Measurement of the boosted WW+WZ production cross section in the semileptonic decay channel with ATLAS

DISSERTATION

vorgelegt von

Fabian Spettel

Erstellt am

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

in Kollaboration mit der

Universität Tel Aviv

TECHNISCHE UNIVERSITÄT MÜNCHEN

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Measurement of the boosted WW+WZ production cross section in the semileptonic decay channel with ATLAS

Fabian M. Spettel

Vollständiger Abdruck der von der Fakultät für Physik
der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Dr. A. Ibarra
Prüfer der Dissertation:	
1.	Hon.-Prof. Dr. S. Bethke
2.	Univ.-Prof. Dr. St. Paul

Die Dissertation wurde am 11. 01 2017 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 01. 02 2017 angenommen.

Für Mama.

Kurzfassung

In dieser Arbeit wird die erste Messung des kombinierten WW+WZ Wirkungsquerschnitts mit starkem Lorentz-Boost vorgestellt, wobei ein W-Boson durch den Elektron- oder Muon-Kanal und das verbleibende W- oder Z-Boson hadronisch zerfällt. Die Ereignisselektion zielt auf Ereignisse ab, bei welchen die hadronischen Zerfallsprodukte, resultierend aus dem hohen Transversalimpuls des zerfallenden Bosons, stark kollimiert sind. Der verwendete Datensatz besteht aus den gesamten Proton-Proton-Kollisionen die im Jahre 2012 bei einer Schwerpunktsenergie von $\sqrt{s}$=8 TeV mit dem ATLAS Detektor am Large Hadron Collider aufgezeichnet wurden. Die integrierte Luminosität dieses Datensatzes beträgt 20.3 fb^{-1}. Die wichtigsten Untergrundprozesse zum WW+WZ Signal stellen W+Jets, Z+Jets, Top-Quarks und QCD-Multijet-Produktion dar. Letzerer wird durch eine spezielle, datenbasierte Methode abgeschätzt, während die anderen Untergrundprozesse durch Monte-Carlo-Simulation modelliert werden. Insgesamt wurden in den Daten 383 Signalereignisse beobachtet, bei einem geschätzten Untergrund von 12452 Ereignissen. Dies entspricht einem gemessenen Wirkungsquerschnitt in dem hier betrachteten, reduzierten Phasenraum von

$$\sigma_{\text{fid}}^{\text{meas}}(\text{WW} + \text{WZ}) = (31.7 \pm 10.4\,(\text{stat.}) \pm 22.1\,(\text{syst.}))\,\text{fb}.$$

Dieses Ergebnis ist innerhalb der Fehlertoleranzen mit der Standardmodellvorhersage in zweiter Ordnung von $\sigma_{\text{WW+WZ}}^{\text{SM}} = (57.9 \pm 2.9)\,\text{fb}$ vereinbar. Das beobachtete Transversalimpulsspektrum des hadronisch zerfallenden Bosons wird benutzt um via anomaler Dreifacheichbosonkopplungen an WWZ- und WWγ-Vertizes nach neuer Physik zu suchen. Die Beobachtungen stimmen mit der Vorhersage des Standardmodells überein und es werden Vertrauensintervalle für die Parameter, welche die anomalen Kopplungen beschreiben, abgeleitet.

Abstract

This thesis presents the first measurement of the combined WW+WZ production cross section in the boosted regime with one W boson decaying via the electron or muon channel and the remaining W or Z boson decaying to hadrons. The event selection targets events where the hadronic decay products are highly collimated as a result of a large transverse momentum of the decaying boson. The data that were used consist of the entire dataset of proton-proton collisions at a center-of-mass energy of $\sqrt{s}$=8 TeV collected with the ATLAS detector at the Large Hadron Collider during the 2012 data taking period. The integrated luminosity of this dataset corresponds to 20.3 fb^{-1}. The main background processes to the WW+WZ signal are W+jets, Z+jets, top quarks and QCD multijet production. The latter is estimated using a dedicated data-driven technique while the other backgrounds are modeled by Monte Carlo simulation. In total 383 signal events are observed in data with an overall estimated background of 12452 events. The corresponding measured fiducial cross section yields:

$$\sigma_{\text{fid}}^{\text{meas}}(\text{WW} + \text{WZ}) = (31.7 \pm 10.4\,(\text{stat.}) \pm 22.1\,(\text{syst.}))\,\text{fb}.$$

This result is compatible within the errors with the Standard Model next-to-leading-order prediction of $\sigma_{\text{WW+WZ}}^{\text{SM}} = (57.7 \pm 2.9)\,\text{fb}$. The observed transverse momentum spectrum of the hadronically decaying boson is used to search for new physics via anomalous triple gauge boson couplings at WWZ and WWγ vertices. The observations are consistent with the Standard Model and confidence intervals are calculated for the parameters describing the anomalous couplings.

Contents

Chapter 1

Introduction

With the first collisions in all four detectors on March 30th in 2010, the Large Hadron Collider (LHC) has opened up a new field of measurements at unprecedented energy reach and accuracy and thus provided an insight into hitherto unexplored energy regimes in particle physics.
The three primary goals of the two largest LHC detectors ATLAS [1] and CMS [2] were the search for the Higgs boson, precision measurements of quantities predicted by the Standard Model of particle physics (SM) and searches for hints to physics beyond the Standard Model.
The first goal was achieved on July 4^{th} 2012 when the discovery of a Higgs-like particle was announced by ATLAS [3] and CMS [4]. Further measurements on couplings and spin properties of this particle have strengthened the evidence that the discovered particle is indeed the Higgs boson predicted by the SM [5].
The discovery of the Higgs boson was the last missing piece to the completion of the SM which is an experimentally well-tested theory with no contradicting experimental evidence. It is, however, known for a fact that the SM is an incomplete theory, valid only at currently accessible energy ranges. Indications for this assumption are numerous unsolved theoretical problems that cannot be explained by the SM in its current form, such as the baryon asymmetry of the universe, the presence of dark matter and dark energy, neutrino masses or the fact that the Higgs mass of 125 GeV [6] is so small compared to the Planck mass – a problem termed the hierarchy problem. All of these yet unexplained observations lead to the conclusion that there have to be as of yet unobserved new physics phenomena. Since the nature of this physics beyond the Standard Model is unknown, it is important to test the SM and constrain its parameters as tightly as possible in order to be able to detect deviations from the expectations.
A measurement of the WW+WZ production cross section offers a unique way to accomplish those two goals. The cross section measurement itself presents a test of the Standard Model predictions while limits on anomalous coupling parameters, which can be derived from the measurement as well, help to further constrain the still allowed SM parameter space.
In particular the WW pair production cross section has generated a large interest in recent years since several measurements of ATLAS and CMS at center-of-mass energies of 7 and 8 TeV [7–10] consistently observed an excess of the WW cross section compared to its next-to-leading order (NLO) Standard Model prediction. The discrepancies between theory and measurement could

be reduced by evaluating the theory cross section at next-to-next-to-leading order (NNLO) [11] and taking into account logarithmic contributions resulting from soft gluon emission at next-to-next-to-leading logarithm accuracy [12–14]. It is therefore interesting to confront the theory with updated measurements.
Furthermore, WW processes represent a major source of background to Higgs bosons decaying to $WW^{(*)}$. This makes it necessary to demonstrate the level of understanding and the correct modeling of these processes in Monte Carlo (MC) simulation.
Another important motivation for the measurement of diboson production is the fact that the scattering cross section of longitudinally polarized gauge bosons of the weak interaction gets regulated by the presence of an SM-like Higgs boson [15]. If the Higgs boson did not exist, the scattering of $W_L V_L \rightarrow W_L V_L$ cross sections, where V stands either for a W or Z boson, would violate the unitarity bound at the TeV scale. This would lead to an increase of the WV production cross section at large center-of-mass energies and could therefore in principle be detected with the cross section measurement presented in this thesis.
With larger center-of-mass energy at the LHC, an increasing fraction of events is produced in so-called *boosted* topologies. This means that due to the high momentum of decaying particles the decay products are highly collimated in the laboratory frame and cannot be resolved with ordinary techniques anymore. Various analysis methods have been developed in recent years in order to be able to analyze and profit from such event topologies. The measurement of the WW+WZ cross section in the boosted topology thus presents a suitable opportunity to test and validate those new methods.
The production of two gauge bosons moreover provides a crucial test of the non-abelian gauge structure of the electroweak sector of the SM. Gauge bosons of the weak interaction can couple to each other due to exactly this non-abelian nature of the underlying electroweak theory, such that vertices with three or four gauge boson couplings are allowed. Some of the parameters describing these couplings vanish in the SM theory but are experimentally not yet well constrained.
The measurement of the WW+WZ production cross section presented in this thesis provides a means to constrain these parameters, the boosted topology being particularly sensitive to deviations of these couplings from the SM prediction.
Hereafter WW/WZ will be referred to as WV for simplification and to denote the fact that either process is possible.
This thesis is organized in the following way:
Chapter 2 gives an overview of the underlying theoretical foundations of the SM, physics at hadron colliders and diboson production. In Chapter 3 the ATLAS detector and the data acquisition infrastructure are described. Chapter 4 provides a summary over the different signal and background processes, their production mechanisms, cross sections and signatures in the detector. In Chapter 5, the concept of particle jets, their calibration, jet grooming and substructure are introduced. Chapter 6 focuses on the physics object definition, the Monte Carlo samples that are used for signal and background modeling and the event selection. The estimation of the different background processes is presented in Chapter 7. The methods for the extraction of the measured cross section and systematics uncertainties are explained in Chapter 8 and 9, respectively. In Chapter 10, the results on the cross section and the limits on anomalous triple gauge coupling parameters are presented. Finally Chapter 11 provides both a summary and an outlook on the prospects for a repetition of this analysis with LHC Run2 data.

Chapter 2

The Standard Model of Particle Physics

The currently accepted theory to describe the (known) fundamental particles and their interactions is the so-called Standard Model of Particle Physics. The SM incorporates three of the four fundamental interactions in nature, namely the electromagnetic, the weak and the strong force. The fourth fundamental interaction, the gravitational force has not yet been successfully incorporated into this theory. From the experimental point of view this is, however, of little importance since its strength is roughly 26 orders of magnitude smaller than that of the weak interaction (WI) which is the weakest of the three SM forces [16]. For this reason it does not play any role at typical scales of particle physics.
The SM is formulated mathematically in the framework of a Quantum Field Theory which represents particles as excitations of fundamental fields.
In general, the theories of the fundamental forces are gauge theories, meaning the mathematical form of the interactions is obtained by applying a local gauge transformation belonging to a certain gauge group on the fundamental fields, yet requiring the entire Lagrangian to be invariant under this transformation. In this way one is forced to introduce additional fields and interaction terms which then describe the dynamics of the interaction. A priori, there is no compelling reason why the form of all fundamental forces should be able to be derived by this gauge principle.
However, one argument why this should be the case is that any theory describing the interactions of particles must deliver finite results when calculating processes occurring in nature.
At first sight some of the theories in the SM actually lead to infinite integrals, e.g. the self-energy of the electron [17]. This fact posed a big challenge to the theory of electromagnetic interactions called "Quantum Electrodynamics" until techniques called *regularization* and *renormalization* were developed.
With these techniques it was possible to overcome the infinities and obtain meaningful results (see Section 2.1.4). Nevertheless, it was not obvious whether a theory was renormalizable at all until in 1972 it was shown by t'Hooft that every locally gauge invariant theory is renormalizable (provided the mass dimension of the operators is not higher than four) [18].
This is why the gauge principle is such a powerful tool to derive the mathematical form of fundamental interactions.
The SM, its particle content and its forces will be elaborated on in the following chapter.

Furthermore, a brief introduction to the Higgs mechanism, which is the only known way to give mass to the gauge bosons and fermions without destroying gauge invariance, will be given. The information in this chapter is based on Ref. [16, 19, 20].

2.1. Particles and Forces in the Standard Model

Particles of the SM are classified into fermions (particles with half-integer spin) and bosons (integer spin) which at the same time offers a distinction between building blocks of matter (fermions) and force carrier particles (bosons). The fermions can further be divided into six quarks and six leptons.
The six quarks are called up (u), down (d), charm (c), strange (s), top (t) and bottom (b). Different types of quarks are also referred to as different *flavours*. Leptons are divided into the charged leptons, which are called Electron, Muon and Tau lepton (e^-,μ^-,τ^-) and electrically neutral neutrinos (ν_e, ν_μ, ν_τ). Quarks and charged leptons are all massive while the neutrinos are treated as massless in the SM. In fact it turned out that also neutrinos are massive, a conclusion drawn from the observation of neutrino oscillations [21]. The magnitude of the neutrino masses could not yet be measured, only upper bounds could be set and the mechanism how they aquire their mass is different from the other fermions.
Quarks and leptons are grouped into three so-called generations. The only difference between a particle in one generation and the corresponding particle in one of the other generations is its mass (and properties depending on the mass such as lifetime, kinematically allowed decay modes for example). Each generation comprises two quarks and two leptons where the particle mass increases from one generation to the next (for neutrinos this is not yet known).
The number of lepton generations has to be equal to the number of quark generations in order to avoid non-renormalizable divergencies in fermion loop diagrams [19]. However, the total number of existing fermion generations is not predicted by theory. Before the discovery of the Higgs boson, a fourth sequential quark generation was still within the bounds of possibility. The discovery of the Higgs boson and the measurement of its production cross section and decay rates yet let to an exclusion of a fourth generation of chiral quarks [22] (the concept of chirality will be explained in Section 2.1.1). A fourth generation of vector-like quarks [23] would, however, still be allowed.
Table 2.1 shows the fundamental fermions together with several important properties.
The bosons, also referred to as gauge bosons, are mediating the forces between the fermions. Each of the fundamental forces has its associated force carrying particles. The electromagnetic force is mediated by the massless photon (γ), the carriers of the strong force are the equally massless gluons (g) which come in eight different colour combinations, where the term *colour* refers to the so-called colour charge of quarks and gluons (see Section 2.1.4). The weak force is mediated by three massive bosons called W^+, W^- and Z^0 with the experimentally measured masses of [6]:

$$m_W = 80.385 \pm 0.015\,\text{GeV}, \tag{2.1}$$
$$m_{Z^0} = 91.1876 \pm 0.0021\,\text{GeV}. \tag{2.2}$$

Associated to each of the aforementioned particles is a corresponding antiparticle with opposite quantum numbers such as electric charge or lepton number, where the photon, the Z^0 and the gluons are their own antiparticles and the W^+ is the antiparticle of the W^-.
The photon couples to all particles that carry an electric charge, thus all charged particles

	Quarks			Leptons		
Generation	Flavour	Charge (e)	Mass (MeV)	Flavour	Charge (e)	Mass (MeV)
1	u	2/3	$2.3^{+0.7}_{-0.5}$	e	-1	0.511
	d	-1/3	$4.8^{+0.5}_{-0.3}$	ν_e	0	$< 2 \cdot 10^{-6}$
2	c	2/3	1275±25	μ	-1	105
	s	-1/3	95±5	ν_μ	0	< 0.19
3	t	2/3	$(173.21 \pm 0.87) \cdot 10^3$	τ	-1	1776.86
	b	-1/3	$(4.18 \pm 0.03) \cdot 10^3$	ν_τ	0	< 18.2

Table 2.1.: The fundamental fermions known up to date are shown. Numbers are taken from Ref. [6]. On the left-hand side are the quarks, on the right-hand side the leptons. It can be seen that the mass is increasing with the generation. It is interesting to note that the heaviest quark is around 35000 times as heavy as the lightest one. For the quark masses, one has to keep in mind that with the exception of the top quark, the quarks are confined in hadrons and are not observed as physical particles which makes their masses dependent on the particular mass definition in the relevant theoretical framework.

experience the electromagnetic force. Whether a particle participates in the weak and strong interaction depends on other quantum numbers such as the weak isospin and hypercharge for the weak and the colour charge for the strong interaction which will be discussed in Section 2.1.4. Generally, the strong force acts only on quarks and gluons while it depends on a property referred to as chirality (see Section 2.1.1) whether a particle participates in the weak interaction or not. An overview of the gauge bosons is given in Table 2.2. One last particle, the Higgs boson has

Interaction	Gauge Boson	Mass (GeV)	El. Charge (e)	Range (m)
Electromagnetic	Photon (γ)	0	0	∞
Weak	$W^\pm$	80.4	± 1	10^{-18}
	Z^0	91.2	0	
Strong	8 Gluons (g)	0	0	10^{-15}

Table 2.2.: The three fundamental interactions together with their mediating particles.

not yet been mentioned because it is somewhat different from all the others and plays a special role in the SM. Despite being a boson, it cannot be counted to the classic gauge bosons that transmit forces, but it is the excitation of a *Higgs field* which is necessary to give masses to the fundamental particles and which will be discussed in more detail in Section 2.1.2.
The fundamental fermions are mathematically described by four component Dirac spinors which are functions of the spacetime coordinates x_μ.
So far, electromagnetic and weak force have been treated as two completely separate phenomena but they can be unified and both be seen as two aspects of the same *electroweak force*. The unification of the two theories was first accomplished by Glashow, Weinberg and Salam [24–26] and is therefore referred to as the GWS-theory.

2.1.1. The Electroweak Interaction

The electroweak theory is a gauge theory based on the symmetry group $SU(2)_L \otimes U(1)_Y$. The fact that the gauge group is a product of two groups causes the theory to manifest itself as two different interactions: the electromagnetic force, corresponding to the $U(1)_Y$ symmetry and the weak force, corresponding to the $SU(2)_L$ symmetry.
The generators of the two gauge groups are the weak isospin operator $\mathbf{T}$ for the $SU(2)_L$ group and the weak hypercharge operator Y for $U(1)_Y$. The weak isospin operator is a three-component operator where the components can be expressed in terms of the Pauli matrices τ_i as $T_i = \frac{1}{2}\tau_i$. Since those matrices do not commute, the $SU(2)_L$ is a non-abelian group. Both the third component of the weak isospin T_3 as well as the weak hypercharge Y are conserved in all interactions. The weak isospin and the weak hypercharge are connected to the electric charge Q via the Gell-Man-Nishijima relation:

$$Q = T_3 + \frac{Y}{2}. \tag{2.3}$$

It becomes apparent in this equation that the electromagnetic and the weak interaction must be two parts of the same higher theory since the electric charge is closely related to fundamental quantum numbers of the weak interaction.
One outstanding feature of the weak interaction is that it strictly distinguishes between so-called right-handed and left-handed particles. This feature of the WI can be understood by considering the mathematical description of fermions. Fermions are represented by four component Dirac-spinors $\mathbf{\Psi}$ which are functions of the spacetime coordinates x_μ. The handedness of a fermion refers to an intrinsic property of the particle called *chirality*. For a Dirac fermion, the eigenvalues of the γ^5 operator acting on the spinor are ± 1. The chirality of the fermion is then right-handed for the eigenstate associated to the positive eigenvalue and left-handed for the negative case.
The chirality is closely related to another quantity called *helicity* h which is defined as the projection of a particle's spin $\vec{S}$ on its direction of motion:

$$h = \vec{S} \cdot \vec{e_p}. \tag{2.4}$$

For a massless particle which has only two spin states, helicity is equal to chirality (except for a factor S). For a particle with a negligibly small mass, this is true to a good approximation. However, since massive particles can never travel at the speed of light, it is always possible to find a boosted reference frame (travelling faster than the particle but in the same direction), in which the particle's direction is reversed while the direction of its spin stays the same. The particle thus changed its helicity in this reference frame.
Summarizing, chirality is an intrinsic quantum mechanical particle property that is equal to its helicity in case of vanishing particle mass. For massive particles, the chirality is still either positive or negative, while its helicity is actually a quantum admixture of positive and negative.
In order to describe the chirality of particles, one has to introduce the projection operators P_L and P_R. With these operators left-handed and right-handed spinors can be constructed:

$$\mathbf{\Psi}_L = P_L \mathbf{\Psi} = \frac{1-\gamma_5}{2}\mathbf{\Psi}, \qquad \mathbf{\Psi}_R = P_R \mathbf{\Psi} = \frac{1+\gamma_5}{2}\mathbf{\Psi}. \tag{2.5}$$

The peculiar behaviour of the WI regarding the distinction between left- and right-handed particles can be formulated more precisely in the context of parity violation. The parity transformation is a symmetry transformation where all space coordinates are reversed, whereas time coordinates do not change: $(x,t) \to (-x,t)$.
It had been taken for granted that all laws of nature must be symmetric under this transformation

Table 2.3.: Multiplets of the weak interaction with quantum numbers **T**, T and the electric charge Q for the fundamental fermions. d′, s′ and b′ denote eigenstates of the weak interaction (see Section 2.1.3).

Q	**T**	T_3	Y			
0	$\frac{1}{2}$	$\frac{1}{2}$	-1	$\begin{pmatrix}\nu_e\\ e^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau^-\end{pmatrix}_L$
-1	$\frac{1}{2}$	$-\frac{1}{2}$	-1			
$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\begin{pmatrix}\mathrm{u}\\ \mathrm{d}'\end{pmatrix}_L$	$\begin{pmatrix}\mathrm{c}\\ \mathrm{s}'\end{pmatrix}_L$	$\begin{pmatrix}\mathrm{t}\\ \mathrm{b}'\end{pmatrix}_L$
$-\frac{1}{3}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{3}$			
-1	0	0	-2	e^-_R	μ^-_R	τ^-_R
$\frac{2}{3}$	0	0	$\frac{4}{3}$	u_R	c_R	t_R
$\frac{1}{3}$	0	0	$-\frac{2}{3}$	d_R	s_R	b_R

until it was shown by Wu in 1956 [27] that the weak interaction actually does violate this symmetry. Not only is the parity violated but it is actually violated in a maximal way. This was shown by Goldhaber already in 1957 [28]. It was measured that the helicity of neutrinos is always -1 while antineutrinos only exist with a helicity of +1[1]. If parity were conserved in the weak interaction, one would expect an equal distribution of left- and right-handedness amongst neutrinos and antineutrinos.

Due to the parity-violating nature of the weak interaction it is convenient to group the fundamental fermions according to their chirality. The left-handed fermions form doublets of the weak isospin whereas the right-handed ones form singlets (see Table 2.3). Once the correct gauge group and its generators are known, the mathematical form of the interactions, here the electroweak one, is obtained by demanding local gauge invariance of the Lagrangian for free particles

$$\mathcal{L}_0 = \overline{\mathbf{\Psi}}_L i\gamma_\mu \partial^\mu \mathbf{\Psi}_L + \overline{\mathbf{\Psi}}_R i\gamma_\mu \partial^\mu \mathbf{\Psi}_R, \tag{2.6}$$

under the local phase-transformations

$$\mathbf{\Psi}_L \to \mathbf{\Psi'}_L = e^{i(g\boldsymbol{\alpha}(x)\cdot\mathbf{T} + g'\beta(x)Y)}\mathbf{\Psi}_L, \tag{2.7}$$

$$\mathbf{\Psi}_R \to \mathbf{\Psi'}_R = e^{ig'\beta(x)Y}\mathbf{\Psi}_R, \tag{2.8}$$

where the phases $\boldsymbol{\alpha}(x)$ and $\beta(x)$ are arbitrary, real-valued functions of spacetime x_μ.
$\mathbf{\Psi}_L$ stands for a left-handed weak isospin doublet ($\mathbf{T} = \frac{1}{2}$) and $\mathbf{\Psi}_R$ stands for a right-handed singlet ($\mathbf{T} = 0$). It is implicitly understood that it is summed over all quarks and leptons. Here it is assumed that the fermions are massless. The masses are going to be introduced at a later stage (see section 2.1.2). In order to leave the Lagrangian invariant under this transformation, one is forced to introduce an isospin triplet

$$\mathbf{W}^\mu = (W_1^\mu, W_2^\mu, W_3^\mu)^T$$

[1] This statement has a subtlety to it. As it is certain now that neutrinos do indeed have a non-vanishing mass, there must be also right-handed neutrinos since the helicity cannot be exactly 1 anymore. However, in a π^--decay the antineutrino is always right-handed in the center-of-mass (CM) frame which already represents a violation of parity.

of vector fields for the $SU(2)_L$ group and one singlet B^μ for the $U(1)_Y$ group. Adding these generator terms to the derivative ∂^μ results in the so-called *covariant derivative*:

$$D^\mu = \partial^\mu + ig\mathbf{T}\cdot\mathbf{W}^\mu + i\frac{g'}{2}YB^\mu. \tag{2.9}$$

The coupling constants g and g' for the gauge fields $\mathbf{W}^\mu$ and B^μ, respectively, are not yet related in any sense. Up to this point g (g') only stands for the coupling strength between fermions and the $SU(2)_L$ ($U(1)_Y$) bosons. Equation 2.9 is inserted in the initial Lagrangian and for full generality a free kinetic term for the gauge fields is be added:

$$\mathcal{L}_{\text{Gauge}} = -\frac{1}{4}\mathbf{W}_{\mu\nu}\cdot\mathbf{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}\cdot B^{\mu\nu}. \tag{2.10}$$

With the field strength tensors given by:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.11}$$

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon_{ijk}W^j_\mu W^k_\nu. \tag{2.12}$$

This results in a Lagrangian with the following form:

$$\mathcal{L}_1 = \overline{\mathbf{\Psi}}_L i\gamma^\mu\left[\partial^\mu + ig\mathbf{T}\cdot\mathbf{W}^\mu + i\frac{g'}{2}YB^\mu\right]\mathbf{\Psi}_L + \overline{\mathbf{\Psi}}_R i\gamma^\mu\left[\partial^\mu + i\frac{g'}{2}YB^\mu\right]\mathbf{\Psi}_R + \mathcal{L}_{\text{Gauge}}. \tag{2.13}$$

$\mathcal{L}_{\text{Gauge}}$ describes the self couplings and kinetic energies of the $\mathbf{W}^\mu$ and B^μ fields whereas the first term describes the couplings of the fundamental fermions to the force carrying particles of the electroweak interaction. The $SU(2)_L$ field strength tensors (2.12) contain a bilinear term in the W^i fields which is responsible for the trilinear and quadrilinear coupling of gauge bosons described more thoroughly in section 2.3.1.
The physical fields of the $\mathrm{W}^\pm$ bosons can be constructed from the $\mathbf{W}^\mu$ as

$$W^{(\pm)\mu} = \frac{1}{\sqrt{2}}(W^\mu_1 \mp iW^\mu_2), \tag{2.14}$$

which can be regarded as weak-isospin creation or annihilation operators.
It is tempting to identify W^μ_3 with the Z^0 and the B^μ with the photon-field, but if that were the case, neutrinos could couple to the photon. However what can be done is to express the photon field A^μ as linear combination of the two fields where the coefficients are chosen in exactly such a way that the coupling of the photon field to the neutrino field vanishes. Once this linear combination is determined, the Z^μ field which has to be orthogonal to the photon field can be expressed as a rotation of the latter. These considerations lead to:

$$\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix}\cdot\begin{pmatrix} B^\mu \\ W^\mu_3 \end{pmatrix} \tag{2.15}$$

where θ_W is the so-called Weinberg or weak-mixing angle, defined by:

$$\sin(\theta_W) = \frac{g'}{\sqrt{g^2 + g'^2}}. \tag{2.16}$$

The coupling of leptons to the W and Z fields must be the same for all lepton flavours, a trait of the theory that results from the $SU(2)_L$ symmetry and that is known as *lepton universality*.

Another important aspect of the weak interaction regarding the three generation of fermions is, that transitions can only occur within one generation, no generations can be crossed. This feature is called lepton flavour conservation in case of lepton transitions. The same observation also holds for quarks as far as the so-called *weak eigenstates* are considered. In reality, however, also cross-generation quark decays can occur which is possible due to a difference of the eigenstates of the weak interaction and the actual *mass eigenstates* in the quark sector (see Section 2.1.3).

2.1.2. The Higgs Mechanism: Masses for Gauge Bosons and Fermions

Up to this point the gauge bosons have been treated as massless, as the introduction of an explicit mass term would inevitably destroy the gauge invariance. However, as mentioned before it is an experimental fact that these particles do have non-vanishing masses (Eq. 2.1, 2.2).
One potential solution to this dilemma in the theoretical description of the weak interaction was proposed in 1970 by Higgs [29], Brout and Englert [30].
The mechanism which permits to give masses to the gauge bosons (and also the fermions) without breaking gauge invariance is called *spontaneous symmetry breaking* and has a very close connection to the *Higgs mechanism*. The general idea of spontaneous symmetry breaking is to have a Lagrangian which is invariant under transformations of the symmetry groups in question, but then an explicit ground state which breaks the symmetry is chosen. As will be discussed below, this mechanism leads to massive weak gauge bosons, a massless photon and an additional massive boson, called Higgs boson. For this, a scalar weak isospin doublet Φ is introduced:

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix}, \text{ with } \mathbf{T} = \frac{1}{2}, Y = 1. \tag{2.17}$$

Now, another Lagrangian $\mathcal{L}_{\text{Higgs}}$ which is invariant under $SU(2)_L \otimes U(1)$ transformations as well is added to $\mathcal{L}_1$:

$$\mathcal{L}_{\text{Higgs}} = (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi), \tag{2.18}$$

with the covariant derivative

$$D^\mu = \partial^\mu + i\frac{g}{2}\boldsymbol{\tau} \cdot \mathbf{W}^\mu + i\frac{g'}{2} \cdot B^\mu, \tag{2.19}$$

and the Higgs-potential

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \text{ with } \mu^2 < 0 \text{ and } \lambda > 0. \tag{2.20}$$

Since the shape of this potential vaguely resembles a Mexican hat (see Figure 2.1), it is often referred to as “Mexican Hat Potential”. The potential has a minimum at

$$\Phi\Phi^\dagger = \frac{-\mu^2}{2\lambda}. \tag{2.21}$$

In order to break the symmetry of the ground state, a point on this circle is chosen as vacuum expectation value :

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \qquad \text{with} \qquad v = \sqrt{\frac{-\mu^2}{\lambda}}. \tag{2.22}$$

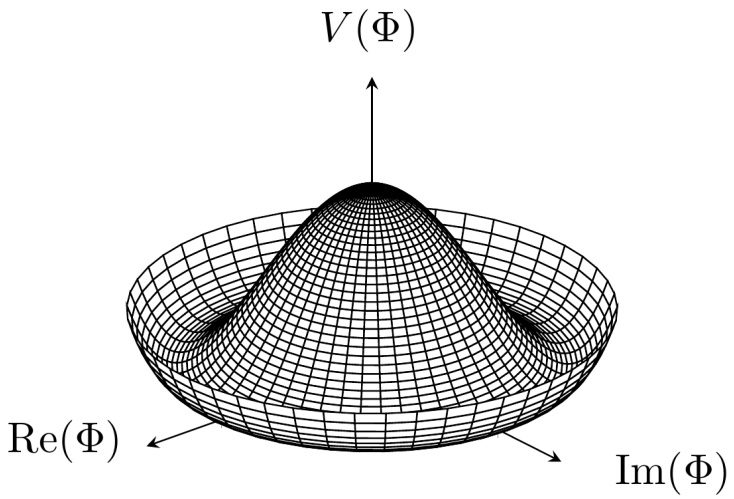


Figure 2.1.: Mexican hat potential with the global minima at $\Phi\Phi^\dagger = \frac{-\mu^2}{2\lambda}$.

Now that the minimum of the potential is known, it can be expanded around Φ_0:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta(x) \end{pmatrix}. \tag{2.23}$$

Inserting the ground state (Equation 2.23) into this equation, yields after neglecting higher order η-mixed terms:

$$\mathcal{L}_{Higgs} = \left[\frac{1}{2}(\partial^\mu \eta)(\partial_\mu \eta) + \mu^2 \eta^2\right] + \frac{1}{2} \cdot \frac{g^2 v^2}{4}(|W_\mu^+|^2 + |W_\mu^-|^2) + \frac{1}{2} \cdot \frac{g^2 v^2}{4\cos^2\theta_W}|Z_\mu|^2. \tag{2.24}$$

Where Equation 2.14 was used to express the physical $W^{(\pm)\mu}$ fields by the W_1, W_2 fields and Equations 2.15 and 2.16 were used to express A^μ and Z^μ as admixture of the W_3^μ and B^μ fields. Looking at the second term, it becomes obvious that the $\mathrm{W}^\pm$ and Z^0-bosons have now acquired a mass given by

$$m_W = \frac{gv}{2} \text{ and } m_Z = \frac{gv}{2\cos\theta_W} = \frac{m_W}{\cos\theta_W}. \tag{2.25}$$

However, for A^μ there is no mass term, so the photon remains massless as desired. The first term represents a new particle with a mass of $m = \sqrt{-2\mu}$. This is the Higgs boson whose existence had been predicted already in 1964, but since μ is not determined by theory, its mass was unknown, making it difficult to be searched for. In 2012 the particle was finally discovered by ATLAS and CMS. The currently highest precision mass measurement yields [6]:

$$m_H = (125.09 \pm 0.21 \pm 0.11)\,\mathrm{GeV}. \tag{2.26}$$

The last missing ingredients for a complete description of the electroweak interaction are now the masses of the fermions which are generated in a quite similar way, i.e. also via spontaneous symmetry breaking. A nice feature of the SM is that the same Higgs doublet we used to generate the masses of the weak gauge bosons can be employed to generate masses for the fermions. One last Lagrangian with the Yukawa couplings of the fermions to the Higgs field is added, here split up into quarks and leptons:

$$\mathcal{L}_{\text{Fermion}}^{\text{Yukawa}} = \mathcal{L}_{\text{Lepton}}^{\text{Yukawa}} + \mathcal{L}_{\text{Quark}}^{\text{Yukawa}}. \tag{2.27}$$

The Lagrangian $\mathcal{L}_{\text{Lepton}}^{\text{Yukawa}}$ can directly be written as:

$$\mathcal{L}_{\text{Lepton}}^{\text{Yukawa}} = \sum_{j=1}^{3} -G_j \left[(\overline{\nu}_j, \overline{l}_j)_L \Phi l_{jR} + \text{hermitian conjugate (h.c.)} \right], \tag{2.28}$$

with the sum over all charged leptons l_j and neutrinos ν_j of the three generations. After the symmetry is broken exactly as in Equation 2.23, mass terms and couplings to the Higgs field for the charged leptons appear:

$$\mathcal{L}_{\text{Lepton}}^{\text{Yukawa}} = \sum_{j=1}^{3} -\frac{G_j}{\sqrt{2}} v \left(\overline{l}_{jL} l_{jR} + \overline{l}_{jR} l_{jL} \right) - \frac{G_j}{\sqrt{2}} \left(\overline{l}_{jL} l_{jR} + \overline{l}_{jR} l_{jL} \right) + \text{h.c.} \tag{2.29}$$

To generate the upper member of a quark doublet, a modified Higgs doublet must be constructed:

$$\Phi_c = -i\tau_2 \Phi^* = \begin{pmatrix} -\overline{\phi^0} \\ \phi^- \end{pmatrix}, \text{ with } \mathbf{T} = \frac{1}{2}, Y = -1, \tag{2.30}$$

which after spontaneous symmetry breaking becomes:

$$\Phi_c \rightarrow \Phi_c(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} v + \eta(x) \\ 0 \end{pmatrix}. \tag{2.31}$$

The quark mass generating Lagrangian is then given by:

$$\mathcal{L}_{\text{Quark}}^{\text{Yukawa}} = \sum_{i,j=1}^{3} -G_{ij}^{d} (\overline{u}_i \overline{d'}_i)_L \Phi d_{jR} - G_{ij}^{u} (\overline{u}_i \overline{d'}_i)_L \Phi_c u_{jR} + \text{h.c.} \tag{2.32}$$

where the u_i' and d_i' are the eigenstates of the weak interaction of the up- and down-type quark-flavours, connected to the mass eigenstates via the complex G-matrices.
This difference between weak eigenstates and mass eigenstates is an important feature of the quark model and is explained in more detail in Section 2.1.3. Equation 2.32 then yields the quark mass terms similar to Equation 2.29.
Summing up all previously derived components leads to the final Lagrangian of the weak interaction, including the kinetic energies of the fermions and gauge bosons, the self interactions of the gauge bosons, the interactions of the fermions with the gauge bosons and the Higgs-couplings and masses of fermions and gauge bosons:

$$\mathcal{L}_{\text{WI}} = \mathcal{L}_1 + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Fermion}}. \tag{2.33}$$

2.1.3. Quark Mixing and the CKM matrix

As stated earlier, the $\text{W}^{\pm}$-bosons can be interpreted as creation/annihilation operators of the weak isospin, implying that quarks absorbing or emitting a W boson have to change their flavour. In the quark sector, flavour transitions can occur within one generation or also cross-generationally. The only condition for a transition in order to be allowed (except for kinematic constraints) is, that the final quark must have different electric charge than the initial one. This so-called *quark mixing* can only occur because the eigenstates of the weak interaction (d', s', b') are not exactly the same as the experimentally observed mass eigenstates of the down-type quarks (d, s, b). The

two sets are connected via a matrix which was introduced in 1973 by Cabibbo, Kobayashi and Maskawa [31, 32] and is therefore called the "CKM-matrix":

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.34}$$

The transition probability for a certain up-type quark U to a certain down-type quark D (and vice versa) is proportional to the squared matrix element $|V_{UD}|^2$. This quark-mixing formalism has the advantage that all up-type quarks couple with the same strength to the weak eigenstates of their generation. Without the CKM matrix, different coupling constants would have to be introduced for every transition.
By convention, the weak eigenstates of the down-type quarks are expressed as mixed mass eigenstates whereas in the up-type case, both eigenstates are equal. It would be equally possible to attribute all the mixing to the up-type quarks and leave the down-type quarks untouched. The CKM matrix has to be unitary, meaning

$$\sum_k V_{ik} V_{jk}^* = \delta_{ij}, \tag{2.35}$$

implying the conditions for the sums of the squared elements in any given single column or row:

$$\sum_k |V_{ik}|^2 = \sum_n |V_{nj}|^2 = 1. \tag{2.36}$$

The elements of the matrix can be determined experimentally or at least strongly constrained through the application of the unitarity condition, for the sake of simplicity the individual elements are quoted without errors here [6]:

$$V_{CKM} = \begin{pmatrix} 0.97427 & 0.22536 & 0.00355 \\ 0.22522 & 0.97343 & 0.0414 \\ 0.00886 & 0.0405 & 0.99914 \end{pmatrix}. \tag{2.37}$$

Upon observation of the numerical values, it becomes apparent that transitions between quarks still preferably occur within one generation and transitions are increasingly suppressed with increasing number of generations that have to be crossed. The CKM formalism shows that the weak interaction does in fact also constrain transitions in the quark sector to happen within a given generation. The reason why here cross-generation transitions are observed nevertheless is that the experimentally observed quark eigenstates are just a mixture of different eigenstates of the weak interaction.

2.1.4. Quantum Chromodynamics

The last one of the fundamental forces that shall be discussed here is the strong interaction which is responsible for the interactions between quarks (additionally to the electromagnetic interaction) and for holding the nucleons in the atomic nucleus together despite its positive electric charges.
The theory of the strong interaction can also be obtained via the gauge principle where the gauge group is the $SU(3)$ instead of the $SU(2)_L \otimes U(1)_Y$ group. The term *colour* is used to designate the quantum numbers associated with this group and this is why this theory of the

dynamics of quarks is also called quantum chromodynamics and the gauge group is referred to as $SU(3)_C$. The three colour charges are called red (r), green (g) and blue (b), but since the $SU(3)_C$ symmetry is an exact symmetry, the assignment of one specific colour to a quark is arbitrary and can be rotated at will. In order to derive the Lagrangian for the strong interaction, one can start again with the Lagrangian for free quark fields:

$$\mathcal{L}_0 = \overline{q_j}(i\gamma_\mu \partial^\mu - m)q_j, \tag{2.38}$$

with the three color fields q_1, q_2, and q_3. This Lagrangian incorporates only one quark type, for a full description of the quark sector, one would have to sum over all flavours. $\mathcal{L}_0$ is now required to be invariant under the local $SU(3)_C$ transformation:

$$q(x) \to q'(x) = \exp\left[-i\theta_a(x)\frac{\lambda_a}{2}\right] q(x), \tag{2.39}$$

where a goes from one to eight and the $\theta_a(x)$ again are real valued functions of x. The λ_a are the so-called Gell-Mann matrices which are the generators of the $SU(3)_C$ group analogously to the Pauli matrices for the $SU(2)_L$.
The covariant derivative in this case is given by be the following equation:

$$D_\mu = \partial_\mu - ig\frac{\lambda_a}{2}A^a_\mu(x). \tag{2.40}$$

The $A^a_\mu(x)$ are the eight gauge fields of the QCD which are called *gluons*. Gauge invariance can here only be achieved, if the $A^a_\mu(x)$ exhibit the following transformation behaviour:

$$A^a_\mu(x) \to A^a_\mu(x) - \frac{1}{g}\partial^a_\mu\theta^a(x) + f^{abc}\theta^b A^c_\mu(x), \tag{2.41}$$

with the f^{abc} being real constants, the so-called structure constants of the $SU(3)_C$. The form of the gluonic field tensor is also determined by the demand for gauge invariance:

$$G^a_{\mu\nu}(x) = \partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) + g f^{abc} A^b_\mu(x) A^c_\nu(x). \tag{2.42}$$

Again a kinetic term for the gluonic fields is added:

$$\mathcal{L}_{gluon} = -\frac{1}{4}G^a_{\mu\nu}(x)G^{\mu\nu}_a(x), \tag{2.43}$$

yielding the final QCD Lagrangian:

$$\mathcal{L}_{QCD} = \overline{q_j}(i\gamma_\mu D^\mu - m)q_j - \frac{1}{4}G^a_{\mu\nu}(x)G^{\mu\nu}_a(x). \tag{2.44}$$

A number of important properties of the strong interaction can be derived by analyzing the form of the Lagrangian which shall be done here in the following.
Phenomenologically, QCD can be described as the interaction of six different quarks, carrying one colour charge with a set of eight gluons where each of the gluons carries one unit of colour and one unit of anti-colour. Antiquarks are carrying anti-colours ($\overline{g}$,$\overline{r}$,$\overline{b}$). Following from the fact that the $SU(3)_C$ is an exact symmetry, colour charge is strictly conserved in the strong interaction, e.g. if a red quark radiates a gluon and becomes a blue quark, the gluon must carry away one unit of red and one of anti-blue, such that the net-colour is still red:

$$q(r) \to q(b) + g(r, \overline{b}).$$

The fact that gluons carry colour charge themselves is a direct consequence of the gauge group $SU(3)_C$ being non-abelian. This fact, in contrast to the electromagnetic interaction where the photon is uncharged, leads to several interesting properties of QCD.
Because of their colour charge, gluons can interact with each other such that there are three-gluon and four-gluon vertices. The self-interaction of the gluons leads to a phenomenon which is called *quark confinement*. This means no free particle carrying net colour charge can be observed, all free particles have to be "white", that is to say either carrying the same amount of colour and corresponding anti-colour charges or carrying all three (anti-) colours at the same time.
Because every quark is colour charged, this fact explains why quarks cannot be observed freely. A single quark q must either be coupled to another anti-quark $\bar{q}'$ of opposite colour charge (mesons) or to two other quarks of the remaining two colours (baryons) to appear white.
If the quark-antiquark pair of a meson were to be separated, there would build up a "colour-tube" caused by the self-interacting gluons in between the two quarks. This colour-tube would contain an increasing amount of energy with increasing distance. Once the energy is large enough, a new quark-antiquark pair emerges in between the initial one and both final mesons are colourless again. This mechanism is called string-breaking and is illustrated in Figure 2.2. The contrary

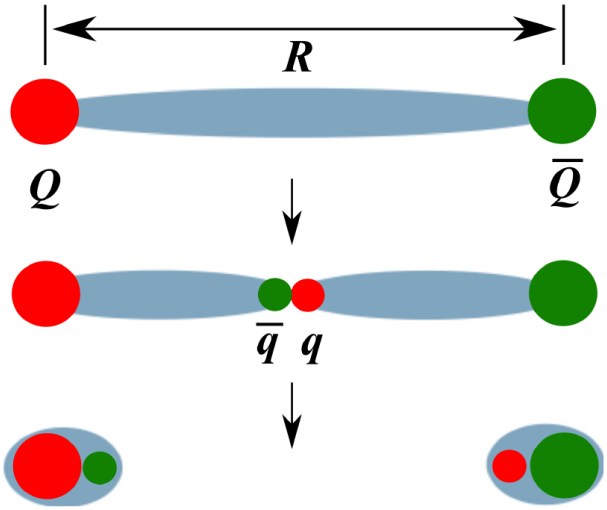


Figure 2.2.: Illustration of string breaking, borrowed from [33].

of this behaviour is what is called *asymptotic freedom*. As the coupling between quarks (α_s) increases with the distance (smaller momentum transfer Q^2), it decreases with smaller distances and higher Q^2 so that in the limit of $Q^2 \to \infty$ the quarks can be treated as bare quasi-free particles:

$$\lim_{Q \to \infty} \alpha_s(Q^2) \to 0.$$

This is why QCD is only applicable as a perturbative theory in case of high energies, because only here the assumption that the coupling constant α_s satisfies $\alpha_s \ll 1$ holds. The fact that the coupling constant is not really a constant but changes with momentum transfer is referred to as *running coupling constant*. The measured running of the strong coupling constant is shown in Figure 2.3.
Although the mediating bosons of the strong interaction are massless, which would in principle allow for an infinite range, their self coupling causes the effective range of the strong force to be finite (see Table 2.2). Calculations of physical observables described by Feynman diagrams involving at least one quark or gluon loop in QCD have to make use of renormalization in order to deliver sensible and finite results. Otherwise, as in the case of all other quantum field theories,

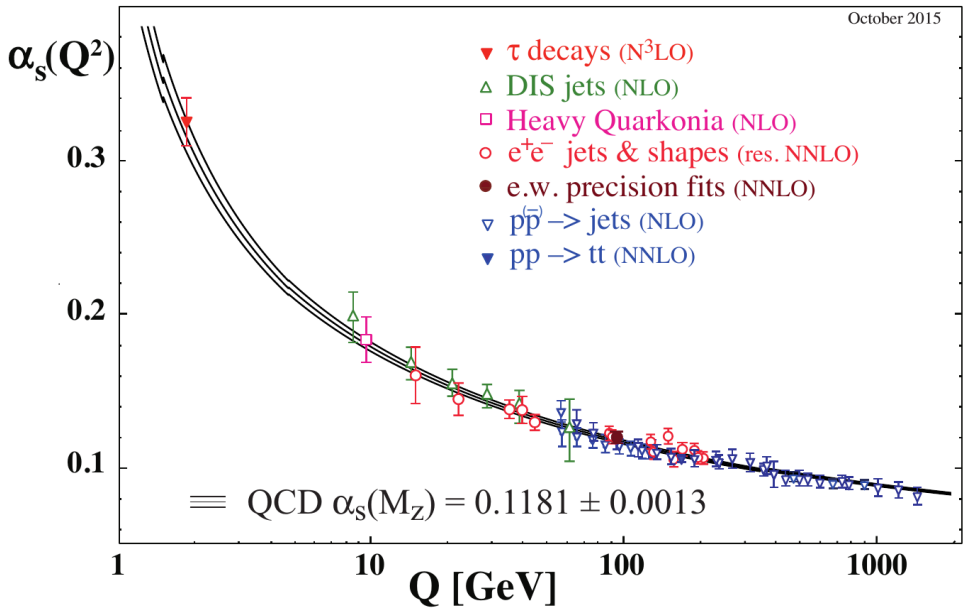


Figure 2.3.: 2015 world average of α_s [34].

divergencies occur in the calculation of physical quantities when integrating over the internal loop momenta.
The cause for these infinities is that the perturbation theory used to compute physical observables is only an effective theory, valid only in a certain finite energy range. In the calculation of e.g. production cross sections, terms resulting from outside of this energy range can contribute as well, leading to divergent results.
The divergencies can be removed by reparametrizing (or "renormalizing") the involved coupling constant from its "bare" value to its experimentally measured physical value which is dependent on the scale of the momentum transfer Q^2. The divergencies in this approach are absorbed in the bare coupling constant allowing to carry out the integration over internal loops with finite results. This approach demands the introduction of a new energy scale parameter μ_R, referred to as the *renormalization scale* at which the physical coupling constant is evaluated. The price to be paid for eliminating the singularities in the theory is exactly this dependency of the obtained results on the artificially introduced scale parameter μ_R.
The renormalization scale is an unphysical scale and in principle the result of a calculated observable χ should not depend on it, a fact expressed by the *renormalization group equation*:

$$\frac{\partial \chi(Q^2, \mu_R^2, ...)}{\partial \ln \mu_R^2} = 0. \tag{2.45}$$

The renormalization group equation will not be satisfied in reality unless the perturbative expansion is carried out up to all orders which is naturally impossible. The strength by which the results vary with a variation of the renormalization scale reflects the lack of knowledge of the exact behaviour of the underlying theory at energies far away from the typical renormalization scale. The renormalization scale dependence can, however, be used as a measure of the accuracy of a particular result, as a vanishing μ_R dependence signifies that sufficient terms have been included in the perturbative expansion.
Although the specific choice of μ_R is somehow arbitrary, it is customarily chosen to correspond to typical energy scales in the calculated process, e.g. the mass of the produced particles or the momentum transfer Q^2.

In leading order perturbation theory, the dependence of the strong coupling constant α_S on Q^2 and μ_R^2 can be expressed as:

$$\alpha_S(Q^2) = \frac{\alpha_S(\mu_R^2)}{1 + \beta_0 \alpha_S(\mu_R^2) \ln(Q^2/\mu_R^2)}, \text{ with } \beta_0 = \frac{11N_C - 2N_f}{12\pi}. \tag{2.46}$$

N_C is the number of colours (3) and N_f the number of active quark flavours q with mass $m_q \ll \mu_R$ (5 at energy scales below the top mass threshold) present in the SM. Equation 2.46 shows that, as stated before, the coupling becomes large at sufficiently low Q^2. The point at which this happens can be denoted more explicitly with the definition of Λ_{QCD} via:

$$\Lambda_{QCD}^2 = \mu_R^2 \exp\left[-\frac{1}{\alpha_S(\mu_R^2)\beta_0}\right], \tag{2.47}$$

with this, α_S becomes:

$$\alpha_S(Q^2) = \frac{1}{\beta_0 \ln(Q^2/\Lambda_{QCD}^2)}. \tag{2.48}$$

At Q^2 values close to Λ_{QCD}^2 the coupling constant becomes large and perturbative QCD breaks down, Λ_{QCD}^2 is measured experimentally to have a value of around 300 MeV.

2.2. Physics at Hadron Colliders

2.2.1. The Structure of the Proton

At hadron colliders the initial colliding particles are not elementary but composite particles, a fact that has various consequences for the calculation of production cross sections.
A naïve description of the proton would be that it is a composite particle, consisting of three pointlike quarks (u, u, d). If this picture were correct, each of the three quarks would carry exactly 1/3 of the total proton momentum. This description is, however, heavily simplified. In reality the three quarks, termed *valence quarks*, are bound by a permanent gluon exchange between each other. Additionally, the gluons can couple to other gluons and split into virtual quark-antiquark pairs which recombine again. These additional quarks are called *sea quarks*. In general, gluons and quarks inside a hadron are commonly referred to as *partons*.
Due to the constant emission and absorption of gluons, the momentum fraction of the valence quarks of the total proton momentum becomes a smeared-out, broad distribution. Furthermore, since the sea quarks and gluons are also carrying some of the proton momentum, the average fraction of a valence quark of the total proton momentum is less than 1/3. Experimental data have shown that the momentum fraction of up quarks, down quarks and gluons are 36 %, 18 % and 46 %, respectively, if the contribution of heavier sea quarks is neglected [19].
As the hard scattering process is initiated by single partons and the kinematics of the event is determined by their momentum distribution, an understanding of such parton distribution functions (PDFs) is essential for the calculation of production cross sections at hadron colliders. However, because the interactions inside the proton happen at a low momentum transfer Q^2, corresponding to a large strong coupling α_S, the PDFs cannot be calculated perturbatively. Instead, they have to be determined by global fits to data in deep inelastic scattering (high energetic electron-proton collisions) processes. A PDF ($f_{a/A}(x_a, Q^2)$) describes the probability to find a parton a inside a proton A with a particular fraction of the total proton momentum x at

a momentum transfer Q^2. The PDF dependence on x has to be determined from measurements. The evolution with Q^2, however, can be computed with the so-called Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [35–37]. PDF sets are provided by several collaborations, e.g. MSTW [38] and CTEQ [39, 40] making use of different parametrizations. One set of PDFs, the CT10 set [41] can be seen in Figure 2.4.
A consequence of the composite nature of the proton is that not the entire proton energy is available in the hard scattering process, but rather only a fraction of the CM energy $\sqrt{s}$. This fraction ($\sqrt{\hat{s}}$) can be calculated as $\sqrt{\hat{s}} = \sqrt{x_1 x_2 \cdot s}$, with the momentum fractions x_1 and x_2 of the two incident partons involved in the hard scatter. The minimal x value (assuming symmetric x for the partons in both protons) for the production of a WW pair is then given as $x^2 = 4m_W^2/s$. The minimal x values for WW production for different CM energies are indicated in Figure 2.4.

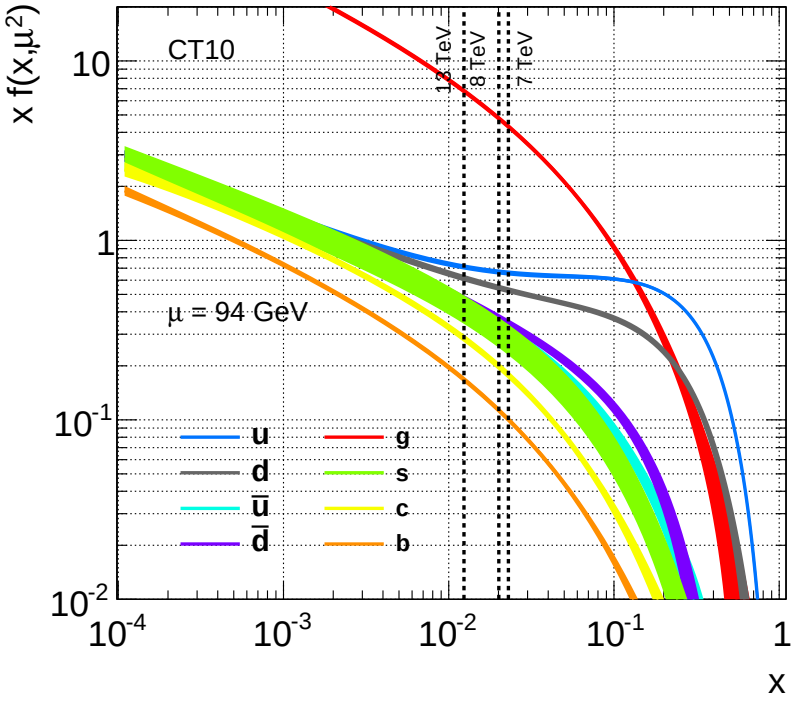


Figure 2.4.: CT10 parton distribution functions with associated systematic uncertainties. PDFs are evaluated at a momentum transfer of $Q = 94\,\text{GeV}$. Minimal x values for WW pair production for different relevant center-of-mass energies at the LHC are indicated as dotted vertical lines. At larger x values, the valence quarks are dominating, according to the simple proton model. At low x values, however, the difference in terms of probability to carry a certain momentum is strongly reduced between most quark flavours since they are all produced as sea quarks with comparable probability. By far dominating at small x are the gluons in the "sea", enhancing the probability of gluon-gluon induced processes.

2.2.2. QCD Factorization

Cross sections for hadronic initial states can be divided into the long-distance, non-perturbative part described above and a short-distance part of the actual hard scatter where the partons can be treated as free particles. With the latter being process-dependent while the former being universal. According to the QCD factorization theorem [42], these two contributions factorize,

such that a leading order cross section of the two protons A and B producing the final state X, $\sigma(AB \to X)$, can be computed by convoluting the PDFs with the partonic cross section:

$$\sigma(AB \to X) = \sum_{a,b} \int dx_a dx_b f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \times \hat{\sigma_0}(ab \to X), \tag{2.49}$$

where the sum is performed over all possible partons a and b, which can contribute to the final state and $\hat{\sigma_0}(ab \to X)$ is the partonic cross section, obtained from the matrix element and the Lorentz invariant phase space factor.

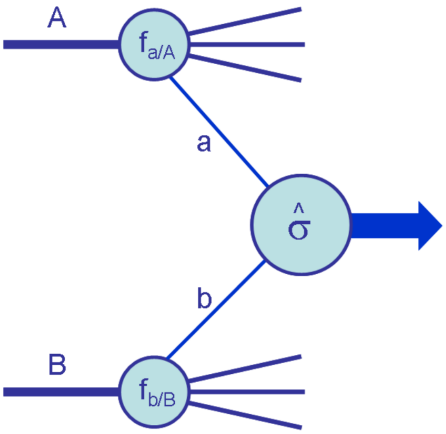


Figure 2.5.: Graphical illustration of a proton-proton hard scattering process [43]. Proton A and B collide and the partons a and b with the momentum fractions $f_{a/A}$ and $f_{b/B}$, respectively undergo a hard scattering process and produce a certain final state. The non-perturbative part of the proton structure can be separated from the hard partonic interaction and the two contributions factorize in the calculation of the total cross section.

The energy scale at which the transition between perturbative QCD and the non-perturbative long-distance effects occurs is denoted as *factorization scale* μ_F. Parton radiation below this scale is attributed to the PDF part of the cross section. Above that scale it is accounted for in the partonic part. As this scale is unphysical and only shifts contributions of one part of the calculation to another, the choice of its value is somewhat arbitrary. It has been shown that a reasonable choice is the scale of the momentum transfer Q in the considered process [44]. Customarily, μ_F and μ_R are chosen to be equal and close to typical momentum scales of the process, e.g. the W boson mass in W production.
Contributions of higher order diagrams to Equation 2.49 will show up as corrections to $\hat{\sigma_0}$, because of involved loop diagrams now also depending on the renormalization scale μ_R:

$$\sigma(AB \to X) = \sum_{a,b} \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \times [\hat{\sigma_0} + \alpha_S(\mu_R^2)\hat{\sigma_1} + ...]_{ab \to X}. \tag{2.50}$$

2.2.3. Coordinates at Hadron Colliders

As a result of the composite nature of the proton, the CM frame of the colliding partons in the hard scatter is not necessarily the CM frame of the colliding protons and therefore also does not coincide with the laboratory frame. Since angular separation and energies are not invariant under Lorentz boosts in the z-direction, a set of specialized coordinates is commonly used at hadron colliders. The most important of them being the *pseudorapidity* η, a coordinate which describes the polar angle θ of the particle relative to the beam axis and has the desired feature

of being invariant under longitudinal Lorentz boosts:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{2.51}$$

For a centrally produced particle directly on the xy-plane, η takes the value of zero and a for a forward particle, η goes to infinity. The relation between the pseudorapidity and the polar angle is visualized in Figure 2.6.

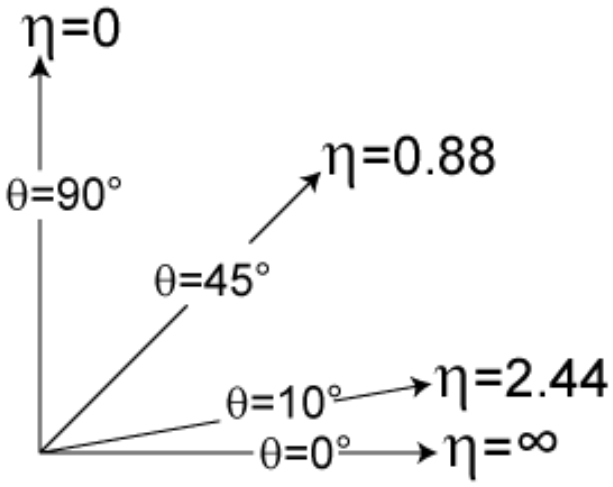


Figure 2.6.: Values of of the pseudorapidity η for different directions of the polar angle θ [45].

For massless particles, the pseudorapidity is the same as the *rapidity* y:

$$y = \frac{1}{2}\ln\frac{E + p_Z}{E - p_Z}, \tag{2.52}$$

which is also invariant under longitudinal boosts but more complicated to measure than the pseudorapidity. Particle production at hadron colliders is roughly constant as function of y.
The azimuthal angle is denoted as ϕ and is unchanged by longitudinal boosts. Angular distances between particles are usually measured in $\eta\phi$-angle space via the dimensionless quantity ΔR given by:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{2.53}$$

Another important quantity is the transverse momentum of a particle p_T, which is its momentum measured in the plane perpendicular to the z-axis (xy-plane):

$$p_T = \sqrt{p_x^2 + p_y^2}. \tag{2.54}$$

2.3. Gauge Bosons of the Weak Interaction

2.3.1. Triple Gauge Boson Couplings

The gauge boson sector of the SM can be investigated very well at hadron colliders. Of particular interest looking in the direction of new physics are the self interactions of vector bosons resulting

from the non-abelian structure of the $SU(2)_Y$. Self couplings of W and Z bosons and photons that can be probed in WV production exist in the form of WWX couplings (X = Z, γ), named triple gauge boson couplings (TGCs). Those couplings could possibly deviate from their predicted SM values since up to date they are not yet highly constrained by measurements [46].
If they actually do deviate, those couplings are referred to as anomalous triple gauge boson couplings (aTGCs). Particularly interesting for the boosted WV topology is the fact that that the deviations from the SM values become more enhanced with increasing center-of-mass energy. For this reason a measurement in the boosted topology should be especially sensitive to such anomalous couplings.
From the Feynman diagrams in Figure 2.7a and 2.9a, it becomes clear that the s-channel production of WW and WZ pairs is sensitive to aTGCs.
The TGCs are customarily described by an effective Lagrangian of which the most general form under the condition that C and P are conserved separately can be expressed as [47]:

$$\frac{\mathcal{L}_{eff}^{WWX}}{g_{WWX}} = -i\left[g_1^X X^\mu (W^-_{\mu\nu}W^{\mu+} - W^+_{\mu\nu}W^{\nu-}) + \kappa_X W^+_\mu W^-_\nu X^{\mu\nu} + \frac{\lambda_X}{M_W^2} X^{\mu\nu} W^{\rho+}_\nu W^-_{\rho\mu}\right], \quad (2.55)$$

with the newly introduced coupling parameters g_1^X, κ_X, λ_X, the photon or Z^0-field X^μ and the field tensors defined as $A_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$.
The overall coupling constants g_{WWX} are $g_{WW\gamma} = e$ and $g_{WWZ} = e \cdot \cot\theta_W$ with the positron charge e and the weak mixing angle θ_W.
In the SM Lagrangian, it is the bilinear term in the field strength tensors in Equation 2.10 which is responsible for trilinear and quadrilinear gauge boson couplings. This term can be rewritten using the physical fields (2.15) and omitting quadrilinear terms as [48]:

$$\begin{aligned} -\frac{1}{4}\mathbf{W}_{\mu\nu}\cdot\mathbf{W}^{\mu\nu} = &- ig[W^-_{\mu\nu}W^{\mu+} - W^+_{\mu\nu}W^{\mu-}](\cos\theta_W Z^\mu + \sin\theta_W A^\mu) \\ &- \frac{ig}{2}(\cos\theta_W Z_{\mu\nu} + \sin\theta_W A_{\mu\nu})[W^{\mu-}W^{\nu+} - W^{\mu+}W^{\nu-}]. \end{aligned} \quad (2.56)$$

Comparing the effective Lagrangian for TGCs (Equation 2.55) with the tree-level SM-Lagrangian (Equation 2.56), it can be deduced that $g_1^X = \kappa_X = 1$ and $\lambda_X = 0$. In the literature, deviations from the SM are commonly parametrized as

$$\Delta g_1^X = g_1^X - 1, \quad (2.57)$$
$$\Delta\kappa_X = \kappa_X - 1, \quad (2.58)$$
$$\lambda_X, \quad (2.59)$$

which are all vanishing in the SM. Imposing electromagnetic gauge invariance for photons fixes Δg_1^γ to $\Delta g_1^\gamma = 0$, so that only five couplings, namely $\Delta g_1^Z, \Delta\kappa_Z, \Delta\lambda_Z, \kappa_\gamma$ and λ_γ, are left. Furthermore, two simplifying assumptions motivated by $SU(2)_L \otimes U(1)_Y$ gauge invariance, the so-called "LEP constraints" [49], are often made:

$$\Delta\kappa_Z^Z = \Delta g_1^Z - \Delta\kappa_\gamma \tan^2\theta_W, \quad (2.60)$$
$$\lambda \equiv \lambda_\gamma = \lambda_Z. \quad (2.61)$$

Applying these constraints leaves three independent aTGC parameters, Δg_1^Z, $\Delta\kappa_\gamma$ and λ. If no physics beyond the SM (BSM) is present, these anomalous terms would cause a violation of unitarity bounds in interaction amplitudes. This violation can be prevented by optionally applying a form factor dependent on the invariant mass of the diboson system $\sqrt{\hat{s}}$ to the generic

anomalous couplings κ :

$$\kappa(\sqrt{\hat{s}}) = \frac{\kappa(0)}{(1 + \frac{\sqrt{\hat{s}}}{\Lambda_{co}^2})^2}, \tag{2.62}$$

with a cutoff energy scale Λ_{co} at which BSM physics should manifest. Upper bounds on the aTGCs can be derived as function of Λ_{co} [50]. The case where $\Lambda_{co} \to \infty$ corresponds to a no-form-factor case. In the literature limits are often quoted using this choice as benchmark as well, even though unitarity is not preserved in this case. The same approach will be followed in this analysis.
Experimentally, anomalous triple gauge couplings would manifest themselves as an increased WV cross section, especially as an increased yield of events at high transverse momentum of the V bosons and at high invariant mass of the WV system. By choosing the boosted channel one takes into account exactly these considerations since here, if present, the anomalies should be best visible.
Parallel to the just described approach, aTGCs can also be parametrized in the context of an effective field theory (EFT) which is preferred by a part of the theory community. Two theoretical complications are eliminated in this approach. It respects $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge invariance and does not introduce arbitrary form factors.
In the EFT approach, the SM is considered to be an effective field theory valid for present energies $(\sqrt{s})$ that is embedded in a higher theory valid up to an energy scale Λ significantly larger than $\sqrt{s}$ [51]. In the EFT, the SM is extended by the introduction of dimension-six operators in addition to the usual SM dimension-four operators in the SM Lagrangian. Those operators receive coupling factors $1/\Lambda$ for each mass dimension above four. The new operators can thus be taken to correspond to terms associated with new physics at a scale of Λ in an expansion in powers of $1/\Lambda$ around the nominal SM operators based on a low-energy approximation [52]. Three of these operators can contribute to triple gauge boson vertices, usually chosen as:

$$\mathcal{O}_{WWW} = \mathrm{Tr}[W_{\mu\nu} W^{\nu\rho} W^{\mu}_{\rho}], \tag{2.63}$$

$$\mathcal{O}_{W} = (D_\mu \Phi)^\dagger W^{\mu\nu} (D_\nu \Phi), \tag{2.64}$$

$$\mathcal{O}_{B} = (D_\mu \Phi)^\dagger B^{\mu\nu} (D_\nu \Phi), \tag{2.65}$$

with the Higgs doublet Φ and D_μ, $W_{\mu\nu}$ and $B^{\mu\nu}$ defined similarly as in Equations 2.9, 2.11 and 2.12, respectively. For more details, refer to Ref. [51]. The coefficients of these operators are commonly called c_{WWW}, c_W and c_B. A simple relation between the aTGC parameters from the LEP scenario and the EFT parameters can be established:

$$c_{WWW}/\Lambda^2 = \frac{2}{3g^2 m_W^2}\lambda\,, \tag{2.66}$$

$$c_W/\Lambda^2 = \frac{2}{m_Z^2}\Delta g_1^Z\,, \tag{2.67}$$

$$c_B/\Lambda^2 = \frac{2}{m_W^2}\Delta\kappa_\gamma - \frac{2}{m_Z^2}\Delta g_1^Z\,, \tag{2.68}$$

with the electroweak coupling constant g. In that sense, the EFT parameters and aTGC parameters from the LEP scenario can be considered equivalent. Those relations are only valid if no form factors are applied to the aTGC parameters which is the case in this analysis.

2.3.2. Standard Model like WW/WZ-production

WW production is the most abundant diboson process in the standard model. The only leading order (LO) production mechanism is that of quark-antiquark annihilation or scattering, occurring via the s-, t- and u channel, depicted in Figure 2.7.

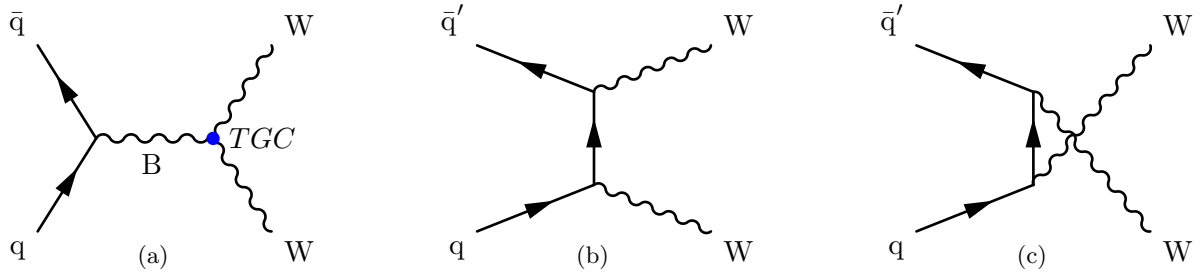


Figure 2.7.: The three tree-level diagrams for WW production, B can either be a Z^0 or a γ: (a) s-channel, (b) t-channel, (c) u-channel.

In the s-channel diagram, a photon or Z boson (denoted as B here) couples directly to the two outgoing W bosons making the channel sensitive to TGCs. In principle the WW production could also be mediated by a Higgs boson instead of a B but the small quark masses and the resulting small coupling in the Higgs production vertex render the amplitude of this diagram negligible.
The inclusive cross section for the quark-antiquark annihilation/scattering production mechanism at $\sqrt{s} = 8$ TeV was calculated in Ref. [11] up to NNLO:

$$\sigma^{\mathrm{LO}}_{\mathrm{q\bar{q}}\to\mathrm{WW}} = 35.50^{+2.4\,\%}_{-3.5\,\%}\,\mathrm{pb}, \tag{2.69}$$

$$\sigma^{\mathrm{NLO}}_{\mathrm{q\bar{q}}\to\mathrm{WW}} = 54.77^{+3.7\,\%}_{-2.9\,\%}\,\mathrm{pb}, \tag{2.70}$$

$$\sigma^{\mathrm{NNLO}}_{\mathrm{q\bar{q}}\to\mathrm{WW}} = 59.84^{+2.2\,\%}_{-1.9\,\%}\,\mathrm{pb}. \tag{2.71}$$

The MC simulation used for this thesis for WW and WZ production was performed with MC@NLO [53] at next-to-leading order. More information on the Monte Carlo WW and WZ production is given in Section 6.2.1.
Besides the quark-antiquark annihilation, the other important WW production mechanism is that of gluon fusion that can only occur via loop diagrams, such as those portrayed in Figure 2.8.

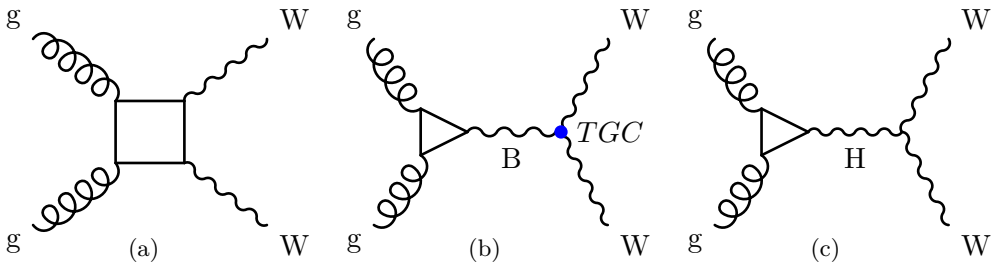


Figure 2.8.: The three simplest gluon fusion loop diagrams for WW production.

The triangle diagram in Figure 2.8b is again sensitive to TGCs. The same diagram with a Higgs boson instead of a B in Figure 2.8c can here have a much larger amplitude than the diagram in Figure 2.7a because virtual heavy quarks with a large Higgs coupling can now circulate in the triangle.

The inclusive cross section for this process with on-shell Higgs bosons (resonant production) was evaluated to

$$\sigma_{gg\to H\to WW} = 4.14^{+7.2\,\%}_{-7.8\,\%}\,\text{pb}\ [54], \tag{2.72}$$

corresponding to about 8 % of the NLO $q\bar{q} \to WW$ cross section. However, an on-shell Higgs boson (with an assumed mass of $m_H = 125\,\text{GeV}$) signifies that one of the W bosons must be off-shell. The event selection in this analysis targets on two reconstructed on-shell W bosons which strongly suppresses contributions from resonant Higgs production to less than 1 % of $\sigma^{NLO}_{q\bar{q}\to WW}$. Due to this small effect compared to the precision of the measurement this channel is neglected in this analysis.
In Ref. [55], the cross section for the non-resonant gg→WW production has been computed at NLO for gg→WW→ $e^+\nu_e\mu^-\bar{\nu_\mu}$. Dividing the resulting cross section by BR(WW→ $e^+\nu_e\mu^-\bar{\nu_\mu}$), $\sigma^{NLO}_{gg\to WW}$ results in

$$\sigma^{NLO}_{gg\to WW} = 2.76^{+0.07\,\%}_{-0.10\,\%}\,\text{pb}, \tag{2.73}$$

corresponding to 5.2 % of $\sigma^{NLO}_{q\bar{q}\to WW}$. The contribution from these processes is taken into account in the calculation of the fiducial cross section but no dedicated MC samples are produced.
Very similar to the WW production mechanism are those of the WZ and ZZ production. For this analysis the ZZ production is of practically no relevance (see Table 10.2) because of the veto on a second lepton in the event selection (see Section 6.3).
Feynman diagrams for WZ production are shown in Figure 2.9, here, due to charge conservation no gluon fusion processes exist. Since the Z boson is only identified by its decay products which are in principle the same particles that can emerge from a photon decay and the photon also couples to the W boson, the actual process which is measured in this case is not pure WZ production but the interference of WZ and $W\gamma^*$. However, the nominal MC samples for WZ production only include on-shell WZ production and no $W\gamma^*$, for more details, refer to Section 6.2.1. For this reason and because charge conservation also restricts the mediating particle in Figure 2.9a to be a W, the MC prediction for aTGCs in WZ production is in this analysis only sensitive to the WWZ and not to the WWγ vertex.
The theory cross sections for WZ production have been evaluated in Ref. [56] and yield:

$$\sigma^{LO}_{q\bar{q}\to WZ} = 13.654^{+1.3\,\%}_{-2.1\,\%}\,\text{pb}, \tag{2.74}$$

$$\sigma^{NLO}_{q\bar{q}\to WZ} = 22.750^{+5.1\,\%}_{-3.9\,\%}\,\text{pb}, \tag{2.75}$$

$$\sigma^{NNLO}_{q\bar{q}\to WZ} = 24.690^{+1.8\,\%}_{-1.9\,\%}\,\text{pb}. \tag{2.76}$$

The WZ production cross sections correspond to approximately 40 % of the WW production cross section.

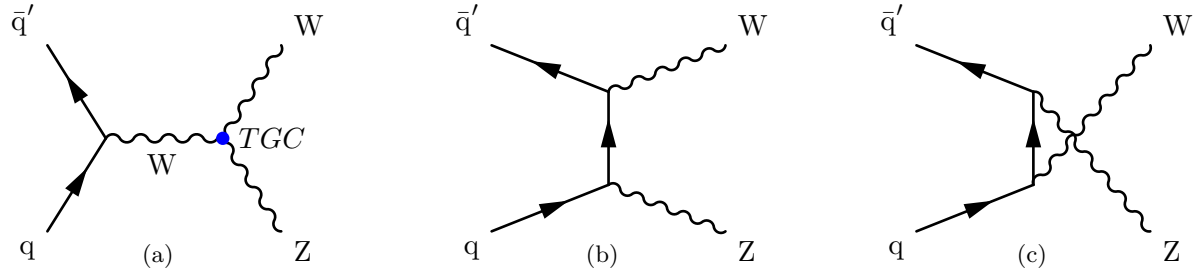


Figure 2.9.: All tree-level WZ production diagrams: (a) s-channel, (b) t-channel, (c) u-channel.

Chapter 3

The ATLAS Experiment at the LHC

3.1. The Large Hadron Collider

The Large Hadron Collider is presently the most powerful particle accelerator on earth. It is located on the French-Swiss border close to Geneva and is run by the European Laboratory for Particle Physics (CERN). The LHC was built in order to investigate the fundamental structure of matter and to search for new particles or phenomena at unprecedentedly high energy scales. The search for the Higgs boson, which had already been searched for since the 1960s, was in particular one of the major reasons to build this giant project which was approved in the 1990s, completed in 2008 and started taking data in March 2010. The LHC is housed in the tunnel that was previously used for an e^+e^--collider, the Large Electron-Positron Collider (LEP). It is located 100 m below the surface and its circumference amounts to 26.7 km. The LHC was designed to collide either protons or heavy-ions which are circulating in opposite directions in two different beam pipes at high energy and frequency.
The particles are made to collide at four of the eight interaction points called IP1 - IP8, at which the four main LHC experiments are located. ATLAS, which used to be an acronym for "**A** **T**oroidal **L**HC **A**pparatu**S**" but is nowadays used as the proper name for the experiment is located in IP1. In IP2 there is ALICE (A Large Ion Collider Experiment), then there is CMS (Compact Muon Solenoid) which can be seen as the complementarty partner to ATLAS at IP5, and LHCb which stands for "LHC beauty" in IP8. LHCb was designed for physics measurements concerning the b-quark sector, especially focusing on CP violation. ALICE is intended to first directly prove the existence of and to further investigate quark-gluon plasma in heavy-ion collisions.
The design center-of-mass energy for pp collisions is 14 TeV (7 TeV per beam) but due to machine safety concerns, the LHC ran at a center of mass energy of 7 TeV until the end of 2011 and at 8 TeV during the 2012 run. Since 2015, the data from proton-proton collisions at $\sqrt{s} = 13$ TeV are taken. The protons are not accelerated in the LHC alone but fly through a number of consecutive pre-accelerators which were mostly part of older CERN experiments and which were altered to serve LHC purposes. A schematic view of the LHC and the acceleration chain is shown in Figure 3.1.

The number of pp collisions per time unit can be calculated via:

$$N_{Coll} = \mathcal{L} \times \sigma_{pp}, \tag{3.1}$$

with σ_{pp} being the proton-proton inclusive cross section and $\mathcal{L}$ the instantaneous luminosity which can be expressed by LHC run parameters as follows:

$$\mathcal{L} = \frac{N_b^2 \cdot n_b \cdot f_{rev}}{A}, \tag{3.2}$$

where N_b is the number of particles per bunch and n_b the number of bunches circulating with the revolution frequency f_{rev}. A is the cross sectional area of the two intersecting beams.
By replacing σ_{pp} with the cross section of the process of interest, Equation (3.1) can be used to compute the rate of any process.
The LHC was designed to operate at an instantaneous luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, while in 2012 the peak instantaneous luminosity was $\mathcal{L}_{Peak} = 7.73 \cdot 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, delivering an integrated luminosity of $\int \mathcal{L}dt = 23.3\,\mathrm{fb}^{-1}$ [57]. This high luminosity is obtained by letting the proton bunches collide at a rate of one collision per 50 ns (25 ns since 2015). In addition, in every bunch crossing, there does not occur only one proton-proton collision but on average there were 20.7 collisions for 2012 run conditions [57].

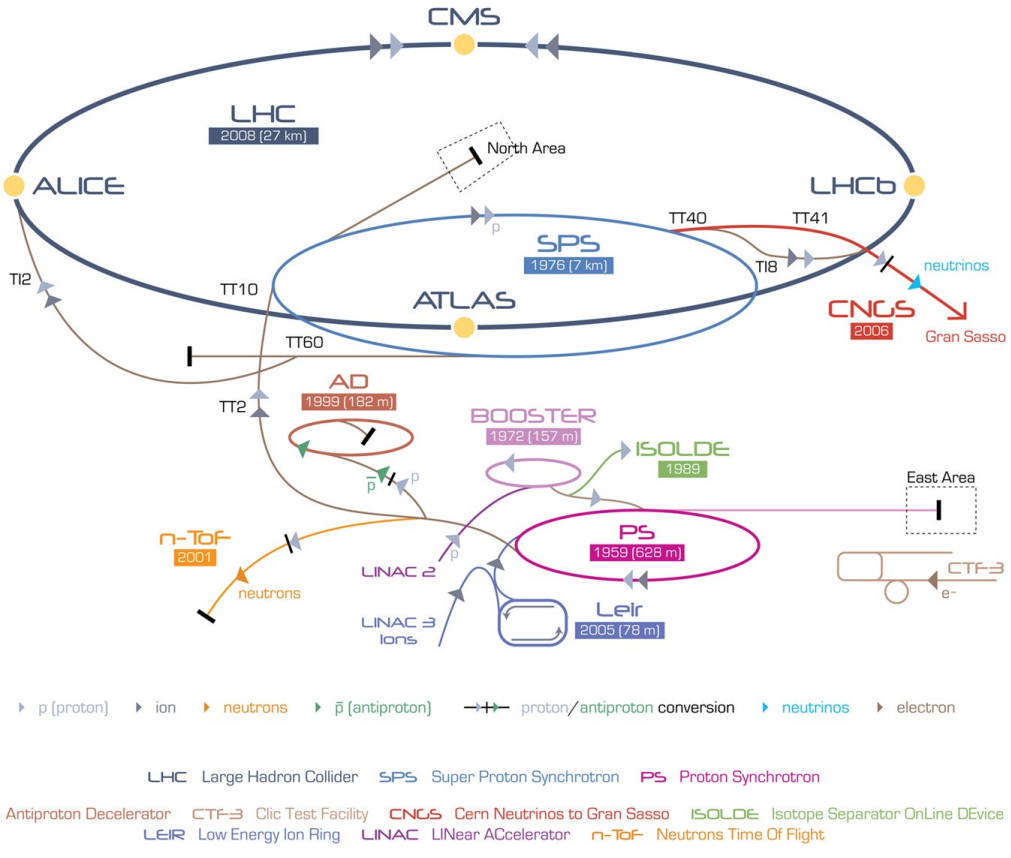


Figure 3.1.: Schematic view of the LHC, its pre-accelerators and experiments [58]. Until they reach their final energy of 4 TeV (in 2012 run conditions), the protons are successively boosted by different accelerators before being injected into the next sequential module. The protons, obtained by ionizing hydrogen atoms, start at the linear accelerator LINAC2 after which they are injected into the PS booster followed by the proton synchrotron (PS). After the PS, the protons are transferred to the super proton synchrotron before finally reaching the LHC ring.

Assuming a bunch spacing of 50 ns, there are 1782 slots throughout the LHC ring which can potentially be filled with proton bunches. These slots are, however, not filled uniformly but the proton bunches are organized in so-called *bunch trains*. These bunch trains consist of 72 or 144 filled bunch slots which are interrupted by smaller and larger gaps of around 10 and 30 empty bunch slots [59]. This is necessary to allow for the rise times of SPS and LHC injection and extraction magnets and has several implications for the so-called pile-up (see Section 5.4).

3.2. The ATLAS Detector

The ATLAS detector is located directly below the CERN main area. More than 3000 scientists are involved in operating the detector, acquiring and analyzing the obtained data. It is one of two multipurpose detectors (the other counterpart being CMS) that were designed in a way that they be capable of measuring a large number of different processes, the most important one at the time of detector design being the Higgs boson production. As it is common in high energy physics, two detectors were built offering complementarity with regard to systematic uncertainties.
In the ATLAS collaboration, the coordinate system was chosen such that the nominal interaction point is the origin of coordinate system and the direction of the beam was chosen to be the z-axis. Naturally, perpendicular to this direction is the xy-plane, with the positive x-axis pointing towards the center of the LHC ring and the positive y-axis pointing upwards.
The ATLAS detector consists of several sub-detectors, each one designed to serve a particular purpose which will be described in the following as they were being used during the 2012 run, starting from the beam pipe going outwards.

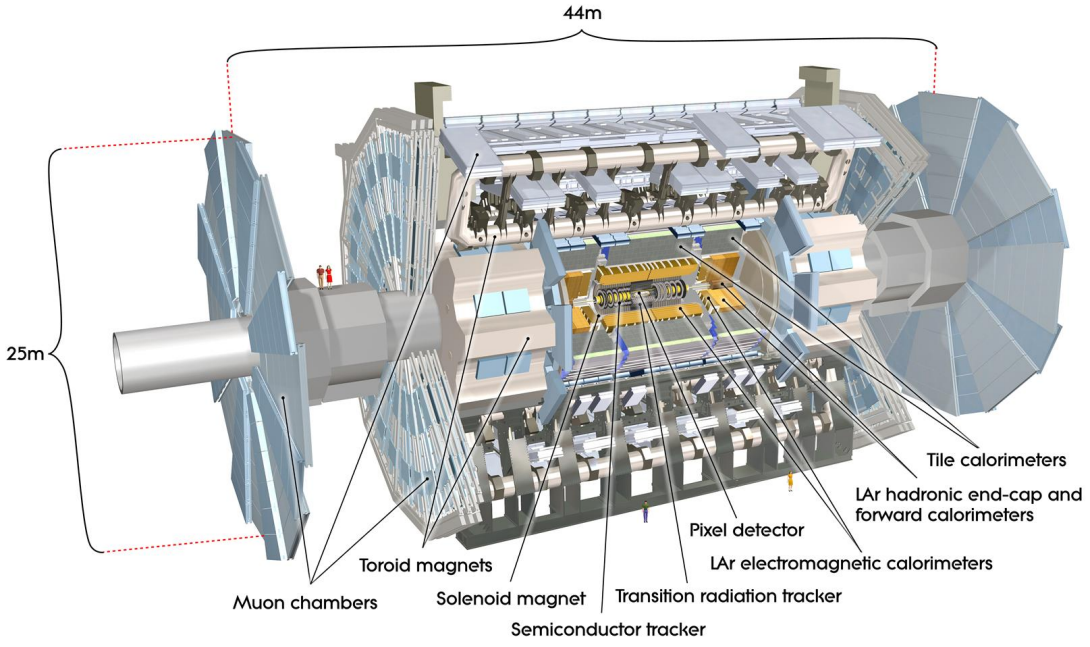
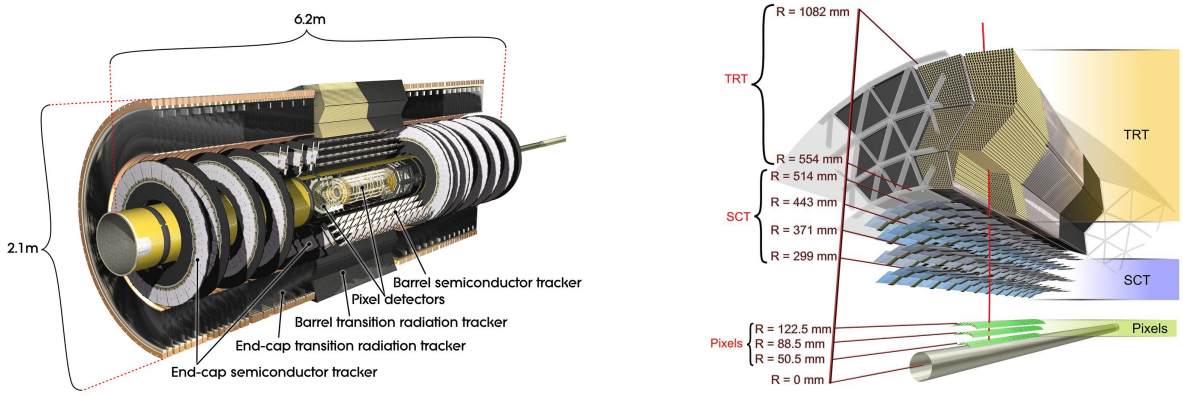


Figure 3.2.: View of the ATLAS detector with all subsystems [60].



(a) (b)

Figure 3.3.: (a) View of the inner detector with its three subsystems.
(b) Zoomed-in view with measures [61].

3.2.1. The Inner Detector

Due to the high luminosity, the ATLAS detector has to deal with an extremely high particle density, especially in the inner region.
In order to still reconstruct particle momenta and vertices accurately, the inner detector has to have a sufficiently fine granularity. For optimal vertex reconstruction, the tracking has to start as close to the interaction point as possible. Therefore the inner detector begins directly on the outside of the beam-pipe which has a diameter of 5 cm and extends up to a radius of 115 cm away from the interaction point. It consists of three subsystems all of which are optimized for different requirements, depending on the distance from the beam axis. The whole inner detector is immersed in a magnetic field parallel to the z-axis which is generated by the inner solenoid (see Section 3.2.3). This magnetic field causes the charged particles emerging from the interaction point to describe a circular trajectory in ϕ. The curvature of this trajectory can be used to determine the particle's charge and momentum with high accuracy. In total, the η-region covered by the inner detector corresponds to $|\eta| < 2.5$. A schematic view of the inner detector is given in Figure 3.3.

Pixel Detector

The innermost system is the pixel detector, followed by a silicon semiconductor tracker (SCT) and a transition radiation tracker (TRT). The pixel detector consists of three concentric tubes at radii of 5, 9 and 12 cm, respectively in the barrel region and three discs in the forward (and backward) direction which serve as end-cap tracker.
Altogether the detector is comprised of 1744 modules of the size of 62.4 mm times 21.4 mm, each one having 46080 pixels to be read out. This makes up for ∼80 million rectangular pixels with the size of 50 times 400 μm with a spatial resolution of 50 μm. Precision reconstruction of primary and also secondary vertices is made possible by this huge amount of channels close to the interaction point.

SCT

Going further outward the next subdetector is the SCT, it is also a precision tracker, delivering eight measurements per track which can be used for momentum, vertex position and impact parameter measurement. It makes use of double silicon strip layers that are arranged at a stereo angle of 40 mrad. In the barrel region there are four of those double layers arranged as concentric cylinders, from the innermost at a radius $R = 30\,\text{cm}$ to the outermost at $R = 51\,\text{cm}$. The end-caps consist of two times nine discs perpendicular to the beam axis in a range of $z = 85\,\text{cm}$ to $z = 273\,\text{cm}$ which allow an η coverage up to $|\eta| < 2.5$. Altogether there are around 6.3 million readout channels in the SCT and it has a nominal resolution of 17 μm in the $R - \phi$-plane and 580 μm in the z-direction.

TRT

The outermost component of the inner detectors is the transition radiation tracker. It works as a combination of a straw tube detector and a transition radiation detector and provides on average 36 measurements per particle in addition to the 3+8 measurements from the pixel detector and the SCT. It consists of 73 layers of straw tubes in the barrel and 160 straw plates in the end-cap region, each tube with a diameter of 4 mm and a gold plated tungsten-rhenium wire in its center. The tubes in the barrel are divided at $\eta = 0$ with a readout at both ends to reduce occupancy. In the barrel region, there are 2× 50 000 tubes and in the end-cap region an additional 32 000 tubes so the number of readout channels amounts to a total sum of 42 0000.
Due to its architecture, the barrel TRT can only provide useful information in the R - ϕ plane and not along the z-axis. The spatial resolution of the TRT is $\sim 130\,\mu$m per straw tube.
Straw tubes are filled with a gaseous mixture of xenon, oxygen and carbon dioxide and the area between the tubes is filled with polypropylene fibers that serve as transition radiation material. Particles traversing a tube ionize the gas, the ionized atoms cause an electric current which is read out as a signal. Additionally, they emit transition radiation when entering the area between the tubes where the dielectric material constant changes.
This effect can be used for discrimination of electrons and different hadronic particles. Lighter relativistic particles like electrons with a certain energy have a higher γ-factor than heavier particles with the same energy. That means an electron emits more transition radiation than for example a pion with the same energy. By applying a threshold value on the transition radiation which is surpassed by an electron but not by a pion, one can accurately distinguish the two particle types.

3.2.2. Calorimeters

Aside from the position and direction of the particles, it is also crucial for physics analyses to determine the energy of decay products as accurately as possible. In ATLAS this is done with several calorimeters which can be roughly divided into two subsystems, namely the liquid argon (LAr) calorimeters and the tile calorimeters. The calorimeter systems cover the full range up to $|\eta| < 4.9$. Different technologies have to be used depending on the η-region according to different conditions in pile-up (see Section 5.4) and requirements of the physics processes of interest.
Furthermore, the calorimeters can be classified as hadronic or electromagnetic calorimeters, which are each optimized to measure corresponding particle types. The tile calorimeter is a completely

hadronic calorimeter whereas the liquid argon calorimeters cover both types of interaction. In order to measure as much of the energy of a certain particle as possible and to avoid punch through of particles to the muon systems on the outside of the calorimeters, the calorimeters have to have a certain thickness. The thicker a calorimeter extends, the higher is the probability that a particle's energy is completely contained inside.
Two important quantities in this context are the radiation length X_0 and the hadronic interaction length λ which are both characteristic material constants.
The radiation length is a measure of the distance travelled in a certain material after which a relativistic electron has lost all but $1/e$ of its initial energy due to bremsstrahlung.
The interaction length is a similar quantity for hadronic particles, it is defined at the longitudinal extent in a certain material after which only $1/e$ of the original incoming particles are present.

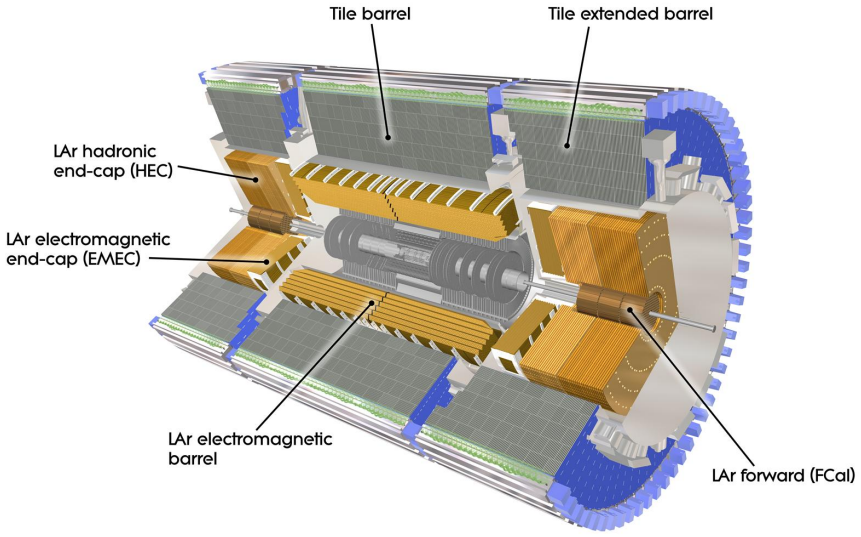


Figure 3.4.: Combined view of LAr- and Tile calorimeters. The crack region between the EMB and the HEC and between the Tile barrel and extended barrel is also visible [62].

Electromagnetic Calorimeters

The electromagnetic calorimeters consist of four parts, two electromagnetic barrels (EMB) and two electromagnetic end-caps (EMEC) (one of each on either side of the detector). A schematic view of the ATLAS calorimeter system is given in Figure 3.4.
The EMB is a sampling calorimeter with an accordion-shaped lead absorber structure and liquid argon (LAr) as the active material. The accordion shape, illustrated in Figure 3.5, makes it possible to construct the calorimeter without azimuthal cracks and therefore with a complete coverage in ϕ. The barrel part extends to an η-range of $|\eta| < 1.475$ with a tiny gap of 4 mm at $\eta = 0$, dividing the barrel part into two smaller parts. The two end-caps cover a region of $1.375 < |\eta| < 3.2$ and are housed in an own cryostat each.
Electron identification is particularly difficult in the transition region between barrel and end-cap usually referred to as crack region. This is the region where the two cryostats for the EMB and the hadronic end-cap calorimeter are positioned. It extends over an η range of $1.37 < |\eta| < 1.52$. Energy measurements in this range are therefore rather limited. This is the reason why in most

analyses involving the identification of electrons, events where the reconstructed electrons fall into this crack region are excluded.
In the rest of the calorimeter, a good containment of electromagnetic showers is guaranteed by a minimum depth of $22\,X_0$ in the barrel and $24\,X_0$ in the end-cap. The EMB is segmented into three layers in R, numbered from one to three going from the beam axis outwards.
A useful feature of the EMB is its fine granularity in its innermost layer, the so-called strip layer ($\Delta\eta \times \Delta\phi = 0.0031 \times 0.1$), which makes it possible to distinguish the two photons from a π^0 decay up to a π^0 transverse momentum of $p_T \approx 45\,\text{GeV}$. The second and third layer have a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$ and 0.05×0.0245, respectively.
The electromagnetic calorimeter in the barrel part is preceded by a presampler layer located inside the liquid argon, directly behind the cryostat wall. Its task is to deliver a rough measurement of a given particle's energy before it enters the calorimeter in order to be able to correct for any potential energy deposited upstream.

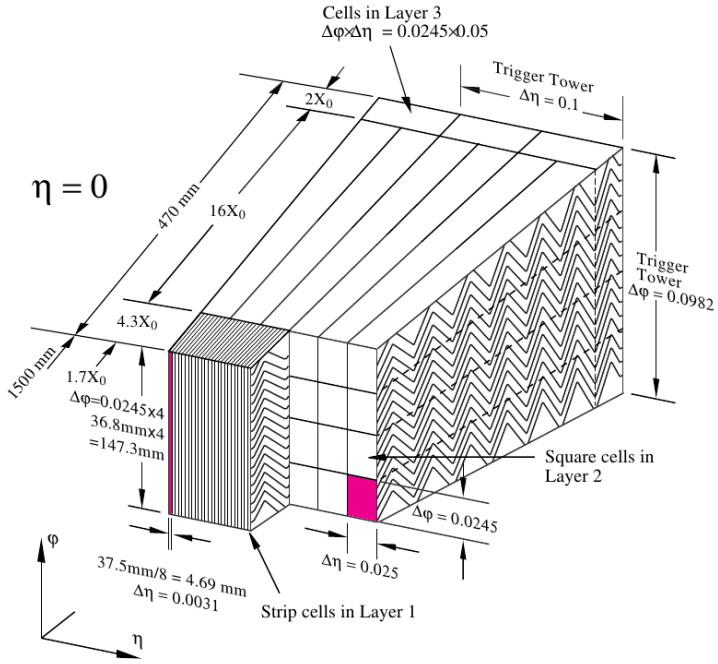


Figure 3.5.: View of a barrel module of the LAr calorimeter showing cell segmentation and accordion structure [1].

Hadronic Calorimeters

The **LAr hadronic calorimeter** is called hadronic end-cap calorimeter (HEC) and covers an η-range of $1.5 \leq |\eta| \leq 3.2$. It partly overlaps with the EMEC and the Tile calorimeter. In contrast to the electromagnetic LAr calorimeters its passive material is not lead but copper. It is located directly behind the EMEC, inside the same cryostat and consists of two independent wheels, each of which consists again of two layers. Each individual wheel consists of 32 wedges, each of which having an opening angle of 11°. The radii of the inner and outer border of the wheels are at R=0.475 m and R =2.03 m. The wheel closest to the interaction point is made from 25 mm copper plates and the other wheel from 50 mm copper plates that are separated by 8.5 mm wide gaps filled with liquid argon. Its total thickness corresponds to ten hadronic interaction lengths. The **Tile calorimeter** consists of two parts on either side, the barrel part which is

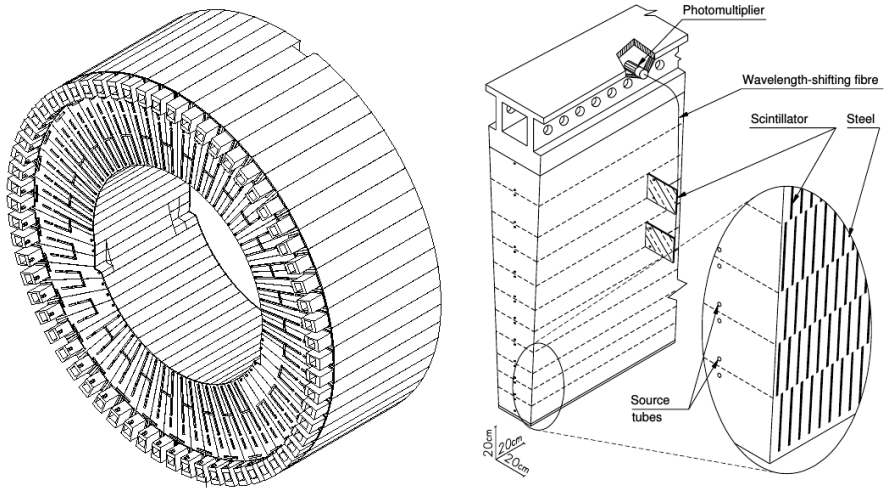


Figure 3.6.: Detailed view of a Tile barrel module and a zoom in into one wedge [63].

called long barrel (LB) and the extended barrel (EB). The LB part encloses the EMB while the EB encloses the EMEC. Its functioning as a hadronic calorimeter in the barrel region as well as for the transition region between the barrel and the end-cap. Radially it extends from R =2.28 m to R =4.25 m. Going from the center outwards the barrel portion of the Tile calorimeter consists of three layers with a thickness of 1.5, 4.1 and 1.8 λ, respectively. In the extended barrel the corresponding layer thicknesses are 1.5, 2.6 and 3.3 λ.
The absorber material of the Tile calorimeter is steel, interleaved with tile scintillator plates that provide the active material. The signal from the scintillator tiles is read out from each end by wavelength-shifting fibers and sent to two separate photomultiplier tubes, as can be seen in Figure 3.6. The granularity of the Tile calorimeter is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$.

Forward Calorimeters

Each of the two forward calorimeters (FCal), one on either side of the detector, consists of one electromagnetic module and two hadronic modules. It covers an η-range from $3.2 \leq |\eta| < 4.9$ and its depth is around 10 interaction lengths with a total number of 3524 readout channels [64]. The FCal operates in a challenging environment since all the proton fragments apart from the hard scatter are flying in the forward direction, meaning that the FCal is exposed to a high radiation level. However, calorimetry is also important at these very forward regions for the investigation of processes with forward jets, e.g. events induced by vector boson fusion processes. In general calorimetric hermiticity is necessary for a precise determination of missing transverse momentum. The FCal is housed inside the same cryostat as the HEC and EMEC, as can be seen in Figure 3.7.
Because of the limited space and in order to avoid punch through to the forward muon system, the FCal is built in a high density design. It consists of three wheels on each side, the inner one made of copper and being optimized for electromagnetic calorimetry and the outer ones using tungsten for hadronic calorimetry. Each wheel is built of a hexagonal matrix of metal rods, which are positioned parallel to the beam axis and are surrounded by a tiny LAr - filled gap (250, 375, 500 μm for FCal 1, 2, 3) which works as active medium. In order to guarantee equidistant spacing between the rods and the matrix over length of the rods and operating time, the rods are surrounded by a helically wound radiation hard plastic fiber.

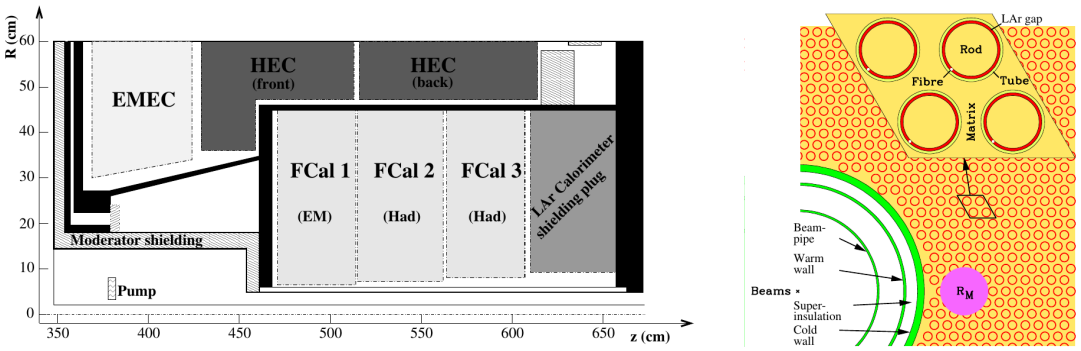


Figure 3.7.: Schematic side- and head-on view of the Forward Calorimeter with zoom in into the absorber matrix [65].

3.2.3. Muon Spectrometer and Magnet System

Muons, in stark contrast to electrons, can traverse the entire inner detector and the calorimeters without significant energy loss. Due to their relatively high mass compared to electrons, energy losses via bremsstrahlung are much smaller and since leptons do not participate in the strong interaction, they also do not initiate hadronic showers.
For this reason the Muon Spectrometer (MS) is situated on the outer perimeter of the ATLAS detector. A schematic overview can be found in Figure 3.8. It measures the direction and momentum of muons with high precision and is also used for triggering on Muons.
The MS consists of four subsystems:
The first two are the **resistive plate chambers** (RPC) and **thin gap chambers** (TGC) which are mainly used for fast triggering. Secondly there are the **monitored drift tubes** (MDT) and **cathode strip chambers** (CSC). These are used to measure precisely the momenta and trajectories of muons which are deflected in the magnetic field provided by the ATLAS magnet system.

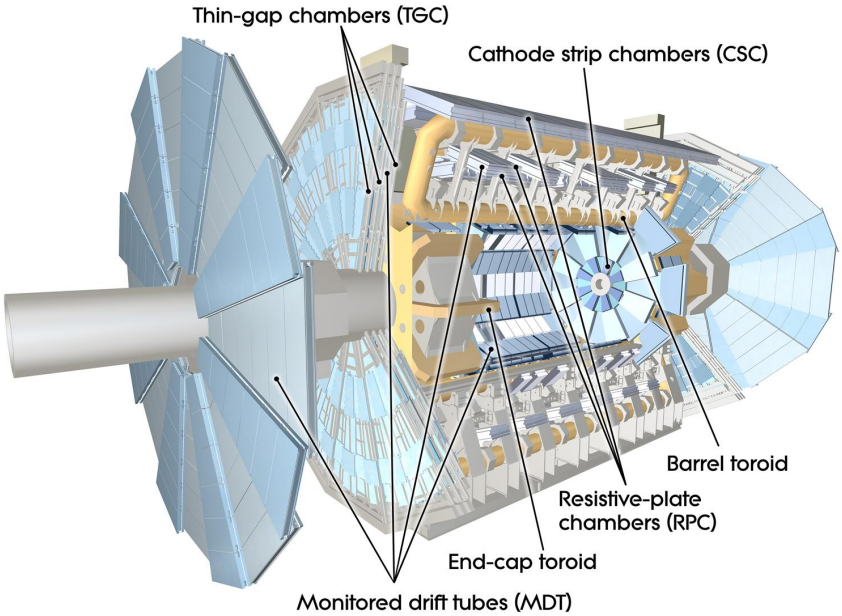


Figure 3.8.: Overview of the muon spectrometer [66].

The magnet system is made up from three toroidal and one solenoidal subsystems, as can be seen in Figure 3.9.
The innermost magnet system is the central solenoid, providing the magnetic field for the inner detector. It is located outside of the inner detector but in front of the LAr calorimeter sharing the same cryostat with the latter. At the nominal current of 7.7 kA it generates a magnetic field of more than 2 T.
The barrel toroid consists of eight super-conducting toroidal magnetic coils with a nominal current of 20.5 kA and with a maximum magnetic field of 4 T up to $|\eta| < 1.4$. For the two end caps, the magnetic field is provided by two smaller end-cap toroids, allowing for a uniform magnetic field from $1.6 < |\eta| < 2.7$.
In the transition region of $1.4 < |\eta| < 1.6$ the field is provided by the combination of the two but here it is not as uniform as in the two dedicated regions. The directions of the fields are chosen such that they be as orthogonal as possible to the direction of the muons to allow for maximal bending and in order to provide a better momentum resolution. The total energy stored in the four subsystems amounts to 1600 MJ [1].

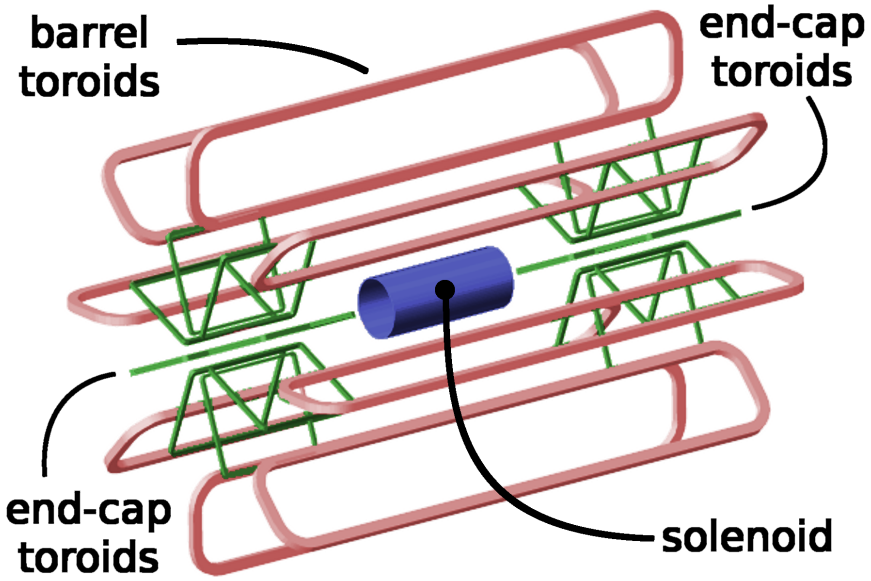


Figure 3.9.: Schematic view of the ATLAS magnet system [67].

The Trackers MDT and CSC

The MDT sub-detector consists of altogether 354 000 aluminum-manganese alloy tubes with a diameter of 29.97 mm packed into 1150 chambers. The tubes are stored parallel to the z-axis in the barrel region and perpendicular to the beam axis in the end-cap region. Inside each tube there is a central 50 μm thin gold-plated tungsten-rhenium wire with a voltage of 3080 V applied between the wire and the tube itself. The tubes are filled with a mixture of argon and carbon dioxide which has been chosen because of its good ageing properties and the fact that it does not adhere to the wires.
When a muon penetrates the tube, the wire acts as anode and the tube as cathode, causing a cascade of electrons drifting to the inner wire and thus generating a signal for the read out system.

The $|\eta|$ coverage of the MDTs extend up to $|\eta| < 2.7$ with the restriction of the innermost layer which only covers the region $|\eta| < 2$. The average spatial resolution per tube is 80 μm.
In the more forward region, the flux of charged particles is too high for the safe operating limit of the MDTs (150 Hz/cm) which also has to struggle with high occupancy and buildup of positive ions at high particle fluxes. Therefore the first layer of MDTs at higher pseudorapidities ($2 < |\eta| < 2.7$) is replaced by multi-wire proportional chambers with segmented cathodes called "Cathode Strip Chambers" (CSC). The safe-operation limit of these modules of 1000 Hz/cm is almost a factor seven higher than for the MDTs.
The CSCs are multiwire proportional chambers, filled with a gas mixture of argon and carbon dioxide. A CSC end cap is formed by a disc which are further divided in eight small and eight large chambers. In every chamber are four CSC planes contained which deliver four independent ϕ and η measurements per track with a spatial resolution of 60 μm in the bending η-plane and 5 mm in the non bending ϕ-plane. The total number of channels amounts to 30 720 [1].

Resistive Plate Chambers

The Muon trigger system in the barrel is based on three concentric layers of RPCs with a coverage of up to $|\eta| < 1.05$. As suggested by its name, no wires are used in the RPC but charged parallel plates at a separation of 2 mm from each other with a gas filled gap in between. The electric field in the gap amounts to 4.9 kV/mm causing avalanches of electrons to occur once a muon crosses the gap. Each RPC consists of two partly overlapping units where each of the units is made of two detection layers. They are either used to trigger on high p_T muons (9-35 GeV) by using the lever arm between the inner and the outer layer or on low p_T muons (6-9 GeV) by using the inner and the middle layer.
It is possible to identify fake muons from noise hits by using (anti-)coincidence between the different layers. The time resolution with $\sim$2 ns being very small enables the RPCs to reject cosmic muons based on timing measurements between the several layers.

Thin Gap Chambers

Complementary to the RPCs in the barrel, the Thin Gap Chambers are used in the end caps to provide fast trigger decisions. These are also multi-wire proportional chambers of 50 μm thick wires with a spacing of 1.8 mm between two parallel cathodes in a distance of 2.8 mm from each other [1]. The wires operate in a gas mixture of carbon dioxide and n-pentane. The nominal spatial resolution of the Thin Gap Chambers is 2-6 mm in R and 3-7 mm in ϕ.

3.2.4. Trigger System

In 2012 the frequency of bunch crossings was at 20 MHz with an average number of interactions per bunch crossing of around 21, at design luminosity it will even be at 40 MHz.
Every recorded event requires storage space of about 2 MB summing up to a data rate of $\sim$ 80 TB/s which is why it is technically impossible to record all events. The frequency of recording events which is feasible is around 100 Hz only, meaning the initial event rate has to be reduced by a factor of $4 \cdot 10^5$ [68].
In order to obtain this enormous reduction while keeping a high efficiency for the rare, interesting

physics processes, a three level hard and software trigger system has been developed which is schematically shown in Figure 3.10 and will be briefly described in the following.

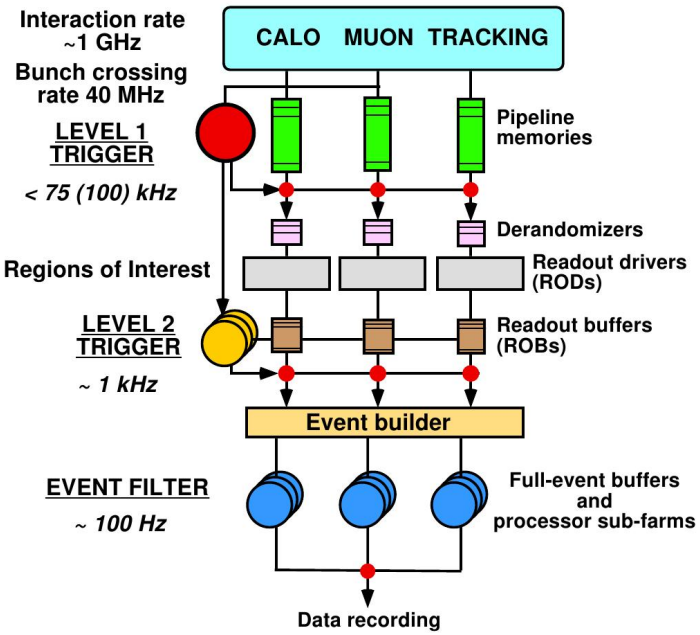


Figure 3.10.: Overview of the three-step trigger architecture [68].

Level-1 Trigger

The first level trigger (L1) is a hardware-based trigger. It collects information from the muon trigger chambers (TGC and RPC) related to high p_T muons and from the calorimeters associated with different quantities and objects such as (missing) transverse momentum, high p_T photons, electrons, jets or τ-leptons decaying into a neutrino and hadrons. The trigger decision also depends on combinations of different physics objects which can be either required or forbidden to coincide. The L1 decision has to be made in $\sim 2.5\,\mu$s and reduces the event rate to 75 kHz meaning a rejection factor of ~530 (99.81 %). In every event, the L1 trigger defines regions of interest (RoIs) which correspond to the η and ϕ regime of potentially interesting physics objects. The RoI and further information about them, e.g. the type of the object and certain criteria passed, usually some threshold are then passed to the level-2 trigger. The L1 trigger needs around $2\,\mu$ s to reach its decision. This time includes propagation delays on cables between the detector and the underground room where the trigger logic is located. In order to minimize data loss caused by trigger dead time, the entire detector information is buffered in so-called pipeline memories until the L1 trigger decision is available.

Level-2 Trigger

Events passing the L1 trigger are then passed on to the level-2 (L2) trigger, which together with the event filter (EF) forms the high level trigger (HLT). The two HL triggers are purely software based triggers, running on dedicated computing clusters. The L2 trigger looks more specifically

at the RoIs, within which it has access to all subdetector data at full precision, not only from the calorimeters and the muon systems. It decides which specific data from the read out buffers are needed for a decision. With this mechanism only about ∼2 % of all event data are necessary for the L2 trigger decision. The L2 trigger allowed a time of around 40 ms per event and reduces the event rate by an additional factor of about 20 (95 %) to the order of ∼3.5 kHz.

Event Filter

The third integral part of the trigger system is the offline Event Filter (EF) which uses fully reconstructed events including among others calibration and alignment information. For the EF the processing time per event is on the order of seconds and it reduces the final output event rate to the order of around 200 Hz corresponding to a data rate of some hundred MB/s. This data can then be written out to a permanent storage system.

3.2.5. Data Processing and Computing Tools

Clearly a data rate of several hundred MB per second and an accumulated data on the order of 10 PB per year, having to be accessed by thousands of scientists all over the world is not feasible to handle with conventional server farms. The data should furthermore be stored for several decades to maintain the possibility of re-analyzing it in case of new findings.
To serve these purposes, the Worldwide LHC Computing Grid (WLCG) was developed. It is a decentralized analysis and storage model where all of the data are stored as multiple copies in computing centers all around the world. This model has several advantages: Firstly it is not as vulnerable to failure or data loss as one centralized storage system and secondly the workload and financial burden for maintenance can be more easily and justly allocated amongst the participating institutes. On the other hand, such an large, newly developed grid computing system naturally poses a number of challenges to overcome. For the efficient exchange of data a large-bandwidth network between the contributing institutes has to be established and it has to be made sure that everyone has the same chance and technical capabilities to access the data. Furthermore, it has to be ensured that the same analysis software can be used by the different participants independently of the locally employed hardware. To satisfy all these boundary conditions the computing grid was designed as a system with four hierarchical “layers”, called tiers.
At the **Tier-0** center, located at CERN, the unprocessed raw data are stored but also event reconstruction and data reprocessing is done here as well as a first calibration of the various physics objects. The reconstructed data is then replicated to at least one of the thirteen Tier-1 centers.
The main purpose of those **Tier-1** centers, as illustrated in Figure 3.11, is the permanent storage and the large-scale reprocessing of experimental data as well as Monte Carlo simulation. Furthermore the reduction of data formats containing detailed event information to smaller data formats containing a more manageable amount of information which can be easily used by analysts is to a large part performed at the Tier-1 centers. Data generated or stored at the Tier-1 sites can be accessed from the large number (currently around 160) of **Tier-2** sites which are typically scientific institutes or universities.
Finally there are the **Tier-3** sites which are typically computing clusters, at universities for example, or can also be just a normal desktop PCs in the office of a scientist. The Tier-3s are used for conducting individual analyses. Members of the various LHC experiments have access

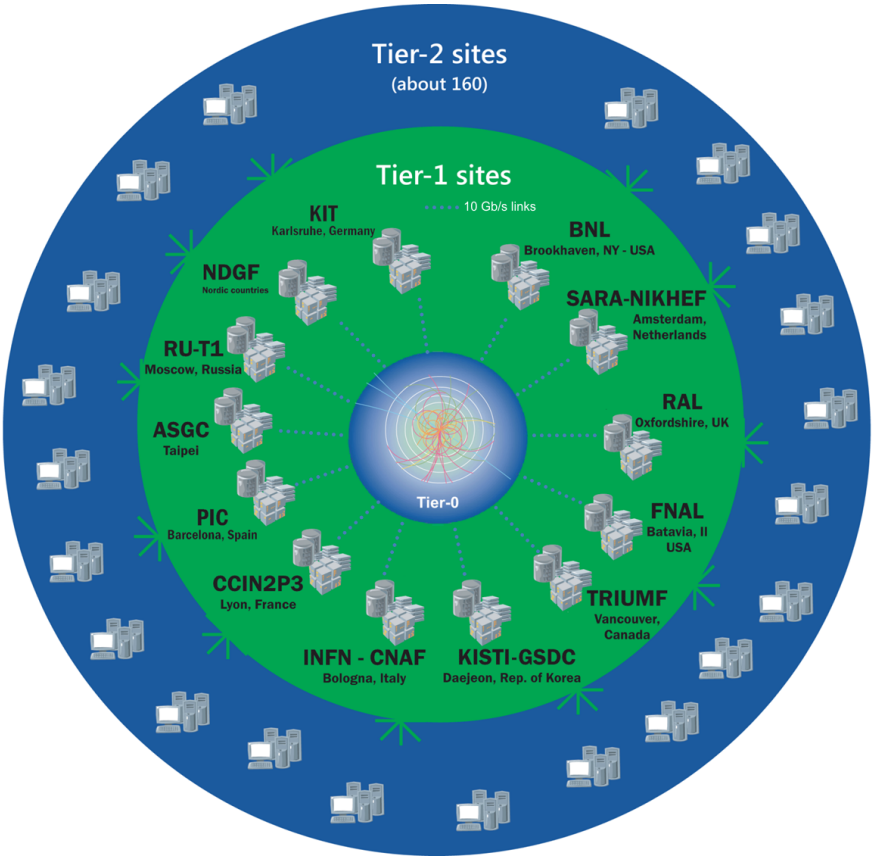


Figure 3.11.: Schematic overview of the Grid architecture up to Tier-2 [69].

rights to the data of their experiment only.
The common interface for using the grid is the "**P**roduction **AN**d **D**istributed **A**nalysis system" (PANDA) which has been developed since 2005 and is the middleware for preparing analysis jobs at a local PC and submitting it to the grid. The philosophy here is to send the analysis jobs to the centers where the data are stored rather than having to transfer a huge amount of data to one central processing location. This of course increasing the processing speed while at the same time saves bandwidth in the data transfer.

Chapter 4

WW+WZ and Background Processes

In this chapter, a short overview over the production and decay mechanisms of the individual signal and background processes will be given. The focus will be on the expected signatures in the detector of the various processes and the possibilities to suppress background events.

4.1. WW+WZ Signal

The production mechanisms for W pairs and WZ were already described in Chapter 2. The two vector bosons can now decay via several channels with different branching ratios, leading to highly different signatures in the detector. A hadronically decaying W boson can in principle decay to any possible combination of up-type and down-type quark except to combinations involving top quarks which are kinematically forbidden. Therefore the decay to bottom quarks is also strongly CKM suppressed. With the exception of the decay to $t\bar{t}$ which is impossible due to kinematic constraints as well, the Z boson can also decay to all quark-antiquark pairs. Using the branching ratios [6],

$$\mathrm{BR}(\mathrm{W} \to \mathrm{q}\mathrm{q}') = 67.41\,\%, \tag{4.1}$$
$$\mathrm{BR}(\mathrm{W} \to \mathrm{l}\nu) = 10.86\,\%, \tag{4.2}$$
$$\mathrm{BR}(\mathrm{Z} \to \mathrm{q}\bar{\mathrm{q}}) = 69.91\,\%, \tag{4.3}$$
$$\mathrm{BR}(\mathrm{Z} \to \mathrm{l}\mathrm{l}) = 3.37\,\%, \tag{4.4}$$

where l indicates each type of lepton (not the sum over them), all other relevant branching ratios can be computed. It is assumed that the branching ratio is exactly the same for all lepton flavours.

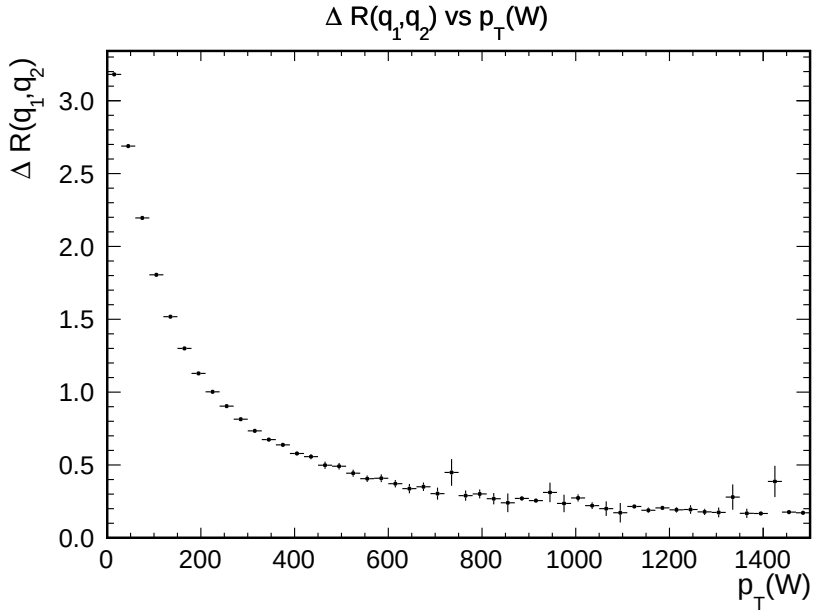


Figure 4.1.: ΔR between the two quarks in the hadronic decay of a W boson depending on its p_T. The plot is obtained from simulated WW events. The average spatial separation between the decay products of the W boson starts to drop below $\Delta R(\mathrm{q}_1, \mathrm{q}_2) = 1$ at p_T (W)>200 GeV, a motivation for the choice of the combination of the large-R jet radius and minimum p_T-cut value used in this analysis (see Section 6).

4.1.1. WW Signal

In case of WW production, the fully hadronic decay channel has the highest branching ratio of $\mathrm{BR}(\mathrm{WW} \to \mathrm{qq'qq'}) = 45.44\,\%$, which is closely followed by the semileptonic decay channel with $\mathrm{BR}(\mathrm{WW} \to \mathrm{l}\nu\mathrm{qq'}) = 43.92\,\%$. Far behind follows the fully leptonic mode with a branching ratio of $\mathrm{BR}(\mathrm{WW} \to \mathrm{l}\nu\mathrm{l}\nu) = 10.61\,\%$.
In the experimental context, decay modes involving τ-leptons are usually not counted to the leptonic modes since the τ-lepton immediately decays again, possibly to hadrons. From this point of view there remain merely:

$$\mathrm{BR}(\mathrm{WW} \to (\mathrm{e}/\mu)\nu\mathrm{qq'}) = 2 \cdot (\mathrm{BR}(\mathrm{W} \to \mathrm{e}\nu) + \mathrm{BR}(\mathrm{W} \to \mu\nu)) \cdot \mathrm{BR}(\mathrm{W} \to \mathrm{qq'}) = 29.28\,\%, \quad (4.5)$$

$$\mathrm{BR}(\mathrm{WW} \to (\mathrm{e}/\mu)\nu(\mathrm{e}/\mu)\nu) = (\mathrm{BR}(\mathrm{W} \to \mathrm{e}\nu) + \mathrm{BR}(\mathrm{W} \to \mu\nu))^2 = 4.71\,\%. \quad (4.6)$$

The electron and muon channels investigated here thus have a branching ratio in each of the two channels of:

$$\mathrm{BR}(\mathrm{WW} \to \mathrm{l}\nu\mathrm{qq'}) = 2 \cdot \mathrm{BR}(\mathrm{W} \to \mathrm{l}\nu) \cdot \mathrm{BR}(\mathrm{W} \to \mathrm{qq'}) = 14.64\,\%, \quad (4.7)$$

where l stands either for an electron or a muon. The electron or muon from the leptonic W can be well detected while the corresponding neutrino does not interact with the detector, thus leaving its signature as large missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{Miss}}$) which will be introduced in more detail in Section 6.1.5.
The other W boson decays to two quarks which, after hadronizing, form two particle jets (see Chapter 5) with a spatial separation that depends inversely on the transverse momentum of the W. The dependence of the separation on the p_T of the W boson is shown in Figure 4.1.

In the event of a gluon radiation in the initial or final state, there could also be one or even several additional (typically low p_T) jets.
For highly boosted hadronically decaying W bosons, the two quarks become so collimated that they cannot be reconstructed as two separate small-R jets anymore. Instead, what can be done is to reconstruct jets with a much larger radius, so-called "large-R jets", and to assume that the entire decay products of the W have been included inside this jet. The invariant mass of this large-R jet should then in principle be the W mass.
Summarizing, the signature expected from a boosted semileptonic WW or WZ decay (where the Z decays hadronically) is the presence of exactly one central lepton (e, μ), large missing transverse momentum and exactly one central large-R jet with an invariant mass in the W/Z mass range and a high transverse momentum.

4.1.2. WZ Signal

In the WZ case, the decay mode that is in principle measured in this analysis corresponds to cases where the W decays leptonically and the Z decays hadronically. Events with a leptonic Z and a hadronic W are highly suppressed by a veto on a second lepton and the requirement of large missing transverse momentum. The two remaining combinations of decay modes of the W and Z boson (both hadronic and both leptonic) are even more suppressed by either the lack of or too many leptons in the event. The relevant branching ratios here for a given lepton channel amount to:

$$\mathrm{BR}(\mathrm{WZ} \to \mathrm{l}\nu\mathrm{q}\bar{\mathrm{q}}) = \mathrm{BR}(\mathrm{W} \to \mathrm{l}\nu) \cdot \mathrm{BR}(\mathrm{Z} \to \mathrm{q}\bar{\mathrm{q}}) = 7.59\,\%, \tag{4.8}$$

$$\mathrm{BR}(\mathrm{WZ} \to \mathrm{qq}'\mathrm{ll}) = \mathrm{BR}(\mathrm{W} \to \mathrm{qq}') \cdot \mathrm{BR}(\mathrm{Z} \to \mathrm{ll}) = 4.54\,\%. \tag{4.9}$$

The sum of the leptonic branching fraction of the electron and muon channel is thus given as two times the result of Equation 4.8:

$$\mathrm{BR}(\mathrm{WZ} \to (\mathrm{e}/\mu)\nu\mathrm{qq}) = 15.81\,\%. \tag{4.10}$$

Topologically, $\mathrm{WZ} \to \mathrm{l}\nu\mathrm{qq}$ events naturally resemble $\mathrm{WW} \to \mathrm{l}\nu\mathrm{qq}'$ events to a high degree, with the primary difference being the invariant mass of the large-R jet. This implies that the only way to distinguish such events would be via a measurement of this jet mass, which is expected to be around 90 GeV in WZ events and around 80 GeV in WW events. However, the jet mass resolution of the detector is not small enough to accurately distinguish such a small mass difference, which is why in this analysis, the sum of the WW and WZ cross section are measured. A comparison of the expected invariant jet mass shapes for W/Z-induced jets obtained from MC simulation can be seen in Figure 4.2.

4.2. W/Z + Jets Background

By far the most dominant background arises from so-called W+jets events which are events with a real W boson and jets from additional quarks or gluons in the final state, as can be seen in Figure 4.3. The W decays leptonically, providing for the lepton and $E_{\mathrm{T}}^{\mathrm{Miss}}$ and the jet makes up for the high p_T large-R jet. This combination makes this process an irreducible background.
The predicted SM cross section for W+jets production, computed with MCFM [70], is furthermore

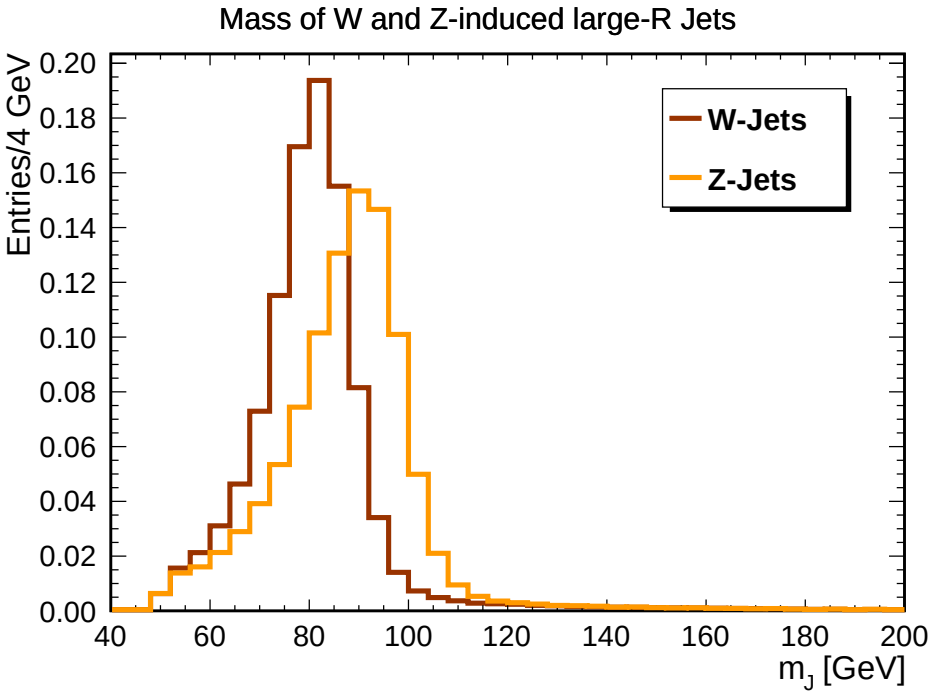


Figure 4.2.: The plots, obtained from MC simulations, show the invariant mass of large-R jets as they are used in this analysis, induced by W or Z bosons. Both distributions are normalized to unit area. The two distributions are overlapping to a large fraction, rendering a distinction of W and Z-induced large-R jets essentially impossible for this analysis.

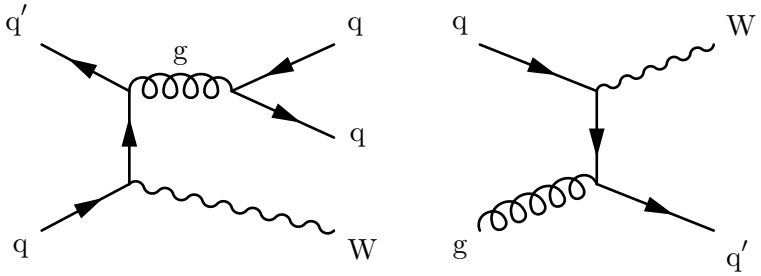


Figure 4.3.: Two representative Feynman diagrams of W+jet production.

several orders of magnitude higher than the cross section for WW production:

$$\sigma_{\mathrm{W}\to l\nu} = 12.087^{+0.57}_{-0.51}\,\mathrm{nb}. \tag{4.11}$$

The quantity $\sigma_{W\to l\nu}$ already takes into account the leptonic branching ratio of the W, such that Equation 4.11 represents the cross section per lepton flavour.
The same process can occur with a Z decaying to two charged leptons instead of a W decaying to a lepton and a neutrino. These Z+jets events (shown in Figure 4.4) can, however, be efficiently suppressed by a veto on a second lepton and the requirement of large $E_{\mathrm{T}}^{\mathrm{Miss}}$. Besides, the cross section for such processes which was computed with MCFM as well is around a factor ten lower than for W+jets:

$$\sigma_{\mathrm{Z}\to\mathrm{ll}} = 1.112\,\mathrm{nb},\ \text{where}\ 66 < m_{\mathrm{ll}} < 116\,\mathrm{GeV}. \tag{4.12}$$

Summarizing, W+jets events are expected to constitute a large background, while Z+jets events are strongly suppressed. It will be shown later that the ratio of Z+jets in V+jets events is only

2.2 % after the event selection is applied (as shown in Figure 8.4b and Table 10.2).

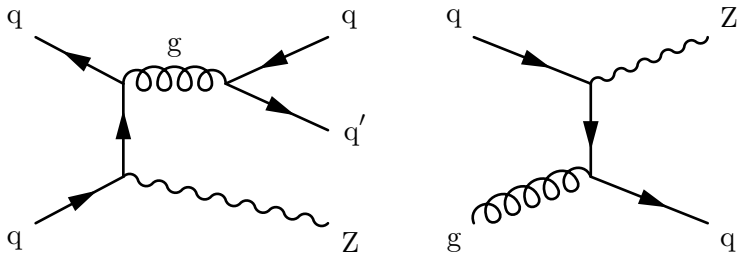


Figure 4.4.: Two representative Feynman diagrams of Z+jet production.

4.3. Top Quark Background

The relatively high production cross section and the actual W bosons emerging from every top quark decay make the top quark background channel the second most important background process after W+jets. Top quarks can be produced either singly via weak processes or as quark-antiquark pairs via the strong interaction at leading order. The top quark pair production is therefore enhanced compared to single top production.
An important property of the top quark is the fact that its decays almost invariably includes a bottom quark as a result of the large value of V_{tb} in the CKM matrix (see Equation 2.37). For this reason the final state quark from a top decay is always labeled with "b" in Figure 4.6 and Figure 4.7.
Kinematically the top quark could decay to down or strange quark as well but these decays are CKM suppressed by a factor of $8 \cdot 10^{-5}$ and $2 \cdot 10^{-3}$, respectively. This fact proves very useful in the context of identifying top quark decays. A technique helping to distinguish jets originating from a b-flavoured quark from other jets was therefore developed.
The bottom quark can kinematically only decay into u/c-flavoured quarks but because these transitions are cross-generation transitions, both decay modes are CKM suppressed. The resulting long lifetime of hadrons containing b-flavoured quarks can be exploited for b-jet identification via the detection of secondary vertices, displaced with respect to the primary vertex, that arise from a late bottom quark decay. This feature and several other features likely to be exhibited by b quark-induced jets are combined in a multivariate technique in order to identify b-jets, a procedure referred to as b-tagging [71].

4.3.1. Top Quark Pair Production

The theoretical prediction for the $t\bar{t}$ production cross section is

$$\sigma_{t\bar{t}} = 252.9 \pm 15\,\text{pb}, \tag{4.13}$$

as calculated with the Top++2.0 program to next-to-next-to-leading order in perturbative QCD, including soft-gluon resummation to next-to-next-to-leading-log order (see Ref. [72] and references therein), and assuming a top quark mass $m_{top} = 172.5$ GeV. One example of $t\bar{t}$ production with a semileptonic decay is shown in Figure 4.5. Figure 4.6 shows more examples of $t\bar{t}$ production

mechanisms, here without the $t\bar{t}$ decay.

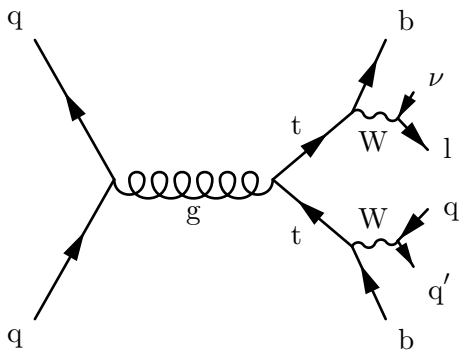


Figure 4.5.: $t\bar{t}$ production with semileptonic decay.

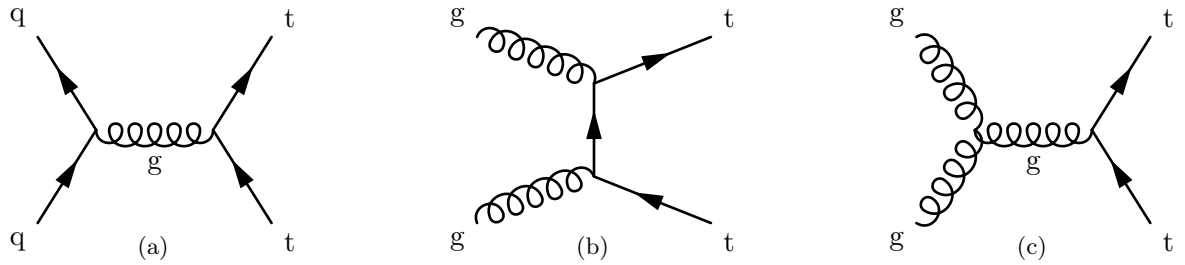


Figure 4.6.: Main Feynman diagrams for top pair production:
(a) s-channel, (b) t-channel, (c) gluon fusion.

Despite its large cross section, the $t\bar{t}$ background can be suppressed relatively easily by vetoing events with a high jet multiplicity (see Figure 6.2 and 6.3). In case of a fully leptonic $t\bar{t}$ decay, the event will likely be discarded due to the second lepton whereas in a semileptonic or all hadronic decay, there will be at least four jets (one jet from the leptonic top quark and three jets from the hadronic top quark). This is entirely different to the WV event topology that in the ideal case has a maximum of two or three jets and can therefore be distinguished very well.

4.3.2. Single Top Quark Production

Single top quark production can proceed via three different channels which are illustrated in Figure 4.7: s-channel, t-channel and W-associated (Wt) production. Because of its small production cross section at the LHC, the s-channel is of little relevance for this analysis. The single top quark production cross sections for the three different channels cumulatively lead to $\sigma_{\text{single top}} = 115.74$ pb with the respective contributions of the three channels [73–75]:

$$\sigma_{\text{s}} = (5.61 \pm 0.22)\,\text{pb}, \tag{4.14}$$

$$\sigma_{\text{t}} = (87.76^{+3.44}_{-1.91})\,\text{pb}, \tag{4.15}$$

$$\sigma_{\text{W}t} = (22.37 \pm 1.52)\,\text{pb}. \tag{4.16}$$

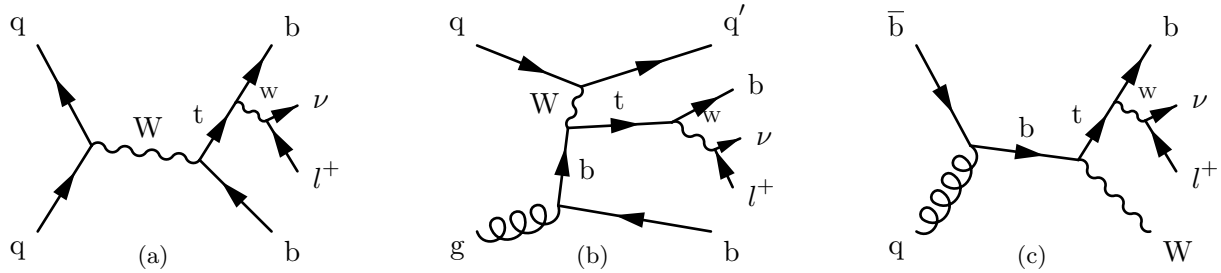


Figure 4.7.: Single top production in the s-channel (a), t-channel (b) and the W-associated channel (c).

The relevant events for this analysis for all single top processes are the ones where the top quark decays leptonically, providing a lepton and $E_{\mathrm{T}}^{\mathrm{Miss}}$. The W-associated production is especially important as a background because there are not only one but two genuine W bosons present in the final state, making this process much harder to suppress than the other single top channels. Despite having a cross section roughly four times smaller than the t-channel (compare Equation 4.15 and 4.16), the event yield after event selection is 13 times higher in the Wt-channel than in the t-channel, see Figure 8.3b. Since the single top quark contribution to the total background is rather small compared to $\mathrm{t\bar{t}}$ (see Figure 8.3a), the two processes are merged in the plots hereafter and labelled as "Top" only.

4.4. QCD Multijet Background

In order to become a background process, QCD multijet events that by definition do not contain any W or Z boson, have to exhibit several uncharacteristic features. First of all, there must be sizable $E_{\mathrm{T}}^{\mathrm{Miss}}$ which in a QCD event is not present since no neutrinos are involved. However, on reconstruction level there can be considerable $E_{\mathrm{T}}^{\mathrm{Miss}}$ due to mismeasurement, pile-up or miscalibration.
Second, the lepton requirement can only be met if a light flavour jet is being misidentified as an electron or muon (which is extremely rare in the muon case) or if there is a real lepton from a semileptonic heavy flavour decay in a jet.
The third requirement, the presence of a high p_T large-R jet can well be met by the QCD processes itself without the necessity of any signature to be faked by another object. Any high p_T quark or gluon can form a large-R jet, however, exhibiting a different substructure than a vector boson-induced one.
In fact, it will be shown that the QCD multijet background is completely negligible in the muon channel and also rather insignificant in the electron channel (see Section 7.1). The only reason why still a contribution from those events is present is because the multijet production cross section is several orders of magnitude higher than the one for WW and WZ. The inclusive cross section for two or more jets with $p_T > 30$ GeV in $|\eta| < 2.8$ was measured at $\sqrt{s} = 7$ TeV [76] to be

$$\sigma_{\text{QCD-jets}} \approx 3000\,\text{nb}. \tag{4.17}$$

The rate of QCD-jets passing lepton identification criteria is not well predicted by Monte Carlo simulation. In addition, an enormous amount of Monte Carlo generated events would have to be generated in order to obtain adequate statistics after event selection, since the lepton requirement strongly suppresses the QCD multijet background.
For those two reasons the multijet background is estimated from a data-driven method (see Section 7.1).

Chapter 5

Jets and Jet Substructure Methods

5.1. Introduction to hadronic Jets

Since quarks and gluons carry colour charge, it is inherently impossible to observe them experimentally as single, free particles. What can be detected, however, are the colourless hadrons generated in the aftermath of the production of a coloured parton.
When a quark or gluon is produced in the hard scatter it will radiate further virtual gluons (q→q+g and g→g+g), followed by gluon splitting into quark-antiquark pairs as described in Section 2.1.4 – a process known as "*parton showering*".
The radiation of particles predominantly occurs at low angles, meaning the shower develops mainly in the direction collinear to the incident parton. The parton shower (PS) process is technically impossible to be computed by matrix element (ME) calculations (fixed order perturbative QCD calculations) since those diagrams would be of rather high order which would be extremely computationally expensive. Additionally, matrix element calculations are valid only for a limited number of particles and when the partons are hard and well separated. Instead of ME calculations, parton shower algorithms are used to describe the momentum-transfer evolution from the hard scatter scale down to a point where perturbative QCD is not valid anymore and hadronization takes over. Those algorithms have the advantages of being computationally inexpensive and having no limitations in terms of particle multiplicity. Parton shower algorithms are valid in regimes of soft and collinear parton radiation. The two most important approaches here are the dipole approach [77] which is for example employed in the MC generator Pythia 8 [78, 79] and the collinear final state evolution [80]. Since PS and ME calculations perform best in complementary regions of phase space, the best descriptions in MC are thus obtained by the combination of the two approaches.
Therefore the outgoing partons of the matrix element calculation are interfaced to the parton shower algorithm together with information about those particles such as mother-daughter relationships and spin correlations. A possible complication here is presented by the potential double-counting of topological configurations in regions where the two approaches overlap. So-called "matching" techniques have been developed in order to overcome this problem, the most important one for this analysis being the CKKW technique [1] [81]. With this technique, the phase

[1] Named after the four first authors Cantani, Krauss, Kuhn and Webber

space is divided into a perturbative fixed order calculation and a parton shower dominated regime by an energy scale matching parameter d_{CKKW}. At energies below d_{CKKW}, parton showering is used to model parton branchings, radiation at larger energies is described by perturbative calculations.
After the PS, when the partons have reached low energies and perturbative QCD breaks down ($Q \sim 1\,\text{GeV}$), confinement sets in and the produced particles start to form observable hadrons, a process referred to as "*hadronization*". As this step takes places in the non-perturbative regime, the hadronization cannot be calculated from first principles. However, it can be parametrized and modeled by different phenomenological approaches such as the Lund string model [82] which is employed in Pythia or the cluster model [83] used in HERWIG [84].
Similar to the parametrization of the hadron structure in PDFs, the hadronization is usually parametrized in terms of fragmentation functions $D_{h\to H}(z, \mu_F^2)$ that describe the probability of a certain parton h to hadronize into a hadron H which carries the fraction z of the incidental parton momentum.
A graphical representation of the different underlying processes in jet formation can be seen in Figure 5.1.

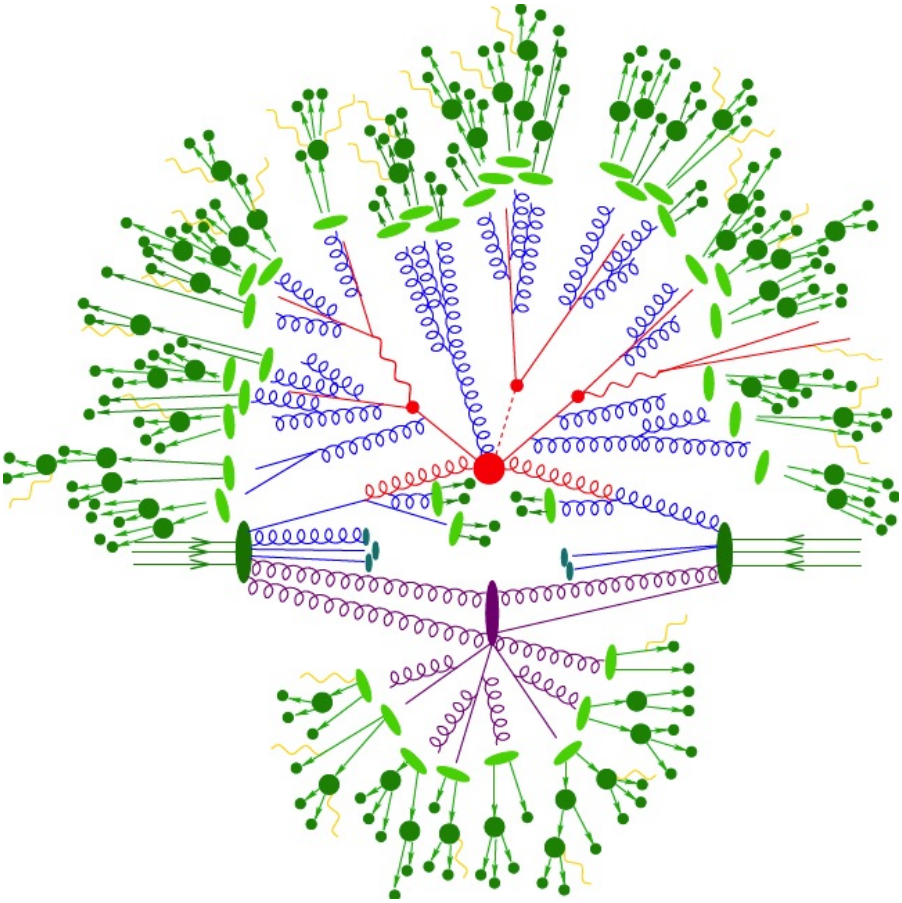


Figure 5.1.: Pictorial representation of a proton proton collision with jet final states as produced by an event generator [85]. The hard interaction (large red circle) produces a $t\bar{t}$ + Higgs final state (small red circles). The left top quark decays hadronically as does the Higgs boson while the right-hand side top quark decays to leptons. QCD radiation is depicted in blue, light green circles are symbolizing the hadronization of final state partons and the dark green circles show hadron decays. A secondary interaction inside between the two protons also takes place (purple circle). Photon radiation depicted by yellow wavy lines can occur at all stages.

Considering the aforementioned processes, the final state expected from the production of a hard parton is a collimated spray of hadrons, referred to as a *jet*.
The internal structure of a jet can significantly differ depending on the type of particle it originated from. Essentially three sorts of jets can be distinguished:

1. Light-quark-initiated jets
2. Gluon-initiated jets
3. Boosted heavy particle-initiated jets

Boosted heavy particle jets are strictly speaking quark jets as well but are called boosted heavy particle jets only if the final state products of two (or more) quarks are merged into one jet. This type of jet is rather different from the other two types which are much more difficult to be distinguished from one another.
However, there are some differences between gluon and quark jets which make an experimental distinction possible [86]. Gluon emission is the dominant process in the parton shower and is proportional to a colour factor determined by the type of emitting particle. When the emitter is a gluon, the factor is $C_A = 3$ and $C_F = 4/3$ when it is a quark. As a consequence, the average multiplicity of (any kind of) particles in a gluon jet is 9/4 times as high as in quark jets (if $E_{\text{particle}} \ll E_{\text{jet}}$) and because of the greater radiation of soft gluons, a gluon jet is broader. The difference in angular width δ (using the definition from Ref. [87]) is given at leading order as $\delta_{\text{g}} = \delta_{\text{q}}^{C_F/C_A}$ [88]. The larger particle multiplicity in gluon jets implies for a given jet energy that a gluon jet must contain fewer hard particles than a quark jet of the same energy. In summary, gluon jets are broader and contain more and softer particles than quark jets at a comparable energy.
The big difference between boosted heavy particle jets and ordinary quark jets is the presence of at least two quark jets (with a symmetric energy share in case of a decaying vector boson) instead of only one and its invariant mass which is the mass of the decaying particle.
In general the jet mass can be calculated as the invariant mass of a multibody system with massless particles. The invariant mass of an N body system with four-vectors p_i of the individual particles can be written as:

$$M_{\text{N}} = (p_1 + p_2 + ... + p_{\text{N}})^2. \tag{5.1}$$

This evaluates to

$$M_{\text{N}} = M_{12}^2 + ... + M_{1\text{N}}^2 + M_{2\text{N}}^2 - m_1^2 - m_2^2 - ... - m_{\text{N}}^2, \tag{5.2}$$

which simplifies for massless particles to

$$M_{\text{N}} = M_{12}^2 + ... + M_{1\text{N}}^2 + M_{2\text{N}}^2, \tag{5.3}$$

with the invariant masses of the two body sub-systems M_{ij}. This in turn can be expressed by the particle momenta and their opening angle θ as [89]:

$$M_{\text{ij}} = |\vec{p_i}||\vec{p_j}| \sin^2 \theta. \tag{5.4}$$

Summarizing, the jet mass is the sum of the invariant masses of all possible combinations of constituents. Considering Equation 5.4, it becomes clear that a jet can aquire mass by either a high momentum of or by large angles between the constituents. This is the reason why gluon jets consisting of rather soft but far-spread particles can still aquire a large mass.
It is desirable to be able to draw as precise conclusions as possible about the initial parton from a jet, without having to consider all underlying parton shower and fragmentation processes in

detail. This consideration inevitably leads to the concept of grouping the final state particles together again, following a certain pattern in order to obtain an object, kinematically closely resembling to the initial parton. The prescription for how this grouping is performed is then called a "jet algorithm" and can be chosen somehow arbitrarily provided it fulfills certain minimal criteria.
A jet algorithm should be able to start with a large number of hadron-like objects that can be described as a four-vector, e.g. simulated particles from MC or calorimeter clusters in the detector, and produce as output a small number of parton-like objects by clustering the initial constituents together.
The inclusion of an additional soft constituent should not alter the final jet in any way – a property which is termed *infrared safety*. Furthermore, the final set of jets should remain unchanged if one constituent is replaced by a pair of two collinear constituents which is referred to as *collinear safety*.
Depending on the particular algorithm in use, the same input constituents can lead to a different set of jets with different direction, energy etc. However, if the same algorithm is used for all jets, they become comparable.
In that sense a jet is not as straightforward to define as a lepton for example. A possible general definition is that a jet is the outcome of a certain clustering algorithm executed on a set of input objects.
A large variety of different jet algorithms exist. Historically, so-called cone-type algorithms were favored at hadron colliders because of their increased computational performance with sequential recombination algorithms (see Section 5.2) at that time.
In general, cone algorithms start with a seed particle in the event and put a cone of a certain radius around it. The momenta of all particles inside the cone are summed up and a jet axis is calculated. If the jet axis coincides with the seed axis (within some precision), the jet is finished. Otherwise the jet axis is used as new seed and the process is iterated until it converges to a stable jet.
The great disadvantage of the cone-type algorithms was that they were neither infrared nor collinear safe (except SISCone [90]) and furthermore they involved several non-physical parameters. Nowadays, after considerable improvement of the algorithmic implementation of sequential recombination algorithms, they have practically replaced cone-type algorithms entirely. Different types of these sequential recombination algorithms have been used in the present analysis and will be described in the following.

5.2. Sequential Recombination Jet Algorithms

A jet can be imagined as being the product of a repeated $1 \to 2$ splitting of quarks or gluons with the outcome of a large number of low-energy particles. The recombination algorithms seek to work this way backwards, beginning from the final end of the decay chain to the start by consecutively combining two constituents into one, in principle leading to the original parton.
The sequential recombination algorithms make use of a generalized measure of distance between the constituents i and j (d_{ij}) and between constituent i and the beam direction (d_{Bi}):

$$d_{ij} = \min(p_{Ti}^{2m}, p_{Tj}^{2m}) \cdot \frac{(\Delta R_{ij})^2}{R^2}, \qquad d_{Bi} = p_{T,i}^{2m}, \tag{5.5}$$

$p_{T,n}$ being the transverse momentum of constituent n. R is a parameter which controls the radius of the jet in the $y - \phi$-plane. Generally a recombination algorithm works as follows:

It starts with a constituent i and makes a list of all d_{ij} values that are theoretically possible, including d_{Bi}. If the smallest entry of the list is part of the d_{ij} set (and not d_{Bi}), i and j are combined to a new constituent with four-momentum $p_i + p_j$ which is added to the list while i and j are removed. This process continues until the smallest entry is not anymore a d_{ij} but d_{Bi}. At this point i is taken out of the list and added to the list of inclusive jets of the event. With this approch all initially present constituents end up in a jet though usually a minimum p_T is required for a jet to be counted as such.
In addition to this inclusive formulation of the sequential algorithms, there is also the exclusive mode. The distance measures here are the same and the only difference is that in the exclusive algorithm, if a d_{Bi} is found to be the smallest entry, i is removed from the list and attributed to the beam jet. This procedure is repeated until the smallest d_{ij} or d_{Bi} is above a certain cut value d_{cut}. All constituents not yet attributed to the beam jet are now classified as the event's non-beam jets. Historically jet reconstruction with such algorithms used to be computationally intensive until in recent years the FastJet software package including a much more efficient implementation was developed [91, 92]. FastJet was also extensively used for jet reconstruction in this analysis.
In practice the different types of sequential recombination algorithms are determined by the value the parameter m takes in equation 5.5:

$$m = \begin{cases} 1 & k_T \text{ Algorithm [93]} \\ 0 & \text{Cambridge/Aachen Algorithm (C/A) [94]} \\ -1 & \text{Anti - } k_T \text{ Algorithm [95]} \end{cases}$$

All three of these algorithms are infrared and collinear safe by construction but each of them exhibits certain advantages or disadvantages for particular analysis purposes.

5.2.1. k_T Algorithm

The distance measure for this algorithm type is closely related to the structure of the underlying QCD branching. Since $m = 1$ here, constituents with low transverse momentum are merged first and the hardest constituents last, which is an approximate inversion of the QCD fragmentation process. One downside of this procedure is that k_T-jets are not *soft-resilient*, meaning that soft particles at the outer rim of the jet can modify its shape which can be a problem since jets with an irregular shape are harder to calibrate and to correct for pile-up. The irregular jet shapes produced in this algorithm can be seen in Figure 5.2.

5.2.2. Cambridge/Aachen Algorithm

Since here $m = 0$, the C/A algorithm uses a very simple distance measure which is just the actual spatial distance in y-ϕ-space and is therefore entirely independent of energies and momenta. The closest constituents are merged first, motivated also by the underlying QCD fragmentation which occurs preferentially at low splitting angles. Since the C/A algorithm is consequently particularly sensitive to the angular ordering of the constituents, it is in principle well-suited for the analysis of jet substructure that seeks to distinguish compact hard subjets from soft, spread out large-angle contributions. The fact that the transverse momentum of the input constituents

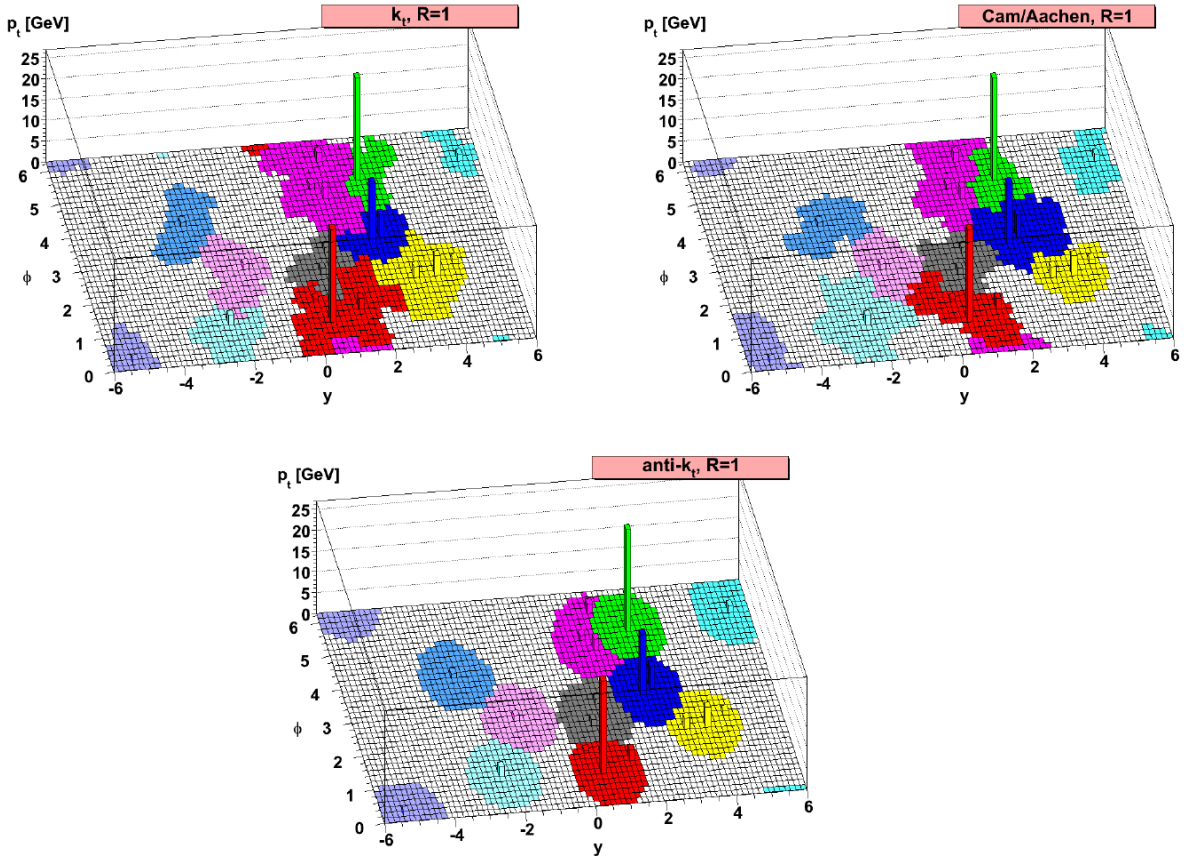


Figure 5.2.: Illustration of the three most important sequential jet algorithms. The same partonic event together with a large number of extremely soft "ghost" particles is clustered with the respective algorithm, with a radius parameter of $R = 1$. The resulting jet shapes can be determined by the ghost particles clustered into the jet. Noteworthy are the irregular shapes of the C/A and k_T jets, as well as the exactly circular shaped anti-k_T jets. Also the typical behaviour of the anti-k_T algorithm of attributing all overlapping constituents between two jets to the harder of two jets is visible. [95]

is irrelevant implies a sensitivity to soft objects which can cause the jets to be irregularly shaped as well.

5.2.3. Anti - k_T Algorithm

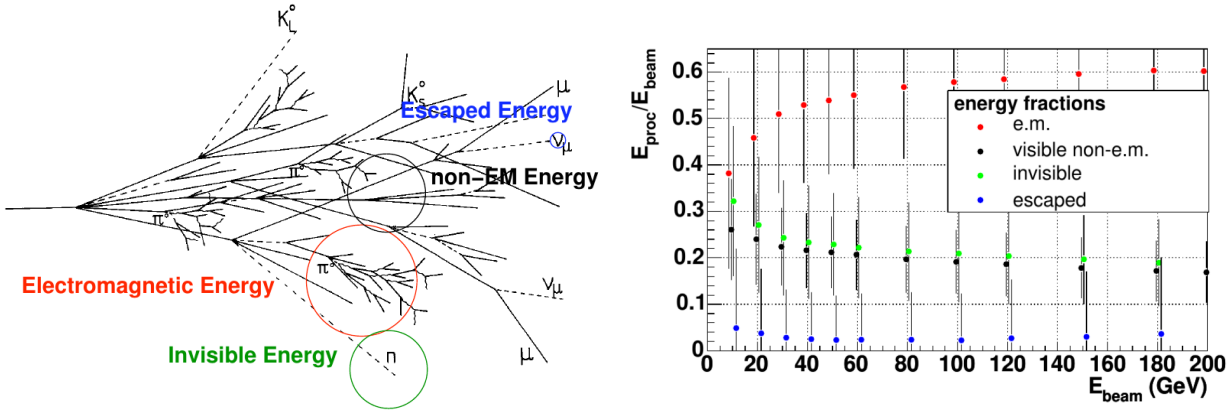
This algorithm uses $p = -1$, therefore $d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \cdot \frac{(\Delta R_{ij})^2}{R^2}$ will always be governed by the harder of two particles and the separation ΔR_{ij}. For that reason the anti - k_T algorithm favours the combination of one hard particle and another spatially close particle, whereas clustering of two soft objects is strongly disfavoured. This means the algorithm tends to agglomerate soft material into an existing hard subjet instead of combining soft contributions first. In consequence, unlike for the C/A and the k_T algorithm, the clustering of the anti - k_T algorithm cannot be appropriately related to the sequential branching of the QCD processes that formed the jet.
The course of the anti-k_T algorithm is furthermore not as straightforward to imagine as that of the other sequential recombination algorithms but can be well illustrated by considering an

event with some well separated hard particles and many soft contributions around those.
As stated earlier, the separation d_{ij} between a hard and a soft particle i and j is almost exclusively determined by the p_T of the harder of the two and much smaller than the separation between two soft particles at the same spatial distance ΔR_{ij}. If in a distance of $2R$ of the initial hard object no other hard constituent is found, the jet will be considered to be final with the soft particles inside a radius of R clustered around the hard *seed*. For this reason the anti-k_T jets usually have perfectly round cone-like shapes.
However, if there is another hard constituent in the range between R and $2R$ of the first jet, a second jet will be formed with the overlapping part being attributed to the jet with the higher transverse momentum. The softer jet in this case cannot have a round shape anymore but rather has to form a circle with a cut-out lens shaped part, as is shown in Figure 5.2.
The important feature here is that soft particles at the jet border do not alter its shape but only hard contributions do so (soft resilience).
The anti-k_T algorithm is the default jet-algorithm in ATLAS since it has several advantages over other algorithms like the regular shapes which make the jets easier to calibrate. In addition, the reconstruction and trigger efficiency as well as performance under pile-up conditions are enhanced compared to other algorithms [96].

5.3. Topological Clustering and Local Hadronic Calibration

5.3.1. Particle Showers

Depending on the type of initial particle, there are two rather different sorts of particle showers which develop during the interaction of the particle with matter.
An electromagnetic shower (a shower caused by electrons or photons) in matter starts with a particle, e.g. a photon undergoing electron-positron pair production in the field of an atomic nucleus. The two leptons will emit further photons (bremsstrahlung) which can again form electron-positron pairs. In this way, the number of secondary particles grows exponentially, signifying for the individual particles an exponential decrease in energy. The process continues until the energy of the produced photons is too small to create further e^+e^- pairs. All the present (now low energy) leptons deposit their energy in the surrounding material via ionization and Compton scattering. In general, because of the exponential nature of the shower development, electromagnetic showers are relatively short and therefore dense objects.
Hadronic showers are usually more complex than electromagnetic showers.
A highly energetic hadron interacts with matter via different nuclear reactions, the most important reason for energy deposition being the production of π^0 or η^0 mesons decaying into two photons and initiating electromagnetic subshowers. The energy released by such processes is called *visible electromagnetic energy* because it can be measured in non-compensating calorimeters.
Another source of energy loss is the production of charged particles other than electrons or muons (mostly charged pions and protons) which deposit their energy via ionization, (*visible non-electromagnetic energy*).
A similar fraction as the visible non-EM energy of the incident particle is lost in processes of nuclear excitation or the break-up of nuclear bindings. In calorimetry, this energy deposition is usually referred to as *invisible energy*, since it is not measurable at all or, in the case of nuclear excitation, only measurable with a time delay such that it cannot be attributed to the original shower anymore.
The last type of energy deposition is the so called *escaped energy*, which is the kinetic energy of



(a) (b)

Figure 5.3.: (a) Illustration of the different types of energy which are either measured in the calorimeter or are undetected [97, 98]. (b) Fractions of the four types of energy deposition processes E_{proc} for charged single pions, summed over active and inactive calorimeter regions in EMEC+HEC, depending on the initial pion energy E_{beam} [99].

neutrinos that emerge during a shower and escape the detector.
Figure 5.3a and 5.3b show an illustration of the different types of energy depositions and their relative importance depending on the incident particle's energy. The involved processes are non-perturbative and cannot be calculated easily. Therefore MC based simulations are used for an empirical description of the longitudinal and lateral shower development. In comparison to electromagnetic showers, hadronic showers develop later and have a larger lateral extension.

5.3.2. Topological Clustering

In ATLAS, jets are generally reconstructed from calorimeter clusters. These clusters are larger groups of calorimeter cells that have been combined according to a particular selection scheme. The reasons for building clusters first instead of using the plain calorimeter cells are manifold. The two most important motivations for clustering are that cells which receive pile-up and noise contributions can easily be suppressed in the clustering process and that the final cluster can be well calibrated to the actual energy of the incoming particle shower. This would be impractical on a cell-level.
Several different clustering schemes exist with the default in ATLAS being the so-called topological clustering [100]. This clustering scheme makes use of the actual shape of the energy deposit by a particle shower in contrast to a fixed-size approach and performs very well in terms of noise suppression and correspondence of the cluster to the initial particle properties [101]. Any calorimeter cell with a signal over noise ratio of more than 4 is a potential cluster seed. The noise in this case is the quadratic sum of electronics noise and estimated pile-up noise and the signal is the absolute value of the measured energy.
All the neighbouring cells are checked for their signal over noise ratio and in case it is above 2, the cell is added to the cluster. This procedure is repeated for every cluster cell until no more neighbouring cells above 2 standard deviations are found. Then, in a final step, one more layer of neighbouring cells at the border of the cluster is added, independently of their signal to noise ratio in order to include low energetic tails of the deposition.
Because energy depositions from two spatially close particles are highly likely to be merged into

one cluster, clusters are split around local energy maxima in a subsequent step. This splitting helps to distinguish energy depositions caused by different particles which in turn facilitates the calibration of the individual clusters and the analysis of jet substructure.
Since basically the entire energy of electromagnetic showers and electromagnetic subshowers in hadronic showers is contained and also measured in the cluster, the cluster is termed to be at the *electromagnetic scale* at this point.
However, in the case of hadronic showers a large part of the initial partons' energies can remain undetected (see Figure 5.3b). In order to recover the different energy types which are not measured in the calorimeter, a dedicated calibration scheme, the *local hadronic calibration* (LC) has been developed.

5.3.3. Local Hadronic Calibration

The local hadronic calibration [99, 102], consists of one classification step and three consecutive weighting steps. The rather large differences between electromagnetic and hadronic showers make it necessary to distinguish between the two kinds of corresponding clusters. This is done in a classification step which looks at certain cluster properties such as the energy density and the cluster's depth in the calorimeter. As described earlier, electromagnetic clusters tend to exhibit a high energy density and a shallow depth inside the calorimeter while hadronic clusters develop deeper inside the calorimeter and have a lower energy density. Based on those properties, a given cluster gets assigned a probability p^{EM} of being electromagnetic. Its probability of being hadronic is then given as $p^{\mathrm{Had}} = 1 - p^{\mathrm{EM}}$. In order to reduce the dependency of the calibration on the classification, the weights applied in all calibration steps are of the following form:

$$\omega^{\mathrm{tot}} = p^{\mathrm{EM}} \cdot \omega^{\mathrm{EM}} + (1 - p^{\mathrm{EM}}) \cdot \omega^{\mathrm{Had}}. \tag{5.6}$$

The individual weights are derived from single pion simulations by comparing the reconstructed energy of calorimeter cells with the actual energy that was deposited inside these cells. Following classification, the actual energy calibration is applied in three steps.
The first weight which is applied is a hadronic weight in order to compensate for invisible and escaped energy in hadronic showers. Since the energy of an electromagnetic shower is measured completely, a calorimeter cluster with an electromagnetic probability of 1 does not receive any hadronic weight. Clusters with a hadronic probability different from zero, however, receive a weight proportional to is hadronic probability because of the fact that in hadronic showers a large part of the energy is not detected.
Following the hadronic weighting, the out-of-cluster step corrects clusters for energy deposited in calorimeter cells which are eventually not attributed to the cluster due to noise thresholds.
Finally, a dead material correction is applied to the cluster. As the name suggests, a calibration factor is evaluated for every cluster, which accounts for energy in the shower that is not deposited in the calorimeter but in inactive, dead material either upstream, beyond or in between calorimeters.
Now the cluster is calibrated to the *hadronic scale* and can be used as input for jet making.

5.4. Pile-Up

Since the LHC is designed to deliver unprecedented instantaneous luminosity, one of the major challenges for physics measurements is the so called pile-up (PU). Every proton bunch contains around $115 \cdot 10^9$ protons [103] from which around 20 collide in every bunch crossing with a proton of the other bunch (in-time pile-up).
The products of all those collisions enter the calorimeters at approximately the same time and hence cannot be associated anymore to the vertex from which they emerged. The amount of this in-time pile-up can be characterized by N_{PV} which is the number of primary vertices actually present in the event. A typical event containing a large number of in-time pile-up events is shown in Figure 5.4.

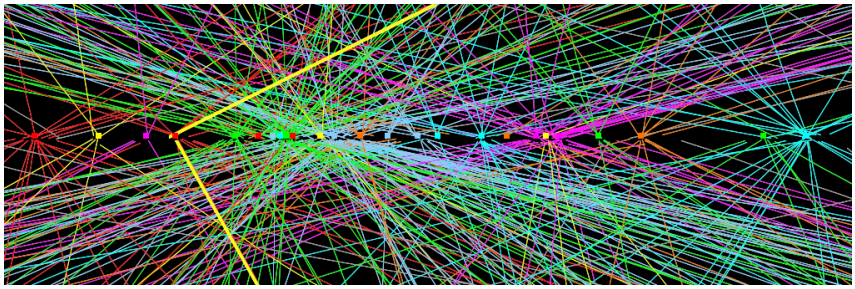
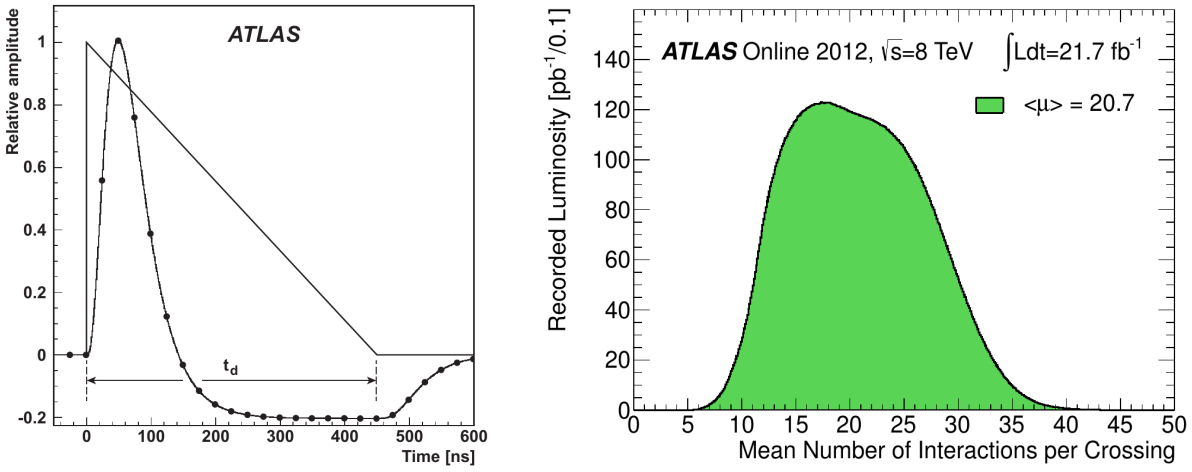


Figure 5.4.: A candidate Z$\rightarrow \mu\mu$ event with 25 reconstructed primary vertices together with inner detector tracks with a p_T above 400 MeV. Lines of a certain colour correspond to tracks reconstructed as stemming from the same vertex. The two thick yellow lines are the reconstructed tracks of the two muons from the Z decay [104]. The distance from the left to the right border of the image corresponds to 20 cm.

Besides the in-time pile-up, there is another sort of pile-up referred to as out-of-time pile-up. This is not such an obvious quantity as the in-time pile-up. The out-of-time pile-up reflects the sensitivity of the signal readout to the signal from previous bunch crossings.
Figure 5.5a shows a typical pulse shape of a calorimeter cell as it appears in the electromagnetic barrel. The original unipolar triangular pulse is generated by an ionizing particle and then shaped by a software algorithm to yield a bipolar shape. The charge collection time (or drift time) t_d amounts to $t_d \approx 450\,\mathrm{ns}$ and the shaped pulse needs around $t_{shaped} \approx 600\,\mathrm{ns}$ to return to the baseline. As t_{shaped} is much larger than the bunch crossing rate of one crossing per $t_{bc} = 50\,\mathrm{ns}$, a signal pulse that is read out can potentially contain the residual signal of $t_{shaped}/t_{bc} = 12$ earlier events. This effect is the so-called out-of-time (OOT) pile-up).
As a consequence of the integral of the shaped pulse amounting to zero, the average in-time pile-up is cancelled by the out-of-time pile up over time. However, a cell in a given event can still be in a negative signal state resulting from the previous bunch crossings when it is read out. Thus the cell exhibits a negative signal which can also lead to entire calorimeter clusters being negative. The on-average cancelling of in-time and out-of-time (OOT) pile-up only happens when the cell history is actually filled with the remainders of previous events. At the beginning of a bunch train this history is naturally empty and gets filled event by event until after 12 events, the out-of-time pile-up signal contribution becomes stable (except for event-to-event fluctuations).
A good measure for the out-of-time pile-up is the number of average interactions per bunch crossing μ which is evaluated as $\mu = \mathcal{L} \cdot \frac{\sigma_{inel}}{n_{bunch}} \cdot f_R$. Here, $\mathcal{L}$ is the instantaneous luminosity, σ_{inel} is the inelastic proton-proton cross section, n_{bunch} is the number of bunches in the collider and f_R is the LHC revolution frequency. This number is not a constant but rather depends on the



(a) (b)

Figure 5.5.: (a) Cell pulse shape in the liquid argon EMB calorimeter. The bipolar shaped pulse is superimposed to the original unipolar triangular pulse [100] (originally appeared in Ref. [105]). The black dots correspond to the times of successive bunch crossings at a rate of one crossing per 25 ns. (b) Luminosity-weighted distribution of the number of interactions per bunch crossing in the 2012 dataset [106].

instantaneous luminosity.
The μ distribution with the mean value of $\langle\mu\rangle = 20.7$ in 2012 data can be seen in Figure 5.5b. For the simulation of pile-up in MC, one simulated hard-scatter event is overlaid with a number of generated *minimum bias* events. Minimum bias events are very soft interaction processes, recorded using a rather inclusive trigger, therefore the selection represents an "unbiased" sample of collisions. The aforementioned number of minimum bias events is randomly drawn from the μ distribution from data taking, if this is already known. If the MC is produced before the actual data are taken, an attempt is made to anticipate the μ distribution as accurately as possible and the produced MC events will subsequently have to be reweighted to the true μ distribution once the final data are collected. The aforementioned effect of incomplete cancelling of the two types of pile-up at the beginning of a bunch train is accounted for by dividing simulated events into four bunch trains as well. Closely related to pile-up is the so called *underlying event* (UE). The UE commonly denotes all activity measured in the detector in a single pp collision which is not resulting from the primary hard scatter. This definition usually includes products from the collision of the proton remnants, initial and final state radiation, noise and multiparton interactions.

5.5. Jet Corrections

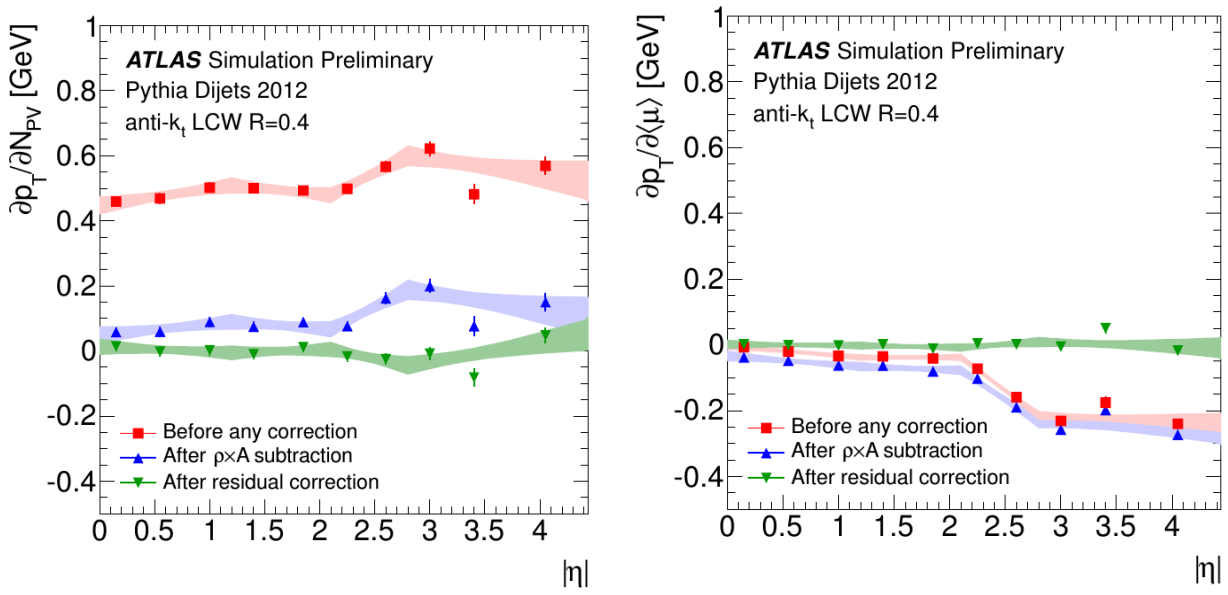
5.5.1. Jet Calibrations for small-R jets

Pile-up plays a major role in the reconstruction and calibration of jets in general. The situation for analysts is the following: There is a hard scatter event which contains the interesting signal, but this is accompanied by a number of "overlaid" pile-up events which can hardly be distinguished from the hard scatter. The pile-up is especially harmful for jets as they can receive contributions

originating from vertices other than the hard scatter vertex without the possibility of telling apart the two contributions. In a high pile-up environment every jet will thus receive a certain pile-up contamination, which in general increases with the jet's effective area.
Furthermore, jet reconstruction has to rely on somewhat arbitrary choices such as the jet radius parameter for example which introduces a certain layer of arbitrariness. In order for a jet to still represent the initial parton from which it originated as accurately as possible, several corrections are applied to jets, the most important of which will be briefly described here [107].
The first step is a jet area based pile-up correction. The idea is to predict the amount of pile-up contributing to a jet based on the average global pile-up activity in the event and to subtract this estimated contribution from the original jet. The pile-up estimation is performed by assessing the average p_T density given by $\rho = \text{median}\left\{p_{T,i}^{\text{jet}}/A_i^{\text{jet}}\right\}$ of the event, with the jet area A^{jet} and i running over all (inclusive) jets found with the k_T algorithm. The median is used instead of the mean here to reduce the influence of the hard scatter jets and obtain a more accurate description of the actual pile-up density. The expected average pile-up-p_T contribution to a given jet is then obtained as $p_T^{PU} = \rho \cdot A^{\text{jet}}$, since the jet area provides a good estimate of a jet's sensitivity to pile-up. This estimated pile-up contamination is then subtracted from the jet's measured transverse momentum.
As discussed in Section 5.4, out-of-time pile-up can lead to negative cell response that, when overlapping with the positive jet signal, will result in a negative modulation of the jet energy. This phenomenon can be regarded as a negative dependence of the jet momentum on out-of-time pile-up. In the forward region, the calorimeter granularity is enhanced in comparison to the central region which increases the probability of overlap with OOT pile-up. In addition the bipolar pulse shape is shorter here resulting in a larger negative amplitude. Consequently forward jets have a higher sensitivity to OOT pile-up, a behaviour which can be seen in Figure 5.6b.
This (negative) dependence is also not removed by the first pile-up correction. For this reason, and because of jets still showing a residual N_{PV} dependence after the first PU correction (see Figure 5.6a), a second PU correction, called residual correction, is applied. The correction is parametrized as function of N_{PV} and $\langle\mu\rangle$ (for in- and out-of-time pile-up, respectively) and obtained by comparing particle jets from simulation with reconstructed jets from simulation. The final pile-up corrected jet p_T is then obtained as

$$p_T^{\text{Corr}} = p_T - A \cdot \rho - p_T^{\text{Res. Corr.}}. \tag{5.7}$$

However, only changing a jet's p_T while leaving its energy unchanged would change the jet direction which is undesirable. Therefore the entire jet four-vector is scaled down by the ratio of $p_T^{\text{Corr}}/p_T^{\text{Orig}}$.
The next important correction is the jet energy scale (JES) correction [108]. This is a simple correction aiming to relate a jet's measured energy to the true energy of the original particles which initiated the jet. For this correction, the measured jet energy is multiplied with a factor obtained by comparing reconstructed jets with simulated particle jets. In order to bring back a jet's energy to the expectation from simulation, it has to be multiplied by the inverse of the average jet response $R = \langle E_{\text{Jet}}/E_{\text{Jet}}^{\text{simul.}}\rangle$. Also here the correction is applied to the entire jet four-vector in order not to change its direction.
The last correction is the residual in-situ calibration which is another multiplicative correction factor that is only applied to the jets in data, not to those in MC. This is done in order to account for effects present in data but not in simulation. The response is obtained by performing a p_T balance between jets and reference objects, where balancing can be between jets and Z, γ



(a) (b)

Figure 5.6.: (a) Dependence of the p_T of small-R jets built from local hadronic calibrated topological clusters on N_{PV}, before PU correction and after each of the two correction steps. While some residual dependence is still present after the jet area correction, it is brought to zero over the full η range after the residual correction. (b) Dependence of the same jets' p_T on $\langle\mu\rangle$, averaged over all N_{PV} bins. The negative dependence in the forward region is well visible as well as its suppression by the residual correction. [107].

or multijets for instance. A given jet's p_T in data is then scaled by the double ratio:

$$\frac{R_{\mathrm{MC}}}{R_{\mathrm{Data}}} = \frac{\langle p_T^{\mathrm{Jet}}/p_T^{\mathrm{Ref}}\rangle_{\mathrm{MC}}}{\langle p_T^{\mathrm{Jet}}/p_T^{\mathrm{Ref}}\rangle_{\mathrm{Data}}} \tag{5.8}$$

The total systematic uncertainty for the jet energy scale broken down into individual contributions can be seen in Figure 5.7.

5.5.2. Jet Vertex Fraction

Somehow orthogonal to the jet calibration methods that aim to correct a given jet for expected pile-up contributions is the *jet vertex fraction* (JVF) variable [107]. It can be used to reject complete jets that are most likely stemming from PU. This pile-up suppression method is only available for small-R jets since large-R jets are unlikely to be produced by pile-up only (especially those with high p_T–a common analysis requirement for large-R jets). Usually the primary vertex in the event with the highest sum of the p_T^2 of all tracks associated to this vertex is considered the hard scatter vertex and the other primary vertices are considered to be pile-up vertices. The JVF is a variable designed to quantify the probability that a given jet emerged from a particular primary vertex v. This is achieved by comparing the p_T of ID tracks stemming from v associated

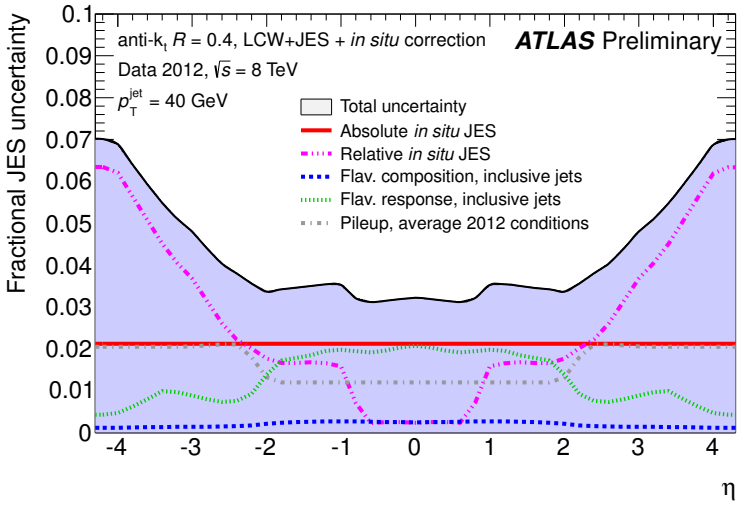


Figure 5.7.: Fractional jet energy scale systematic uncertainty components as function of the jet's pseudorapidity. The plot is obtained for the same small-R jets as used in this thesis: Jets reconstructed with the anti-k_T algorithm using a radius parameter of $R = 0.4$. Inputs to jet reconstruction are calorimeter clusters calibrated with the local hadronic calibration. The total uncertainty is shown as filled blue region. [109]

to this jet with the p_T of all tracks associated to the jet:

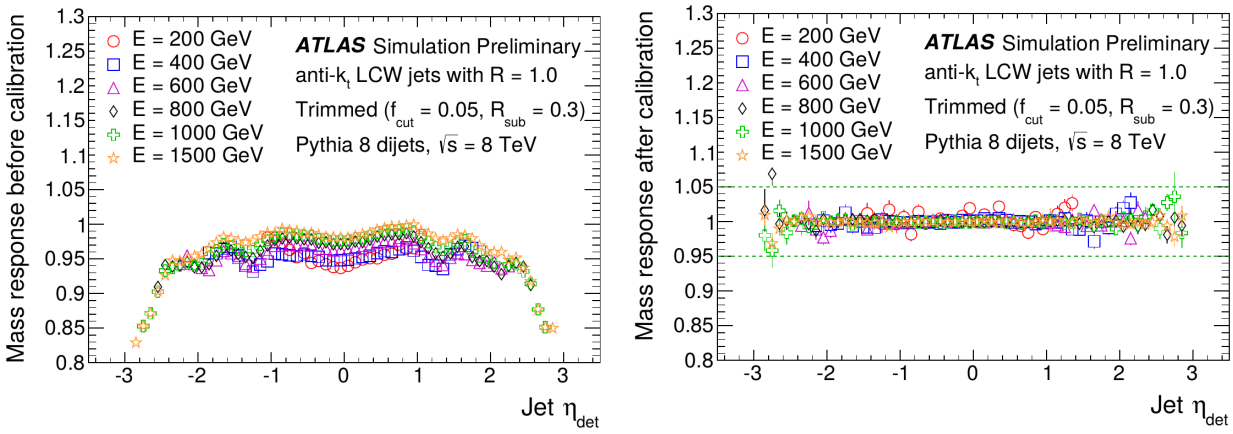
$$\mathrm{JVF}(\mathrm{jet}_i, v) = \sum_{j_{\mathrm{track}} \in v \text{ and jet}} p_T^{j_{\mathrm{track}}} / \sum_{j_{\mathrm{track}} \in \mathrm{jet}} p_T^{j_{\mathrm{track}}}, \tag{5.9}$$

where j runs over all reconstructed primary vertices in the event. The JVF can be calculated for each PV in the event but is useful only if calculated for the hard scatter vertex. It is bound to lie between 0 and 1, with a value of 0 indicating that the jet does not contain any track from the hard scatter vertex and is therefore designated as pile-up jet. A value of 1 signifies that the entire jet-p_T is made up by tracks from the hard scatter vertex and the jet is likely a signal jet. Jets without associated tracks are assigned a JVF of -1. Because the calculation of the JVF needs as input the ID tracks, it is only defined within the tracking acceptance range, i.e. in $|\eta| < 2.5$.

5.5.3. Jet Calibrations for large-R Jets

The calibration procedure applied to large-R jets is slightly different than for small-R jets [110, 111]. The calibration of the JES is derived in the same way as for the standard JES calibration for small-R jets [108]. Calibration factors are derived comparing reconstructed jets with simulated particle jets from a Pythia di-jet sample.
In contrast to small-R jet calibration, no explicit pile-up correction, neither depending on $\langle\mu\rangle$ nor on N_{PV} is applied, since here pile-up mitigation is already achieved by the grooming procedure (see Section 5.6).
An additional correction step is applied to large-R jets which is not done in the standard small-R jet calibration, that is an explicit jet mass calibration procedure. Since one of the major goals of working with large-R jets is to accurately reconstruct their invariant mass which is supposed to be the mass of the decaying particle, this calibration is of larger importance here than for small-R jets. Furthermore, the large-R jet mass is particularly sensitive to soft wide-angle contributions

that have a very small impact on the jet energy but considerably alter its mass.
The jet mass response ($m^{\text{Jet}}_{\text{true}}/m^{\text{Jet}}_{\text{reco}}$) is therefore evaluated on MC in bins of jet energy and η and its inverse is applied as calibration factor to reconstructed jets. Figure 5.8 shows the reconstructed trimmed (see Section 5.6.2) large-R jet mass (here for a subjet radius parameter used in the trimming procedure of $R_{\text{sub}} = 0.3$) before and after calibration. The response before calibration is found to deviate strongly from one even at high p_T and in the central part of the detector, which can be explained by pile-up and noise contributions. After the calibration factors are applied, good closure is obtained.



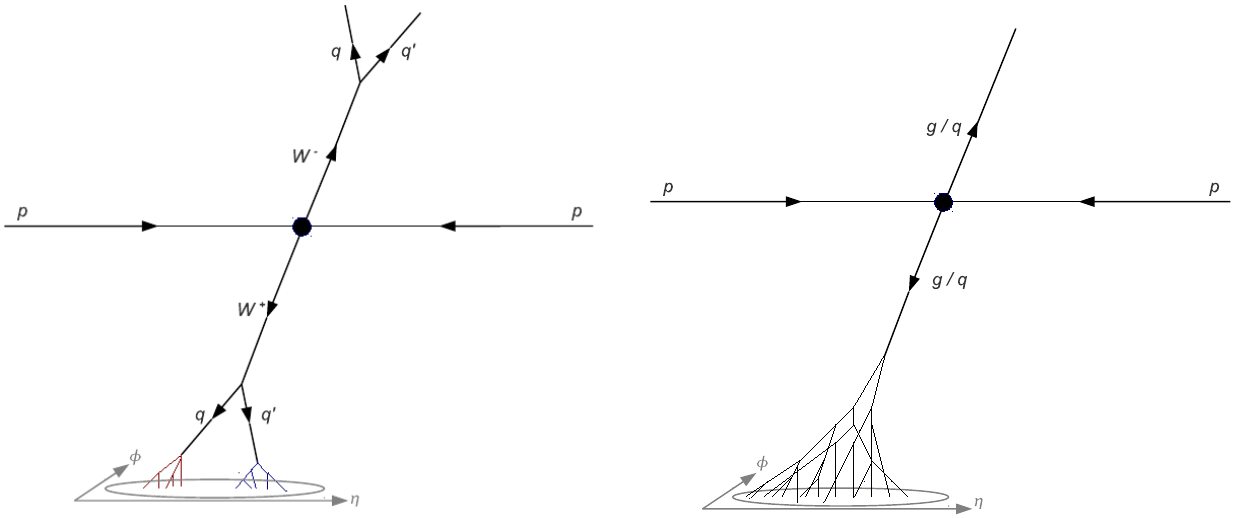
(a) (b)

Figure 5.8.: (a) Jet mass response for different jet energy bins before the calibration is applied. The response drops towards higher η values due to the reduced calorimeter granularity in this region. (b) After calibration the response can be seen to be consistent with one with a largest deviation of around 4 % for the entire energy and η range. [111]

5.6. Jet Grooming

Since the pile-up contamination of a jet depends directly on the area it covers, pile-up plays a much larger role for large-R jets with a radius of the order of $R \sim 1.0$ than for a small-R jets with $R \sim 0.4$ ($A^{\text{large}} = 16 \cdot A^{\text{small}}$) which are relatively robust under pile-up and can be corrected with the methods described in Section 5.5.1. In order to still be able to use the large-R jets, methods had to be developed which aim to actively dispose of the pile-up contribution in the jet while retaining as much of the true signal as possible. The generic term for these kinds of methods is *jet grooming*-methods all of which try to exploit the difference in topology and energy of pile-up contributions and signals from the hard scatter. It is assumed that pile-up primarily consists of soft particles, dispersed approximately uniformly over a large area in the detector. Jets resulting from a hard scatter, however, tend to consist of a small number of very hard, localized constituents that are covering only a small area. In case that a particle radiates a gluon for example, there can be further contributions at small angles to the hard, "main" jet, but they will typically also be hard compared to pile-up contributions.
A large-R jet resulting from a boosted W boson decay for example is expected to consist of two hard subjets – one resulting from each quark, which are relatively symmetric in energy (each quark carries approximately half of the W boson p_T) and little further activity between the two subjets. The mass of the large-R jet (which should amount to 80 GeV) mainly arises from the

high energy of the two subjets, rather than from the angle between them. In contrast to this,



(a) (b)

Figure 5.9.: (a) Internal structure of a W induced large-R jet. The jet mainly consists of two localized hard sub-contributions with little activity in between. (b) A typical QCD induced jet consists of many soft, large angle sub-contributions. [112]

a large-R jet with a comparable mass, resulting from a parton produced in a QCD event has a very different internal structure. Since the original partons carry a colour charge, there will be a colour flow in between two particles in the shower, resulting in many soft, but large angle splittings. It is by these large angles between the individual constituents that a QCD jet acquires its mass and only to a minor extent by the energy of its constituents (see Equation 5.4).
Figure 5.9 shows an illustration of the internal structure of a typical jet resulting from a boosted W boson and a typical quark or gluon jet. Since the pile-up is predominantly consisting of such soft QCD events, the difference in the internal energy flow pattern of the two jet classes can be exploited in order to:

1. Remove pile-up contributions inside a signal jet (jet grooming)
2. Tag jets as either QCD jets or boosted heavy particle jets (jet substructure)

Usually jets have to undergo a certain grooming procedure after which the substructure of the remaining jet is examined in order to draw further conclusions on the nature of the primary particles (partons, gauge bosons) which initiated the jet. The three most important grooming methods will briefly be explained in the following.

5.6.1. Filtering and Mass-Drop Tagging

Filtering is a very simple but efficient technique to extract only the hard subjets from a large-R jet [113]. In principle, it is also possible to apply filtering to a jet without previous mass-drop tagging but this is done rarely, usually the two techniques go together, also referred to as mass-drop filtering. The general way the mass-drop filtering works is by trying to isolate symmetric concentrations of energy inside a large-R jet where each of the two subjets has a significantly smaller mass than the complete jet. The use of filtering together with mass drop tagging will be

referred to simply as filtering from this point on.
The mass-drop technique is illustrated in Figure 5.10 and it can be used to find out whether a large-R jet is consistent with the hypothesis that it consists of two or more hard subjets. This technique already gathers information concerning the jet substructure and is not considered as grooming but rather as tagging already. The assumption is that during clustering of a large-R jet resulting from a boosted heavy particle, at some point the algorithm must have merged the two hard subjets from the two quarks. After this merging step, the mass of the jet will increase dramatically with respect to to the previous step.
The mass-drop tagger seeks to walk this way exactly backwards by unclustering the large-R jet according to its original history until it finds a step where the mass significantly drops in a de-clustering step. If this is found and the splitting between the two subjets j_1 and j_2 found in this step is sufficiently symmetric in energy, the jet is tagged as stemming from a boosted heavy particle. If no such mass-drop is found or if it is very asymmetric, the jet is considered to have originated from a QCD process and is discarded.
Mass drop tagging must not be applied to jets clustered with the anti-k_T algorithm since its history is not QCD motivated (starting with the hardest constituent and adding neighbouring soft contributions). Therefore the de-clustering is not expected to lead to the step where the two hard subjets were merged originally.

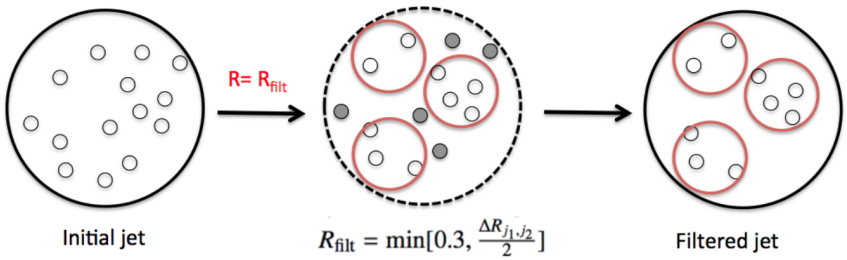


Figure 5.10.: Illustration of the filtering method [114]. A jet is reclustered with a smaller jet radius parameter R_{filt} and the three hardest subjets are chosen to make up the final filtered large-R jet.

In case that the jet passes the tagger, it proceeds to the filtering stage which means all of its original constituents are reclustered, but with a much smaller radius parameter. This is motivated by the ambition to keep only the hard scatter subjets discarding the soft, large angle contributions from pile-up which are assumed to be uniformly scattered below the signal. A typical choice of R_{filt} is

$$R_{\text{filt}} = \min(0.3, \frac{\Delta R(j_1, j_2)}{2}). \tag{5.10}$$

By this procedure, only the hard components of the jet are clustered and the soft part from pile-up is left aside. Since an inclusive jet algorithm will cluster all of the original constituents into jets (mostly with low p_T), a cut has to be placed on the number of hardest subjets to be counted as a filtered large-R jet. Usually the two highest p_T jets are selected, or the hardest three if potential radiation of a hard parton are to be included.

5.6.2. Trimming

Trimming is another important jet grooming method and is closely related to filtering which can be seen in Figure 5.11. The trimming method is essentially a jet reclustering with a smaller jet

radius and a veto on soft contributions [115] and is the method of choice for this analysis. The idea behind this procedure is to take advantage of the fact that contributions from pile-up or the UE are usually much softer than the products of the hard scatter and associated final-state radiation. Therefore soft contributions inside a large-R jet are vetoed while high energy subjets are kept. Preferentially the pile-up is removed from the the large-R jet by trimming but due to spatial overlap with the hard scatter, it is also possible that a fraction of the hard scatter signal gets rejected. This share of energy has to be recovered by a calibration procedure later on.
Trimming typically reduces the mass of light quark or gluon induced jets with a mass below 100 GeV by 30-50 % [114]. The fraction is much smaller for jets induced by boosted heavy bosons (see Figure 5.13b).

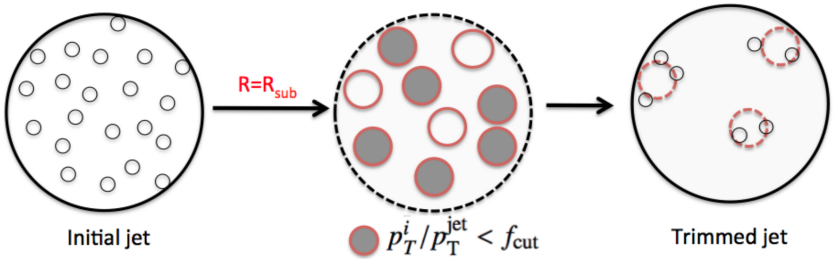


Figure 5.11.: Illustration of jet trimming [114]. The large-R jet is reclustered with a smaller radius parameter R_{sub} and all subjets satisfying the p_T selection criterion are added to one trimmed jet.

As in the filtering method, the entire constituents of an initial large-R jet are reclustered with a much smaller radius parameter R_{sub}. The reclustering algorithm does not necessarily have to be the same was initially used for clustering of the large-R jet. Each resulting subjet i is considered separately and it is checked if its momentum fulfills the criterion: $p_T^i > f_{\text{cut}} \cdot p_T^{jet}$, where f_{cut} is a fixed parameter. The subjet is kept if such a condition is satisfied and is discarded otherwise. The final trimmed jet is then obtained by adding all kept subjets.
During the development of the analysis, filtering was compared to other jet grooming methods and proved to be the most efficient (see Section 5.6.4). Particularly advantageous is the fact that the trimmed background jets receive a very small mass and do not tend to peak in the signal region (jet mass between 65 and 95 GeV) which is the case for mass-drop tagged and filtered jets. This behaviour can be observed in Figure 5.13 and 5.14.

5.6.3. Pruning

As in the case of the two other grooming methods, jet pruning is an "outside-in" method [116], meaning it starts with an already clustered jet which is then reclustered using a subjet finding scheme. A schematic illustration of the pruning method is given in Figure 5.12. Contrary to the other two described methods it does not reject entire subjets but rather single constituents during the clustering process, therefore it does not rely on the identification of individual subjets. It resembles the trimming algorithm in the sense that low energy contributions are rejected while it differs from trimming in the sense that the pruning procedure additionally vetoes wide angle radiation.
For this, two parameters z_{cut} and d_{cut}, corresponding to a minimum p_T fraction and a maximum allowed angle between two constituents, respectively, have to be fixed.

When the chosen jet clustering algorithm is applied to the constituents, it is checked in each recombination step $(i, j \to k)$ whether

$$\Delta R(i,j) > d_{\text{cut}} \text{ and } \min(p_T^i, p_T^j)/p_t^k < z_{\text{cut}}. \tag{5.11}$$

If the two conditions are met, the recombination is not carried out and the softer one of the two subjet candidates is rejected. Otherwise, the jet algorithm adds the next potential jet constituent and the procedure starts anew. In this way the pruned jet is built as a large-R jet where expected soft, large angle constituents from pile-up or the underlying event are "pruned" away.

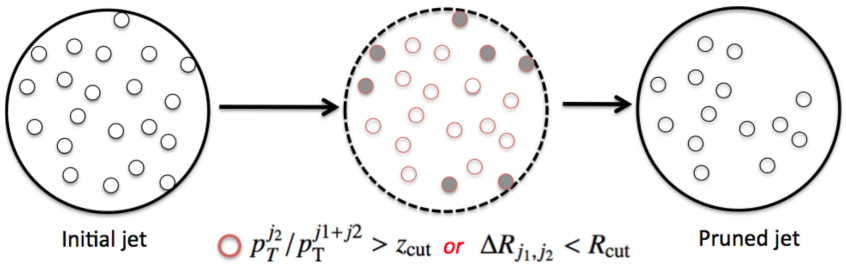


Figure 5.12.: Illustration of the pruning method [114]. Soft, large angle contributions are vetoed in the reclustering process.

5.6.4. Comparison of Grooming Methods

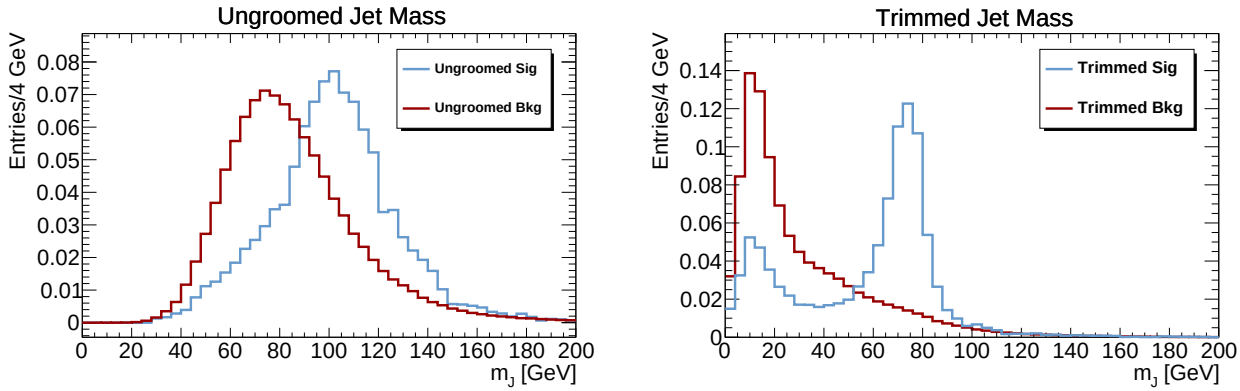
The different grooming methods together with the parameters they depend on were tested thoroughly in Ref. [117]. However, the signal and background composition as well as the investigated phase space in this publication differed from those in the present analysis.
In Ref. [117], the performance comparison of grooming method was based on the direct comparison of QCD-induced and W-induced large-R jets. In order to account for correlations between the W boson p_T and jet substructure features, events were pre-selected by running the C/A algorithm with $R = 1.2$ on stable simulated particles and requiring an ungroomed C/A jet in a p_T range of $200\,\text{GeV} < p_T^{\text{jet}} < 350\,\text{GeV}$ and within an η range of $|\eta| < 1.2$.
Such a selection differs from the selection in this analysis. The background for the presented WW+WZ cross section measurement only consists to a very small fraction of QCD multijet events (see Section 7.1). However, the major background for the analysis presented here is W+jets where the large-R jets originate from the QCD part in the event and the W boson decays leptonically. In that sense the large-R jets here and the jets in Ref. [117] become comparable. Nonetheless, a significant fraction (11%, see Table 10.2) of the background investigated here is created by top quark events that can contain a genuine hadronic W boson. The groomed large-R jets from a true W boson can potentially distort the shape of the groomed background jets with respect to QCD-only background.
Furthermore, the p_T cut used in Ref. [117] is applied on ungroomed jets from simulated particles while in this analysis the cut is placed on groomed jets built from calorimeter clusters.
Because of those differences, a study was carried out again in order to check which of the three grooming methods is most suitable for the purpose of the boosted WW+WZ cross section measurement.
For better comparison, the anti-k_T algorithm was the large-R jet algorithm employed for all grooming methods. This algorithm was found to perform more robustly under pile-up conditions than other jet algorithms at a comparable W-tagging efficiency [117].

Mass-drop tagging cannot be applied to jets clustered with the anti-k_T algorithm. This is due to the fact that the history of the jet clustering is used to uncluster the jet successively in this method. Since the C/A algorithm only uses the spatial distance between constituents and does not depend on their p_T, the unclustering is well defined and infrared and collinear safe. A similar argument also holds for the k_T algorithm whose clustering history is also motivated by QCD fragmentation. This is not the case for the anti-k_T algorithm. In order to still be able to compare mass-drop tagging to the other grooming methods for the same jets, the anti-k_T jets are re-reconstructed with the C/A algorithm such that it obtains a QCD compatible history.
For this, the constituents of the large-R anti-k_T jet are given as input to an exclusive C/A algorithm with a radius parameter of $R = 2$ and a requirement that it returns exactly one jet is applied. In this way a new large-R jet is obtained, kinematically and topologically equivalent to the initial jet, only with a C/A instead of an anti-k_T history and thus suitable for mass-drop tagging.
In searching for the optimal grooming method, the individual methods are tested on exactly those events that passed the analysis event selection which will be introduced in detail in Chapter 6, especially because the background does not consist of only one particular type of process but rather of a mixture of different event classes. The decision on a certain grooming method is therefore based on the performance in precisely the investigated phase space and background composition present in this analysis.
The event selection has been applied here up to the cut on the $E_{\mathrm{T}}^{\mathrm{Miss}}$. This is the stage at which the large-R jet is selected and the respective grooming method is applied.
For the comparison of the different techniques, exactly one ungroomed large-R jet with a p_T of greater than 200 GeV within $|\eta| < 2.0$ and which is not overlapping with a selected electron is required. Furthermore, the default small-jet veto is applied, meaning all events containing a small-R jet with a $\Delta R > 1$ to the large-R jet are discarded. This is the baseline selection for all of the three grooming methods.
Then the ungroomed large-R jet is treated with either one of the three grooming methods using the parameters that proved best in [117] and which are shown in Table 5.1 for a comparable selection. A common subjet algorithm (k_T) and subjet radius ($R = 0.2$) is chosen for all test cases for best comparability. The final groomed jet still has to pass a boost requirement of having a p_T of more than 200 GeV. Since calibrations are not available for all kinds of the large-R jets investigated here, all large-R jets used for this study are left uncalibrated.

Grooming Method	**Parameter 1**	**Parameter 2**
Trimming	$f_{cut} = 5\%$	–
Filtering	$\mu = 1$	$y = 0.09$
Pruning	$z_{cut} = 15\%$	$d_{cut} = 0.5$

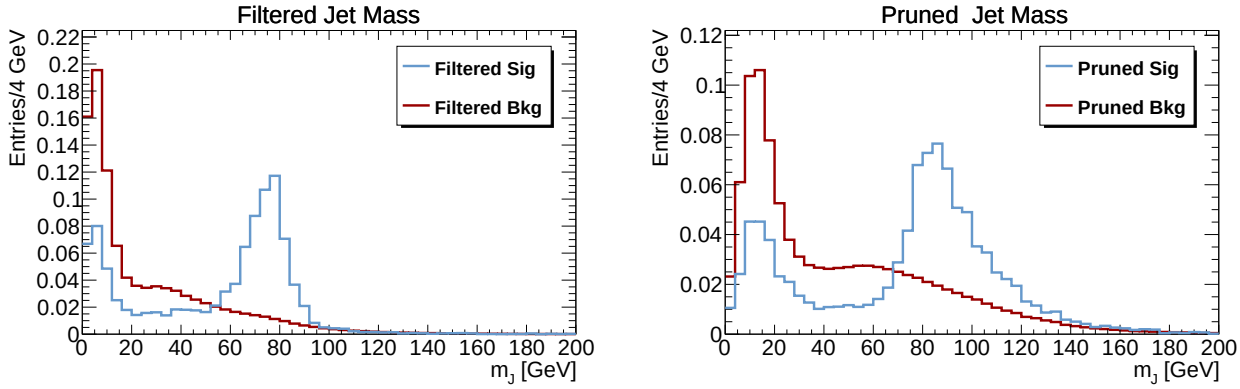
Table 5.1.: Summary of parameter values used for the grooming methods.

Figure 5.14b shows that pruning is the least suitable grooming method since the signal peak is very broad, extending up to far more then 100 GeV and additionally a large fraction of the background shows up in the signal region between 60 and 100 GeV. It also has a shoulder at 60 GeV which would make it difficult to distinguish it from the signal in a fit of the two templates to a data distribution. Table 5.2 also indicates that the signal to background ratio here is by far the worst of all grooming methods.
The performance of trimming and filtering is comparable. With filtering a slightly higher signal to background ratio is achieved. However, when looking at the background shapes for the two methods, in the trimming case a smoothly falling distribution is observed while the filtered background distribution has a shoulder at around 30 GeV and in the signal region is not as



(a) (b)

Figure 5.13.: (a) Mass of the ungroomed large-R jet for signal and background, the large-R jets receive a large pile-up contamination. Without pile-up, the large-R jets from the signal would be expected to peak at a mass of around 80-90 GeV (W and Z bosons). (b) Mass of the trimmed jet for signal and background. The background now peaks at very low masses and exhibits a smoothly falling distribution. The signal is peaking slightly below 80 GeV and also in a second peak at around 10 GeV. The fact that the main peak is below 80 GeV is due to the trimming procedure also inevitably removing a part of the actual signal.



(a) (b)

Figure 5.14.: (a) Mass of the filtered large-R jet for signal and background. The filtered large-R jet is chosen to include the three hardest subjets found in the reclustering procedure in order to include possible jets from final-state radiation. The resolution of the signal peak is larger than for trimmed jets. Furthermore, the background shape exhibits a shoulder at around 30 GeV.
(b) Mass of the pruned jet for signal and background. The signal distribution is rather broad and has a huge tail on the high mass end. Moreover, the background shows a shoulder which is even closer to the signal peak than in the filtering case.

smoothly falling as the trimmed distribution. Furthermore, the resolution is improved for trimmed in comparison to filtered jets. These considerations suggest that a potential signal is better identifiable for trimmed than for filtered jets which is why for this analysis it was decided to use trimming as default grooming method.

Grooming Method	S/B in %
No Grooming	2.10
Trimming	8.34
Filtering	9.44
Pruning	4.76

Table 5.2.: Expected signal over background ratio in the signal region of groomed jet mass between 65 and 95 GeV. Numbers are obtained from the same plots as shown in Figures 5.13 and 5.14, normalized to 2012 integrated luminosity.

It is interesting to note that all grooming methods produce a second signal peak at low jet mass values. This is due to the grooming being too aggressive such that one of the two hard subjets is discarded, resulting in a low mass of the groomed jet.
Another observation common to all grooming methods is the fact that the groomed signal peak is below 80 GeV even though the signal jets are a mixture of W and Z jets which should on average have a mass between 80 and 90 GeV. The lower value found here is a result of all grooming methods inevitably also removing parts of the actual signal. For this study this fact was accepted, also in order to evaluate which of the methods is grooming too strictly. However, in the actual cross section measurement, a dedicated calibration for the used trimmed anti-k_T jets which corrects for this "overgrooming" is applied (see Section 5.5.3).

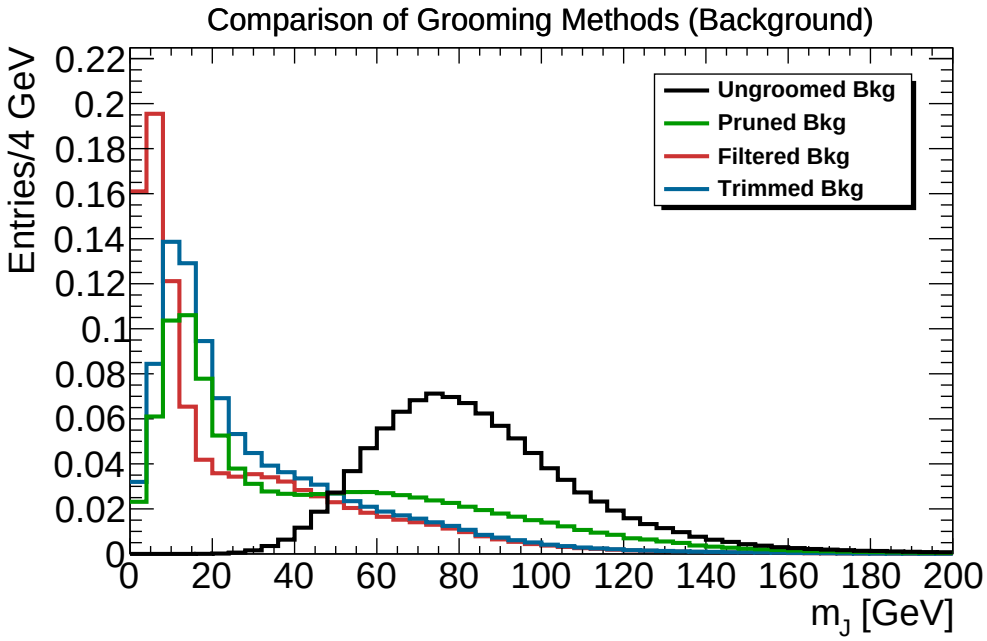


Figure 5.15.: The effect of different grooming methods on the background mixture present in this analysis is shown. The ungroomed large-R jets have a large mass of around 80 GeV on average. Since most of them are QCD jets, grooming greatly reduces their masses.

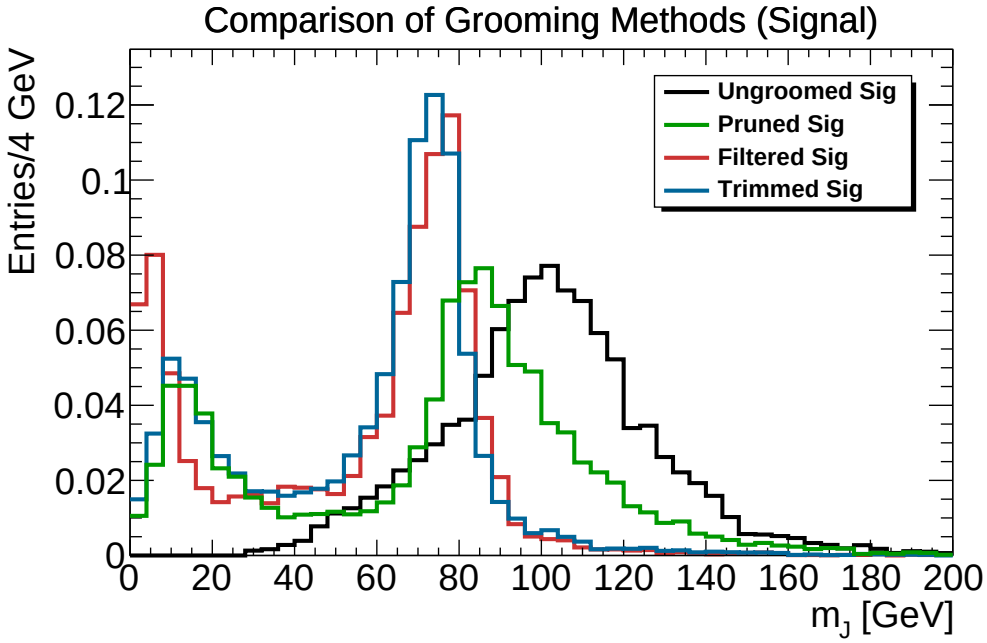


Figure 5.16.: The signal shape of ungroomed large-R jets is compared to jets with all different grooming methods. All grooming methods show a two-peak structure, one peak at low masses and one at the W/Z mass. The peak close to zero results from cases where the grooming is too aggressive, i.e. one of the two hard subjets is not selected in the grooming. Trimming and filtering exhibit a similar performance while pruning is clearly performing worst.

5.7. Jet Substructure Variables

After a jet is groomed, its substructure still contains valuable information about its nature. A groomed QCD jet and a groomed W-induced jet with comparable kinematics still exhibit a rather different internal energy distribution which can be exploited to tag a boosted boson jet and suppress backgrounds from QCD.
The expectation is that a boosted boson jet consists of two hard and narrowly localized subjets with a symmetric energy share and little activity in between while a groomed QCD jet is built of many softer and more asymmetric far-spread contributions.
A rich set of substructure variables has been invented in recent years in order to exploit these differences and a large number of these variables have been tested for their separating power for this analysis. It became apparent that two classes of variables were most promising. The first is N-subjettiness and the second is the a class of energy-correlation variables, each of which will now be described.

5.7.1. N-Subjettiness

One set of discriminating variables relate to a quantity referred to as N-subjettiness [112]. The idea of the N-subjettiness is to quantify to what a degree a large-R jet is compatible with the assumption of being composed of N (or fewer) subjets. This is done by identifying a certain number of subjets inside a large-R jet and checking how well the large-R jet can be described by consisting of exactly these subjets.
For this, these variables need an explicit identification of subjets inside the large-R jet which introduces some layer of arbitrariness.
The subjet identification is usually achieved by reclustering the jet with some exclusive algorithm, which is forced to return exactly N subjets.
The N-subjettiness can be defined *recoil-free* meaning it is insensitive to a possible displacement of the hardest jet constituent with respect to the jet momentum axis due to low energy, large angle radiation. This can be achieved by shifting the subjet axes found by the exclusive algorithm from the jet-axis to the direction of the hardest constituent. The axes found with this procedure are called "winner takes all (wta)" axes in the literature and the N-subjettiness calculated with this axis definition showed an increased performance with respect to the conventional axes definition [118, 119]. For this reason the wta-axes definition has been used throughout the following studies.
After having fixed exactly N subjet directions, the N-subjettiness is computed in the following way:

$$\tau_N^{(\beta)} = \sum_{k \in J} p_{T_k} \min\left(\Delta R_{1,k}^{\beta}, \Delta R_{2,k}^{\beta}, ..., \Delta R_{N,k}^{\beta}\right), \tag{5.12}$$

with k running over all constituents of the original large-R jet and $\Delta R_{i,k}$ being the spatial distance between subjet axis of jet i and the constituent k. The above equation simplifies for the cases investigated in this analysis ($\beta = 1, N = 1, 2$) to:

$$\tau_1 = \sum_{k \in J} p_{T_k} \Delta R_{1,k}, \tag{5.13}$$

$$\tau_2 = \sum_{k \in J} p_{T_k} \min(\Delta R_{1,k}, \Delta R_{2,k}). \tag{5.14}$$

If a jet is mainly composed of N or fewer subjets, τ_N stays small since the ΔR factor is small in each summand. Instead, if the subjet consists of N+1 or more subjets the constituents of the additional subjet lying far away from the N other subjets it will give large contributions to τ_N. So the general observation is, that for an N-prong case τ_N will be small and τ_{N-1} will be large.

5.7.2. Energy Correlation Functions

Another type of discriminating variables are the so-called energy correlation functions. Simplified it can be said that energy correlation functions look at the energies and pair-wise angles between jet constituents in order to identify a present N-prong substructure inside a given jet. These types of variables have proven to be very powerful probes of jet substructure, which combine several advantages compared to other substructure variables [120, 121]. First of all, no explicit identification of subjets is necessary, allowing a very general application without introducing a source of arbitrariness by fixing particular subjets. Secondly, these variables are recoil-free by definition, similar to the N-subjettiness employing the wta axes definition. The insensitivity to recoil is a consequence of all angles entering the energy correlation functions being pair wise angles such that the hardest constituent does not play a special role in their calculation. Because of being recoil-free, these variables do not show a decrease in separation power in such recoil-affected topological configurations [118, 119].
Finally, energy correlation functions showed the best background suppression at a particular signal efficiency for this analysis which is shown in Figure 5.20 and 5.24a.
The general form of the energy correlation functions for a jet composed of N constituents is the N-point correlator:

$$E_{CFN}(\beta) = \sum_{i_1<i_2<\ldots<i_N\in J} \left(\prod_{a=1}^{N} p_{T_{ia}}\right)\left(\prod_{b=1}^{N-1}\prod_{c=b+1}^{N} \Delta R_{i_b i_c}\right)^{\beta}, \tag{5.15}$$

which, for the first four correlation functions reads explicitly:

$$E_{CF0}(\beta) = 1, \tag{5.16}$$

$$E_{CF1}(\beta) = \sum_{i\in J} p_{T_i}, \tag{5.17}$$

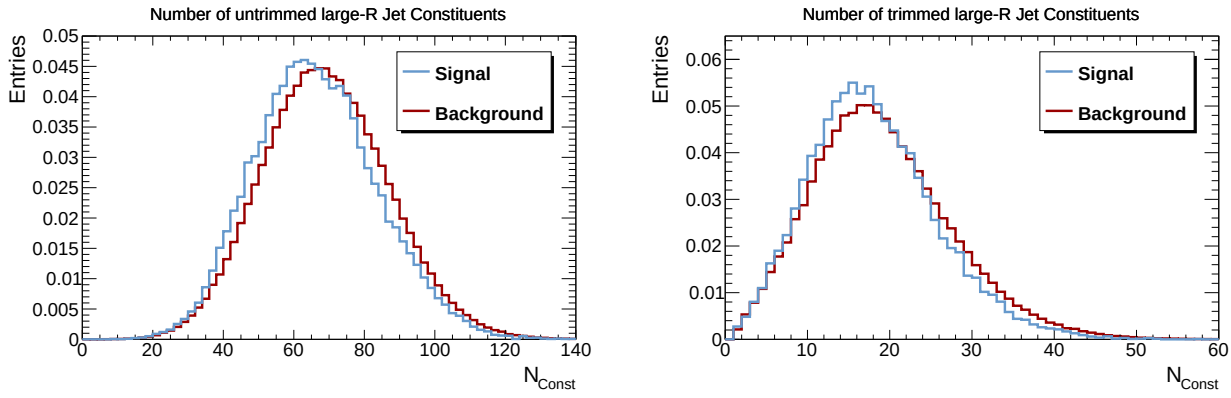
$$E_{CF2}(\beta) = \sum_{i<j\in J} p_{T_i} p_{T_j} (\Delta R_{ij})^{\beta}, \tag{5.18}$$

$$E_{CF3}(\beta) = \sum_{i<j<k\in J} p_{T_i} p_{T_j} p_{T_k} (\Delta R_{ij}\Delta R_{ik}\Delta R_{jk})^{\beta}. \tag{5.19}$$

The important property of these functions is the following: For a jet consisting of N subjets, the (N+1)-point correlator will be much smaller than the N-point correlator. In fact, if there were only exactly N particles observable, the (N+1)-point correlator would be exactly zero. However, in reality, the studied jets only rarely consist of a small number of constituents. The average number of constituents of a trimmed W-induced large-R jet is on the order of 18 (see Figure 5.17b). For this reason the (N+1)-point correlator for an N-prong large-R jet is not expected to yield zero but a small finite value.
In order to exploit this property, a dimensionless set of ratios of the $E_{CFN}(\beta)$ can be constructed:

$$e_2^{(\beta)} = \frac{E_{CF2}(\beta)}{E_{CF1}(\beta)^2} = \frac{1}{p_{T_J}^2} \sum_{i<j\in J} p_{T_i} p_{T_j} (\Delta R_{ij})^{\beta}, \tag{5.20}$$

$$e_3^{(\beta)} = \frac{E_{CF3}(\beta)}{E_{CF1}(\beta)^3} = \frac{1}{p_{T_J}^3} \sum_{i<j<k\in J} p_{T_i} p_{T_j} p_{T_k} (\Delta R_{ij}\Delta R_{ik}\Delta R_{jk})^{\beta}. \tag{5.21}$$



(a) (b)

Figure 5.17.: Number of constituents of the untrimmed (a) and trimmed (b) large-R jet for signal and background. As expected, the (mainly gluon) background jets exhibit on average a larger number of constituents than the boosted boson jets. The number of constituents of the trimmed jet is in both cases well above three so that the 3 point correlator never evaluates exactly to zero. In both cases the difference between signal and background is extremely small – a phenomenon which is also observed for the jet substructure variables (see Section 5.8).

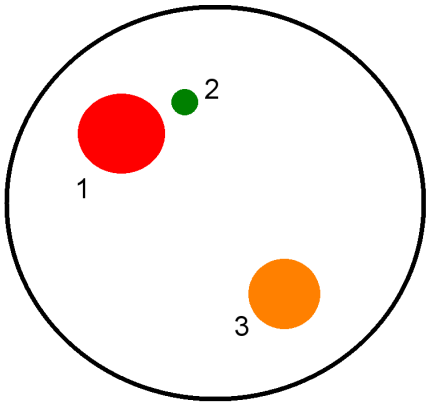


Figure 5.18.: Illustration of an approximately two-prong jet. The three coloured circles represent subjets, their size corresponding to their p_T. Assuming $\beta = 1$, the 2- and 3-point correlator evaluate to:
$E_{CF2} = p_{T_1}p_{T_2}\Delta R_{12} + p_{T_2}p_{T_3}\Delta R_{23} + p_{T_1}p_{T_3}\Delta R_{13}$ and $E_{CF3} = p_{T_1}p_{T_2}p_{T_3}\Delta R_{12}\Delta R_{23}\Delta R_{13}$.
The first term in E_{CF2} is small because of the small angular distance and the low p_T constituent while the second and third term are increasingly large in value leading to a large result for E_{CF2}. In E_{CF3}, however, p_{T_2} and ΔR_{12} contribute as multiplicative factor, leading to a small overall product.

Here the approximation was used that E_{CF1}, the scalar sum of the p_T of the individual constituents, is equal to the total transverse jet momentum [121]. These are the 2- and 3-point correlators normalized to the p_T of the original jet and could already be used as substructure variable. Motivated by the observation that N-subjettiness ratios yield a better discrimination power than the plain N-subjettiness variables on their own (see Section 5.7.1), it was found in [120] and [118] that superior discriminating variables can be constructed by using the double ratios of e_3 and e_2:

$$C_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^2}, \tag{5.22}$$

$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}. \tag{5.23}$$

$C_2^{(\beta)}$ and $D_2^{(\beta)}$ should yield significantly smaller values for 2-prong jets from a boosted heavy particle than for a QCD jets with a much larger number of prongs. The parameter β can take all values $\beta > 0$, however, only values of $\beta = 1$ were studied here and therefore in the following, the explicit dependence on β will be omitted.

5.8. Performance Comparison

The above mentioned substructure variables have, similarly to the grooming methods, been tested in a general way concerning their discrimination power between decaying boosted heavy particles and QCD jets [117]. They have also been re-tested for this thesis, specifically for the signal and background composition and the kinematic phase space present in this analysis.
Since for the WV cross section measurement it is important to have a signal to background ratio which is as high as possible in the signal region ($60\,\text{GeV} < m_{\text{large-R jet}} < 100\,\text{GeV}$), the discrimination power of the investigated variables in this region is of particular importance.
Therefore the study is divided into two parts, the first part considering the performance for the inclusive signal and background samples ($m_{\text{large-R jet}} > 50\,\text{GeV}$) and the second one only looking at the signal mass window of ($60\,\text{GeV} < m_{\text{large-R jet}} < 100\,\text{GeV}$).
Plots shown here are obtained for the one large-R jet present in the event, after applying the complete event selection but before applying any cut on a substructure variable.
In order to be able to establish a ranking between classifiers, a highly useful tool are receiver-operator-characteristics (ROC) [122] curves. These types of curves directly show the background suppression rates that are associated to the corresponding signal efficiency for the whole range of possible cut values on the classifiers.
During the course of this analysis, it revealed itself that the final sensitivity of the measurement actually improves when no substructure cut is applied, more details can be found in Section 5.8.3. Since this is a new and somewhat unexpected result that could prove useful for future analyses, the performance of the substructure variables will nevertheless be presented in the following.
In general, rather small differences between signal and background are observed (compare Figure 5.17 and 5.19 for example). This observation hints to a potential gain in discrimination power by combining several substructure variables into one multivariate method. This option was investigated for this analysis as well but was abandoned due to considerations concerning systematic uncertainties (see Section 11.2).

5.8.1. Performance in entire large-R Jet Mass Range

The performance of the individual N-subjettiness variables is found to be inferior than the energy correlation functions, see Appendix B for more details. The ROC curve in Figure 5.20 exhibits that the separation power of τ_2 is better than that of τ_1. This result is not unexpected since τ_N measures to what a degree a large-R jet is compatible with the assumption of being composed of exactly N subjets. For N=1, this is neither expected for QCD jets (many soft subjets) nor for boosted boson jets (two hard subjets). For this reason τ_2 performs better since gluon jets still do not fulfil the assumption (two subjets) but the jets induces by W or Z bosons do. In [112], it was found that the best performance was not shown by any of the τ_N variables alone, but by the ratio of τ_{N+1}/τ_N.

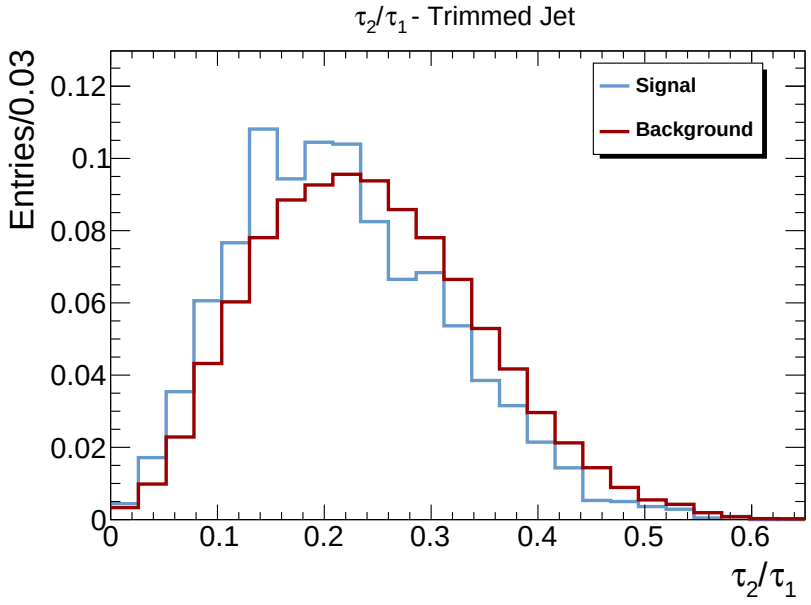


Figure 5.19.: Comparison of WW+WZ and background distribution shapes of τ_2/τ_1. The signal tends to have lower values than the background as it is expected from theoretical considerations.

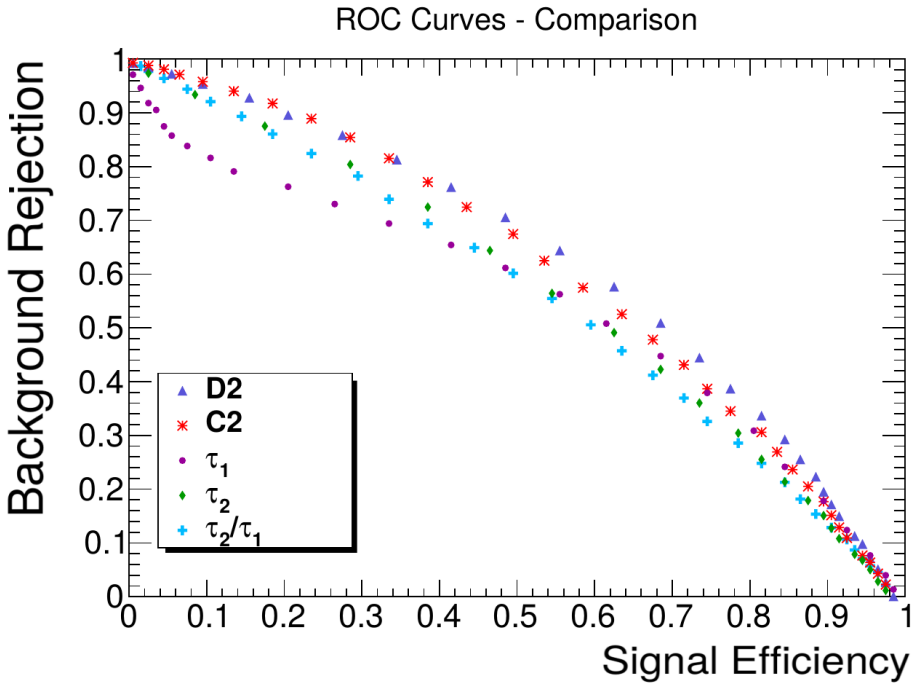
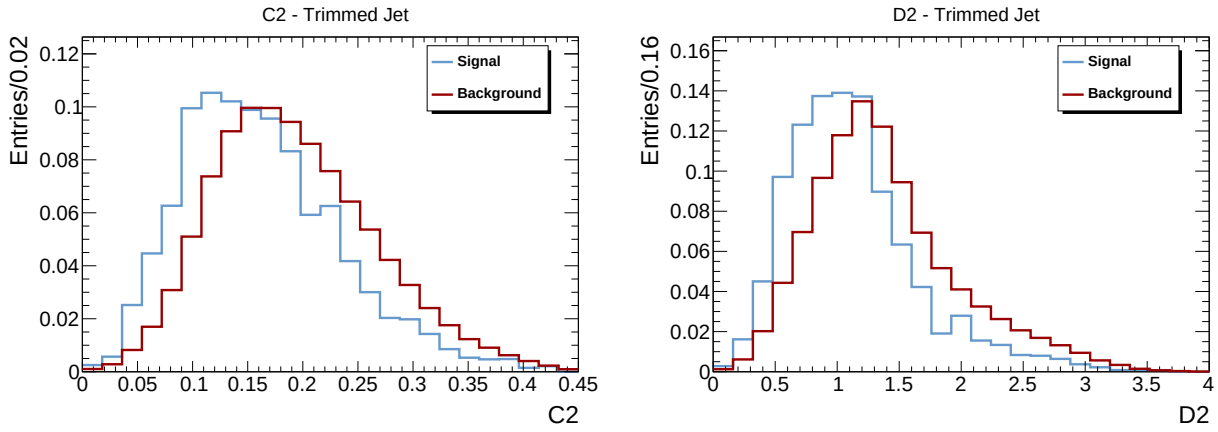


Figure 5.20.: ROC curve comparing background rejection to the corresponding signal efficiency for the five potential cut variables C_2, D_2, τ_1, τ_2 and τ_2/τ_1. D_2 and C_2 outperform the N-subjettiness variables over almost the entire phase space. The performance of C_2 and D_2 is similar, even though C_2 exhibits a slightly worse separation power at the high signal efficiency end.

For the phase space and signal/background composition investigated here, this is not the case, as can be seen in Figure 5.19 and 5.20. The variable τ_2 alone has a slightly better rejection power than τ_2/τ_1 over the complete signal efficiency range. Because of the correlations between

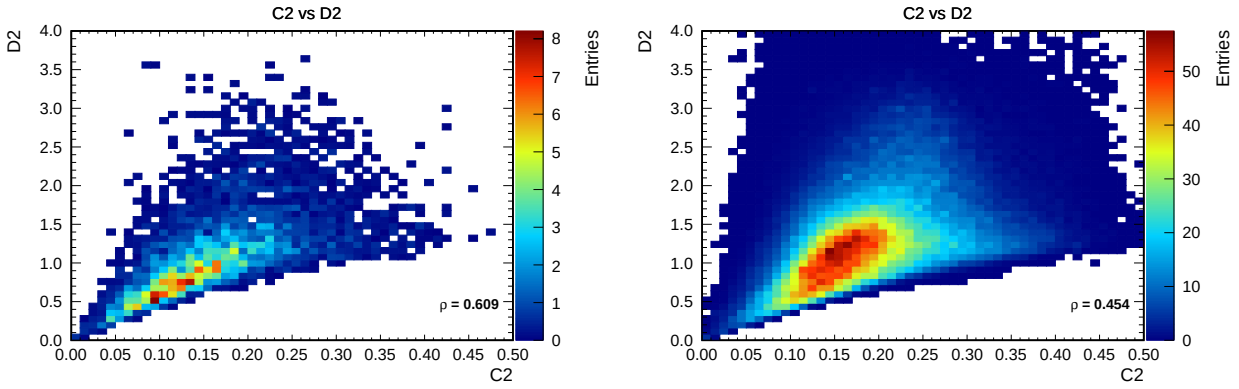
τ_1 and τ_2 being almost identical for signal and background (Figure B.2 in Appendix B) there is no indication either that the ratio of the two variables should lead to a better separation.
Figure 5.21 shows the variables C_2 and D_2. As it is expected, the signal is gathered at smaller values than the background for the two variables. Here the two variables exhibit little difference



(a) (b)

Figure 5.21.: Comparison of signal and background distribution shapes of C_2 (a) and D_2 (b). In both cases the signal peaks at lower values than the background. The separation is slightly better for D_2.

concerning the separation of the two event classes, D_2 can be seen to perform slightly better which is confirmed by the ROC curve in Figure 5.20.
Comparing the N-subjettiness to the energy correlation functions, it is difficult to draw a final conclusion when looking at the entire mass range. D_2 performs best for most of the signal efficiency range. Only at very small signal efficiencies it is on the same level as C_2 and τ_2.
The quantity τ_2/τ_1 shows a slight decrease in performance followed by τ_1 with a considerably worse performance.
A potential difference in the correlation between two variables for signal and background could be exploited by employing a multivariate separation method. This idea is not encouraged by Figure 5.22 since the correlations are generally rather similar for signal and background. Using a two-dimensional cut approach would introduce an additional systematic uncertainty on a second substructure variable, which is expected to be large and would therefore more than outweigh any potential gain in sensitivity. For this reason the multivariate approach has not been followed in this analysis.



(a) (b)

Figure 5.22.: The correlations of C_2 and D_2 for signal (a) and background (b). As expected from the definition of the two variables, C_2 and D_2 are highly correlated. However, since the correlation is similar for the two event classes, the use of a multivariate method is not encouraged by these plots.

5.8.2. Performance in Signal Mass Window

In order to decide which substructure variable works best in discriminating signal and background, the separating power of the variables are re-tested only in the signal mass window between 65 and 95 GeV. This results in a much clearer picture since it becomes apparent in Figure 5.24a, that D_2 performs significantly better than C_2 and τ_2 over almost the complete efficiency range. The ratio τ_2/τ_1 and τ_1 show the worst separation power here as well. For this reason, D_2 was chosen for the final substructure cut variable for the WW+WZ cross section measurement.

5.8.3. Sensitivity with and without Substructure Cut

In the final stage of this analysis, it emerged that the total sensitivity of the cross section measurement does actually not improve but it rather deteriorates when a cut on D_2 is applied. This is understood by looking at Figure 5.24b which shows that the expected sensitivity improves only slightly when a cut is applied, together with the fact that the systematic uncertainty on the D_2 modeling vanishes when the cut is not applied. The approximately constant expected sensitivity for cut values above 1.3 can be explained by the increasing statistics that compensate for the decreasing signal to background ratio when loosening the cut value. Even though applying a cut at $D_2 < 1.3$ yields a background rejection of 51 % at a signal efficiency of 69 % (see Figure 5.20), the associated reduction in statistics almost completely outweighs the gain in signal to background ratio. Since the D_2 cut actually decreased the final sensitivity and furthermore complicates a comparison of theoretical calculations with the measured fiducial cross section, it was decided to not use the D_2 cut.

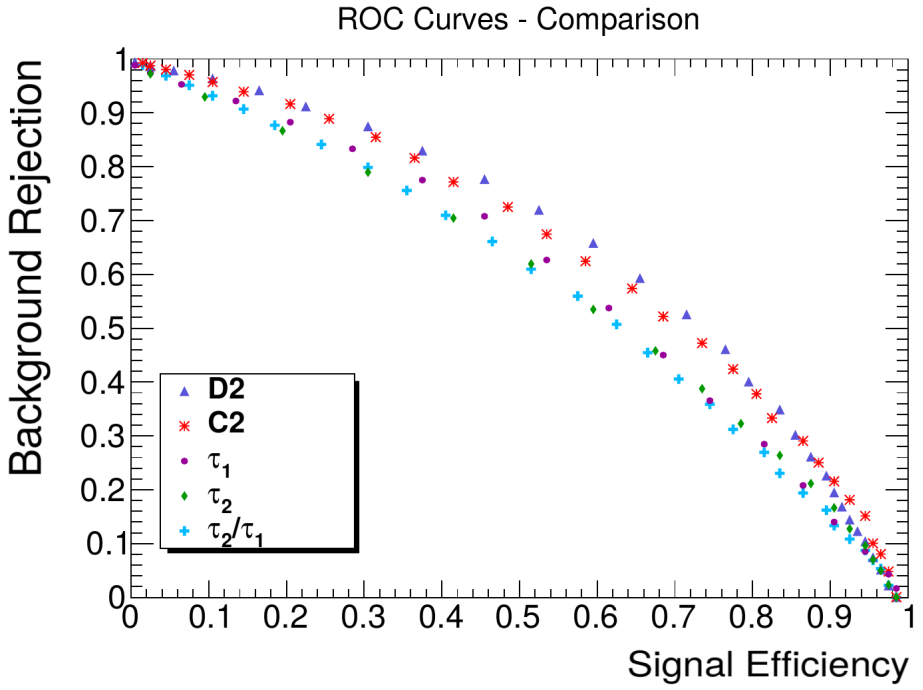
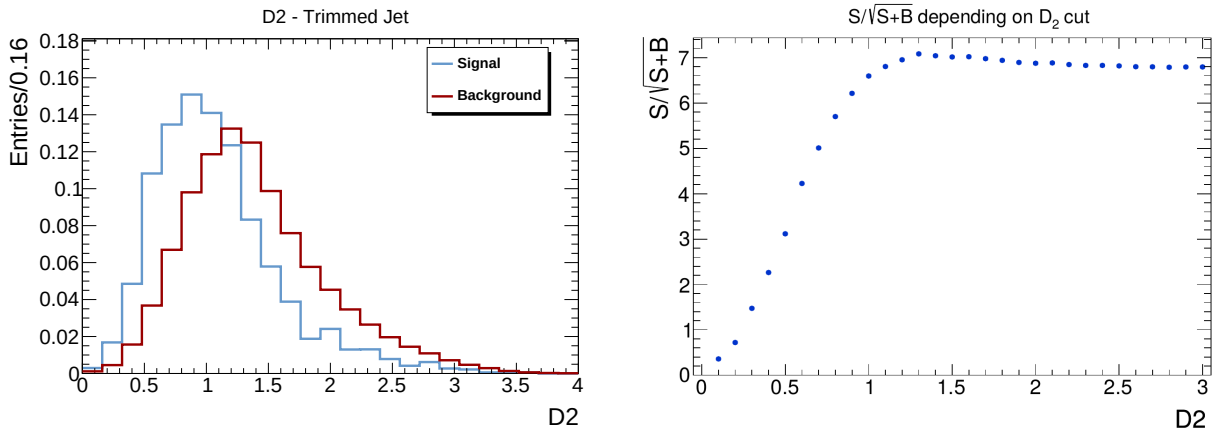


Figure 5.23.: ROC curve comparing C_2, D_2, τ_1, τ_2 and τ_2/τ_1 in the signal mass window. D_2 clearly outperforms all other variables.



(a) (b)

Figure 5.24.: (a) D_2 distributions in signal mass window. The separation is larger here than in the entire mass range which is due to the more confined phase space in this selection. When probing the full mass range, the phase space is looser and both signal and background can diversify compared to their assumed ideal topology which makes them harder to separate.
(b) Expected sensitivity depending on the cut value x for a D_2 cut value of $D_2 < x$ obtained from MC simulation. The sensitivity has a global maximum at $D_2^{\text{Cut}} = 1.3$. However, the difference compared to $D_2^{\text{Cut}} = 3$, which is approximately equivalent to no cut at all (see Figure 5.21b) is marginal.

Chapter 6

Physics Object Definitions and Event Selection

The boosted regime in this analysis is defined by a large transverse momentum of the hadronically decaying vector boson (W or Z). Since the two bosons are produced in a back-to-back topology, this implies a transverse boost for the leptonically decaying W boson as well. The high p_T of the hadronic boson is ensured by requiring a high p_T large-R jet as will be detailed in Section 6.3. Except for a p_T cut on the lepton of $p_T^{\mathrm{l}} > 30\,\mathrm{GeV}$ and on the missing transverse momentum of $E_{\mathrm{T}}^{\mathrm{Miss}} > 50\,\mathrm{GeV}$, there is no direct requirement on the transverse momentum of the leptonic W. However, the high transverse momentum requirement for the hadronic vector boson automatically leads to a similar p_T also for the leptonic W, as can be observed in Figure 6.1 which depicts the situation for WW production.
In this figure the true p_T of the two W bosons obtained from Monte Carlo for events that passed the complete event selection is shown. The vast majority of hadronic and also leptonic bosons are above 200 GeV in p_T. The tail at low p_T of the hadronic W boson which is not present for the leptonic boson is rather an artifact of how the MC generator works than actually physics related. These are events where the p_T of the hadronic W boson is reduced by couplings of the daughter quarks to the initial parton shower or initial-state radiation. These radiative corrections, even though affecting only the decay products of the W boson, are in hindsight also applied to the W boson itself in order to preserve energy and momentum conservation. The quarks together with the radiation still create a high p_T large-R jet and the event passes the event selection. The leptons emerging from the other W boson cannot couple to QCD radiation and therefore the tail here is strongly reduced and still present only because of real final-state photon radiation.

6.1. Physics Objects Definition

Physics objects such as electrons or photons are well defined particles but in the context of collider physics, these objects are only characterized by the specific signatures they leave in the detector. For this reason, in every analysis it has to be clearly specified how exactly all involved physics objects are defined, which shall be done in the following.

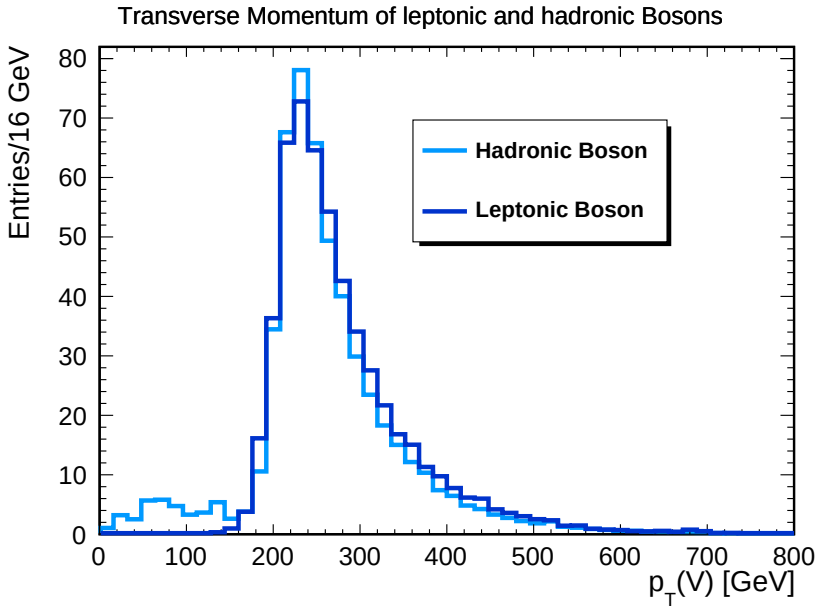


Figure 6.1.: Transverse momentum of the hadronically and leptonically decaying boson at MC generator level, following the application of the full event selection.

6.1.1. Electrons

For this analysis three definitions of electrons are used. One loose baseline selection called "veto electrons" and one selection with stricter quality criteria that is the nominal electron selection. Furthermore, there is one QCD selection which is used for the data driven estimation of the QCD multijet background.
Electron candidates are clusters in the electromagnetic calorimeter with an associated ID track. Additionally, they have to pass a set of quality criteria named `ATLAS MediumPP` which are explained in Ref. [123].
The electron candidate clusters are restricted in η to $|\eta| < 2.47$, excluding the crack region at $1.37 < |\eta| < 1.52$ (see Section 3.2.2).
Electron candidates are required to originate from the hard scatter vertex. Thus, the significance of the track's transverse impact parameter calculated with respect to the beam line, $|d_0/\sigma(d_{0,\text{track}})|$, must be less than five, and the longitudinal impact parameter, z_0 (the difference between the value of z of the point on the track at which d_0 is defined and the longitudinal position of the hard scatter vertex), is required to satisfy $|z_0/\sin(\theta_{\text{track}})| < 0.5\,\text{mm}$.
To guarantee electron isolation, it is required that the transverse momentum measured in the calorimeter within a cone of radius R around the electron is below a certain fraction of the electron's own transverse momentum. For the baseline selection this corresponds to $p_T(R = 0.2)/p_T(\text{el}) < 0.1$. The minimum electron p_T is required to be above 15 GeV.
In addition to this baseline selection, an electron has to pass the following criteria to be counted as a nominal electron. First, it has to pass the `ATLAS TightPP` criteria which is a stricter set of quality requirements than `ATLAS Medium`. Second, the minimum p_T requirement is 30 GeV and there are stricter cuts applied on the E_T and p_T isolation:

- $E_T(R = 0.3)/E_T(\text{el}) < 0.07$
- $p_T(R = 0.3)/p_T(\text{el}) < 0.14$

The electron selection for the QCD estimation is orthogonal to the nominal electron selection but consists of a subset of the veto electrons. Additionally to the veto electron specifications QCD electrons have to satisfy the following requirements:

- $p_T > 30\,\text{GeV}$
- Fail `ATLAS TightPP` quality requirements
- $2\,\text{GeV} < E_T(R = 0.3) < 20\,\text{GeV}$
- At least one hit in the inner layer of the pixel detector

6.1.2. Muons

Equally to electrons also muons are selected with three different sets of selection criteria as veto, nominal and QCD selection muons. In general, muons are identified by the "STACO" combined muon reconstruction algorithm [124]. This algorithm uses reconstructed tracks of muons in the inner detector and hits in the muon spectrometer which are first reconstructed separately and then later combined by a statistical algorithm [125]. All muons are required to satisfy a number of inner detector quality criteria [126]. Aside from these, the requirements on muons are similar to the ones for electrons. With the exception of the eta requirement of $|\eta(\mu)| < 2.4$, the veto muons have to pass the following selections:

- $p_T(R = 0.2)/p_T(\mu) < 0.1$
- $|z_0/\sin(\theta_{\text{track}})| < 0.5\,\text{mm}$

Additionally, several stricter conditions have to be met by a muon to pass the nominal selection:

- $|d_0/\sigma(d_{0,\text{track}})| < 3$
- $p_T(\mu) > 30$ GeV
- $E_T(R = 0.3)/E_T(\mu) < 0.07$
- $p_T(R = 0.3)/p_T(\mu) < 0.07$

The QCD muon selection is again orthogonal to the nominal muon selection. This is achieved by inverting the cut on the transverse impact parameter significance $|d_0/\sigma(d_{0,\text{track}})|$. Additionally to the veto muon selection, those muons have to pass the following requirements:

- $p_T > 30\,\text{GeV}$
- $|d_0/\sigma(d_{0,\text{track}})| > 4$

6.1.3. Jets

Jets in general have already been described in Chapter 5, here it will be focused on the specific jets used in this analysis, which are of two different kinds: "small-R jets" and "large-R jets". Both are reconstructed from topological calorimeter clusters that were calibrated with the Local Hadronic Calibration as described in Section 5.3.3.
Small-R jets are reconstructed with the anti-k_T algorithm and a radius parameter of $R = 0.4$. In order to distinguish hard scatter jets from pile-up jets, a cut on the jet vertex fraction of

$|\mathrm{JVF}(j)| > 0.5$ is applied if the jet lies in $|\eta| < 2.4$ and has a p_T of less than 50 GeV. Small-R jets have to lie inside an η-range of $|\eta| < 4.5$. Such jets are not actively used in the analysis but they are merely used as veto objects in order to suppress background events (see Section 6.3).
The large-R jets are built in a two step procedure. In the first step, they are built with the anti-k_T algorithm with a radius parameter of $R = 1.0$ using as input all topological clusters present in the event.
In order to eliminate pile-up contributions inside the large-R jets, the trimming procedure ($f_{\mathrm{cut}} = 5\%$) using the k_T algorithm with $R_{\mathrm{Subjet}} = 0.2$ as subjet algorithm is applied to them.
As subjet algorithm for trimming, anti-k_T could in principle be used as well. Although its advantage, which is producing circular, regularly-shaped jets, becomes a disadvantage when looking for subjets in a larger jet because the space is limited here. An algorithm that is able to adapt its shape to the busy subjet environment performs better in this case [127] which is why the k_T algorithm was chosen.

6.1.4. Overlap Removal

It is important to minimize misreconstruction/miscalibration of physics objects caused by other objects that are spatially close to them.
Therefore muons overlapping in $\Delta R < 0.4$ with a selected jet are removed from the selected muons – a procedure which helps rejecting muons from heavy flavour decays in jets.
Since every selected electron causes a calorimeter cluster which is also identified as a jet, it is crucial to remove those electron-induced jets from the actual hadronic jets. For this reason, all selected jets within $\Delta R < 0.2$ of a selected electron are removed from the selected jets.
In case that hadronic jets are close to electrons, it is safer not to use these electrons since there is a possible overlap of the respective calorimeter clusters and therefore the possibility of energy misreconstruction. Consequently, all selected electrons that have a close-by jet in $0.2 < \Delta R(\mathrm{el}, \mathrm{jet}) < 0.4$ are removed from the selected electrons.

6.1.5. Missing Transverse Momentum

The missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{Miss}}$) is an important quantity for all analyses involving neutrinos. Since neutrinos hardly interact with matter, they escape the detector without depositing energy inside it. This makes it impossible to reconstruct them in general. However, there is the concept of missing transverse momentum which allows to at least draw conclusions about the neutrino momentum in x and y-direction.
The idea is to use the fact that the transverse momentum of the two head-on colliding protons is exactly zero before the collision and as such it also has to be afterwards.
Assuming all other physics objects in the event can be well measured, one can construct the vectorial sum of the transverse momentum of all those objects. The $E_{\mathrm{T}}^{\mathrm{Miss}}$ is then obtained as the vector that has to be added to bring the sum back to zero. With this method, the transverse momentum of the neutrino and its direction in ϕ can be reconstructed.
The reason why only the momentum in the transverse plane and not the total missing momentum is used is twofold. First, in a hadron-hadron collision, the actual fraction of the total proton momentum of the quarks or gluons that are responsible for the the hard scatter is unknown (see Section 2.2). Second, a large fraction of the proton fragments disappears in forward and backward direction in the beam pipe and thus is not measured, preventing the determination of

the total momentum of the particles present in the event.
The missing momentum is calculated by summing the momentum of all calibrated and pile-up corrected physics objects in x and y-direction plus the sum of p_x (p_y) of calibrated topoclusters which are not associated to any physics object (soft term). The magnitude is of the $E_{\mathrm{T}}^{\mathrm{Miss}}$ is then obtained as

$$E_{\mathrm{T}}^{\mathrm{Miss}} = \sqrt{(E_x^{\mathrm{Miss}})^2 + (E_y^{\mathrm{Miss}})^2}. \tag{6.1}$$

The physics objects used for calculation of the $E_{\mathrm{T}}^{\mathrm{Miss}}$ are pile-up corrected already whereas the soft term, which is highly susceptible to pile-up, is at this point uncorrected.
The default method to correct for this term is the so-called soft term vertex fraction (STVF) method [128]. The STVF is the sum of the p_T of all tracks from the primary vertex divided by the sum of the p_T of all vertices in the event. This ratio provides a good estimation of the signal contribution of the overall activity in the event and is therefore multiplied by the soft term to correct for pile-up. In this analysis the standard $E_{\mathrm{T}}^{\mathrm{Miss}}$ definition in ATLAS [129] was used.

6.1.6. Triggers

As outlined in Section 3.2.4, it is impossible to record all events at the LHC which is why a highly sophisticated trigger system was developed. When conducting a physics analysis, one or several particular triggers to be used for the event selection have to be decided upon. Typically a trigger selects events containing a certain physics object with a p_T above a certain threshold. Of utmost importance is the fact whether the trigger exhibits a prescale over parts of or the complete data taking period. A prescale is a factor P by which the event recording rate is purposefully decreased in order to limit the data rate. For example if a given trigger has a prescale of 20, 19 out of 20 events in which the trigger fired are discarded while only the 20th is recorded. This is a common technique to reduce the rate of highly abundant event types while keeping the possibility to conduct an analysis of such events, with the downside of gathering lower statistics.
For this analysis the unprescaled triggers for single electrons and muons with the lowest p_T threshold in 2012 data are used as baseline triggers. The two triggers require an isolated electron (muon) with $p_T > 24\,\mathrm{GeV}$. Additionally a second set of unprescaled single lepton triggers is used which require an electron (muon) with a $p_T > 60$ (36) GeV without any isolation requirement. At least one of these four triggers must have fired in order for an event to pass this first stage in the event selection.

6.2. Monte Carlo and Data Samples

The dataset used for this analysis corresponds to all ATLAS data taken in 2012 from period A to L, with a total integrated luminosity of $20.3\,\mathrm{fb}^{-1}$. When data were taken in 2012, events were classified into so called “streams”, depending on the triggers that fired in the event.
The three most important streams are the “Egamma”, the “Muons” and the “JetEtMiss”-stream. Each stream is defined by a list of trigger chains. If at least one of these triggers fired in a given event, the event is attributed to that stream. It is also possible that an event is associated to more than one stream. Potential double counting of events present in two streams is avoided by an event veto method explained in Section 6.3.

The only streams used in this analysis are the Egamma and the Muons stream providing for the electron and muon events, respectively. Numbers of collected events per period for the two streams can be found in Appendix A.1.
With the exception of the QCD multijet background, all background processes are estimated by Monte Carlo simulation.
Monte Carlo events of a particular process are generated by a certain event generator and possibly interfaced to a different generator to obtain the modeling of the underlying event, parton shower and hadronization.
The interactions of the MC particles with the detector material and the magnetic field in the detector are simulated with GEANT4 [130]. After the simulation of a particle's passage through the detector, the same identification and reconstruction algorithms are used on the simulated data as in the case of real data. An exhaustive list of all MC samples used in the present analysis is given in Appendix A.

6.2.1. WW+WZ Signal Samples

Nominal WW and WZ samples are generated with MC@NLO v4.07 as $q\bar{q} \to WW/WZ$ events. The CT10 set of parton distribution functions is used to model the momentum of the incoming partons. The events are interfaced to Herwig for the parton shower and hadronization processes and to Jimmy [131] for underlying event simulation, using the AUET2 tune [132]. A filter is applied on particle level, requiring at least one electron or muon that stems either directly from a W boson decay or from the decay of a tau-lepton which in turn stems from a W decay. The reference cross sections used for normalization of these samples stem from the Monte Carlo generator itself and amount to:

$$\sigma^{\mathrm{NLO}}_{q\bar{q}\to WZ}(\mathrm{MC@NLO}) = 21.5\,\mathrm{pb}, \tag{6.2}$$

$$\sigma^{\mathrm{NLO}}_{q\bar{q}\to WW}(\mathrm{MC@NLO}) = 53.2\,\mathrm{pb}. \tag{6.3}$$

In addition to the nominal SM-like WW and WZ samples, another set of WW/WZ samples is produced for an anomalous triple gauge coupling point with the following parameters:
$\Delta g_1^Z = -0.3$, $\Delta g_1^\gamma = 0$, $\Delta\kappa_Z = 1$, $\Delta\kappa_\gamma = 0$, $\lambda_Z = 0.3$ and $\lambda_\gamma = 0$.
These samples are produced in a way that allows them to be reweighted event by event to any desired aTGC point [133]. Because the vector bosons in the nominal sample are generated and decayed by two different generators, the intrinsic width of the bosons and spin correlation effects are not taken into account.
Furthermore, four additional MC@NLO samples with the same settings have been produced in order to evaluate the systematic uncertainties associated to the particular choice of the renormalization and factorization scale used in the simulation. Those samples are produced on particle level only, no detector simulation is performed. Besides that, the only difference to the nominal samples are the varied values of renormalization and factorization scale. In the four samples, μ_F (μ_R) is varied independently by a factor 0.5 and 2 compared to its nominal values while the respective other scale is kept constant.
For the evaluation of systematic uncertainties stemming from the MC generator modeling, two additional diboson samples have been generated, the first using Sherpa v1.4.1 [85] and the second using Powheg v2 [134–136]. For both samples the CT10 PDF set was used.
The Sherpa sample is generated at leading order with up to three additional partons present in the perturbative expansion. The Powheg sample is generated at NLO and (after the decay of the W/Z) interfaced to Pythia 8 for parton showering. In both of these samples, the decay

of the vector bosons happens directly in the matrix element generator which is why here, spin correlation effects and the intrinsic width of the massive bosons are taken into account correctly. Furthermore, the two samples also contain the $W\gamma^*$ contributions in the WZ sample which are not present in the nominal MC@NLO WZ sample. A list of all signal samples is assembled in Appendix A.2.

6.2.2. W/Z+Jets Samples

Nominal W/Z+jets samples are generated with Sherpa using the CT10 PDF set. The vector bosons are forced to decay leptonically. Heavy quarks (c, b) are treated as massive and the generation includes up to four additional partons in the matrix element. Sherpa also performs the parton showering, hadronization and the underlying event simulation. To obtain sufficient statistics at high p_T(V) with reasonable computational effort, the samples are generated in the following slices of W/Z-p_T (units in GeV):

Z+jets p_T-slices:

1. $0 < p_T(\mathrm{Z})$
2. $40 < p_T(\mathrm{Z}) < 70$
3. $70 < p_T(\mathrm{Z}) < 140$
4. $140 < p_T(\mathrm{Z}) < 280$
5. $280 < p_T(\mathrm{Z}) < 500$
6. $p_T(\mathrm{Z}) > 500$

W+jets p_T-slices:

1. $0 < p_T(\mathrm{W})$
2. No corresponding sample
3. $70 < p_T(\mathrm{W}) < 140$
4. $140 < p_T(\mathrm{W}) < 280$
5. $280 < p_T(\mathrm{W}) < 500$
6. $p_T(\mathrm{W}) > 500$

Except for the two $p_T(\mathrm{V}) > 0$ samples which are inclusive W/Z samples, all other samples are exclusive in p_T. In order to avoid overlap of the inclusive and exclusive samples, a cut is placed on the truth W (Z) p_T at $p_T < 70$ (40) GeV for the two inclusive samples.
Each p_T slice is moreover divided into three mutually exclusive subsamples depending on the content of heavy quarks in the fixed order matrix element calculation:

1. c-veto, b-veto
2. c-filter, b-veto
3. b-filter, c-veto

A complete list of nominal W/Z+jets background samples is given in Appendix A.3.1 and A.4. For the evaluation of their associated systematic uncertainties, further W+jets samples including up to five partons in the perturbative calculation are generated with Alpgen [137] using the CTEQ6L1 PDF set [39]. These events are then interfaced to Pythia 6.426 with the Perugia2011C [138] underlying event tune and the CTEQ6L1 PDF set (see Appendix A.3.2) for the parton shower modeling. In contrast to the Sherpa samples, the Alpgen samples are generated separately for events in which the perturbative matrix element calculation contains:

1. Only light quarks (u, d, s)
2. Heavy quarks (c, b)

In the light quark samples, W+cc and W+bb can nevertheless be produced in the parton shower. However, the same processes are produced in the perturbative part of the heavy flavour samples

as well. In the ME heavy flavour samples, the heavy quarks are preferentially produced at higher p_T and larger splitting angles than in the PS of the light quark samples, yet the produced topologies are not completely orthogonal. To avoid double counting a tool was developed by the ATLAS collaboration [139] which is referred to as "heavy flavour overlap removal" tool. It decides based on the ΔR between the two heavy flavour quarks if an event is to be kept or rejected and as such removes the overlap between the two samples.
In all W/Z+jets samples, the sum of all cross sections from the different subsamples is re-scaled to the inclusive W/Z production cross section section quoted in Equations 4.11 and 4.12. For the sake of simplicity, the sum of W+jets and Z+jets will hereafter be referred to as V+jets.

6.2.3. Top Quark Pair and Single Top Quark Samples

The nominal single top quark samples in the s-, t- and Wt-channel and the nominal $t\bar{t}$ sample are generated at NLO with Powheg v1 with the CT10 PDF set and an assumed top quark mass of 172.5 GeV. The samples are interfaced to Pythia 6.42 with the Perugia2011C underlying event tune and the CTEQ6L1 PDF set. The theoretical cross section used for the normalization of the $t\bar{t}$ sample is 252.9 pb (see Equation 4.13). The single top quark samples in the s-, and t-channel are normalized to the NLO theoretical cross sections obtained using the Hathor v2.1 program [140, 141], making use of the MSTW2008 NNLO PDF set. The cross section for the Wt-channel is based on a NNLO prediction and is taken from Ref. [142].
For the evaluation of systematic uncertainties, four further $t\bar{t}$ samples are generated, all of which are using the CT10 PDF set for the perturbative part of the generation. The first one is generated with MC@NLO, the second one with Powheg, both of them are then interfaced to Herwig 6.520.2 with the AUET2 tune and the CT10 PDF set.
The MC@NLO and Powheg samples are used for the evaluation of generator and parton shower uncertainties, respectively. The last two systematic uncertainty samples are generated by AcerMC [143] and interfaced to Pythia 6.426 using the CTEQ6L1 PDF set and the AUET2B tune. Those two samples are produced with less and more initial- and final-state radiation (ISR/FSR), respectively, in order to be used for the assessment of associated systematic uncertainties. A list of all $t\bar{t}$ and single top quark samples used in this analysis is given in Appendix A.5 and A.5.2.

6.3. Event Selection

All events in data have to pass a list containing all good data taking periods, excluding intervals with malfunctioning detector components or corrupted events – the so-called "good runs list" for 2012 data.
Furthermore, a vertex with at least three associated tracks with $p_T(\text{track}) > 400\,\text{MeV}$ is required, as well as exactly one well reconstructed lepton fulfilling all quality requirements for the nominal selection. Events containing a veto lepton of any flavour in addition to the selected nominal lepton are rejected. A matching between the trigger decision and the lepton is required, meaning that the selected lepton and no other lepton in the event must actually have caused the trigger to fire.
For data, there is the additional requirement that electron events are only kept if they are in the Egamma stream and likewise for muon events. If an event belonging to a given lepton channel stems from the opposite lepton flavour stream, it is vetoed. This prevents potential double counting of events as explained in Section 6.2.

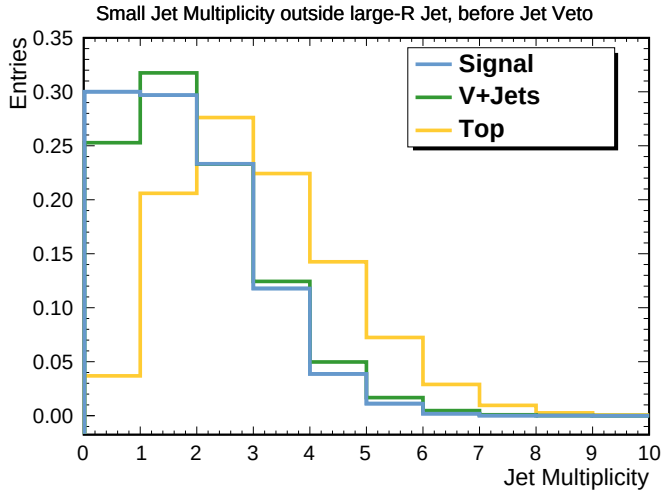


Figure 6.2.: Distribution of the small-R jet multiplicity outside the large-R jet, before the jet veto is applied. All distributions are normalized to unity. All events with a multiplicity > 0 are vetoed. The efficiency of this cut is largest for the signal with 30 %, for V+jets it is less with 25.2 %, while for top quark production it is only 3.7 %. The jet veto thus suppresses top by a factor of around eight with respect to the signal.

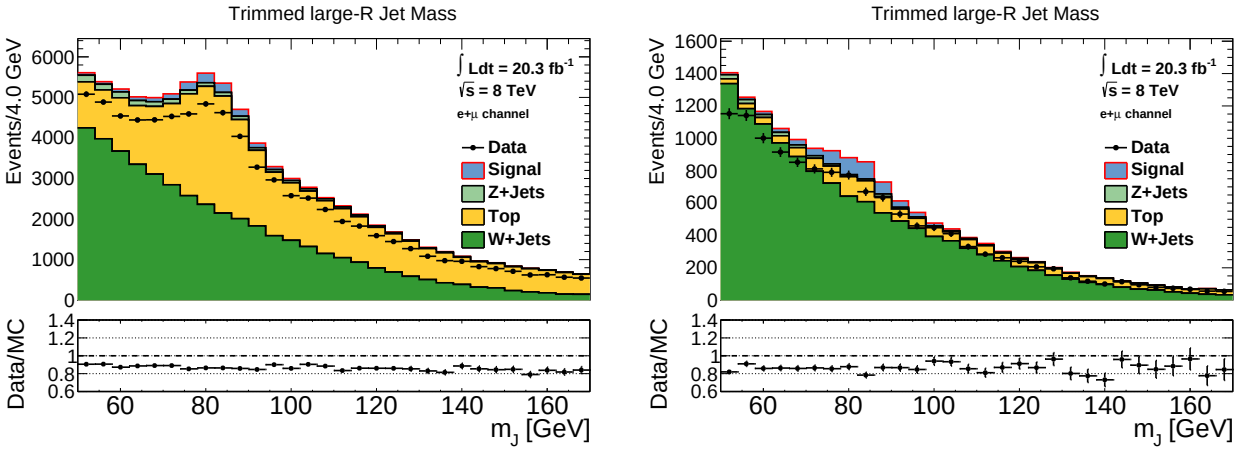
Cut	Name	Description	#Events
C1	Initial	WV event with at least one lepton (el, μ)	620345
C2	Trigger	Pass electron or muon trigger	312426
C3	>0 Leptons	At least one selected lepton	289796
C4	Overlap Removal	Lepton-Jet-overlap removal	229387
C5	Exactly 1 Lepton	Not more than 1 nominal lepton	210092
C6	No Veto Lepton	No additional veto lepton	194302
C7	Trigger Matching	Lepton-Trigger matching	193843
C8	E_T^{Miss}	$E_T^{Miss} > 50$ GeV	72991
C9	Exactly 1 large-R Jet	Exactly 1 selected large-R jet	3677
C10	Jet Veto	No additional small-R jets	977
C12	Large-R Jet Mass	m(large-R Jet) > 50 GeV	672

Table 6.1.: The table shows the expected number of remaining WW+WZ events in 2012 data which are containing at least one electron or muon following the application of each successive cut.

The missing transverse momentum in the event has to satisfy $E_{\mathrm{T}}^{\mathrm{Miss}} > 50\,\mathrm{GeV}$. The hadronic W or Z is reconstructed as a large-R jet of which exactly one with the following properties is required:

1. $p_T > 200\,\mathrm{GeV}$
2. $|\eta| < 2.0$
3. No overlap with the selected lepton in $\Delta R = 1.0$

In order to suppress $\mathrm{t\bar{t}}$ and single top quark background events which typically contain a large number of small-R jets (see Figures 6.2 and 6.3), any event with an additional jet outside of the large-R jet is vetoed as well. The effect of the jet veto can be seen in Figure 6.3.
Since the mass calibration for large-R jets is not valid for jet masses below 50 GeV, a cut on the large-R jet mass of $m_J > 50$ GeV is applied. The effect of each cut on the signal selection can be seen in the cutflow diagram in Figure 6.4.



(a) (b)

Figure 6.3.: (a) Large-R jet mass distribution after the complete event selection except for the jet veto. (b) The same distribution with the jet veto applied. The top background can be seen to be strongly suppressed. There is a visible difference in normalization between simulation and data. This is due to the fact that scale factors derived for W+jets and top background (see Section 7.2) are not yet applied in the two plots. The QCD multijet contribution (see Section 7.1) is neglected here for both scenarios.

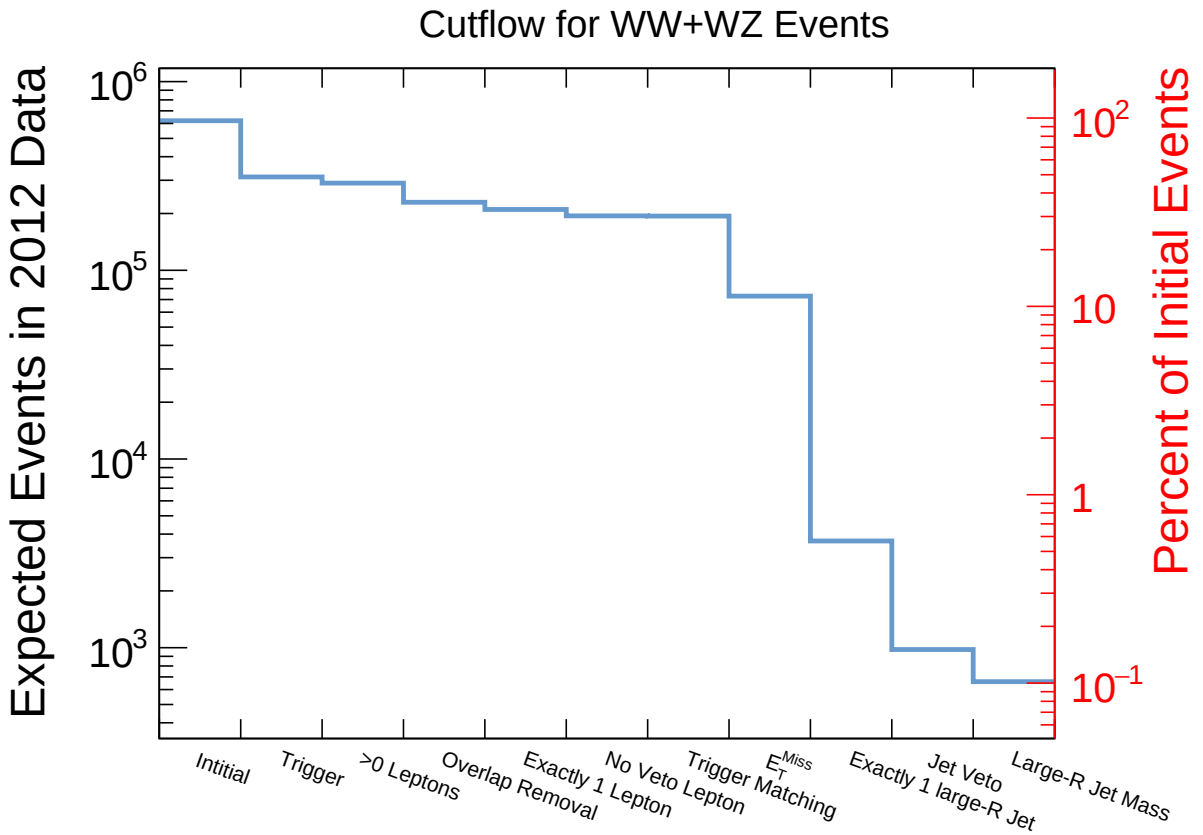


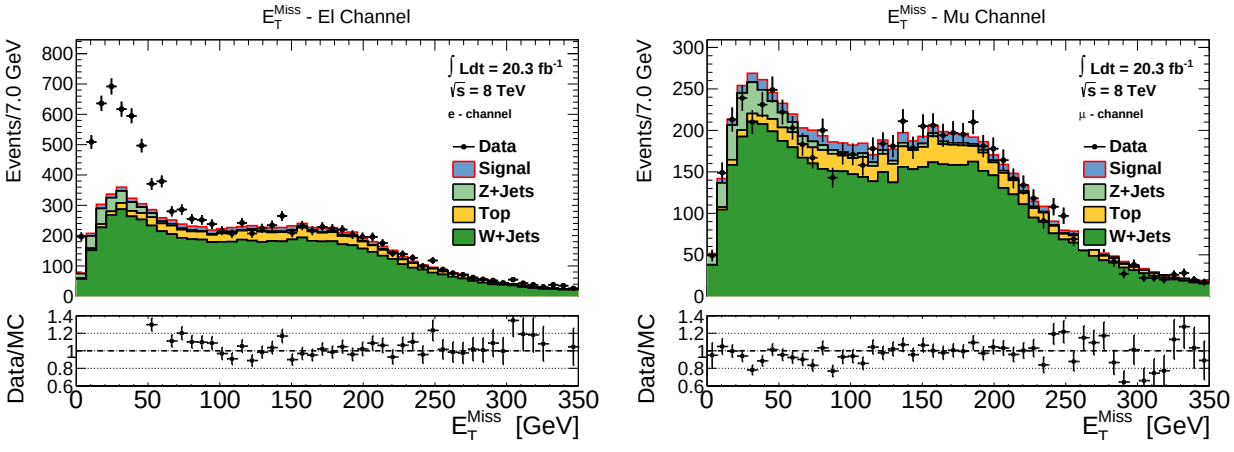
Figure 6.4.: Cutflow diagram for WW+WZ events showing the remaining events on MC after each cut. The precise number of entries for each bin and a description of the cut names is listed in Table 6.1.

CHAPTER 7

Background Estimation

7.1. Data-driven Estimation of QCD Multijet Background

The QCD multijet contribution has to be estimated by a data-driven method mainly due to the very limited MC statistics available for this background.
The crucial variable for the QCD estimation is the missing transverse momentum. As explained in Section 4.4, the QCD multijet processes contain little missing transverse momentum while all other signal and background processes exhibit large $E_{\mathrm{T}}^{\mathrm{Miss}}$ due to the presence of a neutrino in a leptonic W boson decay. This difference can be exploited in order to suppress multijet processes and to distinguish them from the other signal/background processes. A rough estimate



(a) (b)

Figure 7.1.: $E_{\mathrm{T}}^{\mathrm{Miss}}$ distribution obtained without application of the $E_{\mathrm{T}}^{\mathrm{Miss}}$ cut for the electron channel (a) and the muon channel (b). In the muon channel the data are well described with the set of nominal signal and background processes whereas in the electron channel there is evidently a contribution missing – the QCD multijet background.

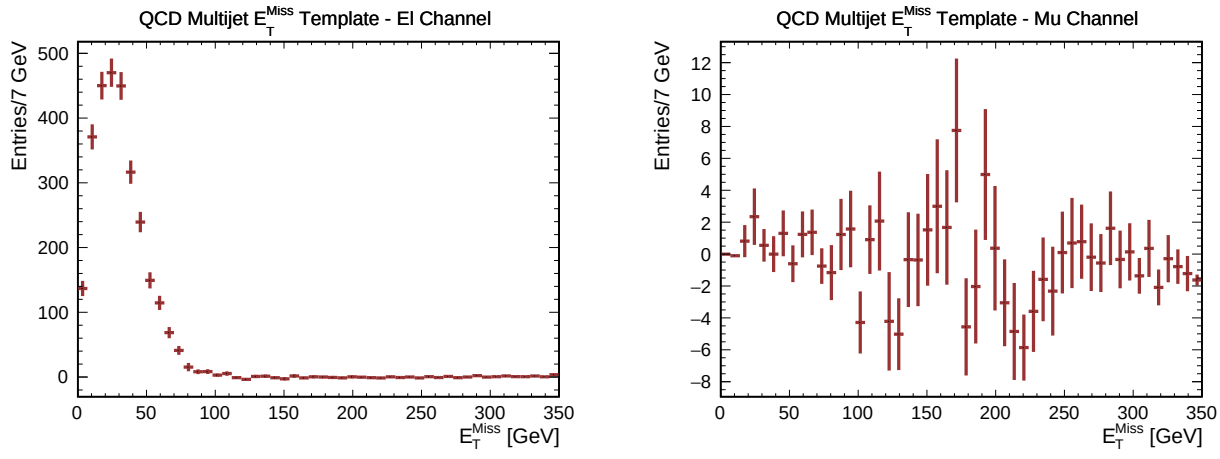
of the QCD multijet contribution can be obtained by applying the default event selection except for the $E_{\mathrm{T}}^{\mathrm{Miss}}$ cut. Since QCD multijet events are expected to contain little $E_{\mathrm{T}}^{\mathrm{Miss}}$, the absence of such a cut leads to a relative enhancement of the QCD fraction present in data after event selection. This additional contribution would show up as "empty" areas in a stacked plot of all expected background processes and the data. For this estimation, approximate scale factors (SF) to correct the top quark and W+jets normalization have already been applied. The exact value of these scale factors depends on the outcome of the QCD multijet contribution and will be described in Section 7.2.

Figure 7.1 shows that in the muon channel, the data are well described by the usual background processes but a contribution is missing in the electron channel. Although most of the QCD contribution is rejected by the $E_{\mathrm{T}}^{\mathrm{Miss}}$ cut and only the white area above 50 GeV between the data histogram and the upper stacked plot border enters the final event selection, the contribution is large enough that it cannot be neglected.

For this reason a multijet template histogram is estimated from data for the electron channel only. The rough assessment of the QCD contribution already suggests that there is no such contribution in the muon channel, an assumption which will prove correct in the following data-driven QCD estimation method.

The reason for this asymmetry between the two lepton channels is the following. While it is relatively common for a QCD jet to be classified as an electron since both produce calorimeter clusters, the same happens much more rarely for muons. To be identified as a muon, a hadronic jet would have to punch through to the muon spectrometer. In the limited cases that this happens the signature of a multiparticle jet in the muon system is naturally highly different from that of a genuine muon. For this reason QCD events are greatly suppressed in the muon channel while a small fraction of them are still misclassified as electron events.

For the QCD multijet estimation, a modified ("inverted") lepton selection which is as close to the nominal selection as possible but orthogonal to it, is defined to obtain a QCD-enriched control region. See Section 6.1 for more details on the lepton selection.



(a) (b)

Figure 7.2.: Data driven estimation of the missing transverse momentum template expected from QCD multijet events for the electron channel (a) and the muon channel (b). The statistics in the muon channel is extremely small which is expected from the muon mistag efficiency considerations. The bin entries in this channel are furthermore mostly consistent with zero for the whole mass range. In the electron channel the $E_{\mathrm{T}}^{\mathrm{Miss}}$ is small with a mean value of 31 GeV.

The invariant mass template is derived in several steps:

1. Producing a QCD shape template for $E_{\mathrm{T}}^{\mathrm{Miss}}$ using the inverted lepton selection.
2. Calculating the QCD normalization by a fit of the $E_{\mathrm{T}}^{\mathrm{Miss}}$ distribution to data.
3. Deriving a QCD shape template for invariant jet mass and scale it to fitted normalization.

In the first step, a QCD enriched control region is obtained by applying the inverted selection criteria to electrons and muons. For the muon channel, the main difference between the nominal and the inverted selection is the cut on the transverse impact parameter. This is inverted for the QCD selection, i.e. $|d_0/\sigma(d_{0,track})| > 4\,\mathrm{mm}$ instead of $|d_0/\sigma(d_{0,track})| < 3\,\mathrm{mm}$. In that way, a sample of muons which do not stem from the primary vertex, as is expected for muons from heavy flavour decays is obtained. In the electron case, except for additional requirements on

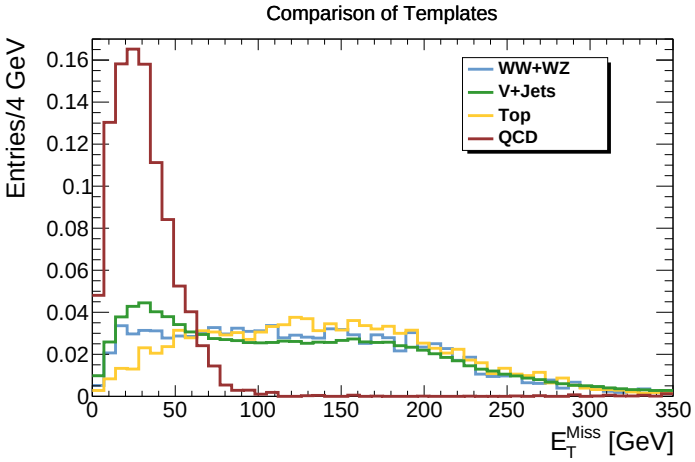
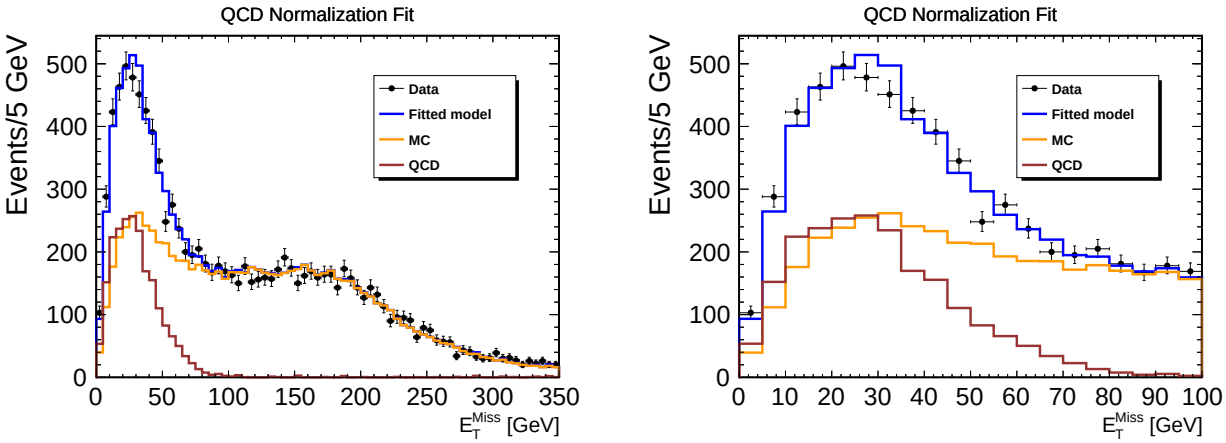


Figure 7.3.: A comparison of $E_{\mathrm{T}}^{\mathrm{Miss}}$ templates for all signal and background processes is shown. The QCD template is obtained by running the inverted lepton selection on data and subtracting the estimated contribution of the MC-modeled processes. The largest $E_{\mathrm{T}}^{\mathrm{Miss}}$ is found in the top template because of potentially two neutrinos in the event. As QCD multijet events do not contain any true missing transverse momentum, this template peaks at low values as expected.

the isolation of the pixel-layer hit (which is applied in order to suppress photon conversion backgrounds), the same criteria as for the nominal selection are applied, with the difference that the QCD electrons have to fail the `ATLAS TightPP` criteria. Thus the sample is enriched with events where jets fake an electron as anticipated for QCD multijet events entering the event selection. A template histogram of the $E_{\mathrm{T}}^{\mathrm{Miss}}$ in the full range is produced for the electron and muon channel, respectively, by running the analysis with the inverted lepton selection and not applying the $E_{\mathrm{T}}^{\mathrm{Miss}}$ cut.
This $E_{\mathrm{T}}^{\mathrm{Miss}}$ template distribution is still contaminated by the other standard background processes like V+jets, $\mathrm{t\bar{t}}$ and single top quark production and also to a small fraction by WV signal. To obtain the template for QCD multijet only, the expected distributions from all these processes are subtracted from the data histogram. Figure 7.2 shows the templates for the two lepton channels.
In the second step, the correct normalization of these $E_{\mathrm{T}}^{\mathrm{Miss}}$ templates is evaluated by a two-component χ^2-fit to data. For this the analysis is re-run with the nominal lepton selection, except that the $E_{\mathrm{T}}^{\mathrm{Miss}}$ cut is not applied (same selection as was used for Figure 7.1).
Then a fit of the $E_{\mathrm{T}}^{\mathrm{Miss}}$ templates of the QCD multijet background and the sum of all other backgrounds and the WV signal is performed to the produced data histogram for each channel separately. All templates used in the normalization extraction fit are shown in Figure 7.3.
However, for the actual fit WW+WZ, V+Jets and top quark template are added according to their SM-expected normalization to form one template. This template is called “MC” and can be seen in Figure 7.4a. For the fit, a smaller range of 0-100 GeV is used since above that, the QCD template is approximately zero. The fitted number of QCD events in the electron channel with this selection amounts to $N_{QCD} = 2104 \pm 100$ for the full range of the $E_{\mathrm{T}}^{\mathrm{Miss}}$. The result of

the fit in the muon channel is consistent with zero. Therefore a QCD template histogram is only used in the electron channel.
In the third step, the actual, final QCD multijet invariant large-R jet mass template is constructed. This is done by running the analysis with the inverted lepton selection, this time with the $E_{\mathrm{T}}^{\mathrm{Miss}}$ cut applied. The resulting invariant mass template is then scaled to $N_{QCD} \cdot \varepsilon_{\mathrm{MET}}$, with $\varepsilon_{\mathrm{MET}}$ being the efficiency of the $E_{\mathrm{T}}^{\mathrm{Miss}}$ cut at 50 GeV on the $E_{\mathrm{T}}^{\mathrm{Miss}}$ QCD template.



(a) (b)

Figure 7.4.: (a) Two component fit of the QCD template and the sum of signal and background templates ("MC") over the full $E_{\mathrm{T}}^{\mathrm{Miss}}$ range. The QCD template only contributes in the low $E_{\mathrm{T}}^{\mathrm{Miss}}$ region below ∼80 GeV. Above that, the $E_{\mathrm{T}}^{\mathrm{Miss}}$ is dominated by the processes with real $E_{\mathrm{T}}^{\mathrm{Miss}}$ from a neutrino in the event. (b) The same plot in the range used for the extraction of the QCD normalization (0-100 GeV).

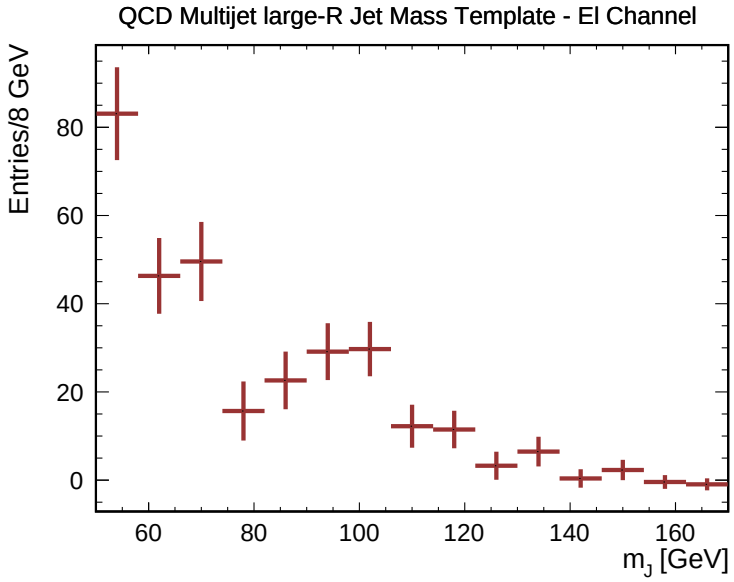


Figure 7.5.: Large-R jet invariant mass template for QCD multijet production in the electron channel, obtained as described in the text. The statistical uncertainty and bin by bin fluctuations are relatively large due to the small remaining statistics in data after applying the inverted electron selection.

7.2. Normalization of Top and W+Jets Background

Figure 7.6 shows the invariant large-R jet mass distribution as obtained from data and from MC using the nominal NLO cross sections for signal and background processes specified in Section 6.2. The data/MC agreement is not optimal, there is an overshoot in MC here which has been observed before in other analyses as well [144, 145]. It will reveal itself that the overshoot is present

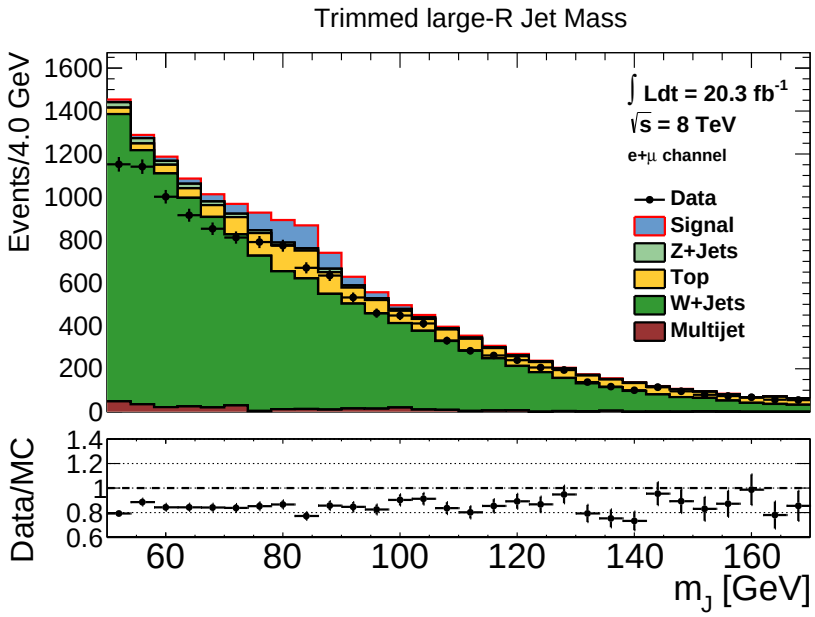
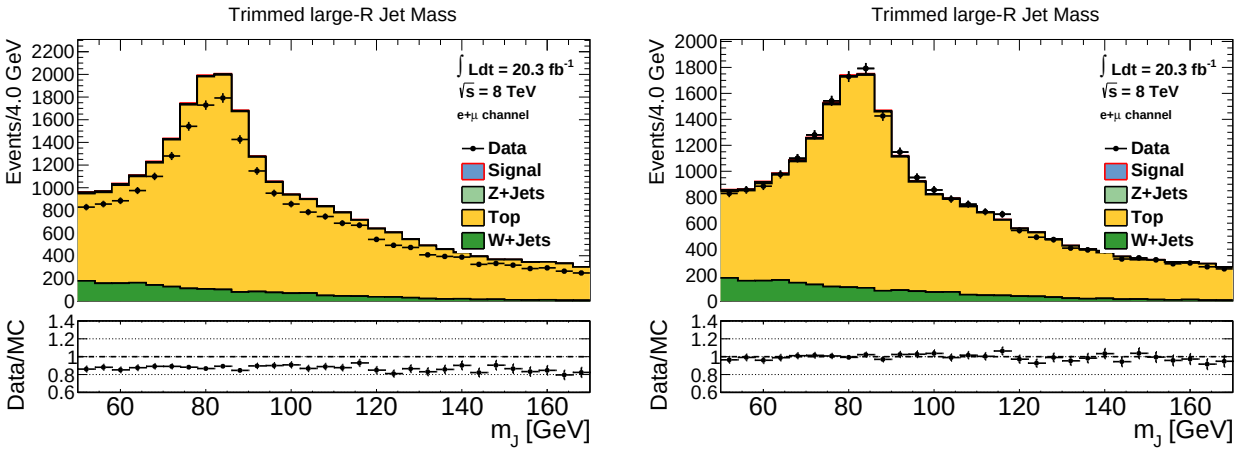


Figure 7.6.: Large-R jet mass distribution after event selection, without application of normalization scale factors to the top quark and W+Jets background. The MC prediction overestimates the data over the complete mass range.

for top quark production and also for W+jets. In order to obtain the correct normalization, a control region is defined for each of the two background processes. Such a control region should be enriched in the respective background and depleted in any other processes.
Assuming the data in this control region consists of a large fraction of this particular background process, a scale factor to correct the normalization can then be obtained by comparing the measured number of events in data with the number expected from MC simulation.
For the top quark background, this control region is obtained by requiring a b-tagged small-R jet outside of the large-R jet instead of the default jet veto.
The working point chosen here corresponds to a purity of b-tagged jets of 92.28 % and a corresponding efficiency of 70 %.
The b-tag requirement strongly suppresses all other relevant processes as can be seen comparing Figure 7.6 with Figure 7.7a. As expected, the large-R jet mass of the top quark background peaks also in the signal region due to the genuine W bosons from the decay of the second top quark (see Figure 8.2).
In the W+jets case, the control region is defined by the nominal event selection where events with a large-R jet falling into the signal mass region ($65\,\text{GeV} < m_{\text{Jet}} < 95\,\text{GeV}$) are vetoed. This selection has the advantages that the signal is strongly depleted and the top quark background is suppressed as well because of its mass peak in this region.



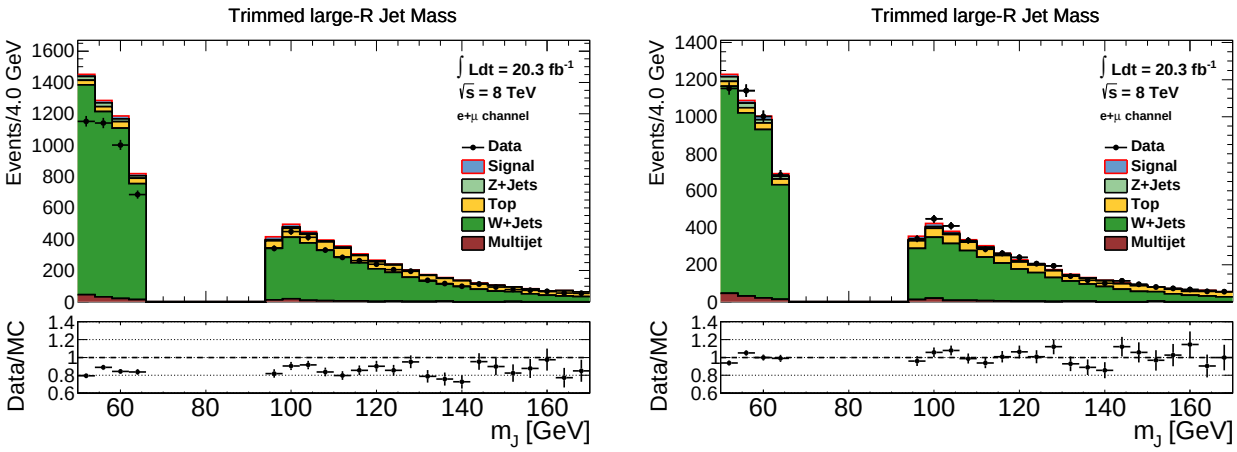
(a) (b)

Figure 7.7.: Large-R jet mass distribution in the top quark control region: Before (a) and after (b) the application of the normalization scale factor. The QCD multijet contribution is negligible in this control region and therefore not included in the plots.

Let n_X denote the total number of events stemming from process X in the control region. Then the normalization scale factor for $t\bar{t}$ and single top quark production is computed as:

$$SF(\text{Top}) = \frac{n_{\text{Data}} - (n_{\text{W+jets}} + n_{\text{Z+jets}} + n_{\text{WW+WZ}})}{n_{\text{Top}}} = 0.86 \pm 0.01. \tag{7.1}$$

The estimated contribution of QCD multijets is neglected in the calculation. This is justified by its small event yield and the additional high purity b-tag requirement which strongly suppresses events not containing a b-jet.



(a) (b)

Figure 7.8.: Large-R jet mass distribution in the W+jets control region: (a) before and (b) after the application of the normalization scale factor.

For the W+jets normalization, the same procedure is used, with a difference being that the newly available knowledge about the top template normalization can already be used as input in order to improve the accuracy of the estimation. Furthermore, the QCD multijet contribution is

also taken into account here. The W+jets normalization scale factor is thus calculated as:

$$SF(\mathrm{W}+\mathrm{jets}) = \frac{n_{\mathrm{Data}} - (SF(\mathrm{Top}) \cdot n_{\mathrm{Top}} + n_{\mathrm{Z+jets}} + n_{\mathrm{WW+WZ}} + n_{\mathrm{QCD}})}{n_{\mathrm{W+jets}}} = 0.84 \pm 0.02. \quad (7.2)$$

The W+jets normalization will be left to float freely in the actual WV cross section extraction fit, so that the scale factors obtained from the control region fit have no impact on the fit result. The scale factors are merely applied to data/MC comparison histograms to get a rough estimation of the data/MC agreement. In case of the top quark background, the SF will be used in the

Process	Top	W+jets
Scale factor	0.87 ± 0.01	0.84 ± 0.02

Table 7.1.: Summary of normalization scale factors extracted from the fit to data in the different control regions, including statistical uncertainties

fit but the full systematic uncertainty resulting from its application is taken into account (see Section 9.1.2). Figure 7.7 and 7.8 show the invariant large-R jet mass for top and W+jets in the respective control regions before and after application of the scale factors. The same distribution for the nominal selection after application of the scale factors can be found in Figure 7.9.

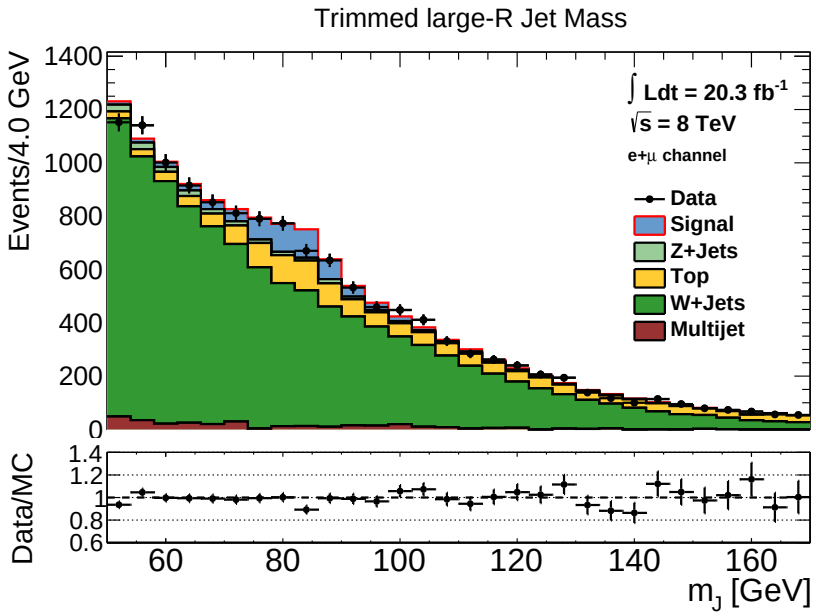


Figure 7.9.: The large-R jet mass distribution after event selection and following the application of normalization scale factors to top quark and W+Jets background. The data/MC agreement is greatly improved with respect to Figure 7.6.

Chapter 8

Extraction of the WW+WZ Cross Section

The fiducial WV production cross section is extracted as a ratio μ of measured WV events, relative to the SM-expected number of WV events in the fiducial volume obtained from MC simulation. The quantity μ is also referred to as "signal strength". This ratio is extracted by means of a binned maximum likelihood fit of the different signal and background templates of the large-R jet mass shown in Figure 8.2 to the corresponding histogram obtained from data. The absolute number of measured WV events N^{WV} can then be computed as:

$$N^{\mathrm{WV}} = \mu \cdot \sum_{b} \nu_b^{\mathrm{sig}}, \tag{8.1}$$

with ν_b^{Sig} being the expected numbers of signal events in bin b.
Both lepton channels are summed in the fit as a single channel. This approach increases the statistics for the individual systematic uncertainties and simplifies the fit procedure. It has been verified that a simultaneous fit in both channels would improve the sensitivity by a negligible amount only.
Because of the small contribution of Z+jets and the similar shape (see Figure 8.4a), the templates from W+jets and Z+jets processes are summed and used as one template for the fit. The ratio of the cross sections of W+jets/Z+jets is taken to be the expected SM value of $\sigma^{\mathrm{NLO}}_{\mathrm{W+jets}}/\sigma^{\mathrm{NLO}}_{\mathrm{Z+jets}}$=10.7 (see Section 6.2.2).
For the fit, a large mass range of 50-170 GeV is used, allowing the contributions of the V+jets and top quark backgrounds to be constrained in the almost signal-free sideband regions. The normalization of V+jets is allowed to float freely in the fit. The fiducial cross section $\sigma_{\mathrm{fid}}^{\mathrm{WW+WZ}}$ is then extracted from the measured number of signal events with the following equation:

$$\sigma_{\mathrm{fid}} = \frac{N^{\mathrm{WV}}}{\mathcal{L}_{\mathrm{Int}} \cdot D_{\mathrm{fid}}}, \tag{8.2}$$

where D_{fid} is a factor that accounts for the difference of events that are actually produced in the fiducial phase space and the events passing all selection criteria on reconstruction level.

8.1. Fiducial Volume

The fiducial volume is a phase space region which is optimized to yield a maximum signal to background ratio after event selection.
It is obtained for Monte Carlo by applying an event selection to particle level quantities as close as possible to the event selection applied on reconstruction level.
The first requirement for the fiducial volume is that the event really be a true semileptonic WV event, with the V decaying hadronically and the W decaying to lν, where l is either an electron or a muon. Tauonic decays are excluded from the fiducial volume. However, tauonic W decays can contribute to the reconstruction level event selection if an electron or muon is created in the tau decay.
For the simulated lepton (electron or muon), the following conditions are applied:
The minimum lepton p_T must be greater than 30 GeV and its η-range is restricted to be within $|\eta| < 2.47$. The lepton p_T here is calculated as the vectorial sum of the transverse momentum of the lepton itself and the transverse momenta of all photons in a cone of $\Delta R \leq 1$ around the lepton which are not stemming from hadronic decays.
The simulated missing transverse momentum is computed as the vectorial sum of the transverse momentum of all non-interacting particles (neutrinos). Its magnitude is required to be greater than 50 GeV, corresponding to the reconstruction level cut on $E_{\mathrm{T}}^{\mathrm{Miss}} > 50\,\mathrm{GeV}$.
Simulated particle jets in general are reconstructed by running the anti-k_T algorithm on all stable MC particles (particles with an average decay length of more than 1 cm) with the exception of muons and neutrinos.
The small-R jets are reconstructed with a radius parameter of $R = 0.4$ and the large-R jets with a radius parameter of $R = 1$. Small-R jets with a p_T of less than 25 GeV are removed from the jet collection.
The overlap removal procedure between small-R jets, electrons and muons is the same as is applied on reconstruction level (see Section 6.1.4).
Exactly one large-R jet with $p_T > 200\,\mathrm{GeV}$, which does not overlap with a selected lepton, and which is inside an η-range of $|\eta| < 2.0$ is required.
Events containing a small-R jet which is not overlapping with the large-R jet in $\Delta R \leq 1$ are rejected. Finally the event is rejected if the selected large-R jet has a mass of less than 50 GeV.

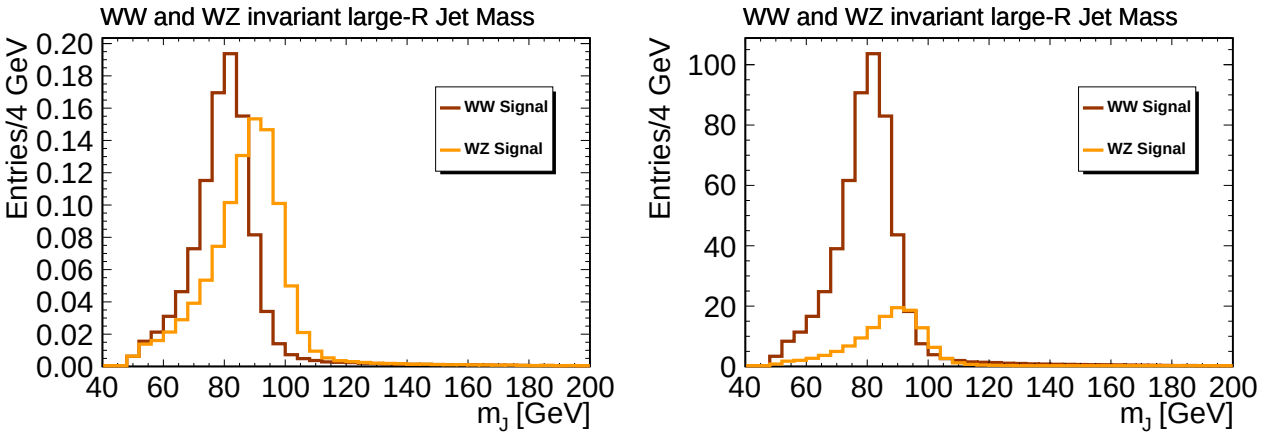
8.2. Definition of D_{fid}

The fact that the signal yield in this analysis is the sum of two different processes with different individual cross sections, acceptances and correction factors makes it necessary to introduce a factor D_{fid} to take this into account. D_{fid} is defined as:

$$D_{\mathrm{fid}} = f^{\mathrm{WW}} \cdot C^{\mathrm{WW}} + (1 - f^{\mathrm{WW}}) \cdot C^{WZ}, \tag{8.3}$$

with C^{WV} being the WV-reconstruction efficiency in the fiducial volume:

$$C^{\mathrm{WV}} = \frac{N_{\mathrm{sel}}^{\mathrm{WV}}}{N_{\mathrm{fid}}^{\mathrm{WV}}}. \tag{8.4}$$



(a) (b)

Figure 8.1.: (a) Shape comparison of WW and WZ templates after full event selection. Both distributions are normalized to unit area. (b) Comparison of the actual contribution of the two processes to the total expected signal events. Both distributions are normalized to the expected yield in 2012 data.

The quantity f^{WW} expresses the fraction of WW events of all signal events (WW+WZ) in the fiducial volume:

$$f^{\mathrm{WW}} = \frac{\sigma_{\mathrm{WW}} \cdot \mathrm{BR}^{\mathrm{WW}}_{\mathrm{l}\nu\mathrm{qq}} \cdot \mathrm{A}^{\mathrm{WW}}}{\sigma_{\mathrm{WW}} \cdot \mathrm{BR}^{\mathrm{WW}}_{\mathrm{l}\nu\mathrm{qq}} \cdot \mathrm{A}^{\mathrm{WW}} + \sigma_{\mathrm{WZ}} \cdot \mathrm{BR}^{\mathrm{WZ}}_{\mathrm{l}\nu\mathrm{qq}} \cdot \mathrm{A}^{\mathrm{WZ}}}, \tag{8.5}$$

with the acceptance:

$$A^{\mathrm{WV}} = \frac{N^{\mathrm{WV}}_{\mathrm{fid}}}{N^{\mathrm{WV}}_{l\nu\mathrm{qq}}}. \tag{8.6}$$

The value of f^{WW} is obtained under the assumption that the ratio of σ_{WW} to σ_{WZ} is equal to the SM expectation. The numerical value of the cross section ratio, computed with the nominal WW/WZ signal samples, amounts to 2.47.

8.3. Fit Procedure

The number of measured signal events in the fiducial volume is extracted via a binned maximum likelihood fit of MC histogram templates to data. Systematic uncertainties are introduced into the fit via a set of nuisance parameters $\vec{\theta}$. The likelihood function is defined as:

$$L(\mu,\ \vec{\theta}) = \prod_b \mathrm{Poisson}(n_b|(\nu_b^{\mathrm{Bkg}} + \mu\nu_b^{\mathrm{Sig}})(\vec{\theta})) \cdot \prod_i f_i(\theta_i), \tag{8.7}$$

with the parameter of interest μ, which is the ratio of the observed to the expected SM signal yield. The parameters ν_b^{Bkg} and ν_b^{Sig} are the expected numbers of background and signal events in bin b, respectively. The quantity n_b is the observed number of events in data in bin b.
The nuisance parameters θ_i are a priori constrained by Gaussian profiles f_i and can potentially be further constrained during the fit.
The magnitude of the uncertainty of the normalization of the top quark background is obtained as the deviation of the extracted scale factor quoted in Table 7.1 from one. The relative contribution

Parameter	Value (in %)
A^{WW}	0.293 ± 0.004 (stat.)
A^{WZ}	0.302 ± 0.009 (stat.)
C^{WW}	58.8 ± 1.5 (stat.)
C^{WZ}	63.4 ± 3.2 (stat.)
f^{WW}	83.0 ± 2.7 (stat.)
D_{fid}	59.6 ± 1.3 (stat.) ± 7.6 (syst.)

Table 8.1.: Results of correction factors necessary for the calculation of the measured and theory fiducial cross section via Equation 8.2 and 8.10, respectively. Numbers were evaluated using the nominal WW and WZ Monte Carlo samples.

of $\mathrm{t\bar{t}}$ and single top processes to the entire top quark background can be seen in Figure 8.3a. Figure 8.3b shows the single top quark background broken down further into the individual production channels.
In order to avoid any bias from fixing the V+jets normalization to a certain value, the normalization is, contrary to the top quark normalization, allowed to float completely freely in the fit. This is implemented by treating the normalization of the V+jets template as an unconstrained normalization factor $NF_{\mathrm{V+jets}}$. Since this background is the largest, it can be efficiently constrained during the fit by the sideband regions .
The template histograms used for modeling signal and background processes are based on finite-statistics MC samples or a finite-statistics control region in case of the QCD multijet template. To account for the associated bin-by-bin statistical uncertainties, additional nuisance parameters are introduced, one for each histogram bin. Each of these nuisance parameters represents the statistical errors in its associated bin from each of the components added in quadrature (Barlow-Beeston method [146, 147]). The *significance* of the WW+WZ process observation, i.e. the confidence level with which the background-only hypothesis can be rejected is obtained by a profile likelihood ratio method. For this, two maximum likelihood fits are performed, in one of them the signal is fixed to zero and in the other the signal can float freely. The nuisance parameters are all allowed to float within their gaussian constraints. The likelihood ratio λ is then given as:

$$\lambda = \frac{L(\mu = 0, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})}, \tag{8.8}$$

where the hat-notation signifies that $L(\hat{\mu}, \hat{\theta})$ is the likelihood function maximized by the signal and all nuisance parameters free to float whereas $L(\mu = 0, \hat{\hat{\theta}})$ is the likelihood function which is maximized by fixing the signal to zero and leaving only the nuisance parameters to float.
It is known as Wilks's theorem [148] that the quantity $-2\ln\lambda$ is, for the case of no signal, asymptotically distributed like a χ^2-distribution with one degree of freedom. This means the signal significance can be estimated as [149]:

$$\mathrm{sig(WW + WZ)} = \sqrt{-2\ln\lambda}. \tag{8.9}$$

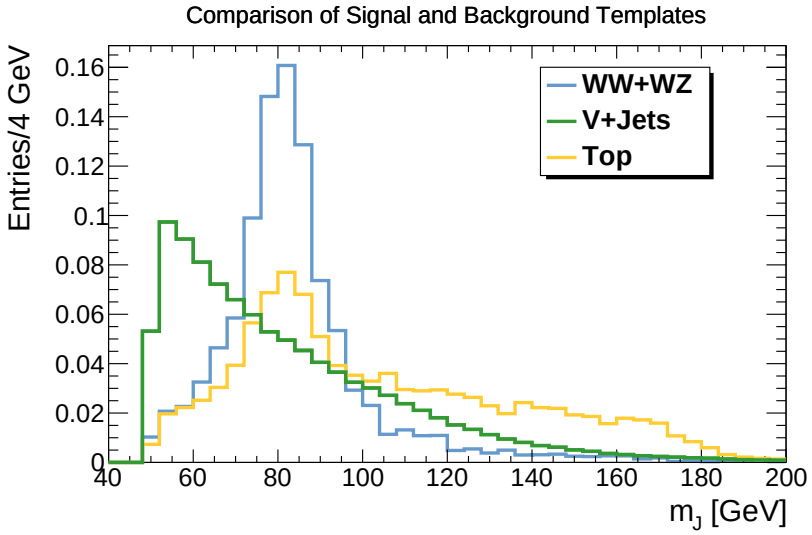
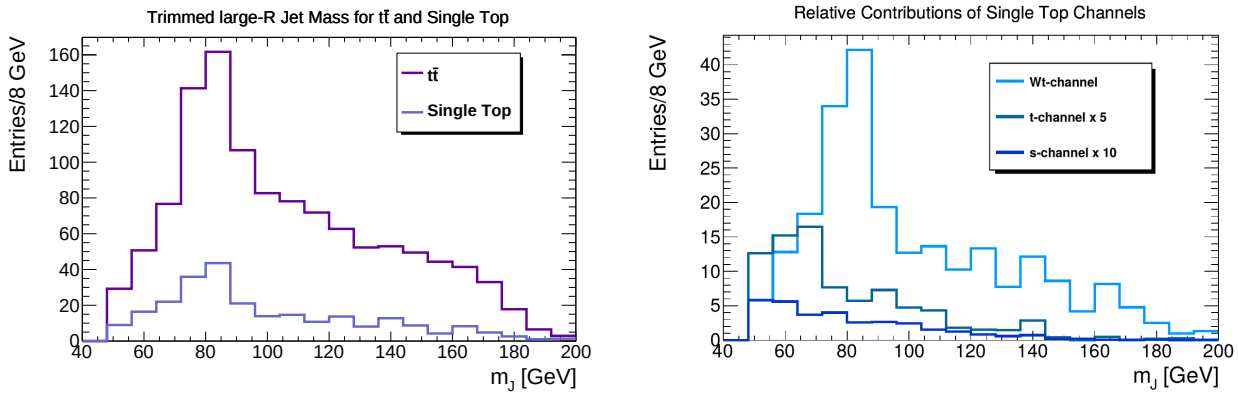


Figure 8.2.: Shape comparison of the signal template and the two most important background templates. For better clarity and because it plays a sub-leading role, the QCD multijet templates is shown separately in Figure 7.5. The top quark background can be seen to peak directly in the signal region because of the real W bosons that are present here. This makes it an important background despite its relatively small normalization.



(a) (b)

Figure 8.3.: (a) Estimated contribution of $t\bar{t}$ and single top processes to the total background. The fraction of single top quark production in the entire top background is $\sim 18\,\%$. (b) Relative contribution of single top quark processes broken down into individual production mechanisms.

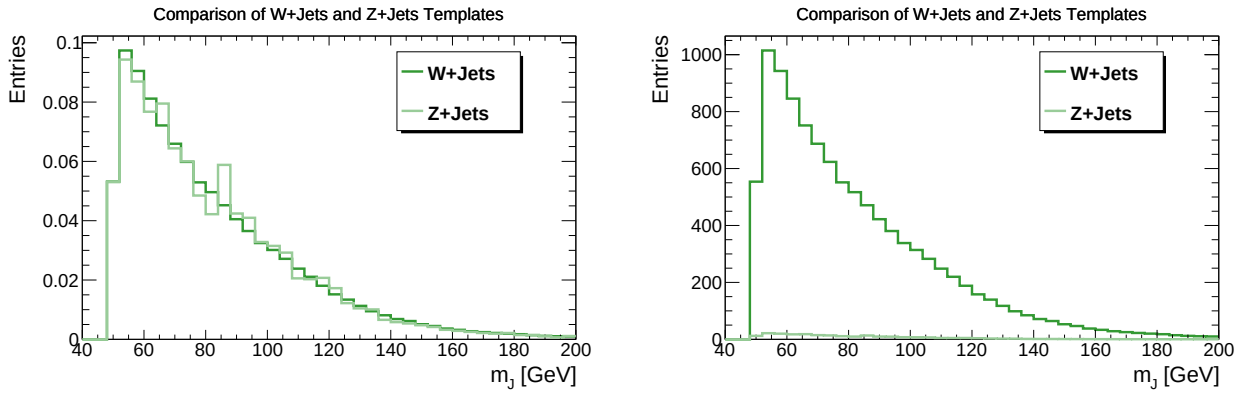
8.4. Theoretical fiducial Cross Section

With l=e, μ, the theoretical fiducial cross section can be expressed as:

$$\sigma_{\text{fid}} = (\sigma_{\text{qq}\to\text{WW}} + \sigma_{\text{gg}\to\text{WW}}) \cdot \text{BR}(\text{WW} \to \text{l}\nu\text{qq}) \cdot \text{A}^{\text{WW}} + \sigma_{\text{WZ}} \cdot \text{BR}(\text{WZ} \to \text{l}\nu\text{qq}) \cdot \text{A}^{\text{WZ}}. \quad (8.10)$$

The numerical values of the input parameters can be found in Table 8.1 and Sections 4.1 and 6.2.1. With these values, the theoretical fiducial cross section evaluates to:

$$\sigma_{\text{fid}}^{\text{theo}}(\text{WW} + \text{WZ}) = (57.9 \pm 2.9)\,\text{fb}. \quad (8.11)$$



(a) (b)

Figure 8.4.: (a) Shape comparison of W+jets and Z+jets templates, both distributions are normalized to unity. Evidently the two templates are nearly identical. (b) Relative contribution of W+jets and Z+jets processes to the final event yield. The Z+jets contribution is only 2.2 % of the one expected from W+jets. Both distributions are normalized to the expected event yield in 2012 data.

The uncertainty on the theory cross section results from the choice of the renormalization and factorization scale and the theory uncertainty on the PDFs that were used to generate the MC sample. All of these uncertainties affect the inclusive cross section as well as the acceptance. The last uncertainty results from the limited MC statistics. A breakdown of these uncertainties is given in Table 8.2. The total uncertainty is evaluated by adding the single contributions in quadrature.

$(\sigma_{qq\to WW} + \sigma_{gg\to WW}) \cdot BR(WW \to l\nu qq) \cdot A^{WW} + \sigma_{WZ} \cdot BR(WZ \to l\nu qq) \cdot A^{WZ}$ [fb]							
Generator	Cross Section	Total Unc.	Incl. Cross Section		Acceptance		
	fb	fb	scale	PDF	scale	PDF	MC stat.
MC@NLO	57.9	2.7	3.1 %	1.8 %	1.2 %	2.4 %	1.3 %

Table 8.2.: Breakdown of the individual contributions to the theory uncertainty on the fiducial WW+WZ cross section. The fiducial volume includes electron and muon final state without intermediate τ-leptons. Scale uncertainty includes the scale uncertainty on the gg→WW cross section, as provided by [55].

Chapter 9

Systematic Uncertainties

The total systematic uncertainty on the measured WW+WZ cross section consists of three contributions:

1. **Uncertainty on the measured signal yield**
2. **Uncertainty on the integrated luminosity**
3. **Uncertainty on D_{fid}**

The uncertainties associated to all of those three contributions are implemented as nuisance parameters in the cross section extraction fit. In this way, correlations between different sources of systematic uncertainty are taken into account correctly.

9.1. Systematic Uncertainty on the Signal Yield

The signal yield uncertainty is estimated with the following method:
For each of the relevant systematic uncertainties, additional varied invariant mass MC template histograms for all signal and background processes are produced by running the complete analysis again, varying the respective uncertainty-affected parameter by one standard deviation up and down. The effect of the varied templates on the maximum likelihood fit is then obtained via the introduction of Gaussian constrained nuisance parameters.
Usually the sources of uncertainties, e.g. the jet energy scale, can be varied in two directions, up and down. In this case the two varied templates are later directly used in the fit. In cases for which this is not true, e.g. MC generator uncertainties which by construction exist only for one direction, the second template is obtained by a symmetrization procedure where the symmetrized histogram is obtained as:

$$h_{\text{Symmetrized}} = 2 \cdot h_{\text{Nominal}} - h_{\text{Varied}}.$$

Since many of the systematic MC samples suffer from rather limited statistics, a template smoothing method [150] which helps to reduce the effects from statistical fluctuations is applied.

The smoothing is a sequence of several consecutive steps. First, systematic uncertainties which are not obtained via symmetrization of a template are averaged over up and down variations, leading to symmetric distributions as well.
In a second step, for a given uncertainty, the varied template histogram is scaled to the same integral as the nominal template. Then the ratio histogram of the varied divided by the nominal histogram is built and a smoothing algorithm (353QH - algorithm [151]) is applied to the histogram entries.
The smoothed ratio histogram is finally multiplied again with the nominal histogram and then scaled back to the original integral.
The effect of the single steps can be seen in Figure 9.1. The left plot shows the original histograms of the up and down variation divided by the nominal histogram. In the middle plot, the same two histograms are shown after averaging, with the original ones still depicted as dashed lines. The right plot shows the final smoothed histograms. The statistical uncertainty of the original templates is indicated as yellow band around zero. The numbers in the upper left corner show the difference in the integral of the up and down variation compared to the nominal selection. Statistical uncertainties of the histograms used for the evaluation of systematic uncertainties are accounted for by the introduction of an additional nuisance parameter for each bin.

9.1.1. Systematic Uncertainties of Physics Objects

All uncertainties on physics objects are evaluated by re-running the analysis for each systematic effect. Particular recommended ATLAS tools then provide the variation of the parameter in question (e.g. the large-R jet p_T) within one standard deviation. This varied quantity is used throughout the rest of the analysis instead of its nominal value.
Uncertainties related to both lepton flavours arise from the trigger and reconstruction efficiency as well as the energy scale and resolution. Furthermore, the electrons are affected by another uncertainty on the isolation and identification [123, 152–155].
The determination of the $E_{\mathrm{T}}^{\mathrm{Miss}}$ present in an event suffers from uncertainties on the resolution and energy scale of the soft term (see Chapter 6.1.5). Therefore these two quantities are varied within their respective one standard deviation uncertainty [129].
Systematic uncertainties of jets are different for small-R jets and large-R jets.
Small-R jets are affected by a large number of systematic uncertainties on different properties of the jet. The most important contribution results from the uncertainty on the jet energy scale (JES) [108, 156], arising from 18 different sources, 14 of them related to calorimetric measurements, 3 related to quark flavours involved and 1 to punch through to the muon system. A summary of all the 18 components can be found in Appendix C.
Another uncertainty is introduced by applying a cut on the jet-vertex fraction (see Section 6.1.3), which is taken into account by varying the cut value up and downwards in a certain, recommended range [107].
The last uncertainty on small-R jets comes from the fact that the jet energy resolution (JER) is too good in MC when compared to data. To account for this fact, the JER on MC is artificially worsened by the application of a p_T and η dependent, recommended smearing factor to the jet four-vector [156].
As discussed in Section 5.4, the pile-up distribution in MC is reweighted to the final distribution observed in data. The uncertainty associated to this weighting procedure is included by varying the nominal pile-up scale factor ($1/1.09 \approx 0.92$) to 0.88 and 0.94.
In general, small-R jets have an effect on the event selection only via the jet veto requirement. For that, the actual value of the jet energy is essentially irrelevant. The only place where the JES

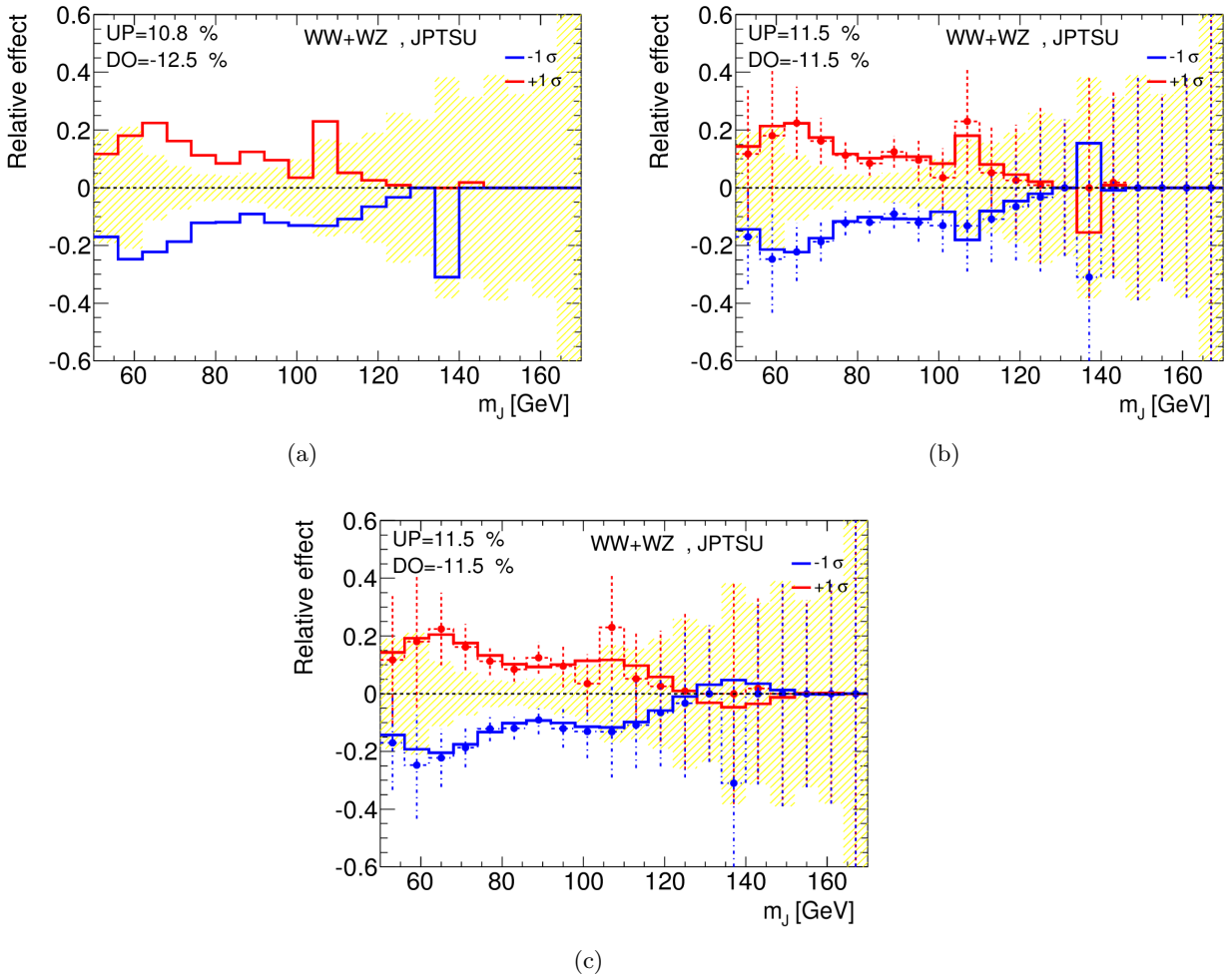
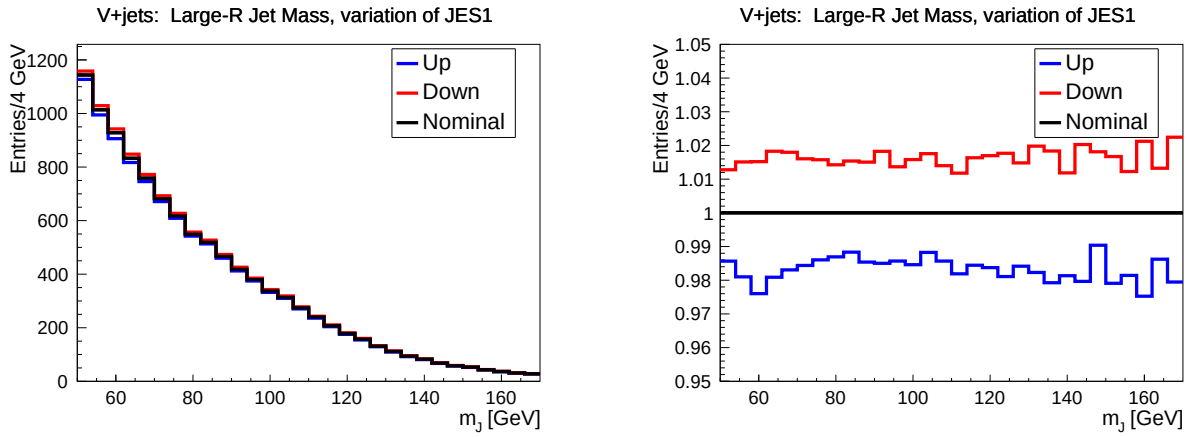


Figure 9.1.: Comparison of the three treatments applied to varied templates for the WW+WZ large-R jet p_T uncertainty.
(a) Raw ratio histograms, neither averaging nor smoothing applied.
(b) Ratio histograms after averaging.
(c) After averaging and smoothing, the original distribution can still be seen as dashed line and points with dashed error bars.

variation of small-R jets could affect the jet veto, is in the object definition of jets. As described in Section 5.5.2, a jet is only counted as jet, if its p_T exceeds 25 GeV. That means if a given jet is only narrowly below or above the threshold of 25 GeV, it could be counted or dismissed as a jet, depending on the JES variation. For this reason the systematic uncertainty attributed to small-R jets is in general very small in this analysis. This can also be seen in Figure 9.2 that shows the largest JES uncertainty component together with the nominal histograms for the V+jets background. All uncertainties on small-R jets are also propagated to the evaluation of missing transverse momentum in the event.
In contrast to the small-R jets, uncertainties related to large-R jets have a large and direct impact on the final invariant mass distribution.
As for any other uncertainty, uncertainties associated to large-R jets are assessed by varying the respective parameter by one standard deviation. As in the small-R jets case, there are



(a) (b)

Figure 9.2.: (a) Varied template histograms associated to the largest JES uncertainty component together with the nominal template. (b) Ratio of the variations and the nominal template.

the systematic uncertainties on the scale and resolution of the jet energy to be considered. In addition, there is a systematic uncertainty considered for the jet mass as well. The magnitude of the mass and energy scale uncertainty has been determined based on a data/MC double ratio method, described in Ref. [117] and is applied in an η and p_T dependent way to the jets. The uncertainties on the mass and energy resolution are obtained by artificially smearing the resolution obtained in the measurement by 20% of its nominal value.

9.1.2. Modeling and Normalization Uncertainties

WW/WZ Modeling

Uncertainties resulting from the choice of the renormalization and factorization scale are taken into account by the comparison of the nominal MC@NLO sample with other dedicated MC@NLO samples. These additional samples were generated with the renormalization scale varied by a factor two and by a half with respect to their nominal values (see Chapter 6.2.1).
MC generator uncertainties are assessed by comparison of the nominal sample with the Sherpa sample and the Powheg+Pythia 8 sample. Since MC@NLO and the two other generators differ in the implementation of the matrix element, as well as the parton shower- and hadronization model, this comparison can be used as uncertainty for those three generation uncertainties at once.
For this, the Sherpa template is compared directly to the nominal MC@NLO sample. The Powheg+Pythia 8 sample, however, is obtained as the sum of the Powheg WW sample and the Sherpa WZ sample, since the Powheg WZ sample is lacking some of the Z boson decay modes ($l\nu\nu\nu$, $l\nu ll$). The systematic variation is not obtained by directly comparing Powheg to the nominal sample but by dividing the Powheg sample by the Sherpa sample and applying the resulting difference to the nominal (MC@NLO) histogram. This is done in order to separate the effect of the narrow width approximation (NWA) which is used in the nominal MC@NLO sample but not in the Powheg and Sherpa samples (compare Section 6.2.1) from the rest of the generator differences and to avoid double counting the difference associated to the NWA.

PDF uncertainties on the shape and normalization of the signal templates have been assessed by comparing the nominal m_J template with templates obtained by variation of the PDFs using LHAPDF6 [157].
For that, the signal events, originally generated using the CT10 PDF-set, were reweighted to the 52 (26 up/down each) CT10-error sets aside from the central set. From these variations, a single so-called envelope histogram, describing the uncertainty at 68 % C.L. was extracted for up- and down-variations, respectively. The uncertainty on the shape as well as the normalization is less then 3 % over the complete mass range. For this reason, it is neglected in the fit.

V+Jets Modeling

Since the relative contribution of Z+jets to the V+jets background is so small (see Figure 8.4b), a separate evaluation of associated systematic uncertainties is not performed. This is justified by the fact that the expected systematic uncertainty from this contribution is much smaller than already its statistical uncertainty. The nominal Z+jets templates are added to the varied W+jets templates for the fit procedure. Important systematic uncertainties for the modeling of W+jets production arise from several sources:

- MC generator modeling
- Renormalization scale
- Factorization scale
- CKKW matching scale
- Parton shower modeling

Uncertainties on the generator modeling are assessed by comparing the default Sherpa sample with the Alpgen+Pythia sample mentioned in Section 6.2.2. This Alpgen generated sample is known to suffer from a severe bug. This bug results in an underpopulated phase space corresponding to a ΔR between the two hardest partons in the range of $0.4 < \Delta R < 0.7$. In order to compensate for this mismodeling, the Alpgen sample is reweighted to the nominal Sherpa sample using as the reweighting variable the ΔR between the first two subjets found in the trimming procedure.
The comparison of the Alpgen+Pythia sample with the Sherpa sample already accounts for a differences in parton shower modeling which is why no additional, dedicated sample is used for this uncertainty. For the evaluation of systematic uncertainties resulting from the choice of the CKKW matching, renormalization and factorization scale, six additional Sherpa samples have been generated on particle-level only (no detector simulation).
In these samples, the renormalization and factorization scale are varied by a factor 0.5 and 2 with respect to their nominal value and the CKKW matching scale is changed to 15 GeV and 30 GeV from its nominal value of 20 GeV.
Since no full detector simulation samples are available here, the varied templates are obtained by comparing the particle-level templates from these samples with the particle-level nominal template. Events are selected by the requirement to pass the selection criteria defining the fiducial phase space. The varied templates used in the fit are then obtained as the nominal (full detector simulation) Sherpa histograms multiplied by the ratio of the particle-level varied over nominal histograms.

Top Modeling

As it can be seen in Figure 8.3a, by far the largest part of the top quark background is made up of $t\bar{t}$ events. For this reason the modeling uncertainties are obtained for $t\bar{t}$ samples only and, like in the V+jets case, the nominal single top quark template is added to the varied $t\bar{t}$ templates.
Three major uncertainties contribute to the top background uncertainty. The first one is the modeling uncertainty depending on the MC generator which is employed. This is estimated by comparing the MC@NLO+Herwig sample to the Powheg+Herwig sample and applying the difference between the two to the nominal Powheg+Pythia sample.
The second large modeling uncertainty arises from parton shower modeling which is assessed by comparing the nominal Powheg+Pythia sample to the Powheg+Herwig sample. The last uncertainty is the uncertainty on the amount of initial and final state radiation. Two dedicated samples are used for the estimation of this uncertainty (see Section 6.2.3).
As described in Chapter 7, a scale factor for the top quark background normalization had to be introduced. It is accounted for the full systematic uncertainty resulting from this procedure by associating an uncertainty to the normalization of the top quark template corresponding to the deviation of the scale factor from one (14 %), instead of just using the the uncertainty from the most recent theoretical cross section calculations of only 6 % (compare Equation 4.13).

QCD Multijet Modeling

In addition to an overall estimated uncertainty on the normalization of the QCD multijet contribution of 50 %, three further sources of systematic uncertainties are considered. The first uncertainty results from the choice of the particular fit range in order to extract the normalization of the QCD contribution. This uncertainty has been assessed by varying the fit range from its nominal value of [0,100] GeV to [5,100] GeV and [20,150] GeV, the binning itself remains unchanged. Since this variation only has an effect on the normalization and not on the shape of the QCD multijet template which is moreover on the level of only 3 %, this uncertainty is neglected in the fit.
The second uncertainty results from the choice of the parameters in the definition of the inverted lepton selection. This is estimated by re-deriving the multijet template for two varied QCD lepton definitions where the calorimeter isolation is varied in the electron case and the cut on the d_0 significance is varied in the muon case:

Lepton Variation 1:

- $1.5\,\text{GeV} < E_T(R = 0.3) < 17.5\,\text{GeV}$
- $|d_0/\sigma(d_{0,\text{track}})| > 4.5$

Lepton Variation 2:

- $3\,\text{GeV} < E_T(R = 0.3) < 25\,\text{GeV}$
- $|d_0/\sigma(d_{0,\text{track}})| > 3.5$

The third source of systematic uncertainty is the size of the contribution of real lepton (not faked by jets or emerging from heavy flavour decays) events from the known backgrounds and the signal to the data template obtained with the inverted lepton selection. This uncertainty is evaluated by varying this contribution by 20 % up and down.

9.2. Systematic Uncertainties on the integrated Luminosity and on D_{fid}

The uncertainty on the integrated luminosity amounts to ±1.9%. Details about its derivation can be found in Ref. [158]. The uncertainty on the integrated luminosity is incorporated in the fit as overall uncertainty on the normalization of the top quark template, all other normalizations are already free floating and as such not affected by the luminosity uncertainty.
Systematic uncertainties on D_{fid} are evaluated by calculating D_{fid} using the varied $N_{\text{sel}}^{\text{WV}}$ for each source of systematic uncertainty.
The PDF uncertainty on D_{fid} is estimated by recalculating D_{fid} for each of the 52 CT10 PDF-error members. For this, $N_{\text{fid}}^{\text{WV}}$, $N_{\text{sel}}^{\text{WV}}$ and $N_{\text{l}\nu\text{qq}}^{\text{WV}}$ are varied simultaneously. The correlation of the PDF uncertainties is taken to be 100 % between WW and WZ which results in the most conservative possible error. For each pair of PDF-members the symmetrized deviation from the nominal D_{fid} value is summed in quadrature, leading to a relative error on D_{fid} of 1.6 %. This contribution is very small compared to other uncertainties and for this reason is neglected in the fit.
The impact of the systematic uncertainties of D_{fid} on the fiducial cross section is obtained by implementing the uncertainties on D_{fid} directly in the fit procedure. For this, the gaussian constrained nuisance parameters associated to a particular source of uncertainty S receive an additional overall uncertainty corresponding to the range of:

$$[\frac{D_{\text{fid}}^{\text{Down}}(S)}{D_{\text{fid}}^{\text{Nom}}},\ \frac{D_{\text{fid}}^{\text{Up}}(S)}{D_{\text{fid}}^{\text{Nom}}}]. \tag{9.1}$$

This procedure has the advantage that correlations between the signal shape and the value of D_{fid} are automatically taken into account correctly.

Chapter 10

Results

10.1. Cross Section Results

10.1.1. Expected Performance

The performance of the cross section fit can be estimated without making use of real data by using so called "Asimov data" [149]. This means that the data distributions are mimicked by the sum of the nominal MC signal and background distributions instead of stemming from real measured events. The result is an artificial data distribution that looks exactly like the SM expectation for signal plus background.
The great advantage of Asimov data is that the general sensitivity of the measurement and fit procedure including all systematic uncertainties can be evaluated without having to rely on a particular measured dataset and the possibility of being biased by the data result.
The fit to Asimov data naturally has to deliver exactly the SM expectation as result for the signal strength μ. The important quantity here is the associated uncertainty to this result which is a measure for the expected sensitivity. The general form of the result of the fit to Asimov data is the following:

$$\mu = 1 \pm t, \tag{10.1}$$

with t being the quadratic sum of statistical and systematic uncertainty obtained from the fit. The nuisance parameters can potentially be constrained further during the fit to Asimov or observed data. This is for example the case when a particular systematic uncertainty is estimated too large. During the fit to Asimov data the fit procedure finds that it is impossible (or highly unlikely) to obtain the data distribution with the nuisance parameter associated to this uncertainty shifted within its predefined uncertainty. For this reason the nuisance parameter is then constrained to a smaller than the predefined range. This behaviour helps to constrain for example the top quark and V+jets normalization by taking into account the jet-mass sideband regions.
Typically no large constraints on the nuisance parameters should be observed on Asimov data since this would signify that the systematic uncertainties provided by the dedicated ATLAS combined performance groups are too large.

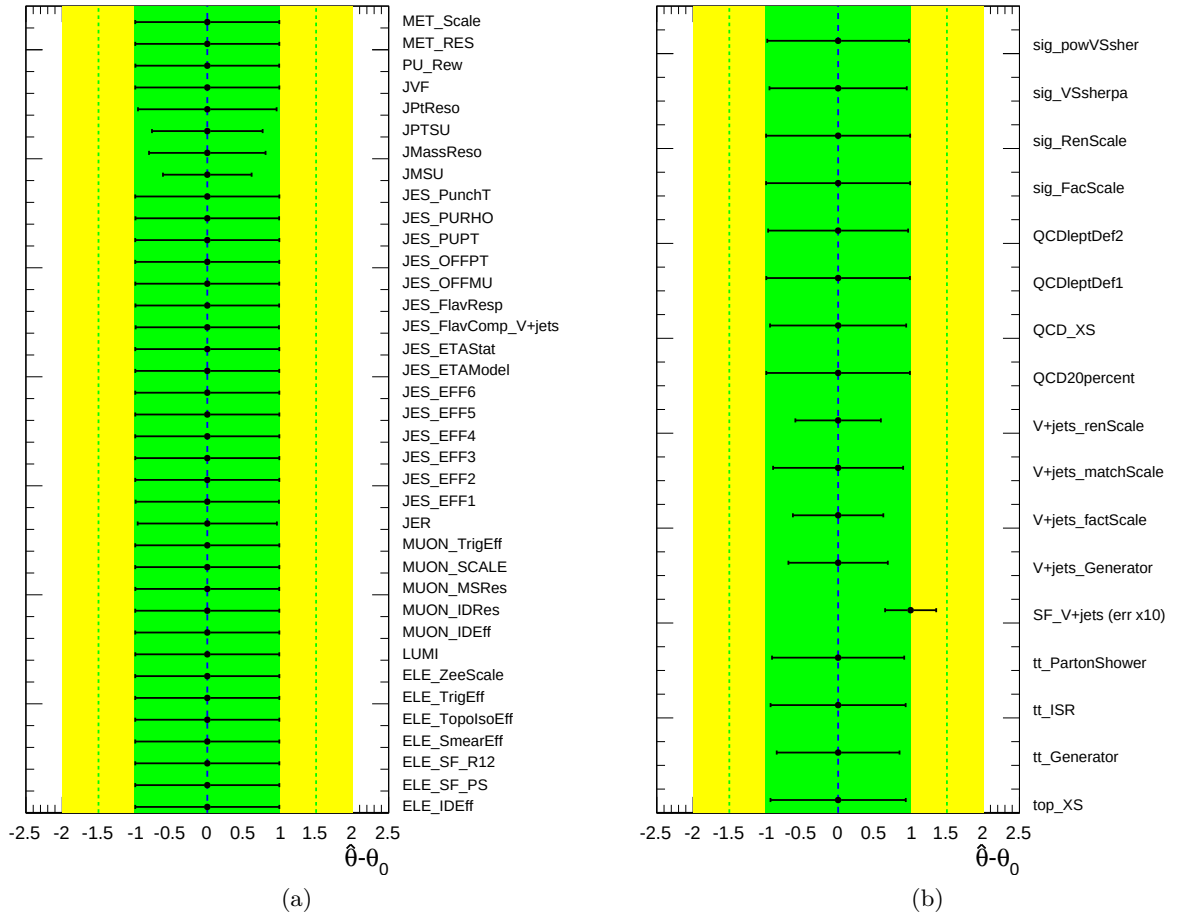


Figure 10.1.: Constraints on nuisance parameters systematic uncertainties obtained from a fit to Asimov data split up into detector (a) and modeling uncertainties (b). The plot shows the deviation of the best-fit value $\hat{\theta}$ from the nominal value θ_0 for all different nuisance parameters θ_i. Most systematic uncertainties are unconstrained, the central value is zero for all nuisance parameters as expected for Asimov data. The scale factor for the V+jets normalization (labeled SF_V+jets) is not a nuisance parameter and is expected to be 1.

Quantity	Expected	Observed
μ	1	0.57
Total error on μ	0.46 (46 %)	0.44 (77.2 %)
Error on μ (stat only)	0.20 (20 %)	0.20 (33 %)
Significance	2.4 σ	1.4 σ

Table 10.1.: Results for the expected and observed sensitivity and significance. The expected quantities are obtained by a fit to Asimov data while the observed values stem from a fit to real data.

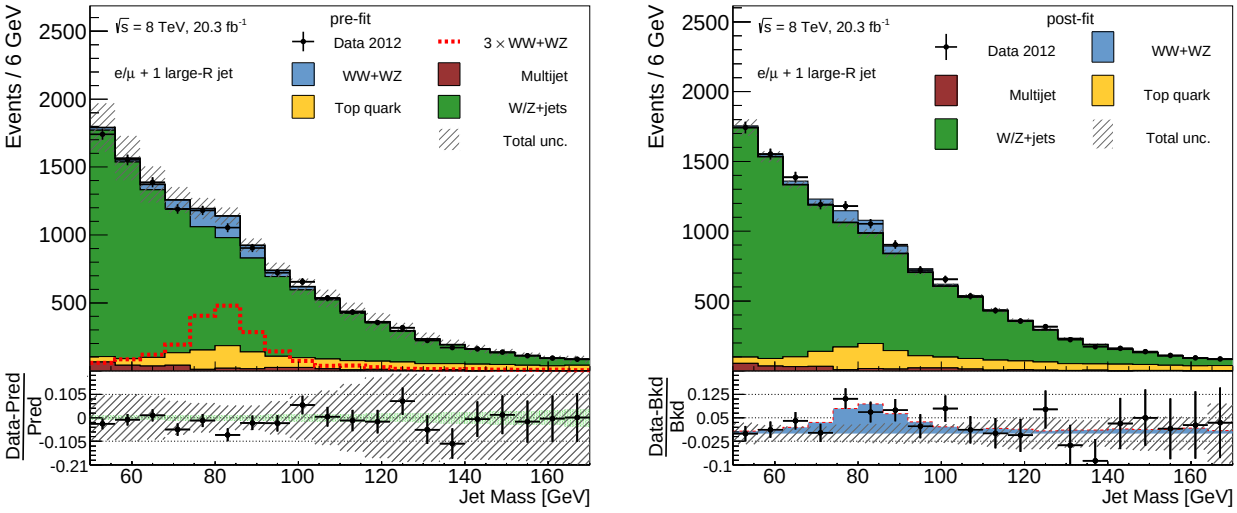
The error on the signal strength is evaluated using the same method as is used for the estimation of the significance. The likelihood ratio $\lambda = L(\mu, \hat{\hat{\theta}})/L(\hat{\mu}, \hat{\theta})$ is scanned over many different values of μ in the numerator which results in a curve shown in Figure 10.4. The lower and upper bounds of μ are then obtained as the intersections of this curve with the line at $\Delta[-\log(L)] = 0.5$, corresponding to a significance of one standard deviation. Figure 10.1 shows the constraints on the different nuisance parameters obtained from the fit to Asimov data. The exact meaning of

the names of the nuisance parameters is explained in Appendix C. There are no large constraints observed and all central values are at zero as is expected for Asimov data.

10.1.2. Result on Data

The fit to observed data yields a signal strength of $\mu = 0.57 \pm 0.44$ with an observed significance of $1.4\,\sigma$. A summary of the results obtained from Asimov and data is given in Table 10.1. Figure 10.2 shows the observed data and the MC prediction before and after the fit is performed, including statistical and systematic uncertainties.
Using Equation 8.1 to calculate the measured signal events yields N^{WV}=383±294 which translates via Equation 8.2 to an acceptance and efficiency-corrected fiducial cross section of:

$$\sigma_{\text{fid}}^{\text{meas}}(\text{WW} + \text{WZ}) = (31.7 \pm 10.4\,(\text{stat.}) \pm 22.1\,(\text{syst.}))\,\text{fb}. \tag{10.2}$$



(a) (b)

Figure 10.2.: (a) Distribution of the large-R jet invariant mass before the fit is performed. The bottom plot shows the ratio of the data minus the total SM signal plus background prediction over the observed data. The green hatched region in the ratio plot corresponds to the statistical uncertainty and the gray hatched region to the systematic uncertainty.
(b) Postfit distribution of the large-R jet mass. The bottom plot shows the relative excess of the data over the background-only prediction together with the fitted signal template. The plot is obtained by shifting all templates according to the best-fit values of all nuisance parameters and normalization factors. The uncertainties shown take into account both statistical uncertainties and systematic uncertainties from the fit.

A plot of the best-fit values of the nuisance parameters can be seen in Figure 10.3 for two different fit scenarios. The first scenario (black markers) corresponds to the unconditional fit, where μ is floating freely.
The second scenario corresponds to the background-only hypothesis ($\mu = 0$) and is shown in red. Some nuisance parameters receive a shift in the unconditional fit the largest of which is exhibited by the V+jets generator uncertainty with a shift of $\sim 0.5\,\sigma$. In the background-only fit, the largest shifts observed are for the nuisance parameters associated to the top quark cross section and modeling. This can be explained by the top quark distribution, which is also peaking in the mass region, having to compensate for the left out MC signal sample in this area.

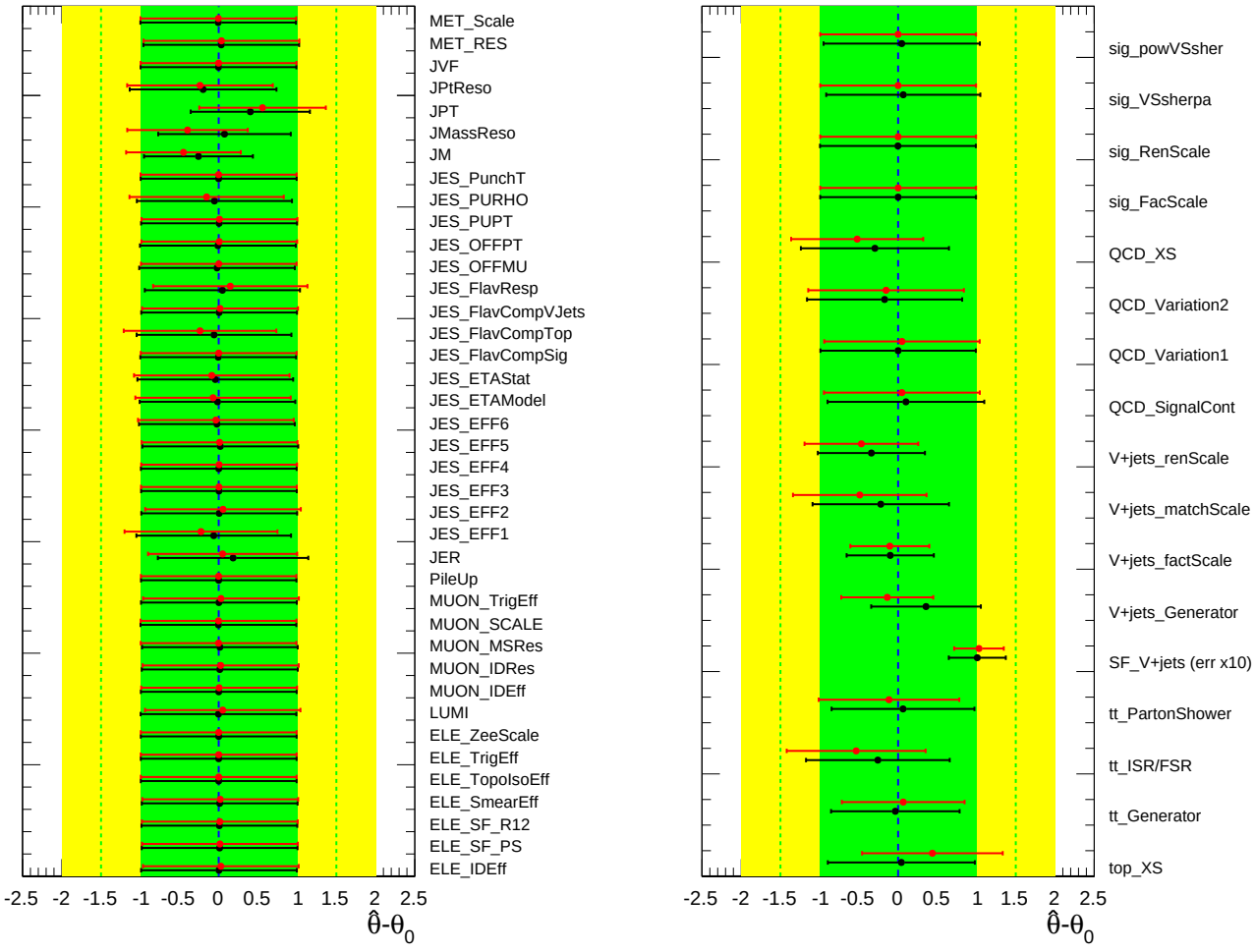


Figure 10.3.: Best-fit values of nuisance parameters after the fit to data split up into modeling uncertainties (a) and detector uncertainties (b). Black lines are showing results for the unconditional fit where the signal is free floating and red lines for the background-only fit ($\mu = 0$). In the background only fit, the largest pulls are received by the top quark background modeling related nuisance parameters. In the unconditional fit, the nuisance parameter associated to the V+jets normalization is pulled most strongly.

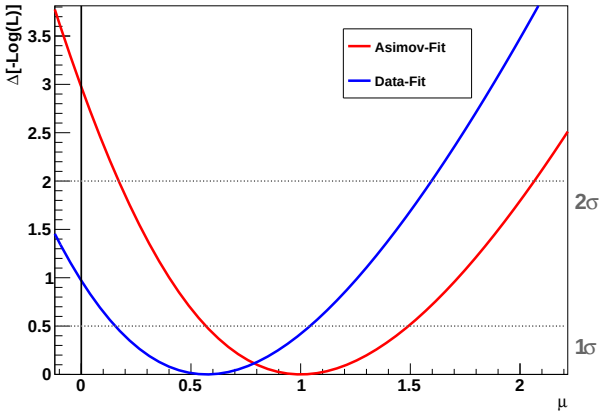


Figure 10.4.: One-dimensional scan of the negative log likelihood in dependence of μ for Asimov and real data. The expected (observed) significance of the WW+WZ process can be read off as the intersection of the Asimov (data) curve with the y-axis as sig $= \sqrt{2 \cdot y}$ (compare Equation 8.9). Significance levels are indicated as dashed grey lines.

The relative importance of the different systematic uncertainties is evaluated by a ranking procedure. For this, many fits to the dataset are performed, where in each fit iteration one nuisance parameter is fixed at a value of $\pm 1\,\sigma$. In this way it becomes possible to rank the individual contributions based on the overall impact they have on the result of the final measurement. The resulting ranking plot is shown in Figure 10.5. It shows that the highest-ranking uncertainties are the V+jets generator uncertainty and the V+jets normalization. In general the different background modeling and cross section uncertainties have a large impact on the result while the small-R jet energy scale and lepton uncertainties show, as expected, a negligible effect. Another large source of systematic uncertainty is associated to the scale and resolution of the large-R jet p_T and mass (labeled as JPT, JM, JPtReso and JMassReso).

Process	Electron Channel	Muon Channel	Summed
ZZ	1 ± 1	3 ± 2	4 ± 3
WW	293 ± 7	251 ± 7	543 ± 10
WZ	69 ± 3	60 ± 3	128 ± 5
WW+WZ	361 ± 8	310 ± 8	672 ± 11
W+jets	5689 ± 23	4821 ± 21	10510 ± 31
Z+jets	108 ± 5	137 ± 3	245 ± 6
$t\bar{t}$	647 ± 7	489 ± 6	1136 ± 9
Single Top	140 ± 9	109 ± 8	249 ± 12
QCD Multijet	311 ± 18	0	311 ± 18
Total SM prediction	7256 ± 32	5867 ± 25	13124 ± 40
Total Data	7327 ± 86	5672 ± 75	12999 ± 114
S/B	0.054	0.050	0.05
S/B (sig. window)	0.10	0.090	0.10
$S/\sqrt{B}$ (sig. window)	7.1	5.0	4.9

Table 10.2.: Expected and measured event yields for the different signal and background processes rounded to integer numbers. The signal window refers to $65\,\text{GeV} < m_\text{J} < 95\,\text{GeV}$.

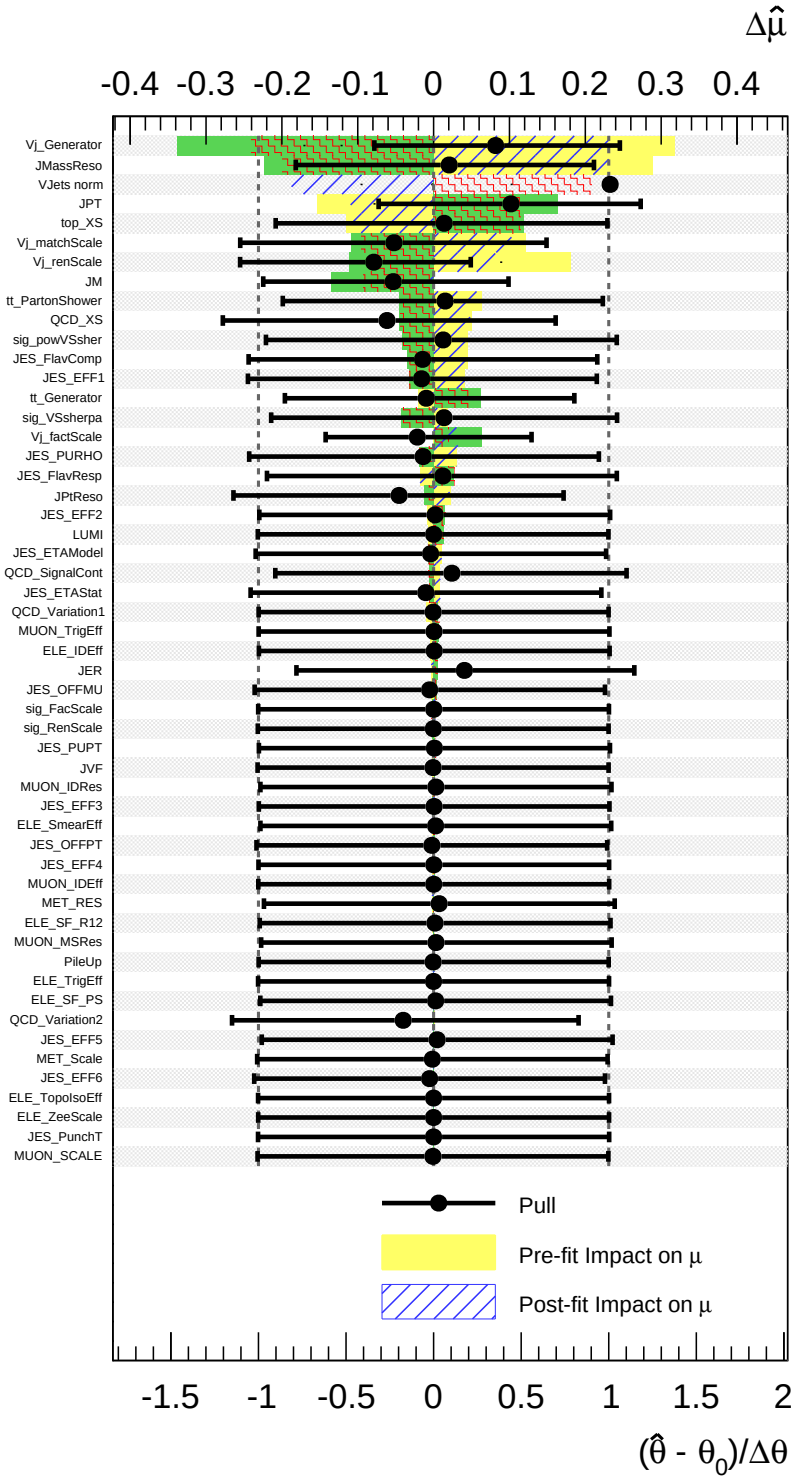


Figure 10.5.: Ranking of the different sources of systematic uncertainty obtained from the fit to data. The solid-yellow (blue-hatched) bands refer to the pre-fit (post-fit) impact on the signal strength μ when the nuisance parameter is shifted up by one standard deviation (upper x-axis). Solid-green (red-hatched) bands refer to the pre-fit (post-fit) impact on μ when the nuisance parameter is shifted down by one deviation. Fitted nuisance parameter values are shown as points with error bars (lower x-axis). The x-value of each point corresponds to the deviation of a given nuisance parameter θ from its nominal value θ_0 in units of standard deviations. The normalization factor on the V+jets template (labeled SF_V+jets) is not a nuisance parameter in the fit, so the expected value is 1.

10.2. Limits on Anomalous Triple Gauge Couplings

If not stated otherwise the results presented in this chapter are based on a collaborative effort within a working group of which the final results will be published in Ref. [159].
Limits on anomalous triple gauge coupling parameters are derived by comparing the measured p_T spectrum of the large-R jet with the p_T spectrum that would be expected if anomalous couplings were present. The derivation is performed by a binned maximum-likelihood fit of signal and background templates to the p_T^J spectrum in data. The event selection used for the limit setting is the same as is applied for the cross section measurement except for an additional requirement on the large-R jet mass to lie inside the signal window of $65 < m_J < 95\,\text{GeV}$. This selection has proven to enhance the sensitivity to anomalous couplings. Figure 10.6 shows the measured and expected large-R jet p_T in the signal region together with the expectation for an anomalous TGC parameter.

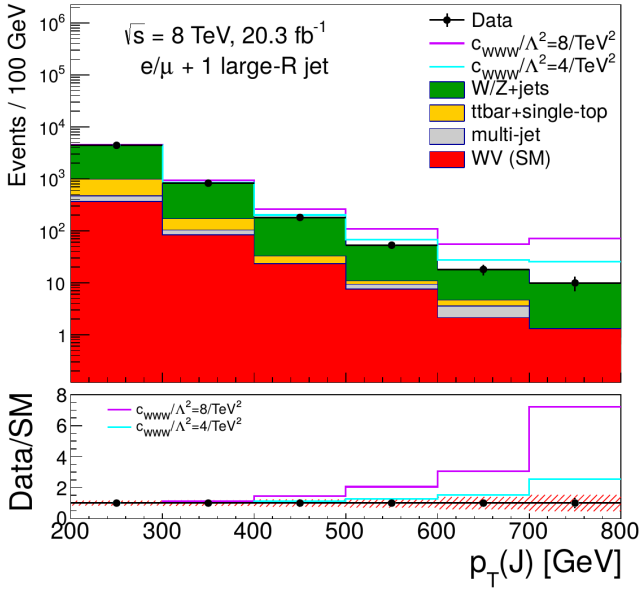


Figure 10.6.: Observed p_T^J distribution in the signal region for the sum of the muon and electron channels. The points represent the data with Poisson errors. The white-filled histogram shows the expected signal in the presence of an anomalous triple gauge coupling of: $c_{WWW}/\Lambda^2 = 4$ TeV^{-2} and 8 TeV^{-2}.

The fit procedure itself is carried out similarly to the cross section fit with one major difference. Systematic uncertainties here are also included via nuisance parameters but contrary to the cross section fit, where each nuisance parameter represented a particular source of systematic uncertainty, in the aTGC fit each nuisance parameter represents the total systematic uncertainty resulting from all sources on one particular p_T^J bin. There is one nuisance parameter for signal and one for background, such that the total number of nuisance parameters amounts to $2m$, when m is the number of bins.
Let $\vec{\rho}$ be the vector of aTGC parameters (e.g. $\rho_1 = \lambda$, $\rho_2 = \Delta\kappa_\gamma$, $\rho_3 = \Delta g_1^Z$), and $\vec{\theta}$ be a vector of nuisance parameters. The likelihood function is then defined as:

$$L(\vec{\rho},\vec{\theta}) = \prod_{i=1}^{m} \text{Poisson}(N^i_{\text{data}}, \mu^i(\vec{\rho},\vec{\theta})) \times \frac{1}{(2\pi)^m} e^{-\frac{1}{2}(\vec{\theta}\cdot C^{-1}\cdot\vec{\theta})}\,, \tag{10.3}$$

with the expected number of events in bin i denoted by μ^i:

$$\mu^i(\vec{\rho},\vec{\theta}) = N^i_{\text{sig}}(\vec{\rho})\, h(1+\theta_i) + N^i_{\text{bkg}}\, h(1+\theta_{i+m})\,, \tag{10.4}$$

and with the expected number of signal and background events N^i_{sig} and N^i_{bkg} and the log-normal uncertainties on the event yields being described by the function h:

$$\begin{aligned} h(1+\theta_i) &= (1+\sigma_i)^{\theta_i/\sigma_i} \text{ for } \theta_i >= 0 \\ &= (1-\sigma_i)^{-\theta_i/\sigma_i} \text{ for } \theta_i < 0\,, \end{aligned} \tag{10.5}$$

with $\sigma_i = \sqrt{C_{i,i}}$. The covariance matrix C describes systematic uncertainties on the signal and background in each bin as well as bin-to-bin correlations between these uncertainties including correlations between signal and background. It is computed by linearly summing the individual matrices C^s which are formed for each source of systematic uncertainty s in the following way:

$$C^s_{i,j} = \frac{(N^s_i - N^0_i)(N^s_j - N^0_j)}{N^0_i N^0_j}\,, \tag{10.6}$$

where

$$\begin{aligned} N^{s/0}_i &= N^{i,s/0}_{\text{sig}} \text{ for } i < m\,, \\ &= N^{i-m,s/0}_{\text{bkg}} \text{ for } m <= i < 2m\,, \end{aligned} \tag{10.7}$$

and with the nominal expected signal or background yield $N^{i,0}_{\text{sig/bkg}}$ in bin i and the according expected varied yield $N^{i,s}_{\text{sig/bkg}}$ for the "up" variation of the systematic uncertainty s. Since all systematic variations are symmetric, using "up"or "down" variations leads to the same result. The observed and expected 95 % confidence intervals are determined employing a frequentist method also used and explained in Ref. [160]. The method shall be outlined briefly in the following. As a test statistic, the profile likelihood ratio $\lambda(\rho)$ is used:

$$\lambda(\rho) = \frac{\mathcal{L}(N_{\text{data}}|\rho,\hat{\hat{\theta}})}{\mathcal{L}(N_{\text{data}}|\hat{\rho},\hat{\theta})}\,, \tag{10.8}$$

with the same conventions for the notation as used in Section 8.3. The observed likelihood ratio $\lambda_{obs}(\rho)$ is then evaluated for a variety of different ρ values. A large number (10000) of "toy" datasets are generated for each test-value of ρ and the likelihood ratio $\lambda_{toy}(\rho)$ is calculated for each of them. The probability of obtaining a result which is at least as unlikely as the observed one (p-value) is computed as the fraction of toy dataset that have a value $\lambda_{toy}(\rho) < \lambda_{obs}(\rho)$. Confidence intervals (95 % confidence level) are derived by scanning over values of ρ to determine the ρ intervals which have a p-value of more than 5 %. Since this method is computationally intensive and computation time is rapidly increasing with the number of bins, the binning has been optimized in order to reduce computing time without loosing much sensitivity. It was found that a number of six bins where the last bin includes overflow as is shown in Figure 10.6 was an optimal choice.

Results for the expected and observed 95 % C.L. limits for the LEP scenario without form factor and for the EFT parametrization are shown in Table 10.3. The limits on each parameter are calculated while fixing the other two parameters to zero. The Asimov (observed) dataset is used for the calculation of the expected (observed) limits.

The WW+WZ cross section was also measured recently in the non-boosted (resolved) regime [159]. The corresponding analysis uses the exact same event selection as is used in the present analysis with the major difference between boosted and resolved selection being the jet selection. In the resolved regime, instead of selecting a large-R jet with high p_T, two central ($|\eta| < 2.4$) small-R jets with $p_T > 25\,\text{GeV}$ are selected. The cross section extraction, aTGC limit setting and evaluation of systematic uncertainties are otherwise identical to the boosted analysis.
In order to quantify the potential gain in sensitivity to aTGCs by using the boosted topology, it is of high interest to compare the results of the two topologies.

	Boosted (this analysis)		**Resolved**	
Parameter	Expected Limit	Observed Limit	Expected Limit	Observed Limit
$\lambda_Z = \lambda_\gamma$	[-0.015, +0.015]	[-0.013, +0.013]	[-0.029, +0.029]	[-0.025, +0.025]
$\Delta\kappa_\gamma$	[-0.070, +0.074]	[-0.063, +0.066]	[-0.160, +0.180]	[-0.085, +0.125]
Δg_1^Z	[-0.024, +0.027]	[-0.021, +0.024]	[-0.038, +0.054]	[-0.029, +0.050]
c_{WWW}/Λ^2	[-3.5, +3.5]	[-3.2, +3.1]	[-7.0, +6.9]	[-6.0, +6.0]
c_B/Λ^2	[-22, +23]	[-19, +20]	[-49, +55]	[-39, +46]
c_W/Λ^2	[-5.7, +6.6]	[-5.2, +5.8]	[-9.3, +13.4]	[-6.9, +12.1]

Table 10.3.: The expected and observed 95 % C.L. limits on the aTGC parameters λ, $\Delta\kappa_\gamma$ and Δg_1^Z in the LEP scenario without form factor and the EFT parameters c_{WWW}/Λ^2, c_B/Λ^2 and c_W/Λ^2 for both, the boosted (this analysis) and resolved selection. Figure 10.7 shows a graphical representation of the limits on EFT parameters, the units of which are given in TeV^{-2}.

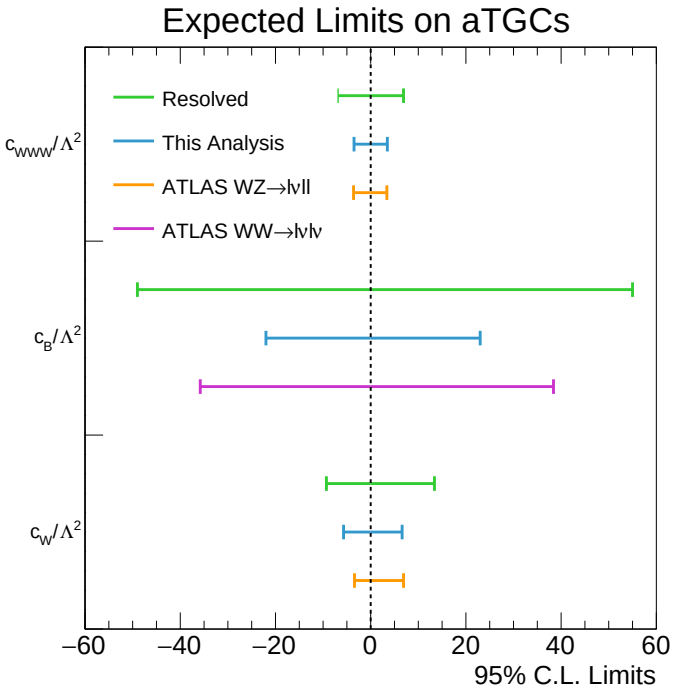


Figure 10.7.: Expected limits on EFT parameters from the presented boosted and resolved WW+WZ measurement together with the currently strictest expected limits of other publications referenced in Table 10.4.

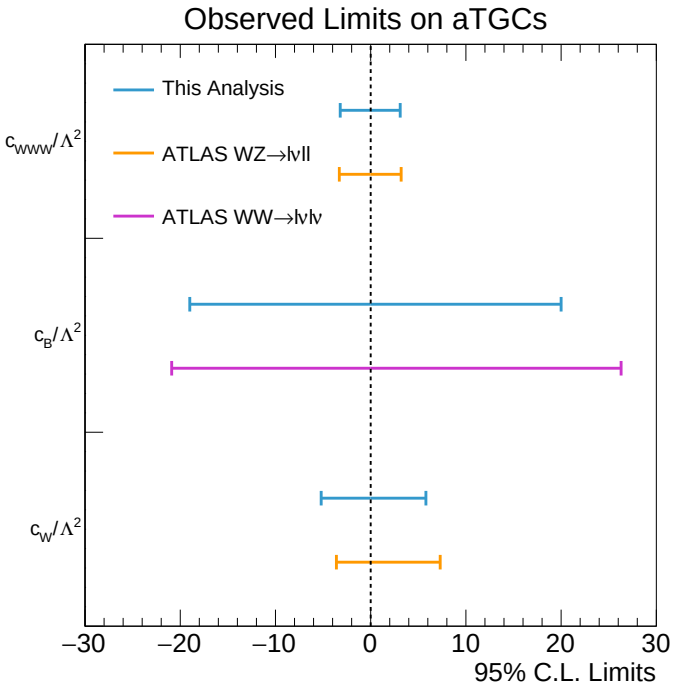


Figure 10.8.: Observed limits in the present analysis compared to currently most stringent limits. References are detailed in Table 10.4.

A summary of the expected and observed limits obtained from the resolved selection can be found in Table 10.3 as well. The difference in sensitivity is visualized in Figure 10.7 for the EFT parameters. It is evident that the sensitivity is strongly improved in the boosted topology in comparison to the resolved topology for all parameters. The largest reduction of the allowed range is found for $c_B/\Lambda^2$with a reduction of 57 %.
The currently strongest published limits on the couplings investigated in this analysis are collected in Table 10.4 and are also included in Figure 10.8. It is visible that limits presented in this analysis are competitive and in case of c_B/Λ^2even superior to currently best limits.

Parameter	Expected Limit	Observed Limit	Process	Reference
λ	[-0.031, +0.031]	[-0.019, +0.019]	WW	[8]
$\Delta\kappa_\gamma$	Not published	[-0.130, +0.095]	LEP comb.	[46]
Δg_1^Z	[-0.033, +0.037]	[-0.016, +0.027]	WW	[8]
c_{WWW}/Λ^2	[-3.6, +3.4]	[-3.3, +3.2]	WZ	[161]
c_B/Λ^2	[-35.8, +38.4]	[-20.9, +26.3]	WW	[8]
c_W/Λ^2	[-3.4, +6.9]	[-3.6, +7.3]	WZ	[161]

Table 10.4.: Most stringent limits up to date for the for the EFT parameters and for the LEP scenario aTGC parameters without any form factor applied. Units for the EFT parameters are in TeV^{-2}.

Chapter 11

Summary and Outlook

11.1. Summary

A measurement of the WW+WZ fiducial cross section in the semileptonic decay mode in the boosted regime has been presented in the exclusive zero jet channel. The presence of WW+WZ signal has been established with an observed significance of $1.4\,\sigma$. The result of $\sigma_{\text{fid}}^{\text{meas}}(\text{WW} + \text{WZ}) = (32.8 \pm 25)\,\text{fb}$ is compatible within its systematic and statistic uncertainty with the Standard Model expectation of $\sigma_{\text{fid}}^{\text{theo}}(\text{WW} + \text{WZ}) = (57.7 \pm 2.9)\,\text{fb}$.
Furthermore, the observed data has been tested for the presence of anomalous triple gauge couplings. The observed results are all in good agreement with the SM prediction such that limits have been set on the parameters describing those anomalous couplings for two parametrizations, namely the LEP parametrization and the EFT parametrization. Compared to the limits derived using the resolved topology the results of this analysis show an improvement of up to 57 %.
Limits are also competitive with the currently strongest limits extracted from fully leptonic WW and WZ measurements, despite much smaller statistics and signal to background ratio.
The reason for the comparatively large sensitivity considering the generally low statistics and large systematic uncertainties is the higher sensitivity to aTGCs at high jet transverse momentum in the boosted regime – one of the original motivations for this channel. However, systematic uncertainties associated to the large-R jet and the signal and background modeling are large which is why the extracted limits are still only comparable and not actually better than the best presently existing limits.

11.2. Prospects for LHC Run 2

Since the significance observed for boosted WW+WZ production is relatively low, it is worth contemplating the prospects of a repetition of this measurement using LHC Run 2 data. The amount of pile-up is slightly larger in this data which was taken in 2015 and 2016 at $\sqrt{s}$=13 TeV compared to the Run 1 data at $\sqrt{s}$=8 TeV used in this analysis (see Figure 11.1). This can be potentially harmful to a reiteration of the WW+WZ cross section measurement with this dataset. There are, however, several considerations that make an analysis using Run 2 data very attractive.

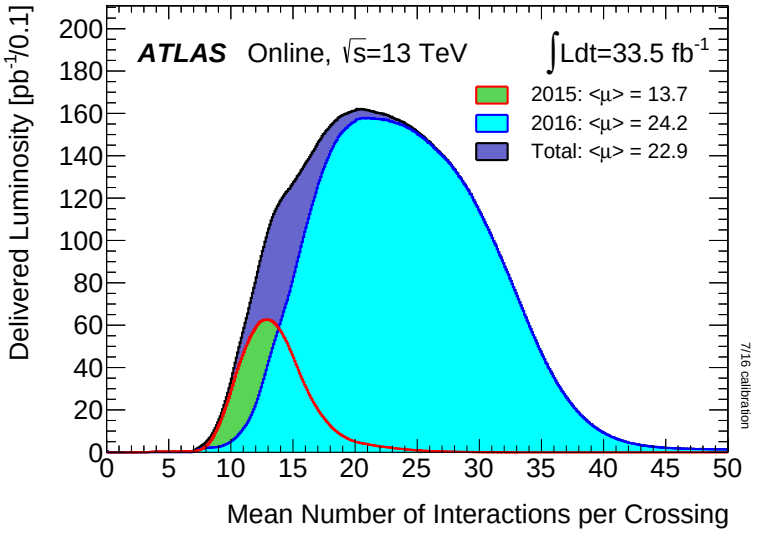


Figure 11.1.: Distribution of the average interactions per bunch crossing for the LHC Run 2 data split up in the two data taking periods of 2015 and 2016. The mean value is with 22.9 around 10.6 % higher than in the 2012 dataset (compare Figure 5.5b)

In the present analysis, the systematic error of the cross section measurement is indeed larger than the statistical error but not to a high degree. Moreover, large effort is also made in order to minimize systematic uncertainties associated to large-R jets which are amongst the dominant sources of systematic uncertainty in this analysis (compare Figure 10.5). The systematic uncertainty is thus likely to be reduced sizably in a future analysis. This implies that it is possible to gain considerably in precision by using a larger dataset. The WW and WZ cross section at a center-of-mass energy of $\sqrt{s} = 13$ TeV are known to a high accuracy at the theory level.
For WW, the quark-antiquark annihilation mechanism it has been computed at NNLO ($\mathcal{O}(\alpha_S^2)$) [11, 162], non-resonant gluon fusion production has been evaluated at NLO ($\mathcal{O}(\alpha_S^3)$) [55] and the resonant gg→H→WW production is known at $\mathcal{O}(\alpha_S^5)$ [54, 163].

	σ_{WW}^{NNLO} [pb]	σ_{WZ}^{NNLO} [pb]
$\sqrt{s} = 8$ TeV	$59.84^{+2.2\,\%}_{-1.9\,\%}$	$24.69^{+1.8\,\%}_{-1.9\,\%}$
$\sqrt{s} = 13$ TeV	$118.7^{+2.5\,\%}_{-2.2\,\%}$	$51.11^{+2.2\,\%}_{-2.0\,\%}$

Table 11.1.: Summary of the production cross section of WW [11] and WZ [56] (excluding resonant Higgs production for WW) at center-of-mass energies of 8 and 13 TeV.

These high precision predictions would be interesting to be tested experimentally. Table 11.1 shows a summary of WW and WZ cross sections at 8 and 13 TeV. In both cases, the cross-section is approximately a factor of two larger at $\sqrt{s} = 13\,\text{TeV}$ compared with that at $\sqrt{s} = 8\,\text{TeV}$.

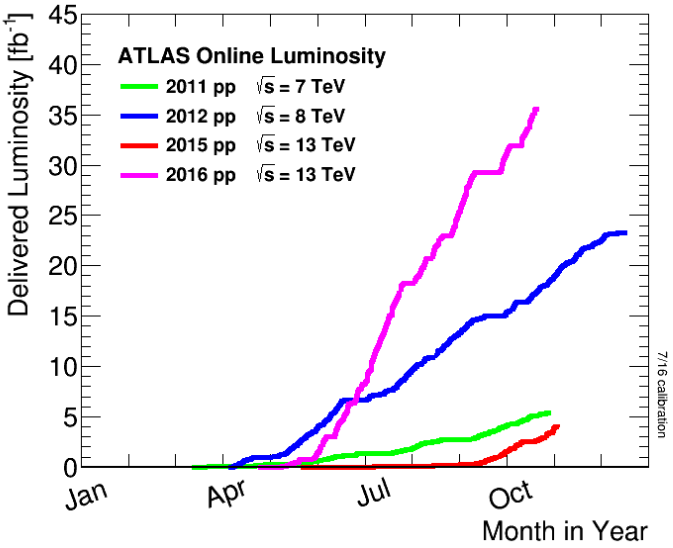


Figure 11.2.: Development of the integrated luminosity collected over one year for the last four data taking periods. The total integrated luminosity collected at a center-of-mass energy of 13 TeV is $39.8\,\text{fb}^{-1}$, at 8 TeV it is $20.3\,\text{fb}^{-1}$ [164].

Together with the larger integrated luminosity collected for the Run 2 dataset (compare Figure 11.2) which adds up to $39.8\,\text{fb}^{-1}$ and is thus also almost a factor two larger than the Run 1 dataset 8 TeV dataset of $20.3\,\text{fb}^{-1}$, this leads to an expected increase of roughly 400 % in the WW and WZ event yield. The increase in statistics becomes even larger when taking into account the fact that a larger fraction of events will be in the boosted regime at $\sqrt{s} = 13\,\text{TeV}$.
On the other hand, events that are highly boosted such that the two subjets of the W or Z boson are merged into one subjet can still be well distinguished from QCD background jets by the jet substructure methods. This is due to the decisive criterion for suitable substructure variables such as the 2-subjettiness or the D_2 variable being how well the jet can be described as being composed of two *or less* subjets. This would still well be the case for highly boosted heavy particle jets whereas it is still not true for typical spread-out QCD jets.
The gain in statistics would therefore result in a reduction of the statistical error of at least 50 %. Furthermore, substructure analysis techniques are a very active field which is constantly progressing, meaning a future analysis could profit from new results or variables. The abundance of those jet substructure variables also suggests the use of a multivariate technique such as an artificial neural network or a boosted decision tree. In this way, one could exploit the separating power of several variables at once instead of relying on only one of them.
This approach was also followed in the present analysis and showed promising results [165] but was dismissed due to the anticipated difficulties in estimating systematic uncertainties. However, for a potential Run 2 data based analysis this possibility might be worth to be investigated again. This is especially the case when considering the fact that for the presented analysis it was found that the use of one substructure variable alone is in total not beneficial for the sensitivity.
In summary a repetition of this analysis with Run 2 data appears to be a worthwhile venture that could further reduce uncertainties on the WW+WZ cross section and help to constrain aTGC parameters.

APPENDIX A

Monte Carlo and Data Samples

A.1. Data

Data Taking Period	Events
Period A	43229382
Period B	177476843
Period C	51322976
Period D	113174461
Period E	86651806
Period G	43857431
Period H	53569883
Period I	36739595
Period J	95315250
Period L	30689178
Sum	732026805

Table A.1.: Number of events in the Egamma stream per data taking period.

Data Taking Period	Events
Period A	43656420
Period B	163532962
Period C	80935368
Period D	112743915
Period E	81667914
Period G	41182058
Period H	49412952
Period I	34765048
Period J	88410196
Period L	29040147
Sum	725346980

Table A.2.: Number of events in the Muons stream per data taking period.

A.2. WW+WZ Signal

A.2.1. WW+WZ Nominal Samples

Physics Process	Generator	σ·BR·$\varepsilon_{\text{Filter}}$ [pb]	Statistics [10^6]
$q\bar{q} \rightarrow WW$ (lepton filter)	MC@NLO	23.62	4.98
W^+Z (lepton filter)	MC@NLO	4.47	1.00
W^-Z (lepton filter)	MC@NLO	2.49	0.49
$q\bar{q} \rightarrow WW$ (aTGC, lepton filter)	MC@NLO	70.46	3.96
W^+Z (aTGC, lepton filter)	MC@NLO	19.60	0.98
W^-Z (aTGC, lepton filter)	MC@NLO	9.36	0.50

Table A.4.: Nominal samples used for SM-like signal modeling and signal modeling including anomalous triple gauge couplings.

A.2.2. WW+WZ Systematics Samples

Physics Process	Generator	σ·BR·$\varepsilon_{\text{Filter}}$ [pb]	Statistics [10^6]
$q\bar{q} \rightarrow WW$ (all decays)	Powheg	54.43	10.00
$WZ \rightarrow l\nu q\bar{q}$	Powheg	4.87	10.00
$WZ \rightarrow q\bar{q}ll$	Powheg	1.59	2.39
$q\bar{q} \rightarrow WW \rightarrow e\nu q\bar{q}$	Sherpa	7.72	3.16
$q\bar{q} \rightarrow WW \rightarrow \mu\nu q\bar{q}$	Sherpa	7.74	3.16
$q\bar{q} \rightarrow WW \rightarrow \tau\nu q\bar{q}$	Sherpa	7.71	3.16
$q\bar{q} \rightarrow WW \rightarrow l\nu l\nu$	Sherpa	5.61	0.40
$WZ \rightarrow e\nu q\bar{q}$	Sherpa	2.00	0.84
$WZ \rightarrow \mu\nu q\bar{q}$	Sherpa	2.00	0.84
$WZ \rightarrow \tau\nu q\bar{q}$	Sherpa	2.01	0.84
$WZ \rightarrow q\bar{q}ee$	Sherpa	1.54	0.18
$WZ \rightarrow q\bar{q}\mu\mu$	Sherpa	1.54	0.18
$WZ \rightarrow q\bar{q}\tau\tau$	Sherpa	1.54	0.18
$WZ \rightarrow l\nu ll$	Sherpa	10.23	2.70
$WZ \rightarrow l\nu\nu\nu$	Sherpa	1.47	0.40

Table A.6.: Signal samples used for the evaluation of systematic uncertainties.

A.3. W+Jets

A.3.1. W+Jets Nominal Sample

Physics Process	Generator	σ·BR·$\varepsilon_{\text{Filter}}$ [pb]	Statistics [10^6]
W→eν(p_T >0, b-filter)	Sherpa	154.38	15.00
W→eν(p_T >0, c-filter/b-veto)	Sherpa	591.62	10.00
W→eν(p_T >0, c-veto/b-veto)	Sherpa	11324.49	49.87
W→$\mu\nu$(p_T >0, b-filter)	Sherpa	154.43	14.99
W→$\mu\nu$(p_T >0, c-filter/b-veto)	Sherpa	513.12	9.99
W→$\mu\nu$(p_T >0, c-veto/b-veto)	Sherpa	11404.48	49.78
W→$\tau\nu$(p_T >0, b-filter)	Sherpa	154.38	14.96
W→$\tau\nu$(p_T >0, c-filter/b-veto)	Sherpa	557.10	9.99
W→$\tau\nu$(p_T >0, c-veto/b-veto)	Sherpa	11360.17	49.88
W→ eν(40<p_T <70, b-filter)	Sherpa	24.80	1.10
W→ eν(40<p_T <70, c-filter/b-veto)	Sherpa	123.34	0.90
W→ eν(40<p_T <70, c-veto/b-veto)	Sherpa	570.00	16.93
W→$\mu\nu$(40<p_T <70, b-filter)	Sherpa	24.81	1.10
W→$\mu\nu$(40<p_T <70, c-filter/b-veto)	Sherpa	119.03	0.90
W→$\mu\nu$(40<p_T <70, c-veto/b-veto)	Sherpa	574.57	16.95
W→$\tau\nu$(40<p_T <70, b-filter)	Sherpa	24.82	1.10
W→$\tau\nu$(40<p_T <70, c-filter/b-veto)	Sherpa	121.47	0.89
W→$\tau\nu$(40<p_T <70, c-veto/b-veto)	Sherpa	571.87	16.95
W→ eν(70<p_T <140, b-filter)	Sherpa	12.66	2.00
W→ eν(70<p_T <140, c-filter/b-veto)	Sherpa	55.43	3.00
W→ eν(70<p_T <140, c-veto/b-veto)	Sherpa	207.49	15.00
W→$\mu\nu$(70<p_T <140, b-filter)	Sherpa	12.66	2.00
W→$\mu\nu$(70<p_T <140, c-filter/b-veto)	Sherpa	53.76	3.00
W→$\mu\nu$(70<p_T <140, c-veto/b-veto)	Sherpa	209.29	14.99
W→$\tau\nu$(70<p_T <140, b-filter)	Sherpa	12.66	2.00
W→$\tau\nu$(70<p_T <140, c-filter/b-veto)	Sherpa	54.83	3.00
W→$\tau\nu$(70<p_T <140, c-veto/b-veto)	Sherpa	208.02	15.00
W→ eν(140<p_T <280, b-filter)	Sherpa	2.16	5.00
W→ eν(140<p_T <280, c-filter/b-veto)	Sherpa	7.61	2.00
W→ eν(140<p_T <280, c-veto/b-veto)	Sherpa	24.47	2.00
W→$\mu\nu$(140<p_T <280, b-filter)	Sherpa	2.17	4.99
W→$\mu\nu$(140<p_T <280, c-filter/b-veto)	Sherpa	7.42	1.99
W→$\mu\nu$(140<p_T <280, c-veto/b-veto)	Sherpa	24.70	1.99
W→$\tau\nu$(140<p_T <280, b-filter)	Sherpa	2.17	4.00
W→$\tau\nu$(140<p_T <280, c-filter/b-veto)	Sherpa	7.54	2.00

W→$\tau\nu$(140<p_T <280, c-veto/b-veto)	Sherpa	24.56	2.00
W→ eν(280<p_T <500, b-filter)	Sherpa	0.17	0.90
W→ eν(280<p_T <500, c-filter/b-veto)	Sherpa	0.47	0.99
W→ eν(280<p_T <500, c-veto/b-veto)	Sherpa	1.38	2.49
W→$\mu\nu$(280<p_T <500, b-filter)	Sherpa	0.17	0.90
W→$\mu\nu$(280<p_T <500, c-filter/b-veto)	Sherpa	0.46	1.00
W→$\mu\nu$(280<p_T <500, c-veto/b-veto)	Sherpa	1.39	2.50
W→$\tau\nu$(280<p_T <500, b-filter)	Sherpa	0.17	0.90
W→$\tau\nu$(280<p_T <500, c-filter/b-veto)	Sherpa	0.47	0.20
W→$\tau\nu$(280<p_T <500, c-veto/b-veto)	Sherpa	1.38	0.50
W→ eν(p_T >500, b-filter)	Sherpa	0.01	0.10
W→ eν(p_T >500, c-filter/b-veto)	Sherpa	0.03	0.01
W→ eν(p_T >500, c-veto/b-veto)	Sherpa	0.07	0.01
W→$\mu\nu$(p_T >500, b-filter)	Sherpa	0.01	0.09
W→$\mu\nu$(p_T >500, c-filter/b-veto)	Sherpa	0.03	0.61
W→$\mu\nu$(p_T >500, c-veto/b-veto)	Sherpa	0.07	1.05
W→$\tau\nu$(p_T >500, b-filter)	Sherpa	0.01	0.09
W→$\tau\nu$(p_T >500, c-filter/b-veto)	Sherpa	0.03	0.01
W→$\tau\nu$(p_T >500, c-veto/b-veto)	Sherpa	0.07	0.05

Table A.7.: All sub-samples of the nominal W+jets background sample split up into p_T-slices and quark content in the perturbative matrix element calculation.

A.3.2. W+Jets Generator Systematics Sample

Physics Process	Generator	σ·BR·$\varepsilon_{\text{Filter}}$ [pb]	Statistics [10^6]
W→ eν+0p	Alpgen+Pythia	9208.23	29.46
W→ eν+1p	Alpgen+Pythia	2031.13	47.94
W→ eν+2p	Alpgen+Pythia	614.20	17.50
W→ eν+3p	Alpgen+Pythia	167.29	4.86
W→ eν+4p	Alpgen+Pythia	42.75	2.47
W→ eν+5p	Alpgen+Pythia	13.55	0.80
W→$\mu\nu$+0p	Alpgen+Pythia	9208.00	31.97
W→$\mu\nu$+1p	Alpgen+Pythia	2031.36	43.62
W→$\mu\nu$+2p	Alpgen+Pythia	614.36	17.61
W→$\mu\nu$+3p	Alpgen+Pythia	167.30	4.80
W→$\mu\nu$+4p	Alpgen+Pythia	42.77	2.55
W→$\mu\nu$+5p	Alpgen+Pythia	13.56	0.79
W→$\tau\nu$+0p	Alpgen+Pythia	9208.00	31.88
W→$\tau\nu$+1p	Alpgen+Pythia	2030.56	48.07
W→$\tau\nu$+2p	Alpgen+Pythia	614.39	17.59
W→$\tau\nu$+3p	Alpgen+Pythia	167.28	4.98
W→$\tau\nu$+4p	Alpgen+Pythia	42.81	2.55
W→$\tau\nu$+5p	Alpgen+Pythia	13.55	0.79
Wbb+0p	Alpgen+Pythia	59.18	1.60
Wbb+1p	Alpgen+Pythia	51.70	1.40
Wbb+2p	Alpgen+Pythia	27.14	0.70
Wbb+3p	Alpgen+Pythia	15.45	0.40
Wc+0p	Alpgen+Pythia	859.87	23.00
Wc+1p	Alpgen+Pythia	310.97	8.20
Wc+2p	Alpgen+Pythia	81.19	2.09
Wc+3p	Alpgen+Pythia	18.67	0.50
Wc+4p	Alpgen+Pythia	5.42	0.20
Wcc+0p	Alpgen+Pythia	169.26	4.30
Wcc+1p	Alpgen+Pythia	163.04	4.14
Wcc+2p	Alpgen+Pythia	95.43	2.39
Wcc+3p	Alpgen+Pythia	50.17	0.99

Table A.8.: All sub-samples of the W+jets background sample used for the evaluation of the generator uncertainty split up into light quark and heavy flavour samples. Light quark samples are sub-divided into samples containing different numbers of additional partons in the fixed order matrix element calculation. Heavy flavour samples are divided according to the quark flavours and numbers of additional partons present in the matrix element.

A.4. Z+Jets Sample

Physics Process	Generator	σ·BR·$\varepsilon_{\text{Filter}}$ [pb]	Statistics [10^6]
Z → ee(p_T >0, b-filter)	Sherpa	34.77	4.00
Z → ee(p_T >0, c-filter/b-veto)	Sherpa	351.97	3.00
Z → ee(p_T >0, c-veto/b-veto)	Sherpa	856.14	4.98
Z →$\mu\mu$(p_T >0, b-filter)	Sherpa	34.76	4.00
Z →$\mu\mu$(p_T >0, c-filter/b-veto)	Sherpa	352.58	3.00
Z →$\mu\mu$(p_T >0, c-veto/b-veto)	Sherpa	856.22	4.99
Z →$\tau\tau$(p_T >0, b-filter)	Sherpa	34.73	4.00
Z →$\tau\tau$(p_T >0, c-filter/b-veto)	Sherpa	352.23	3.00
Z →$\tau\tau$(p_T >0, c-veto/b-veto)	Sherpa	856.32	4.98
Z → ee(70<p_T <140, b-filter)	Sherpa	2.73	1.40
Z → ee(70<p_T <140, c-filter/b-veto)	Sherpa	11.72	1.00
Z → ee(70<p_T <140, c-veto/b-veto)	Sherpa	18.58	2.00
Z →$\mu\mu$(70<p_T <140, b-filter)	Sherpa	2.73	1.39
Z →$\mu\mu$(70<p_T <140, c-filter/b-veto)	Sherpa	11.70	1.00
Z →$\mu\mu$(70<p_T <140, c-veto/b-veto)	Sherpa	18.57	2.00
Z →$\tau\tau$(70<p_T <140, b-filter)	Sherpa	2.73	1.40
Z →$\tau\tau$(70<p_T <140, c-filter/b-veto)	Sherpa	11.73	1.00
Z →$\tau\tau$(70<p_T <140, c-veto/b-veto)	Sherpa	18.58	2.00
Z → ee(140<p_T <280, b-filter)	Sherpa	0.43	1.00
Z → ee(140<p_T <280, c-filter/b-veto)	Sherpa	1.65	0.40
Z → ee(140<p_T <280, c-veto/b-veto)	Sherpa	2.39	0.60
Z →$\mu\mu$(140<p_T <280, b-filter)	Sherpa	0.43	1.00
Z →$\mu\mu$(140<p_T <280, c-filter/b-veto)	Sherpa	1.65	0.40
Z →$\mu\mu$(140<p_T <280, c-veto/b-veto)	Sherpa	2.38	0.60
Z →$\tau\tau$(140<p_T <280, b-filter)	Sherpa	0.43	0.80
Z →$\tau\tau$(140<p_T <280, c-filter/b-veto)	Sherpa	1.65	0.40
Z →$\tau\tau$(140<p_T <280, c-veto/b-veto)	Sherpa	2.38	0.60
Z → ee(280<p_T <500, b-filter)	Sherpa	0.03	0.18
Z → ee(280<p_T <500, c-filter/b-veto)	Sherpa	0.10	0.55
Z → ee(280<p_T <500, c-veto/b-veto)	Sherpa	0.14	0.55
Z →$\mu\mu$(280<p_T <500, b-filter)	Sherpa	0.03	0.17
Z →$\mu\mu$(280<p_T <500, c-filter/b-veto)	Sherpa	0.10	0.55
Z →$\mu\mu$(280<p_T <500, c-veto/b-veto)	Sherpa	0.14	0.55
Z →$\tau\tau$(280<p_T <500, b-filter)	Sherpa	0.03	0.18
Z →$\tau\tau$(280<p_T <500, c-filter/b-veto)	Sherpa	0.10	0.55
Z →$\tau\tau$(280<p_T <500, c-veto/b-veto)	Sherpa	0.14	0.55
Z → ee(p_T >500, b-filter)	Sherpa	<0.01	0.09

Z → ee(p_T >500, c-filter/b-veto)	Sherpa	0.01	0.06
Z → ee(p_T >500, c-veto/b-veto)	Sherpa	0.01	0.15
Z →$\mu\mu$(p_T >500, b-filter)	Sherpa	<0.01	0.10
Z →$\mu\mu$(p_T >500, c-filter/b-veto)	Sherpa	0.01	0.01
Z →$\mu\mu$(p_T >500, c-veto/b-veto)	Sherpa	0.01	0.01
Z →$\tau\tau$(p_T >500, b-filter)	Sherpa	<0.01	0.09
Z →$\tau\tau$(p_T >500, c-filter/b-veto)	Sherpa	0.01	0.06
Z →$\tau\tau$(p_T >500, c-veto/b-veto)	Sherpa	0.01	0.15

Table A.9.: All sub-samples of the nominal Z+jets background sample split up into p_T-slices and quark content in the matrix element.

A.5. $t\bar{t}$ + Single Top

A.5.1. $t\bar{t}$ + Single Top Nominal Samples

Physics Process	Generator	σ·BR·$\varepsilon_{\text{Filter}}$ [pb]	Statistics [10^6]
$t\bar{t}$ (not all hadronic)	Powheg+Pythia	137.32	15.00
Single top t-channel l^+	Powheg+Pythia	18.39	4.99
Single top t-channel l^-	Powheg+Pythia	9.96	5.00
Single top s-channel (leptonic)	Powheg+Pythia	1.82	6.00
Single top W t-channel	Powheg+Pythia	22.36	1.00

Table A.10.: Nominal samples used for $t\bar{t}$ and single-top modeling.

A.5.2. $t\bar{t}$ + Single Top Systematics Samples

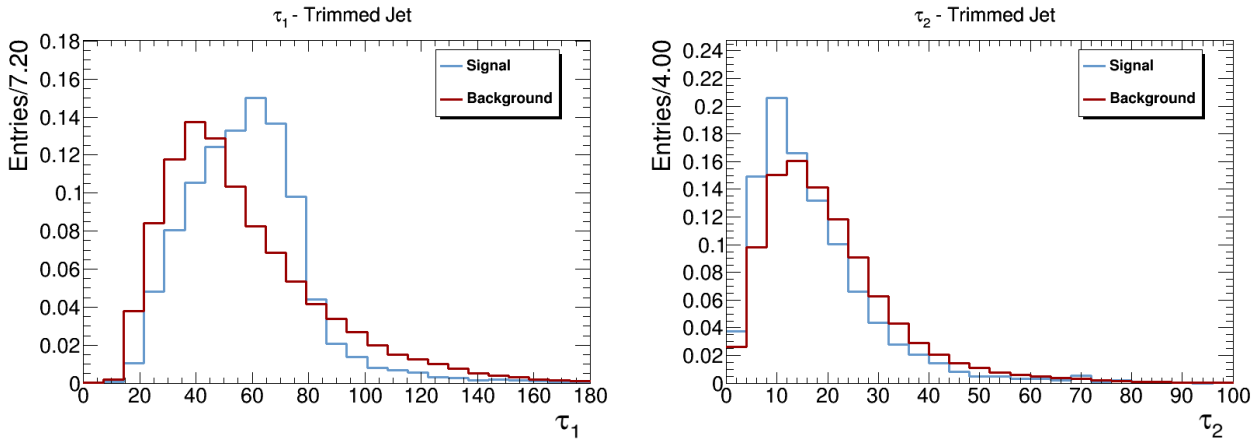
Physics Process - Systematic	Generator	σ·BR·$\varepsilon_{\text{Filter}}$ [pb]	Statistics [10^6]
$t\bar{t}$- Generator (lepton filter)	MC@NLO+Jimmy	137.32	15.00
$t\bar{t}$- Parton Shower (lepton filter)	Powheg+Jimmy	137.32	30.00
$t\bar{t}$- ISR/FSR up (not all hadronic)	AcerMC+Pythia	137.32	15.00
$t\bar{t}$- ISR/FSR down (not all hadronic)	AcerMC+Pythia	137.32	15.00

Table A.12.: $t\bar{t}$ and single-top MC samples used for the evaluation of systematic uncertainties.

APPENDIX B

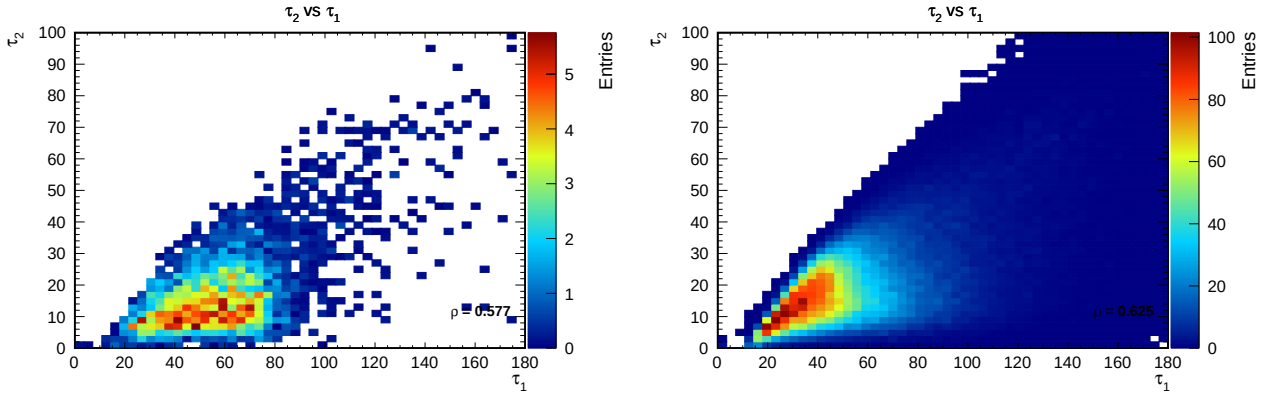
Additional Jet Substructure Plots

In the following, additional plots of jet substructure variables and their mutual correlations are shown for the entire mass range. Figure B.1 shows τ_1 and τ_2 for signal and background events after event selection in the full mass range, both distributions are normalized to unit area.
τ_1 does not perform well in terms of discrimination power which is not unexpected. This variable can be used to describe to what a degree the large-R jet is compatible with the assumption of being composed of exactly one subjet. This is not predicted to be the case for a large-R jet resulting from decaying vector boson which should in principle have two subjets. For a quark or gluon-induced QCD jet, this is not expected either since those jets tend to consist of many soft, large angle contributions.
The situation looks rather different when going to τ_2, shown in Figure B.1b. Here, small values are expected for signal and larger values for the background which is also clearly observed.



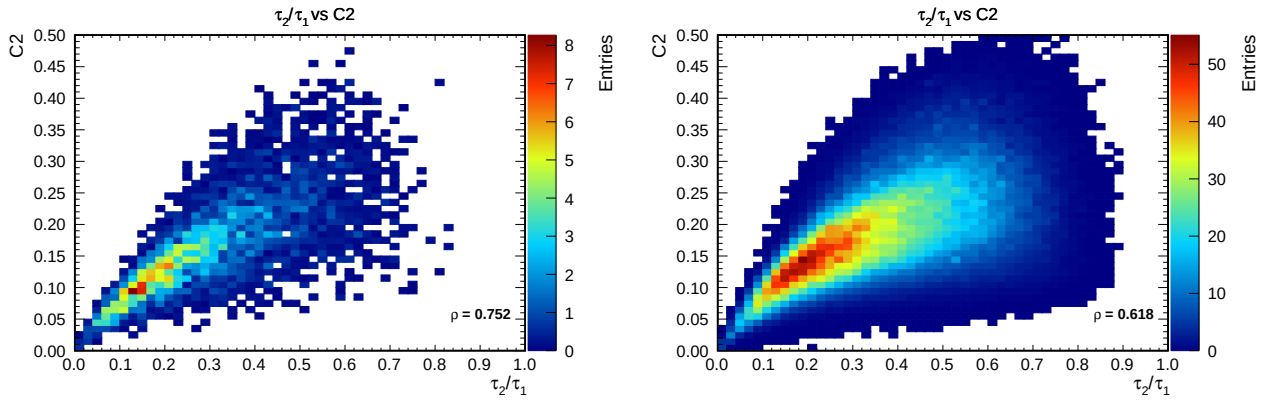
(a) (b)

Figure B.1.: Comparison of WW+WZ and background distribution shapes of τ_1 (a) and τ_2 (b).



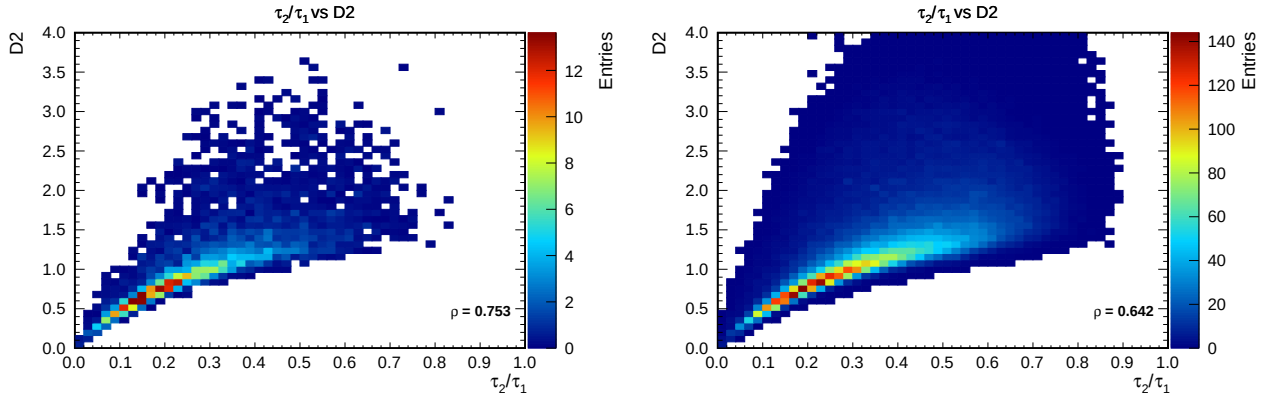
(a) (b)

Figure B.2.: Correlation between τ_2 and τ_1 for signal (a) and background (b).



(a) (b)

Figure B.3.: Correlation between τ_2/τ_1 and C_2 for signal (a) and background (b).



(a) (b)

Figure B.4.: Correlation between τ_2/τ_1 and D_2 for signal (a) and background (b).

APPENDIX C

Glossary of systematic uncertainties

Tables C.1 and C.2 give an overview of the various systematic uncertainties and their respective names in the fit uncertainties from background and signal modeling / for the detector uncertainties, respectively.

Fit naming convention	Description
V+jets_Generator	Comparison of Sherpa with Alpgen+Pythia V+jets
V+jets_factScale	V+jets factorization scale uncertainty
V+jets_renScale	V+jets renormalization scale variation
V+jets_matchScale	V+jets CKKW matching scale variation
SF_V+jets	V+jets normalization scale factor
top_XS	Top quark production cross section uncertainty
tt_Generator	Comparison of Powheg+Pythia6 with MC@NLO $t\bar{t}$
tt_PartonShower	Comparison of Powheg+Pythia6 with Powheg+Herwig $t\bar{t}$
tt_ISR/FSR	AcerMC ISR/FSR samples with more/less PS
QCD_XS	QCD multijet cross section uncertainty
QCD_Variation1	Inverted lepton selection variation 1
QCD_Variation2	Inverted lepton selection variation 2
QCD_SignalCont	20% variation of real lepton contribution
sig_VSsherpa	Comparing MC@NLO with Sherpa signal
sig_powVSsher	Comparing Sherpa with Powheg signal
sig_factScale	WV factorization scale variation
sig_renScale	WV renormalization variation

Table C.1.: List of all nuisance parameters from modeling uncertainties and their names in the fit.

Fit naming convention	Description
JES_EFF1	Effective *in situ* component 1
JES_EFF2	Effective *in situ* component 2
JES_EFF3	Effective *in situ* component 3
JES_EFF4	Effective *in situ* component 4
JES_EFF5	Effective *in situ* component 5
JES_EFF6	Effective *in situ* component 6
JES_ETAModel	Eta intercalibration: MC modeling
JES_ETAStat	Eta intercalibration: stat uncertainties
JES_OFFMU	Pile-up, mu term
JES_OFFPT	Pile-up, NPV term
JES_PUPT	Pile-up, p_T term
JES_PURHO	Pile-up, rho topology
JES_FlavComp	Flavor composition of jets
JES_FlavResp	Flavor response
JES_bjet	b-jet uncertainty
JES_PunchT	Punch through jets
JES_Closeby	Closeby jets
JER	Jet Energy Resolution uncertainty
JVF	Jet Vertex Fraction
JPT	Large-R jet p_T uncertainty
JM	Large-R jet mass uncertainty
JPTReso	Large-R jet p_T resolution
JMReso	Large-R jet mass resolution
MET_Scale	$E_{\mathrm{T}}^{\mathrm{Miss}}$ scale - soft terms
MET_RES	$E_{\mathrm{T}}^{\mathrm{Miss}}$ resolution - soft terms
PileUp	Pile-up uncertainty
MUON_TrigEff	Uncertainty on muon trigger efficiency scale factor
MUON_IDEff	Uncertainty on muon identification efficiency scale factor
MUON_SCALE	Muon energy scale
MUON_MSRes	Muon energy resolution (muon spectrometer)
MUON_IDRes	Muon energy resolution (inner detector)
LUMI	Uncertainty on luminosity
ELE_TrigEff	Uncertainty on electron trigger efficiency scale factor
ELE_IDEff	Uncertainty on electron identification efficiency scale factor
ELE_TopoIsoEff	Uncertainty on electron isolation efficiency scale factor
ELE_SmearEff	Energy calibration difference in data/MC
ELE_SF_Zee	Electron energy scale component 1
ELE_SF_R12	Electron energy scale component 2
ELE_SF_PS	Electron energy scale component 3

Table C.2.: List of all nuisance parameters from detector uncertainties and their names in the fit.

Bibliography

[1] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[2] The CMS Collaboration. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (2008), S08004. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[3] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 1 –29. DOI: http://dx.doi.org/10.1016/j.physletb.2012.08.020. URL: http://www.sciencedirect.com/science/article/pii/S037026931200857X.

[4] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[5] *ATLAS Higgs Physics Public Results*. [Online; accessed 27.10.2016]. 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HiggsPublicResults.

[6] A. K. Olive and others (Particle Data Group). "Review of Particle Physics". In: *Chin. Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.

[7] Atlas Collaboration. "Measurement of W^+W^- production in pp collisions at $\sqrt{s}$=7 TeV with the ATLAS detector and limits on anomalous WWZ and WWγ couplings". In: *Phys. Rev.* D87.11 (2013). [Erratum: Phys. Rev.D88,no.7,079906(2013)], p. 112001. DOI: 10.1103/PhysRevD.87.112001,10.1103/PhysRevD.88.079906. arXiv: 1210.2979 [hep-ex].

[8] Atlas Collaboration. "Measurement of total and differential W^+W^- production cross sections in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector and limits on anomalous triple-gauge-boson couplings". In: *JHEP* 09 (2016), p. 029. DOI: 10.1007/JHEP09(2016)029. arXiv: 1603.01702 [hep-ex].

[9] CMS Collaboration. "Measurement of the W^+W^- Cross section in pp Collisions at $\sqrt{s} = 7$ TeV and Limits on Anomalous $WW\gamma$ and WWZ couplings". In: *Eur. Phys. J.* C73.10 (2013), p. 2610. DOI: 10.1140/epjc/s10052-013-2610-8. arXiv: 1306.1126 [hep-ex].

[10] CMS Collaboration. "Measurement of W+W- and ZZ production cross sections in pp collisions at sqrt(s) = 8 TeV". In: *Phys. Lett.* B721 (2013), pp. 190–211. DOI: 10.1016/j.physletb.2013.03.027. arXiv: 1301.4698 [hep-ex].

[11] T. Gehrmann et al.
"W^+W^- Production at Hadron Colliders in Next to Next to Leading Order QCD".
In: *Phys. Rev. Lett.* 113.21 (2014), p. 212001. DOI: 10.1103/PhysRevLett.113.212001.
arXiv: 1408.5243 [hep-ph].

[12] Prerit Jaiswal and Takemichi Okui.
"Explanation of the WW excess at the LHC by jet-veto resummation".
In: *Phys. Rev.* D90.7 (2014), p. 073009. DOI: 10.1103/PhysRevD.90.073009.
arXiv: 1407.4537 [hep-ph].

[13] Massimiliano Grazzini et al.
"Transverse-momentum resummation for vector-boson pair production at NNLL+NNLO".
In: *JHEP* 08 (2015), p. 154. DOI: 10.1007/JHEP08(2015)154.
arXiv: 1507.02565 [hep-ph].

[14] Patrick Meade, Harikrishnan Ramani and Mao Zeng.
"Transverse momentum resummation effects in W^+W^- measurements".
In: *Phys. Rev.* D90.11 (2014), p. 114006. DOI: 10.1103/PhysRevD.90.114006.
arXiv: 1407.4481 [hep-ph].

[15] Michal Szleper.
"The Higgs boson and the physics of WW scattering before and after Higgs discovery".
In: *ArXiv Physics e-prints* (2014). arXiv: 1412.8367 [hep-ph].

[16] David J Griffiths. *Introduction to elementary particles; 2nd rev. version.*
Physics textbook. New York, NY: Wiley, 2008.
URL: https://cds.cern.ch/record/111880.

[17] J. Bovy.
"The self-energy of the electron: a quintessential problem in the development of QED".
In: *ArXiv Physics e-prints* (Aug. 2006). arXiv: physics/0608108.

[18] G. 't Hooft and M. Veltman. "Regularization and renormalization of gauge fields".
In: *Nuclear Physics B* 44.1 (1972), pp. 189 –213.
DOI: DOI:10.1016/0550-3213(72)90279-9.
URL: http://www.sciencedirect.com/science/article/B6TVC-4719KYC-173/2/d6e222ccf600624017080c27f2774d62.

[19] F. Halzen and A. D. Martin.
Quarks and Leptons: An Introductory Course in Modern Particle Physics.
John Wiley & Sons, 1984.

[20] Peter Schmüser. *Feynman-Graphen und Eichtheorien für Experimentalphysiker.*
Berlin Heidelberg: Springer, 1995.

[21] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos".
In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562.
arXiv: hep-ex/9807003 [hep-ex].

[22] Otto Eberhardt et al. "Impact of a Higgs Boson at a Mass of 126 GeV on the Standard Model with Three and Four Fermion Generations".
In: *Phys. Rev. Lett.* 109 (24 2012), p. 241802. DOI: 10.1103/PhysRevLett.109.241802.
URL: http://link.aps.org/doi/10.1103/PhysRevLett.109.241802.

[23] J. A. Aguilar-Saavedra et al.
"Handbook of vectorlike quarks: Mixing and single production".
In: *Phys. Rev.* D88.9 (2013), p. 094010. DOI: 10.1103/PhysRevD.88.094010.
arXiv: 1306.0572 [hep-ph].

[24] S. L. Glashow. "Partial Symmetries of Weak Interactions".
In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[25] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[26] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Elementary particle theory. Relativistic groups and analyticity.* Proceedings of the Eighth Nobel Symposium. (Aspenäsgarden, Lerum, 19th–25th May 1968). Ed. by Nils Svartholm. Stockholm: Almquist & Wiksell, 1968, pp. 367–377.

[27] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay".
In: *Phys. Rev.* 105.4 (Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413.

[28] M. Goldhaber, L. Grodzins and A. W. Sunyar. "Helicity of Neutrinos".
In: *Phys. Rev.* 109.3 (Feb. 1958), pp. 1015–1017. DOI: 10.1103/PhysRev.109.1015.

[29] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons".
In: *Phys. Rev. Lett.* 13.16 (Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[30] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons".
In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.
URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[31] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays".
In: *Phys. Rev. Lett.* 10.12 (June 1963), pp. 531–533.
DOI: 10.1103/PhysRevLett.10.531.

[32] Makoto Kobayashi and Toshihide Maskawa.
"CP-Violation in the Renormalizable Theory of Weak Interaction".
In: *Progress of Theoretical Physics* 49.2 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.
URL: http://ptp.ipap.jp/link?PTP/49/652/.

[33] M. N. Chernodub. "Background magnetic field stabilizes QCD string against breaking".
In: *ArXiv Physics e-prints* (2010). arXiv: 1001.0570 [hep-ph].

[34] Siegfried Bethke, Gunther Dissertori and Gavin P. Salam.
"World Summary of α_s (2015)". In: *Proceedings, High-Precision α_s Measurements from LHC to FCC-ee: Geneva, Switzerland, October 2-13, 2015.* 2015, pp. 7–10.
URL: http://inspirehep.net/record/1415403/files/C15-10-12.1_p7.pdf.

[35] V.N. Gribov and L.N. Lipatov.
"Deep inelastic electron scattering in perturbation theory".
In: *Physics Letters B* 37.1 (1971), pp. 78 –80. ISSN: 0370-2693.
DOI: http://dx.doi.org/10.1016/0370-2693(71)90576-4.
URL: http://www.sciencedirect.com/science/article/pii/0370269371905764.

[36] Guido Altarelli and G. Parisi. "Asymptotic Freedom in Parton Language".
In: *Nucl. Phys.* B126 (1977), pp. 298–318. DOI: 10.1016/0550-3213(77)90384-4.

[37] Yuri L. Dokshitzer. "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics."
In: *Sov. Phys. JETP* 46 (1977). [Zh. Eksp. Teor. Fiz.73,1216(1977)], pp. 641–653.

[38] A. D. Martin et al. "Parton distributions for the LHC".
In: *Eur. Phys. J.* C63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5.
arXiv: 0901.0002 [hep-ph].

[39] J. Pumplin et al.
"New generation of parton distributions with uncertainties from global QCD analysis".
In: *JHEP* 07 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012.
arXiv: hep-ph/0201195 [hep-ph].

[40] Hung-Liang Lai et al. “New parton distributions for collider physics”. In: *Phys. Rev. D* 82 (7 2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. URL: http://link.aps.org/doi/10.1103/PhysRevD.82.074024.

[41] Hung-Liang Lai et al. “New parton distributions for collider physics”. In: *Phys. Rev.* D82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

[42] John C. Collins, Davison E. Soper and George F. Sterman. “Factorization of Hard Processes in QCD”. In: *Adv. Ser. Direct. High Energy Phys.* 5 (1989), pp. 1–91. DOI: 10.1142/9789814503266_0001. arXiv: hep-ph/0409313 [hep-ph].

[43] John M. Campbell, J. W. Huston and W. J. Stirling. “Hard Interactions of Quarks and Gluons: A Primer for LHC Physics”. In: *Rept. Prog. Phys.* 70 (2007), p. 89. DOI: 10.1088/0034-4885/70/1/R02. arXiv: hep-ph/0611148 [hep-ph].

[44] Fabio Maltoni et al. “Choosing the Factorization Scale in Perturbative QCD”. In: *ArXiv Physics e-prints* (2007). arXiv: hep-ph/0703156 [HEP-PH].

[45] Wikimedia Commons. *A graphic showing the relationship between angle and pseudorapidity.* 2007. URL: https://en.wikipedia.org/wiki/File:Pseudorapidity2.png.

[46] S. Schael et al. “Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP”. In: *Phys. Rept.* 532 (2013), pp. 119–244. DOI: 10.1016/j.physrep.2013.07.004. arXiv: 1302.3415 [hep-ex].

[47] K. Hagiwara et al. “Low energy effects of new interactions in the electroweak boson sector”. In: *Phys. Rev. D* 48 (5 1993), pp. 2182–2203. DOI: 10.1103/PhysRevD.48.2182. URL: http://link.aps.org/doi/10.1103/PhysRevD.48.2182.

[48] J. B. Hansen. “Triple Gauge boson couplings in W pair production via e+ e- annihilation”. PhD thesis. Bohr Inst., 1996. URL: http://preprints.cern.ch/cgi-bin/setlink?base=preprint&categ=cern&id=thesis-99-029.

[49] G. Gounaris et al. “Triple gauge boson couplings”. In: *AGS / RHIC Users Annual Meeting Upton, New York, June 15-16, 1995*. [,525(1996)]. 1996, pp. 525–576. arXiv: hep-ph/9601233 [hep-ph]. URL: http://alice.cern.ch/format/showfull?sysnb=0215385.

[50] H. Aihara et al. “Anomalous gauge boson interactions”. In: *Electroweak Symmetry Breaking and New Physics at the TeV Scale* (1995). DOI: 10.1142/9789812830265_0009. arXiv: hep-ph/9503425 [hep-ph].

[51] Celine Degrande et al. “Effective Field Theory: A Modern Approach to Anomalous Couplings”. In: *Annals Phys.* 335 (2013), pp. 21–32. DOI: 10.1016/j.aop.2013.04.016. arXiv: 1205.4231 [hep-ph].

[52] Matthias U. Mozer. *Electroweak Physics at the LHC*. Springer, 2016.

[53] Stefano Frixione and Bryan R. Webber. “Matching NLO QCD computations and parton shower simulations”. In: *JHEP* 06 (2002), p. 029. DOI: 10.1088/1126-6708/2002/06/029. arXiv: hep-ph/0204244 [hep-ph].

[54] J R Andersen et al. "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties". In: *ArXiv Physics e-prints* (2013). Ed. by S Heinemeyer et al. DOI: 10.5170/CERN-2013-004. arXiv: 1307.1347 [hep-ph].

[55] Fabrizio Caola et al. "QCD corrections to W^+W^- production through gluon fusion". In: *Phys. Lett.* B754 (2016), pp. 275–280. DOI: 10.1016/j.physletb.2016.01.046. arXiv: 1511.08617 [hep-ph].

[56] Massimiliano Grazzini et al. "$W^{\pm}Z$ production at hadron colliders in NNLO QCD". In: *Phys. Lett.* B761 (2016), pp. 179–183. DOI: 10.1016/j.physletb.2016.08.017. arXiv: 1604.08576 [hep-ph].

[57] *"ATLAS Experiment - Luminosity Public Results"*. [Online; accessed 10.06.2016]. 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#2012_pp_Collisions.

[58] *TE-EPC-LPC in LHC*. [Online; accessed 2016-01-19]. 2016. URL: http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/pagesources/Cern-Accelerator-Complex.jpg.

[59] Michael Hostettler and Giulia Papotti. "Observations from LHC proton-proton physics operation". In: *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*. 2014, pp. 199–202. DOI: 10.5170/CERN-2014-004.199. arXiv: 1409.5216 [physics.acc-ph]. URL: https://inspirehep.net/record/1317918/files/arXiv:1409.5216.pdf.

[60] ATLAS Collaboration. "Studies of the performance of the ATLAS detector using cosmic-ray muons". In: *Eur.Phys.J.* C71 (2011), p. 1593. DOI: 10.1140/epjc/s10052-011-1593-6. arXiv: 1011.6665 [physics.ins-det].

[61] Joao Pequenao. *"Computer generated image of the ATLAS inner detector"*. [Online; accessed 07.02.2016]. 2008. URL: http://cds.cern.ch/record/1095926.

[62] *"A computer generated image of the full calorimeter"*. [Online; accessed 19.01.2016]. 2008. URL: http://atlas.ch/photos/calorimeters-combined-barrel.html.

[63] *The Scintillator Tile Calorimeter*. [Online; accessed 19.01.2016]. 2016. URL: http://hedberg.web.cern.ch/hedberg/home/atlas/calorimeters_9.pdf.

[64] S.A. Majewski. "Electronic readout of the ATLAS Liquid Argon calorimeter: Calibration and performance". In: *Real Time Conference (RT), 2010 17th IEEE-NPSS*. May 2010, pp. 1–5. DOI: 10.1109/RTC.2010.5750339.

[65] *The Forward Calorimeter*. [Online; accessed 2016-02-15]. 2011. URL: http://lunvis.web.cern.ch/lunvis/taskforce/main.html.

[66] *Overview of the ATLAS muon spectrometer components, labeled*. [Online; accessed 19.01.2016]. 2008. URL: http://www.atlas.ch/photos/muons-combined.html.

[67] ATLAS Collaboration. *ATLAS magnet system: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/338080.

[68] ATLAS Collaboration. "ATLAS detector and physics performance: Technical Design Report, 1". In: Technical Design Report ATLAS (1999).

[69] *WLCG, Worldwide LHC Computing Grid*. [Online; accessed 2016-01-17]. 2016. URL: http://wlcg-public.web.cern.ch/tier-centres.

[70] John M. Campbell and R. Keith Ellis.
"An Update on vector boson pair production at hadron colliders".
In: *Phys. Rev.* D60 (1999), p. 113006. DOI: 10.1103/PhysRevD.60.113006.
arXiv: hep-ph/9905386 [hep-ph].

[71] *Calibration of ATLAS b-tagging algorithms in dense jet environments.*
Tech. rep. ATLAS-CONF-2016-001. Geneva: CERN, 2016.
URL: http://cds.cern.ch/record/2127958.

[72] Michal Czakon and Alexander Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders".
In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021.
arXiv: 1112.5675 [hep-ph].

[73] Nikolaos Kidonakis. "NNLL resummation for s-channel single top quark production".
In: *Phys. Rev.* D81 (2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028.
arXiv: 1001.5034 [hep-ph].

[74] Nikolaos Kidonakis. "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-". In: *Phys. Rev.* D82 (2010), p. 054018.
DOI: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451 [hep-ph].

[75] Nikolaos Kidonakis. "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production". In: *Phys. Rev.* D83 (2011), p. 091503.
DOI: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792 [hep-ph].

[76] ATLAS Collaboration. "Measurements of multijet production cross sections in proton-proton collisions at 7 TeV center-of-mass energy with the ATLAS Detector".
In: ATLAS-CONF-2010-084 (Oct. 2010). URL: http://cds.cern.ch/record/1298854.

[77] Gosta Gustafson and Ulf Pettersson. "Dipole Formulation of QCD Cascades".
In: *Nucl. Phys.* B306 (1988), pp. 746–758. DOI: 10.1016/0550-3213(88)90441-5.

[78] Torbjörn Sjöstrand, Stephen Mrenna and Peter Z. Skands.
"PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026.
DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[79] Torbjörn Sjöstrand, Stephen Mrenna and Peter Z. Skands.
"A Brief Introduction to PYTHIA 8.1".
In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867.
DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[80] Andy Buckley et al. "General-purpose event generators for LHC physics".
In: *Phys. Rept.* 504 (2011), pp. 145–233. DOI: 10.1016/j.physrep.2011.03.005.
arXiv: 1101.2599 [hep-ph].

[81] S. Catani et al. "QCD matrix elements + parton showers". In: *JHEP* 11 (2001), p. 063.
DOI: 10.1088/1126-6708/2001/11/063. arXiv: hep-ph/0109231 [hep-ph].

[82] Bo Andersson et al. "Parton Fragmentation and String Dynamics".
In: *Phys. Rept.* 97 (1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[83] G. Marchesini and B.R. Webber.
"Simulation of QCD jets including soft gluon interference".
In: *Nuclear Physics B* 238.1 (1984), pp. 1 –29. ISSN: 0550-3213.
DOI: http://dx.doi.org/10.1016/0550-3213(84)90463-2.
URL: http://www.sciencedirect.com/science/article/pii/0550321384904632.

[84] G. Corcella et al. "HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)". In: *JHEP* 01 (2001), p. 010.
DOI: 10.1088/1126-6708/2001/01/010. arXiv: hep-ph/0011363 [hep-ph].

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

[86] ATLAS Collaboration. "Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C74.8 (2014), p. 3023. DOI: 10.1140/epjc/s10052-014-3023-z. arXiv: 1405.6583 [hep-ex].

[87] George Sterman and Steven Weinberg. "Jets from Quantum Chromodynamics". In: *Phys. Rev. Lett.* 39 (23 1977), pp. 1436–1439. DOI: 10.1103/PhysRevLett.39.1436. URL: http://link.aps.org/doi/10.1103/PhysRevLett.39.1436.

[88] Jason Gallicchio and Matthew D. Schwartz. "Quark and Gluon Jet Substructure". In: *JHEP* 04 (2013), p. 090. DOI: 10.1007/JHEP04(2013)090. arXiv: 1211.7038 [hep-ph].

[89] A. R. Baden. "Jets and kinematics in hadronic collisions". In: *Int. J. Mod. Phys.* A13 (1998), pp. 1817–1845. DOI: 10.1142/S0217751X98000809.

[90] Gavin P. Salam and Gregory Soyez. "A Practical Seedless Infrared-Safe Cone jet algorithm". In: *JHEP* 05 (2007), p. 086. DOI: 10.1088/1126-6708/2007/05/086. arXiv: 0704.0292 [hep-ph].

[91] Matteo Cacciari and Gavin P. Salam. "Dispelling the N^3 myth for the k_t jet-finder". In: *Phys. Lett.* B641 (2006), pp. 57–61. DOI: 10.1016/j.physletb.2006.08.037. arXiv: hep-ph/0512210 [hep-ph].

[92] Matteo Cacciari, Gavin P. Salam and Gregory Soyez. "FastJet User Manual". In: *Eur. Phys. J.* C72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[93] Stephen D. Ellis and Davison E. Soper. "Successive combination jet algorithm for hadron collisions". In: *Phys. Rev.* D48 (1993), pp. 3160–3166. DOI: 10.1103/PhysRevD.48.3160. arXiv: hep-ph/9305266 [hep-ph].

[94] Yuri L. Dokshitzer et al. "Better jet clustering algorithms". In: *JHEP* 08 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323 [hep-ph].

[95] Matteo Cacciari, Gavin P. Salam and Gregory Soyez. "The Anti-k(t) jet clustering algorithm". In: *JHEP* 0804 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[96] Paola Giovannini. "Studies on the top quark mass measurement in the all-hadronic top-antitop decay channel with ATLAS". Presented 14 Jul 2011. PhD thesis. Munich: Tech. U., 2011. URL: https://cds.cern.ch/record/1376462.

[97] C. Grupen and B. Shwartz. *Particle Detectors*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2008. ISBN: 9781139469531. URL: https://books.google.de/books?id=XCPlJTu3GQkC.

[98] Sven Menke. private communication. 7th Sept. 2016.

[99] T. Barillari et al. *Local Hadronic Calibration*. Tech. rep. ATL-LARG-PUB-2009-001-2. Geneva: CERN, June 2008. URL: https://cds.cern.ch/record/1112035.

[100] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *ArXiv Physics e-prints* (2016). arXiv: 1603.02934 [hep-ex].

[101] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance.* Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, Apr. 2008. URL: https://cds.cern.ch/record/1099735.

[102] *Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV.* Tech. rep. ATLAS-CONF-2010-053. Geneva: CERN, 2010. URL: https://cds.cern.ch/record/1281310.

[103] *"Beam".* [Online; accessed 10.06.2016]. 2016. URL: http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/beam.htm.

[104] *"ATLAS Stand-Alone Event Displays".* [Online; accessed 10.06.2016]. 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayStandAlone#2012_Z_mu_mu_event_with_high_pil.

[105] ATLAS Collaboration. "Monitoring and data quality assessment of the ATLAS liquid argon calorimeter". In: *JINST* 9 (2014), P07024. DOI: 10.1088/1748-0221/9/07/P07024. arXiv: 1405.3768 [hep-ex].

[106] *Luminosity Public Results.* [Online; accessed 17.01.2016]. 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults\#2012_pp_Collisions.

[107] *Pile-up subtraction and suppression for jets in ATLAS.* Tech. rep. ATLAS-CONF-2013-083. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1570994.

[108] Atlas Collaboration. "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV". In: *Eur. Phys. J.* C73.3 (2013), p. 2304. DOI: 10.1140/epjc/s10052-013-2304-2. arXiv: 1112.6426 [hep-ex].

[109] *Jet Etmiss Approved 2013 JES Uncertainty.* [Online; accessed 2016-09-07]. 2013. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013JESUncertainty.

[110] *Performance of large-R jets and jet substructure reconstruction with the ATLAS detector.* Tech. rep. ATLAS-CONF-2012-065. Geneva: CERN, 2012. URL: https://cds.cern.ch/record/1459530.

[111] *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS.* Tech. rep. ATLAS-CONF-2015-037. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2044941.

[112] Jesse Thaler and Ken Van Tilburg. "Identifying Boosted Objects with N-subjettiness". In: *JHEP* 03 (2011), p. 015. DOI: 10.1007/JHEP03(2011)015. arXiv: 1011.2268 [hep-ph].

[113] Jonathan M. Butterworth et al. "Jet Substructure as a New Higgs-Search Channel at the Large Hadron Collider". In: *Phys. Rev. Lett.* 100 (24 June 2008), p. 242001. DOI: 10.1103/PhysRevLett.100.242001. URL: http://link.aps.org/doi/10.1103/PhysRevLett.100.242001.

[114] ATLAS Collaboration. "Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector". In: *JHEP* 09 (2013), p. 076. DOI: 10.1007/JHEP09(2013)076. arXiv: 1306.4945 [hep-ex].

[115] David Krohn, Jesse Thaler and Lian-Tao Wang. "Jet Trimming".
In: *JHEP* 1002 (2010), p. 084. DOI: 10.1007/JHEP02(2010)084.
arXiv: 0912.1342 [hep-ph].

[116] Stephen D. Ellis, Christopher K. Vermilion and Jonathan R. Walsh. "Recombination algorithms and jet substructure: Pruning as a tool for heavy particle searches".
In: *Phys. Rev. D* 81 (9 May 2010), p. 094023. DOI: 10.1103/PhysRevD.81.094023.
URL: http://link.aps.org/doi/10.1103/PhysRevD.81.094023.

[117] ATLAS Collaboration. "Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV".
In: *Eur. Phys. J.* C76.3 (2016), p. 154. DOI: 10.1140/epjc/s10052-016-3978-z.
arXiv: 1510.05821 [hep-ex].

[118] Andrew J. Larkoski, Duff Neill and Jesse Thaler. "Jet Shapes with the Broadening Axis".
In: *JHEP* 04 (2014), p. 017. DOI: 10.1007/JHEP04(2014)017.
arXiv: 1401.2158 [hep-ph].

[119] Andrea Banfi, Gavin P. Salam and Giulia Zanderighi.
"Principles of general final-state resummation and automated implementation".
In: *JHEP* 03 (2005), p. 073. DOI: 10.1088/1126-6708/2005/03/073.
arXiv: hep-ph/0407286 [hep-ph].

[120] Andrew J. Larkoski, Gavin P. Salam and Jesse Thaler.
"Energy Correlation Functions for Jet Substructure". In: *JHEP* 06 (2013), p. 108.
DOI: 10.1007/JHEP06(2013)108. arXiv: 1305.0007 [hep-ph].

[121] Andrew J. Larkoski, Ian Moult and Duff Neill.
"Power Counting to Better Jet Observables". In: *JHEP* 12 (2014), p. 009.
DOI: 10.1007/JHEP12(2014)009. arXiv: 1409.6298 [hep-ph].

[122] John A. Nevin.
"SIGNAL DETECTION THEORY AND OPERANT BEHAVIOR: A Review of David M. Green and John A. Swets' Signal Detection Theory and Psychophysics.1".
In: *Journal of the Experimental Analysis of Behavior* 12.3 (1969), pp. 475–480.
ISSN: 1938-3711. DOI: 10.1901/jeab.1969.12-475.
URL: http://dx.doi.org/10.1901/jeab.1969.12-475.

[123] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton–proton collision data.* ATLAS-CONF-2014-032. 2014.
URL: http://cdsweb.cern.ch/record/1706245.

[124] Bernardo Resende. "Muon identification algorithms in ATLAS".
In: *PoS* EPS-HEP2009 (2009), p. 431.

[125] ATLAS Collaboration. "Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data".
In: *Eur. Phys. J. C* 74 (2014), p. 3130. DOI: 10.1140/epjc/s10052-014-3130-x.
arXiv: 1407.3935 [hep-ex].

[126] *"ATLAS Muon Combined Performance"*. [Online; accessed 27.06.2016]. 2012.
URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MCPAnalysisGuidelinesData2012.

[127] Mario Campanelli. private communication. 24th Aug. 2016.

[128] *Pile-up Suppression in Missing Transverse Momentum Reconstruction in the ATLAS Experiment in Proton-Proton Collisions at sqrts = 8 TeV.*
Tech. rep. ATLAS-CONF-2014-019. Geneva: CERN, 2014.
URL: https://cds.cern.ch/record/1702055.

[129] *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*. Tech. rep. ATLAS-CONF-2013-082. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1570993.

[130] S. Agostinelli and J. Allison. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250 –303. ISSN: 0168-9002. DOI: http://dx.doi.org/10.1016/S0168-9002(03)01368-8. URL: http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[131] J. M. Butterworth, Jeffrey R. Forshaw and M. H. Seymour. "Multiparton interactions in photoproduction at HERA". In: *Z. Phys.* C72 (1996), pp. 637–646. DOI: 10.1007/BF02909195,10.1007/s002880050286. arXiv: hep-ph/9601371 [hep-ph].

[132] *New ATLAS event generator tunes to 2010 data*. Tech. rep. ATL-PHYS-PUB-2011-008. Geneva: CERN, 2011. URL: https://cds.cern.ch/record/1345343.

[133] Brian Lindquist. private communication. 15th June 2016.

[134] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [hep-ph].

[135] Stefano Frixione, Paolo Nason and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[136] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[137] Michelangelo L. Mangano et al. "ALPGEN, a generator for hard multiparton processes in hadronic collisions". In: *JHEP* 07 (2003), p. 001. DOI: 10.1088/1126-6708/2003/07/001. arXiv: hep-ph/0206293 [hep-ph].

[138] Peter Zeiler Skands. "Tuning Monte Carlo Generators: The Perugia Tunes". In: *Phys. Rev.* D82 (2010), p. 074018. DOI: 10.1103/PhysRevD.82.074018. arXiv: 1005.3457 [hep-ph].

[139] ATLAS Collaboration. "Study of heavy-flavor quarks produced in association with top-quark pairs at $\sqrt{s} = 7$ TeV using the ATLAS detector". In: *Phys. Rev.* D89.7 (2014), p. 072012. DOI: 10.1103/PhysRevD.89.072012. arXiv: 1304.6386 [hep-ex].

[140] P. Kant et al. "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions". In: *Comput. Phys. Commun.* 191 (2015), pp. 74–89. DOI: 10.1016/j.cpc.2015.02.001. arXiv: 1406.4403 [hep-ph].

[141] M. Aliev et al. "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR". In: *Comput. Phys. Commun.* 182 (2011), pp. 1034–1046. DOI: 10.1016/j.cpc.2010.12.040. arXiv: 1007.1327 [hep-ph].

[142] Nikolaos Kidonakis. "Top Quark Production". In: *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)*. 2014, pp. 139–168. DOI: 10.3204/DESY-PROC-2013-03/Kidonakis. arXiv: 1311.0283 [hep-ph]. URL: https://inspirehep.net/record/1263209/files/arXiv:1311.0283.pdf.

[143] Borut Paul Kersevan and Elzbieta Richter-Was. "The Monte Carlo event generator AcerMC version 1.0 with interfaces to PYTHIA 6.2 and HERWIG 6.3". In: *Comput. Phys. Commun.* 149 (2003), pp. 142–194. DOI: 10.1016/S0010-4655(02)00592-1. arXiv: hep-ph/0201302 [hep-ph].

[144] ATLAS Collaboration. "Search for production of WW/WZ resonances decaying to a lepton, neutrino and jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C75.5 (2015). [Erratum: Eur. Phys. J.C75,370(2015)], p. 209. DOI: 10.1140/epjc/s10052-015-3593-4,10.1140/epjc/s10052-015-3425-6. arXiv: 1503.04677 [hep-ex].

[145] ATLAS Collaboration. "A search for $t\bar{t}$ resonances using lepton-plus-jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 08 (2015), p. 148. DOI: 10.1007/JHEP08(2015)148. arXiv: 1505.07018 [hep-ex].

[146] J. S. Conway. "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra". In: *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*. 2011, pp. 115–120. DOI: 10.5170/CERN-2011-006.115. arXiv: 1103.0354 [physics.data-an]. URL: https://inspirehep.net/record/891252/files/arXiv:1103.0354.pdf.

[147] Roger J. Barlow and Christine Beeston. "Fitting using finite Monte Carlo samples". In: *Comput. Phys. Commun.* 77 (1993), pp. 219–228. DOI: 10.1016/0010-4655(93)90005-W.

[148] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Ann. Math. Statist.* 9.1 (Mar. 1938), pp. 60–62. DOI: 10.1214/aoms/1177732360. URL: http://dx.doi.org/10.1214/aoms/1177732360.

[149] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an].

[150] Valerio Dao. private communication. 9th Mar. 2016.

[151] Jerome H. Friedman. "Data Analysis Techniques for High-Energy Particle Physics". In: *1974 CERN School of Computing, Godoysund, Norway, 11-24 Aug 1974: Proceedings*. 1974, p. 271. URL: http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=slac-r-176.html.

[152] ATLAS Collaboration. "Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data". In: *Eur. Phys. J.* C74.11 (2014), p. 3130. DOI: 10.1140/epjc/s10052-014-3130-x. arXiv: 1407.3935 [hep-ex].

[153] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data". In: *Eur. Phys. J.* C74.10 (2014), p. 3071. DOI: 10.1140/epjc/s10052-014-3071-4. arXiv: 1407.5063 [hep-ex].

[154] *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at sqrts = 7 TeV in 2011*. Tech. rep. ATLAS-CONF-2012-048. Geneva: CERN, 2012. URL: https://cds.cern.ch/record/1450089.

[155] ATLAS Collaboration. "Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV". In: *Eur. Phys. J.* C75 (2015), p. 120. DOI: 10.1140/epjc/s10052-015-3325-9. arXiv: 1408.3179 [hep-ex].

[156] ATLAS Collaboration. "Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector". In: *Eur. Phys. J.* C73.3 (2013), p. 2306. DOI: 10.1140/epjc/s10052-013-2306-0. arXiv: 1210.6210 [hep-ex].

[157] Andy Buckley et al. "LHAPDF6: parton density access in the LHC precision era". In: *Eur. Phys. J.* C75 (2015), p. 132. DOI: 10.1140/epjc/s10052-015-3318-8. arXiv: 1412.7420 [hep-ph].

[158] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC". In: *ArXiv Physics e-prints* (2016). arXiv: 1608.03953 [hep-ex].

[159] ATLAS Collaboration. *Measurement of WW/WZ production with a hadronically-decaying boson reconstructed as one or two jets, and a search for anomalous gauge couplings in pp collisions at $\sqrt{s} = 8$ TeV at ATLAS*. Tech. rep. Geneva: CERN, In Preparation.

[160] ATLAS Collaboration. "Measurement of WZ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C72 (2012), p. 2173. DOI: 10.1140/epjc/s10052-012-2173-0. arXiv: 1208.1390 [hep-ex].

[161] *Measurement of $W^{\pm}Z$ boson pair-production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector and confidence intervals for anomalous triple gauge boson couplings*. Tech. rep. ATLAS-CONF-2016-043. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2206093.

[162] Massimiliano Grazzini et al. "W^+W^- production at the LHC: fiducial cross sections and distributions in NNLO QCD". In: *JHEP* 08 (2016), p. 140. DOI: 10.1007/JHEP08(2016)140. arXiv: 1605.02716 [hep-ph].

[163] E. Bagnaschi et al. "Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM". In: *JHEP* 02 (2012), p. 088. DOI: 10.1007/JHEP02(2012)088. arXiv: 1111.2854 [hep-ph].

[164] *Luminosity Public Results Run2*. [Online; accessed 26.10.2016]. 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[165] Elizabeth Fons. "Improvement of the event selection for W-pairs in the semi-leptonic channel with a boosted hadronic W". MA thesis. University of Buenos Aires, Dec. 2015.

Acknowledgements

During the last four years in which this thesis has been developed, I received support from many different people and this is the place to express my gratitude to all of them.
First of all, I'd like to thank Prof. Dr. Bethke for giving me the opportunity to work in this very nice environment on a topic that couldn't be more interesting.
Thanks also go to Dr. Béla Majorovits for his useful advice during the Advisory Panel meetings.
I would like to express my special thanks to Sven for his great support and supervision during the whole time of working on this thesis. Even though I know he was rather busy himself, he always took the time for me to explain things and even go through the deepest parts of my code.
Big thanks I owe to Andi, without whose knowledge of Athena and C++ I would have been completely lost many times and who also let me stay in his nice flat in Sergy during my time at CERN. Numerous expeditions to Aumeister won't be forgotten either.
I would like to thank Tom for the small breaks sitting on the grass behind the institute, the interesting physics discussions, after-work activities, countless English questions and his help concerning coding that got me started in my PhD time.
Many thanks also go to the rest of our group Denis, Giselher, Horst, Jörg, Peter and Teresa for their support during loong meetings. In particular I want to thank Andrei for the very nice time I had sharing an office with him.
Thanks and best wishes also go to Eli with whom I had a great time working and sharing the office as well: good luck for your future in Great Britain.
Furthermore, I want to express my gratitude towards Halina, Aharon and Orel for making my times in Israel so educational and enjoyable and for all of the very helpful support concerning physics, multivariate analysis methods and coding in general.
It was a great pleasure to work with the group of Brian Lindquist, Chiara Roda, Dmitri Tsybyshev, Felix Bührer, Mario Campanelli, Sasha Solodkov and Valerio Dao during my last year. I particularly wish to thank David Freeborn, Margharita Spalla and Milène Calvetti for their support implementing the analysis and tracking down bugs in many hours of detailed kinematics comparisons via Skype.
There are many other members of the institute who deserve to be mentioned here:
I would like to thank the "lunch-crew", Frau Träder, Hans, Barbara Wankerl, Christiane Winter, Peter Seyboth, Stefan Stonjek and Stefan Kluth who (except for the two Stefans) are no ATLAS members but still kindly joined the ATLAS table as often as they could and made lunch an interesting time where it was possible to talk about something other than physics.
My grateful thanks also go to Anja Schielke, Dieter Fischer and Corina Brunnlechner who were always very nice and really helpful when it came to administrative matters.
On the private side, I'd like to thank my father, whose explanations about stars and atoms during my childhood surely contributed to my interest in physics, my sister for the always funny phone calls and joint vacations and especially my mother who sadly passed away last year and to whom this work is dedicated.

I also owe many thanks to all of the current and former flatmates I had during my time in Munich: Philipp, Lisa and Irmi who received me very warmly and welcoming when I moved here, Myri, Aharon, Lupi, Anne, Sarah and Jonas with whom I had an awesome time and living together has always been a pleasure.
Last but not least, without being able to name them all, I want to thank my friends in Munich, Nuremberg, Würzburg, Hamburg, Berlin somewhere else in Germany or in the world, just for being there.