MASTERARBEIT

Study of Higgs boson production via Vector Boson Fusion in the $H{\to}W^{\mp}W^{\pm(*)}{\to}\ell^{-}\bar{\nu}\ell'^{+}\nu'$ decay mode with the ATLAS Detector

Carsten D. Burgard

Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität Freiburg

Study of Higgs boson production via Vector Boson Fusion in the $H{\rightarrow}W^{\mp}W^{\pm(*)}{\rightarrow}\ell^{-}\bar{\nu}\ell'^{+}\nu'$ decay mode with the ATLAS Detector

MASTERARBEIT

vorgelegt von

Carsten D. Burgard

Prof. Dr. Karl Jakobs

Fakultät für Mathematik und Physik

ALBERT-LUDWIGS-UNIVERSITÄT

Freiburg im Breisgau

3. Oktober 2013

Dekan: Prof. Dr. Michael Růžička
Betreuer: Prof. Dr. Karl Jakobs

"If I have seen further it is by standing on ye shoulders of giants"

– Isaac Newton

Eidesstattliche Erklärung

Hiermit versichere ich, die eingereichte Masterarbeit selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben. Alle Zitate sind gekennzeichnet und alle Abbildungen enthalten nur die originalen Daten und sind keiner inhaltsverändernden Bildbearbeitung unterzogen worden. Ich erkläre weiterhin, dass die vorliegende Arbeit noch nicht anderweitig als Masterarbeit eingereicht wurde.

3. Oktober 2013

....
Ort, Datum   Unterschrift

Abstract

The search for the elusive Higgs Boson, the missing piece of evidence for the Englert–Brout–Higgs–Guralnik–Hagen–Kibble mechanism of Electroweak Symmetry breaking, has been ongoing for several decades. The ATLAS and CMS experiments at the Large Hadron Collider at CERN in Geneva have brought a remarkable amount of public attention to this quest, culminating with the announcement of discovery of a new boson with a mass slightly above $125\,\mathrm{GeV}$ on July 4th, 2012. Studies investigating whether this newly discovered particle is indeed the Higgs boson that would to current knowledge complete the particle content of the Standard Model of Particle Physics are ongoing. So far, however, no evidence to the contrary has been found.

This thesis presents a search for evidence of Higgs boson production via Vector Boson Fusion in the $H \to W^{\mp} W^{\pm(*)} \to \ell^- \bar{\nu} \ell'^+ \nu'$ decay mode, using data collected by the ATLAS detector in 2012 in proton-proton collisions with a centre-of-mass energy of $\sqrt{s} = 8\,\mathrm{TeV}$. Earlier studies by the ATLAS collaboration are revisited and improved upon. A mechanism for efficient exploitation of Monte Carlo samples used for background estimation is presented and evaluated. An optimization procedure for the event selection based on a full likelihood fit including systematic uncertainties is developed, presenting and discussing the implementation, results and performance. Employing these methods, a measurement of the convoluted cross section and branching ratio for VBF Higgs production in the $H \to W^{\mp} W^{\pm(*)} \to \ell^- \bar{\nu} \ell'^+ \nu'$ decay mode is conducted and presented.

Zusammenfassung

Die Suche nach dem Higgs Boson, der experimentell nachzuweisenden Manifestation des Englert–Brout–Higgs–Guralnik–Hagen–Kibble-Mechanismus' der Elektroschwachen Symmetriebrechung, dauert schon einige Jahrzehnte an. In den letzten Jahren ist das öffentliche Interesse dank der Experimente ATLAS und CMS am Großen Hadronenbeschleuniger in Genf stetig angewachsen, um zuletzt bei der Ankündigung der Entdeckung eines neuen Teilchens mit einer Masse von $m \approx 125\,\text{GeV}$ durch die beiden Kollaborationen am 4. Juli 2012 einen vorläufigen Höhepunkt zu erreichen. Umfangreiche Studien, welche die Natur des neuen Teilchens untersuchen und überprüfen, ob es sich tatsächlich um dasjenige Higgs-Teilchen handelt, welches den Teilcheninhalt des Standardmodells nach heutigem Stand der Forschung vervollständigen würde, wurden seither durchgeführt. Bisher gibt es jedoch keine Anzeichen für Abweichungen von den vorhergesagten Eigenschaften eines Standardmodell-Higgs-Teilchens.

In dieser Arbeit werden Studien zur Suche nach der Produktion von Higgs-Teilchen durch Vektor-Boson-Fusion im Zerfallskanal $H \to W^{\mp}W^{\pm(*)} \to \ell^- \bar{\nu} \ell'^+ \nu'$ vorgestellt. Hierfür werden vom ATLAS-Detektor im Jahr 2012 in Proton-Proton-Kollisionen mit einer Schwerpunktsenergie von $\sqrt{s} = 8\,\text{TeV}$ aufgezeichnete Daten verwendet. Frühere Studien durch die ATLAS-Kollaboration werden hierfür überarbeitet und optimiert. Ein Mechanismus zur effizienteren Ausnutzung von Monte-Carlo zufallsgenerierten Datensätzen zur Untergrundabschätzung wird vorgestellt und ausgewertet. Eine Studie zur Optimierung der Ereignisselektion basierend auf einem Maximum-Likelihood-Verfahren einschließlich systematischer Unsicherheiten wird ausgearbeitet und diskutiert. Die Umsetzung und Leistungsfähigkeit des Verfahrens werden gemeinsam mit den aus der Studie erhaltenen Ergebnissen vorgestellt. Unter Ausnutzung dieser neuartigen Verfahren wird eine Messung des Produkts aus Wirkungsquerschnitt und Verzweigungsverhältnis im untersuchten Zerfallskanal durchgeführt und deren Ergebnisse präsentiert.

Preface

In some regard, young particle physicists joining a big collaboration like ATLAS or CMS today can consider themselves the *lucky generation*. Many decades of inexplicable effort, incorporating the hard work and creativity of thousands of scientists, have brought us to where we are today. The LHC and the ATLAS and CMS detectors, each one of them among the largest and most complex machines that have ever been built by mankind, have been providing a steady stream of collision data at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ for the better part of a year, ready to be analysed. For many years, the scientific community and even the general public have been waiting eagerly for the first results.

I certainly consider myself lucky – not only for the opportunity to join the amazing quest of unravelling the cosmic mysteries, but also for enjoying the company of some truly fine colleagues and friends along the way.

Acknowledgements

Table of Contents

1 Introduction

Mankind's quest to understand the inner workings of nature and the universe might be as old as mankind itself. Along the way, several milestones have radically changed our understanding of the universe and even the concept of reality. Newton's description of the gravitational force, Maxwell's famous equations, Einstein's discovery of general relativity, and quantum mechanics with the discovery and description of the weak and strong nuclear forces have revolutionized science and technology. A re-occurring trend in these course-changing milestones seems to be the discovery of new natural forces, or their unification, that is, their description through a common theoretical concept and formalism.

The unification of the electromagnetic and the weak force by Glashow, Salam and Weinberg and the postulation of the weak mediator particles in 1967/68 [1] with subsequent attribution of the Nobel prize in 1979 provides a candidate for the next milestone on this quest. The postulated particles were discovered in 1983 by the UA1 and UA2 experiments [2] at the European Organization for Nuclear Research (CERN, acronym resembles former name *Conseil Européen pour la Recherche Nucléaire*), and were the first gauge bosons with a non-zero invariant mass to be discovered. A group of problems arising from their non-vanishing masses could be solved by the original authors, employing the mechanism of spontaneous symmetry breaking, which was proposed by François Englert, Robert Brout, Peter Higgs, Gerald Guralnik, Carl Hagen and Tom Kibble as early as 1964 [3, 4].

This theoretical mechanism gave rise to the postulation of an additional particle, which was named *Higgs boson* after one of the leading authors of the aforementioned mechanism. However, the discovery of this newly postulated particle proved more difficult than initially expected.

In the meantime, mankind's understanding of the forces of nature was summarized into a coherent theoretical framework, which is today called the Standard Model of Particle Physics (SM). The Standard Model shows a remarkable amount of precision in the description and prediction of processes and experimental outcomes. Nevertheless, some questions remain unanswered, such as the nature of dark matter or the perceived problem of the hierarchy between the fundamental scales incorporated by the Standard Model. The simplest known solution to the cosmic

puzzle of the origin of the fundamental particles' masses is the mechanism of electroweak symmetry breaking (EWSB), for which the discovery of the Higgs boson has long been a missing key ingredient.

On July 4th 2012, the decade long search for the previously unobserved Higgs boson was met with the announced discovery of a new boson with a mass of $m \approx 125\,\text{GeV}$ by the ATLAS and CMS experiments [5, 6]. These are worldwide renowned, high-profile experimental facilities at the Large Hadron Collider (LHC) in Geneva, dedicated to the test and extension of knowledge in particle physics, and – among other subjects – especially to the discovery of the elusive Higgs boson.

This discovery did not only gather public attention even outside the scientific community [7], but also marks the end of an era, possibly completing the observed particle content of the Standard Model. However, whether the newly announced particle is indeed the boson predicted by the Higgs mechanism remains to be seen and is subject of ongoing research.

Apart from direct property measurements such as of the spin of the new particle, a thorough investigation must also include the observation of the particle in less prominent decay channels. Thereby, it is possible to provide an indirect measurement of the particle's couplings to the decay products. Different production modes can be revealed through varying detector signatures of the corresponding collision events.

For an investigation of the Higgs-like nature of the new boson, the decay channel in which the boson decays to a pair of charged weak bosons and further to a pair of oppositely charged leptons and their corresponding neutrinos is of particular interest, as it provides a relatively high chance of discovery. Corresponding theories predict that the Higgs boson decay into a pair of charged weak bosons is the dominant decay mode for intermediate and high Higgs boson mass hypotheses, theorizing a mass value above $m_H \approx 135\,\text{GeV}$. Indeed, an excess in this decay channel has contributed to the announced discovery in summer 2012.

The main production modes for a Higgs particle at the Large Hadron Collider include, most prominently, the production by interaction of two strong gauge bosons (commonly labelled ggF, for *gluon-gluon-Fusion*, but referred to as *gluon fusion*), but also the production by interaction of two vector bosons (labelled VBF, for *Vector Boson Fusion*). The latter would incorporate a detector signature with two additional hadronic jets, emerging from the primary interaction vertex into opposing directions close to the beam pipe. Combined with the requirement of two opposite-sign leptons and a large imbalance in the transverse momenta of the collision products (originating from the almost-invisible neutrinos) yields a very rare and characteristic signature that allows for a dedicated search for the VBF Higgs production in the $H \to W^{\mp} W^{\pm (*)} \to \ell^- \bar{\nu} \ell'^+ \nu'$ decay channel.

This thesis is structured as follows: Chapter 2 presents a brief introduction to the Standard Model of Particle Physics and the Higgs Mechanism of Electroweak Symmetry Breaking. Chapter 3 contains a description of the Large Hadron Collider and the ATLAS Detector, the experimental setup that provides the data used for the studies presented. Some peculiarities and general issues associated with the acquisition and the processing of this data are briefly explained in Chapter 4. The signal process for which the present analysis seeks to present evidence and potential sources of background are detailed in Chapter 5. Chapter 6 gives an overview over the data samples and the respective data taking conditions and also lists the Monte Carlo simulated samples employed to obtain background predictions. Chapter 7 proceeds by presenting and applying a method to improve the exploitation of these simulated samples in order to increase the validity and statistical significance of the background predictions. A simplified data analysis based on rectangular cuts is presented in Chapter 8. Chapter 9 gives a brief introduction to methods of statistical analysis and provides an approximate estimation formula for discovery significance. In Chapter 10, an optimization technique for the analysis presented in Chapter 8 is motivated and studied, employing the findings from Chapter 9. In Chapter 11, a statistical analysis of the data is carried out, employing the findings from all previous chapters. Chapter 12 presents a summary and conclusions from these studies and an outlook for potential improvements.

2   Theoretical Overview

In this chapter, the Standard Model of Particle Physics is briefly presented, including the particle content as well as the basic concepts of gauge groups and interactions. The Glashow-Salam-Weinberg model of electroweak interactions is introduced, leading over to the Higgs Mechanism of Electroweak Symmetry Breaking.

2.1   Introduction

The Standard Model of Particle Physics (SM) can be formally established in several ways. Some are motivated by the historic course of discovery, others by means of aesthetics or simplicity. The theoretical formalism presented here is using the Euler-Lagrangian description of quantified field densities and follows the argumentation of Ref. [8].

2.1.1   The Euler-Lagrange formalism for free fields

The Euler-Lagrange equation for a Lagrangian density $\mathscr{L}$ of continuously varying fields $\phi(x^\mu)$ is given as

$$\partial_\mu \left(\frac{\partial \mathscr{L}}{\partial \left(\partial_\mu \phi \right)} \right) - \frac{\partial \mathscr{L}}{\partial \phi} \;\; = \;\; 0,$$

where the shorthand ∂_μ is used for $\frac{\partial}{\partial x^\mu}$. In the following, the term Lagrangian will always refer to a Lagrangian density.

Substituting various choices for the Lagrangian $\mathscr{L}$ will yield different well known physical equations. For example,

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

yields the Dirac equation for a free fermionic field ψ of mass m, that is,

$$\left(i\gamma^\mu \partial_\mu - m \right) \psi \;\; = \;\; 0 \tag{2.2}$$

where γ^μ are the Dirac γ matrices, and the shorthand $\bar\psi = \psi^\dagger \gamma^0$ is used. Furthermore, all equations in this chapter use natural units, expressing physical action in units of the reduced Planck constant $\hbar$ and velocities in units of c, the speed of light in vacuum.

2.1.2 Symmetries and gauge interactions

Obviously, Eq. 2.2 is invariant under any global phase (or "gauge") transformation $U(\alpha) = e^{i\alpha}$ such that $\psi \to e^{i\alpha}\psi$. These transformations form the abelian group $U(1)$. By Noether's theorem, this symmetry implies a conserved current, which in this case can be shown [8] to be the electromagnetic charge current density, enforcing that the electrical charge is a globally conserved quantity.

Trying to generalize this approach to transformations of the type $\psi \to e^{i\alpha(x)}\psi$ for any arbitrarily space-time dependent function $\alpha(x)$, it becomes apparent that the Lagrangian is no longer invariant under this type of *local* transformation.

However, local gauge invariance can be enforced by replacing the derivative ∂_μ with the covariant derivative

$$\mathcal{D}_\mu = \partial_\mu + igA_\mu(x) \tag{2.3}$$

where g is the charge of the Dirac particle and the vector field A_μ transforms under $U(\alpha(x))$ as

$$A_\mu(x) \to A_\mu(x) - \frac{1}{g}\partial_\mu\alpha(x).$$

With this covariant derivative, the Lagrangian becomes

$$\begin{aligned}\mathscr{L} &= i\bar\psi\gamma^\mu\mathcal{D}_\mu\psi - m\bar\psi\psi \\ &= i\bar\psi\gamma^\mu\partial_\mu\psi - m\bar\psi\psi - g\bar\psi\gamma^\mu\psi A_\mu.\end{aligned}$$

The gauge field A_μ couples to (i.e. interacts with) the Dirac particle in exactly the same way as the photon field couples to charged particles. The prefactor of the last term can be identified with the electromagnetic current density. In order to regard A_μ as the physical photon field, the corresponding kinetic energy term needs to be added. Due to the desired gauge invariance, it can only involve the field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.4}$$

Combining these insights and setting g to the electron charge yields the Lagrangian of Quantum Electrodynamics

$$\mathscr{L} \;=\; \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu}\psi - m\right)\psi + e\bar{\psi}\gamma^{\mu}\psi A_{\mu} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \qquad (2.5)$$

Since mass terms of the form $\frac{1}{2}m^2 A_\mu A^\mu$ are forbidden by gauge invariance, the photon must be massless.

2.2 The Standard Model of Particle Physics

In analogy to the previous section, the Lagrangian for the strong and electroweak forces can be constructed and used to derive equations of motion for all known physical interactions besides gravity. Several decades of research have led up to what is today considered the Standard Model of Particle Physics. This section will give a brief and by no means complete summary on these insights.

2.2.1 Particle content

The Standard Model distinguishes two fundamentally different types of elementary particles, depending on their spin. *Bosons* have integer spin. They act as force carriers and are representations of gauge fields similar to the photon field A_μ that was introduced to enforce local gauge invariance in the previous section. These interactions and their mediators will be detailed in the next section.

Fermions, on the other hand, have half-integer spin. Due to spin conservation and the Pauli principle, fermions and composite fermionic structures can be stable and thus capable of forming matter.

Some fermions carry colour charge and are referred to as quarks, which carry an absolute electric charge of either two-thirds or one-third. Colourless fermions are called leptons, some of which carry unit charge, whereas neutrinos carry no electric charge and are thus only subject to the weak interaction.

So far, three generations of fermions have been discovered, that is, for each unique combination of quantum numbers, there is not one, but three fermions, carrying exactly these numbers and only differing in mass. However, since the heavier family members can decay into their corresponding lighter siblings, only the lightest fermion generation is stable.

The fermion fields consist of two different chiral components, which are by convention called left- and right-handed. Chiral theories like the theory of electroweak interaction treat the chiral field components differently. The most prominent exam-

Fermions	Leptons			Quarks			
colour charge	none			red, green or blue			
electrical charge	1		0	$+2/3$		$-1/3$	
Generation	Flavour	Symbol	Symbol	Flavour	Symbol	Flavour	Symbol
1.	electron	e	ν_e	up	u	down	d
2.	muon	μ	ν_μ	charm	c	strange	s
3.	tau	τ	ν_τ	top	t	bottom	b
"family" name	charged leptons		neutrinos	up-type quarks		down-type quarks	

Table 2.1: Fermions of the Standard Model

ple of this asymmetry is the right-handed neutrino (or left-handed anti-neutrino) which, if assumed massless[1], does not couple to any known type of interaction and is hence unobservable.

All Standard Model fermions and their quantum numbers are summarized in Table 2.1. For each particle, there is a corresponding anti-particle with equal quantum numbers, but opposite electrical charge and parity.

2.2.2 Gauge groups and interactions

The complete gauge group of the Standard Model[2] is $SU(3)_C \otimes SU(2)_T \otimes U(1)_Y$. From this, particle interactions can be derived similarly to the procedure presented in Section 2.1.2.

Glashow-Weinberg-Salam model of electroweak interactions

The subgroup $SU(2)_T \otimes U(1)_Y$ of the complete Standard Model gauge group contains electroweak symmetry transformations of the form

$$U(x) \quad = \quad e^{i\frac{Y}{2}\alpha(x)+iT\vec{\sigma}\cdot\vec{\beta}(x)},$$

where Y is the hypercharge and T the weak isospin of the fermion field on which the transformation acts. The values of these quantum numbers can be derived

---

[1]The observation of neutrino oscillations requires non-zero neutrino masses. However, the numerical value of these masses is still unknown. So far, only upper bounds have been measured, proving the neutrino masses to be several orders of magnitude smaller than masses of other fermions [9].

[2]The $SU(2)$ group is commonly displayed with the index L in the literature. However, to keep consistency with the choice of T for the weak isospin, the notation $SU(2)_T$ is used for identification throughout this thesis.

from the electric charge and chirality of the field in question[3]. Furthermore, $\alpha(x)$ and $\vec{\beta}(x)$ are arbitrary space-time dependent functions, and $\vec{\sigma}$ is the three-vector of Pauli matrices. In analogy to Eq. 2.3, the corresponding covariant derivative is

$$\mathcal{D}_\mu \;\;=\;\; \partial_\mu + ig_Y \frac{Y}{2} B_\mu(x) + ig_T T\vec{\sigma}\cdot\vec{W}_\mu(x), \tag{2.6}$$

where g_Y is the $U(1)_Y$ gauge coupling, g_T is the $SU(2)_T$ gauge coupling, and B_μ and W_μ^j ($j = 1,2,3$) are the new gauge fields associated with the groups $U(1)_Y$ and $SU(2)_T$, respectively.

In order to allow for $SU(2)_T$ gauge symmetry, all left handed fermions are arranged in doublets with $T = 1/2$, whereas all right handed fermions are singlets with $T = 0$. However, the explicit fermion mass terms are dropped, since the mass of both left-handed isospin components would need to be identical.

The newly introduced gauge fields B_μ and W_μ^j are not indeed physical, but instead mix to the physical fields

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

which correspond to the photon and the weak Z and W bosons, respectively. Here, θ_W is the weak mixing angle, an experimentally determined parameter that connects the two electroweak coupling strengths g_Y and g_T with the electromagnetic unit charge e via

$$e \;\;=\;\; g_T \sin\theta_W = g_Y \cos\theta_W.$$

The free Lagrange density for the fields B_μ and W_μ is given by

$$\mathcal{L}_{\text{free}} \;\;=\;\; -\frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{4}\sum_{j=1}^{3}W_j^{\mu\nu}W_{\mu\nu}^j$$

with

$$B_{\mu\nu} \;\;=\;\; \partial_\mu B_\nu - \partial_\nu B_\mu$$

---

[3]Right handed field components have weak isospin $T = 0$ and $T^3 = 0$, left handed ones have $T = 1/2$ and $T^3 = \pm 1/2$, where those fields with negative electric charge also have negative third weak isospin component. The hypercharge can then be calculated via the Gell-Mann–Nishijima formula $Q = T^3 + 1/2 \cdot Y$.

in analogy to Eq. 2.4 and

$$W_{\mu\nu}^j = \partial_\mu W_\nu^j - \partial_\nu W_\mu^j - g_T \sum_{k,l=1}^{3} f^{jkl} W_\mu^k W_\nu^l,$$

for $j = 1, 2, 3$ and with the structure constants f^{jkl} of the SU(2) group.

Explicit mass terms for the W and Z bosons are again forbidden by gauge invariance. Hence, this model cannot (yet) completely describe the weak interaction. Weak gauge boson masses as well as fermion masses can, however, be introduced by the Higgs mechanism as described in Section 2.3.

Quantum chromodynamics

The subgroup $SU(3)_C$ consists of transformations

$$U(x) = e^{i \sum_{j=1}^{8} \gamma_j(x) \frac{\lambda_j}{2}},$$

acting on colour multiplets, that is, one triplet for each quark flavour. Here, local gauge symmetry can again be obtained by substituting a covariant derivative, this time of the form

$$\mathcal{D}_\mu = \partial_\mu + i g_C \sum_{j=1}^{8} \frac{\lambda_j}{2} G_\mu^j(x)$$

with the gluon fields G_μ^j, the Gell-Mann-Matrices λ_j, and the strong coupling constant g_C. The Lagrange density of the gluon field is

$$\mathcal{L}_{\text{free}} = -\frac{1}{4} \sum_{j=1}^{8} G_j^{\mu\nu} G_{\mu\nu}^j$$

with the gluon field strength tensor

$$G_{\mu\nu}^j = \partial_\mu G_\nu^j - \partial_\nu G_\mu^j - g_C \sum_{k,l=1}^{8} g^{jkl} G_\mu^k G_\nu^l,$$

where g^{jkl} are the structure constants of the SU(3) group.

2.3 Spontaneous symmetry breaking and the Higgs mechanism

From the previous sections, it became apparent that none of the Standard Model particles can have explicit mass terms after electroweak unification. This strongly contradicts experimental evidence for non-vanishing masses of most Standard Model particles. In the following, the Higgs mechanism of spontaneous symmetry breaking is used to generate masses for all massive Standard Model particles, that is, the weak gauge bosons as well as the fermions and also the newly introduced Higgs boson itself.

In order to achieve this, the Higgs field $\phi(x)$ is proclaimed as a new weak isospin doublet of complex scalar fields

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.7}$$

The corresponding Lagrange density is postulated as

$$\mathscr{L} = |\mathcal{D}_\mu \phi|^2 + V(\phi),$$

where the shorthand notation $|\psi|^2 = \psi^\dagger \psi$ has been used. The additional term $V(\phi)$ is the (external) Higgs potential

$$V(\phi) = \mu^2 |\phi|^2 + \lambda |\phi|^4$$
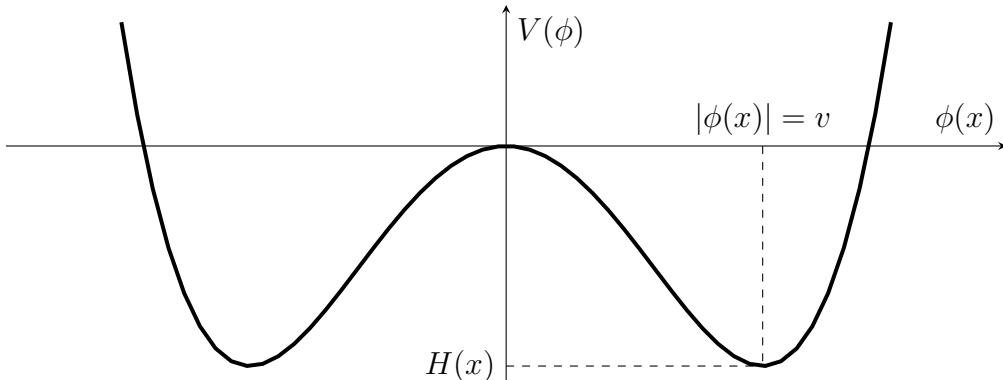
Figure 2.1: Illustration of the Higgs potential shape $V(\phi)$. Shown is a one-dimensional projection onto $|\phi|$. The full shape can be shown in two dimensions, since V only depends on $|\phi|^2$, exhibiting the famous "Mexican hat"-shape.

with two complex parameters μ and λ. The configuration $\mu^2 < 0$ and $\lambda > 0$ has the peculiar feature that $|\phi|^2 = 0$ does not correspond to a local minimum. Instead, a local minimum can be found at $|\phi|^2 = -\frac{\mu^2}{2\lambda} \overset{\text{def}}{=} v$, illustrated in Fig. 2.1. Hence, $\phi(x)$ has a non-vanishing vacuum expectation value ϕ_0. Since V exhibits a rotational symmetry, ϕ_0 can be chosen as

$$\phi_0 \;\; = \;\; \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.8}$$

With this choice, the representation of the Higgs field is

$$\phi(x) \;\; = \;\; \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix},$$

where $H(x)$ is the physical Higgs field.

Inserting the covariant derivative $\mathcal{D}_\mu$ from Eq. 2.6 yields the Lagrangian after spontaneous symmetry breaking

$$\mathcal{L} \;\; = \;\; \frac{1}{4} \left(2\partial_\mu \partial^\mu + \lambda v^2 - 2\lambda v H(x) - \lambda H^2(x) \right) (v + H(x))^2 \tag{2.9}$$
$$+ \frac{g_T^2}{8} \left(2 W_\mu^- W_+^\mu + \frac{Z_\mu Z^\mu}{\cos^2 \theta_W} \right) (v + H(x))^2.$$

2.3.1 The masses of the weak gauge bosons

The terms of primary interest from Eq. 2.9 are those containing only the $W_\mu^\pm$ and Z_μ fields, but not the Higgs field. These are

$$\frac{g_T^2}{8} \left(2 W_\mu^- W_+^\mu + \frac{Z_\mu Z^\mu}{\cos^2 \theta_W} \right) \cdot v^2 \;\; = \;\; \frac{1}{2} \cdot \left(\frac{v g_T}{2} \right)^2 \cdot \left(2 W_\mu^- W_+^\mu + \frac{Z_\mu Z^\mu}{\cos^2 \theta_W} \right),$$

yielding the mass terms

$$m_W \;\; = \;\; \frac{v g_T}{2} \tag{2.10}$$

$$m_Z \;\; = \;\; \frac{v g_T}{2 \cos \theta_W} = \frac{m_W}{\cos \theta_W}. \tag{2.11}$$

However, the $W_\mu^\pm$ and Z_μ fields are now massive, requiring an additional longitudinal degree of freedom. These have been absorbed by choice of Eq. 2.8.

The mixed terms from Eq. 2.9 combining the weak boson fields with the Higgs field describe their corresponding coupling, which one finds to be proportional to g_T as well as m_Z and m_W, respectively.

2.3.2 The mass of the Higgs boson

The terms from Eq. 2.9 that are quadratic in the Higgs field and contain no other fields add up to

$$-\tfrac{1}{2}\left(2\lambda v^2\right)H^2\left(x\right).$$

This is a mass term for a Higgs boson of mass $m_H = \sqrt{2\lambda v^2}$. The value of $v \approx 246$ GeV can be experimentally determined from the known masses of the W and Z bosons using Eq. 2.10. The only missing piece for the Higgs boson mass is then the value of the Higgs self-coupling λ. Theory, however, does not predict the value of λ. This value (or, equivalently, the mass of the Higgs boson) can only be determined experimentally.

On July 4$^{\text{th}}$ 2012, the ATLAS and CMS collaborations have announced the discovery of a Higgs-like boson with a mass near 125 GeV, using data taken at the Large Hadron Collider. This value is in good agreement with exclusion limits derived earlier. At the Large Electron Positron Collider LEP, an exclusion limit of $m_H > 114.4$ GeV was found at 95% confidence level [10]. Electroweak precision measurements at LEP exclude $m_H > 152$ GeV at 95% confidence level [11]. TeVatron searches have excluded a Higgs boson mass in the range $149\,\text{GeV} < m_H < 182\,\text{GeV}$ [12]. General considerations on the consistency of the Standard Model demand a Higgs boson mass below approximately 1 TeV in order to conserve unitarity in $WW \to WW$ scattering [13].

2.3.3 Fermion masses

Fermion mass terms are not an intrinsic part of the Higgs mechanism. They can, however, be generated by postulating Yukawa couplings of the fermions to the Higgs doublet. Considering a left handed fermion doublet L and the corresponding right handed singlet R, the masses for the lower component ψ_L of the doublet can then be generated by extending the Lagrangian by

$$\mathscr{L} \;=\; -g\left(\bar{L}\phi R + \bar{R}\bar{\phi}L\right). \tag{2.12}$$

Substituting the Higgs field after symmetry breaking from Eq. 2.7 and choosing the coupling g such that $vg = m_\psi \sqrt{2}$ then yields

$$\mathscr{L} = -m_\psi \bar{\psi}\psi.$$

The mass terms for the upper component can be generated similarly, only using $\phi_c = -i\sigma_2 \phi^*$ instead of ϕ in Eq. 2.12.

3   The ATLAS Experiment at the Large Hadron Collider

Experiments in High Energy Physics (HEP) nowadays require highly sophisticated and complex instrumentation, usually specifically designed for the particular experiment. Specialized, high-performance instruments like the Large Hadron Collider, introduced in Section 3.1, or the ATLAS detector, introduced in Section 3.2, would not be possible without a tremendous amount of technological research, large-scale industry cooperations and international funding. Due to their immense complexity, only a very brief overview is given at this point.

3.1 The Large Hadron Collider

The *Large Hadron Collider* (LHC) is currently (as of 2013) the worlds largest and highest-energy particle collider, operated by the *European Organization for Nuclear Research* (CERN, fr. *Conseil Européen pour la Recherche Nucléaire*) and located in a centre for particle physics experiments northwest of Geneva, also often referred to as "CERN". The LHC is a circular accelerator with a circumference of approximately 26.7 km [14], located underground in the tunnel of the former *Large Electron-Positron Collider* (LEP), whose dismantling began in 2000 in order to make room for the LHC. A more complete description, which also served as a guideline and source for this brief introduction, can be found in Ref. [14].

The LHC machine consists of 1 232 superconducting dipole magnets with a nominal magnetic field strength of up to 8.3 T. Two beams of protons counter-rotate in the LHC ring in opposite directions. Unlike particle-antiparticle colliders, where the opposing beams can share one ring due to their opposite charge, the high design luminosity of 10^{34} cm^{-2}s^{-1} disfavoured the use of anti-protons. Thus, also due to the limited space in the tunnel, a twin bore magnet (or "two-in-one") design for the superconducting ring magnets was chosen.

Both beams consist of a large number of bunches (up to 2 808 each) of approximately 10^{11} protons each. One of the main features of the LHC is the high kinetic energy of these protons. The collider was designed for a beam energy of 7 TeV, and
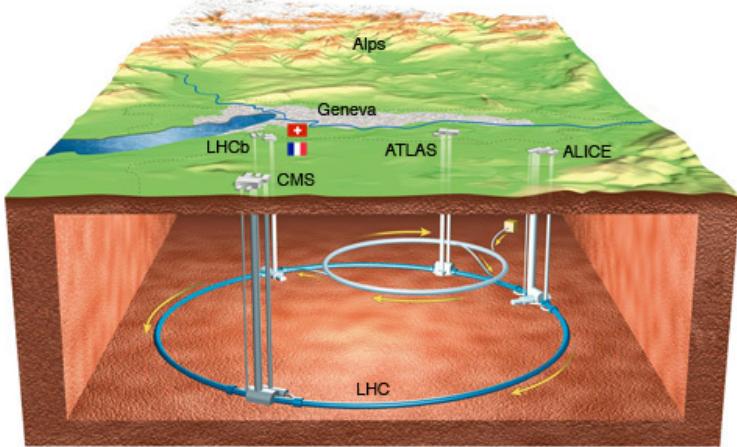
Figure 3.1: CERN overview and LHC tunnel. Taken from Ref. [15].

although this energy will not be reached before 2015, the beam energy of 4 TeV used for data taking in 2012 [16] still makes the LHC the highest-energy man-made particle accelerator ever built.

The LHC is also capable of a second operational mode, acting as a collider for heavy ions. Collisions of lead nuclei with each other or with protons can be induced at a beam energy of 2.76 TeV (per nucleon) and an instantaneous luminosity of $10^{27}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

The LHC has four beam intersection points, around which large detector systems have been built. Each of these experimental sites is operated by an international collaboration of scientists and technicians, and each of the detectors was designed in a unique way in order to fit the needs and special purpose of the respective experiment.

Two of these experiments, ATLAS and CMS, located on opposite sides of the ring, were designed as omni-purpose detectors, whilst the other two, ALICE and LHCb, concentrate on rather specialized fields of research. While LHCb is a single-arm spectrometer with a special focus on *b*-quark activity close to the beam pipe, ALICE focuses on the physics of heavy ion collisions. A number of other, much smaller experiments is sharing the intersection points with the four mentioned above.

3.2 The ATLAS experiment

The ATLAS (*A Toroidal LHC ApparatuS*) experiment is one of the two omni-purpose detector experiments at the LHC. The spectrum of research objectives of the ATLAS collaboration is wide and contains among others the

- precise measurement of the properties of the newly discovered Higgs boson [5] and the exploration of electroweak symmetry breaking,

- experimental validation of the Standard Model of Particle Physics at the TeV energy scale, and

- search for evidence for different "beyond-Standard-Model" theories such as Supersymmetry or the Randall-Sundrum-model [17] predicting the existence of large extra dimensions and microscopic black holes.

More than 3 000 scientists from 174 institutes in 38 countries participate in the ATLAS experiment [18].

The high beam energy and luminosity provided by the LHC provide an excellent environment to measure rare processes like the production of heavy particles. However, the high luminosity comes at the price of a high frequency of collisions, posing a severe challenge for the detector components, event reconstruction algorithms and data analysis techniques. If operated at the design specifications, bunch crossings happen approximately every 25 ns, resulting in a total of 10^9 inelastic scattering events every second due to multiple interactions per bunch crossing.

The ATLAS detector, as depicted in Fig. 3.2, has a nominally forward-backward-symmetric cylindrical layout. Among the main (and name-giving) design features are the three large superconducting toroid magnets, arranged in an eight-fold azimuthal symmetry, surrounding the calorimeters. A brief summary of the detector layout will be given in the following sections. A more comprehensive description, which has also served as the main source for the information and graphic illustrations in this chapter, can be found in Ref. [19, 20].

The ATLAS coordinate-system with the coordinates r, ϕ and η is briefly explained in appendix 6.

3.2.1 Inner detector

The inner detector is dedicated to the measurement of the trajectories and momenta of charged particles. Highly efficient tracking algorithms allow the reconstruction of tracks by connecting information from different layers. As the tracker is immersed in a well-known magnetic solenoid field of 2 T, track curvature information can be used to measure the momenta and electric charges of reconstructed particles and to reconstruct the position of the interaction vertices with great precision, aiding particle identification and disentanglement of tracks from other primary interaction vertices, respectively. The ATLAS inner detector is capable of achieving an intrinsic momentum resolution of $\sigma_{p_\mathrm{T}}^{\mathrm{rel}} = 0.05\% \cdot pT \oplus 1\%$. [19]
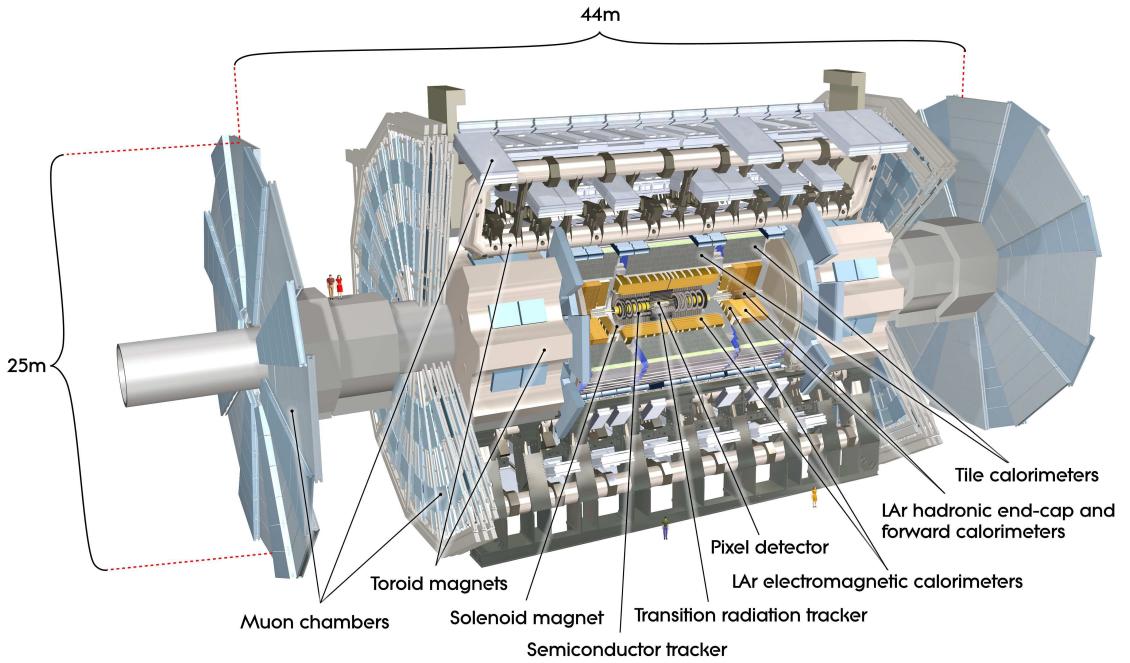
Figure 3.2: ATLAS Detector, length 44 m, radial dim. 25 m, weight approx. 7 000 t.
Reproduced from Ref. [19] with kind permission from IOP Publishing.

Pixel detector

The innermost detector is the silicon pixel detector, consisting of three layers of silicon pixel cells, providing coverage for the region of $|\eta| < 2.5$. The 1 744 pixel sensors have a thickness of 250μm and nominal sizes of $50 \times 400\ \mu$m^2 or $50 \times 600\ \mu$m^2, depending on the position within the containing module. Each pixel sensor yields 46 080 readout channels, amounting to a total of 80 million readout channels for the pixel detector.

The geometry ensures that particles traversing the tracker will typically pass through at least three pixel layers. These charged particles generate roughly 25 000 e^+e^- pairs in the silicon, yielding an electrical signal.

Silicon microstrip detector

Proceeding outward, four double-layers of silicon microstrip (SCT) detectors follow in the barrel region. Each of these 15 912 sensors has a thickness of $285\ \mu$m and consists of 768 active, 12 cm long readout strips. The silicon strip detector yields approximately 6.3 million readout channels.

Combining information from both silicon trackers, reconstruction of the primary and secondary vertices is possible within $|\eta| < 2.5$ for charged tracks way below $p_\mathrm{T} < 500\,\mathrm{MeV}$. However, the latter value is used as a cutoff due to limited computing resources for track reconstruction.

Transition radiation tracker

The third component of the inner detector right before the solenoid is the transition radiation tracker (TRT), a thick layer of polyamide drift tubes, specifically designed to provide robust electron identification up to $|\eta| < 2.0$ and over a wide energy range between $0.5 - 150\,\mathrm{GeV}$. The total number of readout channels is approximately $351\,000$.

The working principle of this detector component relies on the observation that charged particles emit electromagnetic radiation when crossing an interface between two media with different dielectric constants. Since the intensity depends on the particles' Lorentz factor, it can be used to distinguish between particle masses.

3.2.2 Calorimeter system

The calorimeters are located outside the inner detector. The dense absorber material is likely to interact with particles escaping the inner detector and induce showering, forcing the particles to deposit energy in the calorimeter by ionization or excitation of electrons or interaction with the electromagnetic field of the nuclei. Displaced charge carriers in the active material are multiplied and collected by applying a gradient voltage to the ionized material, allowing for a measurement of the energy deposition.

Electromagnetic calorimeter

The inner, electromagnetic calorimeter is divided into a barrel part as well as two end-caps and is designed as a sampling calorimeter with lead as the absorber and liquid argon as the active material. The electrodes as well as the absorber plates offer a unique, accordion-shaped design in order to provide full azimuthal symmetry without interrupting cracks, allowing for several active layers in depth, three in the precision measurement region ($|\eta| < 2.5$) and two in the more outward regions ($2.5 < |\eta| < 3.2$). The central region ($|\eta| < 1.8$) is equipped with an additional presampler layer. The electromagnetic calorimeter achieves an overall energy resolution of $\sigma_E^{\mathrm{rel}} = \sqrt{E/\mathrm{GeV}} \cdot 10\% \oplus E/\mathrm{GeV} \cdot 0.7\%$. [19]

Hadronic calorimeter

The hadronic calorimeter surrounds the electromagnetic one. It uses steel absorbers and scintillating tiles in the barrel region ($|\eta| < 1.7$) and again liquid argon with plates of copper as absorber material in the end cap ($1.5 < |\eta| < 3.2$). In combination with the liquid argon forward calorimeters, a full coverage of the range $|\eta| < 4.9$ is achieved. The hadronic calorimeter provides an energy resolution of $\sigma_E^{\mathrm{rel}} = \sqrt{E/\mathrm{GeV}} \cdot 50\% \oplus 3\%$ in the barrel region. [19]

3.2.3 Muon spectrometer

The outermost detection layer is the muon chamber system, which mainly consists of monitored drift tubes. Bypassing muons ionize electrons within the gas-filled tubes, where a gradient voltage amplifies and collects the electric signal. In the end-cap regions, cathode strip chambers are used. Furthermore, the muon spectrometer contains a separate trigger system, using resistive plate chambers and thin gap chambers.

Magnetic deflection of tracks by the large superconducting toroid magnets allows to measure charge and momentum, similar to the working principle of the semiconductor tracker. Additionally, since most other particle types have already been absorbed by the calorimeter system, the system provides muon identification.

The muon chambers are arranged in three concentric cylindrical shells, providing coverage for the barrel and end-cap regions up to $|\eta| < 2.7$. [19]. The momentum resolution achieved by the muon spectrometers is roughly $\sigma_{p_\mathrm{T}}^{\mathrm{rel}} = 25\% \oplus p_\mathrm{T}/\mathrm{GeV} \cdot 3.3\% p_\mathrm{T}^2/\mathrm{GeV}^2 \oplus 0.17\%$ [21].

3.2.4 Data acquisition and reconstruction

In the previous sections, a brief overview over the physical structure and the functionality of the hardware components of the ATLAS detector was given. However, an equivalently elaborate data acquisition and processing architecture is employed to allow the filtering, storage, reconstruction and management of the data recorded by the detector systems. The components and functionality of that architecture are briefly presented in this section.

Trigger

In full operation, the ATLAS detector is subject to a collision frequency of approximately 400 MHz, while the data storage and offline computing resources are capable of treating event data with a frequency of roughly 300-600 Hz.

This discrepancy is resolved by a highly efficient, multi-layered trigger system. Dedicated hardware components host the Level 1 (L1) triggers which can achieve up to 40 million trigger decisions per second due to direct access to the raw event data recorded by the detector, reducing the event frequency to approximately 75 kHz.

The second and third filtering layers are formed by the Level 2 Trigger system (L2) and the Event Filter (EF), which are hosted on standard computing facilities. The Level 2 trigger uses Level 1 Trigger information to identify regions of interest and performs a refined analysis of the associated detector regions, reducing the event rate to approximately 2 kHz.

The Event Filter as a final step employs reconstruction and identification algorithms similar to those used in the offline analysis, which will be described in the following section. Using these complex decision algorithms, the event rate can be effectively reduced below the critical threshold.

It is worthwhile to note that some triggers achieve an additional reduction in event rate by picking events at random with a certain probability. This technique is commonly referred to as *trigger prescaling*.

3.2.5 Reconstruction and identification

Any analysis of experimental data in high energy physics relies on objects like electrons, muons or hadronic jets. The raw detector output, on the other hand, consists of single-point measurements like particle hits in the tracker or the muon system, and energy depositions in the calorimeter cells. The post-processing of this raw data in order to reconstruct and identify physical objects employs complex algorithms. A selection of these algorithms relevant to the analysis presented in this thesis will be briefly presented in the following sections.

Track reconstruction

Tracking information forms a key ingredient to the interpretation of data from other detector subsystems. Track candidates are reconstructed by connecting clusters of particle hits in the pixel and semiconductor tracker layers subsequently, starting with the innermost layer and proceeding outward. These candidates are then refined and purged by resolving ambiguities in the cluster-to-track association and applying track fit quality criteria. After extrapolation to the transition radiation tracker, high-momentum tracks passing certain quality criteria are refitted with a refined detector model.

A second method called backtracking proceeds in the reverse order, from outermost to innermost layer, thereby improving the efficiency for secondary tracks from photon conversions and decays of long-lived particles.

Vertex reconstruction

Using the reconstructed tracking information, the reconstruction of vertices (that is, interaction points of the hard scatter) proceeds by locating the global maximum of the longitudinal impact parameters z_0 over all reconstructed tracks. The exact definition of the parameter z_0 can be found in appendix 6.3. After this maximum is identified as a vertex, new vertices are seeded from tracks incompatible by at least 7σ with all known vertices in an iterative procedure, refitting the track-vertex-associations and vertex positions in each iteration until either all tracks are matched or no additional vertex can be seeded.

Electron reconstruction

Electrons are reconstructed starting from energy depositions in the electromagnetic calorimeter. These are clustered employing a rectangular *sliding-window* technique in $(\phi \times \eta)$-space [22]. These clusters are then associated with reconstructed tracks, preferring tracks with silicon hits and a small ΔR between the cluster position and the extrapolated track impact point on the calorimeter layer.

However, these electron candidates still have a low purity and are likely to be induced not by physical, high-energy electrons from a collision, but from jets or other detector activity. Hence, three sets of electron identification criteria have been defined and are commonly used in various analyses. They are labelled `loose++`, `medium++` and `tight++`, increasing in purity, but decreasing in efficiency in this order [23, 24].

Muon reconstruction

Based on the combined information from the muon spectrometer, the calorimeter system and the inner detector, three categories for muon candidates exist [25].

Standalone muons are constructed from muon spectrometer hits yielding a trajectory compatible with the interaction point. A significant fraction of these standalone muons does not originate from a primary vertex, but from the decay of some heavy hadron in the calorimeter.

Segment tagged muons are seeded by high-p_{T} muon candidate tracks in the inner detector that can be extrapolated to match small track *segments*, consisting of at least two hits in separated layers of the muon spectrometer.

Calorimeter tagged muons are likewise seeded by high-p_T muon candidate tracks in the inner detector, but required to match a deposition in the calorimeter system which is compatible with a minimum ionizing particle[1].

Combined muons are obtained from matching pairs of complete tracks from the muon spectrometer and the inner detector. The kinematic properties of the track are in this case obtained by partially refitting the track to the individual segments, depending on the algorithm used.

Jet reconstruction

Jets are observed as groups of topologically related energy deposits in the calorimeter system, associated with reconstructed tracks of charged particles. The uncertainty on the jet energy measurement is the leading systematic uncertainty for many ATLAS analyses, including the one presented in this thesis.

The reconstruction starts from topological clusters, which are seeded from calorimeter cells with a measured energy deposit that exceeds the medium expected measurement noise by at least 4σ. Subsequently, all surrounding cells with energy deposits exceeding the noise by at least 2σ are added recursively. Finally, one iteration of nearest neighbour cells is added without application of a threshold. [22] The anti-k_T-algorithm [26] is then applied to these clusters to form jets.

Apart from actual particle jets, a number of physical objects and effects can lead to a jet-like energy deposition in the calorimeter, causing *fake jets* [27]. An additional complication arises from the need to attribute jets to vertices, since multiple tracks can contribute to a single jet. Hence, the jet vertex fraction (JVF) of a pair of one jet and one vertex is defined as the total momentum sum of tracks contributing to the jet that are compatible with the respective vertex, divided by the total jet momentum. The JVF is bounded by $0 \leq \mathrm{JVF} \leq 1$, where 0 corresponds to full incompatibility and 1 to full compatibility of the jet to the respective vertex[2] [28]. However, since the JVF relies on accurate track and vertex reconstruction, it can only be sensibly defined in the central detector region, making the handling of forward- and backward-jets especially challenging.

Since jets are composite objects, consisting of multiple hadrons as well as their showering and decay products, the reconstruction of their respective energy based on calorimeter clusters is non-trivial. Different particles may interact with the detector material by various means, and the relative signal yield in the calorimeter as a function of the energy of the particle in general depends on the type of

---

[1]For particles travelling through matter, the mean rate of energy loss depends on the type of matter and particle, but also on the particles' momentum. Minimum ionizing particles or MIPs are particles whose mean energy deposition rate is close to the minimum.

[2]If no track was found to contribute to the jet, it is assigned a jet vertex fraction of JVF$= -1$.

interaction undergone. Some calorimeters are designed to *compensate* these effects and provide a uniform energy scale for jets. The ATLAS calorimeter, however, is non-compensating, and various techniques exist to estimate and apply jet energy scaling. On the one hand, only electromagnetically interacting particles like electrons or photons deposit almost their entire energy in a way visible to the calorimeter, while hadronically interacting particles and especially neutrons deposit a large fraction of their energy invisibly. To correct for these effects, the method of *local hadron calibration* (LC) was introduced [29]. Roughly speaking, the origin of each cluster is estimated based on the cluster shape and position in the detector, and a corresponding *local* correction factor is applied to account for invisible energy. Other scaling schemes exist, but the jets used in the analysis presented here were reconstructed using the approach of *local hadron calibration*.

It is important to note that the jet energy scale (JES) [30] and jet energy resolution (JER) uncertainties are among the leading systematic uncertainties for many physics analysis.

b-Tagging

Usually, the type of parton that initiated a jet cannot be determined reliably. However, due to the relatively long-lived nature of b-hadrons, identification of b-jets is to some extent possible by relying on the precision of vertex reconstruction to disentangle primary scattering vertices from secondary vertices of b-hadron decays. Since b-jets are of great interest for many analyses, a wealth of different algorithms has been developed to classify jets, a procedure referred to as b-tagging. Many of these b-tagging algorithms employ neural networks or boosted decision trees to achieve an optimum of classification power [31].

The analysis presented in this thesis uses b-tagging mainly to reject events with top-quarks. In this case, the MV1 algorithm is used, which employs a neural network combining the results of other b-tagging algorithms [32].

4   Phenomenology at the LHC

This chapter gives details on how theories in particle physics and especially the Higgs Mechanism of Electroweak Symmetry Breaking can be probed using data taken by collider experiments.

4.1 General aspects of proton-proton collisions

Particle colliders have proven extremely successful in the cause of probing and extending our knowledge about the physical processes on the smallest scales accessible to observation, gradually extending the frontier of accessibility towards higher energies, or, equivalently, smaller scales.

The range of observable processes is not only determined by the centre-of-mass energy of the collision and the collision rate, but also by the types of the colliding particles. Hence, each collider and each experiment has a distinct agenda, closely related to the design specifications. Lepton-lepton colliders, for example, typically provide well-defined measurement conditions with known initial and final states, allowing measurements to be conducted with a high precision. Lepton-hadron colliders, on the other hand, allow to conveniently probe hadronic structures. Hadron-hadron colliders, finally, allow to reach much higher energies and event rates at the expense of higher uncertainties due to imperfectly known and highly complex initial and final states, qualifying them as *discovery machines*. The LHC is such a discovery machine, one of the main objectives being the discovery of the Higgs boson. However, due to the sophisticated technology used for the high-profile experiments at the LHC, precision measurements are nevertheless possible.

This section contains a brief overview over the general approach in connecting experimental measurements with theoretical predictions, also mentioning the complications arising from the fact that the LHC is a hadron collider.

4.1.1 Cross sections and luminosity

Particle colliders are experimental setups allowing to probe theoretical predictions. The connection between theory, which typically predicts a cross section σ of a certain physical process, and the measurement of N candidate events for this process is made by the equation

$$dN \;=\; \sigma \cdot L \, dt \tag{4.1}$$

where L is the instantaneous luminosity of the experimental setup, at each point in time given as

$$L \;=\; f \frac{n_1 n_2}{4\pi \sigma_x \sigma_y}. \tag{4.2}$$

Here, n_1 and n_2 are the respective numbers of particles in the colliding bunches, σ_x and σ_y are the root mean square deviation of the transverse beam size in horizontal and vertical direction, and f is the frequency at which bunches collide at the interaction point. Integration of Eq. 4.1 over the entire measurement time yields

$$N \;=\; \sigma \cdot \int L \, dt = \sigma \cdot \mathcal{L}$$

where $\mathcal{L}$ is the integrated luminosity, a measure for the total amount of collision data recorded by some experiment.

4.1.2 Matrix elements

The cross section for the interaction between a set of free elementary particles can be calculated via

$$\sigma \;=\; \int \frac{1}{F} \, |\mathcal{M}|^2 \, dQ,$$

where F is the incident particle flux in the laboratory frame and dQ is a Lorentz-invariant phase space volume element. The scattering matrix element $\mathcal{M}$ can be determined from the Lagrangian of the underlying theory. While it is in general not possible to derive an analytical expression for $\mathcal{M}$, the decrease of the coupling constants α_s towards larger values of momentum transfer Q^2 allow for approximate expressions to be derived from a perturbative expansion in powers of the coupling constants.

However, the above formula is not directly applicable to processes at hadron colliders, since the colliding particles are composite objects, consisting of quarks
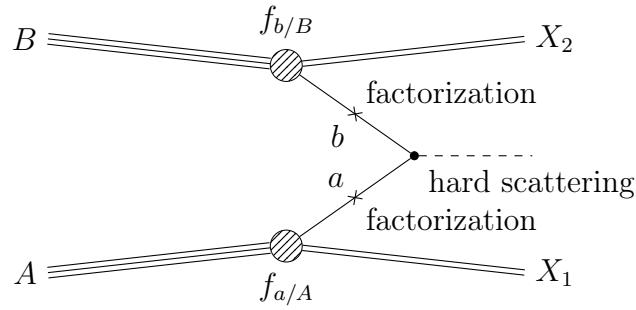
Figure 4.1: Schematic view of the factorization of a high p_T hadron collision into a soft part absorbed into PDFs ($f_{a/A}$ and $f_{b/B}$) and a hard scattering. Adapted from Ref. [33].

and gluons. When considering these collisions, the two-particle interaction with the highest four-momentum transfer is referred to as hard scattering, whereas the remnant of interactions is summarized under the notion of the underlying event.

4.1.3 Factorization

While the calculation of the cross section for such a scattering process involves long-distance, low energy ("soft") effects that are inaccessible for perturbative QCD, the short-distance, high energy interaction of the hard scattering is singled out by factorization, illustrated in Fig. 4.1. In the depicted process, two hadrons A and B collide, where the hard scattering takes place between the partons a and b. Here, the factors $f_{a/A}$ and $f_{b/B}$ reflect the probabilities to encounter a parton a with the required longitudinal momentum within A, and respectively for a parton b in the hadron B. These factors $f_{a/A}$ and $f_{b/B}$ depend on the required longitudinal momentum fractions x_a and x_b of the partons a and b with respect to the hadrons A and B and cannot be derived from first principles, but must be measured in a dedicated experiment, for which lepton-hadron colliders have proven to be important tools. The universal functions $f_{a/A}$ and $f_{b/B}$ are commonly referred to as parton distribution functions. Using this approach, the cross section σ of some interaction involving two hadrons h_1 and h_2 can be calculated as

$$\sigma = \sum_{a,b} \int dx_1 \, dx_2 \, f_{a/h_1}\left(x_1, \mu_F^2\right) \, f_{b/h_2}\left(x_2, \mu_F^2\right) \, \widehat{\sigma},$$

where μ_F denotes the factorization scale and $\widehat{\sigma}$ denotes the parton-level cross section of the hard scattering [33].

decay

hadronization

parton
shower

hard
scattering

parton
distributions
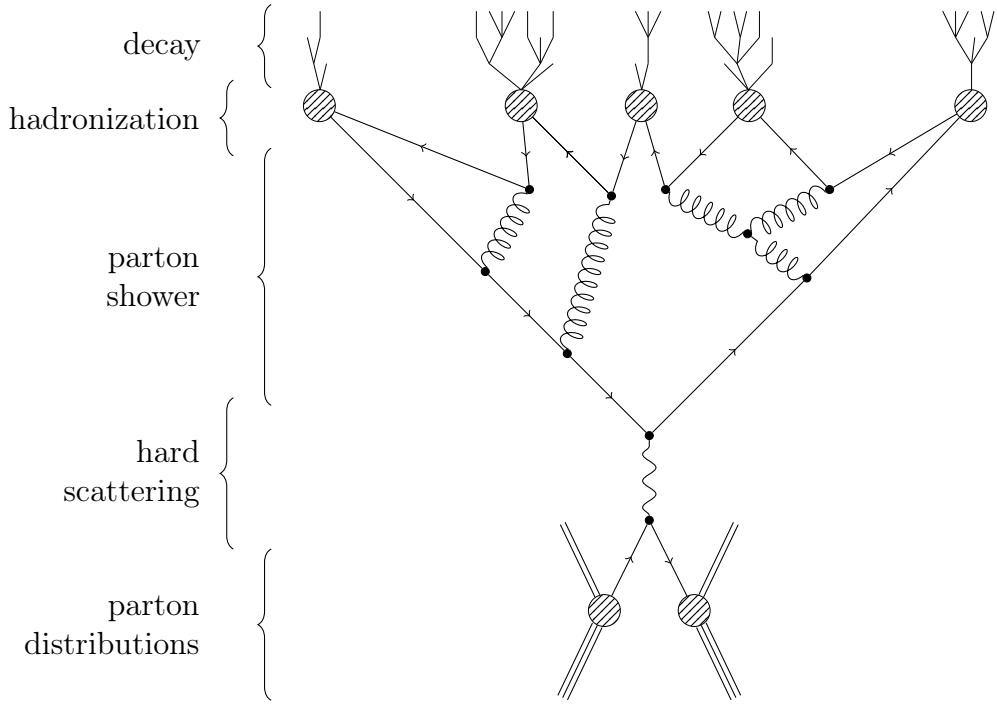
Figure 4.2: Schematic illustration of the basic structure of events simulated with a shower-
ing and hadronization generator. The time evolution of the event goes from
bottom to top. Adapted from Ref. [34].

4.1.4 Hadronization

Coloured objects like quarks and gluons undergo a complex process of hadronization
before becoming observable in the detector as colourless hadrons. This hadroniza-
tion process falls into the regime of soft quantum chromodynamics and is inaccessible
to perturbative calculations. From a phenomenological point of view, however,
the hadronization process can be measured and parametrized, giving rise to ex-
perimentally obtained fragmentation functions. The complete process, including
factorization, hard scattering, hadronization and subsequent decay, is illustrated
by Fig. 4.2.

4.1.5 Underlying event and pile-up

As already mentioned in Section 4.1.2, the hadronic remnants of the colliding
particles are treated separately from the hard scattering, the latter typically being
the main point of interest at discovery machines like the LHC. In fact, the underlying
event contributes to the hadronic background and poses a source of experimental
uncertainties. While the resulting hadronic debris mainly ends up in the forward

and backward detector regions, the proton remnants carry away an unknown amount of longitudinal momentum from the primary interaction vertex. This impedes the interpretation of the absolute longitudinal momentum of the hard scatter.

A second source of hadronic background is the simultaneous collision of other beam particles from the same bunch, referred to as in-time pile-up. Similarly, collision products from the preceding and succeeding collisions are referred to as out-of-time pile-up, the latter becoming important primarily due to the tight bunch spacing in the collider. However, since the collision events are governed by chance and probability, these additional collisions are also beneficial since they increase the instantaneous luminosity and hence the chance of occurrence of "interesting" (rare) events.

4.1.6 Event Simulation

In the context of quantum mechanics and quantum field theory, the occurrence of physical processes is described in a purely probabilistic fashion. Since methodology described in the previous sections, combined with the complex interplay between the final-state particles and the detector, is too complex for an analytical description, a comparison between experimental data and theoretical predictions is only possible on the basis of a finite set of simulated collision events for specific physical processes.

Dedicated computer programs perform the simulation of the various steps mentioned in the previous sections in order to obtain a Monte Carlo sample that is directly comparable to experimental data taken in a real experiment.

Event Generation

The first step in the event simulation is the event generation, based on the calculated cross sections and matrix elements for the process in question. A wide variety of event generators exist, many built in a mutually compatible fashion to enable interfacing them to each other between the different steps of event generation.

- PYTHIA [35, 36] is a general high-energy physics event generator for hard processes, initial- and final-state parton showers, multiple parton-parton interactions, beam remnants, string fragmentation and particle decays.
- HERWIG(++) [37] is a general-purpose Monte Carlo package for **H**adron **E**mission **R**eactions **W**ith **I**nterfering **G**luons with versatile interfacing capabilities and particular emphasis on the detailed simulation of QCD parton showers.

- SHERPA [38] is a Monte Carlo event generator for the **S**imulation of **H**igh-**E**nergy **R**eactions of **PA**rticles that provides complete hadronic final states in simulations of high-energy particle collisions.
- The POWHEG BOX [39] is a **PO**sitive **W**eight **H**ardest **E**mission **G**enerator at next-to-leading order (NLO).
- MC@NLO [40] is a HERWIG-based scheme for combining event generators with NLO QCD rate calculations.
- MadGraph [41] is a versatile matrix element and Feynman diagram generator, customizable to incorporate any theory that can be expressed as a Lagrangian.
- ALPGEN [42] is a leading-order QCD and EW event generator specialized for hard multiparton processes in hadronic collisions.
- AcerMCis a dedicated Monte Carlo event generator for Standard Model background processes in pp collisions at the LHC.
- gg2WW [43] is a parton-level integrator and event generator for processes of the type $gg \rightarrow H \rightarrow WW$ and $gg \rightarrow H \rightarrow ZZ$
- JIMMY [44] is a multi-parton showering extension for HERWIG.
- TAUOLA [45] is a library for τ-lepton decays.
- PHOTOS [46] is a package for QED radiative corrections in decays of resonances.

The generated events also include the additional interactions from pile-up. This is typically achieved by generating a sample of single-interaction, minimum-bias events that are subsequently overlayed with the preselected simulations of specific physical processes to obtain an accurate simulation of pile-up.

Detector simulation and digitization

The GEANT4 [47] (**Ge**ometry **an**d **t**racking) framework allows the simulation of the passage of particles through matter, capable of simulating arbitrary detector geometries and their interactions with the bypassing particles.

In the digitization process, the detector response is modelled to reproduce the conversion of energy depositions in the detector material to digital signals, received by the data acquisition system.

Since the detector simulation and digitization process resembles the computationally most expensive step in the generation of Monte Carlo samples, fast simulation frameworks like ATLFAST-II [48] are being developed.

4.2 Phenomenology of the Standard Model Higgs boson at ATLAS

While the previous section focused on a general overview of the phenomenology encountered at hadron colliders, the methodology of confronting theoretical predictions with measured data from experiments, this section details to some extent the specific phenomenology of production and decay of a Standard Model Higgs boson in the previously presented experimental environment, and especially within the scope of the ATLAS detector.

4.2.1 Higgs boson production and decay at the LHC

As stated in Section 2.3.2, the Higgs boson mass is the only free parameter of the Higgs Mechanism of Electroweak Symmetry Breaking. Once this parameter is fixed, the couplings to all Standard Model particles can be derived from theory. As a notable result, these couplings are proportional to the mass of the corresponding particle. Thus, processes involving heavier particles coupling to the Higgs boson are expected to occur proportionally more often.

The four leading processes of Higgs boson production at the LHC are gluon fusion (ggF), Vector Boson Fusion (VBF), Higgs-strahlung off massive vector bosons (VH) and production in association with a pair of top quarks (ttH). The respective leading-order Feynman diagrams for these processes are depicted in Fig. 4.3. The relative contribution of the different production modes are illustrated in Fig. 4.4 as functions of the hypothesized Higgs boson mass. The total production cross section decreases towards higher values of m_H, with ggF being the dominant production mode for all values of $m_H < 1\,\text{TeV}$. However, the relative contribution of the VBF production mode with respect to the dominant ggF increases towards higher masses. As visible from Fig. 4.3, Higgs boson production via Vector Boson Fusion is accompanied by two additional quarks, which leads to a very distinct event topology with two highly energetic jets with a large rapidity gap, while the decay products from the Higgs boson are likely to be encountered in the central detector region, between the two primary jets. This distinct topology allows for dedicated searches for evidence of VBF Higgs production, which will be detailed in the remaining chapters of this thesis.

The branching fractions of a Standard Model Higgs boson into various decay modes are shown in Fig. 4.5 as functions of the hypothesized Higgs boson mass. The number of available decay modes and their relative branching fraction differs notably depending on m_H, leading to a large number of possible search channels at the LHC, which will be briefly presented in the following.

(a) Gluon fusion (ggF)

(b) Vector Boson Fusion (VBF)

(c) Higgs strahlung (VH)

(d) Associated production (ttH)
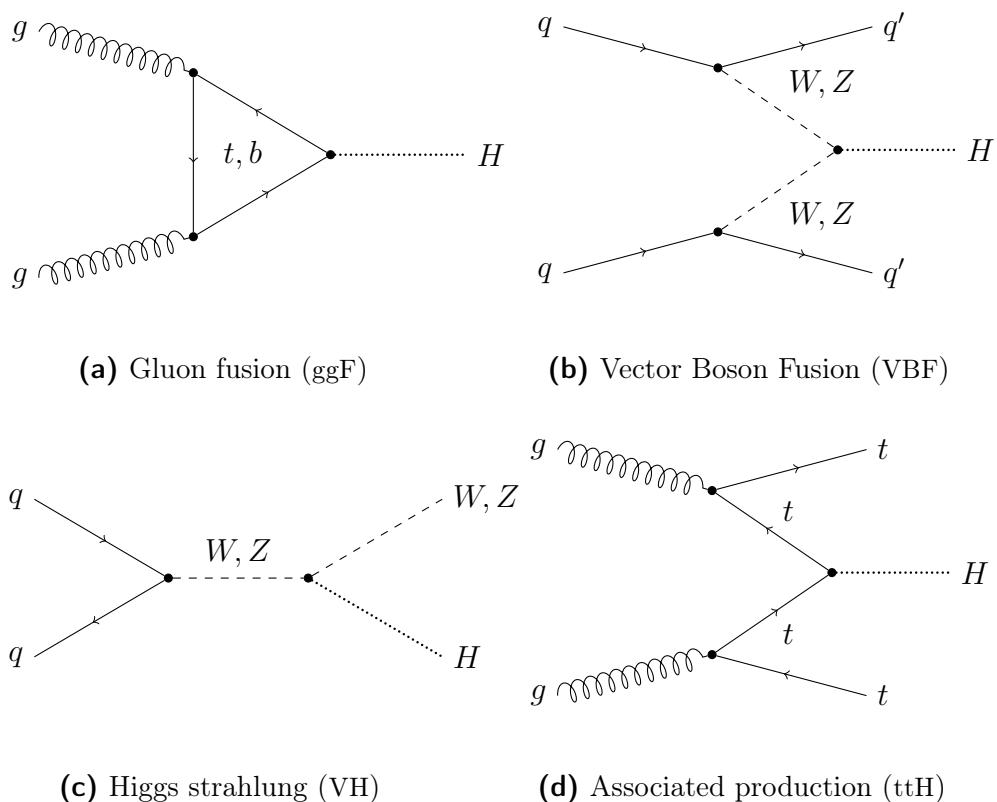
Figure 4.3: Leading order Feynman diagrams depicting the four leading processes of Higgs boson production at the LHC.
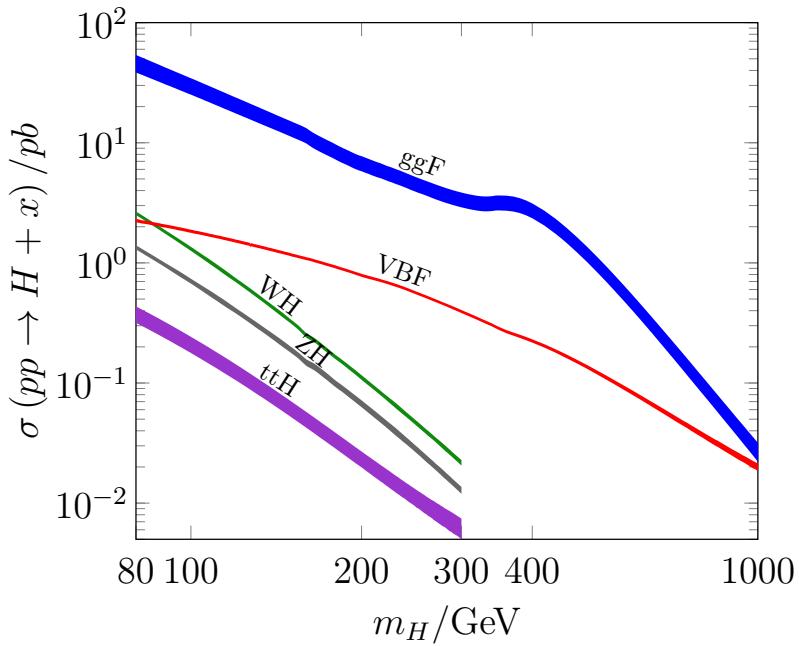
**Figure 4.4:** Standard Model Higgs boson production cross sections at $\sqrt{s} = 8\,\text{TeV}$. Adapted from Ref. [49].
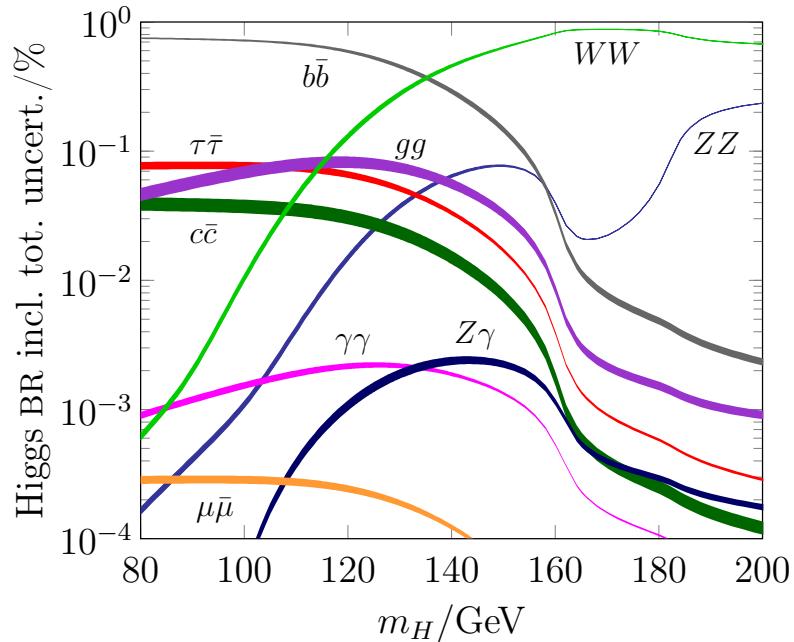
Figure 4.5: Standard Model Higgs boson decay branching ratios as a function of the Higgs boson mass. Adapted from Ref. [49].

4.2.2 Higgs boson search channels and current results

For a Higgs boson mass in the range $120\,\text{GeV} < m_H < 135\,\text{GeV}$, a wealth of different decay channels is available, the dominant decay mode being $H \to b\bar{b}$. However, an analysis in this channel is especially challenging due to the hadronically enriched environment at the LHC. Nevertheless, ambitious analysis are being carried out to reveal potential evidence for a Higgs boson decaying into a pair of two b-quarks. The second highest contribution comes from the $H \to W^+W^{-(*)}$ decay mode, which is the focus of this thesis. Further decay modes are $H \to c\bar{c}$ and $H \to gg$, which are rendered invisible by the large hadronic background. The $H \to \tau\tau$ decay mode poses a challenging analysis as well due to the high probability of taus decaying hadronically, but is intensively investigated. Finally, $H \to ZZ^{(*)}$ and $H \to \gamma\gamma$ provide a very clean signature. The corresponding ATLAS results are currently the leading contributions to the measurement of a Higgs boson and have been continuously updated since the announcement of discovery on July 4th 2012 [5].

Current results include measurements of the Higgs boson mass m_H as well as the spin and parity values J^P. The measured values of m_H and the signal strength μ from the ATLAS and CMS collaborations are listed in Tables 4.1 and 4.2. The signal strength μ is defined as the ratio between the observed and predicted cross sections and is commonly used as a convenient measure to quantify the observed excess. The μ-values fluctuate notably between different search channels and the two experiments, but are compatible with the Standard Model prediction of $\mu = 1$.

	ATLAS combination	CMS combination
m_H	$125.5 \pm 0.2\text{(stat.)}\pm^{+0.5}_{-0.6}\text{(syst.)}$ [50]	$125.8 \pm 0.4\text{(stat.)}\pm0.4\text{(syst.)}$ [51]
$\mu_{\text{comb.}}$	$1.33 \pm 0.14\pm\text{(stat.)}0.14\text{(syst.)}$ [52]	0.88 ± 0.21 [51]

Table 4.1: Summary of recent ATLAS and CMS measurements for the Higgs boson mass m_H and the total observed signal strength μ in combination of all channels.

	ATLAS $H \to W^\mp W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$	CMS $H \to W^\mp W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$
μ_{ggF}	$0.82 \pm 0.24(stat.) \pm 0.28\text{(syst.)}$ [53]	0.79 ± 0.38 [54]
μ_{VBF}	$1.66 \pm 0.67\text{(stat.)}\pm0.42\text{(syst.)}$ [53]	$0.62^{+0.58}_{-0.47}$ [55]

Table 4.2: Summary of recent ATLAS and CMS measurements for the Higgs boson mass m_H and the total observed signal strength μ for the $H \to W^\mp W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ channel, differentiated between ggF and VBF Higgs production modes.
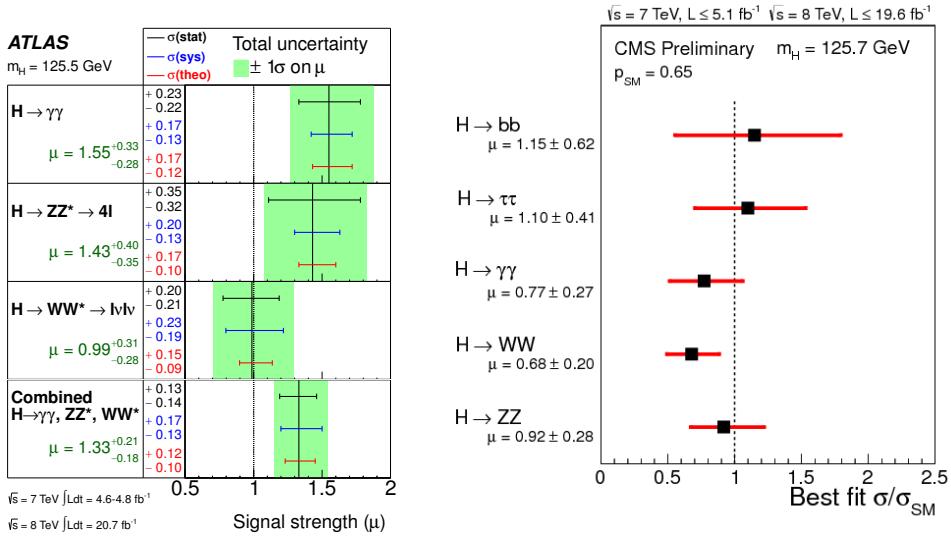
Figure 4.6: Observed and expected μ-values for various Higgs search channels from ATLAS (left, from Ref. [52]) and CMS (right, from Ref. [56], based on Ref. [51]).

Individual μ-values from different channels from both experiments are shown in Fig. 4.6.

The measurement of the spin and parity J^P is somewhat more involved, since the interpretation of the results depend on the assumed production mode. Results are given testing the compatibility of the observed data with various hypotheses. Figure 4.7 shows the ATLAS expected and observed 68% confidence level exclusion limits on hypotheses considered alternatives to the SM prediction of $J^P = 0^+$. Measurements of J^P from CMS only slightly disfavour hypotheses other than 0^+ [51].
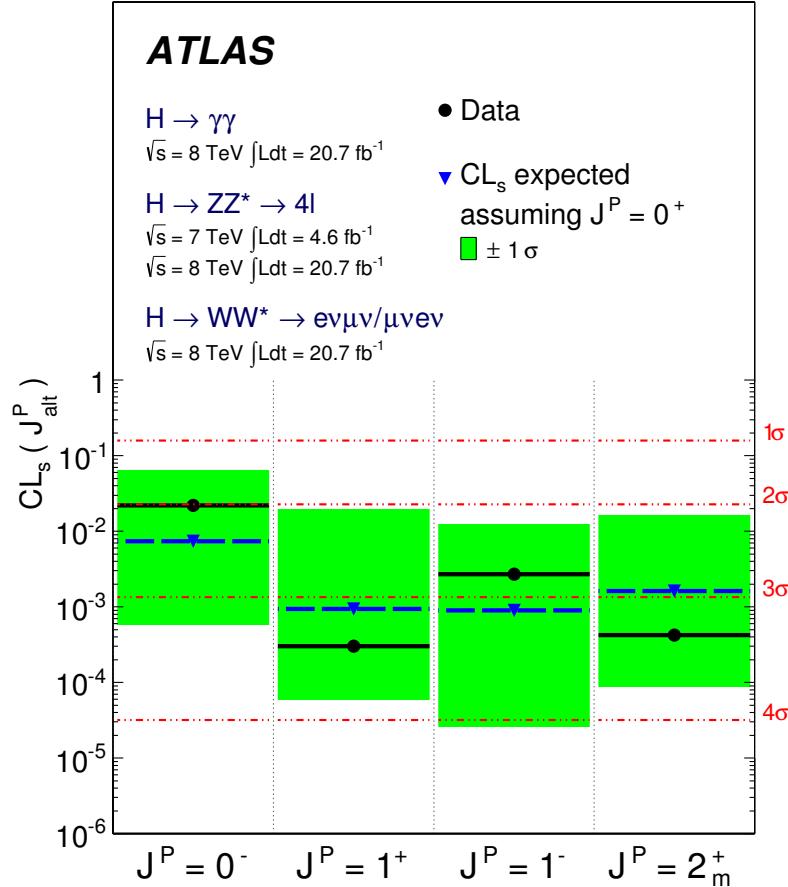
Figure 4.7: Observed (black circles/solid lines) and expected (blue triangles/dashed lines) confidence level CL$_s$ [57] for various spin–parity hypotheses. The green band represents the 68% CL$_s$ expected exclusion range for an assumed 0^+ signal [58]. In the case of spin 0, 100% gluon induced production is assumed, whereas 100% quark-antiquark annihilation is assumed in either spin 1 scenario. For spin 2, a special production scheme of 96% gluon- and 4% quark induced production has been assumed.

5 Signal & Background Processes

This chapter aims at providing a general overview on the physical processes encountered by a dedicated search for production and decay of a Higgs boson with a mass of $m_H \approx 125$ GeV. For the sake of illustration, distributions of simulated collision events over different variables are shown. The technical details of the event simulation used for these plots are described in Chapter 6. The event selections mentioned in the captions of these plots are detailed in Chapter 8.

5.1 The $H \rightarrow W^{\mp}W^{\pm(*)} \rightarrow \ell^{-}\bar{\nu}\ell'^{+}\nu'$ signal process

The different Higgs boson production modes have already been introduced in Section 4.2.1. However, the Higgs boson lifetime is estimated to be short [59], thus only the decay products can be observed. From Figure 4.5, it becomes apparent that the $H \rightarrow WW$ decay mode is the dominant one for intermediate Higgs boson masses beyond $m_H \approx 135$ GeV. Even for lower Higgs boson masses in the range $115 \, \mathrm{GeV} < m_H < 135$ GeV, the decay to a pair of W-bosons is the second most likely mode. Taking into account the discovery announcement on July 4th [5], a dedicated analysis of this decay mode seems justified.

5.1.1 Decay modes of the WW pair

Not only the Higgs boson is expected to be too short-lived for a direct observation, also W bosons as the primary decay products have a lifetime too short to be encountered in the detector directly, giving rise to a wealth of final states, which are detailed and distinguished in this section.

Tables 5.1 and 5.2 illustrate the branching ratios of the W bosons and charged τ leptons that may emanate from these decays. Obviously, the most prominent final state is the fully hadronic one, with a total abundance of approximately 54% with respect to all possible final states of the WW system. However, these purely hadronic final states are desperately concealed within the large hadronic background at the LHC. Alternatively, the fully leptonic and semi-leptonic (mixed) final states can be considered, with relative abundances of approximately 5% and

W decay modes	Fraction Γ_i/Γ in percent
$e\nu_e$	10.75 ± 0.13
$\mu\nu_\mu$	10.57 ± 0.15
$\tau\nu_\tau$	11.25 ± 0.20
hadrons	67.60 ± 0.27

Table 5.1: Decay branching ratios of the W boson [60].

τ decay modes	Fraction Γ_i/Γ in percent
$e\nu_e\nu_\tau$	17.83 ± 0.04
$\mu\nu_\mu\nu_\tau$	17.41 ± 0.04
purely multileptonic	$< 10^{-4}$
purely hadronic	≈ 65

Table 5.2: Decay branching ratios of the τ lepton [60].
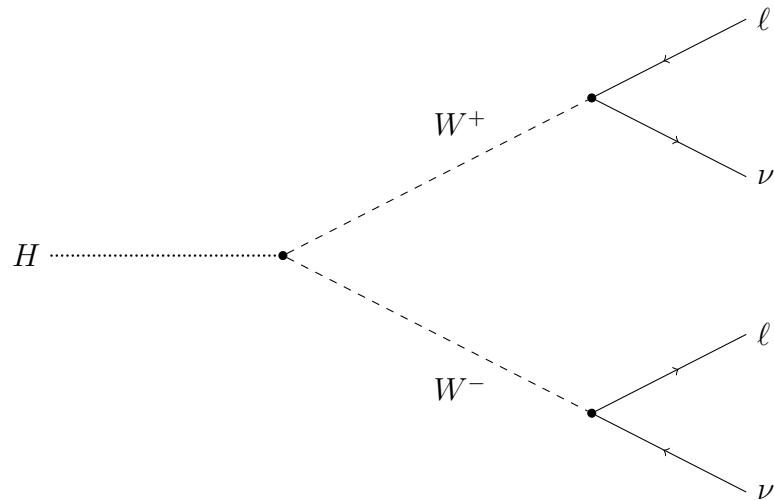
Figure 5.1: Leading order Feynman diagram illustrating the Higgs boson decay in the mode $H \to W^\mp W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$. The final state is characterized by a pair of oppositely charged leptons and a corresponding neutrino/anti-neutrino pair.

41%, respectively. While the latter is more frequent, the former provides a clean and easily recognizable topology. However, while both are the subject of dedicated studies within the ATLAS collaboration, only the former will be elaborated within the scope of this thesis. A Feynman diagram of the $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ decay mode is presented in Fig. 5.1.

5.1.2 Signature of the $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ final state

For many analyses like the search for ggF Higgs boson production, it is customary to assume that the Higgs boson is produced with a small transverse momentum. However, this assumption is clearly not valid when the Higgs boson recoils against other physical objects, as in the case of VBF Higgs production, where transverse momenta of the Higgs boson in the range $50 - 100\,\mathrm{GeV}$ have to be expected [59]. Nevertheless, since the azimuthal direction of the Higgs boson boost is random and the angular correlations detailed in this section are strong, the transverse boost of the Higgs boson is unlikely to alter the angular distribution of the decay products. Hence, for the sake of simplicity of the description, the Higgs boson is in the following sections assumed to be produced at rest. In the case at hand, this approximation is sufficiently accurate for the conclusions drawn to be valid.

Transverse momentum

Using the above approximation, the W bosons are expected to emerge into opposite directions of the transverse ϕ-plain. Due to the large amount of energy contained in each W boson, the comparably lightweight charged leptons in the final state of the $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ decay mode are expected to have a large transverse momentum.

Missing transverse energy

While the above argument holds true for the two final state neutrinos as well, their evanescent interaction probability with the detector material renders any direct measurement of their kinematic properties impossible. However, under the assumption of a small transverse momentum of the Higgs boson, the total sum of transverse momenta of particles emerging from the central interaction vertex allows to indirectly estimate the transverse momentum of the evasive neutrinos. Based on this consideration, missing transverse energy is defined as

$$\vec{E}_\mathrm{T}^\mathrm{miss} = -\sum_i \vec{p}_T^{\,i}$$

where $\vec{p}_T^{\,i}$ denotes the individual transverse momenta of the photons, electrons, muons, jets, and unassociated calorimeter cell clusters. In practice, however, the calculation is often modified to achieve robustness against pile-up induced mismeasurements. Common variants include using only charged tracks originating from the primary vertex ($E_{T,\text{track}}^{\text{miss}}$), or downscaling calorimeter clusters according to their respective (soft-term) vertex fraction ($E_{T,\text{STVF}}^{\text{miss}}$ [61]). Here and in the following, the notation E_T^{miss} (without the vector symbol) is used for the absolute value of the missing transverse energy.

While the missing transverse energy is a quantity of great use for event selection, the fact that neither the individual momenta nor the angular separation between the two neutrinos can be inferred inhibits reconstruction of the kinematic properties of the Higgs boson in question, greatly impeding a precise Higgs boson mass measurement. Nevertheless, means to estimate the Higgs boson mass are presented in Section 5.1.2.

Angular correlations
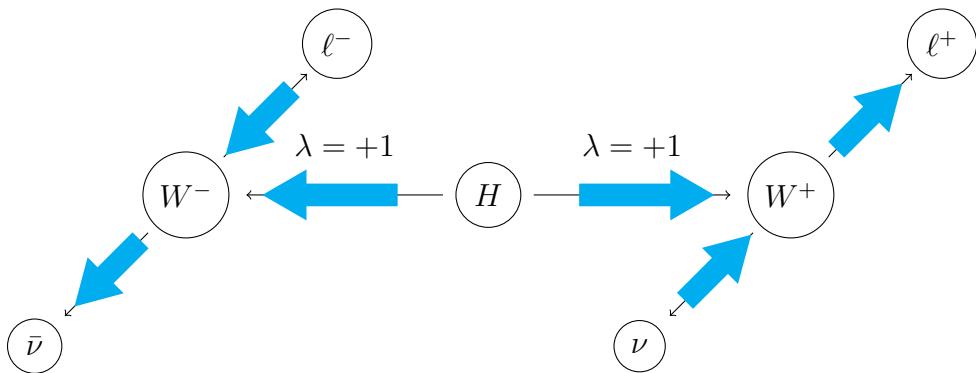
Another important aspect of the event topology is the angular correlation of the emerging particles, namely the leptons and the neutrinos. Due to the parity violating nature of the electroweak theory (see Section 2.2.2) and the scalar nature of the Higgs boson, the neutrinos and the leptons likewise tend to emerge closely aligned, with the corresponding pairs emerging in opposite directions. The topology of these angular correlations is illustrated in Fig. 5.2.

Transverse mass as discriminant variable
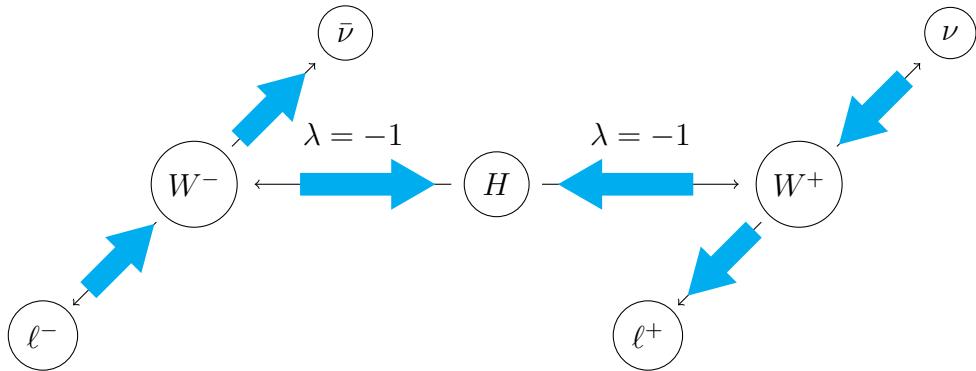
As noted previously, the mass of the Higgs boson cannot be reconstructed from the experimentally accessible event kinematics. However, the transverse mass m_T, defined by

$$m_T^2 \;\; = \;\; \left(E_T^{\ell\ell} + E_T^{\text{miss}}\right)^2 - \left|\vec{p}_T^{\,\ell\ell} + \vec{p}_T^{\,\text{miss}}\right|^2,$$
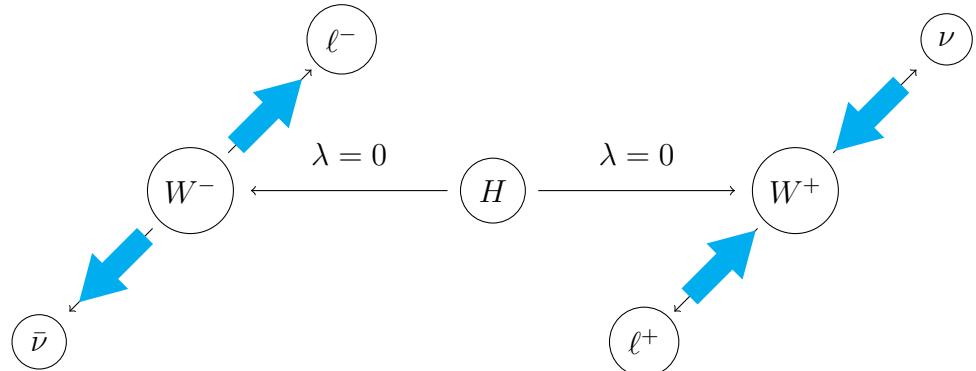
proves to be a highly efficient discriminant variable. Not only is it useful to implement an efficient event selection, simulated m_T-shape templates for different Higgs boson mass values also allow some degree of discrimination between the respective mass hypotheses. Here, the quantities $E_T^{\ell\ell} = m_{\ell\ell}^2 + p_T^{\ell\ell\,2}$ and $\vec{p}_T^{\,\ell\ell}$ denote the kinematic quantities of the combined dilepton system.

(a) Both W bosons have positive helicity, forcing the neutrinos and the leptons likewise to be emitted into the same hemisphere.

(b) Both W bosons have negative helicity, forcing the neutrinos and the leptons likewise to be emitted into the same hemisphere.

(c) Both W bosons have a vanishing longitudinal spin component and thus no helicity. The angular distribution of the decay products is random.

Figure 5.2: Illustration of angular correlations in the WW-system, based on the assumption of a scalar Higgs boson produced at rest. The actual measured distributions may be altered by the shift of reference frame.

final state	momentum configuration	denoted as	paraphrase
$e\nu e\nu$	-	ee	same flavour (SF)
$\mu\nu\mu\nu$	-	$\mu\mu$	same flavour (SF)
$e\nu\mu\nu$	$p_\mathrm{T}^e > p_\mathrm{T}^\mu$	$e\mu$	different flavour (DF)
$\mu\nu e\nu$	$p_\mathrm{T}^\mu > p_\mathrm{T}^e$	μe	different flavour (DF)

Table 5.3: Flavour configurations of the purely leptonic final state.

5.1.3 Differentiation into subchannels

The wealth of decay modes of the WW system introduced in Section 5.1.2 allows to subdivide the analysis into various subchannels, depending on the various possible configurations of lepton flavour and kinematics as well as the number of accompanying jets.

Here and in the following, the term *leptons* only refers to light leptons, that is, electrons or muons. The heavier τ leptons have a much shorter lifetime and are in general not directly visible in the detector.

Lepton flavour

The purely leptonic final state can be subdivided in three to four distinct final states, depending on transverse momenta and flavour of the final state charged leptons. For the sake of notation, the charged lepton with the higher transverse momentum (that is, the "leading" lepton in HEP parlance) is mentioned first, the other ("subleading") charged lepton being mentioned second. The differentiation of the four resulting configurations of lepton flavour and transverse momenta is presented in Table 5.3.

Jet multiplicity

In addition to a differentiation by intrinsic and kinematic properties of the leptons in question, the final state can also be differentiated by the multiplicity of jets encountered.

The dominant Higgs boson production mode at the LHC is the gluon fusion mode. Here, the final state of the hard scattering products does not include any leading-order jets. However, additional hadronic activity may result from initial- or final-state radiation, the underlying event or pile-up.

On the other hand, VBF Higgs boson production results in two additional jets. Since these originate from the initial quark pair, they are expected to have large longitudinal momenta of opposite orientation. Thus, a highly energetic jet pair with

a large separation in the longitudinal η coordinate will be used to *tag* a candidate event as a VBF candidate. Consistently, the two causative jets are referred to as *tagging jets*.

As a trivial consequence, events with a jet multiplicity of $N_{\text{jet}} < 2$ are more likely to originate from the dominating ggF Higgs boson production mode, while VBF production contributes significantly for jet multiplicities of $N_{\text{jet}} \geq 2$, as illustrated by Fig. 5.3. However, some VBF events end up in the lowest two jet multiplicity bins due to the imperfect reconstruction efficiency of the detector.
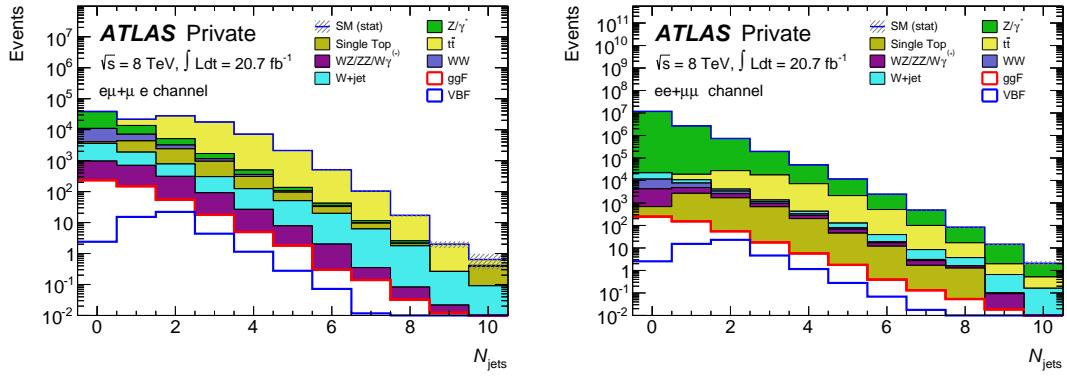
Figure 5.3: Jet Multiplicity distribution in the different flavour (DF, left) and same flavour (SF, right) before the dijet requirement.

A detailed analysis including both jet multiplicity categories, the VBF category with $N_{\text{jet}} \geq 2$ as well as the ggF category with $N_{\text{jet}} < 2$, is presented in Ref. [53].
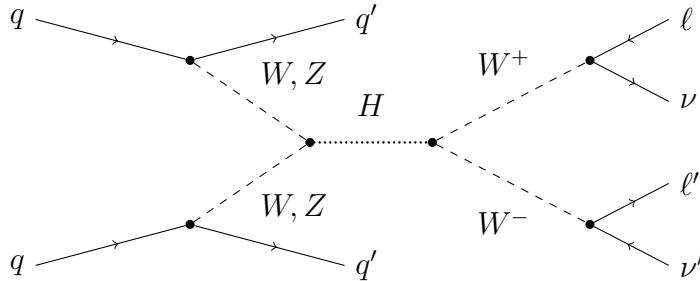
Figure 5.4: Leading order Feynman diagram illustrating the VBF Higgs boson production in a proton-proton collision with decay in the $H \rightarrow W^{\mp}W^{\pm(*)} \rightarrow \ell^{-}\bar{\nu}\ell'^{+}\nu'$ mode. The final state contains the two tagging jets as well as a pair of charged leptons and a corresponding neutrino/anti-neutrino pair.

In this thesis, the analysis of the $N_{\text{jet}} \geq 2$ category is discussed, presenting studies to improve the event selection (see Chapter 10) and the exploitation of the
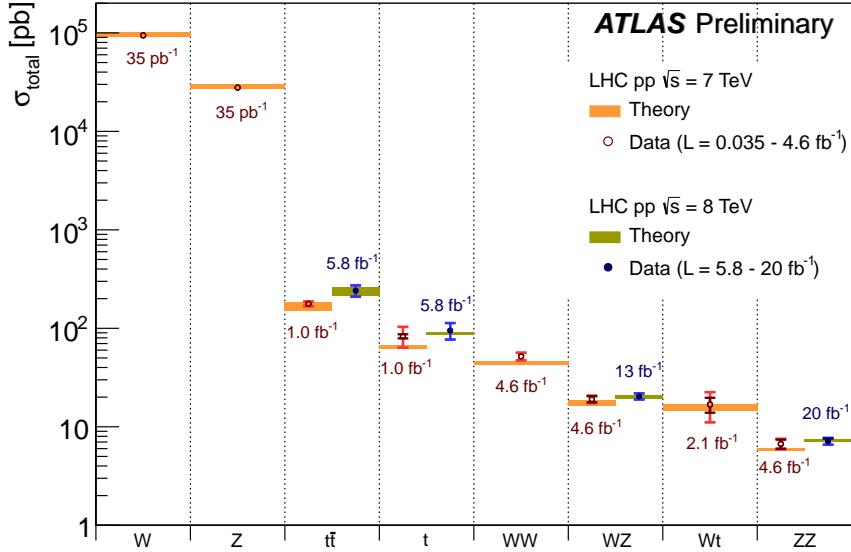
Figure 5.5: Summary of total production cross section measurements and corresponding theoretical predictions at the LHC for several SM processes that contribute as backgrounds to the $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ analysis [62].

finite Monte Carlo samples (see Chapter 7). The leading order Feynman diagram for the production and decay of the Higgs boson is shown in Fig. 5.4.

5.2 Background processes

A number of Standard Model processes is capable of producing a detector signature similar to the one of the $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ decay. Additional jets might either be an intrinsic feature of these processes or originate from initial or final state radiation, the underlying event or pile-up. Hence, these processes contribute as backgrounds to the expected event yield. The most significant contributions are shown in Fig. 5.5. The corresponding processes are discussed in this section. Complying with the general canon, background processes are dubbed irreducible if they result in a final state identical to the signal process, whereas they are labelled reducible if the final state is only experimentally indistinguishable from the signal due to the imperfect identification performance of the experimental setup.

Elaborate techniques have been developed to separate these backgrounds from the signal and to estimate the remaining contribution as precisely as possible. The Monte Carlo generators used to obtain estimates for these backgrounds are presented in Chapter 6. An overview over the event selection and several data-
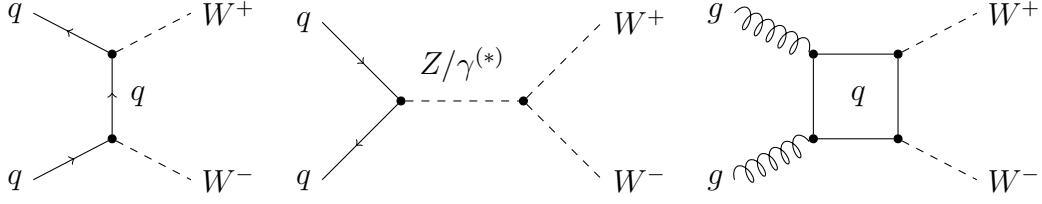
Figure 5.6: Representative leading-order Feynman diagrams illustrating SM processes for WW production in a proton-proton collision.

driven background corrections and estimates are presented in Chapter 8, based on the results presented in Ref. [53].

5.2.1 Standard Model WW production
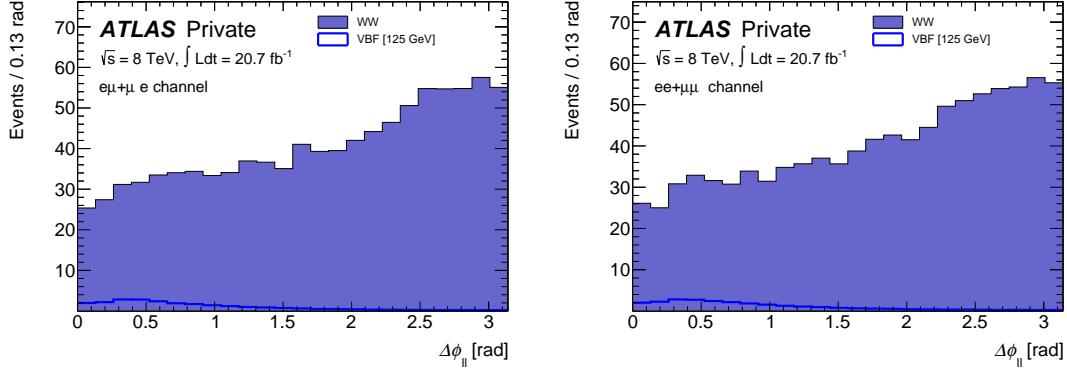
Figure 5.7: Distribution of the dilepton opening angle $\Delta\phi_{\ell\ell}$ in DF (left) and SF (right) after the preselection (see Section 8.2.1).

Production of a pair of oppositely charged W bosons in a proton-proton collision is possible in the Standard Model by various means. Several representative leading-order SM Feynman diagrams are shown in Fig. 5.6. Although the purely leptonic decay mode is disfavoured (as seen in Table 5.1), and the probability of both W bosons decaying leptonically is only about 5%, these events exhibit a final state topology identical to the signal, containing two oppositely charged leptons and a corresponding neutrino/anti-neutrino pair. Standard Model WW production is hence to be considered an irreducible background. However, since the W bosons do not originate from an intermediate scalar particle, the kinematic properties of these events are potentially different, lacking the features of angular correlation illustrated in Section 5.1.2. Furthermore, the WW SM background is mostly resonant, with both W bosons being produced on their respective mass shells, the kinematic

properties of these events are different from the signal, with the subleading lepton on average acquiring a larger transverse momentum. Hence, dedicated cuts on kinematic properties of the dilepton system like the transverse momenta p_{T}^{ℓ} of the leptons as well as the angle $\Delta\phi_{\ell\ell}$ between them allow to efficiently reduce the background contribution arising from Standard Model WW production. Figure 5.7 shows the distributions of the dilepton opening angle for signal and WW background in comparison.

Further complications arise from possible interference effects between WW production via SM processes and via $H \to W^+W^{-(*)}$. The interference arising from the gluon-gluon WW production mode have been shown to reduce the number of signal events in the $H \to W^+W^{-(*)}$ decay channel by up to 10% [43]. However, the impact of these interference effects on the final result of the analysis was found to be negligible [53].

5.2.2 Top quark production

Top quarks are abundantly produced at hadron colliders due to the hadronically enriched environment. Single top quarks can be produced at weak vertices involving a W boson, either as an intermediate exchange particle or in the final state. Top quark pairs can be produced from quark-quark or gluon-gluon interactions, the latter being the dominant production mode at the LHC. Both forms of top quark produced are summarized as *top background*. Representative leading-order Feynman diagrams for both types of processes are shown in Fig. 5.8.

Since top quarks decay in the mode $t \to Wb$ with a branching ratio of more than 99% [60], the W bosons arising from these decays or from Wt associated production can lead to a topology similar to the $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ signal. However, performant b-tagging as introduced in Section 3.2.5 is capable of identifying b-jets originating from the top quark production or decay, allowing for highly efficient background rejection. The distribution of the number of b-tagged jets for top background and VBF signal is shown in Fig. 5.9. Nevertheless, the sheer abundance of these $t\bar{t}$ or single-top events causes a significant fraction to survive the signal event selection.

5.2.3 Z/γ^*+jets production

A pair of oppositely charged leptons with a large transverse momentum can naturally arise from the *Drell-Yan* process (DY) in association with jets, against which the dilepton system can recoil (Z/γ^*+jets). Representative Feynman diagrams for these processes are shown in Fig. 5.10.

(a) Top quark pair production

(b) Single top quark production
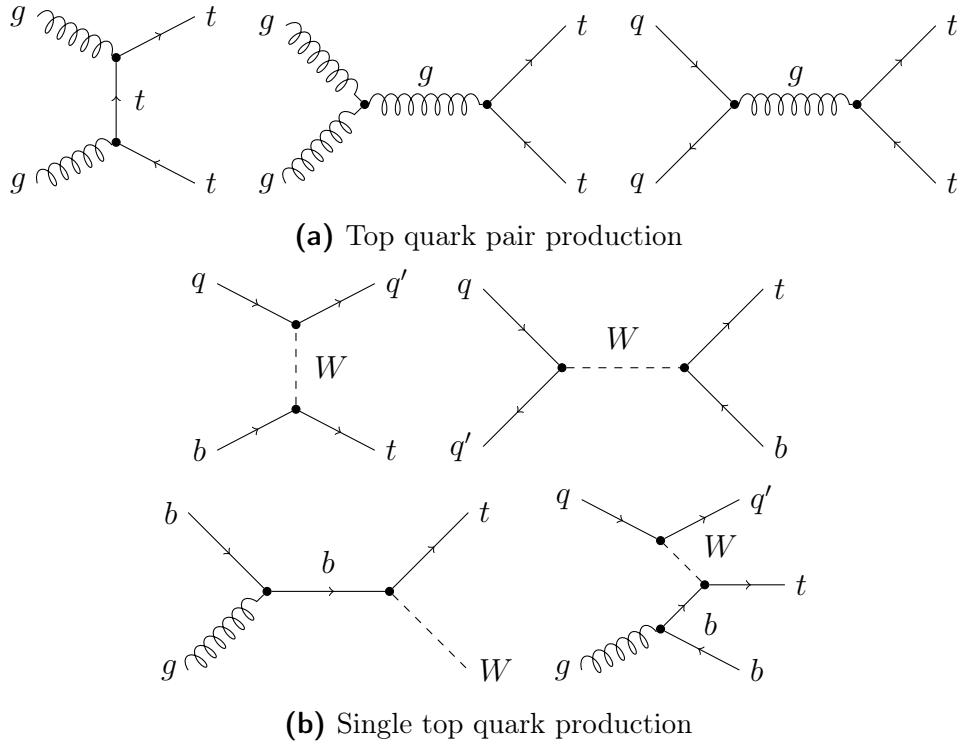
Figure 5.8: Representative leading-order Feynman diagrams illustrating SM production of **(a)** a $t\bar{t}$ pair or **(b)** a single top quark in proton-proton collisions
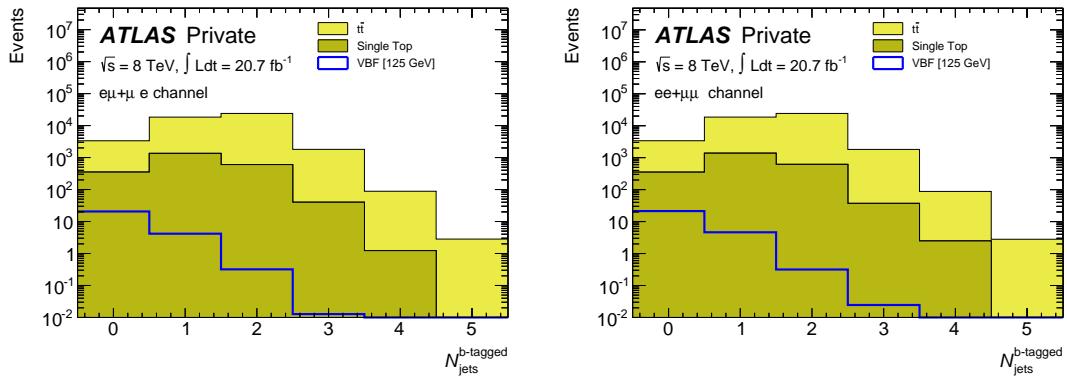

Figure 5.9: Distribution of the number of b-tagged jets for DF (left) and SF (right) after the preselection (see Section 8.2.1).
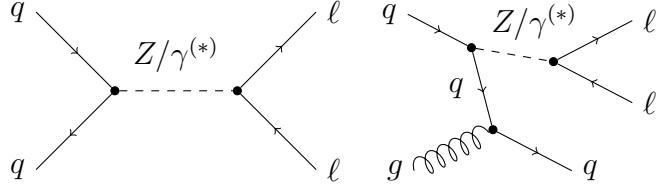
Figure 5.10: Representative Feynman diagrams illustrating the *Drell-Yan* process for the production of a photon or a Z boson singly (left) or in association with jets (right).
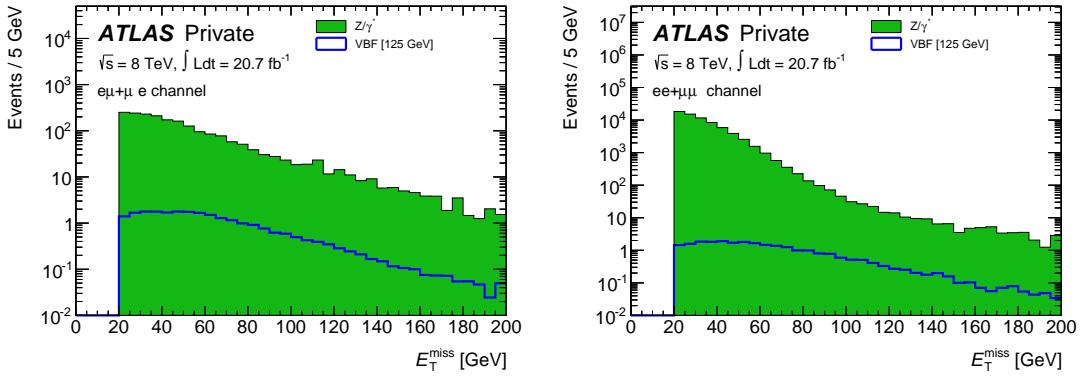
Figure 5.11: Distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ for DF (left) and SF (right) after the preselection (see Section 8.2.1), including an explicit veto on events with $|m_{\ell\ell} - m_Z| < 15\,\mathrm{GeV}$.

This process does not yield an intrinsic component of missing transverse energy, but any mismeasurement of momentum or energy deposition of a physical object such as a jet as well any particles bypassing the detector acceptance can easily lead to a fake $E_{\mathrm{T}}^{\mathrm{miss}}$ component (see Section 5.1.2).

The decay of the Z boson into two oppositely charged leptons can also proceed through $Z \to \tau\tau \to 2\ell 4\nu$. The neutrinos produced in these decays only carry a small amount of transverse momentum because the angular correlations exhibited by these events are quite opposite to the ones of $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ decays. The neutrinos tend to be emitted in opposite hemispheres, causing the missing transverse energy contributions to cancel out. However, since the Z boson is produced in association with jets, it will generally be boosted, forcing the decay products to be aligned more closely and thus leading to a non-vanishing contribution to the missing transverse energy.

Since Z bosons are abundantly produced at the LHC (see Fig. 5.5), $Z+$jets events pose a formidable background to the $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ analysis. In general, the $Z+$jets background can be efficiently reduced by requiring a large
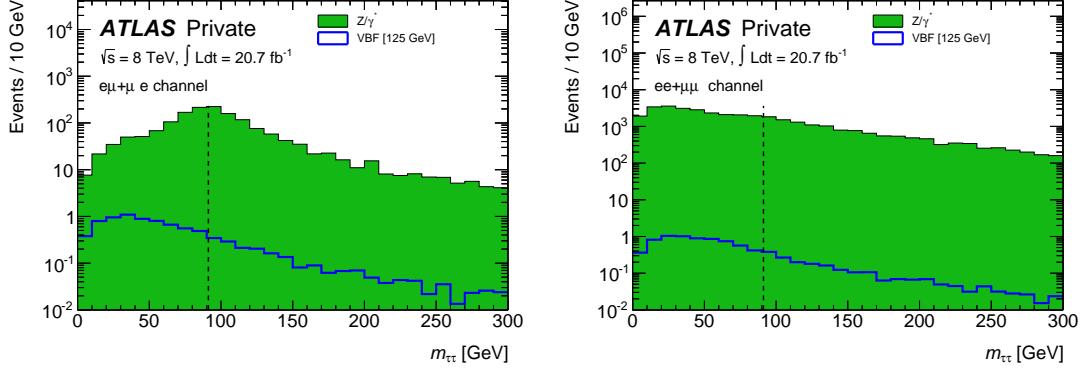
Figure 5.12: Distribution of $m_{\tau\tau}$ for DF (left) and SF (right) after the preselection (see Section 8.2.1), already including an explicit veto on events with $|m_{\ell\ell} - m_Z| < 15\,\mathrm{GeV}$. The dashed line marks the Z-boson mass of $91.2\,\mathrm{GeV}$.
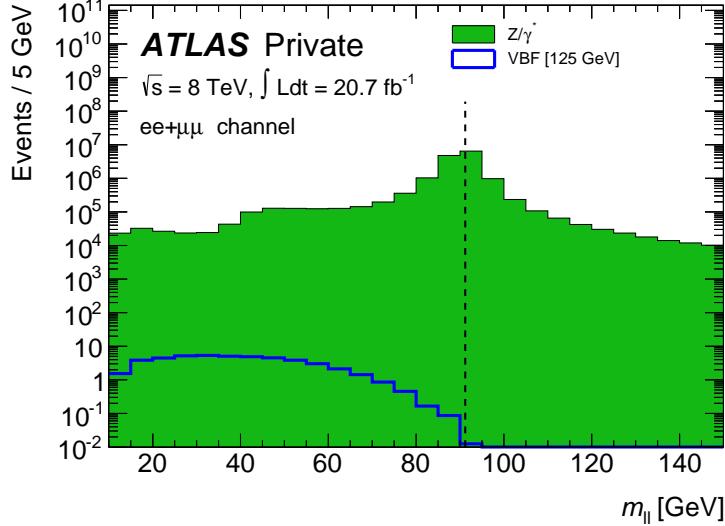
Figure 5.13: Distribution of $m_{\ell\ell}$ for SF after the dijet requirement. The dashed line marks the Z-boson mass of $91.2\,\mathrm{GeV}$.

amount of missing transverse energy. The distributions of missing transverse energy are shown in Figure 5.11.

Additionally, $Z \to \tau\tau$ decays are identified and removed from the event selection by reconstructing the invariant mass $m_{\tau\tau}$ of the $\tau\tau$-system, which can be achieved by either using the collinear approximation [63] for the evading neutrinos, or more elaborate techniques employing knowledge about the Z and τ decay probabilities. The corresponding event distributions are shown in Fig. 5.12.
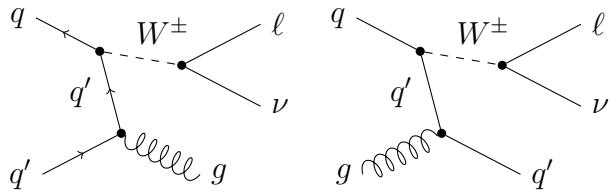
Figure 5.14: Representative leading-order Feynman diagrams illustrating the production
of a W boson in association with jets in a proton-proton collision.

For the same-flavour channel, the Z+jets background is especially challenging,
against which the invariant mass $m_{\ell\ell}$ of the dilepton system is able to discriminate.
For Z+jets events in the same-flavour channel, this variable has a strong peak
around the Z boson mass, which is not present for signal events. Distributions of
$m_{\ell\ell}$ are shown in Fig. 5.13.

5.2.4 W+jets production

Figure 5.14 shows representative leading-order Feynman diagrams for the production
of a W boson in association with a jet. These processes generally do not lead to the
characteristic topology of a pair of oppositely charged leptons with high transverse
momentum. However, the high rate of W boson production at the LHC (compare
Fig. 5.5) yields a large amount of events where the W boson decays leptonically
and one accompanying jet induces the measurement of a lepton. This can proceed
through misidentification of the jet, but is more commonly caused by heavy hadrons
decaying into leptons with sufficient momentum to separate from the remainder
of the jet and pass the lepton isolation criteria. In order to disentangle these fake
leptons, different techniques exist. Identification can be improved by reconstructing
the point of closest approach to the interaction point. Electron identification
can proceed based on shower shapes, and the especially challenging case of fake
electrons from collimated pion jets can be treated by considering the ratio between
the momentum measured by the inner detector and the energy deposition in the
calorimeter. For muons, one can compare momentum information from the inner
detector and the muon system.

Furthermore, the large systematic uncertainties associated with the W+jets back-
ground require efficient suppression techniques. The treatment of this background
is especially challenging since the expected m_{T} shape is similar to the one of the
signal[1]. The corresponding shape predictions from Monte Carlo Simulation are
shown in Fig. 5.15.

---

[1]This shape is generally not well modelled by most Monte Carlo generators, necessitating a data
driven estimation method for the W+jets background, see Chapter 8.
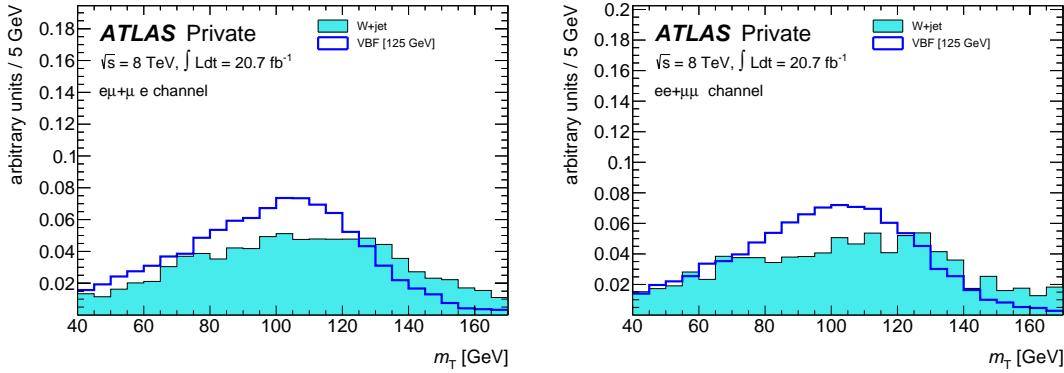
Figure 5.15: Shapes of the distribution of the transverse mass m_T for DF (left) and SF (right) after the preselection (see Section 8.2.1).

While efficient reduction of this background requires highly performant lepton identification and reconstruction, a large fraction can already be suppressed by exploiting the kinematic properties of the event such as the angle between the reconstructed leptons.

5.2.5 $WZ/ZZ/W\gamma$ production

Similarly to the WW background discussed previously, dibosonic background processes like $WW^{(*)}$, $ZZ^{(*)}$ or $W\gamma^{(*)}$ production can yield a signal-like topology via the leptonic decay modes of the bosons.

Events containing WZ or ZZ pairs can exhibit a signal-like topology through leptonic decays of either one or both of the bosons, thus providing isolated charged leptons and missing transverse energy.

The background contribution from the $W\gamma^{(*)}$ process has an m_T shape similar to the expected signal shape and arises from cases where the photon converts to an electron-positron pair.
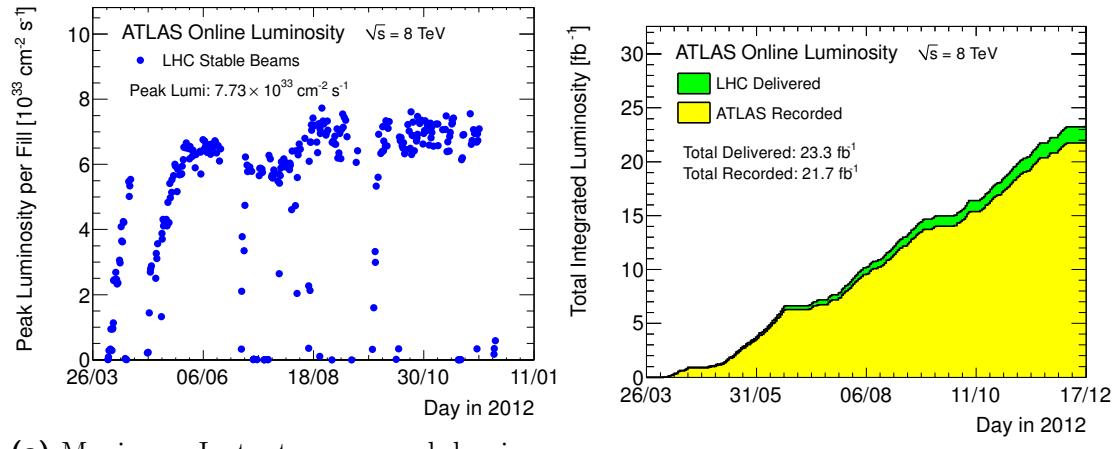
In general, the relatively small production cross section of these dibosonic processes make them less problematic. Furthermore, an efficient background suppression for these modes is possible by rejecting events with three or more leptons, although the finite detector acceptance and reconstruction performance limits the power of this approach.

6 Data and Monte Carlo Samples

This chapter presents an overview over the real and simulated data samples used in the analysis. For the data samples, the trigger criteria are described. For the Monte Carlo simulated samples, the generators used for the different signal and background samples are listed.

6.1 Data samples

The data used in this analysis were taken by the ATLAS detector in the year 2012 at a centre-of-mass collision energy of $8\,\mathrm{TeV}$. It corresponds to an integrated luminosity of $20.7\,\mathrm{fb}^{-1}$ and has been used already for extensive analysis in Ref. [53]. The time evolution of the integrated luminosity during the 2012 data taking is illustrated by Fig. 6.1b.

(a) Maximum Instantaneous peak luminosity versus day, delivered during stable beam periods.

(b) Integrated luminosity delivered and recorded as a function of time in days.

Figure 6.1: Integrated and instantaneous luminosities delivered to and recorded by the ATLAS detector. Taken from Ref. [64].
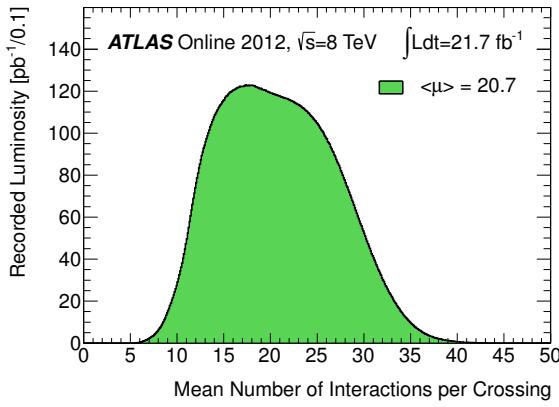
Figure 6.2: Luminosity weighted distributions of the mean number μ of interactions per
bunch crossing for the full data set recorded by the ATLAS detector in 2012.
Taken from Ref. [64].

The data were taken at relatively high values of the instantaneous luminosity L
as illustrated by Fig. 6.1a. As shown in Fig. 6.2, each bunch crossing yielded on
average more than 20 interactions, leading to a large amount of pile-up. During
the 2012 data taking, a bunch spacing of 50 ns was used, leaving empty every
second position in the bunch train (or *bucket* in collider parlance) with the aim of
increasing the total luminosity by increasing the numbers of particles per bunch,
at the price of a higher amount of pile-up.

The data samples were triggered requiring at least one loosely isolated electron
or muon with $p_{\mathrm{T}} > 24$ GeV. The trigger efficiencies were measured as functions of
the pseudorapidity η and the transverse momentum p_{T} using Z candidates. With
respect to the lepton identification criteria described in Section 8.1.3, they are
approximately 70% for muons in the range $|\eta| < 1.05$, 90% for muons in the range
$1.04 < |\eta| < 2.4$ and 90% for electrons [53].

The data analysis in the search for $H \to W^{\mp}W^{\pm(*)} \to \ell^- \bar{\nu}\ell'^+ \nu'$ events originating
from VBF production relies on electrons, muons, the missing transverse energy and
momentum, as well as on hadronic jets and the angular distributions of all these,
requiring virtually every component of the detector system to operate properly.
In the case of any serious detector malfunction during data taking, it is necessary
to exclude the corresponding data from the analysis to ensure an acceptable data
quality. For this purpose, dedicated subgroups within the ATLAS collaboration
exist who monitor the quality of the data taking conditions and publish *good run
lists* (GRL) [65]. This procedure naturally reduces the amount of available data.
However, the data taking conditions in 2012 were remarkable, such that $20.7\,\mathrm{fb}^{-1}$
out of the total 21.3 fb^{-1} achieved the label *all good for physics*, corresponding to
95.5% [66].

Since the data taking conditions typically change from time to time, the data taking is divided into several *periods*, enumerated alphabetically, and further subdivided into different *runs*, which are identified by a six-digit number. The availability of triggers or parameters of the detector operation typically change between different periods, which typically cover several weeks each. For example, period A of the 2012 data taking was from April 4th to 20th. The 8 TeV pp collision runs in 2012 ended with the end of period M on December 16th.

6.2 Monte Carlo-simulated samples

The variety of samples used by the analysis and their corresponding contributions as well as the generators used for simulation are listed in Table 6.1. For most processes, separate programs are used to generate the respective hard scattering process and for the simulation of parton showering (PS), hadronization, and the underlying event (UE). Signal acceptances as well as reconstruction and identification efficiencies are in general obtained from a full simulation of the ATLAS detector using GEANT4 [47], although some background samples use ATLFAST-II [48].

6.2.1 Simulation of Higgs boson production

Simulated samples of Higgs boson production and decay include the dominant gluon fusion process, the Vector Boson Fusion and Higgs-strahlung. The production in association with a top quark pair is neglected due to its smaller cross section. Only the $H \to W^{\mp}W^{\pm(*)} \to \ell^- \bar{\nu}\ell'^+\nu'$ decay mode with electrons or muons in the final state is considered, including the contributions from leptonic tau decays. The Higgs boson decay width and branching fraction are calculated using HDECAY [72] and PROPHECY4F [73, 74], respectively.

The ggF Higgs boson production cross section is calculated including higher order corrections up to next-to-next-to-leading order (NNLO) in QCD [75–80], soft QCD resummations up to next-to-next-to-leading logarithm (NNLL) [81], and electroweak (EW) corrections up to next-to-leading order (NLO) [82, 83]. The calculation of the VBF Higgs boson production cross section involves approximate NNLO QCD corrections [84] and full NLO QCD and EW corrections [85–87]. The cross sections of VH associated production modes are calculated up to NNLO QCD corrections [88, 89] and NLO EW corrections [90].

Process	MC Generator	$\sigma \cdot \mathcal{B}$ [pb]	$\mathscr{L}$ [fb^{-1}]
ggF $H \to W^+W^{-(*)}$	POWHEG [39] + PYTHIA8 [36]	0.44	2 341.2
VBF $H \to W^+W^{-(*)}$	POWHEG+PYTHIA8	0.035	16 610.9
VH $H \to W^+W^{-(*)}$	PYTHIA8	0.013	511.3–2 133.7
$q\bar{q}, gg \to WW$	POWHEG+PYTHIA6 [35]	5.7	~475.4
$q\bar{q}, gg \to WW + 2j$	SHERPA [38] with no $\mathcal{O}(\alpha_s)$-terms	0.039	5 125.6
$gg \to WW$	GG2WW 3.1.2 [43, 67] + HERWIG [37] + JIMMY [44]	0.016	1 271.2–1 442.3
$t\bar{t}$	MC@NLO [40]+HERWIG+JIMMY	240	376.2–2 013.5
	MC@NLO+HERWIG+JIMMY	28	223.4–1 650.1
Single top: tW, tb	AcerMC [68] + PYTHIA6	88	~210
Single top: tqb	ALPGEN [42] + HERWIG (MLM [69]) + JIMMY	16000	9.2–335.0
Z/γ^* (incl.)	ALPGEN+HERWIG+JIMMY (Data-driven estimate)	37000	data-driven
$W^{(*)} \to \ell\nu$ (incl.)	POWHEG+PYTHIA8	0.73	1 785.7–134 328.0
$Z^{(*)}Z^{(*)} \to 4\ell$	POWHEG+PYTHIA8	0.83	377.1–3 029.7
$WZ/W\gamma^{(*)}$ ($m_{Z/\gamma^*} > 7$ GeV)	MadGraph [41, 70, 71] + PYTHIA6	11	394.3–456.1
$W\gamma^{(*)}$ ($m_{\gamma^*} \leq 7$ GeV)	ALPGEN+HERWIG (MLM) + JIMMY	370	128.7–248.9
$W\gamma$			

Table 6.1: Monte Carlo generators used to model the signal and background processes. All W and Z decay channels are included in the corresponding product of the cross section (σ) and branching fraction ($\mathcal{B}$) at $\sqrt{s} = 8$ TeV. Adapted from Ref. [53]. The respective equivalent luminosities are given as ranges or average values when the corresponding contribution consists of several subsamples.

6.2.2 Simulation of Standard Model background

The value quoted in Table 6.1 for the Drell-Yan process excludes the low dilepton invariant mass range of $m_{\ell\ell} < 10\,\text{GeV}$. For $W\gamma \to \ell\nu\gamma$ events, the photon is required to have $p_\text{T} > 8\,\text{GeV}$ and to be isolated from the lepton by at least $\Delta R > 0.25$. For $W\gamma^* \to \ell\nu\ell\ell$ events, at least two leptons must have $m_{\ell\ell} > 2m_e$ and $p_\text{T} > 5\,\text{GeV}$, furthermore $|\eta| < 3$ for same and $|\eta| < 5$ for different flavour.

The interference between $Z^{(*)}$ and $\gamma^\star$ is included for $WZ^{(*)}$ and $W\gamma^*$, with the sample boundary at $m_{\ell\ell} = 7\,\text{GeV}$. For $WZ^{(*)}$, two charged leptons with $p_\text{T} > 5\,\text{GeV}$ and $|\eta| < 2.8$ are required. The $Z^{(*)}Z^{(*)} \to 4\ell$ processes have a $m_{\ell\ell} > 4\,\text{GeV}$. The W and Z bosons are always assumed to decay leptonically. The quoted cross sections include the branching ratios and are summed over lepton flavours except for the top quark production for which the inclusive cross section is quoted.

The cross sections for the $W\gamma$ and $W\gamma^*$ processes are normalized to the MCFM [91] NLO predictions. The normalization factors (K-factors) have been calculated to be 1.15 for $W\gamma$ and 2.01 for $W\gamma^*$ for $0.5\,\text{GeV}< m_{\ell\ell} < 7\,\text{GeV}$, $|\eta|<2.5$ and $p_T > 25\,\text{GeV}$ (15 GeV) for the leading (subleading) lepton, respectively. [53]

For some processes like $t\bar{t}$ and Z+jets background, the analysis uses additional samples with a truth-level filter to select VBF-like events. The filter definition and special treatment of these samples is detailed in Chapter 7.

The parton distribution function (PDF) set from CT10 [92] is used for the POWHEG and MC@NLO samples, and CTEQ6L1 [93] is used for the ALPGEN, Mad-Graph, and PYTHIA6/8 samples.

In some cases, namely for $q\bar{q}, gg \to WW$ and single top processes and the VBF filtered $t\bar{t}$ sample, the ATLFAST-II [48] simulation technique is used. Nature, purpose and use of the VBF filtered samples are detailed in Chapter 7.

6.2.3 Pile-up reweighting

Figure 6.2 shows that each hard scattering event is on average accompanied by approximately 20 additional interactions. As mentioned in Section 4.1.6, a coherent event simulation is typically achieved by randomly overlaying simulated minimum bias events with the simulation of the desired hard scattering event. However, since the exact pile-up conditions are not known until the data taking is complete (and Monte Carlo samples need to be available earlier), the pile-up conditions in the simulated Monte Carlo samples typically do not match those in the data recorded by the detector.

To compensate this shortcoming, the simulated events are *reweighted* according to their number of reconstructed vertices to achieve better agreement between the

respective pile-up distributions. While this procedure is completely valid from a theoretical point of view, it can lead to practical problems since the events in the high pile-up tail of the distribution can acquire large weights, potentially causing large statistical uncertainties. Nevertheless, all Monte Carlo samples used in this analysis are subject to pile-up reweighting to mimic the pile-up conditions in data.

Due to difficulties of the Monte Carlo generators in modelling the complex interplay between the pile-up conditions and the vertex reconstruction efficiency, the pile-up distribution from Monte Carlo is shifted by a *pile-up rescaling* parameter. For the time being, the value of the pile-up rescaling parameter is taken to be $R_\mu = 0.9$. The uncertainty on the precise value of this parameter as well as the general degradation of measurement precision to this extent are reflected by the introduction of an additional systematic uncertainty (see Section 9.2).

6.3 Blind analysis

Any scientific analysis carried out on data that is already present at the time of the analysis is at risk of *overfitting* the data. The term *overfitting* here refers to a situation where an event selection is designed based on the properties of *individual properties* of *given* data events, as opposed to the desirable case of a selection based on *general properties* that are *expected* from signal events on the grounds of theoretical predictions.

In order to reduce the risk of overfitting, the analysis is carried out and optimized *blindly* at first. Here, no signal-like data events are used in the process of optimizing the analysis, and the signal region is only (fully) included after the configuration of the analysis has been fixed.

All optimization studies carried out in this thesis are *blinded* in that they do not consider signal-like data events. However, actual data is used in largely signal-free *control regions* to measure and calculate normalization factors. The corresponding procedures will be presented in Chapter 8.

7   Monte Carlo Sample Merge

Since the production of Monte Carlo samples is computationally expensive, the amount of simulated Monte Carlo events for certain background processes is commonly a limiting factor for optimization studies. This problem is an intrinsic one in the sense that an analysis is considered performant if it is capable of extracting the signal events from the background efficiently. Hence, after the full event selection, only relatively few background events are left, leading to a large relative statistical uncertainty on the remaining background contribution. In this way, the more performant an analysis is, the larger are the amounts of simulated background events it requires.

Since the computationally most expensive step of the sample generation is the detector simulation [94], various attempts have been made to increase the efficiency of the corresponding algorithms, leading to fast-simulation frameworks like ATLFAST-II [48]. Another strategy is to single out and collect signal-like events, employing an analysis-like selection even before the actual detector simulation is carried out, discarding all events failing the selection. This technique is commonly referred to as *filtering* or the production of *filtered* samples.

7.1 Filtered Monte Carlo samples

The efficiency of the filter is arbitrary and can in principle scale linearly with the efficiency of the analysis, which makes this attempt especially promising. However, the implementation of an efficient and unbiased filter is in general not possible since the analysis relies on the properties of *reconstructed* objects, whereas the filter by definition relies on generated *truth* objects. The usage of such filtered samples in any analysis is thus only possible after thorough validation studies and assessment of any possible bias introduced by the filtering algorithm.

Nevertheless, the analysis presented in the following chapter uses samples for the Z+jets and $t\bar{t}$ background processes that are specifically filtered to exhibit a topology similar to events originating from VBF Higgs boson production. The bias introduced by using filtered samples can be eliminated using a *merging* technique, a description and validation of which is presented in the following sections.

7.2 Sample merging

When considering whether to use filtered samples, one would naively have to decide between one of the following scenarios.

- The filtered sample mostly contains events that survive the entire event selection. The event yield in the signal region is large, allowing for a small statistical uncertainty. However, since the filter by definition acts on truth level objects, the sample is potentially biased. While the bias is expected to decrease gradually while applying the event selection, disagreement at early stages might impair general validation and optimization work.

- The unfiltered sample is assumedly unbiased and provides a fully valid background prediction at any stage of the analysis selection. However, the event yield in the signal region is small by construction, giving rise to a large statistical uncertainty.

The choice between the two scenarios can be aided by studies of the filter bias, potentially leading to the introduction of a new systematic uncertainty designed to cover this bias. Under these circumstances, a trade-off has to be decided on between the systematic filter bias uncertainty and the decrease of statistical uncertainty established by the filtered sample. In the following, however, a third possible approach is presented.

- A combination of both samples would provide a large event yield at any stage of the selection, dominated by the contribution from the unfiltered sample at the beginning of the event selection with increasing influence of the filtered sample as the signal selection criteria are subsequently applied.

In order to maintain the validity of the combined sample, events from the filtered sample can be weighted with the filter efficiency ϵ, defined as

$$\epsilon = \frac{n_{\text{passed}}}{n_{\text{total}}} = \frac{n_{\text{passed}}}{n_{\text{passed}} + n_{\text{failed}}}.$$

This approach is fully valid if and only if the two samples are disjoint, that is, no events from the unfiltered sample pass the filtering criteria. In general, this is not true, and so one has to account for the non-empty *overlap* between the two samples. The straight-forward approach is to remove the overlapping events from the unfiltered samples to disjoin the event sets, allowing to successfully merge the samples. The situation is illustrated by Fig. 7.1.

The removal of this overlap is in principle easy to implement, as it is a mere application of the filtering algorithm on the unfiltered sample set. This technique, however, relies on the availability of the truth information at the stage of the posterior filter application. Since the truth information typically corresponds to
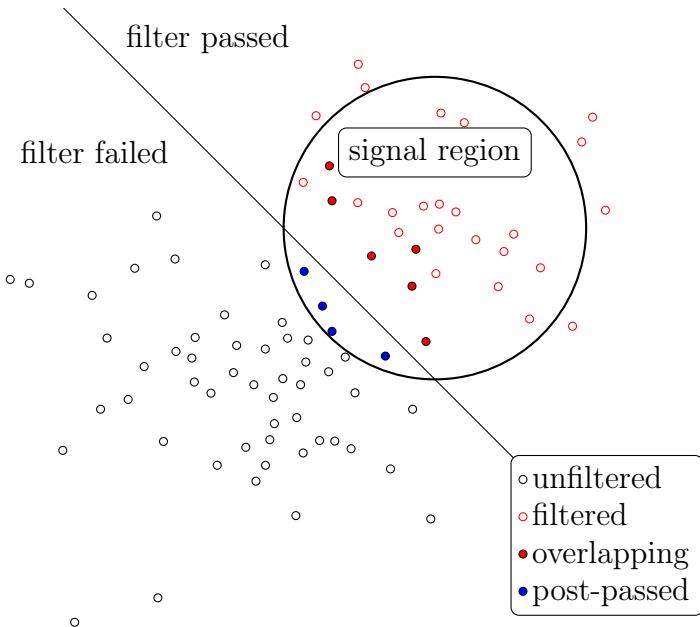
Figure 7.1: Illustration of the merging procedure of filtered and unfiltered samples. The *overlap removal* should remove the events from the unfiltered sample that pass the filter ●, while concatenating all filtered ○ and unfiltered ○ events. In particular, the unfiltered sample includes some events ● that only pass the signal selection due to posterior effects, which will be discussed in Section 7.4.

large amounts of data for each event, channelling the truth information downstream from the sample production stage all the way to the analysis increases the required amount of data storage significantly. However, since the required information can in principle be broken down to the binary *yes/no* decision of the filter, an intermediate approach of executing the filter at an early stage of the data processing and storing the filter decision in a separate variable can provide a solution to this dilemma.

The implementation and validation of the sample merging, including the handling of the filter decision flag, was part of the work performed in association with this thesis. The remaining sections of this chapter present the filter definitions used for the VBF filtered Z+jets and $t\bar{t}$ samples, as well as the corresponding validation of the merging technique including various control plots.

7.3 The VBFForwardJetsFilter

The filter considers truth jets in the pseudorapidity region $\eta < 5.0$ with transverse momenta of $p_\mathrm{T} > 15$ GeV. However, due to the complex nature of hadronization

Pseudo-Jet Removal	truth e, γ or τ with
transverse momentum	$p_\mathrm{T} > 15$ GeV
pseudorapidity	$\eta < 2.5$
distance	$\Delta R < 0.05$
transverse momentum ratio	$p_\mathrm{T}^{e,\gamma,\tau_\mathrm{vis.}}/p_\mathrm{T}^\mathrm{jet} < 0.3$
Jet Selection	genuine jets with
transverse momentum	$p_\mathrm{T} > 15$ GeV
pseudorapidity	$\eta < 5.0$
Filter Requirements	selected jets with
invariant dijet mass	$m_{jj} > 200$ GeV
dijet rapidity gap	$\Delta y_{jj} < 2.0$
not necessarily fulfilled by the same jet pair	

Table 7.1: Configuration of the `VBFForwardJetsFilter` used for the ALPGEN $Z+$jets samples. The configuration used for the MC@NLO $t\bar{t}$ sample described in Section 7.5 uses a different invariant dijet mass threshold of $m_{jj} > 350$ GeV.

(see Section 4.1.4), even truth jets cannot be traced back through the (albeit simulated) hadronization process. Jets might either be genuine in the sense that they directly originate from the hard scatter process, or they may be caused by secondary processes such as hadronic decays of tau leptons or showering interactions of electrons or photons with the detector material[1]. The list of jets surviving the previous cuts is thus purged by removing these *pseudo*-jets.

A jet is considered a pseudo-jet if a truth electron, photon or tau lepton in the pseudorapidity region $\eta < 2.5$ has a transverse momentum of at least $p_\mathrm{T} > 15$ GeV that differs only by at most 30% from the transverse momentum of the jet in question and the distance between both suffices $\Delta R < 0.05$.

Additional complications for the removal of pseudo-jets arise in the case of tau leptons. Since their decays give rise to neutrinos, the momentum of the visible jet might differ from the momentum of the truth tau lepton. In order to accommodate for this effect, only the visible component of the tau lepton four-vector is taken into account for the pseudo-jet matching. Implementation-wise, this is achieved by subtracting the four-vector of all neutrinos from the tau lepton four-vector[2].

---

[1] Considering earlier statements that the use of filtered samples is motivated by the computational expense of the detector simulation, this statement seems contradictory. In fact, even truth level identification of jets requires the previous application of a particle shower generator. These showering generators, however, in general only use a small fraction of the total computation resources consumed by the full detector simulation.

[2] In the context of the work associated with this thesis, the implementation of the `VBFForwardJetsFilter` was found to be erroneous at this point. After the submission of a bug report [95], the calculation has been corrected.

Finally, after the pseudo-jets have been removed from the selection, at least one pair of *real* jets in the event is required to be separated by $\Delta y_{jj} > 2$ in rapidity. Furthermore, one *real* jet pair, although not necessarily the same one, is required to exhibit an invariant dijet mass of $m_{jj} > 200\,\text{GeV}$. For both of these requirements, all possible pairings are considered.

It is worthwhile to note that all numeric values quoted above can be customized when using the filter. The configuration used for the samples at hand is summarized in Table 7.1.

7.4 Alpgen Z+jets sample

The `VBFForwardJetsFilter` was used during the production of an ALPGEN Z+jets sample to select Z+jets background events with a VBF signal like topology.

The sample is split into five subsamples according the number of coloured particles (partons) n_p in the final state of the simulated hard scattering, ranging from $n_p = 0$ to $n_p = 4$. The respective filter efficiencies ϵ and equivalent luminosities of the filtered and unfiltered versions of the subsamples are shown in Table 7.2.

Figure 7.2 shows distributions of the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair, separated by lepton flavour of the Z decay products. The region of phase space selected for these plots is not identical to the signal region. Instead, it corresponds to a Z-enriched control region (defined in Section 8.2.2) to provide larger statistics. Since the event yield of the unfiltered sample is intrinsically poor in the signal region – which was the reason to employ the merging procedure in the first place – the corresponding signal region plots are inconclusive. They are shown in Appendix 4.2. However, Table 7.3 shows the development of the event yield as more cuts are applied, to provide some measure of

		$Z \to ee$			$Z \to \mu\mu$	
n_p	ϵ	$\mathscr{L}$ (unfilt.)	$\mathscr{L}$ (filt.)	ϵ	$\mathscr{L}$ (unfilt.)	$\mathscr{L}$ (filt.)
0	3.35%	9.2	268.5	0.65%	9.3	271.1
1	8.09%	8.6	271.0	3.83%	8.6	267.7
2	24.6%	8.3	261.3	14.1%	8.3	251.9
3	46.9%	7.7	253.0	35.4%	7.7	269.1
4	67.9%	8.0	253.4	59.2%	7.9	238.1

Table 7.2: Filter efficiencies ϵ of the `VBFForwardJetsFilter` as well as corresponding integrated luminosities in units of fb^{-1} for the different ALPGEN Z+jets samples, split by parton multiplicity n_p and lepton flavours of the Z boson decay. The respective uncertainties are negligible.
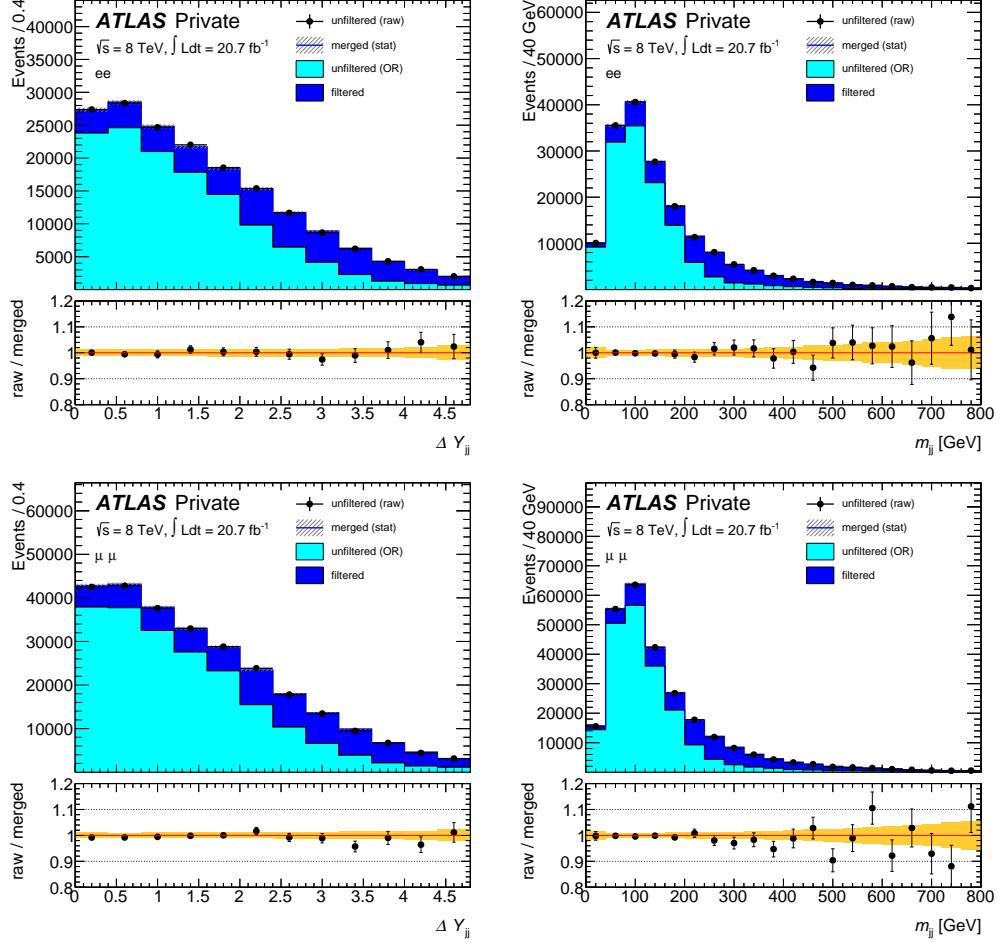
Figure 7.2: Distributions for the dijet rapidity gap Δy_{jj} and the dijet invariant mass m_{jj} of
the leading jet pair in the filtered and unfiltered (raw and overlap-removed)
jet-inclusive $Z \to \ell\ell$ samples for $\ell\ell = ee$ (top) and $\ell\ell = \mu\mu$ (bottom), combined
over all parton multiplicities and normalized to the integrated data luminosity.
The distributions are shown for the Z control region after the preselection
(see Chapter 8). The filtered and overlap-removed unfiltered contributions
are stacked and shown against the raw unfiltered ones. Within the statistical
uncertainties, the respective ratio plots show excellent agreement between the
unfiltered sample before application of the overlap removal and the merged
sum of the filtered and the overlap-removed unfiltered sample.

|	Z control region		signal region	
ee channel	unfiltered	merged	unfiltered	merged
Z-veto/requirement	318128.94 ± 1203.34	318213.85 ± 1042.95	165.13 ± 26.81	148.88 ± 21.72
b-veto	271777.52 ± 1099.16	271991.53 ± 954.93	144.09 ± 24.89	135.35 ± 20.84
$m_{jj} > 600\,\mathrm{GeV}$	5808.47 ± 155.01	5714.61 ± 83.16	5.26 ± 4.77	7.93 ± 4.80
$\Delta y_{jj} > 3.6$	4187.56 ± 130.68	4165.94 ± 75.78	5.26 ± 4.77	6.95 ± 4.79
Central Jet Veto	2241.11 ± 95.68	2254.32 ± 60.76	4.93 ± 4.76	5.99 ± 4.78
Outside Lepton Veto	1835.99 ± 87.28	1845.14 ± 56.67	4.93 ± 4.76	5.84 ± 4.78
$p_{\mathrm{T}}^{\mathrm{tot}} < 50\,\mathrm{GeV}$	1590.79 ± 82.18	1630.70 ± 52.68	4.93 ± 4.76	5.66 ± 4.77
$\mu\mu$ channel	unfiltered	merged	unfiltered	merged
Z-veto/requirement	478712.38 ± 1486.37	481060.53 ± 1303.75	199.73 ± 29.50	210.91 ± 24.54
b-veto	410270.94 ± 1358.63	411906.14 ± 1194.06	156.46 ± 25.96	162.90 ± 20.85
$m_{jj} > 600\,\mathrm{GeV}$	8689.18 ± 187.33	8700.36 ± 107.45	3.74 ± 3.74	9.59 ± 3.85
$\Delta y_{jj} > 3.6$	6501.49 ± 161.60	6497.82 ± 97.99	3.74 ± 3.74	8.09 ± 3.83
Central Jet Veto	3290.79 ± 114.39	3399.01 ± 77.64	0.00 ± 0.00	1.92 ± 0.50
Outside Lepton Veto	2673.00 ± 103.77	2699.30 ± 71.62	0.00 ± 0.00	1.20 ± 0.40
$p_{\mathrm{T}}^{\mathrm{tot}} < 50\,\mathrm{GeV}$	2319.55 ± 97.09	2318.11 ± 64.04	0.00 ± 0.00	1.18 ± 0.40

Table 7.3: Event yields of the raw unfiltered and merged samples in the signal region and Z control region for ee and $\mu\mu$ final states at various cut stages.

the agreement. Since the contributions from these Z samples in the different flavour channels (not including $Z \to \tau\tau$) only arise due to reconstruction and identification deficiencies, they were found to be insignificant and are not shown.

The distribution of the unfiltered sample is shown twice, once in the original (or "raw") state and again after the removal of the overlap with the VBF filtered sample. These distributions show a couple of interesting features, which are discussed in the following.

It is interesting to note the yellow band in the ratio plots at the bottom, which quantifies the relative statistical uncertainty on the event yield from the merged sample, in comparison with the corresponding uncertainty on the event yield from the unfiltered sample alone, depicted by black error bars. Especially towards higher values of Δy_{jj} and m_{jj}, the yellow band increasingly undercuts the black error bars, exposing the total gain in statistical power obtained by merging the samples.

At first sight, one might be surprised that the contribution of the filtered sample is spread out over the full range of the shown variables m_{jj} and Δy_{jj}, since they are included in the filter definition and should naively be restricted to the cut values applied there.

As a possible explanation, one could argue that the filter acts on truth information, while the plots show the variable distributions after the reconstruction stage. While this is certainly true, the spread is too large to be explained with the finite precision of the reconstruction. More importantly, the distributions show the dijet variables based on the leading jet pair, that is, the two jets with the highest transverse momentum. The `VBFForwardJetsFilter`, on the other hand, selects events where any combination of jets passing the quality criteria would pass the selection cuts on their respective dijet variables. Since the jet pair responsible for the event

surviving the filter is not necessarily the leading jet pair, the wide spread of the dijet variables is not entirely surprising.

The apparently large contribution of the unfiltered sample compared to the (statistically more powerful) filtered sample results from the down-weighting of the events from the filtered sample by the efficiencies shown in Table 7.2.

Notably, the overlap removal spares a significant fraction of events from the unfiltered sample even beyond the filter cut values. However, apart from reconstruction effects, this could be explained by the fact that the events in the tail of one distribution (m_{jj} or Δy_{jj}) need not necessarily be in the high tail of the other distribution as well. While this explanation indeed accounts for most of the effect, a small fraction of unfiltered events passes the filter due to posterior effects, as already mentioned in Fig. 7.1. The fraction of such events with respect to the total event yield before the cuts on the dijet variables is a function of the parton multiplicity of the sample, and is with approximately 4% highest in the $n_p = 0$ samples.

In the case at hand, the reason lies in the treatment of pile-up during the simulation process. As stated in Section 4.1.5, pile-up collisions are simulated separately and overlayed with the hard scattering after the full simulation has been completed – much later than the filter needs to become active. Hence, the filter implementation cannot (and does not) take into account truth pile-up jets. Thus, in cases where the filter conditions are met by the event only due to additional pile-up jets, the filter actually rejects the event. These *pile-up passed* events from the unfiltered sample rightfully survive the overlap removal, since this rare type of events is not included in the filtered sample. Naturally, the fraction of such events is highest for samples where the number of jets that originate directly from the hard scattering is small on average.

The pile-up expectation for events in the unfiltered sample is relatively low. Hence, events with a relatively large number of pile-up jets gain a large weight due to the pile-up reweighting, introduced in Section 6.2.3. However, events with a large amount of pile-up jets also have a higher probability of passing the kinematic cuts on dijet variables, leading to a disproportionate fraction of heavily weighted events passing the selection. These events pose a serious problem to the merging procedure as a whole, since their large event weights cause them to dominate the statistical uncertainty compared to events from the filtered sample, which have greatly reduced weights due to the small filter efficiencies. However, up to this point, no efficient solution has been found to tackle this problem.

7.5 MC@NLO $t\bar{t}$ sample

The MC@NLO $t\bar{t}$ sample was filtered using the `VBFForwardJetsFilter` with the m_{jj} threshold shifted to 350 GeV, described in detail in Section 7.3.

Additionally, the `MultiObjectsFilter` was used, which provides a very general filtering interface for arbitrary combinations of truth objects and is highly customizable. In the case at hand, it was used to require at least two truth leptons (electrons, muons, or mixed) with transverse momenta of $p_\mathrm{T} > 5$ GeV in the pseudorapidity range $|\eta| < 10$. The filter efficiency ϵ on this sample is 0.175% with negligible uncertainty. The integrated luminosities of the filtered and unfiltered $t\bar{t}$ samples are 2013.5 and 376.2 fb^{-1}, respectively. The configuration used for the sample at hand is summarized in Table 7.4.

Figures 7.3 and 7.4 show distributions of the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair for the various lepton flavour channels. The region of phase space selected for these plots is a top-enriched control region defined in Section 8.2.2.

Again, since the event yield in the signal region is low, the corresponding plots are only shown in Appendix 5.2. Table 7.5 provides an overview over the event yields of the signal and control regions at various cut stages for the different channels.

Lepton Selection	truth e or μ
pseudorapidity	$\eta < 10.0$
transverse momentum	$p_\mathrm{T} > 5$ GeV
Filter Requirements	two selected leptons

Table 7.4: Configuration of the `MultiObjectsFilter` used for the MC@NLO $t\bar{t}$ sample.
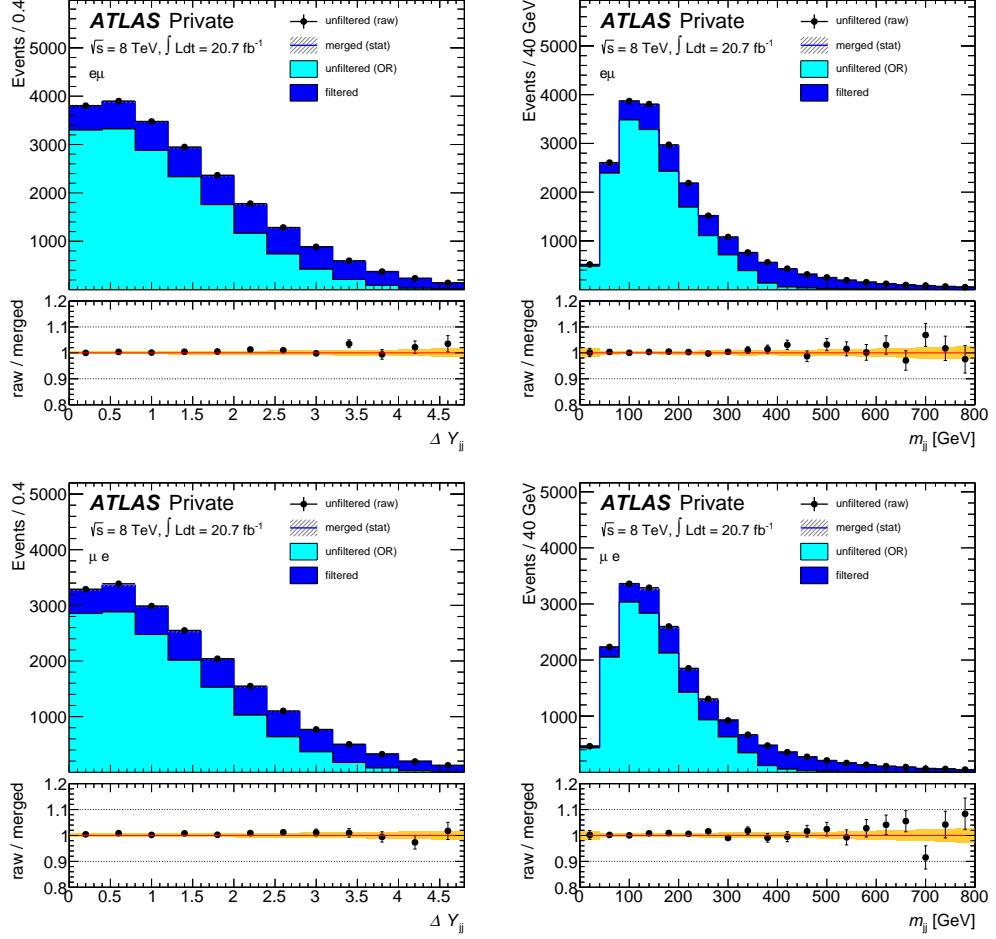
Figure 7.3: Distributions for the dijet rapidity gap Δy_{jj} and the dijet invariant mass m_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $t\bar{t}$ samples for the different flavour channels ($e\mu$ at the top, μe at the bottom). The distributions are shown for the top control region after the preselection (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered ones. The respective ratio plots show good agreement between the unfiltered sample before application of the overlap removal and the merged sum of the filtered and the overlap-removed unfiltered sample within the statistical uncertainties.
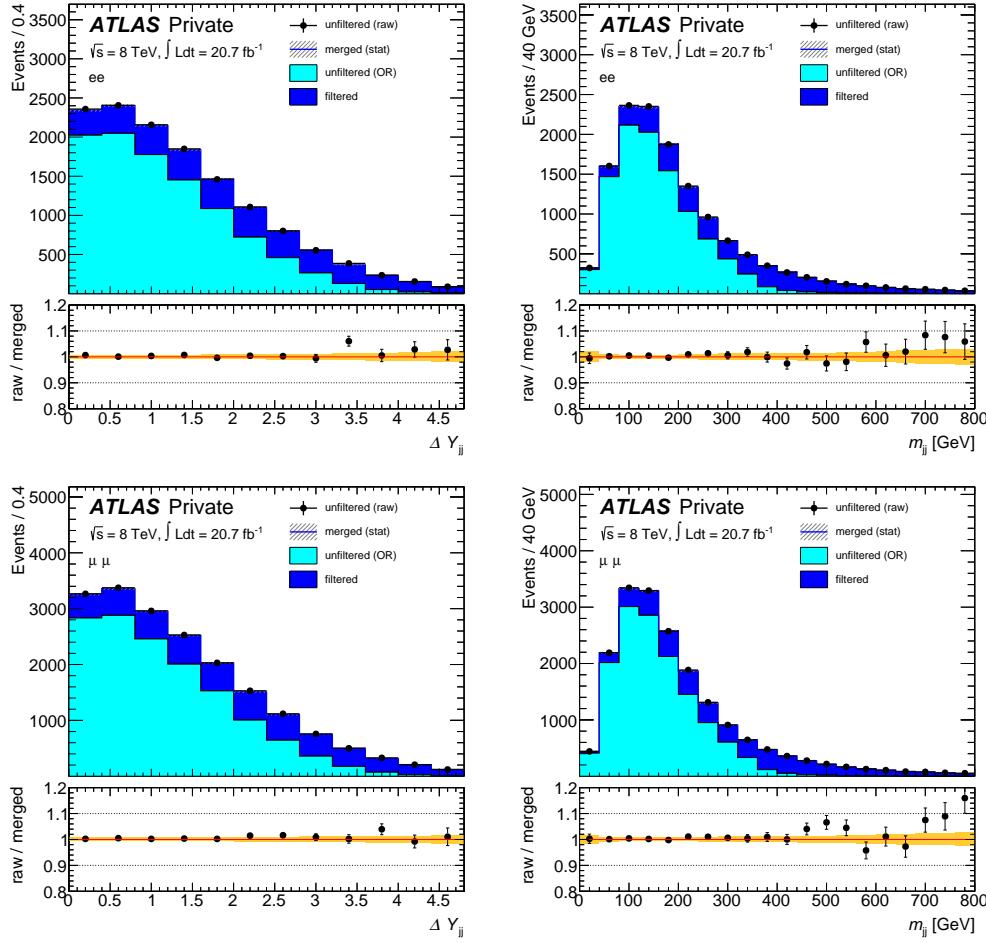
Figure 7.4: Distributions for the dijet rapidity gap Δy_{jj} and the dijet invariant mass m_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $t\bar{t}$ samples for the same flavour channel (*ee* at the top, $\mu\mu$ at the bottom). The distributions are shown for the top control region after the preselection (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered ones. The respective ratio plots show good agreement between the unfiltered sample before application of the overlap removal and the merged sum of the filtered and the overlap-removed unfiltered sample within the statistical uncertainties.

	top control region		signal region	
ee channel	unfiltered	merged	unfiltered	merged
b-jet veto/requirement	13653.63 ± 41.92	13580.62 ± 37.05	1520.16 ± 15.43	1498.84 ± 13.50
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 20\,\mathrm{GeV}, E_{\mathrm{T,STVF}}^{\mathrm{miss}} > 35\,\mathrm{GeV}$	5417.53 ± 26.45	5392.39 ± 23.72	626.16 ± 9.95	616.59 ± 8.83
$m_{jj} > 600\,\mathrm{GeV}$	167.87 ± 4.89	163.50 ± 2.31	28.85 ± 2.06	27.75 ± 0.96
$\Delta y_{jj} > 3.6$	91.99 ± 3.49	92.60 ± 1.65	21.91 ± 1.77	21.39 ± 0.81
Central Jet Veto	19.31 ± 1.51	17.96 ± 0.68	9.12 ± 1.15	8.68 ± 0.52
Outside Lepton Veto	11.37 ± 1.17	10.66 ± 0.52	7.28 ± 1.03	7.60 ± 0.49
$p_{\mathrm{T}}^{\mathrm{tot}} < 50\,\mathrm{GeV}$	10.56 ± 1.13	9.37 ± 0.48	6.20 ± 0.96	6.72 ± 0.47
$p_{\mathrm{T}}^{\ell\ell,\mathrm{jets}} > 25\,\mathrm{GeV}$	10.53 ± 1.13	9.32 ± 0.48	6.20 ± 0.96	6.69 ± 0.47
$Z \to \tau\tau$-Veto	10.52 ± 1.11	9.15 ± 0.47	6.11 ± 0.95	6.55 ± 0.47
$f_{\mathrm{recoil}} < 0.2$	7.97 ± 0.95	6.58 ± 0.40	4.59 ± 0.81	4.91 ± 0.42
$\mu\mu$ channel	unfiltered	merged	unfiltered	merged
b-jet veto/requirement	18857.17 ± 49.80	18752.53 ± 44.15	2108.39 ± 18.25	2074.29 ± 16.08
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 20\,\mathrm{GeV}, E_{\mathrm{T,STVF}}^{\mathrm{miss}} > 35\,\mathrm{GeV}$	7766.88 ± 32.07	7712.90 ± 28.82	908.36 ± 11.96	890.21 ± 10.65
$m_{jj} > 600\,\mathrm{GeV}$	248.19 ± 5.98	233.83 ± 2.72	40.79 ± 2.48	39.73 ± 1.14
$\Delta y_{jj} > 3.6$	148.52 ± 4.48	134.06 ± 1.90	31.32 ± 2.14	31.74 ± 1.00
Central Jet Veto	29.64 ± 1.96	24.88 ± 0.81	11.93 ± 1.21	11.60 ± 0.58
Outside Lepton Veto	17.40 ± 1.41	13.75 ± 0.60	10.05 ± 1.10	9.22 ± 0.52
$p_{\mathrm{T}}^{\mathrm{tot}} < 50\,\mathrm{GeV}$	15.52 ± 1.33	12.73 ± 0.57	9.21 ± 1.05	8.37 ± 0.49
$p_{\mathrm{T}}^{\ell\ell,\mathrm{jets}} > 25\,\mathrm{GeV}$	15.52 ± 1.33	12.72 ± 0.57	9.21 ± 1.05	8.36 ± 0.49
$Z \to \tau\tau$-Veto	15.04 ± 1.31	12.46 ± 0.57	9.11 ± 1.04	8.13 ± 0.49
$f_{\mathrm{recoil}} < 0.2$	9.16 ± 1.03	8.61 ± 0.47	6.22 ± 0.86	5.98 ± 0.42
$e\mu$ channel	unfiltered	merged	unfiltered	merged
b-jet veto/requirement	21940.64 ± 53.38	21833.19 ± 47.26	2451.76 ± 19.63	2424.89 ± 17.36
$m_{jj} > 600\,\mathrm{GeV}$	737.98 ± 10.32	712.31 ± 4.84	116.31 ± 4.26	113.33 ± 2.00
$\Delta y_{jj} > 3.6$	411.01 ± 7.36	394.77 ± 3.34	91.77 ± 3.62	86.65 ± 1.67
Central Jet Veto	70.39 ± 2.91	67.70 ± 1.35	28.91 ± 1.95	28.39 ± 0.95
Outside Lepton Veto	40.37 ± 2.17	37.47 ± 1.01	23.72 ± 1.75	22.69 ± 0.85
$p_{\mathrm{T}}^{\mathrm{tot}} < 50\,\mathrm{GeV}$	35.56 ± 2.03	33.29 ± 0.95	19.98 ± 1.62	19.84 ± 0.78
$Z \to \tau\tau$-Veto	32.89 ± 1.94	31.14 ± 0.91	18.01 ± 1.53	18.28 ± 0.74
μe channel	unfiltered	merged	unfiltered	merged
b-jet veto/requirement	18945.56 ± 49.55	18832.72 ± 43.92	2105.32 ± 18.19	2094.69 ± 16.05
$m_{jj} > 600\,\mathrm{GeV}$	641.32 ± 9.66	628.04 ± 4.58	103.49 ± 4.06	95.52 ± 1.83
$\Delta y_{jj} > 3.6$	345.99 ± 6.80	342.37 ± 3.10	82.21 ± 3.46	74.32 ± 1.53
Central Jet Veto	59.10 ± 2.71	57.97 ± 1.24	26.88 ± 1.93	24.57 ± 0.86
Outside Lepton Veto	36.13 ± 2.07	31.33 ± 0.93	20.61 ± 1.69	19.40 ± 0.75
$p_{\mathrm{T}}^{\mathrm{tot}} < 50\,\mathrm{GeV}$	32.60 ± 1.96	28.09 ± 0.87	17.31 ± 1.59	17.10 ± 0.69
$Z \to \tau\tau$-Veto	30.03 ± 1.89	25.75 ± 0.84	16.00 ± 1.52	15.96 ± 0.67

Table 7.5: Event yields of the raw unfiltered and merged samples in the signal region and top control region for ee, $\mu\mu$, $e\mu$ and μe final states at various cut stages.

8   Search for Evidence of VBF Higgs Boson Production

This chapter presents an analysis of the datasets described in Chapter 6 with the purpose to select possible candidate events for VBF Higgs boson production in the $H \to W^{\mp}W^{\pm(*)} \to \ell^{-}\bar{\nu}\ell'^{+}\nu'$ decay mode, referred to as *signal* in the following. All other processes introduced in Chapter 5 and listed in Chapter 6 are referred to as *backgrounds*, including ggF Higgs boson production for a Higgs mass of $m_H \approx 125\,\text{GeV}$, evidence of which is presented in Ref. [53]. This reference also includes an analysis of the 2-jet channel, henceforth referred to as VBF category for reasons detailed in Section 5.1.3. The analysis presented in Ref. [53] served as a baseline for the analysis presented here and is thus in parts closely resembled.

8.1 Object selection and efficiency corrections

Any HEP data analysis relies on the accurate reconstruction and identification of the physical objects subject to the measurement. However, since the measurement devices and algorithms only provide finite resolutions and efficiencies, the objects on which the analysis operates are not identical with the corresponding physical objects. However, it is convenient and aids the clarity of the description to refer to these reconstructed object candidates by the names of the corresponding physical objects. The quality criteria applied to these candidates before taking them into account for the scope of this analysis, the *object definitions* in HEP parlance, are thus provided in this chapter.

8.1.1 Scale factors

An accurate detector simulation is not only computationally expensive, but also hard to achieve. In general, it cannot be expected that the simulation is capable of reproducing the detector response and trigger decisions with arbitrary accuracy. To account for differences between data and Monte Carlo predictions, efficiency

scale factors are introduced. These are typically defined as

$$s = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}}, \tag{8.1}$$

where ϵ_{data} and ϵ_{MC} are the efficiencies in data and Monte Carlo simulated events, respectively.

However, the measurement of ϵ_{data} is itself nontrivial, and a large variety of methods is employed by dedicated performance groups to assess their respective values and uncertainties. A commonly used method for this purpose is the *tag-and-probe* method. When trying to assess the efficiency scale factor for some particle type, one can look at a region of phase space that is enriched with a specific decay mode with the respective particle in the final state. By posing tight requirements on all other decay products, it is then possible to measure how often the particle of interest is reconstructed and identified successfully.

Scale factors are typically applied in the form of weights to the simulated Monte Carlo events. Object-based scale factors (as opposed to event-based scale factors) in general depend on the kinematic properties of individual objects, and can therefore not only alter the normalization of a distribution, but also the shape.

A large number of scale factors is used in the analysis. In general, they are introduced in the section corresponding to the respective object selection, and the procedure of their assessment is briefly explained. A more general discussion of systematic uncertainties can be found in Section 9.2.

8.1.2 Trigger

For the purpose of searching for a final state with two high-p_{T}, isolated leptons at a hadron collider, the use of lepton triggers is obviously the most sensible choice. Consequently, in order to maximize the event yield, single lepton triggers are used. However, due to limits in the data acquisition rate (see Section 3.2.4), it is unacceptable for single lepton triggers to collect leptons with arbitrarily low p_{T}. Hence, additional dilepton triggers are used to include events with low-p_{T} leptons, since they can be adjusted to lower p_{T} values due to the lower rate of events with two good leptons.

The single lepton triggers employed for this analysis are `EF_e60_medium1` and `EF_e24vhi_medium1` for electrons, as well as `EF_mu36_tight` and `EF_mu24i_tight` for muons. Again, to maximize the yield, these are combined with a logical *or* ($\vee$). The nomenclature of the trigger names is detailed in the following paragraphs.

The `EF` is short for *Event Filter*, then an `e` or a `mu` after the underscore designate the corresponding particle for which these triggers are aimed, that is, electrons or

muons. The number following the particle identifier is the minimum transverse momentum threshold in GeV required for a particle to be selected by the trigger. The next part consists of a set of letters. The string `vh` signifies that the trigger has η and p_T dependent thresholds and a *hadronic leakage*[1] of less than $1\,\mathrm{GeV}$ at the first trigger level. The letter `i` indicates that additional isolation criteria are applied during the online event selection. Finally, the suffixes `medium` or `tight` correspond to the identification criteria similar to the ones introduced in Section 3.2.5, where the addition of the number `1` corresponds to the suffix `++` and thus to an increased tightness of the selection.

The dilepton triggers used are `EF_2e12Tvh_loose1`, `EF_mu18_tight_mu8_EFFS`, `EF_2mu13` and `EF_e12Tvh_medium1_mu8`. The corresponding names are composed of the same building blocks as the single lepton triggers mentioned before, but naturally refer to two leptons, which is signified by a prefix `2` before the lepton identifier, or in case of different lepton identifiers or selection criteria by simply concatenating the information for both leptons. The presence of a `T` denotes that the p_T threshold of the event level filter is only slightly higher than the threshold of the base trigger at L1, such that due to the smooth turn-on curve of the L1 trigger, the trigger efficiency for values close to the threshold might be decreased. The suffix `EFFS`, short for *event filter full scan*, denotes that a full pattern recognition is carried out at the high-level event filter (as opposed to reusing information from earlier trigger decisions) [96].

Trigger scale factors are computed based on offline trigger efficiencies, using the *tag-and-probe* method. For example, single lepton trigger efficiencies can be measured by selecting single-lepton triggered events that have an additional reconstructed lepton, using the first for tagging and the latter for probing. The trigger scale factor for the present analysis has been computed by simple probabilistic combination of the individual single-lepton and dilepton trigger scale factors.

8.1.3 Lepton selection

Leptons are an essential ingredient to the $H \to W^\mp W^{\pm(*)} \to \ell^- \bar\nu \ell'^+ \nu'$ analysis. They constitute the key feature of the signature that allows for efficient triggering and the accuracy and efficiency of their reconstruction is of crucial importance for the present analysis.

Hence, a rather large and complex set of criteria is applied before accepting the respective lepton in the analysis. These criteria are detailed in Table 8.1. Generally

---

[1]The *hadronic leakage* is a measure of the energy puncture from the electromagnetic to the hadronic calorimeter and is measured in a region of 0.2×0.2 in $\eta \times \phi$ around the particle direction behind the EM cluster corresponding to the particle.
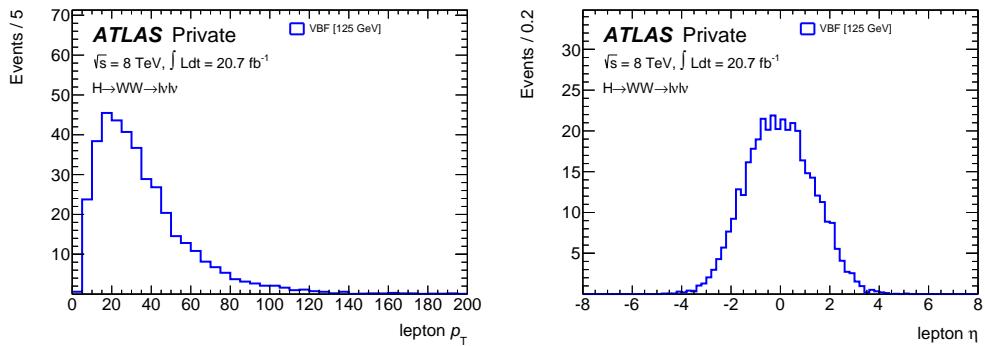
Figure 8.1: Distributions of transverse momentum p_{T} (left) and pseudorapidity η (right) in simulated truth leptons from VBF signal events before application of the object selection. An event filter requiring the leading lepton to have $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ and the subleading lepton to have $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ has already been applied.

speaking, leptons are required to meet a minimum requirement of transverse momentum p_{T}, and to be measured in the central detector region (small $|\eta|$). Furthermore, they are required to be isolated in order to avoid constellations where the lepton originates from the showering process or decay of a heavy hadron. Additional requirements are put on the impact parameters z_0 and d_0 (see Appendix 6.3) in order to reduce contributions from pile-up and secondary leptons from heavy flavor jets, respectively. Distributions of transverse momenta p_{T} and pseudorapidity η in simulated leptons from VBF signal events before application of the identification requirements are shown in Fig. 8.1.

The lepton efficiency scale factors are measured using a *tag-and-probe* method based on decays from Z and W bosons as well as J/Ψ hadrons for low p_{T} [24].

8.1.4 Jet selection

Jets are abundantly produced at hadron colliders. Furthermore, they are of special interest to the analysis, since the sought-after signal topology of VBF Higgs production is supposed to contain two high energy *tagging*-jets.

The requirements posed on jets to be considered for the analysis are listed in Table 8.2. Jets are required to exceed a certain p_{T}-threshold that depends on the pseudorapidity under which the jet was measured. Additionally, the jet is required to meet a threshold on the jet vertex fraction (see Section 3.2.5). However, jets for which the JVF is not well-defined are also included. The excessive amount of pile-up jets in the forward detector region, combined with the lack of tracking information and the disparity of the calorimeter over different detector regions motivate the increase of the p_{T}-threshold for larger values of $|\eta|$. Distributions of transverse

Muons	Selection Criteria
Reconstruction algorithm	STACO combined (see Section 3.2.5, [97])
Transverse momentum	$p_T > 15\,GeV$
Pseudorapidity	$|\eta| < 2.5$
Impact parameters	
transverse	$d_0/\sigma\left(d_0\right) < 3$
longitudinal	$|z_0 \cdot \sin\theta| < 1\,\mathrm{mm}$
Isolation	$\Delta R < 0.3$
Calorimeter	$E_\mathrm{T}^\mathrm{cone,corr.,cell}/p_\mathrm{T} < \min\left(0.014\frac{p_\mathrm{T}}{\mathrm{GeV}} - 0.15, 0.2\right)$
Track	$p_\mathrm{T}^\mathrm{cone}/p_\mathrm{T} < \min\left(0.01\frac{p_\mathrm{T}}{\mathrm{GeV}} - 0.105, 0.2\right)$

Electrons	Selection Criteria
Reconstruction algorithm	EM sliding window [24] & Gaussian Sum Fitter (GSF, [98])
Identification	`tight++` (see Section 3.2.5)
Transverse momentum	$p_T > 15\,GeV$
Pseudorapidity	$|\eta| < 2.47$, excluding $1.37 < |\eta| < 1.52$
Impact parameters	
transverse	$d_0/\sigma\left(d_0\right) < 3$
longitudinal	$|z_0 \cdot \sin\theta| < 0.4\,\mathrm{mm}$
Isolation	$\Delta R < 0.3$
Calorimeter	$E_\mathrm{T}^\mathrm{cone,corr.,topo}/E_\mathrm{T} < 0.16$
Track	$p_\mathrm{T}^\mathrm{cone}/E_\mathrm{T} < \begin{cases} 0.12 & \text{if } 15 < p_\mathrm{T} < 25\,\mathrm{GeV} \\ 0.16 & \text{if } p_\mathrm{T} \geq 25\,\mathrm{GeV} \end{cases}$

Table 8.1: Summary of selection criteria for muons and electrons. The quantities with superscript *cone* refer to the sums of track momenta and transverse energies from clusters found in a cone of $\Delta R < 0.3$ around the lepton candidate. For calorimeter measurements, the addition *cell* denotes that the sum is taken over the calorimeter cells, whereas *topo* denotes that the sum is taken over the respective topological clusters. The definitions of the impact parameters z_0 and d_0 and a motivation for the use of $z_0 \sin\theta$ can be found in Appendix 6.3.
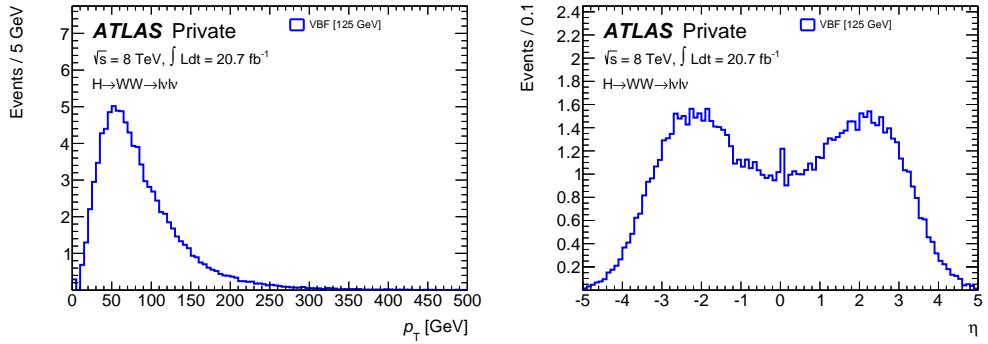
Figure 8.2: Distributions of transverse momentum p_T (left) and pseudorapidity η (right) in simulated truth jets from VBF signal events before application of the object selection.

Jets	Selection Criteria
Reconstruction algorithm	anti-k_T with $R = 0.4$ (see Section 3.2.5, [26])
Jet Cleaning	`looser` [100] (very loose)
Transverse momentum	$p_T > \begin{cases} 25\,\text{GeV} & \text{if } \|\eta\| < 2.5 \\ 30\,\text{GeV} & \text{if } \|\eta\| > 2.5 \end{cases}$
Calibration Scheme	local hadron calibration (see Section 3.2.5, [29])
Jet Vertex Fraction	$\|\text{JVF}\| > 0.5$
b-tagging algorithm	MV1 $- 80\%$ (see Section 3.2.5, [32])

Table 8.2: Summary of selection and b-tagging criteria for jets

momentum p_T and pseudorapidity η in signal event truth jets before application of the selection criteria are shown in Fig. 8.2.

Jet cleaning

Since the reconstruction of jets mainly relies on the calorimeter, occasional noise bursts can easily fake a jet. Thus, a procedure generally referred to as *jet cleaning* is performed, removing all jets from the event that do not meet certain quality criteria [99].

Identification of b-jets

In the present analysis, b-tagging is mainly used to reject top background from the signal region (see Fig. 5.9). For this purpose, the `MV1` b-tagging algorithm is used at a working point of 80% tagging efficiency, as already denoted in Table 8.2 and mentioned in Section 3.2.5.

The scaling factor associated with the b-tagging efficiency is subdivided into three different factors, corresponding to different probabilities.

- The b-tagging efficiency is the probability of correctly identifying and tagging a jet originating from a b-quark.
- The c-tagging efficiency is the probability of incorrectly b-tagging a jet originating from a c-quark.
- The mistag efficiency is the probability of incorrectly b-tagging a jet originating either from a light u-, d- or s-quark or from a gluon.

The event-based b-tagging scaling factor is then the product of the individual scaling factors for all selected jets in the event and is applied as a b-tagging *weight* to the event.

8.1.5 Overlap removal

In order to further ensure the purity of the object selection, a number of overlap procedures is employed. This is done to protect the analysis from potential double-counting of objects and to reject measurements of objects that do not originate from a hard scatter, but instead from the showering or decay of others. Naturally, the outcome of these overlap removal procedures is sensitive to the order in which they are conducted. Hence, the subsequent description of criteria matches the order in which the respective procedures are employed in the analysis.

Although muons are generally considered to interact very little with the detector material, they can still induce an albeit small shower in the calorimeter. If the energy deposition of these matches some track from the central interaction vertex, this might lead to the measurement of a *fake electron*. In rare cases, these fake electrons might even pass the selection criteria listed above (see Table 8.1). Hence, if a muon is found to overlap with an electron within a cone of $R < 0.1$, the electron is removed from the event.

Similarly, if a true electron cluster is matched by some additional track, it is possible that both of these tracks pass the tight isolation criteria listed above. In cases where two electrons are measured within a cone of $R < 0.1$, the electron with the lower transverse momentum is removed from the event.

It is more common for an electron-induced energy deposition in the calorimeter to be misidentified as a jet. If a jet and an electron overlap within a cone of $R < 0.3$, the jet is removed from the event.

Especially heavy-flavour jets have a high probability of yielding muons as decay products. Although these muons are in general efficiently removed from the event by the selection criteria listed in Table 8.1, additional protection against these heavy-flavour muons is achieved by removing muons that overlap with a jet within a cone of $R < 0.3$.

Shape distributions of ΔR between pairs of electrons and jets or muons and jets before application of the respective overlap removal procedures are shown in Fig. 8.3.

8.2 Event selection

The sensitivity of any statistical data analysis in high-energy physics relies to a large extent on an efficient event selection. All previous chapters and sections so far have prepared a setup where conducting a performant analysis is merely a matter of employing a performant event selection to separate events which are considered to be candidates for the sought-for signal process from the (albeit overwhelming) background. The signal process subject to this analysis as well as the numerous sources of background have been introduced in Chapter 5. This section presents an event selection capable of significantly improving the signal-to-background ratio
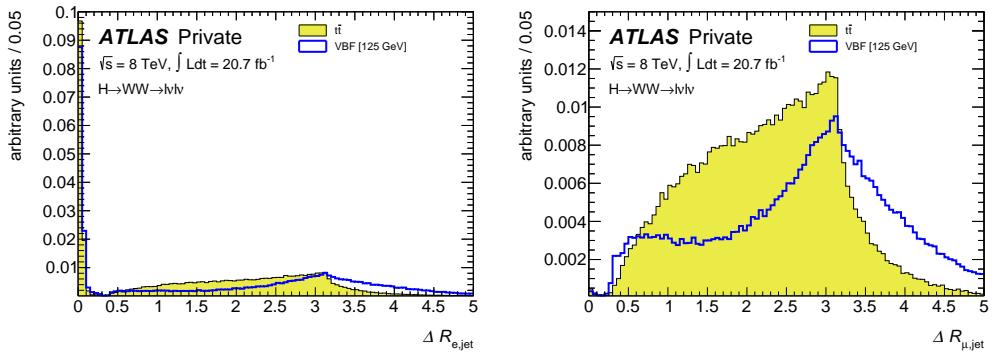
Figure 8.3: Shapes of distributions of distances ΔR between pairs of electrons and jets (left) or muons and jets (right) passing the object selection for VBF signal and $t\bar{t}$ events with at least two `tight`ly identified leptons. While the density of closely aligned leptons and jets harshly decreases towards $\Delta R < 0.5$, a strong peak at $\Delta R = 0$ is visible for electrons since electron showers are likely to be reconstructed as jets.

in the collected data events. Some aspects of this selection are only described qualitatively for the time being – the precise numeric threshold of the selection are discussed in a subsequent optimization study in Chapter 10.

8.2.1 Preselection

The dataset considered for analysis already has a large amount of selection criteria applied, be it the imposed trigger requirements (see Section 8.1.2), or be it data quality criteria imposed (see Section 6.1). Additionally, the event contents have been filtered by applying object selection criteria (see Section 8.1). Furthermore, a set of *preselection* criteria is applied to reduce the amount of data to a manageable quantity. In order to pass the preselection, events are required to have

- a primary vertex consistent with the beam spot position with at least three associated tracks with $p_{\mathrm{T}} > 0.4\,\mathrm{GeV}$,

- exactly two light and oppositely charged leptons passing the object selection criteria, at least one of which must match a triggering object, and at least one on which has a transverse momentum of $p_{\mathrm{T}} > 22\,\mathrm{GeV}$, and

- an invariant dilepton mass of $m_{\ell\ell} > 10\,\mathrm{GeV}$ in the DF channel and similarly $m_{\ell\ell} > 12\,\mathrm{GeV}$ in the SF channel.

- Additionally, at least two jets are required to pass the respective object selection criteria.

The first requirement constitutes a general data quality criterion. The second requirement is specifically designed to select events that exhibit a topology similar to the sought-for $H \to W^{\mp}W^{\pm(*)} \to \ell^{-}\bar{\nu}\ell'^{+}\nu'$ decay. The third requirement serves as a background rejection against low-mass Drell-Yan production with decays into a pair of oppositely charged leptons. Finally, the dijet requirement is used to select events that exhibit a VBF-like topology.

Nevertheless, even after these preselection criteria have been applied, the background still exceeds the signal expectation by more than a factor of 10^{3}. Hence, additional requirements on various kinematic variables are applied to reduce specific types of background contributions. Some of these requirements, or *cuts* in HEP parlance, can be inverted in order to create *control regions*, which are introduced in the next section.

8.2.2 Separation of control regions

It has been a recurring subject throughout this thesis that Monte Carlo predictions cannot be ultimately trusted. In order to investigate the agreement between the data and Monte Carlo predictions, *control regions* (CRs) have been defined. These

correspond to regions in phase-space that contain a (very) pure sample of some special background, while otherwise having similar properties (or, in HEP parlance, being *kinematically close*) to the signal region. These CRs allow a comparison between event yields and shapes of kinematic variables for data and MC, and to calculate correction factors or perform reweighting procedures where necessary.

They are typically defined by using the same event selection as the signal region, but inverting one or more cuts designed to reject one particular background process. Such a region by construction fulfills the requirements mentioned earlier, namely being kinematically close to the signal region while containing a pure sample of some special background process. Practically, however, the low amount of events from the Monte Carlo samples passing the criteria often complicates the construction of such CRs, a problem that has already been addressed in Chapter 7. Hence it is customary to drop some cuts applied on the signal region for one or more of the CRs if investigations show that the corresponding cut does not have a large impact on the shapes of kinematic distributions in the CR in question.

The analysis cuts that are used to define control regions are motivated and discussed in the following sections.

Z/γ^*+jets control region

High-mass Drell-Yan production of Z bosons is an especially challenging background for the SF channel. Based on the earlier discussion of the kinematic properties of such events (see Section 5.2.3 and Fig. 5.13), this background can be efficiently rejected by requiring the invariant dilepton mass $m_{\ell\ell}$ to suffice

$$|m_{\ell\ell} - m_Z| > 15\,\text{GeV}.$$

This cut is only applied in the SF channel and is referred to as *Z-veto* throughout this thesis. It is convenient to define a Z Control region (Z CR) by inverting this particular cut. The distribution of events in $m_{\ell\ell}$ before the separation in control regions is shown in Fig. 8.4.

Top control region

The overwhelming background from $t\bar{t}$ and single top production (see Section 5.2.2) can be efficiently rejected by requiring

$$n_b = 0,$$

where n_b is the number of identified b-jets in the event, employing the MV1 b-tagging algorithm at a working point of 80% b-tagging efficiency. This requirement is
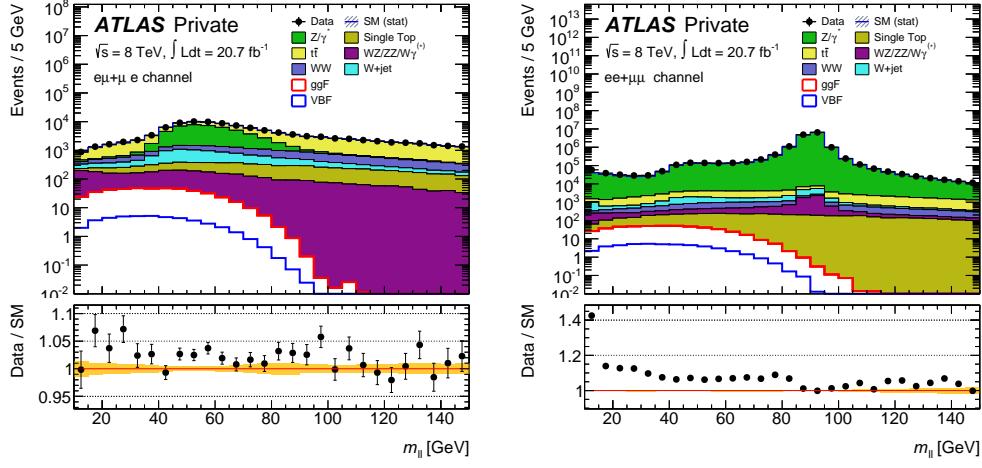
Figure 8.4: Distributions of $m_{\ell\ell}$ after the preselection for DF (left) and SF (right) events.
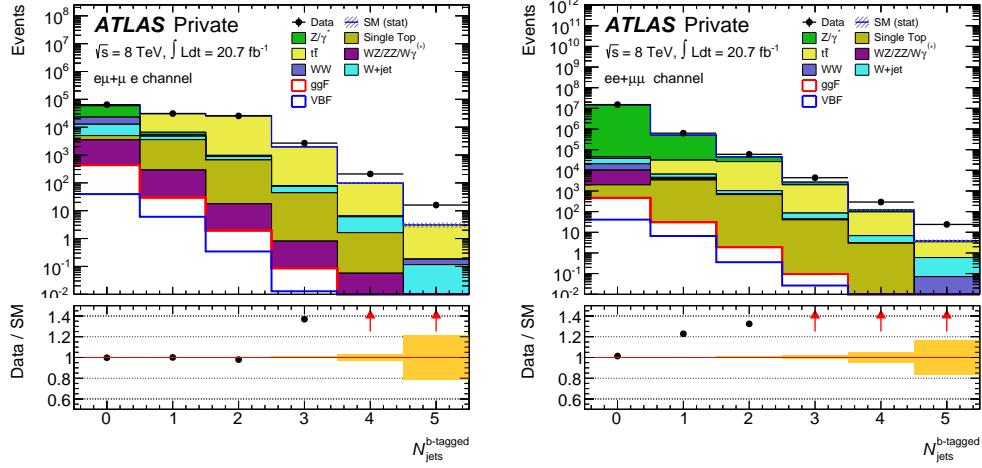
Figure 8.5: Distributions of the number of b-tagged jets after the preselection for DF (left) and SF (right) events. Red arrows indicate the direction of outlying data points in the ratio plot.

referred to as *b veto* or *b-jet veto* throughout this thesis. The inversion of this requirement yields a region enriched by events containing top quarks, which is used as a top control region. The distribution of events over the number of b-tagged jets after the preselection, but before the separation of control regions is shown in Fig. 8.5.

Other control regions

A large-scale analysis like the one presented in Ref. [53] typically employs a large number of control regions to estimate and investigate the various background processes involved in the analysis and to assess the quality of the simulation, often using and comparing different MC generators. However, for the analysis presented in this thesis, the two control regions presented above are sufficient.

Nevertheless, it should be mentioned that in the corresponding analysis presented in Ref. [53], an additional $Z \to \tau\tau$ control region was used in order to calculate a correction factor on the predicted number of $Z \to \tau\tau$ events in the signal region. The procedure employed is, however, fairly complex and performs a precision correction that is not in the focus of this thesis.

8.2.3 VBF-specific selection

The focus on the VBF signature allows a quick and efficient reduction of various kinds of background, since it specifically selects signal-like events.

Dijet kinematics

The distinctive VBF topology includes two highly energetic jets, being emitted oppositely into the forward- and backward-regions of the detector. However, due to the unknown longitudinal boost of the centre-of-mass frame of the collision, the pseudorapidity values of the jets might be shifted. It is thus useful to use their separation in rapidity Δy_{jj} (as the corresponding physical quantity) as well as their invariant dijet mass (as a measure of the recoil energy in the dijet system) to select events with a VBF-like topology. These variables have already been used in Chapter 7 to construct the VBF filter for truth level Monte Carlo generation.

The two tagging jets are identified by selecting the two jets with the highest transverse momentum p_{T}. While these need not necessarily be the same jet pair that was created during the VBF Higgs boson production process, selecting the two highest-p_{T} jets minimizes the probability of selecting pile-up jets.

As an explicit cut, it is required for the tagging jet pair to suffice

$$\Delta y_{jj} > 3.6.$$

As already discussed in Chapter 7, this requirement is different from the requirement imposed by the `VBFForwardJetsFilter` in that it only acts on the leading jet pair. The distribution of Δy_{jj} is shown in Fig. 8.6.

Another explicit cut is placed on the invariant dijet mass m_{jj}. However, the precise value of the optimal cut is studied in Chapter 10. Also the possibility of
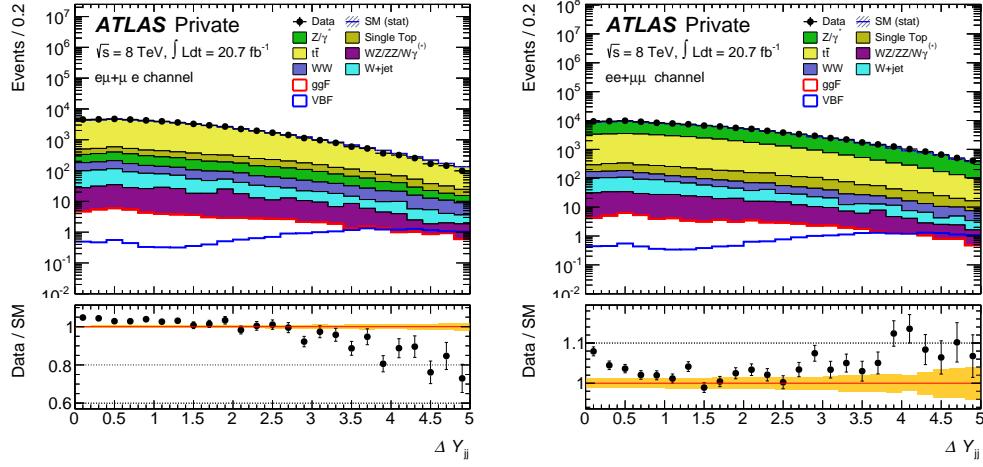
Figure 8.6: Distributions of the rapidity gap Δy_{jj} between the leading jets in the signal region for DF (left) and SF (right) events after the preselection.
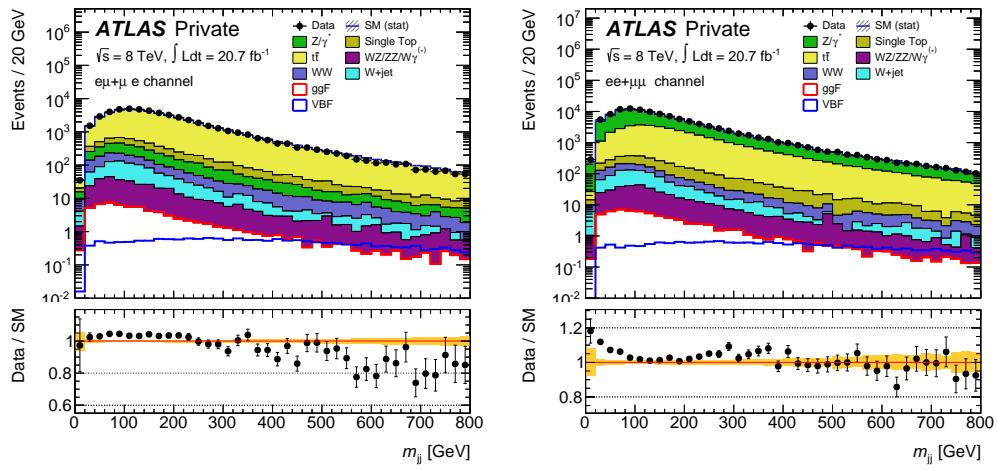
Figure 8.7: Distributions of the dijet invariant mass m_{jj} of the leading jets in the signal region for DF (left) and SF (right) events after the preselection.
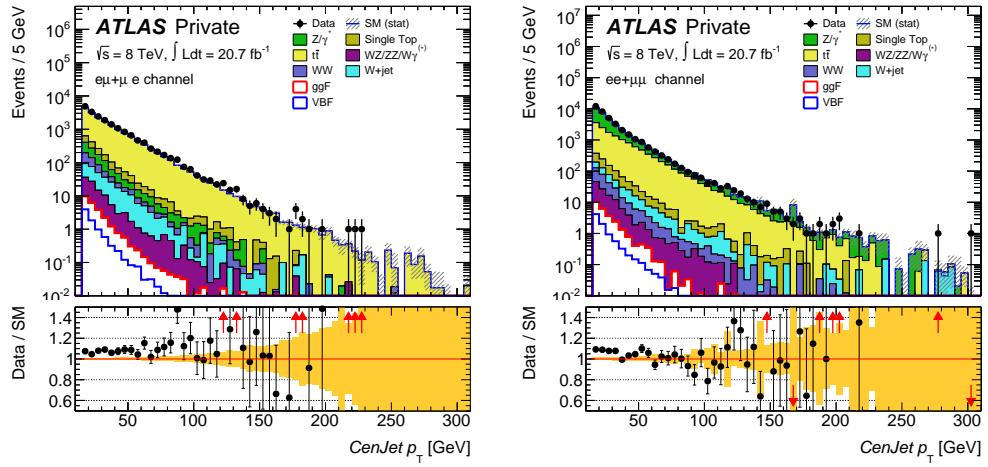
Figure 8.8: Distributions of the transverse momentum p_T of the leading jet contained in the rapidity gap between the two tagging jets for DF (left) and SF (right) events after the preselection. Red arrows indicate the direction of outlying data points in the ratio plot.

splitting the signal region in two by imposing a split boundary in m_{jj} is presented there, including an optimization study of this boundary. The distribution of m_{jj} is shown in Fig. 8.7.

Central detector region

As the two tagging jets are expected to leave a large rapidity gap in the central detector region, only occupied by the two leptons from the W decays, this feature can also be exploited for topological selection cuts.

Since the occurrence of additional leptons has already been excluded by the preselection, the central gap can be ensured by requiring that no jet with a transverse momentum of $p_T > 25\,\mathrm{GeV}$ is measured in the pseudorapidity range spanned by the two tagging jets. This requirement is commonly referred to as the *central jet veto* (CJV). The distribution of the transverse momentum of the leading jet in the central detector region is shown in Fig. 8.8.

Additionally, due to the large mass of the Higgs boson and the spin correlations in the WW system, it is expected that the two leptons are emitted into the central detector region, ideally within the rapidity gap between the tagging jets. This can be transformed into a cut by requiring that the η values of both leptons are within the range spanned by the jets, that is,

$$\max\left(\Delta\eta_{\ell_1 jj}, \Delta\eta_{\ell_2 jj}\right) < \tfrac{1}{2} \cdot \Delta\eta_{jj}$$
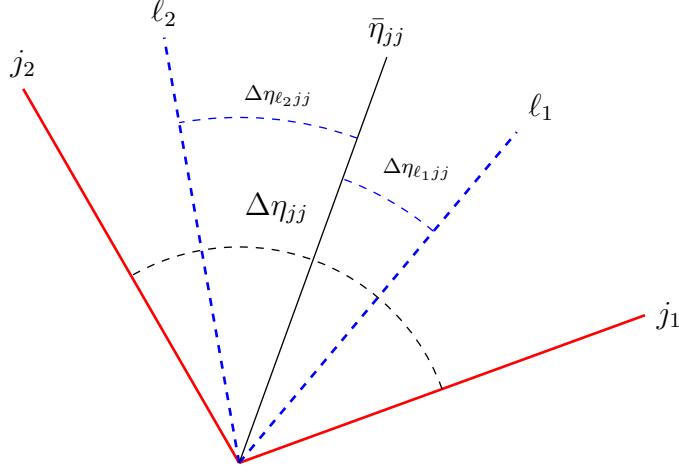
Figure 8.9: Graphical illustration of the (continuous) outside lepton veto.
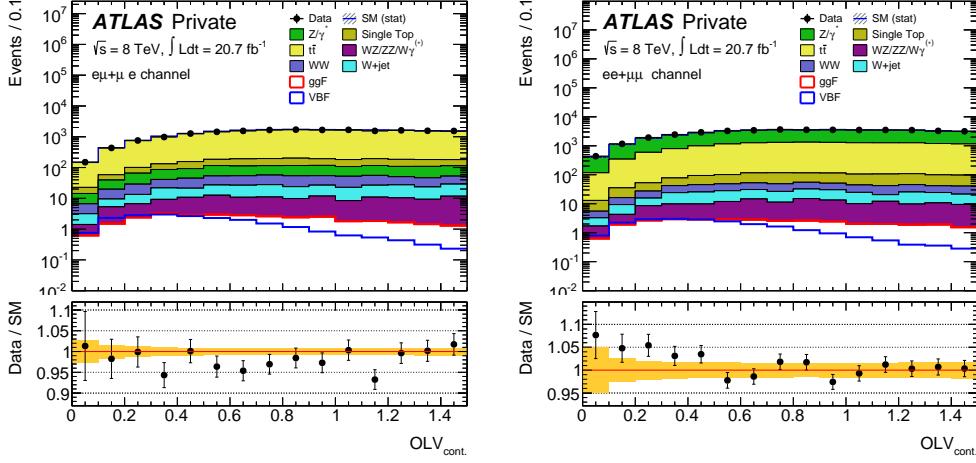
Figure 8.10: Distributions of the continuous outside lepton veto variable $\mathrm{OLV_{cont.}} = max\left(\Delta\eta_{\ell_1 jj}, \Delta\eta_{\ell_2 jj}\right)$ for DF (left) and SF (right) events after the preselection.

where $\Delta\eta_{\ell_i jj} = |\eta_{\ell i} - \bar{\eta}_{jj}|$ is the separation between the pseudorapidity η_{ℓ_i} of the lepton with number i and the average pseudorapidity value $\bar{\eta}_{jj}$ of both jets. The geometry of the situation is illustrated by Fig. 8.9.

However, this binary requirement can easily be transformed into a continuous variable by replacing the fraction $1/2$ with $C/2$, where the parameter C can be tuned to optimize the performance of the event selection. This approach represents an extension as compared to the analysis presented in Ref. [53]. A corresponding optimization study is carried out in Chapter 10. The distribution of the continuous outside lepton veto variable is shown in Fig. 8.10.
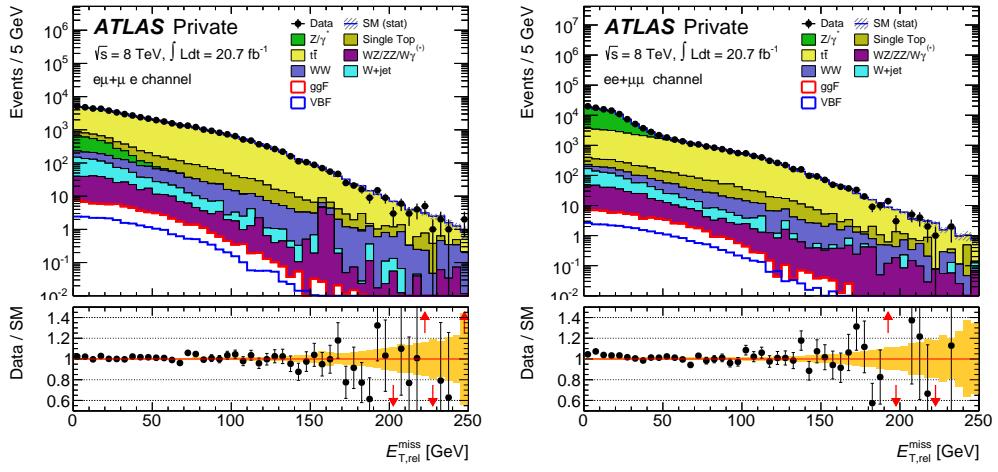
Figure 8.11: Distributions of $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ in the signal region for DF (left) and SF (right) events after the preselection. Red arrows indicate the direction of outlying data points in the ratio plot.

8.2.4 Z/γ^*+jets rejection for the same flavour channel

Even after the Z veto, the same flavour channel suffers from a large contamination from Z/γ^*+jets events. In order to remove these, a number of cuts are applied solely to the SF channel. These are detailed in the following.

Missing transverse energy

As opposed to Z/γ^*+jets events, signal events are expected to exhibit a large amount of missing transverse energy (see Section 5.2.3). However, the large amount of pile-up in data deteriorates the precision of the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, which relies on accurate measurement and reconstruction of the complete set of physical objects in the event. In order to achieve the most robust rejection, separate cuts are placed on $E_{\mathrm{T,STVF}}^{\mathrm{miss}}$ and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$.

The latter is defined as the absolute value of the projection of the missing transverse energy vector onto the direction of the closest object in the transverse plain. For this calculation, leptons and jets are taken into account, even before the requirement on the JVF (see Table 8.2) and the muon-jet overlap removal (see Section 8.1.5) are imposed. If no object is closer than $\pi/2$ to the direction of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ is chosen to equal $E_{\mathrm{T}}^{\mathrm{miss}}$. Studies have shown that this definition of $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ provides increased robustness against pile-up [101].

Optimization studies for the respective values of these cuts are shown in Chapter 10. The distributions of these variables are shown in Figures 8.11 and 8.12.
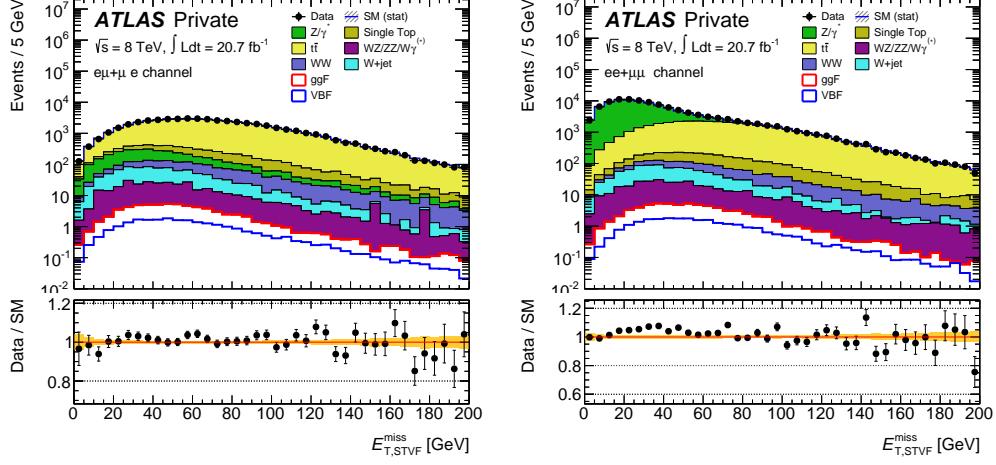
Figure 8.12: Distributions of $E_{\mathrm{T,STVF}}^{\mathrm{miss}}$ in signal region for DF (left) and SF (right) events after the preselection.

Hadronic recoil

An additional property that is useful to reject Z/γ^*+jets background is the momentum balance between the hadronic and the leptonic fractions of the event. Two distinct variables have been defined to quantify this property, $p_{\mathrm{T}}^{\ell\ell,\mathrm{jets}}$ and f_{recoil}.

For $p_{\mathrm{T}}^{\ell\ell,\mathrm{jets}}$, the vectorial sum $\vec{p}^{\,\ell\ell,\mathrm{jets}}$ of all selected jets in the event and both selected leptons is calculated and projected onto the transverse plane. Since Z/γ^*+jets events do not have any (intrinsic) missing transverse momentum or energy due to the lack of neutrinos, events are required to pass a cut of

$$p_{\mathrm{T}}^{\ell\ell,\mathrm{jets}} > 25\,\mathrm{GeV}.$$

The $p_{\mathrm{T}}^{\ell\ell,\mathrm{jets}}$ distribution is shown in Fig. 8.13.

The calculation of the hadronic recoil fraction f_{recoil} is illustrated by Fig. 8.14. The weighted sum of the momenta of all jets within the "pacman"-cone (as defined in the caption of Fig. 8.14) is calculated, the respective values of the jet vertex fraction being used as weights for this purpose. To this extent, all jets with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ are considered. The transverse projection of the resulting vector momentum is then divided by $p_{\mathrm{T}}^{\ell\ell,\mathrm{jets}}$ to obtain f_{recoil} as a measure for the amount of hadronic activity in the angular quadrant opposing the tagging objects. The value of f_{recoil} is dimensionless and bounded in the interval $(0,1)$. A requirement of

$$f_{\mathrm{recoil}} < 0.2$$
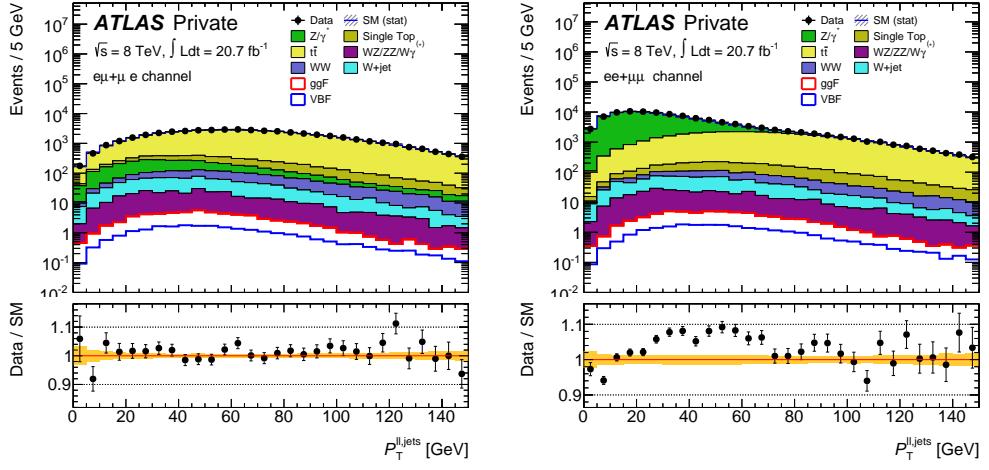
Figure 8.13: Distributions of $p_T^{\ell\ell,\text{jets}}$ for DF (left) and SF (right) events after the preselection.
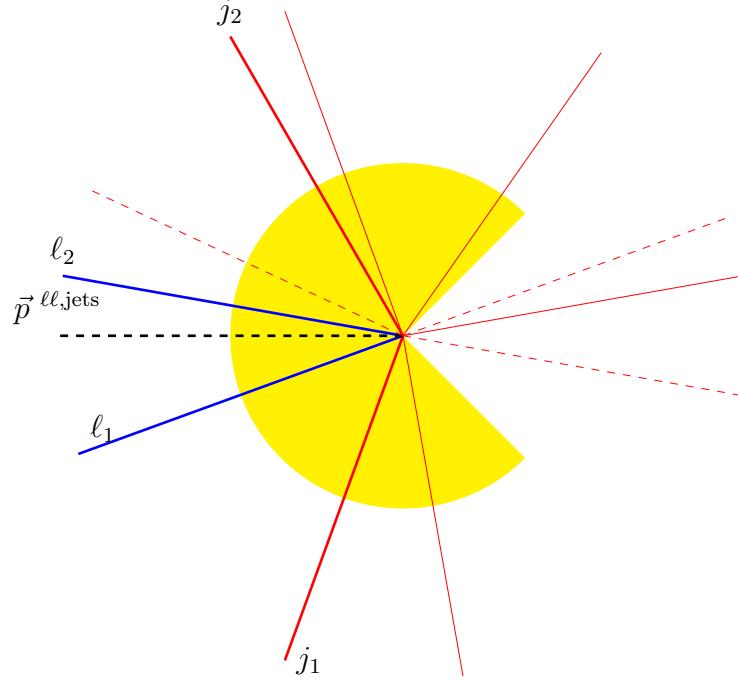
Figure 8.14: Illustration for the definition of the f_{recoil} variable. The blue lines correspond to the leptons, red lines symbolize jets. The leading (tagging) jets are drawn with thick lines, dashed lines symbolize jets not passing the jet selection detailed in Table 8.2. The black dashed line corresponds to the direction of $\vec{p}^{\,\ell\ell,\text{jets}}$, the vectorial sum of all selected jets and leptons in the event (all solid lines). The region within $\Delta\Phi < 3\pi/4$ of the $\vec{p}^{\,\ell\ell,\text{jets}}$ vector is filled in yellow. The remaining quadrant is for obvious reasons referred to as *pacman*-cone.
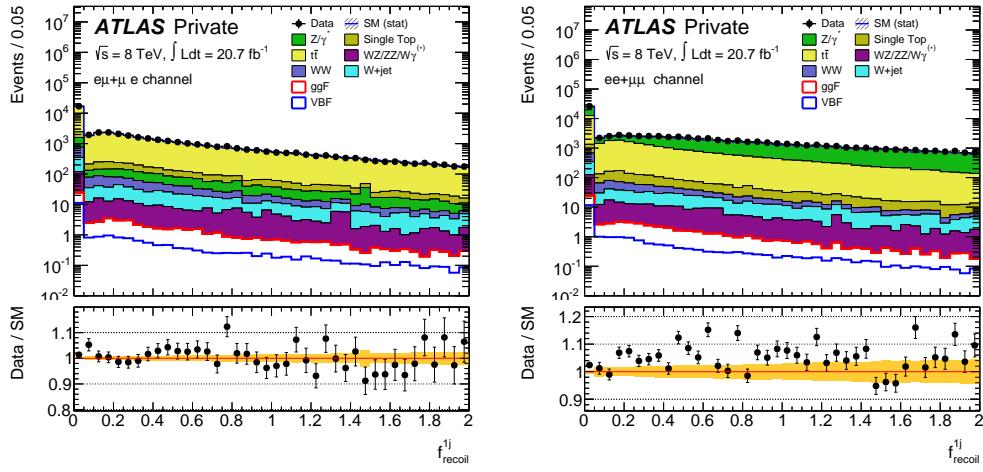
Figure 8.15: Distribution of f_{recoil} for DF (left) and SF (right) events after the preselection.

is imposed, since Z/γ^*+jets events are expected to exhibit a balanced signature, as opposed to the imbalanced angular signature expected from signal events due to the WW spin correlations. The respective distribution is shown in Fig. 8.15.

8.2.5 Additional background rejection

As already discussed in Section 5.2.3, the $Z \to \tau\tau$ poses a challenging background also in the DF channel. An efficient rejection is possible by reconstructing the invariant mass $m_{\tau\tau}$ of the $\tau\tau$-system employing the collinear approximation [63] and requiring it to differ from the Z boson mass by at least

$$|m_{\tau\tau} - m_Z| > 15\,\text{GeV}.$$

This requirement is referred to as $Z \to \tau\tau$ veto throughout this thesis. For illustration, the distributions of $m_{\tau\tau}$ are shown in Fig. 8.16.

Furthermore, an additional rejection of top background events is possible by considering the vectorial sum of the transverse momenta of all selected objects in the event, that is,

$$p_{\text{T}}^{\text{tot}} = \left| \vec{p}_{\text{T}}^{\,\ell_1} + \vec{p}_{\text{T}}^{\,\ell_2} + \sum_{\text{jets}} \vec{p}_{\text{T}}^{\,j} + \vec{E}_{\text{T}}^{\text{miss}} \right|. \tag{8.2}$$

By construction, this variable is sensitive to all sorts of soft activity in the event and mainly sensitive to top background. This type of background can be efficiently rejected by imposing an upper cut of
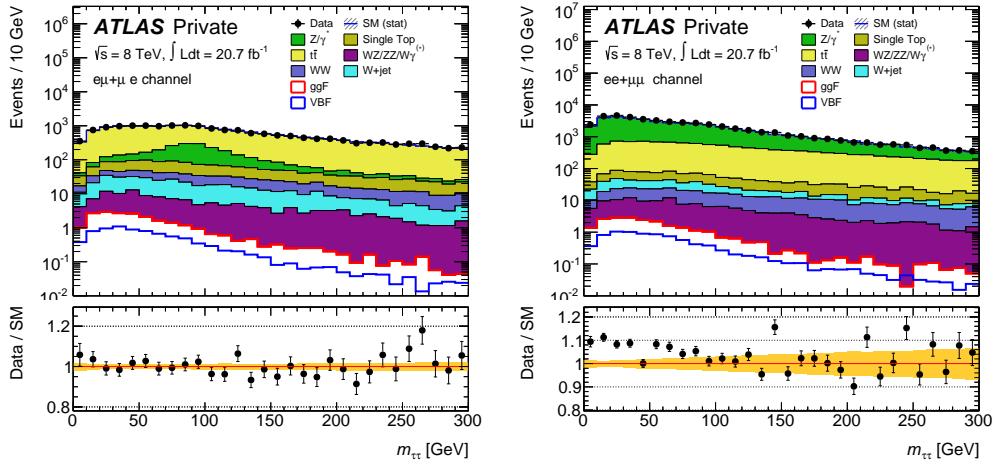
Figure 8.16: Distributions of $m_{\tau\tau}$ for DF (left) and SF (right) events after the preselection
.
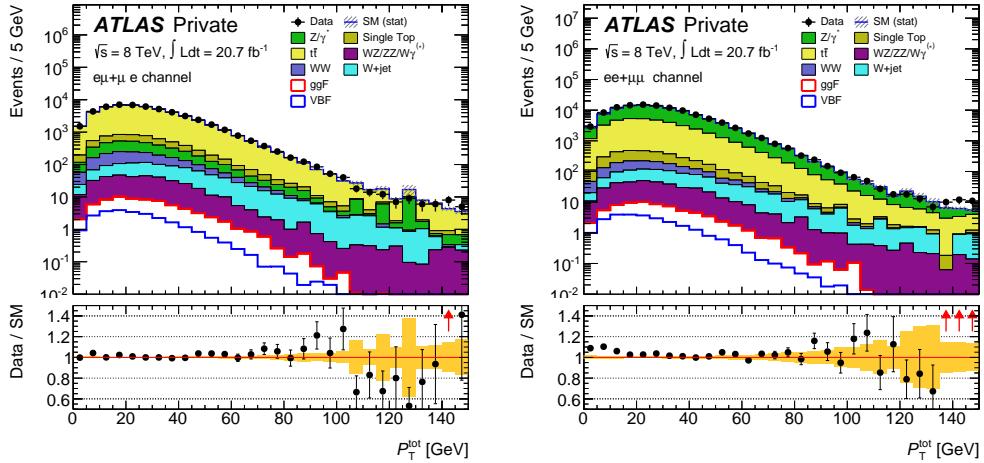
Figure 8.17: Distribution of p_T^{tot} for DF (left) and SF (right) events after the preselection.
Red arrows indicate the direction of outlying data points in the ratio plot.

$$p_T^{tot} < 50\,\mathrm{GeV}$$

on the event. The distribution of p_T^{tot} is illustrated in Fig. 8.17.

8.2.6 Topological selection

Based on the spin correlations in the WW system in Higgs boson decays detailed
in Section 5.1.2, the leptons are expected to be emitted closely aligned. Hence, the
signal topology can be selected by imposing upper cuts on the invariant dilepton
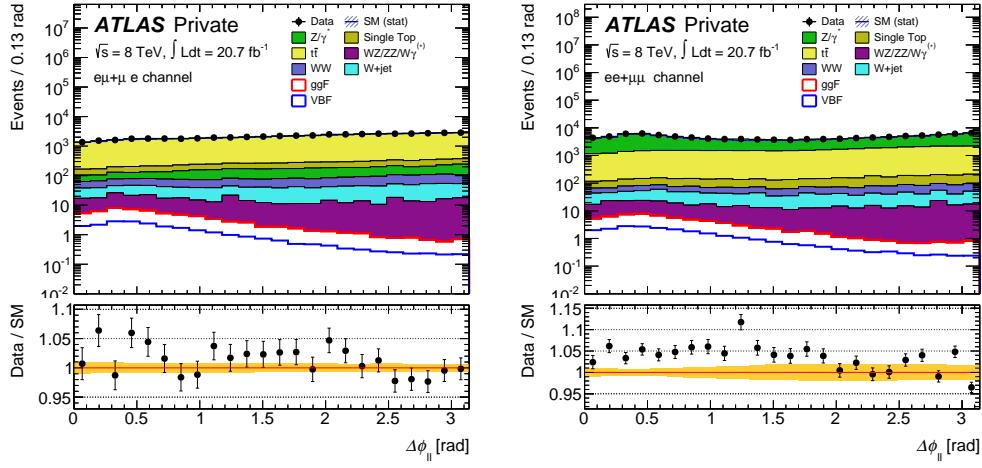
Figure 8.18: Distributions of $\Delta\phi_{\ell\ell}$ in signal region for DF (left) and SF (right) events after the preselection.

mass as well as the angular distance in the transverse plane. These topological cuts are chosen as

$$\Delta\phi_{\ell\ell} < 1.8$$

and

$$m_{\ell\ell} < 50\,\text{GeV}.$$

The respective distributions for both lepton flavour channels are shown in Figs. 8.18 and 8.19.

The application of these topological cuts completes the event selection. The remaining part of this chapter provides details on the cut ordering and additional background estimates. Cut optimization studies are carried out in Chapter 10, leading to a measurement of the VBF Higgs production cross-section in the $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ decay mode in Chapter 11.

8.2.7 Cut ordering

A summary of all the cuts imposed and the respective order of their application in the analysis can be found in Fig. 8.20. While the cut ordering is of no practical relevance for the result, it is often helpful to apply these cuts sequentially in order to investigate their respective effects on the event yield of some signal or control region. This way, distributions can be extracted at various cut stages in order to investigate or illustrate certain effects. The cut ordering presented below has been
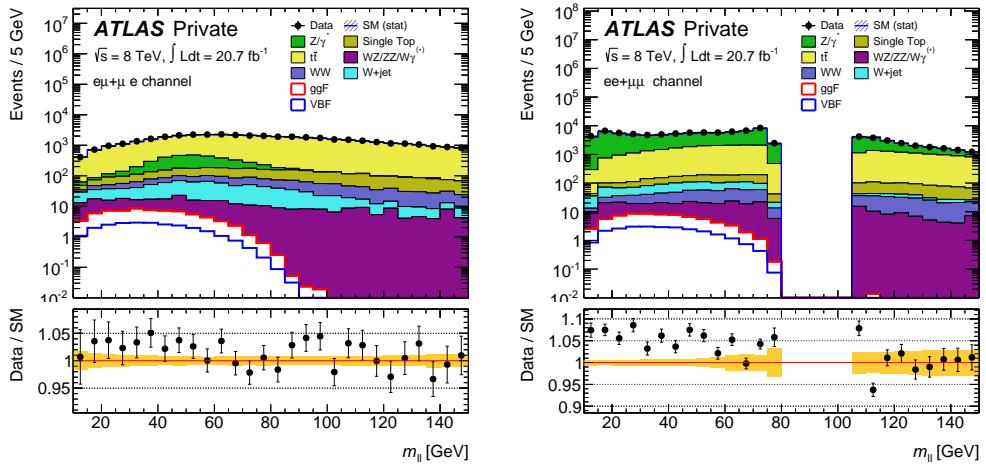
Figure 8.19: Distributions of $m_{\ell\ell}$ in the signal region for DF (left) and SF (right) events after the preselection. The gap in the region of $|m_{\ell\ell} - m_Z| < 15\,\text{GeV}$ is due to the Z veto applied earlier.

used everywhere in this analysis, and corresponding references to cut stages should be clarified by this illustration.

8.3 Background estimation and control samples

It can in general not be expected to achieve a complete rejection of background from all possible processes. An estimation of the amount of background events passing the event selection is thus an integral part of any statistical data analysis.

While the amount of agreement between Monte Carlo predictions for the expected background and signal event yields for virtually any kinematic region is usually striking, the precision of the prediction is limited by theoretical and computational constraints and will also be deteriorated by imperfect detector simulation and limited knowledge about the complex interplay of various aspects of the data taking conditions.

Several commonly used methods can be employed to verify and improve the quality of the predictions for various background processes. However, most of these methods are either incompatible or at least hard to sensibly combine with a statistical analysis using a likelihood fit method (see Chapter 9). A validation of the background predictions has been undertaken earlier and has shown a degree of compatibility sufficient for a statistical analysis of the results [53]. The present analysis employs the *fake factor* method for the W+jets background, which will be briefly described in the following.

8.3.1 Estimation of contributions from W+jets

As already presented in Section 5.2.4, the W+jets background is especially problematic in so far as it is not only abundantly produced, but also mimics the signal shape of the final fit variable m_{T}. The W+jets background does not resemble the signal topology – instead, the background events end up in the signal region due to misidentification of one lepton. However, it cannot be expected from the detector simulation to model the misidentification rate accurately enough to allow a precise background prediction. Consequently, a data-driven estimate is used.

While the majority of these *fake* leptons are indeed true leptons, they are regarded as *fake* because they do not directly originate from the hard scattering, but instead from secondary processes like photon conversion or heavy flavor decays. Especially electrons can also be faked by collimated jets of pions, which can trigger an electromagnetic shower similar to electron showers. To estimate the *fake rate*, one proceeds by constructing two independent control regions.

One of these, furthermore referred to as W+jets CR, is identical to the signal region, but requires one of the leptons to pass looser identification criteria, while failing the tight selection of the signal region, thus obtaining a disjoint, W+jets enriched control region that is disjoint from the signal region by lepton identification. The W+jets control region is depleted from any other backgrounds by subtracting the corresponding Monte Carlo predictions.

For the second control region, one chooses a jet enriched environment, which does not necessarily coincide with any part of the phase space used in the analysis at hand. In principle, the method requires the jets to be similar in nature and kinematics to the jets expected in W+jets events. However, this similarity is practically hard to achieve, for example in terms of properties like the heavy-flavour fraction. Hence, a systematic uncertainty is introduced to cover potential differences, which in turn propagates to become the leading systematic uncertainty on the resulting background estimate [53].

Since the true lepton rate in such events is very small, one can proceed by assuming that indeed all measured leptons are *fake* leptons and thus measure the ratio between the amount of tight leptons and the amount of leptons failing the tight criteria while passing looser ones. This ratio is referred to as the *fake factor*.

This *fake factor* can now be used to extrapolate from the number of W+jets events (with only loosely identified leptons) in the W+jets CR (as defined above) to the corresponding number of W+jets events in the tightly identified signal region. This procedure is commonly referred to as *fake factor method* and is widely applicable beyond the case at hand.
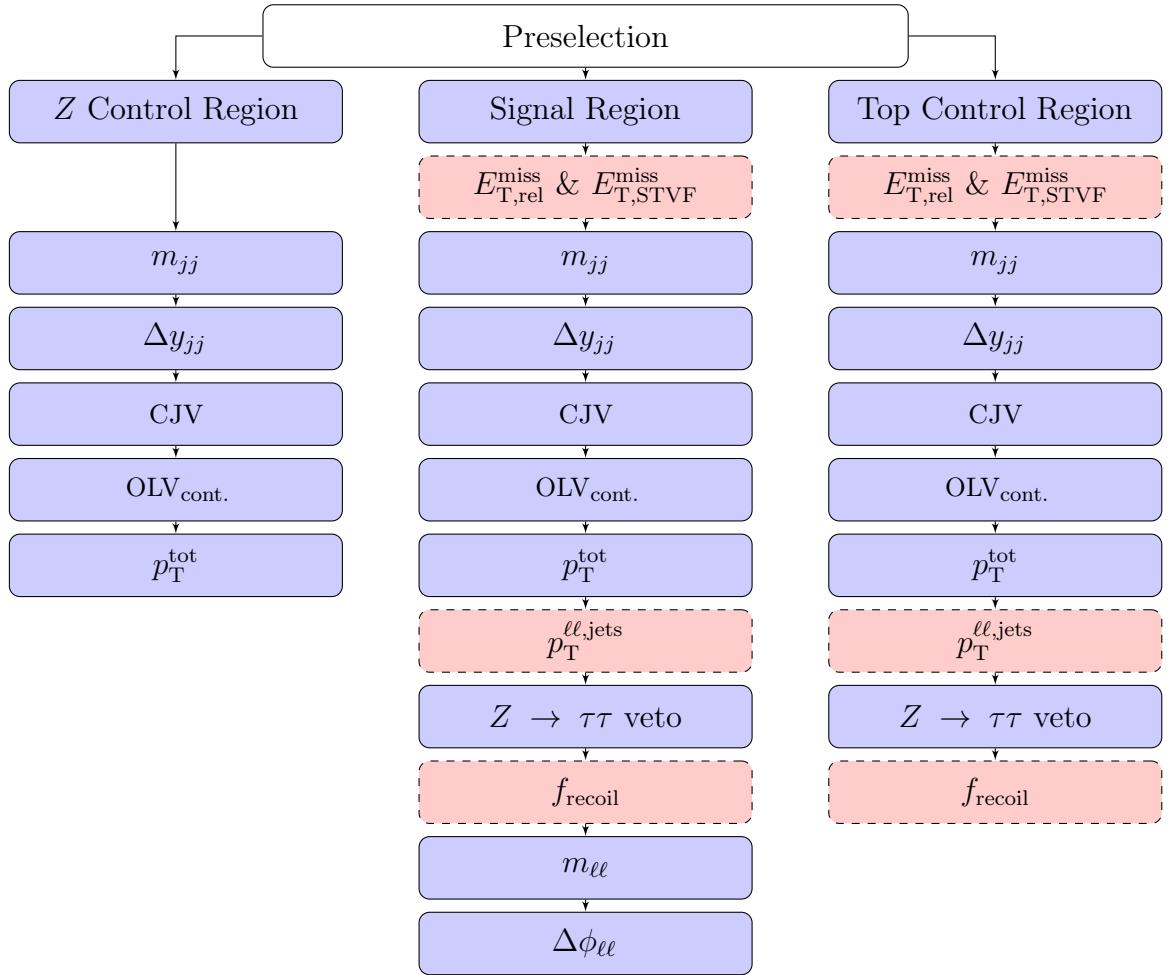
Figure 8.20: Cut ordering of signal and control regions. The cut definitions are given in the text. The [highlighted] nodes contain cuts that are only applied in the SF channel.

9   Statistical Methods and Systematic Uncertainties

The search for a yet unobserved process is a common case in particle physics. It is customary to conclude the analysis with a measured p-value. This terminology is commonly used in a frequentist interpretation to quantify the probability of obtaining the given measurement under the assumption of some hypothesis H. In this section, the statistical methodology used for the case at hand is described and systematic uncertainties are discussed. The cross section measurement in Chapter 11 and the optimization of the event selection described in Chapter 10 make use of the methods and results presented in this chapter. The present description of statistical methodology is largely based on Ref. [102].

9.1 Measures of sensitivity

The statistical significance of some observed signal is commonly quantified by means of a p-value or equivalently the Gaussian significance Z. The latter can be obtained using

$$Z = \Phi^{-1}\left(p - 1\right), \tag{9.1}$$

where Φ^{-1} is the quantile of the standard Gaussian distribution. In particle physics, it is customary to require a significance of $Z \geq 5$ as the appropriate level of confidence to claim a discovery, corresponding to $p \leq 2.87 \cdot 10^{-7}$. The following section describes the maximum likelihood method of hypothesis testing, also yielding the significance of the corresponding test result. After that, a much simpler approximation of significance is presented. Both methods are used and their respective results compared in the remaining part of this chapter.

9.1.1 Maximum likelihood method

Employing the standard glossary of hypothesis testing, the null hypothesis H_0 is in this case the background-only hypothesis, in which no signal is assumed to be present. The alternative hypothesis H_1 on the other hand describes a case where both background and signal are present. It is common to define the *signal strength parameter* μ as a scaling factor for the signal contribution. Using this definition, the null hypothesis H_0 corresponds to the case of $\mu = 0$, whereas the alternative hypothesis H_1 corresponds to $\mu = 1$. In general, however, the hypotheses or models will not only depend on the signal strength parameter μ, which is the parameter of interest in the case at hand, but also on a set of additional *nuisance parameters* θ_j, which can be allowed to float freely within certain physical limits (unconditional), or alternatively be fixed to some known or previously obtained value (conditional). Typically, these nuisance parameters include normalization factors for the various background contributions. For the sake of notation, θ (without an index) refers to the total set $\theta = (\theta_j)$ including all nuisance parameters.

For the purpose of the measurement of the parameter μ and the significance Z, one generally proceeds by defining a set of *regions n_i*. The values n_i here correspond to the respective number of events counted in the bins of one or more histograms. Their expectation values can be written

$$\bar{n}_i = \mu s_i \left(\theta \right) + b_i \left(\theta \right), \tag{9.2}$$

where the s_i are the signal b_i the background event yields predicted for the corresponding region by the alternative hypothesis H_1. These will in general depend on the values of the nuisance parameters θ.

Statistical independence of the regions n_i can in general easily be achieved by constructing the regions in a kinematically disjoint way. The likelihood function $\mathcal{L}$ is then given as

$$\mathcal{L} \left(\mu, \theta \right) = \prod_i P \left(n_i; \bar{n}_i \right) \prod_j G_j \left(\theta_j \right). \tag{9.3}$$

The probability density functions $P \left(k; \lambda \right)$ here encode the probability of measuring a value of k in a region for which a value of λ has been predicted by the hypothesis at hand. For the common case of counting experiments, these are taken to be Poisson distributions

$$P \left(k; \lambda \right) = \frac{\lambda^k}{k!} e^{-\lambda}. \tag{9.4}$$

The functions $G_j(\theta_j)$ are additional constraint terms used to model the impact of nuisance parameters on the likelihood function. It is customary to use Gaussian probability density functions with central values θ_j^0 and widths σ_j^0, both corresponding to the nominal (initial) value and uncertainty chosen for θ_j, typically based on previous measurements.

For the purpose of testing a hypothesized value of μ, the *likelihood ratio* Λ is employed, defined as

$$\Lambda(\mu) = \frac{\mathcal{L}\left(\mu, \widehat{\widehat{\theta}}(\mu)\right)}{\mathcal{L}\left(\widehat{\mu}, \widehat{\theta}\right)} \tag{9.5}$$

The conditional maximum likelihood estimator $\widehat{\widehat{\theta}}(\mu)$ is hereby defined as the set of nuisance parameter values that maximizes the likelihood for some given value of μ. On the other hand, $\widehat{\mu}$ and $\widehat{\theta}$ are unconditional maximum likelihood estimators in the sense that they maximize the likelihood over the full parameter space of μ and θ.

From Eq. 9.5, it can be seen that $0 \leq \Lambda(\mu) \leq 1$, with values of Λ close to one indicating a good agreement between data and the corresponding value of μ. It is customary to define a *test statistic* in order to quantify the agreement with the hypotheses in terms of μ. A common definition is the *log-likelihood* test statistic

$$t_\mu = -2\ln\Lambda(\mu). \tag{9.6}$$

In the case at hand, the presence of signal is expected to coincide with an upward deviation of the number of events with respect to the background-only hypothesis. Hence the occurrence of any signal is associated with a positive value of μ. As a consequence, negative values of the signal strength μ are deemed unphysical, since they would not only coincide with the absence of signal, but also contradict the background-only hypothesis. Consequently, the test statistic is commonly chosen in such a way to reflect this property of the model. A customary choice is

$$q_0 = \begin{cases} t_0 & \text{if } \widehat{\mu} \geq 0 \\ 0 & \text{if } \widehat{\mu} < 0 \end{cases} . \tag{9.7}$$

In most practical cases, the maximization of the likelihood function cannot be carried out analytically. Numeric procedures are hence commonly used for this purpose. The studies shown in this thesis use the MINUIT2 minimization tool [103] to minimize the negative logarithm of the likelihood function. Since earlier studies have shown that a positive value of μ can be expected [53], no special precautions

are taken to anticipate negative μ-values in the studies presented here, and the test statistic $q_0 = t_0$ is used. The case of $\mu < 0$ is neglected.

9.1.2 Poisson approximation

It can be shown [104] that for a single parameter of interest,

$$t_\mu = \frac{(\mu - \widehat{\mu})^2}{\sigma^2} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \tag{9.8}$$

where $\widehat{\mu}$ follows a standard Gaussian distribution with mean μ' and standard deviation σ. Here, N corresponds to the sample size. Considering the relatively simple case of merely testing H_1 with $\mu = 1$ against H_0 with $\mu = 0$ and using this approximation, one can derive [102] that for very large N,

$$p_0 = \Phi\left(\sqrt{q_0}\right), \tag{9.9}$$

where Φ is the Gaussian error function. Accordingly,

$$Z = \sqrt{q_0}. \tag{9.10}$$

Assuming only one region of interest where signal and background event yields s and b are known precisely and ignoring systematic uncertainties and their corresponding constraint terms, the likelihood function from Eq. 9.3 becomes

$$\mathscr{L}(\mu) = \frac{(\mu s + b)^k}{k!} e^{-(\mu s + b)}, \tag{9.11}$$

where k is the observed number of events. Hence,

$$t_0 = -2\ln\frac{\mathscr{L}(\mu = 0)}{\mathscr{L}(\mu = 1)} = -2\left[k\ln\frac{b}{s+b} + s\right] \tag{9.12}$$

and thus

$$Z(\mu = 0) = \sqrt{2\left[k\ln\left(1 + \frac{s}{b}\right) - s\right]}. \tag{9.13}$$

The latter relation for the significance Z is in the following referred to as the *Poisson estimator*.

However, the relation can be somewhat generalized for several disjoint regions by noticing that the likelihood function will simply be the product of the single-region

likelihood functions. Taking the log of the likelihood ratio transforms this product into a sum, of which the significance Z is the square root. Thus, significance estimates from different regions can be added in quadrature, as long as they are statistically independent and correlations are neglected.

$$Z_{\text{tot}} = \sqrt{\sum_i Z_i^2} \tag{9.14}$$

This relation is used in the following whenever Poisson-estimates of significances from different regions are combined.

9.2 Systematic uncertainties

While it is common for a full analysis to include uncertainties on the signal cross sections and background event yield normalizations, a proper discussion and inclusion of these uncertainties in the context of a cut optimization study is a formidable task and well beyond the scope of this thesis. However, a breakdown of the various sources of experiment uncertainties can more easily be achieved with the tools at hand and will thus be presented for the remaining part of this section. The ranking of systematic uncertainties is used as motivation for the choice of systematic variations to incorporate in the cut optimization presented in Chapter 10.

Figures 9.1 and 9.2 provide an overview over all sources of experimental systematic uncertainties considered in the present analysis. The various sources of systematic uncertainties are shown on the horizontal axis and is briefly discussed in the following paragraphs. The vertical axis shows the impact of the corresponding systematic on the uncertainty of the expected value of the signal strength parameter μ. The squares of the values shown were calculated by comparing the uncertainty on $\widehat{\mu}$ obtained from a fully unconditional likelihood fit with the corresponding one from a partially conditional fit where only the nuisance parameter in question was fixed to its best fit value. The uncertainties on the horizontal axis are sorted by their respective impact, largest values beginning on the left hand side. The respective values of some of the rightmost uncertainties seem to be zero. This is due to the influence of the corresponding parameters on the expected value of $\widehat{\mu}$ being too small for the numerical precision of the likelihood fit.
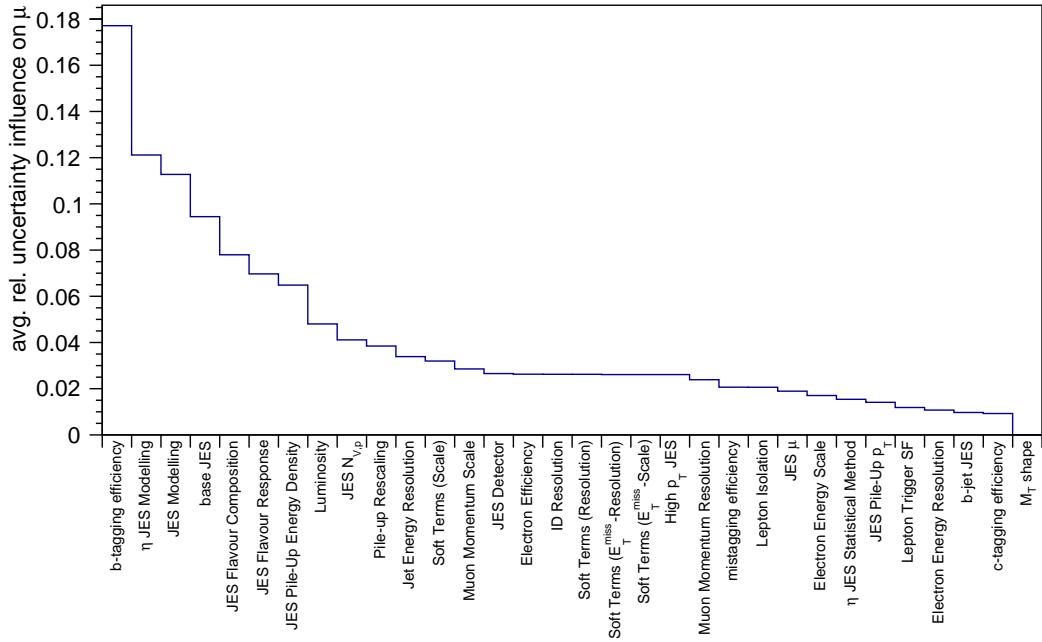
Figure 9.1: Breakdown of Experimental Systematic Uncertainties for the different flavour channel. The vertical axis shows the relative uncertainty on the signal strength parameter μ caused by the respective systematic uncertainty. The breakdown was obtained by comparing the uncertainty on $\widehat{\mu}$ for the unconditional fit with the uncertainty in the case where the corresponding parameter θ_i was fixed at its best fit value $\widehat{\theta}_i$.

9.2.1 Sources of experimental systematic uncertainties

Electron efficiency and the **muon efficiency** refer to the corresponding scaling factors, as introduced in Section 8.1.3. An additional scaling factor arises from the application of the **lepton isolation** criteria, which were detailed in the same section. The uncertainty on the **lepton trigger SF** has already been introduced in the last paragraph of Section 8.1.2. Additionally, the systematic uncertainties on the **electron energy scale** and **resolution** as well as the **muon energy scale** and **resolution** are considered.

The **b-tagging**, **c-tagging** and **mistagging** efficiency refer to the three types of efficiencies associated with b-tagging, as introduced in Section 8.1.4. The uncertainty associated **pile-up rescaling** refers to the uncertainty on the corresponding rescaling parameter, which has been introduced in Section 6.2.3.
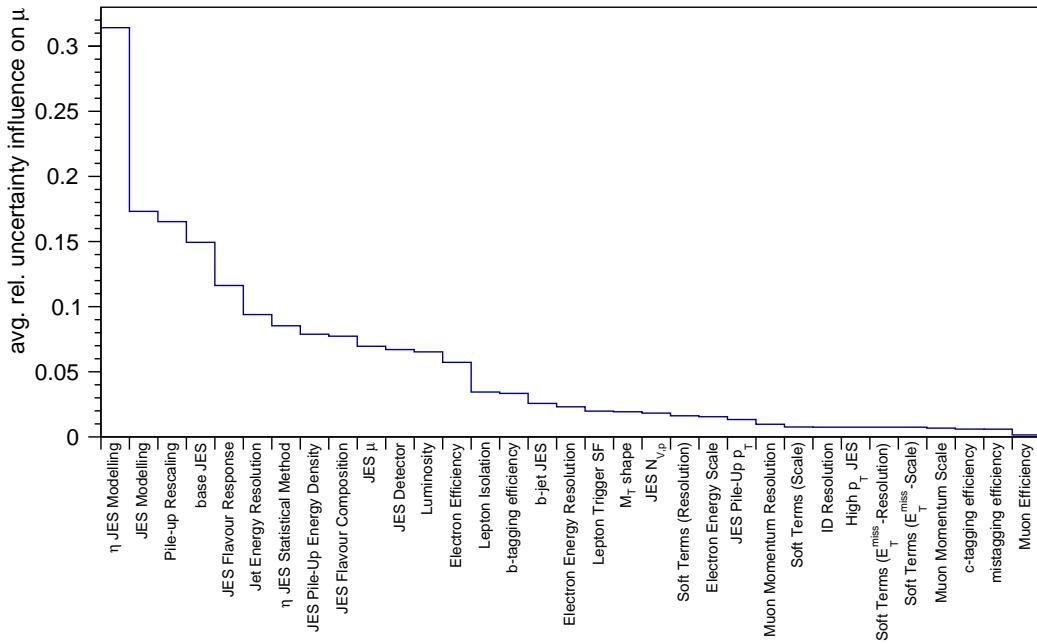
Figure 9.2: Breakdown of Experimental Systematic Uncertainties for the same flavour channel. The vertical axis shows the relative uncertainty on the signal strength parameter μ caused by the respective systematic uncertainty. The breakdown was obtained by comparing the uncertainty on $\widehat{\mu}$ for the unconditional fit with the uncertainty in the case where the corresponding parameter θ_i was fixed at its best fit value $\widehat{\theta}_i$.

The **luminosity** uncertainty originates from the imperfect knowledge of the integrated luminosity of the data set. Luminosity measurements are carried out with special (forward) calorimeters. The luminosity uncertainty is generally considered uncorrelated with any other systematic uncertainty due to its different origin.

The imperfect knowledge of the momentum resolution of the inner detector also contributes as a systematic uncertainty (**ID resolution**).

As the significance is extracted by a fit of the transverse mass shape, the corresponding uncertainty on the m_{T} **shape** prediction from Monte Carlo is also taken into account.

The uncertainties on the **soft term scale** and **resolution** influence the scaling of the cluster energy according to the soft term vertex fraction associated with the calculation of the quantity $E_{\mathrm{T,STVF}}^{\mathrm{miss}}$, see Section 5.1.2.

Numerous systematic uncertainties are related to the imperfect knowledge of the jet energy scale (JES, see Section 3.2.5). These are: the uncertainty on the

baseline value (**base JES**), on the flavour-dependence (**JES flavour composition** and **b-jet JES**), the intercorrelation between the transverse momentum and average density of calorimeter energy deposition of pile-up jets (**JES pile-up $p_\mathbf{T}$** and **JES pile-up energy deposition**) as well as on the general amount of pile-up, either measured by number of reconstructed primary vertices (**JES $N_\mathbf{vp}$**) or by the mean number of interactions per bunch crossing (**JES $\boldsymbol{\mu}$**), furthermore the dependence of the JES on various detector-specific effects (**JES Detector**, **JES Modelling**, $\boldsymbol{\eta}$ **JES Modelling** and $\boldsymbol{\eta}$ **JES Modelling (Statistical Method)**). The **jet energy resolution** uncertainty is considered separately.

For the present analysis, the dominant sources of experimental systematic uncertainties are on the b-tagging efficiency (for different-flavour) and the various modelling parameters of the jet energy scale. The leading role of b-tagging can easily be explained by the dominant nature of the top background, the contribution of which is highly sensitive to the assumed b-tagging efficiency. The large influence of the JES modelling parameters especially as a function of the pseudorapidity η can be explained by the extensive use of jet variables, especially forward jets, in the event selection criteria specific to the VBF topology.

9.2.2 Nuisance parameter pulls

It is customary to not only consider the relative uncertainties arising from the various nuisance parameters, but also their respective *best-fit* values compared to the initial (or *nominal*) values, as well as the corresponding uncertainties in terms of *pulls*. The *pull g* of some nuisance parameter θ is defined as

$$g = \frac{\widehat{\theta} - \theta^0}{\sigma^0} \tag{9.15}$$

where θ^0 and $\widehat{\theta}$ denote the *nominal* and *best-fit* values of the nuisance parameter θ, and σ^0 denotes the nominal uncertainty attributed to the initial value of the corresponding parameter. The uncertainty σ_g on the pull is defined as

$$\sigma_g = \frac{\widehat{\sigma}}{\sigma^0}. \tag{9.16}$$

While some nuisance parameters might be shifted by the fit in one direction or the other, all nuisance parameters are in general expected to exhibit a vanishing pull with an uncertainty close to unity. Nuisance parameters whose pull exhibit an uncertainty significantly smaller than unity are referred to as *overconstrained*. Since the nominal values of most parameters are measured and calculated by dedicated

performance groups, the latter property is generally thought to indicate potential problems with model used for defining the likelihood function.

10 Optimizing the Event Selection

The analysis described in Chapter 8 is based on the event selection already used in Ref. [53]. In order to improve on the results presented previously, this chapter details an optimization technique used to maximize the expected statistical significance of the discovery, leading up to the cross section measurement presented in Chapter 11.

The quantity subject to the optimization is the *expected discovery significance* Z_{exp}. The definition of Z has been introduced in Chapter 9. To protect against overfitting the given data in the signal region, this study is based purely on background estimates obtained from the background estimates presented in Chapter 8 and the Monte Carlo simulated samples listed in Chapter 6.

10.1 Description of methodology

The task of identifying variables with a large separation power between signal and background events is certainly not trivial and requires not only detailed physical knowledge of the matters at hand, but also a great deal of experience in particle physics. However, once a set of variables has been decided upon to exploit their discriminant power, a wide variety of options exist for a successful data analysis.

While multivariate methods like boosted decision trees or neural networks are certainly powerful tools, the inherent complexity of these tools is often listed as an argument against employing these methods. A much simpler approach is the *cut-based* analysis technique presented in Chapter 8. However, while many multivariate techniques have the feature to be self-training in the sense that the optimization of the performance is handled semi-automatically, a simple cut-based analysis lacks this feature and by construction contains a large number of a priori free parameters, the cut values.

10.1.1 Definition of the optimization problem

In its most general form, the optimization problem can be phrased as follows:

Consider an analysis with n parameters $\xi_i \in \mathbb{R}$, where $i \in \mathcal{N} = \{1, \ldots, n\}$.
Which parameter configuration $\vec{\xi}$ from the available configuration space $\mathcal{X} = \mathbb{R}^n$
yields the best value for the target function $f(\vec{\xi}) = Z_{exp}$?

However, there is no obvious way of estimating the significance Z_{exp} based on some configuration $\vec{\xi}$ without actually carrying out the corresponding analysis, which is a computationally expensive task.

Optimizing the value of some computationally expensive function on a high-dimensional input space is a recurring task in many scientific fields, for which efficient techniques have already been developed. However, many of these techniques are inadequate for the task at hand, for a number of reasons.

- Apart from the chosen point in configuration space, the simulated Monte Carlo events contribute as an input to the optimization. They are statistically distributed, but remain constant over the course of the optimization. Thus, the optimization is prone to systematic errors by overfitting the given dataset and should provide means of validating the outcome.

- The input space is by no means abstract, but corresponds to an event selection of an analysis. It should be possible to motivate the outcome of the optimization from a physical point of view. This can be regarded as a possible way of performing a cross-check, as required by the previous bullet.

- The validity of the result is of much higher priority than the runtime of the optimization. A computationally expensive optimization is acceptable, since it is only required to be carried out a small number of times.

- Due to the vast size of the input data used by the analysis, the runtime of the optimization is expected to be limited by the performance of the data storage facility. It is thus preferable to limit the number of times the input data have to be processed as far as possible, even if this comes at the expense of otherwise computationally more expensive algorithms for the optimization.

Based on these arguments, typical algorithms employed for multi-dimensional optimization like the simple yet efficient *steepest-descent* method are impractical for the case at hand, since the evaluation of the target function at a new point in configuration space would always require a full reprocessing of the input data. Additionally, after a local optimum has been found, the surrounding *significance landscape* would have to be investigated in order to decide on the validity and the stability of result, yielding further computational expenses. Hence, a different approach was chosen, which is described in the next section.

10.1.2 Description of the optimization algorithm

For the present studies, the full configuration space is sampled within certain boundaries. These are dictated by physical considerations and the respective number of signal and background events for data and Monte Carlo simulation expected in the phase space region that is selected by the analysis in the corresponding configuration. For this sampling, a rectangular lattice or grid $\mathcal{G} \subset \mathbb{R}^n$ with equidistant spacing in all available dimensions has been chosen, that is,

$$\mathcal{G} = \bigtimes_{i \in \mathcal{N}} \mathcal{R}_i \tag{10.1}$$

where

$$\mathcal{R}_i = \left\{ \xi_i^{\min} + m \cdot \xi_i^{\text{step}} | m \in \{0, \dots, n_i\} \right\} \tag{10.2}$$

is the optimization range with minimum $\xi_i^{\min}$, maximum $\xi_i^{\max} = \xi_i^{\min} + n_i \xi_i^{\text{step}}$, and n_i steps of width ξ_i^{step}. While the range $\mathcal{R}_i$ is only defined in a subspace of the configuration space $\mathcal{X}$, it can be trivially transformed into a respective range in the phase space Φ. Hence, the respective object in phase and configuration space are used synonymously.

The number of points N in the grid is given as the product of steps for all considered dimensions, that is

$$N = \prod_{i \in \mathcal{N}} n_i. \tag{10.3}$$

The analysis proceeds by filling an n-dimensional histogram with a total of N bins, such that the desired regions of phase space are resembled by a rectangular selection of histogram bins. Here, the lowest and highest bin in each dimension is used as an *underflow* or *overflow* and filled with any events that exceed the corresponding range $\mathcal{R}_i$. This way, the input data only needs to be processed once, filling a single (albeit large) histogram. The actual optimization can then proceed by iterating over all possible configurations resembled by the histogram binning, performing the statistical evaluation for each configuration and logging the results to a list of investigated points.

The latter can then be used to conveniently investigate the shape of the significance profile for the different variables at hand, and even correlations between the optimal values of different parameters. In order to visualize the results, a *profiling* procedure is chosen.

10.1.3 Visualization of the results

In one-dimensional significance profiles, the vertical axis shows the expected significance Z_{exp}. The horizontal axis resembles some variable ξ_i subject to the optimization, binned according to the range and step-width used to sample the configuration space. The *optimum* value $Z_{\text{exp}}^{\max_i}$ shown for $\xi_i = x$ then corresponds to the best (largest) value of Z_{exp}, obtained for any configuration X that employs the choice of $\xi_i = x$, that is

$$Z_{\text{exp}}^{\max_i}(x) = \max_{\vec{\xi} \in \mathcal{G}} \{f(\vec{\xi}) \,|\, \xi_i = x\}. \tag{10.4}$$

Analogously, the *top q [%] average* of Z_{exp} is defined for each $\xi_i = x$ as

$$Z_{\text{exp}}^{q,i} = \frac{1}{q} \sum_{\vec{\xi} \in Q_q^i(x)} f(\vec{\xi}). \tag{10.5}$$

Here, $Q_q^i(x)$ is the *constrained quantile set* belonging to the constraint $\xi_i = x$ and the quantile q. This set consists of all configurations $\vec{\xi}$ that belong to the significance quantile q of the *constrained configuration space* $\mathcal{X}_i(x) = \mathcal{G}|_{\xi_i = x}$. For example, the constrained quantile set $Q_{0.1}^i(x)$ contains the best 10% of all configurations $\vec{\xi}$ that employ the choice of $\xi_i = x$. With the notation $\#M$ used for the number of elements in some set M, the constrained quantile set can be written as

$$Q_q^i(x) = \{\vec{\xi} \,|\, \#\mathcal{X}_i(x) \cdot q > \#(\mathcal{X}_i(x) \cap \{\vec{\xi}' | f(\vec{\xi}') > f(\vec{\xi})\})\}.$$

The same visualization technique can be used to visualize the significances $Z_{\text{exp}}^{\max_{i,j}}$ and $Z_{\text{exp}}^{q,i,j}$ as a function of two parameters ξ_i and ξ_j at once. The corresponding values shown on the colour-axis are defined as

$$Z_{\text{exp}}^{\max_{i,j}}(x, y) = \max_{\vec{\xi} \in \mathcal{G}} \{f(\vec{\xi}) \,|\, \xi_i = x \wedge \xi_j = y\}. \tag{10.6}$$

and

$$Z_{\text{exp}}^{q,i,j}(x, y) = \frac{1}{q} \sum_{\vec{\xi} \in Q_q^{i,j}(x,y)} f(\vec{\xi}) \tag{10.7}$$

where, using $\mathcal{X}_{i,j}(x, y) = \mathcal{G}|_{\xi_i = x, \xi_j = y}$, the quantile set $Q_q^{i,j}(x, y)$ is defined as

$$Q_q^{i,j}(x, y) = \{\vec{\xi} \,|\, \#\mathcal{X}_{i,j}(x) \cdot q > \#(\mathcal{X}_{i,j}(x, y) \cap \{\vec{\xi}' | f(\vec{\xi}') > f(\vec{\xi})\})\}.$$

In the *significance profiles* shown in the following sections, the term *profiled optimum* will refer to the definition of $Z_{\exp}^{\max_i}$ or $Z_{\exp}^{\max_{i,j}}$, whereas the term *top q [%] average* refers to the definition of $Z_{\exp}^{q,i}$ or $Z_{\exp}^{q,i,j}$. For the latter two, error bars will be used to quantify the mean variation of the points around this average.

10.2 Topological VBF optimization for the different flavour channel

Employing the methods discussed in the previous sections, an implementation of the optimization is straight-forward. Some technical details are given in Appendix 3. Test runs conducted previously have shown that an optimization of the topological cuts specific to the VBF channel is especially promising. Restricting the optimization of these cuts to the DF channel seems justified, since the latter yields the leading contribution on the final result due to the much lower Drell-Yan background. A separate optimization of the SF specific cuts rejecting this particular background is carried out in the next section.

10.2.1 Region splitting by dijet invariant mass

In the previous chapters, a *cut-based* analysis has been discussed, employing a set of rectangular cuts to select a phase space region with a good ratio of signal to background events. Each cut represents a binary *yes/no* decision on whether to take the event into account for the analysis, or drop it and ignore it completely. However, as far as a statistical analysis is performed, an additional choice is possible. Instead of a *cut*, one can also employ a *split* to separate two phase-space regions with different properties and carry out the statistical analysis on both, either separately or simultaneously. This procedure is similar to the treatment of control regions or to the separate analyses carried out for the various lepton flavour channels.

Even if all other cuts are kept in synchronization between the regions, this method allows to provide the likelihood fit with additional information about the input data. There is no conceptual difference to the case where a fit is performed to a binned distribution in some observable, instead of a single bin representing the entire event yield of the signal region. Using this analogy, a *split* corresponds to the use of a two-bin histogram in the given variable.

The dijet invariant mass m_{jj} is a powerful variable for separating signal and background events, but yields a rapid decrease in event yield towards higher values. Hence, it seems advisable to use an m_{jj} split to separate two signal regions, a *low m_{jj}* region with a potentially higher event yield but lower purity, and a *high m_{jj}*

region with the opposite properties. The split boundary between these two regions is treated as a free parameter of the analysis and optimized along with the cut-type parameters. Of course, the correlation between the optimal values of the m_{jj} cut and the m_{jj} split is of particular interest.

Of course, it is possible to increase the number of bins further. However, a larger number of bins obviously complicates the analysis and decreases the expected event yield per bin. Thus, for the sake of simplicity and feasibility, the present optimization is only concerned with the two-bin case.

10.2.2 Properties of the study

While there is no principle reason prohibiting the use of the full likelihood fit for evaluation of each single point in configuration space, practical reasons like the available disk space and the runtime of the optimization require to restrict the evaluation method somewhat. Based on the ranking of systematic uncertainties presented in Section 9.2, the likelihood fit employed for this optimization study only uses the leading two, which are

- b-tagging efficiency, and
- η JES Modelling.

However, since the luminosity uncertainty is generally assumed to be uncorrelated with any other uncertainty, its inclusion is particularly simple. Hence, it was also included in the likelihood fit. The input to the likelihood fit are binned histograms in the transverse mass m_{T}. A definition of the latter is given in Section 5.1.2.

The present study investigates a simultaneous optimization of the following parameters:

- a lower cut on the dijet invariant mass m_{jj}, labelled m_{jj}^{min}
- a split in two distinct but simultaneously fitted sets of signal and control regions, separated by a boundary in the dijet invariant mass m_{jj}, labelled m_{jj}^{split}
- a lower cut on the dijet rapidity gap Δy_{jj}, labelled $\Delta y_{jj}^{\mathrm{min}}$
- an upper cut on the continuous outside lepton veto variable $\mathrm{OLV}_{\mathrm{cont.}}$ presented in Section 8.2.3, labelled $\mathrm{OLV}_{\mathrm{cont.}}^{\mathrm{max}}$

In the following, significance profiles are shown, as defined in the previous section. For comparison, the result obtained employing the likelihood fit variant as well as the corresponding result obtained employing the Poisson estimation method (as introduced in Section 9.1.2) are presented.

10.2.3 Results of the optimization

Figures 10.1 and 10.2 show the one-dimensional significance profiles of the cut and split values in the dijet invariant mass m_{jj} using the different evaluation methods. These exhibit a couple of interesting features, which are discussed in the following.

For a start, it is interesting to note that the Poisson estimator in general yields higher significances compared to the likelihood fit estimator. This can easily be explained by the fact that the latter includes systematic uncertainties, decreasing the total expected significance. Also, since the Poisson estimator is a simple arithmetic calculation, it is not surprising that the shape is more regular and exhibits less distinct features compared to the fit. Among the most prominent features of the results obtained from the likelihood fit as compared to the ones obtained from the Poisson estimator are a relatively sharp peak at $m_{jj}^{\mathrm{min}}{=}800\,\mathrm{GeV}$ as well as a moderate one at $m_{jj}^{\mathrm{split}}{=}600\,\mathrm{GeV}$. In order to investigate these effects,
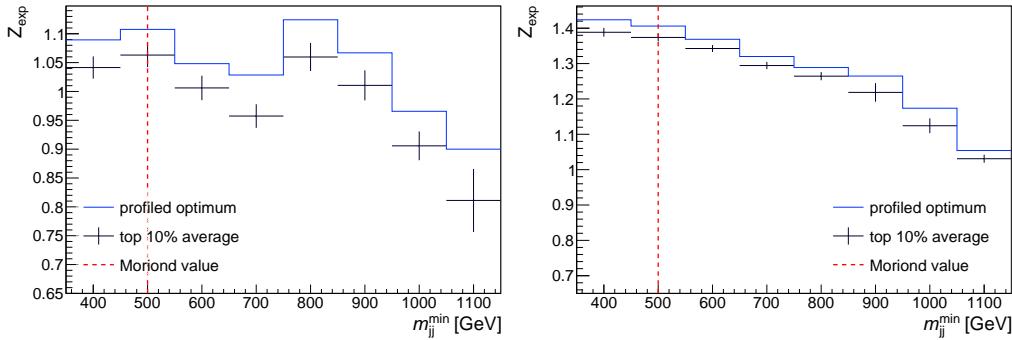
Figure 10.1: Significance profiles for the lower cut on the dijet invariant mass m_{jj}, using the full likelihood fit method (left) and the simple Poisson approximation (right). The dashed red line corresponds to the cut value used for the analysis presented at Moriond 2013 [53].
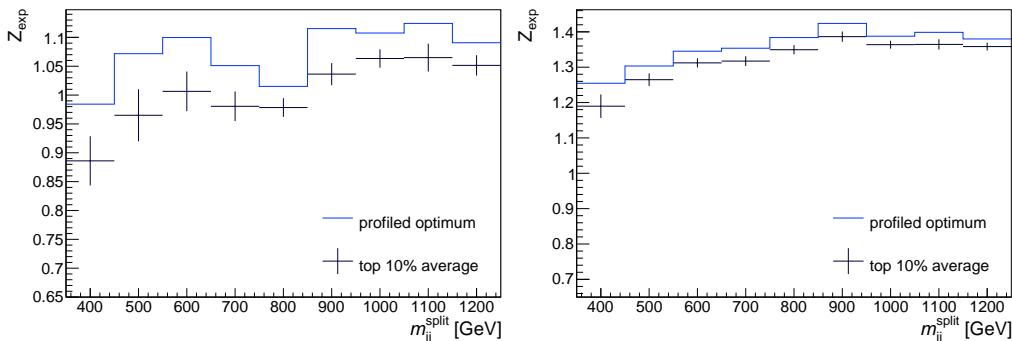
Figure 10.2: Significance profiles for the region split boundary in the dijet invariant mass m_{jj}, using the full likelihood fit method (left) and the simple Poisson approximation (right).
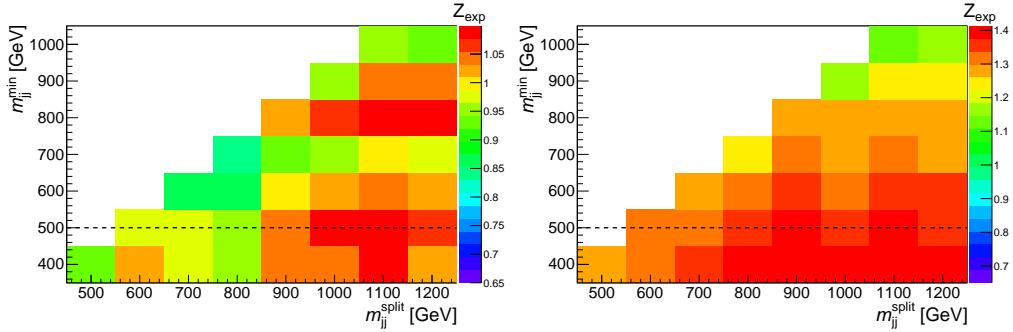
Figure 10.3: Two-dimensional top 10% average significance profiles for the lower cut value and region split boundary in the dijet invariant mass m_{jj}, using the full likelihood fit method (left) and the simple Poisson approximation (right). The significance is colour-coded, values increasing from blue to red. The dashed black line corresponds to the values chosen for the analysis presented at Moriond 2013 [53].

it is interesting to look at the correlation between the respective optimized values in both parameters, m_{jj}^{min} and m_{jj}^{split}, simultaneously, shown in Fig. 10.3.

Firstly, the upper left half of Fig. 10.3 is white (empty). This reflects the property that applying a lower cut value higher than the split boundary is meaningless.

The Poisson estimator generally favours a lower cut value m_{jj}^{min} in combination with a higher split value m_{jj}^{split}. This corresponds to the intuitive approach of applying a relatively loose selection, and splitting off the high-purity region for a separate analysis.

However, in the left (likelihood-fit) plot of Fig. 10.3, a couple of interesting regions can be distinguished. There seems to be a rather interesting configuration employing a very loose and inclusive selection with m_{jj}^{min}=400 GeV and m_{jj}^{split}=600 GeV. The region 400 GeV< m_{jj} < 600 GeV provides a large event yield for signal as well as for background. However, the significance value achieved by this configuration falls short approximately 10% of the optimum configuration in significance. Hence, this region of configuration space is not chosen for the present analysis, but should be kept in mind for future investigations.

The highest significance values are achieved around m_{jj}^{split}=1.1 TeV. Interestingly, there are two localized maxima, one corresponding to a configuration with m_{jj}^{min}=500 GeV, the other corresponding to m_{jj}^{min}=800 GeV. These two local maxima are separated by a local minimum around m_{jj}=700 GeV, where the significance decreases by about 10%. This effect is somewhat surprising and has been investigated in more detail, taking into account correlations with the other variables subject to this optimization study. An explanation is discussed within the next paragraphs.
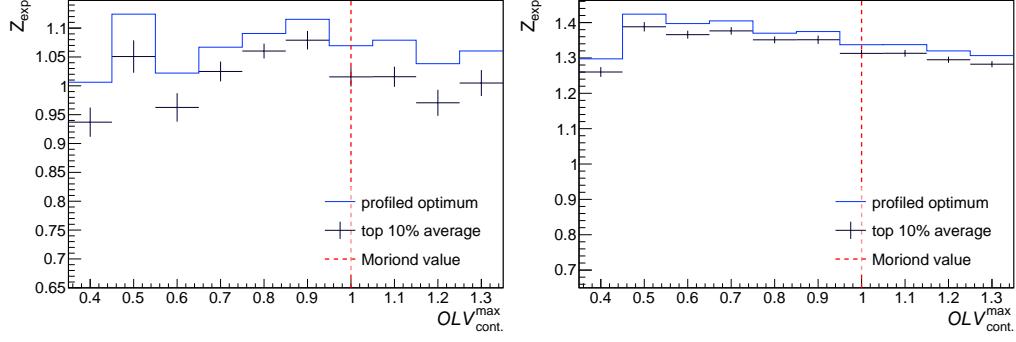
Figure 10.4: Significance profiles for the continuous outside lepton veto $\text{OLV}_{\text{cont.}}^{\text{max}}$, using the full likelihood fit method (left) and the simple Poisson approximation (right). The dashed red line corresponds to the cut value used for the analysis presented at Moriond 2013 [53].
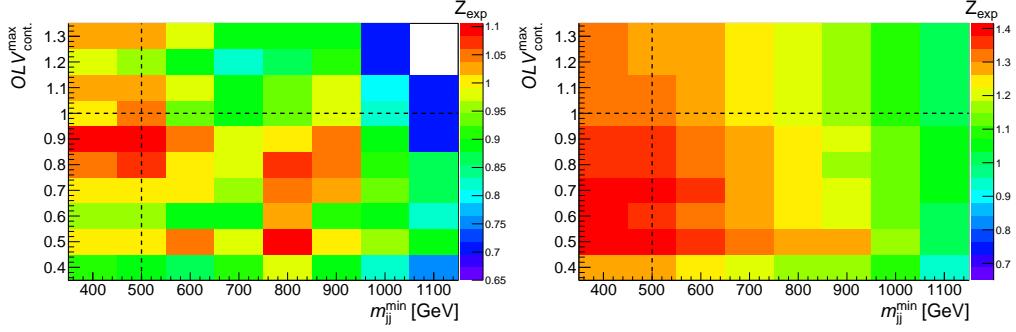
Figure 10.5: Two-dimensional top 10% average significance profiles for the lower cut value on m_{jj} and the continuous outside lepton veto $\text{OLV}_{\text{cont.}}$, using the full likelihood fit method (left) and the simple Poisson approximation (right). The significance is colour-coded, values increasing from blue to red. The dashed black line corresponds to the values chosen for the analysis presented at Moriond 2013 [53].

Figure 10.4 shows the optimization results for the continuous outside lepton veto variable $\text{OLV}_{\text{cont.}}^{\text{max}}$. It is interesting to note that while the Poisson estimator claims a steady rise in significance towards lower values until $\text{OLV}_{\text{cont.}}^{\text{max}}$=0.5, where the significance experiences a sharp drop-off due to insufficient event yields, the likelihood fit estimator does not confirm this trend. Instead, it shows a broad optimum around $\text{OLV}_{\text{cont.}}^{\text{max}}$=0.9, and a very sharp and isolated peak at $\text{OLV}_{\text{cont.}}$=0.5. This two-peak structure seems to mimic the behaviour seen earlier in Fig. 10.1. Indeed, the correlation plot in Fig. 10.5 shows that these are the same points, located at m_{jj}^{min}=800 GeV and $\text{OLV}_{\text{cont.}}^{\text{max}}$=0.5. This configuration corresponds to extremely tight cuts, selecting signal regions very high in purity and extremely low
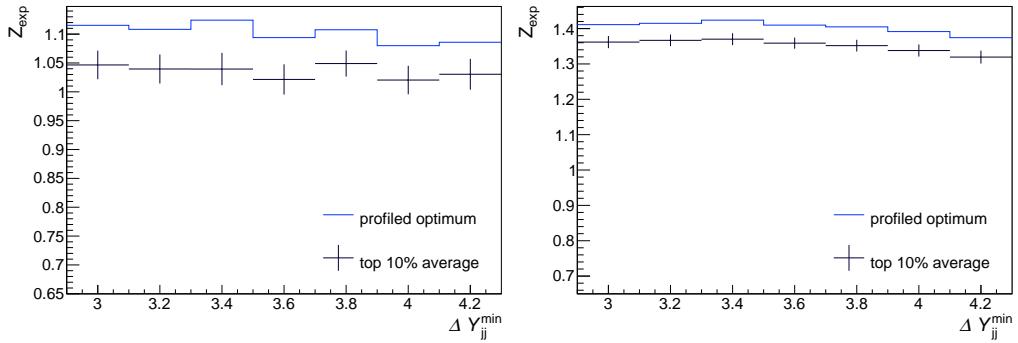
Figure 10.6: Significance profiles for a lower cut on the dijet rapidity gap Δy_{jj}, using the full likelihood fit method (left) and the simple Poisson approximation (right). The cut value chosen for the analysis presented at Moriond 2013 was $\Delta y_{jj}^{\min}{=}2.8$ [53].

in event yield. The total event yield of several Monte Carlo background samples approaches zero in this region, hence it does not seem advisable to use this optimum, let alone the fact that the event yields of the signal regions are well below one event for this configuration. To avoid the risk of selecting an empty signal region, this maximum is ignored for now. However, for future studies, it seems advisable to implement means of protecting the optimization against this effect.

Summarizing the conclusions drawn in the previous paragraphs, the optimal selection should employ

$$m_{jj}^{\min}{=}500\,\mathrm{GeV},\; m_{jj}^{\mathrm{split}}{=}1100\,\mathrm{GeV},\; \mathrm{and}\; \mathrm{OLV}_{\mathrm{cont.}}^{\max}{=}0.9.$$

The remaining variable taken into consideration in the present study is the dijet rapidity gap Δy_{jj}, on which a lower cut is applied. The corresponding optimization results are shown in Fig. 10.6. The observed shape is rather flat and does not exhibit any strong features like the ones presented previously. Additionally, the Poisson and likelihood fit estimators show a similar trend. Hence, the best value at

$$\Delta y_{jj}^{\min}{=}3.4$$

can be chosen without the need of a detailed discussion.

10.3 Optimization for the same flavour channel

The same flavour channel suffers from a large amount of Drell-Yan background and hence poses a more challenging terrain for an analysis. Nevertheless, as discussed in Chapter 8, powerful discriminant variables allow the efficient removal of these

background events. While the cut values on some of these variables, like the $m_{\ell\ell}$-window cut for the Z-veto, are rather obvious, the same is not true for other variables like the missing transverse energy and its derived quantities $E_{\mathrm{T,STVF}}^{\mathrm{miss}}$ and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$. For a same-flavour optimization, a set of these variables has been considered in a dedicated study presented in this section.

10.3.1 Properties of the study

As discussed previously, the number of systematic variations that can be taken into account during the optimization suffers some practical limitations. Hence, based on the ranking of systematic uncertainties presented in Section 9.2, the leading two systematic uncertainties

- η JES Modelling and
- JES Modelling,

and the luminosity uncertainty have been included in the likelihood fit for this optimization study. Additionally, limitations of disk storage space necessitate the application of a cut requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$ on the preselection level. Hence, no cut values below that were considered. Apart from that, the remaining parameters of the analysis were chosen according to the results obtained in the previous section. Especially, also the m_{jj}-splitting was performed at $m_{jj} = 1.1\,\mathrm{TeV}$.

The parameters taken into account were lower cuts on

- the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$,
- the relative missing transverse energy $E_{\mathrm{T,rel}}^{\mathrm{miss}}$, and
- the soft-term corrected missing transverse energy $E_{\mathrm{T,STVF}}^{\mathrm{miss}}$.

Since these observables are expected to be highly correlated, this choice of optimization variables exhibits the strength of the presented approach, allowing the simultaneous optimization of an entire parameter group.

In order to be consistent with the previously used notation, the lower cut values are identified by appending an upper index "min" to the quantities.

The significance profiles shown for this study exhibit an additional feature that was not present in any of the previously presented visualizations: the display of underflow bins. The lowest bin shown on each axis is an *underflow bin*. The significance value shown for these bins does not correspond to a cut value of the central bin coordinate – instead, the contents of these bins correspond to the *removal* of the respective cut, or equivalently a cut value of negative infinity.

10.3.2 Results of the optimization

Figure 10.7 shows the expected significance profiled as a function of $E_\mathrm{T}^{\mathrm{miss,min}}$. First of all, it is interesting to note a clear and distinct optimum at $E_\mathrm{T}^{\mathrm{miss,min}}{=}60\,\mathrm{GeV}$. This is rather surprising, as the analysis presented at Moriond 2013 [53] only employs a preselection cut on $E_\mathrm{T}^{\mathrm{miss}}$, and instead uses the allegedly more pile-up resistant variables $E_\mathrm{T,rel}^{\mathrm{miss}}$ and $E_\mathrm{T,STVF}^{\mathrm{miss}}$. The respective significance profiles of the latter are shown in Figures 10.8 and 10.9. However, the significance profile for the lower cut on $E_\mathrm{T,rel}^{\mathrm{miss}}$ exhibits a sharp decrease in significance towards tighter cuts, questioning the value of this variable as compared to the (standard) missing transverse energy $E_\mathrm{T}^{\mathrm{miss}}$. It is interesting to note that the Moriond analysis cut meets an albeit local optimum. However, the optimization clearly shows that the removal of the $E_\mathrm{T,rel}^{\mathrm{miss}}$ cut is favourable. The cut on $E_\mathrm{T,STVF}^{\mathrm{miss}}$, on the other hand, seems to have very little effect on the overall significance.

The two-dimensional significance profile of $E_\mathrm{T}^{\mathrm{miss,min}}$ and $E_\mathrm{T,rel}^{\mathrm{miss,min}}$ shown in Fig. 10.10 indicates clearly that the a relatively tight cut at $E_\mathrm{T}^{\mathrm{miss}}{=}60\,\mathrm{GeV}$ completely removes the need of applying a cut on $E_\mathrm{T,rel}^{\mathrm{miss}}$. Interestingly, the configuration chosen in Ref. [53] provides a moderate expected significance using the Poisson estimator, but is clearly not optimal when employing the likelihood fit.

Judging from the two-dimensional significance profile of $E_\mathrm{T}^{\mathrm{miss,min}}$ and $E_\mathrm{T,STVF}^{\mathrm{miss,min}}$, shown in Fig. 10.11, the Poisson estimator claims an anticorrelation between the optimal cut values for $E_\mathrm{T}^{\mathrm{miss}}$ and $E_\mathrm{T,STVF}^{\mathrm{miss}}$ in the sense that tightening the cut on one in general requires to loosen the cut on the other in order to maintain a comparable significance. However, the likelihood estimator does not confirm this property. Instead, a tight cut on $E_\mathrm{T}^{\mathrm{miss}}$ seems to outperform any kind of cut on $E_\mathrm{T,STVF}^{\mathrm{miss}}$.

Figure 10.12 shows the two-dimensional significance profile of $E_\mathrm{T,rel}^{\mathrm{miss}}$ and $E_\mathrm{T,STVF}^{\mathrm{miss}}$. While there is no clear preference for any cut value chosen for $E_\mathrm{T,STVF}^{\mathrm{miss,min}}$ to the point that even removing the cut has close to no effect on the final result, loosening or even removing the cut on $E_\mathrm{T,rel}^{\mathrm{miss,min}}$ generates a remarkable increase in expected significance for the final result.

Following a conservative approach, it seems reasonable to remove the $E_\mathrm{T,STVF}^{\mathrm{miss}}$ and $E_\mathrm{T,rel}^{\mathrm{miss}}$ cuts entirely in order to keep the analysis as simple as possible and achieve a maximum of statistical power for the signal region. Hence, both cuts are rejected and only a cut on

$$E_\mathrm{T}^{\mathrm{miss}} > 60\,\mathrm{GeV}$$

is applied for the same flavour channel. The results from these optimization studies are employed in the measurement of the cross section for VBF Higgs production, as presented in the following chapter.
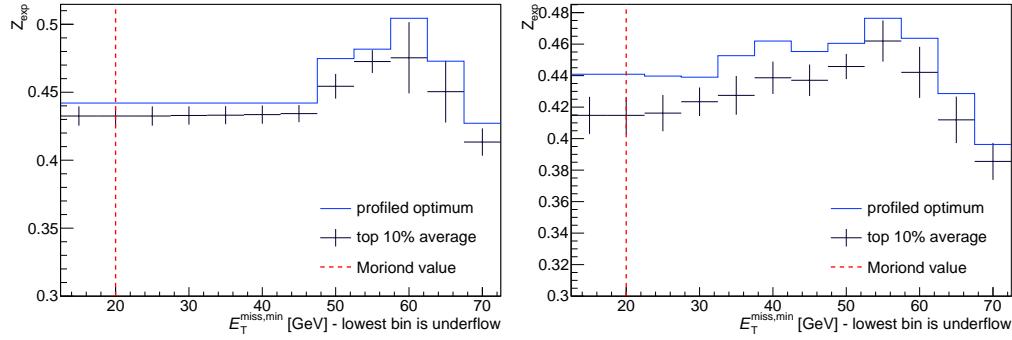
Figure 10.7: Significance profiles for the lower cut on the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, using the full likelihood fit method (left) and the simple Poisson approximation (right). The dashed red line corresponds to the cut value used for the analysis presented at Moriond 2013 [53].
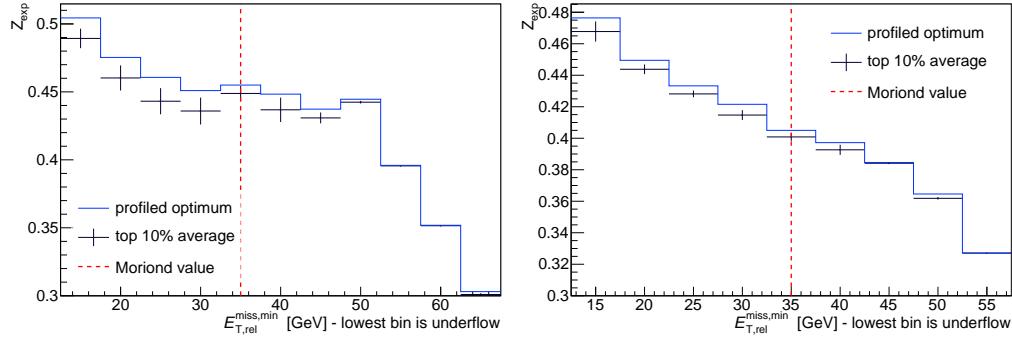
Figure 10.8: Significance profiles for the lower cut on the relative missing transverse energy $E_{\mathrm{T,rel}}^{\mathrm{miss}}$, using the full likelihood fit method (left) and the simple Poisson approximation (right). The dashed red line corresponds to the cut value used for the analysis presented at Moriond 2013 [53].
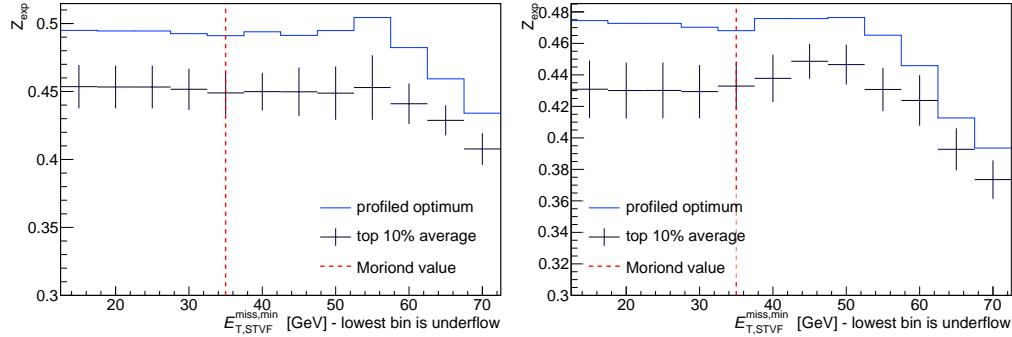
Figure 10.9: Significance profiles for the lower cut on the soft-term corrected missing transverse energy $E_{\mathrm{T,STVF}}^{\mathrm{miss}}$, using the full likelihood fit method (left) and the simple Poisson approximation (right). The dashed red line corresponds to the cut value used for the analysis presented at Moriond 2013 [53].
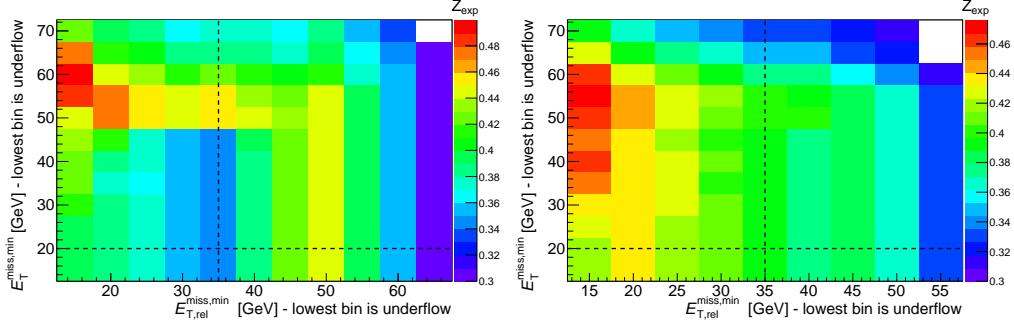
Figure 10.10: Two-dimensional top 10% average significance profiles for lower cuts on the standard and relative missing transverse energy E_T^{miss} and $E_{T,rel}^{miss}$, using the full likelihood fit method (left) and the simple Poisson approximation (right). The significance is colour-coded, values increasing from blue to red. The dashed black lines correspond to the cut values used for the analysis presented at Moriond 2013 [53].
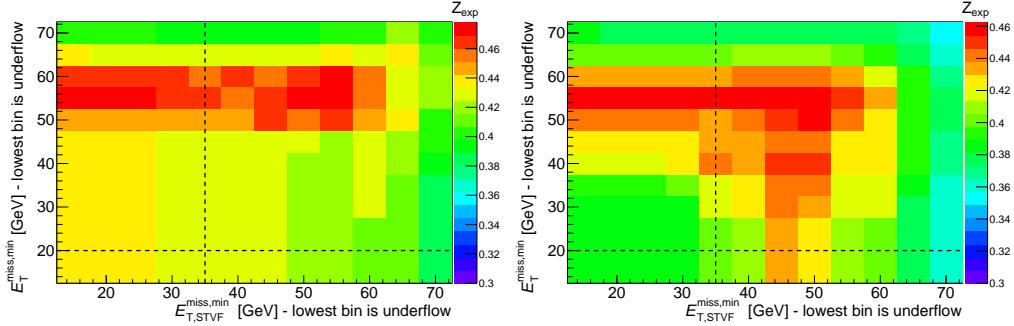
Figure 10.11: Two-dimensional top 10% average significance profiles for lower cuts on the standard and soft-term corrected missing transverse energy E_T^{miss} and $E_{T,STVF}^{miss}$, using the full likelihood fit method (left) and the simple Poisson approximation (right). The significance is colour-coded, values increasing from blue to red. The dashed black lines correspond to the cut values used for the analysis presented at Moriond 2013 [53].
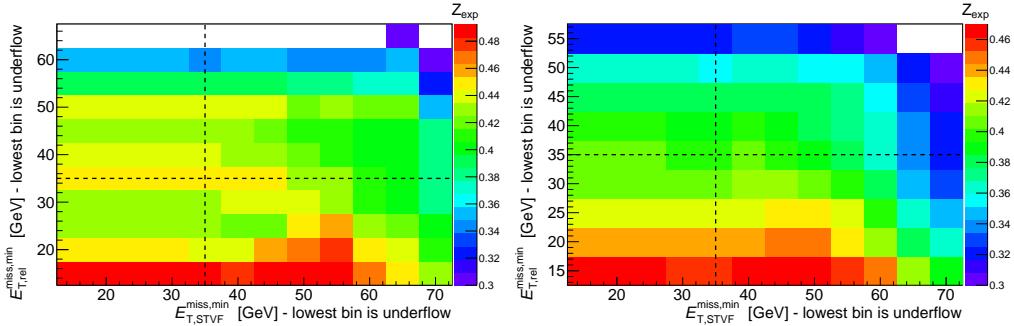
Figure 10.12: Two-dimensional top 10% average significance profiles for lower cuts on the relative and the soft-term corrected missing transverse energy $E_{T,rel}^{miss}$ and $E_{T,STVF}^{miss}$, using the full likelihood fit method (left) and the simple Poisson approximation (right). The significance is colour-coded, values increasing from blue to red. The dashed black lines correspond to the cut values used for the analysis presented at Moriond 2013 [53].

11 Cross Section Measurement of VBF Higgs Boson Production

In this chapter, a measurement of the cross section for VBF Higgs boson production is presented, employing the event selection presented in Chapters 8 and 10. The full event selection including the results from the optimization is displayed in Tables 11.1 and 11.2, together with the event yields at the corresponding cut stages. More detailed tables differentiating between the various background samples can be found in Appendix 12.

The measurement is performed by conducting a likelihood fit including all systematic uncertainties listed in Section 9.2. As previously stated, an inclusion of theoretical uncertainties is not feasible, since they in general depend on the employed event selection. The omitted uncertainties include, for example, the uncertainty on the ggF Higgs production cross-section and extrapolation uncertainties between the signal and control regions.

For the fit, only the signal region and top control regions have been used, and a free normalization factor has been attributed to the simulated Monte Carlo samples for top background to allow the fit to constrain the top contribution with information from the corresponding control region. Tables 11.3 and 11.4 show the respective event yields in the top control region for different and same flavour. It becomes apparent that the top normalization factor is bound to be smaller than unity, so that the top contribution can be adjusted to fit the observations.

A total of eight signal regions and two control regions is fitted simultaneously. The signal region is split into the four lepton flavour sub-channels, each subdivided into low and high m_{jj} regions by the split boundary optimized in Section 10.2. The top control region is also split in m_{jj}, but inclusive with respect to all four lepton flavour sub-channels considered. Contributions from other Higgs boson production modes have been treated as a background with a normalization according to the predicted cross sections.

$e\mu + \mu e$	VBF Signal	Total Bkg.	Observed
preselection	21.85 ± 0.19	8226.91 ± 39.12	7891.00 ± 88.83
$m_{jj} > 500\,\text{GeV}$	10.76 ± 0.14	516.88 ± 6.07	488.00 ± 22.09
$\Delta y_{jj} > 3.4$	10.14 ± 0.13	376.43 ± 4.97	326.00 ± 18.06
CJV	8.19 ± 0.12	164.04 ± 3.35	142.00 ± 11.92
$\text{OLV}_{\text{cont.}} < 0.9$	7.76 ± 0.12	113.78 ± 2.78	103.00 ± 10.15
$p_{\text{T}}^{\text{tot}} < 50$	7.43 ± 0.11	99.64 ± 2.59	97.00 ± 9.85
$Z \to \tau\tau$ veto	6.83 ± 0.11	88.51 ± 2.39	80.00 ± 8.94
$m_{\ell\ell} < 50\,\text{GeV}$	5.41 ± 0.10	19.53 ± 1.21	21.00 ± 4.58
$\Delta\phi_{\ell\ell} < 1.8$	5.16 ± 0.09	16.58 ± 1.04	18.00 ± 4.24
$m_{jj} \leq 1.1\,\text{TeV}$	3.43 ± 0.08	14.26 ± 0.98	16.00 ± 4.00
$m_{jj} > 1.1\,\text{TeV}$	1.74 ± 0.05	2.32 ± 0.37	2.00 ± 1.41

Table 11.1: Expected and observed event yields for the different-flavour channel at various cut stages. A corresponding breakdown of the various background estimates at each cut stage can be found in Table 1 of Appendix 12.

$ee + \mu\mu$	VBF Signal	Total Bkg.	Observed
preselection	22.50 ± 0.20	63972.58 ± 313.48	65242.00 ± 255.43
$E_{\text{T}}^{\text{miss}} > 60\,\text{GeV}$	10.57 ± 0.14	5196.61 ± 52.58	4968.00 ± 70.48
$m_{jj} > 500\,\text{GeV}$	5.65 ± 0.10	391.80 ± 7.57	384.00 ± 19.60
$\Delta y_{jj} > 3.4$	5.26 ± 0.10	241.31 ± 5.60	247.00 ± 15.72
CJV	4.29 ± 0.09	102.51 ± 4.43	113.00 ± 10.63
$\text{OLV}_{\text{cont.}} < 0.9$	4.00 ± 0.08	70.88 ± 4.16	77.00 ± 8.77
$p_{\text{T}}^{\text{tot}} < 50$	3.81 ± 0.08	54.94 ± 1.85	62.00 ± 7.87
$p_{\text{T}}^{\ell\ell,\text{jets}} > 25$	3.80 ± 0.08	54.14 ± 1.83	61.00 ± 7.81
$Z \to \tau\tau$ veto	3.48 ± 0.08	48.98 ± 1.69	54.00 ± 7.35
$f_{\text{recoil}} < 0.2$	2.71 ± 0.07	29.03 ± 1.22	29.00 ± 5.39
$m_{\ell\ell} < 50\,\text{GeV}$	2.14 ± 0.06	7.50 ± 0.66	7.00 ± 2.65
$\Delta\phi_{\ell\ell} < 1.8$	1.95 ± 0.06	6.75 ± 0.64	4.00 ± 2.00
$m_{jj} \leq 1.1\,\text{TeV}$	1.24 ± 0.05	5.40 ± 0.59	3.00 ± 1.73
$m_{jj} > 1.1\,\text{TeV}$	0.71 ± 0.04	1.35 ± 0.24	1.00 ± 1.00

Table 11.2: Expected and observed event yields for the same-flavour channel at various cut stages. A corresponding breakdown of the various background estimates at each cut stage can be found in Table 2 of Appendix 12.

$e\mu + \mu e$	top Background	Other Processes	Observed
preselection	42526.03 ± 67.02	897.62 ± 13.86	44328.00 ± 210.54
$m_{jj} > 500\,\mathrm{GeV}$	2300.06 ± 9.77	30.62 ± 2.54	1892.00 ± 43.50
$\Delta y_{jj} > 3.4$	1196.28 ± 6.63	12.22 ± 1.60	906.00 ± 30.10
CJV	260.52 ± 3.32	5.25 ± 0.86	159.00 ± 12.61
$\mathrm{OLV_{cont.}} < 0.9$	118.07 ± 2.37	3.42 ± 0.55	86.00 ± 9.27
$p_{\mathrm{T}}^{\mathrm{tot}} < 50$	106.67 ± 2.29	3.00 ± 0.47	75.00 ± 8.66
$Z \to \tau\tau$ veto	99.29 ± 2.23	2.55 ± 0.43	68.00 ± 8.25
$m_{jj} \leq 1.1\,\mathrm{TeV}$	93.36 ± 2.18	1.75 ± 0.40	67.00 ± 8.19
$m_{jj} > 1.1\,\mathrm{TeV}$	5.93 ± 0.48	0.80 ± 0.16	1.00 ± 1.00

Table 11.3: Expected and observed event yields for the different-flavour channel in the top control region at various cut stages.

$ee + \mu\mu$	top Background	Other Processes	Observed
preselection	33846.95 ± 59.88	12040.05 ± 145.64	48276.00 ± 219.72
$E_{\mathrm{T}}^{\mathrm{miss}} > 60\,\mathrm{GeV}$	20514.52 ± 46.25	695.56 ± 23.00	21532.00 ± 146.74
$m_{jj} > 500\,\mathrm{GeV}$	1260.66 ± 7.20	52.15 ± 3.15	1057.00 ± 32.51
$\Delta y_{jj} > 3.4$	638.41 ± 4.82	15.97 ± 1.80	483.00 ± 21.98
CJV	134.10 ± 2.32	5.34 ± 0.88	86.00 ± 9.27
$\mathrm{OLV_{cont.}} < 0.9$	58.91 ± 1.61	2.79 ± 0.58	39.00 ± 6.24
$p_{\mathrm{T}}^{\mathrm{tot}} < 50$	52.22 ± 1.53	2.19 ± 0.50	32.00 ± 5.66
$p_{\mathrm{T}}^{\ell\ell,\mathrm{jets}} > 25$	52.08 ± 1.52	2.15 ± 0.50	32.00 ± 5.66
$Z \to \tau\tau$ veto	49.44 ± 1.49	1.54 ± 0.36	27.00 ± 5.20
$f_{\mathrm{recoil}} < 0.2$	28.53 ± 1.12	0.84 ± 0.25	17.00 ± 4.12
$m_{jj} \leq 1.1\,\mathrm{TeV}$	26.75 ± 1.10	0.36 ± 0.21	17.00 ± 4.12
$m_{jj} > 1.1\,\mathrm{TeV}$	1.78 ± 0.21	0.49 ± 0.13	0.00 ± 0.00

Table 11.4: Expected and observed event yields for the same-flavour channel in the top control region at various cut stages.

The resulting expected discovery significance for a Higgs boson of $m_H = 125\,\mathrm{GeV}$ in the VBF production mode is

$$Z_{\mathrm{exp}} = 1.68 \qquad\qquad (p_0 = 0.047),$$

which corresponds to a slight increase of the order of 5% compared to the results presented in Ref. [53]. However, since for the present result, no theoretical uncertainties have been taken into account, the values are not directly comparable.

The nuisance parameter pulls (see Section 9.2.2) encountered during the likelihood fit are visualized by Fig. 11.1. The individual parameters have been briefly introduced in Section 9.2. The only exemption from this is the *top normalization*
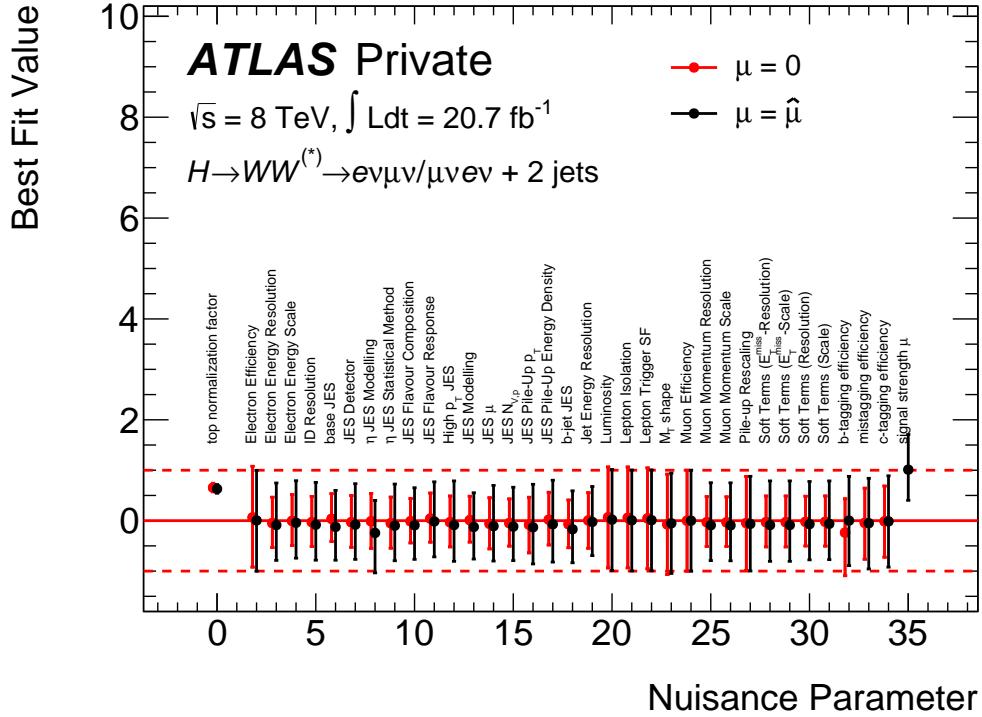
Figure 11.1: Nuisance parameter pulls (see Section 9.2.2) encountered during the likeli-
hood fit conducted for the cross-section measurement. The black markers
correspond to the fully unconstrained likelihood fit with a free signal strength
parameter μ, whereas the red markers correspond to the partially constrained
case where $\mu = 0$ has been fixed.

factor, which is constrained by the fit from the top control region (see Section 8.2.2).
Apart from the latter, no other control regions have been used in the fit, and
no other free normalization factors have been considered. The pulls shown are
generally well-behaved, and apart from the top normalization factor and the signal
strength μ, no other value is significantly pulled or overconstrained.

Interestingly, the value of Z_{exp} is higher than the sum in quadrature of the optimal
values obtained in the optimization studies in Chapter 10. This can be explained
by the fact that for this result, a combined fit of all lepton flavour channels is
performed, potentially containing more information and thus yielding increased
precision.

The corresponding observed values for Z and p are

$$Z_{\mathrm{obs}} = 1.78 \qquad\qquad (p_0 = 0.037).$$

The best fit signal strength μ_{obs} for VBF Higgs boson production is

$$\mu_{\mathrm{obs}} = 1.01^{+0.58}_{-0.52}(\text{stat.})^{+0.35}_{-0.29}(\text{syst.})^{+0.15}_{-0.09}(\text{norm.})$$
$$= 1.01^{+0.70}_{-0.61}$$

where the suffix *stat.* refers to the purely statistical uncertainty of the result, *syst.* refers to the uncertainty resulting from free nuisance parameters due to systematic variations, and *norm.* refers to the uncertainty on the top normalization factor. The numeric values of the respective uncertainties correspond to the one sigma standard Gaussian confidence interval, derived from the profile of the likelihood function with respect the signal strength parameter.

Given that the theory uncertainty on the product of the cross section σ for VBF Higgs boson production and the branching ratio $\mathcal{B}$ for $H \to W^{\mp}W^{\pm(*)} \to \ell^-\bar{\nu}\ell'^+\nu'$ decay is of the order of 5% [105], it can be neglected compared to the experimental uncertainties of up to 70%. The measured value of the convoluted production cross-section and branching ratio is

$$\sigma_{\mathrm{obs}} \cdot \mathcal{B} = 0.036^{+0.025}_{-0.022}, \qquad\qquad [\text{pb}]$$

summarizing all sources of uncertainties. The corresponding expected value is

$$\sigma_{\mathrm{exp}} \cdot \mathcal{B} = 0.035^{+0.028}_{-0.022}, \qquad\qquad [\text{pb}]$$

which is in excellent agreement with the measurement.

12   Summary and Conclusions

On July 4th 2012, the ATLAS and CMS collaborations at CERN in Geneva have announced the discovery of a new boson with a mass of $m \approx 125\,\text{GeV}$ in proton-proton collisions at the LHC. All studies conducted so far exhibit compatibility of the new boson with the sought-for Standard Model Higgs boson, the elusive piece of evidence for the mechanism of electroweak symmetry breaking proposed by François Englert, Robert Brout, Peter Higgs, Gerald Guralnik, Carl Hagen and Tom Kibble in 1964.

This thesis presents the search for evidence of Higgs boson production through Vector Boson Fusion. Events with missing transverse energy, two oppositely charged light leptons and at least two highly energetic jets have been investigated to study VBF Higgs boson production in events where the Higgs boson decays to a pair of oppositely charged W bosons, which in turn decay to one charged lepton and one neutrino each. Earlier studies conducted by a dedicated search group in the ATLAS collaboration have been revisited and optimized.

For this purpose, a *merging* procedure to exploit the finite Monte Carlo samples more efficiently in order to achieve a smaller statistical uncertainty for the analysis has been successfully developed, applied and validated. The *merging* procedure can be extended to other types of simulated samples and different filter definitions. However, especially analyses with a very distinct and rare final state can expect to gain statistical power on background estimates based on Monte Carlo simulations.

The event selection employed for the analysis has been optimized. Observables used for the event selection have been revisited and modified. Groups of threshold values for potentially correlated observables employed for the event selection have been studied, fully sampling the surrounding regions of configuration space and investigating local maxima of the expected statistical significance. For the purpose of evaluating the significance for each configuration, a full likelihood fit including the leading systematic uncertainties has been used and compared to a much simpler approach of estimating the significance using the arithmetic Poisson approximation. Significant discrepancies between both approaches have been found, leading to the conclusion that the likelihood fit method should be used for future optimization studies wherever possible.

A versatile software framework has been implemented to provide general optimization functionality for a vast variety of analyses, employing and extending the functionality of the `HWWAnalysisCode` and using existing solutions for the statistical analysis as far as possible. The `GridScan` optimization framework is widely applicable for many analyses, providing convenient means of optimizing any traditional cut-based analysis. While the simultaneous global optimization of an entire event selection including all systematic uncertainties is still prohibited by performance issues, partial studies already allow insights in the correlations between different observables with respect to their optimal values. Nevertheless, the feasibility of the likelihood fit method including systematic uncertainties for an optimization study has been proven, at the same time quantifying the discrepancies towards simpler means of significance estimation. In the course of these studies, a number of potentially beneficial features for future implementation has been identified, including ways to automatically account for poor event yields, optimization of a higher number of bin boundaries, and generally increased performance with respect to runtime, memory usage and disk space.

The optimized event selection and the aforementioned *merging* procedure have been used for a statistical analysis of data taken by the ATLAS Detector in the year 2012 in *pp*-collisions with a centre-of-mass energy of $\sqrt{s} = 8\,\mathrm{TeV}$ at the LHC. The signal strength and convoluted cross section and branching ratio of VBF Higgs boson production have been measured in the $H \to W^{\mp}W^{\pm(*)} \to \ell^- \bar{\nu}\ell'^+ \nu'$ decay mode and found in excellent agreement with Standard Model predictions. The expected and observed discovery significances were

$$Z_{\mathrm{exp}} = 1.68 \qquad\qquad (p_0 = 0.047)$$
$$Z_{\mathrm{obs}} = 1.78 \qquad\qquad (p_0 = 0.037)$$

The observed signal strength μ was

$$\mu_{\mathrm{obs}} = 1.01^{+0.58}_{-0.52}(\mathrm{stat.})^{+0.35}_{-0.29}(\mathrm{syst.})^{+0.15}_{-0.09}(\mathrm{norm.}) \qquad\qquad = 1.01^{+0.70}_{-0.61}$$

The observed and expected convoluted cross sections and branching ratios were

$$\sigma_{\mathrm{obs}} \cdot \mathcal{B} = 0.036^{+0.025}_{-0.022} \qquad\qquad [\mathrm{pb}]$$
$$\sigma_{\mathrm{exp}} \cdot \mathcal{B} = 0.035^{+0.028}_{-0.022} \qquad\qquad [\mathrm{pb}]$$

The uncertainties derived on these values are slightly lower than the ones obtained in earlier studies by the ATLAS collaboration, but do not include all sources of theoretical uncertainties such as the gluon fusion cross-section uncertainty.

Background Estimates

Tables 11.1 and 11.2 in Chapter 11 present an overview over the expected and observed event yields at various stages of the event selection. In this section, more detailed tables are provided, including details on event yields from individual background estimates. The corresponding numbers are shown in Tables 1 and 2.

$e\mu + \mu e$	ggF	top	diboson	Z+jets	W+jets
preselection	59.46 ± 0.85	5006.73 ± 25.60	1066.42 ± 12.91	1725.07 ± 25.92	353.52 ± 5.88
$m_{jj} > 500\,\text{GeV}$	5.03 ± 0.25	343.25 ± 4.25	73.62 ± 1.93	78.90 ± 3.59	15.78 ± 1.43
$\Delta y_{jj} > 3.4$	4.43 ± 0.23	258.09 ± 3.56	50.40 ± 1.52	51.00 ± 2.84	12.37 ± 1.27
CJV	2.27 ± 0.17	97.19 ± 2.34	33.69 ± 1.17	23.29 ± 1.86	7.51 ± 0.95
$\text{OLV}_{\text{cont.}} < 0.9$	2.02 ± 0.16	66.15 ± 1.93	23.75 ± 0.92	15.62 ± 1.54	6.24 ± 0.86
$p_{\text{T}}^{\text{tot}} < 50$	1.81 ± 0.15	58.46 ± 1.83	21.92 ± 0.88	12.53 ± 1.38	4.93 ± 0.78
$Z \to \tau\tau$ veto	1.71 ± 0.14	54.11 ± 1.77	20.13 ± 0.84	8.48 ± 1.14	4.09 ± 0.72
$m_{\ell\ell} < 50\,\text{GeV}$	1.47 ± 0.13	10.07 ± 0.82	3.24 ± 0.37	2.82 ± 0.66	1.92 ± 0.44
$\Delta\phi_{\ell\ell} < 1.8$	1.38 ± 0.13	9.59 ± 0.81	3.01 ± 0.35	0.84 ± 0.34	1.76 ± 0.42
$m_{jj} \le 1.1\,\text{TeV}$	1.12 ± 0.12	8.21 ± 0.74	2.47 ± 0.34	0.72 ± 0.32	1.75 ± 0.42
$m_{jj} > 1.1\,\text{TeV}$	0.26 ± 0.06	1.37 ± 0.32	0.55 ± 0.09	0.12 ± 0.12	0.01 ± 0.07

Table 1: Expected and observed event yields for the background estimates in the different-flavour channel at various cut stages.

$ee + \mu\mu$	ggF	top	diboson	Z+jets	W+jets
preselection	60.68 ± 0.86	3950.05 ± 22.72	906.86 ± 9.76	58689.36 ± 312.30	350.02 ± 11.16
$E_{\text{T}}^{\text{miss}} > 60\,\text{GeV}$	26.45 ± 0.56	2403.77 ± 17.59	475.55 ± 6.79	2208.93 ± 48.94	74.74 ± 3.73
$m_{jj} > 500\,\text{GeV}$	2.77 ± 0.18	190.79 ± 3.15	40.38 ± 1.34	155.83 ± 6.68	1.98 ± 0.96
$\Delta y_{jj} > 3.4$	2.14 ± 0.16	137.78 ± 2.57	25.30 ± 1.01	75.59 ± 4.81	0.50 ± 0.73
CJV	1.05 ± 0.11	50.94 ± 1.52	16.14 ± 0.75	33.73 ± 4.06	0.66 ± 0.54
$\text{OLV}_{\text{cont.}} < 0.9$	0.82 ± 0.10	34.99 ± 1.25	11.99 ± 0.62	22.64 ± 3.90	0.43 ± 0.43
$p_{\text{T}}^{\text{tot}} < 50$	0.76 ± 0.10	30.67 ± 1.18	10.83 ± 0.57	12.27 ± 1.26	0.41 ± 0.36
$p_{\text{T}}^{\ell\ell,\text{jets}} > 25$	0.76 ± 0.10	30.61 ± 1.17	10.81 ± 0.57	11.61 ± 1.23	0.34 ± 0.35
$Z \to \tau\tau$ veto	0.73 ± 0.09	29.12 ± 1.16	10.44 ± 0.56	8.55 ± 1.04	0.14 ± 0.30
$f_{\text{recoil}} < 0.2$	0.44 ± 0.07	18.84 ± 0.93	7.06 ± 0.46	2.65 ± 0.60	0.03 ± 0.23
$m_{\ell\ell} < 50\,\text{GeV}$	0.30 ± 0.06	4.48 ± 0.44	1.25 ± 0.18	1.43 ± 0.42	0.04 ± 0.16
$\Delta\phi_{\ell\ell} < 1.8$	0.28 ± 0.06	4.06 ± 0.43	1.17 ± 0.17	1.23 ± 0.41	0.01 ± 0.14
$m_{jj} \le 1.1\,\text{TeV}$	0.23 ± 0.05	3.22 ± 0.39	0.85 ± 0.16	1.13 ± 0.40	-0.04 ± 0.11
$m_{jj} > 1.1\,\text{TeV}$	0.05 ± 0.03	0.83 ± 0.20	0.32 ± 0.06	0.09 ± 0.08	0.05 ± 0.09

Table 2: Expected and observed event yields for the background estimates in the same-flavour channel at various cut stages.

Common Analysis Framework

The studies presented in this thesis make use of the large code base of the data analysis framework `ROOT` [106], extended by tools from the sub-projects `RooFit`, `RooStats` and the HSG3 Common Analysis Framework [107]. The author is one of the main contributors to the latter software framework, and the software written for the purpose of these studies has been contributed to the Framework. This chapter gives a brief introduction into the motivation for and functionality of this software framework, with special focus on the implementation of the tools used for this thesis.

1 The `HWWAnalysisCode` package

While the solution to many of the problems faced in a high-profile data analysis are straight-forward, their implementation does not often share this feature. One of the main reasons for this is the need for *meta-data* management. For example, it is not sufficient to store the data and Monte Carlo samples (*datasets* or *nTuples* is HEP-parlance). A successful analysis needs additional information on the Generators used, the respective cross-sections of the processes stored in each sample, and the equivalent luminosities for the sake of normalization.

Over time, a large number of software packages has been spawned by physicists with the aim of providing assistance for data analysis. However, most of these software packages have a focus on data preprocessing steps like reconstruction or object selection. Only comparably few software packages are dedicated to aid the final step of the analysis – the event selection, the optimization thereof, and the statistical evaluation of the resulting data. As a result, the physicist is often left with the task to perform manual metadata management for the final steps of the analysis – a process which is both cumbersome and error-prone.

For this exact reason, the `HWWAnalysisCode` Package, the final processing stage of the Common Analysis Framework (CAF) of the HSG3 was designed as a modular and flexible toolkit, providing an integrated solution for implementing a complex event selection in a simple and highly modular fashion, including many classes for automated metadata management and inspection. The design considerations for this software framework are detailed elsewhere [108], and the documentation

is available online [107] in a fashion similar to the documentation of the analysis framework ROOT, to which the `HWWAnalysisCode` is to be considered an extension.

Despite the name, the `HWWAnalysisCode` is fairly unspecific to the analysis of Higgs decays in the $H \to W^{\mp}W^{\pm(*)} \to \ell^{-}\bar{\nu}\ell'^{+}\nu'$ mode and can be applied to a broad variety of data analysis, implementing virtually any nTuple-based event selection.

As for HSG3, the `HWWAnalysisCode` is the official toolbox employed to produce any publication results. It was also used to produce the histograms shown in Chapters 5 and 8 of this thesis. Like with most scientific software, the code as well as the documentation are under constant development.

2 `Freiburg Statistics Code`

The Freiburg-based Statistics Code is a rather recent extension to the CAF and is not currently used for official ATLAS publications. Nevertheless, it is a powerful tool interfacing the metadata handling intrinsic to the `HWWAnalysisCode` with the statistical evaluation capabilities of the `RooStats` framework and provides a wealth of pre-implemented functionality. It is highly configurable and has been used for the statistical analyses presented in Chapters 9, 10 and 11.

3 `GridScan` optimization framework

The GridScan Optimization Framework is a component of the `HWWAnalysisCode`. However, since its development was part of the work associated with this thesis, it is listed as a separate component here in order to provide additional details on the implementation.

The entire optimization process described in Chapter 10, starting from the nTuples and ending with the production of the optimization plots shown in Sections 10.2 and 10.3, has been implemented in a highly modular and configurable fashion. The GridScan code has been contributed to the `HWWAnalysisCode` with the aim of providing an integrated cut optimization toolking, allowing anyone familiar with `ROOT` to perform similar studies like the one presented in this thesis with minimal effort.

The entire optimization process consists of three separate steps, which are performed by different compiled binaries. All of them are configured via a central configuration file (`GridScan.cfg`), the name of which can be altered and passed to the binary call as an argument if neccessary. These steps are:

1. `PrepareGrid`. The code will read in a `samples.root` file (the typical output format produced by the `HWWAnalysisCode`), locate the input files and perform the task of creating and filling the n-dimensional histograms (`CounterGrid`s), saving them to an output file `grids.root`. Depending on the choice of the dimensionality and binning of the configuration space, the file size of the latter can be excessively large. While the data storage efficiency of the output format is constantly being improved, the intrinsic problem of the excessive size is directly correlated with the increasing volume of high-dimensional spaces and cannot be overcome.

2. `ScanGrid`. The `grids.root` file is read in. This step performs the optimization by looping over all configurations possible given the finite structure of the `CounterGrid`. Each point in configuration space will be evaluated and pushed to a list, including the results of the evaluation. This step is highly configurable in so far as the statistical evaluation procedure is arbitrary. The user can choose to use a simple (and fast) estimation procedure like the Poisson estimator introduced in Section 9.1.2, but an evaluation with the full Likelihood fit machinery is also possible, employing the capabilities of the Freiburg Statistics Code for this purpose[1]. The resulting data structure will be written to another file `results.root`.

3. `MakePlots`. This rather lightweight application will read in the `results.root` from the previous step and produce plots depicting the distributions of points in the configuration space sorted by quality as well as the significance profiles presented in Section 10.1.3.

The code comes with default configuration files, which are heavily documented. The code of the `GridScan` optimization framework are documented as a part of the `HWWAnalysisCode` [107].

---

[1]Since the Freiburg Statistics Code merely provides an interface the `RooFit`/`RooStats` framework, the latter takes care of performing the likelihood fit using the MINUIT2 [103] minimization algorithm. However, an open memory leak [109] in the `HistFactory` tool currently inhibits the computational efficiency of the Freiburg Statistics Code and therefore the GridScan cut optimization framework from.

Monte Carlo Sample Merge

This appendix chapter contains additional validation and control plots for the merging procedure described in Chapter 7.

4 Alpgen Z+jets sample

This section contains auxiliary plots for the discussion of the merging procedure for the ALPGEN Z+jets sample presented in Section 7.4.

4.1 Z control region

Figure 1 the same distributions as Fig. 7.2, but after application of the event selection level jets cuts.

4.2 Signal region

Figure 2 shows the signal region plots corresponding to the ones shown in Fig. 7.2.

Figure 3 shows the signal region plots corresponding to the ones shown in Fig. 1.

Since the event yield of the signal region at this cut stage is intrinsically poor, their significance is limited.

5 MC@NLO $t\bar{t}$ sample

This section contains auxiliary plots for the discussion of the merging procedure for the MC@NLO $t\bar{t}$ sample presented in Section 7.5.

5.1 Top control region

Figures 4 and 5 show the same distributions as in Fig. 7.3 and 7.4, but after application of the dijet kinematic cuts.
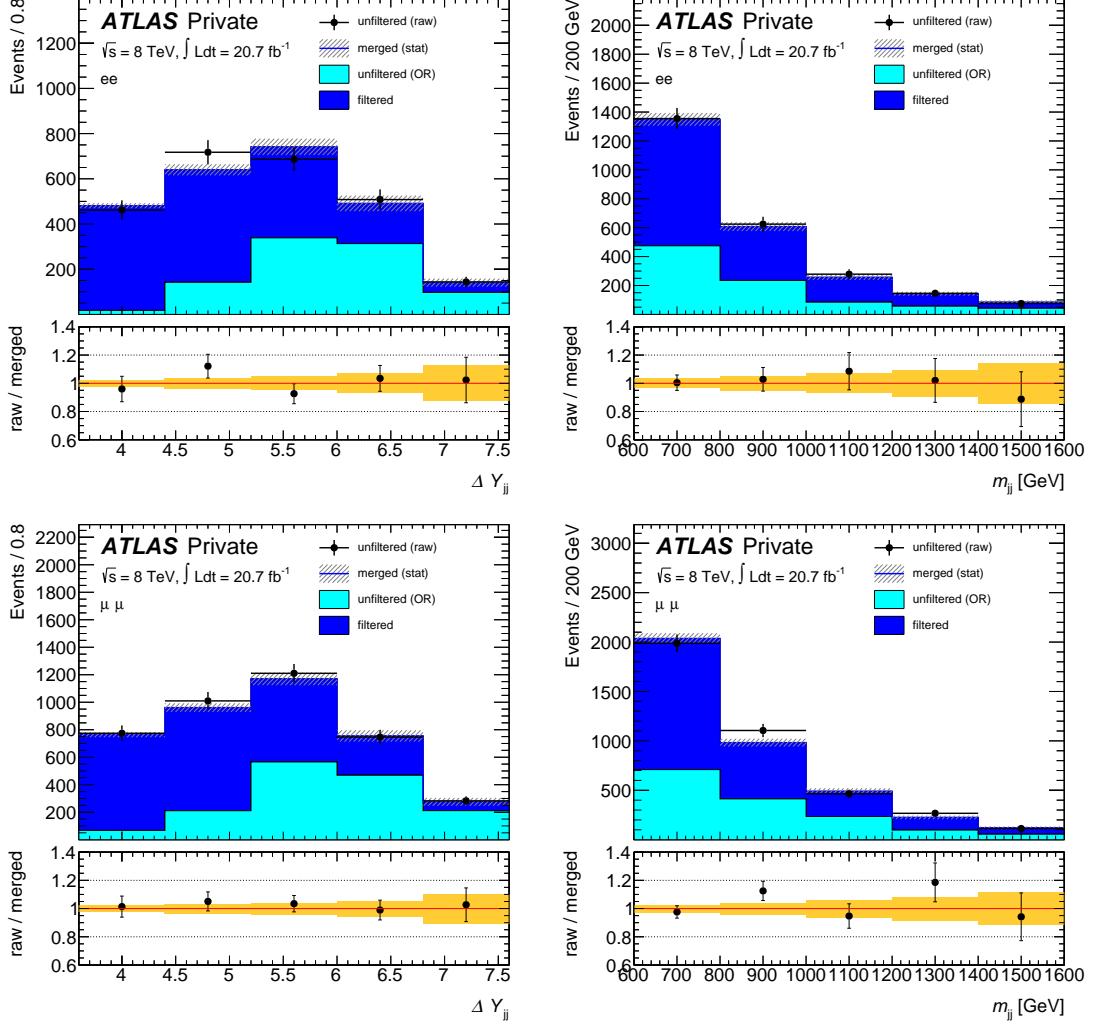
Figure 1: Distributions for the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $Z \to \ell\ell$ samples for $\ell\ell = ee$ (top) and $\ell\ell = \mu\mu$ (bottom). The distributions are shown for the Z control region after application of the dijet kinematic requirements $\Delta y_{jj} > 3.6$ and $m_{jj} > 600\,\text{GeV}$. The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered one. Red arrows indicate the direction of outlying data points in the ratio plot.
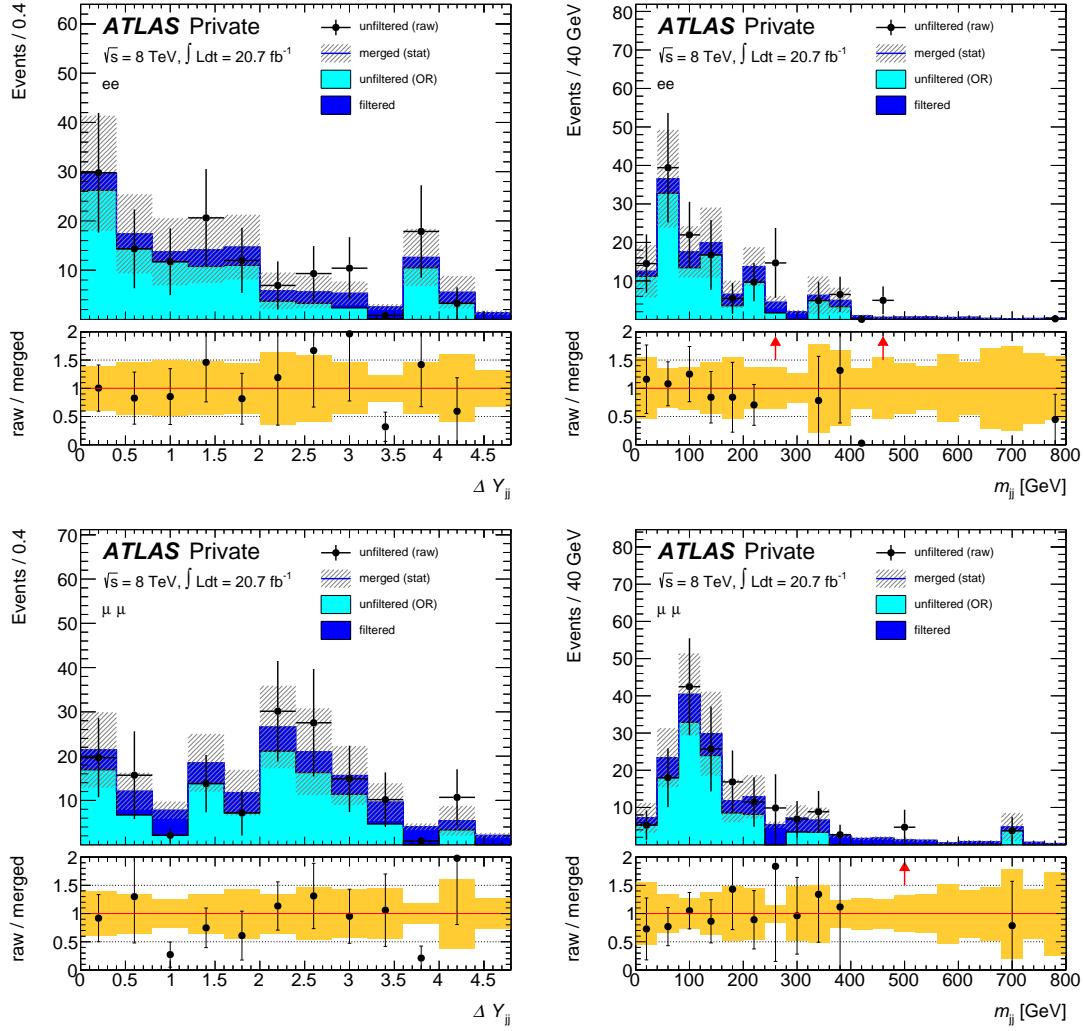
Figure 2: Distributions for the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $Z \to \ell\ell$ samples for $\ell\ell = ee$ (top) and $\ell\ell = \mu\mu$ (bottom). The distributions are shown for the event selection of the signal region after the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered one. Red arrows indicate the direction of outlying data points in the ratio plot.
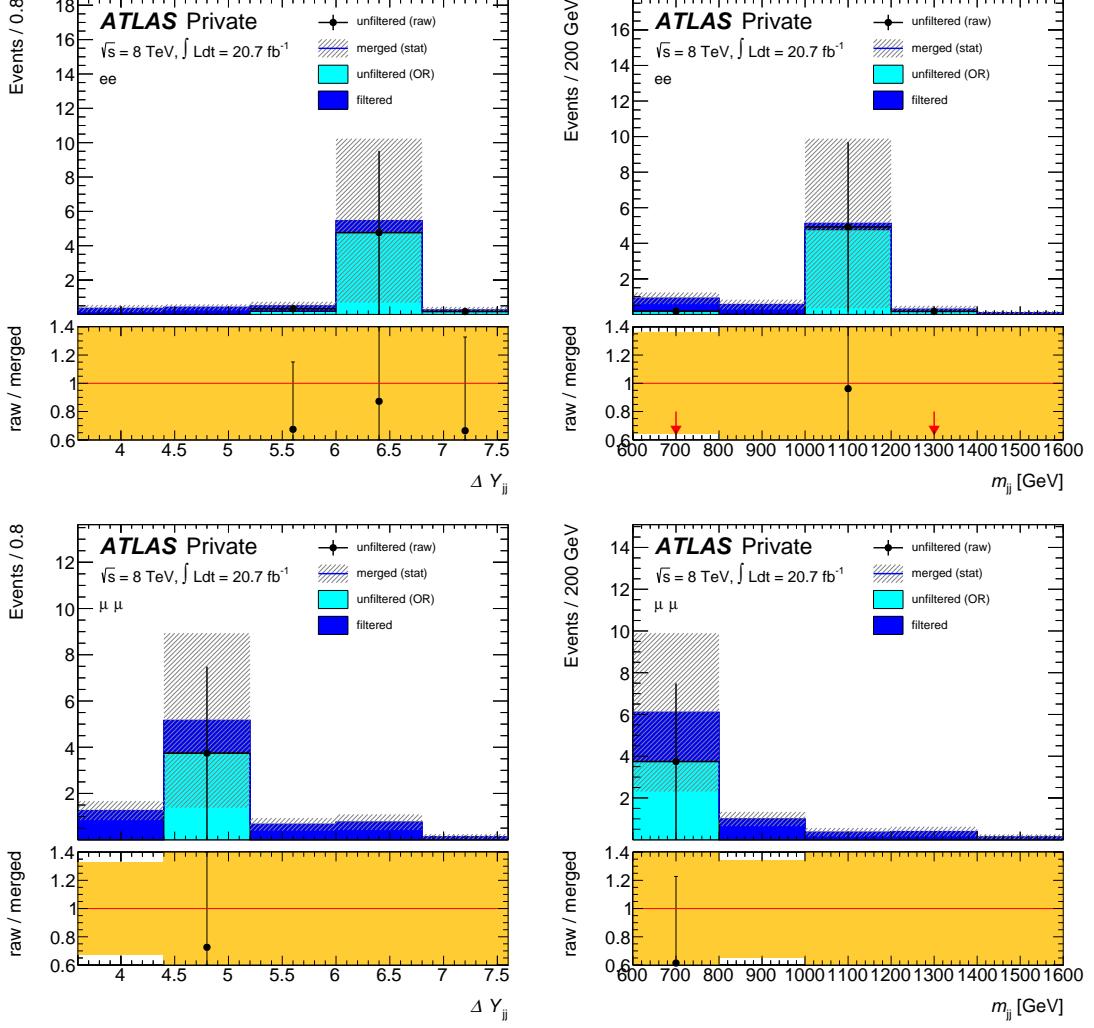
Figure 3: Distributions for the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $Z \to \ell\ell$ samples for $\ell\ell = ee$ (top) and $\ell\ell = \mu\mu$ (bottom). The distributions are shown for the event selection of the signal region after cuts on the dijet kinematic variables (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered one. Red arrows indicate the direction of outlying data points in the ratio plot.
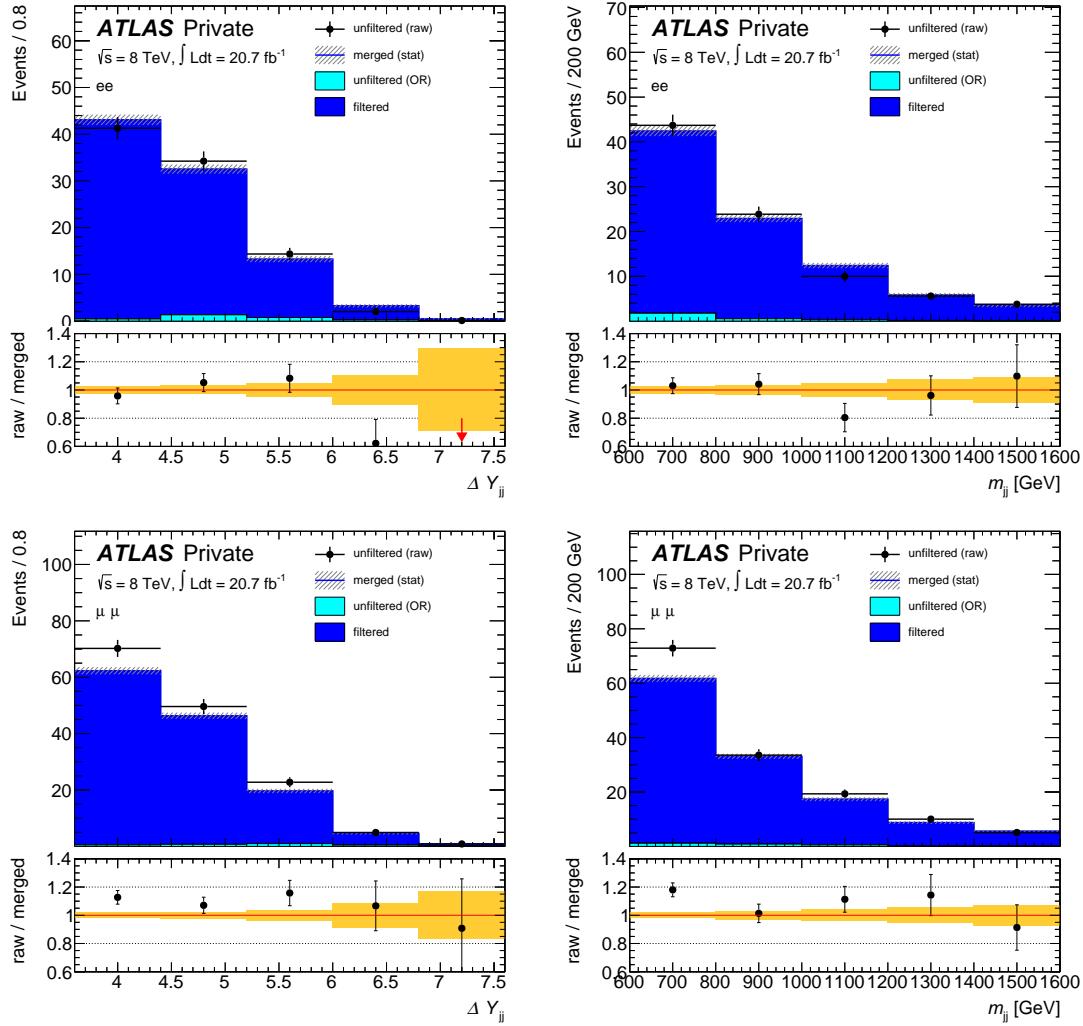
Figure 4: Distributions for the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $t\bar{t}$ samples for the same flavour channel (*ee* at the top, $\mu\mu$ at the bottom). The distributions are shown for the event selection of the signal region after the cuts both dijet kinematic variables (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered one. The respective ratio plots show good agreement between the unfiltered sample before application of the overlap removal and the merged sum of the filtered and the overlap-removed unfiltered sample within the statistical uncertainties. Red arrows indicate the direction of outlying data points in the ratio plot.

Figure 5: Distributions for the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $t\bar{t}$ samples for the same flavour channel (*ee* at the top, $\mu\mu$ at the bottom). The distributions are shown for the event selection of the signal region after the cuts both dijet kinematic variables (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered one. The respective ratio plots show good agreement between the unfiltered sample before application of the overlap removal and the merged sum of the filtered and the overlap-removed unfiltered sample within the statistical uncertainties. Red arrows indicate the direction of outlying data points in the ratio plot.

5.2 Signal region

Figures 6 and 7 show the signal region plots corresponding to the ones shown in Fig. 7.3 and Fig. 7.4. Figures 8 and 9 show the signal region plots corresponding to the ones shown in Fig. 4 and Fig. 5.

Since the event yield of the signal region at this cut stage is intrinsically poor, the statistical significance of these is limited.
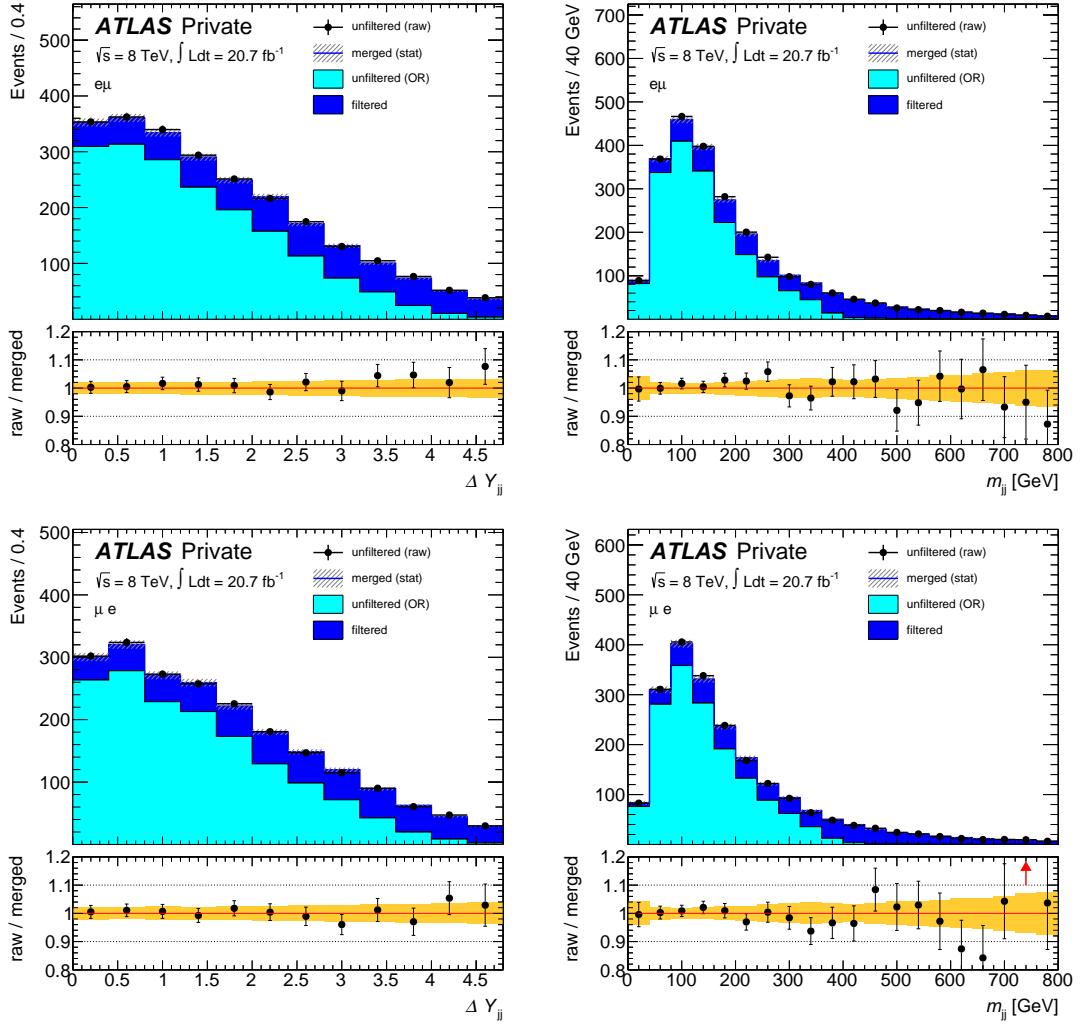
Figure 6: Distributions for the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $t\bar{t}$ samples for the different flavour channel ($e\mu$ at the top, μe at the bottom). The distributions are shown at the entry point of the signal region (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered one. The respective ratio plots show good agreement between the unfiltered sample before application of the overlap removal and the merged sum of the filtered and the overlap-removed unfiltered sample within the statistical uncertainties. Red arrows indicate the direction of outlying data points in the ratio plot.

Figure 7: Distributions for the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $t\bar{t}$ samples for the same flavour channel (*ee* at the top, $\mu\mu$ at the bottom). The distributions are shown at the entry point of the signal region (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered one. The respective ratio plots show good agreement between the unfiltered sample before application of the overlap removal and the merged sum of the filtered and the overlap-removed unfiltered sample within the statistical uncertainties. Red arrows indicate the direction of outlying data points in the ratio plot.

Figure 8: Distributions for the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $t\bar{t}$ samples for the different flavour channel ($e\mu$ at the top, μe at the bottom). The distributions are shown for the event selection of the signal region after the cuts both dijet kinematic variables (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered one. The respective ratio plots show good agreement between the unfiltered sample before application of the overlap removal and the merged sum of the filtered and the overlap-removed unfiltered sample within the statistical uncertainties. Red arrows indicate the direction of outlying data points in the ratio plot.

Figure 9: Distributions for the dijet invariant mass m_{jj} and the dijet rapidity gap Δy_{jj} of the leading jet pair in the filtered and unfiltered (raw and overlap-removed) jet-inclusive $t\bar{t}$ samples for the same flavour channel (*ee* at the top, $\mu\mu$ at the bottom). The distributions are shown for the event selection of the signal region after the cuts both dijet kinematic variables (see Chapter 8). The filtered and overlap-removed unfiltered contributions are stacked and shown against the raw unfiltered one. The respective ratio plots show good agreement between the unfiltered sample before application of the overlap removal and the merged sum of the filtered and the overlap-removed unfiltered sample within the statistical uncertainties. Red arrows indicate the direction of outlying data points in the ratio plot.

Auxiliary Material

6 The ATLAS coordinate system

The coordinate system of ATLAS is right-handed. The x-axis points in radial direction towards the centre of the circle which the LHC tunnel describes. The y-axis is slightly tilted with respect to the vertical due to the general tilt of the tunnel, yielding a deviation from the vertical of approximately 1.5°. The z-axis points along a tangent to the LHC tunnel tube centre. This direction is commonly referred to as *longitudinal* or *forward/backward*, as opposed to the transverse x/y-plain. The z-axis is also referred to as *beam axis*.

6.1 Basic coordinate system

Common coordinates are r_T, η and ϕ. Here, r_T is the distance from the beam axis itself, and ϕ is the angle of the track in the transverse plain. The coordinate η is referred to as *pseudorapidity*, defined as

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

where θ denotes the polar angle with respect to the beam axis. This choice of η as a suitable coordinate is well motivated by the following considerations.

The LHC is a hadron collider. Considering the case of two partons colliding at the central interaction vertex, each of them carries a fraction x of the total momentum of the proton (which is sometimes denoted as x_{Bj} and referred to as *Bjorken-* or *Feynman-x* [8]). The precise value of x is different for both collision partners and randomly distributed according to the parton distribution functions (see Section 4.1.3). Hence, the relative longitudinal velocity of the centre-of-mass frame of the collision is unknown and cannot be measured easily, since a non-negligible fraction of objects emerging the collision will evade detection due to their close alignment with the beam pipe. Thus, it is desirable to choose a coordinate system that is invariant under longitudinal Lorentz-transformations.
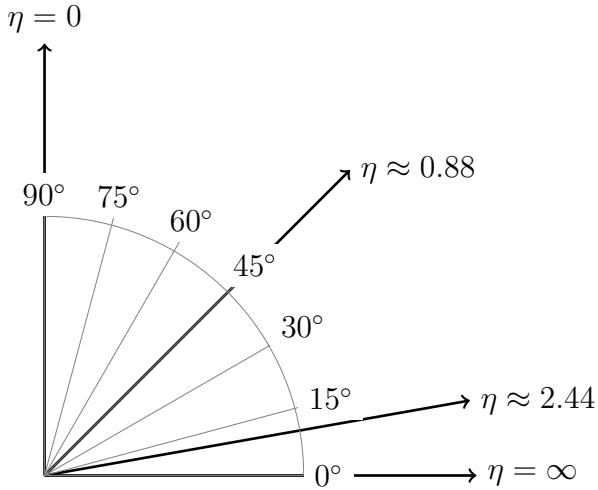
Figure 10: Pseudorapidity η vs. polar angle θ

Although η is not Lorentz-invariant, the use of η (instead of the angle θ) has some advantages. The *pseudorapidity* η is very closely related to the *rapidity y*.

$$
\begin{aligned}
y &= \operatorname{arc\,tanh}\left(\frac{p_z}{E}\right) \\
&\overset{\text{def}}{=} \frac{1}{2}\ln\left(\frac{1+\frac{p_z}{E}}{1-\frac{p_z}{E}}\right) = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) \\
&\overset{E\gg mc^2}{\approx} \frac{1}{2}\ln\left(\frac{|\vec{p}|+p_L}{|\vec{p}|-p_L}\right) = -\ln\sqrt{\frac{|\vec{p}|-p_L}{|\vec{p}|+p_L}} = -\ln\sqrt{\frac{1-\cos\theta}{1+\cos\theta}} \\
&\overset{\text{def}}{=} -\ln\tan\frac{\theta}{2} \\
&= \eta
\end{aligned}
$$

While the rapidity y is not Lorentz invariant either, rapidity *differences* Δy are Lorentz invariant. Hence, pseudorapidity differences $\Delta\eta$ inherit this property in an approximate manner. The use of pseudorapidity as a coordinate is especially convenient since it can be directly calculated from θ, the angle with respect to the beam pipe.

The relation between θ and η is illustrated by Fig. 10. The value $\eta = 0$ here corresponds to the radial direction, i.e. $\theta = \pi/2$. Points along the beam axis, on the other hand, correspond to infinite values of pseudorapidity. While this might seem discomforting, pseudorapidity values one has to deal with in practice rarely exceed $\eta \approx 5$ due to the limited forward and backward coverage of the detector. This value corresponds to $\theta \approx 0.8°$.

6.2 Radial distances

Since object coordinates in an event are expressed in values of η and ϕ, all geometric quantities are consistently defined based on these coordinates. One of the most commonly used derived geometric quantities is the radial *distance* ΔR between two objects like particle tracks or calorimeter clusters, defined as

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

For straight tracks or vectors originating from the interaction point, ΔR is independent of the distance r_T from the beam axis. However, taking into account direction changes of tracks that might result from interaction with the detector material or the magnetic field of the inner detector, it is common for values of ΔR between the same two tracks to deviate between different detector layers.

6.3 Impact parameters

In order to locate vertices in the detector, it is convenient to define the impact parameters d_0 and z_0 of tracks. Considering some track and its point of closest approach to the beam line, the transverse impact parameter d_0 is the radial coordinate r_T of this point, whereas the longitudinal impact parameter z_0 is the z-coordinate of this point.

The z-resolution of the inner detector is approximately proportional to $\sin\theta$. Hence, it is sometimes convenient to use $z_0 \sin\theta$ instead of z_0, since the former achieves a much more evenly distributed resolution as a function of η. For example, posing tight z_0 requirements on tracks can result in approximately random decisions for very forward tracks, which can be avoided by using $z_0 \sin\theta$ instead.

List of Figures

List of Figures

List of Tables

Bibliography

[1] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266. `http://link.aps.org/doi/10.1103/PhysRevLett.19.1264`.

[2] I. Kenyon, "The discovery of the intermediate vector bosons", *European Journal of Physics* **6** no. 1, (1985) 41. `http://stacks.iop.org/0143-0807/6/i=1/a=006`.

[3] W. Higgs, P. "Broken symmetries, massless particles and gauge fields", *Physics Letters* **12** no. 2, (1964) 132–133. `http://www.sciencedirect.com/science/article/pii/0031916364911369`.

[4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Broken Symmetries and the Goldstone Theorem", *Advances in Physics* **2** 567–708. `http://www.datafilehost.com/d/7d512618`. ISBN 978-0470170571.

[5] **ATLAS Collaboration**, G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B* no. 716, (August 31, 2012) 1–29, `arXiv:1207.7214 [hep-ex]`. `http://arxiv.org/abs/1207.7214`.

[6] **CMS Collaboration**, "Observation of a new boson with a mass near 125 GeV", CMS-PAS-HIG-12-020, CERN, Geneva, 2012. `CDS:1460438`.

[7] D. Overbye, *Physicists Find Elusive Particle Seen as Key to Universe*, New York Times, July 5, 2012. `http://www.nytimes.com/2012/07/05/science/cern-physicists-may-have-discovered-higgs-boson-particle.html`

[8] F. Halzen and A. D. Martin, *Quarks and Leptons – An Introductory Course in Modern Particle Physics*. John Wiley & Sons, 1984. ISBN 978-0-471-88741-2.

[9] M. Gonzalez-Garcia and Y. Nir, "Neutrino Masses and Mixing: Evidence and Implications", *Rev. Mod. Phys.* **75** (2003) 345–402, `arXiv:hep-ph/0202058 [hep-ph]`.

[10] **LEP Working Group for Higgs boson searches, ALEPH, DELPHI, L3 and OPAL Collaborations**, R. Barate *et al.*, "Search for the Standard Model Higgs boson at LEP", *Phys. Lett.* **B565** no. CERN-EP-2003-011,

(2003) 61–75, `arXiv:hep-ex/0306033 [hep-ex]`.

[11] LEP Electroweak Working Group, *Combined Results*, Website, 2012. `http://lepewwg.web.cern.ch/LEPEWWG`

[12] **CDF and D0 Collaborations**, R. Roser and D. Denisov, "Higgs Boson Studies at the Tevatron", *Phys. Rev. D.* **N. N.** no. FERMILAB-PUB-13-081-E, (2013) , `arXiv:1303.6346 [hep-ex]`.

[13] A. Djouadi, "The Anatomy of Electro-Weak Symmetry Breaking. I: The Higgs boson in the Standard Model", *Phys. Rept* **457** (2008) 216, `arXiv:hep-ph/0503172 [hep-ph]`.

[14] L. Evans and P. Bryant, "The LHC Machine", *Journal of Instrumentation* **The CERN Large Hadron Collider: Accelerator and Experiments** (August, 2008) 164. `http://jinst.sissa.it/LHC/LHCmachine/2008_JINST_3_S08001.pdf`.

[15] LHCb Collaboration, *the LHCb Detector*, Website, 2008. `http://lhcb-public.web.cern.ch/lhcb-public/en/detector/Detector-en.html`

[16] C. Carli, ed., *Chamonix 2012*, no. CERN-2012-006 in Workshop on LHC Performance. CERN, Geneva, Apr, 2012. `CDS:1424362`.

[17] L. Randall and R. Sundrum, "A Large Mass Hierarchy from a Small Extra Dimension", *Phys. Rev. Lett.* **83** (1999) 3370–3373, `arXiv:hep-ph/9905221 [hep-ph]`.

[18] ATLAS Collaboration, *ATLAS*, Fact Sheet, 2013. `http://www.atlas.ch/fact-sheets-1-view.html`

[19] ATLAS Collaboration, "The ATLAS Experiment at the CERN LHC", *Journal of Instrumentation* **The CERN Large Hadron Collider: Accelerator and Experiments** (August, 2008) 437. `http://jinst.sissa.it/LHC/ATLAS/2008_JINST_3_S08003.pdf`.

[20] **ATLAS Collaboration**, G. Aad *et al.*, "Expected Performance of the ATLAS Experiment – Detector, Trigger and Physics", Tech. Rep., CERN, 2009. `arXiv:0901.0512 [hep-ex]`.

[21] **ATLAS Collaboration**, A. Salvucci, "Measurement of muon momentum resolution of the ATLAS detector", `arXiv:1201.4704 [physics.ins-det]`.

[22] **ATLAS Collaboration**, W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, "Calorimeter Clustering Algorithms: Description and Performance", ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008. `CDS:1099735`.

[23] **ATLAS Collaboration**, G. Aad *et al.*, "Performance of the Electron and Photon Trigger in pp Collisions at $\sqrt{s} = 7\,\mathrm{TeV}$", ATLAS-CONF-2011-114,

CERN, 2011. `CDS:1375551`.

[24] **ATLAS Collaboration**, G. Aad *et al.*, "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data", *Eur. Phys. J.* **C72** (2012) 46, `arXiv:1110.3174 [hep-ex]`.

[25] **ATLAS Collaboration**, G. Aad *et al.*, "Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector", ATLAS-CONF-2011-063, CERN, 2011. `CDS:1345743`.

[26] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **2008** no. 04, (2008) 063. `http://stacks.iop.org/1126-6708/2008/i=04/a=063`.

[27] **ATLAS Collaboration**, G. Aad *et al.*, "Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data", ATLAS-CONF-2012-020, CERN, 2012. `CDS:1430034`.

[28] **ATLAS Collaboration**, D. W. Miller, A. Schwartzman, and D. Su, "Jet-Vertex Association Algorithm", ATLAS-CONF-2008-008, CERN, 2008. `CDS:1082880`.

[29] **ATLAS Collaboration**, P. Giovannini, "Local hadron calibration with ATLAS", *Journal of Physics* **293** no. 012057, (2011) .

[30] **ATLAS Collaboration**, G. Aad *et al.*, "Jet energy measurement with the ATLAS detector in proton-proton collisions at $sqrts = 7$ TeV", `arXiv:1112.6426 [hep-ex]`.

[31] **ATLAS Collaboration**, G. Aad *et al.*, "Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data", ATLAS-CONF-2011-102, CERN, 2011. `CDS:1369219`.

[32] **ATLAS Collaboration**, "Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector", ATLAS-CONF-2012-043, CERN, Geneva, Mar, 2012. `CDS:1435197`.

[33] 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 [hep-ph]`.

[34] A. Dobbs, M. 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]`.

[35] T. Sjostrand, S. Mrenna, and P. Z. Skands, PYTHIA *6.4 Physics and Manual*, 2006. `arXiv:hep-ph/0603175 [hep-ph]`.

[36] T. Sjöstrand, S. Mrenna, and P. SkandsBinoth, *A Brief Introduction to* PYTHIA *8.1*, 2008. `arXiv:0710.3820 [hep-ph]`.

[37] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, *et al.*, HERWIG *6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, 2001. `arXiv:hep-ph/0011363 [hep-ph]`.

[38] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, *et al.*, *Event generation with* SHERPA *1.1*, 2009. `arXiv:0811.4622 [hep-ph]`.

[39] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in* POWHEG, 2010. `arXiv:0911.5299 [hep-ph]`.

[40] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations", *JHEP* **0206** (2002) 029, `arXiv:hep-ph/0204244 [hep-ph]`.

[41] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5: Going Beyond*, 2011. `arXiv:1106.0522 [hep-ph]`.

[42] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, ALPGEN*, a generator for hard multiparton processes in hadronic collisions*, 2003. `arXiv:hep-ph/0206293 [hep-ph]`.

[43] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer, "Gluon-induced W-boson pair production at the LHC", *JHEP* **0612** (2006) 046, `arXiv:hep-ph/0611170 [hep-ph]`.

[44] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Hoche, *et al.*, *General-purpose event generators for LHC physics*, 2011. `arXiv:1101.2599 [hep-ph]`.

[45] Z. Was, "TAUOLA– the library for τ lepton decay", *JHEP* (2000) , `arXiv:hep-ph/0011305 [hep-ph]`.

[46] E. Barberio and Z. Was, "PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0", *Comput.Phys.Commun.* **79** (1994) 201–308, `arXiv:hep-ph/0206293 [hep-ph]`.

[47] S. Agostinelli *et al.*, "GEANT4 – A Simulation Toolkit", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** no. 3, (2003) 250 – 303.

[48] G. Aad *et al.*, "The ATLAS Simulation Infrastructure", *Eur. Phys. J.* **C72** (2010) 823–874, `arXiv:1005.4568 [physics.ins-det]`.

[49] LHC Higgs Cross Section Working Group, *CrossSections*, Website, 2013. `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections`

[50] **ATLAS Collaboration**, G. Aad *et al.*, "Combined measurements of the mass and signal strength of the Higgs-like boson with the ATLAS detector using up to $25\,\text{fb}^{-1}$ of proton-proton collision data", ATLAS-CONF-2013-014, CERN, 2013. `CDS:1523727`.

[51] **CMS Collaboration**, M. Chen, "Combination and Standard Model Scalar Boson Properties in CMS", `arXiv:1305.4775 [hep-ex]`.

[52] **ATLAS Collaboration**, G. Aad *et al.*, "Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC", CERN-PH-EP-2013-103, CERN, 2013. `arXiv:1307.1427 [hep-ex]`.

[53] **ATLAS Collaboration**, G. Aad *et al.*, "Measurements of the properties of the Higgs-like boson in the $WW^{(*)} \to \ell\nu\ell\nu$ decay channel with the ATLAS detector using 25fb^{-1} of proton-proton collision data", ATLAS-CONF-2013-030, CERN, 2013. `CDS:1527126`.

[54] **CMS Collaboration**, "Evidence for a particle decaying to W^+W^- in the fully leptonic final state in a standard model Higgs boson search in pp collisions at the LHC", CMS-PAS-HIG-13-003, 2013. `CDS:1523673`.

[55] **CMS Collaboration**, G. Aad *et al.*, "Update of the search for the Standard Model Higgs boson decaying into WW in the vector boson fusion production channel", CMS-PAS-HIG-13-022, CERN, 2013. `CDS:1590404`.

[56] **CMS Collaboration**, *CMS Higgs Physics Results*, TWiki, 2013. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsHIG`

[57] A. L. Read, "Modified frequentist analysis of search results (the CL$_s$ method)", CERN-OPEN-2000-205, CERN, Jan. 17-18, 2000. `CDS:451614`.

[58] **ATLAS Collaboration**, "Evidence for the spin-0 nature of the Higgs boson using ATLAS data", `arXiv:1307.1432 [hep-ex]`.

[59] **LHC Higgs Cross Section Working Group**, S. Dittmaier *et al.*, "Handbook of LHC Higgs Cross Sections: 2. Differential Distributions", `arXiv:1201.3084 [hep-ph]`.

[60] **Particle Data Group**, Beringer, J. and others, "The Review of Particle Physics", *Phys. Rev.* **D86** (2013) 010001. `http://pdg.lbl.gov/`.

[61] **ATLAS Collaboration**, "Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at $\sqrt{s} = 7\,\text{TeV}$", Tech. Rep., CERN, Geneva, Jul, 2012. `CDS:1463915`.

[62] ATLAS Public Results, *ATLAS Physics Summary Plots*, Website, 2013. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/CombinedSummaryPlots`

[63] K. Ellis, R. *et al.*, "Higgs decay to $\tau^+\tau^-$: A possible signature of intermediate mass Higgs bosons at the SSC", *Nucl. Phys. B* no. 297, (1988) 221.

[64] ATLAS Public Results, *Online Luminosity Plots*, Website, 2013. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`

[65] J. Adelman *et al.*, "ATLAS offline data quality monitoring", *Journal of Physics: Conference Series* **219** no. 4, (2010) 042018.

http://stacks.iop.org/1742-6596/219/i=4/a=042018.

[66] ATLAS Collboration, *ATLAS Data Quality Information*, TWiki, 2012.
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPubli
cResults2010#2012

[67] N. Kauer and G. Passarino, "Inadequacy of zero-width approximation for a
light Higgs boson signal", arXiv:1206.4803 [hep-ph].

[68] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator
AcerMC version 2.0 with interfaces to* PYTHIA *6.2 and* HERWIG *6.5*, 2004.
arXiv:hep-ph/0405247 [hep-ph].

[69] J. Alwall, S. Hoche, F. Krauss, N. Lavesson, L. Lonnblad, *et al.*,
"Comparative study of various algorithms for the merging of parton showers
and matrix elements in hadronic collisions", *Eur. Phys. J.* **C53** (2008)
473–500, arXiv:0706.2569 [hep-ph].

[70] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, *et al.*,
*MadGraph/*MADEVENT *v4: The New Web Generation*, 2007.
arXiv:0706.2334 [hep-ph].

[71] R. C. Gray, C. Kilic, M. Park, S. Somalwar, and S. Thomas, "Backgrounds
To Higgs Boson Searches from $W\gamma^{(*)} \to \ell\nu\ell(\ell)$ Asymmetric Internal
Conversion", arXiv:1110.1368 [hep-ph].

[72] A. Djouadi, J. Kalinowski, and M. Spira, "HDECAY: A Program for Higgs
boson decays in the standard model and its supersymmetric extension",
Comput. Phys. Commun. **108** (1998) 56–74, arXiv:hep-ph/9704448
[hep-ph].

[73] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, "Precise
predictions for the Higgs-boson decay $H \to WW/ZZ \to 4$ leptons", *Phys.
Rev. D.* **013004** (2006) 74, arXiv:hep-ph/0604011 [hep-ph].

[74] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, "Radiative
corrections to the semileptonic and hadronic Higgs-boson decays
$H \to WW/ZZ \to 4$ fermions", *JHEP* **0702** (2007) 080,
arXiv:hep-ph/0611234 [hep-ph].

[75] A. Djouadi, M. Spira, and P. Zerwas, "Production of Higgs bosons in proton
colliders: QCD corrections", *Phys. Lett. B* **264** no. 3–4, (1991) 440 – 446.

[76] S. Dawson, "Radiative corrections to Higgs boson production", *Nucl. Phys. B*
359 no. 2–3, (1991) 283 – 300.

[77] M. Spira, A. Djouadi, D. Graudenz, and P. Zerwas, "Higgs boson production
at the LHC", *Nucl. Phys.* **B453** (1995) 17–82, arXiv:hep-ph/9504378
[hep-ph].

[78] R. V. Harlander and W. B. Kilgore, "Next-to-next-to-leading order Higgs production at hadron colliders", *Phys. Rev. Lett.* **88** (2002) 201801, `arXiv:hep-ph/0201206 [hep-ph]`.

[79] C. Anastasiou and K. Melnikov, "Higgs boson production at hadron colliders in NNLO QCD", *Nucl. Phys.* **B646** (2002) 220–256, `arXiv:hep-ph/0207004 [hep-ph]`.

[80] V. Ravindran, J. Smith, and W. L. van Neerven, "NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions", *Nucl. Phys.* **B665** (2003) 325–366, `arXiv:hep-ph/0302135 [hep-ph]`.

[81] S. Catani, D. de Florian, M. Grazzini, and P. Nason, "Soft gluon resummation for Higgs boson production at hadron colliders", *JHEP* **0307** (2003) 028, `arXiv:hep-ph/0306211 [hep-ph]`.

[82] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini, "Two loop light fermion contribution to Higgs production and decays", *Phys. Lett.* **B595** (2004) 432–441, `arXiv:hep-ph/0404071 [hep-ph]`.

[83] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, "NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders", *Phys. Lett.* **B670** (2008) 12–17, `arXiv:0809.1301 [hep-ph]`.

[84] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, "Higgs production via vector-boson fusion at NNLO in QCD", *Phys. Rev. Lett.* **105** (2010) 011801, `arXiv:1003.4451 [hep-ph]`.

[85] M. Ciccolini, A. Denner, and S. Dittmaier, "Strong and electroweak corrections to the production of Higgs + 2 jets via weak interactions at the LHC", *Phys. Rev. Lett.* **99** (2007) 161803, `arXiv:0707.0381 [hep-ph]`.

[86] M. Ciccolini, A. Denner, and S. Dittmaier, "Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC", *Phys. Rev.* **D77** (2008) 013002, `arXiv:0710.4749 [hep-ph]`.

[87] K. Arnold *et al.*, "VBFNLO: A parton level Monte Carlo for processes with electroweak bosons", *Comput. Phys. Commun.* **180** no. 9, (2009) 1661 – 1670.

[88] T. Han and S. Willenbrock, "QCD correction to the $pp \rightarrow WH$ and ZH total cross sections", *Phys. Lett. B* **273** no. 1-2, (1991) 167 – 172.

[89] O. Brein, A. Djouadi, and R. Harlander, "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders", *Phys. Lett.* **B579** (2004) 149–156, `arXiv:hep-ph/0307206 [hep-ph]`.

[90] M. Ciccolini, S. Dittmaier, and M. Kramer, "Electroweak radiative corrections to associated WH and ZH production at hadron colliders", *Phys. Rev.* **D68** (2003) 073003, `arXiv:hep-ph/0306234 [hep-ph]`.

[91] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 10–15, `arXiv:1007.3492 [hep-ph]`.

[92] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, *et al.*, "New parton distributions for collider physics", *Phys. Rev.* **D82** (2010) 074024, `arXiv:1007.2241 [hep-ph]`.

[93] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, *et al.*, "Implications of CTEQ global analysis for collider observables", *Phys. Rev.* **D78** (2008) 013004, `arXiv:0802.0007 [hep-ph]`.

[94] **ATLAS Collaboration**, J. Apostolakis, A. Buckley, A. Dotti, and Z. Marshall, "Final Report of the ATLAS Detector Simulation Performance Assessment Group", CERN-LCGAPP-2010-01, 2010. `sftweb.cern.ch/AAdocuments`.

[95] C. Burgard and E. M. Lobodzinska, *Bug in* `VBFForwardJetsFilter::sumDaughterNeutrinos`, Savannah Bug Report, March 7, 2013. `https://savannah.cern.ch/bugs/?100777`

[96] **ATLAS Collaboration**, M. Biglietti, S. Armstrong, K. A. Assamagan, J. T. M. Baines, C. P. Bee, M. Bellomo, and J. A. C. Bogaerts, "Muon Event Filter Software for the ATLAS Experiment at LHC", ATL-DAQ-CONF-2005-008, 2004. `CDS:820783`.

[97] **ATLAS Collaboration**, B. Resende, "Muon identification algorithms in ATLAS", ATL-PHYS-PROC-2009-11, 2009. `CDS:1209632`.

[98] **ATLAS Collaboration**, "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung", ATLAS-CONF-2011-063, 2012. `CDS:1449796`.

[99] **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 $\sqrt{s} = 7\,\text{TeV}$", ATLAS-CONF-2010-038, CERN, Geneva, Jul, 2010. `CDS:1277678`.

[100] Jet/E_T^miss working group, *e/γ Physics Calibration*, TWiki, 2013. `https://twiki.cern.ch/twiki/bin/view/Sandbox/EgammaPhysCalib`

[101] **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 $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS detector", *Phys. Rev. Lett.* **108** (2012) 111802, `arXiv:1112.2577 [hep-ex]`.

[102] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", `arXiv:1007.1727 [physics.data-an]`.

[103] F. James and M. Winkler, "MINUIT Users Guide", http://seal.web.cern.ch/seal/documents/minuit/mnusersguide.pdf.

[104] A. Wald, "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large", *Transactions of the American Mathematical Society* **45** no. 3, (1943) 426–482.

[105] **LHC Higgs Cross Section Working Group**, S. Heinemeyer *et al.*, "Handbook of LHC Higgs Cross Sections: 3. Higgs properties", arXiv:1307.1347 [hep-ph].

[106] Rene Brun et. al., *ROOT – An Object Oriented Data Analysis Framework*. http://root.cern.ch/

[107] A. Walz, C. Burgard, *et al.*, *HWWAnalysisCode – HSG3 Common Analysis Framework*, Documentation Webpage. http://wwwhep.physik.uni-freiburg.de/~cburgard/CAF-doc

[108] A. Walz, *Search for $H \to W^\pm W^{\mp(*)} \to \ell^+ \nu \ell^- \bar{\nu}$ Decays in the Gluon Fusion and Vector-Boson Fusion Production Modes at the LHC*. Diploma thesis, Albert-Ludwigs-Universität Freiburg, 2012. http://portal.uni-freiburg.de/jakobs/dateien/abschluss/diplo_walz/view.

[109] C. Burgard and W. Verkerke, *Memory Leak in RooFit/HistFactory*, JIRA Bug Report, June 4, 2013. https://sft.its.cern.ch/jira/browse/ROOT-5236

[110] M. Dührssen, *Study of Higgs bosons in the WW final state and development of a fast calorimeter simulation for the ATLAS experiment*. PhD thesis, Albert-Ludwigs-Universität Freiburg, 2010. CDS:1261372.

[111] E. Schmidt, *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*. PhD thesis, Albert-Ludwigs-Universität Freiburg, 2013. CDS:1562308.