Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität Freiburg

Search for the Standard Model Higgs boson in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay mode in proton-proton collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV with the ATLAS experiment

Evelyn Schmidt

CERN-THESIS-2013-077
20/02/2013

Dissertation

Search for the Standard Model Higgs boson in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay mode in proton-proton collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV with the ATLAS experiment

DISSERTATION

zur Erlangung des Doktorgrades der
Fakultät für Mathematik und Physik der

ALBERT-LUDWIGS-UNIVERSITÄT
Freiburg im Breisgau

vorgelegt von
Evelyn Schmidt

Dekan:	Prof. Dr. Michael Růžička
Referent:	Prof. Dr. Karl Jakobs
Koreferent:	JunProf. Dr. David Groß
Prüfer:	Prof. Dr. Gregor Herten
	JunProf. Dr. Harald Ita
	Prof. Dr. Karl Jakobs

Datum der mündlichen Prüfung:
20. Februar 2013

Contents

"Quand tu veux construire un bateau, ne commence pas par rassembler du bois, couper des planches et distribuer du travail, mais réveille au sein des hommes le désir de la mer grande et large."

Antoine de Saint-Exupéry

1

Introduction

Modern particle physics research is dedicated to study the fundamental constituents of matter and their interactions. To date, it is known that the elementary building blocks of matter exist in form of three *families* of *quarks* and *leptons*. Theories describing the dynamics of these elementary particles and their fundamental interactions have been developed during the past decades and are condensed in the *Standard Model* of particle physics. The Standard Model is a quantum field theory that describes three of the four known fundamental interactions between quarks and leptons, namely the strong, the electromagnetic and the weak interaction. A compatible and complete quantum mechanical description of gravity, the fourth fundamental interaction, is still absent. The behaviour of quarks and leptons under the afore mentioned forces differ significantly. While the strong force acts exclusively upon quarks and the electromagnetic force concerns only charged particles, the weak force affects all fundamental particles. The Standard Model has unified the weak and electromagnetic forces into a common one, the electroweak force. As a matter of fact, all the forces can be described on the basis of the same underlying principle of gauge symmetry. The interactions between fundamental particles are described by gauge fields and the exchange of corresponding gauge bosons. The Standard Model contains several such bosons: the photon, the W bosons (W^+ and W^-) and the Z boson, responsible for the electromagnetic and weak interactions, and the gluons, responsible for the strong interaction.

The Standard Model has been proven to describe all available experimental high energy physics data to high accuracy. However, this formalism has a weakness. As a consequence of the gauge symmetry only massless particles are allowed. Experimentally however, the W and Z bosons are found to be rather massive particles. The *Higgs mechanism* was thus proposed as the simplest solution to attribute mass to the W and Z bosons and also to all quarks and leptons while preserving gauge symmetry. In this theoretical model, the existence of an underlying Higgs field is postulated and the masses arise from the interaction of particles with this field. This mechanism implies the existence of one additional particle, the *Higgs boson* (H), which is the field quantum of the Higgs field. Therefore, the model can be verified by the experimental observation of the Higgs boson.

The search for the Higgs boson has been a major motivation for the construction of the Large Hadron Collider (`LHC`) at the particle physics research laboratory `CERN`[1] near Geneva, Switzerland. It is a circular particle accelerator providing proton-proton collisions at unprece-

[1] Conseil Européen pour la Rechérche Nucléaire

dented centre-of-mass energies of up to $\sqrt{s} = 8$ TeV in the year 2012. In those proton-proton collisions a variety of particles can be produced whose identity can be determined via the reconstruction and identification of their decay products by detectors located at the beam interaction points. Two multi-purpose experiments, ATLAS[2] and CMS[3] are installed at the LHC to allow for independent studies of known processes as well as searches for the Higgs boson and other phenomena beyond the Standard Model.

In the Standard Model, the mass of the Higgs boson is a free parameter. Once a mass value is specified, all physical properties of the Higgs boson, e.g. its production cross section and decay modes, are predicted by the theory. The Higgs boson mass can be constrained from both theoretical considerations and experimental measurements. These constraints prefer a Standard Model Higgs boson with a mass in the range of 115 GeV to 158 GeV [1, 2]. For a Higgs boson mass heavier than 135 GeV, the decay into a pair of W bosons, $H \to WW^*$, becomes the dominant decay channel. This decay mode is highly sensitive to a Standard Model Higgs boson in the mass range around the W boson pair production threshold of 160 GeV.

This thesis was carried out within the ATLAS collaboration and is dedicated to the search for the Standard Model Higgs boson in the decay channel $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$, with ℓ denoting electrons and muons. The analysis is based on the datasets corresponding to an integrated luminosity of 4.7 fb^{-1} collected at $\sqrt{s} = 7$ TeV in the year 2011 and 5.8 fb^{-1} collected at $\sqrt{s} = 8$ TeV in the year 2012 in proton-proton collisions by the ATLAS experiment. The Standard Model diboson W^+W^- production constitutes the largest background to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ signal. Thus, special emphasise was put on studying the agreement of Monte Carlo predictions to data in a signal-free control region enriched in dileptonic W^+W^- events. The analyses and cross-checks performed in the course of this thesis significantly contributed to and strengthened the published $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis [3–5].

Combining the results obtained by analysing the 2011 and 2012 datasets, an excess of events was observed in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channel which can be interpreted as the first evidence for a new particle. This observation complements the results obtained in searches for the Standard Model Higgs boson in other decay modes, in particular $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$, leading to the observation of a neutral boson with a mass of 126 ± 0.4 (stat) ± 0.4 (sys) GeV. The observation has a significance of 5.9 standard deviations [6].

Physics analyses like the ones mentioned above require a detailed understanding of known Standard Model processes which implies an excellent understanding of the detector performance and the reconstruction of basic physics objects like leptons or jets. In the course of this thesis an additional emphasis was placed on a solid understanding of the electron identification capabilities at the ATLAS experiment. The baseline electron identification relies on characteristics of the electromagnetic showers reconstructed in the ATLAS calorimetry system. A set of electron shower shape distributions was extracted using $Z \to e^+e^-$ decays on the dataset taken in the year 2010 corresponding to an integrated luminosity of 35 pb^{-1} collected in proton-proton collisions at $\sqrt{s} = 7$ TeV by the ATLAS experiment. This study served as an important basis for the improvement of the calorimeter simulation and contributed significantly to the optimisation of the electron identification.

This thesis is organised as follows: Chapter 2 provides an overview of the Standard Model of particle physics, the Higgs mechanism and the relevant phenomenological aspects at a hadron collider. The experimental setup, the LHC and the ATLAS detector, is described in Chapter 3. The reconstruction of physics objects that are relevant for the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ anal-

[2] A Toroidal LHC ApparatuS
[3] Compact Muon Solenoid

ysis, namely electrons, muons, jets and missing transverse energy is described in Chapter 4 including details of the corresponding performances and efficiencies.

A dedicated study of electron shower shape variables is presented in Chapter 5. The results obtained are published in Ref. [7]. The search for the Higgs boson in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay mode is discussed in detail in Chapter 6 including the combination with results obtained by investigating other decay modes.

A study dedicated to the modelling of the diboson background is presented in Chapter 7. The Standard Model W^+W^- process is the largest background to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ signal. The present background estimation consists of identifying a signal-free control region enriched in dileptonic W^+W^- events. While this helps to control the overall normalisation of this background in the signal region, it is not suited for more detailed inspections of important kinematic distributions since the control region is not sufficiently pure. Therefore, a new method has been developed within this thesis to address this important question. Based on the kinematic similarity between the dileptonic $W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ and trileptonic $W^\pm Z \to \ell\nu\ell\ell$ decays, this method allows for a comparison of the kinematic distributions between data and Monte Carlo diboson predictions. This also opens the possibility to further improve the understanding of the W^+W^- background. The analysis presented is based on data corresponding to an integrated luminosity of 10.7 fb^{-1} collected at $\sqrt{s} = 8$ TeV in the year 2012 in proton-proton collisions by the ATLAS experiment.

2

Theoretical Overview

This chapter provides a brief introduction to the Standard Model of particle physics and the mathematical description of particles and their interactions. The necessity of the Higgs mechanism is motivated and the predicted production and decay modes of the corresponding Higgs boson are discussed as they are relevant for the analyses presented later in this thesis. The current constraints on the Higgs boson mass are briefly summarised. Furthermore, the phenomenological aspects of high energetic proton-proton collisions are outlined and the predicted cross sections of known Standard Model processes are presented.

Thorough introductions into quantum field theories and the Standard Model of particle physics are given in various text books and lecture notes. The following section is based on the detailed description presented in Refs. [8–11] unless stated otherwise.

It should be noted that throughout this thesis *natural units* are used. This implies in particular that the reduced Planck constant $\hbar$ and the speed of light c are chosen as $\hbar = c = 1$. As a consequence, the dimensions of basic quantities can be written in terms of energy

$$[\text{energy}] = [\text{mass}] = [\text{momentum}]$$

and energies, momenta and masses are given in *electron-volt* (eV) which is the kinetic energy gained by an electron when accelerated by an electric potential of one Volt.

2.1. The Standard Model of Particle Physics

The Standard Model of particle physics is a gauge quantum field theory which describes the elementary particles and their interactions. Developed in the second half of the last century, this elegant theoretical framework incorporating the Glashow-Weinberg-Salam theory of electroweak processes and quantum chromodynamics has come to be called the Standard Model (SM). It provides a unified description of the electromagnetic, weak and strong interactions, based on the symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. The fundamental interactions are derived from the principle of *local gauge invariance* with respect to this symmetry group. Particle masses are introduced by an idea nowadays known as the "Englert-Brout-Higgs" mechanism or simply "Higgs mechanism". A spontaneous symmetry breaking of the electroweak group into the electromagnetic subgroup while leaving the strong interaction part $SU(3)_C$ invariant

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \rightarrow SU(3)_C \otimes U(1)_{QED}$$

generates the masses of the weak gauge bosons and also of the fermions. Furthermore, it gives rise to the appearance of a physical scalar particle which is called *Higgs boson*.

The Standard Model is an exceptionally successful phenomenological model which is able to describe the known experimental facts in particle physics with high precision [12]. It should be noted that gravity eludes its integration into a quantum field theoretical framework so far. However, it is much too weak to play a significant role at present particle collider energies.

2.1.1. Particle Content

The fermionic matter content of the Standard Model is given by the known leptons and quarks. These particles are considered as elementary particles in the sense that within the present experimental resolution they do not appear to possess any further substructure. Each particle is characterised by a unique set of quantum numbers. The particles are organised in a three-fold family structure of increasing mass shown in Table 2.1.

Table 2.1.: Overview of known leptons and quarks in the Standard Model. The quarks and leptons are spin 1/2 fermions. The masses are taken from Ref. [13]. Due to the fact that quarks cannot be observed as free particles, the given masses are model dependent and have large uncertainties. Each particle has a corresponding anti-particle with the same mass but opposite additive quantum numbers (see Section 2.1.5).

Generation	Symbol	Name	Mass	Charge [e]
		Leptons		
1^{st}	ν_e	Electron neutrino	< 2 eV	0
	e	Electron	0.511 MeV	-1
2^{nd}	ν_μ	Muon neutrino	< 0.19 MeV	0
	μ	Muon	105.7 MeV	-1
3^{rd}	ν_τ	Tau neutrino	< 18.2 MeV	0
	τ	Tau	1.777 GeV	-1
		Quarks		
1^{st}	u	Up	1.7 to 3.1 MeV	$+2/3$
	d	Down	4.1 to 5.7 MeV	$-1/3$
2^{nd}	s	Strange	80 to 130 MeV	$+2/3$
	c	Charm	1.18 to 1.34 GeV	$-1/3$
3^{rd}	b	Bottom	4.1 to 4.4 GeV	$+2/3$
	t	Top	172.9 GeV	$-1/3$

There are six types of leptons, six types of quarks and their antiparticles with the same mass but opposite additive quantum numbers. All fermions take part in the weak interaction while only quarks are also subject to the strong interaction since they carry so-called *colour charge*. Quarks and electrically charged fermions participate in electromagnetic interactions.

The fermions are grouped into three families with increasing mass. All known stable matter is made up of the fermions from the first generation, namely up and down quarks (in the form

of protons and neutrons) and electrons. The electrically charged fermions of the second and third generation are unstable and decay into lighter particles. It should be noted that leptons are observed as free particles while quarks are *confined* within hadrons containing either a quark and an anti-quark (mesons) or three quarks (baryons) due to the strong interaction. Hence, the given quark masses in Table 2.1 are model dependent and have large uncertainties.

The Standard Model treats the neutrinos as massless particles, albeit recent experiments have shown that neutrinos have small but non-zero masses [14]. Since the masses are tiny the approximation of vanishing neutrino masses still leads to very precise predictions of the Standard Model.

The strong, weak and electromagnetic interactions are mediated by bosons carrying spin 1. The known bosons of the Standard Model are listed in Table 2.2. The strong interaction is mediated by eight massless gluons that couple to the colour charge of quarks and gluons. The mediator of the electromagnetic interaction is the massless photon while the mediators of the weak interaction are three massive particles, namely the $W^{\pm}$ and Z bosons.

All fermions and gauge bosons discussed so far have been observed experimentally in high energy physics experiments over the last decades [13]. In case the Higgs mechanism turns out to be valid, the last missing piece in the Standard Model is a massive, electrically neutral Higgs boson which is the phenomenological footprint of the Higgs mechanism.

Table 2.2.: Overview of known bosons of the Standard Model that mediate the strong, electromagnetic and weak interactions. Their electrical charges and masses are taken from Ref. [13].

Symbol	Name	Interaction	Mass	Charge [e]	Coupling to
g	gluon	strong	0 (theoretical)	0	quarks and gluons
γ	photon	electromagnetic	0 (theoretical)	$<5\times10^{-30}$	charged particles
$W^{\pm}$	W boson	weak	(80.399±0.023) GeV	±1	all fermions, W, Z, γ, Higgs
Z^0	Z boson	weak	(91.1876±0.0021) GeV	0	all fermions, W, Higgs

2.1.2. Symmetries and Gauge Invariance

It has been known from classical physics that the invariance of a system under certain transformations is related to the conservation of corresponding quantities. For example, if a system is invariant under translations in space, its momentum is conserved. This connection between symmetries and conservation laws is known as Noether's theorem. The same result extends to quantum physics where symmetries are described by groups whose linear representations in the Hilbert space of quantum states appear in form of matrices.

In quantum field theory, particles are described by fields that are linear operators acting on the quantum mechanical Hilbert space of the particle states. The dynamics of the fields (particles) as well as their interactions are determined by a Lorentz-invariant, renormalisable Lagrangian $\mathscr{L}$. The equations of motion can be determined by minimising the action $S = \int d^4x \mathscr{L}$. The structure of the Lagrangian can be derived using symmetry arguments under which the equations of motion are invariant and which are consistent with experimental data. The interaction Lagrangian of the Standard Model may be written as the sum of two

parts:

$$\mathscr{L}_{SM} = \mathscr{L}_{QCD} + \mathscr{L}_{EW}$$

The term $\mathscr{L}_{QCD}$ representing the gauge theory of quantum chromodynamics (QCD) can be determined by its local gauge invariance with respect to a $SU(3)$ symmetry group. The electroweak part is defined by the invariance with respect to a $SU(2) \times U(1)$ symmetry group. The method to determine the form of the Lagrangian is briefly explained using the example of pure electromagnetic interactions. The electroweak model and quantum chromodynamics are discussed in the following sections.

2.1.3. Electromagnetic Interactions

The Lagrangian describing the electromagnetic interaction, denoted as $\mathscr{L}_{QED}$, can be derived starting from the Lagrangian describing a free Dirac fermion with mass m:

$$\mathscr{L}_{Dirac} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi, \tag{2.1}$$

where ψ is the spinor field representing fermions, $\bar{\psi} = \psi^{\dagger}\gamma^{0}$ is its adjoint and γ^{μ} denotes the *Dirac γ-matrices*. The corresponding equation of motion is the so-called *Dirac equation*

$$\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi = 0.$$

The Lagrangian $\mathscr{L}_{Dirac}$ is invariant under a *global* $U(1)$ gauge transformation:

$$\psi(x) \rightarrow \psi'(x) = e^{iQ\theta}\psi(x)$$

where $Q\theta$ is an arbitrary real constant. However, the free Lagrangian is no longer invariant under local phase transformations where the phase is allowed to depend on the space-time coordinate such that $\theta = \theta(x)$. The necessary and sufficient condition for the system to be also invariant under so-called local $U(1)$ gauge transformation is to introduce a new vector field $A_{\mu}(x)$, transforming as

$$A_{\mu}(x) \rightarrow A'_{\mu}(x) = A_{\mu}(x) - \frac{1}{e}\partial_{\mu}\theta \tag{2.2}$$

and to define the covariant derivative as

$$\partial_{\mu} \rightarrow D_{\mu} = \partial_{\mu} + ieQA_{\mu}(x)$$

which has the required property of transforming like the field itself. The Lagrangian

$$\begin{aligned}\mathscr{L} &= i\bar{\psi}(x)\gamma^{\mu}D_{\mu}\psi(x) - m\bar{\psi}(x)\psi(x) \\ &= \mathscr{L}_{Dirac} - eQA_{\mu}(x)\bar{\psi}(x)\gamma^{\mu}\psi(x)\end{aligned}$$

is then invariant under local phase transformations. Thus, requiring local gauge invariance has imposed an interaction between the gauge field A_{μ} and the spinor fields describing the fermions, and moreover, imposed the topological structure of the interaction term itself. The corresponding electromagnetic charge Q remains, however, completely unconstrained by the gauge symmetry. In order to describe space-time motion of the gauge field, identified as the physical photon field, an extra term, itself again gauge invariant, must be added to the full Lagrangian as follows:

$$\mathscr{L}_{kin} = -\frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x)$$

involving the gauge invariant field strength tensor defined as $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Thus imposing the requirement of local phase invariance on the free fermion Lagrangian led to the Lagrangian of quantum electrodynamics (QED):

$$\mathscr{L}_{QED} = \underbrace{\bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x)}_{\text{Kinetic energy and mass of } \psi(x)} \underbrace{- eQA_\mu(x)\bar{\psi}(x)\gamma^\mu\psi(x)}_{\text{Interaction}} \underbrace{- \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x)}_{\text{Kinetic energy of } A_\mu} \tag{2.3}$$

The addition of a mass term $\frac{1}{2}m^2 A_\mu A^\mu$ is prohibited since it would violate gauge invariance. Hence the gauge particle, the photon, is predicted to be massless.

In summary, demanding local gauge invariance led to the correct interacting field theory of QED. Thus, it has become one of the most basic and essential ingredients.

2.1.4. Strong Interactions

The strong interaction between quarks is described by quantum chromodynamics (QCD) [15–17]. Experimental observations of mesonic and baryonic states in combination with requirements according to Fermi-Dirac statistics have led to the assumption of the existence of a new quantum number called *colour*. Under this hypothesis, quarks possess an additional quantum number called colour. Each quark has three colour values referred to as *red*, *green* and *blue*. The anti-quarks possess anti-colours ($anti-red$, $anti-green$ and $anti-blue$). The underlying local-gauge group for QCD is a $SU(3)$ group. All observed particles are colourless and thus the observed particle states transform like colour singlets under $SU(3)$.

Quantum chromodynamics is based on the generalisation of the same aforementioned idea with the $U(1)$ gauge group being replaced by the $SU(3)$ group acting on the colour charged quarks. The Lagrangian is required to be invariant under local SU(3) transformations of the form:

$$q(x) \to Uq(x) = e^{i\alpha_a(x)\lambda_a/2}q(x)$$

where q denote one of the three colour fields and U is an arbitrary 3×3 unitary matrix. The fundamental representation of the SU(3) matrices is provided by the *generators* $\lambda_a/2$ (a=1,2,...,8) known as the Gell-Mann matrices. The corresponding covariant derivative:

$$D_\mu = \partial_\mu - ig_s\frac{\lambda^a}{2}G^a_\mu$$

involves now eight different gauge fields G^a_μ, called *gluon fields*, and g_s denote the strong coupling constant. The SU(3) matrices, being non-commutative, give rise to an additional term when transforming the gluon fields which involves the gluon fields themselves:

$$G^a_{\mu\nu} \to (G^a_{\mu\nu})' = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc}G^b_\mu G^c_\nu$$

where f^{abc} denotes the *structure constants* of $SU(3)$ that satisfy the commutation relations $[\frac{\lambda^a}{2}, \frac{\lambda^b}{2}] = i\ f^{abc}\lambda^c/2$. This last term does not have an analogue in QED and shows a remarkable new property that induces self-interaction between the gauge bosons. It reflects the fact that gluons themselves carry colour charge. The non-Abelian nature of the $SU(3)$ gauge group leads to another extremely important property namely that strong interactions between quarks and gluons become weaker at short distances (or higher energies), known as *asymptotic freedom*, and stronger at large distances (low energies) which is called *confinement*. The final Lagrangian of QCD is then given by:

$$\mathscr{L}_{\text{QCD}} = \sum_q i\bar{q}\gamma^\mu\partial_\mu q - g_s\sum_q \bar{q}\gamma^\mu\frac{\lambda^a}{2}G^a_\mu q - \frac{1}{4}G^a_{\mu\nu}G^{a,\mu\nu} \tag{2.4}$$

where the sums run over all quark flavors q. Just as for photons, local gauge invariance requires the gluons to be massless.

2.1.5. Electroweak Unification

The unified description of the electromagnetic and weak interactions was developed in the 1960s, primarily by Sheldon Glashow, Abdus Salam and Steven Weinberg, and recognised with the Nobel Prize in physics in 1979 [18–20]. It is based on the gauge symmetry group $SU(2)_L \otimes U(1)_Y$ where L refers to left-handed fields and Y is the weak hypercharge. All fermions in the Standard Model are subject to the electroweak interaction. The fermions are spinor fields ψ and there are three generations of left- and right-handed chiral fermions $\psi_{L,R}$ which are given by the projections $\psi_{L,R} = \frac{1}{2}(1 \mp \gamma^5)\psi$. Experimental observations substantiate that weak interactions distinguish between left-handed and right-handed fermions. Under the action of the $SU(2)_L$ group, the left-handed fermions transform as doublets of two leptons or quark weak eigenstates of the same generation. The right-handed fermions transform as singlets and are not affected by actions of the $SU(2)_L$ group. All fermions carry weak hypercharge Y.

The weak isospin and hypercharge fulfil the Gell-Mann-Nishima relation, which links them to the physically observable electric charge via $Q = I_3 + Y/2$. An overview of all Standard Model fermions and their electroweak quantum numbers, namely the charge Q, the weak isospin T, its third component T_3 and the hypercharge Y, are given in Table 2.3. The quarks with a prime symbol are eigenstates of the weak interaction, which are related to the mass eigenstates by a unitary rotation in flavour space that is described by the CKM matrix [21].

The left-handed isospin doublets ψ_L and the right-handed isospin singlet ψ_R transform under the action of the $SU(2)_L \otimes U(1)_Y$ direct product group as follows:

$$\begin{aligned} \psi_L \to \psi'_L &= e^{i\alpha^a(x)\cdot T_a + i\beta(x)Y}\psi_L, \;\; a = 1,2,3 \\ \psi_R \to \psi'_R &= e^{i\beta(x)Y}\psi_R \end{aligned}$$

where $\alpha^a(x)$ and $\beta(x)$ are local phases and $T_a/2$ and Y are the generators of the $SU(2)_L$ and $U(1)_Y$ groups of gauge transformations, respectively. Since there are now four gauge parameters, $\alpha^a(x)$ and $\beta(x)$, there are as well four different gauge fields needed and the covariant derivative is defined as:

$$D_\mu = \partial_\mu + i\frac{g}{2}W^a_\mu T_a + i\frac{g'}{2}B_\mu Y, \tag{2.5}$$

where W^a_μ $(a = 1,2,3)$ and B_μ denote the gauge fields related to the $3+1$ degrees of freedom of the $SU(2)_L \otimes U(1)_Y$ symmetry group. The coupling constants g and g' determine the strength of the coupling to the $SU(2)_L$ gauge fields and the $U(1)_Y$ gauge field, respectively. The corresponding field strength tensors of the gauge fields are given by:

$$\begin{aligned} W^{\mu\nu}_a &= \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\varepsilon_{abc}W^b_\mu W^c_\nu, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \end{aligned}$$

where ε_{abc} denotes the totally antisymmetric tensor. The properly normalised kinetic Lagrangian that is invariant under local $SU(2)_L \otimes U(1)_Y$ transformations is found to be:

$$\mathscr{L}_{\mathrm{EW}} = \sum_j i\bar{\psi}^j_L\gamma^\mu D_\mu\psi^j_L + \sum_k i\bar{\psi}^k_R\gamma^\mu D_\mu\psi^k_R - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \tag{2.6}$$

where the sum in i and k runs over all doublets and singlets listed in Table 2.3. Since the field strengths $W^a_{\mu\nu}$ contain a quadratic piece, the Lagrangian gives rise to cubic and quartic

self-interactions among the gauge fields.

Experiments show that the gauge fields $W_a^{\mu\nu}$ and B_μ do not carry the right quantum numbers to be directly identified with the experimentally observed $W^\pm$ and Z bosons and the photon. Instead, a linear combination of these gauge fields leads to the physically observable states according to:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right), \tag{2.7}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu, \tag{2.8}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \tag{2.9}$$

where $W_\mu^\pm$ and Z_μ denote the fields of the weak gauge bosons, A_μ the photon field and θ_W the weak mixing angle. The combinations above allow to relate the electric charge e and the electroweak couplings by:

$$e = g\sin(\theta_W) = g'\cos(\theta_W). \tag{2.10}$$

The Lagrangian formulated in terms of the transformed fields shows that the charged-current interactions of the W bosons to quarks and leptons are universal to all fermions because of the gauge symmetry. The W bosons decay into fermion anti-fermion pairs of different flavours that still belongs to the same $SU(2)_L$ doublet, e.g $W^+ \to e^+\nu_e$. The coupling of quarks and leptons to W bosons have the form:

$$\frac{-ig}{\sqrt{2}}\gamma^\mu(1-\gamma^5). \tag{2.11}$$

The factor $(1-\gamma^5)$ is crucial since γ^μ alone would yield a vector coupling whereas $\gamma^\mu\gamma^5$ gives an axial vector. A theory that adds a vector to an axial vector is bound to violate the conservation of parity, and this happens in weak interactions. The violation here is maximal in the sense that the two terms are equally large.

The neutral-current interactions of the Z boson exhibit couplings to fermion anti-fermion pairs of the same flavour, e.g. $Z \to e^+e^-$. Here, the coupling of the Z boson has the more complex form of:

$$\frac{-ig}{\cos(\theta_w)}\gamma^\mu\frac{1}{2}(c_V^f - c_A^f\gamma^5). \tag{2.12}$$

While the photon has the same interaction with left- and right-handed fermions, the Z boson couples differently depending on the fermion charge and weak isospin. This is expressed by the coefficients c_V^f and c_A^f that depend on the particular quantum numbers of the quark or lepton (f) involved in the following way:

$$c_V^f = I_3^f - 2q_f\sin^2(\theta_w) \text{ and } c_A^f = I_3^f. \tag{2.13}$$

2.2. The Higgs Mechanism

It should be noted that the Lagrangian of the Standard Model $\mathscr{L}_{SM} = \mathscr{L}_{QCD} + \mathscr{L}_{EW}$ described so far does not involve any mass terms. Explicit gauge boson mass terms such as $W_\mu^a W^{\mu a}$ violate gauge invariance. Furthermore, left- and right-handed fields transform differently under $SU(2)_L$ and $U(1)_Y$ gauge transformations. Therefore, adding by hand fermion mass terms is also not possible as these terms would couple the left- and right-handed fields which behave as doublets and singlets under the gauge symmetry. Nevertheless,

Table 2.3.: Overview of the quantum numbers of the Standard Model fermions in the Glashow-Weinberg-Salam model. The indices L and R refer to left-handed and right-handed helicity states. For quarks, the dashed symbols indicate eigenstates of the electroweak interaction which are orthogonal superpositions of the mass eigenstates, where the mixing is described by the CKM-matrix. Right-handed neutrinos are decoupled from interactions of the SM and are therefore not considered (adapted from Ref. [8]).

	Generation			Quantum numbers			
	1st	2nd	3rd	I	I_3	Y	$Q[e]$
Leptons	$\begin{pmatrix}\nu_e\\ e^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau^-\end{pmatrix}_L$	1/2	1/2	-1	0
				1/2	$-1/2$	-1	-1
	e^-_R	μ^-_R	τ^-_R	0	0	-2	-1
Quarks	$\begin{pmatrix}u\\ d'\end{pmatrix}_L$	$\begin{pmatrix}c\\ s'\end{pmatrix}_L$	$\begin{pmatrix}t\\ b'\end{pmatrix}_L$	1/2	1/2	1/3	2/3
				1/2	$-1/2$	1/3	$-1/3$
	u_R	c_R	t_R	0	0	4/3	2/3
	d_R	s_R	b_R	0	0	$-2/3$	1/3

experimental observations prove that the fermions and weak gauge bosons do have mass.

In order to generate masses, the Lagrangian has to be extended without violating the fundamental principle of gauge invariance. This can be achieved via the *Higgs mechanism* which exploits the principle of spontaneous symmetry breaking to generate the desired mass terms [22–25]. The basic idea is that while the Lagrangian will always remain gauge invariant, the ground state of the system is not. The symmetry breaking of the ground state is driven by the existence of the scalar field called the *Higgs field*.

2.2.1. Spontaneous Symmetry Breaking

The Higgs field Φ, a weak isospin doublet of complex scalar fields with hypercharge $Y = 1$, is postulated:

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

which contains four real scalar fields ϕ_i. It is subject to an external potential:

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda\left(\Phi^\dagger\Phi\right)^2. \tag{2.14}$$

In combination with the covariant derivative as given in Eq. 2.5, this leads to the Lagrangian

$$\mathscr{L}_H = (D_\mu\Phi)^\dagger(D_\mu\Phi) - V(\Phi) \tag{2.15}$$

$$= |(i\partial_\mu - gT\cdot W_\mu - g'\frac{Y}{2}B_\mu)\Phi|^2 - \mu^2\Phi^\dagger\Phi + \lambda\left(\Phi^\dagger\Phi\right)^2 \tag{2.16}$$

where $|\ |^2 = (\ )^\dagger(\ )$. $\mathscr{L}_H$ is invariant under local $SU(2)_L \otimes U(1)_Y$ transformations. The first term describes the coupling of the electroweak gauge fields and the Higgs field. By choosing $\lambda > 0$, which is required by vacuum stability, and $\mu^2 < 0$, a potential of the form illustrated in Fig. 2.1 is obtained. There is an infinite set of degenerate states with minimum energy, satisfying:

$$\Phi^\dagger\Phi = -\frac{\mu^2}{2\lambda}.$$

By choosing a particular ground state, the symmetry gets spontaneously broken. This is because the vacuum state has a lower symmetry than the potential $V(\Phi)$ itself. An appropriate choice of a particular minimum is:

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

Obviously, this ground state is no longer gauge invariant but the Lagrangian itself still obeys the full symmetry. The $SU(2)_L \otimes U(1)_Y$ symmetry gets spontaneously broken to the electromagnetic subgroup $U(1)_{QED}$, which by construction still remains a true symmetry of the vacuum. The excitations over the ground state can be parametrised in the general form:

$$\Phi(x) = \frac{e^{iT_a\theta^a(x)}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.17}$$

with four real fields denoted as $\theta^a(x)$ and $H(x)$. The important point is that the local $SU(2)_L$ invariance of the Lagrangian allows to rotate away any dependence on $\theta^a(x)$. These three fields are the would-be massless Goldstone bosons associated with the spontaneous symmetry breaking. Taking the physical, unitary gauge $\theta^a(x) = 0$ and substituting the resulting parametrisation of $\Phi(x)$ along with the covariant derivative as defined in Eq. 2.5 in the Lagrangian of Eq. 2.15, it takes the form:

$$\mathscr{L}_H = \frac{1}{2}\partial_\mu H \partial^\mu H + (v+H)^2\left(\frac{g^2}{4}W_\mu^\dagger W^\mu + \frac{g^2}{8\cos^2\theta_w}Z_\mu Z^\mu\right) - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4}H^4. \tag{2.18}$$

In this representation, the fields are expressed in the physical weak boson fields $W_\mu^\pm$ and Z_μ. Due to the non-zero value of v, the second term of the Lagrangian contains bilinear terms in the weak boson fields. These terms can be interpreted as mass terms for the electroweak gauge bosons as follows:

$$m_W = m_Z \cdot \cos\theta_W = \frac{vg}{2} \quad \text{and} \quad m_\gamma = 0. \tag{2.19}$$

In summary, by introducing the Higgs field along with a potential that induces spontaneous symmetry breaking, a gauge invariant Lagrangian is obtained that contains mass terms for the electroweak bosons. The photon is still massless, in accordance with experimental observations, because $U(1)_{QED}$ remains an unbroken symmetry by the particular choice of the ground state. Furthermore, a new scalar boson is introduced into the model with mass $m_H = \sqrt{(2\lambda v^2)}$ which interacts with the massive electroweak bosons. The last two terms of $\mathscr{L}_H$ can be interpreted as self-interaction terms via trilinear and quartic couplings.

The masses of the $W^\pm$ and Z bosons have been measured experimentally (see Table 2.2). The relation for m_W/m_Z as predicted in Eq.2.19 has been confirmed [13]. The vacuum expectation value has been estimated to be $v = 246$ GeV by comparison to the coupling constant of the Fermi model G_F. The weak interactions are treated as point-like interactions in this model and the coupling constant can be obtained from muon decay experiments.

2.2.2. Fermion Masses

An attractive feature of the newly introduced Higgs field is that it can generate masses for the $W^\pm$ and Z bosons as well as for leptons and quarks. The Lagrangian has to be extended once more by adding the so-called Yukawa-terms of the following form:

$$\mathscr{L}_{Yukawa} = -G_l^{ij}\bar{L}_L^i \Phi l_R^j - G_d^{ij}\bar{Q}_L^i \Phi d_R^j - G_u^{ij}\bar{Q}_L^i \Phi_C u_R^j + \text{hermitian conjugate}, \tag{2.20}$$

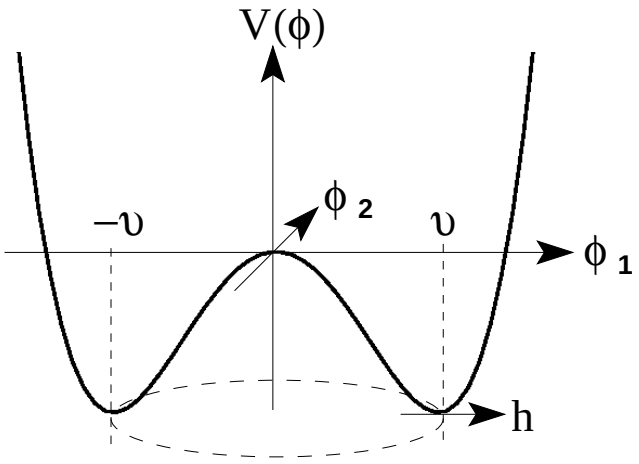


Figure 2.1.: Two-dimensional illustration of the Higgs potential $V(\Phi)$ as defined in Eq. 2.14 with parameters $\lambda > 0$ and $\mu^2 < 0$. There is an infinite set of degenerate states with minimum energy indicated by the dashed circle. The vacuum expectation value of the field Φ is $v = \frac{\mu^2}{2\lambda}$ (adapted from Ref. [8]).

where $\bar{L}_L^i$ and $\bar{Q}_L^i$ denote the lepton and quark isospin doublets, respectively, and l_R^j (d_R^j, u_R^j) are the corresponding lepton (up-type, down-type quark) isospin singlets. The second term involves the charge-conjugate Higgs doublet $\Phi_C = i\sigma_2\Phi^*$ with quantum numbers opposite to Φ. The matrices G_l^{ij}, G_d^{ij} and G_u^{ij} define the coupling constants and mixing between the generations, which affects the quarks, where the eigenstates of the weak interaction are a mixture of the mass eigenstates.

By spontaneously breaking the symmetry and substituting the expansion of Φ around its vacuum expectation value, the term in the Lagrangian relevant for e.g. electrons has the form:

$$\begin{aligned} \mathscr{L}_{electron} &= -\frac{G_e}{\sqrt{2}} v(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{G_e}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L)H \\ &= -m_e \bar{e}e - \frac{m_e}{v}\bar{e}eH, \end{aligned}$$

where $m_e = \frac{G_e v}{\sqrt{2}}$ is the electron mass. The coupling constant G_e is arbitrary and the actual mass of the electron is not predicted. Besides the mass term, the Lagrangian contains an interaction term coupling the Higgs scalar to the electron which is proportional to the mass of the electron. The other fermion masses are generated in the same way, except for neutrinos, which are treated as massless particles.

In summary, the final Lagrangian of the SM comprises the following terms:

$$\mathscr{L}_{SM} = \mathscr{L}_{QCD} + \mathscr{L}_{EW} + \mathscr{L}_H + \mathscr{L}_{Yukawa}$$

as defined in Eq. 2.4, Eq. 2.6, Eq. 2.18 and Eq. 2.20.

2.2.3. The Higgs Boson

The Higgs mechanism gives rise to the existence of an electrically neutral, CP-even boson with spin zero. The Higgs boson couplings to fermions have the form:

$$G_f \propto \frac{m_f\sqrt{2}}{v} \quad , \tag{2.21}$$

and the Higgs boson coupling to the gauge bosons have the form:

$$g_{HVV} \propto \frac{M_V^2}{v} \quad and \quad g_{HHVV} \propto \frac{M_V^2}{v^2}. \tag{2.22}$$

Hence, the couplings grow with the particle masses. These predictions can be tested as soon as the Higgs boson is experimentally observed. The mass of the Higgs boson itself is connected with the vacuum expectation value and the parameter λ by $m_H = \sqrt{(2\lambda v)}$. Since λ is unknown, the mass of the Higgs boson itself is not predicted by the model. Constraints on the Higgs boson mass determined from theoretical and phenomenological aspects are briefly discussed in the following section.

2.2.4. Constraints on the Higgs Boson Mass

The Higgs boson mass can be constrained by several theoretical considerations. A characteristic example is the scattering process $W^+W^- \to W^+W^-$ at high energies where the Higgs boson contributions guarantee a finite forward scattering amplitude. Thus, the requirement of unitarity leads to the upper bound $m_H \lesssim 870$ GeV [26].

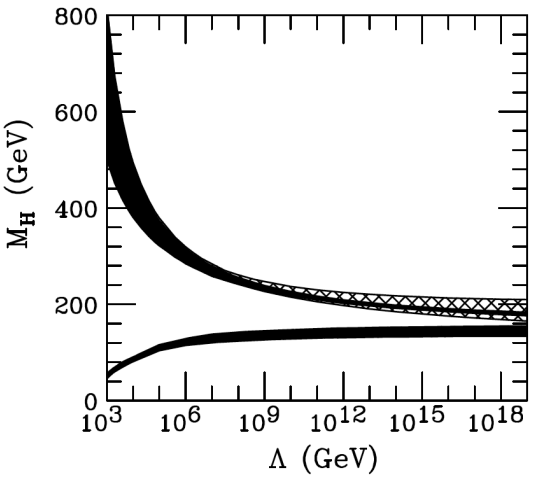
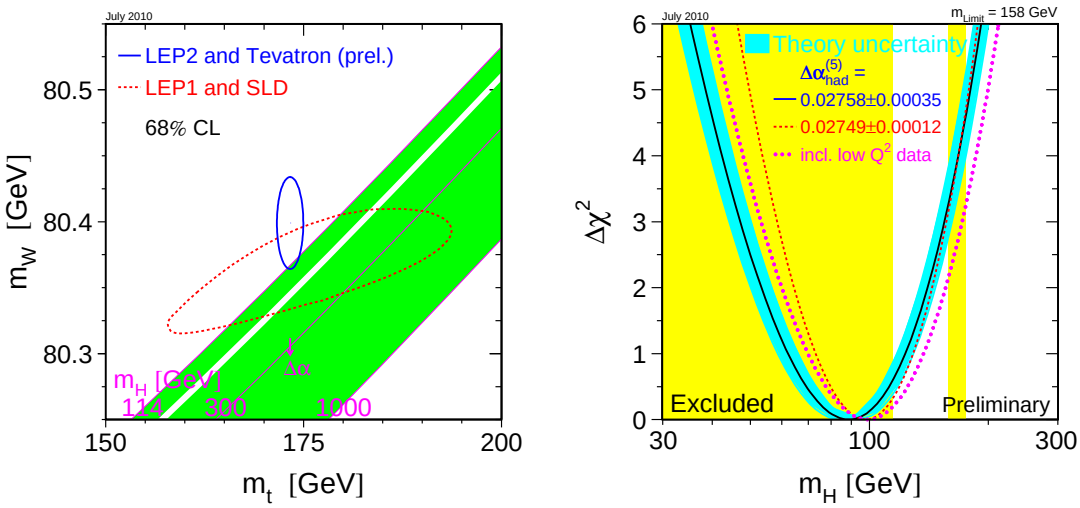


Figure 2.2.: Theoretical constraints on the Higgs boson mass [27]. The upper bound corresponds to Higgs boson masses for which the SM Higgs sector ceases to be meaningful as function of the validity scale Λ of the Standard Model for a top-quark mass of (175±6) GeV and the strong coupling $\alpha_s(m_Z) = (0.118 \pm 0.002)$. The lower bound indicates a value of m_H for which perturbation theory is certainly expected to be reliable at scale Λ. The allowed region lies between the bands and the hashed and solid bands illustrate the impact of various uncertainties.

Even tighter upper and lower bounds on the Higgs boson mass can be extracted from the quartic Higgs boson coupling which increases monotonically with the energy scale Q. In

fact, the coupling $\lambda(Q^2)$ varies logarithmically with the squared energy Q^2 and diverges at a certain energy, called Landau pole. To prevent this divergence, an upper limit on the Higgs boson mass is needed which depends on the energy cutoff Λ up to which the SM is valid. For instance for the value of $\Lambda \approx 10^{16}$ GeV a rather light Higgs boson with $m_H \lesssim 200$ GeV is needed. In the regime where λ is low, additional contributions of top-quark and massive gauge bosons have a sizeable effect in the running of the quartic coupling. In case the coupling λ gets too small, the top-quark contributions can be dominant and could cause negative values $\lambda(Q^2) < 0$. This would lead to a scalar potential $V(Q^2) < V(v)$ which would not be stable anymore since it has no minimum. Therefore, the stability argument leads to the requirement $\lambda(Q^2) > 0$ which implies also a lower limit on the Higgs boson mass which depends again on Λ. For relative low values of $\Lambda \approx 10^3$ GeV and relative high values of $\Lambda \approx 10^{16}$ GeV the lower bounds of $m_H \gtrsim 70$ GeV and $m_H \gtrsim 130$ GeV are obtained, respectively [26]. The scale dependence of the bounds on the Higgs boson mass imposed by these arguments are illustrated in Fig. 2.2.



(a) (b)

Figure 2.3.: Experimental constraints on the Higgs boson mass determined from electroweak precision measurements at (a) `LEP` and the (b) `Tevatron` [28]. The plot on the left-hand side shows the contour curves of 68% probability in the (m_t, m_W)-plane obtained from measurements at `LEP`, `Tevatron` and `SLD`. The shaded band shows the SM relationship for these masses as a function of m_H. The plot on the right-hand side shows the $\Delta\chi^2$ curve obtained by performing a global fit to precision electroweak measurements carried out at `LEP`, `SLC` and at the `Tevatron` experiments DØ and `CDF`. The solid line is the result of the fit using all data while the band represents an estimate of the theoretical uncertainty due to missing higher order corrections. The vertical band shows the 95% CL exclusion limit on $\mathrm{m_H}$ from the direct Higgs boson searches at `LEP` (up to 114.4 GeV) and `Tevatron` (158 GeV to 175 GeV) [2].

Experimental limits on the Higgs boson mass are deduced both from direct searches and indirect measurements. The Higgs boson mass is indirectly related to the mass of the top-quark (m_t) and the mass of the W boson (m_W) through higher order loop corrections. The three masses have a linear relationship to each other under the assumption that the Standard

Model is the correct theory of nature. The masses of the top-quark and of the W boson have been measured very precisely at LEP, Tevatron and SLD[1] tightening the range in possible values of the mass of the Higgs boson as shown in Fig. 2.3.

Indirect experimental constraints on the Higgs boson mass can be deduced from fits to the precision measurements of electroweak observables performed, most notably, at LEP, SLC and Tevatron. Taking the theoretical uncertainties into account, the 95% one-sided confidence level upper limit on m_H is 158 GeV. The vertical band shows the 95% CL exclusion limit on m_H from the direct Higgs boson searches at LEP (up to 114.4 GeV) and Tevatron (158 GeV to 175 GeV) [2].

In summary, both theoretical and experimental constraints prefer a relatively "light" Standard Model Higgs boson with a mass in the range of 115 GeV to 158 GeV.

2.3. Phenomenology of Proton-Proton Collisions

The QCD Lagrangian introduced in Section 2.1.4 describes the strong interaction between quarks and gluons. This alone is not yet sufficient to make experimentally testable predictions about proton-proton collisions. It has to be taken into account that protons are composite objects constituted of quarks and gluons. The following sections briefly summarise the most important techniques that aim for a realistic description of the phenomenology at hadron colliders. A detailed description can be found in Ref. [29].

2.4. General Aspects of Proton-Proton Collisions

Protons are not fundamental particles but are composed of more basic constituents called partons which are valence quarks, sea quarks and gluons. In proton-proton collisions, these constituents interact with each other. The scattering processes at high energy hadron colliders can be classified as either *hard* or *soft*. The latter appear with small momentum transfer between the interacting partons and constitute the majority of proton-proton interactions. On the contrary, hard scattering reactions of individual partons are characterised by a large momentum transfer and e.g. Higgs boson or W and Z boson production belong to these processes. The underlying theory for all such processes is QCD but the approach and level of understanding are very different for the two cases. For hard processes, the strong coupling strength is small and thus the scattering cross section can be reliably calculated using perturbation theory. However, for soft processes the strong coupling parameter becomes significantly larger and thus the scattering cannot be calculated pertubatively.

Figure 2.4 illustrates the interaction of two partons a and b that are constituents of the protons A and B. The interacting partons produce some final state X via the partonic process $ab \to X$. Due to confinement, the proton remnants hadronise into colour-neutral states. These additional hadrons are referred to as the *underlying event*.

The incoming partons carry colour charge, thus bremsstrahlung can occur. Emissions related to the incoming partons are called *initial state radiation* (ISR). Similarly, emissions associated to the outgoing objects are called *final state radiation* (FSR). Both effects can lead to additional jets being reconstructed in the detector which might significantly modify the event final state topology.

Since collisions are made to occur between bunches containing billions of protons, multiple proton-proton interactions can occur in one bunch-crossing. These additional interactions, which are mainly soft inelastic hadronic interactions, are called *minimum bias events*. They

[1] **S**tandford **L**inear Accelerator Center Large **D**etector

constitute the so-called *in-time pile-up* contributions. In case the response time of individual detector components is larger than the spacing between two bunches, also contributions from previous bunch crossings can overlay with the current bunch crossing. This contribution is called *out-of-time* pileup.

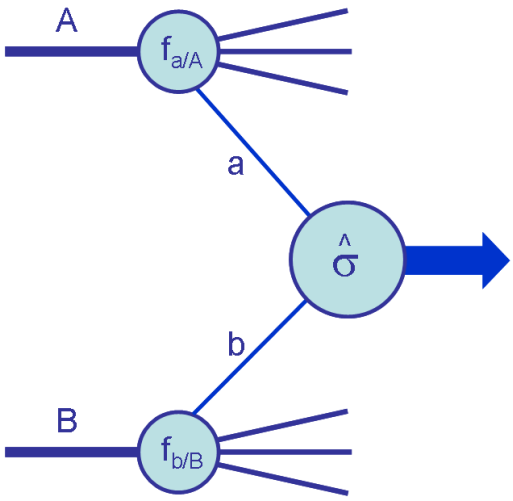


Figure 2.4.: Diagrammatic structure of a proton-proton collision, where the partons a and b interact and lead to the final state X [29].

2.4.1. Factorisation

The underlying principle of *factorisation* rests on the fact that one can separate dynamics in terms of different momentum scales. In view of proton-proton collisions, the soft processes including long-distance effects can be separated from the hard scattering reaction of individual partons at large momentum transfer Q. An observable that involves strong interactions can be parametrised by a product of at least two functions:

$$\text{Observable}(Q^2, p_{had}) = \mathscr{C}(Q^2, |p_{parton}| > \mu_F) \otimes \mathscr{P}(Q^2, |p_{parton}| < \mu_F, p_{had}) \tag{2.23}$$

where p_{had} and p_{parton} denote the hadron and parton momenta, respectively [30]. The function $\mathscr{C}$ describes the short distance part which can be computed in perturbation theory. All long distance physics is contained in the function $\mathscr{P}$ that cannot be described with perturbative QCD. By introducing the auxiliary *factorisation scale* μ_F, the dependence on the hard scale Q is factorised from the dependence on soft momenta. This approach is beneficial since it usually happens that the same soft functions $\mathscr{P}$ enter different physical processes. Thus, once extracted experimentally, it can be used for several processes. It should be noted that in case the observable is calculated to all orders of perturbation theory, it is invariant under changes to the factorisation scale. The observable of interest is the cross section of various processes occurring in proton-proton collisions as explained in the following sections.

2.4.2. Hadronic Cross Section and Parton Distribution Functions

The calculation of cross sections involving hadrons in the initial state has to take their compositeness into account. Applying the factorisation theorem, the hadronic cross section for a hard scattering process of hadrons A and B leading to the final state X can be expressed by [29, 31]:

$$\sigma_{AB} = \sum_{a,b} \int \mathrm{d}x_a \mathrm{d}x_b f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \hat{\sigma}_{ab\to X}, \tag{2.24}$$

corresponding to the structure depicted in Fig. 2.4. All partons a and b that can contribute to the final state X are considered in the summation. $\hat{\sigma}_{ab\to X}$ denotes the partonic cross section involving the initial state partons a and b. The parton distribution functions (PDFs) are given by $f_{a/A}(x_a, Q^2)$ ($f_{b/B}(x_b, Q^2)$) and describe the probability to find a parton a (b) carrying the momentum fraction x_a (x_b) of the hadron A (B). They are not predicted by perturbation theory and have to be determined from experimental data. Much of the experimental information originates from deep inelastic scattering data on structure functions, such as obtained by experiments at the electron-proton collider HERA at DESY, as well as of direct measurements of W boson and Z boson production at hadron colliders [29]. While the PDFs itself are not directly accessible in QCD perturbation theory, their evolution depending on the momentum transfer Q^2 is calculable. A system of coupled integro-differential equations, the so-called Altarelli-Parisi (DGLAP) equations, describe the evolution of the PDFs. Thus, given an input distribution determined at a certain momentum fraction x and momentum transfer Q^2, the DGLAP equations can be solved to predict the PDFs at a higher scale.

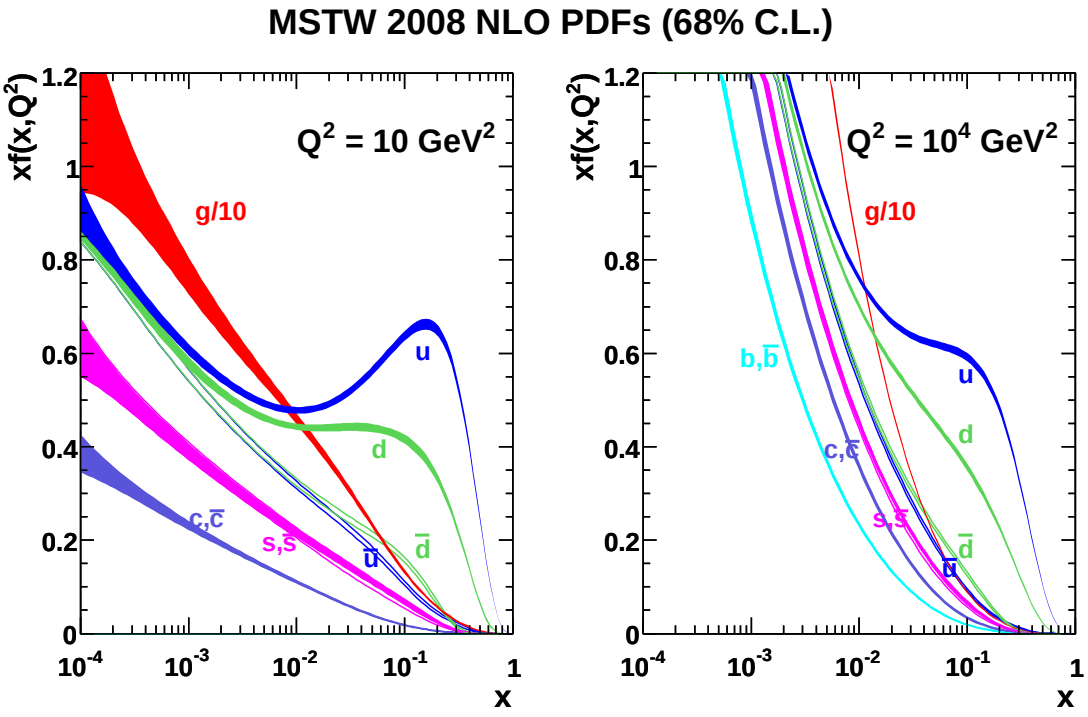


Figure 2.5.: Parton distribution functions of the proton as determined for the MSTW08 PDF set for momentum transfer scales of $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2, respectively (taken from Ref. [32]).

Several PDF sets are made public by different phenomenology working groups from their work on the available experimental data. The PDFs of the proton determined by the MSTW group [32] are shown in Fig. 2.5 for two different scales $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2, including the associated one-sigma (68%) confidence level (C.L.) uncertainty bands. It should be noted that the uncertainties on the distributions related to gluons are the largest. The reason is that the gluon distribution is the one least constrained by experimental data. A common technique to estimate PDF uncertainties is the Hessian method in which orthonormal eigenvectors are used as basis to propagate the uncertainties from the input data [29, 31]. A set of orthogonal error eigenvectors is obtained for each PDF and the uncertainties result from an excursion along the "+" and "-" directions for each eigenvector. Uncertainties on physical observables such as cross sections can be evaluated by the variations resulting from these one-sigma error eigenvector sets. It should be noted that an additional uncertainty in the determinations of the PDFs results from the uncertainty on the value of $\alpha_s(M_Z^2)$ used in the global fits.

2.4.3. Partonic Cross Section and Higher Order Corrections

Predictions for the partonic cross sections can be achieved with perturbative QCD calculations. The partonic cross section can be expanded in power series in the coupling constants as follows:

$$\hat{\sigma}_{ab\to X} = \hat{\sigma}_0(1 + c_1(\mu_F, \mu_R)\alpha_s(Q^2, \mu_R^2) + c_2(\mu_F, \mu_R)\alpha_s^2(Q^2, \mu_R^2) + ...), \tag{2.25}$$

where $\hat{\sigma}_0$ denotes the leading-order partonic cross section, α_s is the strong coupling constant depending on the momentum transfer Q^2 and the renormalisation and factorisation scales μ_F and μ_R, respectively. The renormalisation scale μ_R is the scale for the absorption of further divergences in the partonic cross sections in the running coupling constant. The coefficients c_i depend on these scales as well.

The simplest prediction is obtained with the lowest order in the expansion series, $\hat{\sigma}_0$. This is obtained by calculating the squared matrix element represented by tree-level Feynman diagrams for the involved particles and couplings and integrated over the appropriate phase space. The integration must often be carried out numerically and restrictions on the phase space have to be imposed in order to avoid divergences.

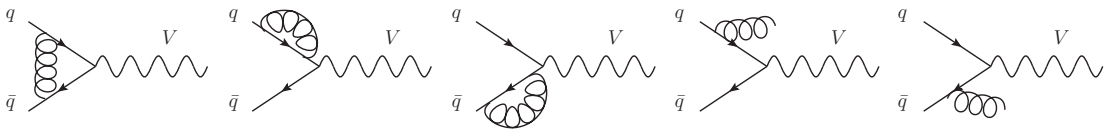


Figure 2.6.: Feynman diagrams with gluon corrections arising at next-to-leading order for the process $q\bar{q} \to V$ (with $V = W$ boson or Z boson). The three diagrams on the left-hand side represent virtual corrections while the two diagrams on the right-hand side represent real emissions.

The terms of higher orders in α_s in Eq. 2.25 are related to radiations of gluons. These so-called *next-to-leading order* contributions are illustrated for gluon corrections to the process $q\bar{q} \to V$ (with $V = W$ boson or Z boson) in Fig. 2.6. The results achieved by including higher order contributions are more precise, but the calculations are getting more complicated and singularities arise e.g. from virtual corrections. It was demonstrated that at each order in the perturbation series, the singularities from virtual corrections and the soft and collinear

singularities cancel such that the correction to the cross sections results in a finite value. One of the benefits of performing calculations to higher orders in perturbation theory is the reduction of the dependence of related predictions on the factorisation μ_F and renormalisation μ_R scales. The dependence vanishes in case the calculation is performed up to all orders in perturbation theory. In the absence of a complete set of higher order corrections, it is necessary to make a specific choice for the two scales in order to make predictions for cross sections. It is sensible to choose μ_F and μ_R values of the order of the typical momentum scales of the hard scattering process, e.g. for the process $q\bar{q} \to Z \to \ell^+\ell^-$ the standard choice is $\mu_F = \mu_R = M_{\ell\ell}$, the invariant mass of the lepton pair. The uncertainty on partonic cross sections is determined by variations of these scales.

The so-called *K-factor* for a given process is a useful short notation which represents the ratio between the next-to-leading order (NLO) and the leading order (LO) cross sections:

$$K_{NLO} = \frac{\sigma_{NLO}}{\sigma_{LO}}. \tag{2.26}$$

It should be noted that the K-factor may be different for various kinematic regions of the same process.

2.4.4. Luminosity

Events of a given type occur at a collider at a rate R which is related to the cross section σ for this event type and the instantaneous luminosity L delivered by the accelerator as follows:

$$R = \frac{\mathrm{d}N}{\mathrm{d}t} = \sigma \cdot L \tag{2.27}$$

L can be determined from the properties of the colliding bunches. Considering two such bunches containing n_1 and n_2 particles that collide head-on with a frequency f, the instantaneous luminosity is given by:

$$L = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \tag{2.28}$$

under the assumption of bunches with identical Gaussian profiles with standard deviations σ_x and σ_y in the horizontal and vertical directions with respect to the beam axis, respectively. The instantaneous luminosity is measured in units of $[\mathrm{cm}^{-2}\mathrm{s}^{-1}]$.

The delivered luminosity at a proton-proton collider can be monitored by measuring the observed interaction rate per bunch-crossing μ_{vis} which is related to the luminosity as follows:

$$L = \frac{\mu_{\mathrm{vis}} n_b f_{rev}}{\sigma_{\mathrm{vis}}} \tag{2.29}$$

where $\sigma_{\mathrm{vis}} = \varepsilon\sigma_{\mathrm{inel}}$ is the total inelastic cross section multiplied by the efficiency ε of a particular detector, n_b is the number of bunches per beam and f_{rev} is the frequency of complete turns around the ring. Detectors capable to perform such measurements are presented in Section 3.2.5.

The total number of events in a given period of data taking is obtained by integration of the rate R over time:

$$N = \sigma \cdot \int L \, \mathrm{d}t = \sigma \cdot \mathcal{L} \tag{2.30}$$

where the quantity $\mathcal{L}$ is called *integrated luminosity* which is used to quantify the amount of proton-proton collisions. Cross sections are usually measured in units of *barn*[2]. At the LHC, typical cross section of physics processes are in the range of nanobarn (pb) or femtobarn (fb).

[2] 1 barn = 10^-28 m^2=10^-24 cm^2

2.4.5. Expected Cross Sections at Hadron Colliders

The formalism described in the previous sections is used to obtain predictions for some important Standard Model cross sections at anti-proton-proton and proton-proton colliders. An overview of cross sections of some benchmark processes as a function of the centre-of-mass energy at the `Tevatron` and `LHC` colliders is shown in Fig. 2.8. It is evident that the total inelastic proton-proton cross section is six or seven orders of magnitude higher compared to the production cross section of W or Z bosons. The Higgs boson production cross section, which strongly depends on the Higgs boson mass, is predicted to be about ten to eleven orders of magnitude smaller compared to the total inelastic proton-proton cross section. Thus, very high luminosities are needed to produce a sufficient rate of these rare processes.

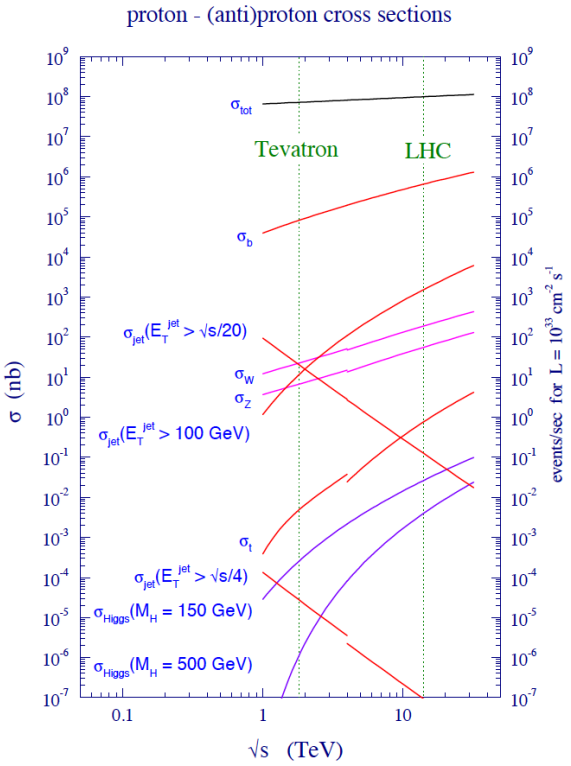


Figure 2.7.: Predicted Standard Model cross sections at the `Tevatron` and `LHC` colliders as function of the centre-of-mass energy $\sqrt{s}$. In case of the `LHC` the energy at $\sqrt{s} = 14$ TeV is marked. The collider operated at $\sqrt{s} = 7$ TeV during the data taking periods in 2010 and 2011 and at $\sqrt{s} = 8$ TeV during the data taking period in 2012 (taken from Ref. [29]).

2.4.6. Monte Carlo Event Generation

An accurate simulation of physics processes is indispensable in order to compare predictions of the Standard Model to events produced in proton-proton collisions. This is an essential requirement to gain a proper understanding of the processes involved in the analyses presented in this thesis in Chapter 6.

Due to the quantum mechanical nature of particle physics, it is only possible to predict the occurrence of a predefined final state with a given kinematical configuration produced in inelastic proton collisions on a probabilistic basis. The probability for such a process is proportional to the predicted cross section. The expected probability distributions emerge

only when averaging over large event samples. So-called *event generators* are used to simulate a possibly large number of events and the four-momenta of the final state particles are provided. In those generators, random numbers are involved to make choices intended to reproduce the probabilities for different outcomes at various stages of the process. Thus the event generators are referred to as *Monte Carlo* event generators.

The complexity of the event structure requires a split up of the simulation into several stages as illustrated in Fig 2.8. In terms of time, the simulation includes the hard interaction, followed by *parton showering* which incorporates higher order effects by allowing for branching of the quarks and gluons. The event now temporarily consists of a number of elementary particles, possibly including some with net non-zero colour charge. A special simulation program then regroups all these coloured partons into coulour-neutral composite hadrons as dictated by the colour confinement using a phenomenological model commonly referred to as *hadronisation*. The resulting particles can decay further. Finally, features of the underlying event and pile-up are added in. This steps are briefly discussed in the following sections and further details can be found in Ref. [33].

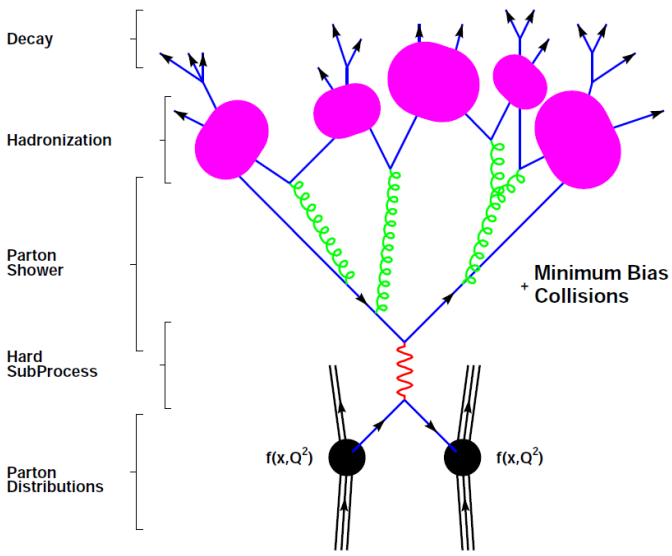


Figure 2.8.: Schematic illustration of the basic structure of a events simulated with a showering and hadronisation generator, such as PYTHIA [34] and HERWIG [35]. The time evolution of the event goes from bottom to top (taken from Ref. [33]).

In a second step, the four-momenta of the final state particles can be passed as inputs to a program simulating the response of the detector. This step will be presented in Chapter 4 after introducing the ATLAS detector in Chapter 3.

To date, there is a variety of simulation programs available and the various steps used in the simulation are common to most of the programs but differ in the exact implementation. Among the Monte Carlo programs used in this thesis are MC@NLO [36], POWHEG [37, 38], PYTHIA [34] and HERWIG [35] (see Section 6.4). The predictions provided by PYTHIA and HERWIG are based on leading-order matrix elements while those for MC@NLO and POWHEG involve next-to-leading order QCD corrections. Although the present Monte Carlo generators are capable to simulate the momentum and angular distributions of the resulting objects to a high accuracy, it should always be kept in mind that generators are not perfect and provide only an approximation of the real process taking place in nature.

Hard Scattering

The starting point of event generation is the hard-scattering process and its cross section is calculated according to Eq. 2.24. The generated events are selected according to the probability of the particular kinematic four-momentum and spin configurations of the final state particles to be produced. Most of the available Monte Carlo generators allow only the computation of tree-level matrix elements with a fixed number of partons in the final state. These parton-level generators describe a specific final state to lowest order in perturbation theory.

Parton Showers

Higher order effects in perturbation theory are added by evolving the event using the *parton shower* method. It is an iterative process involving basically the `DGLAP` splitting functions in so-called *Sudakov form factors* as probability distributions for the parton branching such as $q \to qg$,$g \to q\bar{q}$ and $g \to gg$. For the simulation of FSR, the evolution of branchings is performed starting from the energy scale of the hard interaction. Algorithms going backwards in time allow to include ISR processes with a similar approach ensuring that the energy distribution of the incoming partons is compatible with the measured PDFs. This is achieved by weighing the Sudakov form factors with the PDFs at the corresponding scale. The iterative method is justified for soft and collinear emissions of partons but has its limitations for the emission of hard and wide-angle partons. The final outcome of successive branchings is a parton shower in which each initial parton from the hard process is replaced by a number of partons moving in roughly the same direction. The typical scale of relative transverse momenta at the end of the shower is set by a cutoff which is usually chosen around 1 GeV. Thus, parton shower algorithms are run to model parton emissions down to the non-pertubative soft regime of hadronisation.

Parton shower algorithm are implemented in generators such as `PYTHIA` or `HERWIG`. Many other Monte Carlo generators do not have their own parton shower modelling but their output can be interfaced to the aforementioned programs for this purpose.

Hadronisation

The methods described so far involve particles at the partonic level while only colour-neutral hadrons are observed on the experimental side. In between exists an important phase called *hadronisation* where all the outgoing partons end up confined inside hadrons. This processes is not described from first principles since perturbation theory is not applicable at long distances. Thus, the description of this phase has to involve some approximate modelling. Several of such models are available.

One example is the Lund string model [39] as implemented in `PYTHIA`, a widely used framework to model the hadronisation process. It treats quark-anti-quark pairs as connected by colour strings and gluons are regarded as energy and momentum carrying kinks on the string. The colour string acquires more and more potential energy as the quarks move apart. Once the maximum energy is reached the string breaks apart by the production of new $q\bar{q}$ pairs. A quark from one break can combine with an antiquark form an adjacent one to form a colour singlet meson.

This approach assumes that the hadronisation process is mostly independent of the perturbative processes since the hadronisation scale is much smaller than the hard scales. Therefore,

the parameters of the models can be tuned to data. The simpler initial state at an electron-positron collider, such as LEP implies that this is the logical place to tune the parameters of the hadronisation framework to data [40]. Afterwards, those parametrisation functions can be applied in predictions of processes at other experiments, like those at the LHC.

Underlying Event and Multiple Interactions

As already mentioned in Section 2.4, in events containing a hard parton-parton interaction, the underlying event represents the additional activity not directly associated with that interaction. In the present discussion, the underlying event is defined to represent the additional activity after hard bremsstrahlung of the hard interaction has already been taken into account.

Furthermore, more than one parton per proton may interact in proton-proton collisions. Thus, so-called *multiple interactions* can occur. Both processes usually appear at low transverse momentum scale and can therefore not be treated perturbatively and must be described by phenomenological models based on experimental results.

Routines to simulate the additional activity are implemented in PYTHIA and JIMMY [41]. It is possible to tune the parameters of the underlying event model to data. For this purpose properties of the underlying event have been measured by the ATLAS experiment in proton-proton collisions at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV [42, 43]. For example, the measurements of charged particle distributions provide sensitivity to the parameters describing the multi-parton interactions and colour reconnection of the hadronic final state. These dedicated measurements have led the ATLAS experiment to adopt sets of tuned parameters to be used in physics event generation.

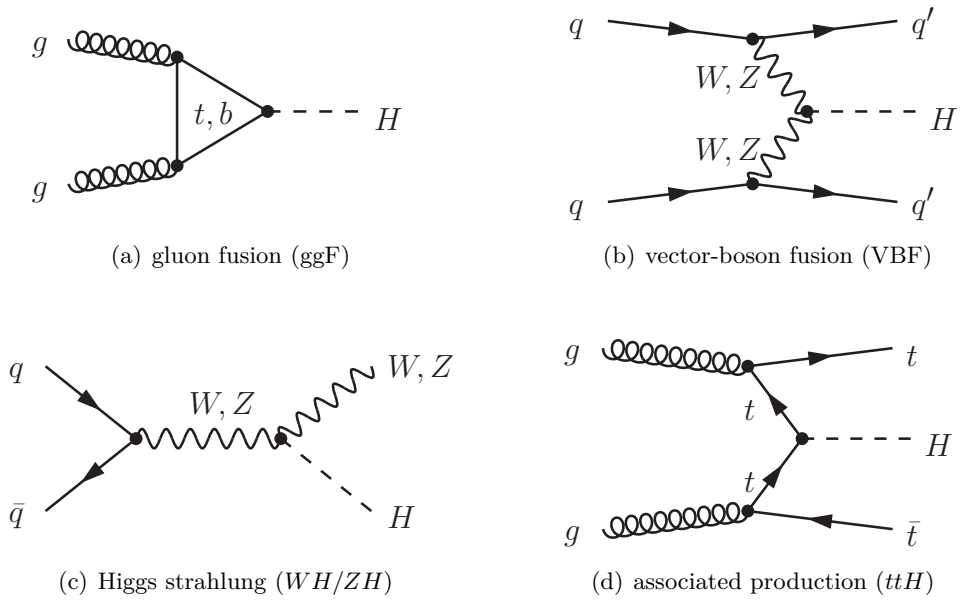
2.5. Phenomenology of the Standard Model Higgs Boson at the LHC

The search for the Standard Model Higgs boson is one of the key aspects for the experimental programs at the LHC. The Higgs mechanism and the Higgs boson along with its coupling properties are introduced in Section 2.2. The following sections focus on the production and decay modes of a Standard Model Higgs boson in proton-proton collisions at the LHC [44]. Special emphasis is put on the decay channels that have a high discovery potential at the ATLAS experiment.

2.5.1. Higgs Boson Production at the LHC

There is only a limited number of processes relevant for the production of a Standard Model Higgs boson in proton-proton collisions at the LHC. These are the gluon fusion ($gg \to H$), the vector-boson fusion ($qq' \to qq'H$), the associated production with vector-bosons called *Higgs strahlung process* (qq' $\to$ WH or ZH) and the associated production with top-quarks ($gg \to t\bar{t}$H). The leading order Feynman diagrams illustrating these production processes are shown in Fig 2.9.

The gluon fusion production mode via an intermediate heavy quark loop is the dominant production mechanism up to Higgs boson masses of about 1 TeV. The next important production process arises from vector-boson fusion which becomes comparable to the gluon fusion mode for very large Higgs boson masses. A characteristic feature of these events is the presence of two forward jets arising, in leading order, from the two outgoing quarks. The



(a) gluon fusion (ggF)

(b) vector-boson fusion (VBF)

(c) Higgs strahlung (WH/ZH)

(d) associated production (ttH)

Figure 2.9.: Leading order Feynman diagrams illustrating the four main production mechanisms of a Standard Model Higgs boson at the LHC, namely the gluon fusion (a) and the vector-boson fusion production process (b). The Higgs strahlung process (c) and the associated production with a $t\bar{t}$ pair (d) are shown on the bottom row.

associated production modes with vector-bosons or a $t\bar{t}$ pair exhibit smaller cross sections. The vector-bosons or the top-quark pair in the final states may be used as an additional handle to identify Higgs bosons arising from these production modes. The gluon fusion and the vector-boson fusion production modes are of particular interested for the analysis presented in Chapter 6.

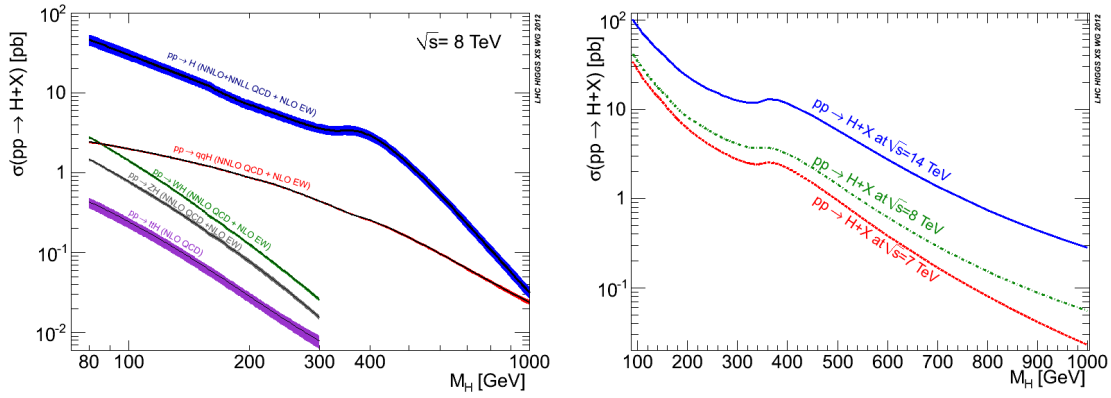
The cross sections for the individual production channels as a function of the Higgs boson mass are shown in Fig. 2.10 for a centre-of-mass energy of $\sqrt{s} = 8$ TeV. The total production cross section for different centre-of-mass energies of $\sqrt{s} = 7$ TeV, $\sqrt{s} = 8$ TeV and $\sqrt{s} = 14$ TeV are shown in Fig. 2.10 as well as a function of the Higgs boson mass.

2.5.2. Higher Order Corrections to Higgs Boson Production at the LHC

This section briefly presents the current status of higher order corrections available to the gluon fusion production mode following Refs. [44, 46].

In QCD perturbation theory, the leading order contribution to the gluon fusion cross section is proportional to α_s^2. The gluon fusion production mode contains an intermediate heavy quark loop and the main contribution arises from top-quarks due to the large Yukawa coupling to the Higgs boson. It is possible to work in the infinite top-quark mass limit to reduce the one-loop $gg \to H$ vertex to a tree-level effective vertex. Next-to-leading order QCD corrections can then be calculated as corrections to the effective Hgg vertex and the complexity of the calculation is reduced.

Higher order corrections to the gluon fusion process arise from virtual corrections, propagator corrections and emissions of real particles as illustrated in Fig. 2.11. The NLO corrections are large and increase the leading-order cross section by about 80-100% at the LHC. The NLO order of QCD corrections have been calculated both with [47, 48] and without [49, 50] the infinite top-quark mass limit and the exact calculation is very well approximated by the



(a) (b)

Figure 2.10.: (a) Cross sections for the dominant Standard Model Higgs boson production processes in proton-proton collisions at the LHC at a centre-of-mass energy of $\sqrt{s} = 8$ TeV as a function of the Higgs boson mass. (b) Total cross section for the production processes of a Standard Model Higgs boson at the LHC for centre-of-mass energies of $\sqrt{s} = 7$ TeV, $\sqrt{s} = 8$ TeV and $\sqrt{s} = 14$ TeV (both taken from Ref. [45]).

large-m_{top} limit. The NNLO corrections to the total cross section have been calculated using the infinite top-quark mass limit [51–57], leading to an additional increase of the cross section of about 25%. In addition to these corrections, large leading logarithms have been summed up that appear at small transverse momenta [58]. By including the resummation of soft gluon contributions up to NNLL, the cross section increased again by about 7-9% at $\sqrt{s} = 7$ TeV.

Considerable work has also been done to determine the size of electroweak corrections. Two-loop electroweak effects are known [59–63] and their size strongly depends on the Higgs boson mass. The cross section is increased by +5% for m_H =125 GeV and reduced by about -2% for m_H =300 GeV [63].

The current best calculation of the gluon fusion cross section, including all known effects from NNLO fixed order calculations, NNLL soft gluon logarithmic resummation and NLO electroweak corrections are shown in Fig. 2.10. Full NLO QCD and EW corrections [64–66] and approximate NNLO QCD corrections [67] are used to calculate the cross sections for VBF signal production. The cross sections of the associated WH/ZH production processes are calculated up to NNLO QCD corrections [68,69] and NLO EW corrections [70]. The cross section for the associated production with a $t\bar{t}$ pair are estimated at NLO in QCD [71–75]

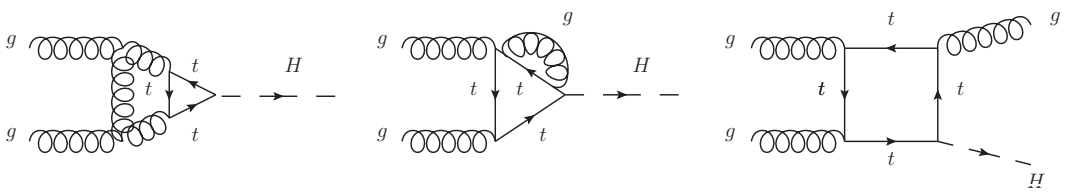
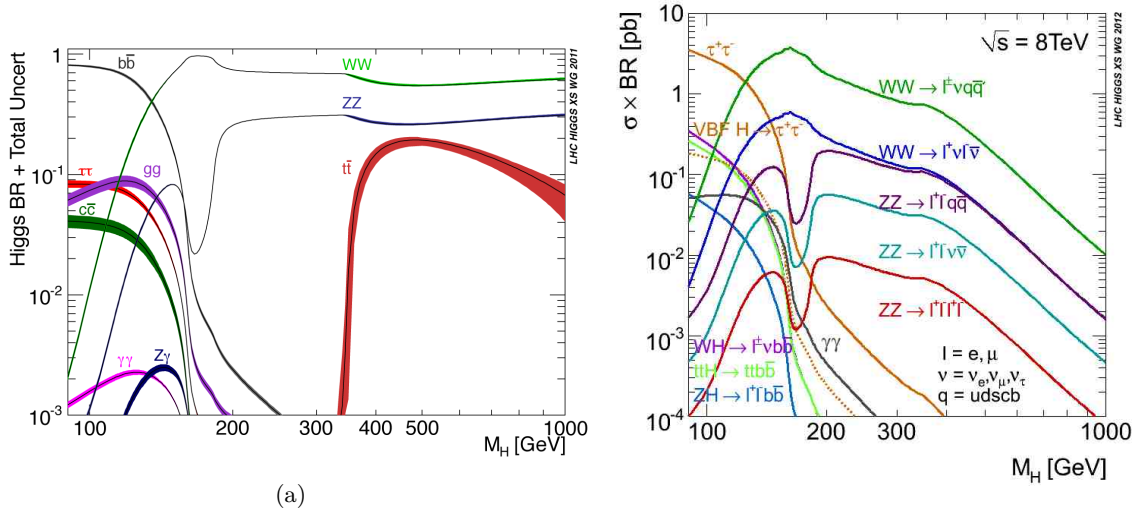


Figure 2.11.: Feynman diagrams illustrating next-to-leading order processes to the gluon fusion Higgs boson production mode arising from vertex corrections (left), propagator corrections (middle) and emissions of real particles (right).

2.5.3. Higgs Boson Decay Modes

Since the couplings of the Higgs boson are always proportional to some mass scale, the most probable decay mode of the Higgs boson is the one into the heaviest kinematically available particles. Once the Higgs boson mass is given, its coupling to the fermions and gauge bosons is fixed by the Standard Model and thus the production cross section, decay width and branching ratios are fully determined. The branching ratios of the Standard Model Higgs boson for different decay modes as function of the Higgs boson mass are shown in Fig. 2.12. The $H \to b\bar{b}$ decay mode is by far the dominant channel below the threshold for the production of a pair of W bosons. Once the mass of the Higgs boson is large enough to allow the production of a pair of gauge bosons, $H \to W^+W^-$ and $H \to ZZ$ become the dominant decay channels. For $m_H > 2m_{top}$, the branching fraction for the decay $H \to t\bar{t}$ becomes also sizeable, although it stays smaller than the boson decay modes.



(a) (b)

Figure 2.12.: (a) Branching ratios of the different Standard Model Higgs boson decay modes as a function of the Higgs boson mass. (b) The branching ratios times the Higgs boson production cross section at the LHC at a centre-of-mass energy of $\sqrt{s} = 8$ TeV as a function of the Higgs boson mass (both taken from Ref. [45]).

2.5.4. Higgs Boson Search Channels at ATLAS

Various decay channels are subject to the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. The combination of the various decay channels features sensitivity to a SM Higgs boson over a broad range of the Higgs boson mass [76, 77]. However, due to the environment of hadronic collisions, some of the decay modes are less accessible than others. For example, considering the product of the branching ratios times the Higgs boson production cross section shown in Fig. 2.12 it is evident that at relatively low Higgs boson masses, decays to $b\bar{b}$ are dominant. However, it is very difficult to perform a search for a Standard Model Higgs boson produced in the gluon fusion mode which decays to $b\bar{b}$ due to the overwhelming direct QCD production of $b\bar{b}$ pairs in proton-proton collisions. The search becomes feasible in case of the vector-boson fusion or associated production mode since they provide additional characteristics that help to discriminate against the dominant multi-jet background but it still remains challenging.

Processes with leptons, e.g. electrons and muons, or photons in the final state are more beneficial in the environment of proton-proton collisions since they provide a clean signature. Thus, the highest sensitivity is expected for searches focusing on $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay modes. It should be noted that the $H \to \gamma\gamma$ decay mode is mainly sensitive to low Higgs boson masses in the range of 110 to 140 GeV while the $H \to ZZ^* \to 4\ell$ and $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channels are sensitive over a much broader range. The $H \to ZZ^* \to 4\ell$ channel is interesting since the Higgs boson mass can be fully reconstructed using the four leptons in the final state. However, the low branching ratio for the $ZZ \to \ell^+\ell^-\ell^+\ell^-$ decay mode leads to a low number of expected events and a significant amount of data must be analysed to identify a possible signal.

The signature of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channel is characterised by the decay products of the W bosons, namely two isolated, oppositely charged leptons and large missing transverse energy caused by the neutrinos escaping undetected the ATLAS detector. This decay mode is the most sensitive channel over a wide range of the Higgs boson mass and the search for the Standard Model Higgs boson in this channel is discussed in detail in Chapter 6.

The design of the LHC detectors has taken into account all these very characteristic properties in order to optimise the search potential for the Higgs boson. The individual detector components are discussed in Chapter 3.

2.6. Status of Experimental Observations and Theoretical Predictions

The framework of the Standard Model has been presented in the previous sections. It gives a detailed description of the fundamental particles and the interacting forces. Taking the compositeness of protons into account, the cross sections of processes occurring in proton-proton collisions can be predicted. The measurements of several important benchmark Standard Model production cross sections performed by the ATLAS collaboration are compared to the corresponding theoretical expectations in Fig. 2.13.

The production cross sections of the inclusive Drell-Yan processes $W \to \ell\nu$ and $Z \to \ell^+\ell^-$ ($\ell = e, \mu$) have been measured in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector [78]. The measurements are based on data corresponding to an integrated luminosity of about 35 pb^{-1} collected in 2010 by the ATLAS experiment. The cross sections are integrated over a fiducial kinematic range and also extrapolated to the full kinematic range to obtain the total integrated W and Z/γ^* cross sections. The integrated W and Z/γ^* cross sections in the electron and muon channels are combined. The precision of the integrated cross section measurements has reached the percent level. Comparisons with QCD predictions therefore are made at next-to-next-to-leading order in perturbation theory using recent NNLO PDFs. The theoretical predictions and the experimental results, shown in Fig. 2.13, are in good agreement.

A measurement of the $t\bar{t}$ production cross section in the all-hadronic channel was performed based on proton-proton collision data taken in 2011 at $\sqrt{s} = 7$ TeV by the ATLAS experiment [79]. The data corresponds to an integrated luminosity of 1.02 fb^{-1}. The measured $t\bar{t}$ cross section is consistent with the predictions for the cross section determined at next-to-next-to-leading order.

The WW production cross section has been measured in proton-proton collisions at $\sqrt{s} =$ 7 TeV [80] in purely leptonic decay channels $W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ with $\ell = e, \mu$. The measurement is based on data corresponding to an integrated luminosity of 4.7 fb^{-1} collected with the ATLAS detector. The measured WW production cross section is compatible with the

Standard Model cross section prediction determined at next-to-leading order.

It is evident from Fig. 2.13 that the theoretical predictions, all calculated at NLO or higher, and the experimental results are in good agreement for both cross sections measured at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV for all displayed processes.

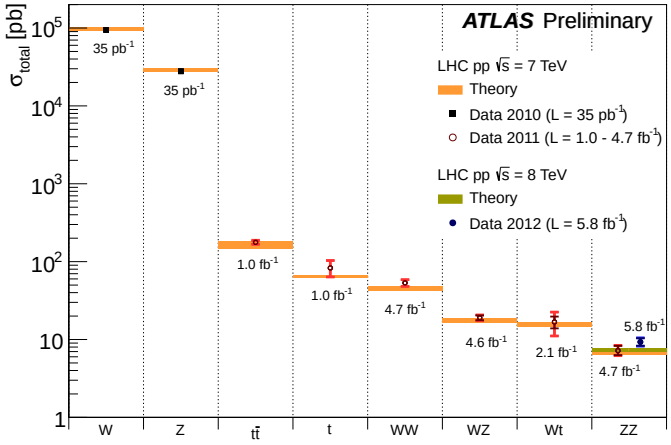


Figure 2.13.: Summary of several Standard Model total production cross section measurements compared to the corresponding theoretical expectations. The W and Z boson inclusive cross sections were measured with 35 pb^{-1} of integrated luminosity collected in 2010 while all other measurements were performed using the larger datasets collected in 2011 and 2012. The dark error bar represents the statistical uncertainty. The red error bar represents the full uncertainty, including systematic and luminosity uncertainties. All theoretical expectations are calculated at NLO or higher (taken from Ref. [81]).

3

The Large Hadron Collider and the ATLAS Detector

The datasets used in this thesis were collected with the ATLAS detector, a multi purpose particle detector placed at the LHC, a large hadron collider ring at CERN near Geneva in Switzerland. This chapter provides an overview of the experimental setup. First, a description of the LHC is given in Section 3.1 following closely Refs. [82, 83]. Afterwards, the main components of the ATLAS detector, namely the tracking detector, the electromagnetic and hadronic calorimeters and the muon chambers, are briefly discussed in Section 3.2. A much more detailed description can be found in Refs. [84, 85].

3.1. The Large Hadron Collider

The LHC is today the largest particle collider ring operating at highest energies ever reached. The accelerator is designed to collide proton-proton beams at a centre-of-mass energy of 14 TeV and has been operating in 2012 at a centre-of-mass energy of 8 TeV. There are four interaction points at the LHC and the general purpose detectors ATLAS (A Torodial LHC ApparatuS) and CMS [86] (Compact Muon Solenoid) are located at two of them. Both of them were designed and optimised for new physics studies such as the search for the Higgs boson or new particles predicted by supersymmetry models. There are two more experiments called LHCb [87] and ALICE [88] (A Large Ion Collider Experiment). The former is dedicated to the study of B-hadrons and CP violation in hadron decays while the latter focuses on lead ion collisions aiming for studies of the quark-gluon plasma which is postulated to have existed during the early universe.

The LHC is hosted in the tunnel that was built previously for the LEP collider which operated between 1989 and 2000 [89]. The tunnel was built at a mean depth of 100 m and has a circumference of 27 km. The accelerator complex at CERN is a succession of machines with increasingly higher energies as illustrated in Fig. 3.1. Protons are injected into the PS Booster at an energy of 50 eV from the linear accelerator *Linac2*. The booster accelerates the protons up to 1.4 GeV and the proton beams are fed to the Proton Synchrotron (PS) where they are accelerated to 25 GeV. The protons are then sent to the Super Proton Synchrotron (SPS) where they reach an energy of 450 GeV. Finally, the beams are transferred to the LHC where beams are circulated both in clockwise and anticlockwise direction while reaching their

final energy.

Protons arrive at the LHC in bunches which are prepared in the smaller machines beforehand. The bunches are circulated in a vacuum tube and electromagnetic devices are used to control their movement. The main components of the accelerator are superconducting dipole magnets, operating at a temperature of 1.9 K and designed for producing magnetic fields of 8.33 T that keep the particles in their nearly circular orbits. Superfluid helium is used to cool the more than 1200 dipole magnets. Furthermore, quadrupole magnets focus the beam, and accelerating cavities that are electromagnetic resonators keep the bunches at a constant energy by compensating for energy losses.

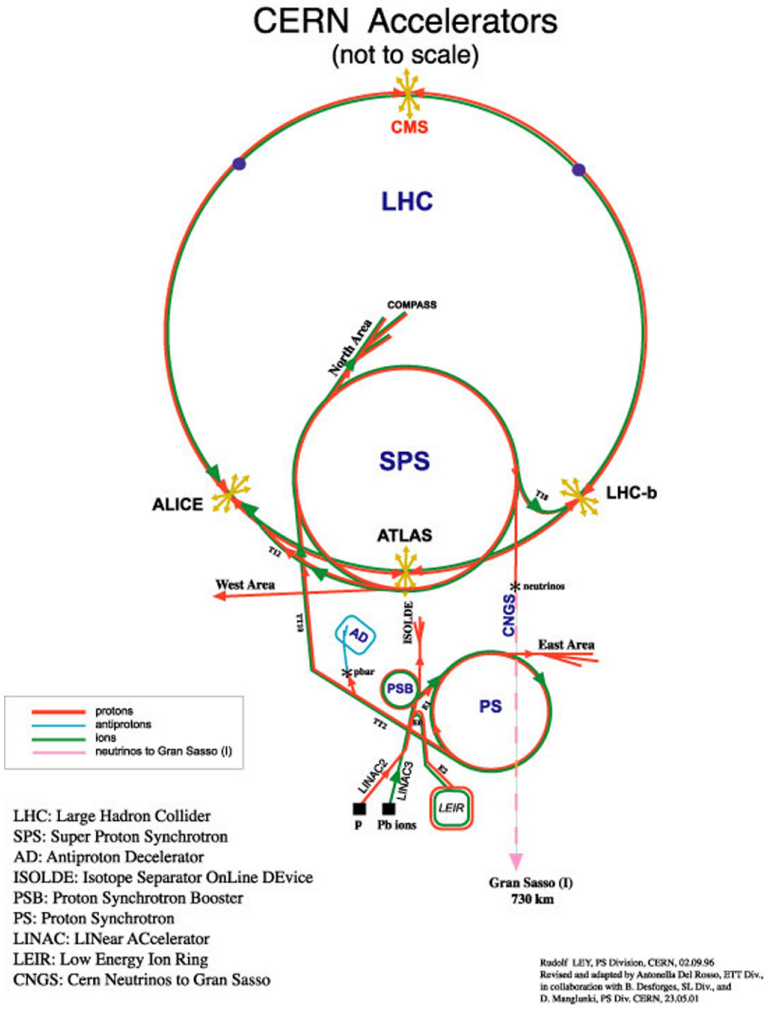


Figure 3.1.: Accelerator complex used to achieve the 7 TeV centre-of-mass energy of the colliding protons [90].

Under nominal operating conditions, each proton beam has 2808 bunches, with each bunch containing about 10^{11} protons. The bunch size is not constant since the bunches get squeezed and expanded as they circulate around the LHC. While the bunches of particles are designed to be 7.55 cm long and a millimetre wide when they are far from a collision point, they are squeezed to about 16 μm as they approach the collision point to allow for a greater chance of proton-proton collisions. The LHC is designed to operate with up to 2808 bunches with a spacing of about 7.5 m which corresponds to a time interval of 25 ns between two bunch

crossings or a frequency of 40 MHz. This configuration is designed to achieve an instantaneous luminosity of 10^{34} cm^{-2}s^{-1}. The machine luminosity depends solely on beam parameters and can be calculated as follows [91]:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \varepsilon_n \beta^*} F \tag{3.1}$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the frequency of complete turns around the ring, γ_r is the relativistic gamma factor for particles in the beam, ε_n is the beam emittance which is a measure of how much the particles depart from the ideal trajectory, β^* is the beta function at the collision point giving the envelope for the particle motion and F is the luminosity reduction factor due to the crossing angle at the interaction point. At a nominal centre-of-mass energy of 14 TeV and with the given number of bunches, the average number of interactions per bunch-crossing would amount to $\mu \approx 25$.

3.1.1. Data Taking at the Large Hadron Collider

The first operation of the LHC started in autumn 2008 but was interrupted by a severe incident only a few days afterwards. The repair of the damage, caused by faulty connections between the superconducting magnets, took more than one year [92]. In November 2009, the operation was resumed by proton-proton collisions at a centre-of-mass energy of 900 GeV followed by collisions at $\sqrt{s} = 7$ TeV starting in March 2010 after a winter shutdown. The LHC provided proton-proton collisions at $\sqrt{s} = 7$ TeV in 2010 and 2011 and the centre-of-mass energy was even further increased to 8 TeV in 2012. The values of some characteristic parameters for the designed operation as well as for the operations at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV are summarised in Table 3.1.

Table 3.1.: Overview of machine parameters of the LHC operation during the data taking periods in 2010, 2011 and 2012 compared to the design values (taken from Refs. [13,93,94]).

	2010	2011	2012	Design
Centre-of-mass energy	7 TeV	7 TeV	8 TeV	14 TeV
Peak luminosity [cm^{-2}s^{-1}]	2×10^{32}	3.65×10^{33}	7.73×10^{33}	10^{34}
Protons per bunch ($\times10^{11}$)	0.1-1.2	0.6-1.2	1.48	1.15
Number of bunches	7 TeV	200-1380	1380	2808
Average collisions per bunch-crossing	≤ 3	9.1	20	22
Time between bunches [ns]	$\geq$150	75 and 50	50	25

The integrated luminosity of proton-proton collisions delivered by the LHC and recorded by the ATLAS experiment in 2010, 2011 and 2012 are shown in Fig. 3.2. The total integrated luminosities of proton-proton collisions at $\sqrt{s} = 7$ TeV amount to 45.0 pb^{-1} and 5.25 fb^{-1} collected by the ATLAS experiment in 2010 and 2011, respectively. The increase in luminosity is mainly due to the higher number of bunches and protons per bunch in combination with a shorter time interval between the bunches. In 2012, the total integrated luminosity of proton-proton collisions at $\sqrt{s} = 8$ TeV recorded by the ATLAS experiment sums up to 21.7 fb^{-1}.

The full dataset taken in 2010 is used for the determination of electron shower shapes presented in Chapter 5. The complete 2011 dataset and the dataset collected between the beginning of April and the middle of June 2012, corresponding to 5.25 fb^{-1} and 6.3 fb^{-1}, respectively, are included in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis presented in Chapter 6.

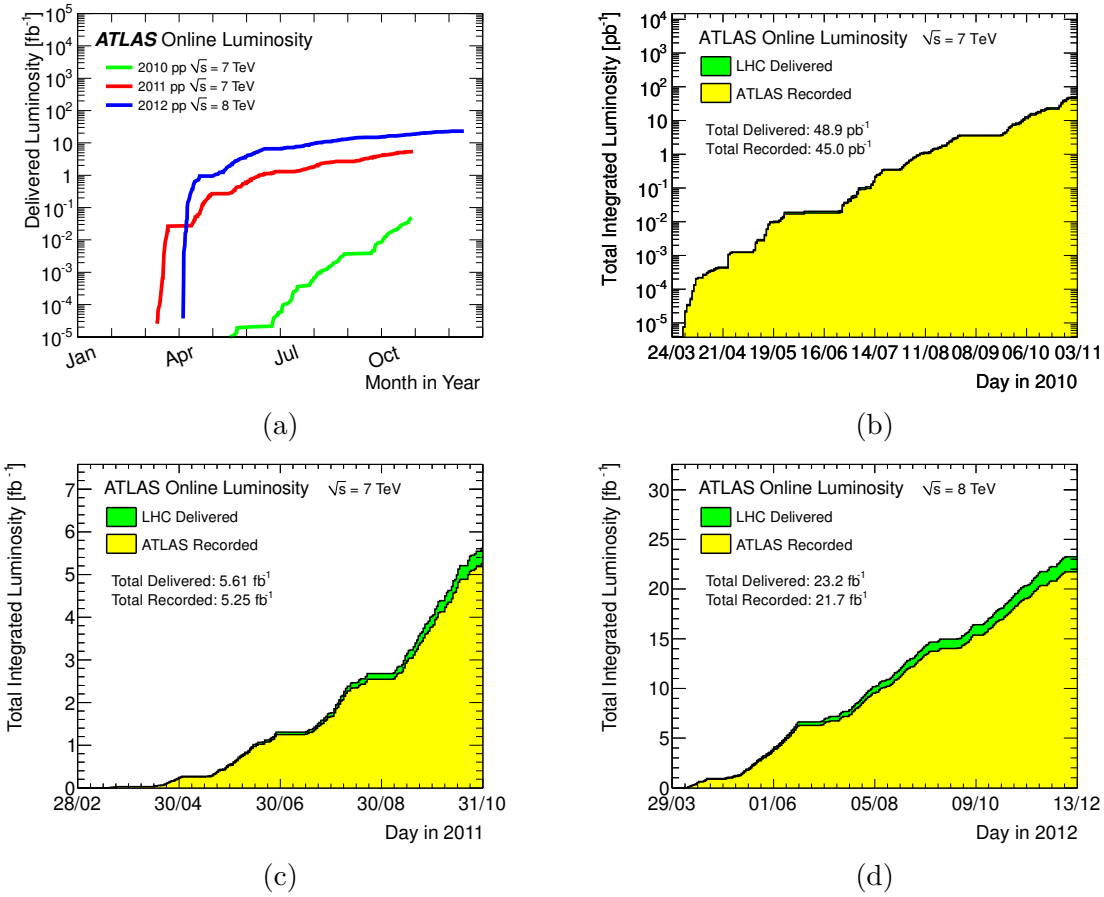


Figure 3.2.: (a) The cumulative luminosity versus day delivered to and recorded by the ATLAS experiment in proton-proton collisions are shown for the 2010 (green), 2011 (red) and 2012 (blue) data taking periods. The individual integrated luminosities of proton-proton collisions at $\sqrt{s} = 7$ TeV versus day in 2010 (b) and 2011 (c) and at $\sqrt{s} = 8$ TeV in 2012 (d) are shown as well. The plots in the top row are shown in logarithmic scale to emphasise the fast ramp up due to the huge increase of the instantaneous luminosity [93].

3.2. The ATLAS Detector

Technically, the ATLAS detector is designed in view of the high interaction rates, particle multiplicities and energies expected to emerge from proton-proton and heavy ion collisions. Furthermore, the requirements to perform precision measurements of known processes as well as searches for new physics phenomena and particles imply the ability to identify the full range of particles which may be produced in these processes. Thus, all individual components of the ATLAS detector must be able to efficiently reconstruct and identify basic experimental objects such as electrons, photons, muons, τ-leptons and hadronic jets. The resulting requirements

to meet all these goals are [84]:

- Radiation-hard electronics and sensor elements in combination with fast readout of high granularity detector elements to handle the particle flux and reduce the influence of pile-up.
- Large acceptance in pseudorapidity (defined below) with almost full azimuthal coverage.
- Good momentum resolution and reconstruction efficiency of charged particles in the inner detector to allow for vertex detection and reconstruction of secondary vertices.
- Very good electromagnetic calorimetry to identify electrons and photons in combination with full-coverage hadronic calorimetry to facilitate accurate measurements of jets and missing transverse energy.
- Good muon identification and momentum resolution over a wide range of momenta.
- Fast and highly efficient trigger on events of interest to achieve an acceptable rate of events to be stored.

The ATLAS detector consists of several concentric layers and has a forward-backward symmetric structure around the beam pipe. A cut-away view of the overall layout is shown in Fig. 3.3. From the inside to the outside, the main detector components are the inner detector (ID), the calorimeter system and the muon system (MS). The various subdetectors are briefly discussed in the following sections. The descriptions are based on Refs. [84, 85] if not stated otherwise. It should be emphasised that this detector represents the work of a large collaboration of several thousand physicists, engineers, technicians and students over a period of more than 20 years of dedicated design, development, fabrication and installation.

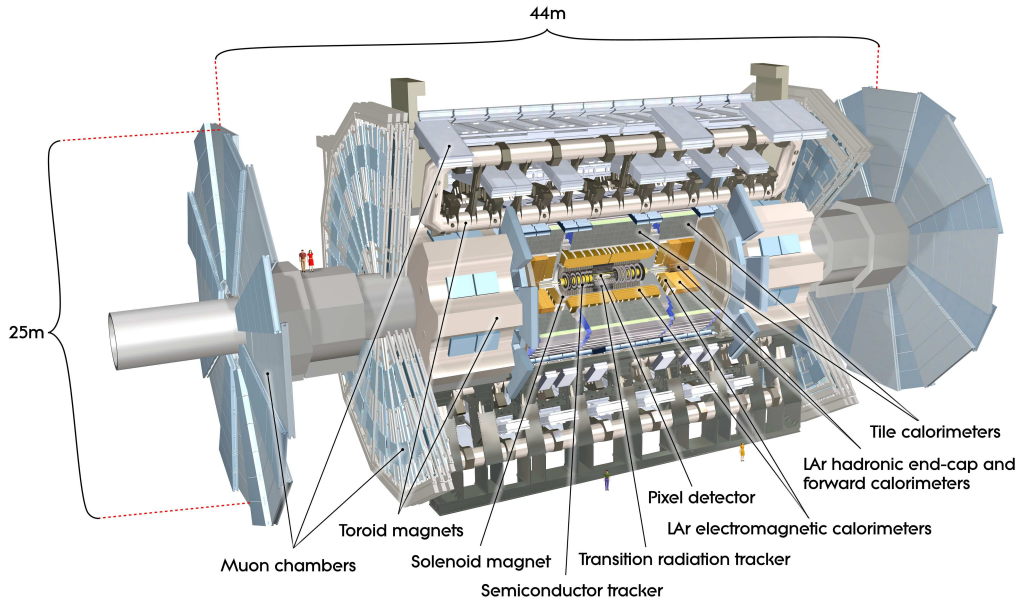


Figure 3.3.: Cut-away view of the ATLAS detector. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tonnes [84].

The right-handed coordinate system used to describe the ATLAS detector is briefly summarised here as it is repeatedly used throughout this thesis. The origin of the coordinate system is defined by the nominal interaction point and the beam that goes counter-clockwise in the LHC ring (looked from above) defines the z-axis. The x-y-plane is transverse to the beam axis and the positive x-axis is defined as pointing from the interaction point towards the centre of the LHC ring while the y-axis is defined as pointing upwards. The azimuthal angle ϕ is measured in the x-y-plane around the beam axis and defined as:

$$\phi = 1/\tan(x/y) \tag{3.2}$$

The polar angle θ is measured between the z-axis and the x-y-plane and defined as:

$$\theta = 1/\tan(\sqrt{x^2+y^2}/z). \tag{3.3}$$

The pseudorapidity is defined as

$$\eta = -\ln(\tan(\theta/2)) \tag{3.4}$$

which is equal to the rapidity $y = -\ln((E+p_z)/(E-p_z)))$ in case of massless objects ($E = |\vec{p}|$). The transverse momentum p_{T}, the transverse energy E_{T} and the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ are defined in the x-y-plane, if not stated otherwise. The distance ΔR between two objects in the pseudorapidity-azimuthal angle space is defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{3.5}$$

where $\Delta\eta$ and $\Delta\phi$ are the differences between the object coordinates in η and ϕ, respectively.

3.2.1. Inner Detector

The ATLAS inner detector is designed to provide robust pattern recognition, excellent momentum resolution and both primary and secondary vertex measurements for charged tracks above a given p_{T} threshold of typically 0.5 GeV. The inner detector consists of three independent and complementary subsystems with decreasing spatial resolution, namely the pixel detector followed by the Semi Conductor Tracker (SCT) and the Transition Radiation Tracker (TRT). A cut-away view of the ATLAS ID is shown in Fig. 3.4. The high-resolution semiconductor pixel and strip detectors cover the region $|\eta|< 2.5$ while the TRT reaches up to $|\eta|< 2.0$. The whole system is surrounded by a magnetic field of 2 Tesla generated by the central solenoid, which extends over a length of 5.3 m with a diameter of 2.5 m.

Pixel Detector

The innermost part of the tracking volume consists of three finely granulated cylindrical layers of pixel detectors providing high-resolution pattern recognition. The pixel modules are arranged on concentric cylinders around the beam axis in the barrel region and are located on disks perpendicular to the beam axis in the endcap region. A schematic view of a quarter-section of the ATLAS ID is shown in Fig. 3.5. The arrangement of the detectors is chosen such that each charged track originating from the interaction region crosses at least three pixel layers. Each pixel detector comprises 1744 pixel sensors with a nominal size of 50×400 $\mu\mathrm{m}^2$ in $(R-\phi)\times z$ and 250 μm thickness. The intrinsic accuracies for the measurement of single space points is 10 μm in $(R-\phi)\times z$ and 115 μm in z. The innermost layer of pixels, the so-called *B-layer*, is as close as 5 cm to the beam pipe. It is responsible for the measurement of the

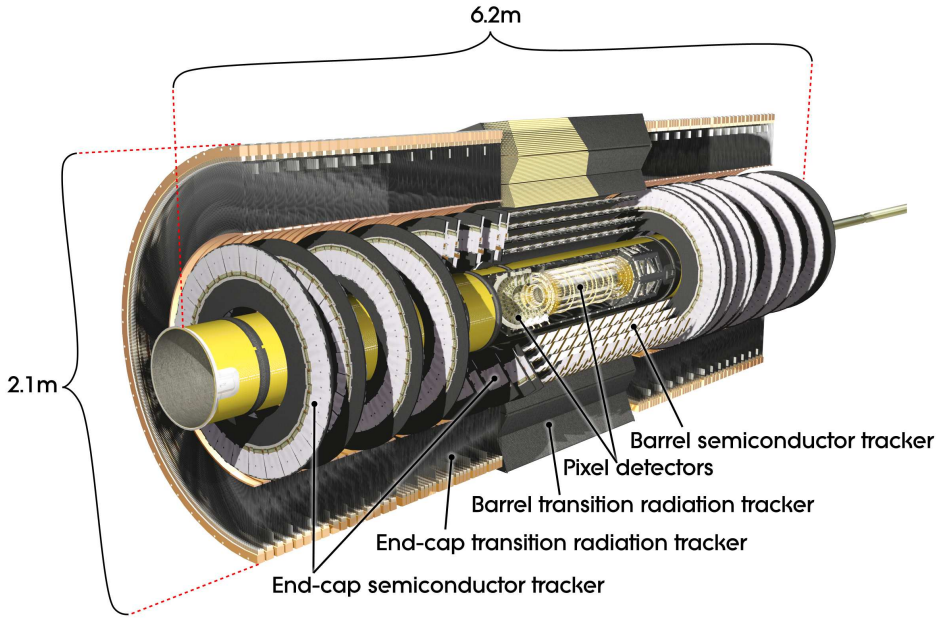


Figure 3.4.: Cut-away view of the ATLAS inner detector [84].

secondary vertices and the identification of jets originating from b-quarks (see Section 4.2.6). In addition, this layer also helps to suppress electrons originating from photon conversions (see Section 4.2.4). Due to radiation damage it is expected that the B-layer will be replaced after approximately three years of operation at design luminosity.

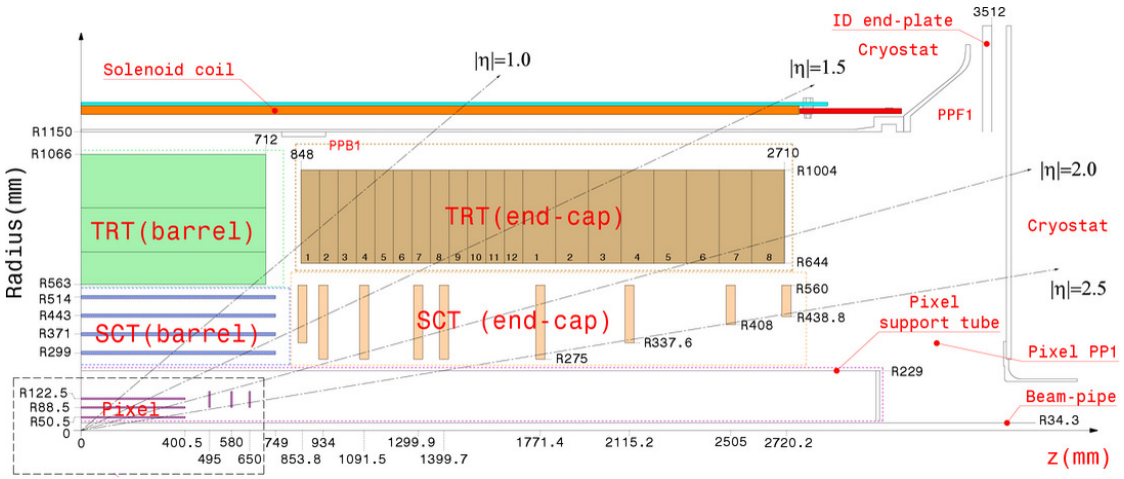


Figure 3.5.: Plan view of a quarter-section of the ATLAS inner detector showing each of the major elements of Pixel, SCT and TRT detectors with its active dimensions and the support structure [84].

Silicon Strip Detector

The Silicon Strip Detector is the one next to the pixel detector and comprises four cylindrical layers in the barrel region and nine discs in the endcap region. The modules consist of two silicon strip sensors of 258 μm thickness which are glued together with one at a stereo angle of

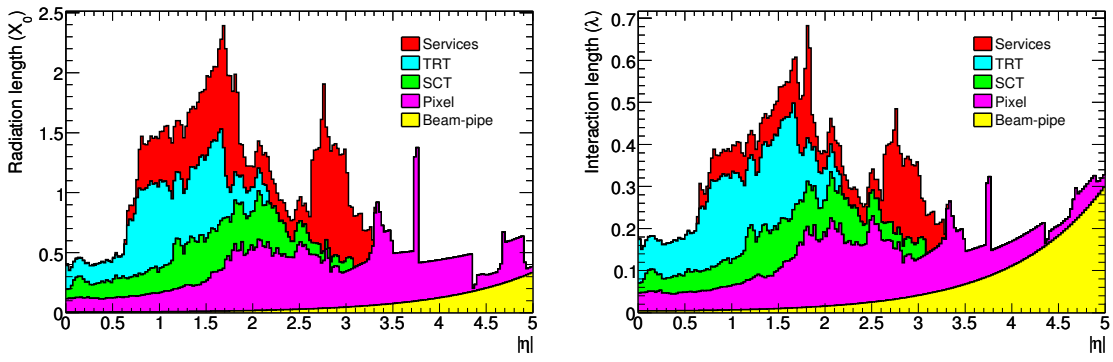
40 mrad with respect to the other. This allows the precise determination of three dimensional space points in the $R - \phi$ direction. In the barrel region one set of strips in each layer is oriented parallel to the beam direction while another set of strips is running radially in the endcap region. The mean pitch of the strips is 80 μm. The modules are arranged such that each charged particle crosses at least eight strip layers which result in four space points. The intrinsic accuracies per module are 17 μm in $R - \phi$ and 580 μm in z for barrel layers and in R for endcap discs. Due to the high-radiation environment it is necessary to operate the silicon sensors at low temperatures of approximately -5 to -10°C to maintain adequate noise performance after radiation damage.

Transition Radiation Tracker

The outermost part of the ID is the Transition Radiation Tracker which is composed of many layers of polyimide drift tubes of 4 mm diametre interleaved with transition radiation material. The straws are mechanically stabilised by carbon fibres and filled with a gas mixture of 70% Xenon, 27% CO_2 and 3% O_2. Charged particles crossing a boundary between different dielectric media emit transition radiation with an intensity proportional to the Lorentz factor $\gamma = E/m$. These photons have the energy of typically several keV and are absorbed in the Xenon-based gas mixture of the straw tubes. Electrons have a lower mass compared to pions and thus emit a significant amount of transition radiation starting at much lower momenta than hadrons. This effect is used to enhance the identification capabilities for electrons and thus achieve a better discrimination of electrons and pions. The readout of the TRT comprises a low threshold to detect ionisation and a high threshold to identify transition radiation. The TRT is divided into a barrel detector where the straws of 144 cm length are oriented parallel to the beam axis and a endcap region where the 37 cm long straws are arranged radially in wheels. The TRT only provides information in $R - \phi$ direction, for which it has an intrinsic accuracy of 130 μm per straw. Since the straws are arranged to give on average 36 hits per track, the large number of hits compensates for the lack of intrinsic precision.

Material Effects in the Inner Detector

The hits detected in the pixel, SCT and TRT detectors serve as input for the reconstruction of charged particle tracks (see Section 4.2.1). The reconstruction is feasible for tracks with transverse momenta $p_{\mathrm{T}} > 0.5$ GeV and within $|\eta| < 2.5$. However, the efficiency at low momenta is limited due to material effects in the inner detector. The overall material budget of the ID in terms of radiation length X_0 and interaction length λ is shown in Fig. 3.6. The contributions are split up for the individual sub-detectors, including services in their active volume, and contributions of external services which are most striking at the interface of the barrel and endcap regions. Charged particles at relatively low momenta are particularly affected by the amount of material present in the inner detector. For example, electrons can lose a significant amount of energy through bremsstrahlung and many photons convert into an electron-positron pair before reaching the electromagnetic calorimeter. Even a significant fraction of low-energy charged pions will undergo inelastic hadronic interactions inside the inner detector volume. Thus, a detailed modelling of the ID material has been implemented in the simulation to take these effects into account. In case of electrons, the track reconstruction algorithm explicitly considers the effects due to bremsstrahlung (see Sections 4.2.4 and 6.12.1).



(a) (b)

Figure 3.6.: Material distribution in terms of radiation length X_0 (a) and interaction length λ (b) at the exit of the inner detector envelope, including the services and thermal enclosures. The distributions are shown as a function of $|\eta|$ and averaged over ϕ. The breakdown indicates the contributions of external services and of individual sub-detectors, including services in their active volume [84].

3.2.2. Calorimeter System

The ATLAS calorimeter system is designed to measure the amount and the position of energy deposited by particles as precisely as possible. All ATLAS calorimeters are sampling calorimeter made of absorbers and a detection medium. Whenever a particle hits the absorber it interacts with the material and produces so-called *showers* of secondary particles which are detected in the active material. At each stage in the showering, the emerging particles have less and less energy until finally the shower is completely absorbed. Different types of calorimeter are used to provide good resolution of electromagnetic showers as well as good containment of hadronic showers to limit the punch-through into the muon system. Hence, the total thickness of the calorimeter is an important parameter in the design consideration in order to fully contain the particle shower over a large energy range. The ATLAS calorimeter system covers the range up to $|\eta| < 4.9$ and is composed of a finely granulated electromagnetic calorimeter providing precise measurements of electrons and photons covering the pseudorapidity range $|\eta| < 3.2$. It has a total thickness of more than 22 radiation lengths (X_0) in the barrel region and more than 24 radiation lengths in the endcap region. This part is surrounded by a hadronic calorimeter of coarser granularity in the range up to $|\eta| < 3.2$ which is sufficient to perform measurements for jet reconstruction and $E_{\mathrm{T}}^{\mathrm{miss}}$. It features a total thickness of about 10 to 11 interaction lengths and provides therefore a sufficient reduction of punch-through. The setup is complemented by forward calorimeters providing both electromagnetic and hadronic energy measurements up to $|\eta| < 4.9$. A cut-away view of all the sampling calorimeters is presented in Fig. 3.7.

Electromagnetic Calorimeter

The electromagnetic calorimeter is designed to be projective in η and homogeneous in azimuthal coverage without cracks. Thus, it is constructed from accordion-shaped kapton electrodes and lead absorber plates in combination with liquid argon serving as active material. It is divided into a barrel part ($|\eta| < 1.47$) consisting of two identical half-barrels separated by a small gap of 4 mm at $z = 0$ and two endcaps comprising two coaxial wheels covering the region $1.375 < |\eta| < 3.2$. Over the η region matched to the inner detector

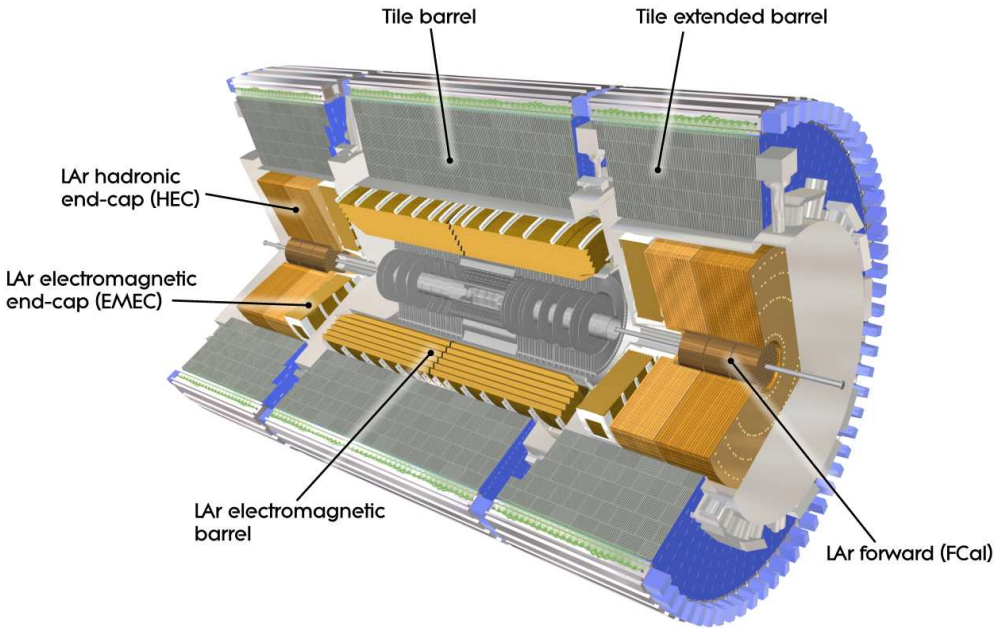


Figure 3.7.: Cut-away view of the ATLAS calorimeter system containing electromagnetic liquid Argon (LAr) and hadronic scintillator-tile calorimeter as well as LAr forward calorimeter [84].

($|\eta| < 2.5$), the electromagnetic calorimeter is segmented in depth into three longitudinal layers with decreasing granularity as illustrated in Fig. 3.8. The fine granularity of the EM calorimeter is ideally suited for precision measurements of electrons and photons. The first layer features a very fine granularity in η an is called *strip layer*. It allows for the precise measurement of the structure of electromagnetic showers and thus allows to disentangle even close-by photon pairs originating from π^0 and η decays. The cells in the second layer have granularities of 0.025×0.025 and 0.050×0.025 in $\Delta\eta \times \Delta\phi$ in the barrel and endcap regions, respectively. Most of the energy of the electromagnetic showers originating from electrons and photons is deposited in the second layer which has a thickness of about 16 X_0. The third layer owns a depth of about 2 X_0 and a granularity of 0.050×0.025 in $\Delta\eta \times \Delta\phi$. This layer helps to distinguish electromagnetic and hadronic showers. In the endcap region, the calorimeter is segmented in two sections in depth and has a coarser lateral granularity than for the rest of the acceptance. In the region of $|\eta| < 1.8$, a presampler detector consisting of active LAr layers is used to correct for the energy lost by electrons and photons upstream of the calorimeter.

The measured energy depositions in all the layers allow to extract so-called *shower shape variables* which can be used to separate electrons and photons from hadronic objects such as pions. The definition of the variables is discussed in Section 4.2.4. The distributions of some characteristic electron shower shape variables are studied in detail both on data and Monte Carlo simulation in Chapter 5.

Hadronic Calorimeter

The hadronic calorimeters surround the electromagnetic ones. The hadronic calorimeter in the range $|\eta| < 1.7$ uses steel as absorber material and scintillating tiles as active material.

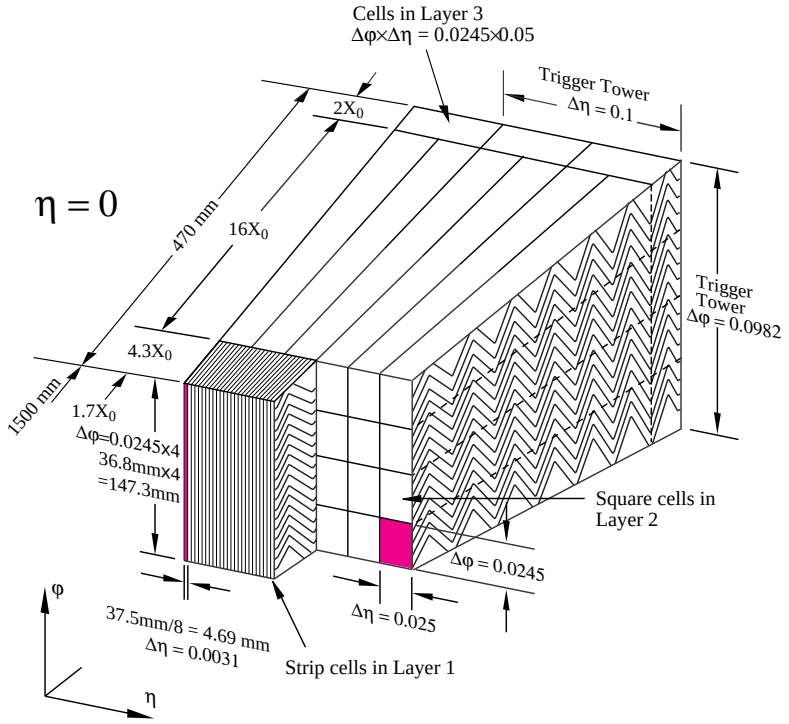


Figure 3.8.: Sketch of the electromagnetic calorimeter displaying the accordion-shaped electrodes. The calorimeter is segmented into three longitudinal layers and the granularity in η and ϕ of the cells in each layer as well as the size of the trigger towers are shown [84].

The scintillator tiles are oriented radially and normal to the beam line which allows for almost seamless azimuthal calorimeter coverage. The light produced in the scintillating material is collected at the edges of each tile and read out into two separate photomultiplier tubes by wavelength-shifting fibres. The hadronic calorimeters are coarser compared to the electromagnetic ones but still possess at least three longitudinal segments. The typical granularity is 0.1×0.1 in $\Delta\eta \times \Delta\phi$ in the first and second layer and $\Delta\eta \times \Delta\phi = 0.1 \times 0.2$ in the third layer. The Hadronic Endcap Calorimeter (HEC) covers the region $1.5 < |\eta| < 3.2$ and the active material is liquid argon while flat copper plates are used as absorber material. The HEC consists of two cylindrical wheels per endcap, each divided into two longitudinal sections. The granularity of the readout cells is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the region $|\eta| < 2.5$ and 0.2×0.2 for larger values of η.

The transition region between the barrel and endcap calorimeters ($1.37 < |\eta| < 1.52$) is called *crack region* and inner detector services as well as services for the barrel liquid argon calorimeter are located there. These regions exhibit reduced energy resolution and thus they are excluded for the studies involving electrons as presented in Chapters 5 and 6.

Forward Calorimeter

The Forward Calorimeter (FCal) covers the region $3.1 < |\eta| < 4.9$. It is a sampling calorimeter consisting of three layers that uses liquid argon as active material. The first layer is optimised for measurements of electromagnetic showers and uses copper as absorber material while the second and third layer are dedicated to the measurements of hadronic showers and thus are made of tungsten. Since these calorimeters are located at high η with a distance of

approximately 4.7 m from the interaction point, they are exposed to high particle fluxes. Thus, very small liquid argon gaps, with an electrode readout structure of small-diameter rods, centered in tubes oriented parallel to the beam direction provide highest possible density and fast readout at the same time.

3.2.3. Muon System

The calorimeter system is surrounded by the muon spectrometer which covers the pseudo-rapidity range $|\eta| < 2.7$ and allows for identification of muons with momenta above 3 GeV and precise determination of the transverse momentum p_{T} up to about 1 TeV. The air-core toroid system, consisting of a long barrel in the region $|\eta| < 1.4$ and two smaller endcap magnets in $1.6 < |\eta| < 2.7$, provides the magnetic field that allows the muon momentum measurement by deflecting the muon track. The bending power is characterised by the field integral $\int B_{\perp} \mathrm{d}l$ of the field component $B_{\perp}$ orthogonal to the muon direction and amounts to 1.5 to 5.5 Tm in the barrel region and to approximately 1.0 to 7.5 Tm in the endcap region. It is evident from Fig. 3.3 that the muon spectrometer defines the overall dimensions of the ATLAS detector.

The muon system consists of precision-tracking chambers located between and on the eight coils of the superconducting toroid magnets in the barrel region, and in front and behind the two smaller magnets (each consisting again of eight coils) in the endcap region. The chambers in the barrel region are arranged in three concentric cylindrical shells at distances of approximately 5 m, 7.5 m and 10 m from the beam axis. In the endcap regions, the muon chambers form large wheels oriented perpendicular to the z-axis that are located at distances of $|z| \approx 7.4$ m, 10.8 m, 14 m and 21.5 m from the interaction point. Particle hit positions are precisely measured by monitored drift tube chambers (MDTs) in the whole range up to $|\eta| < 2.7$ with the exception of the innermost endcap layer where the coverage is limited to $|\eta| < 2.0$. For the innermost tracking layer in the region $2.0 < |\eta| < 2.7$, Cathode-Strip Chambers (CSC) are used which are multi-wire proportional chambers with cathode planes split into strips in orthogonal directions. They provide higher rate capability and time resolution which is beneficial in the forward region.

The muon instrumentation includes, as a key component, independent trigger chambers in the pseudorapidity range up to $|\eta| = 2.4$. Resistive Plate Chambers (RPCs) are used in the barrel region while Thin Gap Chambers (TGCs) are used in the endcap regions. They provide not only trigger information but also bunch crossing-information and measurements of the muon coordinates in the direction orthogonal to those of the precision-tracking chambers. The various strategies to reconstruct and identify muons with the ATLAS experiment are described in Section 4.2.3.

3.2.4. Trigger System

A major challenge for experiments at the LHC is the online selection of interesting events which requires a highly efficient trigger system to reduce the high event rates originating from initial bunch crossing rates of 40 MHz down to 200 Hz compatible with offline computing power and storage capacity. The ATLAS trigger system is composed of three levels of event selection where each level refines the decision made at the previous level by applying additional selection criteria. The three distinct levels, as illustrated in Fig. 3.9, are briefly described in the following.

The Level 1 (L1) trigger stage is hardware-based and uses a limited amount of the total detector information to reach a decision whether to keep an event in less than 2.5 μs, reducing the event rate from up to 40 MHz to about 75-100 kHz. The L1 trigger uses reduced-granularity

information from the calorimeter and muon systems and searches for high transverse momentum signatures originating from electrons, photons, jets and hadronically decaying τ-leptons as well as large missing and total transverse energy. The possible trigger objects within the event identify so-called *regions of interest* (RoI's) that are defined by their geographical coordinates in η and ϕ. They are passed on to the next level.

The Level 2 (L2) trigger is software-based. The selection is largely based on the full-granularity information of all sub-detectors in the RoI's. A sequence of dedicated L2 algorithms is executed for each L1 RoI to compute event feature quantities to determine if the candidate object should be retained. The average processing time available for L2 algorithms is 40 ms and a reduced rate of approximately 3 kHz is achieved.

The final online selection is performed by the event filter (EF) that typically uses the same algorithms as the offline reconstruction taking the full detector information into account. An average processing time of 4 s per event is available to achieve the additional rejection to reduce the output rate to about 200 Hz. The events selected by the EF are finally stored in the CERN computer centre for further offline processing.

Dedicated muon and electron triggers as used for the studies presented in Chapter 5 and Chapter 6 are described in Section 4.2.3 and Section 4.2.4, respectively.

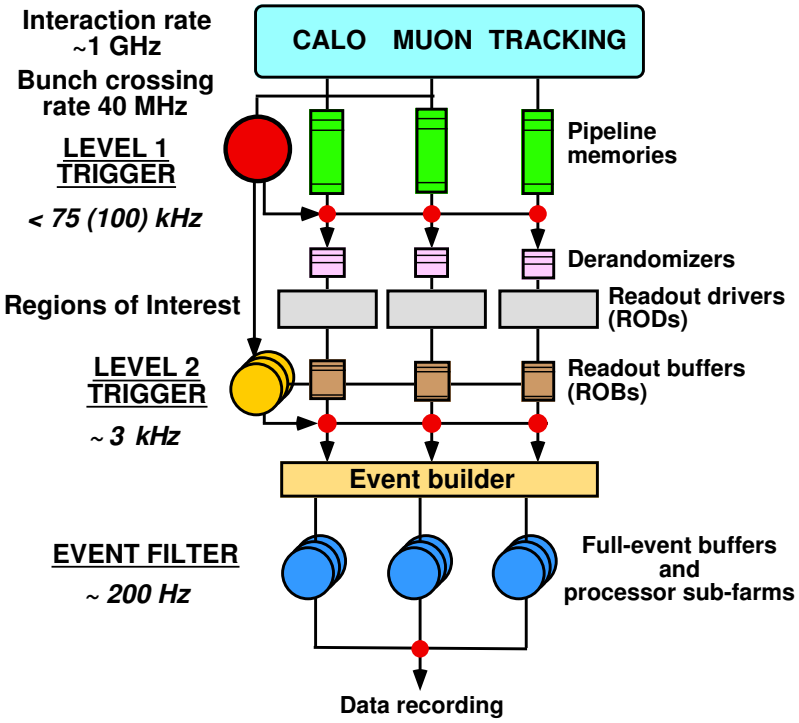


Figure 3.9.: Schematic view of the ATLAS trigger system taken from Ref. [95] with numbers adapted from Ref. [96].

3.2.5. Luminosity Measurements

There are two primary detectors, namely LUCID and BCM used to determine bunch-by-bunch luminosity measurements. LUCID is a Cherenkov detector consisting of sixteen aluminium tubes filled with C_4F_{10} gas surrounding the beam pipe at a distance of $z = \pm 17$ m on each side of the interaction point of the ATLAS detector. It is specifically designed to monitor

the inelastic collision rate by counting the number of charged particles produced in each collision within the LUCID acceptance. This information can be used to determine the relative luminosity in the ATLAS experiment according to Eq. 2.29.

The Beam Conditions Monitor (BCM) [97] comprises four small diamond sensors located on each side of the ATLAS detector at $z = \pm 1.84$ m. The sensors are arranged around the beam pipe in a cross pattern. The BCM is primarily designed to detect early signs of beam instabilities and to issue a beam-abort request in case beam losses start to risk damage to ATLAS detectors. Additionally, the fast readout of the BCM provides a measurement of bunch-by-bunch luminosities in the ATLAS detector at $|\eta| = 4.2$ by counting the number of collisions. It provides an independent cross-check to the results obtained with the LUCID detector.

4

ATLAS Detector Simulation and Event Reconstruction

The aim of any detector simulation is essentially to obtain predictions from Monte Carlo simulation taking detector effects into account. This allows for a direct comparison (on a statistical basis) of simulated events to data events measured with the ATLAS detector. Both simulated and real data events can then be passed to the same ATLAS trigger, reconstruction and identification algorithms. The reconstructed and identified objects, obtained both in data and from Monte Carlo simulation, serve as basis for physics analysis as presented in Chapters 5 and 6.

Two different types of ATLAS detector simulation are presented in Section 4.1, followed by a discussion of the event reconstruction and object identification algorithms in Section 4.2 which are developed by the ATLAS collaboration.

4.1. Event Simulation

This section comprises a brief summary of two different approaches, called the full detector simulation and ATLFAST II, to provide a simulation of the ATLAS detector effects. A detailed description can be found in Ref. [98].

4.1.1. Full Detector Simulation

Events obtained from Monte Carlo event generators, as described in Section 2.4.6, need to be modified such that detector effects are taken into account. Therefore, a dedicated simulation program, carrying events from the event generation stage to an output format which is identical to the one of the true data recorded by the ATLAS detector, is integrated into the ATLAS software framework ATHENA [99]. The generated events are propagated through the full ATLAS detector using the GEANT4 simulation toolkit [100]. An exact description of the material distribution in the ATLAS detector is used in case of the full detector simulation to obtain a detailed description of the detector response. All particles are propagated through the detector using an accurate physics model of particle interactions in the detector material, followed by a simulation of the energy measurement process and the electronics used for read-out. Such a detailed simulation is able to give very precise results. However, because

of the complicated detector geometry and the detailed physics description of interactions, the simulation time for an average ATLAS event is in the order of several minutes. Almost 75% of the full simulation time is spend simulating electromagnetic particles traversing the calorimetry.

4.1.2. Fast Detector Simulation ATLFAST II

ATLFAST II is a fast simulation developed to provide a large number of simulated events to supplement full simulation studies. The aim is to simulate events as fast as possible while still reaching for a sufficiently good detector description. There is a fast tracking simulation for the inner detector and muon system simulation available as well as a fast calorimeter simulation. By default, ATLFAST-II uses full simulation for the inner detector and muon system and the fast calorimeter simulation. Instead of simulating the particle interactions with the detector material, the energy of single particle showers is deposited directly by the fast calorimeter simulation using parametrisations of the longitudinal and lateral energy profile. The parametrisation of electromagnetic and hadronic showers is obtained from fully simulated single photon and charged pion events, respectively. A fine binning of the parametrisation in the particle energy (between 200 MeV and 500 GeV) and pseudorapidity ($|\eta| < 5.0$) is needed to allow for an accurate description of the distribution of active and inactive material in the calorimeter. The shower development depends strongly on the position of the starting point in the calorimeter. Thus, also a binning in the longitudinal shower depth is included in the parametrisation. This parametrisation approach leads to a typical reduction in computing time per event by a factor of 10 to 20 compared to full simulation. ATLFAST II provides the same output format as the full simulation and the simulated events can be passed through the standard ATLAS object reconstruction and identification chain.

4.2. Event Reconstruction

The physics events recorded by the ATLAS detector passing the online trigger selection are stored on tape for later processing. The event reconstruction and object identification algorithms are implemented in the software framework ATHENA [100]. This section describes the methods used to transform the digital output of the ATLAS detector into physics objects including the four-momenta and the complete kinematics of the event.

Dedicated reconstruction algorithms were developed to ensure high efficiency and accuracy for the reconstruction and identification of each type of physics object. The measurements presented in this thesis rely on tracks, electrons, muons, jets and missing transverse energy. The individual algorithms will be described in the following sections. Special emphasis is placed on the reconstruction of electrons to set the stage for a detailed study on the electromagnetic showers reconstructed in the ATLAS calorimeter system presented in Chapter 5.

4.2.1. Track Reconstruction

The reconstruction of charged particle tracks is important as input for the reconstruction of e.g. muons, electrons, τ-leptons and jets. The track reconstruction is basically divided into three stages and described in detail in Ref. [101]. Charged particle tracks can be measured up to $|\eta| < 2.5$ in the ATLAS detector.

In a first step, the raw data from pixel and double-sided SCT detectors are converted into clusters which are transformed into three-dimensional space-points using the known position of the modules. The TRT raw timing information is translated into calibrated drift-circles.

The next step involves running various track-finding algorithms using an inside-out scheme which exploits the high granularity of the pixel and SCT detectors to find prompt tracks originating from the vicinity of the interaction region. Track seeds are formed from a combination of space-points in the three pixel layers and the first SCT layer. These seeds are extended throughout the SCT to form track candidates which are used in a first fit. Ambiguities in the cluster-to-track association are resolved and fake tracks are rejected by applying quality cuts. The selected tracks are extended into the TRT and the measured drift-circles are associated to the track candidates in a road around the extrapolation. Finally, the extended tracks with a transverse momentum above 100 MeV passing certain quality cuts are refitted with the full information of all three detectors [85]. An additional track finding strategy is called back-tracking and starts from unused track segments in the TRT which are extended back into the SCT and pixel detectors. This leads to an improvement of the tracking efficiency for secondary tracks from conversions or decays of long-lived particles.

The third step is a post-processing stage in which a dedicated vertex finder is used to reconstruct primary vertices of the events.

A reasonably good description of the first collision data at $\sqrt{s} = 7$ TeV by the Monte Carlo simulation was found for track seed and track candidate properties [102].

4.2.2. Primary Vertex Reconstruction

The knowledge of the position of the primary interaction point of the proton-proton collision is important to achieve a precise measurement of the track parameters of charged particles. Furthermore, the number of reconstructed vertices provides a direct measure of in-time pile-up on an event-by-event basis due to the high time resolution of the silicon detectors. In the analysis presented in this thesis, the number of primary vertices is used to adjust the pile-up conditions in Monte Carlo simulation to match those in data (see Section 6.3.4).

The reconstruction of primary vertices is organised in two steps [103]. The primary vertex finding algorithm is run first followed by the vertex fitting algorithm. The reconstructed tracks are associated to the vertex candidates by looking for a global maximum in the distribution of z-coordinates of the tracks, computed at the point of closest approach at the beam spot center. The adaptive vertex fitter is designed to reconstruct the vertex position and its corresponding error matrix [104]. The algorithm is a robust χ^2-based fitting algorithm which deals with outlying track measurements by down-weighting their contribution to the overall vertex χ^2. The beam spot position is used as a three-dimensional constraint in the fit as well. This step includes a refitting of the associated tracks constraining them to originate from the reconstructed interaction point. In the first iteration, exactly one vertex is fitted from all the pre-selected tracks. Afterwards, tracks incompatible with the vertex by more than approximately seven standard deviations are used to seed a new vertex. This process is repeated until all available tracks are used or no new vertex seed can be created. The primary vertex reconstruction efficiency is measured in data to be close to 100% for events with at least three selected tracks with $p_{\mathrm{T}} > 100$ GeV [105].

4.2.3. Muon Reconstruction

Muons are the only charged particles which traverse the calorimeter without loosing most of their energy. This unique characteristic is used to detect them in the muon chambers. There are various strategies available to reconstruct and identify muons in the ATLAS detector for events that have passed the trigger selection.

Muon Trigger

The general structure of the trigger system of the ATLAS experiment is described in Section 3.2.4. The muon trigger system consists of three levels [106]. The Level-1 (L1) hardware trigger is based on signals in the muon trigger chambers using Resistive Plate Chambers (RPC) in the barrel ($|\eta| < 1.05$) and Thin Gap Chambers (TGC) in the endcaps ($1.05 < |\eta| < 2.4$). The trigger searches for patterns of hits consistent with high-p_{T} muons originating from the interaction point. In addition, the L1 trigger performs a rough estimate of the position coordinates in η and ϕ of the muon candidates which defines the 'regions of interest' (RoIs) (see Section 3.2.4).

The L2 trigger uses refined p_{T} measurements exploiting the full granularity of the detector within the RoI's. Furthermore, the hits measured in the Monitored Drift Tubes (MDT) of the Muon Spectrometer (MS) stations are used in a pattern recognition algorithm. A linear track fit is performed and the tracks found by the L2 muon stand-alone algorithm can be combined with the reconstructed tracks in the inner detector (ID) by a fast track combination algorithm. The refined transverse momentum is calculated as a weighted average of the ID and L2 muon stand-alone transverse momenta.

The muon event filter uses offline analysis procedures and has access to the full event information to confirm or discard the L2 candidates. There are two alternative offline algorithms implemented. One of them is an 'outside-in' strategy starting from tracks reconstructed in the MS extrapolating them back to the Interaction Point (IP), while the complementary 'inside-out' algorithm starts from the inner detector tracks and performs muon identification outward.

During the 2011 data taking period, the p_{T} threshold of the lowest unprescaled single muon trigger chains was kept at 18 GeV. The corresponding trigger chain called $mu18\_medium$ is used in the analysis presented in Chapter 6.

Muon Trigger Efficiency

The single muon trigger performance is evaluated primarily with a tag-and-probe method using $Z \to \mu^+\mu^-$ decays [107]. Muon pairs are selected by requiring two oppositely charged muons with a dimuon invariant mass near the mass of the Z boson. The 'tag' muon must pass certain isolation criteria and must have an associated event filter (EF) trigger track. The efficiency of a muon trigger chain with respect to offline isolated muons is the fraction of probe muons which have an associated EF trigger track in a cone of $\Delta R < 0.15$ around the probe muon. The efficiency for muon triggers is calculated separately in the barrel and endcap regions. This is due to the geometric coverage of the L1 muon trigger which is about 99% in the endcap regions and about 80% in the barrel region. This is caused by a crack around $\eta = 0$ which is necessary to provide space for services of the ID and the calorimeter.

Figure 4.1 shows measured efficiencies extracted from data and Monte Carlo simulation in the barrel and endcap regions as a function of muon p_{T} for the the *outside-in* and the *inside-out* algorithms. The efficiencies in the plateau region are about 70% in the barrel and 90% in the endcap regions. The triggers are fully efficient for $p_{\mathrm{T}} > 20$ GeV and the overall efficiency in data in the plateau region is well described by the Monte Carlo simulation. The p_{T} threshold is placed at 25 GeV for offline muons used in the analysis presented in Chapter 6.

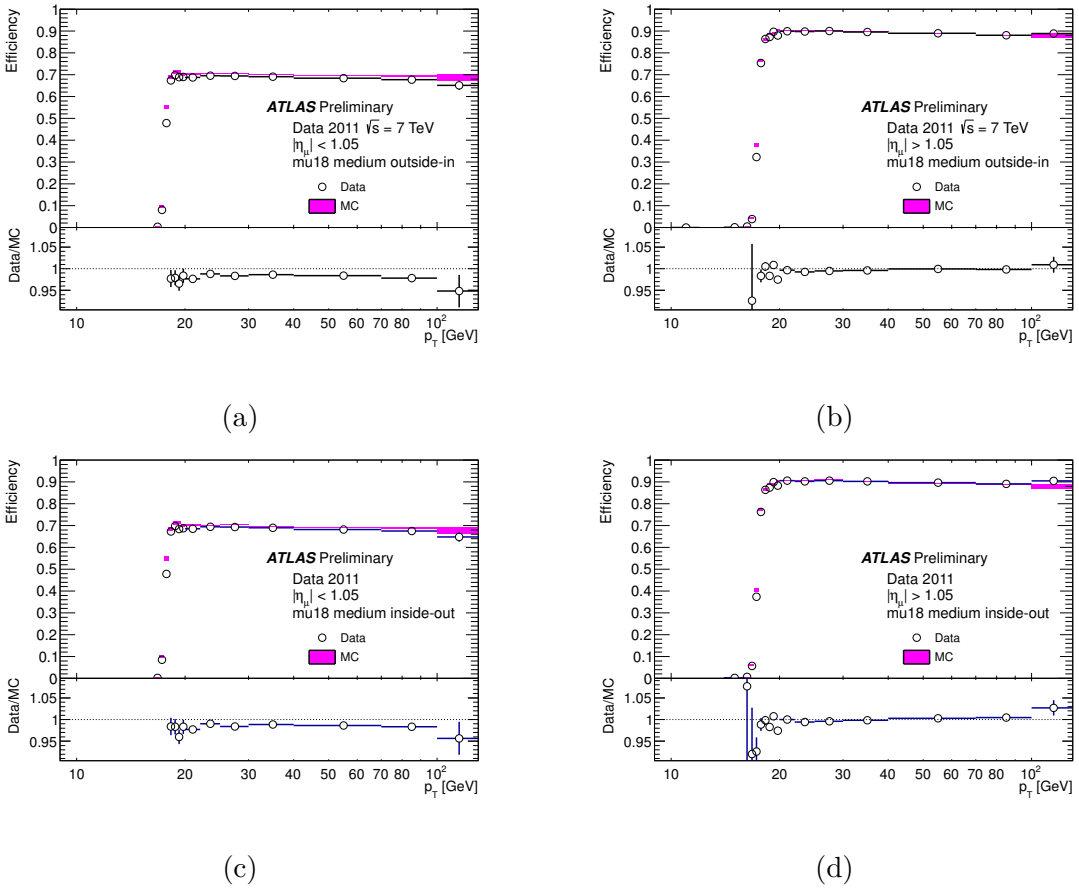


Figure 4.1.: Efficiencies derived from $Z \to \mu^+\mu^-$ events for the $mu18\_medium$ trigger chains in terms of the offline reconstructed muon p_{T}. In the upper part of each plot, the circles and the rectangles show data and Monte Carlo simulation, respectively. The lower part of each plot shows the ratio of data efficiencies compared to those of the Monte Carlo simulation. The vertical error bars and the vertical width of the rectangles in each plot represent the statistical uncertainties. (a) and (b) show efficiencies of the triggers with the muon spectrometer track based algorithm at EF level (*outside-in*) in the barrel and endcap regions, respectively. (c) and (d) show the trigger efficiencies using the inner detector track based algorithm at EF level (*inside-out*) in the barrel and endcap regions, respectively. The efficiencies include the geometric acceptance of the L1 trigger chambers. The amount of data used for the trigger efficiency measurements correspond to an integrated luminosity of 2.8 fb^{-1} recorded by the ATLAS experiment in proton-proton collisions at $\sqrt{s} = 7$ TeV in 2011 [107].

Muon Reconstruction and Identification

The ATLAS detector was designed for efficient muon detection and dedicated reconstruction algorithms are available. There are multiple ways to classify the identified muons [108]:

- **Stand-alone (SA) muons** are reconstructed from tracks solely measured in the muon spectrometer. The direction of flight and impact parameter with respect to the interaction point of the muon are determined by extrapolating the MS track back to the beam axis. In order to give an accurate measurement the impact of multiple scattering and energy loss of the muon in the calorimeter is taken into account.
- **Combined (CB) muons** are based on tracks reconstructed independently in the inner detector and the muon system. The matching of the tracks is performed by a statistical combination algorithm using a χ^2- minimisation. This algorithm tests the compatibility of the tracks based on the reconstructed track parameters and their covariance matrices. The combined measurement improves the momentum resolution and allows the rejection of muons from secondary interactions as well as the ones originating from decays in flight of pions or kaons.
- **Segment tagged (ST) muons** are identified in case the tracks measured in the ID and extrapolated to the MS are associated to straight track segments in the precision muon chambers. The ST algorithm is useful to recover low-p_{T} muons which only reach the inner layer of the muon chambers.

The CB muons are the candidates with highest purity. The ATLAS experiment uses two algorithm chains for reconstructing ST and CB muons, namely Staco [109] (chain 1) and MuID [110] (chain 2). These algorithms follow different pattern recognition strategies to define the CB muons. Staco Combined muons are used in the analysis presented in this thesis.

Muon Reconstruction Efficiency

The ability to reconstruct CB muons depends on the η and ϕ coordinates of the muon and is mainly restricted by the possibility to form the independently reconstructed MS track. This is mostly limited in the following two regions: at $\eta \sim 0$, where the MS is only partially equipped with muon chambers to leave space for services of the ID and the calorimeter, and in the transition region between the barrel and the endcaps at $|\eta| \sim 1.2$. Here, only one muon chamber is traversed by muons in the MS due to staged endcap chambers and therefore no stand-alone momentum measurement is available.

The reconstruction efficiency for CB muons is the product of the muon reconstruction efficiency in the ID, the reconstruction efficiency in the MS and the matching between the ID and MS measurements. The full reconstruction efficiency can be determined by measuring the individual efficiencies.

That is done with a tag-and-probe method using $Z \to \mu^+\mu^-$ decays selected by requiring two oppositely charged isolated tracks with a dimuon invariant mass close to the mass of the Z boson. One of the tracks must be identified as a CB muon and this track is called *tag* muon. The other track, called *probe* muon must be either a stand-alone muon if the ID efficiency is to be measured or an inner detector track to measure the MS and matching efficiency.

The ID reconstruction efficiency is the fraction of stand-alone muon probes which can be associated to an inner detector track. Accordingly, the MS and matching efficiency is defined

as the fraction of ID probes which can be associated to an inner detector track. The average efficiencies and scale factors are summarised in Table 4.1. The scale factors are very close to unity due to an excellent agreement between data and Monte Carlo simulation.

Table 4.1.: Average muon reconstruction efficiencies in Monte Carlo simulation and data [111].

Data:	
MS and Matching Efficiency:	96.93 ± 0.12 %
Inner Detector Efficiency:	$98.84^{+0.14}_{-0.16}$ %
Monte Carlo simulation:	
MS and Matching Efficiency:	96.83 ± 0.04 %
Inner Detector Efficiency:	99.02 ± 0.04 %
Scale Factor:	
MS and Matching Efficiency:	1.001 ± 0.001 %
Inner Detector Efficiency:	0.998 ± 0.002 %

The combined muon reconstruction efficiency as a function of η for muons with $p_{\mathrm{T}} > 20$ GeV is shown in Fig. 4.2. The distribution illustrates that - despite the excellent overall agreement - there is a loss of efficiency around $|\eta| < 0.1$ and $|\eta| \sim 1.2$ due to detector specific reasons mentioned above. The deviations between data and Monte Carlo simulation are noticeable in this regions. The combined scale factors to correct Monte Carlo simulation to match data are determined in bins of p_{T} and η of the reconstructed muons.

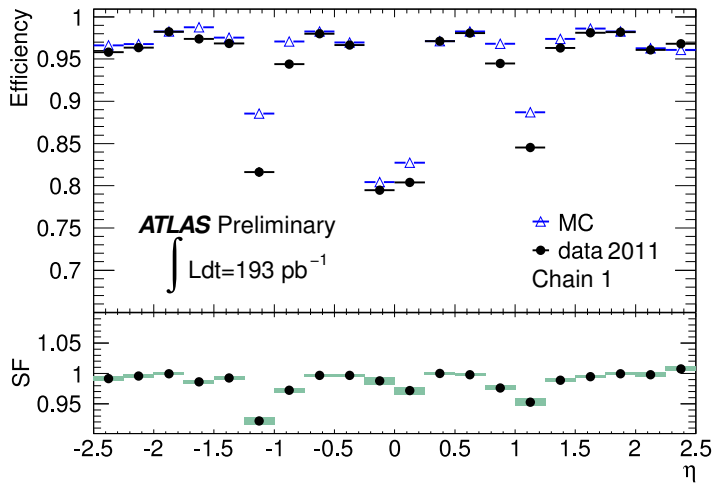


Figure 4.2.: Combined muon reconstruction efficiency with respect to the inner tracking efficiency as a function of the pseudorapidity of the muon for muons with $p_{\mathrm{T}} > 20$ GeV. The lower part of the plot shows the ratio between the measured and predicted efficiencies [112].

Muon Isolation

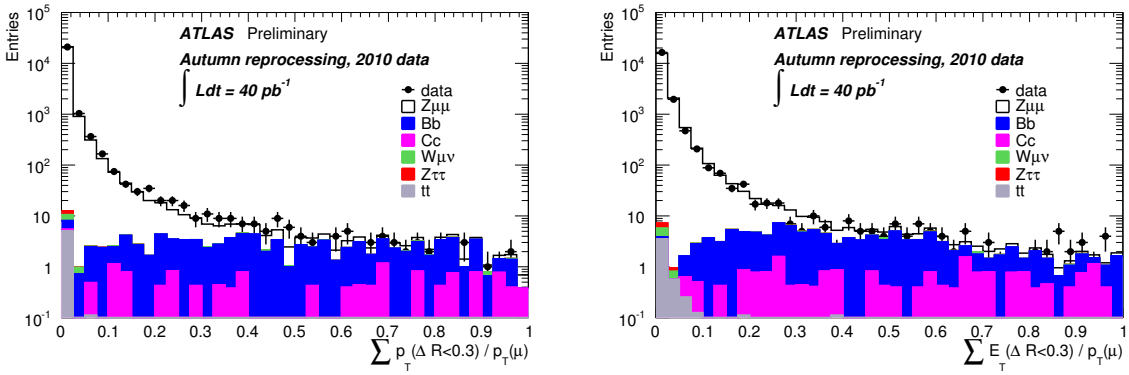
A powerful quantity to reject muons originating from hadron decays is the isolation. Muons from decays of W and Z bosons are not accompanied by other particles in close vicinity. On the contrary, muons from heavy-quark decays or hadron decays are accompanied by additional

particles from the same decay. Both, calorimeter-based and tracking-based isolation variables utilise this fact in their conception.

For the tracking-based isolation the scalar sum of the transverse momenta of the tracks in a cone of R_{iso} around the muon, excluding the muon candidate track itself, is used.

The calorimeter-based isolation is calculated by adding up the energy measured in the calorimeter cells within a cone of R_{iso} around the muon, excluding the energy in the cells associated to the muon after corrections for leakage and pile-up effects.

The measured relative isolation variables using a cone size of $R_{iso} = 0.3$ divided by the p_{T} of the probe muons are compared to the Monte Carlo prediction in Fig. 4.3. The isolation distributions are well modelled by the simulation which leads to reliable prediction of the isolation efficiency defined as the fraction of probe muons passing a given set of isolation cuts. Muon candidates selected in the analysis presented in Chapter 6 are required to be isolated using both relative tracking and calorimeter isolation variables with a cone size of $R_{iso} = 0.3$.



(a) (b)

Figure 4.3.: Comparison of the measured track and calorimeter isolation distributions of the probe muon. (a) Track isolation $\sum p_{\mathrm{T}}(\Delta R < 0.3)/p_{\mathrm{T}}(\mu)$ and (b) calorimeter isolation $E_{\mathrm{T}}(\Delta R < 0.3)/p_{\mathrm{T}}(\mu)$ [108].

Muon Energy Scale and Momentum Resolution

The energy scale and resolution in data is extracted from the width of the dimuon mass distribution in $Z \to \mu^+\mu^-$ decays [113]. The relative resolution of the momentum measurement depends on different effects related to the amount of material that the muon traverses, the spatial resolution of the individual track points and the knowledge of internal alignment of the two subsystems. The ATLAS MS is designed to provide an uniform momentum resolution as a function of the pseudorapidity. The dimuon invariant mass distribution obtained for combined muons is shown in Fig. 4.4. The distribution is fitted using a convolution of the Z boson lineshape and two Gaussian functions modelling the detector resolution effects. The value of the fitted invariant mass at the Z boson pole is sensitive to possible shifts in the momentum scale. Therefore, the accuracy of the momentum scale is probed by measuring the average deviation of the measured invariant mass from the Z boson mass world average value. The resulting resolutions obtained by performing the fit on data and Monte Carlo simulation in different pseudorapidity regions and averaged over all muon p_{T} values are shown in Fig. 4.4.

Data and Monte Carlo prediction show a similar behaviour but the performance in simulation is better compared to data. In order to account for this difference and the observed shift in energy scale, the transverse momentum is smeared in the Monte Carlo simulation.

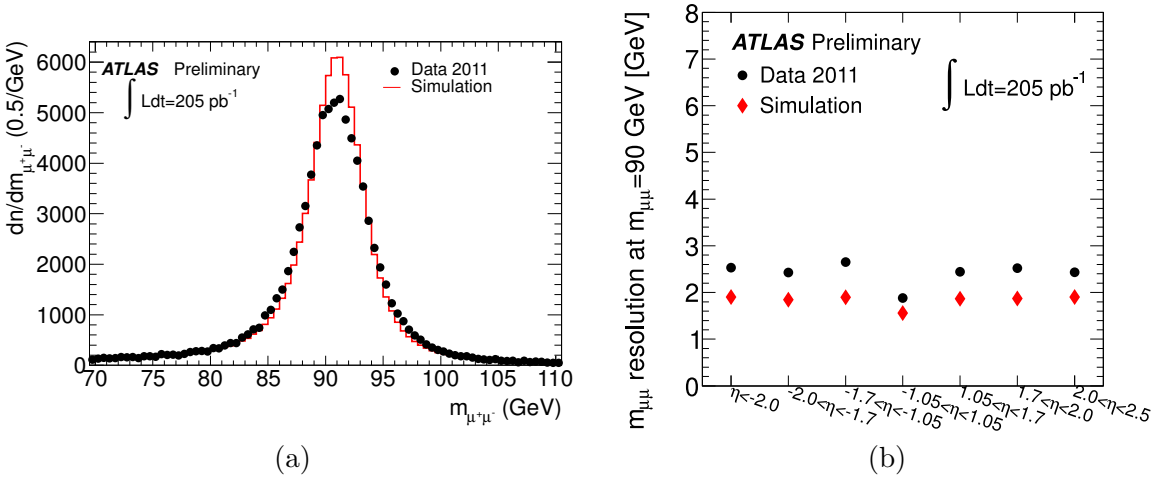


Figure 4.4.: (a) Dimuon invariant mass distribution for oppositely charged muon pairs with transverse momentum above 20 GeV, $|\eta| < 2.5$ and calorimeter isolation (sum of calorimeter cell energies < 2 GeV in a cone of $\Delta R = 0.3$) and (b) dimuon mass resolution of combined muons in different pseudorapidity regions (taken from Ref. [112]).

4.2.4. Electron Reconstruction

Electrons are reconstructed using information from both the calorimeter and the inner detector. The algorithm is designed to provide a high and uniform efficiency to detect electrons over a wide range in p_{T} and η and a powerful rejection against jets and background electrons. The background to isolated electrons include electrons from hadronic jets, heavy flavour decays and Dalitz decays or photon conversions originating from neutral pion decays.

The following sections will address the individual steps which belong to the reconstruction and classification of electrons, namely trigger, reconstruction, identification and isolation requirements. The corresponding efficiency of each step measured based on Z, $W^{\pm}$ and J/ψ decays.

Electron Trigger

The ATLAS trigger system comprises three levels as described in Section 3.2.4. Most of the energy of the electromagnetic showers originating from electrons and photons is deposited in the electromagnetic calorimeter. The first trigger stage (L1) performs a fast selection based on calorimeter information with reduced granularity of so called Trigger Towers covering a region of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. Regions of interest (RoI) are formed for each identified electromagnetic object based on the Trigger Towers. The transverse energy E_T of the electromagnetic cluster is computed with a precision of 1 GeV. A given trigger is satisfied if the core region of the central 2×2 Trigger Towers contains two adjacent towers with a combined energy deposit passing the threshold.

The L2 trigger stage starts from the RoIs identified by the L1 and uses the full granularity of the calorimeter. Fast calorimeter and dedicated tracking reconstruction algorithms are deployed and the L2 decision includes a track-to-cluster matching. In order to comply with

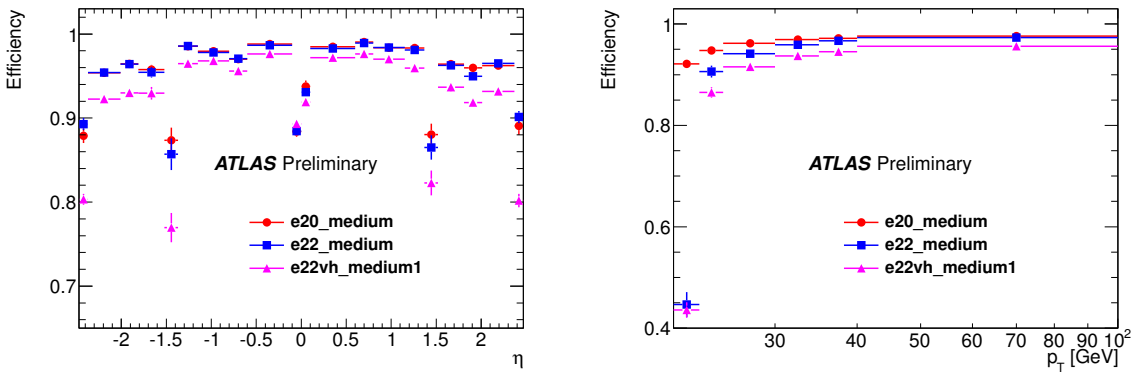
the strict timing requirements, only the second layer of the electromagnetic (EM) calorimeter is used to find the cell containing the largest energy deposit. This cell is used as seed to the cluster algorithm which uses the same definition as the offline electron reconstruction: 3×7 cells in the barrel ($|\eta| < 1.4$) and 5×5 cells in the endcaps ($1.4 < |\eta| < 2.47$).

At the EF stage, the offline reconstruction and identification algorithms, which will be described in detail in the following sections, are used. Looser selection cuts are applied in order to remain fully efficient for offline identified objects.

Three operation points are defined for the standard offline selection with increasing levels of background rejection called *loose, medium* and *tight*. A re-optimised menu was defined for electrons, referred to as the *plus-plus* menu, providing three additional operating points (*loose++, medium++* and *tight++*) with improved performance with respect to the standard menu for a higher pile-up environment. The electron selections called *loose1, medium1* and *tight1* include tighter cuts on cluster quantities and tracks and were implemented to trigger on *plus-plus* offline electrons.

Some of the trigger thresholds had to be adjusted from September 2011 on to keep the overall L1 acceptance rate within the limit. An optimisation of $\eta-$dependent energy thresholds and a veto on hadronic energy of more than 1 GeV deposited in the hadronic layers behind the EM clusters were applied. The corresponding EF triggers have the letter *vh* added in the trigger name (e.g.*e22vh_medium1*).

The EF triggers used in the analysis presented in Chapter 6 are called *e20_medium*, *e22_medium* and *e22vh_medium1*. The EF transverse momentum threshold for the lowest unprescaled single electron trigger was raised from 20 GeV to 22 GeV when the instantaneous luminosity exceeded $2 \times 10^{33} cm^{-2} s^{-1}$. Later on at L2, the re-optimised electron identification criteria were applied, moving from *medium* to *medium1*.



(a) (b)

Figure 4.5.: Event Filter trigger efficiencies as functions of the offline*medium++* electron p_{T} and η for the *e20_medium*, *e22_medium* and *e22vh_medium1* triggers. The vertical bars represent statistical and total systematic uncertainties. Integrated luminosities of 1.8 fb^{-1}, 0.6 fb^{-1} and 2.5 fb^{-1} were recorded by the ATLAS experiment when lowest unprescaled triggers were *e20_medium*, *e22_medium* and *e22vh_medium1*, respectively [114].

Electron Trigger Efficiency

The efficiencies of the High Level Trigger (HLT) electron selection is measured with a tag-and-probe method using $Z \to e^+e^-$ decays. The tag electron is defined as the offline electron

matching to an online electron passing the lowest unprescaled single electron trigger within $\Delta R < 0.15$. In addition, the tag electron is required to have $p_{\mathrm{T}} > 25$ GeV, to satisfy the *tight++* offline electron identification and to be reconstructed within $|\eta| < 2.47$ excluding the transition region between the barrel and the endcap regions (see Section 3.2.2). A second electron is classified as probe if the invariant mass of the electron pair is close to the mass of the Z boson (80 GeV $< m_{ee} <$ 100 GeV). The trigger efficiency is defined as the fraction of probe electrons matching an online electron passing the trigger selection at the EF level. The efficiencies of the *e20_medium*, *e22_medium* and *e22vh_medium1* triggers measured relative to *medium++* offline electrons are shown in Fig. 4.5. The plateau of the trigger efficiency is reached for electrons with $p_{\mathrm{T}} > 35$ GeV. Inefficiencies of the triggers arise mainly from differences in the resolution of reconstruction and identification variables at the HLT with respect to the offline algorithm.

The efficiencies obtained from Monte Carlo simulation of $Z \to e^+e^-$ events agree with the ones measured in data within 1-2%. The data to Monte Carlo simulation ratios (hereby called scale factors) are derived to correct the Monte Carlo predictions for the small residual difference. The efficiencies and scale factors are summarised in Table 4.2. These scale factors are used in the analysis presented in Chapter 6.

Table 4.2.: Trigger efficiencies (in percent) and scale factors (data to Monte Carlo simulation efficiency ratios) including statistical and systematic uncertainties as function of *tight++* offline selection electrons, in different electron p_{T} ranges, for the *e*20*_medium*, *e*22*_medium* and *e*22*vh_medium*1 triggers. Statistical and systematic uncertainties are summed in quadrature (adapted from Ref. [114]).

trigger	offline selection	p_{T} bin [GeV] (21, 23)	(23, 25)	(25, 30)	(30, 35)	(35, 40)	(40, ∞)
		Trigger efficiencies in percent					
e20_medium	*tight++*	93.52 ± 0.70	95.45 ± 0.25	96.82 ± 0.38	97.48 ± 0.28	97.69 ± 0.10	98.02 ± 0.22
e22_medium	*tight++*	-	91.37 ± 1.08	94.90 ± 0.54	96.58 ± 0.50	97.19 ± 0.16	97.75 ± 0.08
e22vh_medium1	*tight++*	-	88.09 ± 1.38	92.92 ± 0.83	94.86 ± 0.62	95.63 ± 0.33	96.49 ± 0.10
		Scale factors					
e20_medium	*tight++*	0.985 ± 0.006	0.991 ± 0.003	0.988 ± 0.002	0.992 ± 0.001	0.992 ± 0.002	0.995 ± 0.001
e22_medium	*tight++*	-	0.981 ± 0.019	0.985 ± 0.008	0.992 ± 0.007	0.998 ± 0.002	1.002 ± 0.001
e22vh_medium1	*tight++*	-	0.958 ± 0.007	0.966 ± 0.002	0.972 ± 0.002	0.978 ± 0.005	0.985 ± 0.003

Electron Reconstruction and Identification

An efficient and optimised reconstruction algorithm is essential to extract a pure signal of electrons above the residual background from jets misidentified as electrons. The offline reconstruction of electrons is initiated by cluster-finding in the electromagnetic calorimeter followed by the association of reconstructed tracks in the inner detector. The algorithms are described in detail in Ref. [7] and a brief summary is given in this section.

The reconstruction of electrons, as described in the following, is restricted to objects within the acceptance range of the inner detector ($|\eta| < 2.5$) [1]. The electron reconstruction algorithm starts from clustering energy deposits in neighbouring calorimeter cells. The clusters are formed by a *sliding-window* algorithm based on towers of 3×5 cells corresponding to the size of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, the granularity of the second calorimeter layer. A window with fixed size of 5×5 towers is moved across the ($\Delta\eta$-$\Delta\phi$)-plane and seed clusters are formed

[1] There is a possibility to identify so-called *forward electrons* outside the tracking coverage up to $|\eta| < 4.9$ (details in Refs. [7, 115]) but they are not subject to this thesis and will be omitted.

Table 4.3.: Definition of variables used for *loose*, *medium* and *tight* electron identification cuts for the central region ($|\eta| < 2.47$) of the detector [7].

Type	Description	Name
Loose selection		
Acceptance	$\|\eta\| < 2.47$	
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ and $\|\eta\| > 1.37$)	R_{had1}
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| > 0.8$ and $\|\eta\| < 1.37$)	R_{had}
Middle layer of EM calorimeter	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centered at the electron cluster position	R_η
	Lateral width of the shower	$w_{\eta 2}$
Medium selection (includes loose)		
Strip layer of EM calorimeter	Total shower width	w_{stot}
	Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	E_{ratio}
Track quality	Number of hits in the pixel detector (≥ 1)	n_{pixel}
	Number of total hits in the pixel and `SCT` detectors (≥ 7)	n_{Si}
	Transverse impact parameter ($\|d_0\|$ <5 mm)	d_0
Track–cluster matching	$\Delta\eta$ between the cluster position in the strip layer and the extrapolated track ($\|\Delta\eta\| < 0.01$)	$\Delta\eta$
Tight selection (includes medium)		
Track–cluster matching	$\Delta\phi$ between the cluster position in the middle layer and the extrapolated track ($\|\Delta\phi\| < 0.02$)	$\Delta\phi$
	Ratio of the cluster energy to the track momentum	E/p
	Tighter $\Delta\eta$ requirement ($\|\Delta\eta\| < 0.005$)	$\Delta\eta$
Track quality	Tighter transverse impact parameter requirement ($\|d_0\|$ <1 mm)	d_0
TRT	Total number of hits in the TRT	n_{TRT}
	Ratio of the number of high-threshold hits to the total number of hits in the TRT	f_{HT}
Conversions	Number of hits in the B-layer (≥ 1)	n_{BL}
	Veto electron candidates matched to reconstructed photon conversions	

if the transverse energy inside the window is above 2.5 GeV and is a local maximum. The efficiency for cluster reconstruction tested on Monte Carlo simulation is 100% for electrons with E_{T} above 15 GeV.

The electron candidates are defined by the existence of one or more reconstructed tracks matching to the cluster. The matching is performed on tracks extrapolated to the middle layer of the calorimeter. The effect of losses due to bremsstrahlung is accounted for when testing the compliance of the track impact point and the cluster position. The size of the sign corrected $\Delta\phi$ window is enlarged to 0.1 on the side where the extrapolated track is bended while it traverses the magnetic field. The $\Delta\phi$ window is 0.05 on the other side and in addition the requirement $\Delta\eta < 0.05$ has to be satisfied for a successful match.

The electromagnetic cluster is rebuilt using a 3×7 (5×5) sliding window size in $\Delta\eta$-$\Delta\phi$ middle layer cell units in the barrel (endcaps). The cluster energy is determined taking into account the energy loss in the material in front of the EM calorimeter, the lateral energy leakage outside the cluster and the energy deposited beyond the EM calorimeter. The four-momentum of the electron is computed using both cluster and track information. The energy is given by the cluster energy and the η and ϕ directions are taken from the corresponding track parameters extrapolated to the vertex.

The baseline electron identification relies on a cut-based selection using variables that include

calorimeter, tracker and combined information. These variables provide a good separation between isolated electrons and jets faking electrons. Three reference sets of cuts are defined with increasing background rejection power: *loose*, *medium* and *tight* [115], where each set contains the preceding one. The expected jet rejection estimated from Monte Carlo simulation is of the order of 1/500, 1/5000 and 1/50000, respectively. The cuts are optimised in 10 bins of cluster η and 11 bins of cluster E_{T} in a range of 5 GeV to above 80 GeV.

- The ***loose*** identification requirements are based on shower shape variables of the second calorimeter layer and hadronic leakage variables.
- The ***medium*** identification requirements furthermore include variables from the EM calorimeter strip layer, track quality and tighter track-to-cluster matching constraints.
- The ***tight*** identification requirements add tighter track-to-cluster matching and track quality constraints. Photon conversions are rejected by requiring a hit in the inner-most tracking layer (B-layer). The `TRT` variables provide additional particle identification information and help to suppress hadron fakes further.

The list of all variables used in the three sets of cuts is shown in Table 4.3.

The efficiency for each set is studied on Monte Carlo simulation using a $Z \to e^+e^-$ inclusive sample. The rejection numbers for $E_{\mathrm{T}} > 20$ GeV are determined using a dedicated filtered dijet sample. In case of simulated events, the origin of each electron can be determined using truth information. The following categories are established to classify reconstructed electron candidates as:

– Isolated electrons in case they are matched to a true electron originating from W bosons or Z bosons.

– Non-isolated electrons if they are matched to a true electron originating from mesons or baryons containing b- or c-quarks.

– Background electrons in case they are matched to a true electron originating from photon conversions, Dalitz decays or the decay of light mesons.

– Hadron fakes if they are matched to true hadrons.

In order to understand the performance of the cut-based identification, it is worth to examine the fraction of surviving electron candidates at each selection level. It can be seen in Table 4.4 that hadron fakes are dominant at reconstruction level. However, after the tight selection, the fraction of non-isolated electrons is significantly increased. The total rejection of jets after applying tight identification requirements is predicted to be about 4.9×10^4 with an efficiency of $72.77 \pm 0.03\%$ [115].

Electron Identification Efficiencies

The measurements of electron selection efficiencies are based on tag-and-probe methods using $Z \to e^+e^-$, $W \to e\nu$ and $J/\psi \to e^+e^-$ decays. The $Z \to e^+e^-$ events provide a clean environment to study the performance while the statistical power of the results is improved using $W \to e\nu$ decays. $J/\psi \to e^+e^-$ decays allow to reach towards lower transverse energies.

The purpose of the tag-and-probe method is the selection of a clean and unbiased sample of *probe* electrons. Therefore, the selection cuts are primarily applied to the other object, the *tag* electron, in the event. In case of $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ events, a well-identified electron is required to tag the event while high missing transverse energy is used for $W \to e\nu$ decays.

Table 4.4.: Expected jet rejection and electron efficiencies for the three reference selections *loose*, *medium* and *tight*. The total jet rejection includes hadron fakes and background electrons from photon conversions and Dalitz decays. The last four columns give the fraction of surviving electron candidates in the dijet sample after each selection level (adapted from Ref. [115]).

Cuts	$E_{\mathrm{T}} > 20$ GeV					
	Efficiency (%)	Jet rejection	surviving candidates (%)			
	$Z \to e^+e^-$		iso	non-iso	bkg	had
Reco	96.25 ± 0.03	87.3 ± 0.1	0.0	1.0	26.0	73.0
Loose	94.68 ± 0.03	614.3 ± 1.5	1.0	2.0	58.0	39.0
Medium	89.61 ± 0.03	4435 ± 30	3.0	9.0	57.0	31.0
Tight	72.77 ± 0.03	$(4.9 \pm 0.1) \times 10^4$	16.0	39.0	22.0	23.0

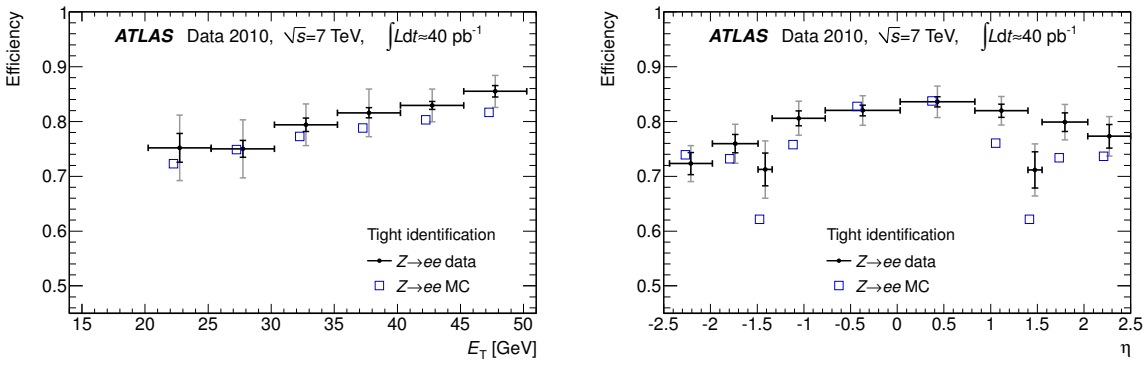
For the efficiency measurements presented here, the contamination of the probe sample by non-negligible background requires dedicated estimation techniques. The background in the $W \to e\nu$ measurement is evaluated with a template fit method based on isolation variables. For the $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ decay channels the reconstructed dielectron mass is the most efficient discriminating variable to estimate the signal and background contributions. A side-band method exploiting the invariant mass distributions of same-sign and opposite-sign pairs is used to subtract the background.

The identification efficiency is defined by the fraction of probe electron candidates after the background subtraction passing a specific set of cuts. The efficiencies measured from $Z \to e^+e^-$ events of tight electron identification requirements as a function of E_{T} and η are shown in Fig. 4.6. The tight cuts rely on tracking information and their performance is sensitive to interactions of the electrons with the ID material which results in a non-uniform efficiency versus η. The efficiencies measured in data and predicted by Monte Carlo simulation manifest some noticeable differences. Electron shower shapes were extracted from $Z \to e^+e^-$ decays in the course of this thesis to examine the discrepancies in electron identification variables. The extraction of the distributions is described in detail in Chapter 5.
Scale factors are estimated by dividing the measured efficiencies by the Monte Carlo predictions for $W \to e\nu$, $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ decays. The combined scale factors are applied to events containing electrons that are selected in the analysis presented in Chapter 6.

Electron Isolation

The electron identification described so far does not include dedicated isolation cuts. The tightness of the isolation criteria depends strongly on the specific analysis needs. Physics analysis including isolated leptons from decays of W or Z bosons usually have the advantage of small background expectations compared to channels with hadronic final states. Nevertheless, these channels can suffer from jet background processes including jets falsely reconstructed as lepton candidates or from real leptons from semi-leptonic heavy-quark decays that mimic the isolated leptons of the signal. As well as for muons, there are calorimeter- and track-based isolation variables which help to separate isolated electrons from non-isolated electrons and fakes. Detailed studies of the variables can be found in Refs. [115, 116].

The track-based isolation variable is sensitive to contributions from additional charged particles. Since all types of electrons can be accompanied by low quality tracks or tracks from



(a) (b)

Figure 4.6.: Efficiencies measured from $Z \to e^+e^-$ events and predicted by Monte Carlo simulation for tight identification criteria as function of (a) the reconstructed electron E_{T} and integrated over $|\eta| < 2.47$ excluding the transition region $1.37 < |\eta| < 1.52$ and (b) of η and integrated over $20 < E_{\mathrm{T}} < 50$ GeV [7].

pile-up events, the discriminator becomes more powerful by applying appropriate quality cuts to the tracks that should be considered. The scalar sum of the transverse momenta of selected tracks in a cone of size R_{iso} around the electron candidate defines the track-based isolation variable:

$$P_T^{cone\ R_{iso}} = \sum_{\substack{tracks \\ tracks \neq electron \\ \Delta R < R_{iso}}} p_T^{track}. \tag{4.1}$$

The electron track itself is excluded from the summation.

The calorimeter isolation variable is sensitive to energy contributions from additional charged and neutral particles. Isolated electrons deposit their energy in a narrow cluster and the amount of energy deposited in cells around the electron cluster gives a handle to distinguish signal electrons from non-isolated electrons and jets. The calorimeter isolation variable is computed from the energy deposited in a cone of radius R_{iso} around the electron candidate, where the energy in 5×7 cells of the electron cluster is removed,

$$E_T^{cone\ R_{iso}} = \Big(\sum_{\substack{cells \\ \Delta R < R_{iso}}} E_T^{cells} \Big) - E_T(5 \times 7\ cells). \tag{4.2}$$

It is found that the following two effects mainly contribute to the isolation determination: first, lateral leakage of the electron cluster into the isolation cone and second, the energy deposited from other collisions during and before the bunch crossing of interest (i.e. pile-up). Dedicated corrections were derived to account for this effects and are documented in Ref. [117].

While a larger cone size contains more energy in case of misidentified jets, a smaller cone size is more robust against energy deposits from pile-up events. A cone of $R_{iso} = 0.3$ around the electron is found to give the best trade-off between high discrimination power and robustness against pile-up [115]. Electron candidates selected in the analysis presented in Chapter 6 are required to be isolated using both relative tracking and leakage corrected calorimeter isolation variables with a cone size of $R_{iso} = 0.3$.

Electron Energy Scale and Resolution

The energy scale and resolution can be determined both from data using methods based on the dielectron mass of $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ decays and from E/p measurements in $W \to e\nu$ decays. The various methods are briefly described below and in detail in Ref. [7].

The well-known masses of the Z boson and the J/Ψ resonance can be used to test the linearity of the response of the EM calorimeter. Alternatively, the ratio of the energy measured by the EM calorimeter is compared to the momentum measured by the inner detector. This allows to determine the energy scale taking advantage of the larger statistics of $W \to e\nu$ decays. Any residual miscalibration in some η bin i is parametrised by

$$E_i^{meas} = E_i^{true}(1 + \alpha_i), \tag{4.3}$$

where E_i^{true} is the true electron energy, E_i^{meas} is the energy measured by the calorimeter including Monte Carlo based energy scale corrections and α_i measures the residual miscalibration in the η bin i. The α_i energy scale correction factors are determined by a fit to the invariant mass distribution or E/p distribution for $Z \to e^+e^-$, $J/\psi \to e^+e^-$ and $W \to e\nu$ events. The correction factors are extracted in 58 η bins over the full calorimeter coverage of $|\eta| < 4.9$ and are considered as the *baseline calibration* method. The resulting α values are within $\pm 2\%$ and $\pm 5\%$ in the barrel and forward regions, respectively.

The fractional energy resolution in the calorimeter is parametrised as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{4.4}$$

with η-dependent parameters. The parameter a is the sampling term and b is the noise term. Both were evaluated from Monte Carlo simulation. The constant term c is determined from fits to the simulated and measured invariant mass distributions of $Z \to e^+e^-$ decays using a Breit-Wigner distribution convoluted with a Crystal Ball function. The width of the Breit-Wigner distribution is fixed to the measured Z boson width and the experimental resolution is described by the Crystal Ball function. The resolution predicted by Monte Carlo simulation is slightly better compared to the measured Gaussian components of the experimental resolution. Therefore, the transverse momentum of electrons is smeared in Monte Carlo simulation to model the data distributions correctly.

4.2.5. Jet Reconstruction

Jets are observed as groups of topologically related energy deposits in the ATLAS calorimeters associated with tracks of charged particles measured in the inner detector. The uncertainty of the jet energy measurement is the dominant experimental uncertainty for various physics results including the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis presented in Chapter 6 of this thesis. There are different algorithms and methods available in the ATLAS experiment to reconstruct and calibrate jets. They are described in detail in Refs. [118,119]. Jets based on the $anti-k_T$ algorithm are used in the context of the analysis presented in Chapter 6 and will be described briefly below.

The reconstruction of jets starts from topological clusters [120] as inputs which are seeded by calorimeter cells with an energy content that is at least four times higher than the root-mean-square (RMS) of the noise distribution. Cells neighbouring the seed with a signal-to-RMS-noise-ratio of two are added iteratively and finally all nearest neighbour cells are added to the cluster without any threshold. The $anti-k_T$ algorithm is then applied to these clusters to form the jets [118].

The jet finding procedure is designed to provide jets whose shape is not influenced by soft radiation. One introduces distances d_{ij} between objects i and j and d_{iB} between object i and the beam (B). The clustering proceeds by identifying the smallest of the distances and if it is d_{ij} the objects i and j are recombined, whereas it is d_{iB} i is called a jet and removed from the list. The cycle goes on until no objects are left over. The $anti - k_T$ algorithm is characterised by the specific definition of the distance measures:

$$d_{ij} = min(1/p_{T,i}^2, 1/p_{T,j}^2)\frac{\Delta R_{ij}^2}{R^2} = min(1/p_{T,i}^2, 1/p_{T,j}^2)\frac{\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2}{R^2}, \tag{4.5}$$

where $p_{T,i}$ is the transverse momentum of object i and R is a fixed distance parameter which is chosen to be 0.4 for the jets used in this thesis.

The functionality of the algorithm may be understood more intuitively by considering the following example of an event that has a few well-separated particles with high p_{T} and many particles with low p_{T}. The d_{1j} between a hard particle 1 and a soft particle j is exclusively determined by the transverse momentum of the hard particle and the ΔR separation of the particles. The d_{ij} between two particles with low p_{T} and a similar separation will instead be much larger. Thus, particles with low p_{T} will tend to cluster with particles with high p_{T} long before they cluster among themselves. If there is a second hard particle in the event, two jets can be formed that will share a boundary which depends on the ratio of energy of the two particles. The behaviour of the jet algorithm for a parton-level event together with $\sim 10^4$ soft particles is illustrated in Fig. 4.7.

Within the ATLAS collaboration, several jet calibration schemes have been developed [85] with different levels of complexity and sensitivities to systematic effects that are complementary in their contribution to the jet energy measurement. The calorimeter have initially been calibrated using test-beam measurements with electrons to provide a correct response for electromagnetic showers [121]. Therefore, each calibration scheme starts from the measured calorimeter energy at the electromagnetic energy scale. Hadrons have a lower energy density and thus a lower response in the calorimeter compared to electrons. Energy corrections are applied to clusters classified as being of hadronic origin which are derived at first order from single pion Monte Carlo simulation. Furthermore, the topology of the calorimeter energy depositions are exploited to correct for calorimeter non-compensation (e.g. nuclear energy losses), for signal losses due to the noise thresholds used in the topological clustering and for the energy lost in non-instrumented regions. This way, the final jet energy calibration is a simple correction relating the calorimeter response directly to the true jet energy. The resulting jets are referred to as EM+JES jets which are used in the analysis presented in Chapter 6.

The uncertainty on the jet energy scale (JES) is derived combining measurements of single hadron response and single pion test-beam data. Additional uncertainties due to the amount of material of the ATLAS detector, the description of the electronic noise and Monte Carlo modelling used in the event are considered as well. The JES uncertainty varies as a function of the jet p_{T} and pseudorapidity and is about $\pm 4\%$ ($\pm 8\%$, $\pm 14\%$) for $anti - k_T$ jets with $p_{\mathrm{T}} > 20$ GeV in the central (endcap, forward) region [119]. The increase is dominated by the uncertainty on the modelling of the soft physics in the forward region which is estimated conservatively. A global JES uncertainty of about $\pm 12\%$ for jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$ is assigned to the jets used in the analysis in Chapter 6. Additional systematic uncertainties of about $\pm 5\%$ arise from the dependence of the jet response on the number of expected interactions per bunch crossing and on the jet flavour.

Uncertainties related to the jet energy resolution (JER) are obtained with an in-situ measurement of the jet response asymmetry in dijet events [122]. The JER uncertainty amounts

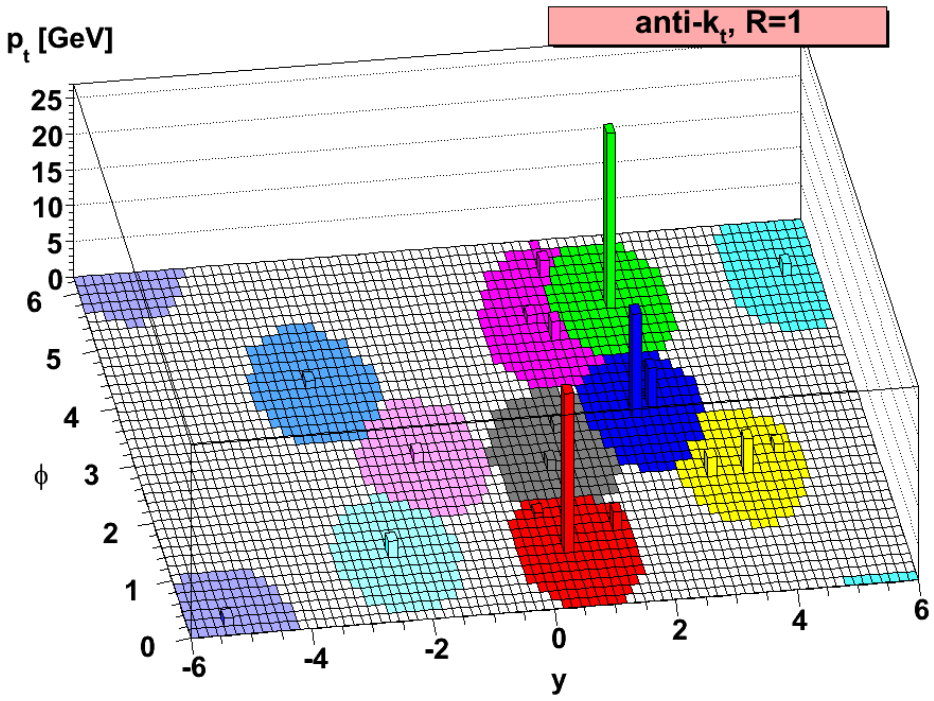


Figure 4.7.: A parton-level event generated with HERWIG [35], together with many soft radiations clustered with the $anti - k_T$ algorithm. The jet boundaries are resilient with respect to soft radiation but flexible with respect to hard radiation. The figure is taken from Ref. [118].

to about $\pm 10\%$ and the JER predicted by Monte Carlo simulation agrees with the data results within this uncertainty for jets with $p_{\mathrm{T}} > 20$ GeV. The jet resolution is smeared in the simulation using a parametrisation depending on the jet transverse momentum and pseudorapidity.

4.2.6. B-tagging Algorithms

Several algorithms were developed to identify jets originating from b-quarks [123]. The tagging is applied to jet candidates reconstructed by the $anti - k_T$ algorithm with a distance parameter of $\Delta R = 0.4$, $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$. The subset of tracks in the event which are associated to the individual jet are taken as inputs for the tagging. The algorithms rely on the relatively long lifetime of hadrons containing a b-quark which is of the order of 1.5 ps. A B-hadron with transverse momentum of 50 GeV will have a significant flight path length and travels typically about 3 mm in the transverse plane before decaying. Therefore, the measurement of the locations of the tracks and displaced vertices with respect to the point where the hard-scatter collision occurred are the main ingredients of b-tagging.

A tagging algorithm is characterised by the efficiency ε_b by which a jet originating from a b-quark is tagged and by the probability of mistakenly tagging a jet originating from a light-flavour parton (u-, d-, s-quark or gluon) as a b-jet, referred to as mistag rate. Both quantities were measured from data. The b-tagging efficiency is determined in an inclusive sample of jets with muons inside and in top-quark pair events with one or two leptons in the

final state while the mistag rate is measured in an inclusive jet sample.

One of the vertex- and lifetime-based tagging algorithms is called *JetFitter* and aims to reconstruct the complete decay chain of B-hadrons. The topological structure of the hadron decays inside the jet are exploited and a Kalman filter is used to find a common line on which the primary vertex and the secondary vertices lie.

One of the impact-parameter based algorithms, called *IP3D*, relies on both the transverse and longitudinal impact parameters as well as their correlations. These measured quantities are compared for each track to pre-defined two-dimensional probability functions obtained from simulation for both the b-jet and light jet hypotheses. The ratio of probabilities defines the so-called track weight.

The b-tagging efficiency for vertex-based algorithms is limited by the secondary vertex finding efficiency but their advantage over the impact-parameter based algorithms is a much lower mistag rate. Therefore, it is of interest to combine both approaches to define robust and powerful tagging algorithms. One combination technique is the use of an artificial neural network which can take advantage of complex correlations between input values. The *JetFitterCombNN* algorithm is defined that way by combining the *IP3D* and *JetFitter* algorithms. It is tuned to achieve a b-jet identification efficiency of 80% in top-quark pair events while providing a light-jet tagging rate of approximately 6% [124]. The b-tagging identification efficiency is measured in data in a sample of jets containing muons with two different methods, described in detail in Ref. [124]. The Monte Carlo prediction is in good agreement with the measurements and the derived scale factors, shown in Fig. 4.8, are used in the analysis presented in Chapter 6. The mistag rate is measured in data and compared to Monte Carlo simulation, too. The estimated scale factors depending on the jet transverse momentum and pseudorapidity, shown in Table 4.5, are of the order of 20 to 30%. They are applied to events where no identified b-jet is found.

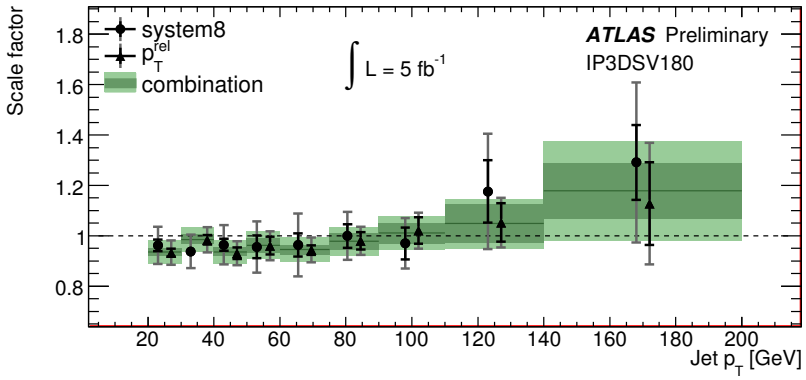


Figure 4.8.: The individual and combined data-to-simulation scale factors for the *JetFitterCombNN* tagging algorithm at 80% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty [125].

4.2.7. Missing Transverse Energy

Protons participating in the collisions at the LHC carry very little momentum in the plane transverse to the beam axis. Therefore, momentum conservation is expected in a collider

Table 4.5.: Mistag rate scale factors measured as a function of jet p_{T} (adapted from Ref. [125]).

Tagger	ε_b(%)	range	Jet p_{T} [GeV]					
			[20,30]	[30,60]	[60,140]	[140,300]	[300,450]	[450,750]
JetFitterCombNN	80	$\lvert\eta\rvert < 1.2$	1.20 ± 0.18	1.07 ± 0.14	1.12 ± 0.16	1.12 ± 0.19	1.16 ± 0.22	1.07 ± 0.25
JetFitterCombNN	80	$1.2 < \lvert\eta\rvert < 2.5$	1.20 ± 0.12	1.14 ± 0.13	1.21 ± 0.13	1.22 ± 0.15	1.30 ± 0.31	1.12 ± 0.39

event in the transverse plane and the total transverse momentum of the system is anticipated to be zero. The missing transverse energy is defined as the momentum imbalance in the system which can arise from not detectable particles such as neutrinos or other stable, weakly-interacting particles.

The negative vector sum of the momenta of all particles detected in an event is referred to as missing transverse momentum $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the symbol $E_{\mathrm{T}}^{\mathrm{miss}}$ is used for its magnitude. An accurate measurement of the missing transverse energy is crucial for the Higgs boson search in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay channel, presented in Chapter 6, due to the direct impact on the reconstruction of the transverse mass calculated from the measurable decay products of the W bosons (see Section 6.5.2). Furthermore, it is essential to minimise the effects of limited detector coverage, finite detector resolution, the presence of dead regions and different sources of noise, as well as cosmic-ray and beam-halo muons crossing the detector that can produce fake $E_{\mathrm{T}}^{\mathrm{miss}}$. The performance of the missing transverse momentum reconstruction is described in detail in Ref. [126] and will be briefly discussed below. The $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is built up of contributions from energy deposits in the calorimeters and muons reconstructed in the muon spectrometer:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,calo} + E_{x(y)}^{miss,\mu}. \tag{4.6}$$

Tracks with low p_{T} are used to recover low energetic particles which are missed in the calorimeters. Muons reconstructed from the inner detector are used to recover muons lost due to limited muon spectrometer coverage. The value of $E_{\mathrm{T}}^{\mathrm{miss}}$ is then calculated as:

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2}. \tag{4.7}$$

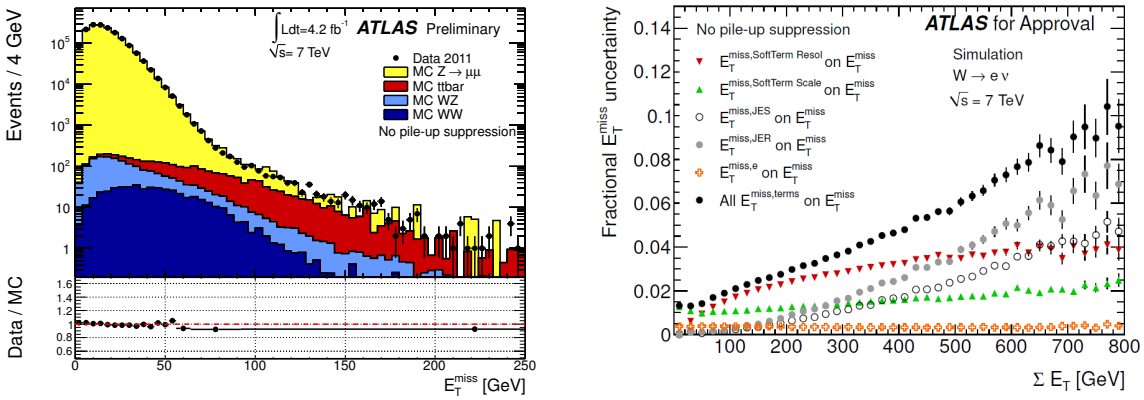
In order to ensure the accurate calculation of the missing transverse energy, the calorimeter cells used in the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are calibrated according to the reconstructed physics object to which they are associated. The chosen order of the association is as follows: electrons, photons, hadronically decaying τ-leptons, jets and muons. Cells not associated with any such objects are also considered in the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation and their contribution is important for the resolution of the missing transverse energy measurement. The $E_{\mathrm{T}}^{\mathrm{miss}}$ can therefore be written as:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,Jets} + E_{x(y)}^{miss,SoftJets} + E_{x(y)}^{miss,\mu} + E_{x(y)}^{miss,CellOut} \tag{4.8}$$

where the *Jets* term contains reconstructed jets with $p_{\mathrm{T}} > 20$ GeV and the *SoftJets* term contains reconstructed jets with $7 < p_{\mathrm{T}} < 20$ GeV. In order to suppress noise contributions which can arise due to the high granularity of the calorimeter, only cells with a significant signal are considered. This is achieved by using only cells belonging to topological clusters which may then be merged into physics objects like jets as described in the previous section.

Event samples used to assess the quality of the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are, among others, events with leptonically decaying W and Z bosons. No genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected in $Z \to \ell^+\ell^-$ events and thus the measured $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ in these events is a direct result of imperfections in

the reconstruction process or in the detector response. The distribution of reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ observed in $Z \to \mu^+\mu^-$ events is shown in Fig. 4.9. The Monte Carlo simulation expectations from $Z \to \mu^+\mu^-$ events and the dominant backgrounds are superimposed and a good agreement with data is observed. The systematic uncertainties on the $E_{\mathrm{T}}^{\mathrm{miss}}$ are evaluated on $W \to \ell\nu$ events as described in detail in Ref. [127]. The results are shown in Fig. 4.9. The overall uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale increases with $\sum E_T$ from about $\pm 2\%$ to $\pm 10\%$.



(a) (b)

Figure 4.9.: (a) Distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ as measured in a data sample of $Z \to \mu^+\mu^-$ events [128]. The expectation from Monte Carlo simulation is superimposed and normalised to data after each Monte Carlo sample is weighted with its corresponding cross section. The lower part of the figure shows the ratio between data and Monte Carlo simulation. (b) Fractional systematic uncertainty on contributions of different term uncertainties on $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty as a function of $\sum E_T$ in Monte Carlo $W \to e\nu$ events. The overall fractional systematic uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale obtained combining the contributions from various terms is shown [127].

"Success is not final, failure is not fatal: it is the courage to continue that counts."

Winston Churchill

5

Determination of Electron Shower Shapes

Many physics processes of prime interest at the LHC, such as Higgs boson decays, are expected to contain electrons with transverse momenta between a few GeV and several GeV in the final state. A solid understanding of the electron identification capability is important to ensure high efficiency as well as high jet rejection rate over a broad energy range. The efficiency to trigger, reconstruct and identify electrons has been measured using $Z \to e^+e^-$, $W \to e\nu$ and $J/\psi \to e^+e^-$ events observed in the data collected by the ATLAS experiment in the year 2010 at a centre-of-mass energy of $\sqrt{s} = 7$ TeV [7]. The baseline electron identification relies on characteristics of the clusters reconstructed in the electromagnetic calorimeter as described in Section 4.2.4. The efficiencies measured in data and predicted by Monte Carlo simulation exhibit some noticeable differences which are related to discrepancies in electron identification variables. In order to understand the sources of the observed inefficiencies, calorimeter shower shape variables have been extracted using $Z \to e^+e^-$ events. The measurement is carried out with a *tag-and-probe* approach by requiring one well-identified electron in the decay (called *tag* electron) and by examining the quantities of the second *probe* electron candidate forming a pair with an invariant mass close to the Z boson mass. The residual background to $Z \to e^+e^-$ events affects the derived distributions and is therefore removed on a statistical basis using the *sPlot* technique [129]. The obtained results serve as a basis to improve the calorimeter simulation in order to achieve a better description of the shower shape distributions. This is an important ingredient to further optimise the electron reconstruction and identification performance.

This chapter is organised as follows: the data and Monte Carlo samples used for the study are summarised in Sections 5.1 and 5.2. The electron shower shapes examined in the context of this thesis are briefly summarised in Section 5.3 followed by the details about the electron candidate selection from $Z \to e^+e^-$ decays in Section 5.4. The method used to extract the shower shapes is discussed in Section 5.5. The validation of the method performed with a closure test on Monte Carlo simulation, the assignment of systematic uncertainties and the shower shape distributions extracted from data compared to MC simulation are shown in Section 5.6. The chapter is closed by presenting a first set of shower shapes obtained with improved calorimeter simulations in Section 5.7.

5.1. Data Samples

The studies were performed based on proton-proton collision data collected with the ATLAS detector in 2010. The LHC delivered the first proton-proton collisions ever at $\sqrt{s} = 7$ TeV in March 2010 and the data taking period lasted until October. The analysis described here comprises the full 2010 dataset. It is divided into nine periods corresponding to different running conditions of the detector and the data-acquisition.

5.1.1. Data Quality

The reconstruction and identification of electrons as well as missing transverse energy, are essential for this analysis. In order to ensure the reliability of these objects, a good alignment and well understood calibration of the inner detector tracking and the electromagnetic calorimetry systems are required. Therefore, only data which were taken when all the relevant detector components were up and running under nominal conditions are considered. These conditions are defined and monitored by the ATLAS Data Quality group [130]. Events are rejected if, for example, the inner detector tracking systems, the electromagnetic calorimetry system or the trigger were in unusual conditions.

5.1.2. Trigger Selection

The dataset was collected using inclusive single electron triggers. At the beginning of the data taking only the hardware-based L1 trigger was used while the high-level trigger was needed in addition for the later periods to keep the output rate within the allocated bandwidth (see Section 4.2.4). The L1 trigger used for about 2% of the dataset (periods A to E3), is calorimeter-based and requires one electron or photon candidate with a transverse energy above 14 GeV. The high-level trigger used for the later periods is called *EF_e15_medium* and requires one electron candidate with transverse energy above 15 GeV that passes the *medium* identification requirements (see Section 4.2.4). Both triggers were not prescaled in the corresponding periods and the selected data sample corresponds to an integrated luminosity of almost 40 pb^{-1}.

5.2. Simulated Event Samples

The measurements are compared to expectations from Monte Carlo simulation. The $Z \to e^+e^-$ events are produced by the PYTHIA [34] generator and processed through the full ATLAS detector simulation [98] based on GEANT4 [100].

The background is modelled by a simulated sample of QCD dijet events generated with the PYTHIA event generator. In order to provide a sufficient number of events in which a high p_{T} electron is reconstructed, a *jet filter* (JF) is applied. The simulated background events are needed to study a possible bias on the extracted shower shapes caused by the background subtraction method. The sample is generated with a cut of $E_{\mathrm{T}} > 15$ GeV on the hard-scattering process. The samples are not pure QCD samples and contain heavy-flavour production, prompt photon production and single W/Z boson production according to their cross sections. A filter is applied at generator level to increase the probability that a selected jet passes the electron identification criteria. As outlined in Section 4.2.4, electrons tend to deposit their energy in a narrow cluster in the electromagnetic calorimeter. Therefore, the filter requires the summed transverse energy of all stable particles (excluding muons and neutrinos) in a window of $\Delta\eta \times \Delta\phi = 0.12 \times 0.12$ to be above 17 GeV. Hence, the filtered

dijet sample is called *JF17* and retains about 8.3% of the simulated events.

All Monte Carlo events are simulated with GEANT4 [100] and are reconstructed with the same software as used for data.

5.2.1. Pile-Up Simulation and Reweighing

Multiple proton-proton interactions can occur per bunch-crossing due to the high number of protons per bunch. These additional interactions are called *in-time pile-up* events and additional particles can occur that overlap with the hard-scattering event. Another source of additional activity in the detector arises due to the operation of the LHC with bunch trains. So called *out-of-time pile-up* is caused by the small temporal spacing between the bunches in a train that results in a bunch crossing frequency which induces interactions with a temporal separation less than the response time of the detector. Both the in-time pile-up and the out-of-time pile-up contributions are modelled in the Monte Carlo simulation by overlaying simulated minimum bias events. The number of additional minimum bias events per bunch crossing follows a Poisson distribution with a mean value of 2.2. Since the conditions changed rapidly during the 2010 data taking, the simulated pile-up distributions do not model equally well all data taking periods. In order to account for the difference between the simulation and the real amount of pile-up as observed in data, the Monte Carlo events are reweighed according to the distribution of multiplicity of primary vertices measured in data.

5.3. Electron Shower Shape Variables

The reconstruction of electrons has been described in detail in Section 4.2.4. The shower shape variables are calculated based on the energy deposited in the various calorimeter layers. The cluster-based discriminating variables used for *loose* and *medium* electron identification are given in Table 5.1. The corresponding distributions are studied in detail in the following sections.

5.4. $Z \rightarrow e^+e^-$ Event Selection

The decay of the Z boson into an electron-positron pair provides a clean sample of events to examine the response of the detector using a tag-and-probe method. For the selection of $Z \rightarrow e^+e^-$ candidate events, a single electron trigger is used as described in Section 5.1.2. To reject contributions from cosmic-ray and beam-halo events, at least one reconstructed primary vertex (PV) with at least three associated tracks is required for every event.

The quality of the reconstructed energy of an electron object depends on the conditions of the EM calorimeter. Problematic regions in the calorimeter arose during data taking. Most of them are caused by failures of the optical transmitters (OTx) to the readout boards. As this can have a severe impact on the reconstructed energy, electrons with part of the cluster falling into these regions are excluded. The number of non-functional OTx increased during the 2010 data taking period. The main part of the dataset was recorded with the condition of defective OTx as of the end of the 2010 data taking. Therefore, a map with all dead regions corresponding to this condition is applied to both Monte Carlo simulation and the whole dataset to exclude these areas. The loss of acceptance due to this exclusion is about 6% per electron.

In the pre-selected events, a tag electron candidate is required to satisfy the following criteria:

Table 5.1.: Definition of shower shape variables used for the electron identification in the central region of the detector in the range $|\eta| < 2.47$ [7].

Shower shape variables		
Name	Description	Type
R_{had}	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used in the ranges $\|\eta\| < 0.8$ or $\|\eta\| > 1.37$)	Hadronic leakage
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used in the range $\|\eta\| > 0.8$ and $\|\eta\| < 1.37$)	
R_η	Ratio of the energy deposited in the middle layer of the EM calorimeter in 3×7 cells and the energy deposited in 7×7 cells centered at the electron cluster position	lateral shower shape
$w_{\eta 2}$	Energy-weighted lateral width of the shower, $\sqrt{(\sum E_i \eta_i^2)/((\sum E_i) - (\sum E_i \eta_i)/(\sum E_i))^2}$ where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×7 cells	
E_{ratio}	Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	profile of shower in strip layer of EM calorimeter
w_{stot}	energy-weighted shower width, $\sqrt{(\sum E_i (i - i_{max})^2)/(\sum E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η, and i_{max} is the index of the highest-energy strip	

- transverse momentum above 20 GeV,
- pseudo-rapidity within the acceptance of the electromagnetic calorimeter ($|\eta| < 2.47$) excluding the barrel-endcap transition region, $1.37 < |\eta| < 1.52$,
- pass *tight* electron identification requirements (see Table 4.3 in Section 4.2.4).

Also, in order to avoid any bias due to the trigger identification requirements on the probe electron quantities, the tag candidate has to match the online electron that triggered the acquisition of the event.

The probe electron candidates are reconstructed by the standard cluster-based algorithm and have to pass the following requirements:

- transverse momentum above 20 GeV,
- pseudo-rapidity within the acceptance of the electromagnetic calorimeter ($|\eta| < 2.47$) excluding the barrel-endcap transition region, $1.37 < |\eta| < 1.52$,
- good track quality: number of pixel hits > 0 and number of Silicon hits > 6,
- opposite charge compared to tag electron.

All the combinations of tag and probe electrons in the event are considered in the following measurements. The total number of events passing the selection criteria stated above are

Table 5.2.: Data events observed in 2010 for the selection described in Section 5.4 using tag electrons that pass the *tight* identification requirements. The number of events quoted in the middle row are before any background subtraction.

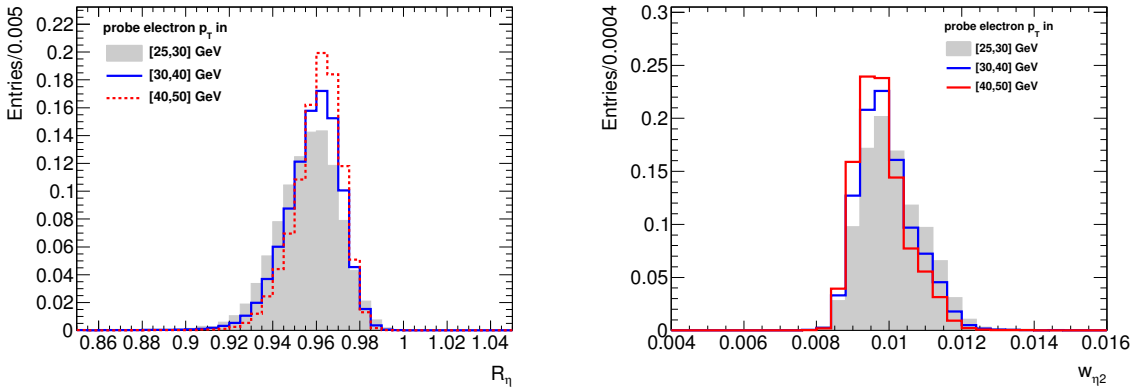
Requirement	number of events	relative cut efficiency [%]
e/γ stream	240085677	
single electron trigger	13927853	5.80
PV with $\geq$3 associated tracks	13920144	99.95
electron reconstructed by standard algorithm	8569686	61.56
$\lvert\eta\rvert <$ 2.5 excluding the region $1.37 < \lvert\eta\rvert < 1.52$	8181191	95.47
$E_{\mathrm{T}} > 20$ GeV	412091	5.04
cluster not in region of dead OTx	359540	87.25
tight tag electron	32832	9.13
matching to trigger object	32762	99.70
probe electron with opposite charge compared to tag electron	23020	70.26

shown in Table 5.2. The efficiency of each requirement with respect to the previous cut stage is given as well. There are 23020 events left at the end of the selection that contain at least one electron pair satisfying all requirements.

5.5. Method to extract the Shower Shapes

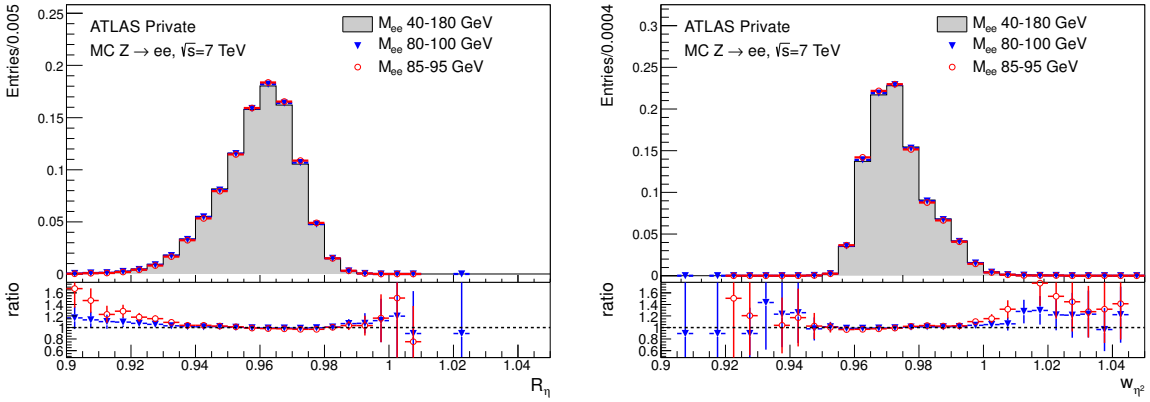
The selection described in the previous section has been used to obtain electron-positron pairs. The invariant mass of the selected pairs is shown in Fig. 5.5 for probe electrons with a transverse momentum of 25 to 30 GeV, 30 to 40 GeV and 40 to 50 GeV. The invariant mass peak of the Z boson is clearly visible. Only kinematic and track quality cuts are applied to the probe electron candidates in order to study their shower shape distributions. A residual background under the Z boson invariant mass peak is present and its level and shape depends on the transverse momentum of the probe electron candidates. One of the reasons is the fact that the width of the shower shape variables depends on the transverse momentum of the reconstructed electron. It is evident from Fig. 5.1 that the shower shape variables in the second layer of the EM calorimeter are broader for electrons with a transverse momentum in the range of 25 to 30 GeV compared to those with higher transverse momentum. As described in Section 4.2.4, the dominant source of background to electron candidates at reconstruction level are hadrons. They tend to cause broader shower shape distributions in the EM calorimeter since the energy is deposited by multiple particles evolving from the hadron decays. Therefore, the probability to falsely identify a hadron as an electron decreases when the transverse momentum of the electron candidate increases.

As detailed in Section 4.2.4, the requirements on the variables involved in the baseline electron identification are optimised in ten bins of cluster η and eleven bins of cluster E_{T} in the range from 5 GeV to above 80 GeV. Due to the limited number of events in the 2010 dataset, it is not possible to study electron shower shape properties in such many bins. Nevertheless, in order to take the dependency of the shower shape variables on the transverse momentum of the probe electrons into account, they are extracted in three bins for probe



(a) (b)

Figure 5.1.: Distributions of electron shower shape variables from $Z/\gamma^* \to e^+e^-$ Monte Carlo simulation: (a) R_η and (b) w_{η^2} for probe electrons with E_T in the interval [25,30] GeV (grey histogram), [30,40] GeV (blue line) and [40,50] GeV (red line).



(a) (b)

Figure 5.2.: Distributions of electron shower shape variables from $Z/\gamma^* \to e^+e^-$ Monte Carlo simulation: (a) R_η and (b) w_{η^2} for probe electrons with E_T in [25,50] GeV and three different invariant mass ranges: 40 to 180 GeV (grey histogram), 80 to 100 GeV (blue markers) and 85 to 95 GeV (red markers).

electrons with E_T in [25,30] GeV, [30,40] GeV and [40,50] GeV.

The residual background under the Z boson invariant mass peak could distort the derived shower shape distributions. The effect of the background can be removed on a statistical basis using the *sPlot* technique. This technique is able to unfold the contributions of the different sources to the distribution of a data sample in a given variable. It is developed in the context of a data sample analysed using a maximum likelihood method making use of the discriminating variable which is the invariant dielectron mass in this case. The technique divides the variables into two categories: discriminating variables whose shape is known and control variables of which one wishes to extract the unknown distribution. In this study, there are two contributions present in the invariant mass distribution that should be separated, namely the $Z/\gamma^* \to e^+e^-$ signal and one combined background component. Templates or

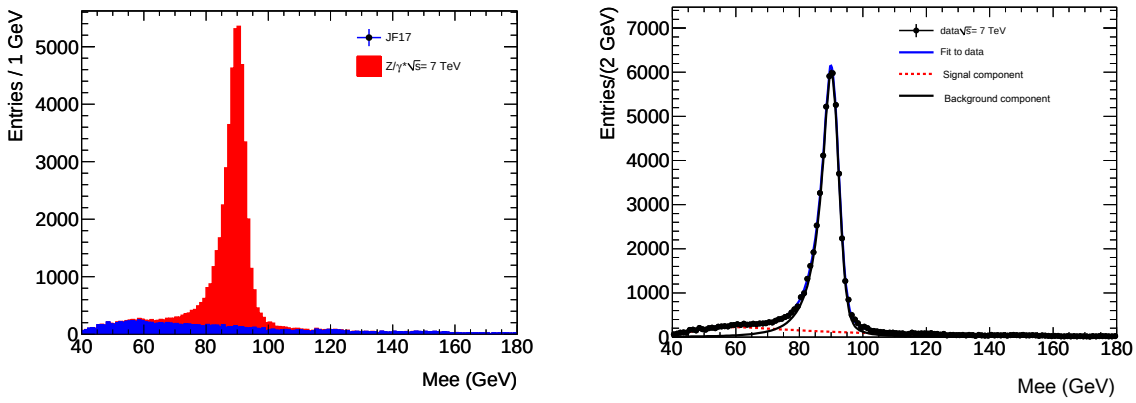
fit functions are used in a likelihood fit to model the signal and background distributions. Each event gets a weight assigned depending on its invariant mass value that reflects the probability to belong either to the signal or background component. These weights are taken into account when checking the shower shape variables of the probe electrons. The method allows to keep all signal events while getting rid of all background events in the resulting weighted distributions.

In order to obtain unbiased results, the correlation between the discriminating variable and the control variables needs to be negligible. This was verified using Monte Carlo simulation by comparing the shower shape variables R_η and w_{η^2} (control variables) for different invariant dielectron mass regions as shown in Fig. 5.2. There is no strong dependence of the shower shapes on the invariant mass visible.

5.6. Validation of the Method (Monte Carlo Studies)

In order to validate the background subtraction method, a closure test is performed on Monte Carlo simulation to check for a possible bias on the electron shower shape variables that could be caused by the background subtraction method due to remaining correlations of the invariant mass and the extracted shapes. The signal sample is constructed using $Z/\gamma^* \to e^+e^-$ events simulated with `PYTHIA`. The filtered dijet sample as described in Section 5.2 is used to model the Monte Carlo background distribution. The number of simulated Monte Carlo events is limited, especially when electron candidates are asked to pass electron identification criteria. Therefore, in order to keep a sufficient number of events, the invariant mass distribution of the background is extracted after relaxing the requirements on E_T and η of the probe electron.

The number of signal and background Monte Carlo events is chosen in the same relative



(a) (b)

Figure 5.3.: (a) Distribution of the reconstructed dielectron mass of probe electrons with E_T in the interval [30,40] GeV. The number of signal and background Monte Carlo events is chosen in the same relative proportions as observed in data. (b) The invariant mass distribution fitted with a Breit-Wigner function convoluted with a Crystal Ball function used as signal model and a single-sided exponential convoluted with a Gaussian used as background model is shown.

proportions as observed in data. The invariant mass distribution for probe electrons with a transverse energy in the range of 30 to 40 GeV is shown in Fig. 5.3. The background

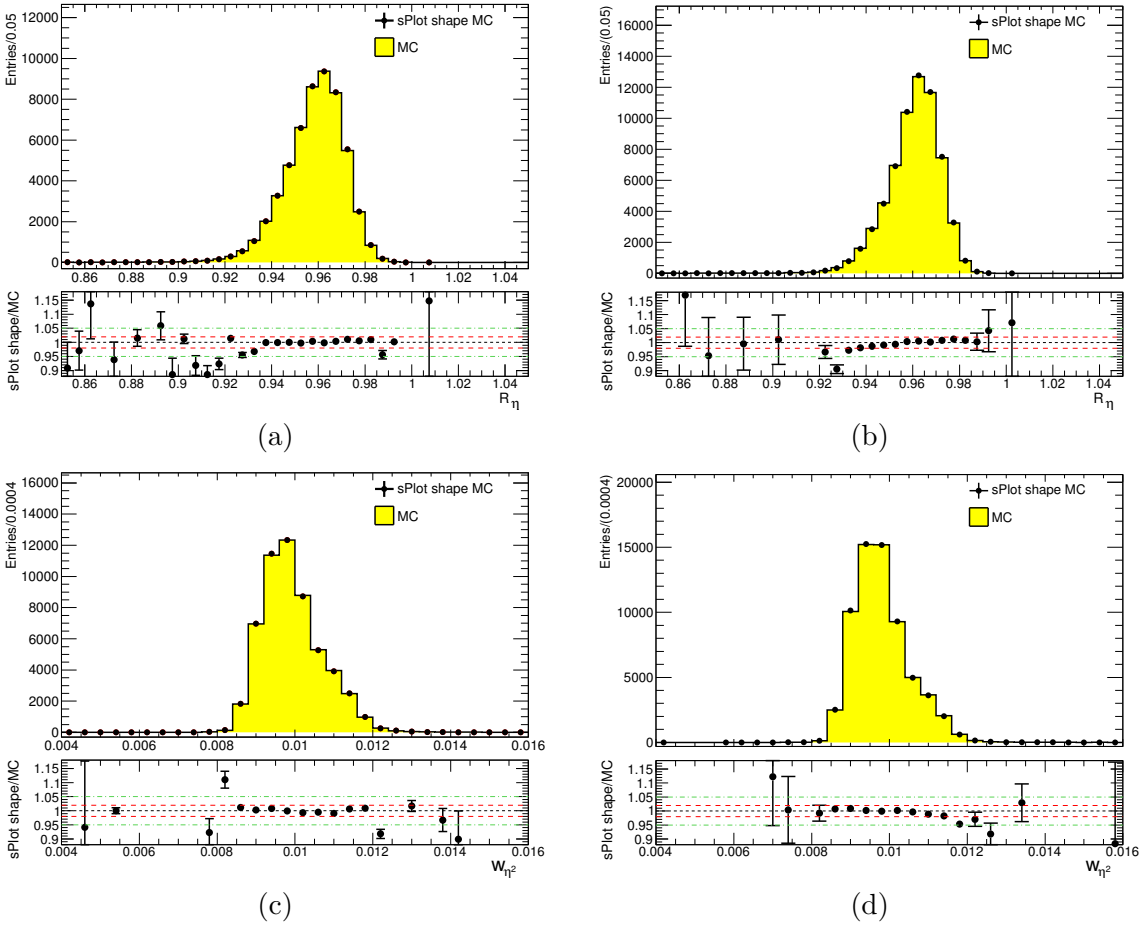


Figure 5.4.: Distributions of electron shower shapes from $Z/\gamma^* \to e^+e^-$ events for probe electrons with E_T in the interval [30,40] GeV (a) R_η and (c) w_{η^2} and E_T in the interval [40,50] GeV (b) R_η and (d) w_{η^2}. The extracted Monte Carlo shower shapes using the sPlot technique for background subtraction are plotted as full circles with error bars, representing the total statistical uncertainties. The Monte Carlo signal shower shape predictions, normalised to the number of entries, are shown by filled histograms.

subtraction with the *sPlot* technique involves a simultaneous fit of the signal and background components to the invariant mass distribution:

$$N(M_{ee}) = N_S \times P_S(M_{ee}) + N_B \times P_B(M_{ee}) \tag{5.1}$$

where N_S and N_B are the numbers of signal and background events, respectively. They are constrained such that their sum is equal to the total number of events in the data sample. $P_S(M_{ee})$ and $P_B(M_{ee})$ are the probability density functions for the signal and background component, respectively. The following functions are used in the fit:

- $P_S(M_{ee})$: a Breit-Wigner distribution convoluted with a parametrisation of the low-mass tail, arising mostly from radiations due to material effects, given by a Crystal Ball function.
- $P_B(M_{ee})$: a single-sided exponential convoluted with a Gaussian.

The fitted components for the invariant mass distribution of probe electrons with a transverse energy in the range of 30 to 40 GeV are shown on the right-hand side in Fig. 5.3. The electron

shower shape variables are extracted after the background subtraction by the *sPlot* technique. These distributions are compared directly to the Monte Carlo signal electron shower shape variables.

As an example, the resulting shapes for the middle layer variables R_η and w_{η^2} for probe electrons with a transverse energy in the range of 30 to 40 GeV and 40 to 50 GeV are shown in Fig. 5.4. The full set of variables for probe electrons with a transverse energy in the intervals [25,30], [30,40] and [40,50] GeV can be found in Appendix B. The electron shower shapes after the background subtraction agree with the signal Monte Carlo shapes within 2-3% for the bulk of the shapes. The deviations rise towards the tails of the shower shape variables. The observed deviations can arise on the one hand from imperfections of the chosen fit models and on the other hand from remaining correlations between the invariant mass and the extracted shower shapes. The discrepancies will be taken into account as systematic uncertainties on the extracted shower shapes as detailed in Sections 5.6.4 and 5.6.5.

The total systematic uncertainties comprise the uncertainties due to the background subtraction method itself estimated on MC simulation and the uncertainties arising from the choice of the fit models to distinguish the signal and background components. The two components are comparable in size. The total systematic uncertainty amounts to at most ±2% for the bulk of the shower shape distributions and increases to about ±5% in the tails.

5.6.1. Statistical Uncertainties

The statistical uncertainties arise from a finite number of events a measurement or prediction is based on. In case of the electron shower shapes extracted from data (that are presented in Section 5.6.2), the number of events is limited by the amount of recorded collision data. The statistical uncertainty is estimated in each bin of the electron shower shape distributions by $1/\sqrt{N_{observed}}$. It should be pointed out that the total uncertainty on the extracted shower shape distributions is dominated by the statistical uncertainty due to the limited number of events in data recorded by the `ATLAS` experiment in 2010.

5.6.2. Electron Shower Shapes extracted from Data

The shower shape distributions are extracted in three bins of the probe electron E_T using the 2010 dataset. The background level and shape varies depending on the probe electron E_T since the probability to falsely identify a hadron as an electron is decreasing when the transverse momentum of the electron candidate is increasing (see Fig. 5.5). A simultaneous fit of the signal and background component is performed to each invariant mass distribution in order to subtract the background component. The global probability density function used in the fit model for the signal and background components are the same as defined in Equation 5.1. The fit is performed in each bin to the invariant mass distribution over a typical range of $40 < M_{ee} < 180$ GeV as shown in Fig. 5.5.

It should be noted that, as a result of the chosen fit models, the Drell-Yan component is subtracted together with the background. Therefore, the extracted electron shower shape distributions on data are compared directly to the $Z \to e^+e^-$ Monte Carlo predictions. The Monte Carlo shower shapes are obtained by a full tag-and-probe analysis applying all cuts as described in the previous section to $Z \to e^+e^-$ events simulated with `PYTHIA`. In addition, the selected simulated electrons are required to match to true electrons originating from Z bosons or mesons and baryons containing b- or c-quarks.

The extracted electron shower shapes from data are compared to the Monte Carlo prediction for probe electrons with E_T in the range of 40 to 50 GeV in Fig. 5.6. These distributions have been published in Ref. [7]. The total shower width in the strip layer w_{stot}, also included in the

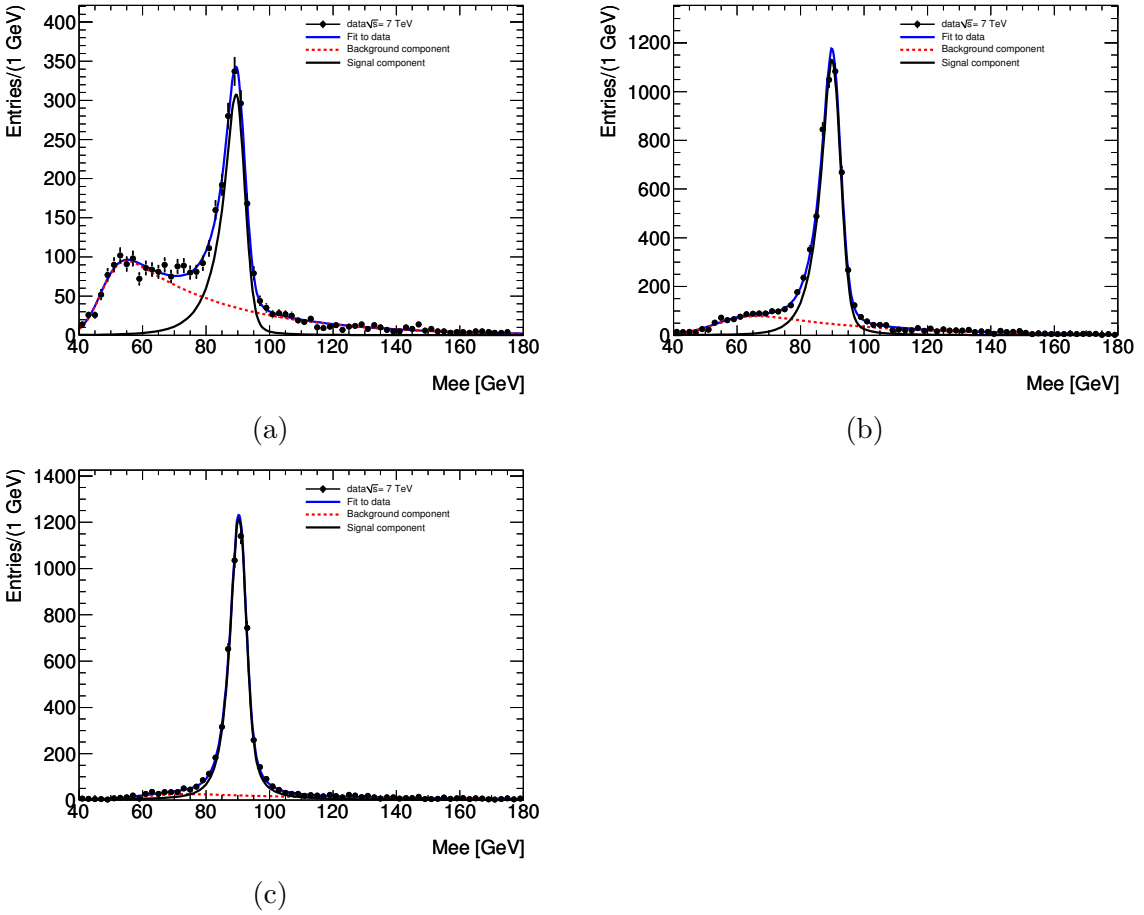


Figure 5.5.: Distributions of reconstructed dielectron mass of electron candidate pairs passing the tag-and-probe selection for probe electrons with (a) E_{T} in the interval [25,30] GeV, (b) E_{T} in the interval [30,40] GeV and (c) E_{T} in the interval [40,50] GeV.

medium identification requirements, is shown in Fig. 5.7 (a). Since many physics analyses performed with the ATLAS experiment are based on final states with isolated leptons, the isolation variable EtCone40/E_T (as defined in Section 4.2.4) has been studied as well and the distribution is shown in Fig. 5.7 (b).

There are significant differences visible for all extracted variables. The hadronic leakage distribution in data is narrower compared to the Monte Carlo simulation. This behaviour indicates an earlier start of the electromagnetic shower which leads to less energy deposited in the third layer of the EM calorimeter. It results in a lower electron identification efficiency for the *medium* criteria observed in data compared to Monte Carlo simulation.

The isolation distribution predicted by Monte Carlo simulation shows smaller isolation energies compared to data. The reasons for the discrepancies observed for the isolation and the shower shape variables could be an insufficient modelling of the pile-up and of the detector material.

The electron showers shapes in the first and second layer of the calorimeter as well as the isolation distribution have been extracted in all three intervals of probe electron E_{T} ([25,30], [30,40] and [40,50] GeV). All distributions are presented in Appendix A. The observed dis-

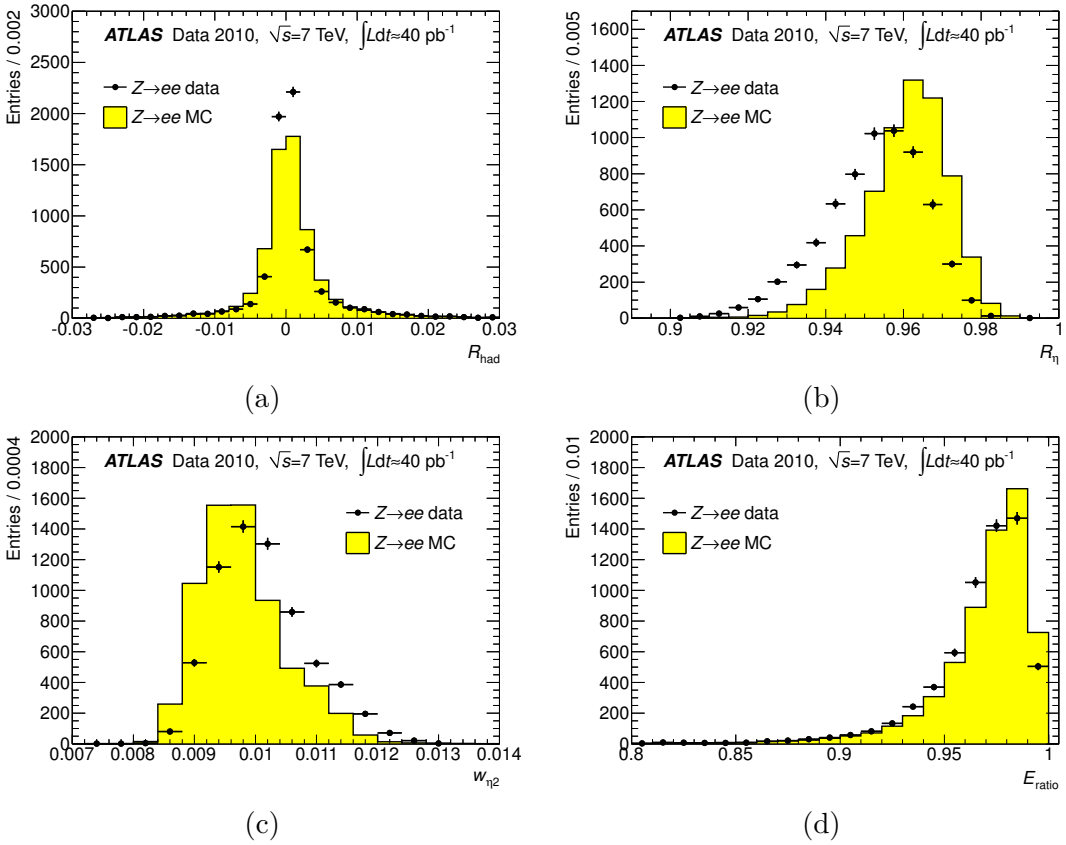


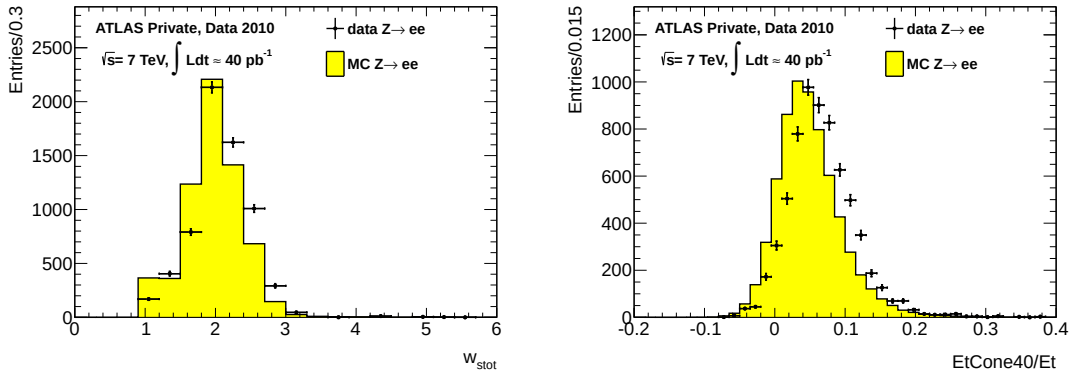
Figure 5.6.: Distributions of electron shower shape variables from $Z \to e^+e^-$ events for probe electrons with E_{T} in the interval [40,50] GeV: (a) hadronic leakage R_{had}, (b) R_{η}, (c) w_{η^2} middle-layer variables and (d) E_{ratio} strip-layer variable. The data points are plotted as full circles with error bars, representing the total statistical and systematic uncertainties. The Monte Carlo predictions, normalised to the number of data entries, are shown by filled histograms. The distributions are published in Ref. [7].

crepancies and shifts of data distributions with respect to the Monte Carlo predictions are alike in all intervals.

5.6.3. Systematic Uncertainties due to the Choice of the Fit Models

The signal and background models used in the fits for the E_{T}-dependent shower shapes are varied to estimate the uncertainty due to the choice of the fit functions or templates. The following models have been investigated:

- $P_S(M_{ee})$:
  - modelled by a Breit-Wigner distribution convoluted with a Crystal Ball function.
  - modelled by a template obtained from $Z \to e^+e^-$ MC simulations.
- $P_B(M_{ee})$:
  - modelled by a single-sided exponential convoluted with a Gaussian.
  - modelled by a Landau function.



(a) (b)

Figure 5.7.: Electron shower shape variables from $Z \rightarrow e^+e^-$ events for probe electrons with E_{T} in the interval [40,50] GeV: (a) w_{stot} total shower width in strip layer, (b) EtCone40/E_T isolation variable. The data points are plotted as full circles with error bars, representing the total statistical and systematic uncertainties. The MC predictions, normalised to the number of data entries, are shown by filled histograms.

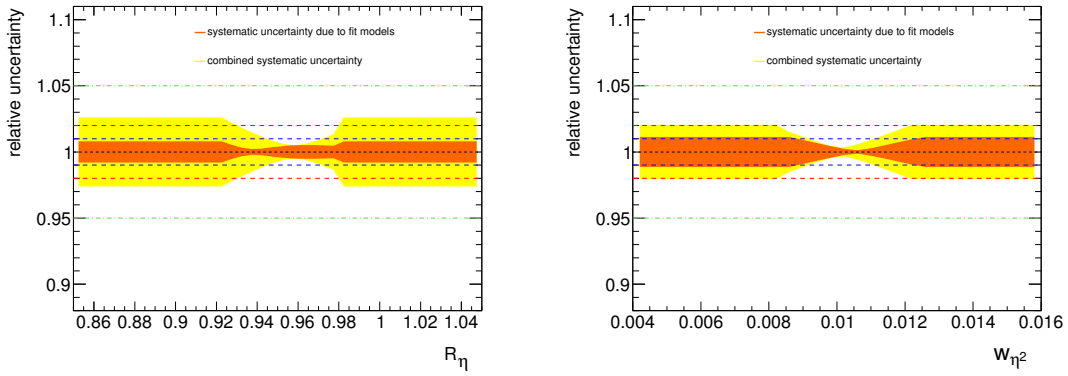
The models for the signal and background contributions are alternatively combined with each other and four different two-component fits are performed to the invariant mass distribution over a typical fit range of $40 < M_{ee} < 180$ GeV in each E_{T} bin. The electron shower shapes extracted using a Breit-Wigner distribution convoluted with a Crystal Ball function as signal model and a single-sided exponential convoluted with a Gaussian as background model are taken as reference distributions. This choice is motivated by the fact that, contrary to the template model, these models do not rely on Monte Carlo simulation for either the signal or the background shape.

In order to study the effect of the variations, the sets of electron shower shapes determined with the different fit models are compared to the reference set. The ratios of the extracted shapes versus the reference shapes are fitted with polynomials up to second order. All bins that contain more than 0.1% of the total entries are taken into account for the fit. The uncertainty values determined at the edges of the fit range are assigned to the bins beyond the fitted region. The different choices of fit models lead to independent sets of shower shape distributions. A conservative approach is used to determine the total uncertainty bands by summing up the contribution of each variation in quadrature. The resulting uncertainty band of the order of $\pm 1\%$ is shown in orange in Fig. 5.8 for the middle layer variables R_η and w_{η^2} for probe electrons with a transverse energy in the range of 40 to 50 GeV. The estimated uncertainty bands for the full set of variables and the three E_{T} bins can be found in Appendix C.

5.6.4. Systematic Uncertainties due to Background Subtraction with the *sPlot* Technique

In order to evaluate the systematic uncertainties due to the background subtraction, the ratios of extracted shapes versus Monte Carlo signal shapes are fitted with polynomials up to second order following the same procedure as described in Section 5.6.3. The ratios of extracted shapes versus Monte Carlo signal shapes for the full set of shower shape variables under investigation and the three intervals of probe electron E_{T} ([25,30], [30,40] and [40,50] GeV) are shown in Appendix B. The systematic uncertainty is about $\pm 2\%$ for the bulk of the

distributions and increases to about ±5% in the tails.



(a) (b)

Figure 5.8.: Distributions of combined systematic uncertainties for shower shape variables from $Z \to e^+e^-$ events for probe electrons in the E_T interval of [40,50] GeV: (a) R_η and (b) w_{η^2} middle-layer variables. The uncertainties due to the choice of the fit models (in orange) and the uncertainties due to the background subtraction with the *sPlot* technique estimated on Monte Carlo simulation are added in quadrature (yellow band).

5.6.5. Combined Systematic Uncertainties

The systematic uncertainty due to the background subtraction method itself estimated on Monte Carlo simulation as described in Section 5.6 is not correlated to the uncertainties observed by varying the fit models on data. Therefore, both uncertainties are summed up in quadrature to get the final uncertainty bands that are shown in yellow in Fig. **??**. The estimated uncertainties for the full set of shower shapes and all three intervals of probe electron E_T ([25,30], [30,40] and [40,50] GeV) are shown in Appendix C. The size of the uncertainties are E_T dependent and they amount to ±1% to ±5% for probe electrons with E_T in the interval of [25,30] GeV, and to ±1% to ±3% in the interval of [40,50] GeV, depending on the shape variable. It should be pointed out once again that the total uncertainty is dominated by the statistical uncertainty for electron shower shapes extracted on the dataset collected in 2010 by the `ATLAS` experiment.

5.7. Improved Electron Shower Shapes

The presented study of the extracted electron shower shapes serves as a basis for investigations to improve the Monte Carlo simulation. Possible sources of discrepancies are imperfections in the description of additional material in front of the calorimeter, misalignment of the detector modules, simplifications of the complex calorimeter geometry and shortcomings in the models used for bremsstrahlung and multiple scattering. They have an influence on the lateral electromagnetic shower profile. A large part of the disagreement has been traced down to an insufficient description of the absorber material in the EM calorimeter used in Monte Carlo simulations. The real absorber is arranged in tiers of Iron, Glue and Lead but is included as a blended material in the absorber geometry. This simplification has the advantage to reduce the time needed to simulate electromagnetic showers of particles. The Monte Carlo predictions obtained with the simplified, blended material description (called G4.9.2) and a

new, detailed description (referred to as G4.9.4, new geometry) are compared to the electron shower shapes extracted from data in Fig. 5.9. The usage of the detailed absorber description in the simulation increases the computing time per electromagnetic shower by 30-60 % but leads to a significant improvement of the electromagnetic shower shape modelling. The remaining discrepancies are still under investigation.

A fast simulation [98] (referred to as `AFII`) composed of a full inner detector and muon spectrometer simulation and a fast calorimeter simulation based on parametrisations of electromagnetic and hadronic shower distributions is also examined (see Section 4.1.2). The parametrisation which is used in the fast simulation has been extracted on shower shapes of single electrons, photons and pions obtained with the full simulation. The lateral shower shape profile parametrisation, which is based on the old geometry, was reweighed to data based on $W \to e\nu$ candidate events selected in the 2011 dataset. The Monte Carlo prediction for $Z \to e^+e^-$ events simulated with the tuned fast simulation shows a good agreement to data (see Fig. 5.9).

The study of electron shower shapes extracted from data as part of this thesis led to a deeper understanding of the sources for differences in the main discriminating variables. In addition, the calorimeter simulations could be refined and a better description of the shower shape distributions is achieved.

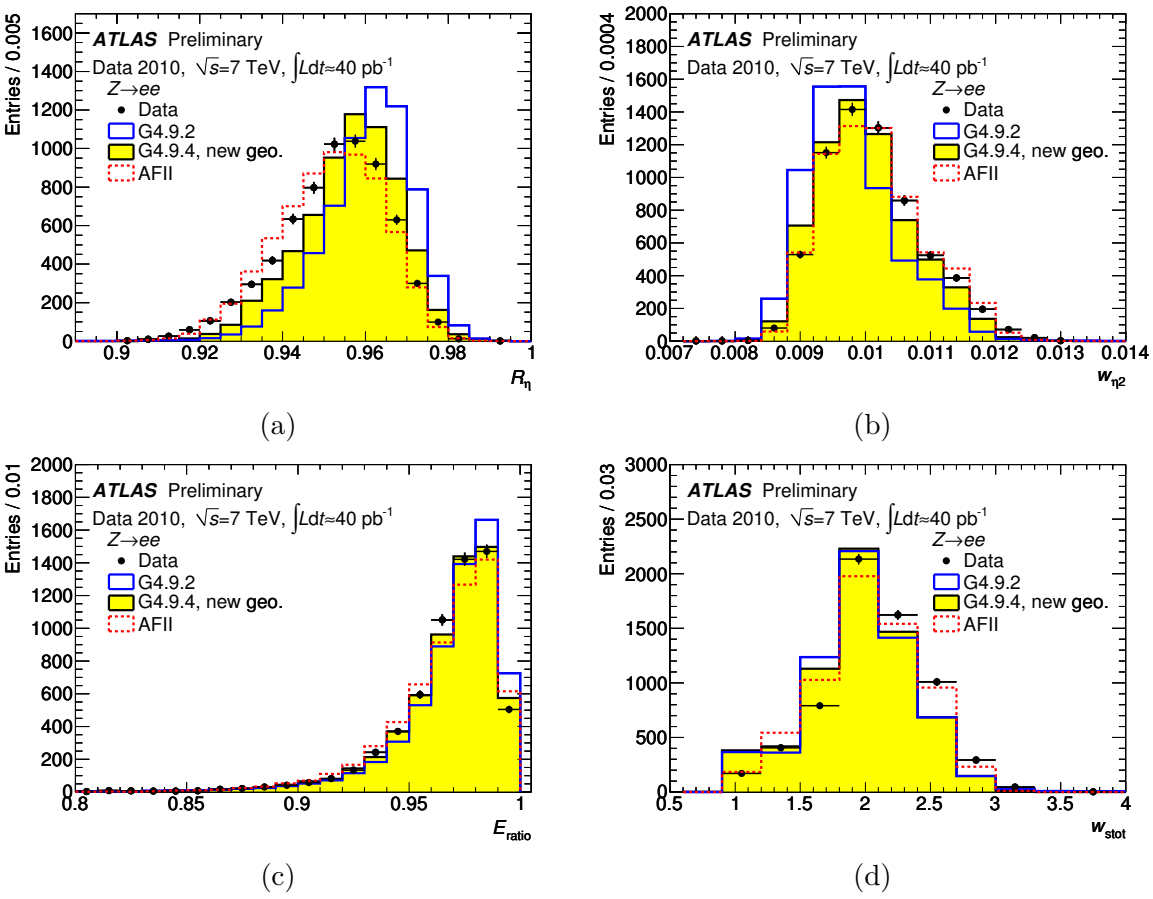


Figure 5.9.: Distributions of electron shower shapes from $Z \to e^+e^-$ events for probe electrons with E_{T} in the interval [40,50] GeV: (a) R_η and (b) w_{η^2} middle-layer variables, (c) E_{ratio} and (d) w_{stot} strip-layer variables. The data points are plotted as full circles with error bars representing the total statistical and systematic uncertainties. The Monte Carlo predictions (G4.9.2 and G4.9.4, new geometry) and the fast simulation (`AFII`), all normalised to the number of data entries, are shown as a blue histogram, filled yellow histogram, and a dashed red histogram, respectively. The plots have been published as public performance plots [131].

"Coming together is a beginning; keeping together is progress; working together is success"

Henry Ford

6

Search for the Higgs Boson in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ Decay Mode

The Higgs boson as predicted by the Standard Model (SM) of particle physics has been searched for since many years. Indirect limits on the Higgs boson mass have been obtained by global fits to electroweak precision results [2]. The masses of particles such as the W boson and the top-quark have been measured very precisely at LEP and the Tevatron and can be combined with other electroweak parameters, resulting in a constraint of $m_H < 158$ GeV at 95 % confidence level (CL). Furthermore, more stringent constraints on the mass of the Higgs boson are set by direct searches performed at these accelerators. The combined results of the four LEP experiments have set a lower limit on the mass of the SM Higgs boson of 114.4 GeV [1] and the combined Tevatron experiments exclude the region $158 < m_H < 175$ GeV [132], both at 95 % CL.

The Standard Model Higgs boson is searched for in various decay channels with the ATLAS experiment since the LHC started to operate at a centre-of-mass energy of 7 TeV in 2010. The leading production mechanism for a SM Higgs boson at the LHC is the gluon fusion (ggF). The $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay mode dominates over a wide Higgs boson mass range and is only superseded by the decay into $b\bar{b}$ pairs in the low mass range with $m_H < 135$ GeV. The presented analysis is focused on the production of a Higgs boson via gluon fusion and the subsequent decay into WW pairs which decay in turn leptonically as illustrated by the basic leading order Feynman diagram shown in Fig. 6.1.

This chapter gives an overview of the physics processes related to the search for the Standard Model Higgs boson in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ dilepton final state. The characteristics of the signal process exploited in the analysis are described in Section 6.1 followed by the discussion of the various background processes contributing to the search in this channel. The experimental observations are compared to the predictions from Monte Carlo simulation in order to test the hypothesis of the presence or absence of a SM Higgs boson. A detailed description of the datasets used in the analysis that were collected by the ATLAS experiment in 2011 and 2012 is given in Section 6.3. The Monte Carlo samples which are sets of events generated in dedicated simulations to model the various background contributions are presented in Section 6.4. The object and event selection employed for the analysis based on the data collected by the ATLAS experiment in 2011 is provided in Section 6.5. This section details as well the procedures which are used to estimate the contributions of several background

processes from data. The sources and evaluation of dominant systematic uncertainties considered in the analysis are covered in Section 6.6. The results of the analysis based on the data taken in 2011 by the ATLAS experiment are presented in Section 6.7. The modifications to the analysis of the data taken in 2012 compared to the one of the data taken in 2011 are described in Section 6.9. The combined results based on the datasets taken in 2011 and 2012 by the ATLAS experiment are shown in Section 6.15.

It should be noted that the acceptance numbers and distributions shown in the following sections are obtained with selection requirements optimised for a Higgs boson with $m_H = 125$ GeV. This mass point is chosen as a reference since it is slightly above the limits observed by LEP and within the reach of this analysis.

6.1. Signature of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ Dilepton Final State

The signature of the channel is characterised by the decay products of the W bosons, namely two isolated, oppositely charged leptons and large missing transverse energy caused by the neutrinos escaping undetected the ATLAS detector.

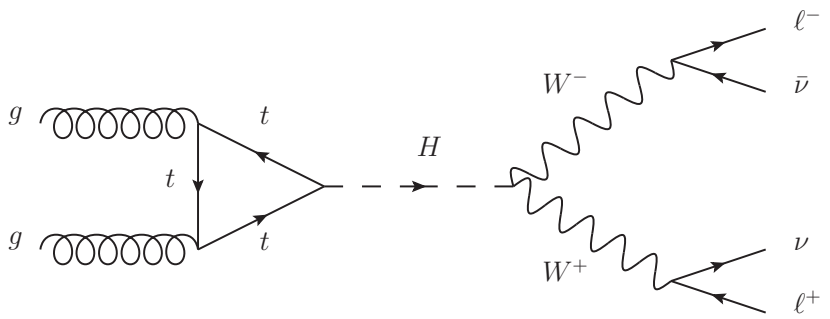


Figure 6.1.: Leading order Feynman diagram of a Higgs boson produced via gluon fusion and the subsequent decay into a WW pair which decay in turn leptonically.

Since the two W bosons in the final state are required to decay leptonically, there are the following three final states considered in this analysis: $e\nu_e e\nu_e$, $\mu\nu_\mu\mu\nu_\mu$ and $e\nu_e\mu\nu_\mu$ which are denoted as ee, $\mu\mu$ and eμ, respectively, throughout the rest of this thesis. The largest part of the signal arises through direct $W \to e\nu$ and $W \to \mu\nu$ decays, however, the small contribution proceeding through an intermediate τ lepton is implicitly included. Additionally, there may be jets in the final state originating from the VBF production mechanism or due to gluon radiation. In order to maximise the signal sensitivity, the search for the Higgs boson is performed in three disjoint channels according to the hadronic jet multiplicity in the event, they are denoted as the H + 0 jets, H + 1 jet and H+ $\geq$ 2 jets channel.

Due to the presence of two neutrinos in the final state, it is not possible to take advantage of an invariant mass reconstructed from all decay products to observe a narrow Higgs boson mass peak. Thus, a transverse mass, m_T [133], computed from the leptons and the missing transverse momentum, is used in this analysis to test for the presence of a signal.

One particular characteristic of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ dilepton final state emerges from the spin-0 property of a SM Higgs boson. The component of the spin in the direction of motion of a particle is called helicity. The W$^\pm$ bosons have three possible values for the

z-component of the spin: -1, 0 and +1. In order to conserve the overall angular momentum of the system, there are three allowed combinations of the longitudinal spin projection for the two W bosons which are travelling in opposite direction in the rest frame of the Higgs boson as shown in Fig. 6.2. The V-A structure of the weak interactions of the W boson decays imposes that neutrinos (anti-neutrinos) only exist in negative (positive) helicity states (see Section 2.1.5). Thus, it can be seen in Fig. 6.2 that, due to the electroweak couplings and the total angular momentum conservation, the charged lepton and anti-lepton in the final state are preferably emitted in the same direction. The two neutrinos tend to go into the opposite direction to the leptons resulting in a sizeable missing transverse energy. Furthermore, the

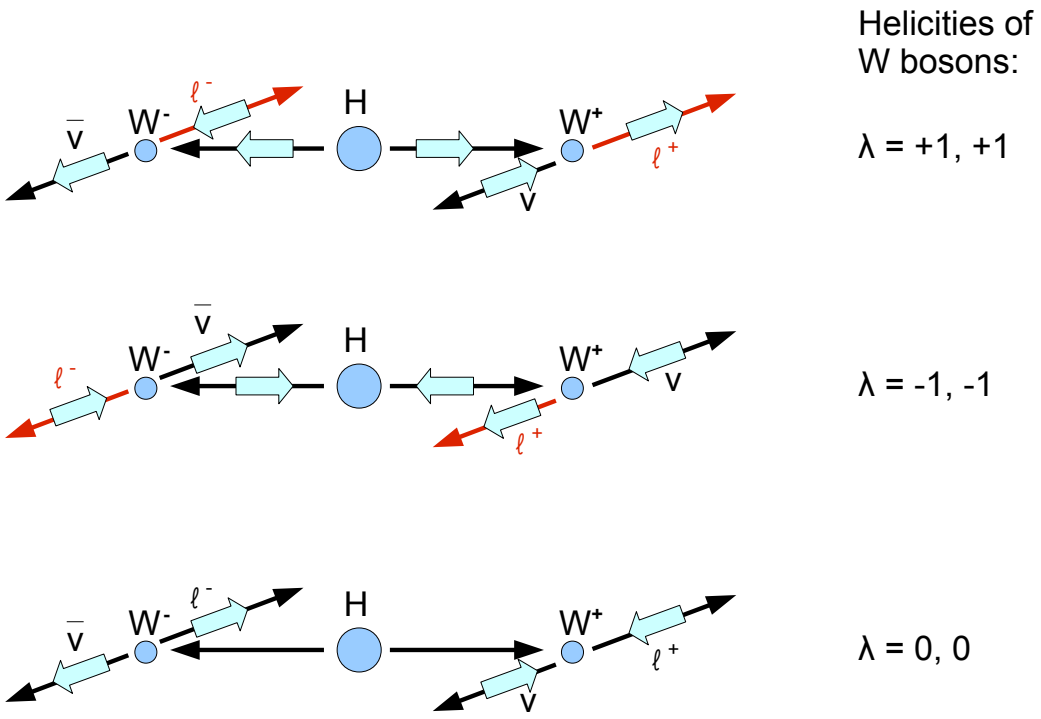


Figure 6.2.: Diagrams showing the decay of a SM Higgs boson with spin 0 into two oppositely charged W bosons with spin 1. The component of the spin in the direction of motion of a particle is indicated in the diagram by short, bold arrows and the direction of motion by the longer arrows. The diagrams illustrate the three possible $H \to W^+W^-$ decays which are allowed by spin conservation (from top to bottom) and both W bosons have: positive helicity, negative helicity or helicity zero. As indicated by the red arrows, the charged lepton and charged anti-lepton in the final state are preferentially emitted in the same direction in case of positive and negative helicity states of both W bosons.

opening angle between the charged leptons tend to be small leading to low values of the invariant mass of the dilepton pair, defined as:

$$\begin{aligned} m_{ll} &= \sqrt{(E_{l_1} + E_{l_2})^2 - (\vec{p}_{l_1} + \vec{p}_{l_2})^2} \\ &= E_{l_1}^2 - p_{l_1}^2 + E_{l_2}^2 - p_{l_2}^2 + 2 \cdot E_{l_1} \cdot E_{l_2} - 2 \cdot p_{l_1} \cdot p_{l_2} \cdot cos\theta \\ &\simeq 2 \cdot E_{l_1} \cdot E_{l_2} \cdot (1 - cos\theta) \\ &\text{for} \quad m_{l_{1,2}} \ll E_{l_{1,2}} \end{aligned}$$

with $m_{l_{1,2}}$, $E_{l_{1,2}}$ and θ denoting the mass, the energy and the opening angle of the leptons, respectively. Hence, selecting events with small invariant mass values will help to separate the signal against background processes, which tend to have larger m_{ll}.

If the Higgs boson has a mass above about 200 GeV, the $W^{\pm}$ bosons are boosted and thus the effect that charged leptons are preferentially emitted in the same direction is reduced.

6.2. Backgrounds to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ Process

There are various background processes that either result in the same final state of two oppositely charged, isolated leptons and missing transverse energy or that mimic the signature due to misidentified objects, limited detector coverage or additional activity arising from pile-up events.

The non-resonant W^+W^- production is one of the most important backgrounds to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ search followed by top-quark pairs, $Z/\gamma^* + \text{jets}$, $W + \text{jets}$ and WZ/ZZ/Wγ productions. A precise knowledge of the background processes is important to achieve an effective reduction of the individual components while retaining most of the expected signal events. The main backgrounds are briefly explained in the following in the order of their impact on the analysis.

6.2.1. Standard Model W^+W^- Production

The dominant background to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ search arises from W^+W^- boson pairs which are mainly produced through quark-antiquark annihilation at the LHC. The leading order Feynman diagrams for the s-channel and t-channel quark-antiquark annihilation and the gluon fusion production mechanism are shown in Fig.6.3. The corresponding next-to-leading order prediction of the total cross section is 45.1±2.8 pb at $\sqrt{s} = 7$ TeV [134] where the s-channel production accounts for $\approx 10\%$ and the gluon fusion through quark loops contributes about 2.9 %. The W^+W^- production cross section has been measured in proton-proton collisions at $\sqrt{s} = 7$ TeV using data corresponding to an integrated luminosity of 4.7 fb^{-1}. The measured cross section is 53.4±2.1(stat)±4.5(syst)±2.1(lumi) pb which is about 10% higher than the NLO prediction quoted above. The SM WW production cross section is expected to increase by approximately 22% when raising the centre-of-mass energy to $\sqrt{s} = 8$ TeV [135].

In the analysis presented in this thesis, the predicted rate of the SM W^+W^- is normalised to data using control regions that are obtained with selections similar to those used in the signal region but with some criteria reversed or modified to create signal-depleted, background-enriched regions. That way, the normalisation is retrieved independently of the NLO prediction and any related data to theory deviations.

The $q\bar{q}$ initial state system that produces the W^+W^- pair has three possible values for the overall z-component of the spin, namely -1, 0 and +1, since the quarks are fermions with spin 1/2. The decay products arising from a system with spin projection state 0 have a topology similar to the one of the $H \to W^+W^-$ signal events. The spin projection state of ±1 can contain one W boson with helicity state 0 and the other W boson with helicity state ± 1. Averaging over all three spin projection states observed in $q\bar{q}$ pairs, significant differences are observed in the invariant mass distribution $m_{\ell\ell}$ and the angular separation $\Delta\phi_{\ell\ell}$ of the two leptons compared to the signal events. These quantities are exploited in more detail in Section 6.5.1.

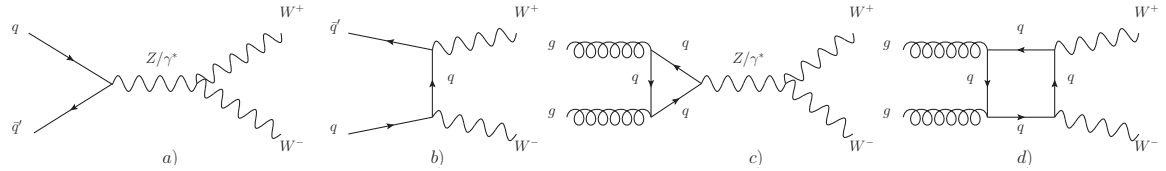


Figure 6.3.: Standard Model leading order Feynman diagrams illustrating the production of W^+W^- pairs at the LHC through $q\bar{q}$ initial state in the t-channel (a) and in the s-channel (b) and through gluon fusion mediated by quark loops ((c) and (d)).

6.2.2. Z/γ^*+jets Production

Two isolated, oppositely charged leptons arising from SM Drell-Yan processes as well as from the leptonic decay of a Z boson produced in association with jets can mimic the signal signature. Even though these events contain only true missing transverse energy in case of $Z/\gamma^* \to \tau^+\tau^-$ decays, falsely reconstructed missing transverse energy can arise from either the mismeasurement of the charged leptons, or the mismeasurement of the associated hadronic jets, or both. Additional energy depositions originating from pile-up events (see Section 6.3.4) can further degrade the energy measurements significantly.

The total Z/γ^* production cross section times the respective leptonic branching ratios for the combined electron and muon channels has been measured within the invariant mass window $66 < m_{ll} < 116$ GeV in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector to be $\sigma^{tot}_{Z/\gamma^*} \cdot BR(Z/\gamma^* \to ll)$=0.82±0.06(stat)±0.05(syst)±0.09(lumi) nb which is in good agreement within the assigned uncertainties with the theoretical prediction of $\sigma^{\mathrm{NNLO}}_{Z/\gamma^*} \cdot BR(Z/\gamma^* \to ll)$=0.96±0.05 nb including next-to-next-to-leading order QCD corrections [136]. The total leptonic production cross section is about four orders of magnitude larger compared to the expected cross section of the signal process.

The Z/γ^*+jets background mostly affects the ee and $\mu\mu$ channels but has a non-negligible contribution to the eμ channel through leptonically decaying τ-leptons in the final state. This background contribution can be significantly reduced by removing events with same flavour lepton pairs that have an invariant mass close to the Z boson mass. In addition, stringent requirements on the missing transverse energy in the reconstructed events are useful to decrease this background further. The predicted rate of Z/γ^*+jets events is normalised to data using control regions. Further details are given in Section 6.5.6.

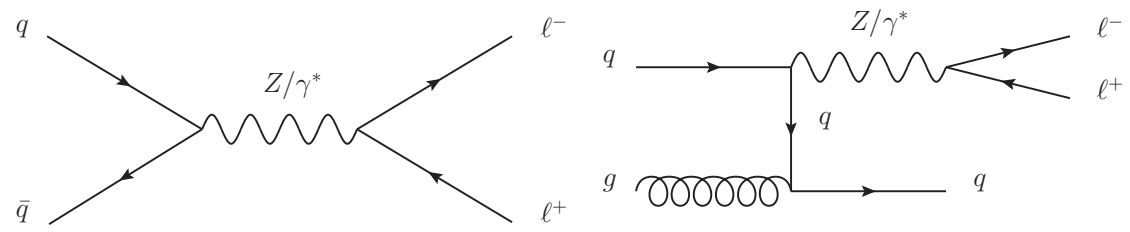


Figure 6.4.: Leading order Feynman diagrams illustrating the Drell-Yan production process without (left) and with the production of an associated jet (right).

6.2.3. Top Quark Production

In proton-proton collisions top-quark pairs are produced through both gluon-gluon and quark-antiquark annihilation as illustrated by the leading order Feynman diagrams in Fig. 6.5. At the LHC, the gluon induced production is by far the dominant production mode with a contribution of 90 % of the total cross section. Single top-quarks are produced via three different mechanisms as shown in Fig. 6.6: the t-channel exchange of a W boson, the s-channel production and decay of a virtual W boson and the associated production of a W boson and a top-quark, denoted as Wt.

Top-quarks decay almost exclusively into a W boson and a b-quark. Consequently, the top-quark pair and Wt production modes can result in signatures containing W^+W^- pairs produced in association with additional jets originating from b-quarks or initial state radiation. In case of leptonically decaying W bosons, these events constitute a background to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ signal.

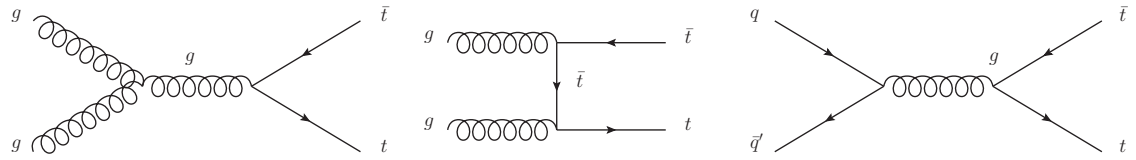


Figure 6.5.: Feynman diagrams illustrating the top-quark pair production processes at lowest order: gluon-gluon annihilation and scattering diagrams (left and middle) and quark-quark annihilation diagram (right).

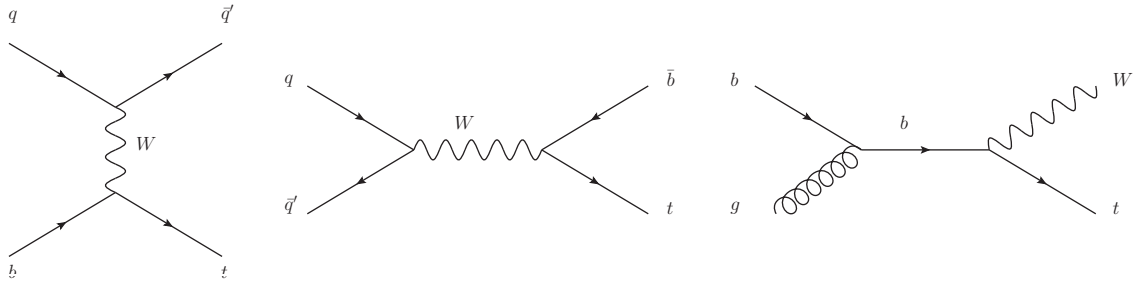


Figure 6.6.: Feynman diagrams illustrating the three production mechanisms of single top-quarks at lowest order in the t-channel (left), in the s-channel (middle) and in association with a W boson (right).

The cross sections of top-quark pair and associated single top-quark production have been measured at a centre-of-mass energy of 7 TeV using events containing two isolated leptons (electrons or muons), large missing transverse momentum, and at least two jets [137] or one jet [139], respectively. The s-channel [138] and t-channel [140] single top-quark production cross section have been measured with a similar selection but requiring only one isolated electron or muon in the final state and jets identified as coming from b-quarks. The production cross sections are summarised in Table 6.1 and the measured values are compatible within the uncertainties with the predicted cross sections.

The top-quark related contribution can be suppressed by vetoing events containing b-tagged

Table 6.1.: Predicted top-quark pair and single top-quark production cross sections for $m_{top} = 172.5$ GeV calculated at NLO with NNLL corrections compared to the measurements performed at a centre-of-mass energy of 7 TeV at the ATLAS detector (adapted from Refs. [137–140]).

process	measured cross section (pb)	predicted cross section (pb)
$t\bar{t}$	176 ± 5 (stat)$^{+14}_{-11}$ (syst) $\pm$ 8 (lumi)	165^{+11}_{-16}
single top		
t-channel	83 ± 4 (stat)$^{+20}_{-19}$(syst)	$64.4^{+2.7}_{-2.0}$
s-channel	< 26.5 pb at 95 % CL	4.6 ± 0.3
single top Wt	16.8 ± 2.9 (stat) $\pm$ 4.9 (syst)	15.7 ± 1.1

jets. The remaining top background contribution is estimated via data-driven methods and normalised to data using control regions.

6.2.4. W+jets and QCD Production

A W boson produced in association with jets as displayed by the leading order Feynman diagrams in Fig. 6.7 arises as background to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ search if the W boson decays leptonically and an accompanying jet is misidentified as a second lepton. These events result in signatures of two leptons and missing transverse energy produced by the neutrino originating from the W boson decay. The rate at which jets are misidentified as high-p_{T} leptons is expected to be of the order of 10^{-4}. However, as the total production cross section of $W \to \ell\nu$ + jets events is about four orders of magnitude higher compared to the expected cross section of the signal process, the W + jets background process is expected to have a similar size as the signal. It is also kinematically very similar to the signal process. The combination of tight lepton track and calorimeter isolation requirements helps to reduce the amount of misidentified leptons originating from hadronic jets or real, non-isolated leptons.

It is not expected that the Monte Carlo simulation models sufficiently well the complicated process of hadronic jets being identified as charged, isolated leptons. Therefore, both the rate and shape of the remaining W + jets background contributions are estimated using a data-driven technique as explained in detail in Section 6.5.6.

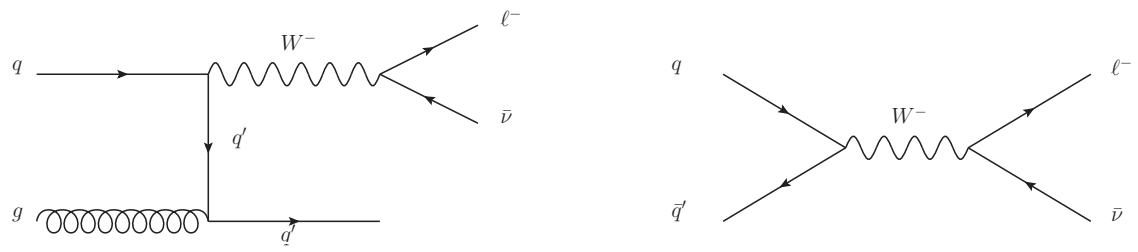


Figure 6.7.: Feynman diagrams illustrating the production of a W boson in association with a jet at a proton-proton collider.

6.2.5. WZ/ZZ/Wγ Production

The last source of background to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ signal comprises WZ, ZZ and Wγ events. These processes contain charged, isolated leptons and missing transverse energy induced by leptonically decaying W bosons. In case of the Wγ process a photon can produce a secondary electron. These background contributions are significantly reduced by vetoing events with three or more identified leptons. At present, these processes are determined using Monte Carlo simulation since they are expected to be small. There are efforts ongoing to check the MC prediction using control regions.

6.3. Data Samples

The data samples used for this analysis were recorded by the ATLAS experiment in 2011 at a centre-of-mass energy of 7 TeV and in 2012 at a centre-of-mass energy of 8 TeV. The data taking period in 2011 lasted from March until October and the analysis described here comprises the full 2011 dataset. In 2012, proton-proton collisions at $\sqrt{s} = 8$ TeV are still ongoing until the end of the year and the dataset collected between the beginning of April and the middle of June 2012 is included in the present analysis.

The data samples are divided in periods corresponding to different running conditions of the detector and the data-acquisition. The 2011 dataset contains eleven periods and the 2012 dataset analysed so far comprises two periods.

6.3.1. Data Quality

The reconstruction and identification of muons, electrons, jets as well as missing transverse energy are essential for this analysis. In order to ensure the reliability of these objects, a good alignment and well understood calibration of the inner detector tracking, the calorimetry and the muon systems are required. Therefore, only data which were taken when all relevant detector components were up and running under nominal conditions are considered. These conditions are defined and monitored by the ATLAS Data Quality group [130]. Events are rejected if, for example, the inner detector tracking systems, the electromagnetic calorimetry system or the trigger system were in unusual conditions.

6.3.2. Trigger Selection

Both datasets were collected using inclusive single muon and single electron triggers. Concerning the 2011 dataset, the single muon trigger required the transverse momentum of the muon to exceed 18 GeV and the single electron trigger p_{T} threshold varied from 20 to 22 GeV. In addition, the trigger object quality requirements were tightened throughout the data-taking period to cope with the increasing instantaneous luminosity. In 2012, the two main triggers require the transverse momentum of the lepton to exceed 24 GeV. In addition, the lepton must be isolated. Therefore, it is required that the scalar sum of the transverse momenta of charged particles within a distance of Δ R = 0.2 of the lepton direction normalised to the lepton transverse momentum be less than 12 % and 10 % for the muon and electron, respectively.

Only unprescaled triggers are used in order to maximise the integrated luminosity of the available dataset. The acceptance of the muon and electron triggers is limited to the range of $|\eta| < 2.4$ and $|\eta| < 2.47$, respectively. The trigger efficiencies are measured as a function of p_T and η of the leptons and data-taking period using Z events as described in detail for muons in Section 4.2.3 and for electrons in Section 4.2.4. The efficiencies in the 2011 and

2012 datasets are approximately 90% for electrons, and 90% (70%) for muons in the endcap (barrel) region of the detector.

6.3.3. Integrated Luminosity

Including the trigger and data quality requirements, the proton-proton collision data taken in 2011 at $\sqrt{s} = 7$ TeV corresponds to an integrated luminosity of 4.7 fb^{-1}. Likewise, the dataset collected between beginning of April to the middle of June 2012 at $\sqrt{s} = 8$ TeV, corresponds to an integrated luminosity of 5.8 fb^{-1} of proton-proton collision data recorded by the ATLAS experiment. The luminosity is independently determined by measuring the average number of inelastic collisions per bunch-crossing using dedicated sub-detectors, especially located close to the beam pipe, as described in Ref. [141]. The relative uncertainty is dominated by the measurement of the LHC beam currents and has been evaluated to be ±3.7% for the 2011 datasets [142]. The preliminary uncertainty on the integrated luminosity of the data taken in 2012 is extracted based on the same methods and estimated to be ±3.6%.

6.3.4. Mean Number of Interactions per Bunch Crossing for 2011 and 2012 Data

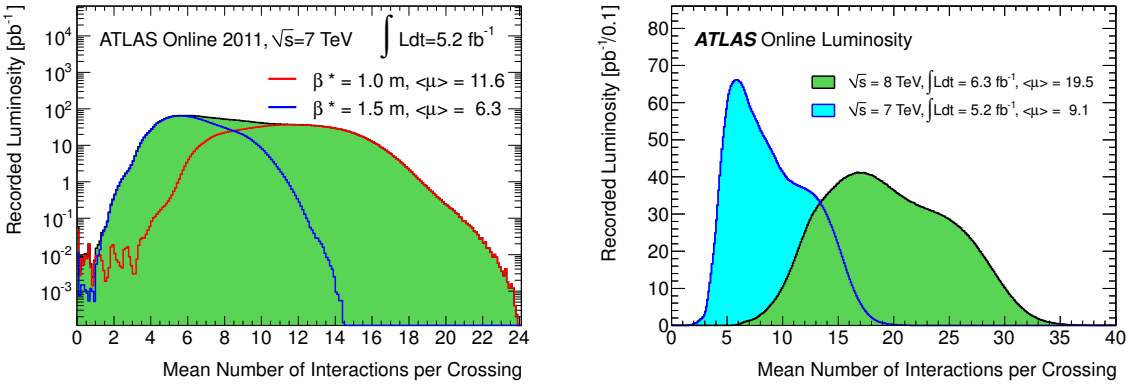
Multiple proton-proton interactions can occur in the same bunch crossing due to the high number of protons per bunch (see Section 2.4). These additional interactions are called in-time pile-up events and additional particles can occur that overlap with the hard-scattering event. The impact of in-time pile-up increases with the particle density of the colliding bunches. Another source of additional activity in the detector arises from the operation of the LHC with bunch trains. So called out-of-time pile-up is caused by the small temporal spacing between the bunches in a train that results in a bunch crossing frequency which induces interactions with a temporal separation less than the response time of the detector.

The mean number of interactions per bunch crossing increased during the data taking period and is shown for the 2011 and 2012 data in Fig. 6.8. The 2011 dataset is split up into two parts. This originates from a change to β^* that denotes the distance to the interaction point where the beam width has doubled. For data taken before and after the technical stop in September 2011, the β^* parameter was reduced from 1.5 to 1.0 meters and the mean number of interactions per crossing increased from 6.3 to 11.6. For the 2012 dataset taken between April 4th and June 18th, the mean number of interactions per crossing increased even further to 19.5.

Both the in-time pile-up and the out-of-time pile-up contributions are modelled in the Monte Carlo simulation by overlaying simulated minimum bias events. Since the conditions changed rapidly during the 2011 data taking, the simulated pile-up distributions do not model equally well all data taking periods. Therefore, in order to adapt the description of the data by the simulation, the Monte Carlo samples are reweighed to the same level of pile-up observed in data using the mean number of interactions per bunch crossing.

6.4. Simulated Samples for the 2011 Analysis

The following sections contain details about the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis performed on the 2011 dataset and published in Ref. [4]. Monte Carlo samples are used to compare the data with theoretical expectations, to estimate background contributions and selection efficiencies. The signal contributions that are considered include the dominant gluon fusion production process, the vector-boson fusion production process ($qq' \to qq'H$, denoted as



(a) (b)

Figure 6.8.: Luminosity-weighted distribution of the mean number of interactions per crossing (μ) for the data taken in 2011 (a) and for the data taken in 2011 and between April 4th and June 18th 2012 (b) by the ATLAS experiment. The integrated luminosities and the mean μ values are given in the figure. The mean number of interactions per bunch crossing corresponds to the mean of the Poisson distribution on the number of interactions per crossing calculated for each bunch (taken from Ref. [93]).

VBF) and the Higgs strahlung process (qq' $\to$ WH, ZH, denoted as WH/ZH). For the decay of the Higgs boson, only the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ mode is considered. The branching ratio for this decay, as a function of m_H, is taken from the HDECAY [143] program.

The signal cross sections are computed up to next-to-next-to-leading order (NNLO) in QCD for the ggF process [47–49, 54–56]. Next-to-leading order (NLO) electroweak (EW) corrections are also applied [60, 63], as well as QCD soft-gluon resummations up to next-to-next-to-leading log (NNLL) [58]. These results improve the NNLO calculations and are detailed in [144–146], assuming factorisation between QCD and EW corrections. Full NLO QCD and EW corrections [64–66] and approximate NNLO QCD corrections [67] are used to calculate the cross sections for VBF signal production. The cross sections of the associated WH/ZH production processes are calculated up to NNLO QCD corrections [68,69] and NLO EW corrections [70].

The ggF and VBF processes are modelled using the POWHEG [37, 38] Monte Carlo (MC) generator, interfaced to PYTHIA [34] for showering and hadronisation. The ggF Higgs boson p_T spectrum is reweighed to agree with the prediction from HqT [147].
Background processes that are taken into account are W + jets, Z/γ^* + jets, top-quark pair and single top production, WW, WZ and ZZ diboson events as well as $W\gamma$ and $W\gamma^*$ events. The Monte Carlo generators used to model signal and background processes are listed in Table 6.2.

The number quoted for the inclusive $Z/\gamma^\star$ process (also referred to in the text as the Drell-Yan process) is obtained using DYNNLO [148] predictions and corresponds to generated dilepton invariant masses exceeding 10 GeV. Kinematic criteria are also applied in the generation of $W(\to \ell\nu)\gamma$ events, namely the photon must have $p_T > 10$ GeV and be separated from the charged lepton by $\Delta R = \sqrt{(\Delta\eta^2) + (\Delta\phi^2)} > 0.1$. Concerning the $W(\to \ell\nu)\gamma^\star(\to \ell'\ell')$ events, the higher and lower transverse momenta of the leptons from the $\gamma^\star$ decay must exceed 15 GeV and 5 GeV, respectively.

For most processes, separate programs are used to generate the hard scattering process and to model the parton showering and hadronisation stages. Wherever HERWIG is used for

showering and hadronisation, JIMMY [41] is used for the simulation of the underlying event.

ALPGEN [149], interfaced to HERWIG [35] with the MLM matching scheme [150] is used to model the production of W and Z/γ^* bosons decaying to charged leptons in association with jets. MC@NLO [36] is used to model $t\bar{t}$ and WW production, using HERWIG for the parton hadronisation; an additional contribution to the continuum WW background from gluon-initiated diagrams is modelled using gg2WW [151] interfaced to HERWIG. SHERPA [152] is used for the generation of ZZ final states while MC@NLO is chosen for WZ production. $W\gamma$ production is modelled with ALPGEN while MADGRAPH [153, 154] is employed for $W\gamma^*$ [155]. AcerMC [156] is used for the generation of single top events in all three production channels (s-channel, t-channel and Wt).

The CT10 PDF set [157] is used for the MC@NLO samples, CTEQ6L1 [152] for the ALPGEN, SHERPA, and MADGRAPH samples, and MRSTMCal [32] for the PYTHIA samples. The ALPGEN Z/γ^* samples are reweighed to the MRSTMCal PDF set since that leads to a better modelling of the lepton p_{T} and η distributions. All generated events are processed with the full simulation of the ATLAS detector which is based on the GEANT4 simulation. This includes a realistic treatment of the event pile-up conditions present in the 2011 data which is modelled by overlaying simulated minimum bias events on top of the original hard-scattering event. The Monte Carlo simulation make use of the knowledge that was gained from the 2011 dataset such as alignment, material distributions, underlying event and minimum bias tunings.

Table 6.2.: Monte Carlo generators used to model the signal and background processes, and corresponding cross sections at $\sqrt{s} = 7$ TeV (given for both $m_H = 125$ GeV and $m_H = 240$ GeV in the case of the signal processes). The number quoted for the inclusive $Z/\gamma^\star$ process is for generated dilepton invariant masses greater than 10 GeV. Leptonic decay modes (charged leptonic decay modes only for $Z/\gamma^\star$ production) are summed over, except for $t\bar{t}$, single-top, WZ and ZZ production; in these cases inclusive cross sections are quoted. The quoted signal production cross-sections include the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ branching ratios.

Process	Generator	m_H (GeV)	$\sigma \cdot$ Br (pb)	Ref.	order
ggF	POWHEG [37, 38]+	125	0.347	[44]	NNLO
	PYTHIA [34]	240	0.265	[44]	NNLO
VBF	POWHEG+	125	$27 \cdot 10^{-3}$	[44]	NNLO
	PYTHIA	240	$34 \cdot 10^{-3}$	[44]	NNLO
WH/ZH	PYTHIA	125	$20 \cdot 10^{-3}$	[44]	NNLO
		240	$6 \cdot 10^{-3}$	[44]	NNLO
$q\bar{q}/qg \to WW$	MC@NLO [36]+HERWIG		4.68	[32, 158]	NLO
$gg \to WW$	GG2WW [151]+HERWIG		0.14	[32, 158]	NLO
$t\bar{t}$	MC@NLO+HERWIG		167	[32, 159]	NNLO
$tW/tb/tqb$	AcerMC [156]+PYTHIA		85	[160, 161]	NNLO
inclusive W	ALPGEN [149]+PYTHIA		$32 \cdot 10^3$	[148]	NNLO
inclusive $Z/\gamma^\star$	ALPGEN [149]+PYTHIA		$15 \cdot 10^3$	[148]	NNLO
ZZ	SHERPA [152]		5.6	[32, 158]	NLO
WZ	MC@NLO		18.0	[32, 158]	NLO
$W\gamma$	ALPGEN		345	[32, 158]	NLO
$W\gamma^\star$ [155]	MadGraph [153, 154]		6.5	[32, 158]	NLO

6.4.1. Reweighing in m_{ll} and $p_{T,ll}$ applied to Drell-Yan Monte Carlo Simulation

Relevant kinematic quantities exploited in the analysis are the invariant mass ($m_{\ell\ell}$) and the transverse momentum ($p_T^{\ell\ell}$) of the dilepton system. Discrepancies between Monte Carlo simulation and data are observed in a region dominated by the Drell-Yan process for low $m_{\ell\ell}$ values. Since the $p_T^{\ell\ell}$ distribution depends on the $m_{\ell\ell}$ values of the events, it is a valid assumption that the observed discrepancies arise from an inaccurate modelling of the $p_T^{\ell\ell}$ distribution for events with low $m_{\ell\ell}$ values. In order to correct for the mismodelling of the simulation, data/MC correction factors are derived in a grid of ($m_{\ell\ell}$,$p_T^{\ell\ell}$) values for ee and $\mu\mu$ final states and zero or one accompanying jet (see Section 6.5.1) in the event, separately. The weights are applied only to simulated $Z/\gamma^\star \to ee$ and $Z/\gamma^\star \to \mu\mu$ processes.

6.5. Event Selection and Background Determination for the 2011 Analysis

The first part of this section contains the selection of $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ candidate events performed on the 2011 dataset. The second part presents the determination and normalisation of background contributions, in particular for $Z/\gamma^* + \text{jets}$, $W + \text{jets}$, top and WW events. These two parts constitute the basis for the statistical analysis of the dataset and its results which are subject of the following sections.

6.5.1. Event Cleaning, Trigger and Object Selection for the 2011 Analysis

This Section describes the event preselection and the selection criteria for the objects used in the analysis. For the selection of $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ candidate events, single electron and muon triggers are used as described in Section 6.3.2. The contamination from pure cosmic-ray and beam-halo events is reduced by requiring at least one primary vertex that is consistent with the beam spot position and that has at least three associated tracks with p_{T} > 400 MeV.

It is essential that the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) is well measured in the selected events and non-collision activity such as cosmic rays, beam-related backgrounds, or noise in the calorimeter can lead to energy depositions that impact the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement. In order to remove such events, a dedicated cleaning algorithm is applied [162, 163]. It removes less than 0.1% of the candidate events and the efficiency is found to be well modelled by the Monte Carlo simulation.

The events are recorded with unprescaled single lepton triggers (see Section 6.3.2). To account for the mismodelling of the trigger performance in simulated samples, event weights are applied. The trigger scale factors per lepton have been determined for muons and electrons as described in Section 4.2.3. Since there are two charged leptons in the selected events, there are two chances to pass a single-lepton trigger. Therefore, it is necessary to define an event-based scale factor which depends on the lepton-based scale factors as follows:

$$\text{Trigger } SF_{event} = \frac{1 - \prod_{N=1}^{2} \varepsilon_{Data,N}}{1 - \prod_{N=1}^{2} \varepsilon_{MC,N}} = \frac{1 - \prod_{N=1}^{2} \varepsilon_{MC,N} \cdot SF_N}{1 - \prod_{N=1}^{2} \varepsilon_{MC,N}} \tag{6.1}$$

where N is the number of offline objects satisfying the selection criteria of the analysis, $\varepsilon_{Data,n}$ and $\varepsilon_{MC,n}$ are the trigger efficiencies for the offline objects determined with data and MC, respectively, and SF_N are the per-lepton scale factors. The calculation assumes the existence of at least one selected lepton which matches to a triggering object within $\Delta R < 0.15$. If both leptons fail the trigger matching, the event is discarded.

$H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ candidate events are pre-selected by requiring exactly two oppositely charged leptons with p_{T} thresholds of 25 GeV and 15 GeV for the leading and subleading lepton, respectively. For muons, the full acceptance is used while for electrons the region $1.37 < |\eta| < 1.52$ is excluded. The selected electron candidates are reconstructed using a combination of tracking and calorimetric information and are required to fulfil tight++ quality, as described in Section 4.2.4. The muon candidates are identified by matching tracks reconstructed in the inner detector and in the muon spectrometer as detailed in Section 4.2.3.

In addition to the set of standard criteria, the analysis further rejects backgrounds, in particular aiming at a reduction of W + jets events, by applying not only impact parameter but also isolation cuts as listed in Table 6.3. Both relative tracking and calorimeter isolation variables with a cone of $R_{iso} = 0.3$ are employed for electrons and muons as defined in Sections 4.2.4 and 4.2.3 in order to further suppress leptons from heavy-flavour decays and jets satisfying the lepton identification criteria. The small differences between data and Monte Carlo simulation arising from these cuts are accounted for with isolation scale factors which are consistent with unity to within 1% for both electrons and muons.

Jets are identified as described in Section 4.2.5, with the additional requirement on the jet vertex fraction (JVF) which is a measure of the probability that the jet belongs to the primary vertex reconstructed in the event and defined as:

$$JVF = \frac{\sum p_{\mathrm{T}}(\text{tracks of jet associated to P.V.})}{\sum p_{\mathrm{T}}(\text{all tracks of jet})} \tag{6.2}$$

That requirement is needed to reduce contributions from pile-up jets. The performance of the JVF algorithm has been studied as a function of the multiplicity of reconstructed primary vertices. It has been found that there is no strong dependence of the JVF on the number of primary vertices and no correction has been applied. The identified jets are used to subdivide the selected events into H + 0 jets, H + 1 jet and H+ $\geq$ 2 jets channels to maximise the sensitivity of the analysis.

As already explained in Section 6.2.3, top-quark events are a substantial background to selected events with one or more than two accompanying jets. A natural way to suppress these events relies on the fact that nearly all top-quarks decays lead to jets containing a B-hadron. For the 2011 analysis, the JetFitterCombNN b-tagging algorithm was used at an operation point leading to 80 % efficiency for the selection of b-jets as described in detail in Section 4.2.6.

6.5.2. Selection of $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ Candidate Events in the 2011 Analysis

The kinematic quantities used in the present $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis to separate signal and background processes are summarised in Table 6.4. One important quantity for the selection is the calorimeter-based missing transverse energy as detailed in Section 4.2.7. A special derived quantity called $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ is used in this analysis and its definition is given in Table 6.4. Compared to $E_{\mathrm{T}}^{\mathrm{miss}}$, the use of $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ increases the rejection of events where the missing transverse energy arises from mismeasurements of jets or leptons. The reason is that in such events the direction in ϕ of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is correlated with the direction of the mismeasured object and therefore these objects get lower $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ values assigned compared to their original $E_{\mathrm{T}}^{\mathrm{miss}}$ values.

Another important variable is the transverse mass m_{T}. It is calculated from the measurable decay products of the W bosons, namely the two leptons and the missing transverse energy.

Table 6.3.: Summary of the object selection criteria for the 2011 analysis.

Object / Requirement	2011 dataset
Muons	
Reconstruction algorithm	Staco combined
Transverse momentum	$p_{\mathrm{T}} > 15$ GeV
Geometrical acceptance	$\lvert\eta\rvert < 2.4$
Impact parameter	
Transverse	$\lvert d_0/\sigma(d_0)\rvert < 3$
Longitudinal	$\lvert z_0\rvert < 1\ mm$
Isolation ($\Delta R < 0.3$)	
Track	$\left(\sum p_{\mathrm{T}}^{\mathrm{trk}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.13$
Calorimeter	$\left(\sum E_{\mathrm{T}}^{\mathrm{cell}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.14$
Electrons	
Identification criteria	tight++
Geometrical acceptance	$\lvert\eta_{\mathrm{cluster}}\rvert < 2.47$, except $1.37 < \lvert\eta\rvert < 1.52$
Transverse momentum	$p_{\mathrm{T}} > 15$ GeV
Overlap removal (e/μ)	$\Delta R(\mathrm{electron}, \mathrm{muon}) < 0.1$
Track reconstruction	Default
Impact parameter	
Transverse	$\lvert d_0/\sigma(d_0)\rvert < 10$
Longitudinal	$\lvert z_0\rvert < 1\ mm$
Isolation ($\Delta R < 0.3$)	
Track	$\left(\sum p_{\mathrm{T}}^{\mathrm{trk}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.13$
Calorimeter	$\left(\sum E_{\mathrm{T}}^{\mathrm{cell}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.14$
Jets	
Reconstruction algorithm	anti-k_t, $R = 0.4$
Geometrical acceptance	$\lvert\eta\rvert < 4.5$
Overlap removal (j, e)	$\Delta R(\mathrm{jet}, \mathrm{electron}) < 0.3$
Calibration scheme	EM+JES
Quality criteria	Looser
Transverse momentum	$p_{\mathrm{T}} > 25$ GeV ($p_{\mathrm{T}} > 30$ GeV for $2.75 < \lvert\eta\rvert < 3.25$)
Jet Vertex Fraction	$\lvert\mathrm{JVF}\rvert > 0.75$
b-tagging	
Tagging algorithm	JetFitterCombNN
Working point	80%

Given that the W bosons originate from a Higgs boson decay, the variable provides information about the mass of the Higgs boson. The kinematic endpoint of the transverse mass distribution of a Higgs boson signal is determined by its mass. Thus, distributions of this variable are expected to exhibit an excess of events in a region related to the Higgs boson mass if a signal is present. The selection requirements presented in the following sections target at the extraction of transverse mass distributions in the H + 0 jets, H + 1 jet and H+ $\geq$ 2 jets channels which have the best possible signal to background ratio.

The pre-selection of $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ candidate events by requiring exactly two

Table 6.4.: Definition of kinematic quantities used in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis to separate signal and background processes.

Analysis specific kinematic quantities	
Kinematic quantity	Definition
Relative missing transverse energy	$E_{\mathrm{T,rel}}^{\mathrm{miss}} = \begin{cases} E_{\mathrm{T}}^{\mathrm{miss}} & \text{if } \Delta\phi \geq \pi/2 \\ E_{\mathrm{T}}^{\mathrm{miss}} \cdot \sin\Delta\phi & \text{if } \Delta\phi < \pi/2 \end{cases}$ here $\Delta\phi$ is the angle between $\vec{p}_T^{\,miss}$ and the p_T of the nearest lepton or jet with $p_T < 25$ GeV
Invariant mass of the two leptons	$m_{\ell\ell} = \sqrt{(E_{\ell_1} + E_{\ell_2})^2 - (\vec{p}\,_{\ell_1} + \vec{p}\,_{\ell_2})^2)}$
Magnitude of the transverse momentum of the dilepton system	$p_T^{\ell\ell} = \lvert\vec{p}_T^{\,\ell\ell}\rvert = \lvert\vec{p}_T^{\,\ell_1} + \vec{p}_T^{\,\ell_2})\rvert$
Azimuthal angular difference between the leptons	$\Delta\phi_{\ell\ell} = \phi_{\ell_1} - \phi_{\ell_2}$ in interval [-π,π]
Transverse mass of dilepton + $E_{\mathrm{T}}^{\mathrm{miss}}$ system	$m_T = \sqrt{(E_T^{\ell\ell} + E_T^{miss})^2 - \lvert\vec{p}_T^{\,\ell\ell} + \vec{p}_T^{\,miss}\rvert^2}$ with $E_T^{\ell\ell} = \sqrt{\lvert\vec{p}_T^{\,\ell\ell}\rvert^2 + m_{\ell\ell}^2}$

oppositely charged leptons leads to a sample that is dominated by $Z/\gamma^* +$ jets events produced via the Drell-Yan process (in the following denoted as Drell-Yan background). The contributions from J/ψ and Υ decays are rejected by requiring the dilepton invariant mass to be greater than 12 GeV for the ee and $\mu\mu$ channels and to be greater than 10 GeV for the $e\mu$ channel.

The distributions of the dilepton invariant mass and the relative missing transverse energy (as defined in Table 6.4) after applying the selection requirements mentioned above are shown for both, data and expectations, in Fig. 6.9. The expectations are decomposed in the Drell-Yan background, top-quark pair and single top-quark production, diboson production (separated into WW and $WW/WZ/W\gamma$ events) and the $W +$jets and QCD multi-jets background which is estimated by a data-driven approach as described in Section 6.5.6. The other background processes and the signal process are normalised to their expected cross sections.

It should be noted that all distributions shown in this chapter have bands indicating the effect of the uncertainties on the normalisation of the contributions. The uncertainties on the shape of the distribution itself are not included in these bands. The shape systematics are taken into account in the expectations shown in the tables and used in the limit extraction.

From the distributions of the dilepton invariant mass ($m_{\ell\ell}$) and the relative missing transverse energy ($E_{\mathrm{T,rel}}^{\mathrm{miss}}$) shown in Fig. 6.9 for the ee and $\mu\mu$ channel and in Fig. 6.10 for the $e\mu$ channel, it is evident that applying cuts on $m_{\ell\ell}$ and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ significantly reduces the Z/γ^*+jets and $W +$ jets backgrounds. Therefore the dilepton invariant mass is required to differ from

Table 6.5.: Summary of the event selection criteria used for the 2011 analysis.

Requirement	$\boldsymbol{H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}}$ **Selection**
Trigger	
ee channel	EF_e20_medium, EF_e22_medium, EF_e22vh_medium1
$\mu\mu$ channel	EF_mu18_MG, EF_mu18_MG_medium
$e\mu$ channel	"or" of electron and muon trigger
Event Cleaning	one primary vertex with ≥ 3 associated tracks
	no "bad jet"
	LAr veto
Candidate selection	**events with exactly two selected, isolated leptons**
p_{T} cut	one leading lepton with $p_{\mathrm{T}} > 25$ GeV
Opposite sign	two leptons with opposite charge
$m_{\ell\ell}$ cut	$m_{\ell\ell} > 10$ GeV for the $ee/\mu\mu$ channels
	$m_{\ell\ell} > 12$ GeV for the $e\mu$ channel
Z veto	$\lvert m_{\ell\ell} - m_Z \rvert > 15$ GeV for the ee,$\mu\mu$ channel
$E_{\mathrm{T,rel}}^{\mathrm{miss}}$ cut	$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45$ GeV for the $ee/\mu\mu$ channels
	$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 25$ GeV for the $e\mu$ channel
0-jet analysis	**events with no additional identified jet**
$p_T^{\ell\ell}$	$p_T^{\ell\ell} > 45$ GeV for $ee/\mu\mu$ channels
	$p_T^{\ell\ell} > 30$ GeV for $e\mu$ channel
1-jet analysis	**events with one additional identified jet**
b-jet veto	reject events in case the jet is b-tagged
p_T^{tot}	magnitude of vector sum $\vec{p}_T^{\,tot} = \vec{p}_T^{\,l,1} + \vec{p}_T^{\,l,2} + \vec{p}_T^{\,jet} + \vec{p}_T^{\,miss} <$ 30 GeV
$Z \to \tau\tau$ veto	veto event if energy fractions carried by visible decay products are positive and $\lvert m_{\tau\tau} - m_{\mathrm{Z}} \rvert < 25$ GeV
2-jet analysis	**events with at least two additional identified jet**
opposite hemispheres	two tagged jets are required to be reconstructed in opposite hemispheres $\Delta\eta_{jet,1} \cdot \Delta\eta_{jet,2} < 0$
central jet veto	veto event if additional central jets with $p_{\mathrm{T}} > 25$ GeV (> 30 GeV) and $\lvert\eta\rvert < 2.75$ ($2.75 < \lvert\eta\rvert < 3.2$) are found
$\lvert\Delta\eta_{jj}\rvert$	pseudorapidity distance of jets is required to be $\lvert\Delta\eta_{jj}\rvert > 3.8$
$\lvert M_{jj}\rvert$	invariant mass of tagged jets is required to be $\lvert M_{jj}\rvert > 500$ GeV
b-jet veto	reject events in case at least one jet is b-tagged
p_T^{tot}	magnitude of vector sum $\vec{p}_T^{\,tot} = \vec{p}_T^{\,\ell,1} + \vec{p}_T^{\,\ell,2} + \vec{p}_T^{\,jets} + \vec{p}_T^{\,miss} <$ 30 GeV
$Z \to \tau\tau$ veto	veto event if energy fractions carried by visible decay products are positive and $\lvert m_{\tau\tau} - m_{\mathrm{Z}} \rvert < 25$ GeV

Requirement	**Common Topological Selection**		
	low m_{H} region	intermediate m_{H} region	high m_{H} region
$m_{\ell\ell}$ for:			
0-/1-jet-events	$m_{\ell\ell} < 50$ GeV	$m_{\ell\ell} < 150$ GeV	-
2-jet-events	$m_{\ell\ell} < 80$ GeV	$m_{\ell\ell} < 150$ GeV	-
$\Delta\phi_{\ell\ell}$	$\Delta\phi_{\ell\ell} < 1.8$	-	-

the Z boson mass by at least 15 GeV for the ee and $\mu\mu$ channel. As already mentioned in Section 6.2.2, the $e\mu$ channel is affected by Z/γ^*+jets events originating predominantly from $Z\gamma^* \to \tau\tau$ production. The small leptonic τ decay branching fractions lead to a much smaller background contribution which is sufficiently suppressed by requiring $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 25$ GeV. Since the Z/γ^*+jets contribution is much larger for the ee and $\mu\mu$ channel, a tighter requirement of $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45$ GeV is used there. Furthermore, after the isolation and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ cuts, the multi-jet background is found to be negligible in all three channels.

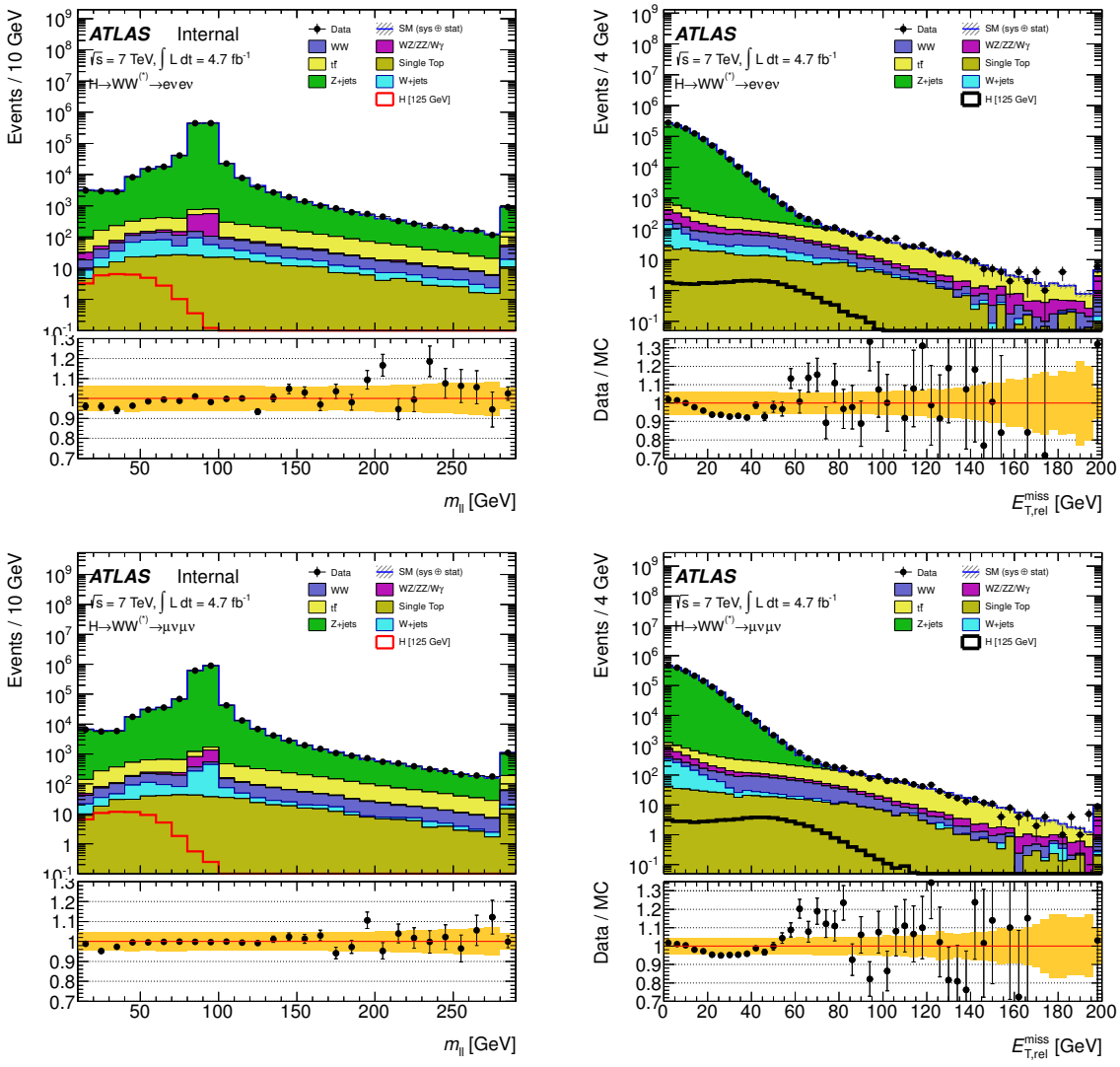


Figure 6.9.: Distributions of the dilepton invariant mass (left) and of the relative missing transverse energy (right), $E_{\mathrm{T,rel}}^{\mathrm{miss}}$, for the ee (top row) and $\mu\mu$ (bottom row) channels with the minimum lepton p_{T} and $m_{\ell\ell}$ requirements applied. The signal and background components are modelled by Monte Carlo simulation and normalised to their expected cross sections. The lower part of each plot shows the ratio between the data and the background expectations, with the yellow band indicating the total systematic uncertainty on the normalisation of the various components. The signal is shown for $\mathrm{m_H} = 125$ GeV. The final bin includes the overflow.

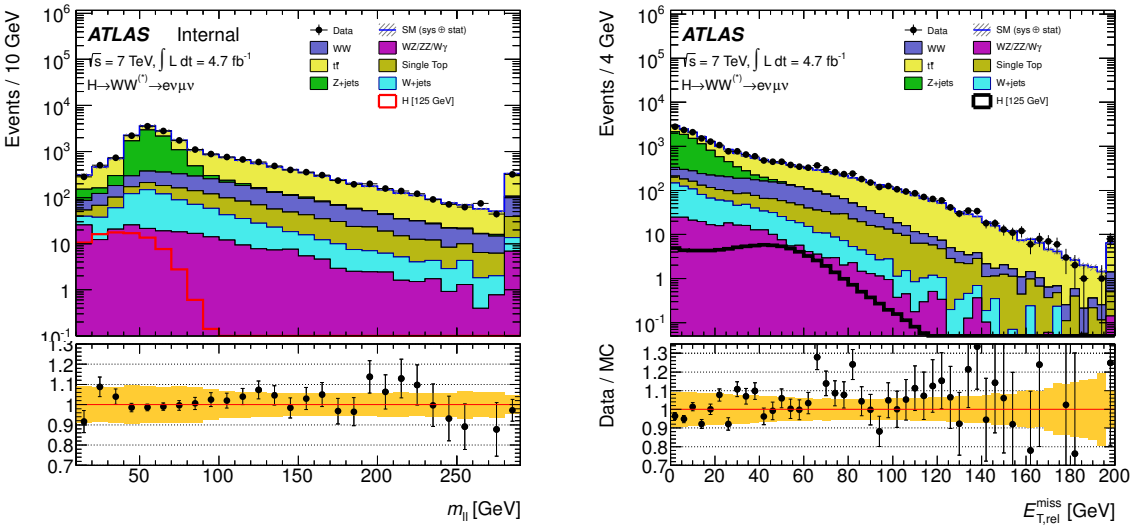


Figure 6.10.: Distributions of the dilepton invariant mass (left) and of the relative missing transverse energy (right), $E_{\mathrm{T,rel}}^{\mathrm{miss}}$, for the $e\mu$ channel with the minimum lepton p_{T} and $m_{\ell\ell}$ requirements applied. The signal and background components are modelled by Monte Carlo simulation and normalised to their expected cross sections. The lower part of each plot shows the ratio between the data and the background expectations, with the yellow band indicating the total systematic uncertainty on the normalisation of the various components. The signal is shown for $\mathrm{m_H} = 125$ GeV. The final bin includes the overflow.

Figure 6.11 shows the multiplicity distribution of jets identified as detailed in Table 6.3 for all lepton flavours combined as well as for each lepton flavour individually. The backgrounds are estimated with Monte Carlo simulation except for the W + jets component which is estimated using a data-driven approach. Furthermore, normalisation factors are applied to the WW, top and Z/γ^* + jets Monte Carlo predictions which are extracted from control regions as described in Section 6.5.6. The composition of the background as well as the signal topology depends significantly on the number of accompanying jets. For events with no identified additional jet, the signal originates almost entirely from the ggF process and the background is dominated by WW and Z/γ^* events in approximately equal proportions. On the contrary, for events with two or more identified jets, the signal contribution contains a much larger fraction of events originating from the VBF production mechanism. In that case, the dominant background process is top-quark pair and single top-quark production. In order to maximise the sensitivity by aiming for the optimal suppression of background components, jet multiplicity dependent criteria are applied in the following. It should be noted that the Z/γ^* background contribution is dominant in the ee and $\mu\mu$ channels for events with no accompanying jet. Therefore, tighter cuts are applied as described in the following section to further suppress this component.

In addition, slightly different selection criteria are used depending on the investigated Higgs boson mass region. The three regions for $\mathrm{m_H} < 200$ GeV, 200 GeV $\leq \mathrm{m_H} \leq 300$ GeV and 300 GeV $< \mathrm{m_H} < 600$ GeV are referred to as low $\mathrm{m_H}$, intermediate $\mathrm{m_H}$, and high $\mathrm{m_H}$ region, respectively.

The selection requirements applied to all events are summarised in Table 6.5. The first part of the table displays the criteria that are applied successively in the beginning to all events while the second part shows criteria depending on the number of accompanying jets in the

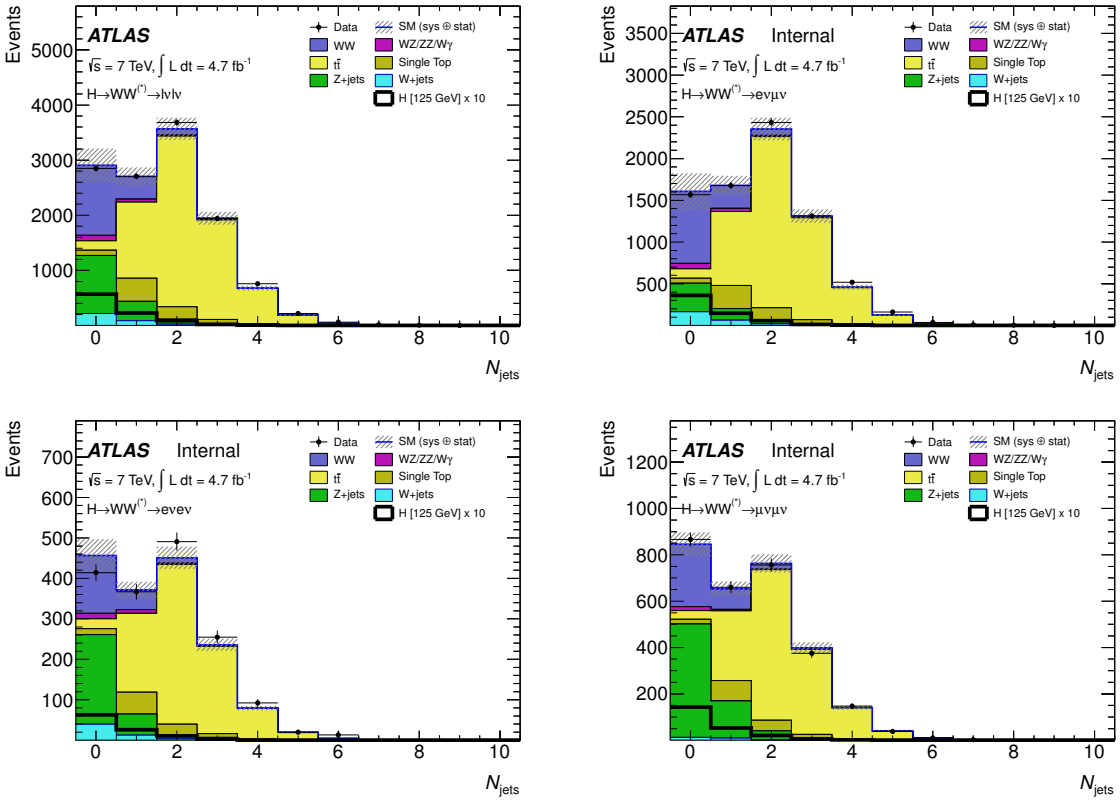


Figure 6.11.: Distributions of the multiplicity of jets for events satisfying the selection criteria up to the cut on $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ for all lepton flavours combined (top left) and separate for the $e\mu$ (top right), ee (bottom left) and $\mu\mu$ (bottom right) channel. The signal and background components are modelled by Monte Carlo simulation and normalised to their expected cross sections. The hatched area indicates the total uncertainty on the background prediction. The expected signal for a SM Higgs boson with $m_H = 125$ GeV is superimposed and multiplied by a factor of 10 for better visibility.

event. The third part of the table covers the selection specific to the investigated Higgs boson mass region. The requirements listed in the second and third part of the table are described in detail in the following sections.

6.5.3. Specific Selection Criteria for the 0-jet Analysis

The selection criteria used for events without any identified, accompanying jet exploit in particular the kinematic properties of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decays as described in detail in Section 6.1. The three quantities of particular interest are the transverse momentum of the dilepton system, the dilepton azimuthal angular difference, $\Delta\phi_{\ell\ell}$, and the dilepton invariant mass. They are shown in Fig. 6.12 after the 0-jet requirement. Due to the spin-0 nature of the Higgs boson, the charged leptons tend to emerge preferentially in the same direction. This kinematic feature is exploited in all jet multiplicities by requirements on $\Delta\phi_{\ell\ell}$ and $m_{\ell\ell}$. For searches in the intermediate and high m_H region, the kinematics of the events are different since the leptons tend to have higher p_T and larger angular separation. Therefore, the events have to satisfy the following requirements depending on the m_H region:

– $\boldsymbol{p_T^{\ell\ell}}$ **requirement:** The transverse momentum of the dilepton system, $p_T^{\ell\ell}$, is required

to be at least 45 GeV for the same flavour channels and 30 GeV for the $e\mu$ channel.

– $m_{\ell\ell}$ **requirement:** The dilepton invariant mass is required to satisfy $m_{\ell\ell} < 50$ GeV for the low m_H selection. For the intermediate m_H selection the requirement $m_{\ell\ell} < 150$ GeV is applied and for the high m_H selection the cut on the invariant mass is dropped.

– $\Delta\phi_{\ell\ell}$ **requirement:** The azimuthal angular difference between the two leptons, $\Delta\phi_{\ell\ell}$, is required to be less than 1.8 radians. The $\Delta\phi_{\ell\ell}$ requirement is not applied for $m_H \geq 200$ GeV.

It should be noted that all these cuts reduce the contribution of $Z/\gamma^* + \text{jets}$ events, too. In particular, the requirement on the transverse momentum of the dilepton system improves the rejection of the $Z/\gamma^* + \text{jets}$ background significantly while still preserving essentially all the signal events.

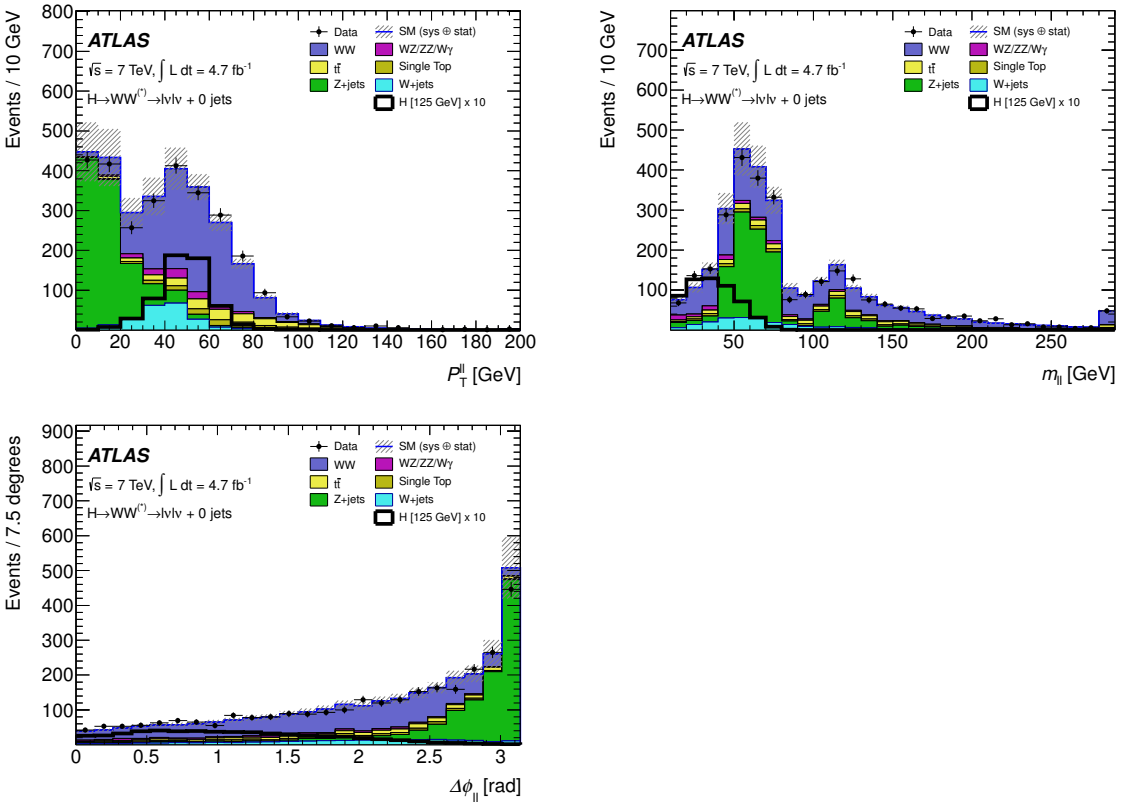


Figure 6.12.: Distributions of the transverse momentum of the dilepton system (top left), the dilepton invariant mass (top right) and the dilepton azimuthal angular difference, $\Delta\phi_{\ell\ell}$,(bottom left) after the 0-jet requirement. The lepton flavours are combined. The signal and background components are modelled by Monte Carlo simulation and normalised to their expected cross sections. The hatched area indicates the total uncertainty on the background prediction. The expected signal for a SM Higgs boson with $m_H = 125$ GeV is superimposed and multiplied by a factor of 10 for better visibility.

6.5.4. Specific Selection Criteria for the 1-jet Analysis

For events with one identified accompanying jet, backgrounds from top-quark decays are the dominant contribution followed by $Z/\gamma^* +$ jets and WW events. In order to suppress these components, the following selection requirements are applied:

- **b-jet veto:** Events are rejected if the jet is identified as originating from the decay of a b-quark or c-quark, using the b-tagging algorithm as described previously in Section 6.5.1. A jet is considered as b-tagged if the b-tagging weight produced by this algorithm is greater than -1.25. This operating point has been chosen to have 80% efficiency for b-jets in top MC events. This requirement suppresses background contributions from $t\bar{t}$ and single top production.

- $\boldsymbol{p_T^{tot}}$ **requirement:** The total transverse momentum, defined as the magnitude of the vector sum $\vec{p}_T^{\,tot} = \vec{p}_T^{\,\ell,1} + \vec{p}_T^{\,\ell,2} + \vec{p}_T^{\,jet} + \vec{p}_T^{\,miss}$, is required to be smaller than 30 GeV. This selection suppresses events with significant soft gluon radiation that recoils against the 2 leptons + 1 jet system but does not leave high p_{T} jets in the detector.

- $\boldsymbol{Z \to \tau\tau}$ **veto:** The $\tau\tau$ invariant mass, $m_{\tau\tau}$, is computed under the assumption that the reconstructed leptons are τ lepton decay products, that the neutrinos produced in the τ decays are collinear with the leptons [164], and that they are the only source of $E_{\mathrm{T}}^{\mathrm{miss}}$. If the energy factions $x_{\tau,1}$ and $x_{\tau,2}$ carried by the visible decay products are in the range [0,1] and the invariant mass of the hypothetical $\tau\tau$ system is within $|m_{\tau\tau} - m_{\mathrm{Z}}| < 25$ GeV, then the event is rejected. This requirement is only applied for events with one identified, accompanying jet because in events without a jet, the leptons tend to have $\Delta\phi_{\ell\ell}$ close to π and the $m_{\tau\tau}$ reconstruction is less accurate.

- $\boldsymbol{m_{\ell\ell}}$ **and** $\boldsymbol{\Delta\phi_{\ell\ell}}$ **requirement:** The event must pass the same $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ requirements as described in Section 6.5.3.

The distributions of the variables used for the topological selection are shown in Fig 6.13 for data and MC simulation after the requirement of only one identified, accompanying jet.

6.5.5. Specific Selection Criteria for the 2-jet Analysis

Although it is not the main subject of this thesis, the 2-jet analysis is included for completeness. Events with two or more accompanying jets are required to pass a set of cuts designed to select events where the Higgs boson is produced via the VBF production mechanism. There are a set of jet-related cuts applied to exploit the special characteristics of the VBF signature in the detector. In addition, since the top background contribution is dominant for these events, the 2-jet selection follows the 1-jet selection described in the previous Section 6.5.4. The following selection criteria are applied:

- **opposite hemisphere requirement:** The two highest-p_{T} jets in the event, so-called "tagging" jets, are required to be reconstructed in opposite hemispheres in η.

- **central jet veto:** No additional jets, besides the tagging jets, with $p_{\mathrm{T}} > 25$ GeV ($>$ 30 GeV) and $|\eta| < 2.75$ ($2.75 < |\eta| < 3.2$) are allowed.

- **b-jet veto:** Events containing one or more jets identified as originating from the decay of a b- or c-quark are rejected. This requirement suppresses top backgrounds.

- **pseudorapidity distance criterion:** The two jets must be separated in pseudorapidity by a distance $|\Delta\eta_{jj}|$ of at least 3.8 units.

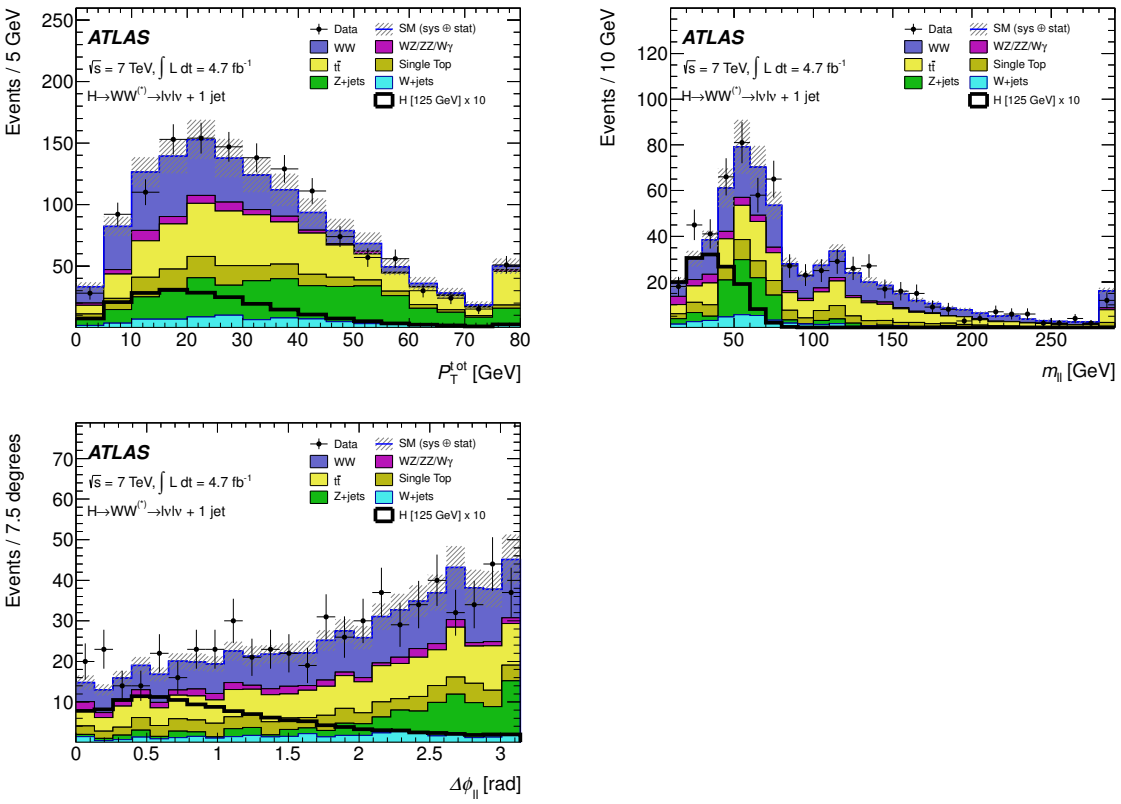


Figure 6.13.: Distributions of the total transverse momentum of the dilepton and jet system (top left), the dilepton invariant mass (top right) and the dilepton azimuthal angular difference, $\Delta\phi_{\ell\ell}$ (bottom left), after the 1-jet requirement. The lepton flavours are combined. The signal and background components are modelled by Monte Carlo simulation and normalised to their expected cross sections except for the W + jets component which is obtained by a data-driven estimate as described in Section 6.5.6 and the WW, top and Z/γ^* + jets backgrounds which are normalised to data in control regions (see Section 6.5.6). The hatched area indicates the total uncertainty on the background prediction. The expected signal for a SM Higgs boson with $m_H = 125$ GeV is superimposed and multiplied by a factor of 10 for better visibility.

- m_{jj} **requirement:** The invariant mass of the two tagging jets, m_{jj}, is required to be at least 500 GeV.
- $\boldsymbol{p_T^{tot}}$ **requirement:** The total transverse momentum definition introduced in Section 6.5.4 is extended to include all selected jets. The events are required to satisfy $p_T^{tot} < 30$ GeV.
- $\boldsymbol{Z \to \tau\tau}$ **veto:** The events satisfying the $Z \to \tau\tau$ selection criteria are rejected.
- $\boldsymbol{m_{\ell\ell}}$ **requirement:** The dilepton invariant mass is required to satisfy $m_{\ell\ell} < 80$ GeV for the low m_H selection. For the intermediate m_H selection the requirement $m_{\ell\ell} < 150$ GeV is applied and for the high m_H selection the cut on the invariant mass is dropped.
- $\boldsymbol{\Delta\phi_{\ell\ell}}$ **requirement:** The event must pass the same requirements as described in Section 6.5.3.

The distributions of the variables used for the topological selection are shown in Fig. 6.14 for data and MC simulation after the jet specific requirements (before the cut on p_T^{tot}). It is evident from the distributions that the number of data events in 4.7 fb^{-1} of data for this channel is very limited. Nevertheless, the 2-jet channel is taken into account to gain sensitivity.

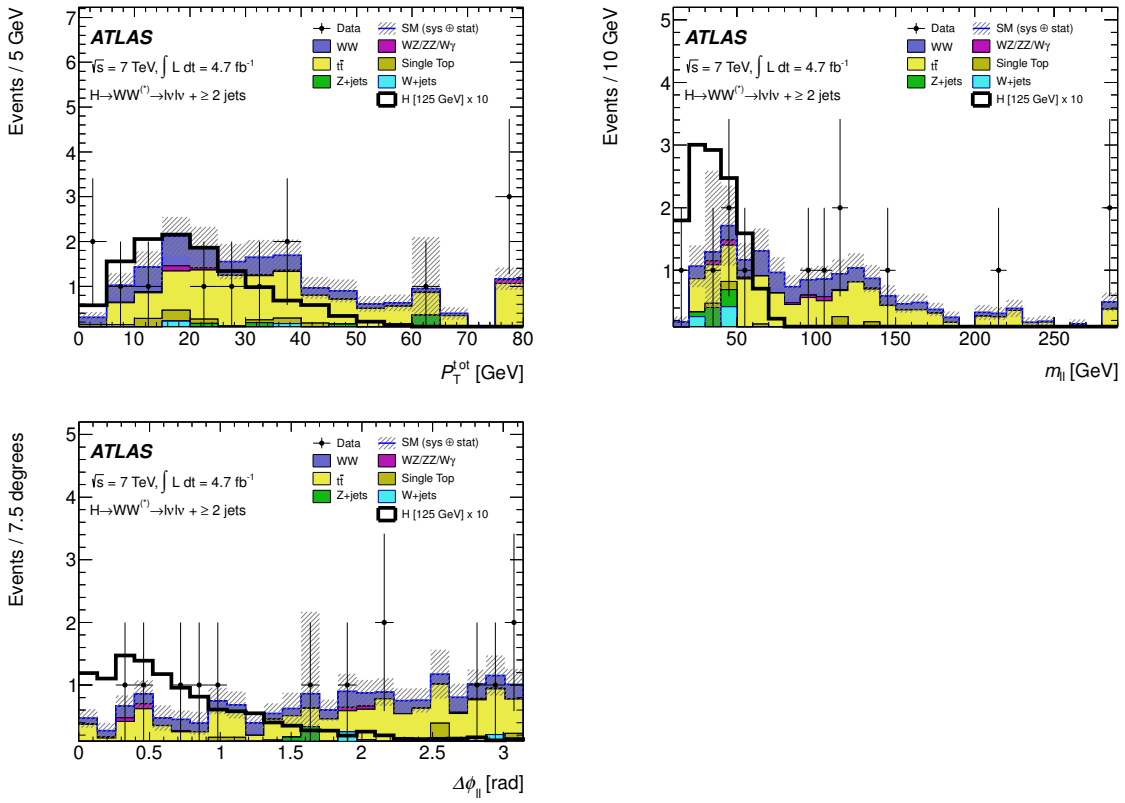


Figure 6.14.: Distributions of the total transverse momentum of the dilepton and jets system (top left), the dilepton invariant mass (top right) and the dilepton azimuthal angular difference, $\Delta\phi_{\ell\ell}$ (bottom left), after the jet specific requirements for the H+ ≥ 2 jets analysis. The lepton flavours are combined. The signal and background components are modelled by Monte Carlo simulation and normalised to their expected cross sections except for the $W+$ jets component which is obtained by a data-driven estimate as described in Section 6.5.6 and the top background which is normalised to data in control regions (see Section 6.5.6). The hatched area indicates the total uncertainty on the background prediction. The expected signal for a SM Higgs boson with $m_H = 125$ GeV is superimposed and multiplied by a factor of 10 for better visibility.

6.5.6. Background Normalisation and Control Samples for the 2011 Analysis

The main background contributions to the selected $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ candidate event samples are the non-resonant WW production, top-quark production, Z/γ^*+jets production and $W+$ jets production. The background components can be taken from MC simulation. However, the estimates may not perfectly describe the production of these processes for various reasons. For example, events containing jets may suffer from theoretical uncertainties due to the hadronisation model, choice of PDFs or the underlying event model. This could

lead to differences in the rate at which jets are falsely identified as leptons when comparing data to MC. Furthermore, mismodelling of the detector response can affect the missing E_{T} resolution and therefore cause differences in the estimate of $Z/\gamma^* + \text{jets}$ events ending up in the signal region in data and Monte Carlo. Finally, the background should be estimated from data whenever possible in order to provide a comparison with the theoretical prediction.

In order to achieve the best possible estimate of the background contributions, control regions are used to determine the correct normalisation factors for the Monte Carlo predictions of $Z/\gamma^* + \text{jets}$, top and WW processes. The backgrounds from fake leptons, which include non-isolated leptons from heavy flavour decays in jets, are directly estimated from data. The control samples and background estimates are described in the following paragraphs.

Z/γ^*+jets Control Samples

The contribution of resonant and non-resonant processes to the selected sample arises when the missing E_{T} in $Z/\gamma^* + \text{jets}$ events is mismeasured (see Section 6.2.2). In the ee and $\mu\mu$ final states, the fraction of Drell-Yan events in the signal region is estimated by a so-called ABCD-method as explained in the following. Four regions, shown in Fig. 6.15, are defined in the $m_{\ell\ell}$-$E_{\mathrm{T,rel}}^{\mathrm{miss}}$ plane:

- **Region A:** 12 GeV $< m_{\ell\ell} <$ 50 GeV and $E_{\mathrm{T,rel}}^{\mathrm{miss}} >$ 45 GeV (signal region)
- **Region B:** 12 GeV $< m_{\ell\ell} <$ 50 GeV and 20 GeV $< E_{\mathrm{T,rel}}^{\mathrm{miss}} <$ 45 GeV
- **Region C:** $|m_Z - m_{\ell\ell}| <$ 15 GeV and $E_{\mathrm{T,rel}}^{\mathrm{miss}} >$ 45 GeV
- **Region D:** $|m_Z - m_{\ell\ell}| <$ 15 GeV and 20 GeV $< E_{\mathrm{T,rel}}^{\mathrm{miss}} <$ 45 GeV

Assuming that $m_{\ell\ell}$ and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ are uncorrelated, the ratios of Z/γ^*+jets between regions A/B and C/D are assumed to be constant. Thus, an extrapolation to the signal region is possible using the number of observed events in data through:

$$A^{\,estimated} = B^{\,observed} \times \frac{C^{\,observed}}{D^{\,observed}} \times \alpha, \quad \text{with } \alpha = \frac{(A^{\,MC}/B^{\,MC})}{C^{\,MC}/D^{\,MC}} \tag{6.3}$$

where $A^{\,MC}, B^{\,MC}, C^{\,MC}$ and $D^{\,MC}$ are the Z/γ^*+jets Monte Carlo event yields in the various regions. The correction factor α is designed to account for any differences in the ratios of Z/γ^*+jets events between these regions in case they are not constant as assumed.

The expected contributions from backgrounds like diboson and top events in the regions B, C and D are generally small compared to the number of $Z/\gamma^* + \text{jets}$ events and are subtracted using Monte Carlo predictions.

The contribution of Z/γ^*+jets events in the signal region is calculated separately in the 0-jet and 1-jet channels. The ABCD estimate can be performed after each cut in the selection but topological requirements with a dependence on either $m_{\ell\ell}$ or $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ can introduce a potential bias. In addition, later topological cuts leave insufficient number of events in the C region for a calculation. Therefore, the number of events in the C and D regions are taken at the $p_T^{\ell\ell}$ and p_T^{tot} cut stages in the 0-jet and 1-jet channels, respectively. Thus, the estimated number of events at the $\Delta\phi_{\ell\ell}$ cut stage is determined by:

$$A'^{\,estimated} = B'^{\,observed} \times \frac{C^{\,observed}}{D^{\,observed}} \times \alpha \tag{6.4}$$

where quantities with a prime indicate that the number of events is taken after the $\Delta\phi_{\ell\ell}$ cut.

The normalisation factors (N.F.) that are applied to the $Z/\gamma^* + \text{jets}$ MC predictions are

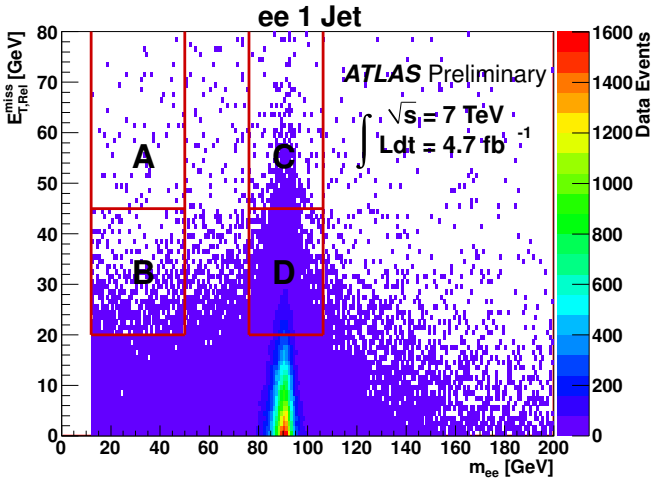


Figure 6.15.: Regions in the ($m_{\ell\ell}$-$E^{\text{miss}}_{\text{T,rel}}$)-plane used to estimate the $Z/\gamma^* +$ jets background in the ee and $\mu\mu$ channels. The distribution is shown for the ee+1-jet channel. The contribution in the signal region A is obtained by scaling the events in the region B (after background subtraction), with the ratio of events in the regions C and D.

determined by the ratio $N.F. = A^{\,estimated}/A^{\,MC}$ and they are listed for the ee and $\mu\mu$ final states and separately for the 0-jet and 1-jet analysis in Table 6.6.

The method presented assumes that the quantities $m_{\ell\ell}$ and $E^{\text{miss}}_{\text{T,rel}}$ are independent for the Z/γ^*+jets background. This assumption has been verified using MC simulation. All these possible biases of the method are evaluated and corrected for using simulated events.

The number of events available in the $Z/\gamma^* +$jets Monte Carlo samples used in this analysis is not large enough to populate the signal region of the 2-jet analysis. Therefore, a cut factorisation approach is used to estimate the fraction of $Z/\gamma^* +$ jets events passing certain requirements. The method is based on the assumption of only weak correlations among the $E^{\text{miss}}_{\text{T,rel}}$ requirement, jet specific criteria and the topological selection. In this way it is possible to extract a conservative estimate of $Z/\gamma^* +$ jets events in the 2-jet analysis.

The $Z/\gamma^* +$ jets background estimate for the $e\mu$ channel is taken from MC simulation. A cross-check has been performed for the 0-jet channel using a control region dominated by $Z \to \tau\tau$ decays. The region is mainly defined by loosening the $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ requirements and reversing the $p_T^{\ell\ell}$ requirement compared to the signal region:

- 10 GeV $< m_{\ell\ell} <$ 80 GeV
- $\Delta\phi_{\ell\ell} < 2.5$
- $p_T^{\ell\ell} <$ 30 GeV

A threshold of $E^{\text{miss}}_{\text{T,rel}} >$ 25 GeV is used to calculate the data/MC normalisation factor which is found to be consistent with unity within the uncertainty of about $\pm$10%. A similar cross-check could not be performed for higher jet multiplicities due to the difficulty of constructing a control region that is pure enough.

W+jets Control Samples

The $W +$jets background contribution is estimated using a control sample of data events with one lepton satisfying the isolation and identification criteria as described in Section 6.5.1.

Table 6.6.: Summary of the normalisation factors applied to the Z/γ^*+jets MC predictions at different stages of the selection in the ee and $\mu\mu$ final states and separately for the 0-jet and 1-jet analysis.

0-jet analysis			1-jet analysis		
cut	ee channel	$\mu\mu$ channel	cut	ee channel	$\mu\mu$ channel
jet veto	0.97	1.02	1 jet	1.05	1.05
$m_{\ell\ell}$ cut	0.95	1.04	p_T^{tot} cut	1.05	0.87
$p_T^{\ell\ell}$ cut	1.04	1.38	$m_{\ell\ell}$ cut	0.87	0.84
$\Delta\phi_{\ell\ell}$ cut	1.05	1.38	$\Delta\phi_{\ell\ell}$ cut	0.88	0.84

A second lepton (denoted "anti-identified") must fail these selection criteria but still pass loosened ones as summarised in Table 6.7.

All other selection criteria are identical to those applied in the signal region. The dominant contribution to this control sample comes from W+jets events with a jet providing an object reconstructed as a lepton. The W + jets background contribution in the signal region is then obtained by scaling the number of events in the data control sample by an extrapolation factor, referred to in the following as a "fake factor". The fake factor is defined as the ratio of the number of fully identified lepton candidates passing all selection criteria to the number which are anti-identified. It is estimated as a function of the anti-identified lepton p_{T} using an inclusive dijet data sample. The residual contributions from leptons produced by leptonic W and Z decays are subtracted. The W candidates are identified by requiring the transverse mass $m_T^W = \sqrt{2p_T^\ell E_T^{miss} \cdot (1 - cos\Delta\phi)}$ to satisfy $m_T^W > 30$ GeV. In this expression, p_T^ℓ is the lepton transverse momentum and $\Delta\phi$ is the difference in azimuth between the lepton and the missing transverse momentum direction. The Z boson candidates are identified by requiring two opposite-sign leptons of the same flavour and their invariant mass close to the Z boson mass: $|m_{\ell\ell} - m_Z| < 15$ GeV. The small remaining lepton contamination, which includes $W\gamma$ and $W\gamma^*$ events, is subtracted using MC simulation.

It should be noted that the W + jets background estimation also includes a prediction of the multi-jet (QCD) background, where both leptons originate either from misidentified jets or particle decays inside heavy flavour jets. The W + jets enhanced region already contains these events. Independent cross-checks for direct QCD contributions to the signal region, in particular to the $\mu\mu$ final state, have been performed using filtered Monte Carlo samples and same-sign events. No evidence for the presence of QCD events is found.

Top Control Samples

Backgrounds from top-quark decays are expected to lead to reconstructed jets identified as originating from b-quarks. However, these jets might either fail to be reconstructed or have $p_{\mathrm{T}} < 25$ GeV and thus will not be considered for the b-jet veto applied to the 1-jet and 2-jet analysis. This results in a non-negligible top-quark contribution to the samples selected in the H + 0 jets, H + 1 jet and H+ $\geq$ 2 jets analyses.

The estimated number of top-quark background events in the H + 0 jets signal region is extrapolated from the events satisfying the candidate selection up to the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ requirement. This dataset is dominated by top-quark background contributions as apparent from Fig. 6.11. The contributions of non-top backgrounds to this sample are subtracted using data-driven estimates for W+jets and Z/γ^*+jets backgrounds, while the other backgrounds are estimated

Table 6.7.: Definition of anti-identified electrons and muons used in the W + jets control sample.

Anti-identified electron	Anti-identified muon
same p_{T} and η range as identified electrons	same p_{T} and η range as identified muons
number of SCT and pixel hits ≥ 4	same ID track requirement as identified muons
$\lvert z_0 \rvert < 1\ mm$	$\lvert z_0 \rvert < 1\ mm$
Isolation ($\Delta R < 0.3$)	Isolation ($\Delta R < 0.3$)
Track: $(\sum p_{\mathrm{T}}^{\mathrm{trk}})/p_{\mathrm{T}} < 0.13$	–
Calorimeter $(\sum E_{\mathrm{T}}^{\mathrm{cell}})/p_{\mathrm{T}} < 0.3$	Calorimeter $(\sum E_{\mathrm{T}}^{\mathrm{cell}})/p_{\mathrm{T}} < 0.3$
Identified electrons removed	Identified muons removed

based on Monte Carlo simulation. In order to propagate the $t\bar{t}$ contribution of this sample to the signal region, the resulting number of events $N_{top}(\ell\ell + E_{\mathrm{T,rel}}^{\mathrm{miss}})$ is scaled by the probability that no jet is reconstructed in these events. This probability is estimated as the square of the efficiency for a top-quark decay to satisfy the jet veto criterion. This efficiency is in turn estimated using another control sample defined by the presence of an additional b-tagged jet. A correction is computed using simulated events to account for single-top background contributions.

The ratio of the number of data-driven top background events over the Monte Carlo prediction for the combined ee, $\mu\mu$ and $e\mu$ channels is estimated to be 1.04±0.07(stat)±0.17(syst).

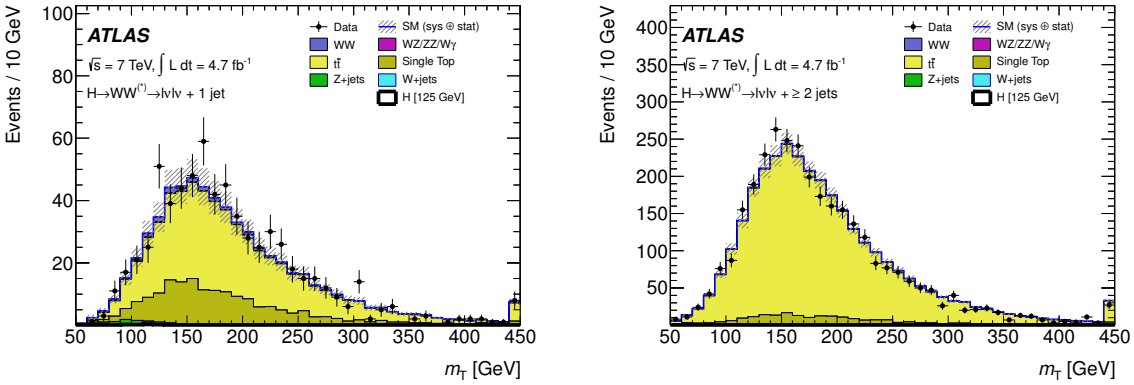


Figure 6.16.: Distributions of the transverse mass, m_{T}, in a top control region defined by requiring a b-tagged jet after the one jet requirement in the H+1 jet analysis (left) and after the central jet veto selection in the H+ $\geq$ 2 jets analysis (right). The lepton flavours are combined. The hatched area indicates the total uncertainty on the background prediction. The final bin includes the overflow.

In the H+1 jet and H+ $\geq$ 2 jets analyses, the top background MC prediction is normalised to data using control samples defined by reversing the b-jet veto and removing the requirement on $\Delta\phi_{\ell\ell}$ and $m_{\ell\ell}$. The resulting samples are dominated by both top-quark pair and single top processes with little contribution from other sources as can be seen in Fig. 6.16 and the expectations can be normalised to data. The resulting normalisation factors in the 1-jet channel are 1.04 ± 0.13 for the ee final state, 1.16 ± 0.10 for the $\mu\mu$ final state and 1.08 ± 0.06

for the $e\mu$ final state. For the 2-jets channel, the normalisation factor has been determined as a function of the dijet invariant mass. In the signal region, defined by $m_{jj} > 500$ GeV, the data tends to be lower than the MC prediction by about 20% although the statistical significance is limited.

WW Control Samples

The WW background Monte Carlo prediction in the H+0 jets and H+1 jet analyses, summed over lepton flavours, is normalised using control regions. These are defined with the nominal selection except that the $\Delta\phi_{\ell\ell}$ and m_{T} requirements are dropped and a lower bound on $m_{\ell\ell} > 80$ GeV is imposed to select a signal depleted region. Figure 6.17 shows the m_{T} distributions for data and Monte Carlo in the WW control regions. The contribution from WW is about 69% of the total events in the zero jet control region and about 47% for the one jet control region. The numbers of events in the WW control regions predicted by Monte Carlo agree well with the numbers observed in data, as can be seen in Table 6.13. The normalisation factors are defined by the ratio between the number of data events, with all backgrounds except the WW background subtracted using MC, to the number of predicted WW events in the control region. The normalisation factor for all lepton flavours combined for the zero and one jet channels is estimated to be 1.052 ± 0.064 and 0.952 ± 0.137, respectively.

This control region is used only for the low $\mathrm{m_H}$ selection in the zero and one jet analyses. For the intermediate and high $\mathrm{m_H}$ region, it is not possible to identify a signal free control region, therefore the WW background is directly taken from Monte Carlo simulation with the associated theoretical uncertainties taken into account. For the H+ ≥ 2 jets analysis, a signal depleted region with a sufficient number of WW events to extract an accurate estimate of this background cannot be determined. Thus, the background contribution is predicted as well using Monte Carlo simulation.

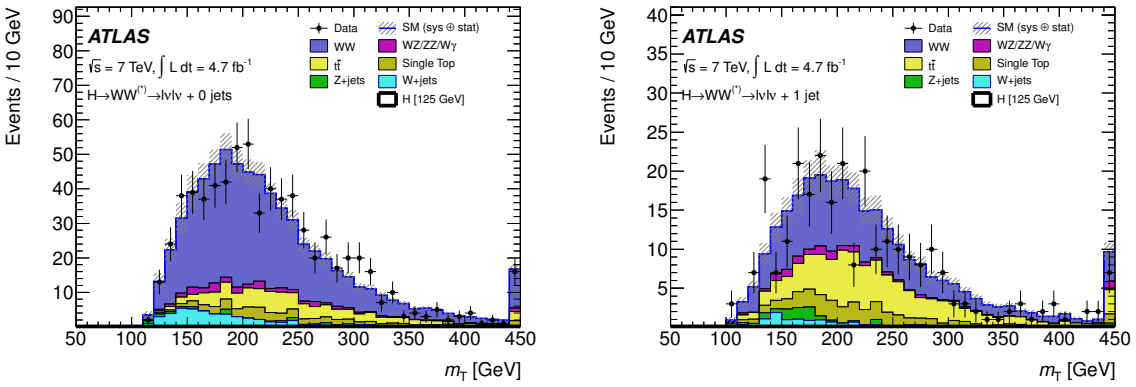


Figure 6.17.: Distributions of the transverse mass, m_{T}, in the WW control region defined by requiring $m_{\ell\ell} > 80$ GeV after the zero jet requirement in the H+0 jets analysis (left) and after the $Z \to \tau\tau$ veto in the H + 1 jet analysis (right). The lepton flavours are combined. The hatched area indicates the total uncertainty on the background prediction. The final bin includes the overflow.

Table 6.8.: Relative systematic uncertainties on the predicted number of signal events [165–167].

Source of Uncertainty on cross section	Signal m_H =125 GeV and m_H =240 GeV	Signal m_H =600 GeV
QCD renormalisation/factorisation scale		
0-jet analysis	± 21%	± 42%
1-jet analysis	± 31%	± 31%
2-jet analysis (VBF signal)	± 4%	± 4%
2-jet analysis (ggF signal)	± 25%	± 25%
Higgs boson lineshape description in POWHEG MC ggF and VBF signal process	150 % × $(m_H/ 1\ \text{TeV})^3$	
PDF uncertainties		
ggF signal process	± 8%	± 8%
VBF signal process	± 2%	± 4%

6.6. Systematic Uncertainties for the 2011 Analysis

There are two essentially different types of uncertainties that can affect the predicted numbers of signal and background events as well as the shape of the m_T distribution. Uncertainties arising from a finite number of events a measurement or prediction is based on are called *statistical uncertainties*. The number of observed events is limited on the experimental side by the amount of recorded collision data. On the contrary, the number of predicted events is limited by the extent of the samples obtained by Monte Carlo simulation. This is mainly restricted by the time that is required to simulate and reconstruct events with the full simulation of the ATLAS detector.

On the other hand, *systematic uncertainties* are caused by a limited knowledge of parameters and objects involved in the analysis. These uncertainties can not be reduced by increasing the number of measurements or events. Systematic uncertainties can arise from both experimental and theoretical sources. The dominant contributions to the systematic uncertainties in the context of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis are described in the following paragraphs.

6.6.1. Theoretical Uncertainties on the Signal Production

Theoretical uncertainties on the signal production cross sections mainly arise from the theoretical calculations and their dependence on parton distribution functions. Uncertainties on the signal production cross sections are determined following Refs. [44] and [165]. The QCD renormalisation and factorisation scales are independently varied up and down by a factor of two. The resulting uncertainties on the inclusive ggF signal production cross section are assumed to be independent of those obtained on the cross sections for production in association with at least one or two jets. These results are used to determine the uncertainties on the cross sections in exclusive jet multiplicity bins including anti-correlations induced by migrations between different jet multiplicities. The relative cross section uncertainties depend on m_H and are given in Table 6.8. The impact of the scale variations on both the VBF signal cross section and the jet veto acceptance, to which the 2-jet analysis is mainly sensitive, is estimated to be ±4%. In the 2-jet channel, around 25% of the signal events are produced via gluon fusion and the uncertainty on the ggF signal cross section is about ±25%. For the high mass range, an additional uncertainty due to the Higgs boson lineshape description in the

POWHEG MC generator (see Refs. [44, 168, 169]) is added in quadrature for both the ggF and the VBF channel. PDF uncertainties are estimated, following Refs. [32, 157] and [170, 171], using the envelopes of error sets as well as different PDF sets, applied separately to quark-quark, quark-gluon and gluon-gluon initiated processes. The relative PDF uncertainty on the dominant ggF process and the VBF uncertainty are given in Table 6.8.

Uncertainties on the modelling of signal and background processes are estimated by using alternative generators, such as MC@NLO for the ggF process, ALPGEN for WW production, POWHEG for $t\bar{t}$ production and PYTHIA for the Z/γ^*+ 2 jets process. The uncertainties associated with the underlying event and parton showering are taken into account in the acceptance uncertainty, but they are negligible compared to the scale uncertainties on the cross sections in exclusive jet bins.

6.6.2. Experimental Uncertainties

The acceptance of data events is affected by various insufficiencies related to the experimental measurements of the objects involved in the signature of the selected events. The observed events in data are compared to predictions from Monte Carlo simulation which are corrected to some extent for an imperfect modelling of the individual objects. These corrections are usually extracted by data-driven techniques and uncertainties are assigned to the correction factors. The uncertainties related to the trigger selection, the lepton identification and the b-tagging efficiencies are determined by varying the corresponding scale factors within their uncertainties. Similarly, the uncertainties on the energy and momentum scale and resolution related to the reconstruction of electrons, muons and jets are assessed by varying the corresponding calibration parameters within their uncertainties. The uncertainties related to pile-up contributions are evaluated by varying their effect on low-energy calorimeter depositions. Finally, to estimate the uncertainties on the missing transverse energy, the contributing terms involved in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation are varied up and down. The impact of the most important sources of experimental systematic uncertainties are summarised in Table 6.9.

The main experimental uncertainties are related to the jet energy scale, which is determined from a combination of test beam, simulation and in situ measurements [119]. The uncertainty on the jet energy scale varies from $\pm 14\%$ to $\pm 2\%$ as a function of jet p_{T} and η for jets with $p_T > 25$ GeV and $|\eta| < 4.5$. An additional contribution to the jet energy scale uncertainty arises from pile-up and is estimated to vary between $\pm 5\%$ and $\pm 0.5\%$. The second largest uncertainty is related to the jet energy resolution and its magnitude rises up to $\pm 5\%$ (see Table 6.9). The reconstruction, identification and trigger efficiencies for electrons and muons, as well as their momentum scales and resolutions, are estimated using $Z \to \ell^+\ell^-$, $J/\psi \to \ell^+\ell^-$ and $W \to \ell\nu$ decays (see Sections 4.2.4 and 4.2.3). All of the resulting uncertainties are smaller than $\pm 1\%$ with the exception of the uncertainty on the electron efficiency, which varies from $\pm 2\%$ to $\pm 5\%$ as a function of lepton p_{T} and η. The uncertainties related to the jet energy and lepton momentum scales are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ computation. Additional contributions arising from jets with $p_{\mathrm{T}} < 20$ GeV and from low-energy calorimeter depositions not associated with reconstructed physics objects are included as well. The uncertainties on the b-tagging efficiency varies from $\pm 4.8\%$ to $\pm 13.7\%$ depending on the p_{T} of the jet. The uncertainty on the measured integrated luminosity is estimated to $\pm 3.9\%$ based on measurements detailed in Ref. [142].

For the purpose of visualisation the effect of systematic uncertainties on Monte Carlo distributions, the individual sources of theoretical and experimental uncertainties are considered to be uncorrelated. The variations are performed independently and their impact on the Monte Carlo event yields is summed up.

Table 6.9.: Experimental sources of systematic uncertainty per object or event.

Source of Uncertainty	Treatment in the analysis
Jet Energy Resolution (JER)	MC jet resolution smeared using jet p_T, η-dependent parametrisation, resolution varies from 25 % to 5 % **uncertainty on JER < 5 %**
Jet Energy Scale (JES)	**global JES:** < 14% for jet $p_T > 25$ GeV and $\|\eta\| < 4.5$ pile-up: < 5% for jet $p_T > 20$ GeV
Electron Selection Efficiency	Separate systematics for electron identification, reconstruction and isolation, added in quadrature Identification: 8% for $p_T < 15$ GeV, decreasing to 1% for $p_T > 30$ GeV in the central region Reconstruction: 0.6 - 1.2% for $p_T > 15$ GeV trigger: 1% uncertainty **Total uncertainty of 2-5%** depending on η and E_T
Electron Energy Scale	Uncertainty **smaller than 1%**, depending on η and E_T
Electron Energy Resolution	Energy varied within its uncertainty, **0.6% of the energy** at most
Muon Selection Efficiency	**0.3-1%** as a function of η and p_T reconstruction smaller than 1%
Muon Momentum Scale and Resolution	Uncertainty **smaller than 1%**
b-tagging Efficiency	p_T dependent scale factor uncertainties, **4.8 - 13.7%**
Missing Transverse Energy	Jet energy scale and lepton momentum scale uncertainties are propagated to $E_{\mathrm{T}}^{\mathrm{miss}}$ computation additional contributions from jets with $p_T < 20$ GeV and low-energy depositions not associated with reconstructed physics objects **Total uncertainty of 1-8%**
Event pile-up	modelling of event pile-up contributions estimated by varying their effect on low-energy calorimeter depositions **Total uncertainty of 0.5-5%**
Luminosity	**3.9%** (taken from Ref. [142])

In addition to the systematic effects covered in Table 6.9, there are uncertainties arising from the data-driven estimation of background processes. The background estimates all follow the same pattern of using a control region in data which is expected to be dominated by the process of interest. It should be noted that correlations between systematic uncertainties on different backgrounds are taken into account in the statistical analysis of the dataset. All the signal and control regions are used as input to the final fit to estimate the signal and background contributions from data as described in Section 6.7.

6.6.3. Dominant Systematic Uncertainties on Different Background Processes

The total uncertainty on the m_{T} distribution predicted by Monte Carlo is dominated by the uncertainty on the normalisation of each background component. The dominant sources of systematic uncertainties affecting the normalisation of the W+jets, top and WW background processes in the signal region are discussed in the following. It should be noted that the presented uncertainties are informative but the correlations between nuisance parameters

and background processes are better taken into account concerning the final uncertainties by the likelihood fit procedure described in Section 6.8.

W+jets Background

The uncertainty on the fake factor is the main uncertainty on the W + jets background estimate. It is dominated by differences in the jet composition between dijet and W + jets samples as observed in Monte Carlo simulation. The total uncertainty comprises the impact of the trigger bias, data sample dependence and the subtraction of the contribution from real leptons from leptonic W boson and Z boson decays. The total uncertainty on the W + jets background comprises the statistical uncertainty in the W + jets control region and the uncertainty on the fake factor and is estimated to amount to approximately ±60 %.

Top and WW Background

The Monte Carlo prediction of the top background component is corrected by the ratio of observed data events to events from Monte Carlo prediction in the control regions. A main component of this method is the correlation between event yields in the signal and control regions. This can potentially be different between Monte Carlo simulation and data. The size of the uncertainty on the transfer of information from the control region to the signal region is obtained by comparing the nominal ratio value to the ratios of signal to control region yields that are varied by different theoretical and experimental sources. The observed differences are added up in quadrature. The resulting relative uncertainties on the normalisation of the WW and top processes in the signal region are mainly due to the transport of information from the control region.

The total uncertainty on the predicted WW background in the signal region, including both statistical and systematic effects, is ±9% for the H + 0 jets which is dominated by the statistical uncertainty on the normalisation to the WW control region. In the H + 1 jet channel, the total uncertainty on the WW background amounts to ±22%. It is dominated by the uncertainty on the normalisation factor which in turn is dominated by the large uncertainty on the top background extrapolated to the WW control region.

The total uncertainties on the estimated top-quark background in the 1-jet and 2-jet signal regions amount to ±23% and ±40%, respectively. These sizeable uncertainties are caused by the extrapolation from the control region characterised by a b-tag requirement to the signal region featuring a veto on b-tagged jets. The non-negligible uncertainty related to the b-jet tagging efficiency in combination with a very efficient tagging working point are the dominant sources. Concerning the H + 0 jets analysis, the total uncertainty on the top-quark background estimate in events with no identified, accompanying jet is ±22%.

It should be pointed out once again that the final uncertainty is assessed more correctly by the likelihood fit procedure used to perform the statistical analysis since correlations are taken into account.

Z/γ^*+jets Background

The estimate of the Z/γ^* + jets background contribution in the signal region relies on the assumption that $m_{\ell\ell}$ and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ are uncorrelated. This assumption has been checked with closure tests comparing the ratio $B^{MC} \times C^{MC}/D^{MC}$ to A^{MC} based on Monte Carlo simulation. The observed discrepancies are used as a measure of the systematic uncertainty on the method. The total uncertainty on the Z/γ^* + jets background in the signal region com-

prises the statistical uncertainty in the individual regions and the systematic uncertainties and amounts to ±38% and ±33% in the H + 0 jets and H + 1 jet channels, respectively.

6.7. Analysis Results of the 2011 Dataset

This Section shows the final results obtained in the search for the Higgs boson in the WW decay mode using the full dataset of 4.7 fb^{-1} taken in 2011 at $\sqrt{s}$ = 7 TeV. Table 6.10 shows the numbers of events expected from signal and background and observed in data at several stages of the selection for the H + 0 jets, H + 1 jet and the H+ ≥ 2 jets channels for all lepton flavours combined. The numbers split up for the ee, $e\mu$ and $\mu\mu$ final state are given in Appendix D. Table 6.11 shows the numbers of events expected from signal and background and observed in data, after the application of all selection criteria and split up into lepton flavours and jet multiplicities. The uncertainties reflect only the limited number of events available in the MC simulation and of the control regions used to normalise the dominant background contributions. Most of the events arise from the H + 0 jets channel in the $e\mu$ final state. The events expected from signal and background and observed in data in the WW and top control regions are shown in Table 6.12 for all lepton flavours combined. The numbers split up into ee,$e\mu$ and $\mu\mu$ final state are given as well in Appendix D. Figure 6.18 shows the transverse mass distribution for the ee, $\mu\mu$ and $e\mu$ channels combined in the H + 0 jets and H + 1 jet channel after all cuts are applied. It is evident from the

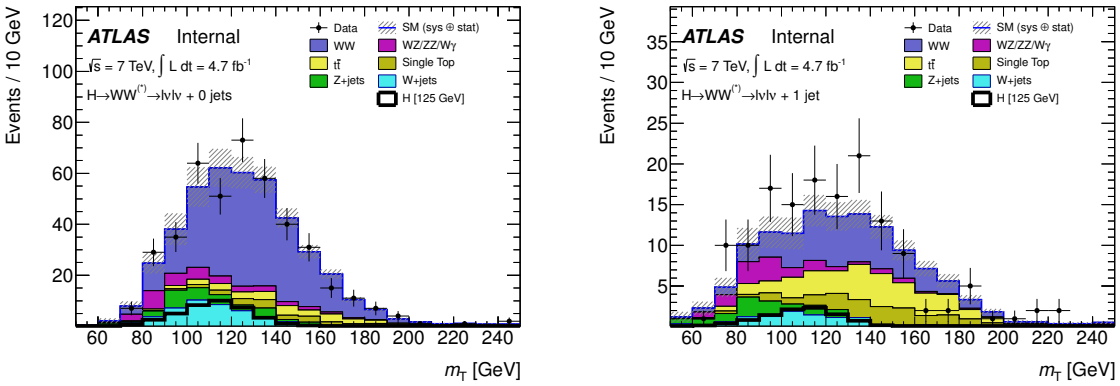


Figure 6.18.: Transverse mass, m_{T}, distributions in the 0-jet (left) and 1-jet (right) channels, for events satisfying all criteria for the low m_{H} selection. The lepton flavours are combined. The expected signal for a SM Higgs boson with m_{H} = 125 GeV is superimposed. The hatched area indicates the total uncertainty on the background prediction.

distributions that the dominant background contribution in the H + 0 jets channel arises from continuum WW production with smaller contributions from top and W + jets events. All the other background sources are of the same size as the expected signal of a Higgs boson with m_{H} = 125 GeV. The similar size of the expected signal and background contributions from Z/γ^* + jets, top and W + jets processes implies that additional efforts will be needed to decrease this contamination and to improve the sensitivity of the analysis. In the H + 1 jet and H+ ≥ 2 jets channels, the dominant backgrounds are top and WW events with equal contributions. In the H+1 jet channel, the background contributions from Z/γ^*+jets, W+jets and $WZ/ZZ/W\gamma$ production give approximately the same number of events as the signal

itself. Increasing the purity of the signal region would equally lead to a better sensitivity in this channel. Similarly to the results in the H + 0 jets channel, the ratio between the observed data and the expected background processes is consistently larger than one, but the deviations are not statistically significant.

To reflect better the sensitivity of the analysis, an additional mass-dependent cut on m_{T} is applied: $0.75 \cdot m_H < m_{\mathrm{T}} < m_H$ for $m_H = 125$ GeV and $0.6\, m_H < m_{\mathrm{T}} < m_H$ for $m_H = 240$ GeV and the results are shown in Table 6.13 for all lepton flavours combined. The uncertainties shown in this table include those of Table 6.9 and Table 6.8 as well as the systematic uncertainties discussed in Section 6.6. They are constrained by the use of the control regions discussed in Section 6.5.6. These uncertainties are also used in the statistical analysis presented in the following. The numbers indicate that the H + 0 jets channel has the largest number of data events and a better signal to background ratio compared to the H + 1 jet channel. Once all cuts are applied, there are zero (two) events observed in data for the low (intermediate) mass selection. Nevertheless, the H+ $\geq$ 2 jets channel has in fact a better signal to background ratio than both the H + 0 jets and H + 1 jet channels. This channel suffers from the small cross section of the VBF Higgs boson production mechanism and thus does not contribute significantly to the sensitivity of the analysis of the current dataset. With a larger dataset, this channel will play an increasingly important role.

Table 6.14 shows the main relative uncertainties on the predicted numbers of signal ($\mathrm{m_H}$ = 125 GeV) and background events for the H + 0 jets and H + 1 jet analyses. Similarly to Table 6.13, the additional m_{T} cut is applied and the constraints from control regions are included. The sources of theoretical and experimental uncertainties are described in Section 6.6 and only the dominant contributions are listed here. The numbers suggest that a reduction of the uncertainty on the fake factor used in the W + jets estimate will improve the sensitivity of the analysis for both the H + 0 jets and H + 1 jet channel.

Table 6.10.: The expected numbers of signal (for $m_H = 125$ GeV) and background events as well as the events observed in data at several stages of the selection. The $W+$jets background is estimated from data, whereas MC predictions normalised to data in control regions are used for WW, Z/γ^*+jets, $t\bar{t}$, and $tW/tb/tqb$ processes. The column "Top" refers to the sum of the contributions from single top and top quark pair production. Contributions from other background sources are taken directly from MC predictions. Only statistical uncertainties associated with the number of events in the MC samples and the data control regions are shown. Starting from jet multiplicity specific selection stages, the numbers for the SM WW and top background contributions are corrected using the normalisation factors derived from the corresponding control samples. In the same flavor channels, the Z/γ^*+jets contribution is corrected using the ABCD method.

	Signal	WW	$WZ/ZZ/W\gamma$	Top	Z/γ^*+jets	$W+$jets	Total Bkg.	Observed
Lepton p_{T}	181 ± 0.40	3871 ± 9.55	3214 ± 20.7	18086 ± 29.9	2800788 ± 1309	3160 ± 37.1	2829118 ± 1310	2823123
Opposite Charge	178 ± 0.39	3856 ± 9.53	2734 ± 15.8	17975 ± 29.8	2793618 ± 1308	2374 ± 36.2	2820557 ± 1309	2816240
$m_{\ell\ell} > 10$ (12) GeV	175 ± 0.39	3848 ± 9.52	2716 ± 15.7	17944 ± 29.8	2790757 ± 1307	2194 ± 35.9	2817460 ± 1308	2806551
Z veto (for $ee, \mu\mu$)	173 ± 0.39	3421 ± 8.97	560 ± 8.00	15999 ± 28.1	278757 ± 404	1403 ± 20.7	300140 ± 406	298691
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 25$ (45) GeV	90.8 ± 0.28	1829 ± 6.62	178 ± 4.82	8162 ± 20.5	1523 ± 27.9	342 ± 5.92	12033 ± 36.1	12231
0j: Jet veto	56.7 ± 0.23	1273 ± 5.58	96.6 ± 3.71	269 ± 3.96	1039 ± 23.7	217 ± 3.63	2894 ± 25.2	2849
0j: $m_{\ell\ell} < 50$ GeV	45.2 ± 0.21	312 ± 2.79	40.8 ± 2.77	48.2 ± 1.76	168 ± 8.81	69.5 ± 2.00	639 ± 10.0	645
0j: $p_T^{\ell\ell} > 30$ (45) GeV	40.1 ± 0.19	282 ± 2.65	34.8 ± 2.51	45.0 ± 1.70	28.0 ± 3.77	49.2 ± 1.52	439 ± 5.72	443
0j: $\Delta\phi_{\ell\ell} < 1.8$ rad	39.0 ± 0.19	276 ± 2.62	33.3 ± 2.43	44.5 ± 1.69	27.7 ± 3.76	44.1 ± 1.42	425 ± 5.64	429
0j: $0.75 \cdot \mathrm{m_H} < \mathrm{m_T} < \mathrm{m_H}$	25.9 ± 0.15	108 ± 1.66	11.7 ± 1.29	11.2 ± 0.85	14.1 ± 2.54	27.4 ± 1.11	172 ± 3.58	174
1j: $N_{\mathrm{jets}} = 1$	22.7 ± 0.13	343 ± 2.58	55.8 ± 2.66	1874 ± 10.3	357 ± 12.7	85.4 ± 2.90	2716 ± 17.0	2706
1j: b-jet veto	20.9 ± 0.13	319 ± 2.48	52.0 ± 2.55	550 ± 5.68	332 ± 12.2	75.9 ± 2.47	1330 ± 14.2	1369
1j: $p_T^{tot} < 30$ GeV	14.0 ± 0.10	226 ± 2.10	34.1 ± 2.16	261 ± 3.96	108 ± 6.23	37.4 ± 1.69	666 ± 8.15	684
1j: $Z \to \tau\tau$ veto	14.0 ± 0.10	220 ± 2.07	33.6 ± 2.16	250 ± 3.88	84.6 ± 5.70	37.7 ± 1.66	627 ± 7.70	644
1j: $m_{\ell\ell} < 50$ GeV	10.9 ± 0.09	48.8 ± 0.97	13.5 ± 1.57	50.4 ± 1.82	23.6 ± 2.73	12.0 ± 0.96	148 ± 3.88	170
1j: $\Delta\phi_{\ell\ell} < 1.8$ rad	10.1 ± 0.09	44.5 ± 0.92	13.1 ± 1.57	48.2 ± 1.78	10.0 ± 1.73	9.9 ± 0.82	126 ± 3.19	145
1j: $0.75 \cdot \mathrm{m_H} < \mathrm{m_T} < \mathrm{m_H}$	6.3 ± 0.07	16.0 ± 0.55	4.6 ± 0.97	12.1 ± 0.90	4.6 ± 1.19	5.0 ± 0.54	42.4 ± 1.94	56
$\geq$ 2j: $N_{\mathrm{jets}} \geq 2$	11.42 ± 0.09	142 ± 1.73	25.5 ± 1.53	6278 ± 18.0	120 ± 6.68	39.6 ± 3.66	6605 ± 19.7	6676
$\geq$ 2j: CJV	8.95 ± 0.08	113 ± 1.55	19.5 ± 1.44	3517 ± 13.4	89.5 ± 5.94	25.5 ± 2.60	3765 ± 15.0	3811
$\geq$ 2j: b-jet veto	7.56 ± 0.07	97.8 ± 1.42	17.5 ± 1.38	404 ± 4.70	77.4 ± 5.48	18.9 ± 1.58	615 ± 7.65	667
$\geq$ 2j: opp. hemispheres	4.17 ± 0.05	46.0 ± 0.97	7.3 ± 0.84	170 ± 3.01	31.6 ± 3.39	8.6 ± 1.06	264 ± 4.83	269
$\geq$ 2j: $\lvert\Delta\eta_{jj}\rvert > 3.8$	1.81 ± 0.02	8.4 ± 0.41	0.9 ± 0.17	25.3 ± 1.05	5.8 ± 1.68	1.7 ± 0.41	42.2 ± 2.07	40
$\geq$ 2j: $m_{jj} > 500$ GeV	1.28 ± 0.02	3.9 ± 0.28	0.4 ± 0.11	11.4 ± 0.66	0.7 ± 0.35	0.9 ± 0.32	17.3 ± 0.87	13
$\geq$ 2j: $p_T^{tot} < 30$ GeV	0.95 ± 0.01	2.5 ± 0.22	0.2 ± 0.08	4.9 ± 0.44	0.3 ± 0.20	0.4 ± 0.22	8.3 ± 0.58	6
$\geq$ 2j: $Z \to \tau\tau$ veto	0.94 ± 0.01	2.4 ± 0.21	0.2 ± 0.08	4.9 ± 0.44	0.2 ± 0.16	0.3 ± 0.21	8.0 ± 0.56	6
$\geq$ 2j: $m_{\ell\ell} < 80$ GeV	0.94 ± 0.01	1.1 ± 0.15	0.1 ± 0.06	1.7 ± 0.28	0.2 ± 0.16	0.2 ± 0.18	3.2 ± 0.40	2
$\geq$ 2j: $\Delta\phi_{\ell\ell} < 1.8$ rad	0.82 ± 0.01	0.8 ± 0.12	0.1 ± 0.06	1.0 ± 0.23	0.0 ± 0.00	(-0.0 ± 0.03)	1.8 ± 0.26	1

Table 6.11.: The expected numbers of signal (for $m_H = 125$ GeV) and background events as well as the events observed in data for each lepton channel after the $\Delta\phi_{\ell\ell}$ cut. Only statistical uncertainties associated with the number of events in the MC samples and the data control regions are shown.

Lepton Channels	0-jet ee	0-jet $\mu\mu$	0-jet $e\mu$	1-jet ee	1-jet $\mu\mu$	1-jet $e\mu$
Total bkg.	60 ± 5	116 ± 10	249 ± 12	19 ± 2	34 ± 4	72 ± 6
Signal	4.0 ± 0.1	9.4 ± 0.1	25.7 ± 0.2	1.2 ± 0.1	2.5 ± 0.1	6.4 ± 0.1
Observed	52	138	239	19	36	90

Table 6.12.: The expected numbers of signal (for $m_H = 125$ GeV) and background events as well as the events observed in data in the control regions. Only statistical uncertainties associated with the number of events in the MC samples and the data control regions are shown. Here, with the exception of W+jets, no normalisation scale factors are applied to the expected numbers. The bottom part of the table lists the number of expected and observed events for each lepton channel after the $\Delta\phi_{\ell\ell}$ cut.

Control Regions	Signal	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
WW 0-jet	0.3 ± 0.1	471 ± 3	26 ± 1	87 ± 2	42 ± 2	7 ± 2	49 ± 2	682 ± 5	697
WW 1-jet	0.1 ± 0.1	128 ± 2	12 ± 1	89 ± 2	34 ± 2	9 ± 2	11 ± 1	282 ± 4	270
Top 1-jet	1.2 ± 0.1	20 ± 1	1.9 ± 0.5	434 ± 4	169 ± 4	7 ± 2	4 ± 1	635 ± 6	676
Top 2-jet	0.1 ± 0.1	0.4 ± 0.1	negl.	10.0 ± 0.7	1.0 ± 0.3	negl.	negl.	11.4 ± 0.7	10

Table 6.13.: The numbers of events expected from signal ($m_H = 125$ GeV and 240 GeV) and background and observed in data after the full low m_H and intermediate m_H selections are shown, including a cut on the transverse mass of $0.75\,m_H < m_T < m_H$ for $m_H = 125$ GeV and $0.6\,m_H < m_T < m_H$ for $m_H = 240$ GeV. The uncertainties shown are the combination of the statistical and all systematic uncertainties, taking into account the constraints from control samples. All numbers are summed over lepton flavours.

	m_H	Signal	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
0-jet	125 GeV	26 ± 7	108 ± 12	12 ± 2	7 ± 2	5 ± 1	14 ± 6	27 ± 16	172 ± 21	174
	240 GeV	61 ± 17	450 ± 49	24 ± 3	73 ± 15	42 ± 9	6 ± 2	36 ± 24	632 ± 63	627
1-jet	125 GeV	6 ± 2	16 ± 5	5 ± 2	8 ± 2	4 ± 2	5 ± 2	5 ± 3	42 ± 7	56
	240 GeV	24 ± 9	95 ± 21	9 ± 1	84 ± 21	39 ± 15	5 ± 1	8 ± 7	241 ± 41	232
2-jet	125 GeV	0.5 ± 0.2	0.2 ± 0.2	negl.	0.2 ± 0.1	negl.	0.0 ± 0.1	negl.	0.4 ± 0.2	0
	240 GeV	2.6 ± 0.4	1.2 ± 0.8	0.1 ± 0.1	2.2 ± 1.0	0.3 ± 0.2	negl.	0.1 ± 0.1	3.9 ± 1.4	2

Table 6.14.: Main relative systematic uncertainties on the predicted numbers of signal (m_H =125 GeV) and background events for the H + 0 jets and H + 1 jet analyses. The same m_T criteria as in Table 6.13 are imposed in addition to the low m_H signal selection criteria. All numbers are summed over lepton flavours.

Source (H + 0 jets)	Signal (%)	Bkg. (%)	Source (H + 1 jet)	Signal (%)	Bkg. (%)
Inclusive QCD scale			1-jet QCD		
ren./fact. scale	19	0	ren./fact. scale	27	0
1-jet QCD			2-jet QCD		
ren./fact. scale	10	0	ren./fact. scale	15	0
W + jets fake factor	0	10	E_T^{miss}	8	3
PDFs	8	2	W + jets fake factor	0	7
WW normalisation	0	6	b-tagging efficiency	0	7
Jet energy scale	6	0	PDFs	7	1

6.8. Statistical Treatment and Results of the 2011 Dataset

All the details about the statistical analysis performed in the context of the search for the Standard Model Higgs boson in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay mode are documented in Ref. [172]. In the context of this thesis, the inputs to the statistical analysis were provided. It should be pointed out that the statistical treatment including the design of the likelihood function, the treatment of the systematic uncertainties and the fitting procedure itself were performed by other members of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis group. Nevertheless, the results are presented in this section to provide a complete picture of the analysis.

As already mentioned, the variable of interest in this analysis is m_T, the transverse mass of the dilepton and E_T^{miss} system. No cuts are applied to this quantity and a fit to the distribution is used instead for testing the agreement with the hypothesis of the presence or absence of a Standard Model Higgs boson. This gives increased sensitivity of the analysis compared to simply cutting on the distribution since the shape information is incorporated. It should be pointed out that the background distributions peak strongly in the middle of the transverse mass distributions shown in Fig. 6.18. The m_T distribution in the final signal region considered in the likelihood function is mapped to five bins in the H + 0 jets channel while in the H+1 jet channel three bins are used and the H+ $\geq$ 2 jets channel uses no binning at all due to the small number of events left over in the signal region. The bins are chosen such that the expected total background is uniformly distributed among them.

To test the statistical significance of the presence or absence of signal events in data, a binned likelihood function is used for the statistical analysis of the data. The procedure used to compute the exclusion limits is based on the modified frequentist method known as CLs [173, 174]. The likelihood function is constructed as the product of Poisson probability terms for the signal and control regions. Rather than using extrapolation coefficients to transfer the background contributions among the control regions and to the signal region, an equivalent parametrisation is used that employs strength parameters as follows:

$$\mathcal{L}(\mu, \mu_b) = P(N|\mu s + \mu_b b_{SR}^{exp}) \times P(M|\mu_b b_{CR}^{exp}) \tag{6.5}$$

where b_{SR}^{exp} and b_{CR}^{exp} are the expected background yields in the signal and control region determined by MC, μ is the signal strength parameter relating the signal cross section to the SM Higgs boson cross section prediction and μ_b is the background strength parameter. N is the measured number of events in the signal region while M is an auxiliary control region measurement.

The simple likelihood function is expanded as a product of Poisson probability terms for the signal and control regions in each lepton flavour channel, jet multiplicity and m_{T} bin. Uncertainties are treated as a set of nuisance parameters (referred to as $\vec{\theta}$) in the distributions with additional constraints. Due to event pile-up conditions changing throughout the data-taking periods (see Section 6.3.4) a progressive degradation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution has been observed. Thus the sensitivity can be improved by splitting the analysis, and linked to this the likelihood in datasets for the first 2.3 fb^{-1} and the remaining 2.4 fb^{-1}. The full likelihood can be written as:

$$\mathcal{L}(\mu,\vec{\theta}) = \{\prod_{l=BK,LM}\prod_{k=ee,e\mu,\mu\mu}\prod_{j=0}^{N_{jets}}\prod_{i=1}^{N_{bins}} P(N_{ijkl}|\mu s_{ijkl}+\sum_{m}^{N_{bg}} b_{ijklm})\} \times \{\prod_{i=1}^{N_{\theta}} N(\tilde{\theta}|\theta)\} \quad (6.6)$$

where $P(N_{ijkl}|\mu s_{ijkl}+\sum_{m}^{N_{bg}} b_{ijklm})$ are the Poisson probability distributions and $N(\tilde{\theta}|\theta)$ are the constraints which are either Gaussian or Poisson distributions in the analysis. The signal and background expectations are also functions of the nuisance parameters θ. These function are parametrised in a way that the response of s and b to each nuisance parameter is factorised from the nominal value of the expected rate, e.g. $s = s_0 \times \prod \nu(\theta)$, where the form of $\nu(\theta)$ depends on the systematic source. Nuisance parameters are treated in four categories:

- **flat systematic uncertainties:** for systematic uncertainties that do not change the shape of the transverse mass distribution, the nuisance parameter is taken as $\nu_{flat}(\theta) = \kappa^{\theta}$ with a Gaussian constraint.
- **shape systematic uncertainties:** for systematic uncertainties that also affect the shape of the transverse mass distribution, the nuisance parameter is split into a flat component affecting only the normalisation and treated as described above, and a shape component with a Gaussian constraint.
- **statistical uncertainty:** due to the limited size of the Monte Carlo (data) samples used to estimate the background contribution from the Monte Carlo (data-driven) prediction. The nuisance parameter is written as $\nu_{stat}(\theta) = \theta$ with a Poisson constraint.
- **uncertainty due to data control region:** originating from the limited size of the data sample in the control region constraining the background normalisation. The contamination due to both the signal and the other backgrounds has to be taken into account as well. The constraint is given by a Poisson distribution.

Due to the fact that each θ represents a different systematic source, the same θ can affect multiple signal and background rates. And similarly each signal or background component may possess a different set of nuisance parameters.

A test statistic q_μ is constructed using the profile likelihood

$$q_\mu = -2\ln\frac{\mathcal{L}(\mu,\hat{\hat{\theta}}_\mu)}{\mathcal{L}(\hat{\mu},\hat{\theta}_{\hat{\mu}})} \quad (6.7)$$

where $\hat{\mu}$ and $\hat{\theta}_{\hat{\mu}}$ are the parameters that maximise the likelihood and $\hat{\hat{\theta}}_\mu$ corresponds to the conditional maximum likelihood of θ given μ and the data. The constraint $0 \leq \hat{\mu} \leq \mu$ is applied to ensure that the signal is positive and to guarantee a one-sided confidence limit. Values are computed for $\hat{\theta}_0^{\mathrm{obs}}$ and $\hat{\theta}_\mu^{\mathrm{obs}}$ which maximise the likelihood for *background-only*

and *signal plus background* hypothesis, respectively. Sampling distributions $f(q|\mu, \hat{\theta}_{\mu}^{\text{obs}})$ and $f(q|0, \hat{\theta}_{0}^{\text{obs}})$ are generated for both scenarios. The p-value expresses the probability of how a given test statistics is compatible with data under the assumption that the *background-only* hypothesis is true. It is calculated by integrating over the according sampling distribution:

$$p_0 = P(q_\mu \leq \tilde{q}_\mu^{\text{obs}} | \text{background} - \text{only}) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(q_\mu|0, \hat{\theta}_0^{\text{obs}}) dq_\mu \tag{6.8}$$

For example, a p-value of 5 % means that the observed data would agree with the SM *background-only* hypothesis in about 5 % of the cases if the experiment would be repeated many times. Thus, the p-value for the signal strength μ is given by:

$$p_\mu = P(q_\mu \leq \tilde{q}_\mu^{\text{obs}} | \text{signal} + \text{background}) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(q_\mu|\mu, \hat{\theta}_\mu^{\text{obs}}) dq_\mu \tag{6.9}$$

A quantity, denoted as CL_s, is defined as the ratio of the two p-values $CL_s = p_\mu / p_0$. The 95 % Confidence Level upper limit on μ is the solution to $CL_s = 0.05$.

The vertical axis in Fig. 6.19 shows the 95% upper limits on the Higgs boson production cross section, divided by the expected cross section for Higgs production in the Standard Model, as a function of the Higgs boson mass (horizontal axis). The dotted black line in Fig. 6.19 indicates the average expected limit in the absence of a Higgs boson based on Monte Carlo simulation. The green and yellow bands indicate the corresponding $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands on the expected limit, respectively. The area where the dotted black line dips below the value of 1.0 marks the Higgs mass range that can be excluded based on the sensitivity of the analysis and the available amount of data. The solid curve reflects the observed experimental limits for the production of the Higgs boson of each possible m_H value. If the solid line dips below the value of 1.0, it can be concluded that the Higgs boson is not produced with the expected cross section for that mass of the Higgs boson. Those values of m_H are excluded with a 95% confidence based on the analysed data. In case the solid black line is above 1.0 and additionally above the dotted black line there might be a hint that a SM Higgs boson exists with a mass at that value. The solid black line in Fig. 6.19 is above 1.0 and at the upper edge of the green band. That means there may be a 68% certainty that the observed rate is above expectations. It could be either a hint for the presence of a Higgs boson of that mass or it could be a sign of background processes or systematic effects that are not well understood.

The observed upper limits shown in Fig. 6.19 exclude a Standard Model Higgs boson with a mass in the range from 133 GeV to 261 GeV at 95% CL, while the expected exclusion range in the absence of a signal is 127 GeV $\leq m_H \leq$ 233 GeV. It should be noted that the H + 0 jets, H + 1 jet and H+ $\geq$ 2 jets analyses and the lepton flavours are combined. Furthermore, Fig. 6.19 displays the observed probability under the assumption that the background-only hypothesis is true as a function of m_H. The dashed line shows the corresponding expectation for the signal+background hypothesis at the given value of m_H. The observed probability of background to produce a signal-like excess for m_H =125 GeV is given by the p_0-value of 0.22. This corresponds to the probability to find a Gaussian distributed variable 0.77 standard deviations above its mean. In summary, no significant excess of events over the expected background is observed over the entire mass range.

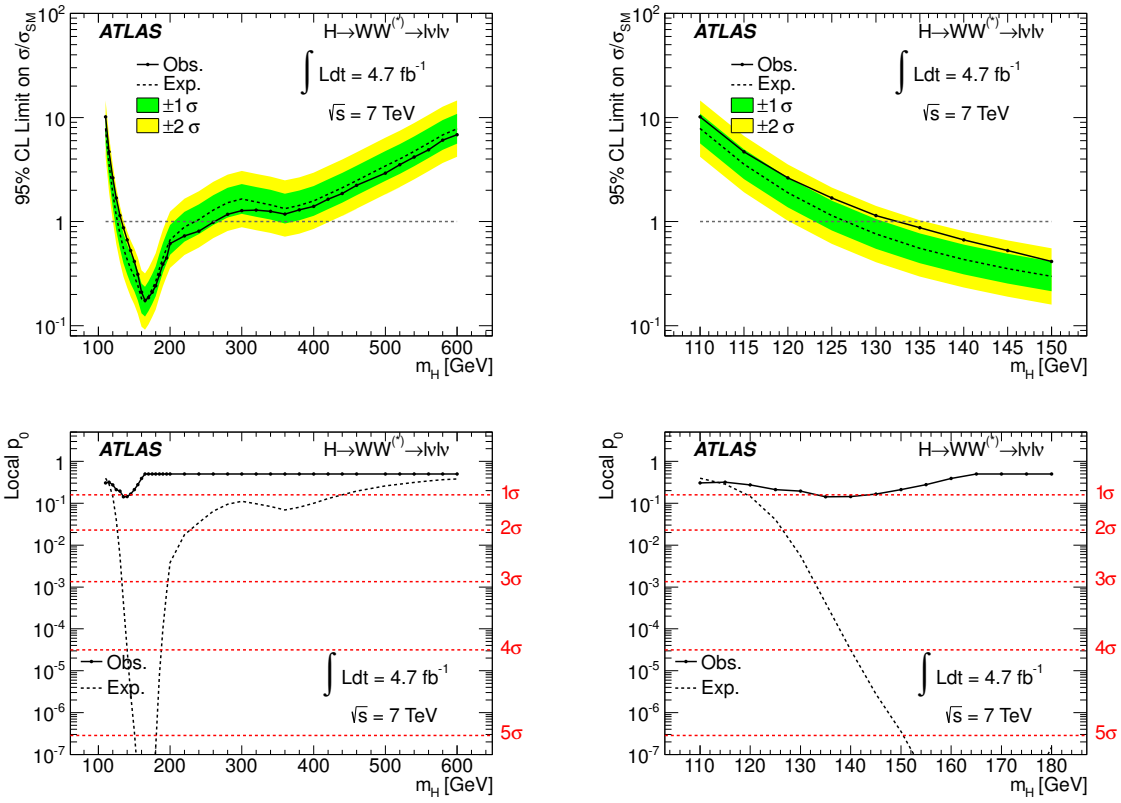


Figure 6.19.: Observed (solid) and expected (dashed) 95% CL upper limits on the Higgs boson production cross section, normalised to the SM cross section, as a function of m_H of the full mass range considered in this analysis (top left) and restricted to the range $m_H <$ 150 GeV (top right). The inner (green) and outer (yellow) regions indicate the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands on the expected limit. respectively. The observed local p_0 as a function of the hypothesised Higgs boson mass m_H (solid line) is shown for the whole mass range (bottom left) and for the mass range $m_H < 150$ GeV (bottom right). The dashed line shows the corresponding expectation for the signal+background hypothesis at the given value of m_H.

6.9. Analysis of the 2012 Dataset

The details about the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis performed on the 2012 dataset as published in Ref. [5] are presented in the following sections. The search focuses on the Higgs boson mass region around 125 GeV that has not been excluded previously by the analysis performed on the 2011 dataset. The analysis described here uses a dataset collected between the beginning of April to the middle of June 2012, which corresponds to 5.8 fb^{-1} of pp collision data at a centre-of-mass energy of 8 TeV (see Section 6.3). The selection requirements are mostly unchanged with respect to the 2011 analysis presented in the previous sections. The most significant modifications affect the trigger (see Section 6.3.2) and lepton isolation criteria that had to be tightened in order to cope with the higher level of pile-up. In the 2011 dataset the average number of interactions per bunch crossing was around 10 while in the 2012 dataset it increased to 19.5 as already mentioned in Section 6.3.4. In addition, this leads to significantly larger $Z/\gamma^* +$ jets background to the ee and $\mu\mu$ final states, since the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution degrades with increasing pile-up. Dedicated procedures are developed to estimate

this background and to keep it under control. But some more time and efforts are needed to put them finally into place. Since the opposite flavour channels feature the major part of the sensitivity of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ search, only this final state is considered in the analysis presented here. Furthermore, the opposite flavour channel is split into two exclusive lepton channels depending on the flavour of the leading lepton. In the following, $e\mu$ and μe will refer to events with a leading electron or muon, respectively.

Motivated by the 2011 combined Higgs boson searches [175], the 2012 early data analysis was modified to *blind* the regions with significant expected signal yields. That way, the selection requirements have been verified by impartial data to Monte Carlo comparison studies in control regions. At the end of June 2012 the agreement between data and the background model was judged to be reasonable and the decision to lift the restrictions and *unblind* the analysis was approved by the ATLAS collaboration.

The details about the *blinding* of the analysis are given in Section 6.10. The signal and background processes used for the 2012 analysis are briefly summarised in Section 6.11. The modifications to the object and event selection for the Higgs boson search in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channel are provided in Section 6.12. The changes concerning the data-driven background estimates are detailed in Section 6.12.3 followed by the impact of the experimental systematics in Section 6.13. The results of the *unblinded* analysis of the 2012 dataset are documented in Section 6.14. Finally, the results obtained at a centre-of-mass energy of 8 TeV on a dataset corresponding to 5.8 fb^{-1} are combined with the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ results extracted on the 2011 dataset in Section 6.15.

6.10. Blinding the 2012 Analysis

In order to exclude the kinematic regions where a signal might be expected, events are ignored if they pass the preselection (see Section 6.12) and the following criteria in addition:

- The azimuthal angular difference between the two leptons, $\Delta\phi_{\ell\ell}$, is required to be less than 1.8 radians.
- The dilepton invariant mass is required to satisfy $m_{\ell\ell} < 50$ GeV.
- The transverse mass m_{T} is required to by in the interval $82.5 < m_{\mathrm{T}} < 140$ GeV.
- There is either no additional jet in the event or the jet is not identified as originating from the decay of a b- or c-quark.

It was verified that the remaining contribution of $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ events to the total expected event yields, assuming a Higgs boson mass of 125 GeV, is below 2 % at every stage of the event selection. After excluding these events from the data and Monte Carlo samples, the simulation quality can be checked in the nominal control regions since the *blinding* criteria are constructed in a way that they do not affect these regions. The agreement between data and Monte Carlo expectations in control regions covering the major background contributions (WW, top , and W+jets/$W\gamma(^*)$) have been found to be sufficient (see Ref. [176]). All results presented in the following sections are obtained on the *unblinded* datasets.

6.11. Simulated Samples for the 2012 Analysis

The Monte Carlo generators used to model the signal and background processes are listed in Table 6.15. As already mentioned in Section 6.4, for most processes, separate programs

are used to generate the hard-scattering process and to model the parton showering, hadronisation, and the underlying events. Either PYTHIA or PYTHIA8 [177] are used for these latter three steps for the signal and some of the background processes. For most of the background processes the same generators are used to model the hard scattering process as for the 2011 analysis (see Section 6.4) except for ZZ, WZ and Wγ^* events where now POWHEG is used as a generator for the former and MADGRAPH for the latter ones.

Table 6.15.: Monte Carlo generators used to model the signal and background processes, and corresponding cross sections at $\sqrt{s} = 8$ TeV (given for both $m_H = 125$ GeV in the case of the signal processes). The number quoted for the inclusive $Z/\gamma^\star$ process is for generated dilepton masses greater than 10 GeV. The $Z^{(\star)}Z^{(\star)} \to 4l$ samples are generated with an invariant mass cut of $m_{\ell\ell} > 4$ GeV. For the $WZ^{(\star)}$ and $W\gamma^\star$ processes, MADGRAPH includes the interference between the $Z^{(\star)}$ and the $\gamma^\star$ and the boundary between the samples is at $m_{\ell\ell} = 7$ GeV. Leptonic decays of W bosons/Z bosons are always assumed, and the quoted cross sections include the branching ratios and are summed over lepton flavours. The exception is top-quark production for which inclusive cross sections are quoted.

Process	Generator	m_H (GeV)	$\sigma \cdot$ Br (pb)	Ref.	order
ggF	POWHEG [37]+PYTHIA8 [177]	125	0.441	[44]	NNLO
VBF	POWHEG [38]+PYTHIA8	125	$35 \cdot 10^{-3}$	[44]	NNLO
WH/ZH	PYTHIA8	125	$25 \cdot 10^{-3}$	[44]	NNLO
$q\bar{q}/qg \to WW$	MC@NLO [36]+HERWIG [35]		5.68	[32, 158]	NLO
$gg \to WW$	GG2WW [151]+HERWIG		0.16	[32, 158]	NLO
$t\bar{t}$	MC@NLO+HERWIG		238.1	[32, 159]	NNLO
tW/tb	MC@NLO+HERWIG		28	[160, 161]	NNLO
tqb	AcerMC [156]+PYTHIA		88	[160, 161]	NNLO
inclusive W	ALPGEN [149]+HERWIG		$37 \cdot 10^3$	[148]	NNLO
inclusive $Z/\gamma^\star$	ALPGEN+HERWIG		$16 \cdot 10^3$	[148]	NNLO
$Z^{(\star)}Z^{(\star)} \to 4l$	POWHEG+PYTHIA8		0.73	[32, 158]	NLO
$WZ^{(\star)}$	MADGRAPH [153, 154]+PYTHIA [34]		1.54	[32, 158]	NLO
$W\gamma^\star$	MADGRAPH [155]+PYTHIA		9.26	[32, 158]	NLO
$W\gamma$	ALPGEN+HERWIG		369	[32, 158]	NLO

It should be noted that the cross sections for WW, ZZ, $WZ^{(\star)}$, $W\gamma$ and $WZ^{(\star)}/W\gamma^\star$ processes are normalised to the MCFM [158] NLO predictions. The CT10 PDF set [157] is used for the POWHEG and MC@NLO samples, and CTEQ6L1 [152] is used for the ALPGEN, MADGRAPH and PYTHIA8 samples. The simulation incorporates a model of pile-up conditions in the 2012 data, including both the effects of multiple pp collisions in the same bunch crossing and in nearby bunch crossings.

Furthermore, concerning the signal processes, only the $H \to WW^{(\star)} \to e\nu\mu\nu$ mode is considered, including the small contribution from leptonic τ decays. The branching ratio for this decay as a function of m_H is calculated using the PROPHECY4F [178, 179] program, with HDECAY also used for calculating the total width.

6.12. Event Selection and Background Determination for the 2012 Analysis

The selection of $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ candidate events is not changed with respect to the one performed on the 2011 datasets (see Section 6.5) with the exception of small changes in the lepton selection that are detailed in Section 6.12.1. The methods to determine and normalise the background contributions for W + jets, top and WW events are the same as in the 2011 analysis. Concerning the Z/γ^* + jets background contribution, a validation region is used to check the Monte Carlo prediction. The extracted normalisation factors and uncertainties on the background estimates are briefly summarised in Section 6.12.3.

6.12.1. Object Selection for the 2012 Analysis

This section describes the modifications to the selection criteria for the objects used in the analysis of the 2012 dataset with respect to those used for the 2011 dataset and detailed in Section 6.5.1.

Electrons

The electron reconstruction algorithm used in the ATLAS experiment (as described in Section 4.2.4) is modified to use tracks that are refitted by the Gaussian Sum Filter (GSF) algorithm [180]. It is based on a non-linear generalisation of the Kalman Filter and helps in particular to account for bremsstrahlung losses. This bremsstrahlung recovery process helps to improve the electron four-momentum reconstruction at the interaction vertex. The electron candidates are identified using the *tight++* selection criteria (see Section 4.2.4) which was optimised for 2012 to cope with the increased level of pile-up and trigger rates. In addition to the set of standard criteria and similar to the 2011 analysis (see Section 6.5.1), not only impact parameter but also isolation cuts are applied in order to reduce in particular contributions from W + jets events. Due to the use of the GSF algorithm in the electron reconstruction, it was possible to tighten the impact parameter criteria with respect to the analysis performed on the 2011 dataset. Both the tracking and calorimeter based isolation criteria are optimised and tightened as well leading to p_{T} dependent requirements for the track isolation criteria. The changes of the 2012 selection criteria with respect to the ones used for the 2011 dataset are summarised in Table 6.16. There are small differences between data and Monte Carlo prediction arising from the identification, impact parameter and isolation selection requirements. They are accounted for with scale factors determined with a tag-and-probe technique performed on Z boson decays. The scale factors are found to be consistent with unity within 1-2 % [181].

Muons

Muons are reconstructed by the Staco combined algorithm (see Section 4.2.3) which was used for the 2011 analysis already. The geometrical acceptance is extended to $|\eta| < 2.5$ since the tracking is still sufficient to allow for a reliable reconstruction. In order to account for larger uncertainties on the longitudinal impact parameter z_0 at high values of pseudorapidity η, a cut on $|z_0 \cdot \sin\theta|$ is used for the 2012 analysis. Just as for the electron selection, both the tracking and calorimeter based isolation criteria are optimised and tightened resulting in p_{T} dependent requirements listed in Table 6.16. The isolation requirements are further enhanced by rejecting a muon candidate in case a reconstructed jet passing the selection criteria described in the following paragraph is present within a cone of $\Delta R = 0.3$ around

Table 6.16.: Summary of the modifications in the object selection criteria for the 2012 datasets compared to the 2011 analysis.

Object / Requirement	2011 dataset	2012 dataset
Muons		
Reconstruction algorithm	Staco combined	
Transverse momentum	$p_{\mathrm{T}} > 15$ GeV	
Geometrical acceptance	$\lvert\eta\rvert < 2.4$	$\lvert\eta\rvert < 2.5$
Impact parameter		
Transverse	$\lvert d_0/\sigma(d_0)\rvert < 3$	$\lvert d_0/\sigma(d_0)\rvert < 3$
Longitudinal	$\lvert z_0\rvert < 1\ mm$	$\lvert z_0 \cdot \sin\theta\rvert < 1\ mm$
Isolation ($\Delta R < 0.3$)		
Track	$\left(\sum p_{\mathrm{T}}^{\mathrm{trk}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.13$	$\left(\sum p_{\mathrm{T}}^{\mathrm{trk}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.15$ and $\left(\sum p_{\mathrm{T}}^{\mathrm{trk}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < \frac{0.01 \cdot p_{\mathrm{T}}}{\mathrm{GeV}} - 0.105$
Calorimeter	$\left(\sum E_{\mathrm{T}}^{\mathrm{cell}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.14$	$\left(\sum E_{\mathrm{T}}^{\mathrm{cell}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.2$ and $\left(\sum E_{\mathrm{T}}^{\mathrm{cell}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < \frac{0.014 \cdot p_{\mathrm{T}}}{\mathrm{GeV}} - 0.15$
Overlap removal (μ/j)	–	$\Delta R(\mathrm{muon}, \mathrm{jet}) < 0.3$
Electrons		
Identification criteria	tight++	
Geometrical acceptance	$\lvert\eta_{\mathrm{cluster}}\rvert < 2.47$, except $1.37 < \lvert\eta\rvert < 1.52$	
Transverse momentum	$p_{\mathrm{T}} > 15$ GeV	
Overlap removal (e/μ)	$\Delta R(\mathrm{electron}, \mathrm{muon}) < 0.1$	
Track reconstruction	Default	Gaussian Sum Filter
Overlap removal (e/e)	–	$\Delta R(\mathrm{electron}, \mathrm{electron}_{p_{\mathrm{T}}>}) < 0.1$
Impact parameter		
Transverse	$\lvert d_0/\sigma(d_0)\rvert < 10$	$\lvert d_0/\sigma(d_0)\rvert < 3$
Longitudinal	$\lvert z_0\rvert < 1\ mm$	$\lvert z_0 \cdot \sin\theta\rvert < 0.4\ mm$
Isolation ($\Delta R < 0.3$)		
Track	$\left(\sum p_{\mathrm{T}}^{\mathrm{trk}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.13$	$\left(\sum p_{\mathrm{T}}^{\mathrm{trk}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.12$ (0.16) for $p_{\mathrm{T}} < 25$ GeV ($p_{\mathrm{T}} \geq 25$ GeV)
Calorimeter	$\left(\sum E_{\mathrm{T}}^{\mathrm{cell}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.14$	$\left(\sum E_{\mathrm{T}}^{\mathrm{cell}}\right)_{\mathrm{C}}/p_{\mathrm{T}} < 0.16$
Jets		
Reconstruction algorithm	anti-k_t, $R = 0.4$	
Geometrical acceptance	$\lvert\eta\rvert < 4.5$	
Overlap removal (j, e)	$\Delta R(\mathrm{jet}, \mathrm{electron}) < 0.3$	
Calibration scheme	EM+JES	
Quality criteria	Looser	
Transverse momentum	$p_{\mathrm{T}} > 25$ GeV ($p_{\mathrm{T}} > 30$ GeV for $2.75 < \lvert\eta\rvert < 3.25$)	$p_{\mathrm{T}} > 25$ (30) GeV for $\lvert\eta\rvert \leq 2.5$ ($\lvert\eta\rvert > 2.5$)
Jet Vertex Fraction	$\lvert\mathrm{JVF}\rvert > 0.75$	$\lvert\mathrm{JVF}\rvert > 0.5$
b**-tagging**		
Tagging algorithm	JetFitterCombNN	MV1
Working point	80%	85%

the muon candidate. The small differences between simulation and data arising from the isolation and impact parameter requirements are accounted for with scale factors which are consistent with unity within 1-2 %.

Jets

Two selection criteria are optimised with respect to the 2011 analysis in order to cope with the higher level of pile-up in the 2012 dataset: the p_{T} threshold used for jet counting and the jet vertex fraction (JVF) requirement. The JVF criterion can only be used within the acceptance region of the tracking system ($|\eta| < 2.5$). The requirement was tightened to $|JVF| > 0.75$ in order to reduce contributions from pile-up jets sufficiently. Consequently, the jet p_{T} threshold was studied separately for central ($|\eta| < 2.5$) and forward ($|\eta| > 2.5$) jets and was set to 25 GeV and 30 GeV, respectively. The tighter requirement for forward jets helps to keep the contribution from jets arising from pile-up events under control.

For the 2011 analysis, the JetFitterCombNN b-tagging algorithm was used at an operation point leading to an efficiency of 80 % for the selection of b-jets (see Section 6.5.1). A neural-network based tagging algorithm called MV1 was developed in the course of 2011 which uses the results from the JetFitterCombNN algorithm in combination with additional impact parameter and secondary vertex information as inputs. This algorithm is used in 2012 as the baseline tagging algorithm for physics analysis. An operation point leading to 85 % efficiency for the selection of b-jets is chosen for the 2012 analysis.

6.12.2. Selection of $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ Candidate Events in the 2012 Analysis

The event selection requirements common to the H+0 jets, H+1 jet and H+ $\geq$ 2 jets channels stayed identical to the ones used for the 2011 analysis (see Table 6.5). Figure 6.20 shows the distribution of $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ in dilepton events satisfying all selection requirements up to but not including the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ requirement. After the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ requirement the Drell-Yan background is reduced further and mainly a concern for the H + 0 jets channel as can be concluded from the multiplicity distribution of jets shown in Fig. 6.20.

The requirements that are specific to the H+0 jets, H+1 jet and H+ $\geq$ 2 jets channels have been discussed in detail for the 2011 dataset in Sections 6.5.3, 6.5.4 and 6.5.5, respectively. They stayed unaltered for the 2012 analysis.

6.12.3. Background Normalisation and Control Samples for the 2012 Analysis

As already described in Section 6.5.6 for the 2011 analysis, control regions are used to determine the correct normalisation factors for the Monte Carlo predictions of $Z/\gamma^* +$jets, top and WW processes. The extracted normalisation factors and the corresponding uncertainties are described briefly in the following paragraphs.

Z/γ^*+jets Control Samples

A control region dominated by $Z/\gamma^* \to \tau\tau$ events is constructed by requiring 10 GeV $< m_{\ell\ell} <$ 80 GeV, $\Delta\phi_{\ell\ell} < 2.5$ and $p_T^{\ell\ell} < 30$ GeV. These are the same requirements that have been used to check the MC simulation for the $e\mu$ channel in 2011 (see Section 6.5.6). To extract the normalisation factor, the additional requirement of $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 25$ GeV is used that is in line with the cut applied to this channel in the signal selection. By comparing data to the Monte

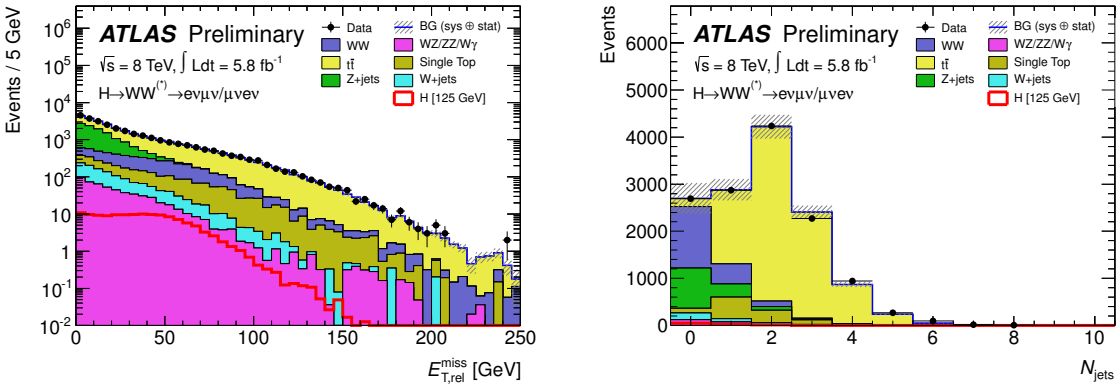


Figure 6.20.: Distributions of $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ (left) and the multiplicity of jets (right) for events satisfying the selection criteria up to the cut on $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ (omitted for the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ distribution). The $e\mu$ and μe channels are combined. The hatched area indicates the total uncertainty on the background prediction. The expected signal for a SM Higgs boson with m_H =125 GeV is superimposed.

Carlo prediction, the scale factor is found to be consistent with unity within the uncertainty of about ±10%.

W+jets Control Samples

The procedure to estimate the W + jets background contribution stayed the same as for the 2011 dataset (see Section 6.5.6). The uncertainty on the fake factor estimate is still the dominant uncertainty on the W + jets background contribution. The fake factors have been determined individually for muons and electrons. The values and the assigned uncertainties depend on the p_{T} of the reconstructed anti-identified lepton. The total relative uncertainty on the W + jets background is approximately ±40% for electron fakes and ±60% for muon fakes.

Top Control Samples

The method to determine the top background contribution to the 0-jet signal region was not changed with respect to the 2011 analysis (see Section 6.5.6). The ratio of the number of estimated top background events to the one obtained by Monte Carlo simulation is 1.11±0.06 (statistical uncertainty) for the combined $e\mu$ and μe channel.

The selection of the top control regions for the H+1 jet and H+ ≥ 2 jets channels is identical to the one used for the 2011 dataset. The resulting samples are dominated by contributions from top-quark pair and single top production processes which is evident from the transverse mass distributions obtained in the 1-jet and 2-jet top control regions and shown in Fig. 6.21. The resulting normalisation factors for the combined $e\mu$ and μe final states are 1.11±0.05 (statistical uncertainty) for the H + 1 jet analysis and 1.05±0.01 (statistical uncertainty) for the H+ ≥ 2 jets analysis, respectively.

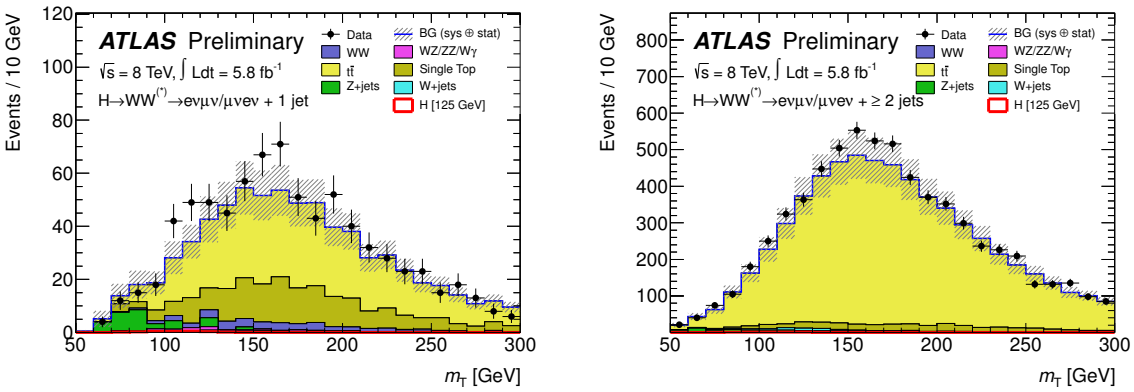


Figure 6.21.: Distributions of the transverse mass, m_{T}, in a top control region defined by requiring a b-tagged jet after the one jet requirement in the H + 1 jet analysis (left) and after the central jet veto selection in the H+ $\geq$ 2 jets analysis (right). The $e\mu$ and μe final states are combined. The hatched area indicates the total uncertainty on the background prediction. The final bin includes the overflow.

WW Control Samples

The definition of the WW control samples used to normalise the WW background Monte Carlo predictions in the H + 0 jets and H + 1 jet analyses stayed the same as for the 2011 dataset (see Section 6.5.6). The transverse mass distributions obtained in the WW control region of the 0-jet and 1-jet channel are shown in Fig. 6.22. The individual contributions from various SM background processes to the WW control sample are given in Table 6.17. The fraction of events originating from SM WW production is about 70 % in the zero jet control region and about 45 % in the one jet control region. The top backgrounds are scaled by the normalisation factors derived from the corresponding control regions as described in Section 6.12.3. The W + jets background is estimated via the data-driven method described in Section 6.12.3. The $e\mu$ and μe final states are combined and the WW normalisation factors are determined to be 1.06±0.06 (statistical uncertainty) for the H + 0 jets analysis and 0.99±0.15 (statistical uncertainty) for the H + 1 jet analysis, respectively.

Table 6.17.: The expected numbers of signal and background events in the WW control region in the H + 0 jets and H + 1 jet analyses. The signal is shown for m_{H} = 125 GeV. The W + jets and top backgrounds are estimated by the data-driven methods described in Section 6.12.3 and 6.12.3. Only statistical uncertainties associated with the number of events in the MC samples are shown.

Control Regions	Signal	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	Z/γ^* + jets	W + jets	Total Bkg.	Obs.
WW 0-jet	0.3±0.0	531±4	43±2	104±3	62±4	11±4	38±1	789±9	820
WW 1-jet	0.1±0.0	112±3	13±1	80±3	34±3	9±4	7.7±0.8	256±6	255

The decision to *unblind* the analysis was based on the satisfactory agreement between the Monte Carlo predictions and the data distributions in the control regions. Shortly afterwards, a WW Monte Carlo sample generated with `Powheg+Pythia8` became available. This sample was used to perform some cross-checks concerning the modelling of the WW background

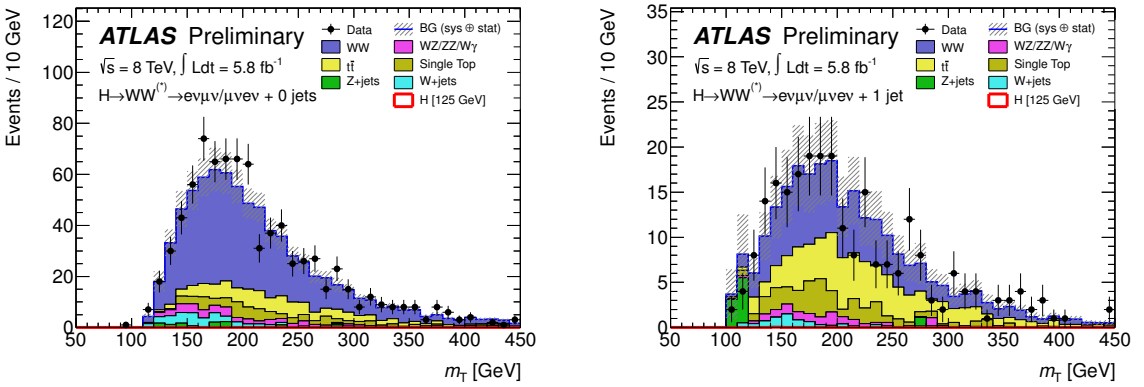


Figure 6.22.: Distributions of the transverse mass, m_{T}, in the WW control region in the H+0 jets analysis (left) and in the H+1 jet analysis (right). The top backgrounds are scaled by the normalisation factors derived from the corresponding control regions as described in Section 6.12.3. The $e\mu$ and μe final states are combined. The hatched area indicates the total uncertainty on the background prediction. The final bin includes the overflow.

component. The m_{T} distributions in the WW control region of the H + 0 jets analysis are shown with the default MC@NLO generator and the POWHEG generator used for modelling the WW background prediction on the top left-hand and right-hand side in Fig. 6.23, respectively. There are both differences in the total number of predicted events and some discrepancies in the shape of the m_{T} distribution observed. A Kolmogorov-Smirnov (KS) test [182] is performed to test the compatibility of the distributions predicted by Monte Carlo to the ones observed in data. It should be noted that the KS test is applied to binned distributions instead of unbinned data which the test is intended for. This will result in upward biases on the reported values. Furthermore, there are only statistical uncertainties included in the test and no bin-to-bin correlations are taken into account which leads to another upward bias on the values. Nevertheless, the KS probability values can be used to compare the compatibility of distributions obtained by two different generators to data. It becomes apparent that the KS probabilities are moderately higher for most of the distributions generated with POWHEG compared to those obtained with MC@NLO. A larger dataset is needed in order to reach a definite decision which generator should be used.

One of the selection criteria to define a control region with very little contribution from a potential Higgs boson signal is the requirement $m_{\ell\ell} > 80$ GeV. On the contrary the signal region selection comprises the requirement $12 < m_{\ell\ell} < 50$ GeV. An extrapolation from this control region to the signal region is performed using a transfer factor $\alpha = N^{\text{signal region}}/N^{\text{control region}}$. The value of this factor is shifted by 7% when POWHEG is used instead of MC@NLO as generator for the WW Monte Carlo simulation. This value is larger than the theoretical uncertainty of ±3.5% that has been assigned to the transfer factor based on generator studies comparing predictions from MC@NLO and MCFM [183]. Consequently, the uncertainty is increased from ±3.5% to ±7% to cover the discrepancy between POWHEG and MC@NLO.

The m_{T} shape distributions of the WW background component obtained by using MC@NLO or POWHEG as a generator was equally studied in the signal region. Figure 6.24 shows the m_{T} distributions in the signal region of the H + 0 jets analysis after the application of the $\Delta\phi_{\ell\ell}$ requirement. The distributions for the combined $e\mu$ and μe channels with the default MC@NLO generator and the POWHEG used for modelling the WW background prediction are shown in

Fig. 6.24, respectively. A direct comparison of the predicted shape of the m_{T} distributions modelled with MC@NLO and POWHEG is shown on the bottom left-hand side. The m_{T} distributions, normalised to the number of predicted events, are shown on the bottom right-hand side. A minor shift in the shape of the WW background between MC@NLO and POWHEG can be observed. The shape of the background is incorporated in the final binned likelihood fit performed to test the statistical significance of the presence or absence of signal events in data. In order to take the observed discrepancies into account, an extra systematic uncertainty has been added to the likelihood function that allows the WW m_{T} shape to continuously float between the boundaries specified by the MC@NLO and POWHEG generators. The default shape is modelled by the MC@NLO generator and the one obtained with POWHEG is considered as a one sigma deviation.

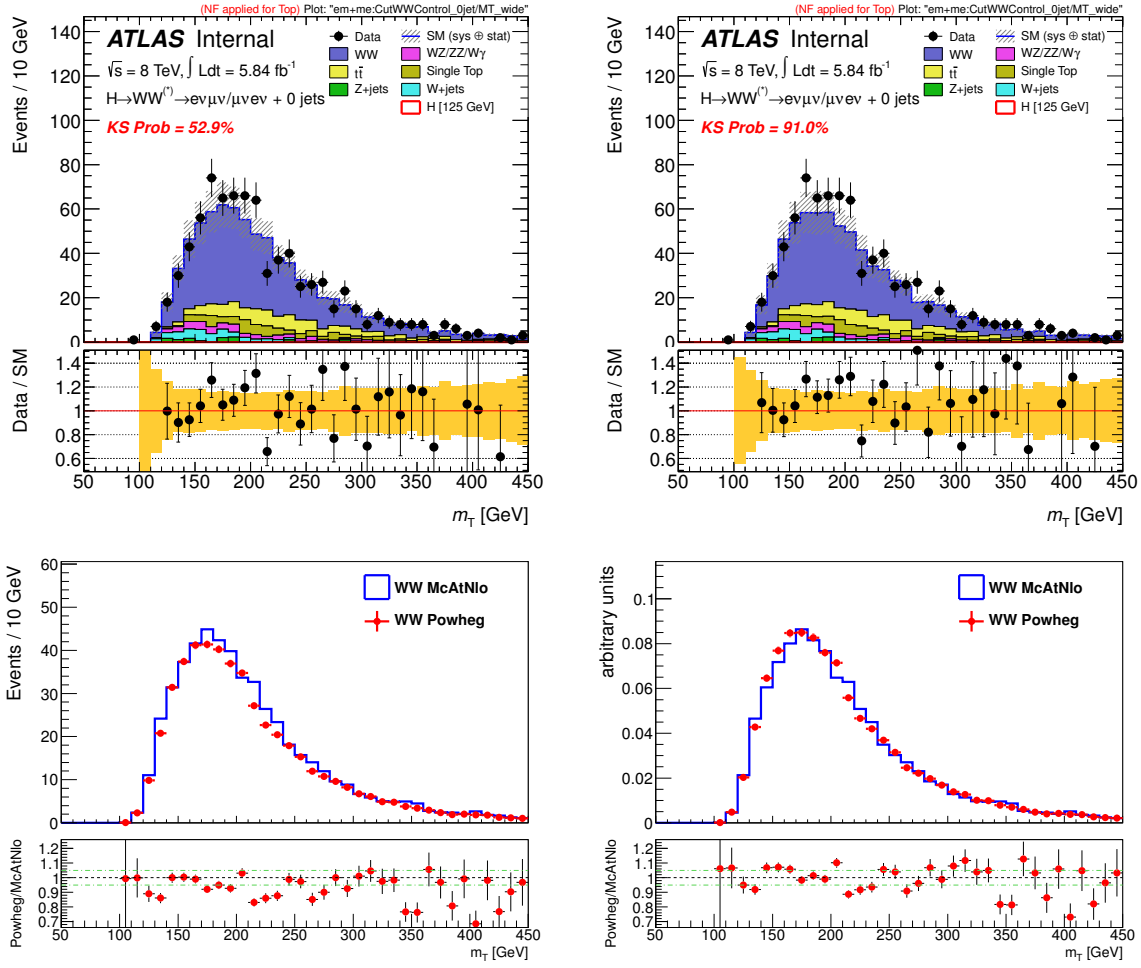


Figure 6.23.: Distributions of the transverse mass, m_{T}, in the WW control region in the H + 0 jets analysis using the default MC@NLO generator (top left) and the POWHEG generator (top right) for modelling the WW background. The $e\mu$ and μe final states are combined. The hatched area indicates the total uncertainty on the background prediction. A comparison of the predicted shape of the m_{T} distribution of the WW background modelled with MC@NLO and POWHEG is shown (bottom left). The m_{T} distributions are normalised to unity individually (bottom right) to check for differences in the shape regardless of the number of predicted events.

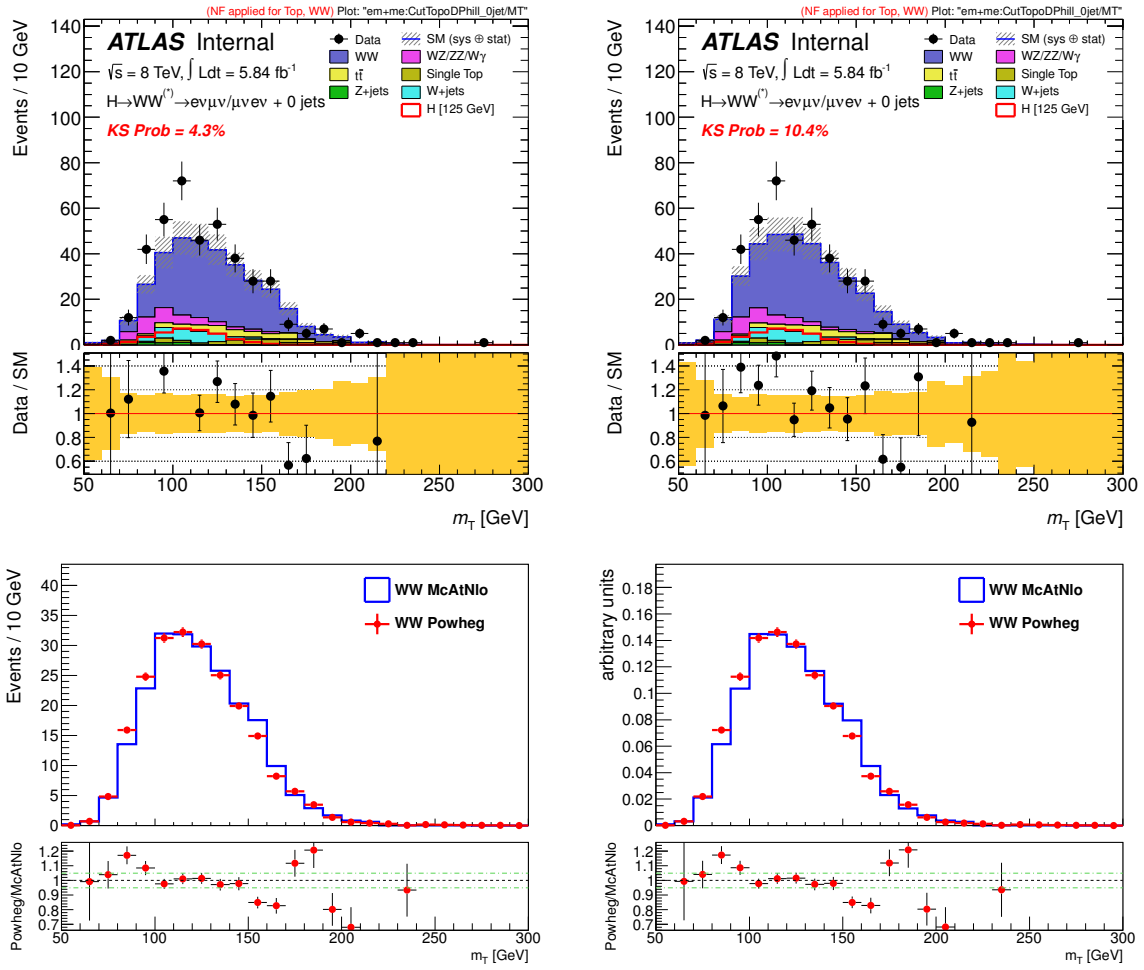


Figure 6.24.: Distributions of the transverse mass, m_{T}, in the H + 0 jets signal region after the $\Delta\phi_{\ell\ell}$ requirement (as defined in Table 6.5) using the default MC@NLO generator (top left) and the POWHEG generator (top right) for modelling the WW background. The $e\mu$ and μe final states are combined. The hatched area indicates the total uncertainty on the background prediction. A comparison of the predicted shape of the m_{T} distribution of the WW background modelled with MC@NLO and POWHEG is shown (bottom left). The m_{T} distributions are normalised to unity individually (bottom right) to check for differences in the shape regardless of the total number of predicted events.

6.12.4. WW m_T Shape Studies

The W^+W^- process is the largest background in this analysis. The present estimation in the control region helps to control the overall normalisation of this background in the signal region. The shape of the transverse mass distribution of the WW process in the signal region is taken from Monte Carlo simulation. An uncertainty on the shape is determined by comparing different Monte Carlo generators as described in the previous section. A precise knowledge of the shape of the WW transverse mass distribution in the signal region is crucial because it directly impacts the sensitivity to a possible signal. Thus, it is important to investigate the m_{T} shape modelling further by comparing the Monte Carlo prediction directly to data.

The signal-free control region enriched in dileptonic W^+W^- events is used to study how well Monte Carlo diboson generators are able to model the distributions observed in data. The checks are focusing on the m_{T} distribution in the WW control region in the H+0 jets analysis. The Monte Carlo background predictions for all components except the WW contribution are subtracted from data to allow a direct comparison of the WW background prediction to data. The m_{T} distributions of the WW background obtained with the `MC@NLO` and `POWHEG` generator compared to the data distribution (after the subtraction) are shown in Fig. 6.25. The respective ratios of the data distribution over the WW MC predictions are shown in Fig. 6.25 as well. The ratio is fitted using a linear function in order to check for a difference in the overall rate and a shift between the WW Monte Carlo distribution and data. A slightly negative slope parameter is observed which is basically driven by the low data to Monte Carlo ratio values with m_{T} values higher than 400 GeV. In this particular region, there are only 30 events in data observed before subtraction of the non-WW MC components. The slope parameter value is -0.0014 ± 0.0007 in case `MC@NLO` is used as a generator to model the WW background prediction and in case `POWHEG` is used the slope parameter value is slightly lower and amounts to -0.0012 ± 0.0008. As the statistical uncertainty on the extracted fit parameters is large, more data events are needed to conclude on the significance of the observed trend.

The study in the WW control region in the $\mathrm{H} + 0$ jets analysis suffers from the fact that there are contributions from other background sources like top, $W +$ jets and $Z/\gamma^* +$ jets production which amount to about 27% in the control region. Thus, it is not suitable for a more detailed inspection of important kinematic distributions.

6.12.5. Diboson Studies using a WZ Control Sample

A purer control region was searched for that allows to study the modelling of diboson events. A selection focusing on three leptons leads to a very pure sample of $W^{\pm}Z \to \ell\nu\ell\ell$ events with only little background contributions from $Z/\gamma^* +$ jets events. Thus, the WZ control sample allows for a direct comparison of the kinematic distributions predicted by diboson Monte Carlo generators and observed in data. Since the dileptonic W^+W^- and trileptonic $W^{\pm}Z$ decay modes feature kinematic similarities it is possible to transfer some of the conclusions gained on distributions in the WZ control region to W^+W^- distributions. To allow for this transfer, the selected $W^{\pm}Z \to \ell\nu\ell\ell$ events are modified to mimic a system of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ as those of $W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay as described in the following.

As the mass of the Z boson is relatively close to that of the W boson, the kinematics of WZ decays is similar to that of WW decays. The main idea of the WZ study is to

- reconstruct the trileptonic decays,
- identify the lepton pairs originating from the Z bosons

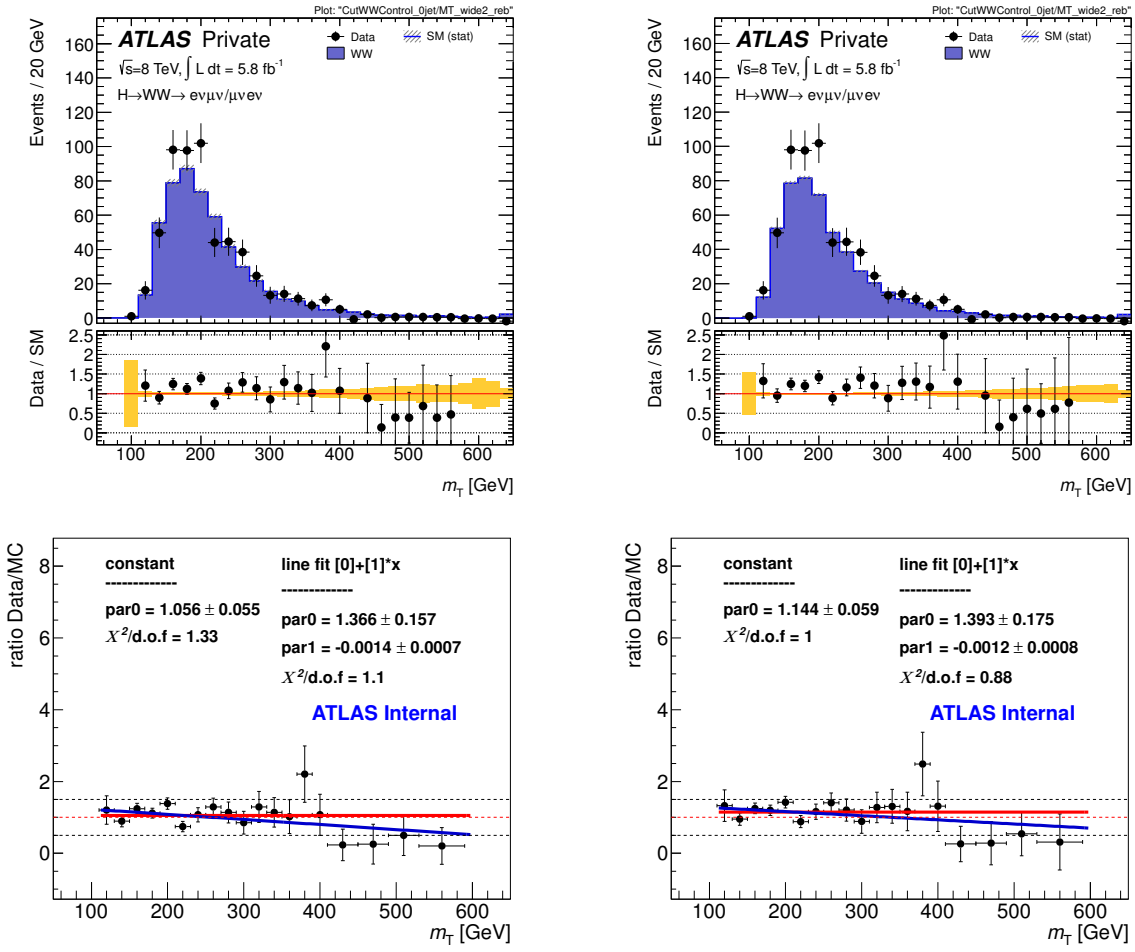


Figure 6.25.: Distributions of the transverse mass in the WW control region in the H+0 jets jets analysis using the default MC@NLO generator (top left) and the POWHEG generator (top right) for modelling the WW background. The $e\mu$ and μe final states are combined. The hatched area indicates the statistical uncertainty on the background prediction. The MC background predictions are subtracted from data with the exception of the WW background component. On the bottom row, fits to the ratio of the resulting data distribution over the WW Monte Carlo prediction are shown. Both a constant (red line) and a linear function (blue line) are used for the fit to check for a difference in the overall rate and the presence of a slope, respectively.

– assign one of the leptons to be a 'neutrino' as if the Z boson was a W boson,

– at a later stage, veto the presence of additional reconstructed jets to obtain events with kinematics as close as possible to the H + 0 jets analysis,

– compare the resulting mocked WW kinematic shapes to data to observe any mismodelling of the diboson production prediction.

This subsection presents the results obtained on a dataset corresponding to an integrated luminosity of 5.8 fb^{-1} collected at a centre-of-mass energy of 8 TeV by the ATLAS detector in 2012. At the LHC, the dominant $W^{\pm}Z$ production mechanism is from quark-antiquark initial states and to a lesser extend from gluon fusion. The leading order Feynman diagrams illustrating the production of $W^{\pm}Z$ pairs through $q\bar{q}$ initial state are shown in Fig. 6.26. The

analysis considers the experimental signature of the $W^{\pm}Z \to \ell\nu\ell\ell$ final state that can be split up in four channels: $e^{\pm}\nu e^{+}e^{-}$, $\mu^{\pm}\nu e^{+}e^{-}$,$e^{\pm}\nu\mu^{+}\mu^{-}$ and $\mu^{\pm}\nu\mu^{+}\mu^{-}$. The small contribution from leptonic τ decays resulting in electrons or muons in the final state are included as well.

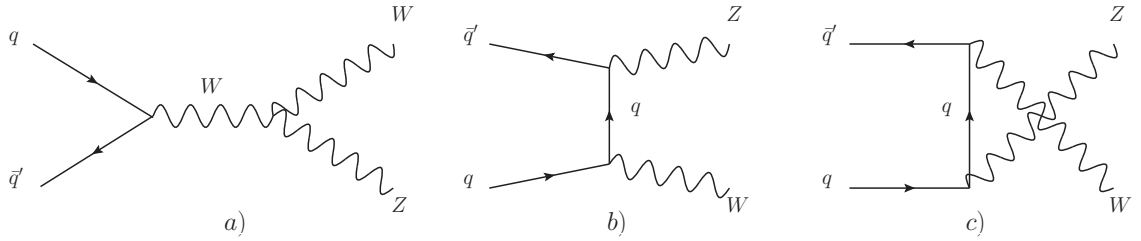


Figure 6.26.: Leading order Feynman diagrams illustrating the production of $W^{\pm}Z$ pairs at the LHC through $q\bar{q}$ initial state in the s-channel (a), in the t-channel (b) and in the u-channel (c).

Given that the study is aiming to test the agreement of data and distributions obtained using Monte Carlo diboson generators, events originating from ZZ decays are considered as signal events. The dominant background contributions arise from jets produced in association with Z bosons or $W^{\pm}$ bosons and top-quark production. The ZZ, $W+\text{jets}$, $Z/\gamma^{*}+\text{jets}$, single top and $t\bar{t}$ Monte Carlo events are modelled by the generators given in Table 6.15. The cross sections for $W(Z/\gamma^{(\star)})$ processes are normalised to the MCFM [158] NLO predictions (see Table 6.18). For $W(Z/\gamma^{(\star)})$ with $m_{Z/\gamma^{\star}} > 7$ GeV, the POWHEG prediction is used and events generated with $m_{Z/\gamma^{\star}} < 7$ GeV are removed from the sample.

Table 6.18.: Monte Carlo generators used to model the $W(Z/\gamma^{(\star)})$ processes. MADGRAPH includes the interference between $Z^{(\star)}$ and $\gamma^{\star}$ and the boundary between the samples is at $m_{\ell\ell} = 7$ GeV. Leptonic decays of W bosons/Z bosons are always assumed, and the quoted cross sections include the branching ratios and are summed over lepton flavours.

Process	Generator	$\sigma \cdot \text{Br}$ (pb)
$W(Z/\gamma^{(\star)})(m_{Z/\gamma^{\star}} > 7)$ GeV	POWHEG+PYTHIA8	12.7
$W(Z/\gamma^{(\star)})(m_{Z/\gamma^{\star}} < 7)$ GeV	MADGRAPH [153, 154]+PYTHIA [34]	14.3

The object selection for electrons, muons, jets and missing transverse energy is identical to the one used for the $H \to W^{+}W^{-} \to \ell^{+}\nu\ell^{-}\bar{\nu}$ analysis of the 2012 dataset as described in Section 6.12.1. The event selection follows the criteria described in Section 6.12.2 and listed in Table 6.5 up to the candidate selection. Thereafter, exactly three selected leptons are required that pass the isolation criteria as well. Two of the leptons must be consistent with a Z boson candidate by satisfying the criteria to have the same flavour and opposite charge. Figure 6.27 shows the $m_{\ell\ell}$ distribution of the Z boson candidate in events that pass the selection. Furthermore, the invariant mass of the two leptons must be close to the mass of the Z boson within a range of 10 GeV, $|m_{\ell\ell} - m_{Z}| < 10$ GeV. The events are required to satisfy $E_{\text{T}}^{\text{miss}} > 25$ GeV. This helps further reduce the contribution from $Z/\gamma^{*}+\text{jets}$ events as evident from the $E_{\text{T}}^{\text{miss}}$ distribution shown in Fig. 6.27. The third lepton is attributed to the decay of a $W^{\pm}$ boson and must have a transverse momentum larger than 20 GeV which suppresses the $Z/\gamma^{*}+\text{jets}$ contribution further. The transverse mass of the $W^{\pm}$ boson

candidate is calculated as

$$m_{T,W} = \sqrt{(2p_T^{\ell} E_T^{miss}(1 - cos(\Delta\phi))} \tag{6.10}$$

where p_T^{ℓ} is the transverse momentum of the third lepton and $\Delta\phi$ is the azimuthal angle between the third lepton and $\vec{p}_T^{\,miss}$. Figure 6.27 shows the distribution of the transverse mass of the $W^{\pm}$ boson candidate in events that pass the selection criteria up to the requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV. In order to reduce the contributions from processes other than the $W(Z/\gamma^{(\star)})$ and ZZ productions, the transverse mass must be greater than 20 GeV. The distributions of the jet multiplicity and the transverse momentum of the Z boson candidate are shown in Fig. 6.28 for events that pass all selection criteria.

Table 6.19.: The expected numbers of signal (WZ and ZZ) and background events as well as the events observed in data are listed at several stages of the selection. The lepton flavours are combined. Only statistical uncertainties associated with the number of events in the MC samples are shown. The event selection criteria are adopted from the analysis that was performed on the 2011 dataset to measure the $W^{\pm}Z$ production cross section [184].

	WZ	ZZ	WW	$W\gamma$	Top	$Z/\gamma^* + \mathrm{jets}$	Total MC	Obs.
lepton p_{T}	510.05 ± 2.87	128.76 ± 0.60	0.46 ± 0.10	0.1 ± 0.1	16.92 ± 1.28	241.10 ± 18.22	897.4 ± 18.5	1068
OS leptons	509.31 ± 2.87	128.58 ± 0.60	0.46 ± 0.10	0.1 ± 0.1	16.92 ± 1.28	240.91 ± 18.22	896.3 ± 18.5	1067
$\lvert m_{\ell\ell} - m_Z\rvert < 10$ GeV	440.50 ± 2.68	87.87 ± 0.49	0.13 ± 0.05	0.0 ± 0.0	5.32 ± 0.64	119.77 ± 13.17	653.6 ± 13.5	756
$E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV	359.78 ± 2.42	40.99 ± 0.33	0.10 ± 0.04	0.0 ± 0.0	4.79 ± 0.60	46.41 ± 8.16	452.1 ± 8.5	488
$m_{\mathrm{T,Wcand}} > 20$ GeV	306.87 ± 2.23	27.68 ± 0.27	0.04 ± 0.02	0.0 ± 0.0	3.36 ± 0.50	19.49 ± 4.96	357.4 ± 5.5	385
fraction in %	85.9	7.7	< 0.01	negl.	0.9	5.5		
jet veto	168.03 ± 1.65	11.88 ± 0.17	0.01 ± 0.01	0.0 ± 0.0	0.12 ± 0.09	13.14 ± 4.28	193.2 ± 4.6	204
fraction in %	87.0	6.1	< 0.01	negl.	< 0.1	6.8		

The expected numbers of events originating from WZ and ZZ decays and the background contributions from W + jets, Z/γ^* + jets and top production processes are listed at several stages of the selection in Table 6.19. The numbers observed in data are higher compared to the total number of events predicted by Monte Carlo simulation. The overal agreement improves significantly after the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is applied. This indicates that an underestimation of the Z/γ^* + jets background component might be the source for the discrepancies at the beginning of the selection. The purity of the WZ and ZZ events is about 87% after all selection requirements. The only remaining background contribution consists in events from Z/γ^* + jets production processes which account for about 6% of the total number of events. This proportions are still observed after requiring no additional, reconstructed jets in the events. There are 385 events observed in data before the zero jet requirement and 204 events are observed in data afterwards. This is a quarter of the events observed in the WW control region for the H + 0 jets analysis performed on the same dataset. However, the sample obtained with the $W^{\pm}Z$ selection has the advantage that it contains much less contribution from non-diboson processes.

The transverse mass of the selected three leptons and the $E_{\mathrm{T}}^{\mathrm{miss}}$ system is defined as:

$$m_{T,WZ} = \sqrt{(E_T^{\ell\ell\ell} + E_T^{miss})^2 - (p_x^{\ell_1} + p_x^{\ell_2} + p_x^{\ell_3} + E_x^{miss})^2 - (p_y^{\ell_1} + p_y^{\ell_2} + p_y^{\ell_3} + E_y^{miss})^2} \tag{6.11}$$

where $E_T^{\ell\ell\ell}$ is the transverse energy and p_x^{ℓ}, p_y^{ℓ} are the transverse momentum components of the three leptons. In order to mimic the two-lepton-and-$E_{\mathrm{T}}^{\mathrm{miss}}$ system as in the $W^+W^- \to$

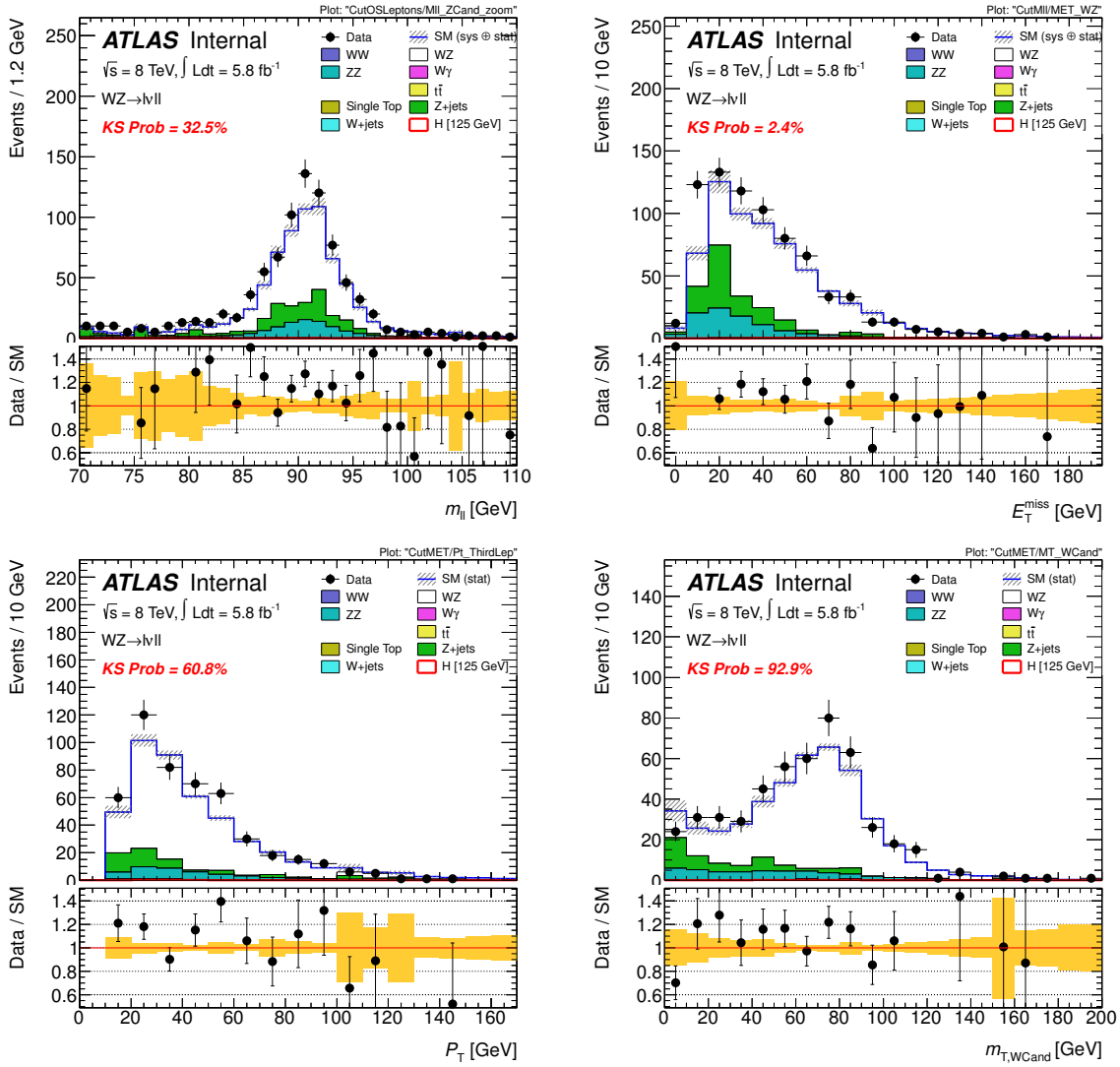


Figure 6.27.: Distributions of the dilepton invariant mass $m_{\ell\ell}$ of the Z boson candidate for events passing the trilepton selection criteria (top left) and missing transverse momentum of the trilepton events that pass the requirement on the Z boson invariant mass $|m_{\ell\ell} - m_Z| < 10$ GeV (top right). Distributions of the transverse momentum of the lepton attributed to the W boson candidate (bottom left) and transverse mass of the W boson candidates (bottom right) in the events that satisfy the selection up to the requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV.

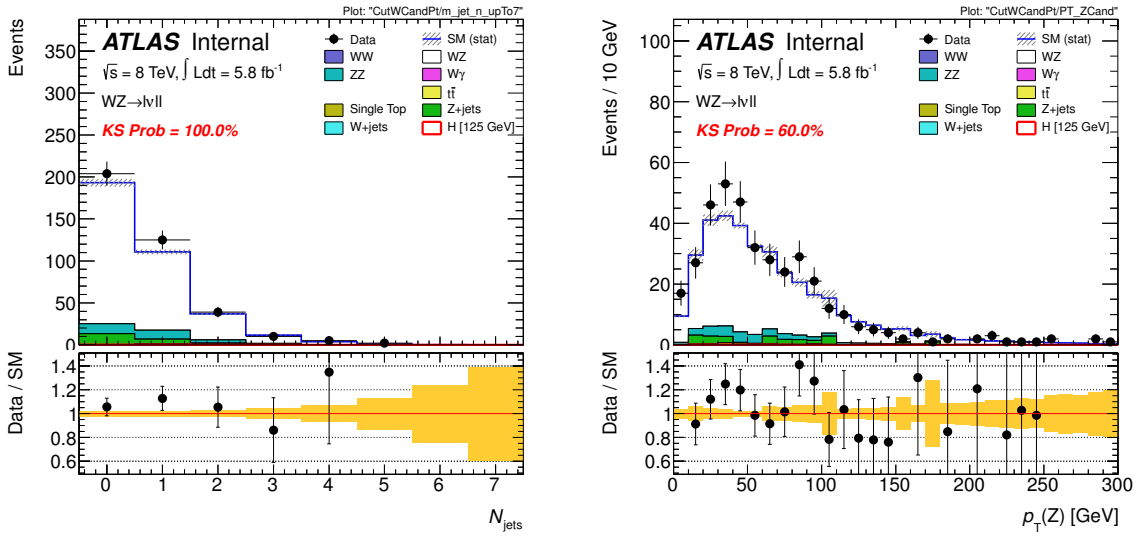


Figure 6.28.: Distributions of the jet multiplicity (left) and the transverse momentum of the Z boson candidate (right) for events that pass all selection criteria.

$\ell^+\nu\ell^-\bar{\nu}$ decays, one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$. It is ensured that the remaining two leptons have opposite charge to be as close as possible to the signature of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ final state. The transverse mass of the modified system is calculated in the same way as for the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis (see Table 6.4). The distributions of the transverse mass $m_{T,WZ}$ and the transverse mass of the modified system, denoted as $m_{T,l\nu l\nu}$, are shown in Fig. 6.29 for events that pass all selection criteria and for events that pass in addition the zero-jet requirement. Even if the observed number of events is limited, the distributions serve as test bench to examine the prediction of the Monte Carlo diboson kinematics, in particular in terms of the transverse mass shapes.

In the following, all the Monte Carlo background predictions but WZ and ZZ are subtracted from data. The resulting distributions are shown in Fig. 6.30 and Fig. 6.31 for events that pass all selection criteria and for events that pass in addition the zero jet requirement, respectively. The ratios of the data distributions over the Monte Carlo predictions are shown in Fig. 6.30 and Fig. 6.31 as well. As in the case of the WW control region, the ratios are successively fitted by a constant and a linear function in order to check for a difference in the overall rate and shifts between the Monte Carlo and data distributions. The bins in the high mass tail of the transverse mass distribution are merged to avoid empty bins that cannot be handled by the χ^2 fit that is performed. The fit range is set by the highest transverse mass bin that has a data contribution larger than zero after the background subtraction is performed. Slightly negative slope parameters are observed for both the transverse mass of the three leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ system and the modified transverse mass of events passing all selection criteria and the additional zero jet requirement. The obtained slope parameter values are summarised in Table 6.20. The trend of a shift between the data and Monte Carlo distributions is more evident in the case of the $m_{T,WZ}$ distribution compared to the modified transverse mass distribution. It is noticeable that the slope parameters are comparable with zero within at most 1.8 standard deviations. The same trend was also observed for the m_T shape study performed in the WW control region as discussed in Section 6.12.4.

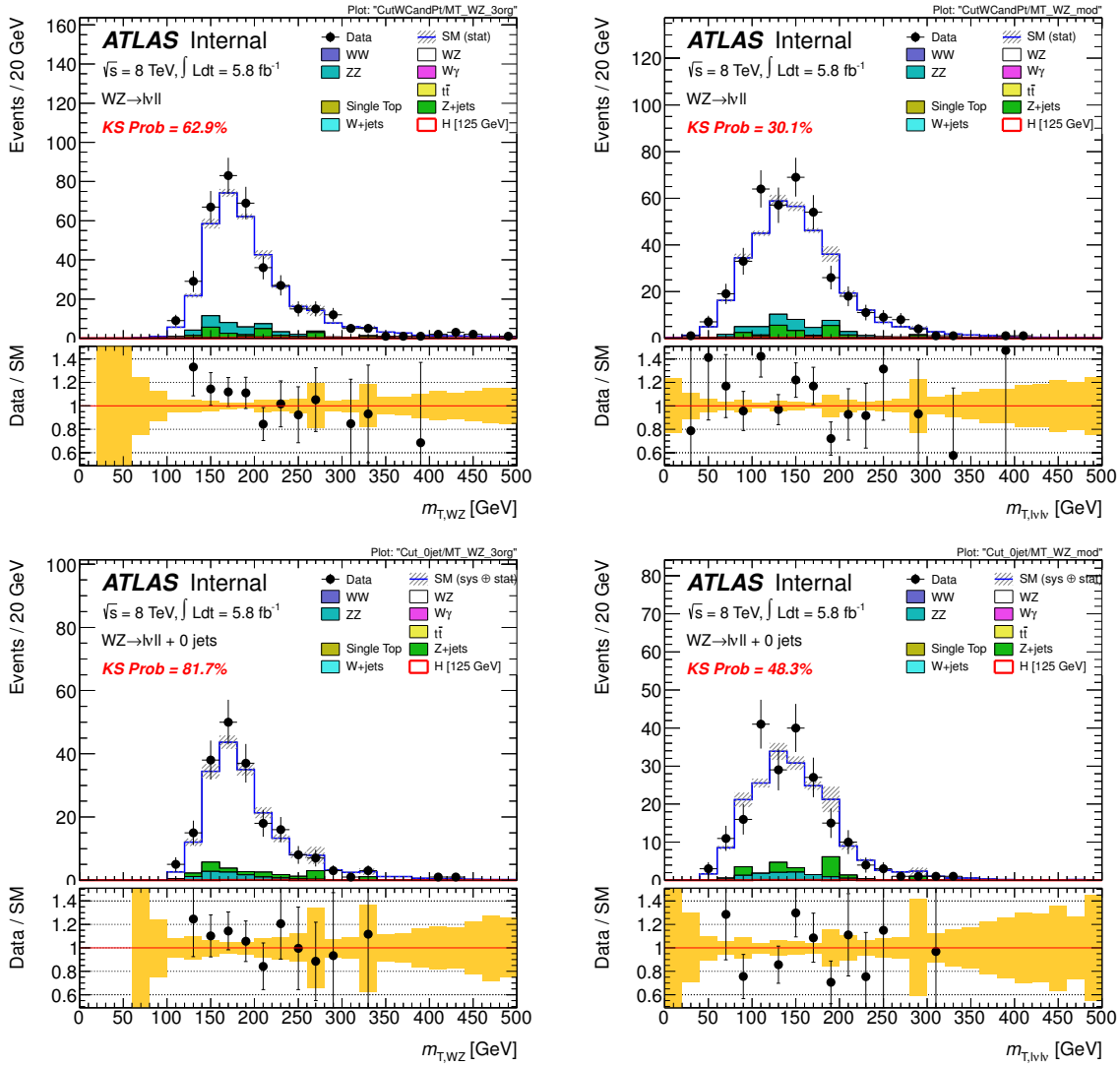


Figure 6.29.: Distributions of the transverse mass of the $W^{\pm}Z$ pair (top left) and the transverse momentum of the modified $W^{\pm}Z$ system (top right), where one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to mimic a system of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$, for events that pass all selection criteria. The distributions of the transverse mass of the $W^{\pm}Z$ pair and the modified $W^{\pm}Z$ system for events that pass all selection criteria and have zero reconstructed jets are shown on the bottom left-hand side and right-hand side, respectively.

Table 6.20.: Summary of slope parameters extracted by linear fits to the ratio of the transverse mass distribution obtained in data over the combined WZ and ZZ Monte Carlo prediction (shown in Fig. 6.30 and Fig. 6.31).

	$m_{T,WZ}$ distribution	$m_{T,l\nu l\nu}$ distribution
slope parameter for events passing all selection criteria	-0.0017 ± 0.0010	-0.0016 ± 0.0010
slope parameter for events passing the jet veto	-0.0025 ± 0.0014	-0.0020 ± 0.0015

The results obtained on a dataset corresponding to an integrated luminosity of 5.8 fb^{-1} collected at a centre-of-mass energy of 8 TeV in 2012 show the first signs of a potential mismodelling of the Monte Carlo diboson prediction compared to data. However, the statistical uncertainty on the extracted slope parameters is quite large and a larger dataset needs to be investigated to confirm the trend with certainty. An extended study of diboson events was performed on a dataset corresponding to an integrated luminosity of 10.7 fb^{-1} and is presented in Section 7.

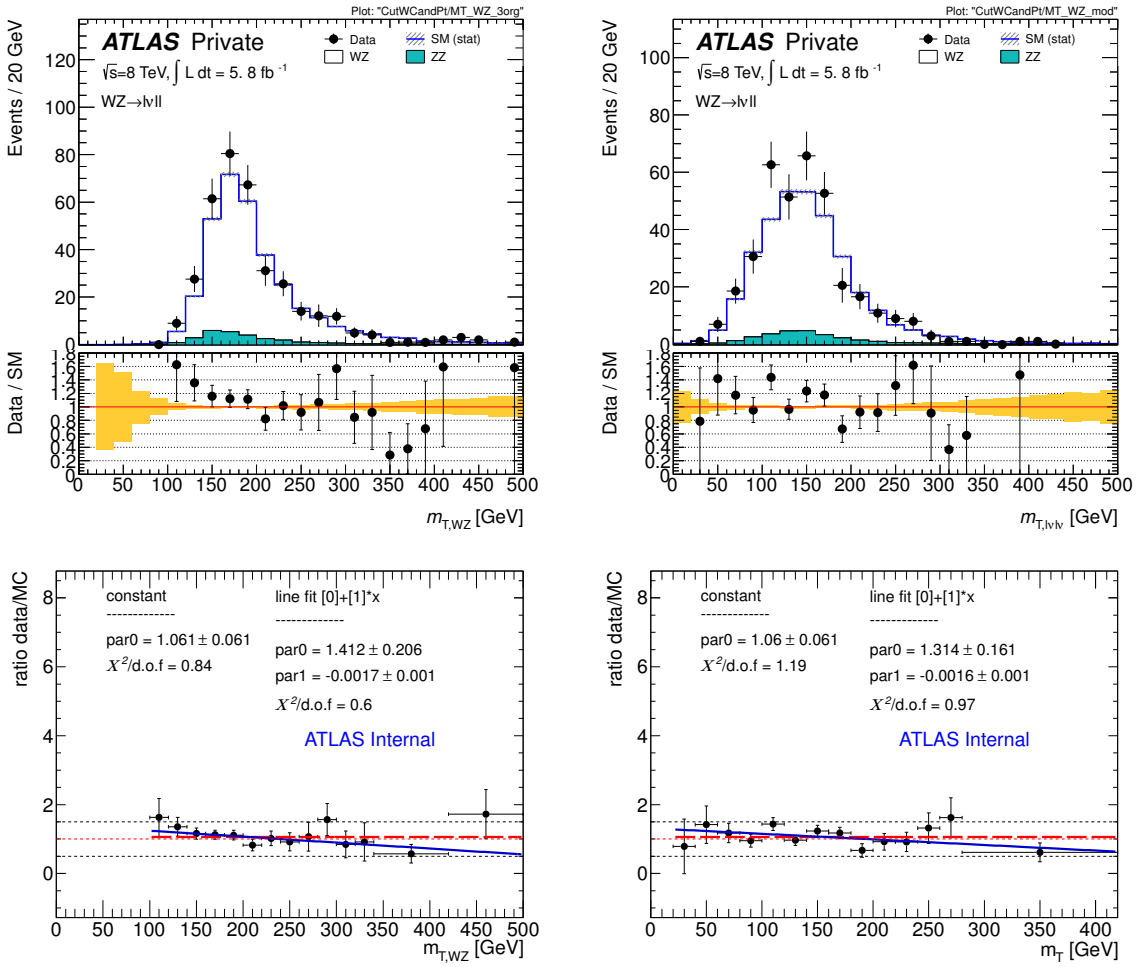


Figure 6.30.: Distributions of the transverse mass of the $W^{\pm}Z$ pair (top left) and the transverse momentum of the modified $W^{\pm}Z$ system (top right), where one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to mimic a system of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$, for events that pass all selection criteria. The MC background predictions are subtracted from data with the exception of the WZ and ZZ Monte Carlo component. The ratio of the resulting data distribution over the WZ and ZZ Monte Carlo prediction is fitted by a constant (red line) and a linear function (blue line) to check for evidence of a slope (bottom left and right).

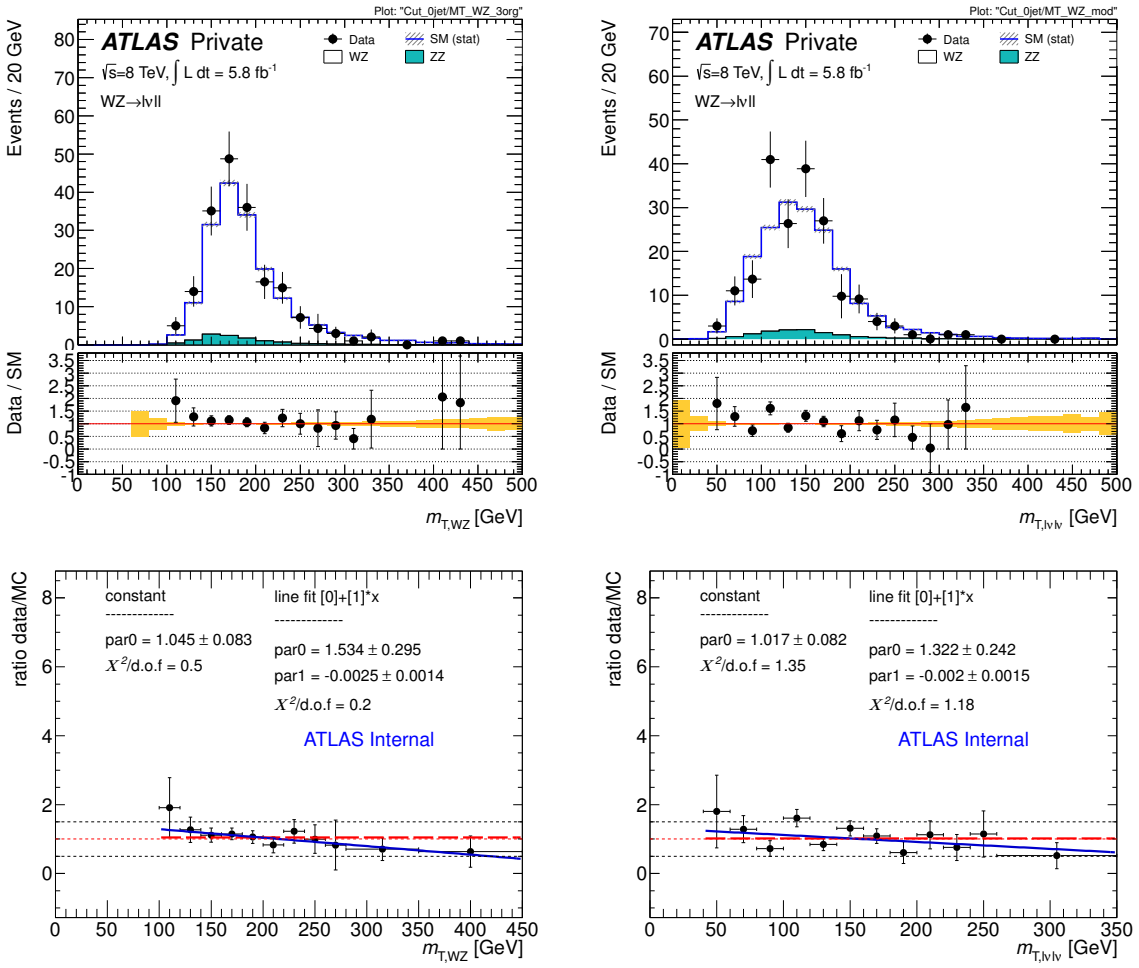


Figure 6.31.: The distributions of the transverse mass of the $W^{\pm}Z$ pair and the modified $W^{\pm}Z$ system for events that pass all selection criteria and have zero reconstructed jets are shown on the top left-hand side and right-hand side, respectively. The MC background predictions are subtracted from data with the exception of the WZ and ZZ Monte Carlo component. The ratio of the resulting data distribution over the combined WZ and ZZ Monte Carlo prediction is fitted by a constant (red line) and a linear function (blue line) to check for evidence of a slope (bottom left and right).

6.13. Systematic Uncertainties for the 2012 Analysis

As already discussed in the context of the 2011 analysis in Section 6.6, systematic uncertainties arise from both experimental and theoretical sources. They are taken into account by considering their effect on the predicted numbers of signal and background events and on the shape of the m_{T} distribution. The dominant contributions to the systematic uncertainties are described in the following.

6.13.1. Theoretical Uncertainties

The systematic uncertainties on the signal production cross sections are determined following the same procedure as described for the 2011 analysis in Section 6.6. The uncertainties obtained by independent variations of the QCD factorisation and renormalisation scales are summed up in quadrature. Thus, the resulting uncertainties on the signal production cross section for $\mathrm{m_H} = 125$ GeV amounts to $\pm 17\%$ for the $\mathrm{H} + 0$ jets analysis and to $\pm 36\%$ for the $\mathrm{H} + 1$ jet analysis [165–167]. The H+ ≥ 2 jets analysis is mainly sensitive to the impact of scale variations on both the VBF signal cross section and the jet veto acceptance. The size of the effect is estimated to be $\pm 4\%$. An additional uncertainty of $\pm 7\%$ is included to account for the effect of the underlying event modelling on the signal acceptance for signal events produced by VBF and passing the jet tagging and central jet veto requirements. In the H+ ≥ 2 jets channel, around 25 % of the signal events are produced via gluon fusion and the uncertainty on the ggF signal cross section is about $\pm 25\%$.

As already has been mentioned for the 2011 analysis in Section 6.6, PDF uncertainties are estimated, following Refs. [32, 157, 170, 171], using the envelopes of error sets as well as different PDF sets, applied separately to quark-quark, quark-gluon and gluon-gluon initiated processes. The relative PDF uncertainty on the dominant ggF process is $\pm 8\%$ and amounts to $\pm 2\%$ for the VBF process for $\mathrm{m_H} = 125$ GeV.

Uncertainties on the modelling of signal and background processes are estimated by using alternative generators, such as `MC@NLO` for the ggF process and `POWHEG` for WW production. As described in Section 6.12.3, the WW Monte Carlo background prediction is normalised to data in signal-free control regions. Depending on whether `MC@NLO+HERWIG` or `POWHEG+PYTHIA8` is used as generator for the WW Monte Carlo background, there are differences in both the total number of predicted events and discrepancies in the shape of the m_{T} distributions observed (see Section 6.12.3). The theoretical uncertainty on the extrapolation from the control regions to the signal region was evaluated in a way to cover the observed discrepancies and amounts to $\pm 7\%$. Furthermore, an additional uncertainty on the m_{T} shape modelling is added. The distribution obtained by the `MC@NLO` generator is taken as default shape and the one obtained by `POWHEG` is considered as a one σ deviation. The uncertainties associated with the underlying event and parton showering are taken into account in the acceptance uncertainty, but they are negligible compared to the scale uncertainties on the cross sections in exclusive jet bins.

Uncertainties on the normalisation of the $W\gamma$ Monte Carlo prediction are evaluated for each jet bin [166]. The relative uncertainty on the predicted number of background events arising from $W\gamma$ processes amounts to $\pm 11\%$ and $\pm 50\%$ for the $\mathrm{H} + 0$ jets and $\mathrm{H} + 1$ jet analysis, respectively. `MadGraph` is used to generated $W\gamma^{\star}$ events and the leading order prediction obtained therefrom is compared to the number of events predicted by the NLO calculation of `MCFM`. Corresponding correction factors and their uncertainties for $W\gamma^{\star}$ events with $m_{\ell\ell} <$ 7 GeV and $m_{\ell\ell} > 7$ GeV are determined to be 1.3 ± 0.3 and 1.51 ± 0.45, respectively.

Table 6.21.: Experimental sources of systematic uncertainties per object or event.

Source of Uncertainty	Treatment in the analysis
Jet Energy Resolution (JER)	MC jet resolution smeared using jet p_T, η-dependent parametrisation, resolution varies from ±22 % to ±7 % **relative uncertainty on JER varies from ±25% to ±17 %**
Jet Energy Scale (JES)	**global JES: ±2 to 9%** depending on jet p_T, η and sample composition for jets with $p_T > 25$ GeV and $\|\eta\| < 4.5$ pile-up: $\pm 0.5 - 10\%$ depending on jet p_T, η and sample composition Gluon fraction uncertainty and uncertainty from close-by jets included
Electron Selection Efficiency	Separate systematics for electron identification, reconstruction and isolation, added in quadrature Identification: ±3% for $p_T < 15$ GeV, decreasing to ±1% for $p_T > 30$ GeV in the central region Reconstruction: ±0.6 - 1.1% for $p_T > 15$ GeV trigger: ±1-2% uncertainty **Total uncertainty of ±2-5%** depending on η and E_T
Electron Energy Scale	Uncertainty **smaller than ±1%**, depending on η and E_T
Electron Energy Resolution	Energy varied within its uncertainty, **±0.6% of the energy** at most
Muon Selection Efficiency	**±0.3-1%** as a function of η and p_T reconstruction smaller than ±1%
Muon Momentum Scale and Resolution	Uncertainty **smaller than ±1%**
b-tagging Efficiency	p_T dependent scale factor uncertainties, **±4.7 - 18%**
Missing Transverse Energy	Jet energy scale and lepton momentum scale uncertainties are propagated to $E_{\mathrm{T}}^{\mathrm{miss}}$ computation additional contributions from jets with $p_T < 20$ GeV and low-energy depositions not associated with reconstructed physics objects **Total uncertainty of ±1-7%**
Event pile-up	modelling of event pile-up contributions estimated by varying their effect on low-energy calorimeter depositions **Total uncertainty of ±0.5-10%**
Luminosity	**±3.6%** [142]

6.13.2. Experimental Uncertainties

As already detailed for the 2011 analysis in Section 6.6, the main experimental uncertainties are related to the jet energy scale, which is determined from a combination of test beam, simulation and in situ measurements. The uncertainty on the jet energy scale varies from ±9% to ±2% as a function of jet p_{T} and η for jets with $p_T > 25$ GeV and $|\eta| < 4.5$ ([119]). An additional contribution to the jet energy scale uncertainty arises from pile-up and is estimated to vary between ±5% and ±1% for in-time pile-up and up to ±10% for out-of-time pile-up. The second largest uncertainty is related to the jet energy resolution and its magnitude is given in Table 6.21. The reconstruction, identification and trigger efficiencies for electrons and muons, as well as their momentum scales and resolutions, are estimated using

$Z \to \ell^+\ell^-$, $J/\psi \to \ell^+\ell^-$ and $W \to \ell\nu$ decays (see Sections 4.2.4 and 4.2.3). All of the resulting uncertainties are smaller than ±1% with the exception of the uncertainty on the electron efficiency, which varies from ±2% to ±5% as a function of lepton p_{T} and η. The uncertainties related to the jet energy and lepton momentum scale are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ computation. Additional contributions arising from jets with $p_{\mathrm{T}} < 20$ GeV and from low-energy calorimeter depositions not associated with reconstructed physics objects are included as well [126]. Their effect on the total signal and background yields is about ±3%. The uncertainty on the b-tagging efficiency is determined using samples containing muons reconstructed in the vicinity of jets and varies between ±5% and ±18% depending on the p_{T} of the jet [124]. The preliminary uncertainty on the integrated luminosity is ±3.6% based on the measurements described in Refs. [141, 142]

As already stated in Section 6.6, the theoretical and experimental uncertainties due to various sources are obtained by independent variations. The uncertainties are considered to be uncorrelated and they are summed up in order to illustrate the effect of the uncertainties on Monte Carlo distributions. However, correlations between uncertainties on different backgrounds are taken into account in the statistical analysis of the dataset.

6.13.3. Dominant Systematic Uncertainties on Different Background Processes

The total uncertainty on the m_{T} distribution predicted by Monte Carlo simulation is dominated by the uncertainty on the normalisation of each background component. The dominant sources of systematic uncertainties affecting the normalisation of the W + jets, top and WW background processes in the signal region are discussed in the following. It should be noted that the presented uncertainties are informative but the correlations between nuisance parameters and background processes are better taken into account concerning the final uncertainties by the likelihood fit procedure already described in Section 6.8.

W+jets background:

The uncertainty on the normalisation of the W + jets background component in the signal region is dominated by the statistical uncertainty in the W + jets control region which amounts to ±3% for the H + 0 jets analysis and to ±17% in the H + 1 jet analysis. An additional uncertainty arises from the uncertainty on the fake factor estimate and is on average of the order of ±50%.

Top and WW background:

The Monte Carlo prediction of the top background component is corrected by the ratio of the number of observed data events to events from Monte Carlo prediction in the control regions. This method relies on the correlations between event yields in the signal and control regions which can potentially be different between Monte Carlo simulation and data. The size of the uncertainty on the transfer of information from the control region to the signal region has to be determined. It is obtained by comparing the nominal ratio value to the ratios of signal to control region yields obtained by varying different theoretical and experimental sources. The observed differences are added in quadrature.

One has to keep in mind that the statistical uncertainties on the number of observed events in the control regions used to derive the normalisation factors and on the number of events predicted by Monte Carlo simulation in the signal region have to be added. These uncertainties are of the order of ±2% for the H + 0 jets and ±7% for the H + 1 jet channel.

The total uncertainty on the predicted WW background component in the signal region is ±13% for the H + 0 jets channel which is dominated by the statistical uncertainty on the

normalisation to the WW control region (about $\pm 6\%$) followed by the uncertainty on the extrapolation from the WW control region to the signal region ($\pm 8\%$). In the H+1 jet channel, the total uncertainty on the WW background amounts to $\pm 42\%$ and is dominated by the uncertainty on the normalisation factor which in turn is dominated by the large uncertainty on the top background extrapolated to the WW control region. The total uncertainties on the estimated top-quark background in the 1-jet and 2-jet signal regions amount to $\pm 36\%$ and $\pm 70\%$, respectively. These sizeable uncertainties are caused by the extrapolation from the control region characterised by a b-tag requirement to the signal region featuring a veto on b-tagged jets. The non-negligible uncertainty related to the b-jet tagging efficiency in combination with a very efficient tagging working point are the dominant sources. Concerning the $\mathrm{H}+0$ jets analysis, the total uncertainty on the top background normalisation in the signal region consists of two pieces. On the one hand the uncertainty due to the data-driven estimate as described in Section 6.5.6 which amounts to $\pm 17\%$. On the other hand, there is the statistical uncertainty on the shape predicted by Monte Carlo simulation in the signal region which is at the level of $\pm 10\%$.

It should be pointed out once again that the final uncertainty is assessed more correctly by the likelihood fit procedure used to perform the statistical analysis since correlations are taken into account.

6.14. Analysis Results of the 2012 Dataset

This section contains the results obtained in the search for the Standard Model Higgs boson in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ dilepton final state using a dataset corresponding to an integrated luminosity of 5.8 fb^{-1} of proton-proton collision data recorded by the ATLAS experiment in 2012 at a centre-of-mass energy of 8 TeV. The expected numbers of signal and background events as well as the numbers of events observed in data are shown in Table 6.22 at several stages of the selection. The uncertainties on the Monte Carlo background prediction are statistical only. After all selection requirements, the dominant background in the $\mathrm{H}+0$ jets channel originates from WW production, followed by smaller contributions from top, non-WW diboson and $W+$jets processes. The dominant background contributions to the $\mathrm{H}+1$ jet and H+ ≥ 2 jets channels arise from WW and top-quark production processes in equal parts. The same behaviour was already observed in the analysis of the 2011 dataset.

The transverse mass distributions of events that pass all selection criteria in the $\mathrm{H}+0$ jets and $\mathrm{H}+1$ jet analyses are shown in Fig. 6.32 for both lepton flavour channels combined. No m_{T} distribution is shown for the H+ ≥ 2 jets channel as there are only two events left over in data after the $\Delta\phi_{\ell\ell}$ requirement is applied. The sum of the transverse mass distributions of the $\mathrm{H}+0$ jets and $\mathrm{H}+1$ jet channels with the expected signal for $\mathrm{m_H}$ =125 GeV stacked on top of the background prediction is shown in Fig. 6.32, too. An excess of events relative to the predicted background is observed in data. In order to better visualise the size of the effect, the total estimated background is subtracted from data and overlaid with the predicted signal for a Standard Model Higgs boson with $\mathrm{m_H}$ =125 GeV. No systematic uncertainties are included but the statistical uncertainties of both the data and the subtracted background are reflected in the uncertainties assigned to the data points. The observed excess of events in data seems to be compatible with the expectation from a Standard Model Higgs boson with $\mathrm{m_H}$ =125 GeV.

Table 6.22.: The expected numbers of signal (for $m_H = 125$ GeV) and background events as well as the events observed in the 2012 dataset at several stages of the selection. The $W+$jets background is estimated from data, whereas MC predictions normalised to data in control regions are used for WW, Z/γ^*+jets, $t\bar{t}$, and $tW/tb/tqb$ processes. The column "Top" refers to the sum of the contribution from single top and top quark pair production. Contributions from other background sources are taken directly from MC predictions. Only statistical uncertainties associated with the number of events in the MC samples and the data control regions are shown. Starting from jet multiplicity specific selection stages, the numbers for the SM WW and top background contributions are corrected using the normalisation factor derived from the corresponding control samples.

	Signal	WW	$WZ/ZZ/W\gamma$	Top	Z/γ^*+jets	$W+$jets	Total Bkg.	Observed
Lepton p_{T}	138 ± 0.74	2820 ± 12.7	1141 ± 10.8	15821 ± 39.4	8557 ± 107	1255 ± 13.8	29594 ± 116	30510
Opposite Charge	136 ± 0.73	2813 ± 12.7	566 ± 7.58	15773 ± 39.1	8396 ± 106	799 ± 12.3	28348 ± 114	29368
$m_{\ell\ell} > 10$ GeV	135 ± 0.73	2810 ± 12.7	559 ± 7.51	15761 ± 39.1	8393 ± 106	797 ± 12.3	28320 ± 114	29339
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 25$ GeV	86.9 ± 0.59	1879 ± 10.7	230 ± 5.02	9752 ± 31.8	1237 ± 39.4	277 ± 6.27	13376 ± 52.4	13410
0j: Jet veto	47.5 ± 0.44	1308 ± 8.98	125 ± 3.94	293 ± 6.85	850 ± 32.0	138 ± 4.00	2714 ± 34.4	2691
0j: $p_T^{\ell\ell} > 30$ GeV	43.4 ± 0.42	1077 ± 8.14	99.2 ± 3.53	264 ± 6.52	46.9 ± 7.69	102 ± 2.29	1589 ± 13.6	1664
0j: $m_{\ell\ell} < 50$ GeV	34.9 ± 0.38	244 ± 3.96	32.6 ± 2.05	45.1 ± 2.84	5.3 ± 2.27	29.0 ± 1.13	356 ± 5.86	421
0j: $\Delta\phi_{\ell\ell} < 1.8$ rad	33.6 ± 0.37	234 ± 3.87	31.6 ± 2.02	44.3 ± 2.82	4.0 ± 1.87	25.0 ± 1.05	339 ± 5.62	407
0j: $0.75 \cdot \mathrm{m_H} < \mathrm{m_T} < \mathrm{m_H}$	20.2 ± 0.29	101 ± 2.56	11.5 ± 1.19	11.4 ± 1.18	1.9 ± 1.31	15.4 ± 0.80	142 ± 3.42	185
1j: $N_{\mathrm{jets}} = 1$	24.9 ± 0.31	396 ± 4.64	74.2 ± 2.70	2131 ± 16.8	283 ± 19.8	68.4 ± 3.05	2953 ± 26.7	2874
1j: b-jet veto	21.1 ± 0.28	334 ± 4.23	56.2 ± 2.34	464 ± 8.31	236 ± 18.1	53.0 ± 2.49	1144 ± 20.6	1115
1j: $p_T^{tot} < 30$ GeV	12.2 ± 0.22	210 ± 3.42	30.5 ± 1.79	203 ± 5.83	124 ± 13.6	23.3 ± 1.61	590 ± 15.4	611
1j: $Z \to \tau\tau$ veto	12.2 ± 0.22	204 ± 3.37	28.8 ± 1.73	193 ± 5.73	97.8 ± 11.7	22.7 ± 1.56	547 ± 13.6	580
1j: $m_{\ell\ell} < 50$ GeV	9.2 ± 0.19	37.3 ± 1.45	10.0 ± 1.10	32.8 ± 2.45	16.1 ± 4.56	8.0 ± 0.91	104 ± 5.56	122
1j: $\Delta\phi_{\ell\ell} < 1.8$ rad	8.6 ± 0.18	33.7 ± 1.37	9.2 ± 1.05	31.2 ± 2.39	3.2 ± 2.28	6.4 ± 0.72	83.6 ± 3.80	106
1j: $0.75 \cdot \mathrm{m_H} < \mathrm{m_T} < \mathrm{m_H}$	4.8 ± 0.14	11.8 ± 0.82	1.9 ± 0.40	9.7 ± 1.31	0.1 ± 0.06	2.3 ± 0.40	25.9 ± 1.64	38
$\geq$ 2j: $N_{\mathrm{jets}} \geq 2$	14.5 ± 0.2	139 ± 2.76	30.4 ± 1.54	7415 ± 26.6	104 ± 11.8	70.8 ± 3.74	7759 ± 29.5	7845
$\geq$ 2j: b-jet veto	9.6 ± 0.2	94.9 ± 2.25	19.2 ± 1.23	400 ± 6.77	61.7 ± 9.00	21.4 ± 1.79	597 ± 11.7	667
$\geq$ 2j: $\lvert\Delta\eta_{jj}\rvert > 3.8$	2.0 ± 0.1	8.3 ± 0.61	2.0 ± 0.40	36.7 ± 2.05	3.6 ± 2.13	1.4 ± 0.51	52.0 ± 3.09	44
$\geq$ 2j: CJV	1.6 ± 0.1	6.5 ± 0.53	1.3 ± 0.31	19.7 ± 1.55	1.2 ± 1.16	0.5 ± 0.34	29.2 ± 2.06	22
$\geq$ 2j: $m_{jj} > 500$ GeV	1.1 ± 0.04	3.2 ± 0.37	0.7 ± 0.23	8.0 ± 0.94	0.0 ± 0.00	(-0.2 ± 0.23)	11.9 ± 1.03	13
$\geq$ 2j: $p_T^{tot} < 30$ GeV	0.8 ± 0.04	1.7 ± 0.29	0.3 ± 0.13	3.3 ± 0.60	0.0 ± 0.00	(-0.2 ± 0.21)	5.3 ± 0.68	6
$\geq$ 2j: $Z \to \tau\tau$ veto	0.7 ± 0.04	1.8 ± 0.29	0.3 ± 0.13	3.2 ± 0.59	0.0 ± 0.00	(-0.2 ± 0.21)	5.2 ± 0.67	6
$\geq$ 2j: $m_{\ell\ell} < 80$ GeV	0.7 ± 0.03	0.6 ± 0.16	0.1 ± 0.10	1.1 ± 0.36	0.0 ± 0.00	(-0.2 ± 0.21)	1.8 ± 0.41	3
$\geq$ 2j: $\Delta\phi_{\ell\ell} < 1.8$ rad	0.6 ± 0.02	0.5 ± 0.15	0.1 ± 0.10	0.8 ± 0.33	0.0 ± 0.00	(-0.1 ± 0.21)	1.4 ± 0.38	2

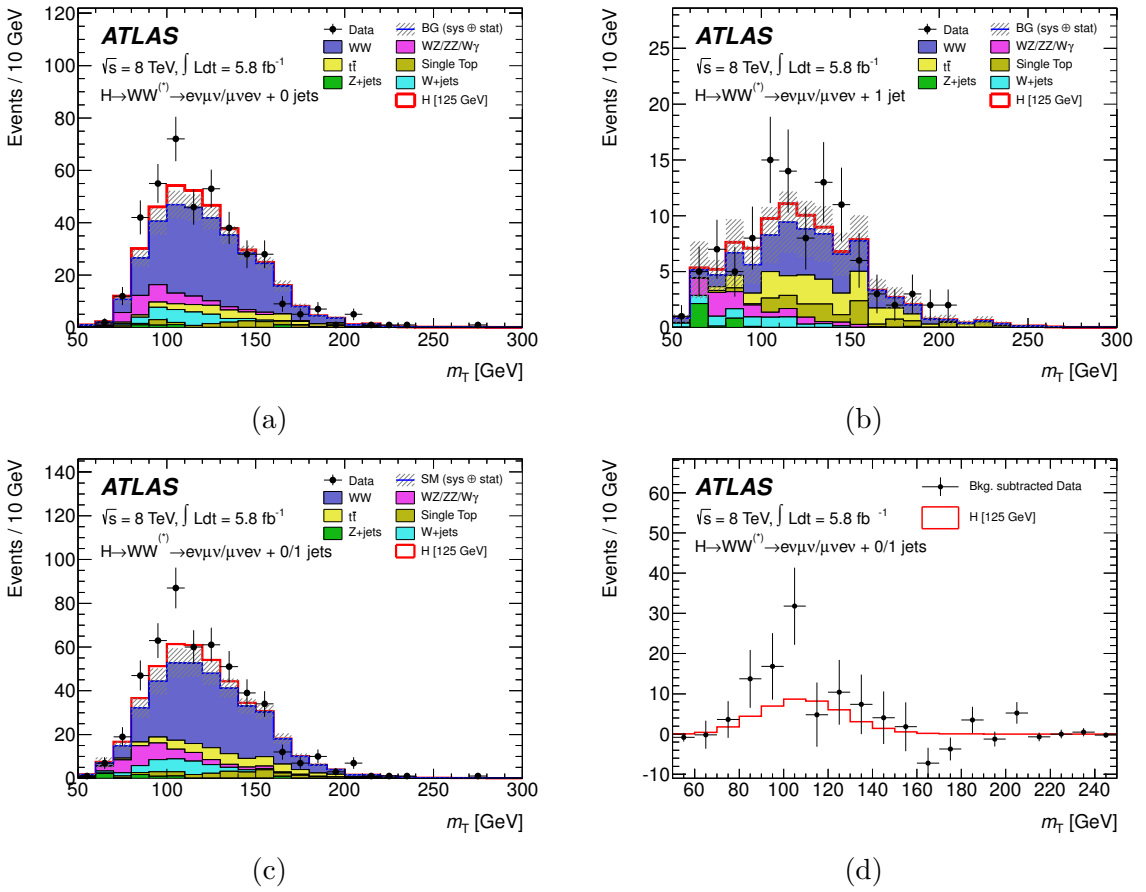


Figure 6.32.: Distributions of the transverse mass, m_{T}, (a) in the 0-jet and (b) 1-jet channels separately and (c) summed up for events satisfying all criteria for the low m_{H} selection. The lepton flavours are combined. The expected signal for a SM Higgs boson with $m_{\mathrm{H}} = 125$ GeV is superimposed. The hatched area indicates the total uncertainty on the background prediction. In order to visualise the excess of events, the estimated background is subtracted from data and compared to the predicted signal for a SM Higgs boson with $m_{\mathrm{H}} = 125$ GeV (d). The statistical uncertainties of both the data and the subtracted background are reflected by the error bars of the data points.

The expected and observed numbers of events in the 2012 dataset after applying all selection requirements including a cut on the transverse mass of $0.75 \cdot m_{\mathrm{H}} < m_{\mathrm{T}} < m_{\mathrm{H}}$ (for $m_{\mathrm{H}} = 125$ GeV) are displayed in Table 6.23. The additional cut on the transverse mass is applied to better reflect the sensitivity of the analysis. The uncertainties comprise the statistical and systematic uncertainties taking into account the constraints from control samples as discussed in Section 6.12.3. The magnitude of the main sources of systematic uncertainties on the signal and background predictions for the H + 0 jets and H + 1 jet analyses are shown in Table 6.24.

Table 6.23.: Expected and observed numbers of events in the 2012 dataset after applying all selection requirements including a cut on the transverse mass of $0.75 \cdot m_H < m_T < m_H$ for m_H =125 GeV. The uncertainties shown comprise the statistical and systematic uncertainties, taking into account the constraints from control samples. All numbers are summed over lepton flavours. For the H+ $\geq$ 2 jets analysis, background processes with fewer than 0.01 events expected are marked by "-" as negligible.

	H + 0 jets	H + 1 jet	H+ $\geq$ 2 jets
Signal	20±4	5±2	0.34±0.07
WW	101±13	12±5	0.10±0.14
$WZ/ZZ/W\gamma$	12±3	1.9±1.1	0.10±0.10
$t\bar{t}$	8±2	6±2	0.15±0.10
$tW/tb/tqb$	3.4±1.5	3.7±1.6	-
$Z/\gamma^* + \text{jets}$	1.9±1.3	0.1±0.1	-
$W + \text{jets}$	15±7	2±1	-
Total Bkg.	142±16	26±6	0.35±0.18
Observed	185	38	0

Table 6.24.: Relative systematic uncertainties on the predicted number of signal and background events (adapted from Ref. [4]).

Source of Uncertainty (0-jet)	Signal uncertainties (%) m_H =125 GeV	Background uncertainties (%)
Incl. QCD renormalisation and factorisation scale	13	-
1-jet incl. QCD renormalisation and factorisation scale	10	-
Parton distribution functions	8	2
Jet energy scale	7	4
WW modelling and shape	-	5
QCD scale acceptance	4	2
WW normalisation	-	4
$W + \text{jets}$ fake factor	-	4
Lepton isolation	3	3
Source of Uncertainty (1-jet)		
Incl. QCD renormalisation and factorisation scale	28	-
1-jet incl. QCD renormalisation and factorisation scale	16	-
WW normalisation	-	14
b-tagging efficiency	-	8
Top normalisation	-	6
Pile-up	5	5

6.15. Combination of the 2011 and 2012 Analysis Results

The statistical analysis of the data employs a binned likelihood function as already discussed in the context of the analysis of the 2011 dataset in Section 6.8. The likelihood function is expanded as a product of Poisson probability terms for the signal and control regions in each lepton flavour channel, jet multiplicity and m_{T} bin. The m_{T} distributions used in the binned likelihood are subdivided into five bins for the $\mathrm{H}+0$ jets signal region and into three bins for the $\mathrm{H}+1$ jet signal region. Concerning the H+ ≥ 2 jets signal region and the WW and top control regions, no binning is used and only the results integrated over m_{T} are included. It should be noted that the structure of the likelihood function changed slightly with respect to the one used for the 2011 analysis since the $e\mu$ and μe channel are treated separately. The full likelihood used for the 2012 analysis can be written as:

$$\mathcal{L}(\mu,\vec{\theta}) = \{\prod_{k=e\mu,\mu e}\prod_{j=0}^{N_{jets}}\prod_{i=1}^{N_{bins}} P(N_{ijk}|\mu s_{ijk}+\sum_{m}^{N_{bg}} b_{ijkm})\} \times \{\prod_{i=1}^{N_\theta} N(\tilde{\theta}|\theta)\} \tag{6.12}$$

where $P(N_{ijk}|\mu s_{ijk}+\sum_{m}^{N_{bg}} b_{ijkm})$ are the Poisson probability distributions and $N(\tilde{\theta}|\theta)$ are the constraints which are either represented as Gaussian or Poisson distributions. The signal and background expectations depend on systematic uncertainties that are parametrised as functions of the nuisance parameters θ. The four general sources of uncertainties are the same as for the 2011 analysis, namely flat systematic uncertainties, shape systematic uncertainties, statistical uncertainties due to the limited size of Monte Carlo samples and uncertainties due to the limited size of the data samples in the control regions.

In the case of shape systematics, only variations that are statistically significant are included. As a matter of fact, the shape variation of the m_{T} distributions of the total background in the $\mathrm{H}+0$ jets and $\mathrm{H}+1$ jet signal regions is dominated by the normalisation variation of the individual backgrounds. The variation in shape of the individual backgrounds has a comparatively small effect on the total shape. For all background processes, shape uncertainties due to b-tagging and lepton identification, trigger and isolation efficiencies are included. A separate shape uncertainty due to the fake rate estimate is included for the $W+$ jets background. It should be noted that for the 2012 analysis, the slight difference between the m_{T} shapes predicted by the `Powheg` and `MC@NLO` generators for the WW background has been taken into account as a shape uncertainty.

The statistical procedure is based on the same test statistic q_μ as defined in Equation 6.7. It is used to compute the p_0-value as defined in Equation 6.8 which is the probability of how a given value of the test statistic is compatible with the Monte Carlo prediction under the assumption that the *background-only* hypothesis is true. Thus, the local p_0-value, shown in Fig. 6.33, displays the probability that background processes can create a fluctuation that can cause an excess that is at least as large as the one observed in data. An excess of events in data over the SM background expectation is observed as reflected by a low observed p_0-value and a fitted signal strength parameter μ deviating from zero as shown in Fig. 6.33. Since the mass resolution is limited for this particular channel, the p_0 distribution is rather flat in the region around $\mathrm{m_H}$ =125 GeV. The p_0-value at $\mathrm{m_H}$ =125 GeV is 8×10^{-4} and corresponds to a significance of 3.1 σ. The minimum value observed for p_0 for the 2012 analysis is found at $\mathrm{m_H}$ =120 GeV and amounts to 6×10^{-4} or, equivalently, a significance of 3.2 σ. Thus, the significance exceeds three standard deviations within the Higgs boson mass range from 110 to 130 GeV. The expected p-value for the signal + background hypothesis for a SM Higgs boson with $\mathrm{m_H}$ =125 GeV is 0.05 which corresponds to 1.6 standard deviations for the 2012

analysis only. The fitted signal strength at m_H =125 GeV amounts to $\mu = 2.1^{+0.8}_{-0.7}$.

For the combination of the 2011 and 2012 analyses, the minimum value observed for p_0 is 3×10^{-3} for the Higgs boson mass m_H =125 GeV hypothesis and corresponds to a significance of 2.8 σ. The expected p-value for the signal plus background hypothesis for a SM Higgs boson with m_H =125 GeV is 0.01 or, equivalently, 2.3 standard deviations for the combined 7 TeV and 8 TeV analysis. The fitted signal strength at m_H =125 GeV is determined to be $\mu = 1.4 \pm 0.5$ and thus is consistent within uncertainties with the expectation for a SM Higgs boson. The observation in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ final state is compared to the results obtained in the $H \to ZZ \to \ell^+\ell^-\ell^+\ell^-$ and $H \to \gamma\gamma$ and channels in Chapter 6.16.

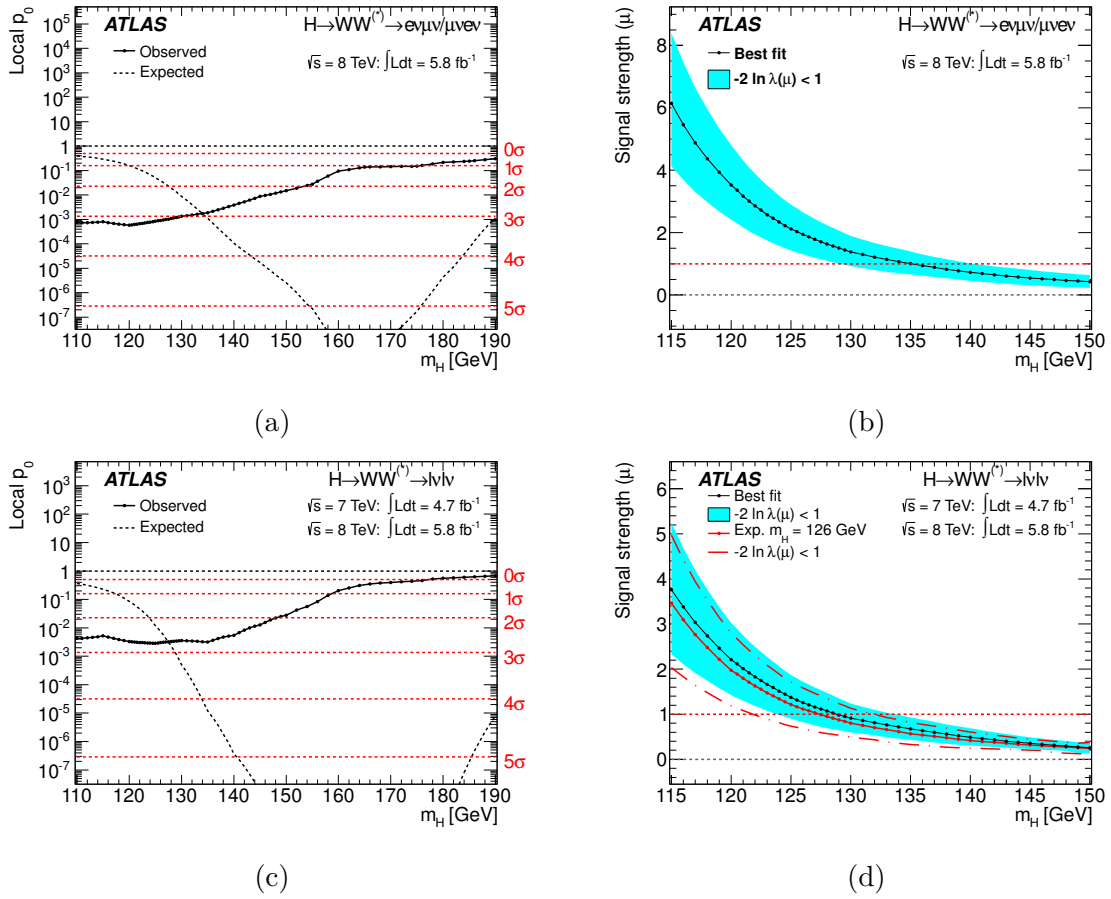


Figure 6.33.: Observed (solid line) local p_0 as a function of m_H (a) for the 8 TeV data and (c) the 7 TeV and 8 TeV combined dataset. The dashed line shows the corresponding expectation for the signal plus background hypothesis at the given value of m_H. The fitted signal strength parameter, μ, as a function of m_H (solid black line with cyan band) is shown on the right-hand side for (b) the 8 TeV data and (d) the 7 TeV and 8 TeV combined dataset. The expected result for a signal hypothesis of m_H = 126 GeV (red line) is included for the combined dataset for comparison.

6.16. Combined Search for the Standard Model Higgs Boson

The Standard Model Higgs boson is searched for in various decay channels in the ATLAS experiment. Especially in the low mass range of 110-130 GeV, non of the individual Higgs

boson decay channel has a discovery potential that would be high enough to discover the Higgs boson or rule out its existence based on the datasets taken in 2011 at $\sqrt{s} = 7$ TeV and between April and June 2012 at $\sqrt{s} = 8$ TeV. Therefore, the statistical combination of different decay channels is essential to reach a higher sensitivity.

In this section, the results obtained by the combination of individual searches are presented as published in Ref. [6]. The analyses in the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels provide the highest mass resolution. The results obtained in these two channels are discussed briefly in the following paragraphs.

6.16.1. Results of the Search in the $H \to ZZ^* \to 4\ell$ Channel

This section presents the search for the Standard Model Higgs boson in the decay channel $H \to ZZ \to 4\ell$, with ℓ denoting electrons and muons, as described in detail in Refs. [6, 185]. The $H \to ZZ^* \to 4\ell$ channel has a relatively low rate of expected signal events. The presence of four leptons in the final state requires an excellent lepton identification efficiency. Especially in case of a low-mass Higgs boson, even leptons with a transverse momentum down to 7 GeV have to be reconstructed well. The analysis is sub-divided into four channels, namely $4e$, $2e2\mu$, $2\mu2e$ and 4μ, arranged by the flavour of the leading lepton pair. The largest irreducible background to the $H \to ZZ^* \to 4\ell$ channel arises from continuum ZZ^* production. For low Higgs boson masses there are also important background contributions arising from Z/γ^*+jets and $t\bar{t}$ processes. The methods to estimate the background contributions are well adapted for the individual sub-channels. Details can be found in Ref. [185].

The expected distribution of the invariant mass of the quadruplet $m_{4\ell}$ for the background processes and for a Higgs boson signal with m_H =125 GeV are compared to the events observed in data in Fig. 6.34. The statistical interpretation of the excess of events near $m_{4\ell}$=125 GeV is presented in Section 6.17. The numbers of expected signal and background events and observed events in data in a window of ±5 GeV around $m_{4\ell}$=125 GeV are presented for the combined 2011 and 2012 datasets in Table 6.25.

Table 6.25.: The numbers of expected signal (m_H =125 GeV) and background events, together with the numbers of observed events in the 2011 and 2012 datasets, in a window of size ±5 GeV around $m_{4\ell}$=125 GeV (taken from Ref. [6]).

	Signal	ZZ^*	$Z/\gamma^* + \text{jets}, t\bar{t}$	Observed
4μ	2.09±0.30	1.12±0.05	0.13±0.04	6
$2e2\mu/2\mu2e$	2.29±0.33	0.80±0.05	0.13±0.04	5
$4e$	0.90±0.14	0.44±0.04	1.09±0.20	2

6.16.2. Results of the Search in the $H \to \gamma\gamma$ Channel

This section presents the search for the Standard Model Higgs boson in the diphoton decay channel, as described in detail in Refs. [6, 187]. The signature of this channel is characterised by two isolated, high energetic photons. The search is performed in the mass range $110 < m_H < 150$ GeV. In this mass range, the branching ratio of Higgs boson decays into two photons is large enough that a search is feasible.

The dominant background to the $H \to \gamma\gamma$ channel arises from SM diphoton production, referred to in the following as $\gamma\gamma$. Background contributions also come from γ+jet and jet+jet production with one or two jets misidentified as photons (called γj and jj in the following)

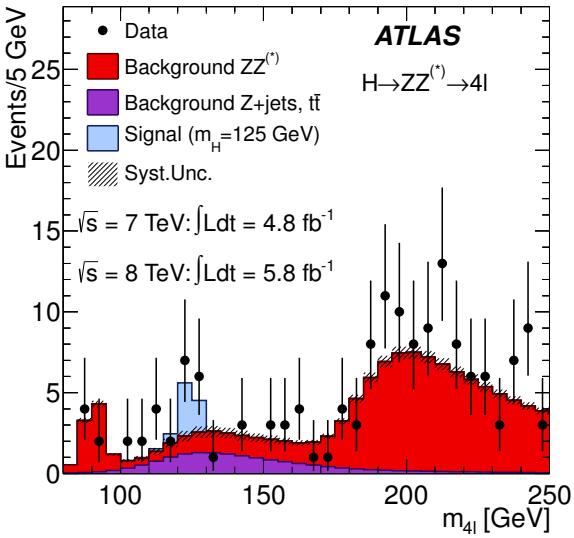


Figure 6.34.: Distribution of the four-lepton invariant mass m_{4l} for the selected candidates, compared to the background expectation in the mass range 80-250 GeV for the combination of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data. The signal expectation for a SM Higgs boson with m_H =125 GeV is also shown. Though a small excess of events is observed for $m_{4l} > 180$ GeV, the measured ZZ^* cross section [186] is consistent with the SM theoretical expectation (taken from Ref. [6]).

and from the Drell-Yan process. The γj and jj processes contribute to the background when one or two jets fragmenting into neutral mesons (mainly π^0) are misidentified as prompt photons. The composition of the selected sample and the expected background yields are estimated by data-driven techniques [187].

The invariant mass of the two photons is then estimated using the photon energies measured in the calorimeter, the ϕ coordinate as determined from the position of the photon in the calorimeter, and η as determined by the identified primary vertex and the photon impact point in the calorimeter. The distributions of the diphoton invariant mass, $m_{\gamma\gamma}$, for the inclusive sample and the combined $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples are shown in Fig. 6.35. The result of a fit including a signal component fixed to m_H =126.5 GeV and a background component described by a fourth-order Bernstein polynomial is superimposed.

Classifying the events into ten mutually exclusive categories with different signal to background ratios and different invariant mass resolutions increases the sensitivity to a Higgs boson signal.

The statistical analysis of the data employs an unbinned likelihood function constructed from the ten categories of the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples. Depending on the category, weights are obtained reflecting the signal to background ratios. The weights are applied to the invariant diphoton mass distribution shown in Fig. 6.35. The statistical interpretation of the excess of events near $m_{\gamma\gamma}$=126.5 GeV is presented in Section 6.17.

6.17. Statistical Treatment

The statistical analysis of the data employs a binned likelihood function similar to the one already discussed in the context of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis in Section 6.8 and

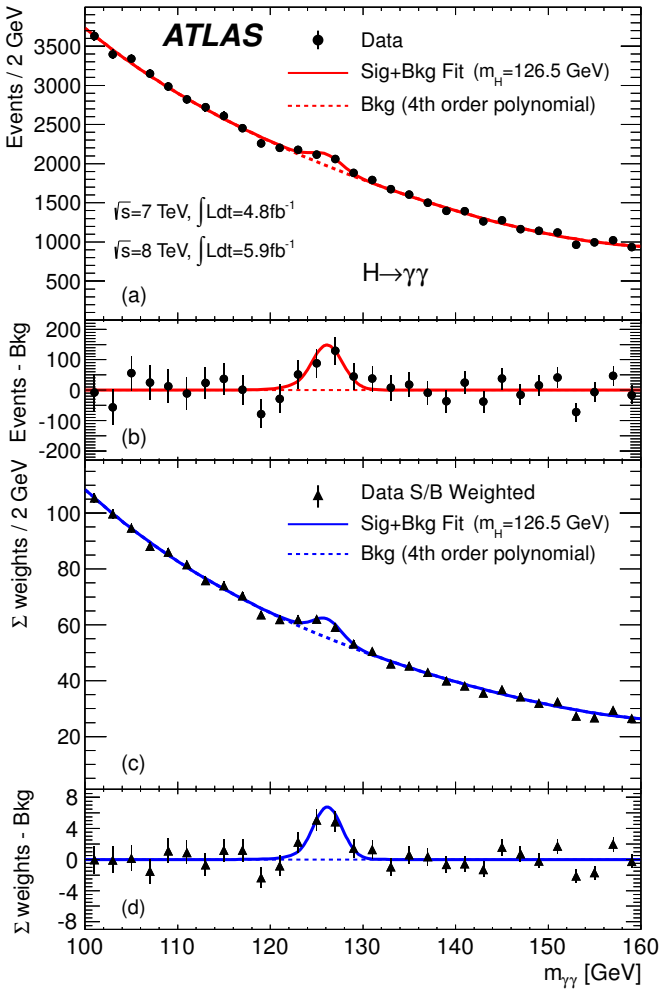


Figure 6.35.: Distributions of the invariant mass of diphoton candidates after all selection requirements are applied for the combination of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data. The inclusive sample is shown in a) and a weighed version of the sample in c). The results of a fit to the data of the sum of a signal component fixed to m_H =126.5 GeV and a background component described by a fourth-order Bernstein polynomial is superimposed. The residuals of the data and weighed data with respect to the respective fitted background components are displayed in b) and d) (taken from Ref. [6]).

described in Refs. [167, 188–191]. For each Higgs boson mass hypothesis the parameter of interest is the overall signal strength factor μ, which acts as a scale factor on the total number of events predicted by the Standard Model for the Higgs boson signal. The signal strength is defined such that μ=0 corresponds to the background-only model and μ=1 corresponds to the SM Higgs boson signal in addition to the background. The likelihood function used for the combination procedure includes all the individual sub-channels given in Table 6.26. Furthermore, all the parameters that describe the systematic uncertainties and their correlations are included. A test statistic $\lambda(\mu)$ is constructed using the profile likelihood ratio [174] in order to test hypothesised values of μ. The test statistic extracts the information on the signal strength from a full likelihood fit to the data.

The test statistic $\lambda(\mu)$ is used to compute the local p_0 value which is the probability that background processes can create a fluctuation that can cause an excess that is at least as large

as the one observed in data. An excess of events observed in data over the SM background expectation is reflected by a low observed p_0-value and a fitted signal strength parameter larger than zero.

The main uncorrelated systematic uncertainties were described in Sections 6.6 and 6.13 for the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channel. The main sources of correlated experimental systematic uncertainties are the uncertainties on the integrated luminosity, the electron and photon trigger identification efficiencies, the electron and photon energy scales, the muon reconstruction and identification efficiencies, jet energy scale and resolution and missing transverse energy. Most of them are considered as fully correlated among the channels. The correlated theoretical uncertainties affect mostly the signal predictions. Sources of systematic uncertainty that affect both the $\sqrt{s} = 7$ TeV and the $\sqrt{s} = 8$ TeV data are taken as fully correlated while the uncertainties on the background estimates based on control samples in data are considered as uncorrelated between the two datasets.

Table 6.26.: Summary of the individual channels contributing to the combination. The transition points between separately optimised m_H regions are indicated where applicable. In channels sensitive to associated production of the Higgs boson, V indicates a W or Z boson. The symbols $\otimes$ and $\oplus$ represent direct products and sums over sets of selection requirements, respectively (taken from Ref. [6]).

Higgs Boson Decay	Subsequent Decay	Sub-Channels	m_H Range [GeV]	$\int L\,dt$ [fb^{-1}]	Ref.
		2011 $\sqrt{s} = 7$ TeV			
$H \to ZZ^{(*)}$	4ℓ	$\{4e, 2e2\mu, 2\mu2e, 4\mu\}$	110-600	4.8	[185]
	$\ell\ell\nu\bar{\nu}$	$\{ee, \mu\mu\} \otimes$ {low, high pile-up}	200-280-600	4.7	[192]
	$\ell\ell q\bar{q}$	$\{b$-tagged, untagged$\}$	200-300-600	4.7	[193]
$H \to \gamma\gamma$	–	10 categories $\{p_{\mathrm{Tt}} \otimes \eta_\gamma \otimes$ conversion$\} \otimes \{2-$jet$\}$	110-150	4.9	[187]
$H \to WW$	$\ell\nu\ell\nu$	$\{ee, e\mu, \mu\mu\} \otimes$ {0-jet, 1-jet, 2-jet} $\otimes$ {low, high pile-up}	110-200-300-600	4.7	[4]
	$\ell\nu qq'$	$\{e, \mu\} \otimes$ {0-jet, 1-jet, 2-jet}	300-600	4.7	[194]
$H \to \tau^+\tau^-$	$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$	$\{e\mu\} \otimes$ {0-jet} $\oplus \{\ell\ell\} \otimes$ {1-jet, 2-jet, VH}	110-150	4.7	[195]
	$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	$\{e, \mu\} \otimes$ {0-jet} $\otimes \{E_T^{\mathrm{miss}} \gtrless 20$ GeV, $E_T^{\mathrm{miss}} \geq 20$ GeV$\}$ $\oplus \{e, \mu\} \otimes$ {1-jet} $\oplus \{\ell\} \otimes$ {2-jet}	110-150	4.7	
	$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$	{1-jet}	110-150	4.7	
$VH \to bb$	$Z \to \nu\nu$	$E_T^{\mathrm{miss}} \in \{120-160, 160-200, \geq 200$ GeV$\}$	110-130	4.6	[196]
	$W \to \ell\nu$	$p_T^W \in \{<50, 50-100, 100-200, \geq 200$ GeV$\}$	110-130	4.7	
	$Z \to \ell\ell$	$p_T^Z \in \{<50, 50-100, 100-200, \geq 200$ GeV$\}$	110-130	4.7	
		2012 $\sqrt{s} = 8$ TeV			
$H \to ZZ^{(*)}$	4ℓ	$\{4e, 2e2\mu, 2\mu2e, 4\mu\}$	110-600	5.8	[185]
$H \to \gamma\gamma$	–	10 categories $\{p_{\mathrm{Tt}} \otimes \eta_\gamma \otimes$ conversion$\}$	110-150	5.9	[187]
$H \to WW$	$e\nu\mu\nu$	$\{e\mu, \mu e\} \otimes$ {0-jet, 1-jet, 2-jet}	110-200	5.8	[5]

6.18. Results of the Combined 2011 and 2012 Datasets

The results obtained for the $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^{(\star)} \to e\nu\mu\nu$ channels based on the 2012 dataset are combined with the improved results obtained for the first two analyses on the 2011 datasets and with previously published results for searches for $H \to ZZ^{(\star)}$, $WW^{(\star)}$, $b\bar{b}$ and $\tau^+\tau^-$ based on the 2011 dataset. It should be noted that especially the addition of the $\sqrt{s} = 8$ TeV data led to a significant gain in sensitivity in the low-mass region with respect to the previous combined search [175].

The observed local p_0 values as a function of the hypothesised Higgs boson mass are shown for the $H \to ZZ^* \to 4\ell$, $H \to \gamma\gamma$ and $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channels in Fig. 6.36 for

the combined 2011 and 2012 datasets. An excess of events observed in data over the SM background expectation near m_H =126 GeV in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels is reflected by a minimum in the p_0 distribution in this region. Both channels provide a high resolution in invariant mass. These excesses are confirmed by the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channel that features a much lower mass resolution. The observed local p_0 for the combination of all channels are shown as a function of the hypothesised Higgs boson mass in the mass range 110 to 150 GeV as well in Fig. 6.36. The minimum value observed for p_0 is found at m_H =126.5 GeV where the corresponding local significance reaches 6.0 σ. The expected value in the presence of a SM Higgs boson signal at that mass amounts to 4.9 σ. This constitutes a conclusive evidence for the production of a new particle. For the 2012 dataset alone, the maximum local significance for the $H \to ZZ^* \to 4\ell$, $H \to \gamma\gamma$ and $H \to WW^{(\star)} \to e\nu\mu\nu$ channels combined is observed at m_H =126.5 GeV and amounts to 4.9 σ, with an expected value of 3.8 σ.

The global significance of a local 5.9 σ excess anywhere in the mass range 110-600 GeV is estimated to be approximately 5.1 σ, increasing to 5.3 σ if the look elsewhere effect is restricted to the mass intervall 110 to 150 GeV. That is the mass range not yet excluded at 99% CL by the LHC combined SM search [197, 198] and the indirect constraints from the global fit to electroweak precision measurements [2].

The mass of the observed new particle is estimated using the likelihood ratio $\lambda(m_H)$ for the two channels with the highest mass resolution, namely $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$. The signal strength parameter is allowed to vary independently. The leading contributions to systematic uncertainties arise form the electron and photon energy scales and resolutions. The resulting estimate for the mass of the new particle is 126.0±0.4 (stat.) ±0.4 (syst.) GeV [6].

The observed excess corresponds to a best-fit signal strength of $\hat{\mu}$= 1.4 ± 0.3 for m_H = 126 GeV which is consistent within the assigned uncertainties with the SM Higgs boson hypothesis of μ=1.0. A summary of the individual and combined best-fit values of the signal strength parameter for a SM Higgs boson mass hypothesis of 126 GeV is shown in Fig. 6.37.

A profile likelihood $\lambda(\mu, m_H)$ is used to test which values of the strength parameter and the mass of a signal hypothesis are simultaneously consistent with the data. Closed contours around the best-fit point $\hat{\mu}, \hat{m}_h$ are expected in case a strong signal is present. In case of the absence of a signal the contours will be upper limits on μ for all values of m_H. Asymptotically, the test statistic -2 $\ln\lambda(\mu, m_H)$ is distributed as a χ^2 distribution with two degrees of freedom. The resulting 68% and 95% CL contours for the $H \to \gamma\gamma$ and $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channels are shown in Fig. 6.38. For the $H \to ZZ^* \to 4\ell$ channel similar contours are shown. The lack of mass resolution in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ final state for low m_H is evident from the contours shown. Nevertheless, the best-fit values for $\hat{\mu}$ and $\hat{m}_h$ are in good agreement for all three analyses.

The observed decays to pairs of vector bosons whose net electric charge is zero identify the new particle as a neutral boson. Furthermore, the observed excess in the diphoton final state disfavours the spin-1 hypothesis [199,200]. Although the presented results based on the datasets taken in 2011 at $\sqrt{s} = 7$ TeV and between April and June 2012 at $\sqrt{s} = 8$ TeV are compatible with the hypothesis that the new particle is the Standard Model Higgs boson, more data are needed to determined its nature in detail.

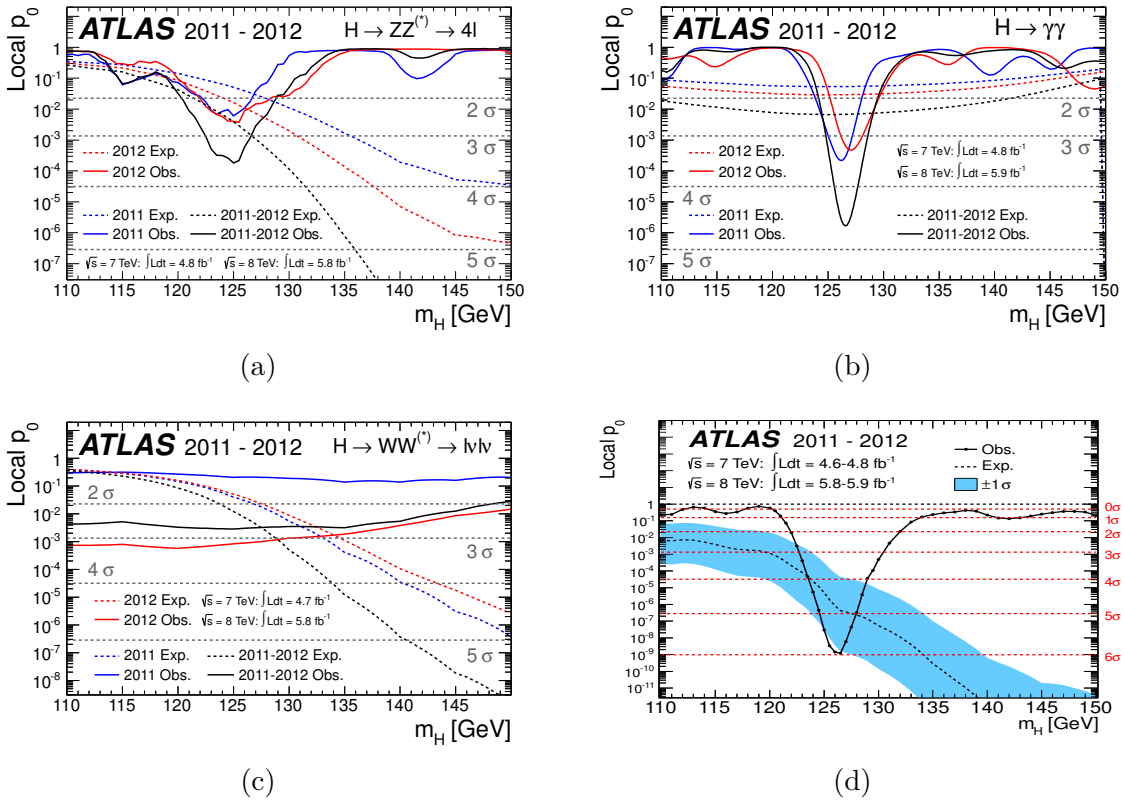


Figure 6.36.: The observed local p_0 as a function of the hypothesised Higgs boson mass for (a) the $H \to ZZ^* \to 4\ell$, (b) the $H \to \gamma\gamma$ and (c) the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channels. Results are shown separately for the $\sqrt{s} = 7$ TeV data (dark blue), the $\sqrt{s} = 8$ TeV data (light red) and their combination (black). The dashed curve shows the expected local p_0 under the hypothesis of a SM Higgs boson signal at that mass. The observed local p_0 as a function of the hypothesised Higgs boson mass for the combination of all channels listed in Table 6.26 is shown in the bottom right subfigure (d). The dashed curve shows the expected local p_0 under the hypothesis of a SM Higgs boson signal at that mass and its $\pm 1\sigma$ band. The horizontal dashed lines indicate the p-values corresponding to significances of 1 to 6 σ (taken from Ref. [6]).

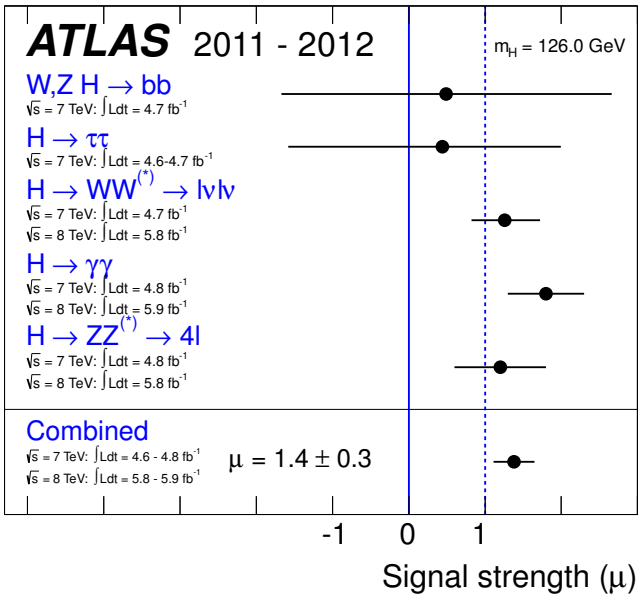


Figure 6.37.: Measurements of the signal strength parameter μ for m_H =126 GeV for the individual channels and their combination (taken from Ref. [6]).

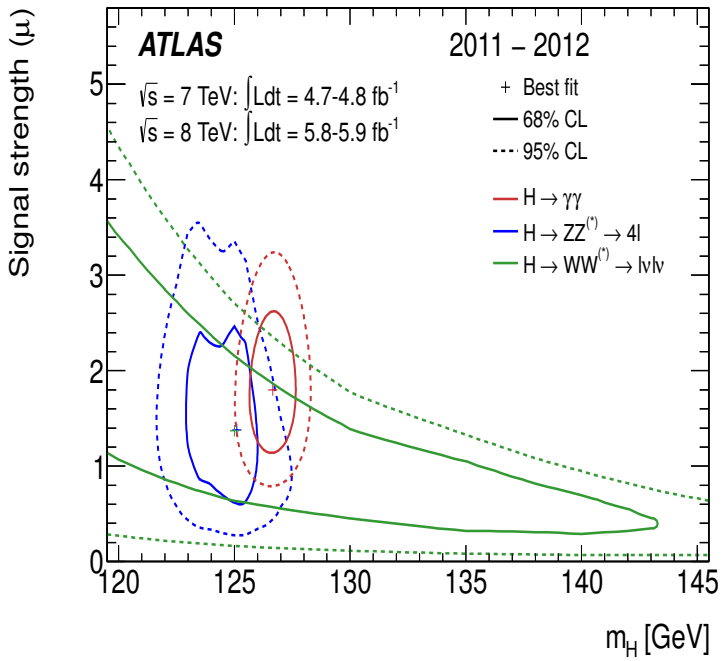


Figure 6.38.: Confidence intervals in the (μ,m_H) plane for the $H \to ZZ^* \to 4\ell$, $H \to \gamma\gamma$ and $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ channels, including all systematic uncertainties. The markers indicate the maximum likelihood estimates ($\hat{\mu}$,$\hat{m}_H$) in the corresponding channels (the maximum likelihood estimates for $H \to ZZ^* \to 4\ell$ and $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ coincide) (taken from Ref. [6]).

7

Extended Diboson Studies using a WZ Control Sample

The Standard Model W^+W^- process constitutes the largest background to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ signal. The present estimation consists of identifying a signal-free control region enriched in dileptonic W^+W^- events. While it helps to control the overall normalisation of this background in the signal region, it is not suitable for inspections of important kinematic distributions since the control region is not sufficiently pure. A totally new method was therefore developed to address this important question based on the kinematic similarity between the dileptonic W^+W^- and trileptonic $W^\pm Z \to \ell\nu\ell\ell$ decays. It allows a direct comparison of the kinematic distributions between data and Monte Carlo diboson predictions and opens the possibility to further improve the understanding of the W^+W^- background.

A first study was performed on a dataset corresponding to an integrated luminosity of 5.8 fb^{-1} (see Section 6.12.5) but the number of events observed in data was very limited. Therefore, the diboson studies were repeated on a larger dataset corresponding to an integrated luminosity of 10.7 fb^{-1}. This section first describes the results obtained in the WZ control region, then investigates the effect on the data to Monte Carlo agreement after a reweighing based on the transverse momentum of the W boson candidate is performed. Finally the transfer of the observations to the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ signal region and the impact on the results are discussed.

7.1. Event Selection and Kinematic Distributions

The object and event selection is identical to the one discussed in Section 6.12.5. The expected number of events originating from WZ and ZZ decays and the background contributions from W + jets, Z/γ^* + jets and top production processes are listed at several stages of the selection in Table 7.1. At the last stage, a veto on the presence of additional reconstructed jets is applied to obtain events with kinematics as close as possible to the zero-jet bin of the H + 0 jets analysis. There are 738 events observed in data before the zero-jet requirement and 390 events are observed in data after applying this requirement, about twice as many events as observed in the dataset corresponding to an integrated luminosity of 5.8 fb^{-1} (see Section 6.12.5). The corresponding distributions of the missing transverse energy, the jet multiplicity,

the dilepton invariant mass $m_{\ell\ell}$ of the Z boson candidate and the transverse mass of the W boson candidates are shown in Fig. 7.1 for events that pass all selection criteria. The overall number of events observed in data is about 15% higher compared to the total number of events predicted by Monte Carlo simulation. A part of the discrepancy is most likely caused by the fact that the $Z/\gamma^* + \text{jets}$ background contribution is not estimated from data. The selected dataset is now divided further according to whether the lepton, which belongs to the W boson candidate, is identified as an electron or muon. It is expected that in the case of an identified muon the contribution of the $Z/\gamma^* + \text{jets}$ background is smaller compared to the case of an identified electron.

Table 7.1.: The expected numbers of signal (WZ and ZZ) and background events as well as the events observed in data are listed at several stages of the selection. The lepton flavours are combined. Only statistical uncertainties associated with the number of events in the MC samples are shown. The event selection criteria are adopted from the analysis performed on the 2011 dataset to measure the $W^{\pm}Z$ production cross section documented in Ref. [184].

	WZ	ZZ	WW	$W\gamma$	Top	$Z/\gamma^* + \text{jets}$	Total MC	Obs.
lepton p_{T}	909.50 ± 5.11	259.29 ± 0.60	0.84 ± 0.17	0.20 ± 0.14	29.81 ± 2.22	439.71 ± 32.90	1639.35 ± 33.43	1935
OS leptons	908.21 ± 5.10	258.93 ± 2.01	0.84 ± 0.17	0.20 ± 0.14	29.81 ± 2.22	439.37 ± 32.90	1637.36 ± 33.43	1934
$\lvert m_{\ell\ell} - m_Z \rvert < 10$ GeV	785.77 ± 4.77	160.19 ± 1.57	0.23 ± 0.09	0.0 ± 0.0	10.16 ± 1.22	211.82 ± 23.34	1168.17 ± 23.90	1362
$E_{\text{T}}^{\text{miss}} > 25$ GeV	641.94 ± 4.31	74.85 ± 1.05	0.19 ± 0.08	0.0 ± 0.0	9.17 ± 1.16	83.98 ± 14.94	810.12 ± 15.63	919
$m_{\text{T,Wcand}} > 20$ GeV	550.24 ± 3.97	53.88 ± 0.89	0.09 ± 0.05	0.0 ± 0.0	7.32 ± 1.05	27.31 ± 7.92	638.85 ± 8.97	738
fraction in %	86.1	8.4	< 0.01	negl.	1.1	4.3		
WCand $\rightarrow \mu\nu$	321.84 ± 3.05	30.02 ± 0.67	0.0 ± 0.0	0.0 ± 0.0	4.66 ± 0.77	5.72 ± 2.66	362.25 ± 4.17	413
WCand $\rightarrow e\nu$	228.40 ± 2.55	23.87 ± 0.58	0.09 ± 0.05	0.0 ± 0.0	2.66 ± 0.70	21.58 ± 7.46	276.60 ± 7.94	325
jet veto	299.49 ± 2.93	23.26 ± 0.59	0.02 ± 0.02	0.0 ± 0.0	0.61 ± 0.34	13.81 ± 5.97	337.25 ± 6.69	390
fraction in %	88.8	6.9	< 0.01	negl.	< 0.2	4.1		

It is evident from the numbers in Table 7.1 that the selected sample is very pure. The only nameable background contribution arises from $Z/\gamma^* + \text{jets}$ events and constitutes about 4% of the total sample. Again, the main objective of this study is a direct comparison of the diboson transverse mass shape as predicted by Monte Carlo to the one observed in data. The transverse mass distributions of the WZ system as defined in Equation 6.11 are shown in Figs. 7.2 to 7.4 for events that pass all selection criteria, for events where the lepton assigned to the W boson candidate is identified as a muon in addition and for events that pass all selection criteria and have zero reconstructed jets, respectively. The transverse mass distributions of the modified system where one of the leptons attributed to the Z boson candidate is assigned to $E_{\text{T}}^{\text{miss}}$ are also shown. The transverse mass of the modified system is calculated in the same way as for the $H \rightarrow W^+W^- \rightarrow \ell^+\nu\ell^-\bar{\nu}$ analysis and detailed in Table 6.4. The Monte Carlo background predictions for all components except the WZ and ZZ contributions are subtracted from data and the resulting transverse mass plots as well as the ratios of the data distributions over the Monte Carlo predictions are shown in Figs. 7.2 to 7.4, too.

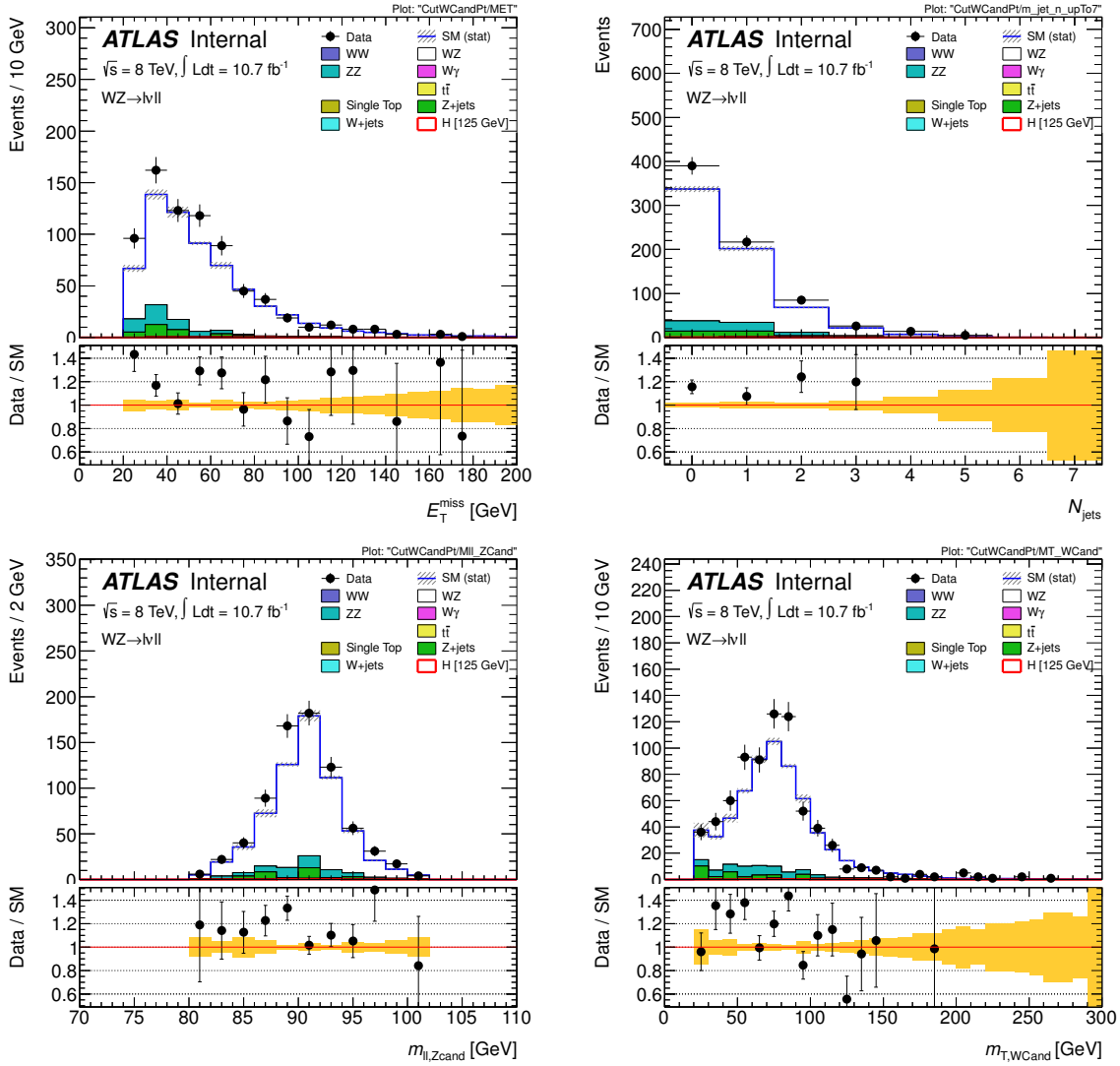


Figure 7.1.: Distribution of the missing transverse energy (top left), the jet multiplicity (top right), the dilepton invariant mass $m_{\ell\ell}$ of the Z boson candidate (bottom left) and the transverse mass of the W boson candidates (bottom right) for events that pass all selection criteria. The lower part of each plot shows the ratio between the data and the Monte Carlo expectations with the yellow band indicating the statistical uncertainty in the normalisation of the various components.

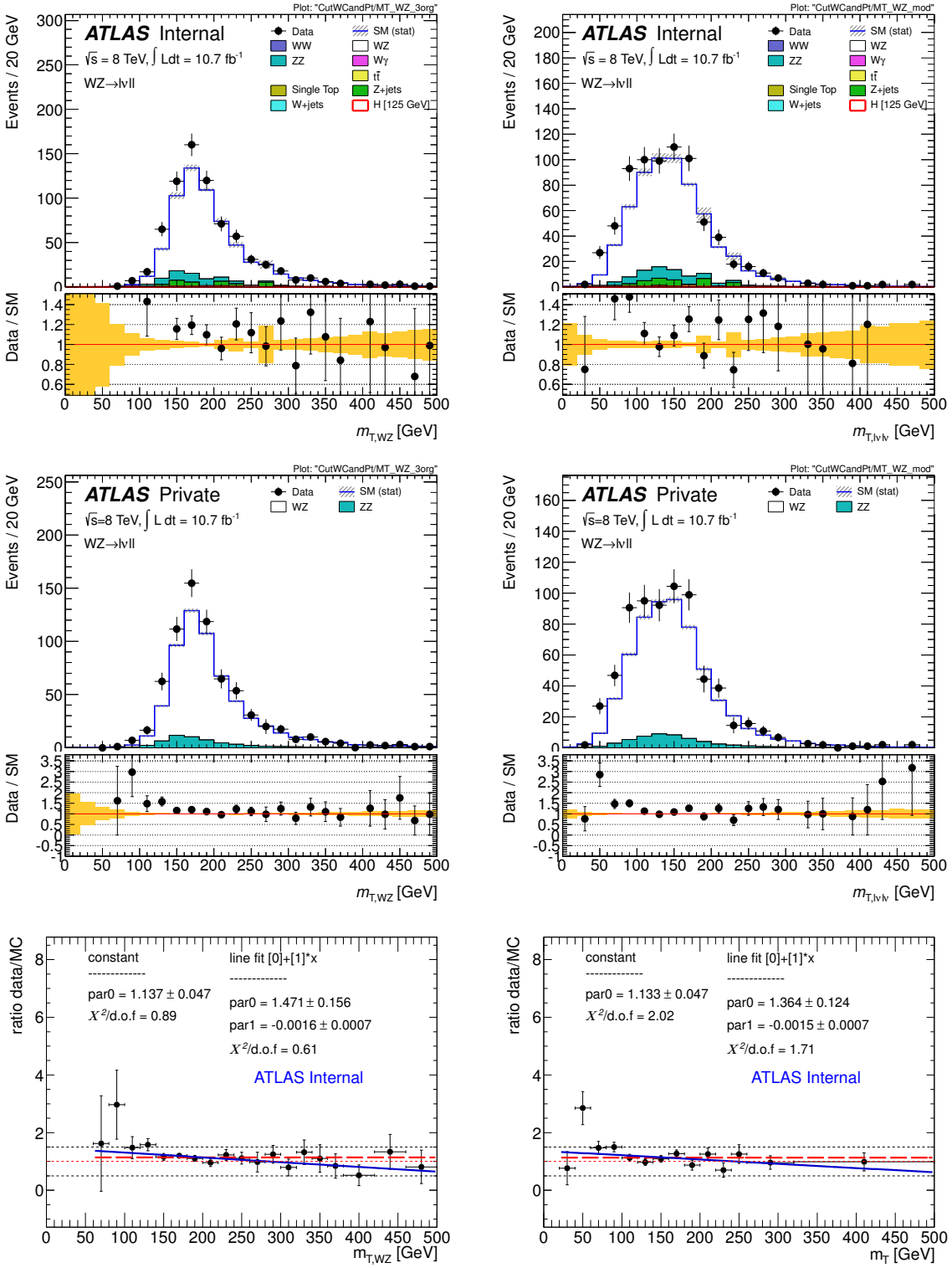


Figure 7.2.: Distributions of the transverse mass of the WZ pair (top left) and of the modified WZ system (top right), where one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to mimic a system of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$, for events that pass all selection criteria. The Monte Carlo background predictions for all components except the WZ and ZZ contributions are subtracted from data in order to compare the MC prediction of WZ and ZZ events directly to data (middle row). The ratios of the data distributions over the Monte Carlo predictions are fitted by a constant and a linear function in order to check for a difference in the overall rate and shifts between the Monte Carlo and data distributions (bottom row).

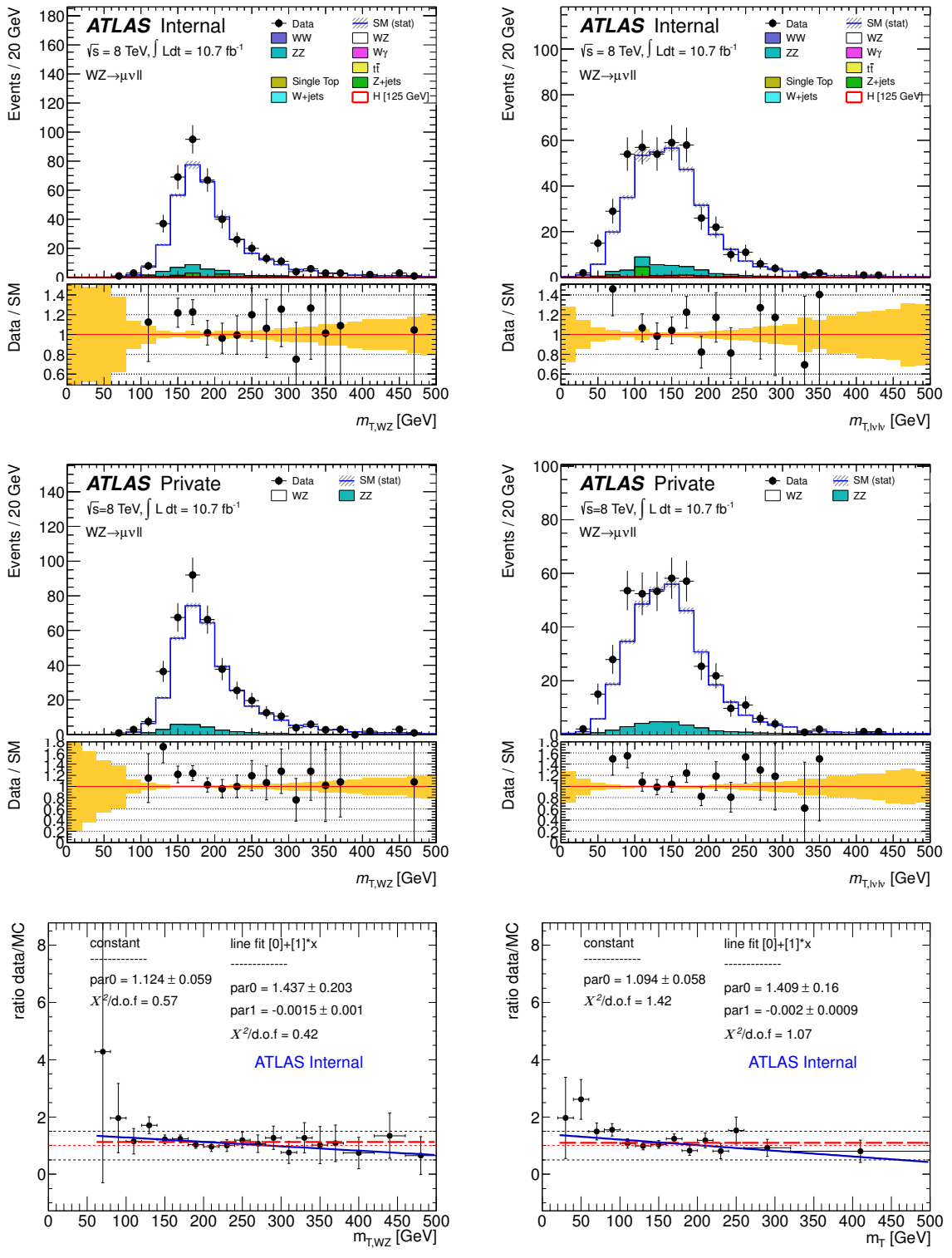


Figure 7.3.: Distributions of the transverse mass of the WZ pair (top left) and of the modified WZ system (top right), where one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to mimic a system of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$, for events that pass all selection criteria and the lepton assigned to the W boson candidate is identified as a muon. The Monte Carlo background predictions for all components except the WZ and ZZ contributions are subtracted from data in order to compare the MC prediction of WZ and ZZ events directly to data (middle row). The ratios of the data distributions over the Monte Carlo predictions are fitted by a constant and a linear function in order to check for a difference in the overall rate and shifts between the Monte Carlo and data distributions (bottom row).

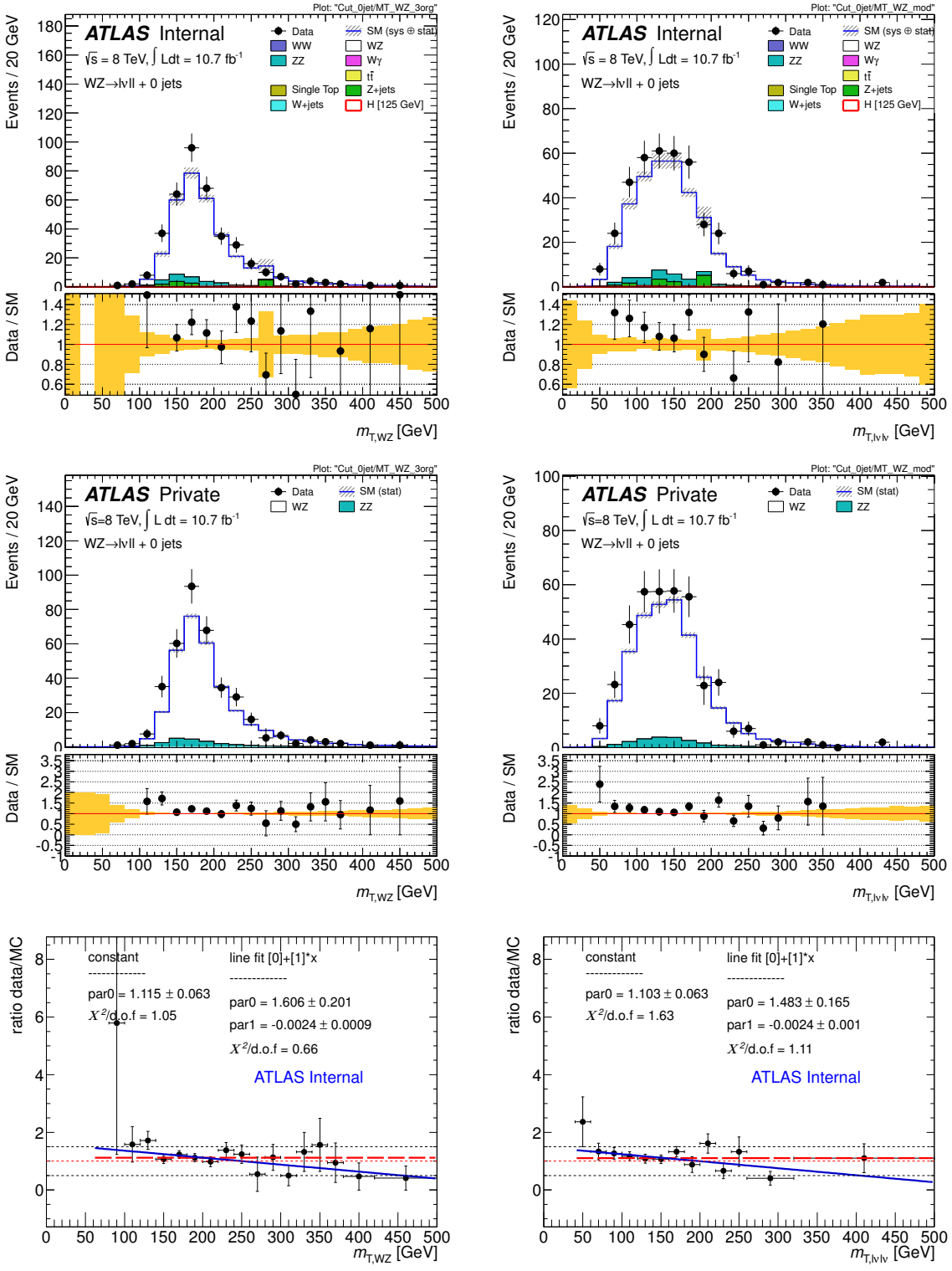


Figure 7.4.: Distributions of the transverse mass of the WZ pair (top left) and of the modified WZ system (top right), where one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to mimic a system of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$, for events that pass all selection criteria and have zero reconstructed jets. The Monte Carlo background predictions for all components except the WZ and ZZ contributions are subtracted from data in order to compare the MC prediction of WZ and ZZ events directly to data (middle row). The ratios of the data distributions over the Monte Carlo predictions are fitted by a constant and a linear function in order to check for a difference in the overall rate and shifts between the Monte Carlo and data distributions (bottom row).

The ratios are again fitted by a constant and a linear function in order to check for a difference in the overall rate and shifts between the Monte Carlo and data distributions. The bins in the high mass tail of the transverse mass distributions are again merged to avoid empty bins that cannot be handled by the χ^2 Fit that is performed. Negative slope parameters are observed for both the transverse mass of the WZ system and the modified system that mimics the signature of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$. The obtained slope parameter values and their statistical uncertainties are summarised in Table 7.2. All of them are different from zero by at most 2.7 standard deviations. The trend of slightly negative slope parameters that was observed already on a dataset corresponding to an integrated luminosity of 5.8 fb^{-1} is confirmed and gets even more pronounced. It should be pointed out that the effect is observed for events that pass all selection criteria and for events passing the additional requirement that the lepton assigned to the W boson candidate is identified as muon. Concerning the latter events, a lower contribution from $Z/\gamma^* +$ jets background is expected. Thus, the observed effect should not be affected by the uncertainty on the expected Z/γ^*+jets contribution that is taken from Monte Carlo prediction for this study. The observed shift between the transverse mass distributions observed in data and predicted by Monte Carlo diboson simulation is of comparable size for events that pass all selection criteria and for events that have zero reconstructed jets in addition.

Table 7.2.: Summary of slope parameters extracted by linear fits to the ratio of the transverse mass distribution obtained in data over the combined WZ and ZZ Monte Carlo prediction (shown in Fig. 7.2, Fig. 7.3 and Fig. 7.4).

	$m_{T,WZ}$ distribution	$m_{T,l\nu l\nu}$ distribution
slope parameter for events passing all selection criteria	-0.0016 ± 0.0007	-0.0015 ± 0.0007
slope parameter for events where the lepton assigned to the W-boson candidate is identified as a muon	-0.0015 ± 0.0010	-0.0020 ± 0.0009
slope parameter for events with zero reconstructed jets	-0.0024 ± 0.0009	-0.0024 ± 0.0010

7.2. Reweighing Studies

The distribution of the transverse momentum of the W boson is shown in Fig. 7.5 for events that pass all selection criteria. There are differences observed between the distribution predicted by Monte Carlo simulation and the one observed in data. The two-dimensional distribution of the transverse mass of the WZ system versus the transverse momentum of the W boson is shown for WZ events predicted by Monte Carlo simulation in Fig. 7.5. It is evident from the plot that the transverse mass of the WZ system and the transverse momentum of the W boson are strongly correlated variables. Similarly, a correlation is observed between the transverse mass of the modified WZ system and the transverse momentum of the W boson candidate. Thus, the mismodelling of the W boson transverse momentum is a possible source of the discrepancy observed between data and Monte Carlo in the WZ transverse mass distribution.

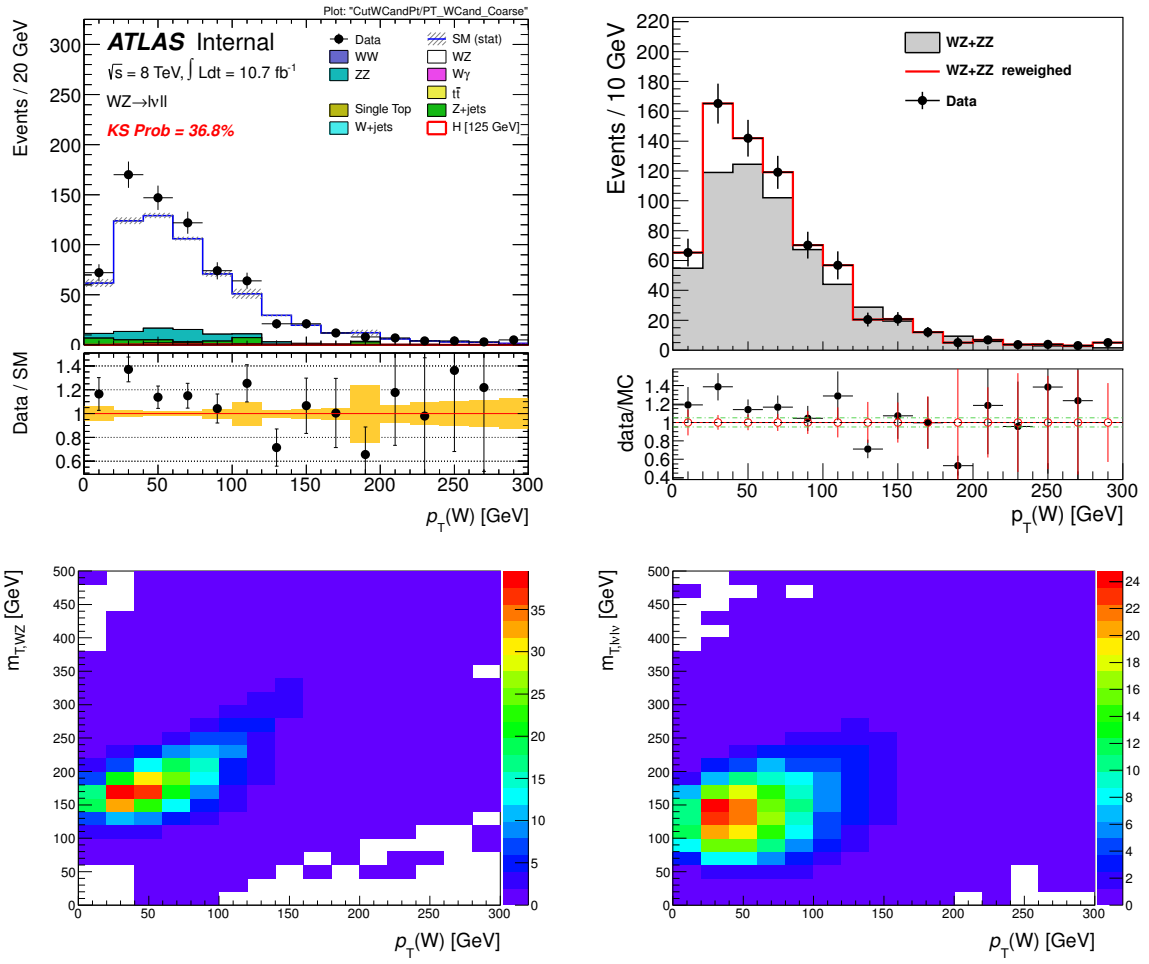


Figure 7.5.: Distributions of the transverse momentum of the W boson candidate (top left) before and after reweighing the Monte Carlo distribution to data (top right) for events passing all selection criteria. Correlation between the transverse momentum of the W boson candidate and the transverse mass of the WZ pair (bottom left) and the transverse momentum of the modified WZ system (bottom right), where one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to mimic a system of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$, for simulated WZ and ZZ events that pass all selection criteria.

The Monte Carlo background predictions for all components except the WZ and ZZ contributions are subtracted from data and the resulting transverse momentum distribution for events that pass all selection criteria is shown in Fig. 7.5. A reweighing of the Monte Carlo distribution to data is performed in each bin of the transverse momentum distribution. In order to check the impact of this reweighing in p_{T} of the W boson on the shape of the transverse mass distribution, the resulting Monte Carlo distribution is normalised to the initial number of simulated events. The distributions obtained after reweighing compared to the original transverse mass distributions are shown in Fig. 7.6. The ratios of the data distributions over the Monte Carlo predictions after the reweighing are once again fitted by a constant and a linear function. Slightly negative slope parameters are observed. However, they are less significant and consistent with zero within at most 1.3 standard deviations. The improvement is more pronounced in case of the transverse mass of the WZ system. This is

probably due to the more distinctive correlation of that quantity to the transverse momentum of the W boson candidate compared to the one observed between the transverse mass of the modified system and the p_{T} of the W boson candidate. These results lead to the conclusion that the observed shift in the transverse mass distributions between data and Monte Carlo are mostly caused by a mismodelling of the transverse momentum of the W boson candidate.

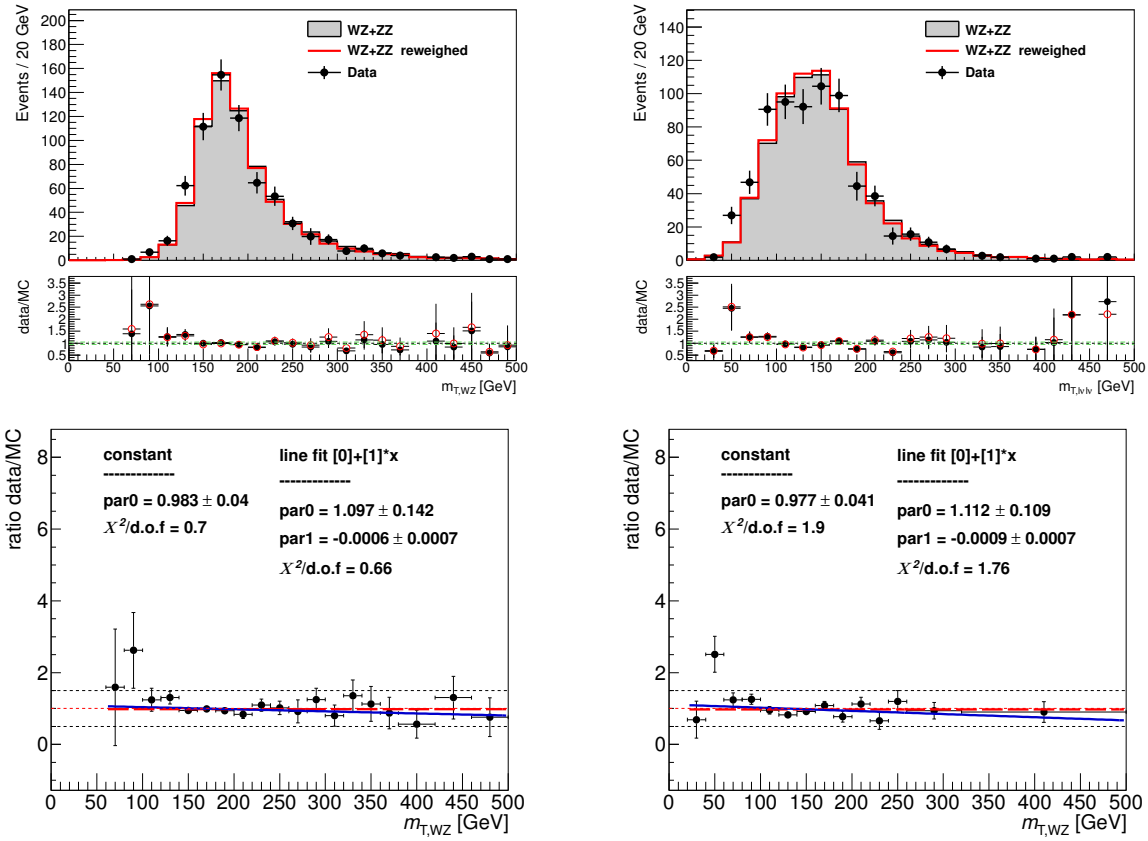


Figure 7.6.: Distributions of the transverse mass of the WZ pair (top left) and of the modified WZ system (top right), where one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to mimic a system of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$, for events that pass all selection criteria and have zero reconstructed jets. The transverse mass distribution of simulated WZ and ZZ events that is obtained after performing a bin-by-bin reweighing of the Monte Carlo simulation to data using the transverse momentum of the W boson candidate is shown in red. The distributions obtained after the reweighing are normalised to the initial number of simulated events in order to study a change in the predicted shape, only. The ratios of the data distributions over the Monte Carlo predictions obtained after reweighing are fitted by a constant and a linear function in order to check for a difference in the overall rate and shifts between the Monte Carlo and data distributions (bottom row).

7.3. Transfer to the $H \to W^+W^- \to \ell\nu\ell\nu$ + 0 jets Signal Region

The observed discrepancy in the transverse mass m_{T} of the WZ and ZZ systems gives rise to the assumption of a similar mismodelling of m_{T} of the WW system in the H + 0 jets region. However, the aforementioned reweighing procedure cannot be easily applied to WW

Monte Carlo events since it is not possible to reconstruct the transverse momenta of the two W bosons individually in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ dilepton final state.

To assess the impact of the observed mismodelling in the transverse mass distribution on the shape of the simulated WW m_{T} distribution, it is attempted to transfer the gained knowledge to the $\mathrm{H}+0$ jets signal region. As a first step, the transverse mass distributions of simulated WW events are compared to WZ and ZZ events that pass the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ and $W^{\pm}Z$ event selection, respectively. Figure 7.7 shows two sets of m_{T} distributions

- of events passing all selection criteria including the zero-jet requirement, and,
- of events passing in addition the requirements of $p_T^{\ell\ell} > 30$ GeV, $m_{\ell\ell} < 50$ GeV and $\Delta\phi_{\ell\ell} < 1.8$.

The transverse mass distributions of events selected in the WZ control region are similar but not identical to the ones obtained for simulated WW events ending up in the $\mathrm{H}+0$ jets signal region. One of the reasons is that in case of the transverse mass distribution of the modified WZ system, one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to mimic a system of two leptons and missing transverse energy. The transverse momentum of the lepton has to amount at least to 15 GeV which consequently leads to higher missing transverse energy values for the modified WZ system compared to the ones observed for simulated WW events. Thus the transverse mass distributions of the modified WZ system are shifted to slightly higher values compared to the distributions of simulated WW events ending up in the $\mathrm{H}+0$ jets signal region. Nevertheless, the selected WZ and ZZ events exhibit kinematic properties that are very similar to those of the WW events. To estimate the size of the effect, the observed shift in the transverse mass distribution of the modified WZ system between data and Monte Carlo simulation is directly transfered to the $\mathrm{H}+0$ jets signal region.

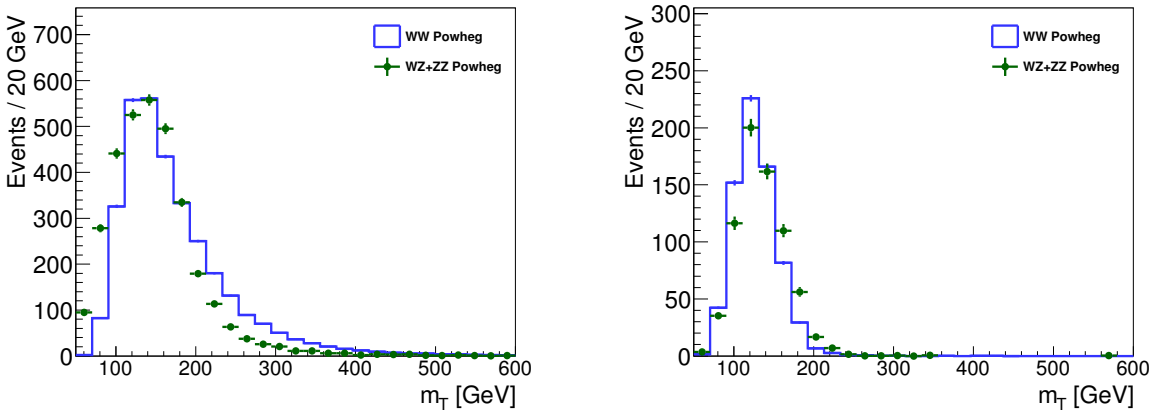


Figure 7.7.: Distributions of the transverse mass of the modified WZ system (where one of the leptons attributed to the Z boson candidate is assigned to $E_{\mathrm{T}}^{\mathrm{miss}}$ in order to mimic a system of two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$) predicted from WZ and ZZ processes generated with `POWHEG` compared to the transverse mass (calculated as for the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis) predicted by WW processes generated with `POWHEG`, for events that pass all selection criteria and have zero reconstructed jets (left) and for events that satisfy the requirements of $p_T^{\ell\ell} > 30$ GeV, $m_{\ell\ell} < 50$ GeV and $\Delta\phi_{\ell\ell} < 1.8$ in addition (right).

The slope parameter that was extracted by a linear fit to the ratio of the transverse mass distribution in data over the one predicted by simulated WZ and ZZ events in the WZ control region including a requirement of zero reconstructed jets amounts to -0.0024 $\pm$ 0.0009 (see Fig. 7.6). The parameters obtained by the fit are included in a function of the form

$$m_T(reweighed) = 1.483 - 0.0024 \times (m_T(original)/1000.) \quad . \tag{7.1}$$

The function is applied to the transverse mass distribution of the WW background processes obtained by using POWHEG as a generator. The resulting reweighed transverse mass distribution is normalised to the initial number of simulated WW events to only exhibit the shape modification. The distributions of the transverse mass in the H + 0 jets signal region for simulated WW events generated with POWHEG before and after the reweighing are compared to the transverse mass distribution predicted by WW events generated with MC@NLO in Fig. 7.8. The difference between the m_{T} shapes predicted by the Powheg and MC@NLO generators for the WW background were taken into account as a shape uncertainty for the 2012 $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ analysis (see Section 6.15). Thus, a comparison to this discrepancy allows for an estimate of the size of the effect caused by the reweighing. The m_{T} distribution redicted by WW events generated with Powheg was taken as default distribution. An extra systematic uncertainty in the likelihood function allows the shape of the WW background component to morph between the shape obtained with POWHEG and MC@NLO (see Section 6.12.3).

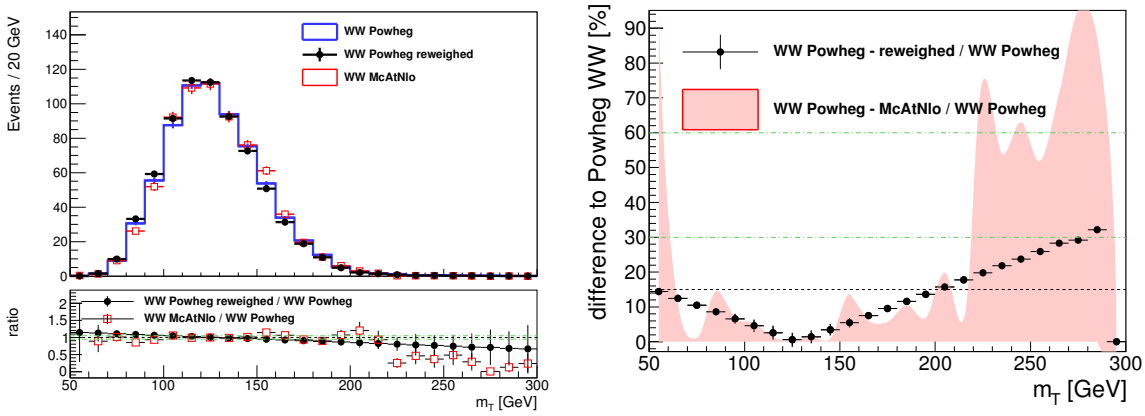


Figure 7.8.: Distributions of the transverse mass in the H + 0 jets signal region for events predicted by WW processes generated with POWHEG before and after a reweighing is applied compared to the transverse mass distribution predicted by WW events generated with MC@NLO (left). The difference of the reweighed distribution compared to the nominal distribution obtained for WW events generated with POWHEG and the difference of the distribution predicted by WW events generated with MC@NLO compared to the one obtained with POWHEG in percent (right).

To get the order of magnitude of the change in shape induced by the reweighing, the difference of the reweighed distribution compared to the nominal distribution obtained for WW events generated with POWHEG are compared to the difference in shaped observed by WW events generated with MC@NLO compared to the one obtained with POWHEG (see Fig. 7.8). The size of the shift in the transverse mass distribution extracted in the WZ control region between data and Monte Carlo simulation is of the order between 10% for low m_{T} values and

30% for high m_{T} values. This discrepancy is well covered by the difference found between the two generators which is already taken into account as a systematic uncertainty in the analysis. Thus, there is no evidence for the neccesity of an additional systematic uncertainty to account for the mismodelling displayed by the reweighed distribution.

7.4. Impact on the Significance in the H + 0 jets Signal Region

The transverse mass distribution of the WW background Monte Carlo prediction is modified according to the function given in Eq. 7.1. In order to get an estimate of the impact of the mismodelling on the expected and observed significance in the H + 0 jets analysis, the modified WW background distribution was put into the analysis as described in Sections 6.8 and 6.18. The resulting observed and expected 95% CL upper limits on the Higgs boson production cross section, normalised to the SM cross section, as a function of $\mathrm{m_H}$ are shown in Fig. 7.9. The local p_0-value is shown in Fig. 7.10. Both results obtained including the original WW Monte Carlo transverse mass distribution and the modified one are displayed. Only a minor change is observed in the curves of the observed and expected 95% CL upper limits on the Higgs boson production cross section. The effect on the local p_0-value is larger since the modified WW Monte Carlo transverse mass distribution is getting more similar to the expected $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ signal distribution. Therefore, an observed excess of events in data would be interpreted as less significant.

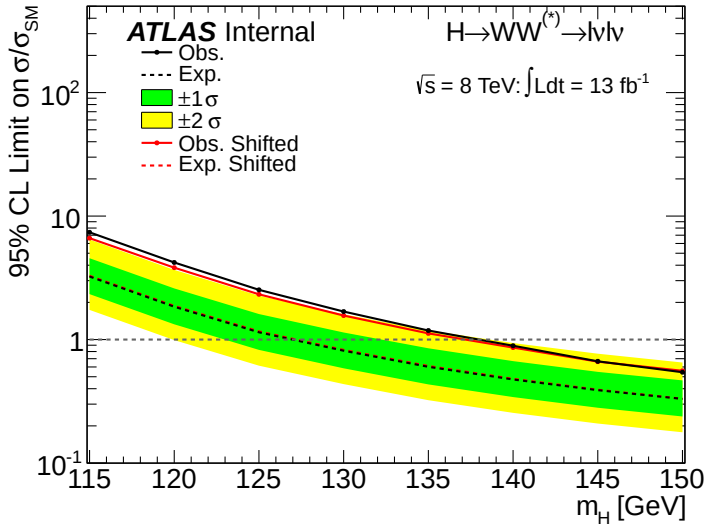


Figure 7.9.: Observed (solid) and expected (dashed) 95% CL upper limits on the Higgs boson production cross section, normalised to the SM cross section, as a function of $\mathrm{m_H}$ restricted to the range $\mathrm{m_H} < 150$ GeV. The inner (green) and outer (yellow) regions indicate the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands on the expected limit, respectively, [201].

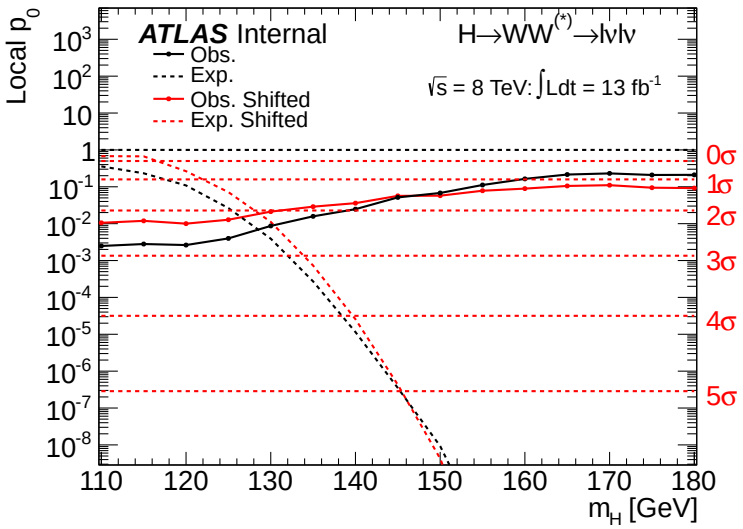


Figure 7.10.: The observed (solid line) probability p_0 for the background-only scenario as a function of m_H for the 8 TeV data. The dashed line shows the corresponding expectation for the signal and background hypothesis at the given value of m_H [201].

Instead of modifying the overall WW Monte Carlo transverse mass distribution, it is more desirable to correct for the underlying source of the difference. More detailed studies are needed on the full dataset taken in 2012 that corresponds to an integrated luminosity of approximately 22 fb^{-1}. The observed effects can be examined more closely since the statistical uncertainties will decrease. A repetition of the present study has the potential to significantly constrain the WW Monte Carlo transverse mass distribution from data. Thus, further studies on a larger data sample will allow to reduce the systematic uncertainty on the m_T shape.

8

Summary

The Standard Model of particle physics is exceptionally successful in describing the known experimental facts in particle physics with high precision. However, the model requires some mechanism that breaks the electroweak symmetry in order to incorporate massive particles into the theory. The experimental verification of such a mechanism is the last missing piece to complete the Standard Model. The Higgs mechanism is the most popular theoretical proposition to spontaneously break the electroweak symmetry. It gives rise to the appearance of a physical scalar particle, the Higgs boson. The hunt for this boson has been a major motivation for the construction of the Large Hadron Collider (LHC) and for the design of the ATLAS and CMS experiments.

The mass of the Higgs boson is not predicted by the model but once a mass value is specified, its production cross section and decay modes are predicted by the theory. The leading production mechanism for a Standard Model Higgs boson at the LHC is the gluon fusion. The Higgs boson couplings to particles increase with their masses. For a Higgs boson mass heavier than 135 GeV, the decay into a pair of W bosons, $H \to WW^*$, becomes the dominant decay channel.

The Standard Model Higgs boson has been searched for in various production and decay channels with the ATLAS experiment since the LHC started to operate at a centre-of-mass energy of 7 TeV in the year 2010. In this thesis, the search for the Standard Model Higgs boson in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay channel, with ℓ denoting electrons or muons, was performed based on proton-proton collision data taken in the years 2011 and 2012 [3–5]. An excess of events relative to the background expectation was observed in data. In combination with results obtained in other individual searches, in particular in the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels, clear evidence for the production of a neutral boson with a measured mass of 126.0 ± 0.4 (stat) $\pm$ 0.4 (sys) GeV was found [6]. The observation has a significance of 5.9 standard deviations and the observed results are compatible with the expected production and decay properties of the Standard Model Higgs boson. The CMS collaboration also announced the observation of a new boson with a mass of 125.3 ± 0.4 (stat) $\pm$ 0.5 (sys) GeV as reported in Ref. [202]. The observed excess of events above the expected background has a local significance of 5.0 standard deviations at a mass near 125 GeV.

This thesis constitutes a major contribution to the search for the Standard Model Higgs boson in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay channel with the ATLAS experiment. The analysis is based on data corresponding to an integrated luminosity of 4.7 fb^{-1} collected at $\sqrt{s} = 7$ TeV

in the year 2011 and 5.8 fb^{-1} collected at $\sqrt{s} = 8$ TeV in the year 2012 in proton-proton collisions by the ATLAS experiment at the LHC. Due to the presence of two neutrinos in the final state, it is not possible to reconstruct an invariant mass involving all decay products. Thus, a transverse mass, m_{T}, computed from the leptons and the missing transverse momentum was used to test for the presence of a signal. The background rate and composition as well as the signal topology depend significantly on the number of jets in the final state. In order to maximise the sensitivity of the search, the analysis is subdivided into the H + 0 jets, H + 1 jet and H+ ≥ 2 jets channels according to the jet multiplicity.

The studies performed in the course of this thesis focus on the H + 0 jets and H + 1 jet channels that provide the best sensitivity. The work consisted in the multi-stage selection of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ final states and the construction of control regions to validate the Monte Carlo simulation of the main background components in general and of the Standard Model W^+W^- process, the dominant background, in particular. Many cross-checks were performed to ensure the robustness of the analysis and to strengthen the confidence in the results. The combined results obtained from the analyses of the datasets taken in the years 2011 and 2012 exhibit an excess of events in data over the background expected from Standard Model processes. The observed excess has a significance of 2.8 standard deviations, corresponding to a background fluctuation probability of $3 \cdot 10^{-3}$ for a Higgs boson mass of $\mathrm{m_H}$ =125 GeV.

Since the Standard Model diboson W^+W^- production is the largest background, it constitutes a dominant source of systematic uncertainties to the discovery as well as the study of the properties of the Higgs boson. Various checks were performed in the signal-free control region enriched in dileptonic W^+W^- events to study the agreement of Monte Carlo predictions to data. However, this control region is not very pure and contains contributions from other background sources like top-quark pair and W + jets production.

An independent and complementary study to constrain the W^+W^- background from data was performed in the course of this thesis. It focuses on the selection of $W^{\pm}Z \to \ell\nu\ell\ell$ events. The dileptonic $W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ and the trileptonic $W^{\pm}Z \to \ell\nu\ell\ell$ decays feature kinematic similarities. The selected $W^{\pm}Z$ control sample has the advantage to be very pure with only little background contributions from Z/γ^* + jets events. The analysis has established a new method to test for differences in the modelling of kinematic distributions by comparing diboson Monte Carlo predictions directly to data. Based on the 10.7 fb^{-1} of data collected by the ATLAS experiment until September 2012, the method uncovered a noticeable shift in the expected diboson kinematic distributions compared to the observed ones. There is some evidence that a mismodelling of the transverse momentum of the W boson is the possible source of these discrepancies. Nonetheless, the deviations observed are well covered by the systematic uncertainties assumed in the present $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ search analysis. The presented method, although currently limited by the number of events observed in data, has the potential to significantly constrain the W^+W^- background kinematic distributions. Thus, further studies on a larger data sample will allow to reduce the systematic uncertainty on the m_{T} shape which is important for precise future measurements of the properties of the Higgs boson.

In view of the discovery of a Higgs-like neutral boson, the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay channel remains very interesting for further studies investigating the properties of this new particle. The determination of spin and CP quantum numbers is essential to establish that the observed excess is indeed caused by a Standard Model Higgs boson. The decay $H \to W^+W^-$ is in particular important to directly determine the coupling strength of the Higgs boson to W bosons.

This thesis also contributed to the understanding of the detector performance and the recon-

struction of physics objects which are very important ingredients to physics analyses like the ones mentioned above. The $Z \to e^+e^-$ resonance peak was used to extract a set of electron shower shape distributions. The study is based on the very first data collected at a centre-of-mass energy of $\sqrt{s} = 7$ TeV in the year 2010 by the ATLAS experiment, corresponding to an integrated luminosity of 35 pb^{-1}. The residual background to $Z \to e^+e^-$ events affects the derived electron shower shape distributions and is therefore subtracted on a statistical basis. The distributions obtained provide important information for the understanding of the electron identification in the ATLAS experiment and are published in Ref. [7]. This study was the first of its kind to be performed on data and led to significant improvements of the calorimeter simulation and the electron identification.

A

Electron Shower Shapes in Bins of E_T

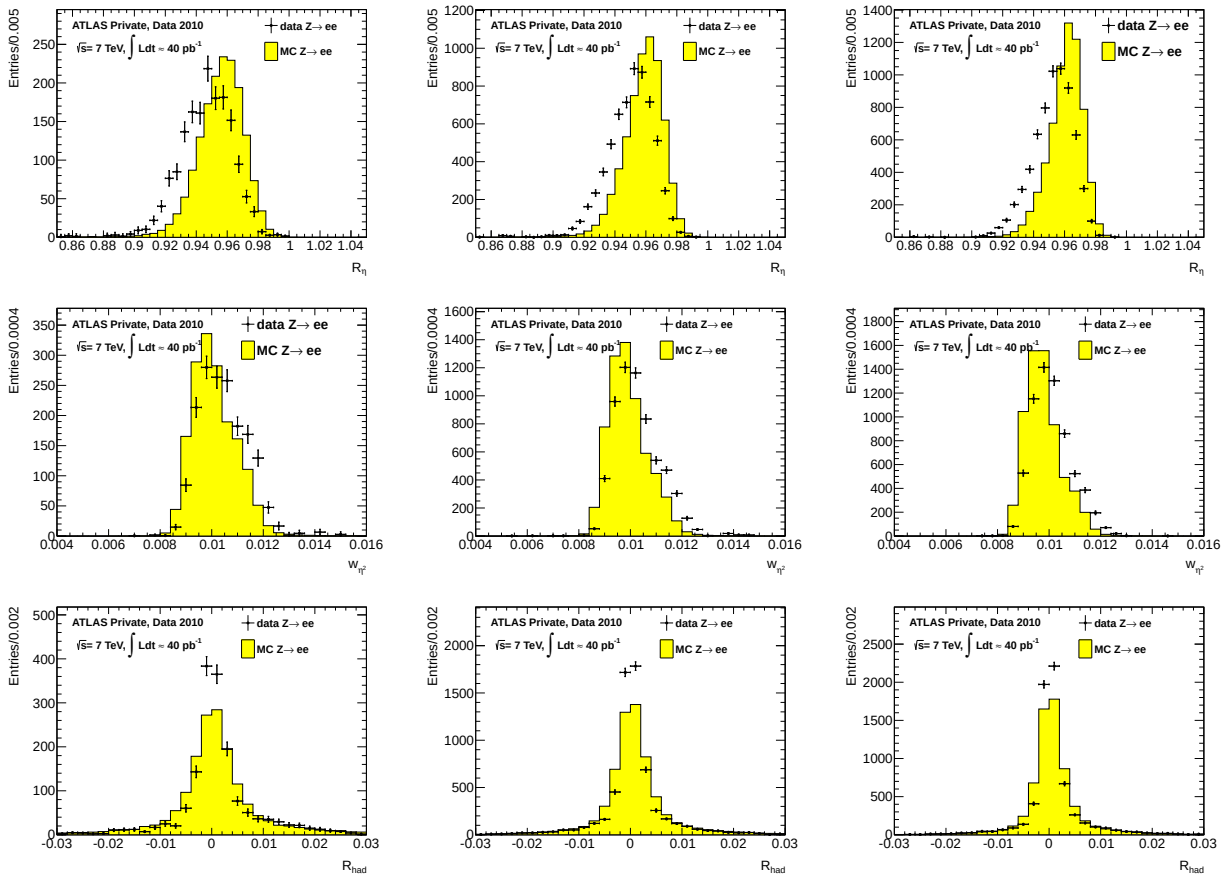


Figure A.1.: Electron shower shapes from $Z \to e^+e^-$ events in the second layer of the calorimeter for probe electrons in three E_{T} intervals ([25-30] GeV, [30-40] GeV, [40-50] GeV from left to right). R_η is shown on the top, w_{η^2} in the middle and the hadronic leakage R_{had} on the bottom row. The data points are plotted as full circles with error bars, representing the total statistical and systematic uncertainties. The MC prediction, normalised to the number of data entries, is shown as a filled yellow histogram.

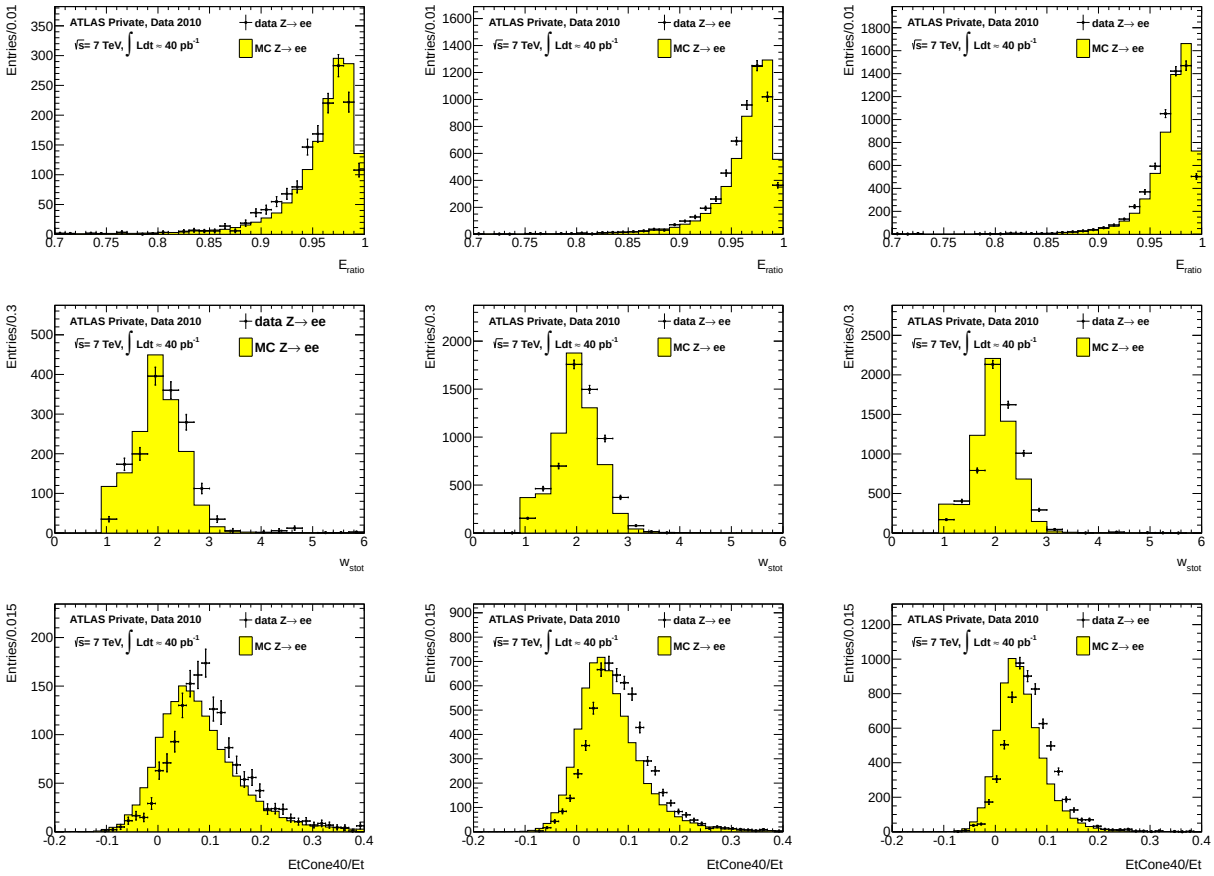


Figure A.2.: Electron shower shapes from $Z \rightarrow e^+e^-$ events in the first layer of the calorimeter and isolation variables for probe electrons in three E_{T} intervals ([25-30] GeV, [30-40] GeV, [40-50] GeV from left to right). E_{ratio} is shown on the top, w_{stot} in the middle and EtCone40/E_T on the bottom row. The data points are plotted as full circles with error bars, representing the total statistical and systematic uncertainties. The MC prediction, normalised to the number of data entries, is shown as a filled yellow histogram.

B

Closure Test for the sPlot Technique

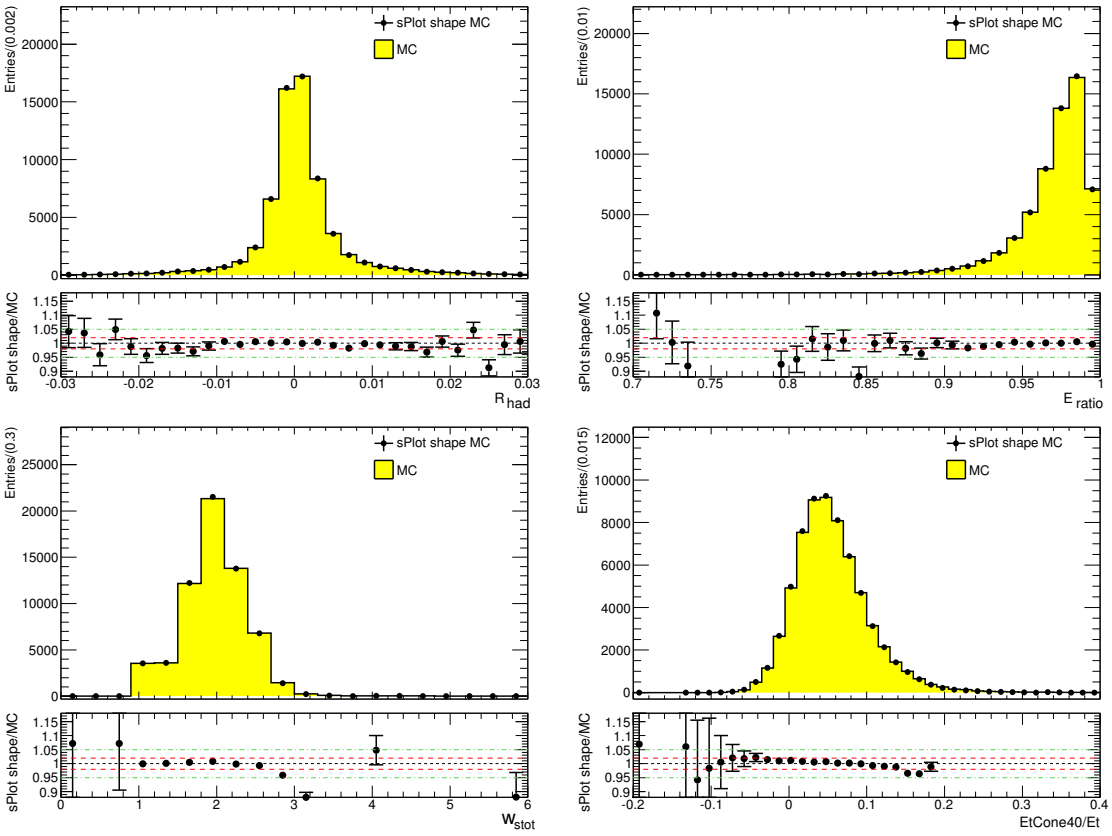


Figure B.1.: Electron shower shapes from $Z \to e^+e^-$ events for probe electrons with E_{T} in the interval [40,50] GeV: (top left) R_{had} hadronic leakage, (top right) E_{ratio} strip-layer variable, (bottom left) w_{stot} total shower width in strip layer, (bottom right) EtCone40/E_T isolation variable. The extracted Monte Carlo shower shapes are plotted as full circles with error bars, representing the total statistical uncertainties. The pure Monte Carlo shower shape predictions, normalised to the number of entries, are shown by filled histograms.

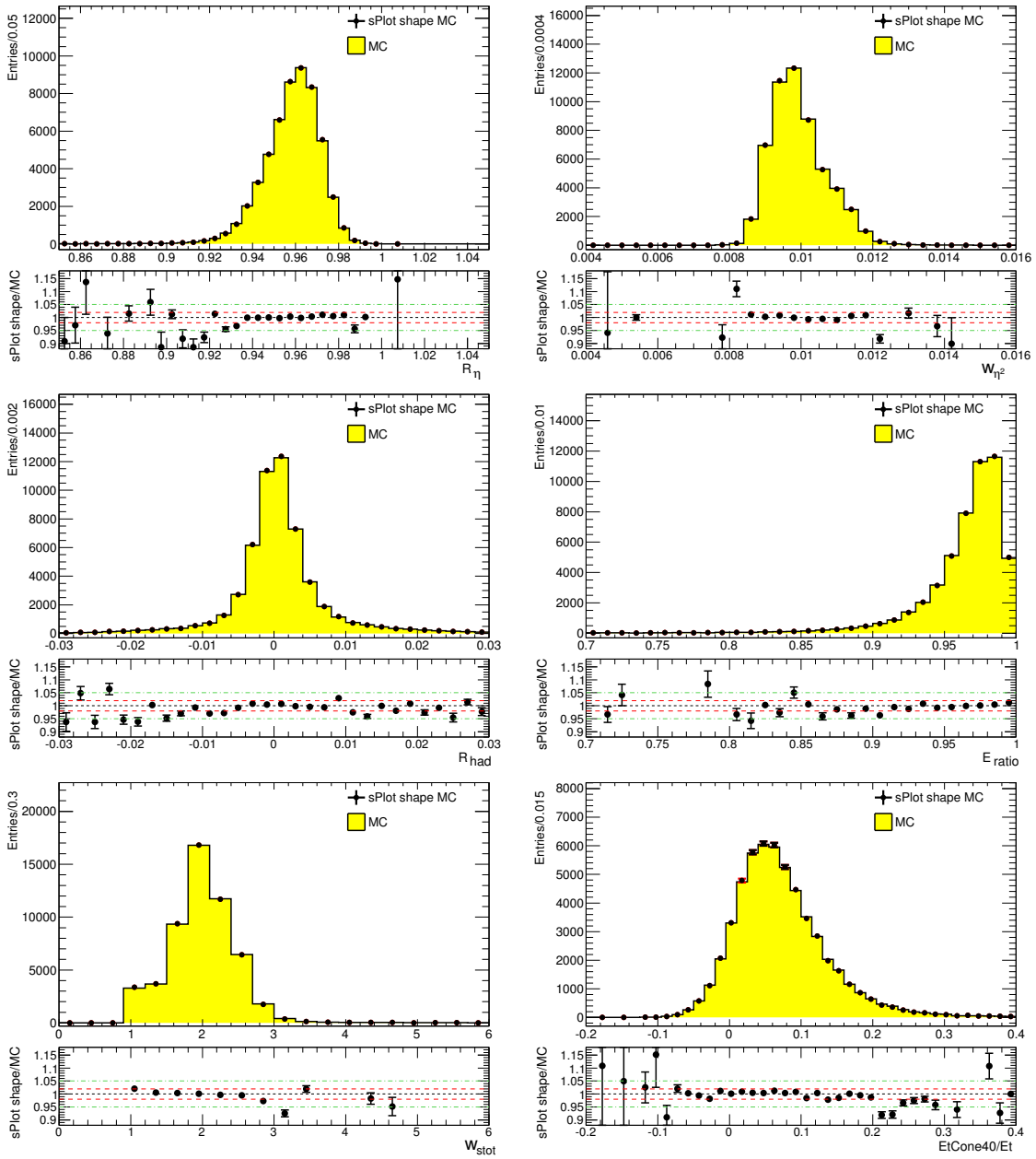


Figure B.2.: Electron shower shapes from $Z \to e^+e^-$ events for probe electrons with E_{T} in the interval [30,40] GeV: (top left) R_η and (top right) w_{η^2} middle-layer variables, (middle left) R_{had} hadronic leakage, (middle right) E_{ratio} strip-layer variable, (bottom left) w_{stot} total shower width in strip layer, (bottom right) EtCone40/E_T isolation variable. The extracted Monte Carlo shower shapes are plotted as full circles with error bars, representing the total statistical uncertainties. The pure Monte Carlo shower shape predictions, normalised to the number of entries, are shown by filled histograms.

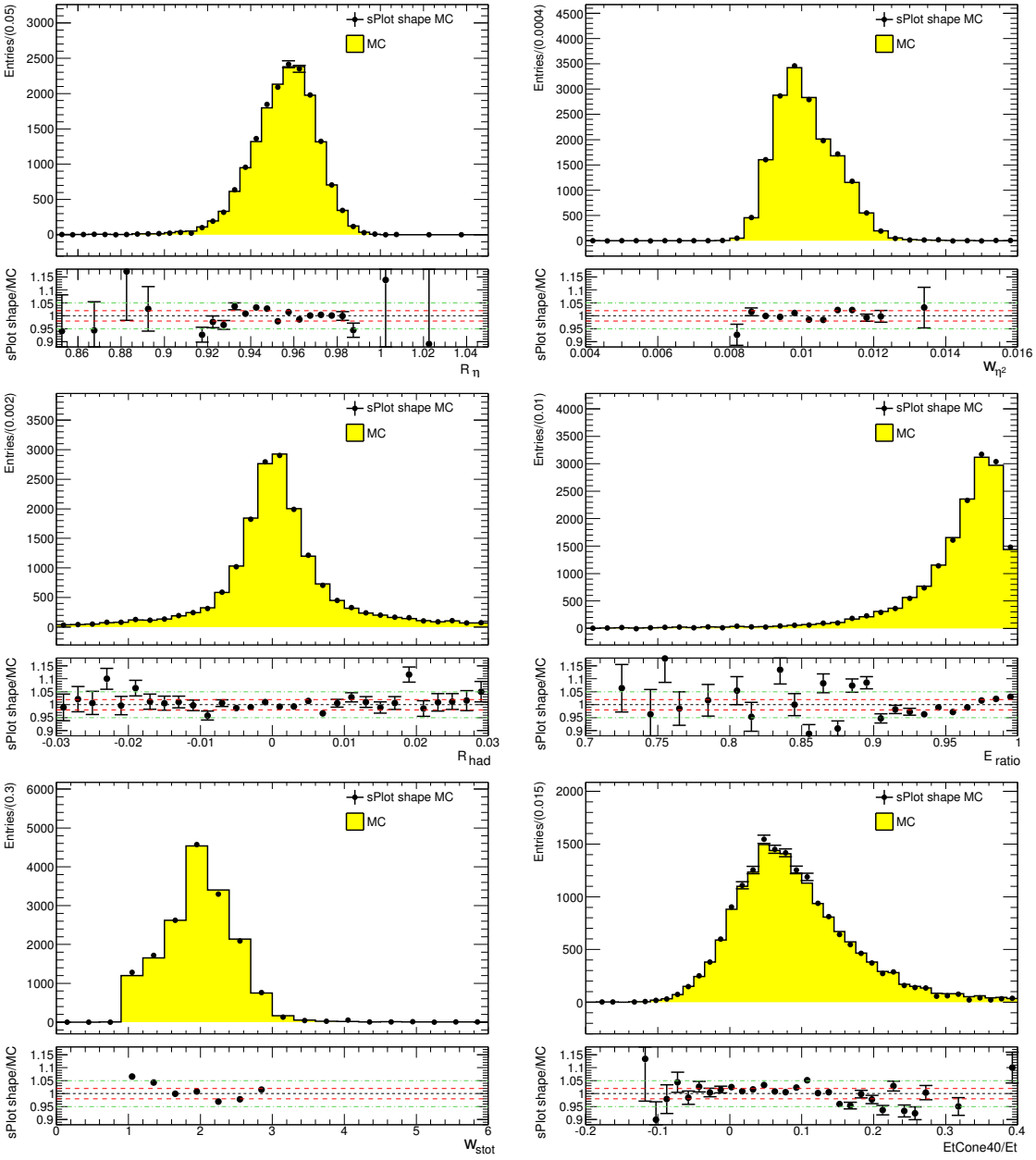


Figure B.3.: Electron shower shapes from $Z \to e^+e^-$ events for probe electrons with E_{T} in the interval [25,30] GeV: (top left) R_η and (top right) w_{η^2} middle-layer variables, (middle left) R_{had} hadronic leakage, (middle right) E_{ratio} strip-layer variable, (bottom left) w_{stot} total shower width in strip layer, (bottom right) EtCone40/E_T isolation variable. The extracted Monte Carlo shower shapes are plotted as full circles with error bars, representing the total statistical uncertainties. The pure Monte Carlo shower shape predictions, normalised to the number of entries, are shown by filled histograms.

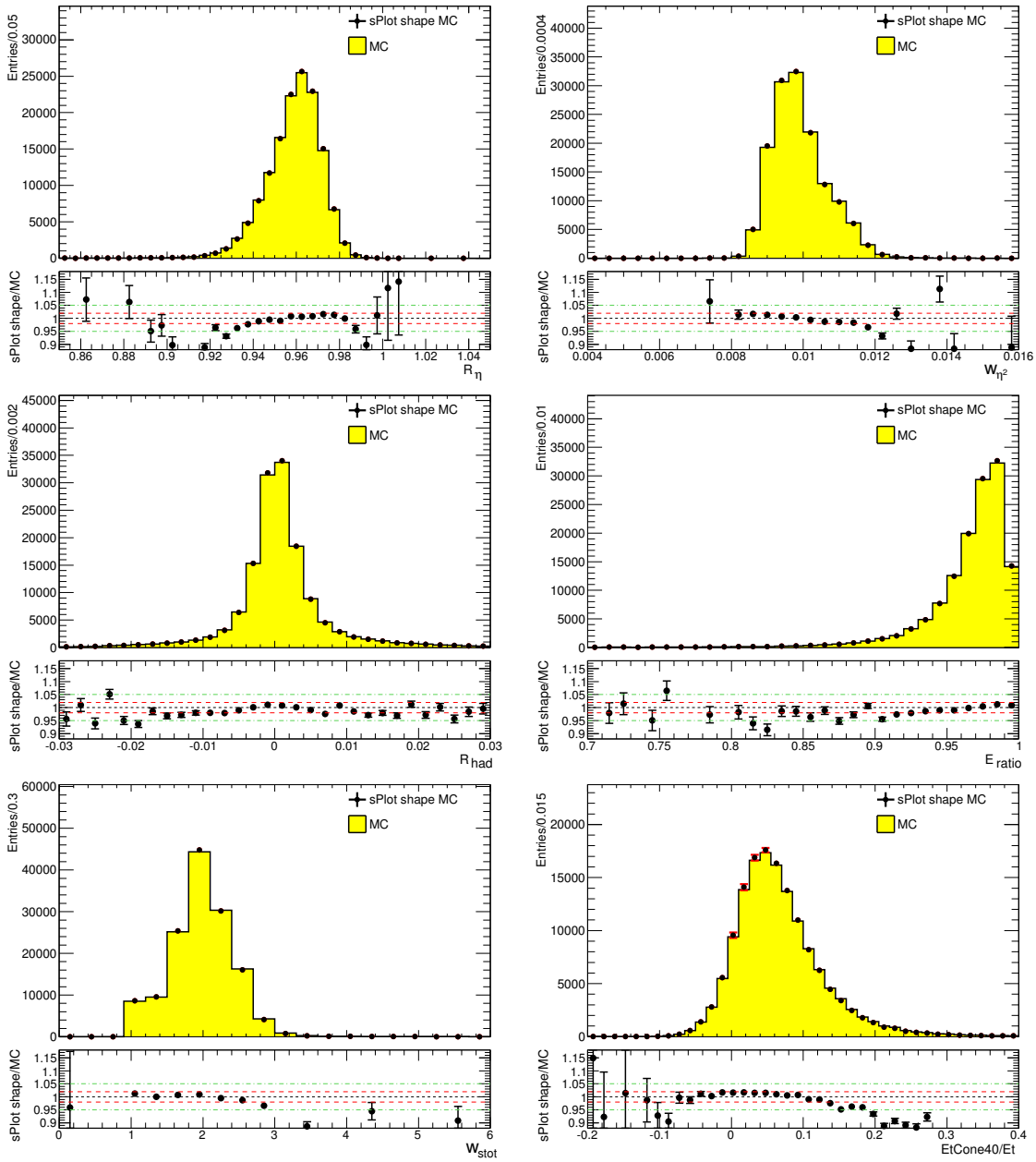


Figure B.4.: Electron shower shapes from $Z \to e^+e^-$ events for probe electrons with E_{T} in the interval [25,50] GeV: (top left) R_η and (top right) w_{η^2} middle-layer variables, (middle left) R_{had} hadronic leakage, (middle right) E_{ratio} strip-layer variable, (bottom left) w_{stot} total shower width in strip layer, (bottom right) EtCone40/E_T isolation variable. The extracted Monte Carlo shower shapes are plotted as full circles with error bars, representing the total statistical uncertainties. The pure Monte Carlo shower shape predictions, normalised to the number of entries, are shown by filled histograms.

C

Systematical Uncertainties on Electron Shower Shapes

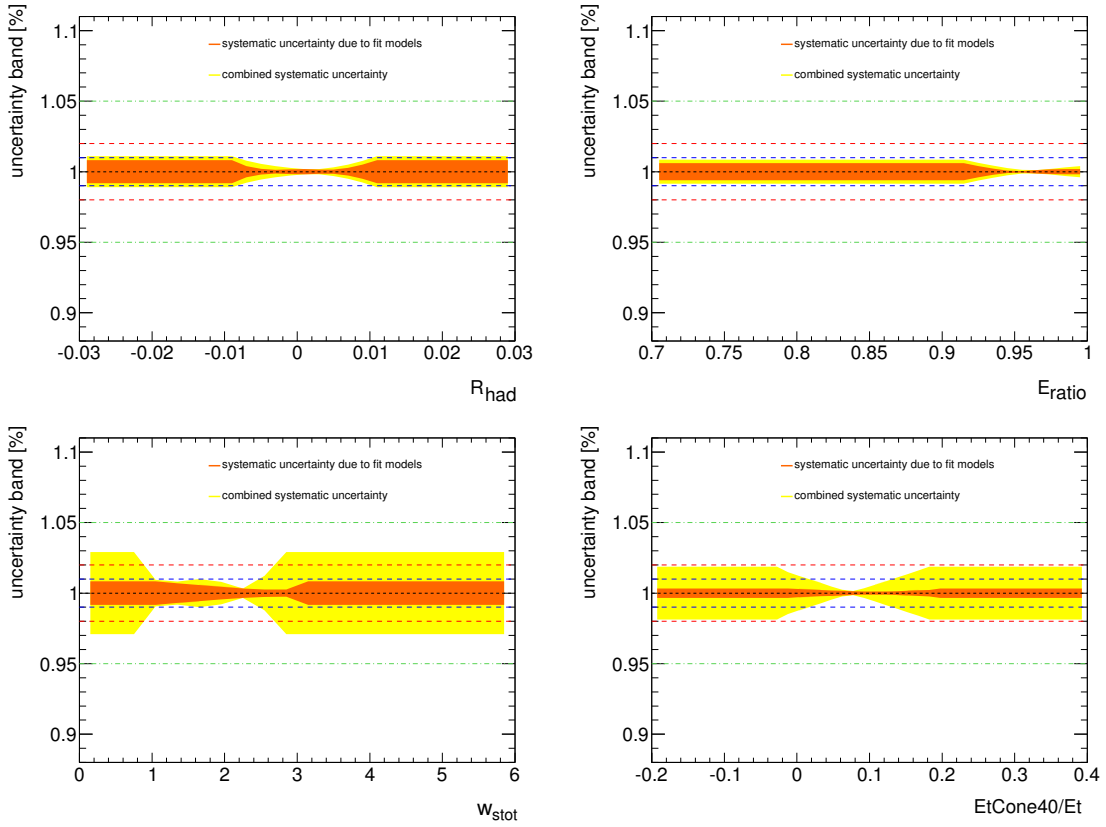


Figure C.1.: Combined systematic uncertainties for shower shapes from $Z \to e^+e^-$ events for probe electrons with E_{T} in the interval [40,50] GeV: (top left) hadronic leakage R_{had}, (top right) strip-layer variable E_{ratio}, (bottom left) total shower width in strip layer w_{stot}, (bottom right) isolation variable EtCone40/E_T. The uncertainties due to the choice of the fit models (in orange) and the uncertainties due to the background subtraction with the sPlot technique estimated on Monte Carlo are added in quadrature (yellow band).

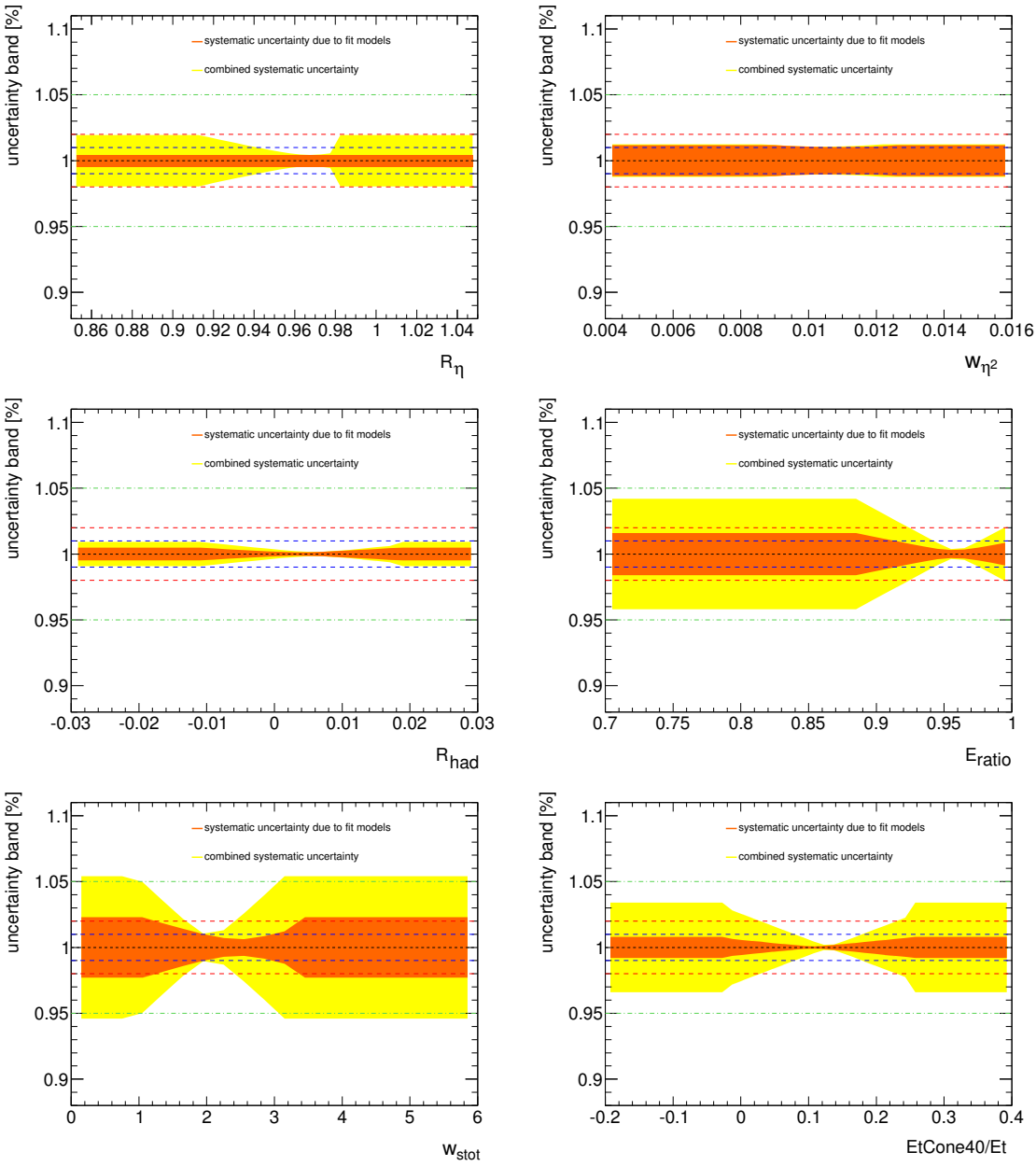


Figure C.2.: Combined systematic uncertainties for shower shapes from $Z \to e^+e^-$ events for probe electrons with E_{T} in the interval [25,30] GeV: (top left) $R_\eta$and (top right) w_{η^2} middle-layer variables, (middle left) hadronic leakage R_{had}, (middle right) strip-layer variable E_{ratio}, (bottom left) total shower width in strip layer w_{stot}, (bottom right) isolation variable EtCone40/E_T. The uncertainties due to the choice of the fit models (in orange) and the uncertainties due to the background subtraction with the sPlot technique estimated on Monte Carlo are added in quadrature (yellow band).

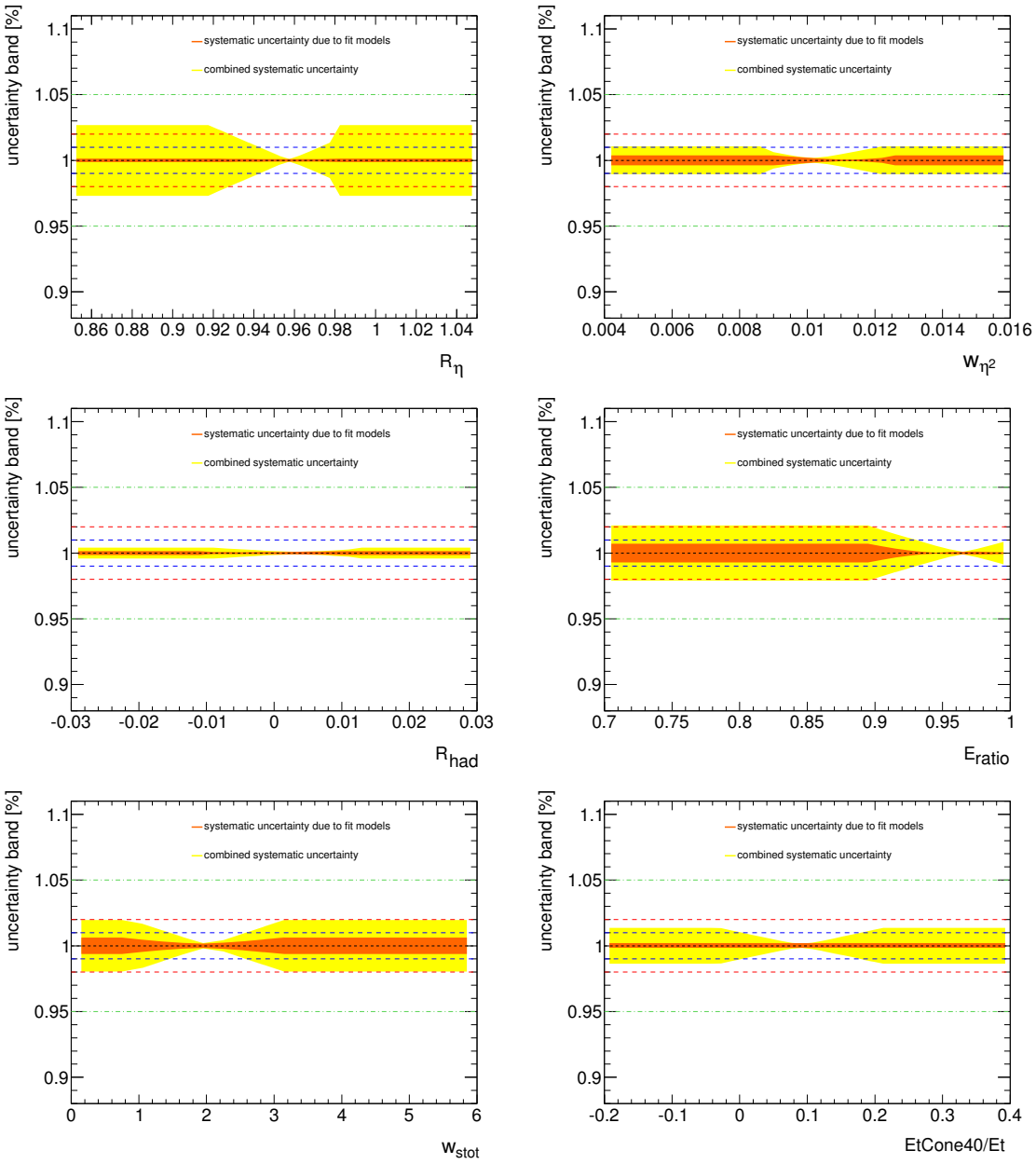


Figure C.3.: Combined systematic uncertainties for shower shapes from $Z \to e^+e^-$ events for probe electrons with E_{T} in the interval [30,40] GeV: (top left) $R_\eta$and (top right) w_{η^2}middle-layer variables, (middle left) hadronic leakage R_{had}, (middle right) strip-layer variable E_{ratio}, (bottom left) total shower width in strip layer w_{stot}, (bottom right) isolation variable EtCone40/E_T. The uncertainties due to the choice of the fit models (in orange) and the uncertainties due to the background subtraction with the sPlot technique estimated on Monte Carlo are added in quadrature (yellow band).

D

Cutflows Split Up into Lepton Channels

Table D.1.: The expected numbers of signal (for $m_H = 125$ GeV) and background events as well as the events observed in data at several stages of the selection for the ee final state. The $W+$jets background is estimated from data, whereas MC predictions normalised to data in control regions are used for WW, Z/γ^*+jets, $t\bar{t}$, and $tW/tb/tqb$ processes. Contributions from other background sources are taken directly from MC predictions. Only statistical uncertainties associated with the number of events in the MC samples and the data control regions are shown. Starting from jet multiplicity specific selection stages, the numbers for the SM WW and top background contributions are corrected using the normalisation factors derived from the corresponding control samples. In the same flavor channels, the Z/γ^*+jets contribution is corrected using the ABCD method.

	Signal [125 GeV]	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	Z/γ^*+jets	$W+$jets	Total Bkg.	Observed	Data/MC
lepton p_{T}	32.42 ± 0.17	726.89 ± 4.18	1166.29 ± 13.15	3191.76 ± 12.36	343.93 ± 5.08	1033774.58 ± 801.13	714.96 ± 12.76	1039918.42 ± 801.46	1034217	0.99 ± 0.00
Opposite Charge	31.88 ± 0.17	721.03 ± 4.16	1015.34 ± 9.94	3162.49 ± 12.31	340.01 ± 5.06	1026746.21 ± 798.53	462.44 ± 12.24	1032447.52 ± 798.81	1028357	1.00 ± 0.00
$m_{\ell\ell} > 10$ GeV	31.14 ± 0.17	719.65 ± 4.16	1008.52 ± 9.87	3156.69 ± 12.29	339.39 ± 5.05	1025871.63 ± 798.36	433.55 ± 12.21	1031529.43 ± 798.64	1025944	0.99 ± 0.00
Z veto (for $ee, \mu\mu$)	30.39 ± 0.16	561.55 ± 3.66	163.52 ± 4.77	2478.50 ± 10.91	267.26 ± 4.49	94754.24 ± 238.17	325.01 ± 5.63	98550.08 ± 238.60	97593	0.99 ± 0.00
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45$ GeV	10.24 ± 0.10	208.11 ± 13.14	26.70 ± 1.88	929.80 ± 63.55	107.98 ± 7.89	289.44 ± 14.91	56.34 ± 1.10	1618.38 ± 91.86	1652	1.02 ± 0.06
Z control region (incl)	0.75 ± 0.03	158.10 ± 1.98	845.00 ± 8.64	678.19 ± 5.68	72.13 ± 2.32	931117.40 ± 762.01	108.54 ± 10.83	932979.35 ± 762.16	928351	1.00 ± 0.00
Top control region (incl)	0.45 ± 0.02	5.47 ± 0.38	0.86 ± 0.21	755.64 ± 6.02	66.56 ± 2.24	6.16 ± 1.78	2.30 ± 0.43	836.99 ± 6.70	910	1.09 ± 0.04
0j: jet veto	6.26 ± 0.08	142.67 ± 9.07	13.39 ± 1.43	25.09 ± 2.04	14.37 ± 1.44	218.00 ± 12.91	39.39 ± 0.87	452.91 ± 21.01	414	0.91 ± 0.06
0j: $m_{\ell\ell} < 50$ GeV	5.30 ± 0.07	43.69 ± 2.91	6.01 ± 1.05	3.98 ± 0.59	3.71 ± 0.59	24.41 ± 3.83	17.05 ± 0.60	98.84 ± 6.03	87	0.88 ± 0.11
0j: $p_{\mathrm{T},\ell\ell} > 45$ GeV	4.01 ± 0.06	37.30 ± 2.52	4.08 ± 0.80	3.63 ± 0.56	3.29 ± 0.55	3.73 ± 1.90	7.75 ± 0.38	59.77 ± 4.64	52	0.87 ± 0.14
0j: $\Delta\phi_{\ell\ell} < 1.8$ rad	4.00 ± 0.06	37.24 ± 2.51	4.08 ± 0.80	3.63 ± 0.56	3.29 ± 0.55	3.75 ± 1.91	7.70 ± 0.38	59.69 ± 4.65	52	0.87 ± 0.14
0j: $0.75 \cdot m_{\mathrm{H}} \leq m_{\mathrm{T}} \leq m_{\mathrm{H}}$	2.93 ± 0.05	13.33 ± 1.02	1.65 ± 0.45	0.64 ± 0.21	0.71 ± 0.23	1.92 ± 1.08	4.80 ± 0.31	23.05 ± 2.02	19	0.82 ± 0.20
0j: Z control region	0.37 ± 0.02	109.34 ± 1.67	189.43 ± 2.98	13.97 ± 0.78	7.59 ± 0.74	739205.38 ± 690.45	9.45 ± 9.52	739535.15 ± 690.53	736673	1.00 ± 0.00
0j: WW control region	0.02 ± 0.01	42.74 ± 1.00	2.93 ± 0.28	12.96 ± 0.75	6.14 ± 0.68	1.70 ± 1.20	4.43 ± 0.22	70.90 ± 1.89	84	1.18 ± 0.13
1j: exactly one jet	2.58 ± 0.05	40.10 ± 6.33	9.21 ± 1.09	202.89 ± 8.94	56.23 ± 3.15	56.07 ± 6.75	12.31 ± 0.51	376.81 ± 21.76	367	0.97 ± 0.08
1j: b-jet veto	2.38 ± 0.04	37.39 ± 5.90	8.89 ± 1.08	57.88 ± 2.94	17.00 ± 1.37	50.84 ± 6.33	11.31 ± 0.46	183.30 ± 13.62	200	1.09 ± 0.11
1j: $p_{\mathrm{T}}^{\mathrm{tot}} < 30$ GeV	1.61 ± 0.04	26.82 ± 4.25	6.57 ± 1.00	25.50 ± 1.54	9.28 ± 0.96	11.21 ± 2.95	5.16 ± 0.32	84.55 ± 7.95	94	1.11 ± 0.16
1j: $Z \to \tau\tau$ veto	1.60 ± 0.04	26.33 ± 4.18	6.51 ± 1.00	24.88 ± 1.51	8.86 ± 0.93	10.59 ± 2.86	5.14 ± 0.32	82.30 ± 7.73	93	1.13 ± 0.16
1j: $m_{\ell\ell} < 50$ GeV	1.33 ± 0.03	6.99 ± 1.16	2.23 ± 0.64	4.55 ± 0.57	2.61 ± 0.48	2.12 ± 0.85	2.03 ± 0.21	20.52 ± 2.18	22	1.07 ± 0.26
1j: $\Delta\phi_{\ell\ell} < 1.8$ rad	1.22 ± 0.03	6.56 ± 1.09	2.17 ± 0.63	4.32 ± 0.56	2.53 ± 0.47	1.93 ± 0.82	1.64 ± 0.19	19.15 ± 2.08	19	0.99 ± 0.25
1j: $0.75 \cdot m_{\mathrm{H}} \leq m_{\mathrm{T}} \leq m_{\mathrm{H}}$	0.87 ± 0.03	2.30 ± 0.42	1.13 ± 0.52	1.11 ± 0.27	0.71 ± 0.24	1.24 ± 0.67	1.18 ± 0.16	7.67 ± 1.14	7	0.91 ± 0.37
1j: Z control region	0.23 ± 0.02	35.64 ± 0.91	287.80 ± 5.16	110.27 ± 2.28	34.49 ± 1.61	144414.33 ± 293.55	73.75 ± 4.51	144956.29 ± 293.64	144564	1.00 ± 0.00
1j: WW control region	0.01 ± 0.00	13.35 ± 0.55	2.21 ± 0.45	11.87 ± 0.71	3.03 ± 0.49	1.74 ± 1.02	1.23 ± 0.13	33.43 ± 1.51	42	1.26 ± 0.20
1j: Top control region	0.13 ± 0.01	2.19 ± 0.24	0.16 ± 0.10	60.78 ± 1.73	22.19 ± 1.29	0.76 ± 0.76	0.64 ± 0.15	86.73 ± 2.31	92	1.06 ± 0.11
$\geq$ 2j: at least 2 jets	1.40 ± 0.03	17.11 ± 0.61	4.10 ± 0.55	731.07 ± 6.13	41.32 ± 1.84	15.36 ± 2.40	4.63 ± 0.46	813.59 ± 6.90	871	1.07 ± 0.04
$\geq$ 2j: central jet veto	1.07 ± 0.03	13.48 ± 0.54	3.07 ± 0.51	418.48 ± 4.58	28.75 ± 1.52	9.12 ± 1.97	3.43 ± 0.35	476.33 ± 5.28	509	1.07 ± 0.05
$\geq$ 2j: b-jet veto	0.90 ± 0.03	11.84 ± 0.50	2.74 ± 0.49	44.37 ± 1.52	6.32 ± 0.73	8.36 ± 1.84	2.68 ± 0.23	76.30 ± 2.60	96	1.26 ± 0.14
$\geq$ 2j: opp. hemispheres	0.49 ± 0.02	5.53 ± 0.34	0.96 ± 0.19	19.45 ± 0.99	3.21 ± 0.51	3.75 ± 1.22	1.06 ± 0.15	33.95 ± 1.71	37	1.09 ± 0.19
$\geq$ 2j: $\Delta\eta_{\mathrm{jj}} > 3.8$	0.22 ± 0.01	1.15 ± 0.15	0.26 ± 0.10	2.88 ± 0.35	0.26 ± 0.13	1.21 ± 0.85	0.17 ± 0.05	5.92 ± 0.95	8	1.35 ± 0.52
$\geq$ 2j: $m_{\mathrm{jj}} > 500$ GeV	0.15 ± 0.01	0.45 ± 0.09	0.03 ± 0.01	1.47 ± 0.24	0.00 ± 0.00	0.00 ± 0.00	0.01 ± 0.02	1.96 ± 0.26	2	1.02 ± 0.73
$\geq$ 2j: $p_{\mathrm{T,tot}} < 30$ GeV	0.11 ± 0.00	0.29 ± 0.07	0.02 ± 0.01	0.69 ± 0.16	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.99 ± 0.17	1	1.01 ± 1.02
$\geq$ 2j: $Z \to \tau\tau$ veto	0.11 ± 0.00	0.29 ± 0.07	0.02 ± 0.01	0.69 ± 0.16	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.99 ± 0.17	1	1.01 ± 1.02
$\geq$ 2j: $m_{\ell\ell} < 80$ GeV	0.11 ± 0.00	0.14 ± 0.06	0.01 ± 0.01	0.15 ± 0.08	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.29 ± 0.09	0	0.00 ± 0.00
$\geq$ 2j: $\Delta\phi_{\ell\ell} < 1.8$ rad	0.10 ± 0.00	0.10 ± 0.05	0.01 ± 0.01	0.03 ± 0.05	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.13 ± 0.07	0	0.00 ± 0.00
$\geq$ 2j: $0.75 \cdot m_{\mathrm{H}} \leq m_{\mathrm{T}} \leq m_{\mathrm{H}}$	0.06 ± 0.00	0.02 ± 0.03	0.01 ± 0.01	0.03 ± 0.03	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.05 ± 0.05	0	0.00 ± 0.00
$\geq$ 2j: Z control region	0.15 ± 0.01	13.13 ± 0.54	367.77 ± 6.26	553.95 ± 5.14	30.05 ± 1.50	47497.68 ± 133.29	25.34 ± 2.52	48487.91 ± 133.57	47114	0.97 ± 0.01
$\geq$ 2j: WW control region	0.00 ± 0.00	0.15 ± 0.05	0.01 ± 0.01	0.54 ± 0.14	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.71 ± 0.15	1	1.41 ± 1.44
$\geq$ 2j: Top control region	0.17 ± 0.01	1.63 ± 0.22	0.32 ± 0.14	373.67 ± 4.34	22.54 ± 1.34	0.55 ± 0.55	0.75 ± 0.27	399.47 ± 4.59	413	1.03 ± 0.05

Table D.2.: The expected numbers of signal (for $m_H = 125$ GeV) and background events as well as the events observed in data at several stages of the selection for the $\mu\mu$ final state. The $W+$jets background is estimated from data, whereas MC predictions normalised to data in control regions are used for WW, Z/γ^*+jets, $t\bar{t}$, and $tW/tb/tqb$ processes. Contributions from other background sources are taken directly from MC predictions. Only statistical uncertainties associated with the number of events in the MC samples and the data control regions are shown. Starting from jet multiplicity specific selection stages, the numbers for the SM WW and top background contributions are corrected using the normalisation factors derived from the corresponding control samples. In the same flavor channels, the Z/γ^*+jets contribution is corrected using the ABCD method.

	Signal [125 GeV]	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	Z/γ^*+jets	$W+$jets	Total Bkg.	Observed	Data/MC
lepton p_{T}	60.14 ± 0.23	1230.73 ± 5.37	1532.07 ± 10.73	5008.65 ± 15.18	532.59 ± 6.19	1759771.97 ± 1034.18	1296.28 ± 32.89	1769372.30 ± 1034.91	1769068	1.00 ± 0.00
Opposite Charge	59.48 ± 0.23	1230.75 ± 5.37	1459.17 ± 10.58	4993.87 ± 15.16	527.21 ± 6.16	1759764.92 ± 1034.18	1173.24 ± 32.56	1769149.16 ± 1034.89	1768817	1.00 ± 0.00
$m_{\ell\ell} > 10$ GeV	57.76 ± 0.22	1226.33 ± 5.36	1455.35 ± 10.58	4978.49 ± 15.14	525.62 ± 6.15	1757795.64 ± 1033.90	1025.32 ± 32.26	1767006.76 ± 1034.60	1761569	1.00 ± 0.00
Z veto (for $ee, \mu\mu$)	56.48 ± 0.22	957.75 ± 4.72	143.70 ± 2.16	3901.30 ± 13.41	408.27 ± 5.43	176912.49 ± 322.36	342.80 ± 17.20	182666.32 ± 323.19	182060	1.00 ± 0.00
$E^{\mathrm{miss}}_{\mathrm{T,rel}} > 45$ GeV	22.31 ± 0.14	389.62 ± 24.42	26.72 ± 0.84	1536.65 ± 104.78	173.20 ± 12.31	698.71 ± 23.59	24.67 ± 2.87	2849.57 ± 154.06	2854	1.00 ± 0.06
Z control region (incl)	1.28 ± 0.04	268.58 ± 2.52	1311.64 ± 10.35	1077.19 ± 7.02	117.35 ± 2.90	1580883.15 ± 982.36	682.52 ± 27.30	1584340.43 ± 982.83	1579509	1.00 ± 0.00
Top control region (incl)	0.87 ± 0.03	10.97 ± 0.51	0.84 ± 0.17	1250.73 ± 7.59	108.64 ± 2.80	15.67 ± 2.92	4.86 ± 1.84	1391.72 ± 8.82	1398	1.00 ± 0.03
0j: jet veto	14.33 ± 0.12	269.08 ± 16.93	17.18 ± 0.66	37.03 ± 2.85	20.07 ± 1.84	480.18 ± 19.26	12.17 ± 1.60	835.72 ± 34.31	866	1.04 ± 0.06
0j: $m_{\ell\ell} < 50$ GeV	12.17 ± 0.11	88.86 ± 5.72	7.42 ± 0.45	8.47 ± 0.88	5.33 ± 0.73	80.40 ± 6.89	4.82 ± 0.91	195.31 ± 11.25	224	1.15 ± 0.10
0j: $p_{\mathrm{T},\ell\ell} >45$ GeV	9.37 ± 0.09	74.27 ± 4.82	6.21 ± 0.41	7.47 ± 0.80	4.83 ± 0.68	22.32 ± 5.13	1.33 ± 0.54	116.43 ± 10.02	138	1.19 ± 0.14
0j: $\Delta\phi_{\ell\ell} < 1.8$ rad	9.35 ± 0.09	74.16 ± 4.81	6.21 ± 0.41	7.42 ± 0.80	4.76 ± 0.68	22.31 ± 5.13	1.33 ± 0.54	116.19 ± 10.01	138	1.19 ± 0.14
0j: $0.75 \cdot m_{\mathrm{H}} \leq m_{\mathrm{T}} \leq m_{\mathrm{H}}$	6.20 ± 0.08	23.84 ± 1.68	2.15 ± 0.25	1.26 ± 0.28	1.08 ± 0.29	12.14 ± 3.17	1.11 ± 0.46	41.59 ± 4.53	54	1.30 ± 0.23
0j: Z control region	0.64 ± 0.03	184.61 ± 2.12	296.58 ± 3.24	19.52 ± 0.95	10.03 ± 0.84	1245363.21 ± 872.42	473.26 ± 21.57	1246347.21 ± 872.70	1242575	1.00 ± 0.00
0j: WW control region	0.02 ± 0.01	72.64 ± 1.27	4.30 ± 0.30	15.58 ± 0.82	7.19 ± 0.72	2.58 ± 1.31	1.04 ± 0.44	103.34 ± 2.19	91	0.88 ± 0.09
1j: exactly one jet	5.33 ± 0.06	74.25 ± 11.66	6.74 ± 0.43	312.81 ± 13.45	91.66 ± 4.62	163.17 ± 11.77	8.54 ± 1.51	657.16 ± 36.88	660	1.00 ± 0.07
1j: b-jet veto	4.92 ± 0.06	68.71 ± 10.80	6.29 ± 0.41	90.86 ± 4.31	29.13 ± 1.93	154.22 ± 11.30	6.08 ± 1.21	355.29 ± 24.33	360	1.01 ± 0.09
1j: $p_{\mathrm{T}}^{\mathrm{tot}} < 30$ GeV	3.30 ± 0.05	48.07 ± 7.57	3.57 ± 0.31	38.81 ± 2.13	18.51 ± 1.43	26.42 ± 4.47	2.91 ± 0.77	138.30 ± 13.36	141	1.02 ± 0.13
1j: $Z \to \tau\tau$ veto	3.30 ± 0.05	47.53 ± 7.49	3.55 ± 0.31	37.83 ± 2.09	18.08 ± 1.41	25.45 ± 4.37	2.97 ± 0.77	135.40 ± 13.11	137	1.01 ± 0.13
1j: $m_{\ell\ell} < 50$ GeV	2.70 ± 0.04	13.86 ± 2.23	1.33 ± 0.20	9.44 ± 0.83	5.01 ± 0.66	8.79 ± 2.04	1.26 ± 0.46	39.69 ± 4.35	38	0.96 ± 0.19
1j: $\Delta\phi_{\ell\ell} < 1.8$ rad	2.49 ± 0.04	12.74 ± 2.06	1.33 ± 0.20	9.05 ± 0.80	4.80 ± 0.65	5.20 ± 1.33	0.97 ± 0.41	34.09 ± 3.58	36	1.06 ± 0.21
1j: $0.75 \cdot m_{\mathrm{H}} \leq m_{\mathrm{T}} \leq m_{\mathrm{H}}$	1.67 ± 0.03	4.37 ± 0.74	0.45 ± 0.12	1.77 ± 0.34	0.86 ± 0.27	3.35 ± 1.08	0.46 ± 0.28	11.24 ± 1.66	16	1.42 ± 0.41
1j: Z control region	0.42 ± 0.02	63.08 ± 1.19	437.32 ± 6.19	157.38 ± 2.69	53.39 ± 1.96	252137.08 ± 391.48	162.12 ± 13.68	253010.37 ± 391.79	253468	1.00 ± 0.00
1j: WW control region	0.01 ± 0.01	21.88 ± 0.71	1.14 ± 0.19	17.05 ± 0.88	6.40 ± 0.69	3.57 ± 1.38	0.89 ± 0.44	50.92 ± 1.97	49	0.96 ± 0.14
1j: Top control region	0.27 ± 0.02	4.49 ± 0.33	0.24 ± 0.08	98.21 ± 2.11	37.24 ± 1.64	1.42 ± 0.88	1.29 ± 0.62	142.89 ± 2.90	158	1.11 ± 0.09
$\geq$ 2j: at least 2 jets	2.65 ± 0.05	31.06 ± 0.81	2.80 ± 0.29	1236.01 ± 7.80	67.85 ± 2.31	48.34 ± 5.05	4.03 ± 1.85	1390.09 ± 9.79	1328	0.96 ± 0.03
$\geq$ 2j: central jet veto	2.10 ± 0.04	24.52 ± 0.72	1.91 ± 0.24	688.80 ± 5.77	48.79 ± 1.94	37.94 ± 4.54	2.18 ± 1.34	804.14 ± 7.75	781	0.97 ± 0.04
$\geq$ 2j: b-jet veto	1.77 ± 0.04	21.15 ± 0.67	1.72 ± 0.22	75.90 ± 1.96	9.83 ± 0.87	32.17 ± 4.16	1.40 ± 0.78	142.16 ± 4.80	170	1.20 ± 0.10
$\geq$ 2j: opp. hemispheres	0.96 ± 0.02	9.98 ± 0.45	0.86 ± 0.14	30.41 ± 1.24	4.54 ± 0.59	11.89 ± 2.46	0.91 ± 0.57	58.59 ± 2.91	71	1.21 ± 0.16
$\geq$ 2j: $\Delta\eta_{\mathrm{jj}} > 3.8$	0.41 ± 0.01	1.66 ± 0.19	0.14 ± 0.06	5.25 ± 0.46	0.76 ± 0.23	3.16 ± 1.28	0.18 ± 0.20	11.16 ± 1.41	10	0.90 ± 0.31
$\geq$ 2j: $m_{\mathrm{jj}} > 500$ GeV	0.29 ± 0.01	0.68 ± 0.12	0.10 ± 0.06	1.95 ± 0.26	0.31 ± 0.14	0.27 ± 0.27	0.25 ± 0.20	3.56 ± 0.47	2	0.56 ± 0.40
$\geq$ 2j: $p_{\mathrm{T,tot}} < 30$ GeV	0.21 ± 0.01	0.54 ± 0.10	0.05 ± 0.05	0.81 ± 0.17	0.14 ± 0.10	0.00 ± 0.00	0.19 ± 0.18	1.74 ± 0.29	1	0.57 ± 0.58
$\geq$ 2j: $Z \to \tau\tau$ veto	0.21 ± 0.01	0.54 ± 0.10	0.05 ± 0.05	0.78 ± 0.17	0.14 ± 0.10	0.00 ± 0.00	0.19 ± 0.18	1.71 ± 0.29	1	0.58 ± 0.59
$\geq$ 2j: $m_{\ell\ell} < 80$ GeV	0.21 ± 0.01	0.31 ± 0.08	0.00 ± 0.01	0.39 ± 0.12	0.07 ± 0.07	0.00 ± 0.00	0.20 ± 0.18	0.97 ± 0.24	0	0.00 ± 0.00
$\geq$ 2j: $\Delta\phi_{\ell\ell} < 1.8$ rad	0.19 ± 0.01	0.20 ± 0.07	0.00 ± 0.00	0.28 ± 0.11	0.07 ± 0.07	0.00 ± 0.00	0.00 ± 0.00	0.55 ± 0.15	0	0.00 ± 0.00
$\geq$ 2j: $0.75 \cdot m_{\mathrm{H}} \leq m_{\mathrm{T}} \leq m_{\mathrm{H}}$	0.12 ± 0.00	0.06 ± 0.04	0.00 ± 0.00	0.06 ± 0.04	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.12 ± 0.06	0	0.00 ± 0.00
$\geq$ 2j: Z control region	0.22 ± 0.02	20.89 ± 0.67	577.75 ± 7.64	900.29 ± 6.42	53.92 ± 1.97	83382.86 ± 225.07	47.14 ± 9.63	84982.85 ± 225.50	83466	0.98 ± 0.00
$\geq$ 2j: WW control region	0.00 ± 0.00	0.24 ± 0.06	0.05 ± 0.05	0.39 ± 0.12	0.08 ± 0.08	0.00 ± 0.00	0.00 ± 0.00	0.74 ± 0.17	1	1.34 ± 1.38
$\geq$ 2j: Top control region	0.33 ± 0.02	3.41 ± 0.29	0.18 ± 0.09	612.47 ± 5.45	38.73 ± 1.74	5.67 ± 1.77	0.75 ± 1.09	661.22 ± 6.09	611	0.92 ± 0.04

Table D.3.: The expected numbers of signal (for $m_H = 125$ GeV) and background events as well as the events observed in data at several stages of the selection for the $e\mu$ final state. The $W+$jets background is estimated from data, whereas MC predictions normalised to data in control regions are used for WW, Z/γ^*+jets, $t\bar{t}$, and $tW/tb/tqb$ processes. Contributions from other background sources are taken directly from MC predictions. Only statistical uncertainties associated with the number of events in the MC samples and the data control regions are shown. Starting from jet multiplicity specific selection stages, the numbers for the SM WW and top background contributions are corrected using the normalisation factors derived from the corresponding control samples.

	Signal [125 GeV]	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	Z/γ^*+jets	$W+$jets	Total Bkg.	Observed	Data/MC
lepton p_T	88.32 ± 0.28	1913.43 ± 6.70	515.29 ± 11.90	8132.26 ± 19.52	876.79 ± 8.04	7241.54 ± 53.00	1148.39 ± 11.56	19827.71 ± 59.79	19838	1.00 ± 0.01
Opposite Charge	87.08 ± 0.27	1904.21 ± 6.69	259.08 ± 6.21	8083.92 ± 19.46	867.63 ± 8.01	7106.87 ± 52.32	738.21 ± 10.00	18959.92 ± 58.00	19066	1.01 ± 0.01
$m_{\ell\ell} > 12$ GeV	86.24 ± 0.27	1902.09 ± 6.69	252.47 ± 6.06	8076.54 ± 19.45	866.83 ± 8.00	7090.23 ± 52.28	735.58 ± 9.98	18923.74 ± 57.94	19038	1.01 ± 0.01
Z veto (for $ee, \mu\mu$)	86.24 ± 0.27	1902.09 ± 6.69	252.47 ± 6.06	8076.54 ± 19.45	866.83 ± 8.00	7090.23 ± 52.28	735.58 ± 9.98	18923.74 ± 57.94	19038	1.01 ± 0.01
$E_{T,rel}^{miss} > 25$ GeV	58.28 ± 0.23	1230.87 ± 76.74	124.45 ± 4.35	4855.83 ± 330.36	558.06 ± 38.49	534.85 ± 13.88	260.93 ± 5.05	7565.00 ± 445.09	7725	1.02 ± 0.06
Z control region (incl)	86.24 ± 0.27	1902.09 ± 6.69	252.47 ± 6.06	8076.54 ± 19.45	866.83 ± 8.00	7090.23 ± 52.28	735.58 ± 9.98	18923.74 ± 57.94	19038	1.01 ± 0.01
Top control region (incl)	2.46 ± 0.05	32.99 ± 0.90	5.48 ± 0.82	3982.03 ± 13.66	352.30 ± 5.11	20.30 ± 2.37	16.84 ± 2.91	4409.94 ± 15.11	4758	1.08 ± 0.02
0j: jet veto	36.15 ± 0.19	860.96 ± 53.74	66.05 ± 3.36	111.58 ± 7.93	60.48 ± 4.64	340.48 ± 11.78	165.42 ± 3.14	1604.96 ± 66.68	1569	0.98 ± 0.05
0j: $m_{\ell\ell} < 50$ GeV	27.75 ± 0.16	179.87 ± 11.39	27.38 ± 2.52	16.48 ± 1.47	10.19 ± 1.12	63.11 ± 4.86	47.65 ± 1.67	344.68 ± 14.42	334	0.97 ± 0.07
0j: $p_{T,\ell\ell} > 30$ GeV	26.76 ± 0.16	170.55 ± 10.81	24.50 ± 2.34	16.17 ± 1.45	9.65 ± 1.08	1.93 ± 0.80	40.08 ± 1.37	262.89 ± 12.92	253	0.96 ± 0.08
0j: $\Delta\phi_{\ell\ell} < 1.8$ rad	25.68 ± 0.15	164.14 ± 10.41	23.03 ± 2.25	15.77 ± 1.42	9.58 ± 1.08	1.60 ± 0.76	35.05 ± 1.26	249.17 ± 12.47	239	0.96 ± 0.08
0j: $0.75 \cdot m_H \leq m_T \leq m_H$	16.72 ± 0.12	70.60 ± 4.59	7.86 ± 1.18	4.64 ± 0.60	2.89 ± 0.51	0.00 ± 0.00	21.47 ± 0.97	107.46 ± 5.36	101	0.94 ± 0.10
0j: Z control region	45.57 ± 0.21	1292.85 ± 5.57	116.12 ± 4.49	152.14 ± 2.66	81.26 ± 2.45	5409.87 ± 47.55	433.84 ± 7.16	7486.08 ± 48.75	7075	0.95 ± 0.01
0j: WW control region	0.24 ± 0.02	355.33 ± 2.85	18.46 ± 1.41	58.70 ± 1.61	29.12 ± 1.47	3.06 ± 1.25	43.32 ± 1.46	508.00 ± 4.31	522	1.03 ± 0.05
1j: exactly one jet	14.76 ± 0.11	228.90 ± 35.83	39.80 ± 2.39	922.59 ± 38.53	288.03 ± 12.78	137.98 ± 6.38	64.54 ± 2.43	1681.84 ± 86.27	1679	1.00 ± 0.06
1j: b-jet veto	13.58 ± 0.10	213.00 ± 33.35	36.81 ± 2.28	262.97 ± 11.42	92.46 ± 4.68	127.33 ± 6.11	58.56 ± 2.10	791.13 ± 48.64	809	1.02 ± 0.07
1j: $p_T^{tot} < 30$ GeV	9.12 ± 0.08	150.77 ± 23.62	23.93 ± 1.89	116.27 ± 5.36	52.37 ± 2.98	70.31 ± 4.60	29.36 ± 1.47	443.00 ± 31.14	449	1.01 ± 0.09
1j: $Z \to \tau\tau$ veto	9.09 ± 0.08	146.34 ± 22.93	23.54 ± 1.88	110.67 ± 5.13	50.09 ± 2.88	48.56 ± 3.91	29.60 ± 1.43	408.79 ± 30.04	414	1.01 ± 0.09
1j: $m_{\ell\ell} < 50$ GeV	6.90 ± 0.07	27.92 ± 4.44	9.98 ± 1.42	18.78 ± 1.29	10.00 ± 0.99	12.73 ± 1.92	8.71 ± 0.82	88.12 ± 6.30	110	1.25 ± 0.15
1j: $\Delta\phi_{\ell\ell} < 1.8$ rad	6.39 ± 0.07	25.16 ± 4.01	9.62 ± 1.42	17.69 ± 1.24	9.77 ± 0.98	2.88 ± 1.02	7.27 ± 0.69	72.39 ± 5.62	90	1.24 ± 0.16
1j: $0.75 \cdot m_H \leq m_T \leq m_H$	3.76 ± 0.05	9.31 ± 1.52	3.07 ± 0.81	4.99 ± 0.57	2.69 ± 0.48	0.00 ± 0.00	3.39 ± 0.43	23.45 ± 2.16	33	1.41 ± 0.28
1j: Z control region	25.20 ± 0.14	442.88 ± 3.18	92.23 ± 3.54	1279.17 ± 7.69	402.13 ± 5.45	1223.55 ± 19.00	191.53 ± 4.62	3631.49 ± 22.22	3609	0.99 ± 0.02
1j: WW control region	0.10 ± 0.02	92.46 ± 1.43	8.17 ± 0.88	60.15 ± 1.64	24.69 ± 1.37	3.34 ± 0.95	9.31 ± 0.72	198.13 ± 2.96	179	0.90 ± 0.07
1j: Top control region	0.76 ± 0.03	12.94 ± 0.57	1.53 ± 0.45	275.20 ± 3.56	109.12 ± 2.84	4.88 ± 1.25	1.94 ± 0.69	405.60 ± 4.83	426	1.05 ± 0.05
$\geq$ 2j: at least 2 jets	7.37 ± 0.08	94.10 ± 1.40	18.60 ± 1.40	3971.67 ± 14.10	229.71 ± 4.29	56.39 ± 3.65	30.98 ± 3.13	4401.45 ± 15.63	4477	1.02 ± 0.02
$\geq$ 2j: central jet veto	5.77 ± 0.07	75.26 ± 1.25	14.56 ± 1.32	2171.33 ± 10.31	160.84 ± 3.56	42.43 ± 3.29	19.86 ± 2.19	2484.28 ± 11.75	2521	1.01 ± 0.02
$\geq$ 2j: b-jet veto	4.88 ± 0.06	64.85 ± 1.15	13.07 ± 1.27	232.94 ± 3.45	34.41 ± 1.65	36.83 ± 3.05	14.78 ± 1.36	396.89 ± 5.36	401	1.01 ± 0.05
$\geq$ 2j: opp. hemispheres	2.72 ± 0.04	30.54 ± 0.79	5.45 ± 0.80	99.50 ± 2.22	13.09 ± 1.02	15.93 ± 1.98	6.65 ± 0.88	171.15 ± 3.45	161	0.94 ± 0.08
$\geq$ 2j: $\Delta\eta_{jj} > 3.8$	1.18 ± 0.02	5.57 ± 0.34	0.53 ± 0.12	15.06 ± 0.79	1.12 ± 0.26	1.42 ± 0.66	1.38 ± 0.36	25.08 ± 1.18	22	0.88 ± 0.19
$\geq$ 2j: $m_{jj} > 500$ GeV	0.83 ± 0.01	2.80 ± 0.23	0.25 ± 0.09	6.98 ± 0.49	0.73 ± 0.21	0.44 ± 0.23	0.62 ± 0.25	11.82 ± 0.68	9	0.76 ± 0.26
$\geq$ 2j: $p_{T,tot} < 30$ GeV	0.63 ± 0.01	1.67 ± 0.18	0.13 ± 0.06	2.79 ± 0.31	0.48 ± 0.17	0.27 ± 0.20	0.20 ± 0.14	5.55 ± 0.47	4	0.72 ± 0.37
$\geq$ 2j: $Z \to \tau\tau$ veto	0.62 ± 0.01	1.61 ± 0.17	0.13 ± 0.06	2.77 ± 0.31	0.48 ± 0.17	0.16 ± 0.16	0.14 ± 0.12	5.30 ± 0.45	4	0.75 ± 0.38
$\geq$ 2j: $m_{\ell\ell} < 80$ GeV	0.61 ± 0.01	0.63 ± 0.11	0.09 ± 0.06	0.82 ± 0.19	0.28 ± 0.13	0.16 ± 0.16	0.00 ± 0.00	1.97 ± 0.31	2	1.02 ± 0.74
$\geq$ 2j: $\Delta\phi_{\ell\ell} < 1.8$ rad	0.53 ± 0.01	0.45 ± 0.09	0.08 ± 0.06	0.53 ± 0.17	0.07 ± 0.07	0.00 ± 0.00	0.00 ± 0.00	1.12 ± 0.21	1	0.89 ± 0.91
$\geq$ 2j: $0.75 \cdot m_H \leq m_T \leq m_H$	0.30 ± 0.01	0.14 ± 0.05	0.00 ± 0.00	0.08 ± 0.06	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.21 ± 0.08	0	0.00 ± 0.00
$\geq$ 2j: Z control region	15.47 ± 0.11	166.36 ± 1.89	44.12 ± 2.01	6645.24 ± 17.67	383.43 ± 5.33	456.81 ± 10.56	110.20 ± 5.20	7806.16 ± 22.06	8354	1.07 ± 0.01
$\geq$ 2j: WW control region	0.00 ± 0.00	0.99 ± 0.14	0.05 ± 0.02	1.95 ± 0.25	0.20 ± 0.11	0.00 ± 0.00	0.15 ± 0.11	3.33 ± 0.32	2	0.60 ± 0.43
$\geq$ 2j: Top control region	0.89 ± 0.03	10.42 ± 0.49	1.43 ± 0.33	1936.38 ± 9.76	126.21 ± 3.16	5.56 ± 1.23	5.03 ± 1.72	2085.04 ± 10.49	2120	1.02 ± 0.02

Bibliography

[1] LEP Working Group for Higgs boson searches, *Search for the Standard Model Higgs boson at LEP*, Phys. Lett. **B565** (2003) 61, arXiv:hep-ex/0306033.

[2] ALEPH, CDF, D0, DELPHI, L3, OPAL and SLD Collaborations, LEP and Tevatron EWWG, SLD Electroweak and Heavy Flavour Groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, arXiv:1012.2367 [hep-ex], CERN-PH-EP-2010-095.

[3] ATLAS Collaboration, *Search for the Standard Model Higgs boson in the $H \to WW^* \to l\nu l\nu$ Decay Channel in pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. Lett. **108** (2012) 111802, arXiv:1112.2577 [hep-ex].

[4] ATLAS Collaboration, *Search for the Standard Model Higgs boson in the $H \to WW^* \to l\nu l\nu$ decay mode with 4.7 fb^{-1} of ATLAS data at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B716** (2012) 62–81, arXiv:1206.0756v2 [hep-ex].

[5] ATLAS Collaboration, *Observation of an Excess of Events in the Search for the Standard Model Higgs Boson in the $H \to WW^* \to l\nu l\nu$ Channel with the ATLAS Detector*, ATLAS note: ATLAS-CONF-2012-098.

[6] ATLAS Collaboration, *Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC*, Phys. Lett. **B716** (2012) 1–29, arXiv:hep-ex/1207.7214.

[7] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. **C72(3)** (2012) 1909, arXiv:1110.3174 [hep-ex].

[8] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984.

[9] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[10] A. Pich, *The Standard Model of Electroweak Interactions*, arXiv:0705.4264 [hep-ph].

[11] S. Dittmaier and M. Schumacher, *The Higgs Boson in the Standard Model - From LEP to LHC: Expectations, Searches, and Discovery of a Candidate*, arXiv:1211.4828 [hep-ph].

[12] G. Altarelli, R. Barbieri and F. Caravaglios, *Electroweak Precision Tests: A Concise Review*, Int.J.Mod.Phys. **A13** (1998) 1031–1058, arXiv:9712368 [hep-ph].

[13] Particle Data Group, K. Nakamura *et al.*, *Review of particle physics*, J. Phys. **G37** (2010) 075021.

[14] 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].

[15] D. Gross and F. Wilczek, *Asymptotically Free Gauge Theories. 1*, Phys.Rev. **D8** (1973) 3633.

[16] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys.Lett. **B47** (1973) 365.

[17] H. Politzer, *Asymptotic Freedom: An Approach to Strong Interactions*, Phys.Rept. **14** (1974) 129.

[18] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl.Phys. **22** (1961) 579.

[19] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys.Lett. **13** (1964) 168.

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

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

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

[23] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys.Rev. **145** (1966) 1156.

[24] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys.Rev.Lett. **13** (1964) 321.

[25] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys.Rev.Lett. **13** (1964) 585.

[26] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, Phys. Rept. **457** (2008) 1–216, arXiv:hep-ph/0503172 [hep-ph].

[27] T. Hambye and K. Riesselmann, *Matching conditions and Higgs mass upper bounds revisited*, Phys. Rev. **D55** (1997) 7255–7262, arXiv:hep-ph/9610272 [hep-ph].

[28] The LEP Electroweak Working Group, *LEP/TEV EW WG Plots for Summer 2010*, LEPEWWG.

[29] 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.

[30] CERN, *European School of High-Energy Physics*. CERN, Geneva, 2001. CERN-2001-003.

[31] R. Devenish and A. Cooper-Sarkar, *Deep inelastic scattering*. Oxford University Press, 2004.

[32] 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].

[33] M. Dobbs, S. Frixione, E. Laenen, K. Tollefson, H. Baer, *et al.*, *Les Houches guidebook to Monte Carlo generators for hadron collider physics*, arXiv:hep-ph/0403045 [hep-ph].

[34] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[35] G. C. et al., *HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010.

[36] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv:hep-ph/0204244.

[37] C. O. S. Alioli, P. Nason and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **0904** (2009) 002, arXiv:0812.0578 [hep-ph].

[38] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **1002** (2010) 037, arXiv:0911.5299 [hep-ph].

[39] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys.Rept. **97** (1983) 31.

[40] I. Knowles and T. Sjostrand, *QCD Event Generators*, arXiv:hep-ph/9601212.

[41] 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.

[42] ATLAS Collaboration, *Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$ and 7 TeV in a diffractive limited phase-space measured with the ATLAS detector at the LHC and new PYTHIA6 tune*, ATLAS note: ATLAS-CONF-2010-031.

[43] ATLAS Collaboration, G. Aad *et al.*, *Measurements of underlying-event properties using neutral and charged particles in pp collisions at $\sqrt{s} = 900\,GeV$ and $\sqrt{s} = 7\,TeV$ with the ATLAS detector at the LHC*, Eur.Phys.J. **C71** (2011) 1636, arXiv:1103.1816 [hep-ex].

[44] G. P. S. Dittmaier, C. Mariotti and R. T. (Eds.), *Handbook of LHC Higgs cross sections: 1. Inclusive observables*, arXiv:hep-ph/1101.0593.

[45] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.

[46] L. Reina, *TASI 2011: lectures on Higgs-Boson Physics*, arXiv:1208.5504 [hep-ph].

[47] S. Dawson, *Radiative corrections to Higgs boson production*, Nucl. Phys. **B359** (1991) 283.

[48] M. S. A. Djouadi and P. Zerwas, *Production of Higgs bosons in proton colliders: QCD corrections*, Phys. Lett. B **264** (1991) 440.

[49] D. G. M. Spira, A. Djouadi and P. Zerwas, *Higgs boson production at the LHC*, Nucl. Phys. **B453** (1995) 17, arXiv:hep-ph/9504378.

[50] M. S. D. Graudenz and P. Zerwas, *QCD corrections to Higgs boson production at proton proton colliders*, Phys. Rev. Lett. **70** (1993) 1372–1375.

[51] R. V. Harlander, *Virtual corrections to $gg \to H$ two loops in the heavy top limit*, Phys. Lett. **B492** (2000) 74–80, arXiv:hep-ph/0007289.

[52] D. d. F. S. Catani and M. Grazzini, *Higgs production in hadron collisions: Soft and virtual QCD corrections at NNLO*, JHEP **05** (2001) 025, arXiv:hep-ph/0102227.

[53] R. Harlander and W. Kilgore, *Soft and virtual corrections to $pp \to H+X$ at NNLO*, Phys. Rev. **D64** (2001) 013015, arXiv:hep-ph/0102241.

[54] R. Harlander and W. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, Phys. Rev. Lett. **88** (2002) 201801, arXiv:hep-ph/0201206.

[55] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, Nucl. Phys. **B646** (2002) 220, arXiv:hep-ph/0207004.

[56] J. S. V. Ravindran and W. van Neerven, *NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions*, Nucl. Phys. **B665** (2003) 325, arXiv:hep-ph/0302135.

[57] J. Blumlein and V. Ravindran, *Mellin moments to the next-to-next-to leading order coefficient functions for the Drell-Yan process and hadronic Higgs boson production*, Nucl. Phys. **B716** (2005) 128–172, arXiv:hep-ph/0501178.

[58] M. G. S. Catani, D. de Florian and P. Nason, *Soft-gluon re-summation for Higgs boson production at hadron colliders*, JHEP **0307** (2003) 028, arXiv:hep-ph/0306211.

[59] A. Djouadi and P. Gambino, *Leading electroweak correction to Higgs boson production at proton colliders*, Phys. Lett. Lett. **73** (1994) 2528–2531, arXiv:hep-ph/9406432.

[60] G. U. Aglietti, R. Bonciani and A. Vicini, *Two loop light fermion contribution to Higgs production and decays*, Phys. Lett. B **595** (2004) 432, arXiv:hep-ph/0404071.

[61] G.Degrassi and F. Maltoni, *Two-loop electroweak corrections to Higgs production at hadron colliders*, Phys. Lett. **B600** (2004) 255–260, arXiv:hep-ph/0407249.

[62] C. S. S. Actis, G. Passarino and S. Uccirati, *NNLO computational techniques: the cases $H \to \gamma\gamma$ and $H \to gg$*, Nucl. Phys. **B811** (2009) 182–273, arXiv:hep-ph/0809.3667.

[63] C. S. S. Actis, G. Passarino and S. Uccirati, *NLO electroweak corrections to Higgs boson production at hadron hadron colliders*, Phys. Lett. B **670** (2008) 12, arXiv:hep-ph/0809.1301.

[64] A. D. M. Ciccolini and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs+2jets via weak interactions at the LHC*, Phys. Lett. B **99** (2007) 161803, arXiv:hep-ph/0707.0381.

[65] A. D. M. Ciccolini and S. Dittmaier, *Electroweak corrections to Higgs production via vector-boson fusion at the LHC*, Phys. Rev. **D77** (2008) 013002, arXiv:hep-ph/0710.4749.

[66] K. A. et al., *VBFNLO: A parton level Monte Carlo for processes with electroweak bosons*, Comp. Phys. Commun. **180** (2009) 1661, arXiv:hep-ph/0811.4559.

[67] F. M. S.-O. M. P. Bolzoni and M. Zaro, *Higgs production via vector-boson fusion at NNLO in QCD*, Phys. Lett. B **105** (2010) 011801, arXiv:hep-ph/1003.4451.

[68] H. T and S. Willenbrock, *QCD correction to the pp → WH and ZH total cross sections*, Phys. Lett. B **273** (1991) 167.

[69] A. D. O. Brein and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. B **579** (2004) 149, arXiv:hep-ph/0307206.

[70] S. D. M. L. Ciccolini and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys. Rev. **D68** (2003) 073003, arXiv:hep-ph/0306234.

[71] W. B. et. al., *Higgs Radiation off Top Quarks at the Tevatron and the LHC*, Phys.Rev.Lett. **87** (2001) 201805, arXiv:hep-ph/0107081 [hep-ph].

[72] W. B. et. al., *NLO QCD corrections to $t\bar{t}H$ production in hadron collisions*, Nucl.Phys. **B653** (2003) 151–203, arXiv:hep-ph/0211352 [hep-ph].

[73] L. Reina and S. Dawson, *Next-to-leading Order Results for $t\bar{t}h$ Production at the Tevatron*, Phys. Rev. Lett. **87** (2001) 01804, arXiv:hep-ph/0107101 [hep-ph].

[74] L. R. S. Dawson, L.H. Orr and D. Wackeroth, *Next-to-leading order QCD corrections to $pp \to t\bar{t}H$ production at the CERN Large Hadron Collider*, Phys.Rev. **D67** (2003) 071503, arXiv:hep-ph/0211438 [hep-ph].

[75] L. O. S. Dawson, C. Jackson, L. Reina, and D. Wackeroth, *Associated Higgs production with top quarks at the Large Hadron Collider: NLO QCD corrections*, Phys.Rev. **D68** (2003) 034022, arXiv:hep-ph/0305087 [hep-ph].

[76] ATLAS Collaboration, *ATLAS Sensitivity Prospects for Higgs Boson Production at the LHC Running at 7, 8 or 9* TeV, ATLAS Note: ATL-PHYS-PUB-2010-015.

[77] ATLAS Collaboration, *ATLAS Sensitivity Prospects for Higgs Boson Production at the LHC Running at 7* TeV, ATLAS Note: ATL-PHYS-PUB-2010-009.

[78] ATLAS Collaboration, G. Aad *et al.*, *Measurement of the inclusive $W^{\pm}$ and Z/γ^* cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. **D85** (2012) 072004, arXiv:hep-ex/1109.5141.

[79] ATLAS Collaboration, *Measurement of $t\bar{t}$ production in the all-hadronic channel in $1.02\,\text{fb}^{-1}$ of pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, ATLAS note: ATLAS-CONF-2011-140.

[80] ATLAS Collaboration, *Measurement of the W^+W^- production cross section at $\sqrt{s} = 7$ TeV with the ATLAS detector and limits on anomalous WWZ and WWgamma couplings*, arXiv:1210.2979 [hep-ex].

[81] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/CombinedSummaryPlots.

[82] O. Brüning *et al.*, *LHC design report. Vol. I-III*, CERN-2004-003-V-1, CERN-2004-003-V-2, CERN-2004-003-V-3.

[83] *LHC: the guide*, http://cdsweb.cern.ch/record/1165534.

[84] ATLAS Collaboration, W. W. Armstrong *et al.*, *ATLAS: Technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*, CERN-LHCC-94-43.

[85] ATLAS Collaboration, A. Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics*, `arXiv:0901.0512 [hep-ex]`.

[86] *CMS, the Compact Muon Solenoid: Technical proposal*, CERN-LHCC-94-38.

[87] LHCb Collaboration, S. Amato *et al.*, *LHCb technical proposal*, CERN-LHCC-98-004.

[88] *ALICE: Technical proposal for a large ion collider experiment at the CERN LHC*, CERN-LHCC-95-71.

[89] *LEP design report*, CERN-LEP-TH-83-29, CERN-LEP-84-01.

[90] *CERN*, `http://cdsweb.cern.ch/record/42384`.

[91] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** no. 08, (2008) S08001.

[92] P. Lebrun, *Interim Summary Report on the Analysis of the 19 September 2008 Incident at the LHC*, 2008-10-16.

[93] `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`.

[94] C. Carli *et al.*, *Chamonix 2012 Workshop on LHC Performance - Chamonix 2011*, CERN-2012-006.

[95] ATLAS Collaboration, A. collaboration, *ATLAS detector and physics performance: Technical Design Report, 1*, CERN-LHCC-99-014.

[96] ATLAS Collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[97] V. Cindro *et al.*, *The ATLAS beam conditions monitor*, JINST **3** (2008) P02004.

[98] ATLAS Collaboration, A. collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823, `arXiv:1005.4568 [physics.ins-det]`.

[99] ATLAS Collaboration, G. Duckeck *et al.*, *ATLAS computing: Technical design report*, CERN-LHCC-2005-022.

[100] GEANT4 Collaboration, S. Agostinelli *et al.*, *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[101] ATLAS Collaboration, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATLAS note: ATL-SOFT-PUB-2007-007.

[102] ATLAS Collaboration, *Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV*, ATLAS note: ATLAS-CONF-2010-072.

[103] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment*, ATLAS note: ATLAS-CONF-2010-069.

[104] R. Frühwirth, W. Waltenberger, and P. Vanlaer, *Adaptive vertex fitting*, J. Phys. **G34** (2007) N343.

[105] ATLAS Collaboration, *Reconstruction of primary vertices in pp collisions at energies of 900 GeV and 7 TeV with the ATLAS detector*, `arXiv:1010.1483v1 [physics.ins-det]`.

[106] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, `arXiv:1110.1530v1 [hep-ex]`.

[107] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2011.*

[108] ATLAS Collaboration, *Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector*, ATLAS note: ATLAS-CONF-2011-063.

[109] ATLAS Collaboration, S. H. et al., *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC at CERN*, IEEE Trans. Nucl. Sci. **51** (2004) 3030–3033.

[110] ATLAS Collaboration, T. L. et al., *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages*, Nucl.Istrum.Meth. **A572** (2007) 77–79.

[111] T. Michael, *Determination of muon reconstruction efficiencies in the ATLAS detector using a tag-and-probe approach in $Z \to \mu\mu$ events*, CERN-THESIS-2011-058.

[112] `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonPerformancePublicPlots`.

[113] ATLAS Collaboration, *ATLAS Muon Momentum Resolution in the First Pass Reconstruction of the 2010 $p-p$ Collision Data at $\sqrt{s} = 7$ TeV*, ATLAS note: ATLAS-CONF-2011-046.

[114] ATLAS Collaboration, *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\sqrt{s} = 7$ TeV in 2011*, ATLAS note: ATLAS-CONF-2012-048.

[115] ATLAS Collaboration, *Expected electron performance in the ATLAS experiment*, ATLAS note: ATL-PHYS-PUB-2011-006.

[116] J. Hartert and I. Ludwig, *Electron isolation in the ATLAS experiment*, ATLAS internal note: ATL-PHYS-INT-2010-052.

[117] H. H. W. M. Hance, D. Olivito, *Performance Studies of e/gamma Calorimeter Isolation*, ATLAS internal note:: ATL-COM-PHYS-2011-1186.

[118] M. Cacciari and G. P. Salam, *The anti-k_t jet clustering algorithm*, JHEP **0804** (2008) 063, `arXiv:hep-ph/0802.1189v2`.

[119] T. A. Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, `arXiv:1112.6426v1 [hep-ex]`.

[120] ATLAS Collaboration, W. L. et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATLAS note: ATL-LARG-PUB-2008-002.

[121] ATLAS Collaboration, E. A. et al, *Combined performance studies for electrons at the 2004 ATLAS combined test-beam*, JINST **5** (2010) P11006.

[122] ATLAS Collaboration, *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy sqrts = 7 TeV*, ATLAS note: ATLAS-CONF-2010-054.

[123] ATLAS Collaboration, *The identification of b-Jets with the ATLAS Detector*, ATLAS note: ATLAS-XXX.PDF.

[124] ATLAS Collaboration, *Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with* $5\,\text{fb}^{-1}$ *of Data from the ATLAS Detector*, ATLAS note: ATLAS-CONF-2012-043.PDF.

[125] ATLAS Collaboration, *Measurement of the Mistag Rate of b-tagging algorithms with* $5\,\text{fb}^{-1}$ *of Data Collected by the ATLAS Detector*, ATLAS note: ATLAS-CONF-2012-040.PDF.

[126] ATLAS Collaboration, G. Aad *et al.*, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS*, Eur. Phys. J. **C72** (2012) 1844, arXiv:1108.5602 [hep-ex].

[127] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS*, ATLAS note: ATLAS-CONF-2012-105.

[128] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtMissPublicCollisionResults.

[129] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[130] J. Adelman *et al.*, *ATLAS offline data quality monitoring*, J. Phys. Conf. Ser. **219** (2010) 042018.

[131] M. Vincter, J. Kretzschmar, and E. Schmidt, *Update of electron shower shapes in EM calorimeter*, ATLAS internal note: ATL-COM-PHYS-2011-1299.

[132] *Combined CDF and DØ Upper Limits on Standard Model Higgs Boson Production with up to* $6.7\,\text{fb}^{-1}$ *of Data*, arXiv:1007.4587 [hep-ex].

[133] B. G. A. J. Barr and C. G. Lester, *Measuring the Higgs boson mass in dileptonic W-boson decays at hadron colliders*, JHEP **0907** (2009) 072, arXiv:0902.4864 [hep-ph].

[134] ATLAS Collaboration, *Measurement of the* W^+W^- *production cross section in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, ATLAS note:ATLAS-CONF-2012-025.

[135] R. K. E. J. M. Campbell and C. Williams, *Vector boson pair production at the LHC*, JHEP **1107** (2011) 018, arXiv:1105.0020 [hep-ph].

[136] ATLAS Collaboration, G. Aad *et al.*, *Measurement of the* $W \to l\nu$ *and* $Z/\gamma^* \to ll$ *production cross sections in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, JHEP **12** (2010) 060, arXiv:1010.2130 [hep-ex].

[137] ATLAS Collaboration, G. Aad *et al.*, *Measurement of the cross section for top-quark pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using final states with two high-pt leptons*, JHEP **1205** (2012) 059, arXiv:1202.4892v3 [hep-ex].

[138] ATLAS Collaboration, G. Aad *et al.*, *Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, arXiv:1205.3130v1 [hep-ex], CERN-PH-EP-2012-082.

[139] ATLAS Collaboration, G. Aad *et al.*, *Evidence for the associated production of a W boson and a top quark in ATLAS at $\sqrt{s} = 7$ TeV*, arXiv:1205.5764v1 [hep-ex], CERN-PH-EP-2012-117.

[140] ATLAS Collaboration, *Search for s-Channel Single Top-Quark Production in pp Collisions at $\sqrt{s} = 7$ TeV*, ATLAS note: ATLAS-CONF-2011-118.

[141] ATLAS Collaboration, G. Aad *et al.*, *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC*, Eur. Phys. J. **C71** (2011) 1630, arXiv:1101.2185 [hep-ex].

[142] ATLAS Collaboration, *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector in 2011*, ATLAS note: ATLAS-CONF-2011-116.

[143] J. K. A. Djouadi and M. Spira, *HDECAY: A program for HIggs boson decays in the Standard Model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56, arXiv:hep-ph/9704448.

[144] R. B. C. Anastasiou and F. Petriello, *Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion*, JHEP **0904** (2009) 003, arXiv:hep-ph/0811.3458.

[145] D. de Florian and M. Grazzini, *Higgs production through gluon fusion: Updated cross sections at the Tevatron and the LHC*, Phys. Lett. B **674** (2009) 291, arXiv:hep-ph/0901.2427.

[146] J. Baglio and A. Djouadi, *Higgs production at the LHC*, JHEP **1103** (2011) 055, arXiv:hep-ph/1012.0530.

[147] D. e. a. de Florian, *Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC*, JHEP **11** (2011) 064, arXiv:1109.2109v1 [hep-ph].

[148] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, *Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO*, Phys. Rev. Lett. **103** (2009) 082001, arXiv:0903.2120 [hep-ph].

[149] M. L. M. et al., *ALPGEN, a generator for hard multi-parton processes in hadronic collisions*, JHEP **0307** (2003) 001, arXiv:0206293 [hep-ph].

[150] J. A. et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. **C53** (2008) 473, arXiv:0706.2569 [hep-ph].

[151] N. K. T. Binoth, M. Ciccolini and M. Krämer, *Gluon-induced W-boson pair production at the LHC*, JHEP **0612** (2006) 046, arXiv:0611170 [hep-ph].

[152] T. G. et al., *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007, arXiv:0811.4622 [hep-ph].

[153] J. A. et al., *MadGraph MadEvent v4: The new web generation*, JHEP **0709** (2007) 028, arXiv:0706.2334 [hep-ph].

[154] F. M. J. Alwall, M. Herquet, O. Mattelaer, and T. Stelzer, *MadGraph5 : Going Beyond*, JHEP **06** (2011) 128, arXiv:1106.0522 [hep-ph].

[155] M. P. R. C. Gray, C. Kilic, S. Somalwar, and S. Thomas, *Backgrounds to Higgs boson searches from $W\gamma^* \to l\nu l(l)$ asymmetric internal conversion*, arXiv:1110.1368 [hep-ph].

[156] E. R.-W. B. P. Kersevan, *The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5*, arXiv:0405247v2 [hep-ph].

[157] H.-L. Lai *et al.*, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[158] J. M. Campbell and R. K. Ellis, *An update on vector boson pair production at hadron colliders*, Phys. Rev. **D60** (1999) 113006, arXiv:hep-ph/9905386.

[159] U. L. M. Aliev, H. Lacker and P. Uwer, *HATHOR, HAdronic Top and Heavy quarks crOss section calculatoR*, 2011.

[160] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, 2010.

[161] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, 2010.

[162] ATLAS Collaboration, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV*, ATLAS note: ATLAS-CONF-2010-038.

[163] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets2011.

[164] R. K. E. et al., *Higgs decay to $\tau^+\tau^-$: A possible signature of intermediate mass Higgs bosons at the SSC*, Nucl. Phys. **B297** (1988) 221.

[165] G. P. S. Dittmaier, C. Mariotti and R. T. (Eds.), *Handbook of LHC Higgs cross sections: 2. Differential distributions*, arXiv:hep-ph/1201.3084.

[166] I. Stewart and F. Tackmann, *Theory uncertainties for Higgs mass and other searches using jet bins*, Phys. Rev. **D85** (2012) 030411, arXiv:1107.2117 [hep-ph].

[167] ATLAS and CMS Collaborations, *Procedure for the LHC Higgs boson search combination in summer 2011*, ATLAS note: ATL-PHYS-PUB-2011-011.

[168] M. H. Seymour, *The Higgs boson line shape and perturbative unitarity*, Phys. Lett. B **354** (1995) 409, arXiv:hep-ph/9505211.

[169] C. S. G. Passarino and S. Uccirati, *Higgs Pseudo-Observables, Second Riemann Sheet and All That*, Nucl. Phys. **B834** (2010) 77, arXiv:hep-ph/1001.3360.

[170] M. B. et al., *The PDF4LHC working group interim recommendations*, arXiv:hep-ph/1101.0538.

[171] R. D. Ball *et al.*, *Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology*, Nucl. Phys. **B849** (2011) 296, arXiv:1101.1300 [hep-ph].

[172] R. Aben *et al.*, *Statistical analysis aspects of the search for the Standard Model Higgs boson in the $H \to WW^* \to l\nu l\nu$ decay mode with 4.7 fb^{-1} of ATLAS data at $\sqrt{s} =$ 7 TeV*, ATLAS internal note: ATL-COM-PHYS-2012-350.

[173] A. L. Read, *Presentation of search results: the CL_s technique*, J. Phys. **G28** (2002) 2693.

[174] E. G. Glen Cowan, Kyle Cranmer and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv:physics.data-an/1007.1727.

[175] ATLAS Collaboration, *Combined search for the Standard Model Higgs Boson using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS Detector at the LHC*, Phys. Lett. **B710** (2012) 49–66, arXiv:hep-ex/1202.1408.

[176] R. Aben *et al.*, *Nominal $H \to WW^* \to l\nu l\nu$ analysis with 5.8 fb^{-1} of data collected with the ATLAS detector at $\sqrt{s} = 8$ TeV*, ATLAS internal note: ATL-COM-PHYS-2012-875.

[177] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, CERN-LCGAPP **04** (2007) 1, arXiv:hep-ph/0710.3820.

[178] S. D. A. Bredenstein, A. Denner and M. Weber, *Precise predictions for the Higgs-boson decay $H \to WW/ZZ \to 4$ leptons*, Phys. Rev. **D74** (2006) 013004, arXiv:hep-ph/0604011.

[179] S. D. A. Bredenstein, A. Denner and M. Weber, *Radiative corrections to the semileptonic and hadronic Higgs-boson decay $H \to WW/ZZ \to 4$ leptons*, JHEP **0702** (2007) 080, arXiv:hep-ph/0611234.

[180] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS note: ATLAS-CONF-2012-047.

[181] R. Aben *et al.*, *Object selection for the $H \to WW^*$ search with the ATLAS detector at $\sqrt{s} = 8$ TeV*, ATLAS internal note: ATL-COM-PHYS-2012-861.

[182] W. Eadie and F. James, eds., *Statistical Methods in Experimental Physics*. North-Holland Publishing Co, 1971. ISBN: 0-444-10117-9.

[183] R. Aben *et al.*, *Theoretical inputs to the $H \to WW^*$ search with the ATLAS detector at $\sqrt{s} = 7$ TeV*, ATLAS internal note: ATL-COM-PHYS-2012-862.

[184] ATLAS Collaboration, G. Aad *et al.*, *Measurement of the $W^{\pm}Z$ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. C. **72** (2012) 2173, arXiv:1208.1390 [hep-ex].

[185] ATLAS Collaboration, *Observation of an excess of events in the search for the Standard Model Higgs Boson in the $H \to ZZ^* \to 4l$ channel with the ATLAS detector*, ATLAS note: ATLAS-CONF-2012-092.

[186] ATLAS Collaboration, *Measurement of the total ZZ production cross section in the four-lepton channel using 5.8 fb^{-1} of ATLAS data at $\sqrt{s} = 8$ TeV* , ATLAS note:ATLAS-CONF-2012-090.

[187] ATLAS Collaboration, *Observation of an excess of events in the search for the Standard Model Higgs Boson in the $\gamma\gamma$ channel with the ATLAS detector*, ATLAS note: ATLAS-CONF-2012-091.

[188] ATLAS Collaboration, *Combined search for the Standard Model Higgs Boson in pp collision at $\sqrt{s} = 7$ TeV with the ATLAS Detector*, Phys. Rev. **D86** (2012) 032003, arXiv:hep-ex/1207.0319.

[189] K. S. C. e. a. L. Moneta, K. Belasco, *The RooStats Project*, 2010.

[190] L. M. e. a. K. Cranmer, G. Lewis, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, CERN note: CERN-OPEN-2012-016.

[191] W. Verkerke and D. Kirby, *The RooFit toolkit for data modeling*, 2003.

[192] ATLAS Collaboration, A. Collaboration, *Search for the Standard Model Higgs boson in the $H \to ZZ \to llνν$ decay channel using 4.7 fb^{-1} of $\sqrt{s} = 7$ TeV data with the ATLAS detector*, Phys. Lett. B (2012) , arXiv:1205.6744v1 [hep-ex].

[193] ATLAS Collaboration, A. Collaboration, *Search for a Standard Model Higgs boson in the mass range 200-600 GeV in the $H \to ZZ \to llqq$ decay channel*, Phys. Lett. B (2012) , arXiv:1206.2443 [hep-ex].

[194] ATLAS Collaboration, A. Collaboration, *Search for the Standard Model Higgs boson in the $H \to WW^* \to \ell\nu jj$ decay channel at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Lett. B (2012) , arXiv:1206.6074 [hep-ex].

[195] ATLAS Collaboration, A. Collaboration, *Search for the Standard Model Higgs boson in the $H \to \tau\tau$ decay mode in $\sqrt{s} = 7$ TeV pp collisions with ATLAS*, JHEP (2012) , arXiv:1206.5971 [hep-ex].

[196] ATLAS Collaboration, A. 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*, arXiv:1207.0210 [hep-ex].

[197] ATLAS and CMS Collaborations, *Combined Standard Model Higgs Boson searches with up to 2.3 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV at the LHC*, ATLAS note: ATLAS-CONF-2011-157.

[198] ATLAS and CMS Collaborations, *Combined Standard Model Higgs Boson searches with up to 2.3 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV at the LHC*, CMS note: CMS-PAS-HIG-11-023.

[199] L. D. Landau, *The moment of a 2-photon system*, Dokl. Akad. Nawk. USSR **60** (1948) 207.

[200] C. Yang, *Selection Rules for the Dematerialization of a Particle into Two Photons*, Phys.Rev **77** (1950) 242–245.

[201] A. Armbruster, *Private communication.*

[202] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30–61, arXiv:hep-ex/1207.7235.

"But I don't want to go among mad people, Alice remarked.
Oh, you can't help that, said the Cat: we're all mad here. I'm mad. You're mad.
How do you know I'm mad ? said Alice.
You must be, said the Cat, or you wouldn't have come here."

Lewis Carroll, *Alice in Wonderland*

Acknowledgments

I would like to take the opportunity to thank those people that I had the pleasure to work with during the time of my PhD program.

First of all I would like to thank Prof. Karl Jakobs for the opportunity to work on such an interesting topic. I am greatful for his guidance and support through all the years of my PhD program. Many thanks for the possibility to spend so many months at CERN which is an experience that I would never want to miss.

A big thanks goes to Christina Skorek for all her help with the smaller and sometimes bigger problems with administrative issues. Your friendly manner and the enjoyable chats in the morning provided a good start to the day.

Special thanks goes to my dear office colleagues Susanne Kühn, Matti Werner, Johanna Bronner and Hannah Arnold for creating a friendly, amusing and cooperative atmosphere. Many thanks to you for your patience, for cheering me up and comforting me! I am very thankful to my flatmates Giacinto Piacquadio, Susanne Kühn und Christian Weiser for the great time we had together in St. Genis. Giacinto, my time at CERN would not have been half as nice without your company. Thanks a lot for the many fruitful discussions while cooking late at night and for taking care of me during the ups and downs of my thesis. A deep thank goes to Susanne for being a true friend to me, encouring and supporting me in many ways during these years. I consider myself unimaginable lucky to share an office with you for all these years. Christian, thanks a lot for being an amiable, enjoyable and even-tempered flat and office mate. I really enjoyed the pleasant discussions over a glass of wine (or two) that was supposed to speed up my productivity.

A big thank you to Jochen Hartert for his ever-willingness to listen to me, for his assistance with technical problems and for providing his very valuable objective view on topics (and problems).

I'm thankful to Stan Lai for his humorous company when sharing an office at CERN. Thanks a lot for your help with difficult phrasings of emails and all the useful discussions.

I would like to thank Matti for the collaboration and exchange of ideas on the fast simulation tasks. I am greatful to Michael Duehrssen for sharing his experience especially on the fast simulation but as well on the physics and computing issues.

During my work on the $J/\psi \to e^+e^-$ analysis in the second year of my PhD, I had the pleasure to collaborate with Marthe Teinturier, Christos Anastopoulos and Lydia Fayard. I would like to thank Lydia and Christos for sharing their experience and providing input to my studies. Thank you Marthe for the detailed comparisons and the friendly cooperation. I'm greatful to Marthe and Marina Billoni for an amazing time at the summer school in Cargese.

Many thanks to Elina Berglund, Sarah Heim and Maria Fiascaris for working together with me on the electron performance studies. I appreciate very much the friendly atmosphere and

the very helpful discussions while cross-checking our results. I'm thankful to Daniel Froidevaux for sharing his experience and providing very useful feedback to my studies.

I want to thank Stefan Winkelmann for the nice company during SCT shift training. A special thanks goes to Stefan Winkelmann, Holger von Radziewski and Markus Warsinsky for taking such good care of me after my accident - especially after a night shift! I would have been lost without you!

I'm very thankful to Ralf Bernhard and Andreas Walz for the great time we had together on the Higgs analysis. You contributed a lot to make my second summer at CERN an unforgettable experience. Many thanks to Ralf for sharing his braod experience on the Higgs analysis and for introducing me to the Higgs group. Andi, it was a great pleasure to work together with you especially during the times when "the sun did not set in Freiburg". Together with you, the work was a lot more fun! Thank you for your never-ending willingness to help solve the minor or major technical problems. I am also deeply thankful to Tuan Vu Ahn for the many fruitful discussions, for taking care of me during this last year of my PhD and for proofreading my thesis. Attending the workshop in Ischia together with you, Andi and Tuan, was one of the highlights during my PhD.

I would like to thank Markus Schumacher for the nice dialogs and his encouragement. Many thanks to Kristin for cheering me up so many times and for your invaluable advice.

I'm thankful to all members of the Freiburg group for the friendly, constructive and positive atmosphere.

A special thank goes to Magda Chelstowska, Elina Berglund, Johanna Bronner and Joana Machado Miguens for the wholehearted cooperation. Thanks to Magda for all the detailed cross-checks, for sharing all the informations and for many nice chats. Johanna, I am very lucky that we met again on this analysis. Thank you for your nice company and the great collaboration.

I'm very thankful to Jianming Qian for his confidence in my work and for the opportunity to take responsibility for important parts of the analysis.

Many thanks to Christoph Anders for his kind and amusing company at the conference in Ann Arbor and many thanks to Jianming Qian for his hospitality.

I want to thank all the members of `HSG3` group for the friendly atmosphere and for their input to my studies.

Last but not least I would like to thank my family for their unwavering support. Ein besonderer Dank gebührt meinen Eltern und meinen Brüdern für ihre vielfältige Unterstützung und die erholsamen Tage zu Hause in denen ich immer wieder neue Kraft schöpfen konnte. Ich bin unermesslich dankbar für die Gewissheit, dass ich mich immer auf euch verlassen kann, egal was kommt! Ein besonderer Dank geht an meine Oma fürs Daumendrücken und für ihr Vertrauen in mich. Ein grosses Dankeschön an meine ganze Familie für die Hilfe bei all den Aufgaben und Arbeiten rund um den Weinberg! Ein ganz besonderes Dankeschön geht an Rüdiger, für seine unendliche Geduld - besonders wenn ein meeting am Sonntag mal wieder länger dauerte oder unbedingt noch ein job abgeschickt werden musste! Vielen Dank für deinen Optimismus und deine Unterstützung während meiner gesamten Doktorarbeit.