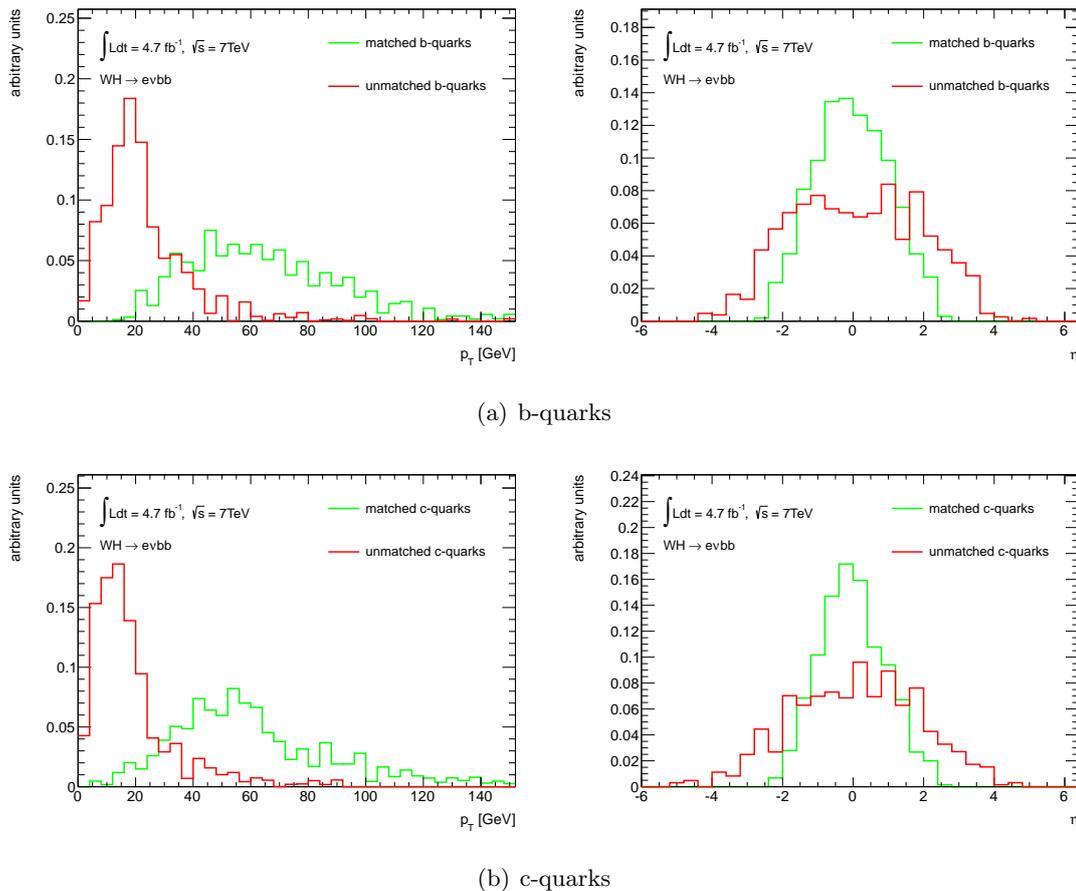


Figure 6.3: p_T and η distributions of the matched (green) and unmatched (red) truth quarks from the $t\bar{t}$ decays after the event selection including b-tagging for the two jets and the additional m_H cut. In (a) the b-quarks of the events where exactly one b-quark is matched to a jet are shown. In (b) the c-quarks from the semi-leptonic $t\bar{t}$ events where one W boson decays into a c-quark in the final state are shown. The unmatched quarks are not reconstructed as jets and therefore cannot be vetoed.

$t\bar{t}$ for $p_T^W > 200$ GeV (about 67% / 76% = 88%) comes from events, where one W boson decays leptonically into an electron and the other hadronically into a c-quark in the final state. This high contribution can be explained by c-quarks that fake b-jets, since in about 53% of the events for $p_T^W > 200$ GeV only one of the two b-tagged jets is actually caused by a b-quark, the other is caused by a c-quark. In these cases the second b-quark is lost.

The spectra of the matched and unmatched b- and c-quarks are shown in Fig. 6.3. One can see that the unmatched quarks have low- p_T and are therefore hard to recover. However, the b-tagged jets that are caused by c-quarks can be used for discrimination by using a tagging algorithm that is sensitive to c-quarks, as shown in the next Section.

6.2 Discriminating variables

In the following variables that yield separation between the signal and the $t\bar{t}$ and other backgrounds are discussed. These variables are used as input variables for Boosted Decision Trees

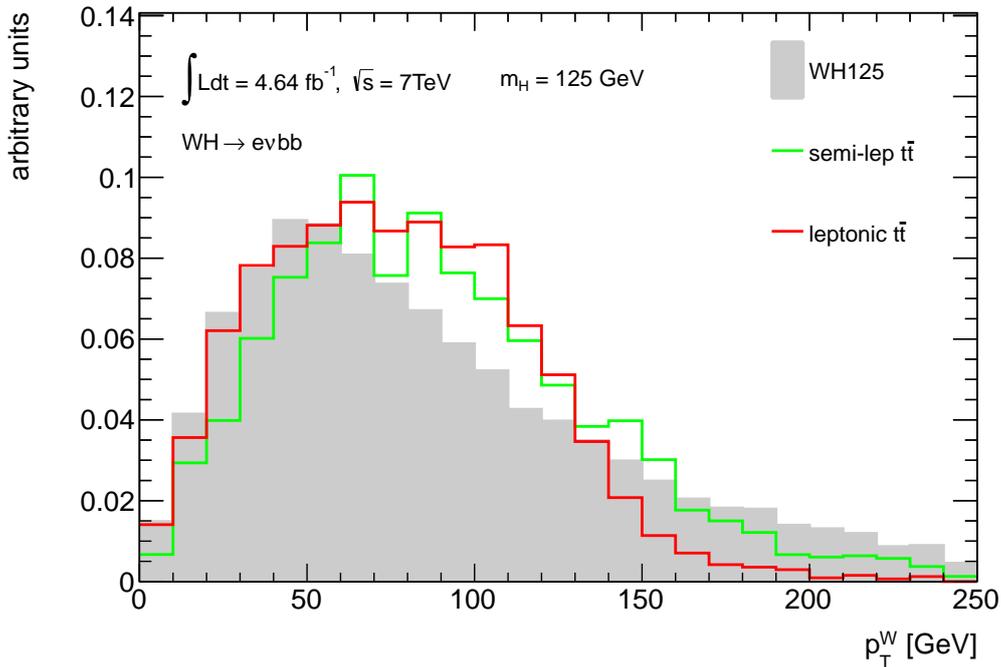


Figure 6.4: Distributions of p_T^W normalised to the same area after event selection and additional m_H -cut for the WH signal, the leptonic and the semi-leptonic $t\bar{t}$ background.

(BDTs). BDTs make use of the separation in the individual variables and their correlations to extract the signal. All input variables for the BDTs, defined in Section 5.10, are discussed in the following. The definition of the BDTs is given in Section 7.

One important variable to separate the signal from the $t\bar{t}$ background is p_T^W , shown in Fig. 6.4. The variable itself has only little discrimination, but its correlations to other variables have a strong effect on the $t\bar{t}$ background.

In the following variable distributions are shown after the event selection and the additional m_H -cut of $80 \text{ GeV} < m_H < 150 \text{ GeV}$. The distributions for the leptonic and the semi-leptonic $t\bar{t}$ are shown separately and split into $p_T^W < 50 \text{ GeV}$ and $p_T^W > 100 \text{ GeV}$ to make the effect of p_T^W visible.

In Fig. 6.5 basic kinematic variables are shown:

E_T^{miss} The semi-leptonic $t\bar{t}$ has a similar E_T^{miss} distribution as the signal, since both processes have one neutrino stemming from a W boson decay. The leptonic $t\bar{t}$ instead has significantly higher E_T^{miss} values for low- p_T^W due to its additional neutrino. This difference gets smaller for high- p_T^W .

m_T^W The m_T^W distribution is again very similar for the signal and the semi-leptonic $t\bar{t}$, but is shifted to higher values for the leptonic $t\bar{t}$ in both p_T^W regions.

m_H The best discriminating variable is m_H for both $t\bar{t}$ and all other backgrounds, see also Fig. 5.3.

$\Delta R(\text{jet1}, \text{jet2})$ The distance ΔR between the two jets, $\Delta R(\text{jet1}, \text{jet2})$, is similar for the signal

and the leptonic $t\bar{t}$, but slightly smaller for the semi-leptonic $t\bar{t}$ for low- p_T^W . It has larger values for both backgrounds than for the signal for high- p_T^W .

Further kinematic variables are shown in Fig. 6.6:

- $\Delta\phi(\mathbf{e}, \mathbf{E}_T^{\text{miss}})$** The signal process has one neutrino and one electron from the same W boson. For high p_T^W the angle between them, $\Delta\phi(\mathbf{e}, \mathbf{E}_T^{\text{miss}})$, is small. The leptonic $t\bar{t}$ instead has two neutrinos, one on the same side of the electron in the transverse plane and one on the other, resulting in larger $\Delta\phi(\mathbf{e}, \mathbf{E}_T^{\text{miss}})$.
- $\Delta\phi_{\text{min}}(\mathbf{e}, \mathbf{jet})$** In the signal process for high p_T^W the two b-jets are close to each other in R and the electron has a large angle to both of them in the transverse plane. Instead for the $t\bar{t}$ background the b-jets are spread more widely and one of the jets stems from the same top quark as the electron. Therefore the distance $\Delta\phi_{\text{min}}(\mathbf{e}, \mathbf{jet})$ is smaller for the $t\bar{t}$ background.
- H_T** Since the two top-quarks have a significantly higher mass than the W boson and the Higgs boson in the signal process, the total energy in the event, H_T , is higher for low- p_T^W . For high- p_T^W this relation is reversed. no obvious explanation for this behaviour was found.
- p_T^{imb}** In an ideal signal event with no additional hard radiation p_T^{imb} is zero. This is not true for $t\bar{t}$ events and this shows up with higher values in the p_T^{imb} distribution, especially for the semi-leptonic $t\bar{t}$.

In Fig. 6.7 the used b-tagging variables are shown separately for the two jets. The jet with the higher p_T is named *jet1* and the other *jet2*. Since the $t\bar{t}$ background has two real b-jets like the signal, the b-tagging variables are not expected to yield strong separation between these two samples. Nevertheless, as was shown in Section 6.1 the two jets of the $t\bar{t}$ background that pass the event selection not always are true b-jets, but there is also a contribution of b-jets faked by c-quarks. This fact might yield some potential for separation. The b-tagging variables are:

- MV1** The MV1 output has only a weak discrimination between the signal and the $t\bar{t}$ background containing two real b-jets. Its separation against the occurring c-quarks is not very strong, see Fig. 4.1. But it has good separation to backgrounds containing only light or c-jets, see Fig. 5.4.
- JFcombNNc** Opposed to MV1 the JFcombNNc output has some discrimination between the signal and the semi-leptonic $t\bar{t}$ background for high- p_T^W . This effect is due its higher c-rejection of this algorithm and the large contribution of b-tagged jets faked by c-quarks in the $t\bar{t}$ background, as discussed before.

In summary this set of variables provides separation between signal and the $t\bar{t}$ background. This separation is especially apparent when taking the correlations between p_T^W and the other variables into account. This effect can best be exploited by a multivariate approach like BDTs.

The variables shown provide also discrimination against other backgrounds, as described in Section 5.10: E_T^{miss} and m_T^W are the most important ones to reject the multijet background and the b-tagging variables separate the backgrounds containing only light or c-jets.

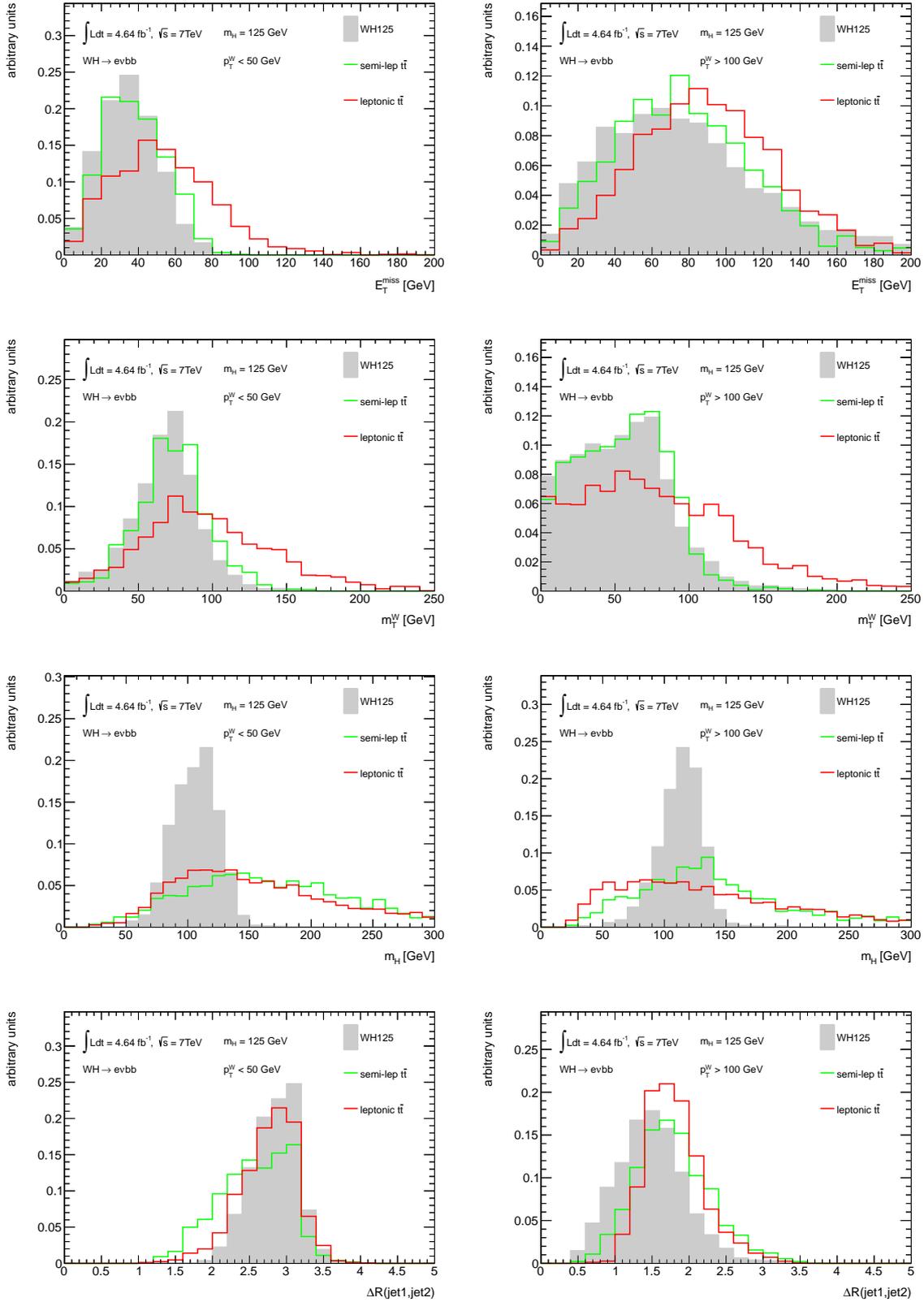


Figure 6.5: Distributions of E_T^{miss} , m_T^W , m_H and $\Delta R(\text{jet1}, \text{jet2})$ normalised to the same area after the event selection. The additional m_H -cut is applied except for the m_H distribution. Plots for $p_T^W < 50 \text{ GeV}$ are shown on the left and for $p_T^W > 100 \text{ GeV}$ on the right.

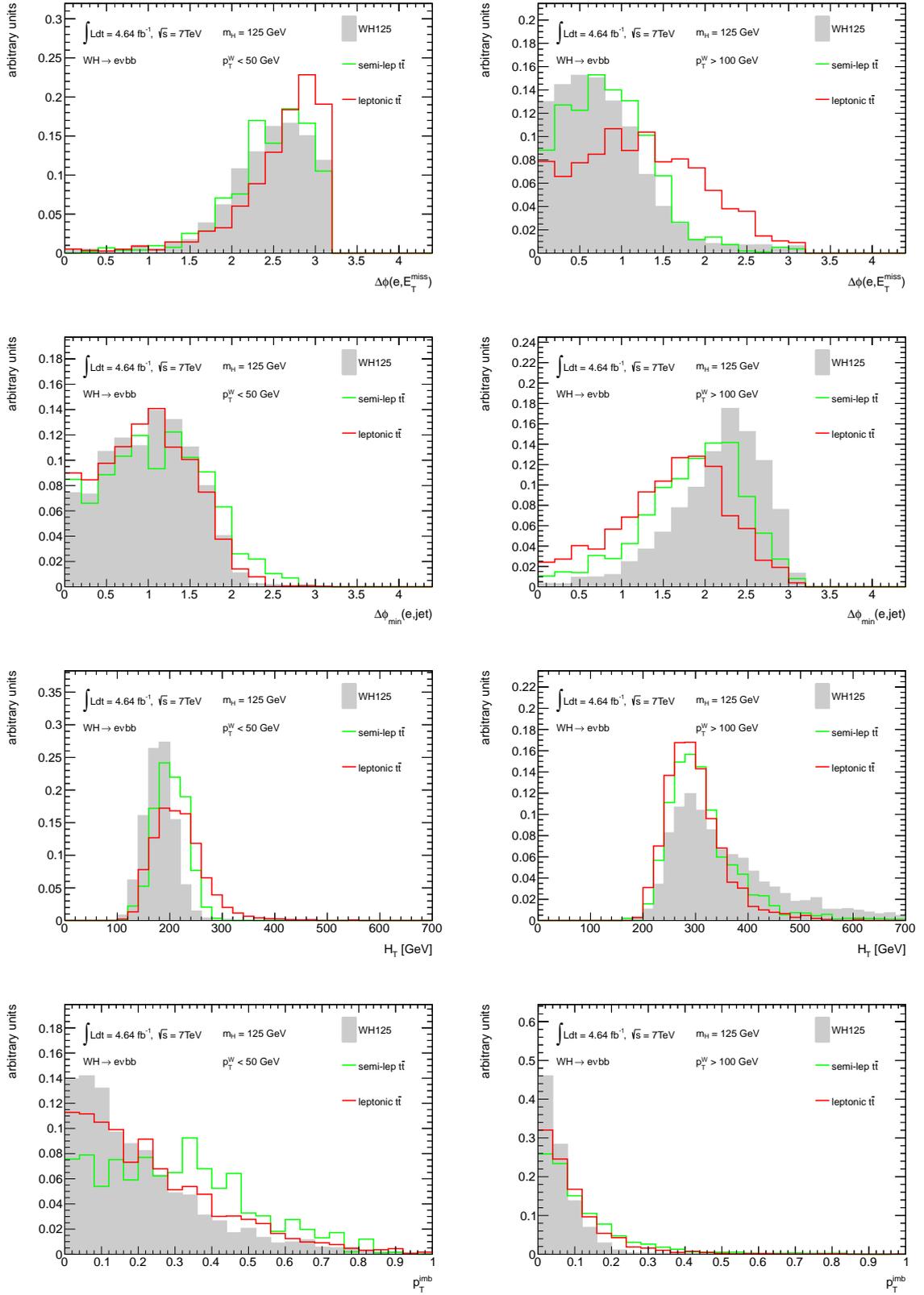


Figure 6.6: Distributions of $\Delta\phi(e, E_T^{\text{miss}})$, $\Delta\phi_{\text{min}}(e, \text{jet})$, H_T and p_T^{imb} normalised to the same area after the event selection and the additional m_H -cut. The distributions for $p_T^W < 50$ GeV are shown on the left and for $p_T^W > 100$ GeV on the right.

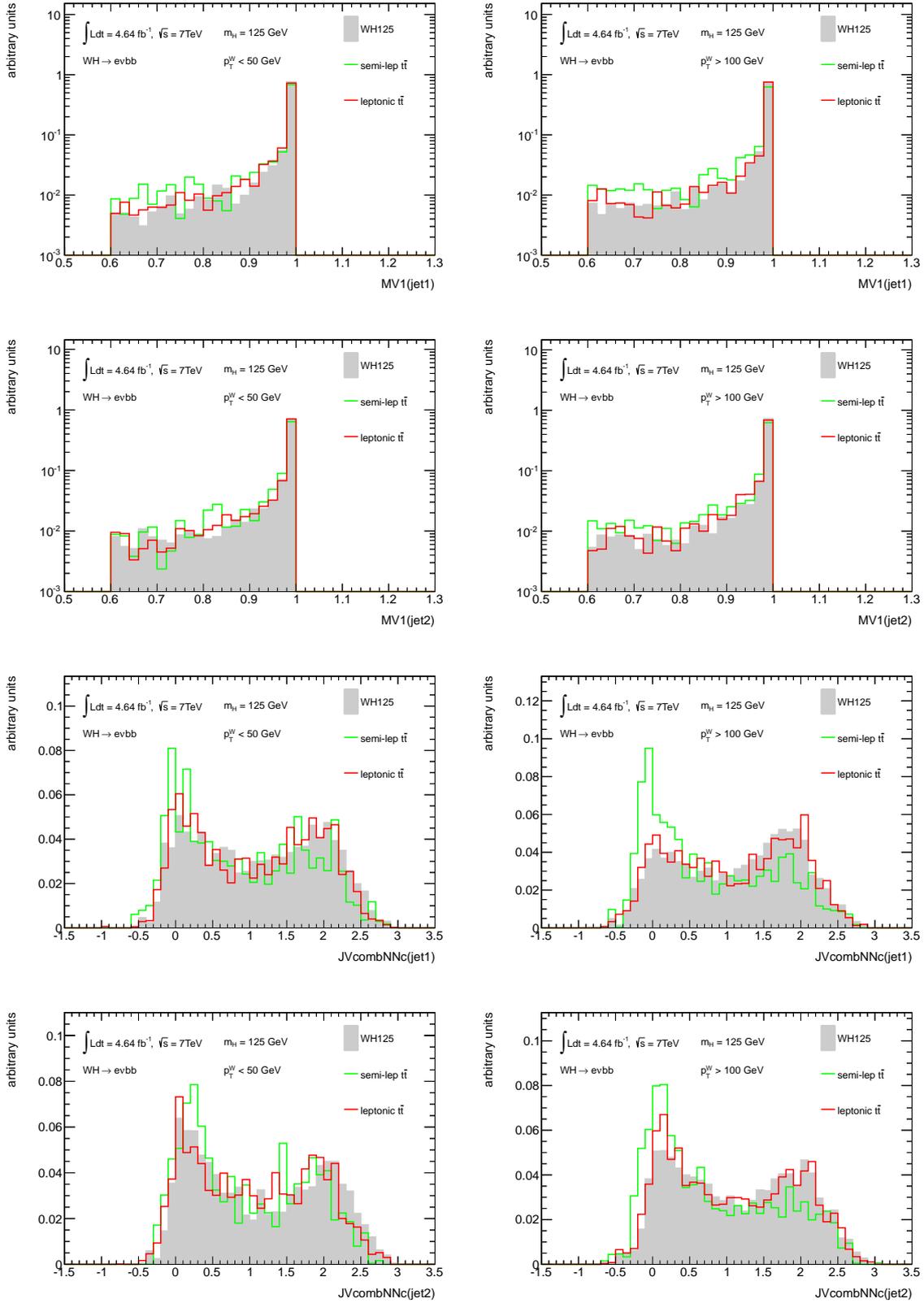


Figure 6.7: *Distributions of MV1(jet1), MV1(jet2), JFcombNNc(jet1) and JFcombNNc(jet2) normalised to the same area after the event selection and the additional m_H -cut. The distributions for $p_T^W < 50$ GeV are shown on the left and for $p_T^W > 100$ GeV on the right. The MV1 distribution is plotted on a logarithmic scale.*

After the basic event selection, presented in Section 5, further discrimination between the signal and the backgrounds is needed to increase the sensitivity. This can be done using additional one- or two-sided cuts on single variables to select one region in the multidimensional space of available variables as signal region. Such a cut-based classifier is not able to take correlations between the variables into account and does not have an optimal signal-to-background separation.

This is different for multivariate analysis techniques. Here multiple variables are combined in one classifier to yield maximum separation. Several techniques exist, e.g. *Neural Networks* (NNs) or *Boosted Decision Trees* (BDTs), both of which are commonly used in particle physics. The use of BDTs yields advantages over NNs in delivering a robust performance, while being easy to configure. Also the time to build the classifier (*training*) is significantly shorter for BDTs.

A BDT is a collection of *decision trees*, combined in one classifier, as described below. Decision trees are built from nodes each containing a binary decision based on a cut on a single variable, dividing the sample into a signal- and a background-like region, as sketched in Fig. 7.1. Opposed to the cut-based classifier each variable can occur multiple times in the tree. This increases the separation performance and makes it possible to take correlations into account. The input variables used here are the ones listed in Section 6.2 and are summarised again in Table 7.1. The implementation of the BDTs described below is performed by the Toolkit for Multivariate Data Analysis (TMVA) [67].

basic kinematics	$E_T^{\text{miss}}, p_T^W, m_T^W, m_H, \Delta R(\text{jet1}, \text{jet2})$
further kinematics	$\Delta\phi(e, E_T^{\text{miss}}), \Delta\phi_{\min}(e, \text{jet}), H_T, p_T^{\text{imb}}$
b-tagging	$\text{MV1}(\text{jet1}), \text{MV1}(\text{jet2}), \text{JFcombNNc}(\text{jet1}), \text{JFcombNNc}(\text{jet2})$

Table 7.1: *Input variables for the BDTs. Definitions are given in Section 5.10.*

7.1 BDT construction and configuration

The construction of a decision tree, called *training*, is done using simulated background and signal samples, that are normalised to the same integral. The training starts at the root node. A separation index¹ is defined for each input, x_i , variable as

$$g(X_i) = p(X_i) \cdot (p(X_i) - 1) \quad (7.1)$$

$$\text{with } p(X_i) = \frac{N_{\text{sig}}(X_i)}{N_{\text{bkg}}(X_i)}, \quad (7.2)$$

where $p(X_i)$ is the purity of the resulting sample, if a cut, $X_i > x_i$, is applied. The variable with the highest maximum of the index, $g(X_i^{\text{max}})$, is chosen and the sample is split at X_i^{max}

¹The index described here is called *Gini index*.

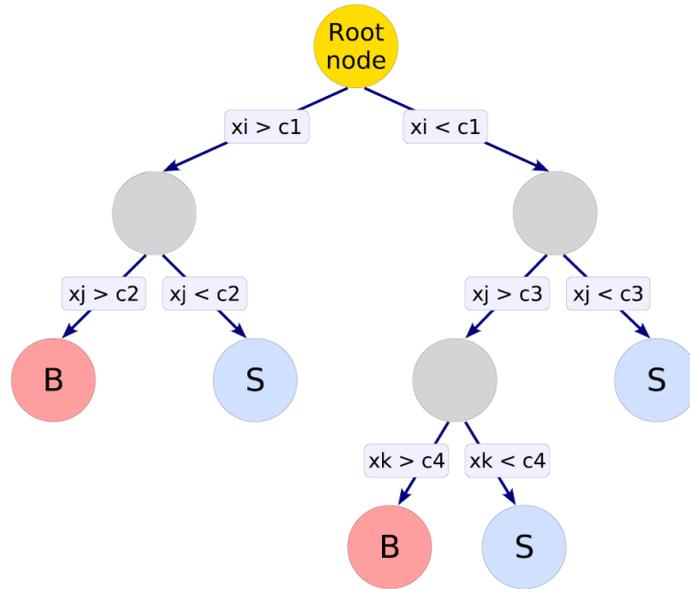


Figure 7.1: Sketch of a decision tree. The root node contains the first decision, based on variable x_i . The sample is split at the cut value $X_i \equiv c_1$ into a signal- and a background-like sub-sample. Each sub-sample is divided again, until a stopping criterion is reached. The depth of this tree is three, which is the maximum number of nodes an event can pass. The final sub-samples are called leaves and are labeled as B for background-like and as S for signal-like.

into a background- and a signal-like sub-sample. This procedure is repeated recursively for each sub-sample, adding decision nodes to the tree. The number of points, $nCuts$, at which g is calculated is limited to 20 for each variable, since larger numbers do not increase the performance of the trees significantly.

One could let such a tree grow until each sub-sample consists of only one event. In this case the tree would be extremely *overtrained*, meaning it responds to the statistical fluctuations in the training sample. Overtraining can be tested: if the separation performance of the tree is significantly better on the training sample than on an independent test sample, it is overtrained.

To reduce overtraining² the minimum number of events in the sub-samples is set to 100. Additionally, the maximum depth of the tree is limited to four. The depth is defined as the maximum number of decision nodes including the root node an event can pass. A decision tree of depth three is sketched in Fig. 7.1. These requirements put strong limitations on the size of the tree. Such a small tree alone does not have great separation performance, but this can be solved with *boosting*.

Boosted Decision Trees (BDTs) are built from a collection of decision trees. Boosting not only increases the performance of single trees, but also make them more resistant against overtraining. The boosting³ is done by training a first tree and then applying a *boost weight* α to the misclassified events. These are the signal (background) events that end up in the

²Overtraining can also be reduced by *pruning* the tree, meaning reducing its size after the training. This technique is not used here.

³The boosting described here is called *adaptive boost*, or short *AdaBoost*.

parameter	value
NTrees	400
nEventsMin	100
MaxDepth	4
BoostType	AdaBoost
SeparationType	GiniIndex
nCuts	20
PruneMethod	NoPruning

Table 7.2: Configuration parameters of the BDT training in TMVA notation [67].

background (signal) sub-samples. The weight for each tree i is calculated as

$$\alpha_i = \frac{1 - \text{err}_i}{\text{err}_i}, \quad (7.3)$$

where err_i is the missclassification rate of the tree. By construction $\text{err} \leq 0.5$, therefore $\alpha \geq 1$. A second tree is trained using the reweighted events and this is repeated until a saturation in performance is reached, in case of this thesis $N_{\text{trees}} = 400$ times.

The individual trees are combined in one output variable,

$$y(x) = \frac{1}{N_{\text{trees}}} \sum_{i=1}^{N_{\text{trees}}} \ln(\alpha_i) \cdot h_i(x), \quad (7.4)$$

where $x = (x_1, \dots, x_n)$ represents the input variables and $h_i(x) = +1(-1)$ for x ending up in an signal-(background-)like subsample in tree i .

A list of all described configuration parameters for the BDT training is given in Table 7.2. These parameters are optimised regarding the signal efficiency versus the background rejection [68].

7.2 Signal BDTs

For each signal sample with m_H from 110 to 135 GeV in 5 GeV steps a dedicated *signal BDT* is trained. Each signal BDT is trained against all backgrounds including the multijet background. The background normalisations derived in Section 8.2 are applied before the training.

To avoid biases due to overtraining the signal and background samples are split into *training* and *test* samples. The splitting is done using the event number: events with odd numbers are used for training, and events with even numbers for testing. For plotting and evaluation only the test samples are used. This applies to all BDTs.

The BDT output for $m_H = 125$ GeV is shown in Fig. 7.2 and for all other mass points in the appendix (see Fig. A.2). The naming convention is *BDTmass*, where *mass* denotes the Higgs boson mass, m_H , in GeV. One can see, that the signal is accumulated around 0, while the backgrounds are distributed towards lower values between -1 and 0. While the signal is well separated from backgrounds containing only light or c-jets (multijet, W +light, diboson, $W+c$, $W+c\bar{c}$), the separation from backgrounds containing real b-jets ($t\bar{t}$, single-top, $W+b\bar{b}$) is less significant. The most important remaining backgrounds in the signal region are $t\bar{t}$ and $W+b\bar{b}$. The shape of the backgrounds shows good agreement with data.

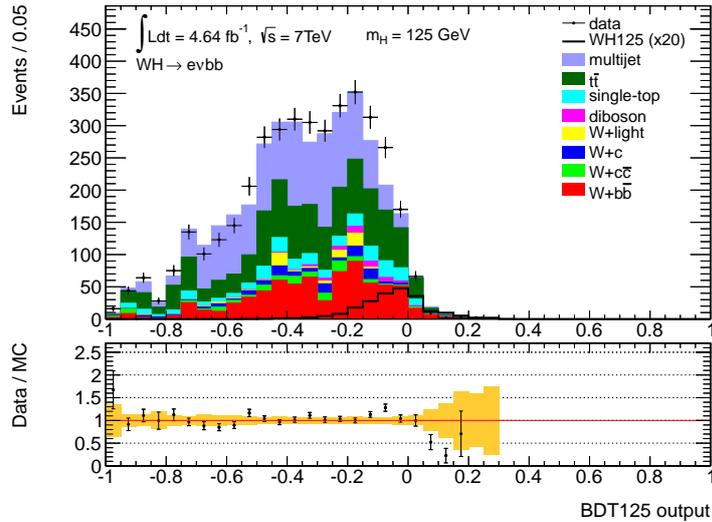


Figure 7.2: Output distribution of the signal BDT trained for a Higgs boson signal with $m_H = 125$ GeV. The backgrounds are normalised using the fit described in Section 8.2. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

7.3 Background BDTs

To separate a particular background from other backgrounds so-called *background BDTs* are trained. These are later used in the background estimation, see Section 8.2, and the evaluation of shape systematics, see Section 9.3.

Their training is accomplished by defining one background as “signal” and all others as background. The naming is $BDT.bkg$, where bkg is the background sample defined as signal. The simulated backgrounds are normalised using Monte Carlo simulations. The multijet background is scaled using the fit described in Section 8.1, before the training.

The resulting BDT outputs, trained by considering the multijet, $t\bar{t}$ and $W+b\bar{b}$ backgrounds as “signal”, are shown in Fig. 7.3 and for all other backgrounds in the appendix (see Fig. A.1). One can see that multijet and $t\bar{t}$ are well separated from the other backgrounds, while $W+b\bar{b}$ is less separated. The WH signal is widely spread in the background-BDT outputs and has a similar shape as the $W+b\bar{b}$ background. Again, the data shape is well described by the backgrounds.

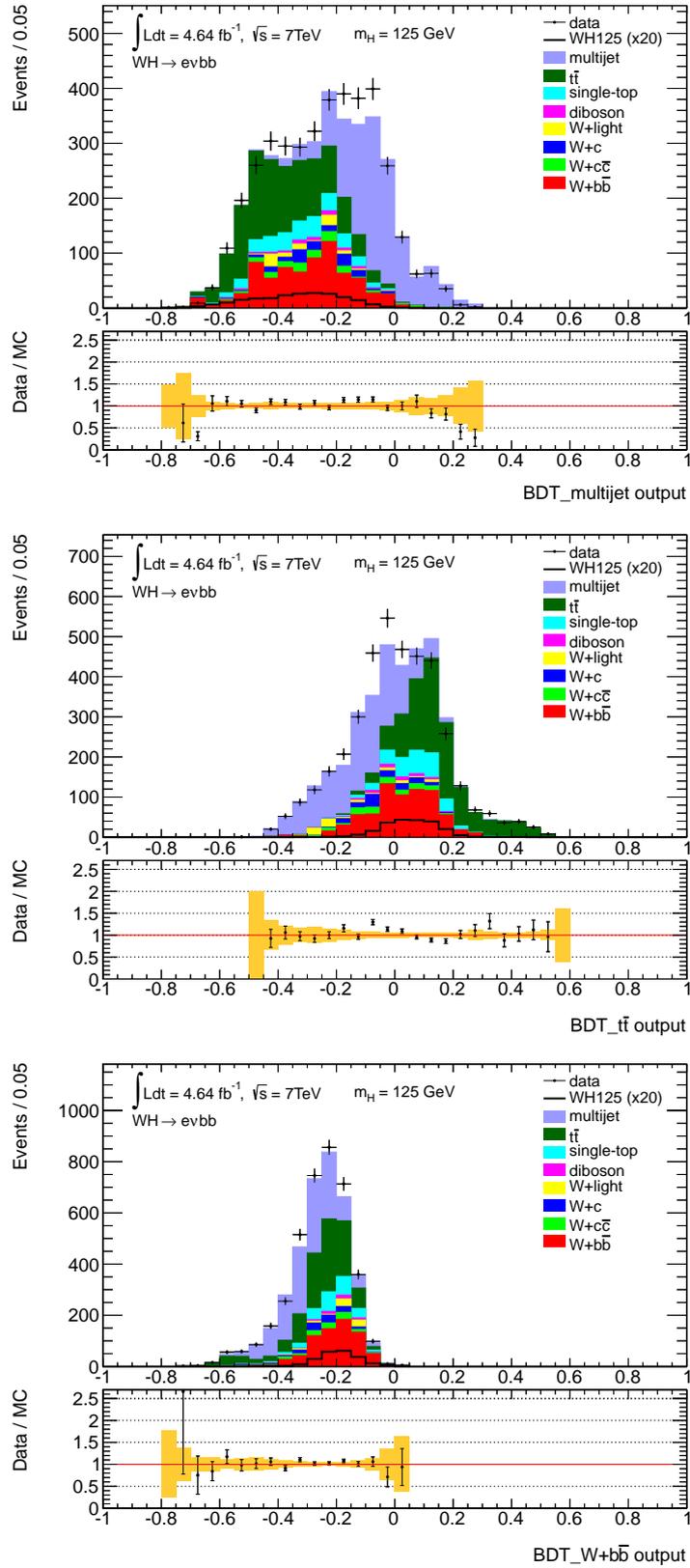


Figure 7.3: Output distributions of the background BDTs trained by defining the multijet, $t\bar{t}$ and $W+b\bar{b}$ backgrounds as signal. The backgrounds are normalised using the fit described in Section 8.2. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

Some background processes modeled by Monte Carlo simulation have large uncertainties on their normalisations and shapes. To reduce these uncertainties one can try to estimate the normalisations from data. The usual approach is to define control regions for the various backgrounds by varying the cuts of the nominal analysis. These regions should be dominated by the corresponding background. This allows to derive scale factors to correct the predictions of the simulation. The derived scale factors are used in the nominal analysis, assuming the normalisation stays the same with the different cuts. To account for this assumption systematic uncertainties have to be assigned.

In this thesis a different approach is used. The control regions are defined within the nominal event selection by training BDTs for each background, as described in Section 7.3. This approach avoids the previous assumption and is technically easier, as each background is handled the same way. Instead, systematic uncertainties on the shapes of the BDT outputs have to be considered.

The various background normalisations are fitted in the BDT-output distributions (Section 8.2) after the estimation of the multijet background is presented in the following Section 8.1.

8.1 The multijet background

As mentioned in Section 5.2, the multijet background has to be estimated using the data. This is done by extracting a multijet sample from data, from which a template in the E_T^{miss} distribution is extracted. Since the E_T^{miss} distribution separates the multijet and the other backgrounds well, it can be used to fit the normalisation of the multijet background.

To define the multijet sample, the electron isolation cuts are inverted and electron identification requirements are modified. This sample should have as little contamination of other backgrounds as possible, otherwise the normalisations from the fit are biased. These other backgrounds are the Monte Carlo simulated ones (W +light, diboson, $W+c$, $W+c\bar{c}$, $W+b\bar{b}$, single-top and $t\bar{t}$) and will be called electroweak (EW) backgrounds in the following¹.

The inverted selection should have only little effect on the shape of the E_T^{miss} distribution, otherwise the fit result is biased and cannot be translated from the inverted selection to the nominal selection. The standard inversion used in this thesis, called *InvIsoPT*, inverts the track isolation by applying the cut $0.1 < p_T^{\text{cone}}(R_0 = 0.2)/p_T^e < 0.5$ and has less requirements for the electron identification as the nominal selection (see below). The same calorimeter isolation as in the nominal selection is applied.

The E_T^{miss} distribution of the multijet template with standard inversion is shown in Fig. 8.1 (left). It has $N_{\text{raw}} = 1094$ events in total and $N_{\text{EW}} = 6.11$ events from the EW processes. The EW processes are normalised using Monte Carlo simulations, as described in Section

¹All these backgrounds are produced by weak interactions, except for the $t\bar{t}$ background. Nonetheless, the $t\bar{t}$ background is included in the EW backgrounds in the following, which can be justified by the weak decays of the top quarks.

name	$E_T^{\text{cone}}(R_0 = 0.3)$	$p_T^{\text{cone}}(R_0 = 0.2)$	ID bits
Nominal	$< 0.14 E_T^e$	$< 0.1 p_T^e$	tight++
InvIsoPTinc	-	$0.1 - 0.5 p_T^e$	medium++
InvIsoPT	$< 0.14 E_T^e$	$0.1 - 0.5 p_T^e$	medium++
Inv	$> 0.14 E_T^e$	-	tight- and (TRT or CM)
InvIso1	$> 1 \text{ GeV}$	-	tight- and (TRT or CM)
InvIso3	$> 3 \text{ GeV}$	-	tight- and (TRT or CM)
InvIso5	$> 5 \text{ GeV}$	-	tight- and (TRT or CM)
InvIso7	$> 7 \text{ GeV}$	-	tight- and (TRT or CM)
InvCM	$> 0.14 E_T^e$	-	tight- and not CM
InvCMpassTRT	$> 0.14 E_T^e$	-	tight- and TRT, but not CM

Table 8.1: Various electron isolation cuts used for the definition of the multijet templates. The isolation cuts in the nominal selection (see also Section 5.4), are denoted by *Nominal* and the inverted cuts by *InvType*. The nominal ID bits *tight++* are loosened by removing the requirements for passing the TRT and the conversion match (CM) bits, which defines *tight-*. Removing even more requirements defines *medium++*, as described in Section 4.2. Single bits are requested again, e.g. *InvTRTpassCM* requires passing of the *tight-* and CM bits, but failing of the TRT bit.

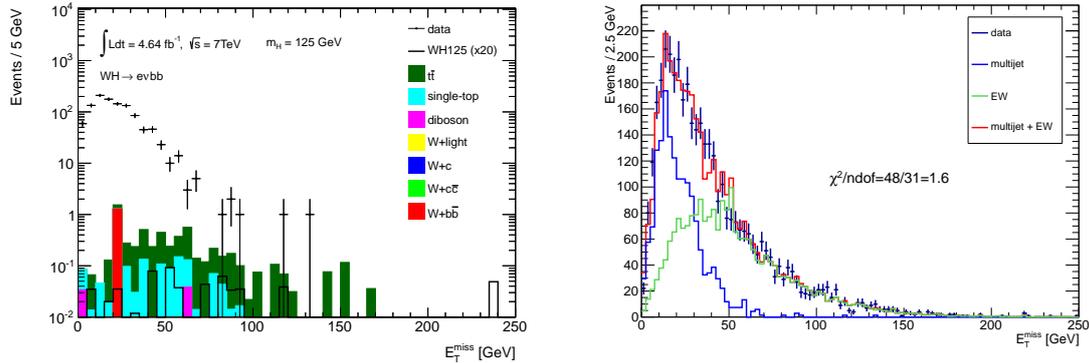


Figure 8.1: Distribution of E_T^{miss} of the multijet template when applying the standard inverted selection *InvIsoPT* (left). The backgrounds in the plot are normalised using Monte Carlo simulation. The signal distribution is scaled by a factor 20. In the right plot the E_T^{miss} distributions of the multijet (blue) and EW (green) templates are shown after the fit to data. The sum of these two distributions is shown in blue and data as black points with error bars.

5.2.6. This gives a negligible EW contamination of

$$\frac{N_{\text{EW}}}{N_{\text{raw}}} = \frac{6.11}{1094} = 0.56 \%. \quad (8.1)$$

The normalisation factors are determined in a fit of the multijet and EW template distributions to data. For the EW template and the data the nominal event selection is applied. The fit of the E_T^{miss} distribution is performed using a log-likelihood approach. This fit takes only the statistical uncertainties from data into account. The sum of the multijet template and the EW template is fitted to data in the range $10 \text{ GeV} < E_T^{\text{miss}} < 90 \text{ GeV}$, letting the

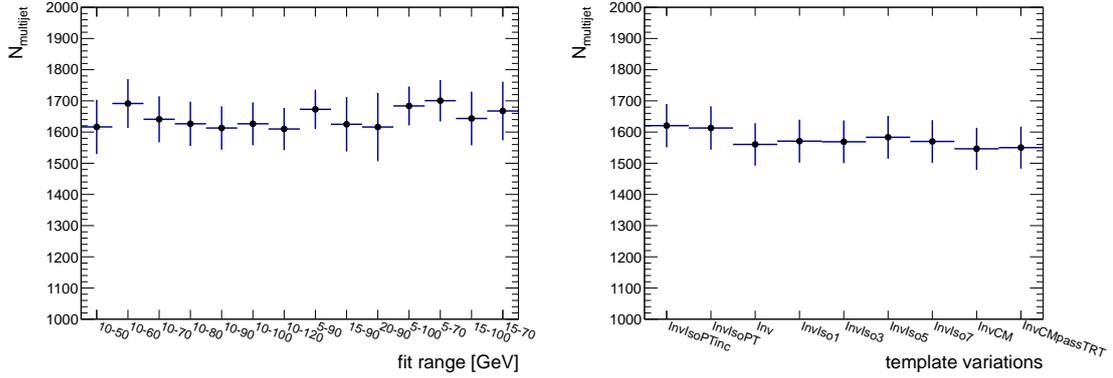


Figure 8.2: Fit results for various fit ranges (left) and template selection (left), as described in the text.

normalisations of both templates float. The fit result is shown in Fig. 8.1 (right) and leads to $N_{\text{multijet}} = 1612 \pm 69$, which corresponds to a scale factor of

$$\frac{N_{\text{multijet}}}{N_{\text{raw}}} = \frac{1612}{1094} = 1.47 \pm 0.06. \quad (8.2)$$

and the quality of the fit is fair with $\chi^2/\text{ndf} = 1.6$. This scale factor is used as event weight for the multijet sample.

The fit range is varied to evaluate its influence on the result and the results are shown in Fig. 8.2 (right). Since the results agree within statistical uncertainties, no systematic uncertainty due to the choice of the fit range is assigned.

In addition to the standard inversion cut described above, also other criteria are applied. These aim to vary the contributions from various objects in the multijet template and therefore to estimate the related systematic uncertainties. All inversions are listed in Table 8.1 and the templates are shown in the appendix (see Figs. A.3 and A.4).

A simple variation is to release the calorimeter isolation, as is done for *InvIsoPTinc*. This increases significantly the number of events in the sample selected to define the multijet template. Another variation is done by varying the the isolation cut from $E_T^{\text{cone}} > 1 \text{ GeV}$ to $E_T^{\text{cone}} > 7 \text{ GeV}$, for *InvIso1* to *InvIso7*. By increasing the isolation threshold the fractions of electrons from heavy flavour objects, such as c- and b-jets, increase in the template. By inverting the CM^2 bit (*InvCM*) the fraction of converted photons is increased.

The fit results using the different inversion criteria are shown in Fig. 8.2 (left). All results agree within statistical uncertainties and no systematic uncertainty due to the definition of the inversion is assigned.

To evaluate the statistical uncertainty arising from the multijet and the EW template toy experiments are performed. For both templates the values of the bins in the E_T^{miss} distribution are randomised at the same time using Gaussian distributions. The means values, μ , of the Gaussian distributions are set to the nominal bin values and their standard deviations, σ , are set to the statistical uncertainty. A better approach would be to use Poissonian distributions, but since most bins in the fit range contain more than ten events, the Gaussian distributions deliver a good approximation. For each of the 10,000 Monte Carlo experiments the randomised templates are fitted to the data and the results are shown in Fig. 8.3. This

²The CM (conversion match) bit is set if a track fits to a converted photon.

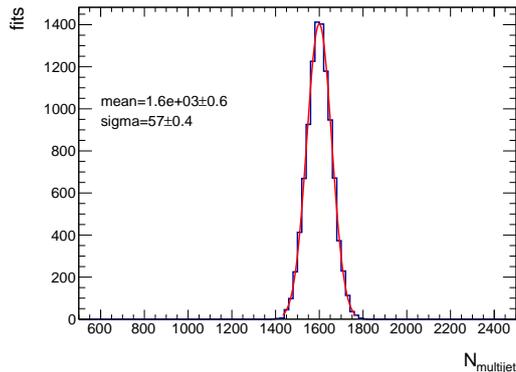


Figure 8.3: Fit results obtained by randomising the bin values in the E_T^{miss} distributions of the multijet and EW templates.

distribution is fitted with a Gaussian distribution and the observed standard deviation of $\sigma = 57.0 \pm 0.4$ is assigned as relative uncertainty of

$$\frac{\sigma}{\mu} = \frac{57}{1600} = 3.6\%. \quad (8.3)$$

The multijet sample used to extract the E_T^{miss} distribution is also used to extract other variables. The variable distributions are normalised using the scale factor derived above. The systematic uncertainties on their shapes have to be evaluated, which is done for the signal BDT outputs in Section 9.3.

8.2 Background Fits

The background estimation presented here uses a somewhat unusual approach: so-called background BDTs are trained to enrich the fractions of single backgrounds within the event selection. The training of these BDTs is described in Section 7.3. The normalisations of the various backgrounds are fitted using the output distributions of these BDTs.

Before the backgrounds can be estimated the contribution of the signal has to be reduced in the samples. This is achieved by training a dedicated BDT, called *BDT125a*. It was trained using the WH sample with $m_H = 125$ GeV against all backgrounds. Before the training the simulated backgrounds are normalised using Monte Carlo simulations, as described in Section 5.2.6. The multijet background is normalised by applying the scale factor derived in the previous section. The resulting output is shown in Fig. 8.4. The difference to the signal BDT125 is, that the background normalisations derived in this Section are not applied before the training.

A cut was placed on this output: $\text{BDT125a} < -0.2$. This reduces the WH signal to $(4.3 \pm 0.4)\%$ with $m_H = 125$ GeV. For the other Higgs boson masses in the given range the signal is reduced to maximum of $(8.5 \pm 2.0)\%$, as shown in Table 8.2. The cut efficiency for the backgrounds is about 60%. An example for the background-BDT output before and after the cut is given for $\text{BDT}_{t\bar{t}}$ in Fig. 8.5. One can see that the signal vanishes, while a significant fraction of background events pass the cut. The agreement of data and Monte Carlo simulation after this cut is still good and the shape of the BDT output is not significantly affected. The latter is expected, since the signal is widely spread in the background BDTs.

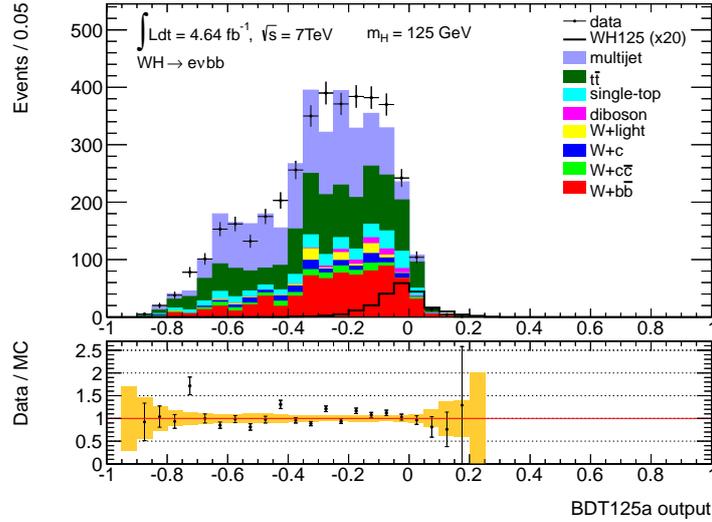


Figure 8.4: Output distribution of the signal *BDT125a*, which was trained for a Higgs boson signal with $m_H = 125$ GeV before applying the background normalisations derived by the fit described. The backgrounds in the plot are normalised using this fit. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

m_H [GeV]	efficiency [%]
110	8.5 ± 2.0
115	5.7 ± 0.5
120	5.0 ± 0.4
125	4.3 ± 0.4
130	5.0 ± 0.4
135	6.0 ± 1.5

Table 8.2: Efficiencies of the cut *BDT125a* < -0.2 on the signal samples for different values of m_H .

After the signal suppression, a fit is performed to estimate the various backgrounds. The fit uses a certain number of background-BDT output distributions simultaneously and lets the scales of defined background samples float. It is a maximum likelihood fit, taking into account the statistical uncertainties from the data sample.

The normalisations of the dominant background processes, multijet, $t\bar{t}$ and $W + b\bar{b}$ have been allowed to float in the fit. The output distributions of their corresponding background BDTs, shown in Fig. 8.5 (right) for the $t\bar{t}$ background and in Fig. 8.6 for the multijet and $W + b\bar{b}$ backgrounds, have been used in the fit. The scale factor of the other backgrounds, single-top, W +light, diboson, $W + c$ and $W + c\bar{c}$, have been fixed to one. The results of this fit, called *nominal* fit, are shown in Table 8.3.

The derived scale factors are always stated relative to the normalisation defined by the Monte Carlo simulation, which depends on the cross section from theory. However, the absolute normalisation from the fit does not depend on the cross section. Therefore, the corresponding uncertainties from theory do not influence the uncertainty on the normalisations

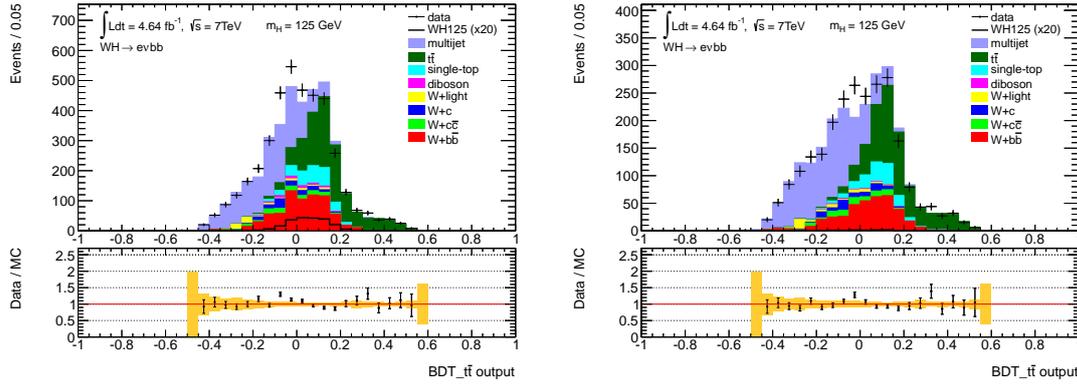


Figure 8.5: Output distributions of the background BDT that was trained by defining the $t\bar{t}$ background as signal. In the left plot no additional cut was applied and in the right plot the signal is reduced by applying the cut $BDT_{125a} < -0.2$. The backgrounds are normalised using the fit described in the text. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

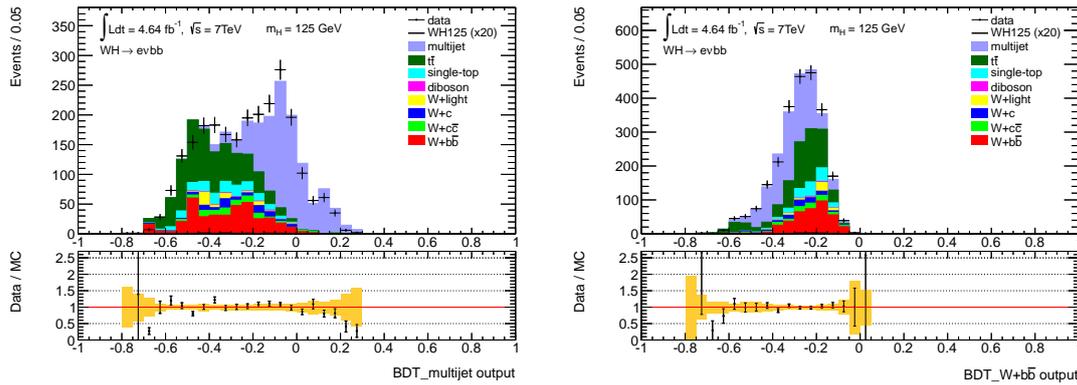


Figure 8.6: Output distributions of the background BDTs that were trained by defining the multijet (left) and $W + b\bar{b}$ (right) backgrounds as signal. The signal is reduced by applying the cut $BDT_{125a} < -0.2$. The backgrounds are normalised using the fit described in the text. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

of the fitted backgrounds.

The fit results derived above include only statistical uncertainties from data. Monte Carlo experiments were performed to estimate the statistical uncertainties from the background samples. Therefore, the shapes of the background BDT output distributions of the backgrounds were randomised. For each of the 1000 randomisations a fit of the scale factors was performed as before. The results are shown in Fig. 8.7.

The randomisation was done by multiplying the weight for each event by a random number. This random number was drawn from a discrete Poissonian distribution with mean $\lambda = 1$. The reweighted events are filled into the background-BDT histograms. By this procedure the shapes of the corresponding distributions are randomised. This is done for each of the 1000 fits.

sample	nominal	MC experiments	$\sigma_{\text{stat}}^{\text{fit}}$
multijet	0.82 ± 0.02	0.82 ± 0.03	0.05
$t\bar{t}$	1.13 ± 0.06	1.18 ± 0.05	0.07
$W + b\bar{b}$	1.58 ± 0.17	1.44 ± 0.21	0.18

Table 8.3: Background normalisation factors derived by the nominal fit and by Monte Carlo experiments. The scales are given relative to the normalisation by cross section from theory. The $\sigma_{\text{stat}}^{\text{fit}}$ is the total relative statistical error on the scale from the fits for the limit calculation, as described in the text.

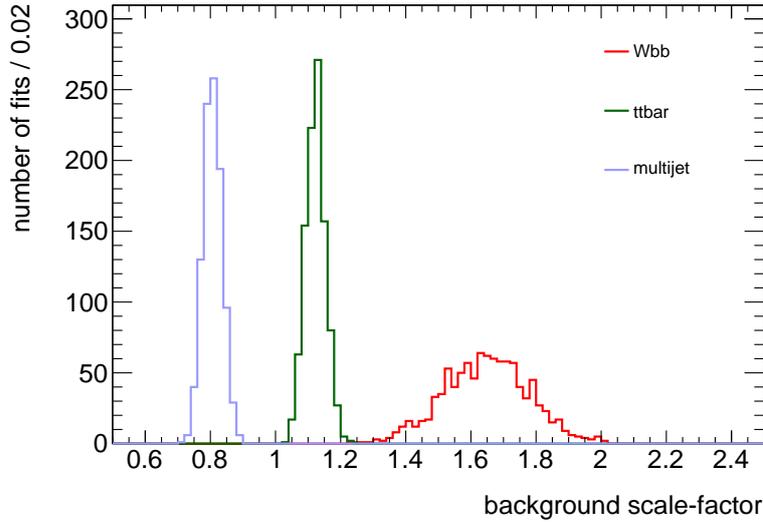


Figure 8.7: Distributions of the scale factors for the $W + b\bar{b}$, $t\bar{t}$ and multijet backgrounds determined in 1000 Monte Carlo experiments, as described in the text.

The randomised fit results are distributed in roughly Gaussian shapes, as shown in Fig. 8.7. The mean values of these distributions correspond to the derived scale factors for this fit and the root mean squares (RMS) of the distributions are taken as uncertainties. The results are listed in Table 8.3.

Although the fit procedure was validated, the derived scale factors do not perfectly agree with the nominal fit. This might be due to the Poissonian distribution being discrete, and might be investigated in a follow-up study. Nonetheless, the derived results from the Monte Carlo experiments are used in total *relative* statistical uncertainties, σ_{stat} . They are calculated for each fitted background as

$$(\sigma_{\text{stat}}^{\text{fit}})^2 = \left(\frac{\delta_{\text{nominal}}}{S_{\text{nominal}}} \right)^2 + \left(\frac{\delta_{\text{MCexp.}}}{S_{\text{MCexp.}}} \right)^2, \quad (8.4)$$

where S_{fit} are the scale factors derived by the fits and δ_{fit} their absolute uncertainties. These uncertainties are shown in Table 8.3. They are used in the limit calculation in Section 9.2, where also the systematic uncertainties on the fit results are estimated.

In this chapter limits on the cross section for the WH signal process as a function of m_H are derived. This is done by fitting the background and signal samples to data. The fit is performed using the output distributions of the BDTs that were trained to separate the $WH \rightarrow e\nu b\bar{b}$ signal from the backgrounds. In this fit systematic uncertainties are treated as nuisance parameters. Then a limit on the Higgs boson cross section is derived using the CL_s method [69].

Before this can be done, systematic uncertainties on the normalisations and shapes of the distributions for the background and signal processes are estimated to define the nuisance parameters. Various techniques are used, depending on the process:

- For the simulated processes not fitted to data (single-top, W + light, diboson, $W + c$ and $W + c\bar{c}$), the uncertainties from reconstruction and identification of the final state objects are estimated by Monte Carlo simulation. Additionally, the uncertainty on the theoretical cross sections and on the luminosity are taken into account.
- For the simulated processes that are normalised using the fit in Section 8.2 ($W + b\bar{b}$ and $t\bar{t}$) systematic uncertainties are derived from the fit results.
- For the multijet background the systematic uncertainty is taken from the analysis in Ref. [1], which is used as reference analysis.

Furthermore, there are systematic uncertainties on the shapes of the BDT-output distributions. These are estimated by two distinct approaches:

- Various parameters of the reconstruction and identification are varied by their corresponding uncertainties.
- The backgrounds are reweighted to data in control regions defined by the background BDTs. This novel technique is described below.

The various sources of systematic uncertainties are listed in the following Section 9.1, their influences on the BDT output distributions are discussed in the Sections 9.2 and 9.3 and the limits are extracted in Section 9.4.

9.1 Sources of systematic uncertainties

Systematic uncertainties arise from a variety of sources. First, there are uncertainties on the normalisation of the backgrounds and the signal: the luminosity delivered by the LHC has an uncertainty of $\sigma_{\text{lumi}} = \pm 3.9\%$ [70]. The theoretical cross sections of the simulated processes have uncertainties of $\pm 5\%$ to $\pm 15\%$, as listed in Table 5.3.

Additional uncertainties arise from the reconstruction and identification of the final state objects, that affect the normalisations and shapes of output distributions. These are derived by varying certain parameters within their respective uncertainties. The various parameters are listed in Table 9.1.

name	uncertainty	description
ElecScale		electron energy scale
ElecSmear		electron energy resolution
ElecSFID		electron scale factor from the identification
ElecSFReco		electron scale factor from the reconstruction
ElecSFTrig		electron scale factor from the triggers
ElecSFIso		electron scale factor from the isolation
JES	$\pm (2.5 - 4) \% [39]$	jet energy scale
JERSyst	$\pm 10 \% [71]$	jet energy resolution
METresolution		E_T^{miss} soft term resolution
METscale		E_T^{miss} soft term scale
FlavTagL		MV1 light jet tagging efficiency
FlavTagC		MV1 c-jet tagging efficiency
FlavTagB	$\pm (5 - 19) \% [41]$	MV1 b-jet tagging efficiency

Table 9.1: Sources of systematic uncertainties due to identification and reconstruction effects. The soft terms refer to calorimeter cells that are not assigned to a physical object in the reconstruction, as explained in Section 4.8. Uncertainties are listed for the parameters that have the strongest effect on the dominant backgrounds ($W + b\bar{b}$, $t\bar{t}$), as shown in Table 9.2.

Also the theoretical descriptions of the simulated processes have uncertainties that effect the shapes, e.g. renormalisation and factorisation scales [56]. One possibility to evaluate this effect is to use different Monte Carlo simulations. But a more inclusive approach, that does not rely completely on simulation, is performed in Section 9.3 by reweighting the background shape to data.

9.2 Normalisation systematics

The systematic uncertainties on the normalisations of the background and signal processes are estimated with various approaches, depending on the type of process:

Processes estimated by Monte Carlo simulation The systematic uncertainties on the normalisation of the processes that are estimated using Monte Carlo (MC) simulation (single-top, $W + \text{light}$, diboson, $W + c$ and $W + c\bar{c}$) are estimated by propagating the uncertainties on the parameters listed in Table 9.1 to the samples. They are calculated as

$$\sigma = \frac{N_{\text{MC}}^{\text{sys}}}{N_{\text{MC}}^{\text{nom}}} - 1, \quad (9.1)$$

where $N_{\text{MC}}^{\text{nom}}$ is the nominal number of events of a certain simulated background and $N_{\text{MC}}^{\text{sys}}$ the number of events when applying a certain variation. The results are shown in Table 9.2.

One can see that the dominant systematic uncertainties on the normalisation arise from the jet-energy scale and resolution and from the flavour tagging scale factors. Electron energy and scale factors cause only small uncertainties. There are no uncertainties on the normalisation coming from the missing transverse energy, since no cut on this variable is applied in the selection.

name		$W+b\bar{b}$	$W+c\bar{c}$	$W+c$	diboson	W +light	single-top	$t\bar{t}$	WH
ElecScale	down	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	up	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
ElecSmear	down	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	up	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
ElecSFID		± 0.01	± 0.01	± 0.01	± 0.01	± 0.01	± 0.01	± 0.01	± 0.01
ElecSFReco		± 0.01	± 0.01	± 0.01	± 0.01	± 0.01	± 0.01	± 0.01	± 0.01
ElecSFTrig		± 0.01	± 0.01	± 0.01	± 0.01	± 0.01	± 0.01	± 0.01	± 0.01
ElecSFIso		± 0.02	± 0.01	± 0.02	± 0.02	± 0.01	± 0.02	± 0.02	± 0.02
JERSyst		± 0.07	± 0.06	± 0.12	± 0.10	± 0.31	± 0.05	± 0.04	± 0.06
JES	down	0.05	-0.03	-0.03	0.02	-0.20	0.07	0.15	0.04
	up	-0.04	-0.03	-0.09	-0.07	-0.29	-0.10	-0.15	-0.06
FlavTagL		± 0.01	± 0.03	± 0.07	± 0.01	± 0.20	± 0.00	± 0.00	± 0.00
FlavTagC		± 0.02	± 0.22	± 0.16	± 0.04	± 0.00	± 0.02	± 0.01	± 0.00
FlavTagB		± 0.11	± 0.00	± 0.00	± 0.09	± 0.00	± 0.12	± 0.13	± 0.13
$\sigma_{\text{sys}}^{\text{MC}}$		± 0.15	± 0.23	± 0.24	± 0.16	± 0.51	± 0.17	± 0.20	± 0.16

Table 9.2: Relative systematic uncertainties on the normalisation of the various samples due to identification and reconstruction effects calculated from Monte Carlo simulation. The WH signal is listed for $m_H = 125 \text{ GeV}$. Where available the variations are listed as up and down separately, the others are symmetrised. The bottom line states the quadratic sum of the individual contributions, as described in the text.

uncertainty	$W+b\bar{b}$	$W+c\bar{c}$	$W+c$	diboson	W +light	single-top	$t\bar{t}$	WH
σ_{cs}	0.05	0.05	0.05	0.07	0.05	0.08	0.11	0.15
$\sigma_{\text{stat}}^{\text{MC}}$	0.06	0.12	0.11	0.03	0.36	0.02	0.01	0.02
$\sigma_{\text{sys}}^{\text{MC}}$	0.15	0.23	0.24	0.16	0.51	0.17	0.20	0.16
$\sigma_{\text{norm}}^{\text{MC}}$	0.17	0.26	0.27	0.18	0.63	0.19	0.23	0.22

Table 9.3: Relative uncertainties on the normalisations of the various backgrounds taken from Monte Carlo simulation. The WH signal is listed for $m_H = 125 \text{ GeV}$. The uncertainties on the theoretical cross sections, σ_{cs} , are taken from Table 5.3. The statistical uncertainties, $\sigma_{\text{stat}}^{\text{MC}}$, are calculated from the number of events in the samples. The systematic uncertainties, $\sigma_{\text{sys}}^{\text{MC}}$, are taken from Table 9.2. Finally, the total uncertainty, $\sigma_{\text{norm}}^{\text{MC}}$, is the quadratic sum of the individual uncertainties above.

The uncertainties on the W +light background are large due to the low number of events in the sample. This unwanted statistical effect is not further evaluated, since this background has only a small contribution in the signal region and will have not a large effect on the final result.

The uncertainties of the fitted backgrounds ($W+b\bar{b}$ and $t\bar{t}$) are only listed for completeness and are not used later.

The various systematic uncertainties are combined in a total relative uncertainty, $\sigma_{\text{sys}}^{\text{MC}}$, for each background. This is done by quadratically summing all systematic uncertainties resulting in up-variations ($\sigma > 0$) and all those resulting in down-variations ($\sigma < 0$) separately and then choosing the larger one:

$$(\sigma_{\text{sys}}^{\text{MC}})^2 = \max \left\{ \sum_{\sigma > 0} \sigma^2, \sum_{\sigma < 0} \sigma^2 \right\}. \quad (9.2)$$

name		$W + b\bar{b}$	$t\bar{t}$
ElecScale	down	-0.04	0.02
	up	0.00	-0.01
ElecSmear	down	-0.04	0.02
	up	-0.01	0.00
ElecSFID		± 0.01	± 0.00
ElecSFReco		± 0.01	± 0.00
ElecSFTrig		± 0.00	± 0.00
ElecSFIso		± 0.01	± 0.00
JERSyst		± 0.17	± 0.05
JES	down	-0.05	0.04
	up	0.05	-0.01
METresolution	down	-0.01	0.01
	up	0.03	-0.01
METscale	down	-0.04	0.02
	up	0.06	-0.03
FlavTagL		± 0.02	± 0.01
FlavTagC		± 0.06	± 0.00
FlavTagB		± 0.05	± 0.02
$\sigma_{\text{sys}}^{\text{fit}}$		± 0.21	± 0.07

Table 9.4: *Relative systematic uncertainties on the $W + b\bar{b}$ and $t\bar{t}$ prediction calculated from the background normalisation fits. Where available the variations are listed as up and down separately, the others are symmetrised. The bottom line states the quadratic sum of the individual contributions, as described in the text.*

The results are shown in Table 9.3. Also shown are the uncertainties on the cross sections, σ_{cs} , and the statistical uncertainties, $\sigma_{\text{stat}}^{\text{MC}}$, taken from the number of events in the Monte Carlo samples. A total relative uncertainty on the normalisation, $\sigma_{\text{norm}}^{\text{MC}}$, is calculated as

$$\sigma_{\text{norm}}^{\text{MC}} = \sqrt{\sigma_{\text{cs}}^2 + (\sigma_{\text{stat}}^{\text{MC}})^2 + (\sigma_{\text{sys}}^{\text{MC}})^2}. \quad (9.3)$$

Backgrounds estimated from data The backgrounds that are estimated from fits to data ($W + b\bar{b}$, $t\bar{t}$) are the dominant ones and their uncertainties will have a strong effect on the final result. Their systematic uncertainty on the normalisation is derived using the fit. The number of events from the fit in the nominal analysis, $N_{\text{fit}}^{\text{nom}}$, are compared with the ones from the fit when applying a variation, $N_{\text{fit}}^{\text{sys}}$. For each variation the fit is redone. These variations alter the shape of the background templates and thus also the fit result. The systematic uncertainty is then calculated as

$$\sigma = \frac{N_{\text{fit}}^{\text{sys}}}{N_{\text{fit}}^{\text{nom}}} - 1 \quad (9.4)$$

and the results are shown in Table 9.4.

These results reflect the effect of the shape variations of the backgrounds on the fit result. The variation of the normalisation by cross section uncertainties instead does not have an effect on the normalisation from the fit. Therefore the uncertainty on the cross section, σ_{cs} , does not have to be taken into account for the fitted backgrounds.

uncertainty	$W+b\bar{b}$	$t\bar{t}$	multijet
$\sigma_{\text{stat}}^{\text{fit}}$	0.18	0.07	0.05
$\sigma_{\text{sys}}^{\text{fit}}$	0.21	0.07	—
$\sigma_{\text{norm}}^{\text{fit}}$	0.28	0.10	0.05

Table 9.5: *Relative uncertainties on the normalisations of the various backgrounds estimated by the fits. The statistical uncertainties, $\sigma_{\text{stat}}^{\text{fit}}$, are taken from Table 8.3. The systematic uncertainties, $\sigma_{\text{sys}}^{\text{fit}}$, are taken from Table 9.4. Finally, the total uncertainty, $\sigma_{\text{norm}}^{\text{fit}}$, is the quadratic sum of the individual uncertainties above.*

uncertainty	$W+b\bar{b}$	$W+c\bar{c}$	$W+c$	diboson	$W+\text{light}$	single-top	$t\bar{t}$	WH	multijet
σ_{norm}	0.28	0.26	0.27	0.18	0.63	0.19	0.10	0.22	0.30

Table 9.6: *Final systematic uncertainties on the normalisation the various backgrounds and the signal. The WH signal is listed for $m_H = 125 \text{ GeV}$.*

The individual uncertainties are combined in a total systematic uncertainty, $\sigma_{\text{sys}}^{\text{fit}}$, as

$$\left(\sigma_{\text{sys}}^{\text{fit}}\right)^2 = \max \left\{ \sum_{\sigma > 0} \sigma^2, \sum_{\sigma < 0} \sigma^2 \right\}. \quad (9.5)$$

The results are listed in Table 9.5 together with the statistical uncertainties, $\sigma_{\text{stat}}^{\text{fit}}$, taken from the fits in Section 8.2. The total relative uncertainty on the normalisation, $\sigma_{\text{norm}}^{\text{fit}}$, is calculated as

$$\sigma_{\text{norm}}^{\text{fit}} = \sqrt{\left(\sigma_{\text{stat}}^{\text{fit}}\right)^2 + \left(\sigma_{\text{sys}}^{\text{fit}}\right)^2}. \quad (9.6)$$

The systematic uncertainty on the normalisation of the multijet background cannot be estimated by this approach, since this background is completely taken from data. Instead the following approach is pursued.

The multijet background The available uncertainty on the normalisation of the multijet background is the statistical uncertainty of $\pm 5\%$ on the BDT-based fit result, as listed in Table 9.5. Since this does not account for systematic effects, a conservative uncertainty of $\pm 30\%$ is assumed, which follows the discussion in Ref. [1]. This has to be improved for a more conclusive study.

Combination The final systematic uncertainties on the normalisations of the various processes are denoted as σ_{norm} . They are taken from the fits for the fitted backgrounds ($W+b\bar{b}$ and $t\bar{t}$) as $\sigma_{\text{norm}} = \sigma_{\text{norm}}^{\text{fit}}$ and from Monte Carlo simulation for the other processes as $\sigma_{\text{norm}} = \sigma_{\text{norm}}^{\text{MC}}$. The only exception is the multijet background, where a fixed value of $\sigma_{\text{norm}} = \pm 30\%$ is taken.

The results are listed in Table 9.6. The uncertainty on the luminosity is not included here and is considered independently in the limit calculation.

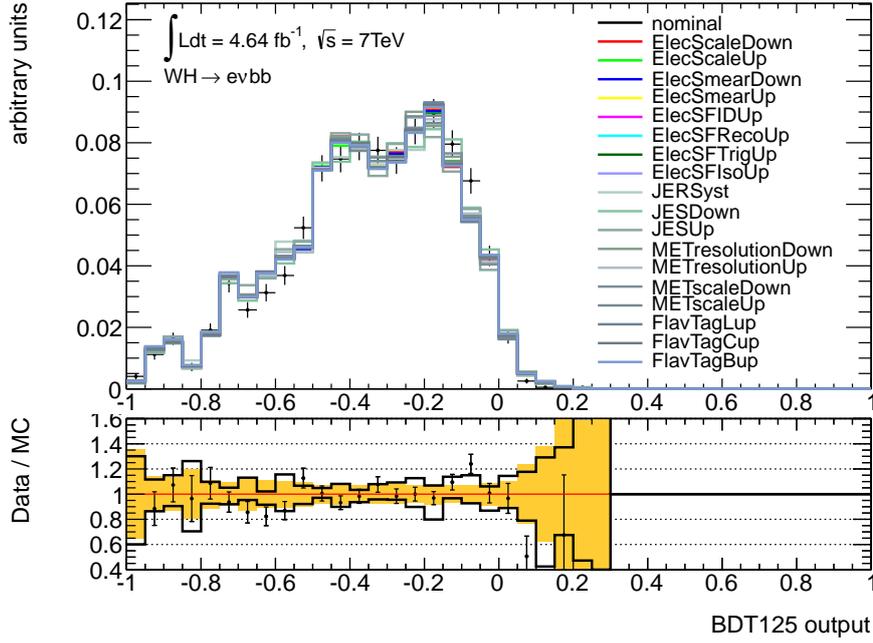


Figure 9.1: Output distributions of the signal BDT trained for a Higgs boson signal with $m_H = 125 \text{ GeV}$. The shape differ resulting from variations of the parameters from reconstruction and identification. The distributions are normalised to the same area. The background normalisations are taken from the fit described in Section 8.2 and are not affected by the variations. The orange band in the ratio plot denotes the statistical uncertainty on the total background without variation. The black line in the ratio shows the sum of squares of the uncertainties, as described in the text.

9.3 Shape systematics

The systematic uncertainties on the shapes of the signal-BDT outputs are estimated using two approaches:

Varying parameters The systematic uncertainties on the shapes of the signal-BDT outputs can be derived from the Monte Carlo simulation. This is done by comparing the BDT-output distribution from the nominal analysis with the distributions obtained when varying the parameters from reconstruction and identification, as listed in Table 9.1.

An example is given in Fig. 9.1: the nominal output distribution of the signal BDT with $m_H = 125 \text{ GeV}$ is shown together with the distributions obtained when applying the variations. The distributions are normalised to the same area, since only the shape difference is of interest. Where only the bin values of the up-variations, n_i^{up} , are shown for each bin i the down-variations are derived by

$$n_i^{\text{down}} = 2n_i^{\text{nom}} - n_i^{\text{up}}, \quad (9.7)$$

where n_i^{nom} are the bin values of the nominal distribution.

The systematic uncertainty for each bin i is calculated as

$$\sigma_i = \frac{n_i^{\text{sys}}}{n_i^{\text{nom}}} - 1. \quad (9.8)$$

The total up and down uncertainties are then derived as

$$\sigma_i^{\text{up}} = +\sqrt{\sum_{\sigma_i > 0} \sigma_i^2} \quad (9.9)$$

$$\text{and } \sigma_i^{\text{down}} = -\sqrt{\sum_{\sigma_i < 0} \sigma_i^2}. \quad (9.10)$$

These are shown as black line in the ratio plot in Fig. 9.1. This total shape uncertainty is not necessary larger than the statistical uncertainty (orange band), since they are a priori independent. However, migration effects play a role for small samples (e.g. $W + \text{light}$), resulting in larger systematic uncertainties.

The up and down uncertainties are combined in a final uncertainty from reconstruction and identification for each bin i as

$$\sigma_i^{\text{reco}} = \max \left\{ \sigma_i^{\text{up}}, -\sigma_i^{\text{down}} \right\}. \quad (9.11)$$

These are used for the limit calculation. This is a simplification of the usual approach of taking each parameter separately into account, as done in Ref. [1].

In Fig. 9.2 the shape differences from the variations of the individual parameters from reconstruction and identification are shown. It is apparent that the uncertainties on the jet energy scale and flavour tagging scales are dominant. The shape uncertainties for other mass points are shown in the appendix (see Fig. A.5).

This approach of summing a large number of variations might overestimate the shape uncertainty, since statistical effects play a role. This can be seen by the large systematic normalisation uncertainties on the $W + \text{light}$ background, which has a small number of events in the sample. On the other, hand this approach does not include the shape uncertainties of the b-tag variables that are used in the BDTs. This is also the case in Ref. [1], since these uncertainties are not available up to now.

Also the shape uncertainties arising from the multijet background and from the theoretical description of the processes simulated by Monte Carlo are not included in this analysis nor in Ref. [1]. Therefore, another approach has been developed, where the shape difference between data and the backgrounds in BDT outputs is evaluated, as discussed in the following.

BDT reweighting The approach of estimating the systematic uncertainty on the shape of the signal-BDT output distributions presented here uses a reweighting of the backgrounds to data. This is done in control regions defined by the background BDTs, which are defined in Section 7.3. These control regions contain the same events as the nominal event selection, but individual backgrounds are separated from the others. From the shape difference between the backgrounds and data in these control regions weights are derived.

The output distributions of the background BDTs are normalised to the same area for data and total background. For each bin i in each output distribution j a weight, w_{ij} , is calculated as

$$w_{ij} = \frac{n_{ij}^{\text{data}}(x)}{n_{ij}^{\text{MC}}(x)}. \quad (9.12)$$

These weights are applied to the events of all backgrounds. This procedure alters the shapes of the background distributions, but their normalisations are untouched.

In an example the BDT $t\bar{t}$ is used to calculate these weights and they are applied to the events of the backgrounds. If now the shape of the total background is compared to data in

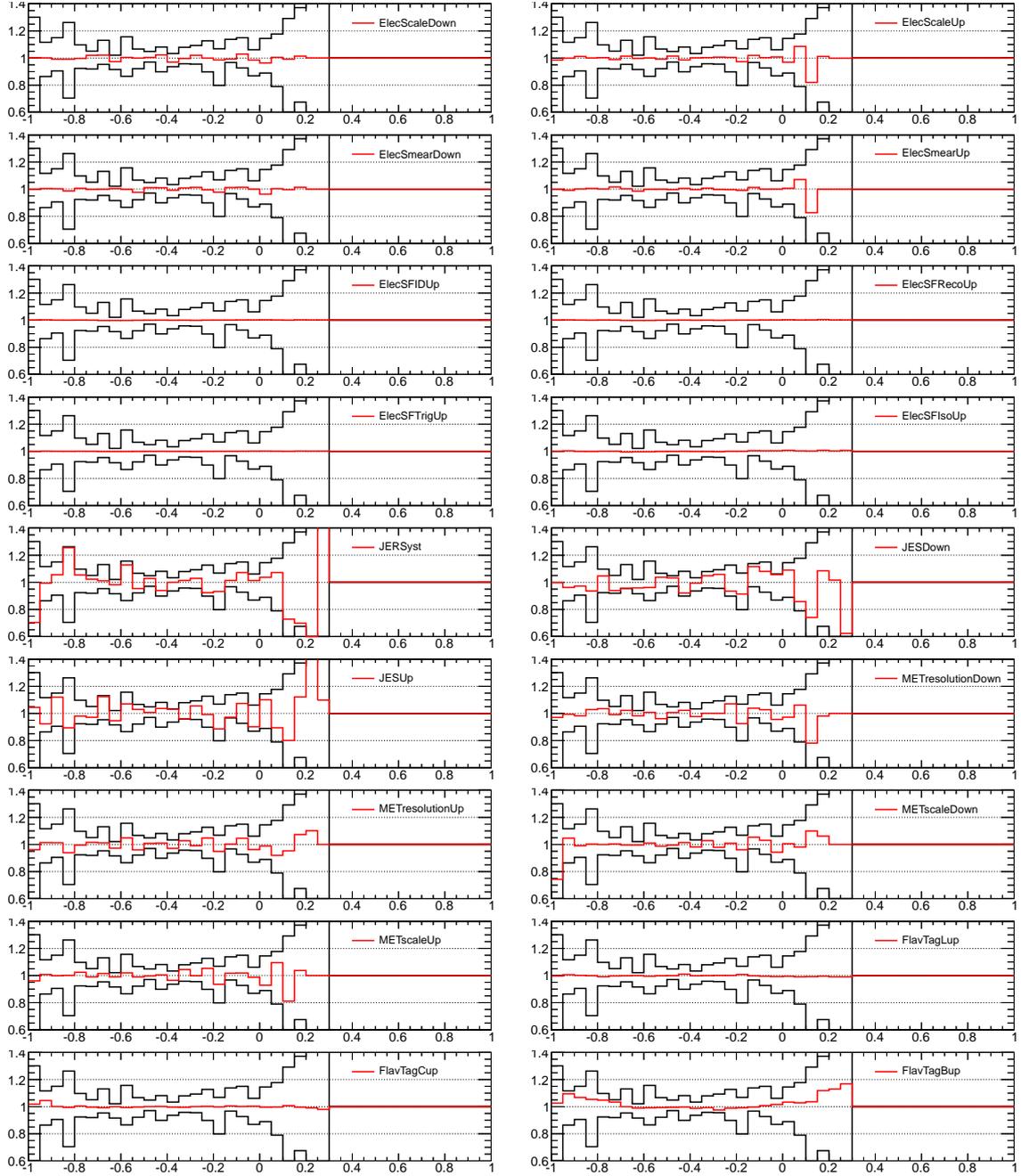


Figure 9.2: Systematic uncertainties on the shape of the signal BDT output trained for a Higgs boson signal with $m_H = 125$ GeV. The BDT output value is plotted on the x-axis and the ratio n^{sys}/n^{nom} is plotted on the y-axis. The shapes differences resulting by varying the parameters from reconstruction and identification are plotted as red line. The black line is the sum of squares of the uncertainties, as also shown in Fig. 9.1.

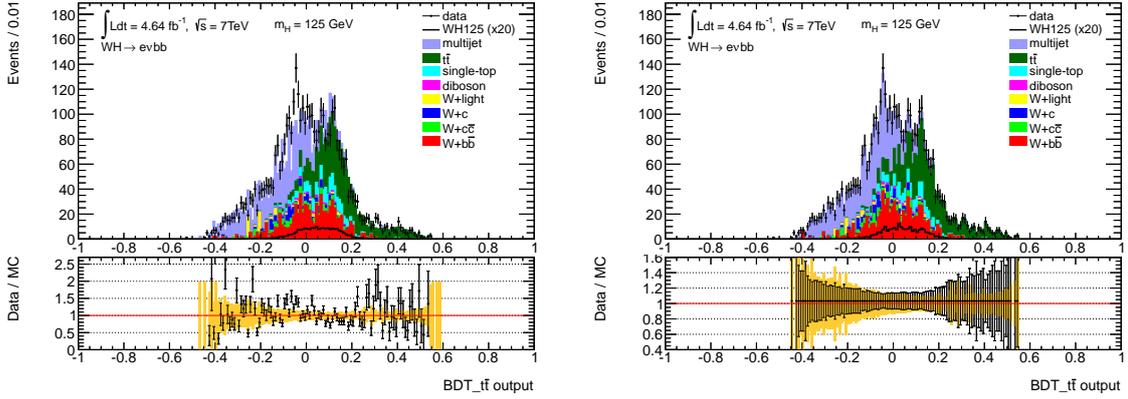


Figure 9.3: Output distributions of the background BDT trained by defining the $t\bar{t}$ background as signal. On the left the nominal distribution is shown and on the right the BDT reweighting is applied. Since the BDT $t\bar{t}$ was used to calculate the weights, the shapes of data and background agree perfectly by definition in the right plot. The backgrounds are normalised using the fit described in Section 8.2. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

the BDT $t\bar{t}$ output, one finds that they agree perfectly by definition, as shown in Fig. 9.3. In other output distributions however, the shapes of the background will be altered too, but do not necessarily agree with data.

This *BDT reweighting* is done using all available background BDTs, resulting each in a different shape. An example is plotted in Fig. 9.4: the nominal distribution of the signal BDT with $m_H = 125$ GeV is shown together with the reweighted distributions.

Unlike for the shape uncertainties from reconstruction and identification the total up and down uncertainties are taken to be the largest variation:

$$\sigma_i^{\text{up}} = \max_{\sigma_{ij} > 0} \{\sigma_{ij}\} \quad (9.13)$$

$$\text{and} \quad \sigma_i^{\text{down}} = \min_{\sigma_{ij} < 0} \{\sigma_{ij}\}, \quad (9.14)$$

where i denotes the bin and j the background BDT that was used for reweighting. These largest up-/down-variations are shown in the ratio plot of Fig. 9.4 and the shape uncertainties for the other mass points are shown in the appendix (see Fig. A.6).

This approach of taking the largest variations is certainly not optimal. But one cannot proceed as before and sum all variations, since each reweighting is done using the same results, thus double counting would occur. A better approach might be to iteratively summarise the shape differences from all background BDTs. This can be considered in a follow-up study.

The final uncertainties on the shapes from the BDT reweighting, σ_i^{BDT} , are calculated as before in Equation 9.11 and are used in the limit calculation.

This approach of the BDT reweighting is not yet validated and it is not clear, if it covers all uncertainties. The validation of this method has to be considered in a follow-up study. However, this approach is sensitive to the shape difference between the estimated backgrounds and data. Therefore, it is expected to include a variety of uncertainties, e.g the ones arising from b-tagging, the theoretical description of the simulated processes and the shape of the multijet background.

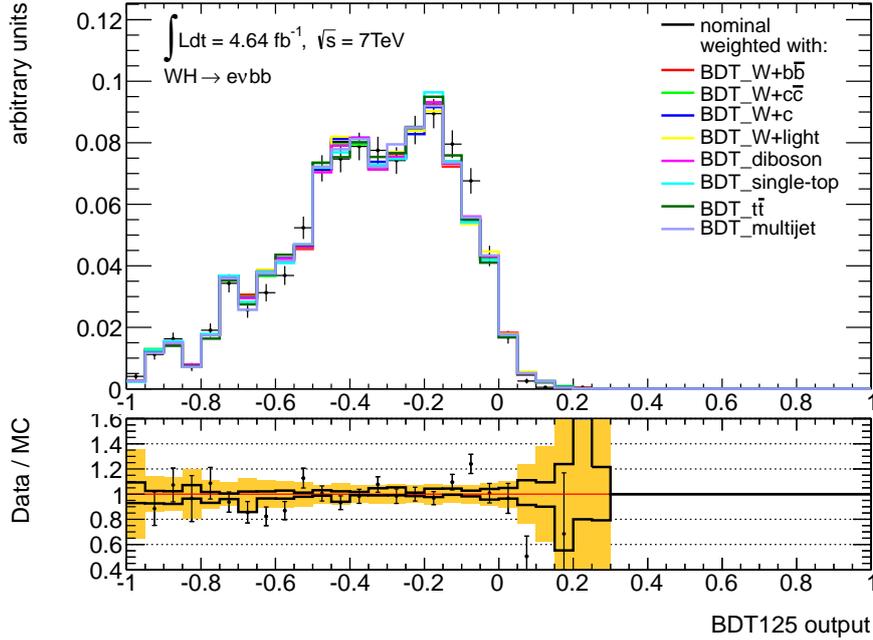


Figure 9.4: Output distributions of the signal BDT trained for a Higgs boson signal with $m_H = 125 \text{ GeV}$. The shape is varied by applying the BDT reweighting using the background BDTs. The distributions are normalised to the same area. The background normalisations are taken from fit described in Section 8.2 and are not affected by the variations. The orange band in the ratio plot denotes the statistical uncertainty on the total background without variation. The black line in the ratio shows the maximum up-/down-variation, as described in the text.

Combination The systematic uncertainties on the shape from reconstruction and identification and from the BDT reweighting are combined in a total shape uncertainty as

$$\sigma_i^{\text{shape}} = \sqrt{(\sigma_i^{\text{reco}})^2 + (\sigma_i^{\text{BDT}})^2}. \quad (9.15)$$

This uncertainty includes some double counting, since the uncertainties from reconstruction and identification are at least partly included in the BDT reweighting. However, this shape uncertainty is expected to be conservative and is taken to be the final shape uncertainty.

The shape uncertainty on the signal process is not considered so far.

9.4 Limits

The limits on the Higgs boson cross-section are calculated using the standard ATLAS procedure with RooFit [72] as fitting algorithm and the CL_s method [69] for limit extraction.

The CL_s method is based on profile likelihood ratios for the signal plus background ($s+b$) and the background only (b) hypotheses [73]. From these *test statistics* the CL_s is calculated as

$$CL_s = \frac{p_{s+b}}{1 - p_b} < \alpha, \quad (9.16)$$

where p_{s+b} and p_b are the probabilities of the data being described by the $s+b$ - and the b -hypotheses, respectively. By setting $\alpha = 5\%$ a confidence level of $CL = 1 - \alpha = 95\%$ for

the upper limit of the signal hypothesis is defined. The divisor $(1 - p_b)$ in Equation 9.16 was introduced to make this limit save against excluding models where $s \ll b$ [69].

The test statistics are derived by fitting the signal and the various backgrounds simultaneously to data in the signal-BDT output distributions. For each mass point from $m_H = 110$ to 135 GeV a dedicated fit is performed using the corresponding BDT output.

Various limits, (a) to (e), have been calculated, taking into account different uncertainties for comparison:

- (a) the statistical uncertainties, σ_{stat} of the signal and each background;
- (b) the total normalisation uncertainties, σ_{norm} , of the signal and each background, which include σ_{stat} ;
- (c) the total normalisation uncertainties, σ_{norm} , and the shape uncertainty from the BDT reweighting, σ^{BDT} ;
- (d) the total normalisation uncertainties, σ_{norm} , and the shape uncertainty from reconstruction and identification, σ^{reco} ;
- (e) the total normalisation uncertainties, σ_{norm} , and the total shape uncertainty, σ^{shape} .

In addition, the uncertainty on the luminosity, σ_{lumi} , is taken into account for each limit. The uncertainties listed are taken to be uncorrelated and are fitted as nuisance parameters. Since eight backgrounds are taken into account separately, 10 nuisance parameters are fitted in (a) and (b) and 11 parameters in (c) to (e). The results for the expected and observed limits are plotted in Fig. 9.5 and listed in Table 9.7. The shown results of the reference analysis [1] are discussed in the next Section 9.5.

Looking at the results one can see that the expected and the observed limits always agree within errors. There is no significant deficit nor excess visible. The limits increase towards higher masses. This is expected, since the branching ratio and the cross section for $H \rightarrow b\bar{b}$ drop in the high mass region.

When comparing the results (a) to (e) one can see, that by adding more systematic uncertainties the expected limits rise, as expected. The observed limits do not follow this trend strictly, but they always agree within uncertainties.

The limits for $m_H = 120$ GeV behave in an unexpected way, since they become smaller when more systematic uncertainties are added. Further studies are needed to understand this effect, e.g. by investigating the fit results of the nuisance parameters.

9.5 Comparison to the reference analysis

In the following the results derived above are compared to the results of the reference analysis [1], which uses a cut-based approach and the $m_{b\bar{b}}$ distribution for the limit calculation. The results are plotted in Fig. 9.5 and listed in Table 9.7 for both analyses.

The comparison of the results has to be done carefully, since there exist differences in the two analyses. The most important is, that the reference analysis also includes the muon channel. Therefore the number of expected signal and background events is about doubled compared to the analysis presented here.

Other differences exist in the background estimation, done with dedicated control regions and a sideband fit of the m_H distribution for certain backgrounds. Also the handling of the systematic uncertainties is somewhat different. More parameters of the identification and reconstruction are taken into account. On the other hand there is no need to estimate the

(1) expected 95 % CL exclusion limits in $\sigma/\sigma_{\text{SM}}$						
m_H [GeV]	(a)	(b)	(c)	(d)	(e)	Hbb paper
110	4.0	4.0	4.4	5.5	6.0	4.2
115	3.7	3.9	4.3	5.1	5.2	4.9
120	4.0	4.5	4.6	4.2	2.7	5.9
125	5.4	5.6	6.2	7.2	7.4	7.5
130	6.8	7.0	7.5	8.7	8.9	9.2
135	11.7	12.9	13.6	16.8	17.5	

(1) observed 95 % CL exclusion limits in $\sigma/\sigma_{\text{SM}}$						
m_H [GeV]	(a)	(b)	(c)	(d)	(e)	Hbb paper
110	6.1	7.2	6.6	7.6	7.8	3.3
115	3.5	3.0	2.6	3.2	3.2	4.0
120	4.0	3.6	3.5	3.4	2.8	4.9
125	5.5	4.0	3.7	4.8	5.0	5.5
130	6.5	5.0	5.3	6.8	7.2	5.9
135	12.2	12.7	13.4	14.1	14.8	

Table 9.7: *Expected (1) and observed (2) 95 % CL exclusion limits in multiples of the Higgs boson cross-section as expected in the Standard Model as function of m_H . The columns (a) to (e) refer to the use of various uncertainties, as listed in the text. The final limit is taken to be (e). The Hbb paper limits are taken from [1], include additionally the muon channel and have different systematic uncertainties, see text.*

shape uncertainty of the b-tagging variables, since only a defined cut on the MV1 distribution is applied, for which the uncertainties are known [41].

Nevertheless, the analysis presented here provides limits on the Higgs boson cross-section that are similar to those of the cut-based reference analysis [1], which uses about twice the number of events. This allows to conclude, that the BDT-based approach is able to deliver results with higher sensitivity than the cut-based one. This has to be validated after carefully reconsidering all sources of systematic uncertainties for the BDT-based analysis.

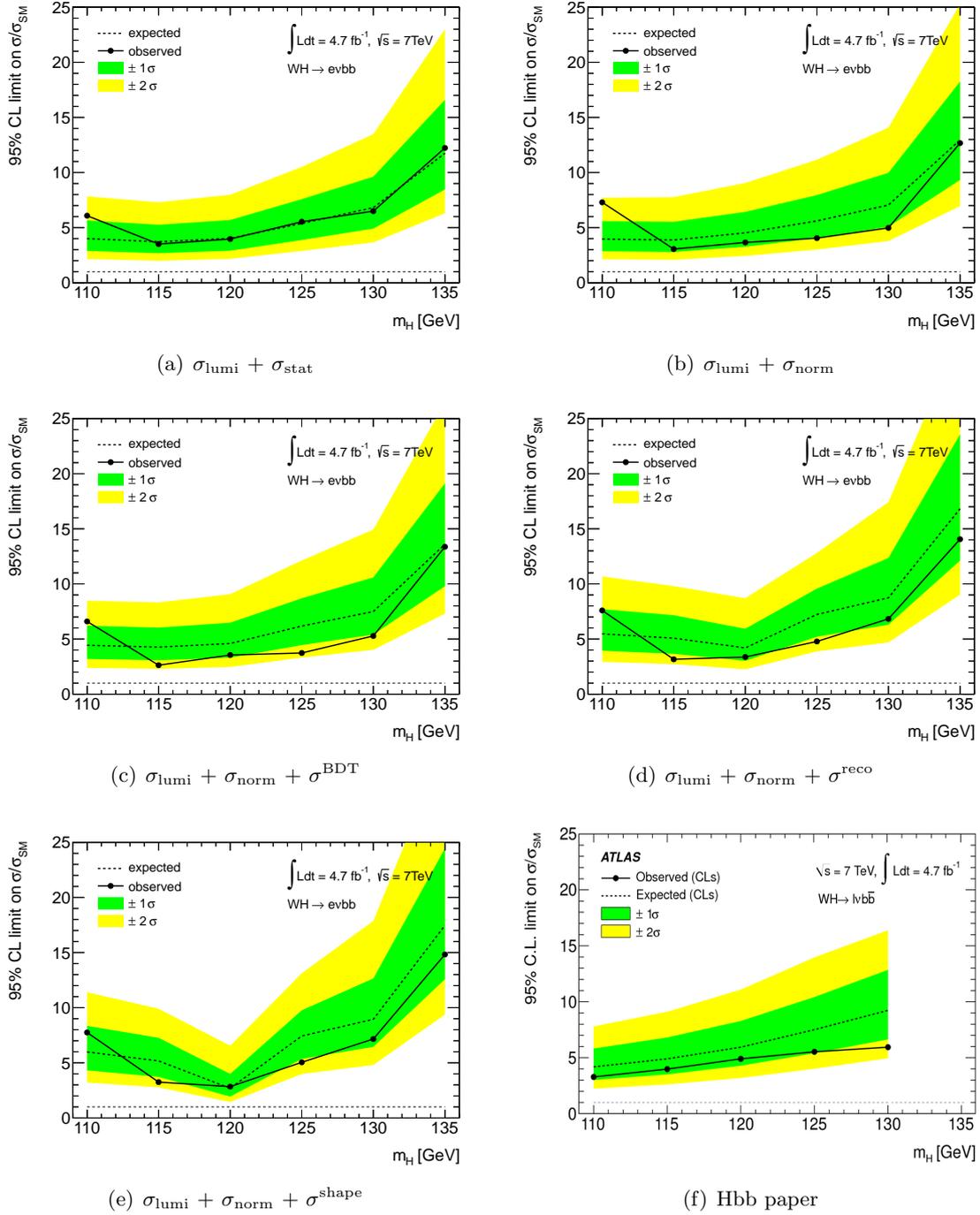


Figure 9.5: Expected and observed 95 % CL exclusion limits in multiples of the Higgs boson cross-section as expected in the Standard Model as function of m_H . Figs. (a) to (e) refer to the use of various uncertainties, as listed in the text. The final limit is shown in (e). The Hbb paper limits (f) are taken from [1], include additionally the muon channel and have different systematic uncertainties, see text.

In this thesis the search for a Higgs boson in the $WH \rightarrow e\nu b\bar{b}$ channel was performed using the data collected by the ATLAS experiment in 2011 at $\sqrt{s} = 7$ TeV. As a preliminary result the sensitivity of the cut-based reference analysis [1], published by the ATLAS collaboration, was exceeded by using Boosted Decision Trees (BDTs) for the extraction of the signal. This BDT-based analysis represents the first multivariate analysis of the $WH \rightarrow e\nu b\bar{b}$ channel in the ATLAS collaboration.

First, a basic event selection was performed. The selection followed the procedure in the cut-based analysis except for releasing some kinematic cuts. Instead, the corresponding variables were used as input variables for the BDTs. The BDTs are expected to optimise the selection based on the input variables. Additionally, by removing cuts more events were preserved for the training of the BDTs.

The most important reducible background, the $t\bar{t}$ background, was investigated in detail. The largest fraction of the $t\bar{t}$ background stems from fully leptonic decays, $t\bar{t} \rightarrow WbW\bar{b} \rightarrow l\nu l\nu b\bar{b}$. They constitute about 66% of the $t\bar{t}$ background. For large transverse momenta of the reconstructed leptonically decaying W boson, $p_T^W > 200$ GeV, the fraction of the semi-leptonic decays, $t\bar{t} \rightarrow WbW\bar{b} \rightarrow qql\nu b\bar{b}$, was found to be dominant. This is due to a large fraction of c -quarks that are misidentified as b -jets. An algorithm specifically designed to optimise the rejection of c -jets, JetFitterCombNNc, was used as input variable for the BDTs.

Other variables were evaluated regarding the separation of the signal from the leptonic and the semi-leptonic $t\bar{t}$ backgrounds. Finally, a set of 13 variables that provide fair separation were selected as input variables for the BDTs. The separation of individual variables was seen to improve when taking correlations into account. Therefore, a multivariate approach is expected to provide a higher sensitivity than a cut-based analysis.

Boosted Decision Trees were trained to extract the signal. For each boson mass in the range from 110 to 135 GeV in 5 GeV steps, a dedicated signal BDT was trained. In addition, background BDTs were trained to separate single backgrounds from all other backgrounds.

The normalisations of the dominant backgrounds, $W + b\bar{b}$, $t\bar{t}$ and the multijet background, were estimated from data. Their normalisation factors were fitted simultaneously in the output distributions of the background BDTs to data. The other backgrounds, single-top, diboson, $W + c$, $W + c\bar{c}$ and $W +$ light jets were normalised using Monte Carlo simulations.

The systematic uncertainties on the normalisations of the backgrounds and the signal were estimated by a standard approach, varying parameters of the reconstruction and identification of the final state objects. The resulting variations on the normalisations are taken as systematic uncertainties. Also the shape uncertainties of the BDT output distributions were estimated by this approach.

In addition, a novel technique was developed to estimate the shape uncertainties of the BDT outputs: by training the background BDTs the individual backgrounds are enriched in control regions within the nominal event selection. By reweighting the shape of the total background to data in these distributions the shape difference to data for each background is estimated. The systematic uncertainty derived by this technique is expected to cover uncertainties that cannot be estimated using the standard approach. This has to be validated in a follow-up study.

Finally, limits on the Standard Model Higgs boson cross section were calculated. This was done by fitting the backgrounds and the signal in the output distributions of the signal BDTs to data. In this fit the systematic uncertainties were treated as nuisance parameters. The limits were extracted using the CL_s method [69].

For the electron channel, $WH \rightarrow e\nu b\bar{b}$, an expected 95 % CL exclusion limit of 7.4 times the expected cross section for a Standard Model Higgs boson with $m_H = 125$ GeV was calculated. The corresponding result of the cut-based reference analysis is 7.5 including the muon channel [1].

Simplifications had to be done to fit a full analysis from event selection to limit calculation into the timescale of this thesis. Nonetheless, the significance of the cut-based analysis was exceeded and the potential of a BDT-based approach was shown. This preliminary result has to be validated by carefully re-checking the systematic uncertainties related to the BDT-based analysis.

This thesis sets the starting point for a more sophisticated multivariate analysis, including the muon channel and the associated production with a Z boson for the data collected by the ATLAS experiment in 2011 and 2012. Novel techniques were developed for the estimation of the background normalisations and the systematic uncertainties. These techniques are promising, but yet have to be validated in further studies.

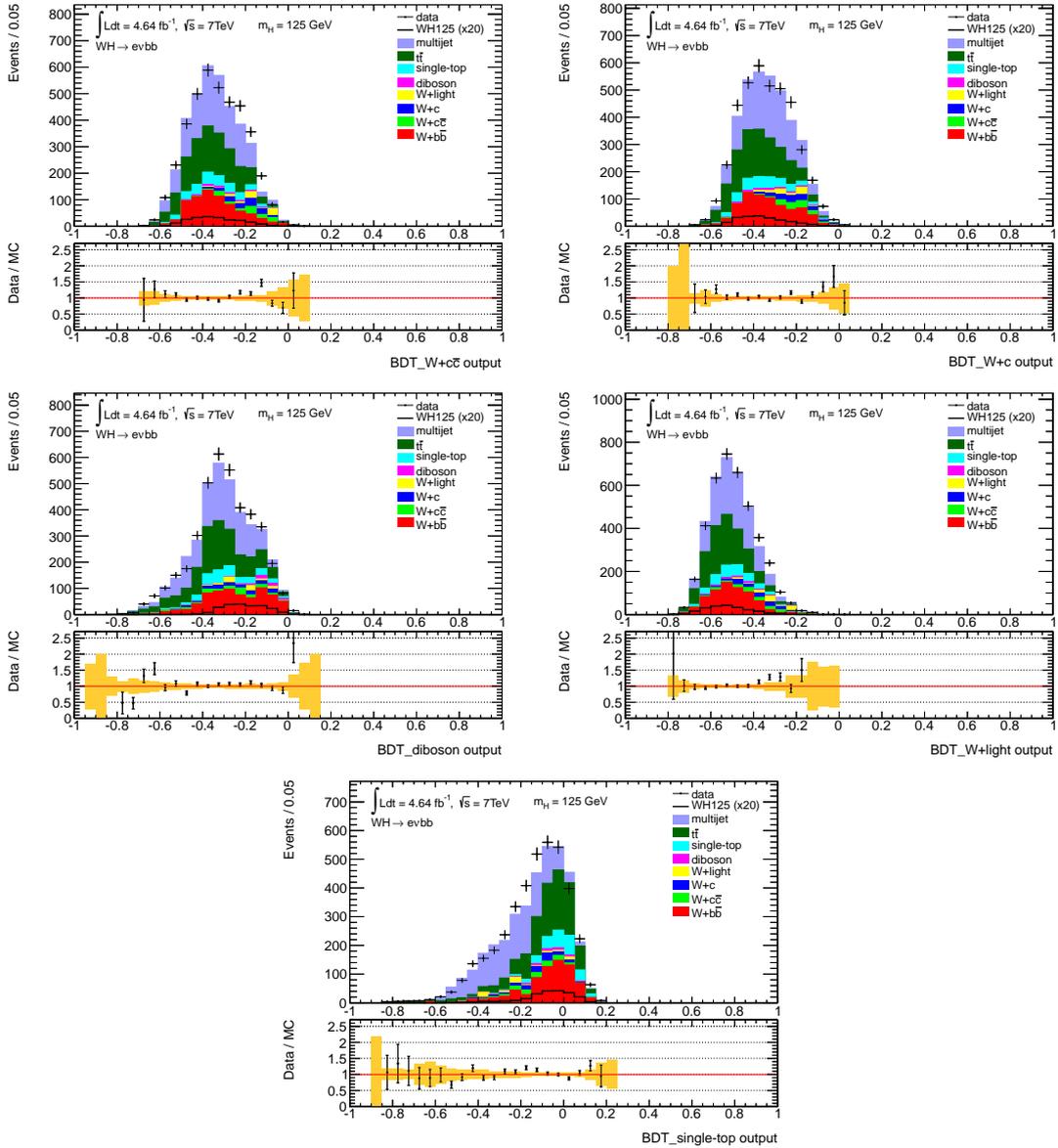


Figure A.1: Output distributions of the background BDTs that were trained with $W+c\bar{c}$, $W+c$, diboson, $W+\text{light}$ and single-top . The backgrounds are normalised using the fit described in Section 8.2. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

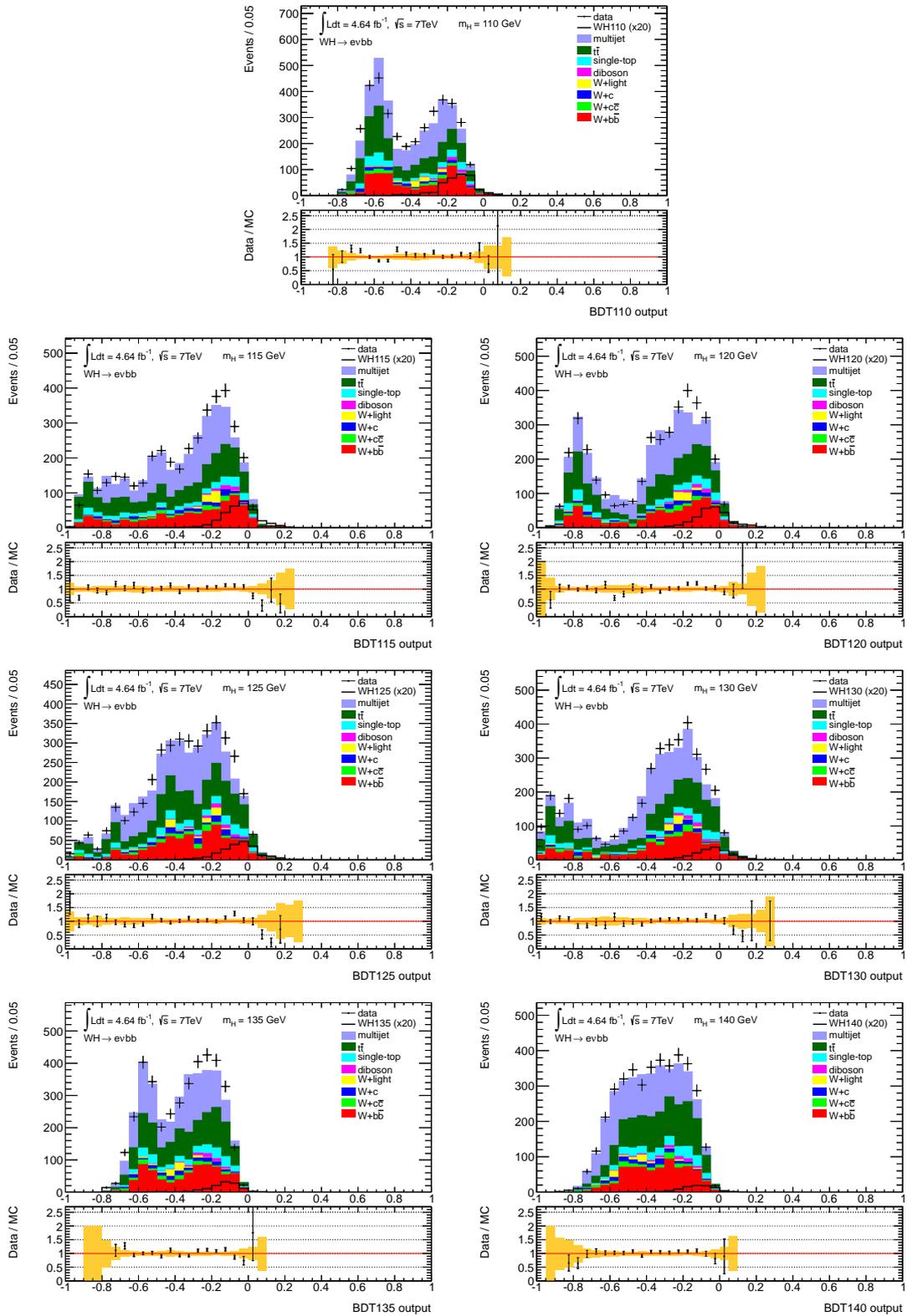


Figure A.2: Output distributions of the signal BDTs trained for the Higgs boson signal with various masses. The backgrounds are normalised using the fit described in Section 8.2. The signal distribution is scaled by a factor 20. The orange band in the ratio plot denotes the statistical uncertainty on the total background.

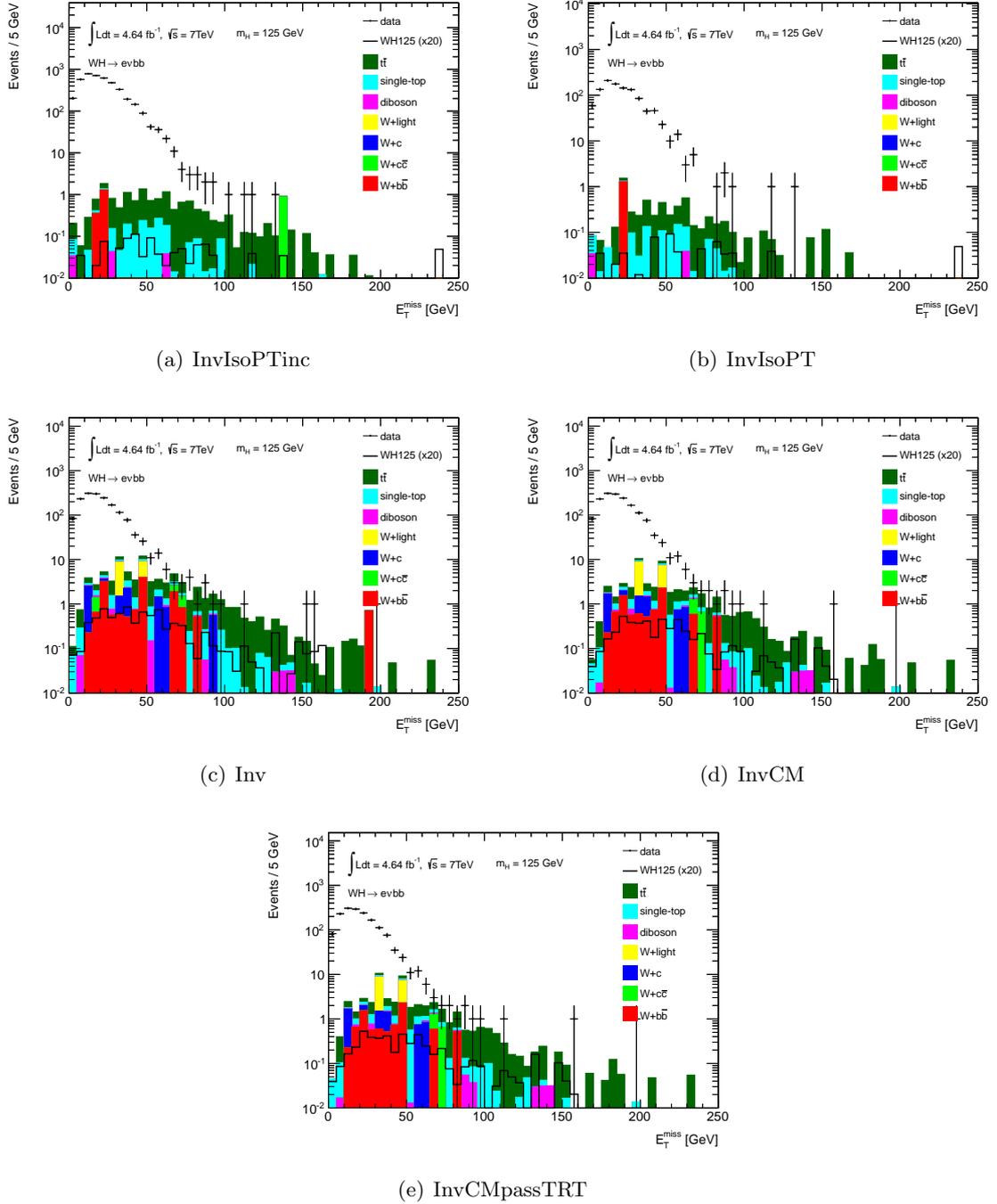
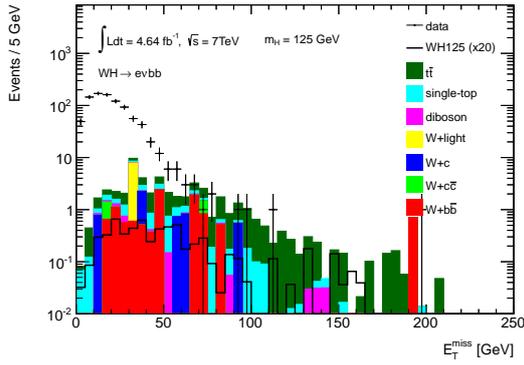
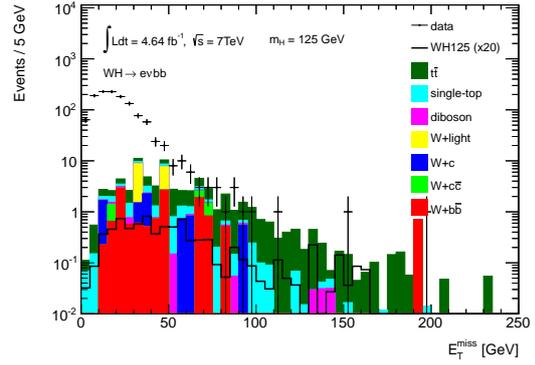


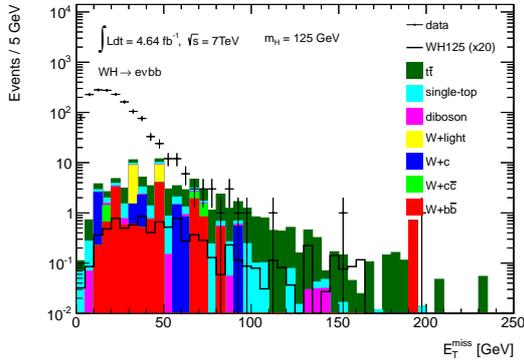
Figure A.3: Various multijet templates in E_T^{miss} . The templates are derived by varying the inverted isolation criterion and identification requirements of the electron. The backgrounds are normalised using Monte Carlo simulation. The signal distribution is scaled by a factor 20.



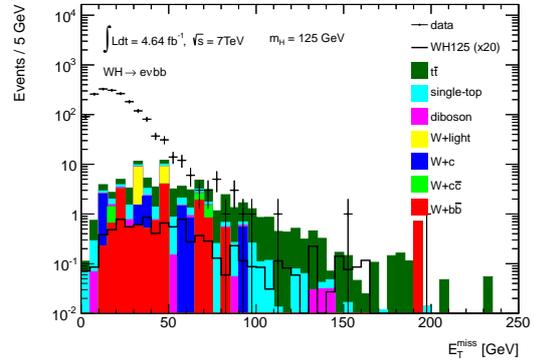
(a) InvIso1



(b) InvIso3



(c) InvIso5



(d) InvIso7

Figure A.4: Various multijet templates in E_T^{miss} . The templates are derived by varying the inverted isolation criterion and identification requirements of the electron. The backgrounds are normalised using Monte Carlo simulation. The signal distribution is scaled by a factor 20.

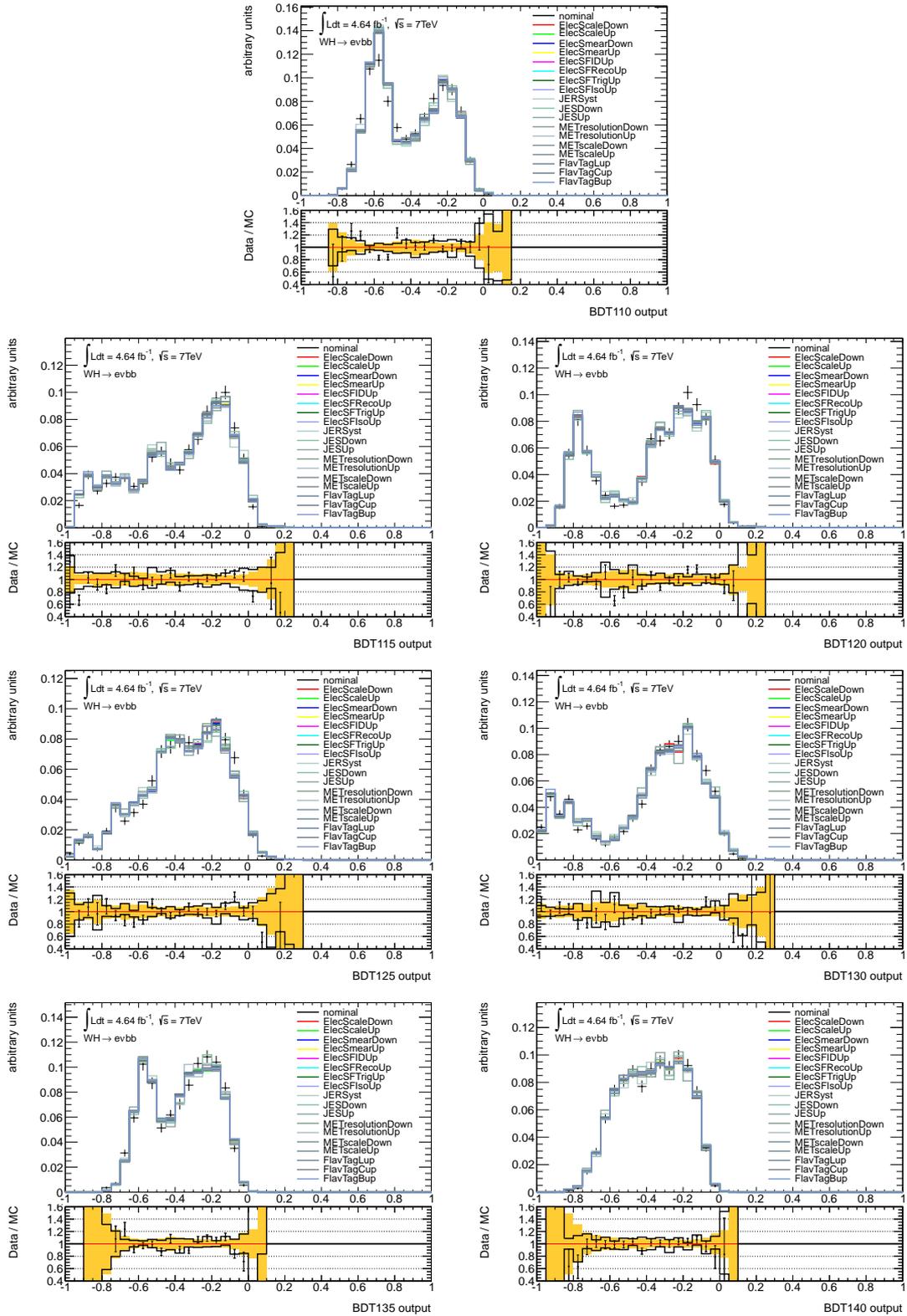


Figure A.5: Output distributions of the signal BDTs trained for the Higgs boson signal with various masses. The shape is varied by applying variations to the parameters from reconstruction and identification. The distributions are normalised to the same area. The background normalisations are taken from fit described in Section 8.2 and are not affected by the variations. The orange band in the ratio plot denotes the statistical uncertainty on the total background without variation. The black line in the ratio shows the sum of squares of the uncertainties, as described in the text.

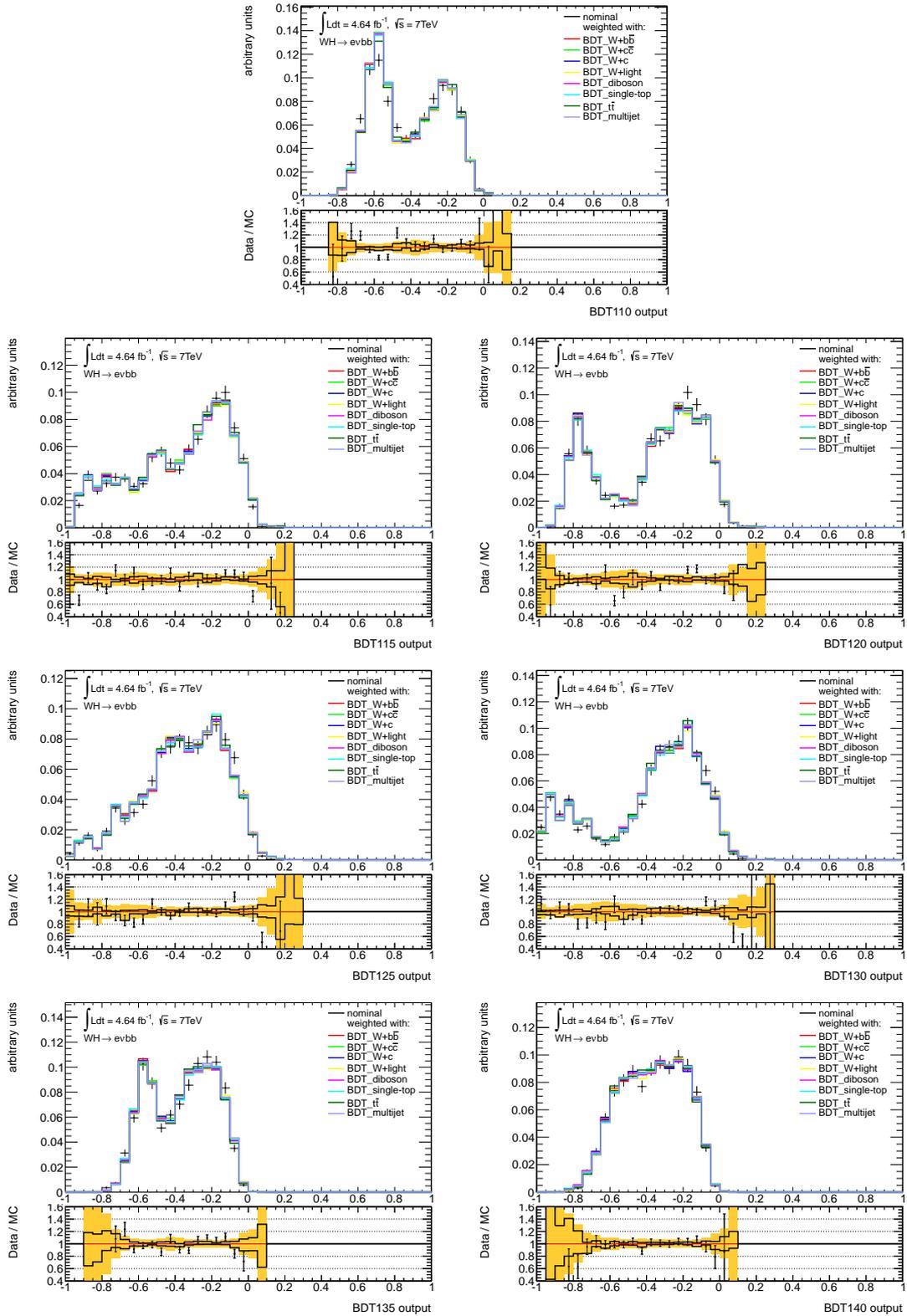


Figure A.6: Output distributions of the signal BDTs trained for the Higgs boson signal with various masses. The shape is varied by applying the BDT reweighting using the background BDTs. The distributions are normalised to the same area. The background normalisations are taken from fit described in Section 8.2 and are not affected by the variations. The orange band in the ratio plot denotes the statistical uncertainty on the total background without variation. The black line in the ratio shows the sum of squares of the uncertainties, as described in the text.

Bibliography

- [1] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a b-quark pair with the ATLAS detector at the LHC*, ATLAS-HIGG-2012-08, 2012 .
- [2] H. Gray and G. Piacquadio and R. Goncalo, *Search for the Standard Model Higgs boson in associated production with a vector boson and decaying to bottom quarks with the ATLAS detector*, ATL-COM-PHYS-2012-1081, 2012 .
- [3] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984.
- [4] D. Griffith, *Introduction to Elementary Particles*. Wiley-VCH, 2. ed., 2008.
- [5] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press, 1995.
- [6] Super-Kamiokande Collaboration, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys.Rev.Lett. **81** (1998) 1562, arXiv:hep-ex/9807003 [hep-ex].
- [7] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl.Phys. **22** (1961) 579.
- [8] A. Salam and J. Ward, *Electromagnetic and weak interactions*, Phys.Lett. **13** (1964) .
- [9] S. Weinberg, *A Model of Leptons*, Phys.Rev.Lett. **19** (1967) 1264.
- [10] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.
- [11] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys.Rev.Lett. **13** (1964) 508.
- [12] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys.Rev.Lett. **13** (1964) 321.
- [13] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys.Rev.Lett. **13** (1964) 585.
- [14] F. Tanedo, *Why do we expect a Higgs boson? (Part II)*, quantumdiaries.org.
- [15] ALEPH, CDF, D0, DELPHI, L3, OPAL and SLD Collaborations and LEP, Tevatron and SLD Electroweak Working Groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, arXiv:1012.2367 [hep-ex], CERN-PH-EP-2010-095.
- [16] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189, arXiv:0901.0002 [hep-ph].
- [17] Y. L. Dokshitzer, Sov.Phys. JETP **46** (1977) 641.

- [18] V. Gribov and L. Lipatov, *Sov.J.Nucl.Phys* **15** (1972) 438.
- [19] G. Altarelli and G. Parisi, *Nucl. Phys.* **B126** (1977) 298.
- [20] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept. Prog. Phys. **70** (2007) 89, [arXiv:hep-ph/0611148](https://arxiv.org/abs/hep-ph/0611148).
- [21] B. R. Webber, *A QCD model for jet fragmentation including soft gluon interference*, *Nucl. Phys. B* **238** (1984) 492.
- [22] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, *Phys.Rept.* **97** (1983) 31.
- [23] <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>.
- [24] T. Rias, http://en.wikipedia.org/wiki/Higgs_boson.
- [25] LHC Higgs Cross Section Working Group, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections>.
- [26] K. Tobe and J. D. Wells, *Higgs boson mass limits in perturbative unification theories*, [arXiv:0204.196](https://arxiv.org/abs/0204.196) [hep-ph].
- [27] The LEP Electroweak Working Group, <https://lepewwg.web.cern.ch/LEPEWWG/plots/winter2012/>.
- [28] *Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to 10.0 fb⁻¹ of Data*, [arXiv:1207.0449](https://arxiv.org/abs/1207.0449) [hep-ex].
- [29] ATLAS Collaboration, *Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC*, CERN-PH-EP-2012-218 .
- [30] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, CERN-PH-EP-2012-220 .
- [31] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.
- [32] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [33] B. Dolgoshein, *Transition radiation detectors*, *Nuclear Instruments and Methods in Physics Research* **A326** (1993) 434–469.
- [34] ATLAS Collaboration, G. Duckeck *et al.*, *ATLAS computing: Technical design report*, CERN-LHCC-2005-022.
- [35] ATLAS Collaboration, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATLAS note: ATL-SOFT-PUB-2007-007.
- [36] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, [arXiv:0901.0512](https://arxiv.org/abs/0901.0512) [hep-ex].
- [37] ATLAS Collaboration, *Performance of Jet Algorithms in the ATLAS Detector*, ATLAS-PHYS-INT-2010-129 (November, 2010) .

- [38] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-kt jet clustering algorithm*, arXiv:0802.1189 [hep-ph] .
- [39] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV taken in 2010*, ATL-COM-PHYS-2011-978, 2011 .
- [40] ATLAS Collaboration, *Commissioning of high performance b-tagging algorithms with the ATLAS detector*, ATLAS-CONF-2011-102, July, 2011 .
- [41] ATLAS Collaboration, *Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb⁻¹ of Data from the ATLAS Detector*, ATLAS-CONF-2012-043, March, 2012 .
- [42] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, PERF-2012-04, 2012 .
- [43] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, ATLAS note:ATLAS-PERF-2010-04-002.
- [44] ATLAS Collaboration, *ATLAS Muon Momentum Resolution in the First Pass Reconstruction of the 2010 p-p Collision Data at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2011-046 .
- [45] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s}=7$ TeV in ATLAS 2010 data*, ATLAS-CONF-2011-032, 2011 .
- [46] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph].
- [47] S. Frixione, P. Nason, and B. R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, JHEP **08** (2003) 007, arXiv:hep-ph/0305252.
- [48] M. L. Mangano, M. Moretti, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001.
- [49] J. Huston *et al.*, *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **07** (2002) 012, arXiv:0201195 [hep-ph].
- [50] *Les Houches Accord PDF Interface*, <http://hepforge.cedar.ac.uk/lhapdf/>.
- [51] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.
- [52] G. Corcella *et al.*, *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv:hep-ph/0011363.
- [53] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C72** (1996) 637, arXiv:hep-ph/9601371.
- [54] GEANT4 Collaboration, S. Agostinelli *et al.*, *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [55] *ATLAS Metadata Interface*, <https://ami.in2p3.fr>.

- [56] M. Aliev *et al.*, – *HATHOR – HAdronic Top and Heavy quarks crOss section calculator*, Comput. Phys. Commun. **182** (2011) 1034, arXiv:1007.1327 [hep-ph].
- [57] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) , arXiv:1001.5034 [hep-ph].
- [58] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) , arXiv:1103.2792 [hep-ph].
- [59] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-* , Phys. Rev. D **82** (2010) , arXiv:1005.4451 [hep-ph].
- [60] J. Butterworth *et al.*, *Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV*, ATLAS internal note: ATL-COM-PHYS-2010-695.
- [61] S. Ask, D. Malon, T. Pauly, and M. Shapiro, *Report from the Luminosity Task Force*, ATLAS note: ATL-GEN-PUB-2006-002, ATL-COM-GEN-2006-003, CERN-ATL-COM-GEN-2006-003, July.
- [62] P. Sommer, *Impact of z-vertex on efficiencies*, Talk at the ATLAS egamma meeting, 4. Apr 2012.
- [63] ATLAS Twiki, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/AtlasProductionGroupMC11c>.
- [64] <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BTaggingCalibrationDataInterface>.
- [65] ATLAS Twiki, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets2011#Simple_veto.
- [66] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.
- [67] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.
- [68] Update on the BDT analysis for WH to evbb, D. Buescher, ATLAS internal meeting <https://indico.cern.ch/conferenceDisplay.py?confId=167418>.
- [69] Alexander L. Read, *Presentation of search results: The $CL(s)$ technique*, J. Phys., G28, 2002, 2693-2704 .
- [70] ATLAS Collaboration, *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector in 2011*, 2011.
- [71] ATLAS Public Results, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2011JetResolution>.
- [72] Wouter Verkerke and David Kirkby, *The RooFit toolkit for data modeling*, arXiv:physics/0306116 [physics.data-an], 2003 .
- [73] A. S. Forum, *Frequentist Limit Recommendation from*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/StatisticsTools/>.

Erklärung der Selbstständigkeit

Hiermit versichere ich, dass ich die vorliegende Diplomarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Freiburg, den _____

Daniel Büscher

