



Master Thesis

Jochen Jens Heinrich

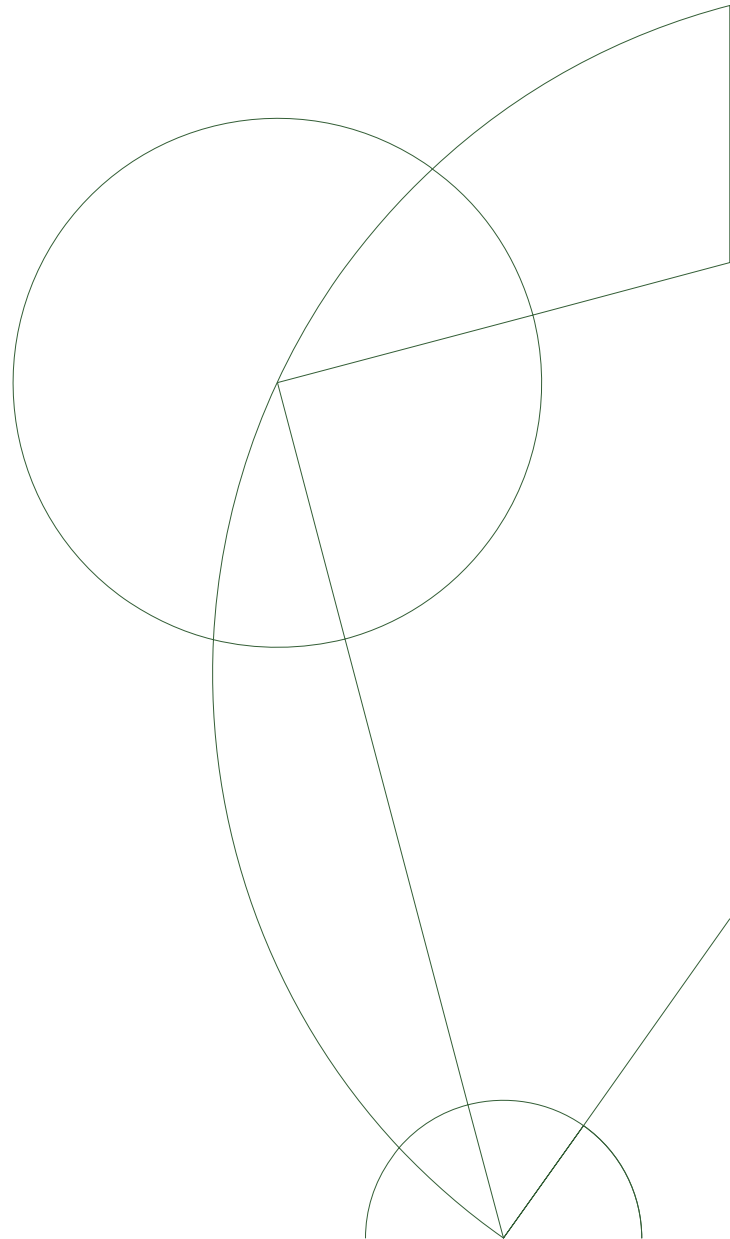
Reconstruction of boosted $W^\pm$ and Z^0 bosons from fat jets

CERN-THESIS-2014-152
12/09/2014



Academic supervision by Peter Henrik Hansen

August 15th 2014



Abstract

We present the reconstruction of heavily boosted $W^\pm$ and Z^0 bosons from large R-parameter jets (fat jets) in all-hadronic proton-proton collisions at $\sqrt{s} = 8$ TeV at the LHC. The electroweak gauge bosons are boosted to a degree at which their hadronic decay products are collimated enough to be reconstructed as a single fat jet. A mass-drop filtering procedure which is validated in studies on Monte Carlo (MC) samples is then applied to the fat jets with $p_T > 420$ GeV to suppress pileup and soft radiation. $W^\pm$ and Z^0 bosons are identified based on their filtered jet mass. The efficiency of common substructure observables and event shape variables, in distinguishing between signal and QCD background is evaluated on MC and the optimized observable selection used for the training of two boosted decision trees (BDT), in order to reduce the dijet background not originating from the decay of an electroweak gauge boson. For the first BDT, signal MC has been trained against background MC, while for the second one the background is taken from sidebands in data. A peak corresponding to the mass of the electroweak gauge bosons can be observed in the filtered fat jet mass spectrum. The production cross section for hadronically decaying electroweak gauge bosons for $p_T > 420$ GeV and $|\eta| < 1.9$ is measured to be $\sigma_1 = 0.557 \pm 0.032$ (stat.) ± 0.054 (sys.) pb and $\sigma_2 = 0.599 \pm 0.031$ (stat.) ± 0.058 (sys.) pb for the two BDTs respectively. The data sample corresponds to 20.3 fb^{-1} collected with the ATLAS detector in 2012.

Preface

I consider myself very lucky. This is because luck was a necessary component for being able to work on my master thesis in a time where generations of particle physicists before me have piled up their knowledge about nature and achieved to form an understanding of the basics of the world surrounding us. To get to this point, an immense technical effort to develop better and more efficient tools was necessary that led to more and more powerful particle accelerators culminating in the construction of the current record holder - the Large Hadron Collider (LHC) at CERN. Just as the ATLAS and CMS detector, it is among the biggest and most complex machines ever built by humankind and it took nearly three decades from first concepts to the arrival of the first data which was eagerly awaited not only by the scientific community. Working at the ATLAS experiment among a collaboration of over 3000 scientist from 177 universities in 38 countries all over the world was always a privilege to me.

The announcement of the discovery of the Higgs boson in July 2012 marked the completion of the Standard Model of particle physics. A model that has been developed, tested, and verified for the past fifty years. Naturally, the approach on particle physics is shifting as a result. Instead of working in the framework of one dominant theory a multitude of different hypothesis beyond the Standard Model have to be considered and many questions can be raised. Will supersymmetry prevail? What is the nature of dark matter? Is there a graviton as force-carrier of the gravitational force? Or is this maybe it and there is nothing more to discover? We can not know until we made every possible effort. There is no clear view on what to expect in the future and in the end nature might surprise us all. What we enter is an open quest to unravel nothing less than the mysteries of the universe.

Is it possible to imagine a more exciting prospect than this?

The work conducted and documented in this thesis is my humble effort to contribute to this fantastic endeavour and help to pave the way for future studies in search for new physics.

This thesis was written with future master students in mind, who just started their master projects on a related topic of particle physics. It is meant to give a summary of some of the insights I have gained in the course of the last year. Obviously, the limited scope of the thesis prohibits to go into much detail but I have put great effort into always citing the reference which helped me most in gaining an understanding of a certain issue. They can always be considered as a suggestion for further reading. The time-consuming search for good reviews and sources of information is hopefully reduced by my work.

Acknowledgements

A lot of people have contributed one way or another to the completion of this thesis. First and foremost, I want to thank Troels C. Petersen for many physical discussions, his constant help and guidance throughout all stages of the project and for reading parts of my thesis in advance giving a lot of helpful feedback. A big thank you goes to Peter H. Hansen for being my supervisor on this project and giving me the possibility to work on this exciting subject with many skilled colleagues. Thank you for your help and ideas in any problem I brought to you. Further, I like to thank Alejandro Alonso for helping me out with any kind of bizarre coding problems I created in the past year and for producing the skimmed samples I used. Björn S. Nilsson has my gratitude for many interesting conversations and never failing to point out a funny story on the internet to me. Thank you to Andreas, Alexander, Gitte, Mikkel, Morten and especially Simon for sharing the office with me over the course of this project and always creating a pleasant and productive working environment. Indeed, I very much enjoyed working at the Niels Bohr Institute for this past year and I am very grateful to all members of the work group for always being helpful and letting me feel very welcome. Jürgen Freund deserves to be mentioned here for igniting my joy for physics. Thank you to all my friends for many relaxing evenings with discussions and philosophy over a beer. For the constant support I got from my family I am deeply grateful. Lastly, this thesis would not have been possible in this form without Laura Oestreich who kept encouraging me in everything I did and gave me the support I sometimes needed.



Contents

i	Abstract	iii
ii	Preface	v
	Contents	vii
1	Introduction	1
I	Theory	3
2	Theoretical Background	5
2.1	The Standard Model of Particle Physics	5
2.1.1	The fundamental particles	5
2.1.2	Quantum Field Theory	7
2.2	Lagrangian and local gauge symmetries	7
2.3	Weinberg-Salam model of electroweak interaction	9
2.4	Quantum Chromodynamics	10
2.5	Spontaneous breaking of symmetry and Higgs mechanism	11
2.6	Perturbation theory and Feynman diagrams	13
3	The ATLAS Experiment at the Large Hadron Collider	15
3.1	The Large Hadron Collider (LHC)	15
3.2	The ATLAS detector	18
3.2.1	Inner Detector	20
3.2.2	Calorimeters	22
3.2.3	Muon Spectrometer	24
3.2.4	Trigger	25
3.3	Event simulation	26
4	Jet Physics at the LHC	27
4.1	Origin of Jets	27
4.2	Jet reconstruction algorithms	28
4.3	Parton Distribution Functions	30
4.4	Contamination by underlying event and pileup	32
4.4.1	Underlying event	32
4.4.2	Pileup	32
4.5	Jet grooming algorithms	32

4.5.1	Mass-drop Filtering	33
4.5.2	Jet Trimming	35
4.5.3	Jet Pruning	36
II	Analysis	39
5	Initial studies on Monte Carlo	41
5.1	Validating the mass-drop filtering procedure	41
5.2	Suppression of pileup	43
5.3	Validation of the filtered mass	44
5.4	Jet charge	45
6	Boosted $W^\pm$ and Z^0 bosons from fat jets	49
6.1	Introduction	49
6.2	Monte Carlo and data samples	49
6.3	Event and jet selection	52
6.4	Unfiltered jet mass spectrum	54
6.5	Discriminating variables	57
6.5.1	Substructure Observables	57
6.5.2	Event shape variables	61
6.5.3	Additional discriminators	64
6.6	MVA training and performance	65
6.6.1	MVA training	65
6.6.2	Removing correlations between the BDT and m_{jet}	69
6.7	Fitting	75
6.7.1	Development of a fit model	75
6.7.2	Fitting BDT(MC)	79
6.7.3	Fitting BDT(DATA)	80
6.7.4	Computation of the production cross sections	82
6.8	Systematic errors	83
6.8.1	Systematic uncertainty of the signal PDF	84
6.8.2	Systematic uncertainty of the background PDF	85
6.9	Results	86
7	Conclusion and outlook	87
	Appendix A Momentum measurement	93
	Appendix B Main parameters of the calorimeter system	95
	Appendix C Boosted Decision Trees (BDTs)	99
	Appendix D BDT correlation plots	103
	Appendix E Additional plots and figures	107
	List of Figures	111
	List of Tables	113
	Bibliography	115

Introduction

The concept of a 'Large Hadron Collider in the LEP Tunnel' was first discussed thirty years ago in 1984 on a workshop in Lausanne hosted by the European Committee for Future Accelerators (ECFA) [1]. At this point, the essential idea of the Standard Model of particle physics, the Higgs mechanism [2, 3, 4], was already twenty years old. But the interest in the theory sparked after the predicted $W^\pm$ and Z^0 electroweak gauge bosons were discovered at the UA1 and UA2 experiments at CERN's¹ Super Proton Synchrotron (SPS) [5, 6, 7, 8]. Thus began the hunt for the elusive Higgs boson. In 1992 the ATLAS and CMS collaborations had published their Letters of Intent for two massive general-purpose detectors that were designed to find the Higgs boson if it existed [9, 10]. The next fundamental breakthrough followed in 1995 when the top quark was discovered at Fermilab [11, 12]. After the ATLAS detector was approved in 1997, the excavation works for the cavern a hundred meters underground began and were completed in 2003 having taken three years of civil engineering. Extensive test beam studies on the ATLAS detector above-ground preceded its assembly in the cavern. A process that took another four years. On September 10th 2008 a proton beam was successfully steered around the 27 km long LHC tunnel for the first time. However, the data taking was delayed for over a year following an accident that occurred on September 19th 2008 [13], but resumed in late 2009. LHC delivered proton-proton collision data throughout the end of 2012, albeit at lower than design energy. On July 4th 2012 both the ATLAS and CMS collaboration announced the discovery of a Higgs-like new boson based on the data collected so far [14, 15]. With the discovery of the Higgs boson the Standard Model of particle physics is complete. However, the particle physicists job is not. There exists a large variety of theories beyond the Standard Model that predict new particles that very often exhibit decay modes to electroweak gauge bosons which cannot be detected directly but can be reconstructed from their decays to detectable particles. Since we have so far not observed any additional particles it is assumed that they have masses that are inaccessible to us with current experimental techniques. To expand our reach as far as possible it is necessary to bid farewell to the clean signatures of leptonic decay channels due to their very limited relative frequency. [16] quotes the branching ratios per lepton flavour as only $\Gamma_i/\Gamma(W^\pm \rightarrow l^\pm \nu) = (10.80 \pm 0.09) \%$ and $\Gamma_i/\Gamma(Z^0 \rightarrow l^+ l^-) = (3.3658 \pm 0.0023) \%$ which makes it very difficult to get statistically significant event numbers at high masses. Instead the hadronic decay modes with branching ratios of $\Gamma_i/\Gamma(W^\pm \rightarrow \text{hadrons}) = (67.60 \pm 0.27) \%$ and $\Gamma_i/\Gamma(Z^0 \rightarrow \text{hadrons}) = (69.91 \pm 0.06) \%$ look very appealing to us. However, there exist some drawbacks besides the unclearer signatures and measurement conditions. Jets are large objects involving many particles from showering processes. The

¹The European Organization for Nuclear Research. The name CERN is an acronym for the former French name 'Conseil Européen pour la Recherche Nucléaire'.

decay of electroweak gauge bosons initiates two jets than can be closely collimated when the decaying particle is boosted. In that case, standard jet clustering algorithms have a hard time resolving two separate jets or fail altogether. If so, the decay channel is lost to us. This scenario becomes even more frequent after LHC restarts with 13 TeV and 14 TeV in early 2015. Consequently the last years have seen a great deal of work on salvaging the boosted hadronic decay channels.

The subject of this master thesis is to validate a fairly new procedure called mass-drop filtering that is based on jet substructure and seeks to identify symmetric subjets within one large R-parameter jet. The validation is carried out by reconstructing boosted $W^\pm$ and Z^0 bosons and measuring their production cross section. In the future, the technique can be used to search for new physics.

The thesis is divided into two main parts. Part I contains the chapters 2 to 4 and focuses on the theoretical basics necessary to follow the work described here. Chapter 2 introduces the Standard Model of particle physics including the particle content. The experimental hardware, meaning the Large Hadron Collider and the ATLAS detector are described in detail in chapter 3. Chapter 4 briefly introduces the physics around jets at the LHC, including descriptions of the function principle of jet reconstruction algorithms, parton distribution functions and contaminating influences on the measurement. The chapter also describes several jet grooming algorithms like the mass-drop filtering which is used in this work. Part II contains the analysis done over the course of one year. Chapter 5 summarizes several smaller studies that were carried out on Monte Carlo samples at the beginning of the project. The main part of the thesis is found in chapter 6 containing the reconstruction of boosted vector bosons from fat jets. A short conclusion and outlook is given in chapter 7.

Part I

Theory

Theoretical Background

In the following chapter we want to introduce the theoretical fundamentals of the present analysis. This includes the centerpiece of modern particle physics - the standard model with its different particles and interactions in 2.1. The remaining part of the chapter will focus on brief descriptions of the theoretical framework from Quantum Field Theory (QFT) like Euler-Lagrange equations, symmetries and gauge invariance in 2.2. Brief discussions of electroweak interactions and Quantum Chromodynamics (QCD) follow in 2.3 and 2.4. In 2.5 the Higgs mechanism is discussed. The chapter concludes with descriptions of perturbative QCD and Feynman diagrams in 2.6.

2.1 The Standard Model of Particle Physics

Over the course of the last decades particle physicists have worked to gain an understanding of the fundamental structure of matter and their interactions. Accompanied by experimental discoveries this work culminated in the creation of the Standard Model (SM) towards the end of the 1960s. The combination of electromagnetic force and weak force into the electroweak theory by Glashow in 1960 [17] was expanded by Weinberg [18] and Salam [19] to incorporate the Higgs mechanism a few years later. Since then the model was put on trial by experimental verification of its predictions over and over again and it is yet to fail a major test. According to the Standard Model, the universe is composed of a very limited number of fundamental particles whose interactions are governed by three principle forces: electromagnetic, weak and strong force. Despite its great success the Standard Model is incomplete as it lacks the incorporation of gravitation as a fourth fundamental force and fails to provide an accurate description of many features of the observable universe like neutrino oscillation, dark matter or baryon asymmetry. Those and many more unanswered questions should keep particle physicists busy for many decades to come.

In the following an overview over the various elementary particles of the standard model is given.

2.1.1 The fundamental particles

The largest group of particles in the Standard Model are the fundamental matter particles, namely leptons and quarks. They are the bricks that form the matter surrounding us. Each of the two groups features three generations with ascending masses and descending stability. The first generation of the leptons contains the electron e^- as the lightest lepton and its corresponding neutrino ν_e . The second and third generation contain the heavier muon μ^- and tau τ^- with their related neutrinos ν_μ and ν_τ , respectively. In the group of the quarks the lightest particles are again found in the first generation with “up quark” u

Three generations
of matter (fermions)

	I	II	III		
mass →	2.4 MeV/c ²	1.27 GeV/c ²	171.2 GeV/c ²	0	? GeV/c ²
charge →	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0	0
spin →	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	0
name →	u up	c charm	t top	γ photon	H Higgs boson
Quarks	4.8 MeV/c ²	104 MeV/c ²	4.2 GeV/c ²	0	
	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	d down	s strange	b bottom	g gluon	
Leptons	<2.2 eV/c ²	<0.17 MeV/c ²	<15.5 MeV/c ²	91.2 GeV/c ²	
	0	0	0	0	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	Z⁰ Z boson	
Gauge bosons	0.511 MeV/c ²	105.7 MeV/c ²	1.777 GeV/c ²	80.4 GeV/c ²	
	-1	-1	-1	±1	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	e electron	μ muon	τ tau	W[±] W boson	

Figure 2.1: Overview over all fundamental particles in the Standard Model with their corresponding spin, charge and mass. The figure was taken from [22].

and “down quark” d . “charm quark” c and “strange quark” s form the second, “top quark” t and “bottom² quark” b the third. A fourth generation of fermions was often speculated about but measurements seem to tend towards an exclusion of this model [20]. Leptons are subject to the electromagnetic and weak force, while Quarks in addition also interact via the strong force. Therefore, Quarks carry a colour charge and come in three different colour flavours. However, observable objects are always colour neutral. Therefore, it is impossible to observe an isolated quark. Several quarks can form bound objects held together by the strong force which either consist of two quarks, in which case we speak of mesons, or three quarks, where we call the object a baryon. Mesons and baryons are collectively called hadrons. Every stable form of matter is composed of particles from the first generation, as the heavier particles are allowed to decay into the lighter particles. All quarks and leptons are fermions, i.e. they carry a half-integer spin. In addition to the matter particles described here, there are also the corresponding anti-particles which are duplicates of the matter particle in respect to mass, magnetic moment and spin, but have an opposite electrical charge and opposite parity. They are either denoted by a bar over the symbol, e.g. $\bar{b}$ and $\bar{\nu}_e$ as anti-particles to b and ν_e , or distinguished by exploiting the opposite charges of the anti-particles, e.g. e^+ for the positron as anti-particle to the electron e^- . The fermion fields have two components of opposite helicity which we call left- and right-handed. The theory of electroweak interaction is a chiral theory and does not treat the two components on the same footing. The right-handed neutrino for example is presumed massless³ and does not interact in any known way. It is therefore unobservable for us.

The second major group of the elementary particles are the gauge bosons, which mediate three of the four known forces and possess an integer spin. The photon γ is the carrier for the electromagnetic interaction. It is a massless, wide-range (inverse-square) particle

²Sometimes also called “beauty quark”

³A non-zero neutrino mass is required to account for neutrino oscillation. So far only upper bounds for this mass have been measured showing that it is indeed extremely small [21].

which renders the electromagnetic force non-negligible not only in subatomic ranges but also on macroscopic scales. The exchange particles for the weak nuclear interaction are W^+ , W^- and Z^0 bosons. In contrast to the photon, the $W^\pm$ and Z^0 exhibit a mass of $m_W = 80.4 \text{ GeV}/c^2$ and $m_Z = 91.2 \text{ GeV}/c^2$ [16]. The weak force is short-ranged and only relevant on atomic and subatomic scales. Lastly, the gluon g mediates the strong nuclear interaction. It carries a colour charge just like the quarks and is responsible for the binding of quarks. The strong force is very short-ranged, but as the name indicates very much stronger than the electromagnetic or weak force on subatomic scales. The fourth fundamental interaction is not included in the Standard Model. The graviton as force carrier was speculated about but remains unobserved to this day, e.g. [23, 24]. Luckily, compared to the other three interactions the gravitational force is orders of magnitude smaller at the relevant minuscule scales and therefore negligible.

The last particle of the Standard Model is the spin-0 Higgs boson. It exhibits no electrical or colour charge and can be understood as an excitation in the Higgs field which adheres to the gauge bosons of the weak interaction and the fermions to give them mass. A particle with compatible characteristics to the Standard Model Higgs boson has recently been discovered at the Large Hadron Collider (LHC) [14, 15, 25, 26].

A graphical overview over all the particles in the Standard Model along with their spin, mass and electrical charge can be seen in figure 2.1. The mass of the Higgs boson is not yet included in the figure, even though it is strongly indicated that the new found particle is the Higgs boson [27]. If so, the mass of the Higgs is given by $m_H = 125.5 \pm 0.2(\text{stat.})^{+0.5}_{-0.6}(\text{sys.}) \text{ GeV}$ [28] according to ATLAS measurements.

2.1.2 Quantum Field Theory

The mathematical foundation of the standard model is called Quantum Field Theory (QFT)[29] and describes the merging of quantum mechanics with classical field theories like for example electromagnetism. The resulting field theory in this case is called Quantum Electrodynamics on which we will have a closer look in the following section. Furthermore, in order to deliver an accurate description of particle physics, the QFT needs to include concepts from the special theory of relativity in what is called a relativistic Quantum Field Theory. In a QFT, particles appear as an excitation in the particles underlying physical field. Their interactions are accounted for by adding interaction terms between the corresponding physical fields. In contrast to classical field theories, where there is a finite number of degrees of freedom for a fixed number of particles, a QFT does not have these restrictions and can therefore also describe systems with varying particle numbers.

In the following chapters the most basic concepts of the QFT underlying the SM are introduced.

2.2 Lagrangian and local gauge symmetries

For the remainder of this work we will follow common practice and work in natural units where the reduced Planck constant $\hbar$ and the speed of light in vacuum c is set to

$$\hbar = c = 1 \quad . \quad (2.1)$$

Factors of $\hbar$ and c can always be restored by dimensional analysis. For this and the following chapters we follow the argumentation of [30] which gives an accessible approach on theoretical particle physics.

All kinds of interactions between particles are governed by local gauge symmetries. From this idea it can be followed that conserved quantities like electric charge are not

just globally, but locally conserved. It is easiest to discuss this using the Euler-Lagrange formalism.

The Euler-Lagrange equation is given by

$$\frac{\partial}{\partial x_\mu} \left(\frac{\partial \mathcal{L}}{\partial \frac{\partial \phi(x_\mu)}{\partial x_\mu}} \right) - \frac{\partial \mathcal{L}}{\partial \phi(x_\mu)} = 0 \quad (2.2)$$

where ϕ is a field parametrised by the continuous function x_μ and $\mathcal{L}$ is called the Lagrangian density related to the Lagrange function that is used in classical mechanics to describe a single particle [31] via

$$L = \int d^4x \mathcal{L}(\phi, \partial\phi/\partial x_\mu) \quad . \quad (2.3)$$

Forthwith, we will adapt common terminology and refer to the Lagrangian density itself solely as Lagrangian.

Clever choices of the Lagrangian will result in well-known equations of motion. As one example the substitution of

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2 \phi^2 \quad , \quad (2.4)$$

where $\partial_\mu \phi$ is shorthand for $\partial/\partial x_\mu$ (and correspondingly $\partial^\mu \phi = \partial/\partial x^\mu$), yields the Klein-Gordon equation

$$\partial_\mu \partial^\mu \phi + m^2 \phi = (\square^2 + m^2)\phi = 0 \quad . \quad (2.5)$$

The same way one can construct the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \quad (2.6)$$

for a free fermionic field ψ by simply using the Lagrangian

$$\mathcal{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi \quad . \quad (2.7)$$

Here $\bar{\psi}$ is defined as $\bar{\psi} = \psi^\dagger \gamma^0$ and γ_μ are the Dirac matrices. Let us now assume that $\psi(x)$ is a complex field describing an electron. We can see at once that eq. 2.7 is invariant under a global phase (gauge) transformation $\psi \rightarrow e^{i\alpha}\psi$ with the real constant α . All of these transformations form a unitary Abelian group called the $U(1)$ group. Noether's theorem now states that there exists a conserved quantity - in our case the electric charge (shown in [30]).

Generalizing this concept to local gauge transformations of the form $\psi \rightarrow e^{i\alpha(x)}\psi$, where $\alpha(x)$ now is an arbitrary function of space and time, we see that the Lagrangian is no longer invariant under this kind of transformation. We define the covariant derivative

$$D_\mu \equiv \partial_\mu - ieA_\mu \quad , \quad (2.8)$$

where A_μ is a vector field called the gauge field that transforms like $A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu \alpha$ and e is the charge of the Dirac particle. Replacing ∂_μ by D_μ in eq. 2.7 yields

$$\mathcal{L} = i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi \quad (2.9)$$

$$= i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi + e\bar{\psi}\gamma^\mu \psi A_\mu \quad (2.10)$$

restoring the local gauge invariance.

The gauge field A_μ couples⁴ to the Dirac particle just like the photon field. Indeed, the final term in eq. 2.10 can be rewritten as $-j^\mu A_\mu$ where j^μ symbolizes the current density. If we are to regard A_μ as the physical photon field, it is necessary to add a new term to the Lagrangian to account for the field's kinetic energy. The constraints on the term in requiring it to be invariant under local gauge transformation lead us to the conclusion that it has to involve the field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad . \quad (2.11)$$

The end-product we get is the Lagrangian of Quantum Electrodynamics (QED)

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad . \quad (2.12)$$

It can be noted that mass terms of the form $\frac{1}{2}m^2 A_\mu A^\mu$ are not allowed due to gauge invariance. Therefore, the photon as the gauge particle must be massless.

The Standard Model can be fully described by the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. Here the $SU(2)_L \times U(1)_Y$ subgroup describes electroweak interactions and the $SU(3)_C$ subgroup originates from Quantum Chromodynamics. In the following both concepts are discussed briefly.

2.3 Weinberg-Salam model of electroweak interaction

If we want to include weak interactions into the Lagrangian from chapter 2.2 we have to take into account that electroweak processes are invariant under weak hypercharge transformation $U(1)_Y$ and under transformation of the weak isospin $SU(2)_L$. We therefore introduce two new interactions - a weak hypercharge current j_μ^Y that couples to a field B^μ with coupling constant g_Y and an isotriplet of weak currents $\vec{J}_\mu$ that couples to three additional fields denoted $\vec{W}^\mu$ with coupling g_T :

$$-\frac{ig_Y}{2}j_\mu^Y B^\mu = -ig_Y\bar{\psi}\gamma_\mu\frac{Y}{2}\psi B^\mu \quad (2.13)$$

$$-ig_T\vec{J}_\mu \cdot \vec{W}_\mu = -ig_T\bar{\chi}_L\gamma_\mu\vec{T} \cdot W^\mu\chi_L \quad (2.14)$$

Here, Y and T are the generators of the $U(1)_Y$ and $SU(2)_L$ transformation respectively. The combined subgroup contains transformations of the form

$$\chi_L \rightarrow e^{i\vec{\alpha}(x)\cdot\vec{T}+i\beta(x)Y}\chi_L \quad (2.15)$$

$$\psi_R \rightarrow e^{i\beta(x)Y}\psi_R \quad , \quad (2.16)$$

where L marks the left-handed components that form isospin doublets with isospin $T = \frac{1}{2}$ and accordingly R the right-handed components that form isosinglets with $T = 0$. Again, $\vec{\alpha}$ and β are arbitrary functions that depend on time and space. The hypercharge can be calculated using the relation

$$Q = T^3 + \frac{1}{2}Y \quad , \quad (2.17)$$

where Q is the electric charge. This means that the electromagnetic current is a combination of the neutral currents j_μ^Y and $\vec{J}_\mu^3$. Plainly,

$$j_\mu^{EM} = j_\mu^Y + \vec{J}_\mu^3 \quad . \quad (2.18)$$

⁴by which we mean "interacts"

From eq. 2.13 and 2.14 one can see that we can define two physical gauge fields A_μ and Z_μ as orthogonal combinations of the unphysical fields B_μ and W_μ^3 . Defining the mixing angle θ_W ⁵, we can express the physical fields as

$$A_\mu = B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W \quad (2.19)$$

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

and in addition

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}} \quad (2.21)$$

which obviously correspond to the photon and the electroweak vector bosons Z^0 and $W^\pm$. Constraints from the electromagnetic interaction make it furthermore possible to fix the coupling constant g_Y and g_T to

$$e = g_Y \cos \theta_W = g_T \sin \theta_W \quad . \quad (2.22)$$

With that the final free Lagrangian containing only the kinetic energies and self-interaction terms for the fields W_μ and B_μ is given by

$$\mathcal{L}_{free}^{WS} = -\frac{1}{4} \vec{W}_{\mu\nu} \cdot \vec{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} \cdot B^{\mu\nu} \quad , \quad (2.23)$$

where

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (2.24)$$

and

$$\vec{W}_{\mu\nu} = \partial_\mu \vec{W}_\nu - \partial_\nu \vec{W}_\mu - g_T \vec{W}_\mu \times \vec{W}_\nu \quad . \quad (2.25)$$

Just like in the previous chapter, gauge invariance forbids to introduce mass-terms to the Lagrangian and the electroweak gauge bosons are therefore massless. This is obviously in contrast to reality and solved by the Higgs mechanism (see 2.5).

2.4 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is rooted in the $SU(3)_C$ gauge subgroup that contains phase transformations of the form

$$Uq(x) \rightarrow e^{i\alpha_a(x)\frac{\lambda_a}{2}} q(x) \quad , \quad (2.26)$$

where λ_a are the Gell-Mann matrices and a sum over $a = 1, 2, \dots, 8$ is implied. Analogous to chapter 2.2, we seek to impose local gauge invariance on the Lagrangian. This we do by introducing eight new gauge fields G_μ^a which transform according to

$$G_\mu^a \rightarrow G_\mu^a - \frac{1}{g_C} \partial_\mu \alpha_a - f_{abc} \alpha_b G_\mu^c \quad (2.27)$$

with the strong coupling constant g_C and the structure constant f_{abc} of the $SU(3)_C$ subgroup. Further, we define the covariant derivative

$$D_\mu = \partial_\mu + ig_C \frac{\lambda_a}{2} G_\mu^a \quad . \quad (2.28)$$

⁵Often called Weinberg angle

The Lagrangian for all gluon fields is given by

$$\mathcal{L}_{free}^{QCD} = -\frac{1}{4}G_a^{\mu\nu}G_{\mu\nu}^a \quad (2.29)$$

with

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_C f_{abc} G_\mu^b G_\nu^c \quad . \quad (2.30)$$

Through requirements of local gauge invariance the gluons are massless.

2.5 Spontaneous breaking of symmetry and Higgs mechanism

In the theory we have discussed so far fermions and electroweak gauge bosons are massless. Since observations from nature suggest differently, a solution needs to be found. If one were to break the symmetry of the Lagrangian by adding mass terms of the form $m^2 W_\mu W^\mu$ by hand, one would encounter divergences that lead to a non-renormalizable theory. Such a theory has no predictive capability and is therefore utterly meaningless. Instead, the symmetry of the Lagrangian is broken spontaneously by what is called the Higgs mechanism.

We want to modify the theory so the $W^\pm$ and Z^0 bosons become massive, while keeping the photon massless. In order to achieve this, four real scalar fields are introduced. Keeping the Lagrangian gauge invariant demands them to belong to $SU(2) \times U(1)$ multiplets and we arrange them in an isospin doublet called the Higgs doublet

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (2.31)$$

with hypercharge $Y = 1$. The Lagrangian now looks like

$$\mathcal{L} = |\mathcal{D}_\mu \phi|^2 - V(\phi) \quad , \quad (2.32)$$

where the notation $|\phi|^2 = \phi^\dagger \phi$ is used and

$$\mathcal{D}_\mu = i\partial_\mu - g_Y \frac{Y}{2} B_\mu - g_T \vec{T} \cdot \vec{W}_\mu \quad . \quad (2.33)$$

$V(\phi)$ is called the Higgs potential and is of the form

$$V(\phi) = \mu^2 |\phi|^2 + \lambda (|\phi|^2)^2 \quad (2.34)$$

with two complex parameters μ and $\lambda > 0$. If $\mu^2 > 0$ the resulting potential corresponds to a self-interacting scalar field with mass μ whose ground state is given by $\phi = 0$. Of much greater interest is the opposite case where $\mu^2 < 0$. The resulting potential has a ring of minima at positions

$$|\phi|^2 = -\frac{\mu^2}{\lambda} \equiv v \quad (2.35)$$

and $\phi = 0$ does not correspond to the vacuum expectation value ϕ_0 of the system which is non-vanishing. The potential that is often referred to as Mexican hat potential is drawn in figure 2.2.

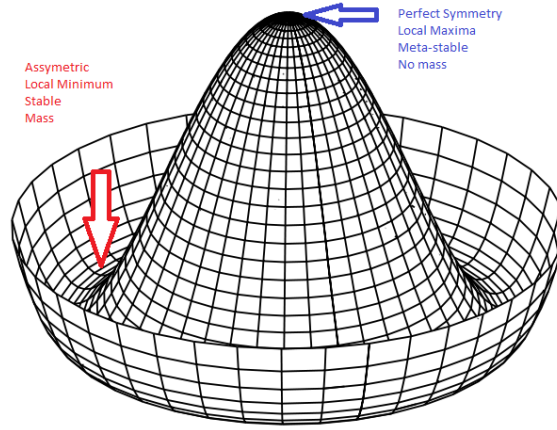


Figure 2.2: The Higgs potential drawn in three dimensional space. It becomes apparent why the potential is often referred to as Mexican hat potential. The local maximum value in the center corresponds to $\phi = 0$. A ring of minima can be seen around it. The figure was taken from [32].

Perturbative calculations describing quantum fluctuations around a minimum require a choice of the minimal value. This we call spontaneous symmetry breaking. We choose

$$\phi_0 = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.36)$$

and the Higgs field is represented by

$$\phi(x) = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} . \quad (2.37)$$

Here, $h(x)$ is the neutral Higgs field from eq. 2.31 where the remaining three fields have been gauged. Substituting the vacuum expectation value into the Lagrangian, we get the relevant term

$$\frac{g^2 v^2}{8} \left((W_\mu^1)^2 + (W_\mu^2)^2 \right) + \frac{v^2}{8} \left(g_Y B_\mu - g_T W_\mu^3 \right) \left(g_Y B^\mu - g_T W^{3\mu} \right) . \quad (2.38)$$

Mass terms are typically of the form $M_W^2 W^+ W^-$. Using eq. 2.21, we can compare this to equation 2.38 and find the mass of the $W^\pm$ boson is given by

$$M_W = \frac{vg_T}{2} . \quad (2.39)$$

Through a little bit of work one arrives at the mass term for the Z^0 boson

$$m_Z = \frac{m_W}{\cos \theta_W} = \frac{vg_T}{2 \cos \theta_W} . \quad (2.40)$$

Furthermore $M_A = 0$, i.e. the photon is still massless as was desired.

It should be noted that the Higgs-mechanism presented here is not the only possibility to achieve electroweak symmetry breaking. Alternative theories are beyond the scope of this thesis but include (minimal) supersymmetric Higgs bosons [33], Higgsless models with extra-dimensions [34], gauge-Higgs unification scenarios [35, 36] or little Higgs models [37].

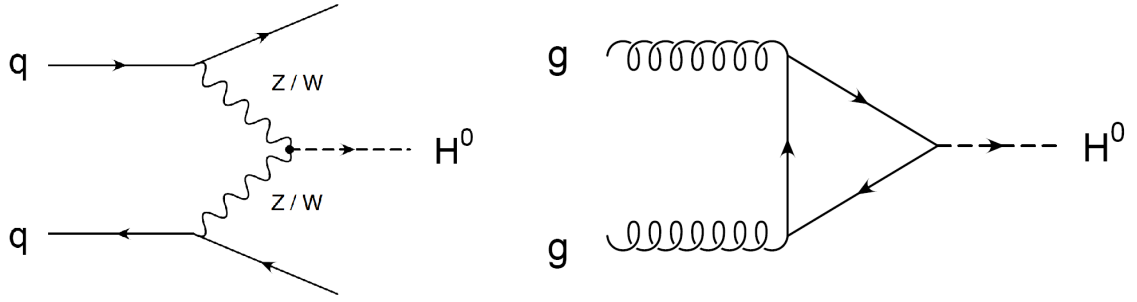


Figure 2.3: Two examples of Feynman diagrams showing two different Higgs production channels. Left: A Higgs boson is produced via vector boson fusion. Time runs from left to right, which makes the lower solid line an anti-quark as it runs backwards. Right: gluon-gluon fusion Higgs production. The Higgs is produced via a loop which indicates a NLO diagram. The figures were taken from [39].

2.6 Perturbation theory and Feynman diagrams

The main interest of particle physicists is the cross section of a process. It can be calculated using the Lagrangian and usually starts from a two-particle initial state $|i\rangle$. The probability of the system to end up in a final state $|f\rangle$ is given by

$$S_{if} = \langle f | \mathcal{T} e^{i \int \mathcal{L} d^4x} | i \rangle, \quad (2.41)$$

where we assume that the particles only interact via the strong force and $\mathcal{T}$ is the time-ordering operator. S_{if} is only one entry of what is called the scattering matrix S . There is no way of solving this equation analytically. Instead we do an approximate calculation using the concepts of perturbative theory. The assumption is that the strong coupling constant g_C is small enough to expand eq. 2.41 in powers of g_C . The result is given by

$$S_{if} \approx \delta_{if} + i \int \langle f | \mathcal{L}(x) | i \rangle d^4x + \mathcal{O}(g_C^2) \quad (2.42)$$

and is called perturbative QCD. $\mathcal{O}(g_C^2)$ contains all remaining terms that are at least of the order g_C^2 . The degree to which this equation has been solved is given by the order. If only the first term is considered, we speak of leading order (LO), while the inclusion of the second term is dubbed next-to-leading order (NLO). Higher terms follow this scheme exactly, naming the third order the next-to-next-to-leading order (NNLO). A way of working with and interpreting this equation is to use a diagrammatic representation that was introduced by Richard P. Feynman, where particles travel along lines (antiparticles in opposite direction) and interact at a vertex with other particles. A set of Feynman rules that are dictated by the Lagrangian govern the way the diagrams are drawn and assign values to the individual features. We will not go into further detail and refer to dedicated literature (for example [38]) instead. Figure 2.3 gives two examples of Feynman diagrams for the production of Higgs bosons through vector boson fusion (VBF) and gluon-gluon-fusion (ggF).

The ATLAS Experiment at the Large Hadron Collider

In the following chapter, we want to introduce the apparatus used to perform measurements in the context of the Standard Model as described in the previous chapter - and beyond. In 3.1 a description of CERN's Large Hadron Collider is given. The data used in this thesis was collected by the ATLAS detector which is introduced in 3.2 and also includes descriptions of the relevant sub-detectors. Monte Carlo event simulation is addressed in 3.3.

3.1 The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) [40, 41, 42] is a proton-proton particle collider located at CERN near Geneva below the border between France and Switzerland. It was built between 1998 and 2008 in the 27 km long tunnel previously occupied by the Large Electron-Positron (LEP) Collider [43], at a mean depth of 100 meters. With a design center of mass collision energy of 14 TeV it is the most powerful particle accelerator ever built, leaving the previous record of $\sqrt{s} = 1.96$ TeV in $p\bar{p}$ collisions at Fermilab's Tevatron [44] far behind. However, since the begin of operation in November 2009 it has so far collided protons at $\sqrt{s} = 7$ TeV in 2010/2011 and $\sqrt{s} = 8$ TeV in 2012, i.e. roughly half of its design energy with the full intensity to be reached after a longer shutdown period in run II starting in early 2015. The design of the LHC allows not only for pp collisions but also for special heavy ion runs where lead nuclei are injected. Those $Pb+Pb$ or $p+Pb$ collisions at $\sqrt{s} = 5.5$ TeV per nucleon pair are for instance used to study the quark-gluon plasma [45]. At four different sites along the circumference the two beams are allowed to interact. At those collision points the four principal experiments are located: Two multi-purpose detectors ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid), as well as two specialised experiments ALICE (A Large Ion Collider Experiment) and LHCb (LHC-beauty). Furthermore three additional smaller detectors were built. TOTEM (TOTal Elastic and diffractive cross section Measurement) sharing the interaction point with CMS, LHCf (Large Hadron Collider forward) at the ATLAS interaction point and MoEDAL (Monopole and Exotics Detector At the LHC) sharing with LHCb.

LHCs predecessor LEP was a lepton-lepton collider which has the distinct advantage of providing excellent measuring conditions because the initial and final states are very well defined. This allows the performance of high-precision measurements at such machines. Lepton-hadron colliders partly exploit this and are able to probe the structure of the hadron. Hadron-hadron colliders like LHC on the other hand suffer from large uncertainties because neither the initial nor the final state is very well known and can be rather

complex (see chapter 4). But it also allows for higher energies and a higher data-taking rate. Thus, LHC is often labeled a discovery machine. Metaphorically speaking, it can be seen as the sledge hammer that cracks open the proton in search for new physics while a lepton machine is the scalpel that precisely determines its properties. The Experiments at the LHC therefore cover an immense physical spectrum and the agenda includes the investigation of the origin of mass, dark matter, gaining insights into the first moments of the universe after the Big Bang, investigation of the matter-antimatter asymmetry in the universe and more. It might even reveal additional dimensions or uncover expansions to the SM like Supersymmetry [46].

One of the most crucial quantities in collider physics besides the beam energy is the luminosity $\mathcal{L}$, as it links the cross section σ_p of a process P to the interaction rate $\frac{dR}{dt}$.

$$\frac{dR}{dt} = \mathcal{L} \times \sigma_p \quad (3.1)$$

To integrate the luminosity for head-on collisions of two bunches of protons in LHC in the coordinate frame (x,y,s) a few parameters have to be taken into account. The most important ones are the two distribution functions of the beam-bunches ρ_1 and ρ_2 , the number of protons per bunch N_1 and N_2 and the distance to the central collision point s_0 . The overlap integral to compute the luminosity is now given by

$$\mathcal{L} = N_1 N_2 f N_b K \int \int \int \int_{-\infty}^{\infty} \rho_1(x, y, s, -s_0) \rho_2(x, y, s, s_0) dx dy ds ds_0 \quad , \quad (3.2)$$

where N_b is the number of colliding bunches, f is the collision frequency and K is a kinematic factor which is defined for example in [47]. The integrals can be evaluated as is done in [48]. The general result for two unidentical, elliptical Gaussian bunches with widths $\sigma_{x/y,1/2}$ is given by

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{2\pi \sqrt{\sigma_{x,1}^2 + \sigma_{x,2}^2} \sqrt{\sigma_{y,1}^2 + \sigma_{y,2}^2}}. \quad (3.3)$$

For an estimate it is possible to assume two identical, round beam-bunches, i.e. $\sigma_{1,x} = \sigma_{2,x} = \sigma_{1,y} = \sigma_{2,y} = R$ and $N_1 = N_2 = N_p$ [49], which yields

$$\mathcal{L} = \frac{N_p^2 f N_b}{4\pi R^2} \quad . \quad (3.4)$$

LHC aims at an unprecedented luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, having reached a peak luminosity of $7.73 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ in 2012 as can be seen in figure 3.1. The design luminosity for heavy ion collisions is $10^{27} \text{ cm}^{-2} \text{ s}^{-1}$. Furthermore, the design specifies 2808 bunches per proton beam with approximately 10^{11} protons in each bunch. This results in a collision event every 25 ns [50]. The actual collision rate in the first three years of physics programme was reduced to one event every 50 ns. The individual bunches are highly collimated within the beam-pipe exhibiting a length of roughly 1 ns and a transverse spread of roughly $15 \mu\text{m}$ [49]. In the years 2011 and 2012 ATLAS has recorded data corresponding to an integrated luminosity of $L_{2011} = \int \mathcal{L} dt = 5.08 \text{ fb}^{-1}$ and $L_{2012} = \int \mathcal{L} dt = 21.3 \text{ fb}^{-1}$. In figure 3.2 the LHC delivered and ATLAS recorder integrated luminosity over the 2011 and 2012 runs are plotted.

There is however a price that has to be paid for the high luminosity of LHC. In almost all the cases there is not only one interaction of protons per bunch crossing but several. This effect is called pileup (see chapter 4.4). Even in run I the peak number of simultaneous pp interactions reached values of 10 – 40 per event [51]. In the 2012 data

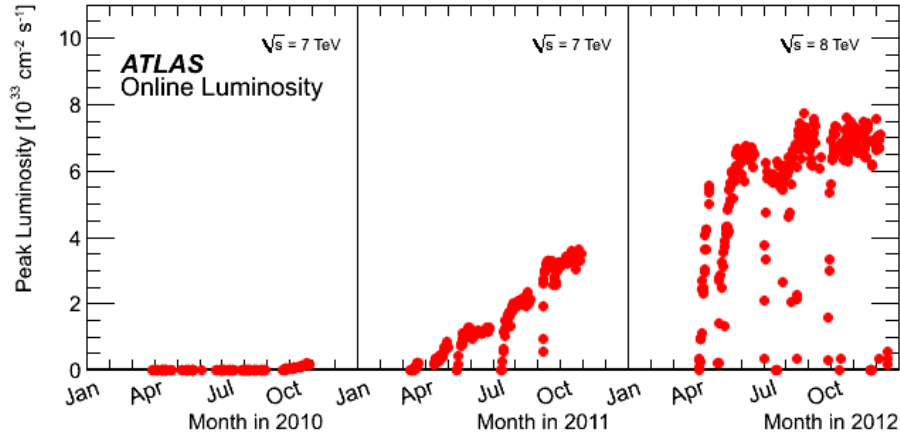


Figure 3.1: Peak luminosity reached over the course of LHC’s first three years of physics programme 2010-2012.

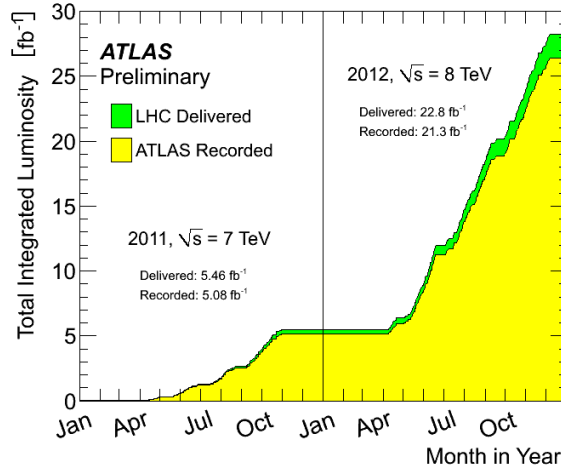


Figure 3.2: LHC delivered and ATLAS recorded integrated luminosity over the 2011 and 2012 runs.

an average number of 20.7 interactions has been measured [52]. In table 3.1 the main properties of the LHC mentioned in the text are summarized.

Before the protons are injected into LHC they undergo a pre-acceleration in several accelerators. Coming from a small bottle of hydrogen gas, protons are stripped of their electrons and accelerated to 50 MeV in LINAC 2. Next in line is the Proton Synchrotron Booster (PSB), reaching an energy of 1.4 GeV before the Proton Synchrotron takes over and accelerates the protons to 25 GeV. The Super Proton Synchrotron (SPS) finally reaches the LHC injection energy of 450 GeV. The whole accelerator complex is sketched in figure 3.3 and more information can be found in [54].

Unlike the Tevatron, where opposite charge particles are accelerated, LHC needs two separated systems for acceleration. In total there are 16 radiofrequency (RF) accelerating cavities, 1232 dipole magnets each 14.3 metres long which are used as bending magnets, and 392 quadrupole magnets which focus the beam. To bend the high energy beams a gigantic magnetic field of up to 8.33 T is produced, which is achieved by cooling down the superconducting magnets to 1.9 K using liquid helium.

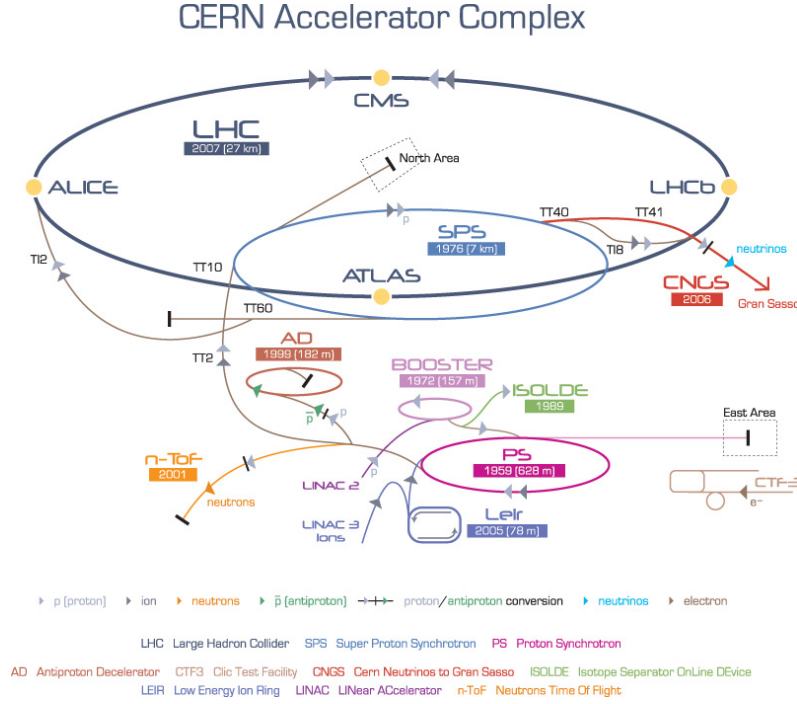


Figure 3.3: Overview over the CERN accelerator complex. Before being injected into LHC protons are pre-accelerated by LINAC 2, BOOSTER, PS and SPS. The figure was taken from [55]

3.2 The ATLAS detector

The ATLAS detector [56, 57] is a general-purpose particle detector at the LHC designed for high precision measurements and searches for new discoveries. The detector is roughly cylindrical with a length of 44 meters and a diameter of 25 meters making ATLAS the biggest particle detector at the LHC. It can be guessed that the whole detector weighs around 7000 tons [58]. A computer-generated model of the complete detector is depicted in figure 3.4. The center of the detector is the proton beam interaction point (IP) with the cylinder axis parallel to the beam pipe around which several layers of concentric detector cylinders are located. The individual sub-detectors will be discussed in the following sections. From the IP outwards particles have to pass the Inner Detector (ID) (section 3.2.1), which tracks charged particles and is itself composed of different types of detectors, then the electromagnetic and hadronic calorimeter (section 3.2.2), that measures energy deposits, and finally the Muon Spectrometer (section 3.2.3). The different sub-detectors are necessary because the different signatures particles leave behind in the sub-detectors

	unit	design value	value for 2012 LHC
proton energy	GeV	7000	4000
peak luminosity	$\text{cm}^{-2}\text{s}^{-1}$	1.0×10^{34}	7.73×10^{33}
protons per bunch		1.15×10^{11}	$1.6 - 1.7 \times 10^{11}$
number of bunches		2808	1374
bunch spacing	ns	24.95	50
average interactions per bunch crossing		23	20.7 (range 10-40)

Table 3.1: Summary of the main properties of the LHC. Listed are the design values and the actual values corresponding to the 2012 data taking runs. The design values were taken from [40, 50], the actual values for the 2012 runs were taken from [52, 53].

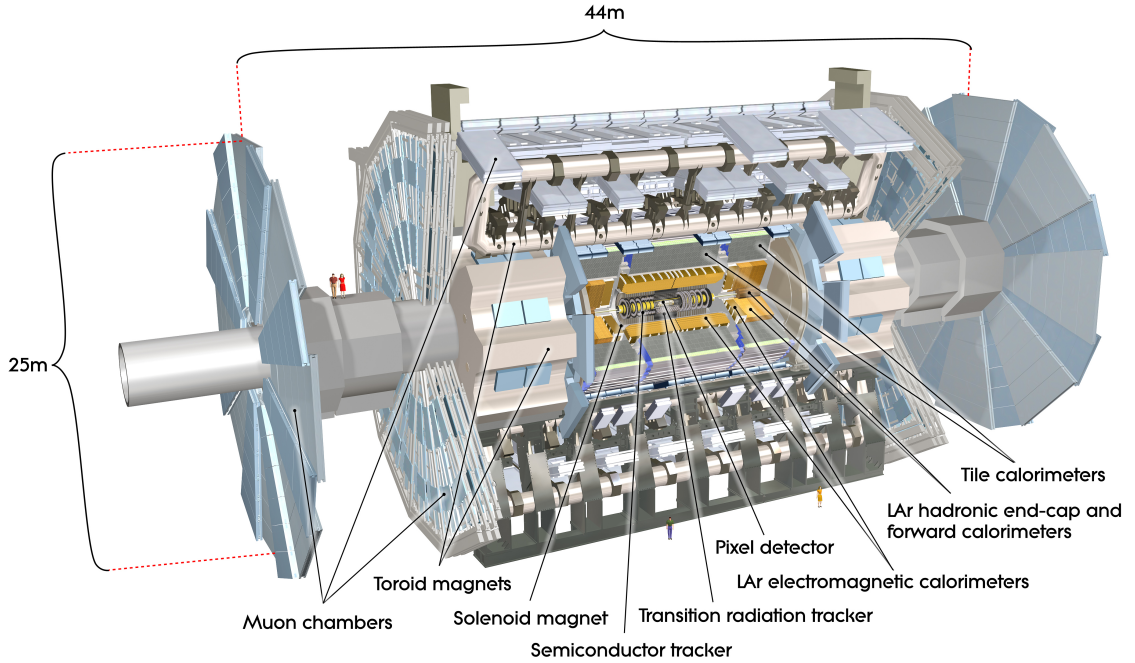


Figure 3.4: A computer-generated model of the ATLAS detector showing all sub-detectors and the magnet systems. The figure was taken from [59]

allow for a very efficient particle identification. Figure 3.5 shows the working principle of the ATLAS detector and we can see how, for example, photons and electrons are distinguished: both leave clusters of energy in the electromagnetic calorimeter, however, since the photon is electrically neutral it is invisible to the trackers, while the electron leaves a reconstructable track leading to the energy deposit. The only known particle that can escape detection in ATLAS is a neutrino. Because of the conservation of energy and momentum neutrinos will manifest themselves in missing energy and their presence is deduced from that.

An essential part of the ATLAS detector are the two vast superconducting magnet systems [60] that provide the large magnetic fields necessary to perform an accurate momentum measurement of charged particles via the Lorentz force induced bending of their tracks. The Inner Detector is addressed by a superconducting solenoid magnet [61] located in a barrel between the ID and the electromagnetic calorimeter. It produces a 2 T magnetic field in the central with a peak of 2.6 T on the windings. The second magnet system is attributed to the Muon Spectrometer and located just outside the calorimeters. It consists out of eight large air-core barrel toroids with a peak field of 3.9 T and two air-cored End-Cap Toroids reaching 4.1 T [62, 63]. This way ATLAS provides the possibility to measure the momentum of muons twice, ensuring a sufficient resolution to a range of several orders of magnitude in muon momenta.

Before discussing the sub-detectors we want to describe the coordinate frame used in ATLAS. The nominal IP is the origin of a right-handed (x, y, z) coordinate system with the z -axis parallel to the beam pipe, the x -axis pointing to the center of the LHC ring and the y -axis aligned towards the surface. Of a more practical use is a spherical coordinate frame in which the plane perpendicular to the beam axis is described by the radial coordinate r and the azimuthal angle ϕ in addition to the polar angle θ . Since θ is

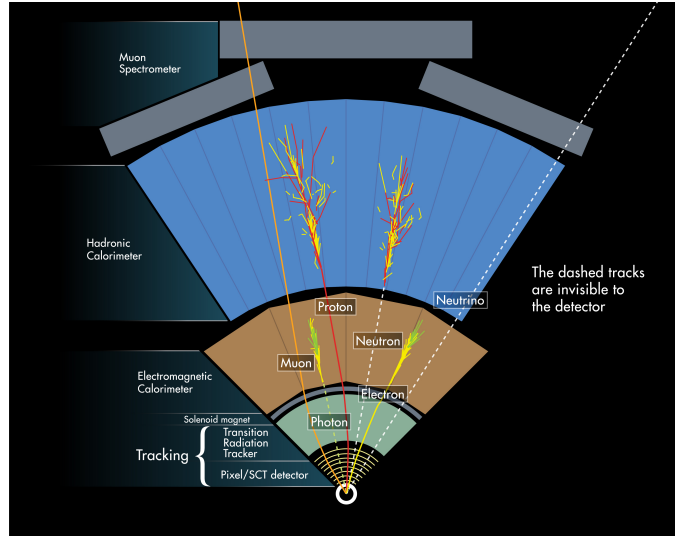


Figure 3.5: Schematic representation of the working principle of the ATLAS detector. The figure was taken from [59]. It should be noted that the figure draws a rather unrealistic picture of the calorimeter system as it shows protons and neutrons depositing their energy completely in the hadronic calorimeter. This is obviously not the case and the fraction of energy deposited in the EM calorimeter by hadrons can be considerable.

not invariant under boosts in z -direction, rapidity, defined as

$$y = \frac{1}{2} \cdot \ln \left(\frac{E + p_L}{E - p_L} \right) \quad (3.5)$$

is used, where E is the energy of a particle and p_L its longitudinal momentum, i.e. momentum along the z -axis. More commonly used is the pseudorapidity

$$\eta = -\ln \left(\tan \left(\frac{\theta}{2} \right) \right) \quad (3.6)$$

which can be understood as the massless limit of the rapidity.

While the momentum along the beam axis carried by the colliding partons is unknown to us, we do know that the total momentum perpendicular to the beam axis has to vanish. The conservation of energy and momentum therefore only works in the transverse plane. As a result we will henceforth only consider transverse quantities indicated by the subscript T . The most important one is the transverse momentum p_T , which is defined as $p_T = p \cdot \sin(\theta)$.

3.2.1 Inner Detector

The purpose of the ATLAS Inner Detector (ID) [64, 65] is the precision tracking of charged particles very close to the IP in the regime $|\eta| < 2.5$. The interaction of particles with the detector material result in a series of discrete points which allow the reconstruction of the particles track and gives information about the type of particle, its electric charge and momentum. A detailed description of the momentum measurement in the ID can be found in Appendix A. The ID has the typical cylindrical shape with a length of about 6.2 m and a radius of 1.15 m, making it the smallest of the three major sub-detector systems. The whole cylinder is submerged in a 2 T magnetic field created by a solenoid magnet sitting just outside of the ID, which causes the bending of charged particles necessary for the determination of their charge and momentum. To allow for optimal particle identification

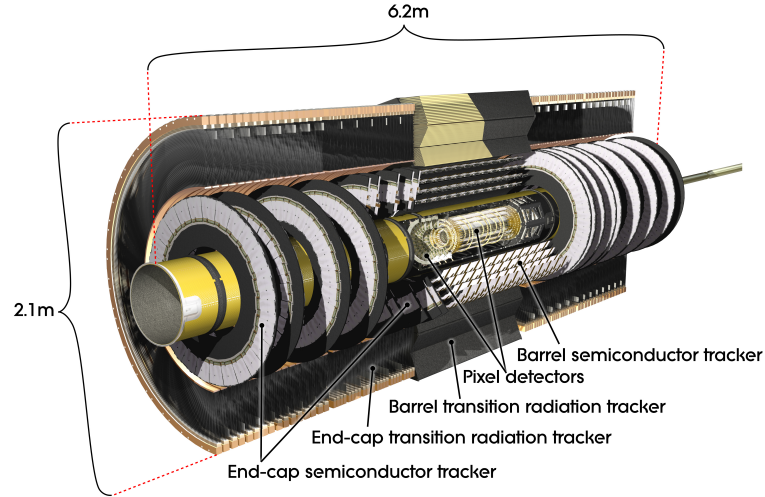


Figure 3.6: Overview of the ATLAS Inner Detector. [59]

the ID consists of three individual sub-detectors that are arranged concentrically around each other and the IP. A schematic view can be seen in figure 3.6.

The innermost sub-detector is the Pixel Detector (PD) [66]. To determine whether a particle originated from the primary collision, or whether it is the product of the decay of a short-lived particle it is necessary to get tracking information from as close as possible to the IP. The three layers of the Pixel detector are located 50.5 mm, 88.5 mm, and 122.5 mm away from the beam pipe. There are furthermore three discs at the end of each barrel to increase the angular coverage. The proximity to the IP exposes the PD to extreme radiation. Therefore, all components have to be radiation hardened to withstand a dose of $5 \cdot 10^5$ Gy [67]. The PD achieves a resolution of $10 \mu\text{m}$ in the transverse ϕ direction and about $115 \mu\text{m}$ in the longitudinal z direction. The 80 million readout channels in the PD alone correspond to about half of the ATLAS readout channel total.

The Semiconductor Tracker (SCT) [68] is located around the PD and consists of silicon micro-strip detectors with dimension $6.36 \times 6.40 \text{ cm}^2$ for high precision tracking. In the barrel region they are arranged into four double-layers at radii between 30.0 cm and 52.0 cm to the IP which allows the position measurement in two dimensions. The barrels are completed by two end-caps consisting of nine discs each on both sides. On average eight precision measurements per track can be expected. Since the area covered by the SCT is much greater than for the PD, it is an essential tool for the momentum measurement and vertex determination. The SCT achieves a spatial resolution of $16 \mu\text{m}$ in the $r - \phi$ plane and $580 \mu\text{m}$ in z direction, where a small angle stereo is used to obtain a measurement in z direction.

The outermost sub-detector of the ID is the Transition Radiation Tracker (TRT) [69]. When relativistic charged particles traverse through the boundary of two materials with different dielectric constants the particle emits electromagnetic radiation called transition radiation (TR). This radiation is detected by the TRT in addition to the signal resulting from the charged particle tracking. The TRT consists of 370000 straws with a length of up to 144 cm and a diameter of 4 mm that are filled with a gas mixture containing Xenon⁶. A tiny gold-plated tungsten wire with a diameter of $30 \mu\text{m}$ in the center of each straw acts as anode for a drift-time measurement. In the barrel region the straws are read out at both ends to minimize the down time. On average 36 hits are detected in the barrel

⁶In fact it is 70%Xe, 27% CO₂ and 3% O₂

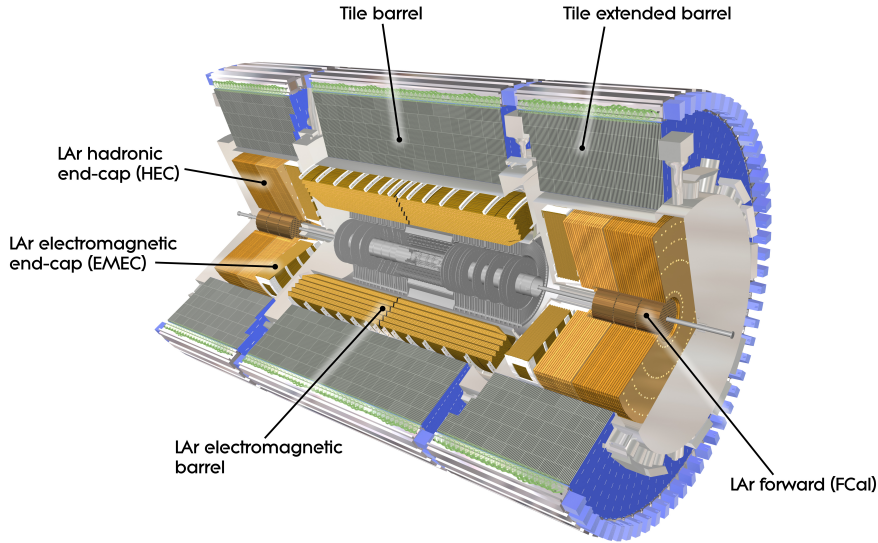


Figure 3.7: A computer generated cutaway diagram of the ATLAS calorimeter systems. The figure was taken from [59].

region. This way the TRT achieves a spatial resolution of $170 \mu\text{m}$. The TR is produced in polypropylene foils that are located between the straws. The signal in the straws due to TR photons is much stronger than the one caused by the passing of charged particles. As a consequence two independent thresholds for TR and passing particles are applied. Furthermore, the magnitude of the TR grows with increasing velocity of the charged particle and for a fixed energy the velocity for lighter particles is higher. Therefore, the TRT allows to distinguish the lightest particles, i.e. electrons and positrons, from heavier ones like pions.

3.2.2 Calorimeters

The cylindrical ATLAS calorimeter system [70, 71, 72] located just outside the solenoid magnet is passed next by the particles. The purpose of the calorimeters is to measure the particles energy by absorbing the particle shower it caused. Since electrons and photons interact electromagnetically with matter, which is fundamentally different from the hadronic interactions of hadrons, two separate calorimeter systems are needed. The electromagnetic calorimeter is nested inside of the hadronic calorimeter. Both are sampling calorimeters meaning that within the calorimeter the active detector material, which is used to measure the energy of the shower, and the passive material, which is responsible for creating the shower, are placed alternatingly. The fine granularity and the wide angular coverage of the calorimeters allow the reconstruction of missing energy from neutrinos or other invisible particles. A cutaway diagram of the calorimeters is depicted in figure 3.7

Electromagnetic Calorimeter

The functional principle of the EM calorimeter is very simple. Electrons or photons entering the passive material of the calorimeter are subject to two processes: Electrons interact with the nuclei in the material and reduce their kinetic energy by emitting a photon. This is called bremsstrahlung. In the presence of a nucleus high energy photons

can undergo pair production processes where a particle and its anti-particle are created. One can see, that the two processes lead to an electromagnetic shower that continues until there is not enough energy left to keep it alive. The total energy of the shower is a measure for the energy of the original incoming photon or electron. It is easy to see why very dense materials such as iron or lead are preferred for the passive layers.

The electromagnetic (EM) Calorimeter consists of a barrel section which covers the pseudorapidity range $|\eta| < 1.475$ and two end-caps (EMEC) on either side extending the coverage to $1.375 < |\eta| < 3.2$. The active material in the EM calorimeter is liquid argon (LAr), while lead is used as absorption material. The electrodes are of an accordion shape as sketched in figure 3.8. The accordion geometry was chosen to ensure full coverage in ϕ , since it naturally prevents the formation of dead towers. Furthermore, it allows a fast read-out at both ends of the electrodes. The foldings are axial in the central region of the calorimeter and radial in the end-cap region. The granularity of the EM calorimeter varies with $|\eta|$. While the resolution in the end-cap region is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, it is much finer in the central region where $\Delta\eta \times \Delta\phi = 0.25 \times 0.25$ is achieved. In this region the calorimeter consists of three channels with varying thicknesses. The innermost layer of the calorimeter has a much finer granularity compared to the outer layers. As is shown in figure 3.8 the inner layer provides an eight times higher resolution in η compared to the remaining layers which enables one to distinguish the photons from a $\pi^0 \rightarrow \gamma\gamma$ decay from prompt photons. The granularity of the outermost third layer is very coarse since the vast part of the energy is absorbed in the second layer and a finer granularity therefore unnecessary. The thickness of the LAr EM barrel is at least $22 X_0$, where X_0 is the radiation length, and at least $24 X_0$ in the EMEC region. The EM calorimeter is housed inside an evacuated cryostat. Since particles have to pass the dead material of the cryostat and the solenoid magnet their basic properties have to be corrected for losses they experienced before entering the calorimeter. For $|\eta| < 1.8$ this is done by the presampler.

The EM calorimeter was designed to gain an energy resolution of

$$\frac{\sigma(E)}{E} = \frac{10\% \sqrt{\text{GeV}}}{\sqrt{E[\text{GeV}]}} \oplus 0.7\% \sqrt{\text{GeV}} \quad , \quad (3.7)$$

where the first term is the stochastic term and the second one is the constant term. Test-beam measurements [58] and tests with cosmic muons [73] have shown that this energy resolution is achievable.

Hadronic Calorimeter

The hadronic calorimeter (H-CAL) is dedicated to particles interacting via the strong force. The particle showers produced in the passive layers are due to strong interactions and by nature much more complicated than the underlying processes of the shower formation in the E-CAL.

The hadronic calorimeter consists of a tile barrel (TileCal) in the central region which covers the range $|\eta| < 1.0$ and has a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ and an extended tile barrel for $0.8 < |\eta| < 1.7$ with the same resolution. They overlap in pseudorapidity so no particle can escape undetected. The active medium are scintillating tiles that are alternated with layers of iron, which acts as the absorbing material. The hadronic end-cap (HEC) covers $1.5 < |\eta| < 3.2$ with a granularity of $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$. As for the EM calorimeter the active material in the HEC is liquid argon with lead as passive material. In the very forward region the liquid argon forward calorimeter (FCal) extends the coverage to $3.1 < |\eta| < 4.9$ with a granularity of $\Delta x \times \Delta y = 3.0 \text{ cm} \times 2.6 \text{ cm}$. However, instead of an accordion shape as one would find in the EM calorimeter, the FCal is composed

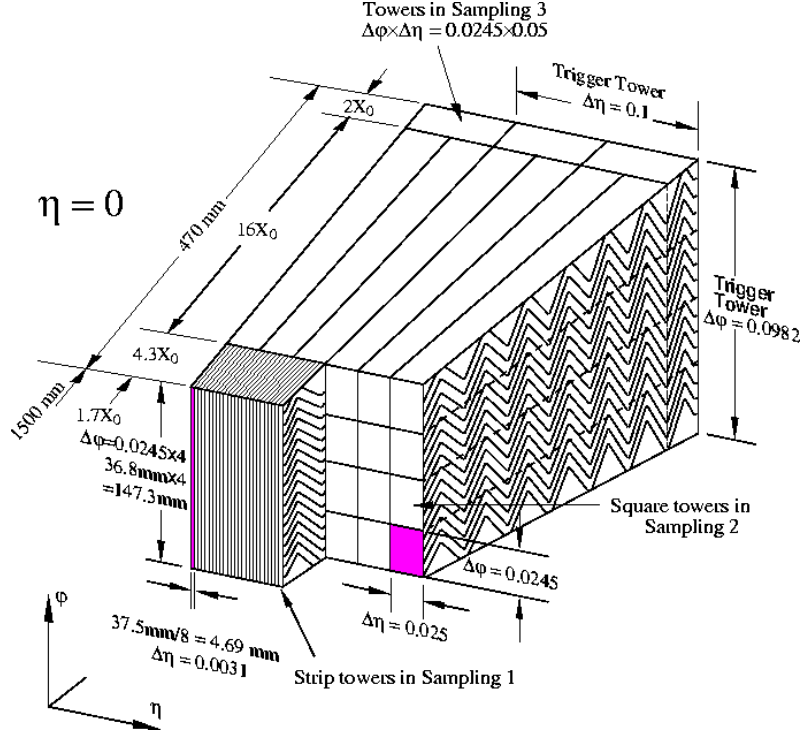


Figure 3.8: Sketch of a module of the ATLAS LAr electromagnetic calorimeter. The figure was taken from [58]

of concentric copper and tungsten tubes. The HEC, FCal and EM calorimeters share the same cryostat. The hadronic calorimeter is a maximum of 11λ thick, where λ is the interaction length. This is necessary to prevent punch-through into the muon system by containing all shower end-products, only allowing muons (and invisible particles) to pass. Pion test-beam measurements [58] reveal an energy resolution of the tile calorimeter of

$$\frac{\sigma(E)}{E} = \frac{(56.4 \pm 0.4) \% \sqrt{\text{GeV}}}{\sqrt{E[\text{GeV}]}} \oplus (5.5 \pm 0.1) \% \sqrt{\text{GeV}} \quad (3.8)$$

and of

$$\frac{\sigma(E)}{E} = \frac{(70.6 \pm 1.5) \% \sqrt{\text{GeV}}}{\sqrt{E[\text{GeV}]}} \oplus (5.8 \pm 0.2) \% \sqrt{\text{GeV}} \quad (3.9)$$

for the HEC [74].

Since the calorimeter system is the most essential sub-detector for jet physics we have listed the main design properties of the ATLAS calorimeter system in a table which can be found in Appendix B. The Appendix also includes diagrams showing the amount of material in front and in the different calorimeter systems. The information and plots were taken from [58].

3.2.3 Muon Spectrometer

The outermost sub-detector system is the ATLAS Muon Spectrometer (MS) [75]. Its huge end-cap wheels are responsible for the characteristic shape of the ATLAS detector. The MS serves a double purpose. On the one side muon tracks are measured with high precision, on the other side it provides a trigger source for a multitude of interesting physics studies involving muons. To accomplish this a variety of different technologies

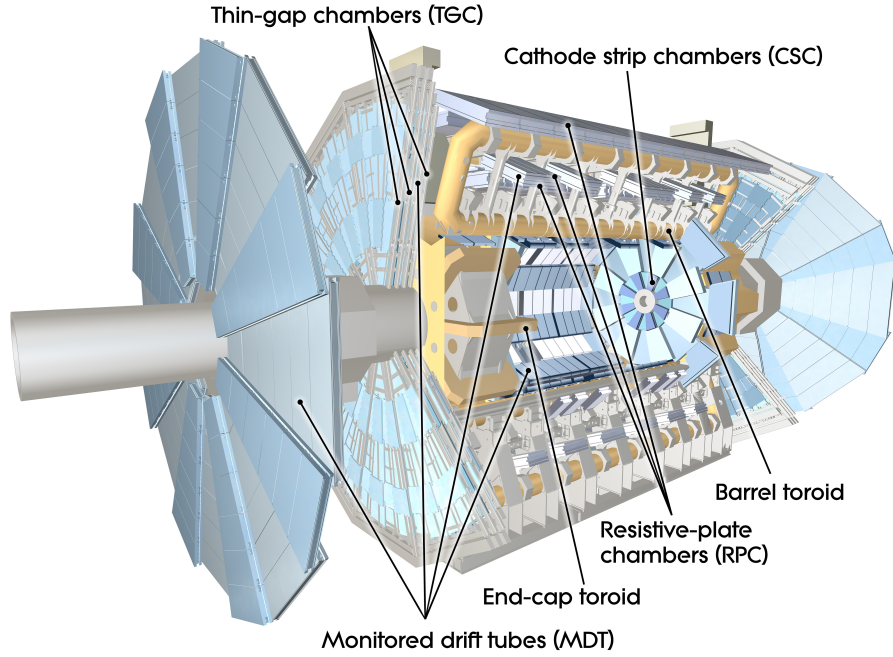


Figure 3.9: Layout of the ATLAS Muon Spectrometer and its subsystems. The figure was taken from [59].

are used in the MS. The high precision tracking measurements are done by Monitored Drift Tubes (MDT) made of gas filled⁷ aluminium tubes with a $50\,\mu\text{m}$ wire in its center. The MDT's are complemented in the innermost plane of the end-caps by Cathode Strip Chambers (CSC), a multiwire proportional chamber. The CSC's have a better time resolution and are therefore better suited for the forward region, i.e. $2.0 < |\eta| < 2.7$ where the background is expected to be largest. The magnetic field of the air-core barrel toroid and the end-cap toroid causes the muon tracks to bend and the momentum can be determined as described in Appendix A.

The second purpose of the MS, the triggering, is addressed by Resistive Plate Chambers (RPC) in the barrel region $|\eta| < 1.05$ and Thin Gap Chambers (TGC) in the end-cap region $1.05 < |\eta| < 2.4$. The RPC is a gas filled detector with two parallel resistive plates at a distance of 2 mm. The readout is done by capacitive coupling to readout strips on both detector plates. The operating principle of the TGC is very similar to the CSC, however the distance between anode and anode is smaller than the distance between anode and cathode. Both systems have to provide a muon momentum measurement in the plane perpendicular to the measurements of the tracking chambers. A schematic overview of the ATLAS Muon Spectrometer can be seen in figure 3.9.

3.2.4 Trigger

The ATLAS-detector uses a highly efficient three level trigger system to reduce the interaction rate of roughly 1 GHz at design luminosity to a more manageable 200 Hz [76]. Otherwise, event storage and offline computing would be unable to comply with the high data taking rates [77].

Level 1 trigger (L1) is a hardware based trigger system, which has direct access to a subset of data from muon and calorimeter system. The speed of the decision is limited

⁷91% Ar, 4% N₂ and 5% CH₄

by delays resulting from cables that connect the trigger logic to the detector. In the end it takes roughly $2\mu\text{s}$ for the trigger to reach its decision. In the meantime all the event data has to be kept in buffers until the trigger decision is known.

Those events which passed the level 1 trigger are then subject to the next layer trigger (L2). L2 takes Regions of Interest (RoIs) that were identified by the level-1 trigger into account and uses the full granularity from all sub-detectors to reach a decision whether the event is worth keeping. The level-2 trigger is software based and many events are evaluated simultaneously on computing farms. The decision process takes in the order of 10 ms. Obviously, all the collision data is still kept in readout buffers. After the level-2 trigger the data taking rate has been reduced to approximately 1 kHz.

A final step in the trigger event selection is the Event Filter (EF). The EF reconstructs the events similarly to the offline analysis described later in this thesis. The processing time for a single event in the event filter is in the order of 1 s. Complex decision algorithms reduce the event rate to the afore mentioned 200 Hz.

Some triggers use a technique called trigger prescaling which essentially seeks to lower the data taking rate further by selecting events on a random basis with a certain probability.

3.3 Event simulation

In chapter 2 we have described the possibilities of extensive sets of particle interactions with complicated final states. This, in combination with the very complex detector responses to these final states makes an analytical treatment of particle collisions virtually impossible. Furthermore, imagine we try to observe electroweak gauge boson by identifying two jets originating from its decay. We can take two jets in close proximity and calculate their invariant mass but we have no reason to assume that these two particular jets originate from the boson decay. In fact, the expected background to this process from pure QCD events is orders of magnitude larger than the signal (see chapter 6). Only looking at the plain spectrum is not enough to make an observation or even discovery with the ATLAS detector. Instead particle physicists simulate a number of events containing the process of interest and expected background processes and compare the simulation to the data gathered in the experiment. These Monte Carlo (MC) samples are produced by elaborate software packages called event generators on huge computing-farms. An extensive amount of work is put into the development of event generators as they have to give an extremely accurate picture of real collisions. There exists a wide variety of different generators which use different models and approximations and yield slightly different results. The result of an analysis thus always depends to a degree on the generator used [78]. A review of some of the most common event generators along with descriptions about their working principle can be found in [79].

The simulated samples have to be overlaid by additional minimum-bias samples to mimic the effects from pileup or underlying event (see chapter 4.4). The number of overlaid interactions is dictated by the number of simultaneous collisions in data.

As a last step the influences from the detector have to be accounted for. The generated events are passed through a full simulation of the detector. This means that the detector response to all final states has to be modelled and regions of dead material or inefficiencies have to be considered. Through the complexity of this whole process the average CPU time needed for the simulation of a single QCD event is around 20 minutes [80]. A typical physics analysis requires millions of MC events.

Jet Physics at the LHC

One of the most frequently appearing objects at particle physics experiments like ATLAS are jets, highly collimated sprays of hadronic particles originating from the hard interaction of partons in the proton-proton collision. Because of their occurrence in many processes of interest, the study of jets is of major importance for the discovery of new physics. In the following chapter, we want to give an introduction to jet and collider physics in general. This includes the origins of jets in 4.1. We will then discuss how a jet is manifested as a physical object in the detector with advanced algorithms in 4.2. 4.3 gives a brief introduction to the concept of parton distribution functions and a discussion of contamination to jets is given in 4.4. The chapter closes with an in-detail description of several jet grooming algorithms in 4.5.

4.1 Origin of Jets

Coloured objects like quarks or gluons (collectively called partons [81]) are produced through a multitude of different processes at particle colliders. When they appear as a result of a scattering process they develop bremsstrahlung cascades through soft and collinear splitting. The result are sprays of collimated gluons and quark-antiquark pairs. This is called a parton shower. Eventually they will hadronise at the colour confinement radius which is of the order $1.0 \cdot 10^{-15} \text{ m}$ [82] and form hadron sprays that are observed in the detectors and are called jets. Figure 4.1 illustrates the formation of a jet through parton shower and hadronisation. The figure interprets the non-perturbative hadronisation process within the string model [83, 84, 85, 86] which sees the softer gluons as field lines that through the gluon self interaction form strings⁸. The strings then break and fragment into quark-antiquark pairs in its intense colour field. Gluons that were produced in the parton shower cause kinks on the string which influences for instance the angular distribution of the hadrons. Besides the particle jets originating from two decay quarks, it especially predicts the production of hadronic sprays in the space between the jets by the string itself [87]. This can be observed in data. The string model is not the only model to describe hadronisation but it was used in the production of parts of the Monte Carlo samples used in this study. Alternative models are the cluster model [88, 89, 90] and independent fragmentation model [91].

The configuration of the initial high energetic quarks and gluons is very well reflected in the characteristics of the observed jet, for example its energy or the angular distribution of the jet constituents. Thus, studying a jet gives hints on the properties of the underlying parton and allows to deduce characteristics such as their flavour or spin.

⁸As opposed to electric or magnetic field lines that are far spread since there is no self coupling term for the photon in the Lagrangian.

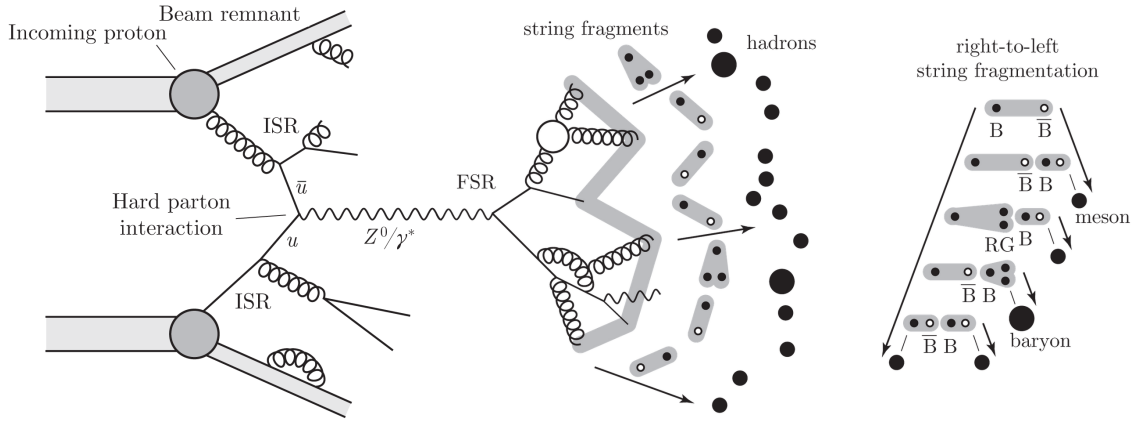


Figure 4.1: Illustration of the formation of a jet originating from the hard interaction. Parton showers and hadronisation are depicted. The process forms one or more jets that can be observed in a detector. The abbreviations ISR and FSR stand for initial and final state radiation. The figure was reproduced from [92].

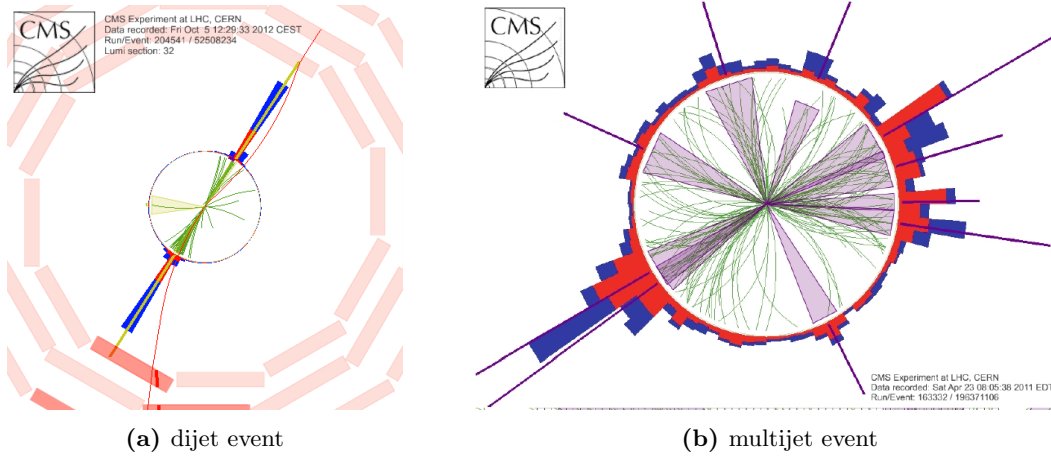


Figure 4.2: (a) shows a dijet event in which both jets can be identified easily, while (b) shows a busy multijet event. Both events were recorded with the CMS detector. The figures were taken from [95] and [96].

The parton shower is not limited to final state partons but can occur for initial state partons as well. The two possibilities are distinguished by referring to the first case as Initial State Radiation (ISR) and Final State Radiation (FSR) for the latter case.

The first evidence for the observation of a quark jet was found at the SPEAR e^+e^- collider where the anisotropy of the hadron distribution was shown [93, 94].

4.2 Jet reconstruction algorithms

In figure 4.2 two events containing jets recorded by the CMS detector are shown. The red and blue bars show the energy deposited in the calorimeter system. Using the eye only we have no problem of distinguishing two back-to-back jets in figure 4.2(a). There can be very little doubt about this conclusion. The situation is different if we look at a more busy event like in figure 4.2(b). Identifying the right number and direction of jets is virtually impossible. Especially on the right side of the detector where we have a hard time deciding whether we see one very large jet or several smaller ones. It becomes clear,

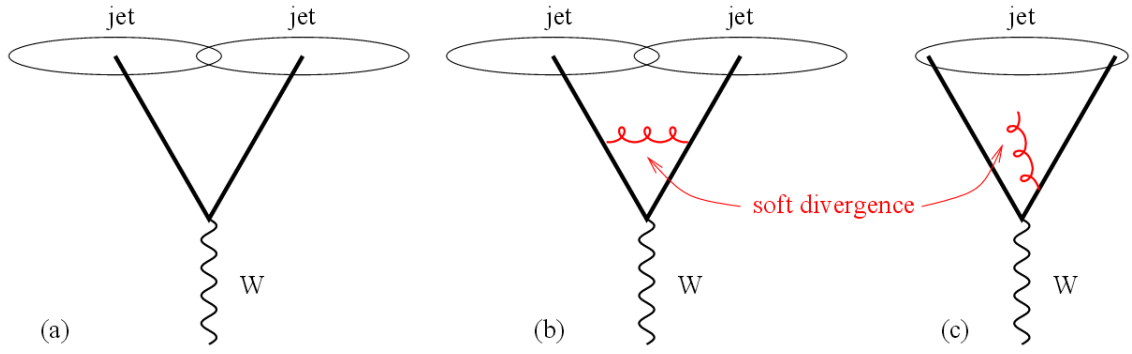


Figure 4.3: Illustration of infra-red and collinear safety. Here, an IRC unsafe algorithm is used to reconstruct the decay jets of a $W^\pm$ boson. The hard partons in all three cases are identical and they only differ in the appearance of a soft gluon. Yet, the final result in figure (b) and (c) is very different as the process results in only one jet in (c) while two were found previously. The figure was taken from [98].

that in order to compute trustworthy cross sections we need a good definition of a jet that can be computed fully automated for billions of events.

The task is fulfilled by dedicated jet reconstruction algorithms. The demands for such an algorithm are tremendous: The jet is required to be stable towards contaminating effects from underlying event, hadronisation or detector noise which should furthermore have a small impact on the jet itself. Pileup will surround hard jets as well and the algorithms should be as insensitive as possible to its effects. In addition, the computation needs to be very fast, because high transverse momentum jets are used for the triggering of events which may contain several jets. Here a fast decision is crucial since in the meantime all event information has to be stored in buffers. Lastly, it is worth bearing in mind that a jet is not a physical object in a sense that it exists outside the definition of the jet algorithm used. It is therefore essential to have a definition that allows to compare the experimental results with the theoretical predictions. This means the jet algorithm is demanded to be infra-red and collinear (IRC) safe. The principle of an IRC unsafe algorithm is illustrated in figure 4.3 where an arbitrary emission of a soft particle or the collinear splitting of a parton leads to a completely different reconstructed jet landscape. If the number of reconstructed jets is affected by the presence of an infinitely soft gluon then perturbation theory fails and the cross sections become infinite [97].

Historically the first jet algorithm were cone algorithms inspired by the first algorithm developed for e^+e^- collisions in the 1970's by Stermann and Weinberg [99]. Algorithms of this type tried to maximize the deposited energy within a cone of fixed radius. In ATLAS neither this nor a seeded version of this algorithm is used as they are either too time-consuming or not IRC safe [97, 100]. Instead three different forms of sequential algorithms⁹ [103] have been regarded as very effective for the use in ATLAS studies: The k_t [104, 105], anti- k_t [106] and Cambridge/Aachen [107, 108] algorithms.

The working principle is as follows¹⁰: For a specific cluster i two quantities d_{iB} and d_{ij} are defined which give the distance of the cluster i to the beam B and to another

⁹The first sequential IRC safe algorithm was introduced by the JADE collaboration for e^+e^- collisions [101, 102].

¹⁰This definition is obviously only valid for hadron machines, as no partons can be emitted from the beam in e^+e^- colliders. A leptonic version of the k_t algorithm was developed in [104]

cluster j . The distance measures are given by

$$d_{iB} = p_{T,i}^{2k} \quad (4.1)$$

$$d_{ij} = \min(d_{iB}, d_{jB}) \frac{\Delta R_{ij}^2}{R^2} \quad (4.2)$$

where p_T is the clusters associated transverse momentum to the beam, $\Delta R_{ij}^2 = \Delta\eta^2 + \Delta\phi^2$. Here $k = 1$ for the k_t , $k = -1$ for the anti- k_t and $k = 0$ for the Cambridge/Aachen algorithm. The parameter R is called the radius-parameter of the jet. The smaller R is chosen in the jet reconstruction the finer the scale of the jet identification. This means that the number of jets in an event depends strongly on this parameter. But there is no reason to be alarmed by that. In fact, it offers the possibility to look at events on different scales and use new techniques to gain an advantage in an analysis. We will continue this thought in chapter 4.5.

Next, the following steps are performed in a loop

- Scan through all distances d_{iB} and d_{ij} . Find the minimal distance and label it d_{min}
- In case d_{min} is a d_{iB} , the cluster is declared a jet and withdrawn from the list.
- On the other hand, if d_{min} is from the d_{ij} list, the clusters i and j are merged.
- Repeat until the list of clusters is empty.

There are several possibilities on how the merging of two clusters is done. The most common is to simply add the clusters four-vectors. This recombination scheme is called the E-scheme [109].

According to this procedure a soft gluon can be labelled a jet which automatically means that the sequential algorithm is IRC unsafe. This is solved by additionally applying a cut on the transverse momentum of the jet.

The main difference between the k_t and anti- k_t algorithm is the order of merging. Soft clusters at close proximity are clustered first in the k_t algorithm. This gives a good correspondence between theoretical predictions for the jet and the actual reconstructed jet. It was the preferred choice at e^+e^- machines [110]. However, at LHC, where the contaminating effects are much larger, the shapes of the jet can become very asymmetric which might cause problems with the jet calibration. Therefore the anti- k_t jet is used more commonly. Here the hardest components are clustered first, which means that the hardest jets are of nicely defined circular shapes.

An additional jet algorithm currently used in studies is called SISCone [100]. However, since it is of no relevance to this study we refer the reader to e.g. [98] for further details. It is also recommended as a review on this chapter.

4.3 Parton Distribution Functions

In the processes we are interested in, the initial state always consists of quarks or gluons, for which we draw Feynman diagrams and compute the cross section for a given final state. However, at LHC we do not collide pure quarks or gluons which would be impossible, but protons. A proton is a composite particle consisting of three valence quarks two of which are up quarks, the other is a down quark giving the proton an overall electric charge of +1. Beside the valence quarks the protons consists of an undetermined number of gluons and quark-antiquark pairs (sea quarks). In order to be nevertheless able to compute experimentally verifiable cross sections we need to introduce parton distribution

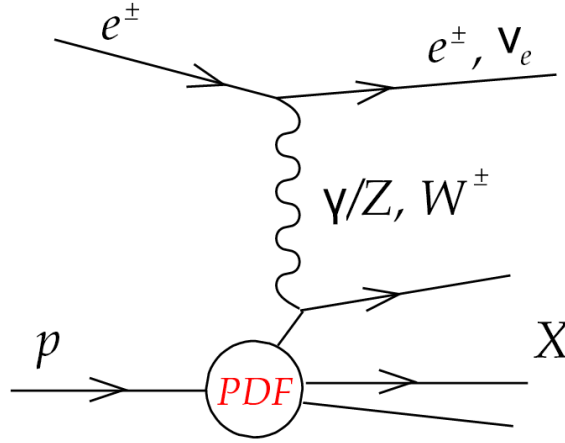


Figure 4.4: Diagram of a neutral current (charged current) deep inelastic scattering experiment where electrons (positrons) are collided with protons. A boson interacts with one of the proton partons which results in the final state X . The figure was taken from [123].

functions (PDF). They were first conceived by Richard P. Feynman to explain Bjorken scaling [111] in deep inelastic scattering experiments [112]. A set of PDFs is specific for one type of hadron, but since this study deals with proton-proton collisions we are only interested in the distributions for the proton. When probing a proton at scale Q^2 , i.e. the energy scale of the hard process, the PDF $f_i(x, Q^2)$ is defined as the probability of finding a parton i that carries the fraction x of the proton momentum. The parton content of a proton is not predicted by QCD and one can only derive evolution equations called the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) [113, 114, 115, 116] equations that can be used to fit experimental observations to determine the shape of the PDFs. This is done by collaborations like MRST/MRSW [117], NNPDF [118] or CTEQ [119]. Figure E.1 shows an example PDF from MSTW 2008 at NLO in which $f_i(x, Q^2)$ was plotted as a function of x .

The cross section of a real proton-proton collision process where the partons i and j from the incoming protons A and B produce a final state X can be calculated by computing

$$\sigma_{A,B \rightarrow X} = \sum_{i,j} \int dx_1 f_a(x_1, \mu^2) \int dx_2 f_b(x_2, \mu^2) \hat{\sigma}_{ij \rightarrow X} \quad (4.3)$$

where $\hat{\sigma}$ is the cross section of the fundamental parton scattering and μ the factorization scale. For further reading [120] is recommended.

Deep inelastic scattering (DIS) experiments like HERA [121], TEVATRON or fixed targets [122] have been the driving data source behind the measurement of parton distribution functions. In DIS experiments the structure of a proton is probed by scattering electrons (neutral current [NC]) or positrons (charged current [CC]) of a proton. The principle is drawn in figure 4.4.

The analogous density functions for final state parton distributions are called fragmentation functions $D_i(z, Q^2)$ [124, 125]. They give the probability of a hadron to carry the momentum fraction z of the parton i momentum. There are different fragmentation functions for all hadrons.

4.4 Contamination by underlying event and pileup

A reconstructed jet at hadron colliders is subject to contaminations from underlying event (UE) and pileup (PU). The concept of both is very simple and straightforward, however the analytical description is much more complicated. This is because UE and PU are soft processes and as a result cannot be described by perturbative theory. Therefore, we will not venture into detail and be content with a description of the principle.

4.4.1 Underlying event

Up until now we assumed that the true nature of a proton-proton collision is the interaction of one parton from the first proton and another parton from the second proton. The probability of such an interaction to occur between two specific partons is given by the PDF. This is what was called the hard scattering. However, this model is only half the story. It is of course possible that more than just one parton from each proton is going to interact. These multiple parton interactions usually result in the radiation of soft particles with low to moderate energies. This is because the probability for a second hard interaction within the same proton-proton collision is diminishingly small, which means that in the detector the clear signatures of the hard interaction are spoiled by soft energy deposits throughout the detector space and makes the jet reconstruction itself as well as the determination of its mass much more challenging. [126] shows that at LHC energy deposits around ≈ 2.5 GeV per unit area in the $y - \phi$ plane from UE contamination are to be expected. Since most studies rely on Monte Carlo simulations of events a well-performing model of the UE is necessary. The last decade has seen a great deal of work in this respect (e.g. [127, 128, 129, 130]) and many MC generators incorporate UE in their simulations¹¹.

4.4.2 Pileup

At LHC the beam does not consist of individual protons but large bunches of protons. As a result there is a large probability that instead of only one proton-proton interaction two or more are happening simultaneously. Just as for UE, the probability for two hard scattering processes in the same event is minuscule. The soft radiation originating from pileup will also result in energy deposits all over the detector. A single pileup interaction would be less severe for a measurement than one from UE. This is straightforward to see: A collision that results in a hard scattering is necessarily a central collision and UE contamination is therefore likely to also appear in the central detector region. On the other hand most pileup interactions are not head-on collisions and only the outermost partons are able to interact. Unfortunately the data used in this work exhibits an average number of $\langle \mu \rangle = 20.7$ interactions per bunch-crossing [59]. Going to higher and higher center-of-mass energies the number of simultaneous interactions is going to increase further. At high-luminosity at LHC the contaminations from PU are expected to reach 10 – 20 GeV per unit area [126]. Hence, the removal of pileup is one of the chief-concerns in many studies and a lot of work is put into the development of advanced algorithms and techniques to rid a measurement of the contaminations from UE and PU.

4.5 Jet grooming algorithms

The significantly higher energies at the LHC result in the production of heavily boosted particles, e.g. $W^\pm$ [132], Z^0 [133] or top quarks [134, 135] that can themselves decay

¹¹Including the ones used in this study [131].

hadronically and result in several boosted jets. If the system was boosted beyond a certain limit the standard jet reconstruction methods start to fail due to high collimation of the jets that brings them closer together than the radius parameter $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ of the reconstruction algorithm. This can be illustrated by a quick example: The separation of the decay products of a hadronically decaying heavy particle can be roughly described by

$$\Delta R \cong \frac{2m}{p_T} \quad (4.4)$$

where m and p_T are mass and transverse momentum of the parent particle, respectively [133, 136]. We will verify this relation in chapter 5.1. The separation of the decay products of a hadronically decaying Z^0 boson with a mass of $m_Z = 91.2 \text{ GeV}$ and $p_T = 360 \text{ GeV}$ is expected to be roughly 0.5. Obviously this can not be resolved with a jet-clustering algorithm of radius parameter $R \gtrsim 0.5$. In that case the result will be the reconstruction of a single massive jet with a large radius parameter. We will call such a large-radius jet that exhibits substructure originating from a particle decay of a massive particle a fat jet [137]. Unfortunately a very similar situation can occur when jets from pileup overlay the hard scattering jets. However, the properties of a jet that concentrates all decay products of a massive particle and those of a jet originating from a gluon or light quark but exhibiting the same p_T are different enough to allow a discrimination of the two cases. While a large-radius jet resulting from the decay of a boosted $W^\pm$, Z^0 or top quark exhibits a hard substructure, namely the collimated jets from the hadronic decay, this is not the case for a high p_T jet originating from a light quark or a gluon, where we expect to see a dense jet-core that is surrounded by soft radiation from underlying event and pileup [138]. In other words: in the first case we expect to identify subjets, i.e. clusters within the reconstructed jet where there is a higher concentration of energy, while in the latter case there are no subjets meaning exactly one region of energy concentration [132]. To utilize those differences one tries to remove the effects of unassociated radiation from underlying event and pile-up from the target jets. This procedure is called jet grooming [139].¹²

Currently there are three different jet grooming algorithms well established in ATLAS: mass-drop filtering [141], trimming [142] and pruning [143]. In the following we will introduce the three algorithms one after another.

4.5.1 Mass-drop Filtering

The mass-drop filtering procedure was initially introduced as a tool to utilize the $H \rightarrow b\bar{b}$ search channel at the LHC [141]. It tries to isolate symmetric subjets that exhibit a significantly smaller mass when compared with their parent jet. To achieve this the algorithm works in two stages. At first, energy clusters within the jet are identified using mass-drop and symmetry criteria. In stage two the reclustered jets are filtered. The mass-drop filtering procedure was considered and employed for a variety of different studies, like the search for a SM Higgs boson [144, 145], a minimal supersymmetric SM (MSSM) Higgs

¹²It should be noted that some studies define ‘jet grooming’ as a procedure that always returns an output jet if an input jet is given. This is contrasted with the definition of a tagger where an input jet does not necessarily lead to an output jet. According to this definition the mass-drop filtering procedure is not a jet grooming, but a tagging procedure [140]. On the other hand the 2010 Boost report [139] defines jet grooming as “elimination of uncorrelated UE/PU radiation from a target jet”. This means a grooming algorithm leaves an input jet unchanged if there is no contamination from showering, UE, pileup or hadronisation. Following this definition the mass-drop filter is the only jet grooming algorithm. In this thesis we will ignore this confusion and adopt the less pedantic definitions from above and treat mass-drop filtering, trimming and pruning as jet grooming algorithms.

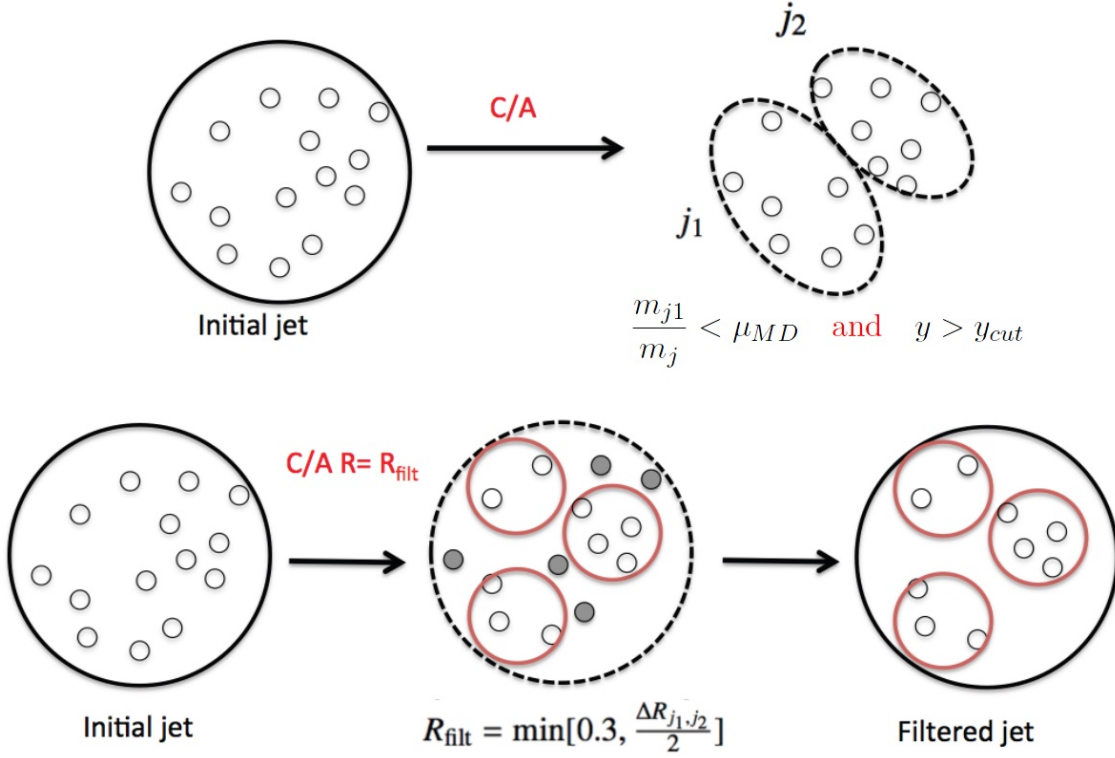


Figure 4.5: Shown are the two stages of the mass-drop filtering procedure. In the upper part the mass-drop and symmetry constraints are sketched, while the lower part depicts the filtering procedure. The figure was taken from [133]

boson [137, 146] or even the search for weakly interacting massive particles (WIMP), to understand the nature of dark matter [147].

Mass-drop and symmetry constraints

The longitudinally invariant Cambridge/Aachen (C/A) algorithm is used since it provides beneficial characteristics for the mass-drop filter. Its sequence of clustering combines the smaller-angle pairs in the beginning resulting in an angular hierarchical structure terminating with the clustering of the widest combinations [133]. The procedure starts off by undoing this final step of the C/A clustering and therefore separating the jet into two subjets j_1 and j_2 with masses m_{j_1} and m_{j_2} . The subjets are labelled in such a way as to ensure that $m_{j_1} > m_{j_2}$ is fulfilled. Next, the mass-drop condition

$$\frac{m_{j_1}}{m_j} < \mu_{MD} \quad (4.5)$$

is applied. The ratio of the mass of the harder subjet and the mass of the parent jet m_j should be significantly smaller than an arbitrary value μ_{MD} given as one of two parameters to the mass-drop filter algorithm. In addition the splitting of the parent jet should not be too asymmetric. Defining the distance between j_1 and j_2 as $\Delta R_{j_1,j_2}$ the constraint can be expressed by requiring

$$\frac{\min[(p_T^{j_2})^2, (p_T^{j_1})^2]}{m_j^2} \times \Delta R_{j_1,j_2}^2 > y_{cut} \quad , \quad (4.6)$$

where $p_T^{j_1}$ and $p_T^{j_2}$ are the corresponding transverse momenta to j_1 and j_2 . The parameter y_{cut} is the second parameter given to the algorithm. In [133] the optimal values for

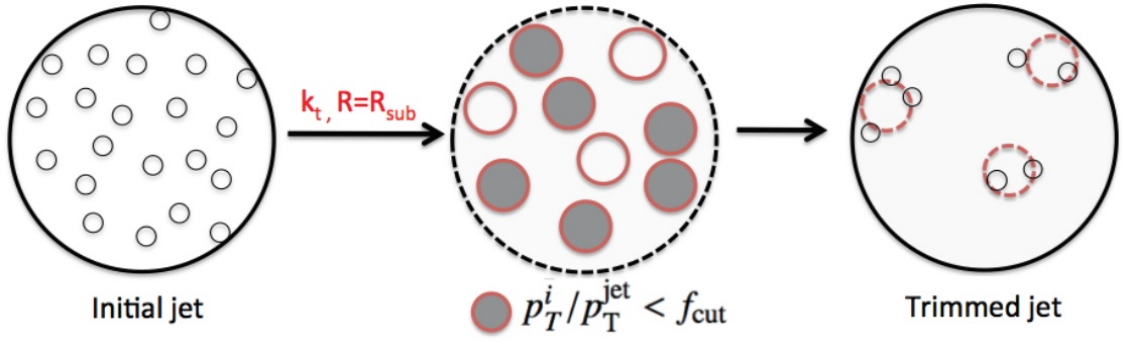


Figure 4.6: Illustration of the jet trimming algorithm. The figure was taken from [133].

$H \rightarrow b\bar{b}$ are given by $y_{cut} = 0.09$ and $\mu_{MD} = 0.67$. However, a subsequent study shows that the more restrictive choice of $\mu_{MD} = 0.25$ leaves the effectiveness of the algorithm in suppressing the background almost unchanged [148]. In case a jet fails to satisfy either criteria the jet is rejected.

Filtering

If the jet j passes the previous stage it is reclustered using the C/A algorithm with a radius parameter $R_{filt} < \Delta R_{j1,j2}$. A value of $R_{filt} = \min[0.3, \Delta R_{j1,j2}/2]$ has been determined as very effective. The filtering eliminates all constituents except the hardest ones and the jet is reclustered into three subjets. This allows for one more subjet caused by (gluon-)radiation in addition to the two-body decay, while filtering out most of the UE/PU contamination. The whole mass-drop filtering procedure is depicted in figure 4.5. Its effectiveness regarding the suppression of pileup effects can be seen for example in [149] or chapter 5. It shows that the dependency of the reconstructed mean jet mass on the number of reconstructed primary vertices has almost disappeared. This dependency can normally be observed in the presence of pileup [136, 149].

4.5.2 Jet Trimming

The jet trimming algorithm first proposed in [142] exploits the fact that contamination to a jet originating from PU, UE or ISR is comparatively soft compared to the products of a hard scattering process. The algorithm selects constituents by comparing their transverse momentum to the one of the reconstructed jet. For this purpose an already reconstructed jet with radius parameter R_{jet} is split up and reclustered into subjets using a k_t algorithm with radius parameter $R_{sub} < R_{jet}$. The k_t algorithm is used, because in opposition to other clustering methods it clusters from softer to harder, which will result in more balanced subjets, where the energy is shared between the subjets. Next, the algorithm loops over all subjets and invokes the softness criterion: Only if the subjets transverse momentum p_T^i fulfills

$$p_T^i > f_{cut} \cdot \Lambda_{hard} \quad , \quad (4.7)$$

where Λ_{hard} is a parameter governed by the kinematics of the hard scattering process and f_{cut} is a parameter of the algorithm, the constituents are kept. Otherwise the subjet is discarded. The remaining subjets form the trimmed jet. This procedure is depicted in figure 4.6.

Naturally, jets lose mass through the removal of pileup contributions during the trimming procedure. For jets from boosted object decays the fraction of lost mass is much

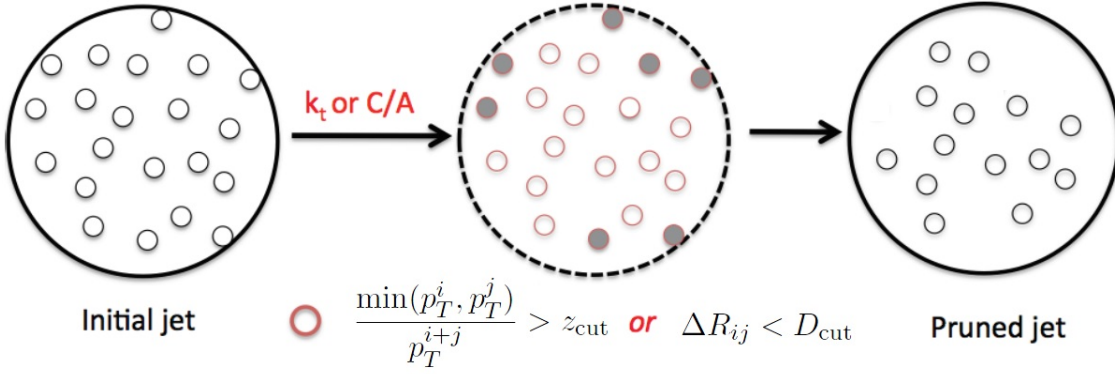


Figure 4.7: Illustration of the jet pruning algorithm. The figure was taken from [133].

smaller than for jets originating from light partons. In these cases the jets can lose between a third and one half of their mass [133]. It is straightforward to see that the fraction of lost mass increases with the number of proton-proton interactions per bunch-crossing. In knowledge about the number of reconstructed vertices the accidental removal of final-state radiation can be compensated by a simple offset of the jet mass as it is demonstrated in [150]. The optimized parameters for the algorithm given in [142] are $f_{\text{cut}} = 0.03$ and $R_{\text{sub}} = 0.2$, however in [133] the values $f_{\text{cut}} = 0.01, 0.05$ and $R_{\text{sub}} = 0.3$ are also explored.

The jet trimming technique was employed in different studies, for example in [151].

4.5.3 Jet Pruning

The jet pruning procedure [143] is very similar to jet trimming by using the relative transverse momentum as a selection criterion. But while the trimming algorithm takes an already reconstructed jet as input and then probes into its substructure by the reconstruction of subjets, the pruning procedures input is a set of protojets and it acts during the reconstruction of a jet. For this we have to remember the working principle of a sequential jet clustering algorithm as described in 4.2. The algorithm browses through the list of all distances d_{iB} and d_{ij} to find the minimal one. If it is a d_{ij} the two protojets i and j are merged and the list of distances is updated. On the other hand if the minimal distance is found to be a d_{iB} the protojet is promoted to a jet and removed from the list. This procedure is continued until the list is empty [143, 152]. The jet pruning algorithm acts when a minimal distance ρ_{ij} is found. Instead of simply merging i and j it is required that the protojets with transverse momenta p_T^i and p_T^j fulfill either

$$\frac{\min(p_T^i, p_T^j)}{p_T^{i+j}} > z_{\text{cut}} \quad (4.8)$$

where z_{cut} is the first parameter given to the algorithm, or

$$\Delta R_{ij} < D_{\text{cut}} \quad (4.9)$$

with D_{cut} the second parameter of the algorithm. Only if the protojets fail to meet both requirements the algorithm puts a veto on the merging and the softer protojet is discarded. The algorithm then continues as described before. The final result is the pruned jet. In figure 4.7 the jet pruning procedure is depicted.

While the motivation for requiring eq. 4.8 is fairly simple and we refer to section 4.5.2 for a detailed explanation, the motivation for eq. 4.9 is to put a veto on wide-angle radiation to only keep the hardest constituents of the jet. It is of little concern to the

pruning algorithm if a k_t or C/A algorithm is used for the jet reconstruction. The choice of the input parameters is investigated in a number of studies [133, 143, 152] and usually leads to optimal values around $z_{\text{cut}} = 0.1$ and $D_{\text{cut}} = m^j/p_T^j$.

The jet pruning technique was used in several searches particularly at CMS, e.g. [153, 154].

Part II

Analysis

Initial studies on Monte Carlo

In 4.5 we have introduced various jet grooming algorithms. For this study the mass-drop filtering technique is utilized. Prior to starting the work on data it is necessary to validate the procedure in a study on MC. It is also a good opportunity to explore the true capabilities of the algorithm and show its power in the recovering of sharp mass peaks even if the unfiltered jet is contaminated heavily by pileup and underlying event. 5.1 through 5.3 summarize this work. The chapter closes with a study on jet charge in 5.4 which can be seen as preparatory work for a subsequent search in the $W^\pm W^\mp$ channel.

5.1 Validating the mass-drop filtering procedure

The large center-of-mass energy at the LHC makes it possible to produce large amounts of heavily boosted electroweak gauge bosons. In chapter 4.5 we described how this leads to a collimation of the resulting hadronic decay products and a failure to resolve two individual jets using standard jet reconstruction algorithms like Anti- k_t with a radius parameter of $R = 0.4$. This claim can be verified by using MC samples that contain hadronically decaying $W^\pm$ and Z^0 bosons.

Two different sets of MC samples are used in this chapter. The first are W+jets and Z+jets samples produced with the HERWIG [155] generator. They are described in detail in chapter 6.2 as they are the signal sample in the main part of this work. The second set are samples of $W' \rightarrow WZ \rightarrow q\bar{q}q\bar{q}$ where the mass of the W' was set to $m_{W'} = 3000$ GeV. The sample was produced using the Pythia 8 [156, 157] generator with the MSTW 2008 [158] leading-order PDF set and the ATLAS underlying event tune AU2 [159].

The W' sample is chosen to plot the distance in $\Delta R = \sqrt{\Delta\eta + \Delta\phi}$ between the two quarks resulting from the decay of the V_h as a function of the transverse momentum of the parent boson. Here we have defined V_h as a combined term for hadronically decaying $W^\pm$ or Z^0 bosons. Figure 5.1 shows the expected behaviour of an increasing collimation with higher momenta. It is not difficult to conclude that jet clustering algorithms, like anti- k_t with an R-parameter of $R = 0.4$ or larger, are going to have a hard time resolving two jets when the p_T of the parent boson exceeds 300 – 400 GeV. Large R-parameter jets, on the other hand, will form a cone that encloses both jets. Before we also claimed that a rough estimation of the expected distance between the decay products is given by

$$\Delta R_{q\bar{q}} \approx \frac{2m}{p_T} \quad . \quad (5.1)$$

We therefore include the law as a dashed line in the figure. It does indeed give a good estimate.

We have now seen that it is essential to consider large R-parameter jets when dealing with boosted $W^\pm$ and Z^0 bosons. We use a Cambridge/Aachen jet algorithm with $R = 1.2$

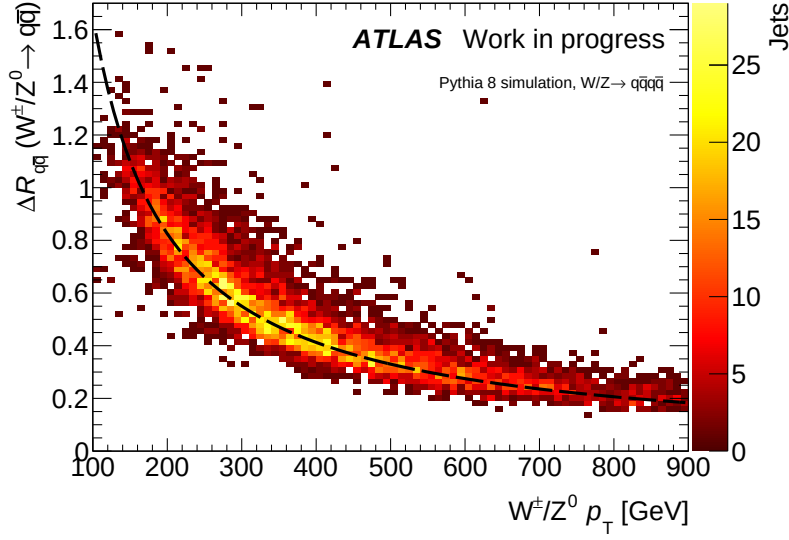


Figure 5.1: Distance in ΔR between the hadronic decay products of boosted $W^\pm$ and Z^0 bosons. The collimation of the quarks increases with higher boson p_T . The dashed line shows the estimation function from eq. 5.1.

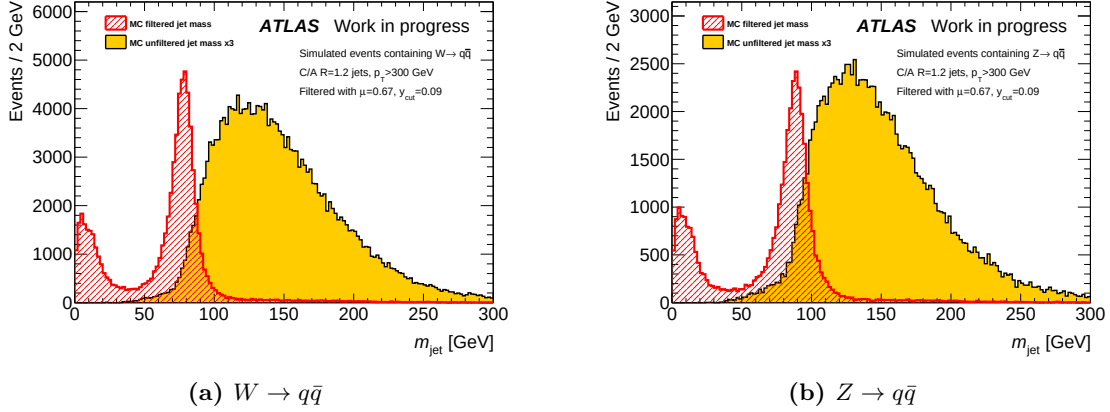


Figure 5.2: Comparison of the unfiltered and filtered jet mass spectrum in MC events for the hadronic decay of (a) $W^\pm$ and (b) Z^0 boson. The unfiltered jet mass spectrum is scaled by a factor of 3 to allow for a better comparison.

to reconstruct fat jets. Unfortunately, the large active area of such a jet makes it very susceptible to influences from pileup, underlying event and detector noise. Hence, the peaks expected at the masses of the gauge bosons are shifted to much higher masses and smeared beyond recognition. This can be seen in figure 5.2 where the distribution of the reconstructed mass of the fat jet has been plotted. Here, the W/Z +jets samples were used and the distributions are drawn separately for $W^\pm$ and Z^0 . It is clear that the unfiltered fat jet is not ideal for application in searches for new physics as it is unable to reproduce resonances in the mass spectrum that correspond to the particles mass. It can also be assumed that the identification of a wide peak is much more difficult than it would be for a sharp peak. In studies with small signal to background ratios and only few signal events in total (like the present study described in chapter 6) the reconstruction of a peak would be virtually impossible.

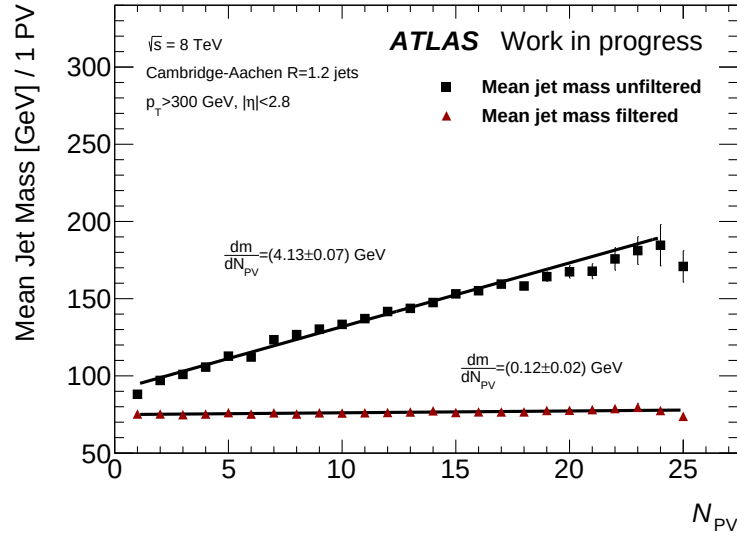


Figure 5.3: Mean reconstructed jet masses for unfiltered and filtered jets as a function of the number of reconstructed good vertices. The data are fitted with linear functions, the slope of which are displayed next to the function.

We now apply the mass-drop filter and choose the parameters $\mu_{MD} = 0.67$ and $y_{cut} = 0.09$. This way the algorithm yields a list of subjets for every fat jet. We reconstruct only those filtered jets for which the unfiltered jet has a transverse momentum of $p_T > 300 \text{ GeV}$ and is composed of two or three individual subjets. The resulting distributions are plotted in figure 5.2 as well. We see at once that the resulting peak is much more narrow than the unfiltered peak and shifted back to the mass of the $W^\pm$ or Z^0 boson.

5.2 Suppression of pileup

The mass-drop filter very effectively removes the contamination of the jet through pileup by reducing the jets active area. A way of demonstrating this is to plot the mean jet mass as a function of the number of reconstructed good vertices both for the filtered and unfiltered jets. A good vertex is defined as a vertex from which at least five tracks with $p_T^{\text{track}} > 500 \text{ MeV}$ originate. Figure 5.3 shows the resulting data points that were fitted with linear functions. As expected the unfiltered jet mass largely depends on the number of vertices as the increasing number of simultaneous interactions deposit more and more energy in the region of the calorimeter from which the fat jet is clustered. The linear fit function has a slope of $\frac{dm_{\text{jet}}}{dN_{vxp}} = (4.13 \pm 0.07) \text{ GeV}$ meaning that every additional vertex increases the reconstructed jet mass by approximately 4 GeV. This is enormous given that the mean number of simultaneous interactions is expected to climb up to 45 in Run II and could reach values above 80 after 2018 [160]. The filtering keeps only the hardest constituents of the jet and discards the remaining softer ones. Therefore, the mean mass of the filtered jets shows only a very weak dependence on the number of vertices where the slope of the linear fit is given by $\frac{dm_{\text{jet}}}{dN_{vxp}} = (0.12 \pm 0.02) \text{ GeV}$.

Mass-drop filtering is a very efficient way of getting rid of pileup.

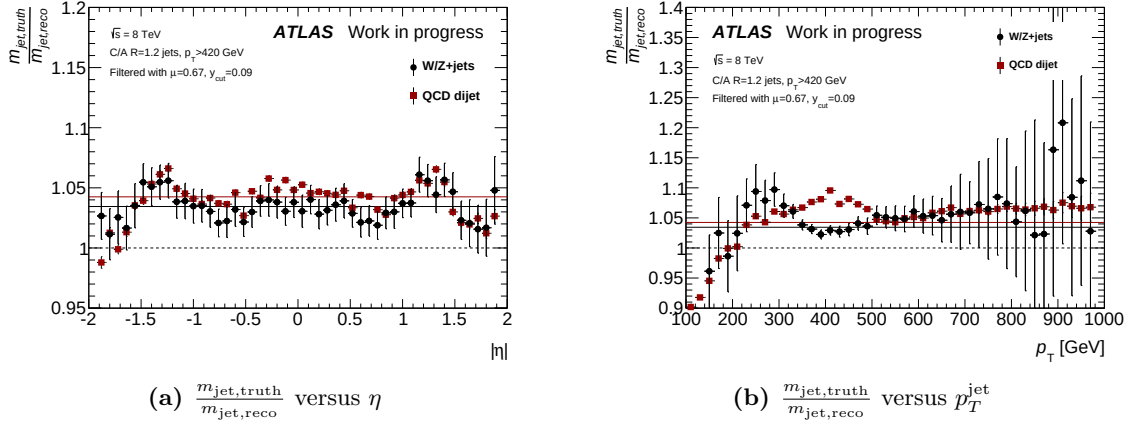


Figure 5.4: The ratio between the jet mass in MC truth and the reconstructed jet mass are plotted as a function of the reconstructed jet (a) η and (b) p_T . As expected the reconstructed jet mass is persistently smaller than the corresponding truth value. The solid lines indicate the global average values which are around 3.4% for W/Z+jets and 4.2% for QCD dijet samples.

5.3 Validation of the filtered mass

We want to return to figure 5.2 for a moment. The mass-drop filtering procedure was praised for restoring the sharp gauge boson mass peak at the right position. However, this is not entirely true. Looking closely, we see that the true mass peaks exhibit masses slightly less than the literature values. We get around 78 GeV for the $W^\pm$ boson and 88 GeV for the Z^0 boson. This comes quite naturally from the filtering. The removal of active area in the fat jet also affects some constituents originating from the gauge boson decay. Usually outliers around the hardest subjets. As a conclusion, the reconstructed mass is slightly less than the actual mass of the jet.

We try to quantify this effect by computing the ratio between the reconstructed jet mass and the truth jet mass in MC. The ratio is then plotted as a function of the reconstructed jets pseudorapidity in figure 5.4(a) and transverse momentum in figure 5.4(b) for two different sets of MC samples both of which are described in detail in chapter 6.2. Only jets with filtered masses in the region of interest, i.e. $40.0 \text{ GeV} < m_{\text{jet}} < 160.0 \text{ GeV}$, are considered if it is possible to match them to a truth jet with a mass of at least 5.0 GeV. As we expected, the reconstructed jet mass is persistently smaller than the truth value. The ratio shows only a weak dependence on η except for slight bumps in the outer regions that could very well originate from the transition between barrel and end-cap calorimeters and is not uncommon. There is, however, some dependency on the p_T of the jet. The higher the momentum of the jet the more severe the negative effects of the filtering get. Since both MC samples were produced mainly involving jets at $p_T > 350 \text{ GeV}$ the discrepancies around this value might be due to the inadequateness of the MC samples. The solid lines in the plots indicate the average values for each individual distribution. From the W/Z+jets sample we judge that the reconstructed jet mass is approximately 3.4% smaller than the mass obtained from truth. This is in accordance to our previous observations. The QCD dijet MC sample yields slightly larger averages and results in an underestimation of the jet mass of 4.2%.

Finally, rather than plotting the ratio between the reconstructed and the true jet mass we want to plot them against each other. This was done in figure 5.5(a). The profile of the scatter plot is drawn in black. In the region of the $W^\pm$ and Z^0 mass a linear behaviour of the algorithm is observed. The linear fit function drawn in white returns the characteristic

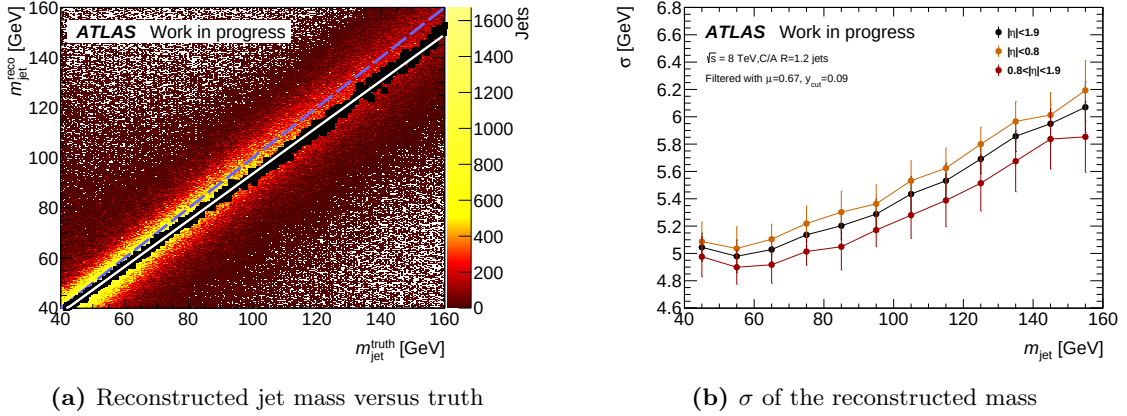


Figure 5.5: In (a) the reconstructed jet mass is drawn versus the jet mass from truth. The profile of the plot is drawn in black and fitted with a white line. The dashed blue line indicates unity. In (b) the widths of the distributions are plotted as a function of the truth jet mass for three different regions in η .

curve

$$m_{\text{jet}}^{\text{reco}} = (1.143 \pm 0.076) + (0.922 \pm 0.001) \cdot m_{\text{jet}}^{\text{truth}} \quad (5.2)$$

for $40.0 \text{ GeV} < m_{\text{jet}}^{\text{truth}} < 160.0 \text{ GeV}$. Looking at y-slices of the plot one can compute the width of the reconstructed jet mass for a given truth jet mass. This was done in figure 5.5(b) for three different regions in η . The width for the jets that are most central in the detector is slightly larger than for the more forward jets in the barrel of the calorimeter. This is because the energy of a more forward jet at the same p_T as a central one is necessarily higher and the calorimeter resolution accordingly better (compare to eq. 3.7 and 3.8). In the mass region of electroweak gauge bosons we expect in average a detector peak width of around 5.2 GeV.

5.4 Jet charge

We can now imagine a study in which we look at diboson events, where at least one of the bosons decays hadronically and is reconstructed using a mass-drop filtering algorithm. It might be necessary to distinguish $W^\pm Z^0$ from $W^\pm W^\mp$ events. Since the masses of $W^\pm$ and Z^0 boson are very close to each other, it is near to impossible to distinguish the two on basis of their mass. A further idea is the exploitation of the differences in electric charge. A naive approach would be to sum over all the tracks within a jet according to

$$Q_{\text{jet}} = \sum_{i \in \text{jet}} q_i \quad , \quad (5.3)$$

where i iterates the tracks. However, due to the difficult measurement conditions at LHC this approach is futile, especially when dealing with fat jets. The influence of experimental noise can be reduced by choosing a weighted definition of the jet charge according to

$$Q_{\text{jet}} = \frac{1}{(p_T)^\kappa} \sum_{i \in \text{jet}} q_i \cdot (p_T^i)^\kappa \quad (5.4)$$

where p_T and p_T^i are the transverse momenta of the jet and the i^{th} track and κ is a weighting constant [161, 162, 163]. The p_T -weighted jet charge observable was first suggested

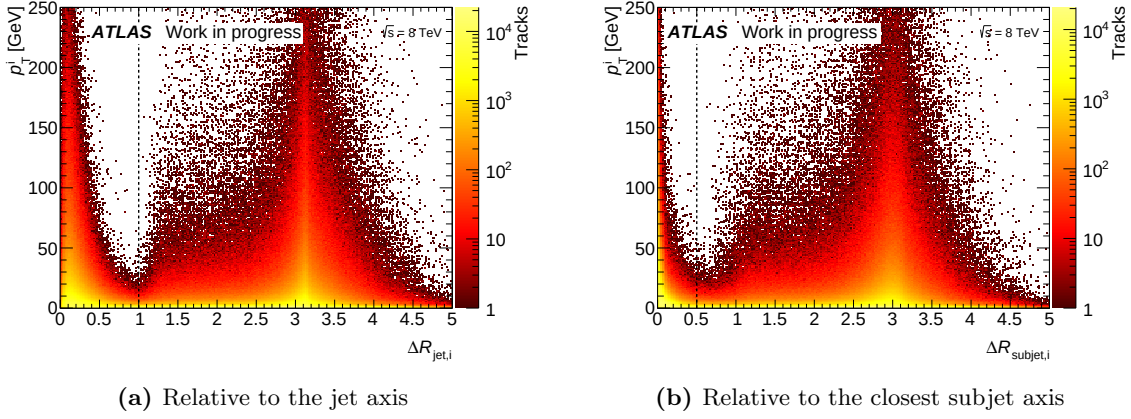


Figure 5.6: Track p_T plotted as a function of the distance in pseudorapidity-azimuth space to (a) the jet axis of the fat jet (b) the subjet axis of the closest leading or sub-leading subjet.

by R.D. Field and R.P. Feynman [164] in 1978 to measure the charge of quarks but other weighting schemes, like pseudorapidity weighting [165], have been used as well.

In the following we will compute the jet charge two different ways using the $W + jets$ and $Z + jets$ MC sample. The first is as suggested by eq. 5.4 and takes into account all tracks within a cone of certain radius R_{jet} around the jet axis (jet charge). Secondly, the mass-drop filter allows us to repeat the procedure but constructing the cone with different R_{subjet} around the two leading subjets instead of the parent jet (subjet charge). The jet charge is then given by the sum of the two subjet charges. In order to find the cone radii R_{jet} and R_{subjet} the track p_T is drawn as a function of the tracks distance to either the jet axis or the closer subjet axis in figure 5.6.

The figure very nicely reveals the recoil-activity on the other side of the detector. As expected, figure 5.6(a) shows that the jet axis is in between the two leading subjets. This can be concluded since the tracks with the largest transverse momentum are located at distances $\Delta R_{jet,i} > 0$ in contrast to figure 5.6(b) where the highest momenta are found parallel to the subjet axis. The dip in activity between the fat jet and the recoiljet is used for the definition of the cone. We choose $R_{jet} = 1.0$ and $R_{subjet} = 0.5$. The values are marked by dashed lines. This arbitrary choice gives by no means the optimized values. However, it serves as a proof of concept.

Next, we try to find the κ for which the separation between the distributions for W^- , W^+ and Z^0 gets maximal. It can be assumed that the reconstructed charge is of symmetric Gaussian shape. It is therefore possible to define the discriminator

$$\xi_{ij} = \frac{|\mu_i - \mu_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}} \quad (5.5)$$

which assumes greater values the smaller the overlap between two distributions i and j with mean μ and standard deviation σ . A track is only considered if it has a $p_T^i > 500$ MeV, at least one hit in the pixel detector, at least six in the silicon microstrip tracker, transverse and longitudinal impact parameter no more than 2.5 mm and 2.0 mm and a maximal $\chi^2/\text{dof} = 2.5$ in the track fit. In figure 5.7 ξ is calculated for all three cases (separation between the distributions of W^+ and Z^0 , W^- and Z^0 and W^+ and W^-) and plotted as a function of κ . Comparing the jet charge with the subjet charge one can see that the subjet charge performs better in all three cases. A maximum of separation is reached at $\kappa = 0.3$ for the subjet charge indicated by the vertical dashed line. The

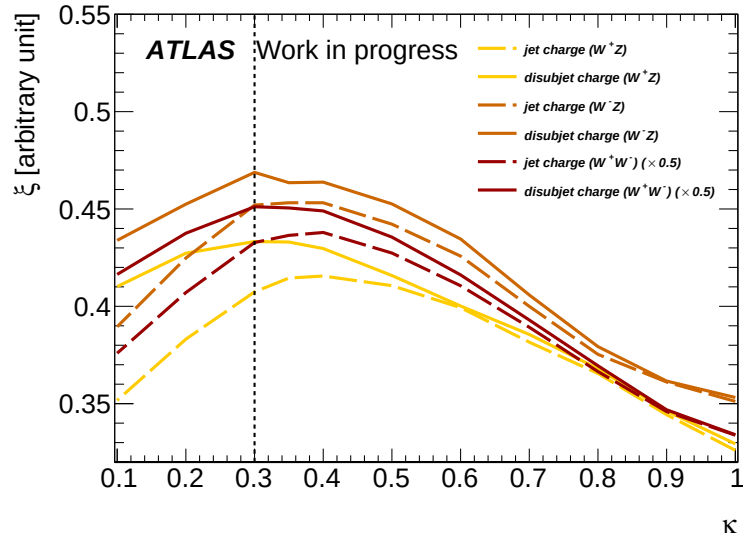


Figure 5.7: Separation power ξ of the jet charge (dashed lines) and subjet charge (solid lines) as a function of the κ parameter. Separation between W^+ and W^- is scaled by a factor of 0.5 for convenience. The vertical dashed line represents the optimal value of $\kappa = 0.3$ for the subjet charge.

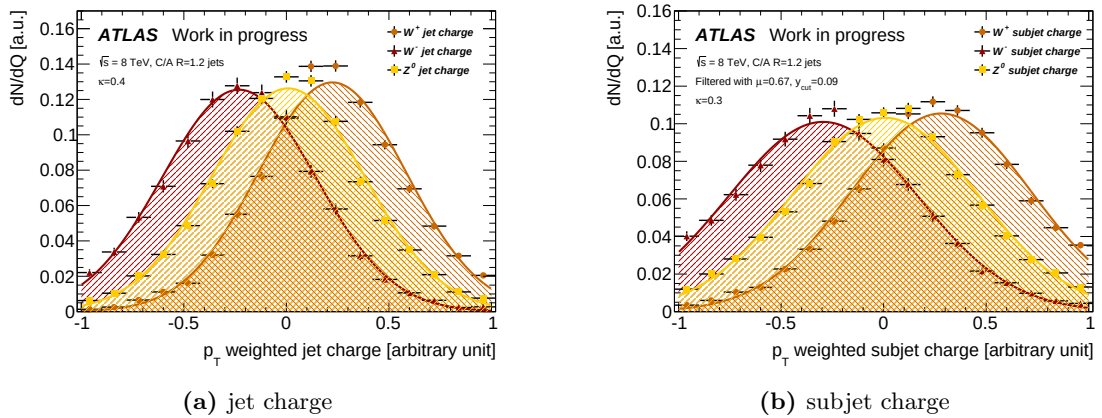


Figure 5.8: Distributions of the p_T weighted (a) jet charge and (b) subjet charge for W^+ , W^- and Z^0 bosons.

jet charge demands a larger κ around 0.4. This result confirms previous observations in which values in the range $0.2 < \kappa < 0.5$ were found to be optimal [163, 166].

Interestingly there seems to be some asymmetry between the W^+ and W^- distributions which lead to $\xi_{W^-Z^0} > \xi_{W^+Z^0}$. In average $|Q_{\text{jet}}|$ is larger for W^- when compared to W^+ . Whether this is due to the charge asymmetry of the collider itself or detector effects is unclear. The effect is not significant on generator level.

Figure 5.8(a) shows the jet charge distributions for $\kappa = 0.4$, the subjet charge with the optimized parameter $\kappa = 0.3$ is plotted in 5.8(b).

The approach chosen here suffers from the forced choice of R_{jet} and R_{subjet} . It is possible to avoid this by introducing an additional ΔR -weight. Inspired by [167] the

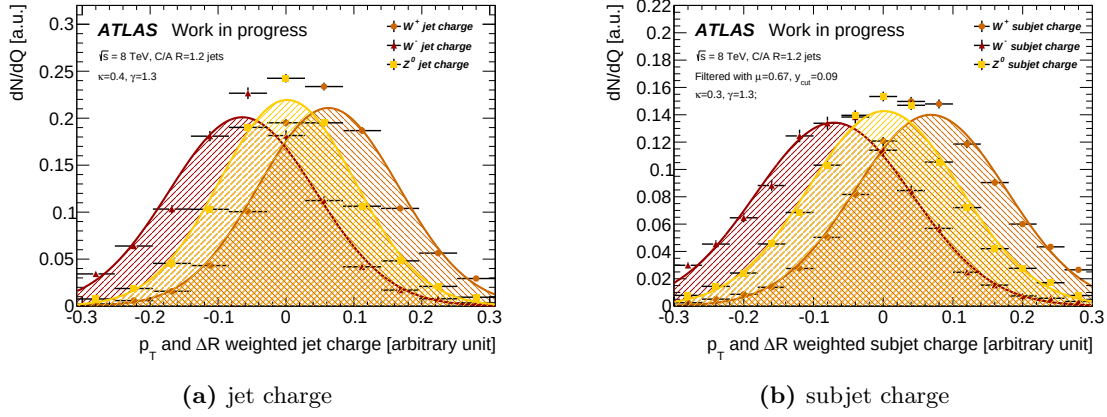


Figure 5.9: Distributions of the p_T and ΔR weighted (a) jet charge and (b) subjet charge for W^+ , W^- and Z^0 bosons.

charge is calculated via

$$Q_{\text{jet}} = \frac{\sum_i q_i \cdot (p_T^i)^\kappa \cdot \exp\left(-\left(\frac{\Delta R_i}{R^*}\right)^\gamma\right)}{\sum_i (p_T^i)^\kappa \cdot \exp\left(-\left(\frac{\Delta R_i}{R^*}\right)^\gamma\right)} \quad (5.6)$$

where ΔR_i is the distance of track i to the jet axis or to the closest subjet axis and R^* is a characteristic distance for which we will use $R^* = R_{\text{jet}} = 1.0$ and $R^* = R_{\text{subjet}} = 0.5$. Note that the sums now run over all i tracks in the event, not only the ones attributed to the jet. The new weighting parameter γ can be treated just as κ . Best separation is achieved with values around $\gamma = 1.3$. Figure 5.9 shows the jet and subjet charge distributions for the p_T and ΔR weighted calculation. The separation achieved by this method is slightly larger than in the p_T -weighted case. Since the jet charge is of small relevance to the actual study carried out in this thesis we will leave an optimization of all parameters κ , γ and R^* for others to do.

Boosted $W^\pm$ and Z^0 bosons from fat jets

6.1 Introduction

$W^\pm$ and Z^0 bosons often appear as final states in the decay of new particles predicted by theories beyond the SM, like additional heavy Higgs bosons, W' or Z' , just to name a few. Many of these new particles have masses on the TeV scale, much larger than the masses of the electroweak gauge bosons themselves. Therefore, we can expect the bosons to be heavily boosted in the laboratory frame. As discussed earlier, we know that the hadronic decay products of the electroweak gauge bosons are collimated and reconstructed into a single fat jet. Because we expect electroweak gauge bosons with high boosts and the cross section for QCD background is steeply falling, the study of fat jets gives a very promising experimental signature to search for new physics in the high energy regime where cross sections are the main limitations. Having the ability to identify boosted $W^\pm$ and Z^0 bosons over the still immense QCD background is crucial to these studies.

It is instructive to try the reconstruction of boosted electroweak gauge bosons in an all-hadronic event first, as the process provides a large cross section. Furthermore, it serves as a proof-of-concept with respect to searches for new particles since it is necessary to show that a new technique can reproduce known resonances in the mass spectrum before using it in search of new physics. While most studies involving electroweak gauge bosons use the clear experimental signature of the leptonic decays and accept to rely on some assumptions resulting from the presence of invisible neutrinos, the all-hadronic analysis is independent of the event-topology and therefore provides a model-independent way to double-check SM predictions.

In the following, we want to investigate a pure two-body process in a blinded study, where a fat jet as the $W^\pm/Z^0$ candidate containing all decay products of the electroweak gauge bosons is found on one side of the detector and a second high p_T narrow recoil jet on the other side. A mass-drop filtering algorithm is employed to suppress pileup and soft radiation. To reduce the fraction of selected fat jets resulting not from the decay of a boosted electroweak gauge boson, a number of jet and event shape variables are used to train two BDTs. The signal yield determined from fit functions is then used to calculate the production cross section of hadronically decaying $W^\pm$ and Z^0 bosons, henceforth collectively called V_h .

6.2 Monte Carlo and data samples

Data sample

The analysis uses all of the proton-proton collisions collected at the ATLAS detector in 2012. The data corresponds to a total integrated luminosity of $\mathcal{L} = 20.3 \text{ fb}^{-1}$. Table

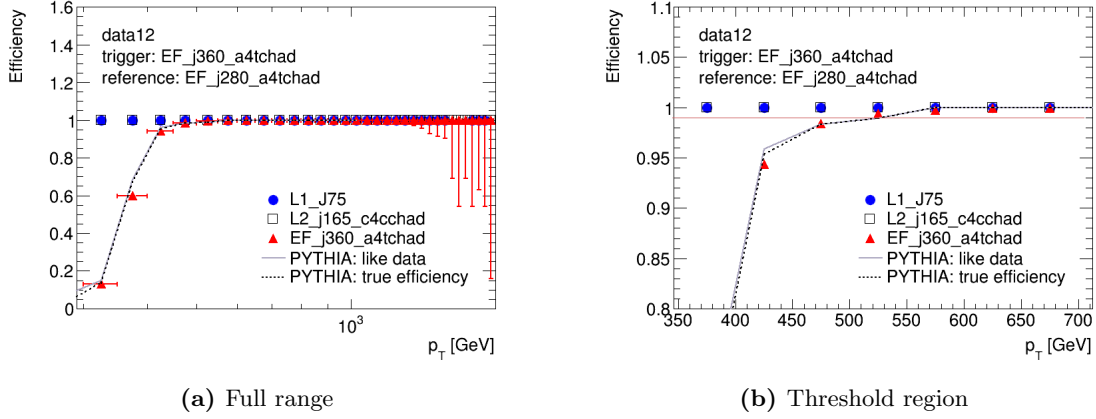


Figure 6.1: Efficiency of the EF_j360_a4tchad trigger as a function of jet p_T for Anti- k_t $R = 0.6$ jets. For the measurement of the efficiency a second trigger with smaller p_T threshold was used as reference. (a) shows the full p_T range, while (b) focuses on the threshold region where the efficiency approaches unity. The figure was taken from [168].

6.1 lists all data taking periods with the associated run numbers and the corresponding integrated luminosity. To capture the data unrescaled high p_T single jet triggers EF_L1J350_NoAlg, EF_j360_a4tchad and EF_j360_a10tcm are used. The trigger efficiency is depending heavily on the transverse momentum of the jet at hand. Measuring the trigger efficiency as a function of the jet p_T reveals very low efficiencies for small momenta that rise to form a plateau for higher jet momenta. In figure 6.1, this is shown exemplary for EF_j360_a4tchad in data, where the trigger EF_j280_a4tchad, which reaches maximum efficiency at much smaller momenta, was used as a reference. Even though the figure was produced for Anti- k_t $R = 0.6$ jets, which are not used here, it reflects the trigger behaviour well. In this study we want to make sure to only select events containing jets for which the trigger efficiency is very close to unity. This requires a harsh cut on the transverse momentum of the jet. The trigger reaches 99% efficiency for (492.5 ± 50.0) GeV [168].

The recommended standard physics good-run list `data12_8TeV.periodAllYear_DetStatus-v61-pro14-02_DQDefects-00-01-00_PHYS_StandardGRL_All_Good.xml` is applied to the data to make sure that only those collisions are analysed for which the detector

Period	Runs	$\mathcal{L}$ [pb $^{-1}$]
A	200604 - 201556	794.02
B	202660 - 205113	5094.66
C	206248 - 207397	1406.02
D	207447 - 209025	3288.39
E	209074 - 210306	2526.28
G	211522 - 212272	1274.81
H	212619 - 213359	1444.93
I	213431 - 213819	1016.26
J	213900 - 215091	2596.34
L	215414 - 215643	839.77
Total		20281.4

Table 6.1: Overview of the ATLAS data taking periods in 2012, associated runs and the corresponding recorded integrated luminosity of the data used in this analysis.

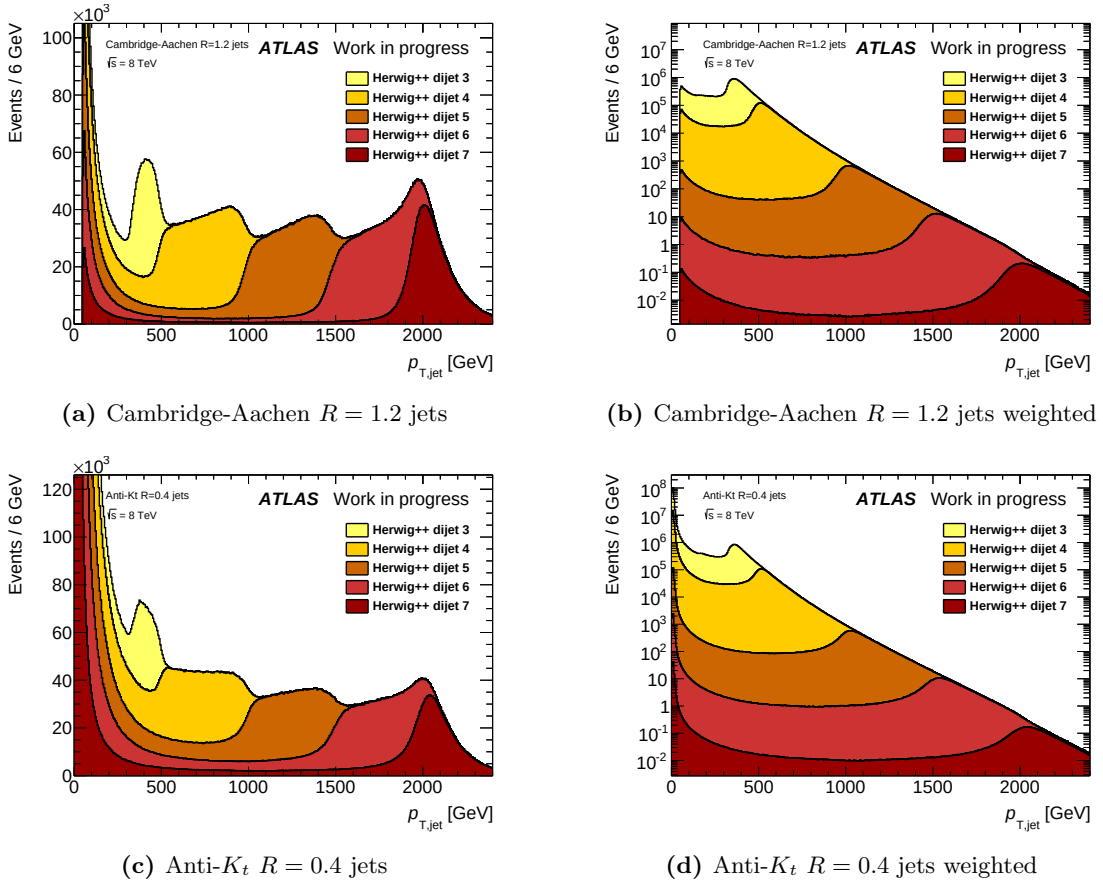


Figure 6.2: p_T spectra of the MC background sample. For production reasons the sample was produced not as a whole but in several segments. To retrieve the smoothly falling background an individual event weight has been applied. (a) shows the Cambridge-Aachen $R = 1.2$ fat jets before application of the weights, (b) after reweighing and scaling to the target luminosity, and (c) and (d) show the corresponding distributions for the Anti- K_t $R = 0.4$ narrow jets.

operated within normal parameters and all subsystems were operational. The analysis uses standard NTUP_COMMON D3PD samples that contain jet substructure information.

Monte Carlo sample

Several MC samples that involve the production of hadronically decaying electroweak gauge bosons and jets with large transverse momenta are used in this study to describe the data. Since multivariate analysis techniques are used on MC, attention was paid during the sample selection to be as consistent as possible in generators between signal MC and background MC.

For the QCD background dijet samples have been simulated using the HERWIG++ [169] generator with the CTEQ6L1-UE-EE-3 tune. The HERWIG algorithms predicate on leading-order (LO) perturbative calculation, which makes corrections for next-to-leading-order (NLO) inevitable.

Signal MC contains samples of hadronically decaying electroweak gauge bosons in combination with jets. Both samples for W+jets and Z+jets are produced with the HERWIG [155] generator expanded by PHOTOS [170] and TAUOLA [171], which are packages for the simulation of radiative correction and τ -lepton decay. HERWIG uses an

MC sample	Generator	p_T range [GeV]	σ [nb]	N_{events}	$\mathcal{L}$ [fb $^{-1}$]
W + jets	HERWIG	> 350	$2.0254 \cdot 10^{-3}$	$2.0 \cdot 10^5$	98.75
Z + jets	HERWIG	> 350	$8.4696 \cdot 10^{-4}$	$1.0 \cdot 10^5$	118.07
Dijet 3	HERWIG++	350 – 500	362.240	$1.4985 \cdot 10^6$	0.004
Dijet 4	HERWIG++	500 – 1000	4.165	$1.4981 \cdot 10^6$	0.36
Dijet 5	HERWIG++	1000 – 1500	$8.3179 \cdot 10^{-2}$	$1.4976 \cdot 10^6$	18.00
Dijet 6	HERWIG++	1500 – 2000	$5.785 \cdot 10^{-3}$	$1.4968 \cdot 10^6$	258.74
Dijet 7	HERWIG++	> 2000	$6.525 \cdot 10^{-4}$	$1.497 \cdot 10^6$	2294

Table 6.2: Summary of all MC samples used in the analysis. p_T ranges give the approximate range for which events have been generated in a particular sample. The Cross sections σ are taken from the ATLAS Metadata Interface (AMI) [174] and the datasets corresponding integrated luminosity is calculated from the cross section and the number of generated events N_{events} . The samples Dijet 0, Dijet 1 and Dijet 2 were produced at very low momenta and are not of interest in this study.

only slightly different hadronisation model than HERWIG++. The CTEQ6L1 [119] PDF set has been utilized for LO calculation, which means that a correction for NLO has to be applied during the analysis to account for the deviating cross-sections. The simulated decay modes for the electroweak gauge bosons are $W^+ \rightarrow u\bar{d}, c\bar{d}, c\bar{s}$, $W^- \rightarrow d\bar{u}, d\bar{c}, s\bar{c}$ and $Z^0 \rightarrow u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}, b\bar{b}$.

All MC samples were simulated with pileup and reweighed to make the average number of interactions per bunch-crossing consistent with the actual value in data. The simulated samples are fed in the GEANT4 [172, 173] simulation of the ATLAS detector and processed using the standard ATLAS reconstruction software.

Since the p_T -spectrum of the simulated dijet sample falls off very steeply, it would be necessary to simulate millions of uninteresting low- p_T events to get a handful of interesting high- p_T events. Since this would be uneconomical, a different approach is chosen. Instead of simulating the spectrum as a whole, it is divided into several regions that are simulated separately. This ensures that even the very high- p_T regions with transverse momenta $p_T > 2 \text{ TeV}$ still hold high statistics. However, this introduces some problems as well, since we need the continuous spectrum to get an accurate simulation. In figure 6.2(a) and 6.2(c) the unweighted p_T -spectrum is depicted for Cambridge-Aachen $R = 1.2$ clustered fat jets and Anti- K_t $R = 0.4$ narrow jets, respectively. One can see the jagged profile of the spectrum instead of the smoothly falling shape. It is therefore necessary to reweigh the samples according to their cross-section and apply an additional weight on an event-by-event basis, to get the spectrum approximately smooth. The p_T spectrum after reweighing and scaling to the target luminosity is shown in figure 6.2(b) and 6.2(d) for fat and narrow jets, respectively. We have only considered the samples dijet 3-7 and ignored the samples produced at lower p_T . The figure shows that this works fine for p_T cuts above 350 GeV.

A summary of all MC samples including the p_T range of the sample, its cross section, number of events and corresponding integrated luminosity can be found in table 6.2.

6.3 Event and jet selection

An event is required to have at least one reconstructed primary vertex with a minimum of five associated tracks with $p_T > 500 \text{ MeV}$. Electroweak gauge boson candidates are found by requiring a reconstructed jet clustered from topological clusters with $R = 1.2$ using the

	isBadLooseMinus selection criteria
HEC spikes	$(f_{HEC} > 0.5 \text{ and } f_{HEC,quality} > 0.5 \text{ and } \overline{LArQ} > 0.8)$ or $ E_{neg} > 60 \text{ GeV}$
EM coherent noise	$f_{EM} > 0.95 \text{ and } f_{LAr,quality} > 0.8 \text{ and } \overline{LArQ} > 0.8$ and $ \eta < 2.8$
Non-collision background	$(f_{EM<0.05} \text{ and } f_{ch} < 0.05 \text{ and } \eta < 2)$ or $f_{EM} < 0.05 \text{ and } \eta \geq 2$ or $f_{max} > 0.99 \text{ and } \eta < 2$

Table 6.3: Quality criteria applied to all narrow jets. For details and explanations of the variable names we refer to [175] and [176].

Cambridge-Aachen algorithm and a minimum transverse momentum of $p_T > 420 \text{ GeV}$ ¹³. A mass-drop filtering procedure with parameters $\mu = 0.67$ and $y_{\text{cut}} = 0.09$ is then applied to the jet. The filtered jet is only reconstructed if it is composed of either two or three subjets after reclustering. Jets for which only a single subjet is found are discarded. Heavy objects are centrally produced thus the filtered jets are required to be within $|\eta| < 1.9$, i.e. in the central region of the detector, where the calorimeter performs best and which is also covered by the ID. This guarantees that the large area of the fat jet is completely contained within the calorimeter. Furthermore, the filtered jet is required to originate at the primary vertex and its mass has to be in the interesting region of $40.0 \text{ GeV} < m_{\text{jet}} < 160.0 \text{ GeV}$. We apply cuts on the subjets as well to reject jets with fake substructure from pileup. The leading and sub-leading subjets are required to have a mass of at least 500 MeV . A filtered jet that fulfills all of the above requirements is labelled a W/Z candidate jet.

On the other side of the detector we demand a second jet, against which the W/Z candidate jet can recoil. The recoil jet is reconstructed from topological clusters using the Anti- K_t algorithm with a radius parameter of $R = 0.4$. The jet has to be a good quality jet, i.e. not contaminated by energy deposits in the calorimeter that are due to hardware effects such as noise in the calorimeter, LHC beam conditions like the collision of a proton with a residual gas molecule in the beam pipe or non-collision background from cosmic showers. The quality criteria suggested by the Jet and Etmiss combined performance group are used in this analysis. For completeness the requirements are listed in table 6.3. We refer to [175] and [176] for a detailed description. The good quality jets are required to be central within $|\eta| < 1.9$ and to originate from the primary vertex. In the ideal case the W/Z candidate jet and the recoil jet are back-to-back. Therefore, we require the recoil jet to be in the opposite hemisphere of the W/Z candidate jet and demand a minimal distance of $\Delta R > 2.5$ to it. In case more than one narrow jet fits this selection, the one which yields the largest invariant mass with the filtered fat jet is chosen as the selected recoil jet.

Only events in which a W/Z candidate jet and an associated recoil jet have been identified are analysed further. The remaining events are discarded.

The event selection is summarized in table 6.4 where the number of events passing a certain cut is listed for signal and background MC next to data.

¹³The choice of $p_T > 420 \text{ GeV}$ corresponds to a trigger efficiency of roughly 95% as can be seen from figure 6.1.

cut	signal MC		background MC		data	
Trigger selection	23366	100.0%	11904083	100.0%	5041839	100.0%
Require good vertex	23366	100.0%	11904083	100.0%	5041834	100.0%
Fat jet with substructure	23235	99.4%	11824836	99.3%	5006639	99.3%
Substructure cuts	19623	84.0%	9176422	77.1%	3666659	72.7%
Mass window	17498	74.9%	6665304	56.0%	2499331	49.6%
Recoiljet identified	13083	56.0%	5419377	45.5%	1997221	39.6%
p_T cut	6278	26.9%	1434939	12.1%	631734	12.5%

Table 6.4: Cutflow of the event selection. The numbers indicate the amount of events passing a certain cut in addition to all previously listed cuts.

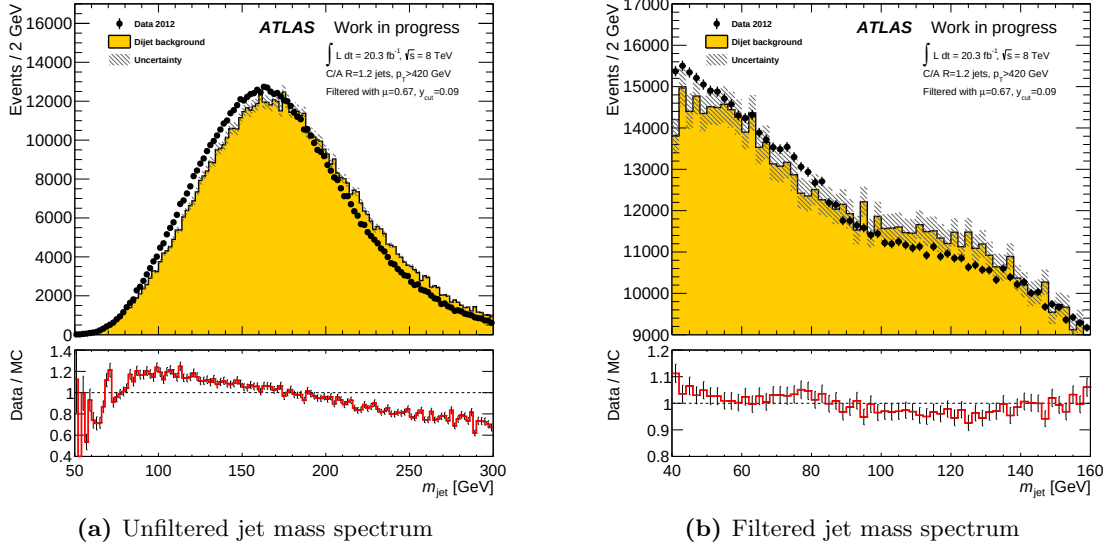


Figure 6.3: Unfiltered (left) and filtered (right) jet mass spectrum in data and dijet background MC. In both plots MC has been normalized to data separately. The lower panels show the ratio of data and MC.

6.4 Unfiltered jet mass spectrum

The jet mass spectrum of the selected jets prior to filtering in data and MC is shown in figure 6.3(a), where the MC prediction has been normalized to data. The large discrepancies between data and MC can be read from the lower panel and is amongst others due to the absence of a jet mass calibration for C/A $R = 1.2$ jets for $\sqrt{s} = 8$ TeV data [177]. It is also a reproduction of results from an earlier study where a correction for $\sqrt{s} = 7$ TeV data was attempted [178]. It was found that even though the agreement was poor, shapes were often produced correctly. This can be confirmed from the figure.

Furthermore, figure 6.3(b) shows the mass spectrum for the same jets after filtering. While the disagreement between data and MC is still significant the principal features of the distribution like the drop-off at low jet masses or the shoulder around 120 GeV are modelled correctly. This is again in agreement with earlier results where it was found that the filtered jet mass of Cambridge-Aachen jets after filtering was the only tested groomed variable for which the MC prediction agreed to data within statistical uncertainties [178]. Due to the limited statistics in MC the histogram shows a jagged shape. In this and all subsequent plots the statistical error of each bin was enlarged by an estimated systematic uncertainty of 2.0%. The total uncertainty per bin assumes values around 3.0%.

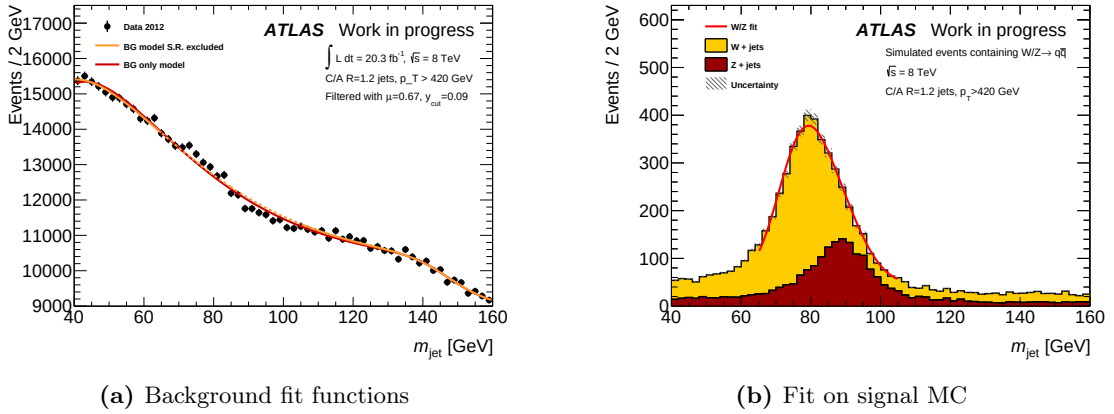


Figure 6.4: (a) shows the fit of the background model to the data both fitted over the full range and with the signal region excluded. The excluded part of the function is indicated as a dashed line. (b) shows the sum of two Gaussians fitted to signal MC.

The conclusion to be drawn from this is that we cannot rely on the background predicted by MC to aid the fitting process of the filtered jet mass spectrum.

Since we are not interested in measuring the V_h mass and rather want to determine the signal yield, we do not attempt to correct the filtered jet mass according to the results from chapter 5.3 to prevent the introduction of additional uncertainties. We can therefore expect to extract masses around 3.5% lower than literature values.

It is legitimate to assume that a background model that describes the data outside the signal region well is also a valid model within the signal region. We define the signal region as $65.0 \text{ GeV} < m_{jet} < 105.0 \text{ GeV}$. The model used to describe the background is driven by two main observations. While in the case of narrow jets without substructure the number of jets decreases potentially with increasing jet masses, this is not the case for filtered jets. The appearance of a high- p_T jet that exhibits substructure and has a very small mass is just as unlikely as a jet exhibiting a very large mass under the same conditions. This results in the formation of a maximum in the spectrum at low jet masses¹⁴ which is followed by the expected exponential reduction. The second observation is the shoulder that appears in the spectrum for jet masses above $\approx 120 \text{ GeV}$. Both characteristics are combined into one fitting function of the form

$$\mathcal{B}(m_{jet}) = A_1 + e^{A_2 - A_3 \cdot (m_{jet} - A_4)} \cdot A_5(m_{jet} - A_6) + A_7 \cdot \frac{2}{\sqrt{\pi}} \int_{A_8(m_{jet} - A_9)}^{\infty} e^{-t^2} dt \quad (6.1)$$

where the first summand is a constant offset, the second one describes the maximum for low jet masses and the subsequent exponential fall-off, while the third summand is the complementary error function and seeks to model the shoulder in the spectrum for higher jet masses. The coefficients A_x are the parameters given to the fit. Fitting this model to the data with the signal region excluded yields a χ^2 -probability of 67.9% ($\chi^2 = 27.81$ with 32 degrees of freedom). The fit can be seen in figure 6.4(a), where the excluded part of the function is indicated by a dashed line. The fit performs reasonably well to be considered a valid background model. Next, we want to explore whether the data favours a model that includes signal or whether the background-only hypothesis is satisfactory. In order to do that the data is fitted twice again and the resulting χ^2 values are compared.

¹⁴Here, the maximum is located at the lower edge of the region of interest at around 35 – 40 GeV and barely visible.

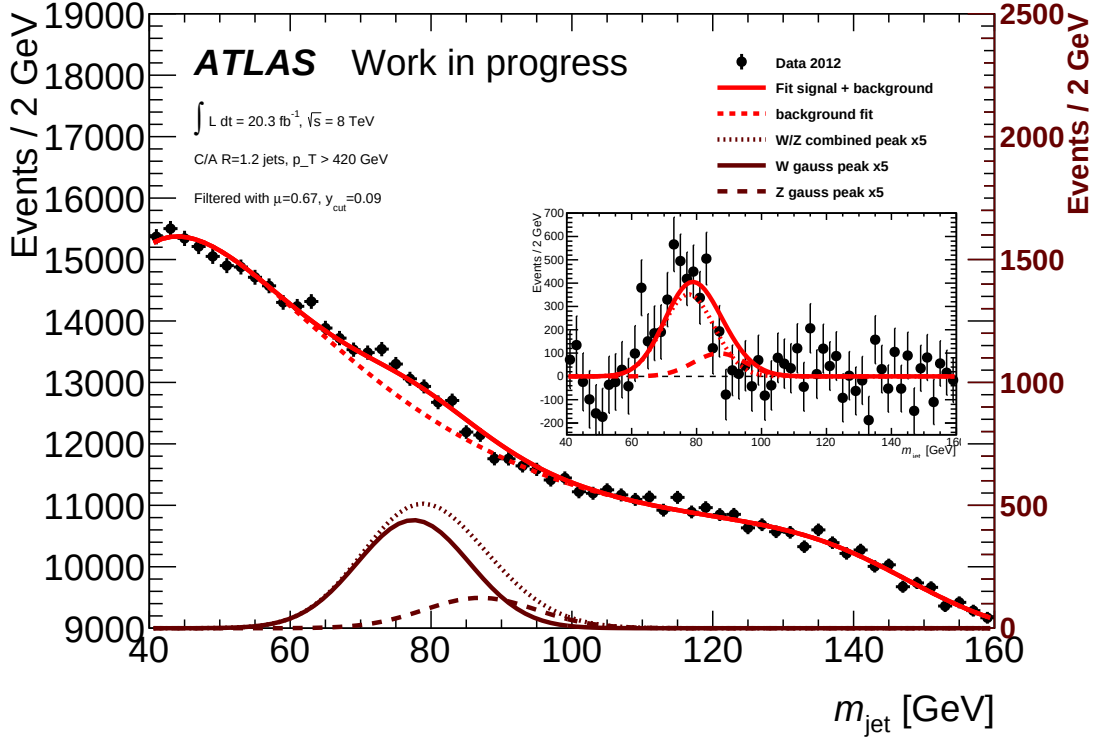


Figure 6.5: Fit of the filtered jet mass spectrum with two Gaussians modelling the signal. The dashed line indicates the background without the fit. The dark red functions at the bottom show the magnified overall peak and the individual contributions from $W^\pm$ and Z^0 . The inlay gives the data with the background part of the fit subtracted. The functions in the inlay are not fitting functions but the signal part of the fit.

The background only model does not require much additional work. Instead of excluding the signal region the model is fitted to the whole range as was drawn in figure 6.4(a) as well. All parameters are left floating. The fit yields $\chi^2 = 81.21$.

Since we expect to see signal from both $W^\pm$ and Z^0 boson it is straightforward to model the signal with the sum of two Gaussians. This is tested by fitting signal MC with this model, as can be seen in figure 6.4(b). For the moment we are only interested in seeing a peak in the jet mass spectrum and do not worry about an accurate signal yield, which is why we are satisfied by only fitting the peak and ignoring the tails of the distribution. The signal+background model is given by the function

$$\Gamma(m_{jet}) = \mathcal{B}(m_{jet}, A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9) + A_{10} \cdot \left[\frac{S_{wz}}{\sigma_w \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{m_{jet} - \mu_w}{\sigma_w} \right)^2} + \frac{1 - S_{wz}}{\sigma_z \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{m_{jet} - \mu_z}{\sigma_z} \right)^2} \right] \quad (6.2)$$

Here, S_{wz} is the relative fraction of $W^\pm$ and Z^0 bosons, $\mu_{w/z}$ is the mean of the $W^\pm/Z^0$ peak and $\sigma_{w/z}$ its width. All of these parameters are fixed to the values obtained in the fit on MC:

$$\begin{aligned} \mu_w &= 78.48 \text{ GeV} & \sigma_w &= 7.95 \text{ GeV} & S_{wz} &= 0.78 \\ \mu_z &= 88.44 \text{ GeV} & \sigma_z &= 7.98 \text{ GeV} & & \end{aligned} \quad (6.3)$$

This leaves only the size of the signal (A_{10}) floating. The fit of the model to the data can be seen as a solid line in figure 6.5. The background component of the fit is shown as

a dashed line for comparison. The signal contribution is plotted five times magnified in dark red colour at the bottom of the figure and also shows the individual contributions from $W^\pm$ and Z^0 . The fit yields $\chi^2 = 60.76$.

The significance of the signal peak can be calculated by

$$\text{sig.} = \sqrt{\Delta\chi^2 - \Delta N_{\text{dof}}} \quad . \quad (6.4)$$

Here ΔN_{dof} is the difference in the number of degrees of freedom between the background-only fit and the signal+background fit. Putting in the numbers gives a peak significance of 4.4σ . The data strongly prefers a contribution from the signal model. This only holds if the background model is valid. The inlay in figure 6.5 shows the data with the background subtracted. A peak is clearly visible. However, the significance of this peak can be improved by employing multivariate data techniques.

Since the energy patterns of jets resulting from the decay of boosted objects are quite different from the ones in pure QCD background jets, it is possible to identify a number of variables which allow to discriminate between QCD background and true V_h jets. These variables are suited as input variables for a multivariate analysis (MVA). Some of the discriminators were later found to be (partly heavily) correlated to the filtered jet mass and therefore not used in the final version of the MVA. It is, however, instructive to discuss them here anyway, since they might be of some relevance to related studies in the future and illustrate the line of action.

6.5 Discriminating variables

We introduce the discriminating variables in three separate groups. 6.5.1 describes the discriminators suggested by the jet substructure and jet-by-jet tagging combined performance subgroup and their performances in the present study. 6.5.2 introduces three event shape variables which are calculated in the rest frame of the jet and 6.5.3 contains descriptions of all remaining discriminators that have been found during this project.

Hereafter, the term QCD jet refers to all types of jets that originate from QCD processes and are associated to the background and do not result from boosted V_h decays.

6.5.1 Substructure Observables

N-subjettiness

The N-subjettiness variables τ_N [179, 180] exploit the differences in energy flow between QCD and $W^\pm/Z^0$ jets by giving a measure of how well a jet is described on the assumption of being composed of N subjets. Jets are clustered with the exclusive- k_t clustering algorithm and forced to cluster the jet in exactly N subjets using all the jet constituents. Each subjet defines an axis around which energy is clustered like lobes. The N-subjettiness variable τ_N is now calculated by computing

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \cdot \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}) \quad . \quad (6.5)$$

The sum considers all jet constituents k with their corresponding transverse momenta $p_{T,k}$. $\Delta R_{S,k}$ symbolizes the distance from the current jet constituent k to a given subjet S . The variable is normalized by the factor

$$\frac{1}{d_0} = \left(\sum_k p_{T,k} \cdot R_0 \right)^{-1} \quad , \quad (6.6)$$

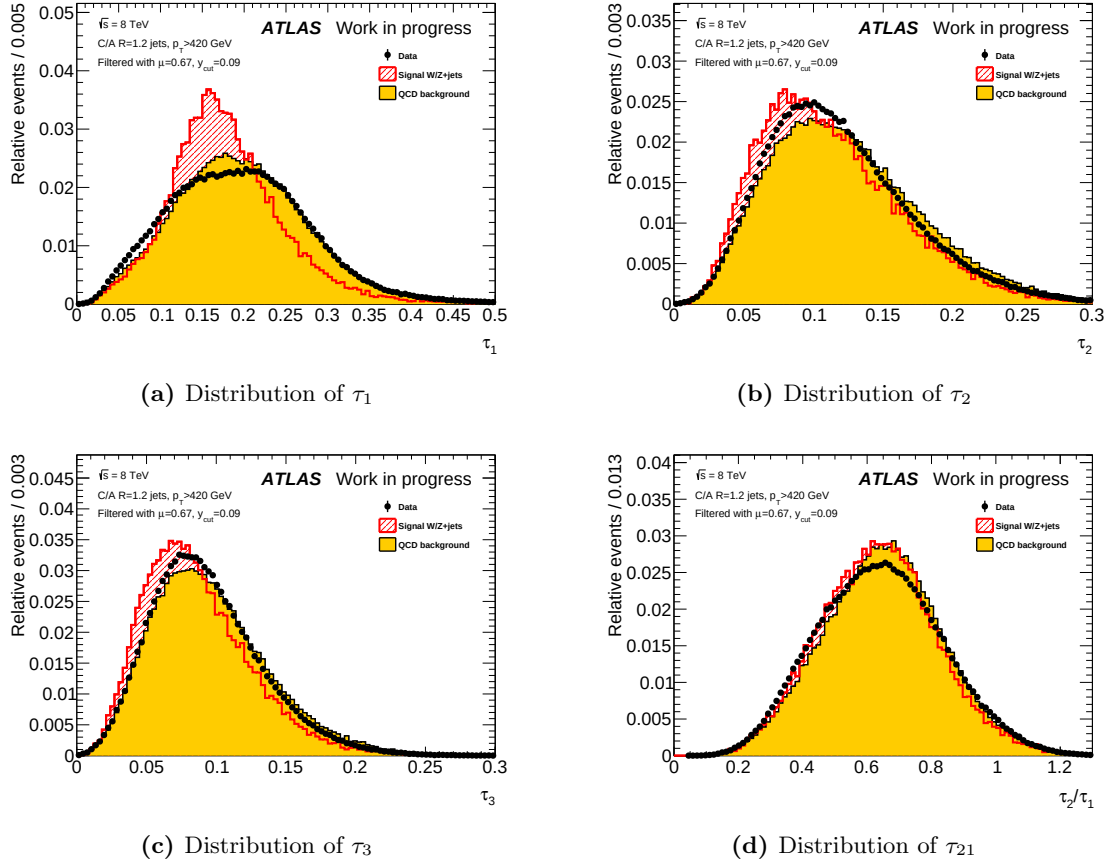


Figure 6.6: Separation power of the N-subjettiness variables. τ_{21} is defined as the ratio τ_2/τ_1 . Plotted are the distributions for signal and background MC and for data with the signal region excluded.

where R_0 is the radius parameter used to cluster the original jet.

It has now become evident why τ_N is referred to as N-subjettiness. Those jets in which the energy flow is parallel to the axis defined by the subjects exhibit a small value for τ_N . This means the jet is well described by N subjects and can therefore be expected to be composed of N or fewer subjects. On the other hand, if τ_N is large, i.e. a significant part of the jets energy flow differs from the defined axis, the jet is better described as being composed of at least $N + 1$ subjects.

Because we want to identify boosted V_h we expect to select jets with a small τ_2 and high τ_1 . In many previous studies, it was found that not the distinct variables τ_1 and τ_2 yield the optimal separation power, but rather the ratio of the two $\tau_2/\tau_1 = \tau_{21}$ [133, 179, 180, 181]. The ratio τ_N/τ_{N-1} is believed to be a robust measure for the identification of N-prong decays [182]. In the present study, however, it has been found that this choice is not optimal in usage after a filtering algorithm has been applied. In fact, it can be easily seen from figures 6.6(a) to 6.6(c) that the plain variables yield a much better separation than their ratio in figure 6.6(d). In the figures, the distributions for signal and background MC (see 6.2) are plotted next to the distributions in data for which the signal region has been vetoed and only the mass ranges of $50.0 \text{ GeV} < m_{\text{jet}} < 65.0 \text{ GeV}$ and $105.0 \text{ GeV} < m_{\text{jet}} < 135.0 \text{ GeV}$ are accepted (we will talk about this in more detail in 6.6.1, but figure 6.14 should give a sufficient impression for now). All three distributions have been individually normalized to one. In the ideal case of a perfect MC description of

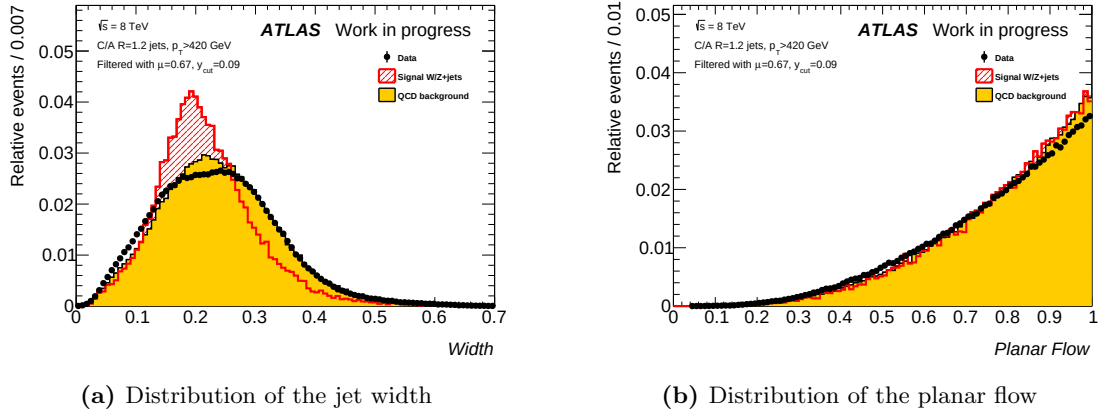


Figure 6.7: Separation power of the jet width and planar flow variables. Plotted are the distributions for signal and background MC and for data with the signal region excluded.

real life events, the distribution for background MC and data should be close to identical.

Jet width

The jet width is a very simple concept. It is defined as

$$W_{\text{jet}} = \frac{\sum_i \Delta R^i p_T^i}{\sum_i p_T^i} \quad , \quad (6.7)$$

where the sums run over all constituents i of the jet. p_T^i corresponds to the transverse momentum of the i^{th} jet constituent and ΔR^i describes its distance from the jet axis in the rapidity-azimuth plane [183]. A jet with a low width has most of its constituents close to the jet axis, while higher values indicate the absence of such a core. A V_h jet would therefore prefer higher values for the jet width. Very small jet widths do not occur since the mass-drop filtering requires substructure, which automatically results in large jet widths. This can be seen in figure 6.7(a).

Planar Flow

The planar flow variable [183, 184] can be used to distinguish multi-body decays from QCD jets, by giving a measure of how linear or planar the energy is deposited in projection on the forward face of the jet cone, i.e. the plane perpendicular to the jet axis. Before computing the planar flow one has to construct the matrix

$$I_w^{mn} = \frac{1}{m_{\text{jet}}} \sum_i \frac{p_m^i \cdot p_n^i}{w_i} \quad (6.8)$$

for a given jet with the constituent particles i , where m_{jet} is the jet mass, w_i the energy of the particle i and p_T^m is the m^{th} component of the particles transverse momentum vector with respect to the axis of the jet. The Eigenvalues of the 2-dimensional matrix I_w^{mn} are labelled λ_1 and λ_2 . The planar flow P_f can now be calculated to be

$$P_f = \frac{4\lambda_1\lambda_2}{(\lambda_1 + \lambda_2)^2} \quad . \quad (6.9)$$

Jets with a fairly linear distribution of energy, consistent with for example a two-body decay like the V_h decays we are interested in, exhibit small P_f , while jets with more

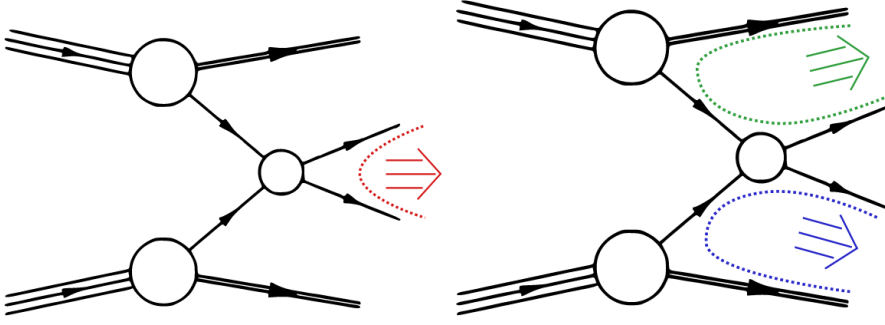


Figure 6.8: Illustration of the possible colour connections in a signal event on the left and a background event on the right. The figure was taken from [186].

uniformly distributed energy patterns approach a value of $P_f = 1$. Therefore, we can expect jets resulting from ≥ 3 -body decays (like e.g. the decay of boosted top quarks) to have values close to one. Those QCD jets that result from the emission of a gluon and exhibit a high mass and high transverse momentum should exhibit a planar flow much smaller than one [185]. However, since the jet was already required to exhibit a substructure inconsistent with this type of jet, the low value peak is suppressed in this study. The pure QCD spectrum is found to be of a rising distribution with peak at $P_f = 1$. Unfortunately, the planar flow variable is sensitive to pileup. This makes the distribution for V_h jets almost indistinguishable from the QCD background as can be seen in figure 6.7(b). As a result the planar flow variable is not used in this analysis.

Dipolarity

Jet dipolarity [187] is a jet observable based on the colour flow. The distribution of radiation in an event is not only governed by the hard parton collision, but also on the colour flow of an event, i.e. the contraction of the colour indices of two partons. We expect a colour-connection of particle strings to form between these two partons resulting in a multitude of particles in the space between them. In the case of a V_h decay where the colour connection exists between the two quarks it is sensible to assume a preference to radiate colour particles in between the two decay products.

This concept is used to formulate the jet dipolarity variable. Let us assume a jet consists of two subjets labelled j_1 and j_2 whose positions are given by coordinates $(\eta_{1/2}, \phi_{1/2})$ as an input to the algorithm. Considering all calorimeter cells i at position (η_i, ϕ_i) with transverse momentum p_T^i of the jet, the jet dipolarity $\mathcal{D}$ can be computed by

$$\mathcal{D} = \frac{1}{\Delta R_{jj}^2} \sum_{i \in jet} \frac{p_T^i}{p_T^{jet}} \cdot \Delta R_i^2 \quad . \quad (6.10)$$

Here, ΔR_{jj} describes the distance between the two subjets in the rapidity-azimuth plane, p_T^{jet} corresponds to the transverse momentum of the parent jet and ΔR_i is the minimal euclidean distance of the calorimeter cell i to the connecting line between (η_1, ϕ_1) and (η_2, ϕ_2) .

The colour information of the parent jet can be deduced from semisoft radiation strayed aside from the cores of the subjet. It is through the ΔR_i weighting that those contribute to a large degree to the calculation of the dipolarity. Therefore, its value will be different depending on the topology of the jet. It will assume lower values if the radiation occurs predominantly between the two subjets and higher values if not, that is if the radiation is rather drawn towards the beam, as it occurs for coloured object jets.

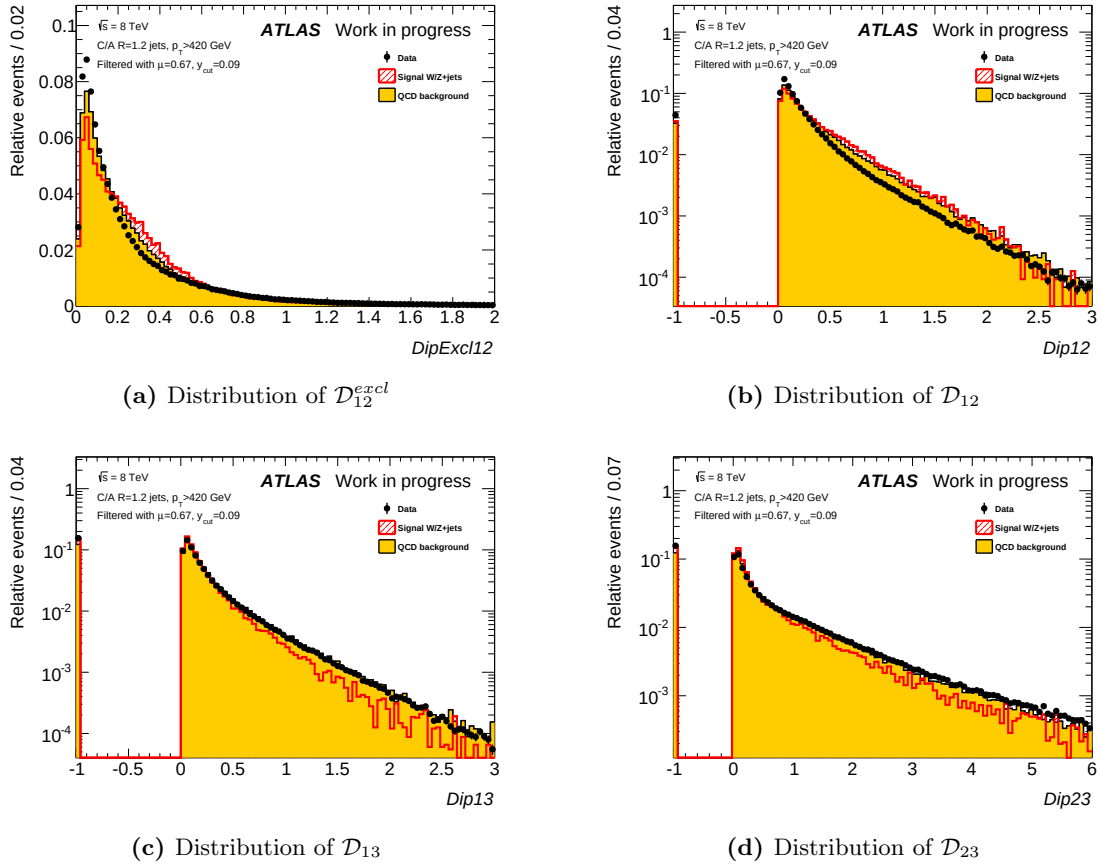


Figure 6.9: Separation power of the dipolarity variables. Plotted are the distributions for signal and background MC and for data with the signal region excluded. $\mathcal{D}_{12}^{excl}$ is exclusively for the case of a fat jet composed of two subjets, while the remaining variables give all permutations of the three-subjet case.

The above definition is exclusive for jets with other than two identified subjets and the dipolarity therefore labelled $\mathcal{D}_{12}^{excl}$. The concept can be generalized to jets with three subjets as well. This leads to the introduction of further jet dipolarities with respect to the third subjet $\mathcal{D}_{12}$, $\mathcal{D}_{13}$ and $\mathcal{D}_{23}$ [188].

Figure 6.9(a) shows the separation power of dipolarity in the case of two subjets, while figure 6.9(b), 6.9(c) and 6.9(d) give the three-subjet case with all combinations of subjets. The accumulation of jets with dipolarity values at $\mathcal{D}_{ij} = -1$ ($i, j = 1, 2, 3, i \neq j$) is due to a failed computation of the variable in the case of a jet composed of only 2 subjets. The figures show that the capability of discriminating between V_h jets and QCD background jets on the basis of jet dipolarity is very limited. Furthermore it reveals some discrepancies between data and MC especially in the $\mathcal{D}_{12}^{excl}$ and $\mathcal{D}_{12}$ cases.

6.5.2 Event shape variables

While the previously discussed variables were calculated in the laboratory frame, the event shape variables allow us to study the jet substructure in the center-of-mass frame (rest frame) of the jet. This is promising since studies have shown that variables calculated in the rest frame of the jet are less correlated to the jet mass [189]. The rest frame of a jet is defined as the system in which the jets four momentum takes on the form $P_{jet} = (m_{jet}, 0, 0, 0)$. In our case this means the subjets get boosted to the center-of-mass

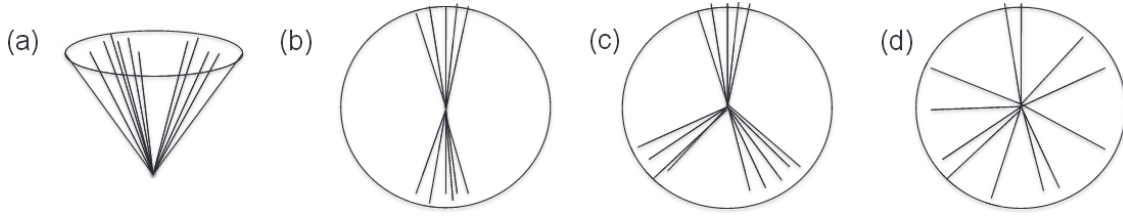


Figure 6.10: Distribution of the jet constituents throughout the jet in different reference frames. (a) shows a jet in the laboratory-frame. The subjects of a jet become apparent when moving to the jets rest frame as was done in (b) for a two-body decay and in (c) for a decay with three particles in its final state. In (d) the random distribution of constituents in a QCD jet is illustrated. The figure was taken from [189]

frame of the jet. In the jets rest frame the topology of a jet from a hadronic two-body decay like $W^\pm/Z^0$ or three-body decay like a top quark looks fundamentally different from QCD jets. As can be seen in figure 6.10, we can identify two (three) clear directional sprays of particles in case of a two(three)-body decay - the subjects. In fact this looks just like a common dijet event in the laboratory frame, where the two jets are back-to-back (compare to figure 4.2(a)). For a QCD jet, which acquires its mass through gluon radiation, no distinct subset configuration can be observed, as the constituents are spread randomly through the jet.

While in the original concept [189, 190, 191] jet energy clusters were used as input for the calculation of the event shape variables, this was not possible here. We instead use the subjects identified by the mass-drop filtering algorithm. This renders the variables less effective but they remain valuable discriminators still.

Sphericity

The sphericity tensor was introduced in the context of electron-positron annihilation models around 1970, i.e. prior to the first hadron-hadron collider experiment at CERN's Intersecting Storage Rings¹⁵ [192]. It is defined as

$$S^{mn} = \frac{\sum_{\iota} p_{\iota}^m \cdot p_{\iota}^n}{\sum_{\iota} |\vec{p}_{\iota}|^2} \quad , \quad (6.11)$$

where the sums run over all subjects ι and m and n are components of the subjects three-momentum vector in the rest frame of the parent jet. Diagonalizing the matrix yields three eigenvalues labelled λ_1 , λ_2 and λ_3 , which are normalized to satisfy $\lambda_1 + \lambda_2 + \lambda_3 = 1$. From the eigenvalues it is simple to calculate the sphericity

$$S = \frac{3}{2}(\lambda_2 + \lambda_3) \quad . \quad (6.12)$$

The sphericity can assume values between 0 and 1 and is a measure of the distribution of subjects. If two subjects are back-to-back and no further subjects identifiable, the sphericity is vanishing. In the opposite case where several subjects are randomly spread the sphericity assumes values greater than zero. The distribution for signal and background jets is plotted in figure 6.11(a). Unfortunately, the variable shows some discrepancies in the agreement between data and MC.

¹⁵Indeed, the shape variables have their origin in e^+e^- collision experiments rather than in hadron colliders like LHC.

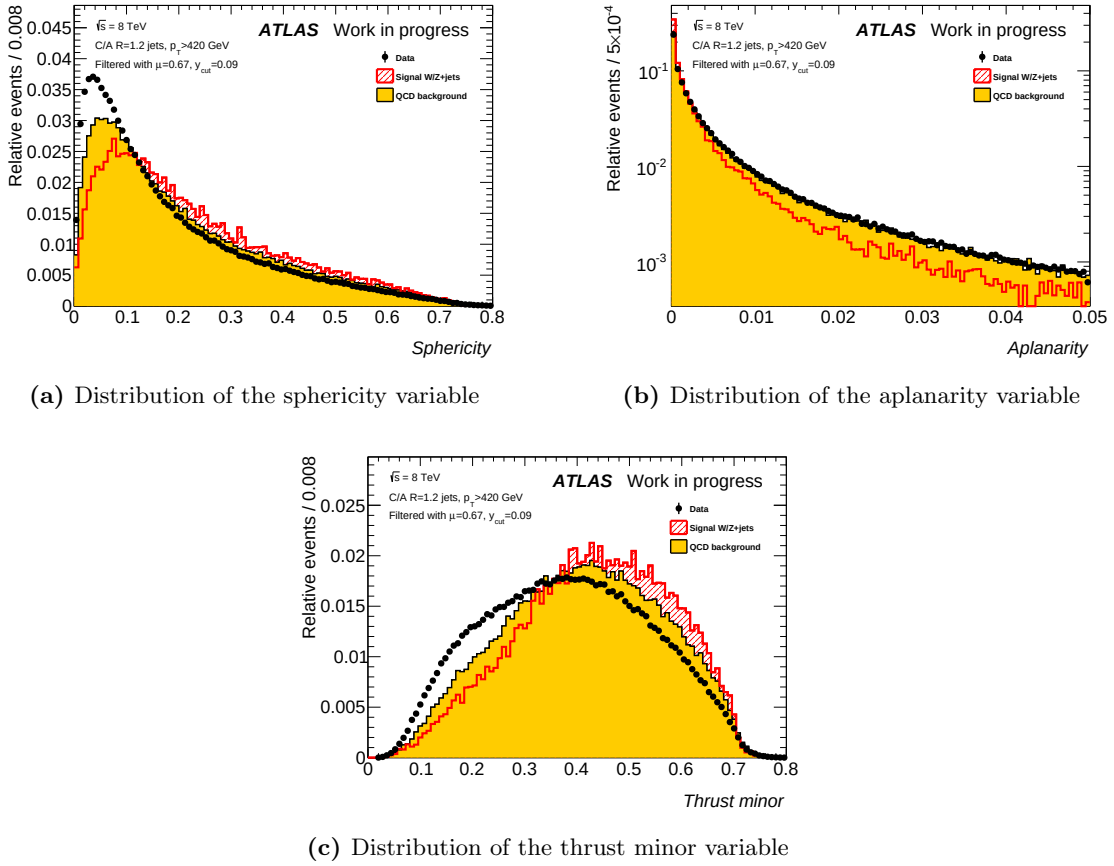


Figure 6.11: Separation power of the event shape variables sphericity, aplanarity and thrust minor. Plotted are the distributions for signal and background MC and for data with the signal region excluded.

Aplanarity

Aplanarity is closely related to sphericity and defined as

$$A = \frac{3}{2} \lambda_3 \quad (6.13)$$

where λ_3 is the eigenvalue calculated previously.

The eigenvectors corresponding to the eigenvalues λ_i are labelled $\vec{v}_i$. $\vec{v}_1$ is called the sphericity axis and together with $\vec{v}_2$ it spans the sphericity event plane [193]. It is now easy to see that aplanarity considers the transverse momentum out of the event plane. It is limited to the range $0 < A < 0.5$. If the subjets are distributed in an extremely directional fashion, i.e. the event is planar and $\lambda_3 = 0$, the aplanarity exhibits very small values. $A = 0.5$ indicates an isotropic subjet distribution. Figure 6.11(b) shows the distribution for signal and background jets.

Thrust minor

The direction within a jet that maximizes the sum of the longitudinal momenta of the jet energy clusters in the jets rest frame is called the thrust axis [194, 195]. It can be used to define the thrust variable

$$T = \max_{|\vec{n}|=1} \frac{\sum_i |\vec{n} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|} \quad (6.14)$$

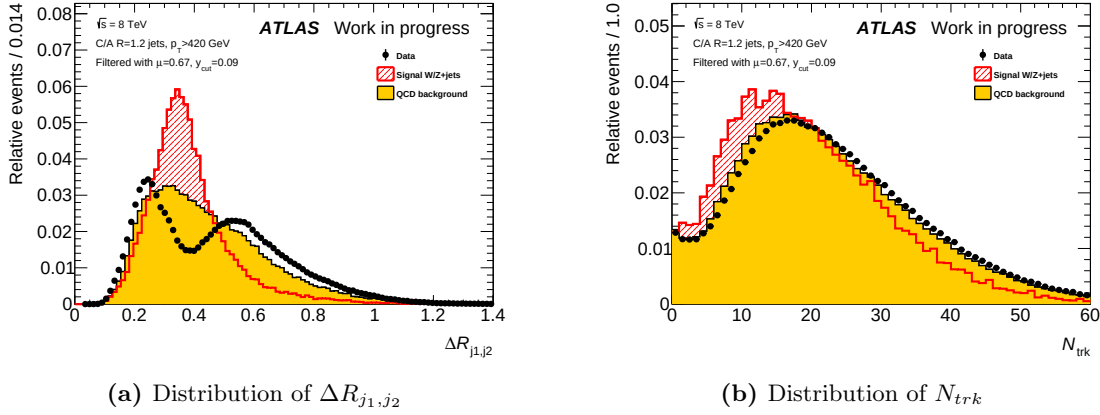


Figure 6.12: Distribution of the distance in ΔR between the two leading subjets and the number of charged tracks between the fat jet and the beam in a narrow segment of $\Delta\phi = 0.4$. Plotted are the distributions for signal and background MC and for data with the signal region excluded.

Thrust is constrained to the interval $0.5 \leq T \leq 1.0$, where large values indicate a directional distribution of subjets, while lower values are associated to isotropic distributions.

Related to thrust is the thrust minor variable which is used in this study. It is defined as

$$T_{min} = \frac{\sum_i |\vec{p}_i \times \vec{T}|}{\sum_i |\vec{p}_i|} \quad , \quad (6.15)$$

where $\vec{T}$ is the thrust axis. Just like for the thrust variable large values correspond to an isotropic distribution, while instead for highly directional jets thrust minor approaches zero. The separation power of the variable can be seen in figure 6.11(c). As for sphericity the variable is not described well in MC. But while sphericity only underestimates the size of the peak, thrust minor shows some discrepancies in the overall shape, as can be seen from the considerable differences between data and background MC.

6.5.3 Additional discriminators

Leading subjet separation

The distance in pseudorapidity-azimuth space between the two leading subjets of the fat jet can be used as a discriminator as well. While it yields a good separation as can be seen from figure 6.12(a), it is heavily correlated to the filtered mass of the fat jet as can be suspected from looking at figure 5.1. This correlation also results in the large discrepancies between data and background MC. The veto of the signal region in data carves out a dip in the continuous distribution at the position of the signal MC peak.

Number of charged tracks between the jet and the beam

Another variable in which we try to exploit the concept of colour flow counts the number of charged tracks in the detector space between the fat jet and the beam in a ϕ -slice of $\Delta\phi = 0.4$ around the jet axis for all $|\eta| > |\eta_{jet}|$. In a QCD background jet where a colour connection exists between the jet initiating parton and the remnants of the hard collision particle strings are formed in between them. Some¹⁶ of these particles appear as charged tracks in the detector. For a track we loosely follow [163] and require the following:

¹⁶Quarks hadronising into a neutral π^0 do not appear as a charged track.

- a minimal transverse momentum of $p_T^{trk} > 500 \text{ MeV}$
- a minimum of one hit in the pixel detector
- a minimum of six hits in the silicon microstrip tracker
- a maximal transverse impact parameter with respect to the primary vertex of $|d_0^{\text{wrtPV}}| < 2.5 \text{ mm}$
- a maximal longitudinal impact parameter with respect to the primary vertex of $|z_0^{\text{wrtPV}}| < 2.0 \text{ mm}$
- a maximal χ^2 per degree of freedom of the track fit of $\chi^2/\text{dof} = 2.5$

Figure 6.12(b) shows the distribution of the variable for signal and background jets. As expected the mean of the distribution for QCD background jets is slightly larger than the mean value for signal jets. It can be noted that the agreement between data and background MC is reasonably well.

Considering the recoiljet

$$\Delta\eta_{\text{jet, recoil}} = |\eta_{\text{jet}} - \eta_{\text{recoil}}| \quad \text{and} \quad (6.16)$$

$$\Delta R_{\text{jet, recoil}} = \sqrt{(\eta_{\text{jet}} - \eta_{\text{recoil}})^2 + (\phi_{\text{jet}} - \phi_{\text{recoil}})^2} \quad (6.17)$$

between the fat jet and the recoiljet as well as the invariant mass

$$m_{\text{jet+recoil}}^2 = m_{\text{jet}}^2 + m_{\text{recoil}}^2 + 2(E_{\text{jet}}E_{\text{recoil}} - \vec{p}_{\text{jet}} \cdot \vec{p}_{\text{recoil}}) \quad (6.18)$$

of the combined fat-jet-recoiljet system, where E is the energy and $\vec{p}$ the three-momentum vector, are considered. The distributions of the three variables for signal and background jets are plotted in figures 6.13(a), 6.13(b) and 6.13(c) respectively.

Subjet mass and fat jet η

The reconstructed masses of the two leading subjets m_{j_1} and m_{j_2} and the pseudorapidity η of the fat jet allow for some signal-background discrimination as well. Their distributions are plotted in figures 6.13(d), 6.13(e) and 6.13(f) respectively. It can be seen that the agreement between data and background MC shows some differences in the distributions for m_{j_1} and m_{j_2} . Since the discrepancies look like they result from a deficit of events in data in the region occupied by signal MC (which was excluded in data), we can suspect that both variables are strongly correlated to the filtered jet mass. Interestingly, the distribution of fat jets in η shows a slight asymmetry. There is no physical reason for this and it can therefore be attributed to detector effects like holes in the calorimeter, dead material and detector inefficiencies. We investigate this further in Appendix E. In figure E.2 the mirrored distribution is plotted on top of the regular distribution to once again show the asymmetry. Figure E.3 shows the distribution of fat jets throughout the η - ϕ plane.

6.6 MVA training and performance

6.6.1 MVA training

Multivariate data analysis techniques like Fisher's linear discriminant, boosted decision trees (BDT) or artificial neural networks (ANN) are a useful aid to decide the affiliation

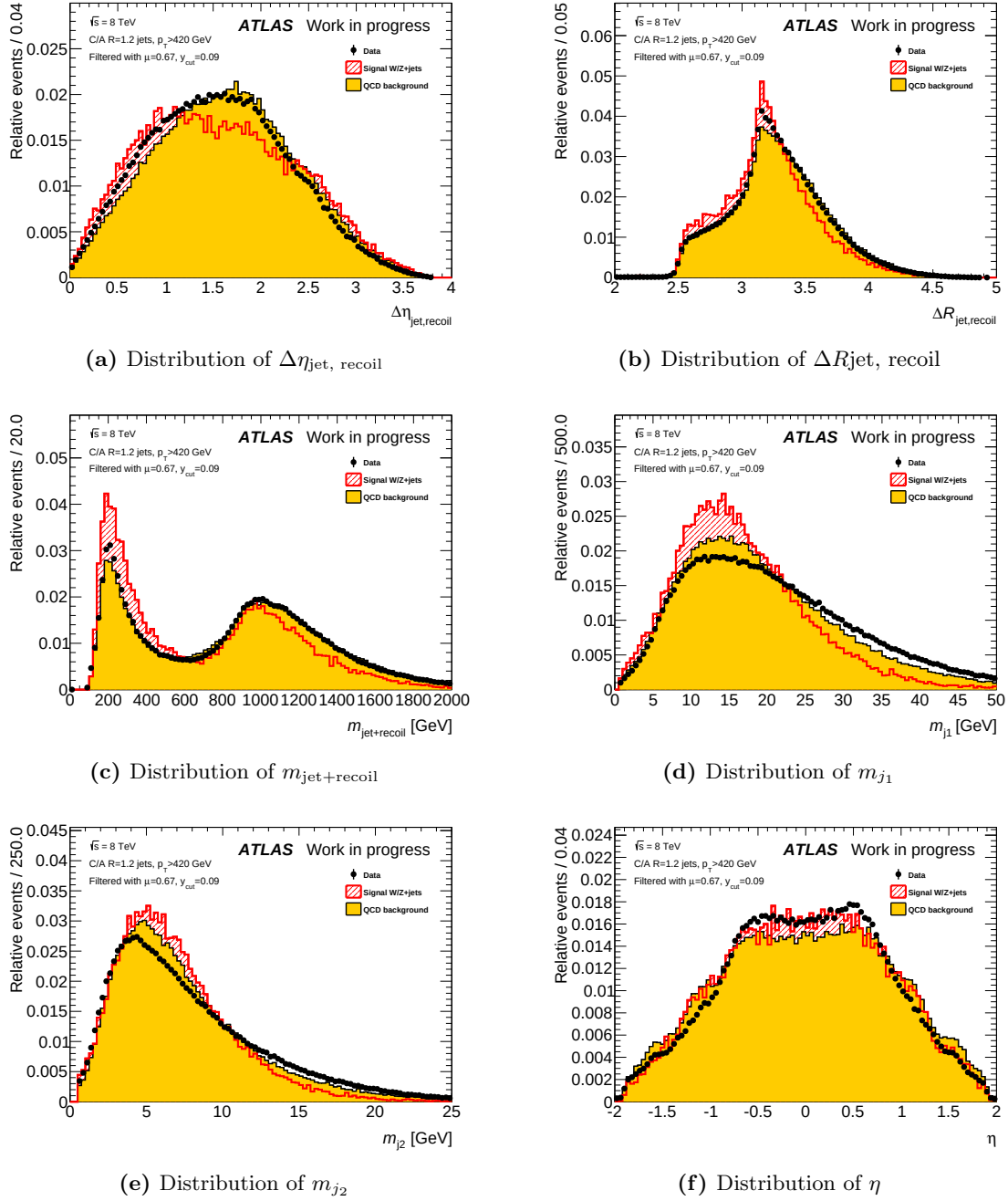


Figure 6.13: Distributions for signal and background MC as well as for data with the signal region excluded of (a) the distance in pseudorapidity between the fat jet and the recoiljet, (b) the distance in ΔR between the fat jet and the recoiljet, (c) the invariant mass of the combined system fat jet and recoiljet, (d) the reconstructed mass of the leading subjet, (e) the reconstructed mass of the sub-leading subjet and (f) the pseudorapidity value of the fat jet.

of a sample to an origin distribution. In our case, of a jet to signal or background. By giving one response value to a random event, it allows us to select a custom fraction of the desired distribution by cutting on the response value. The variables discussed above are input variables for the training of several multivariate data analysis techniques. Roots own multivariate analysis package TMVA [196] is used for the training. In this analysis we pursue the approach of training two separate same flavour MVA methods:

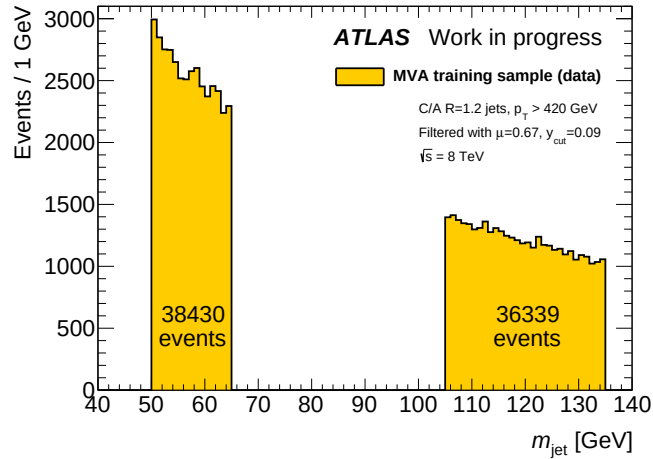


Figure 6.14: Mass spectrum of the filtered fat jets used for MVA training in data. The signal region $65.0 \text{ GeV} < m_{jet} < 105.0 \text{ GeV}$ has been excluded. The widths of the two sidebands were chosen to get an equal number of jets in each sideband.

One in which signal MC is trained against background MC, and a second in which data is used as the background sample. Since data also includes signal we veto the signal region in the training sample and only use the sidebands of $50.0 \text{ GeV} < m_{jet} < 65.0 \text{ GeV}$ and $105.0 \text{ GeV} < m_{jet} < 135.0 \text{ GeV}$. The widths of the sidebands are chosen on purpose to allow for an equal number of jets in either band. This can be seen in figure 6.14, where the masses of the selected W/Z candidate jets in the data background training sample are plotted. The whole signal MC sample has been used for training, which amounts to 82564 events, while only a portion corresponding to 442587 events of the background sample is used.¹⁷ Throughout the analysis the two MVA's are labelled MVA(MC) and MVA(DATA), where the brackets indicate the type of background used for the training of the method.

The indicator of the performance of the MVA method is the ROC-curve in which the signal efficiency, that is the fraction of the signal that passes a cut on the MVA response value, is plotted versus the background rejection, the fraction of background events that do not pass the cut. The better the separation between the signal and background distributions the more the curve moves into the corner of the plot. A straight line going from zero signal efficiency and 100% background rejection to 100% signal efficiency and zero background rejection, on the other hand, shows that the two distributions are identical and no separation can be achieved. In figure 6.15 the ROC-curves for both MVA types are plotted. Five methods have been chosen to show the various performances: Three boosted decision trees with adaptive boost (BDT), gradient boost (BDTG) and one in which a Fisher discriminant is used for node splitting (BDTMitFisher), an artificial neural network (MLP) and Fisher's linear discriminant (Fisher). Comparing the two types reveals that MVA(DATA) yields better results than MVA(MC). This can be explained by two reasons which both show that the MVA(DATA) is not fundamentally superior to the MVA(MC), but rather more susceptible to irrelevant influences. First, the distributions of variables like the separation of the two leading subjets show much less overlap when signal MC is compared to data than when it is compared to background MC. And second

¹⁷It might be noted that the large differences in sample size between MC signal and background are intentional. Referring to figure 6.2 we have to remind ourselves that a large portion of the MC background sample is produced at very high p_T and therefore carries a very small event weight. The sample sizes are therefore more symmetric in the region of interest as it appears by looking at the number of events only.

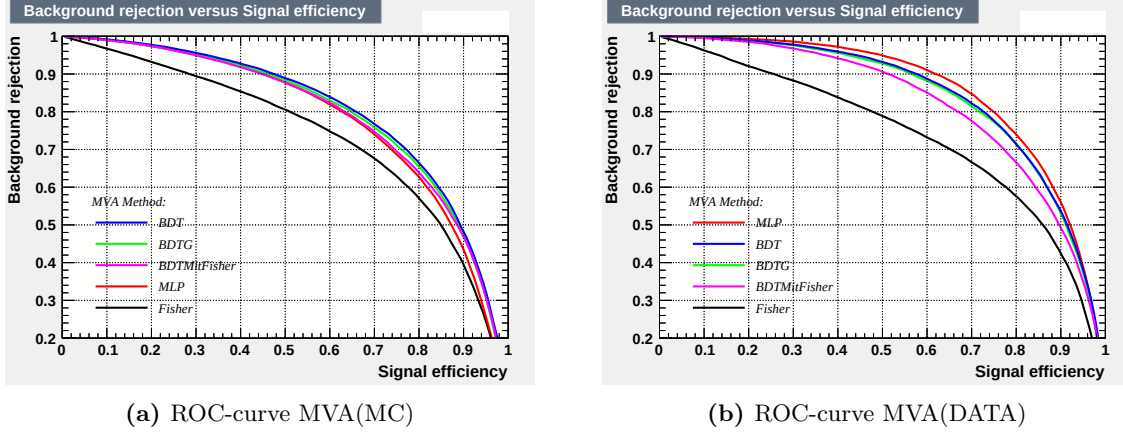


Figure 6.15: Signal efficiency versus background rejection for both MVA types and several MVA methods. The trained methods are three boosted decision trees with adaptive boost (BDT), gradient boost (BDTG) and one in which a Fisher discriminant is used for node splitting (BDT-MitFisher), an artificial neural network (MLP) and Fisher’s linear discriminant (Fisher). For MVA(MC) signal MC was trained against background MC, for MVA(DATA) data sidebands have been used instead of background MC.

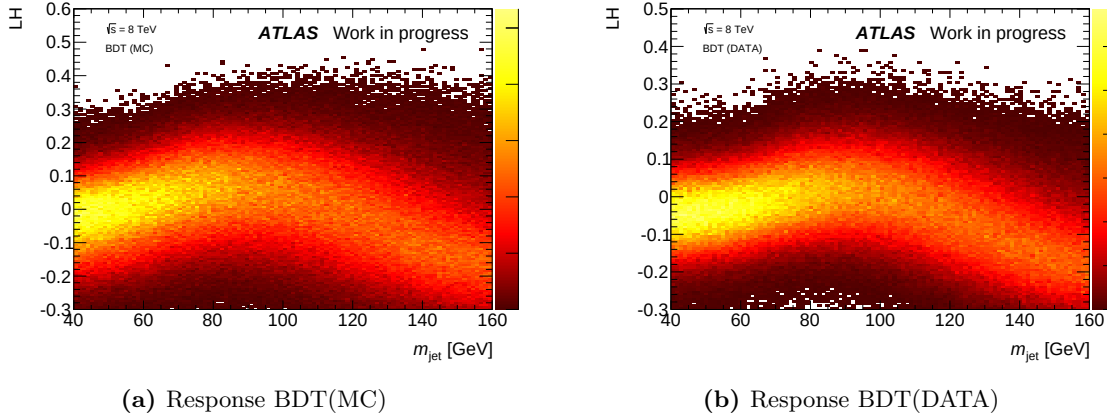


Figure 6.16: BDT response as a function of the filtered jet mass in data for (a) BDT(MC) and (b) BDT(DATA).

it is due to imperfectly modelled shower behaviour in MC. This means in addition to training on the differences in the shower shape of signal and background jets, the MVA also trains on the deficiencies of MC to represent real data. As the agreement of MC and data was not specifically studied in the present work and can only be guessed from the figures in the previous chapter, we are compelled to trust the MVA(MC) more. Looking at which method performs best, we assert that both types agree on the order with the exception of the ANN. While it can’t compete with the BDT methods in the case of MVA(MC), it is the best performing method for MVA(DATA). Because of the reasons stated above we use boosted decision trees from here on, since they perform very well for both types. The BDT trained solely on MC will henceforth be called BDT(MC), the one in which sideband data was used as background sample is labelled BDT(DATA). A detailed description of the working principle of boosted decision trees can be found in Appendix C. The parameters for the BDT training in this specific analysis involved 850 trees with a required minimum number of 150 events per leaf node and adaptive boosting.

The maximum allowed depth of the BDT was set to 3. To find the optimal node splitting cut the Gini-index has been chosen as separation criterion with a minimum number of 20 grid points in variable range. No pruning, i.e. the removal of insignificant BDT branches was allowed in the training.

6.6.2 Removing correlations between the BDT and m_{jet}

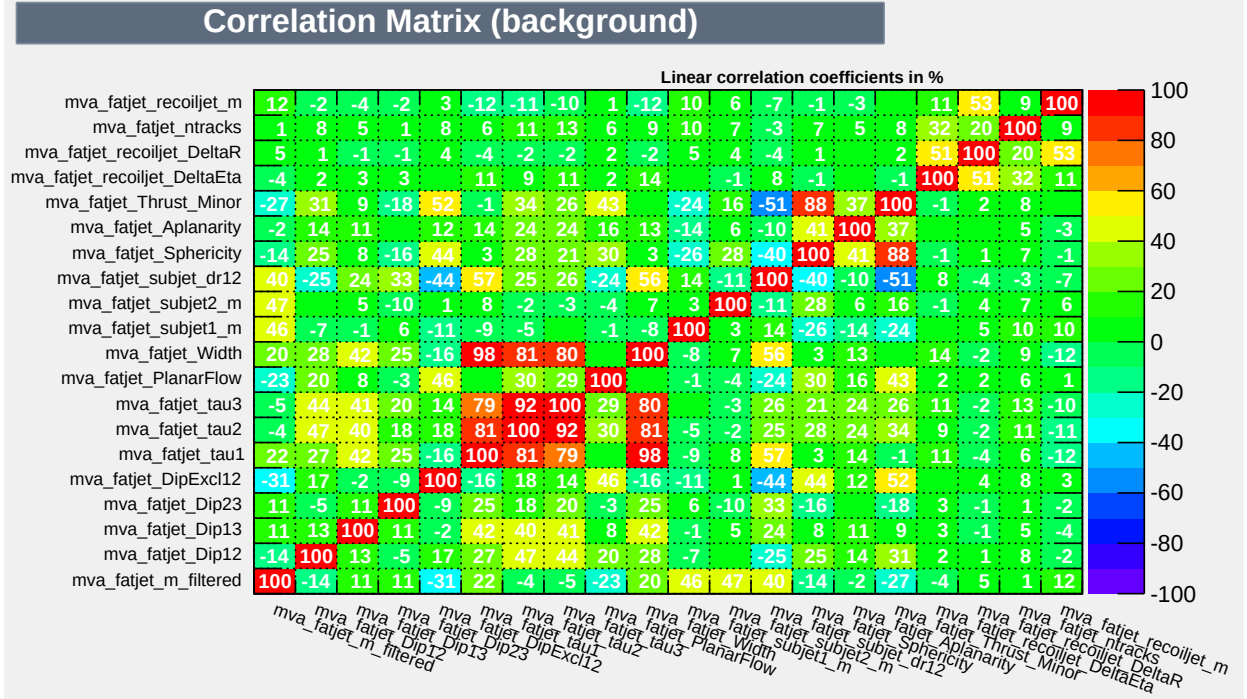
Plotting the BDT response in data as a function of the filtered jet mass as it was done in figure 6.16(a) for BDT(MC) and in 6.16(b) for BDT(DATA) reveals a grave problem. While we would like to have the jet mass and the BDT response to be completely independent, this is obviously not the case. A maximum can be observed around the $W^\pm$ and Z^0 masses. Instead of reducing the background in the whole mass range a cut on the likelihood would reject almost all the background in the sidebands while keeping nearly the full background in the signal region. So besides not gaining anything through the application of the BDT, it would make the fitting even more challenging since it would be necessary to fit a small peak on top of a larger peak. This is obviously not the way to go.

The dependence of the response signal on the filtered jet mass was already introduced by the input variables. If an input variable exhibits some correlation to the filtered jet mass it means that a specific value of the input variable is more likely to result in a jet with a mass in the signal region. The BDT is therefore trained to select jets with masses consistent with the $W^\pm$ and Z^0 masses for signal and background alike.

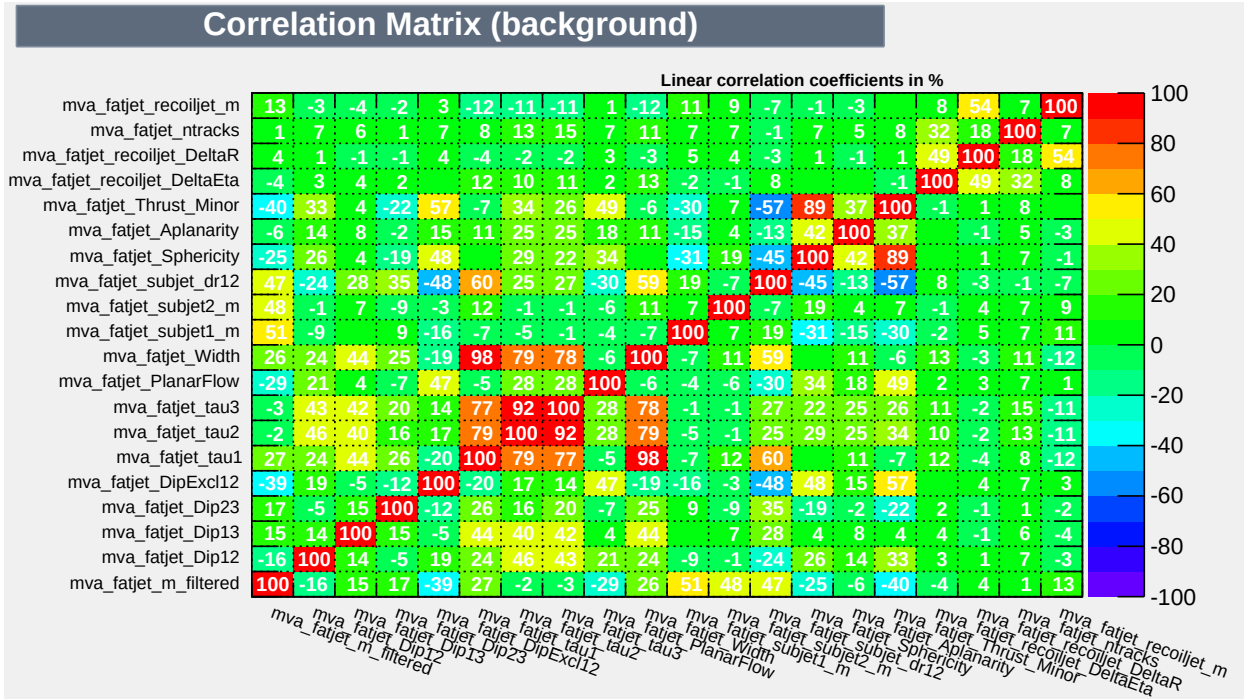
Figure 6.17 shows the correlation matrix for the input variables and the filtered jet mass in the background¹⁸ distribution. The numbers mentioned in the plot are linear correlation coefficients, i.e. the higher the absolute value the stronger the correlation between two variables. The plot reveals some interesting details. For example the three N-subjettiness variables and the jet width are strongly correlated and dropping one or several of these variables would barely influence the efficiency of the BDT. We can see as well that the event shape variables are correlated to each other. For example thrust minor and sphericity are correlated 88% (89%) for BDT(MC) (BDT(DATA)).

The first column in figure 6.17 shows the correlations to the filtered jet mass. Very weakly correlated variables are the jet pseudorapidity, aplanarity, the colour flow variables (with the exception of $\mathcal{D}_{12}^{excl}$), N-subjettiness for $N = 2, 3$ and the variables considering the recoiljet. The problematic variables are the ones that are strongly correlated to the filtered jet mass. Several of this type can be identified: foremost, as was already suspected, the separation between the two leading subjects and their masses. The jet dipolarity variable $\mathcal{D}_{12}^{excl}$ also exhibits some significant correlation. The remaining variables are neither uncorrelated nor strongly correlated, yet they are decisive. While of course the easiest way to get rid of the correlations is to drop the most correlated variables, we do not want to drop too many variables, as this also results in a reduction of the BDT efficiency. But we have to drop some of them to make the correlations manageable. A good place to start is the thrust minor variable. As it is correlated with sphericity we do not lose much efficiency by dropping it. The same argument holds for the jet width. Planar flow not only shows a significant correlation but is also very ineffective in distinguishing signal and background, which makes the variable another ideal candidate for dropping. It is also important to once again note that the matrix only gives linear correlation coefficients and does not incorporate higher order correlations. In figure 6.18(a) the linearly correlated subjet separation is plotted against the filtered jet mass in MC. Clearly higher jet masses result in proportionally larger values for the distance between the two leading subjects.

¹⁸since our sample is background dominated and we want a smooth background, the background correlations are the ones we are interested in.



(a) Correlation matrix BDT(MC)



(b) Correlation matrix BDT(DATA)

Figure 6.17: Linear correlation coefficients in the background distributions for (a) BDT(MC) and (b) BDT(DATA).

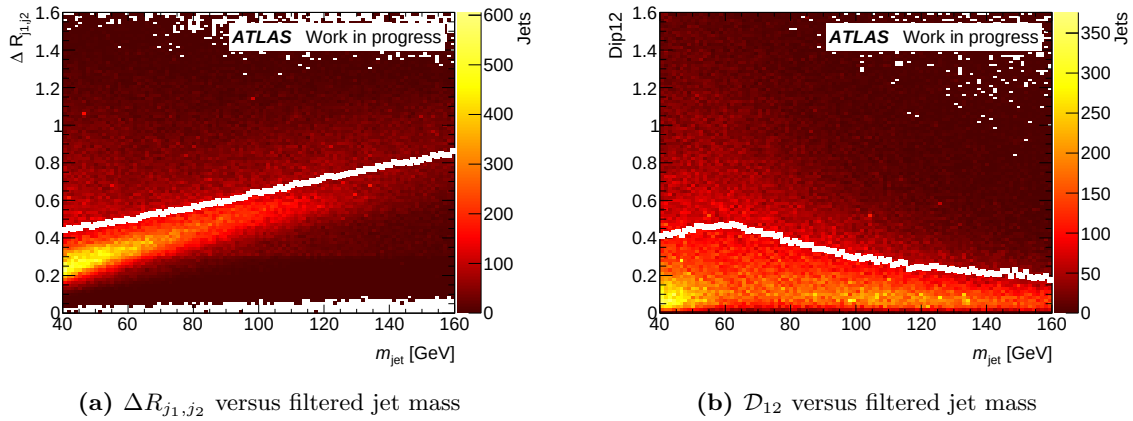


Figure 6.18: (a) $\Delta R_{j_1,j_2}$ and (b) $\mathcal{D}_{12}$ plotted against the filtered jet mass in MC. The white band represents the profile of the plot.

This was evident from the correlation matrix where this variable exhibits a correlation coefficient of 40%. We compare this to figure 6.18(b) in which the same plot is shown for $\mathcal{D}_{12}$ which has a linear correlation coefficient of -14% . The profile is far from flat as could have been expected by the small correlation coefficient. Considering all the correlation scatter plots which are included in Appendix D and the small discrimination power of the variables one comes to the conclusion to drop all jet dipolarity variables.

As a summary table 6.5 lists all variables, their linear correlation coefficients for both BDTs and a remark whether the variable was kept or dropped. Out of the original 20 only 10 variables were used for the final version of the BDT that was trained with the parameters kept unchanged. The performance and characteristics of the reduced BDT are summarized in figure 6.19. The ROC-curves are shown in figure 6.19(a) and 6.19(b). The lines in the plot at 50% signal efficiency roughly indicate the point at which we will be working later on in this study. Naturally, the efficiency of the reduced BDTs decreased. Where at 50% signal efficiency the background rejection rate for the top performing BDT method was 89% (93%) it has now dropped to 87% (88%) for BDT(MC) (BDT(DATA)). While the background rejection was reduced noticeably in the BDT(DATA) it decreased only insignificantly for BDT(MC). Furthermore, we observe that the ROC-curves look

	ρ_{mc}^{lin}	ρ_{data}^{lin}		ρ_{mc}^{lin}	ρ_{data}^{lin}	
$m_{jet+recoil}$	12%	13%	Width	20%	26%	dropped
N_{trk}	1%	1%	PlanarFlow	-23%	-29%	dropped
$\Delta R_{jet,recoil}$	5%	4%	τ_3	-5%	-3%	
$\Delta \eta_{jet,recoil}$	-4%	-4%	τ_2	-4%	-2%	
Thrust minor	-27%	40%	τ_1	22%	27%	
Aplanarity	-2%	-6%	$\mathcal{D}_{12}^{excl}$	-31%	-39%	dropped
Sphericity	-14%	-25%	$\mathcal{D}_{23}$	11%	17%	dropped
$\Delta R_{j_1,j_2}$	40%	47%	$\mathcal{D}_{13}$	11%	15%	dropped
m_{j_2}	47%	48%	$\mathcal{D}_{12}$	-14%	-16%	dropped
m_{j_1}	46%	51%	η_{jet}	0%	0%	

Table 6.5: Overview of all input variables with their linear correlation coefficients ρ^{lin} to the filtered jet mass in percent listed for both types of BDT. Dropped variables were not used for the training of the final version of the BDT.

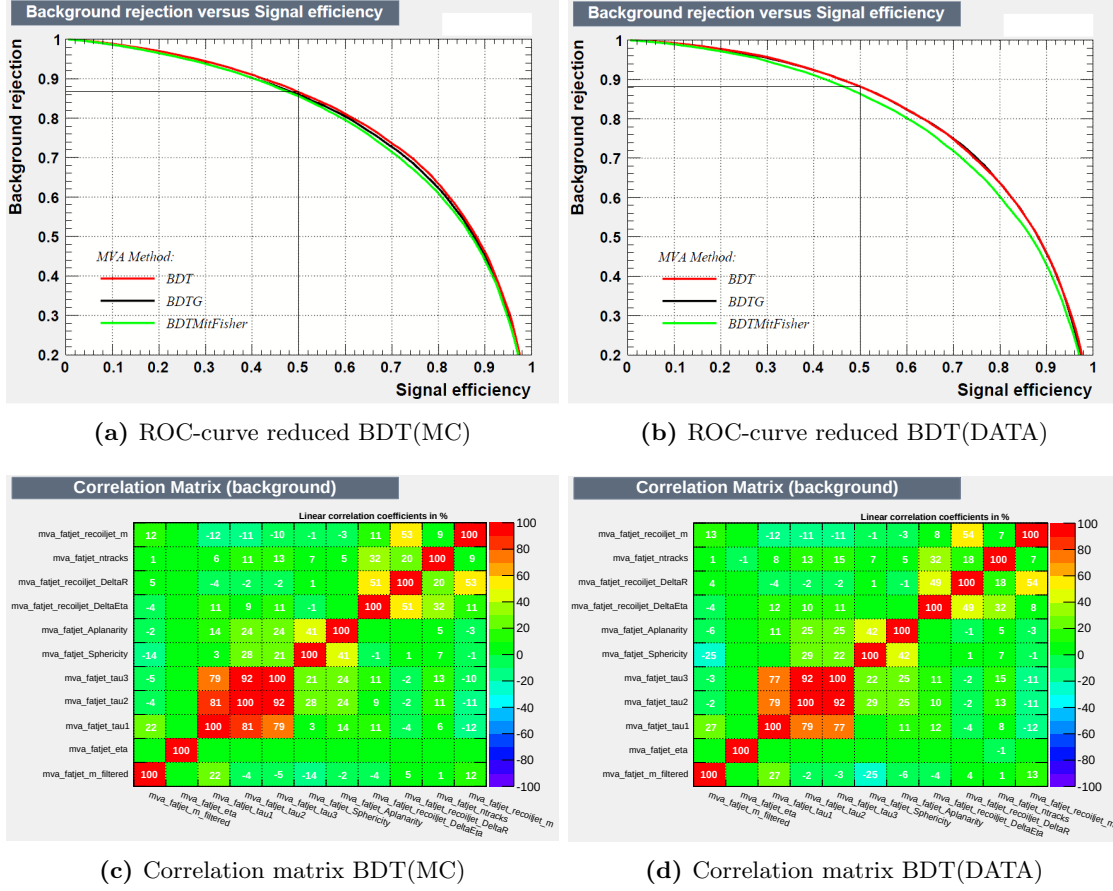


Figure 6.19: Performance of the reduced BDTs. The signal efficiency is plotted against the background rejection for (a) BDT(MC) and (b) BDT(DATA). In (c) and (d) the linear correlation matrix is drawn for BDT(MC) and BDT(DATA) respectively.

much more alike. This increases the trustworthiness of the BDT(DATA) and helps to rehabilitate it. The new correlation matrices for background MC are shown in figures 6.19(c) and 6.19(d) for BDT(MC) and BDT(DATA), respectively. Again the filtered jet mass was added to the matrix and their correlations can be read off in the first column. Event though some correlations remain, the large contributors are gone.

In both cases the BDT with adaptive boosts shows the highest performance. Henceforth we will only consider this BDT method.

Finally, the distributions of the BDT response values for signal and background events show the separation that was achieved through the application of the BDT. They are depicted in figures 6.20(a) and 6.20(b) for BDT(MC) and BDT(DATA) respectively. We observe that the agreement between data and background MC is good for BDT(MC). In the case of the BDT(DATA) the distribution of data seems to be shifted with respect to the background MC one. The shift is probably due to insufficiencies of the MC modelling and could in fact be used to quantify the impact of the imperfect MC modelling on the analysis. The figures also illustrate how a cut on the BDT response value can select a certain fraction of the signal events. In this study the cuts are chosen to keep around 43% of the signal events. The rough position of the cut is marked by a vertical dashed line

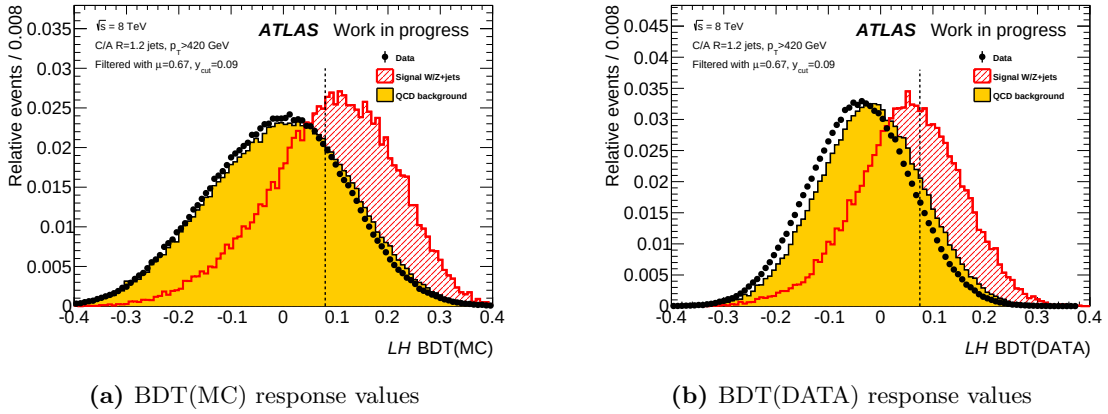


Figure 6.20: The distribution of the BDT response values for signal and background events are plotted in (a) for BDT(MC) and (b) for BDT(DATA). For the data the signal region has been vetoed.

in both plots.¹⁹ Both BDTs have been checked for overtraining. A Kolmogorov-Smirnov test [197] gives the probabilities P^{KS} of the test and training sample originating from the same distribution. It returns

$$\begin{array}{lll}
 \text{BDT(MC):} & P_S^{KS} = 89.9\% & P_B^{KS} = 94.9\% \\
 \text{BDT(DATA):} & P_S^{KS} = 49.1\% & P_B^{KS} = 75.0\%
 \end{array}$$

where the index S and B correspond to signal and background distributions. No indication of overtraining was found for BDT(MC). BDT(DATA) shows a comparatively small probability for the signal distributions but we will consider it nevertheless overtraining free.

It is now possible to again plot the BDT response values LH as a function of the filtered jet mass. This has been done in figure 6.21(a) for the reduced BDT(MC) and in figure 6.21(b) for the reduced BDT(DATA). Comparing these to figure 6.16 we can assert that the curvature of the diagram and therefore the correlation has decreased as was our objective. Still the response is far from uniform and a cut on the BDT response would result in the creation of a peak even in the continuous background spectrum. The reduced correlations, however, open up a way to salvage the situation by correcting the BDT response values for the correlation.

In order to do that, the profile of the two scatter plots is fitted. Looking at the plot we can identify three distinguished regions in the plot. To the left and to the right of the signal region the BDT response shows a linear behaviour. The peak is located within the signal region. Therefore, the fit function

$$\mathcal{C}(m_{\text{jet}}) = \begin{cases} A_1 + A_2 \cdot m_{\text{jet}} & \text{for } m_{\text{jet}} \leq m_1 \\ A_3 + A_4 \cdot m_{\text{jet}} + A_5 \cdot m_{\text{jet}}^2 & \text{for } m_1 < m_{\text{jet}} < m_2 \\ A_6 + A_7 \cdot m_{\text{jet}} & \text{for } m_{\text{jet}} \geq m_2 \end{cases} \quad (6.19)$$

is chosen. Since the function and its first derivative have to be continuous at the sewing points m_1 and m_2 , the number of free fit parameters reduces from 9 to 5, where m_1 and

¹⁹The values marked here are not the actual cut values, as a correction of the BDT response values becomes necessary later on which of course influence the distributions. They are only plotted to show the concept.

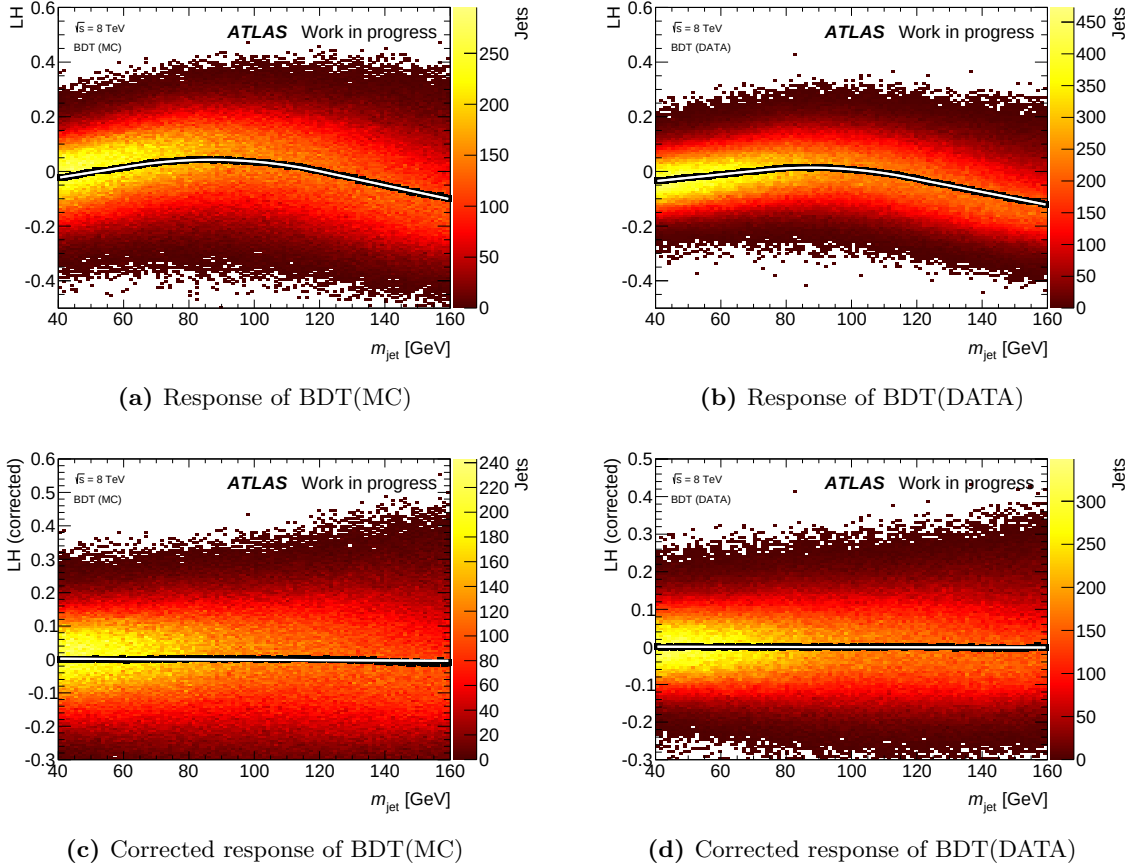


Figure 6.21: Response values of the (a) reduced BDT(MC) and (b) reduced BDT(DATA) plotted as a function of the filtered jet mass. The profile of the plots is indicated by black squares for every column. The white line is the fit function to the profile which is used to correct the response for the correlation to the filtered jet mass. (c) and (d) show the BDT response after application of the correction for BDT(MC) and BDT(DATA) respectively. The linear fit functions to the profile exhibit a negligible slope.

m_2 are free parameters. The fit functions are drawn as white lines on top of the black diagram profile in figures 6.21(a) and 6.21(b). The BDT responses can now be corrected by subtracting the value of the fit function at the corresponding jet mass from the original BDT response values on a jet-by-jet basis. The BDT response after the application of the correction is plotted in figures 6.21(c) and 6.21(d) for BDT(MC) and BDT(DATA) respectively. The response is now much more uniform as indicated by the profile and linear fit function and a cut applied on the BDT response value will remove events in the whole filtered jet mass range instead of selecting a single region.

For the remaining analysis we require a corrected BDT response value of $LH > 0.08$ for the reduced BDT(MC) and of $LH > 0.075$ for the reduced BDT(DATA). The numbers were chosen to get a comparable signal efficiency of 43%. The number of events passing these cuts on the BDT response are listed in table 6.6 which can be seen as the conclusion of table 6.4.

6.7 Fitting

6.7.1 Development of a fit model

Before unblinding the data we want to compare the MC prediction with the data in the sideband regions. We can do this in four different scenarios. The first two show the fraction of data passing the cut on the BDT response values. They are plotted in figures 6.22(a) for BDT(MC) and 6.22(b) for BDT(DATA). Second, it is also illuminating to reverse the cut on the BDT response and show the fraction of data that does not pass the cut, i.e. $LH(MC) < 0.08$ and $LH(DATA) < 0.075$. The plots can be found in figures 6.22(c) and 6.22(d). For all plots background MC has been normalized individually to match the data in the sidebands. In order to do this, it was necessary to scale the MC prediction by factors of 0.52 and 0.45 for BDT(MC) and BDT(DATA) respectively. This is due to missing NLO corrections. Since data in the sideband regions is overwhelmingly dominated by background the error we make in doing so is negligible small. For signal MC we have no way of finding the right scale without using unblinded data. Therefore it is plotted here without any NLO corrections, i.e. with scale one. It serves to get a feeling on the expected order of magnitude and shape of the signal.

The ratio of data and MC can be judged from the lower panels in all four figures. What strikes us immediately is the good agreement for the plots satisfying the BDT cut. This is not at all to be expected, especially after comparing it to figure 6.3(b) in which the agreement before the BDT cut is plotted. Since this looks like the agreement is just happening by pure chance we conduct a simple experiment to study the trustworthiness of MC in the plots fulfilling the BDT cut. The BDT cut value is varied and the influence on the agreement observed. While lowering the cut value leads to a degeneration in agreement, a moderate increase in contrast, has very little effect on the agreement. We should also take into account that the plots for both BDTs are independent and can be considered as two separate experiments. If the good agreement is by chance, we were lucky twice.

Another indication that this is not the case we do get by comparing figure 6.3(b) to the lower plots in figure 6.22 which are failing the BDT cut. We can observe a very similar behaviour in the shapes of data and background MC in all three cases. The amount of jets with small masses is underestimated in MC, while it is overestimated for masses in the range $100 \text{ GeV} < m_{\text{jet}} < 140 \text{ GeV}$.

The observations indicate a simple explanation for the remarkable agreement between data and MC: the presence of a certain type of background-associated jet that is not well described in MC but is luckily assigned a small BDT response value. Since this jet would be largely rejected by a cut, the agreement is good above the cut value. Decreasing the cut allows more mismodeled jets into the sample and results in a decrease of the agreement.

Having a satisfying explanation for the startling conformance, we can exploit it by using background MC to develop a model to describe the background in both measurements.

We are satisfied with using a fifth degree polynomial. It does show some misbehaviour

	signal MC			background MC			data		
Before BDT	6278	26.9%	100.0%	1434939	12.1%	100.0%	631734	12.5%	100.0%
BDT(MC)	2706	11.6%	43.1%	321087	2.7%	22.4%	156535	3.1%	24.8%
BDT(DATA)	2680	11.5%	42.7%	332775	2.8%	23.2%	145065	2.9%	23.0%

Table 6.6: The number of events passing the cut on the BDT response values. The table can be seen as a continuation of table 6.4. The first percentage refers to the total number of events before application of any cut, the second percentage serves to estimate the BDT efficiency.

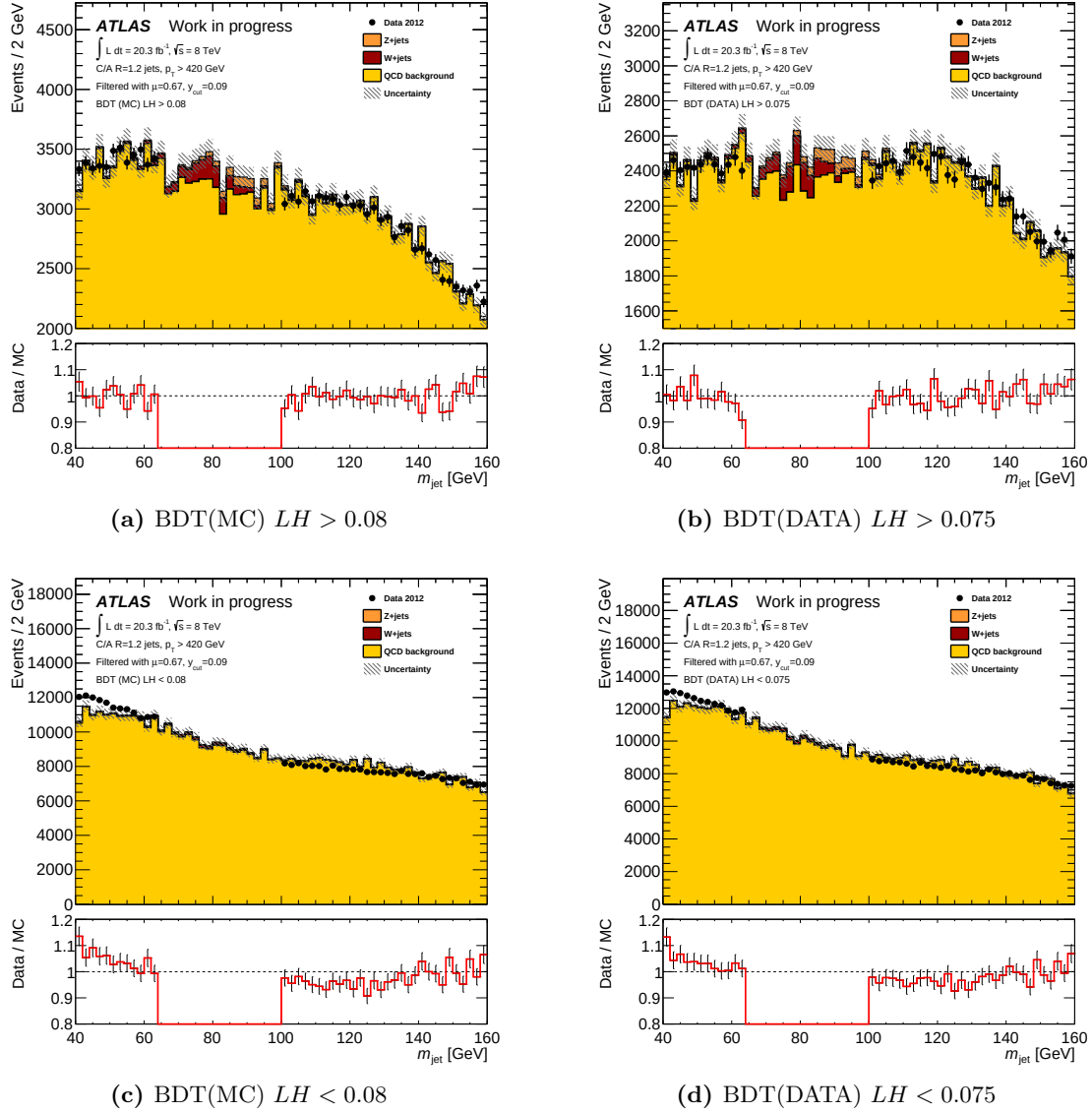


Figure 6.22: Comparison between data and MC in a blinded study. For (a) and (b) a cut on the BDT response of $LH > 0.08$ for BDT(MC) and of $LH > 0.075$ for BDT(DATA) has been applied. In figures (c) and (d) the cut has been reversed. They show the fraction of data that has been rejected by the BDT cut. In all plots background MC has been normalized to data individually. Signal MC is taken uncorrected and is therefore unreliable in scale. The lower panels show the ratio between data and MC.

at the very edges of the plot but fits very nicely in the signal region. Figures 6.23(a) and 6.23(b) show the now unblinded mass spectra for BDT(MC) and BDT(DATA). In both cases a clear peak at a mass consistent with the $W^\pm$ and Z^0 mass is observed. The background fit is drawn as a red line and we can assert that it fits well with data in the sidebands. The lower panel again shows the ratio between data and MC. In the signal region it is noticeably larger than one.

To fit the data over the whole mass range we expand the background fit $\mathcal{B}$ by adding the sum of two Gaussians on top, each representing one of the vector bosons. However, from figure 6.24 we see that this is not enough. The signal exhibits wide tails on both sides of the main peak that are not fitted by the double-Gaussian. One way of accounting

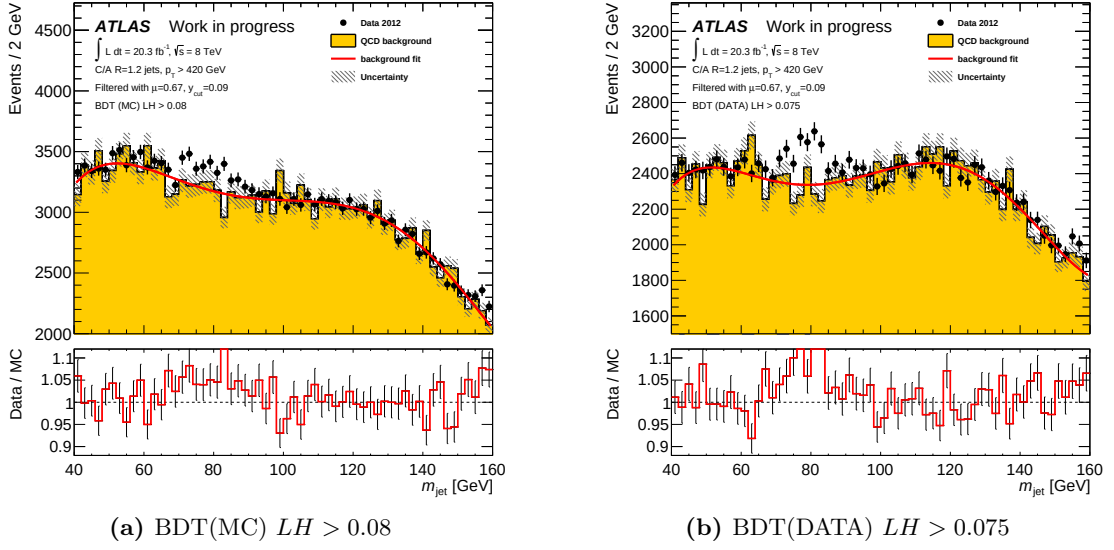


Figure 6.23: Mass spectrum after cutting on the BDT response value for (a) BDT(MC) and (b) BDT(DATA). A peak at masses consistent with the $W^\pm$ and Z^0 mass is observed. The background is estimated by fitting on MC and shown as red line. The lower panels show the ratio between data and MC.

for it, is to add a straight line to the signal fit. That this yields reasonable results can be judged from figure 6.24 where signal MC was fitted with this model.

The plain fit function of the form

$$N(m_{\text{jet}}) = \mathcal{B} + \frac{S_w}{\sqrt{2\pi}\sigma_w} \exp\left[-\frac{1}{2}\left(\frac{m_{\text{jet}} - \mu_w}{\sigma_w}\right)^2\right] + \frac{S_z}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{1}{2}\left(\frac{m_{\text{jet}} - \mu_z}{\sigma_z}\right)^2\right] + (C_l - M_l \cdot m_{\text{jet}}) \quad (6.20)$$

exhibits eight fit parameters. Each Gaussian is described by scale S_x , width σ_x and mean μ_x for $x = w, z$ and the line accounts for an additional two parameters C_l and M_l . However, the number of fit parameters can be reduced drastically by constraining the corresponding parameters for $W^\pm$ and Z^0 to each other. Just like it was done in chapter 6.4 we can estimate the relative scale S_{wz} of the two peaks from MC. As we have seen before, MC will most likely get the absolute normalization of the peaks wrong, but it will exhibit the same discrepancy for $W^\pm$ and Z^0 peak. So instead of using two parameters S_w and S_z , the same situation can be described by only one overall normalization parameter S , where the individual scales of the peaks are constrained to s_{wz} for the $W^\pm$ and $(1 - s_{wz})$ for the Z^0 peak. s_{wz} is determined from a fit on signal MC just like in figure 6.24. For BDT(MC) one finds $s_{wz} = (0.69 \pm 0.04)$. For BDT(DATA) the computed ratio is given by $s'_{wz} = (0.67 \pm 0.05)$

The analogous argument holds for mean and width of the peaks. Especially for the mean we know that the mass-drop filtering will get a lower and therefore wrong value. But it will preserve the ratio between the means of the two distributions which makes it unnecessary to fit using two independent parameters. For BDT(MC) we find from MC:

$$\begin{aligned} \mu_w &= (78.43 \pm 0.22) \text{ GeV} & \sigma_w &= (7.19 \pm 0.23) \text{ GeV} \\ \mu_z &= (88.66 \pm 0.36) \text{ GeV} & \sigma_z &= (7.82 \pm 0.39) \text{ GeV} \end{aligned}$$

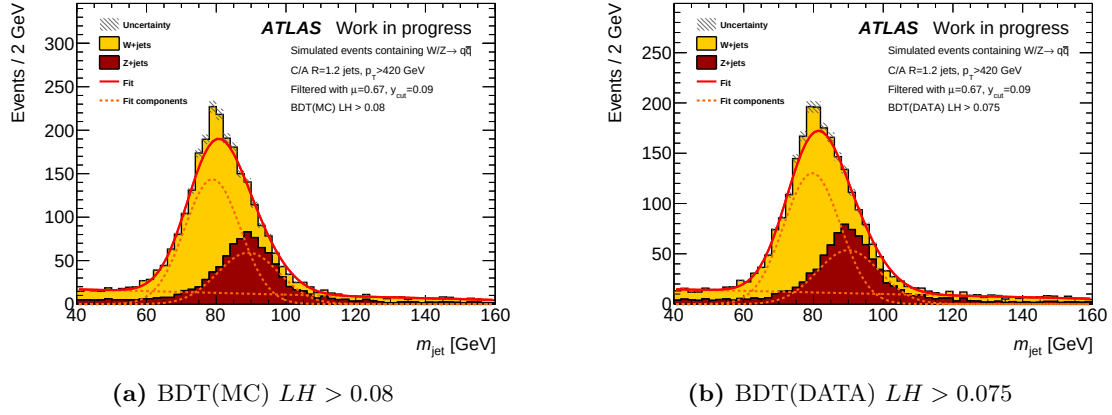


Figure 6.24: Signal MC fitted with the sum of two Gaussians for the central peak and a straight line to describe the tails of the distribution. The individual components of the fit are shown as dashed lines.

From those values we can calculate the ratios

$$\begin{aligned}\mu_{wz} &= \mu_z / \mu_w = (1.1304 \pm 0.0055) \quad \text{and} \\ \sigma_{wz} &= \sigma_z / \sigma_w = (1.089 \pm 0.065) \quad .\end{aligned}$$

In the case of BDT(DATA) we get very similar values:

$$\begin{aligned}\mu'_w &= (78.89 \pm 0.23) \text{ GeV} & \sigma'_w &= (7.24 \pm 0.25) \text{ GeV} \\ \mu'_z &= (89.44 \pm 0.36) \text{ GeV} & \sigma'_z &= (7.68 \pm 0.40) \text{ GeV}\end{aligned}$$

This results in the ratios

$$\begin{aligned}\mu'_{wz} &= (1.1337 \pm 0.0057) \quad \text{and} \\ \sigma'_{wz} &= (1.060 \pm 0.066) \quad .\end{aligned}$$

Out of curiosity we compute the mass ratio again, this time using literature values from [16]. Here we find $\tilde{\mu}_w = (80.385 \pm 0.015) \text{ GeV}$ and $\tilde{\mu}_z = (91.1876 \pm 0.0021) \text{ GeV}$. Using literature values we get $\tilde{\mu}_{wz} = \tilde{\mu}_z / \tilde{\mu}_w = (1.1344 \pm 0.0002)$ which is in very nice agreement to $\tilde{\mu}_{wz}$ and $\mu_{wz}^{(l)}$. This indicates a linear relation between the reconstructed filtered jet mass and the truth jet mass in the signal region. This we have already shown in figure 5.5(a). For the fit the values obtained in MC are used.

Lastly, we fix the parameters of the straight line, since we expect the ratio between the Gaussian peaks and the tails to be constant. Fixing all ratios in the final fit reduces the number of floating parameters from eight to three.

In having scaled background MC to sideband data and fitting the background in MC we have automatically attributed all sideband events to the background. This is of course incorrect. We want to mildly relax the background fit. Therefore, the constant part of the polynomial fit is kept as a floating parameter, while all others are fixed to the values obtained in the fit on background MC.

The final fit function is given by

$$\begin{aligned}N(m_{\text{jet}}) = \mathcal{B}(A_0) + S \left\{ \frac{S_{wz}}{\sqrt{2\pi}\sigma_w} \exp \left[-\frac{1}{2} \left(\frac{m_{\text{jet}} - \mu_w}{\sigma_w} \right)^2 \right] + \right. \\ \left. \frac{(1 - S_{wz})}{\sqrt{2\pi}\sigma_w \cdot \sigma_{wz}} \exp \left[-\frac{1}{2} \left(\frac{m_{\text{jet}} - \mu_w \cdot \mu_{wz}}{\sigma_w \cdot \sigma_z} \right)^2 \right] + C_l - M_l \cdot m_{\text{jet}} \right\} \quad (6.21)\end{aligned}$$

Var.	Description	Usage	BDT(MC)	BDT(DATA)
A_0	0th order polynomial coeff.	floating	$(-7.79 \pm 0.01) \cdot 10^3$	$(-6.21 \pm 0.01) \cdot 10^3$
A_1	1st order polynomial coeff.	fixed	$6.62 \cdot 10^2$	$5.33 \cdot 10^2$
A_2	2nd order polynomial coeff.	fixed	$-1.47 \cdot 10^1$	$-1.24 \cdot 10^1$
A_3	3rd order polynomial coeff.	fixed	$1.53 \cdot 10^{-1}$	$1.36 \cdot 10^{-1}$
A_4	4th order polynomial coeff.	fixed	$-7.56 \cdot 10^{-4}$	$-7.02 \cdot 10^{-4}$
A_5	5th order polynomial coeff.	fixed	$1.41 \cdot 10^{-6}$	$1.37 \cdot 10^{-6}$
S	Signal normalization	floating	$(3.98 \pm 0.57) \cdot 10^3$	$(3.86 \pm 0.43) \cdot 10^3$
μ_w	$W^\pm$ mass	floating	$(7.69 \pm 0.11) \cdot 10^1$	$(7.73 \pm 0.09) \cdot 10^1$
σ_w	$W^\pm$ detector resolution	floating	5.69 ± 1.28	4.69 ± 0.73
S_{wz}	Scale-ratio of $W^\pm$ & Z^0	fixed	$6.87 \cdot 10^{-1}$	$6.70 \cdot 10^{-1}$
σ_{wz}	Width-ratio of $W^\pm$ & Z^0	fixed	1.11	1.06
μ_{wz}	Mass-ratio of $W^\pm$ & Z^0	fixed	1.13	1.13
C_l	Straight line offset	fixed	$5.01 \cdot 10^{-3}$	$4.65 \cdot 10^{-3}$
M_l	Straight line slope	fixed	$2.42 \cdot 10^{-5}$	$2.03 \cdot 10^{-5}$

Table 6.7: Summary of all the fit variables with a short description, usage of the parameter and its value for both BDT cases. Fixed parameters were fixed to the MC prediction.

where only A_0 , S , σ_w and μ_w are left floating.

A summary of all fit parameters with short descriptions and usage can be found in table 6.7. The table also lists the fit values for both cases.

6.7.2 Fitting BDT(MC)

Equation 6.21 is used to fit the final result for the BDT(MC). It can be seen as a solid red line in figure 6.25. The fit yields $\chi^2 = 69.6$ with 56 degrees of freedom. This corresponds to a χ^2 -probability of 10.4% which is a reasonable probability especially since the background was fitted on MC and not on data. The background component of the fit is drawn as a dashed red line, while the signal component is drawn separately at the bottom of the plot magnified by a factor of four for clarity.

The fit parameters give

$$\begin{aligned}
 \mu_w &= (76.9 \pm 1.1)\text{GeV} & \sigma_w &= (5.7 \pm 1.3)\text{GeV} \\
 \mu_z &= (87.0 \pm 1.3)\text{GeV} & \sigma_z &= (6.2 \pm 1.4)\text{GeV}
 \end{aligned}$$

where the reconstructed masses are slightly smaller than what could have been expected. The widths obtained from the fit are also smaller than expectation from MC, but are in the same range as the predicted width from figure 5.5(b) in chapter 5. Furthermore, the signal component of the fit is used to calculate the number of identified bosons. We find

$$\begin{aligned}
 N_W &= (1791.4 \pm 136.0) \\
 N_Z &= (817.7 \pm 59.4) \\
 N_{WZ} &= (2609.1 \pm 148.4) \quad .
 \end{aligned}$$

Since we know that the signal consists of two Gaussians overlaying each other, this signal model was used from the beginning. However, it is also worthwhile to fit a single Gaussian on top of a straight line to see if the data prefers the two-Gaussian model or is better described by only one Gauss function. This fit returns $\chi^2/\text{dof} = 70.4/57$, resulting in a χ^2 -probability of 11.0%. This means the data is indifferent whether it is described by one or two Gaussians. Therefore, our model is a viable model, though not exclusive.

The data with the background component of the fit subtracted is drawn in figure 6.27(a). The excitation in the signal region is clearly observable.

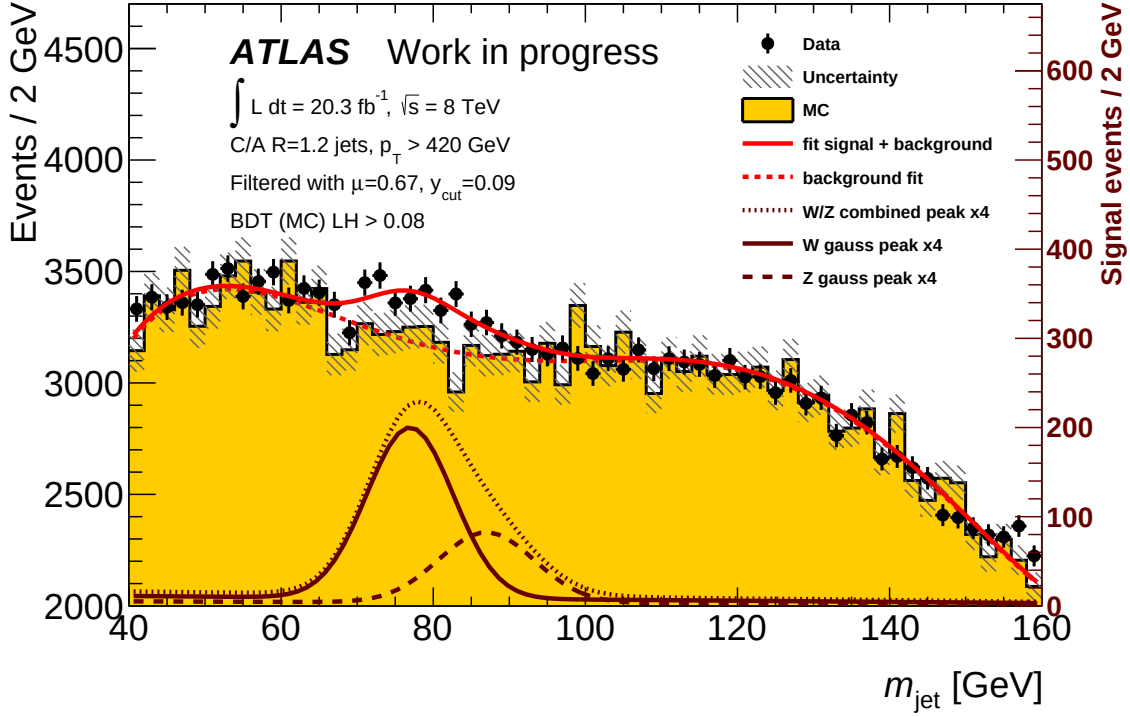


Figure 6.25: Final result for the BDT(MC). The background fit on MC which is shown as a dashed line and is expanded by the sum of two Gaussians on top of a straight line to model the signal. The overall fit is shown as the solid red line. The signal component with contributions from $W^\pm$ and Z^0 of the fit are drawn magnified in dark red colour at the bottom of the plot.

6.7.3 Fitting BDT(DATA)

The procedure for fitting BDT(DATA) is analogous to BDT(MC) and is drawn in figure 6.26. The fit yields $\chi^2 = 64.5$ with 56 degrees of freedom, resulting in a χ^2 -probability of 20.4% which indicates a good fit.

From the fit parameters the mass and width of $W^\pm$ and Z^0 boson are determined. We find

$$\begin{aligned} \mu_w &= (77.3 \pm 0.9) \text{ GeV} & \sigma_w &= (4.7 \pm 0.7) \text{ GeV} \\ \mu_z &= (87.6 \pm 1.2) \text{ GeV} & \sigma_z &= (5.0 \pm 0.8) \text{ GeV} \end{aligned}$$

While the masses are again slightly smaller than MC predictions, the widths turn out to be considerably smaller than the widths of the distribution in signal MC and even slightly smaller than expectation from figure 5.5(b). The number of bosons from the fit are given by

$$\begin{aligned} N_W &= (1698.6 \pm 116.1) \\ N_Z &= (836.6 \pm 61.0) \\ N_{WZ} &= (2571.7 \pm 131.2) \end{aligned}$$

Again we want to test the one-Gaussian signal model as well and find it yields $\chi^2/\text{dof} = 61.5/57$ with a χ^2 -probability of 31.9%. This means in the case of BDT(DATA) the data has a slight preference to the model of only one Gaussian function. This small blemish can be fixed by the availability of more statistics in the future.

The data with the background component of the fit subtracted is drawn in figure 6.27(b). The excitation in the signal region is clearly observable.

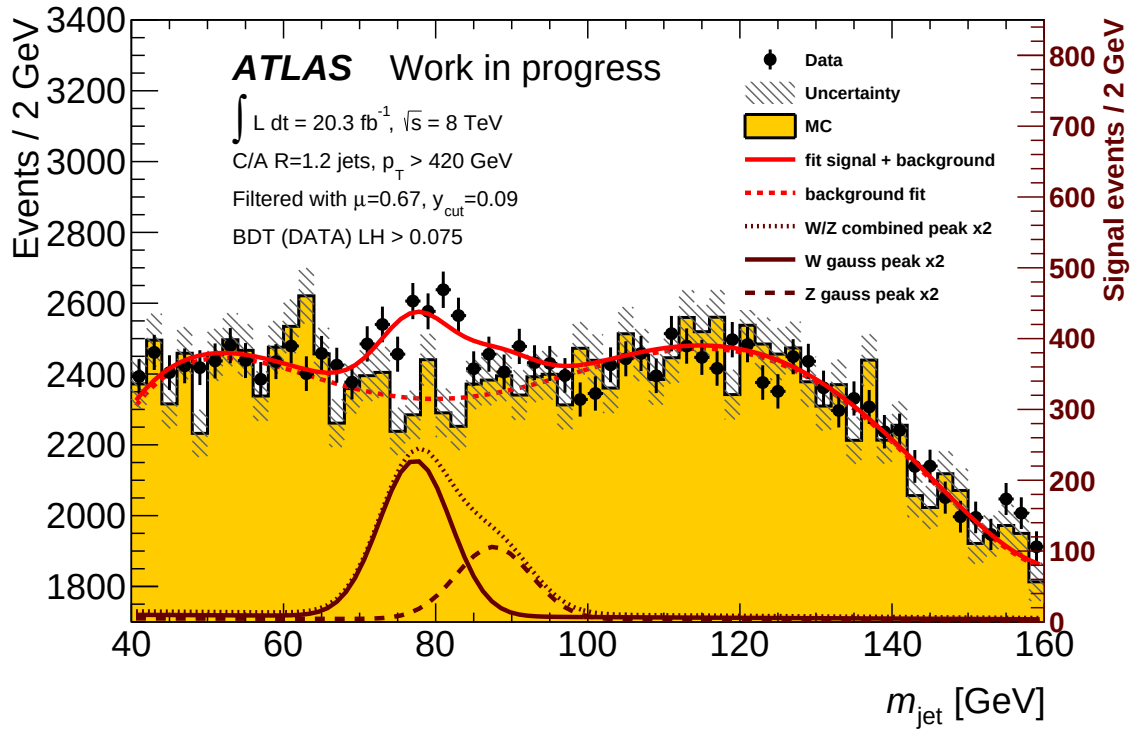


Figure 6.26: Final result for the BDT(DATA). The background fit on MC which is shown as a dashed line and is expanded by the sum of two Gaussians on top of a straight line to model the signal. The overall fit is shown as the solid red line. The signal component with contributions from $W^\pm$ and Z^0 of the fit are drawn magnified in dark red colour at the bottom of the plot.

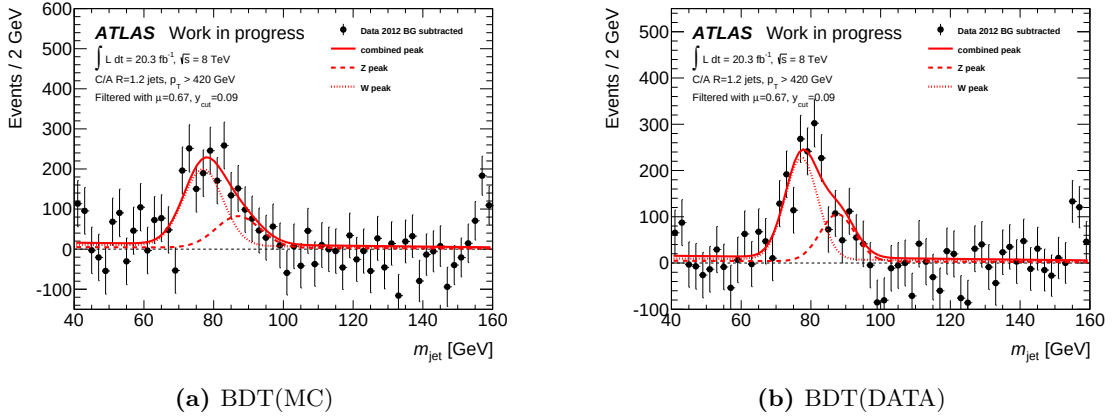


Figure 6.27: Data with the background component of the fit subtracted. The signal component of the fit in chapter 6.7.2 and 6.7.3 is shown as a solid red line. The individual contributions from $W^\pm$ and Z^0 are drawn as dashed lines.

6.7.4 Computation of the production cross sections

The signal yield obtained previously is used to calculate the production cross section of boosted $W^\pm$ and Z^0 bosons with $p_T > 420$ GeV and $|\eta| < 1.9$. It is given by

$$\sigma(pp \rightarrow W^\pm/Z^0 \rightarrow q\bar{q}) = \frac{N_{W^\pm}^{\text{reco}} + N_{Z^0}^{\text{reco}}}{\epsilon_{W^\pm/Z^0} \cdot \mathcal{L}} \quad , \quad (6.22)$$

where $N_{W^\pm}^{\text{reco}}$ and $N_{Z^0}^{\text{reco}}$ are the number of identified signal jets, $\mathcal{L} = \int L dt$ is the integrated luminosity of the data and $\epsilon_{W^\pm/Z^0}$ is the signal reconstruction efficiency. Using signal MC simulation the reconstruction efficiency can be determined using the relation

$$\epsilon_{W^\pm/Z^0} = \frac{\sum_{V_h} N_{V_h}^{\text{reco}}(p_T > 420 \text{ GeV}, |\eta| < 1.9, 40.0 \text{ GeV} < m_{\text{jet}} < 160.0 \text{ GeV}, LH > \Lambda_{BDT})}{\sum_{V_h} N_{V_h}^{\text{truth}}(p_T^{\text{truth}} > 420 \text{ GeV}, |\eta^{\text{truth}}| < 1.9)} \quad . \quad (6.23)$$

Here, the sums over V_h include contributions from both electroweak gauge bosons. p_T , $|\eta|$ and m_{jet} correspond to the reconstructed quantities transverse momentum, pseudorapidity and jet mass. The parameters appearing in the denominator of the fraction p_T^{truth} and η^{truth} are taken from MC truth. The quantity Λ_{BDT} symbolizes the cut on the BDT response value and assumes the value $\Lambda_{BDT} = 0.08$ for BDT(MC) and $\Lambda_{BDT} = 0.075$ for BDT(DATA). From signal MC the efficiencies are computed to be

$$\begin{aligned} \epsilon_{V_h}^{\text{MC}} &= 0.231 \pm 0.002 \\ \epsilon_{V_h}^{\text{DATA}} &= 0.212 \pm 0.002 \quad . \end{aligned}$$

The errors quoted here are statistical only. With the integrated luminosity of $\mathcal{L} = 20.3 \text{ fb}^{-1}$ and the numbers of reconstructed V_h quoted in chapter 6.7.2 and 6.7.3 the cross sections for $\sqrt{s} = 8 \text{ TeV}$ can be calculated and

$$\begin{aligned} \sigma^{\text{MC}}(V_h \mid p_T > 420 \text{ GeV}, |\eta| < 1.9) &= (0.557 \pm 0.032) \text{ pb} \\ \sigma^{\text{DATA}}(V_h \mid p_T > 420 \text{ GeV}, |\eta| < 1.9) &= (0.599 \pm 0.031) \text{ pb} \end{aligned}$$

are found. The production cross sections for both BDTs are in agreement with each other.

It is of course also possible to measure the total production cross section for $W^\pm$ and Z^0 boson. The calculation is according to

$$\sigma(pp \rightarrow V) = \frac{N_V^{\text{reco}}}{\epsilon_V \cdot \mathcal{L} \cdot \text{Br}(V \rightarrow q\bar{q})} \quad V = W^\pm, Z^0. \quad (6.24)$$

Here, $\text{Br}(V \rightarrow q\bar{q})$ is the branching ratio of the gauge boson to decay hadronically. We take the average values from [16] which are given by

$$\begin{aligned} \text{Br}(W^\pm \rightarrow q\bar{q}) &= 0.6760 \pm 0.0027 \\ \text{Br}(Z^0 \rightarrow q\bar{q}) &= 0.6991 \pm 0.0006 \quad . \end{aligned}$$

The efficiencies $\epsilon_{W^\pm}$ and ϵ_{Z^0} are computed analogous to eq. 6.23 and estimated from signal MC. The efficiencies

$$\begin{aligned} \epsilon_{W^\pm}^{\text{MC}} &= 0.228 \pm 0.002 & \epsilon_{W^\pm}^{\text{DATA}} &= 0.206 \pm 0.002 \\ \epsilon_{Z^0}^{\text{MC}} &= 0.236 \pm 0.003 & \epsilon_{Z^0}^{\text{DATA}} &= 0.221 \pm 0.003 \end{aligned}$$

are extracted and the total production cross sections for electroweak gauge bosons with $p_T > 420$ GeV and $|\eta| < 1.9$ at a center of mass energy of $\sqrt{s} = 8$ TeV are determined. For the $W^\pm$ boson we get

$$\begin{aligned}\sigma^{MC}(pp \rightarrow W^\pm \mid p_T > 420 \text{ GeV}, |\eta| < 1.9) &= (0.573 \pm 0.044) \text{ pb} \\ \sigma^{DATA}(pp \rightarrow W^\pm \mid p_T > 420 \text{ GeV}, |\eta| < 1.9) &= (0.600 \pm 0.042) \text{ pb}\end{aligned}$$

and for Z^0 we compute

$$\begin{aligned}\sigma^{MC}(pp \rightarrow Z^0 \mid p_T > 420 \text{ GeV}, |\eta| < 1.9) &= (0.245 \pm 0.018) \text{ pb} \\ \sigma^{DATA}(pp \rightarrow Z^0 \mid p_T > 420 \text{ GeV}, |\eta| < 1.9) &= (0.268 \pm 0.020) \text{ pb} \quad .\end{aligned}$$

The values extracted from both BDTs are consistent with each other.

6.8 Systematic errors

Systematic errors have not been studied here exhaustively. A thorough investigation, however, was done in [191] for a related study on $\sqrt{s} = 7$ TeV. In this study the production cross section of hadronically-decaying boosted $W^\pm$ and Z^0 bosons that were reconstructed using Anti- K_t $R = 0.6$ jets without applying any jet grooming algorithm is measured. Background was rejected using a BDT trained on event shape observables. The summary of the extracted systematic uncertainties on the combined cross section in [191] can be found in table 6.8.

In their fit, the mean and width of the peaks for $W^\pm$ and Z^0 boson were fixed to the values predicted by MC or literature values. This introduces large systematic errors to the measurement. In fact, the jet mass resolution is the largest contributor. Since the fit performed in this study left both, mean and width of the bosons floating the contributions from jet mass resolution and jet mass scale have been avoided. The price we pay for this is a larger statistical error on the measurement. This, however, we do gladly, since statistical errors are much easier to deal with, usually Gaussian and decrease with increasing statistics. Furthermore, we expect that the errors from jet energy scale and jet energy resolution are negligible small.

Sources	2011 data [191] σ_{W+Z}	Present study σ_{W+Z}
MC modelling	4.4%	not investigated
Background PDF	8.8%	6.0%
Signal PDF	5%	2.0%
Jet energy scale	3.7%	not investigated
Jet energy resolution	< 1%	not investigated
Jet mass scale	2.2%	fitted
Jet mass resolution	12.6%	fitted
$t\bar{t}$ contribution	1.3%	not investigated
Single-top and diboson contribution	< 1%	not investigated
$W^\pm$ and Z^0 relative yield	2.9%	not investigated
Luminosity	1.8%	2.8% [198]
Total	17.7%	9.6%

Table 6.8: Summary of the systematic uncertainties in the $W^\pm + Z^0$ production cross section from different sources in [191]. In the present analysis only few sources of systematic errors have been investigated. For the estimation of the total systematic uncertainty results from [191] have been used in the absence of own considerations.

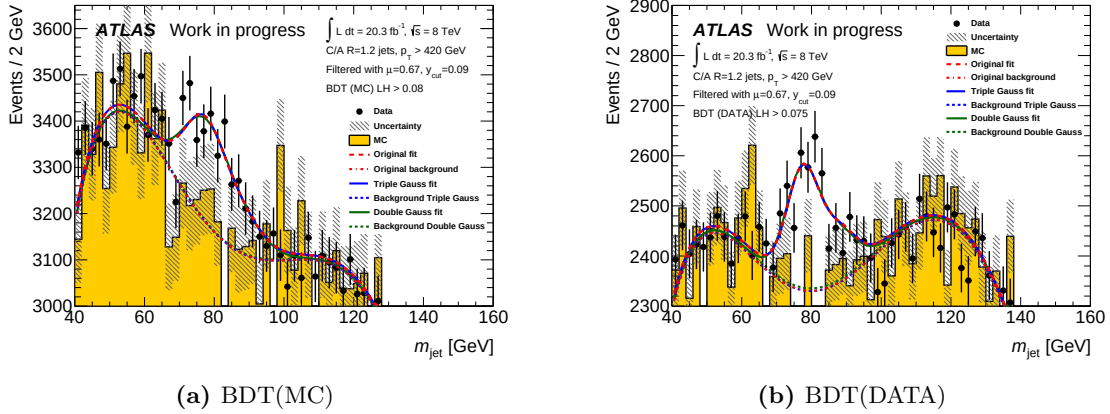


Figure 6.28: To estimate the systematic uncertainty of the signal PDF an alternative fit is performed. The blue and green solid (dashed) lines represent the combined signal and background (background only) fit for a triple- and double-Gaussian fit. The lines in red correspond to the original fits performed in chapters 6.7.2 and 6.7.3.

The rest of the contributors remain. Influences from single top-quark, diboson and $t\bar{t}$ are not investigated due to time constraints and are expected to be small. The systematic uncertainty on the luminosity is obviously present in this study as well. A detailed investigation was done in [199]. The systematic uncertainty on the integrated luminosity for $\sqrt{s} = 8$ TeV data amounts to 2.8% [198].

The uncertainty originating from the MC modelling can be quantified by redoing the study again using a different MC generator. The differences between signal yield and calculated production cross sections from the analysis relying on the HERWIG generator and a study using a different generator allows to estimate the influence of the choice of generator. This remains to be done in the future.

The second and third largest contributors are the signal and background PDF. In the following we want to deal with these two items and get a rough estimate on the contribution in the present study.

6.8.1 Systematic uncertainty of the signal PDF

In fitting the signal with the sum of two Gaussians on top of a straight line we might introduce a systematic error to the measurement as a different signal model might fit the reality more accurately. A good way of investigating the effects of the signal PDF is to fit the data using alternative signal models. We want to investigate two different scenarios.

Instead of using a straight line to model the tails of the distribution, a third much wider Gaussian function is used. This fit-function is given by

$$N'(m_{\text{jet}}) = \mathcal{B}(A_0) + S \left(\Xi + \frac{S_3}{\sqrt{2\pi}\sigma_3} \exp \left[-\frac{1}{2} \left(\frac{m_{\text{jet}} - \mu_3}{\sigma_3} \right)^2 \right] \right) \quad (6.25)$$

where $\mathcal{B}$ is the background with all parameters except the constant A_0 fixed and Ξ is the double-Gaussian already used in the original signal model. The fit parameter S_3 gives the ratio of the third Gaussian, i.e. the tails, to the remaining two. All parameter of the third Gaussian are fixed to the values obtained in MC (see figure E.4(a) in Appendix E). The triple-Gaussian function is now used to fit the data and yields $\chi^2/\text{dof} = 1.24$ [$\chi^2/\text{dof} = 1.14$] corresponding to a fit probability of 10.3% [21.9%] for BDT(MC) [BDT(DATA)]. The fitted function is drawn in figure 6.28(a), the signal component alone can be found in

$\mathcal{B}$	BDT(MC)				BDT(DATA)			
	χ^2	N_{dof}	P_{χ^2}	N_{BG}	χ^2	N_{dof}	P_{χ^2}	N_{BG}
$\mathcal{P}_4$	63.92	54	16.7%	66797	94.60	55	0.1%	50413
$\mathcal{P}_5$	56.69	54	37.5%	66395	72.44	54	4.2%	49726
$\mathcal{P}_6$	57.57	53	31.0%	66456	73.21	53	3.9%	49746

Table 6.9: Fit results for different background models. N_{dof} gives the number of degrees of freedom, P_{χ^2} represents the χ^2 -probability and N_{BG} gives the number of background events in the signal region of $65.0 \text{ GeV} < m_{\text{jet}} < 105.0 \text{ GeV}$.

figure E.5(a). From the new fit function we can calculate the signal yield. For BDT(MC) the new yield is 0.1% larger than the one previously calculated. For BDT(DATA) the corrected signal yield has increased by 0.4%.

In the second model the tails of the distribution are ignored and the signal is fitted only with a double Gaussian function. The used fit function is identical to eq. 6.21 with $C_l = M_l = 0$. See figures 6.28(b), E.4(b) and E.5(a) for the result. The signal yield is expected to change gravely and shows a reduction of 19.0% for BDT(MC) and of 24.6% in the case of BDT(DATA).

The second model is obviously not an adequate model as it disagrees with signal MC. Furthermore, since the triple-Gaussian fit results in almost no change of the signal yield a correction of the measured signal becomes unnecessary. We assign, however, a systematic uncertainty on the signal yield of 2.0%.

6.8.2 Systematic uncertainty of the background PDF

In chapter 6.7.1 we described how a fifth degree polynomial was used to describe the background. Doing so of course allows for a systematic error in the analysis. Since the data is dominated by background it is very important to have a good understanding of the background model. The systematic errors are evaluated by using alternative background fit models and observation of their influences on the signal yield. In our case the procedure seems to be straightforward: If a polynomial of fifth order $\mathcal{P}_5$ performs well, why not try a polynomial of fourth $\mathcal{P}_4$ and sixth $\mathcal{P}_6$ degree? In figure E.6 in Appendix E the background fits have been performed on MC. While the fifth and sixth degree polynomial are almost on top of each other, the fourth order polynomial shows a different behaviour as it does not account for all the bumps and dips very well²⁰. A summary of the χ^2 , degrees of freedom, χ^2 -probability and number of background events in the signal region $65.0 \text{ GeV} < m_{\text{jet}} < 105.0 \text{ GeV}$ for all fits can be found in table 6.9. Since the number of background events in the signal region shows no unreasonable jumps between the models and the χ^2 and corresponding probability of $\mathcal{P}_5$ and $\mathcal{P}_6$ is nearly identical and furthermore $\mathcal{P}_4$ cannot compete with the original model, we feel confirmed in our choice of the background model and it enhances its trustworthiness.

The signal model is the default double-Gaussian on top of a straight line. The combined signal+background function is used to fit the data. The resulting overall fit functions are plotted in figure 6.29(a) for BDT(MC) and 6.29(b) for BDT(DATA). It can be clearly seen, that the fits $\mathcal{P}_5$ and $\mathcal{P}_6$ are almost identical. In fact, the signal yield has only decreased by 3.7% for BDT(MC) and 2.5% for BDT(DATA) when comparing $\mathcal{P}_6$ to the default $\mathcal{P}_5$ model. The fit using $\mathcal{P}_4$ to describe the background, however, gives results that are 17.6% smaller for BDT(MC) and 37.9% for BDT(DATA). This is seen very well in figures E.7 in Appendix E.

²⁰The very first bin was excluded in this fit as it would have drawn the function towards zero at low masses spoiling the fit.

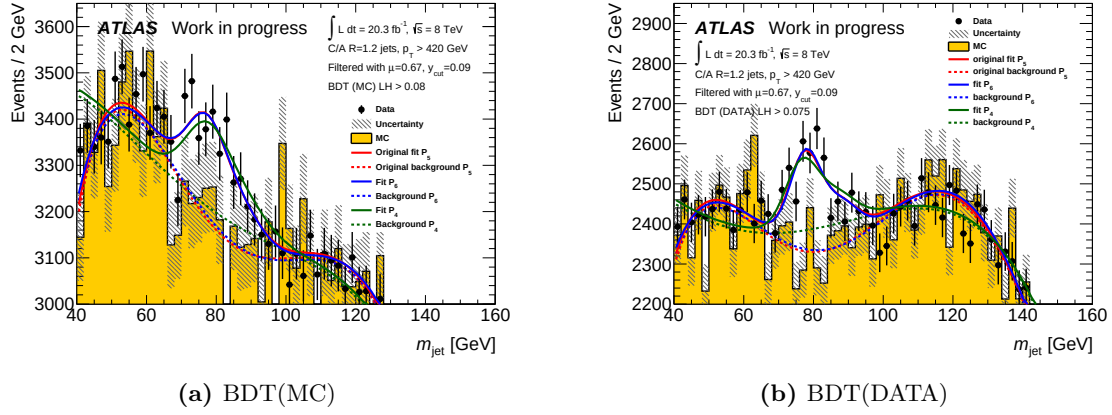


Figure 6.29: Different models for the description of the background are used to estimate the systematic error on the background PDF. The red lines solid (dashed) represent the combined signal and background fit (background only fit) performed in chapters 6.7.2 and 6.7.3 with a fifth degree polynomial representing the background, while the blue and green lines are for the sixth and fourth order polynomial background models respectively.

Since the $\mathcal{P}_4$ fit performs much worse than the other two we do not correct the measured signal yield but assign a systematic uncertainty of 6.0% for both BDTs.

6.9 Results

We assume the uncertainties estimated in chapters 6.8.1 and 6.8.2 are uncorrelated and get the total systematic uncertainty by adding them in quadrature. One can compute the signal yields considering the systematic uncertainties. They are given by

$$\begin{aligned} N_{WZ}^{MC} &= 2609.1 \pm 148.4 \text{ (stat.)} \pm 251.6 \text{ (sys.)} \\ N_{WZ}^{DATA} &= 2571.7 \pm 131.2 \text{ (stat.)} \pm 248.0 \text{ (sys.)} \end{aligned}$$

Using these numbers one can calculate the production cross section of $W^\pm$ and Z^0 bosons with $p_T > 420$ GeV and $|\eta| < 1.9$:

$$\begin{aligned} \sigma_{corr}^{MC}(W^\pm/Z^0 \rightarrow q\bar{q} \mid p_T > 420 \text{ GeV}, |\eta| < 1.9) &= (0.557 \pm 0.032 \text{ (stat.)} \pm 0.054 \text{ (sys.)}) \text{ pb} \\ \sigma_{corr}^{DATA}(W^\pm/Z^0 \rightarrow q\bar{q} \mid p_T > 420 \text{ GeV}, |\eta| < 1.9) &= (0.599 \pm 0.031 \text{ (stat.)} \pm 0.058 \text{ (sys.)}) \text{ pb} \end{aligned}$$

They are clearly different from zero and exhibit a statistical significance of 8.9σ for BDT(MC) and 9.1σ for BDT(DATA).

Conclusion and outlook

Conclusion

In this thesis, the effectiveness of the mass-drop filtering technique to suppress pileup in a jet and recover the sharp mass-peak corresponding to the mass of the initiating particle has been demonstrated. Next the procedure was validated by reconstructing electroweak gauge bosons from MC samples. The technique was then employed on data collected by the ATLAS detector in 2012 to identify boosted $W^\pm$ and Z^0 bosons based on their mass, where the hadronic decay products are collimated and clustered into a single fat jet. To reduce the background from pure QCD events, two boosted decision trees have been used. One in which signal MC was trained against background MC [BDT(MC)] and a second in which sidebands in data have been taken as background sample [BDT(DATA)]. Substructure observables, event shape variables and kinematic characteristics of the jets were found to possess a certain discriminating power exploited here. After removing the small correlation of the jet mass and the BDT response values, the BDTs achieved to reject 87% and 88% of the background while keeping still 50% of the signal. Using this it was possible to observe an excess in the disubjet invariant mass spectrum at masses corresponding to the electroweak gauge bosons. Next, the spectrum was fitted and the signal yield determined from which the production cross section of hadronically decaying boosted $W^\pm$ and Z^0 with $p_T > 420$ GeV and $|\eta| < 1.9$ was calculated to be

$$\begin{aligned}\sigma^{BDT(MC)} &= 0.557 \pm 0.032 \text{ (stat.)} \pm 0.054 \text{ (sys.) pb} \\ \sigma^{BDT(DATA)} &= 0.599 \pm 0.031 \text{ (stat.)} \pm 0.058 \text{ (sys.) pb} \quad .\end{aligned}$$

The two results agree with each other and show a strong statistical significance of 8.9σ for BDT(MC) and 9.1σ for BDT(DATA).

Outlook

The systematic uncertainties in the measurement are possibly underestimated especially for the signal and background PFD since only very few different models were considered. To get a more trustworthy estimate an in-depth analysis of systematic errors has to be performed. It would be very useful to repeat the study using different MC generators to estimate the systematic error caused by the specific choice of Herwig and Herwig++. Here, we have also neglected to investigate contributions from $t\bar{t}$, single-top and diboson events on the mass-spectrum since they are expected to be very small. In future, one could perform additional studies on MC to verify that claim and estimate the systematic error made by ignoring those contributions.

There is still much room to improve the efficiency of the fat jet selection with the BDTs. Variables like the number of charged tracks between the fat jet and the beam have already been proven to be a good discriminator, yet with more fine-tuning and optimization it might be possible to unlock the full potential of the observable. The discriminating power of the event shape variables sphericity, aplanarity and thrust minor can be improved by using D3PDs that contain the full energy cluster information instead of only the limited number of subjects.

Further separation between the signal and background distributions can be achieved by considering new discriminators. It is possible that there exist fundamental differences in the fragmentation of QCD background jets and jets resulting from the decay of boosted colour-neutral objects. Further studies are necessary to find a good estimator that reflects those differences. A very similar concept is already in use at CMS to distinguish jets initiated by gluons and quarks [200].

Finally, in the beginning of the project we have chosen the mass-drop filter since it is the most commonly used jet grooming technique in ATLAS. Other algorithms like jet trimming could be tried as well and their performance compared to the results of this study.

General outlook

The successful observation of resonances in the mass spectrum from $W^\pm$ and Z^0 bosons in the present study shows that the mass-drop filtering procedure is capable of reproducing established physics. This is of central importance in evaluating the usefulness of the procedure in searches for new physics. The distinct advantage of the hadronic channels in terms of higher statistics through large branching ratios makes the mass-drop filtering indeed an extremely valuable tool to probe the spectrum at the very edges of the mass scale where the background from pure QCD processes is much smaller than in this study. This includes obviously not only single-boson channels but in particular fully hadronic diboson channels. While it is most likely not possible to perform high precision measurements in the all-hadronic channels it is nevertheless possible to identify regions of interest, certain mass windows in a specific channel, which show behaviour deviating from SM predictions. The arrival of more data and higher luminosities should then give the possibility to get enough events in the regions of interest in possibly clearer semi-leptonic or even all-leptonic channels as well.

With the begin of Run II in early 2015 the center of mass collision energy is raised to 13 TeV and later 14 TeV. Naturally, this will increase the produced number of hadronically decaying boosted gauge bosons that can not be reconstructed by standard jet clustering. We can also expect the number of simultaneous interactions to rise significantly. The mass-drop filtering should be able to cope with both of these. Hence, the relevancy of the technique is expected to grow. Mass-drop filtering or equally well-performing jet algorithms are indispensable for LHCs Run II.

Outlook on $W^\pm W^\mp \rightarrow l\nu jj$

In the present study, the number of reconstructed V_h over the large background was too small to perform high precision measurements. This is essentially a result from the large cut on the fat jet transverse momentum of $p_T > 420$ GeV which was necessary to ensure the triggers working close to unity efficiency. Much could be gained from lowering the p_T threshold. One way of achieving this without sacrificing the trigger efficiency is to look on $WW \rightarrow l\nu jj$ events instead, where the lepton is either an electron or a muon. The hadronically decaying $W^\pm$ can then be reconstructed using the technique presented here.

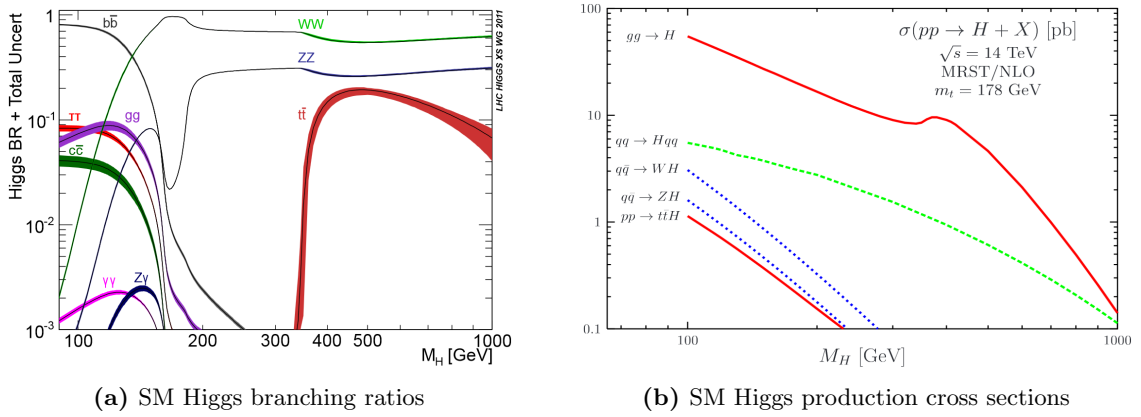


Figure 7.1: (a) shows the different Higgs decay channels and their branching ratios plotted as a function of the Higgs mass m_H . (b) gives the cross sections of the most dominant Higgs production modes as a function of m_H . The figures were taken from [203, 204].

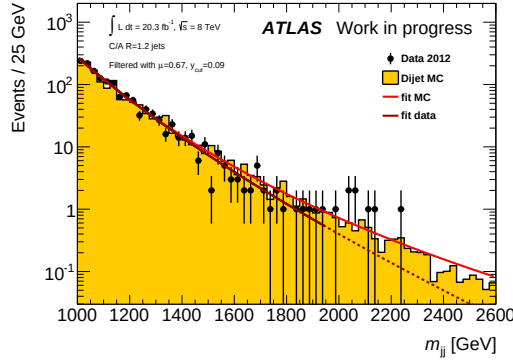
The lepton provides a clean signature and can be used for triggering. Thus, the p_T cut on the fat jet can be lowered significantly. The gain in events would yield a possibly much clearer peak structure that is easier to fit since the background is expected to be of a continuous potentially falling shape. This should also reduce the systematic uncertainty coming from signal and background PDF, which are the largest contributors.

For future projects, it is possible to employ the mass-drop filtering in a search for an additional heavy Higgs boson in the $W^\pm W^\mp$ channel. Figure 7.1(a) reveals that this is actually one of the most promising search channels since the WW decay mode has the largest branching ratios for large Higgs masses. Furthermore, at around $m_H \approx 1000$ GeV the vector boson fusion (VBF) production channel passes the gluon-gluon fusion mode as the channel with the largest cross-section. This can be seen from figure 7.1(b) where the most dominant Higgs production modes are plotted as a function of the Higgs mass m_H for $\sqrt{s} = 14$ TeV. A VBF $H \rightarrow WW \rightarrow l\nu jj$ search (where $l = e, \mu$) seems indeed a suitable continuation of this project. [201, 202] might be helpful as starting points for this new endeavour.

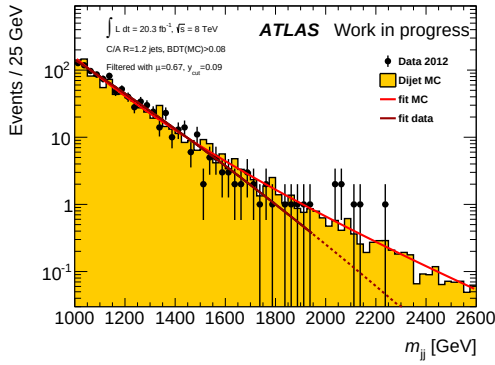
This would also have additional advantages. The charge of the lepton can be determined easily by the bending of the associated track in the magnetic field. However, the charge of a jet is much harder to determine even in the case with no pileup or underlying event contamination since a lot of tracks have to be taken into account and we expect a wide jet charge distribution as we have seen in chapter 5.4. Since we would look for a particle where the electric charge is known to be zero, one expects to find one positive and one negative $W^\pm$ boson. The jet charge can thus be used as an additional discriminator for the retraining of a BDT method and help to reject non-signal jets. There is, however, much more work needed to optimize the calculation of the jet charge for fat jets. But in combination with other variables it might even help to discriminate between different quark flavours or distinguish quark and gluon jets [205, 206].

Outlook on $W^\pm Z^0 \rightarrow jjjj$

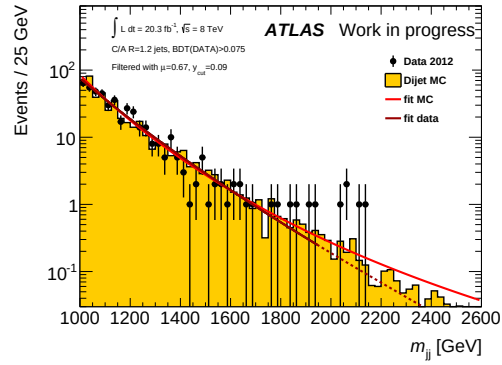
Inspired by [207] it is possible to use the mass-drop filtering to look at all-hadronic WZ final states. We have used the method described in chapter 6 on events where we require two fat jets j_1 and j_2 with $|y_{j_2} - y_{j_1}| < 1.2$. Furthermore, the filtered masses of both jets have to be within 13 GeV of either W or Z boson mass with a combined invariant mass of



(a) Plain m_{jj} spectrum



(b) $BDT(MC) > 0.08$



(c) $BDZ(DATA) > 0.075$

Figure 7.2: Invariant mass spectrum for $WZ \rightarrow 4q$ events. (a) shows the plain spectrum while in (b) a cut on the $BDT(MC)$ response value of 0.08 is applied. In (c), we require $BDT(DATA) > 0.075$. The red function is the fit function of eq. 7.1 fitted to MC, the dark red function fits the data.

$m_{jj} > 1000 \text{ GeV}$. The resulting invariant mass spectrum for the QCD dijet MC samples from the previous chapter and data can be seen in figure 7.2(a). In the plot, both MC and data are fitted separately with a function of the form

$$N(m_{\text{jet}}) = C_0 \cdot \frac{(1-z)^{C_1}}{z^{C_2}} \quad \text{with} \quad z = \frac{m_{\text{jet}}}{8000 \text{ GeV}} \quad (7.1)$$

on the full range for MC and in $1 \text{ TeV} < m_{\text{jet}} < 1.95 \text{ TeV}$ for data. Clearly, we are at the most outer edge of what is still reachable with current data. The agreement between MC and data is reasonably good. Nevertheless, there seems to be a tiny surplus of roughly 6 events at around 2 TeV, which is definitely worth considering further. Since we have two working BDTs trained to distinguish pure QCD jets from V_h jets, we apply a cut on the BDT response value of $BDT(MC) > 0.08$ and $BDT(DATA) > 0.075$. The resulting spectra are found in figures 7.2(b) and 7.2(c). In both cases the impression of a surplus of events remains. For future projects, it is possible to investigate this further, either with new and more data or by switching to a semileptonic channel where the higher signal efficiency could compensate the reduction of events through the smaller branching ratio.

Appendices

Momentum measurement

To measure the momentum of a particle in the ATLAS detector, the sagitta method is used. Since the detector is submerged in a strong magnetic field, the reconstructed tracks of charged particles exhibit a bending which gives information about the particle momentum.

The motion of a charged particle moving in the plane perpendicular to the direction of the magnetic field is governed by the Lorentz force

$$\vec{F}_L = q\vec{v} \times \vec{B} \quad (\text{A.1})$$

where $\vec{B}$ is the magnetic field, $\vec{v}$ the velocity of the particle and q it's charge. The result is a curved trajectory with length L and radius R . The momentum p of the charged particle is given by

$$p = 0.3 \cdot B \cdot R \quad . \quad (\text{A.2})$$

Out of curiosity, we can calculate the radius of the helix trajectory of a 5 GeV charged particle in the 2 T magnetic field of the inner detector and find $R = 8.3\text{m}$. With the definitions from figure A.1 the sagitta s of the track can be expressed as

$$s = R(1 - \cos \alpha) \approx \frac{R\alpha^2}{2} = \frac{L^2}{8 \cdot R} \quad . \quad (\text{A.3})$$

From eq. A.2 we see that $\delta R/R = \delta p/p$ which allows us to write the uncertainty of the sagitta as

$$|\delta s| = \frac{L^2}{8 \cdot R} \cdot \frac{\delta R}{R} = \frac{L^2}{8 \cdot R} \cdot \frac{\delta p}{p} \quad . \quad (\text{A.4})$$

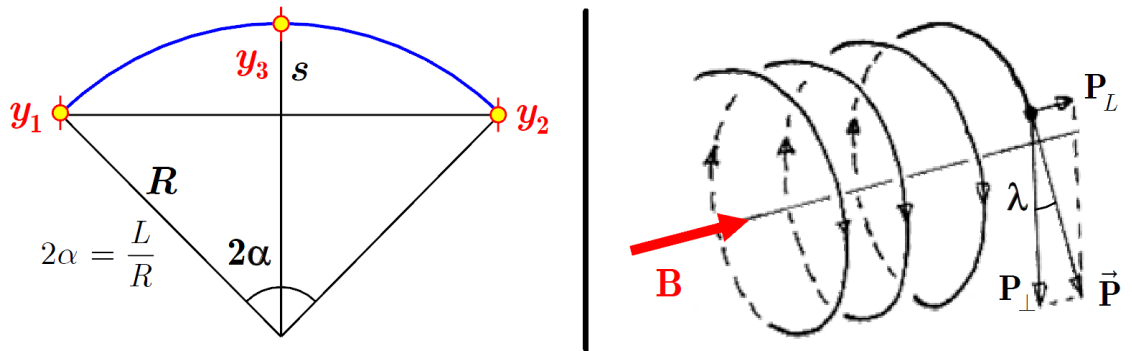


Figure A.1: Left: Definition of the sagitta for a charged particle track (blue line) and three detector hits. Right: Sketch of a helix trajectory of a charged particle in a strong magnetic field. Both figures were taken from [208].

Substituting $R = \frac{p}{0.3 \cdot B}$ and rearranging yields

$$\frac{\delta p}{p^2} = \frac{8}{0.3 \cdot B \cdot L^2} \delta s \quad (\text{A.5})$$

from which we can see that a strong magnetic field, tracking over a wide distance (i.e. L is large) and a large number of tracking points are necessary in order to minimize the uncertainty on the momentum measurement.

Main parameters of the calorimeter system

Table [B.1](#) lists the main parameters of the ATLAS calorimeter system. The figures [B.1](#) and [B.2](#) show diagrams of the amount of material in front and in the calorimeter. The table and the figures were taken from [\[58\]](#).

	Barrel		End-cap	
EM calorimeter				
Number of layers and $ \eta $ coverage				
Presampler	1	$ \eta < 1.52$	1	$1.5 < \eta < 1.8$
Calorimeter	3	$ \eta < 1.35$	2	$1.375 < \eta < 1.5$
	2	$1.35 < \eta < 1.475$	3	$1.5 < \eta < 2.5$
			2	$2.5 < \eta < 3.2$
Granularity $\Delta\eta \times \Delta\phi$ versus $ \eta $				
Presampler	0.025×0.1	$ \eta < 1.52$	0.025×0.1	$1.5 < \eta < 1.8$
Calo. 1st layer	$0.025/8 \times 0.1$	$ \eta < 1.40$	0.050×0.1	$1.375 < \eta < 1.425$
	0.025×0.025	$1.40 < \eta < 1.475$	0.025×0.1	$1.425 < \eta < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \eta < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \eta < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \eta < 2.4$
			0.025×0.1	$2.4 < \eta < 2.5$
Calo. 2nd layer	0.025×0.025	$ \eta < 1.40$	0.1×0.1	$2.5 < \eta < 3.2$
	0.075×0.025	$1.40 < \eta < 1.475$	0.050×0.025	$1.375 < \eta < 1.425$
			0.025×0.025	$1.425 < \eta < 2.5$
Calo. 3rd layer			0.1×0.1	$2.5 < \eta < 3.2$
	0.050×0.025	$ \eta < 1.35$	0.050×0.025	$1.5 < \eta < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	
LAr hadronic end-cap				
$ \eta $ coverage			$1.5 < \eta < 3.2$	
Number of layers			4	
$\Delta\eta \times \Delta\phi$			0.1×0.1	$1.5 < \eta < 2.5$
			0.2×0.2	$2.5 < \eta < 3.2$
Readout channels			5632 (both sides)	
LAr forward calorimeter				
$ \eta $ coverage			$3.1 < \eta < 4.9$	
Number of layers			3	
$\Delta\eta \times \Delta\phi$ (cm)			FCal1: 3.0×2.6	$3.15 < \eta < 4.3$
			FCal1: $\approx 4\times$ finer	$3.10 < \eta < 3.15$
				$4.30 < \eta < 4.83$
			FCal2: 3.3×4.2	$3.24 < \eta < 4.5$
			FCal2: $\approx 4\times$ finer	$3.20 < \eta < 3.24$
				$4.50 < \eta < 4.81$
			FCal3: 5.4×4.7	$3.32 < \eta < 4.60$
			FCal3: $\approx 4\times$ finer	$3.29 < \eta < 3.32$
				$4.60 < \eta < 4.75$
Readout channels			3524 (both sides)	
Scintillator tile calorimeter				
	Barrel		Extended barrel	
$ \eta $ coverage	$ \eta < 1.0$		$0.8 < \eta < 1.7$	
Number of layers	3		3	
$\Delta\eta \times \Delta\phi$	0.1×0.1		0.1×0.1	
	0.2×0.1		0.2×0.1	
Readout channels	5760		4092 (both sides)	

Table B.1: Main parameters of the calorimeter system

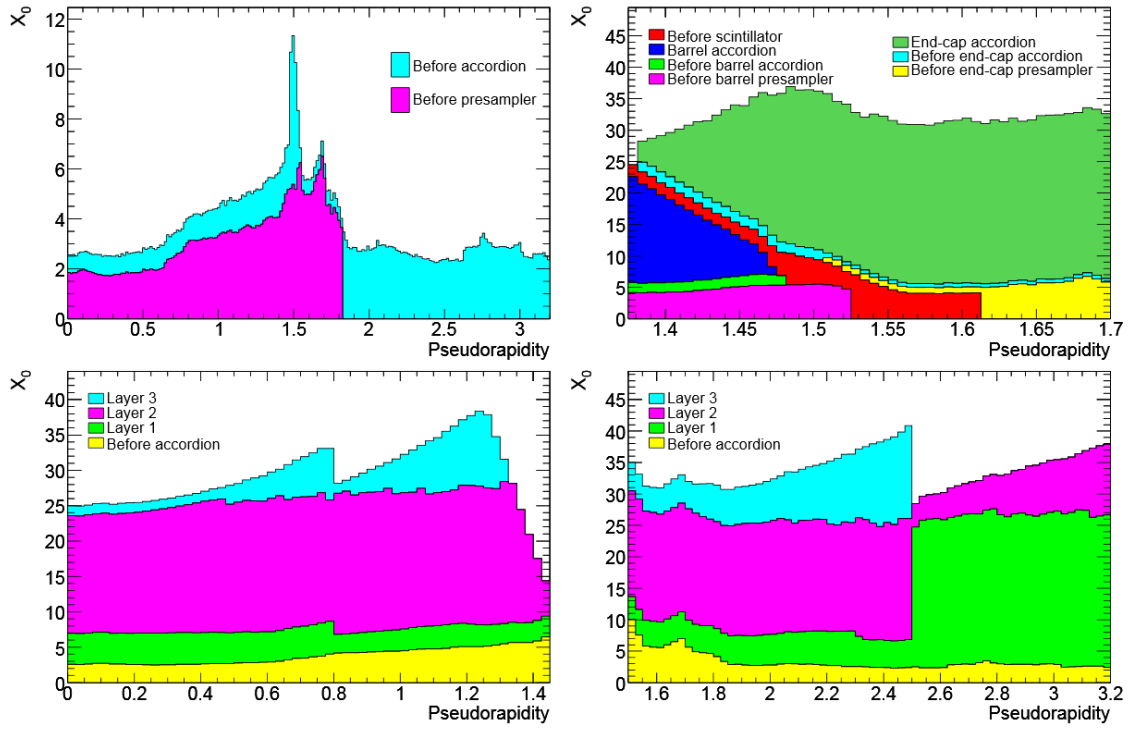


Figure B.1: Material as a unit of the radiation length X_0 in front and in the EM calorimeter as a function of pseudorapidity. Top left: Material before the presampler layer and before the accordion itself. Top right: Transition region between barrel and end-cap cryostat. Material in front is shown as well as the total thickness of the active calorimeter. Bottom: Amount of material in front of the accordion and their total thickness in the barrel (left) and end-cap (right).

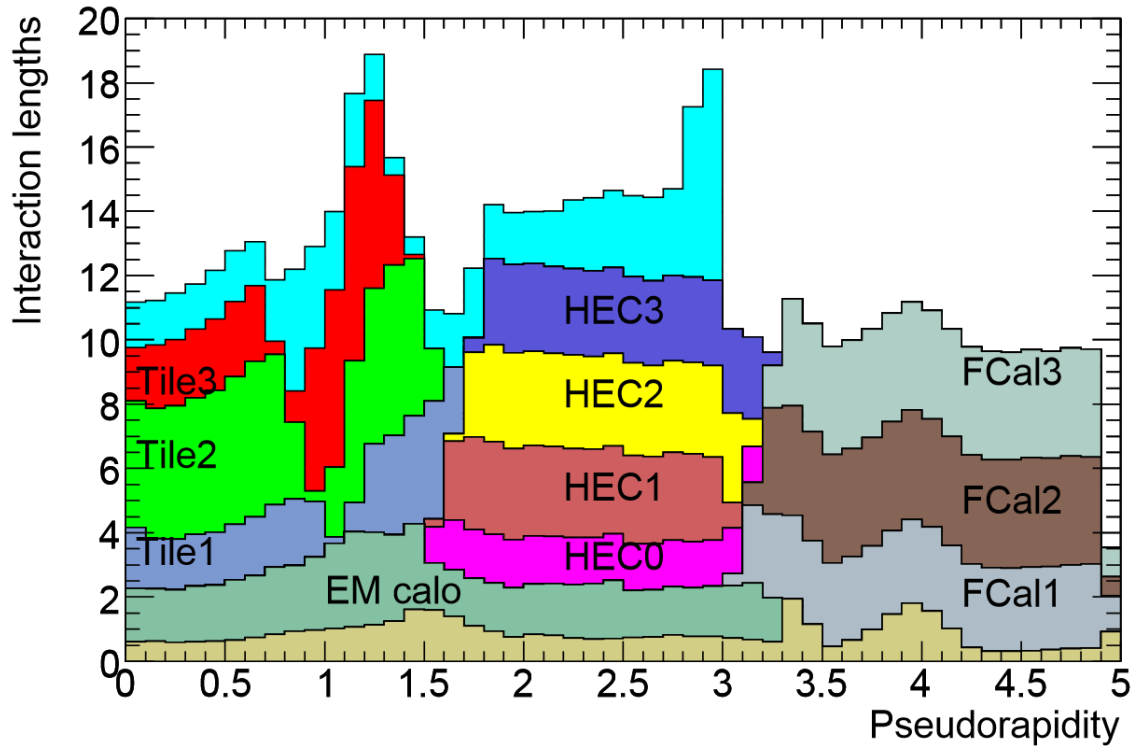


Figure B.2: Amount of material in units of the interaction length in front and in the EM calorimeter, in all hadronic layers and the total thickness of the calorimeter as a function of pseudorapidity.

Boosted Decision Trees (BDTs)

Boosted decision trees (BDTs) have been a popular discrimination instrument for some time [209]. They are a sequential model, which through the comparison of a samples attributes to threshold values classifies unknown events into predetermined groups. For a recent (2013) overview of decision trees see [210]. We want to give a crude picture on the function principle of a BDT. In our case, it serves to classify an input sample into two groups: signal and background. In the process the classification is based on a series of decisions. The tree consists of three elements:

- **Nodes** are points at which the sample undergoes a simple query about a certain characteristic and is divided into two branches.
- **Branches** connect nodes with each other. Branches of trees are always read into the same direction, i.e. from parent-node to daughter-node, not the other way round.
- **Leaves** can be understood as the final nodes of a tree. They are connected to a parent-node by branches but otherwise no more branches divert from it.

The following oversimplified example should demonstrate the working principle and is pictured schematically in figure C.1. Suppose we have a sample consisting of an equal number of tables and stools. The first thing to do is to split up the sample into two parts. The first part is used for the training of the classifiers, the second part later on used for the evaluation of their functionality. For every table and stool we have three variables to work with: height, surface area and weight. At the start the BDT would order all tables and stools according to their height. The algorithm then tries different threshold values to split the training sample into two parts. The one which yields the best separation into mainly tables on one side and predominantly stools on the other is chosen. This same procedure is done for all other available variables, i.e. surface area and weight, and those cutting values are picked that yield the optimal separation. The variable which distinguishes best is selected to form the first node, in our example 'height'. This way out of the initial sample two branches were created. This whole procedure, that is trying all possible cutting values of all remaining variables, is repeated for all branches and all tables and stools in this branch to find the optimal cutting point of the best suited variable for that specific branch. The branch yielding the highest quality increase is chosen next. The algorithm keeps splitting the sample either until the leaves contain solely tables or stools, the number of tables and stools is too small to continue the algorithm, or a given number of leaves is reached.

Of course in our analysis we do not want to distinguish tables from stools, but signal from background and instead of height, surface area and weight we use appropriate classifiers to the analysis. Following [211] the purity of the separation of signal and background

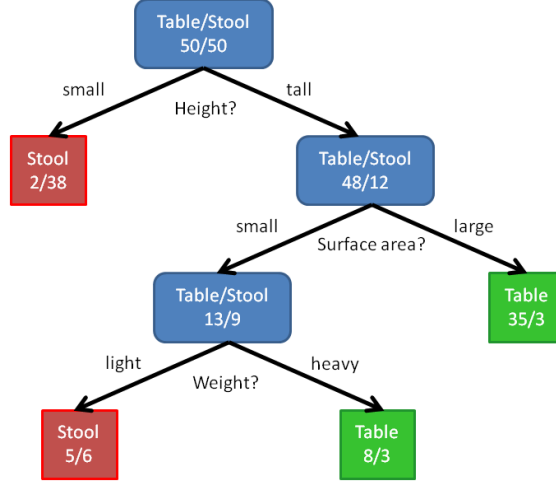


Figure C.1: Schematic illustration of a boosted decision tree to separate a sample of tables and stools according to the three classifiers height, surface area and weight. Boxes with rounded corners represent nodes, square boxes leaves. The branches are indicated by arrows. The figure is inspired by [211]

is computed by

$$p = \frac{\sum_s W_s}{\sum_s W_s + \sum_b W_b} \quad (\text{C.1})$$

where $\sum_{s/b}$ is a sum over all signal/background events and $W_{s/b}$ being the weight of the event. Using this and the number of events n in the branch one can calculate the Gini index G

$$G = \left(\sum_{i=1}^n W_i \right) \cdot p(1-p) \quad . \quad (\text{C.2})$$

In order to get a measure for the quality of the separation between signal and background, the quantity $G_{\text{daughter}}^l + G_{\text{daughter}}^r$, where the indices l and r stand for the left and right branch respectively, is minimized. The quality increase through the splitting of a node is calculated by maximizing

$$\text{criterion} = G_{\text{parent}}^l - G_{\text{daughter}}^l - G_{\text{daughter}}^r \quad . \quad (\text{C.3})$$

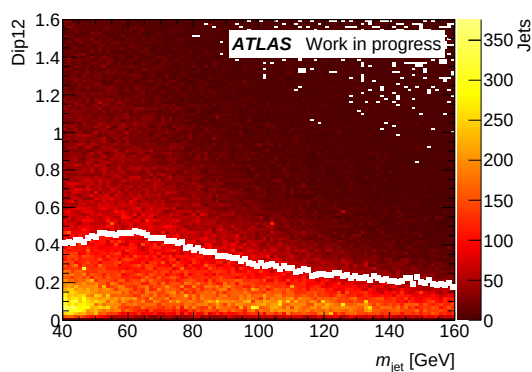
A leaf is labelled a signal leaf, if the leafs purity is larger then 0.5 and a background leaf otherwise. Whether an event travelling through the decision tree is called a signal or a background event depends on the label of the leaf it ends up in. An event in a signal leaf is called a signal event and correspondingly for background. Even though decision trees are a very powerful tool, there remain some problems like the instability problem [212]. This means, that minor changes in the training dataset cause grave changes in the output classification rules. TMVA [196], the toolkit for multivariate data analysis in root, allows the usage of several separation criteria: Gini Index $p \cdot (1-p)$, cross entropy $-p \cdot \ln(p) - (1-p) \cdot \ln(1-p)$, misclassification error $1 - \max(p, 1-p)$, statistical significance $S/\sqrt{S+B}$ and average squared error $1/N \cdot \sum^N (y - \hat{y})^2$, where y is the event-specific target

of the regression and $\hat{y}$ the mean over all events. In the analysis in this thesis, the Gini Index separation method is used.

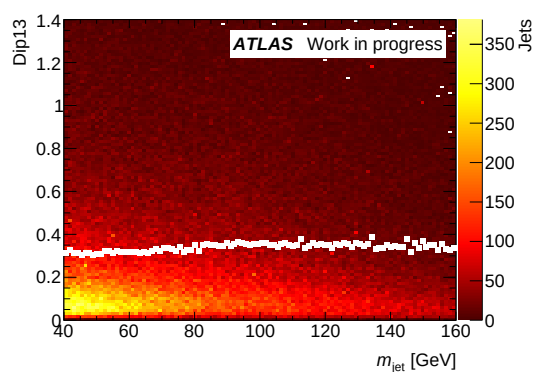
The efficiency of the BDT can be improved by increasing (boosting) the weight of certain events [213, 214]. The procedure is the following: A decision tree from unweighed events is built as explained. Whenever a training event gets classified wrongly, for example in our example above a stool gets labelled as a table, its weight is increased. Using the new weights a new decision tree is built. The procedure is now repeated and the weights of misclassified events is boosted again. It is possible to build thousands of trees during the training of one BDT method.

To make a decision whether an event is signal or background the event travels through all the trees collecting a score. This score is increased by one whenever the event is classified as a signal event by a tree and lowered by one whenever it lands on a background leaf. After reweighing the event score not only allows for a decision whether an event is more likely to be signal or background, but it is possible to select a specific fraction of signal while rejecting a certain amount of background, by cutting on this BDT response value. The measure for the goodness of the BDT is the ROC-curve, for which the background rejection rate is plotted as a function of the signal efficiency.

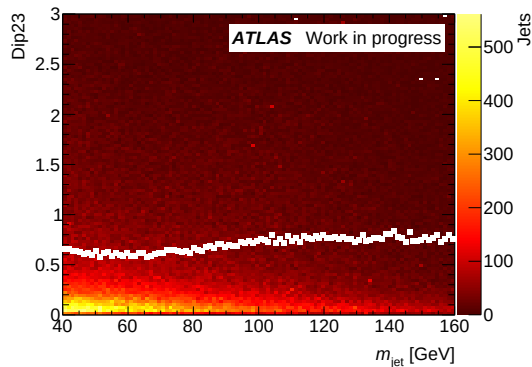
BDT correlation plots



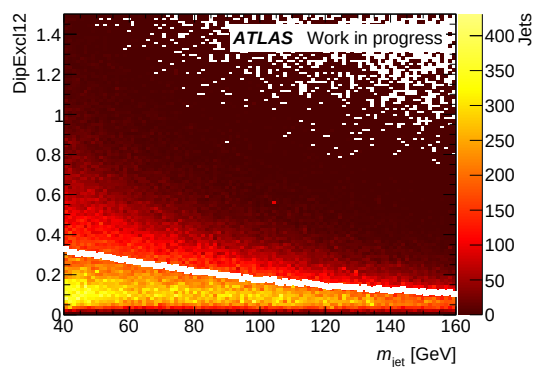
(a) $\mathcal{D}_{12}$ versus m_{jet}



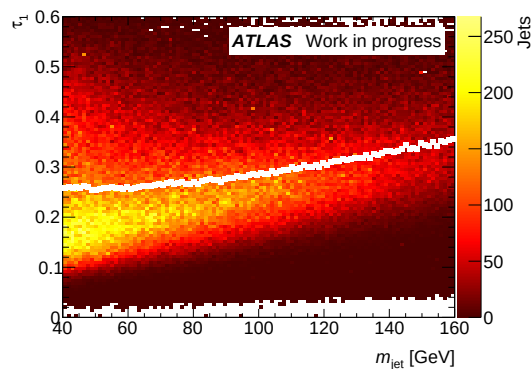
(b) $\mathcal{D}_{13}$ versus m_{jet}



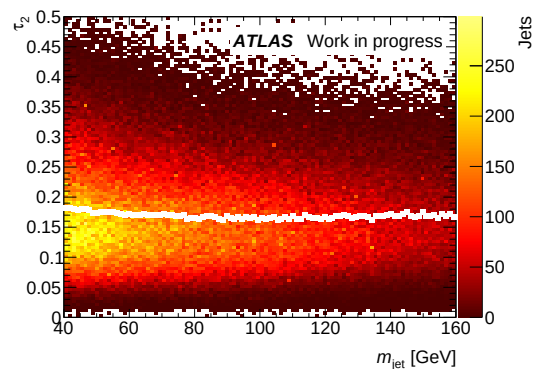
(c) $\mathcal{D}_{23}$ versus m_{jet}



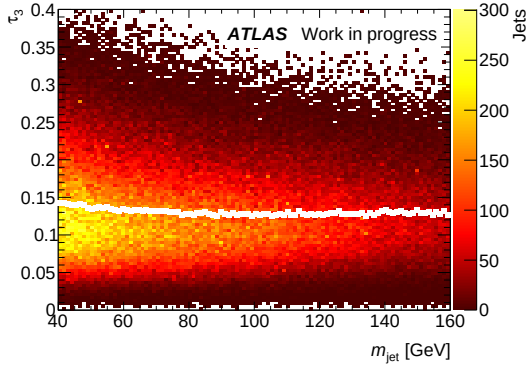
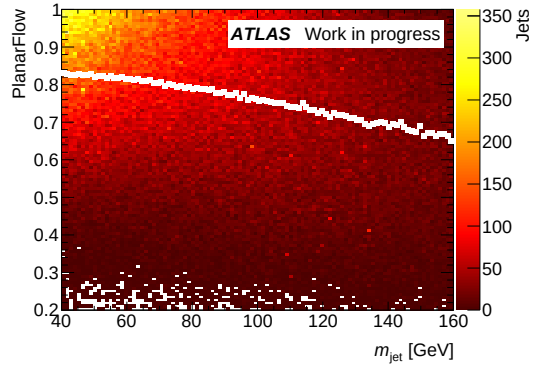
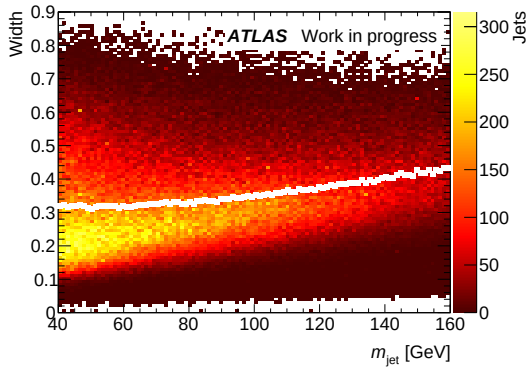
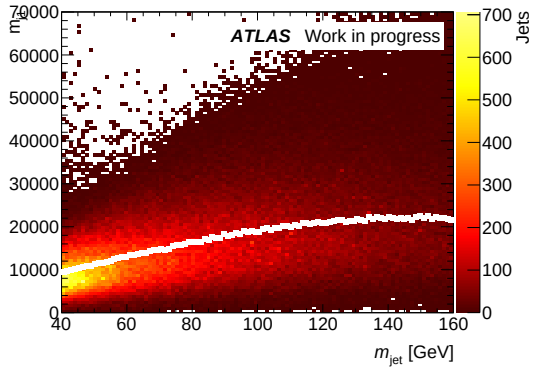
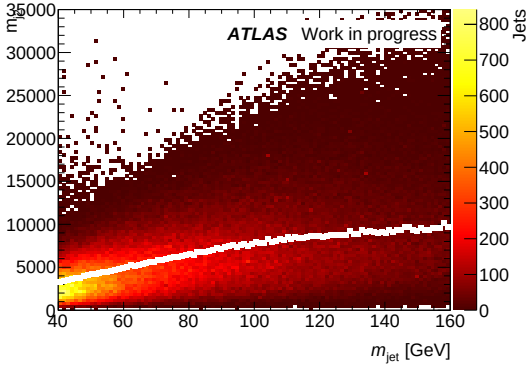
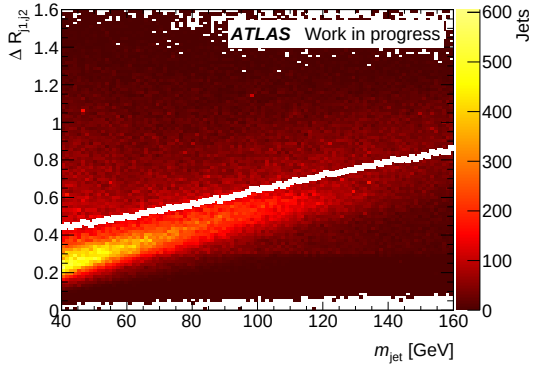
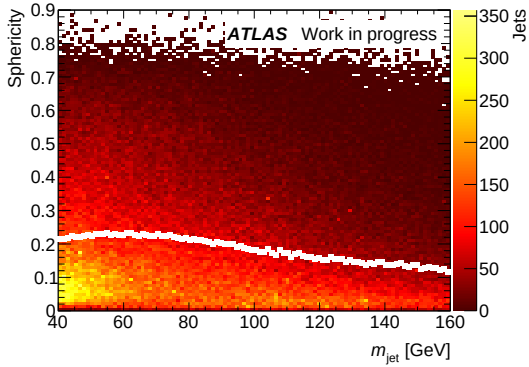
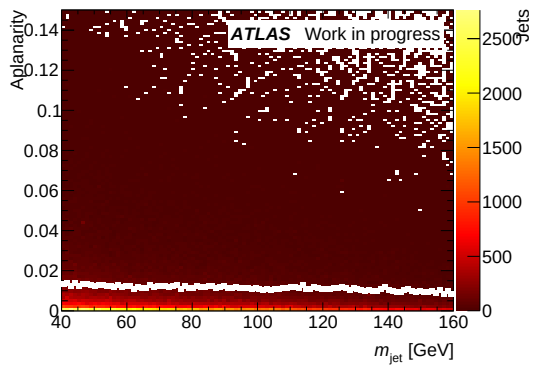
(d) $\mathcal{D}_{12}^{excl}$ versus m_{jet}



(e) τ_1 versus m_{jet}



(f) τ_2 versus m_{jet}

(g) τ_3 versus m_{jet} (h) Planar flow versus m_{jet} (i) Width versus m_{jet} (j) m_{j_1} versus m_{jet} (k) m_{j_2} versus m_{jet} (l) $\Delta R_{j_1,j_2}$ versus m_{jet} (m) Sphericity versus m_{jet} (n) Aplanarity versus m_{jet}

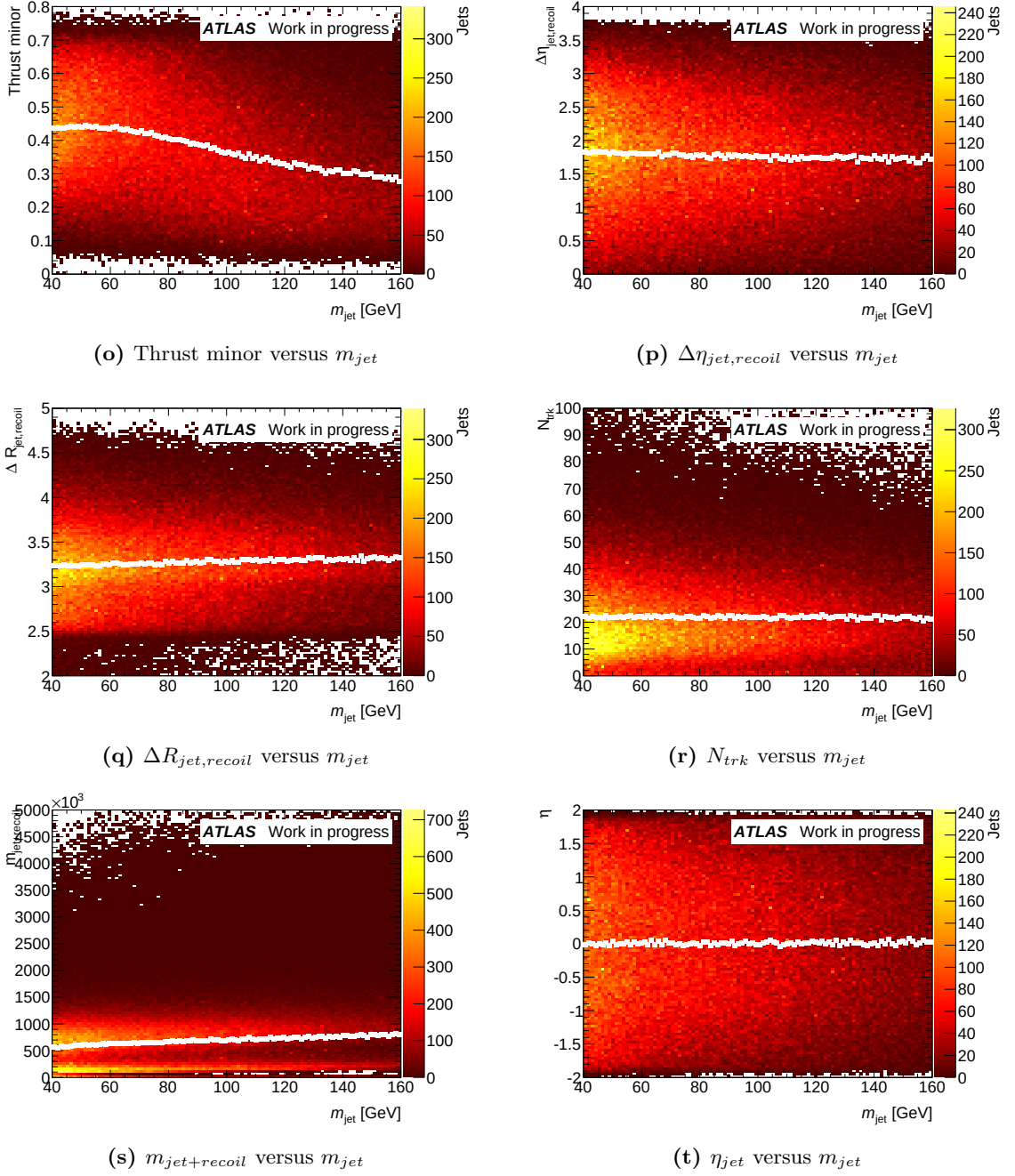


Figure D.1: BDT input variables plotted against the filtered jet mass in background MC to show the correlations. The white dots represent the profile of the scatter plots

Additional plots and figures

Parton Distribution Functions

MSTW 2008 NLO PDFs (68% C.L.)

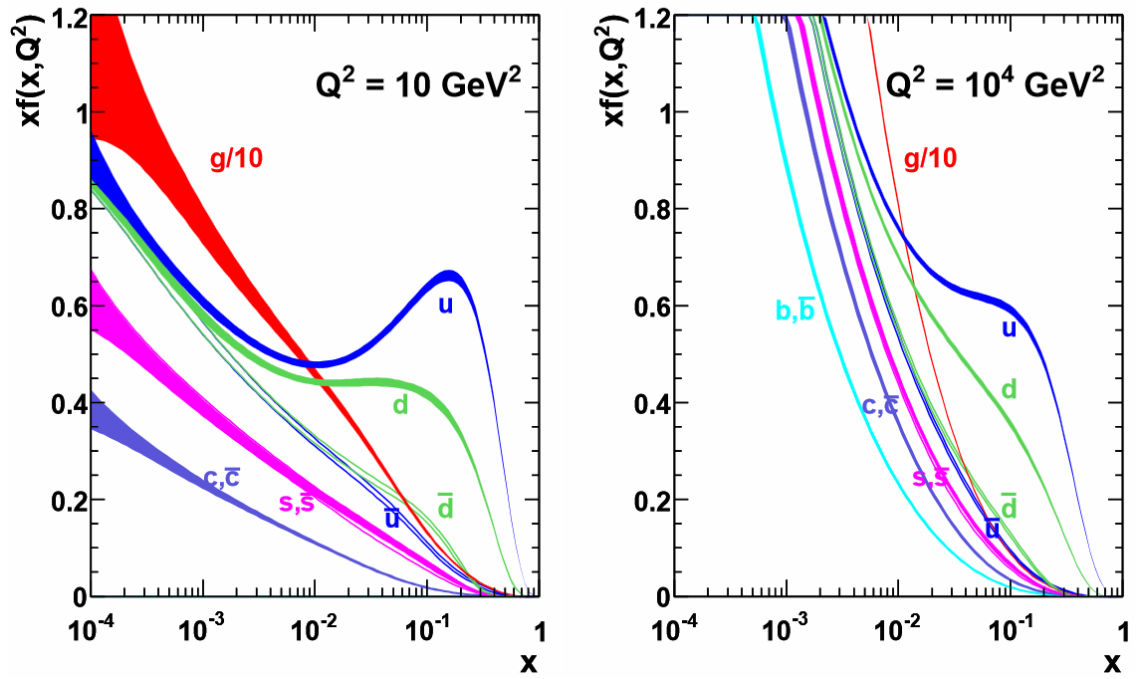


Figure E.1: MSTW 2008 NLO PDFs. The figure was taken from [158].

Asymmetry in the fat jet η distribution

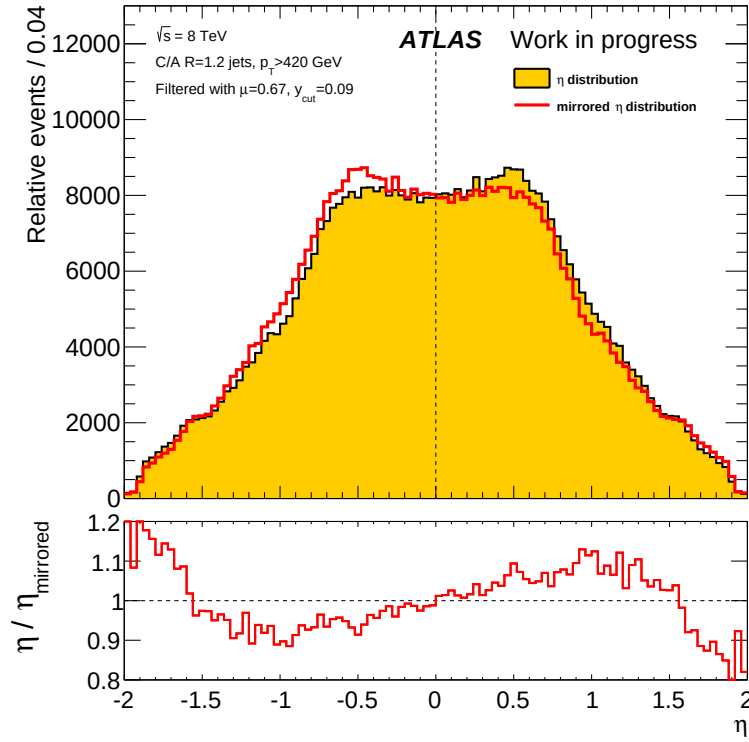


Figure E.2: The asymmetry of the fat jet η distribution is shown by plotting the mirrored distribution on top. The lower panel shows the ratio between regular and mirrored distribution.

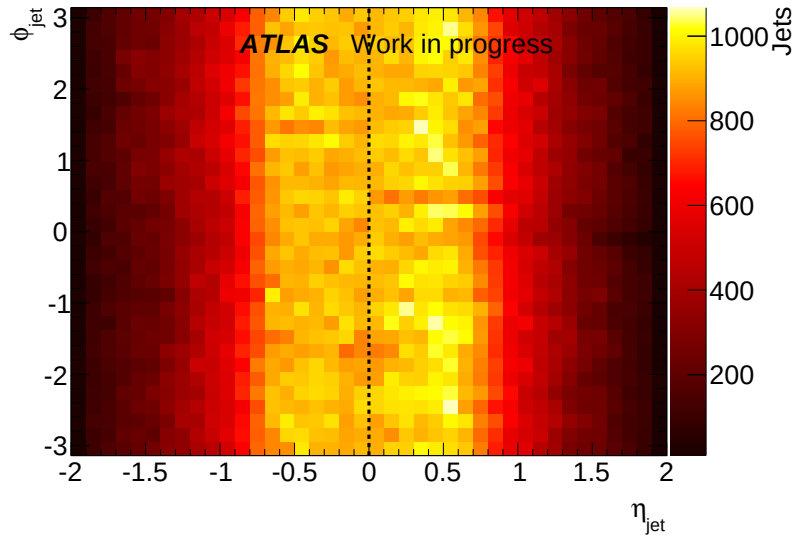


Figure E.3: The distribution of fat jets throughout the η - ϕ plane is shown.

Systematic uncertainty

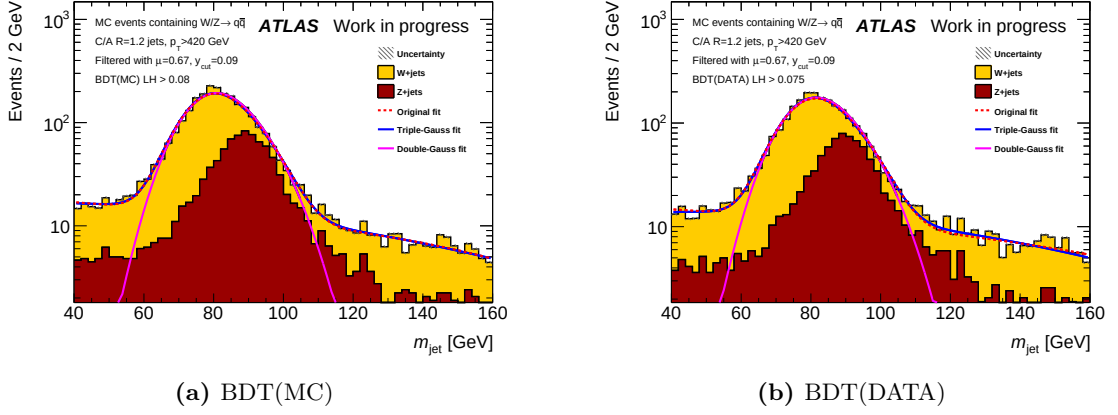


Figure E.4: Signal MC after application of the BDT response cuts. The signal model used to fit the data in chapters 6.7.2 and 6.7.3 is drawn as dashed red line. Two alternative fit models using a third much wider Gaussian function instead of the straight line and a pure double-Gaussian are tried.

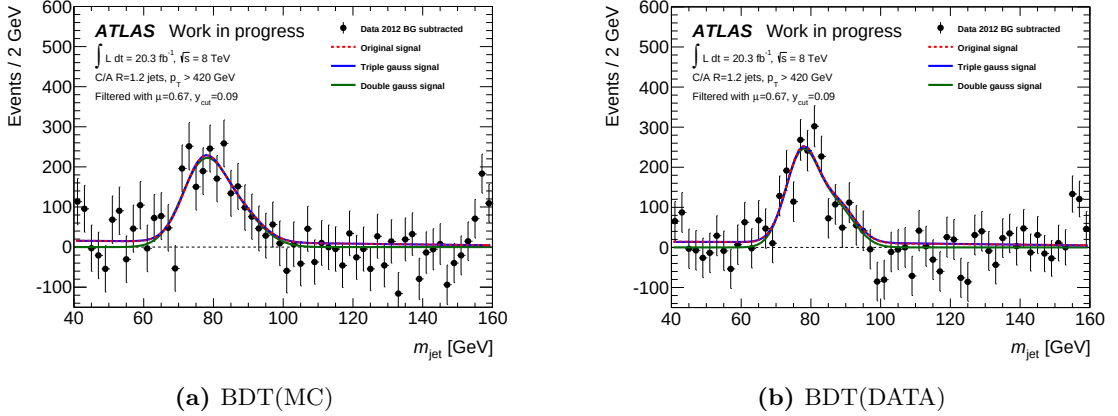


Figure E.5: Data with the background subtracted for the fits performed in chapters 6.7.2 and 6.7.3 as well as the alternative signal models from chapter 6.8.1.

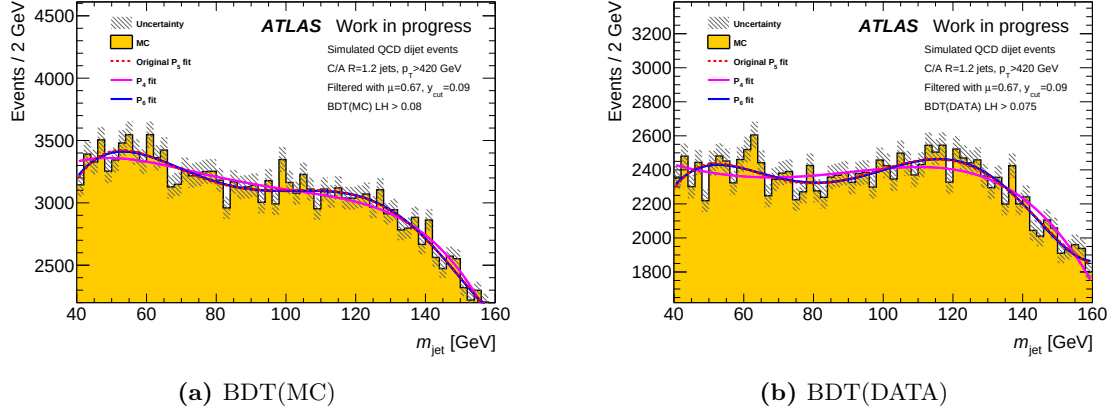


Figure E.6: Background MC after application of the BDT response cuts. While in the original fit a fifth order polynomial function was used to describe the background we use fourth and sixth degree polynomial functions to estimate the systematic error on the signal yield.

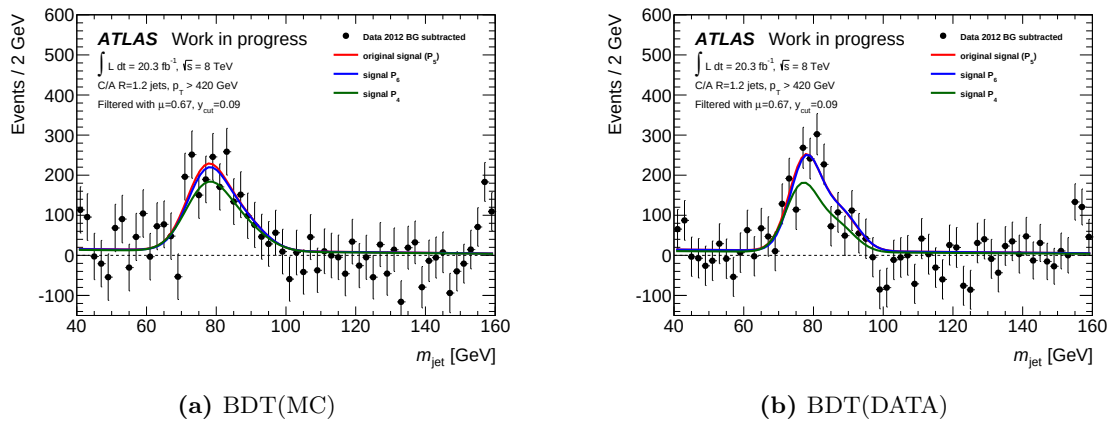


Figure E.7: Data with the background subtracted for three different background models. Where P_n stands for an n -degree polynomial that was used to model the background.

List of Figures

2.1	Overview over all fundamental particles in the Standard Model	6
2.2	The Higgs potential in three dimensional space	12
2.3	Examples of Feynman diagrams	13
3.1	Peak luminosity reached over the course of LHC's first three years	17
3.2	LHC delivered and ATLAS recorded integrated luminosity	17
3.3	Overview over the CERN accelerator complex	18
3.4	Computer-generated model of the ATLAS detector	19
3.5	Working principle of the ATLAS detector	20
3.6	Overview of the ATLAS ID	21
3.7	Cutaway diagram of the ATLAS calorimeter systems	22
3.8	Sketch of a module of the ATLAS LAr electromagnetic calorimeter	24
3.9	Layout of the he ATLAS Muon Spectrometer and its subsystems	25
4.1	Illustration of the formation of a jet	28
4.2	Dijet and Multijet events recorded by CMS	28
4.3	Illustration of infra-red and collinear safety	29
4.4	Diagram of a deep inelastic scattering process	31
4.5	The two stages of the mass-drop filtering procedure	34
4.6	Illustration of the jet trimming algorithm	35
4.7	Illustration of the jet pruning algorithm	36
5.1	Hadronic decay product ΔR versus boson p_T	42
5.2	Comparison of the filtered and unfiltered jet mass spectrum	42
5.3	Filtered and unfiltered jet mass versus number of vertices	43
5.4	Ratio between truth and reconstructed jet mass versus η and p_T	44
5.5	reconstructed jet mass and width versus truth	45
5.6	Track p_T as function of ΔR	46
5.7	Separation power of jet and subjet charge	47
5.8	p_T weighted jet and subjet charge distributions	47
5.9	p_T and ΔR weighted jet and subjet charge distributions	48
6.1	Trigger efficiency as a function of jet p_T	50
6.2	Weighted and unweighted p_T spectra of the MC background	51
6.3	Unfiltered and filtered jet mass spectrum	54
6.4	Filtered jet mass spectrum background fit and signal fit on MC	55
6.5	Fit of the filtered jet mass spectrum	56
6.6	Separation power of the N-subjettiness variables	58

6.7	Separation power of the jet width and planar flow variables	59
6.8	Illustration of colour connections in signal and background	60
6.9	Separation power of the dipolarity variables	61
6.10	Distribution of the jet constituents in different reference frames	62
6.11	Separation power of the event shape variables	63
6.12	Separation power of $\Delta R_{j_1, j_2}$ and N_{trk}	64
6.13	Separation power of recoiljet and subjet mass variables	66
6.14	Mass spectrum of the filtered fat jets used for MVA training in data	67
6.15	ROC-curves for various MVA methods	68
6.16	BDT response versus filtered jet mass	68
6.17	Linear correlation coefficients for BDT(MC) and BDT(DATA)	70
6.18	$\Delta R_{j_1, j_2}$ and $\mathcal{D}_{12}$ versus filtered jet mass	71
6.19	Performance of the reduced BDTs	72
6.20	Performance of the reduced BDTs	73
6.21	Correction of the BDT response values	74
6.22	Comparison between data and MC in a blinded study	76
6.23	Background fits	77
6.24	Signal Fit	78
6.25	Final result for the BDT(MC)	80
6.26	Final result for the BDT(DATA)	81
6.27	Data with the background subtracted	81
6.28	Alternative signal fit for systematics estimation	84
6.29	Alternative background fit for systematics estimation	86
7.1	Higgs production and decay versus Higgs mass	89
7.2	Invariant mass spectrum for $WZ \rightarrow 4q$	90
A.1	Definition of sagitta and helix trajectory	93
B.1	Material in front and in the EM calorimeter	97
B.2	Material in front and in E-CAL and H-CAL	97
C.1	Schematic illustration of a BDT	100
D.1	BDT input variables versus filtered jet mass	105
E.1	MSTW 2008 NLO PDFs	107
E.2	Fat jet distribution in η	108
E.3	Fat jet distribution in the η - ϕ plane	108
E.4	Comparison signal MC fit with different signal models	109
E.5	Background subtracted data with and different signal models	109
E.6	Different background models on BG MC	110
E.7	Background subtracted data with different background models	110



List of Tables

3.1	Summary of the main properties of the LHC	18
6.1	Data taking periods and integrated luminosity	50
6.2	Summary of MC samples	52
6.3	Quality criteria for all narrow jets	53
6.4	Cutflow	54
6.5	BDT input variables and their correlations	71
6.6	Number of events passing BDT cut in data and MC	75
6.7	Fit variables with descriptions and values	79
6.8	Summary of the systematic uncertainties	83
6.9	Fit evaluation of different background models	85
B.1	Main parameters of the calorimeter system	96



Bibliography

- [1] CERN *ECFA-CERN Workshop on Large Hadron Collider in the LEP Tunnel*, Geneva, CERN 1984.
- [2] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (1964) 321–323.
- [3] P. W. Higgs “Broken symmetries, massless particles and gauge fields”, *Phys. Lett.* **12** (1964) 132–133.
- [4] P. W. Higgs “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (1964) 508–509.
- [5] **UA1** Collaboration “Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV”, *Physics Letters B* **122** (1983) 1 103 – 116.
- [6] **UA2** Collaboration “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider”, *Physics Letters B* **122** (1983) 5Ü6 476 – 485.
- [7] **UA1** Collaboration “Experimental observation of lepton pairs of invariant mass around 95 GeV/c² at the CERN SPS collider”, *Physics Letters B* **126** (1983) 5 398 – 410.
- [8] **UA2** Collaboration “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ collider”, *Physics Letters B* **129** (1983) 1Ü2 130 – 140.
- [9] **ATLAS Collaboration** Collaboration *ATLAS: letter of intent for a general-purpose pp experiment at the large hadron collider at CERN*. Letter of Intent. CERN Geneva 1992.
- [10] **CMS** Collaboration “Letter of intent: by the CMS Collaboration for a general purpose detector at LHC”, Tech. Rep. CERN-LHCC-92-003. LHCC-I-1 CERN Geneva 1992. Open presentation to the LHCC 5 November 1992, M. Della Negra/CERN, CMS Spokesman.
- [11] **CDF** Collaboration “Observation of Top Quark Production in $P\bar{p}$ -P Collisions”, *Phys. Rev. Lett.* **74** (1995) 2626–2631 [arXiv:hep-ex/9503002](#) [[hep-ex](#)].
- [12] **DØ** Collaboration “Observation of the Top Quark”, *Phys. Rev. Lett.* **74** (1995) 2632–2637 [arXiv:hep-ex/9503003](#) [[hep-ex](#)].
- [13] M. Bajko *et. al.* “Report of the Task Force on the Incident of 19th September 2008 at the LHC”, Tech. Rep. LHC-PROJECT-Report-1168. CERN-LHC-PROJECT-Report-1168 CERN Geneva Mar, 2009.
- [14] **ATLAS Collaboration** Collaboration “Observation of an Excess of Events in the Search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, Tech. Rep. ATLAS-CONF-2012-093 CERN Geneva Jul, 2012.

- [15] **CMS Collaboration** Collaboration “*Observation of a new boson with a mass near 125 GeV*”, Tech. Rep. CMS-PAS-HIG-12-020 CERN Geneva 2012.
- [16] J. Beringer *et. al.* (Particle Data Group), *Phys. Rev. D* **86** (2012) 010001.
- [17] S. L. Glashow “*Partial-Symmetries of Weak Interactions*”, *Nucl. Phys.* **22** (1961) 579-588.
- [18] S. Weinberg “*A Model of Leptons*”, *Phys. Rev. Lett.* **19** (1967) 1264-1266.
- [19] A. Salam “*Weak and Electromagnetic Interactions*”, *Conf.Proc.* **C680519** (1968) 367–377.
- [20] O. Eberhardt, G. Herbert, H. Lacker *et. al.* “*Impact of a Higgs boson at a mass of 126 GeV on the standard model with three and four fermion generations*”, *Phys. Rev. Lett.* **109** (2012) 241802, [arXiv:1209.1101 \[hep-ph\]](#).
- [21] M. Gonzalez-Garcia and Y. Nir, “*Neutrino Masses and Mixing: Evidence and Implications*”, *Rev. Mod. Phys.* **75** (2003) 345–402 [arXiv:hep-ph/0202058 \[hep-ph\]](#).
- [22] <http://schools-wikipedia.org/images/3071/307168.png>, retrieved: October 31, 2013.
- [23] **DØ Collaboration** “*Search for Randall-Sundrum Gravitons in Dilepton and Diphoton Final States*”, *Phys. Rev. Lett.* **95** (2005) 091801, [arXiv:0505018 \[hep-ex\]](#).
- [24] **DØ Collaboration** “*Search for Randall-Sundrum Gravitons in the Dielectron and Diphoton Final States with 5.4 fb^{-1} of Data from $p\bar{p}$ Collisions at $\sqrt{s} = 1.96\text{ TeV}$* ”, *Phys. Rev. Lett.* **104** (2010) 241802, [arXiv:1004.1826 \[hep-ex\]](#).
- [25] **ATLAS Collaboration** “*Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC*”, *Phys. Lett. B* **716** (2012) 1-29, [arXiv:1207.7214 \[hep-ex\]](#).
- [26] **CMS Collaboration** “*Observation of a New boson at a mass of 125 GeV with the CMS experiment at the LHC*”, *Phys. Lett. B* **716** (2012) 30, [arXiv:1207.7235 \[hep-ex\]](#).
- [27] C. O’Lunaigh “*New results indicate that new particle is a Higgs boson*”, CERN public website March, 2013.
- [28] **ATLAS Collaboration** “*Combined measurements of the mass and signal strength of the Higgs-like boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data*”, CERN , Geneva, March, 2013. ATLAS-CONF-2013-014.
- [29] M. Srednicki *Quantum Field Theory*. Cambridge University Press 2007. sixth edition.
- [30] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley and sons 1984. first edition.
- [31] S. Weinberg *The Quantum Theory of Fields, Volume 1: Foundations*. Cambridge University Press 2005. eighth edition.
- [32] http://laplacian.files.wordpress.com/2010/05/600px-mexican_hat_potential_polar-svg.png, retrieved: July 10, 2014.
- [33] A. Djouadi “*The Anatomy of Electro-Weak Symmetry Breaking. II: The Higgs bosons in the Minimal Supersymmetric Model*”, *Phys. Rept.* **459** (2008) 1–241 [arXiv:hep-ph/0503173 \[hep-ph\]](#).
- [34] C. Csaki, J. Hubisz and P. Meade, “*TASI Lectures on Electroweak Symmetry Breaking from Extra Dimensions*”, Lectures given at the Theoretical Advanced Study Institute, University of Colorado 2004.

- [35] G. Panico, M. Serone and A. Wulzer, “A Model of Electroweak Symmetry Breaking from a Fifth Dimension”, *Nucl. Phys. B* **739** (2006) 186–207 [arXiv:hep-ph/0510373](#) [[hep-ph](#)].
- [36] M. Carena, A. D. Medina, B. Panes, N. R. Shan and C. E. Wagner, “Collider Phenomenology of Gauge-Higgs Unification Scenarios in Warped Extra Dimensions”, *Phys. Rev. D* **77** (2008) 076003, [arXiv:0712.0095](#) [[hep-ph](#)].
- [37] N. Arkani-Hamed, A. Cohen, E. Katz and A. Nelson, “The Littlest Higgs”, *JHEP* **0207** (2002) 034, [arXiv:hep-ph/0206021](#) [[hep-ph](#)].
- [38] K. Kumericki “Feynman Diagrams For Beginners”, Notes for the exercises at the Adriatic School on Particle Physics and Physics Informatics, Split, Croatia 2001.
- [39] <http://www.t2.ucsd.edu/twiki2/bin/view/HEPProjects/HiggsToZZTo4l>, retrieved: July 30, 2014.
- [40] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole and P. Proudlock, *LHC Design Report: Vol. I The LHC Main Ring*. CERN Geneva 2004.
- [41] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole and P. Proudlock, *LHC Design Report: Vol. II The LHC Infrastructure and General Services*. CERN Geneva 2004.
- [42] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole and P. Proudlock, *LHC Design Report: Vol. III The LHC Injector Chain*. CERN Geneva 2004.
- [43] *LEP design report*. CERN Geneva 1983. By the LEP Injector Study Group.
- [44] S. Holmes, R. S. Moore and V. Shiltsev, “Overview of the Tevatron Collider Complex: Goals, Operations and Performance”, *JINST* **6** (2011) T08001, [arXiv:1106.0909](#) [[physics.acc-ph](#)].
- [45] L. Rosselet “The Heavy-Ion Physics Programme with the ATLAS Detector”, *J. Phys.: Conf. Ser.* **110** (2008) 032015.
- [46] D. Lincoln *The Quantum Frontier: The Large Hadron Collider*. Johns Hopkins University Press Baltimore, Maryland 2009.
- [47] M. Furman “The Møller Luminosity Factor”, *LBNL-53553, CBPP Note-543* (2003).
- [48] S. White *Determination of the absolute luminosity at the LHC*. PhD thesis University of Paris-Sud 2010.
- [49] F. L. Sterzo *Search for the Standard Model Higgs Boson in the $H \rightarrow ZZ^* \rightarrow l^+l^-q\bar{q}$ final state with the ATLAS detector at the LHC*. PhD thesis University of Rome 2012.
- [50] L. Evans and P. Bryant, “LHC Machine”, *JINST* **3** (2008) S08001.
- [51] K. Grimm, E. Guido, F. Meloni, S. P. Griso, K. Prokofiev, M. Rudolph and A. Wildauer, “Methods to quantify the performance of the primary vertex reconstruction in the ATLAS experiment under high luminosity conditions”, *J. Phys. Conf. Ser.* **396** (2012) 022041.
- [52] ATLAS Collaboration “Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC”, *Phys. Lett. B* **726** (2013) 88–119 [arXiv:1307.1427](#) [[hep-ex](#)].
- [53] L. Ponce “LHC Machine Upgrades”, Presentation slides, Annecy, France April, 2013 <https://indico.in2p3.fr/getFile.py/access?contribId=81&sessionId=13&resId=0&materialId=slides&confId=6838>.
- [54] <http://home.web.cern.ch/about/accelerators>, retrieved: August 07, 2014.

- [55] <http://bigscience.web.cern.ch/bigscience/Objects/LHC/accelerator.jpg>, retrieved: November 07, 2013.
- [56] **ATLAS** Collaboration *ATLAS detector and physics performance: Technical Design Report Volume I*. Technical Design Report ATLAS. CERN Geneva 1999.
- [57] **ATLAS** Collaboration *ATLAS detector and physics performance: Technical Design Report Volume II*. Technical Design Report ATLAS. CERN Geneva 1999.
- [58] **ATLAS** Collaboration “*The ATLAS Experiment at the CERN Large Hadron Collider*”, *JINST* **3** (2008) S08003.
- [59] Homepage of the ATLAS Experiment: <http://www.atlas.ch>, retrieved: November 08, 2013.
- [60] **ATLAS** Collaboration *ATLAS magnet system: Technical Design Report*. Technical Design Report ATLAS. CERN Geneva 1997.
- [61] **ATLAS** Collaboration *ATLAS central solenoid: Technical Design Report*. Technical Design Report ATLAS. CERN Geneva 1997.
- [62] **ATLAS** Collaboration *ATLAS barrel toroid: Technical Design Report*. Technical Design Report ATLAS. CERN Geneva 1997.
- [63] **ATLAS** Collaboration *ATLAS end-cap toroids: Technical Design Report*. Technical Design Report ATLAS. CERN Geneva 1997.
- [64] **ATLAS Collaboration** Collaboration *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. CERN Geneva 1997.
- [65] **ATLAS Collaboration** Collaboration *ATLAS inner detector: Technical Design Report, 2*. Technical Design Report ATLAS. CERN Geneva 1997.
- [66] **ATLAS** Collaboration *ATLAS pixel detector: Technical Design Report*. Technical Design Report ATLAS. CERN Geneva 1998.
- [67] F. Huegging “*The ATLAS Pixel Detector*”, *IEEE Trans.Nucl.Sci.* **53** (2006) 1732-1736, [arXiv:physics/0412138v2](https://arxiv.org/abs/physics/0412138v2) [[physics.ins-det](#)].
- [68] A. Ahmad, Z. Albrechtskirchinger, P. Allport *et. al.* “*The Silicon Microstrip Sensors of the ATLAS SemiConductor Tracker*”, Tech. Rep. ATL-INDET-PUB-2007-007. ATL-COM-INDET-2007-008. CERN-ATL-COM-INDET-2007-008. 1 CERN Geneva Mar, 2007.
- [69] **ATLAS TRT** Collaboration E. Abat, T. N. Addy, T. P. A. Akesson *et. al.* “*The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance*”, *J. Instrum.* **3** (2008) P02013.
- [70] **ATLAS Collaboration** Collaboration *ATLAS calorimeter performance: Technical Design Report*. Technical Design Report ATLAS. CERN Geneva 1996.
- [71] **ATLAS Collaboration** Collaboration *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN Geneva 1996.
- [72] **ATLAS Collaboration** Collaboration *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN Geneva 1996.
- [73] **ATLAS** Collaboration “*Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions*”, *Eur. Phys. J.* **C70** (2010) 723-753, [arXiv:0912.2642](https://arxiv.org/abs/0912.2642) [[physics.ins-det](#)].

- [74] B. Dowler, J. L. Pinfold *et. al.* “Performance of the ATLAS Hadronic End-Cap Calorimeter in Beam Tests”, Tech. Rep. ATL-LARG-2001-019 Max-Planck Inst. München Oct, 2001.
- [75] **ATLAS Collaboration** Collaboration *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN Geneva 1997. distribution.
- [76] G. N. Hanninger *Observation of $W \rightarrow \tau \nu_\tau$ Decays with the ATLAS Experiment*. PhD thesis Rheinischen Friedrich-Wilhelms-Universität zu Bonn 2011.
- [77] C. D. Burgard “Study of Higgs boson production via Vector Boson Fusion in the $H \rightarrow W^\mp W^\pm (*) \rightarrow l^- \bar{\nu} l'^+ \nu'$ decay mode with the ATLAS Detector”, Master’s thesis Albert-Ludwigs-Universität Freiburg 2013.
- [78] T. Sjöstrand “Monte Carlo Generators”, CERN-LCGAPP-2006-06 (2006) [arXiv:hep-ph/0611247](https://arxiv.org/abs/hep-ph/0611247) [hep-ph].
- [79] A. Buckley, J. Butterworth *et. al.* “General-purpose event generators for LHC physics”, CERN-PH-TH-2010-298 (2011) [arXiv:1101.2599](https://arxiv.org/abs/1101.2599) [hep-ph].
- [80] E. Barberio, J. Boudreau *et. al.* “Fast Simulation of Electromagnetic Showers in the ATLAS calorimeter: Frozen Showers”, *J. Phys.: Conf. Ser.* **160** (2009) 012082.
- [81] R. P. Feynman “The Behavior of Hadron Collisions at Extreme Energies”, in *Special Relativity and Quantum Theory* pp. 289–304 1969.
- [82] A. Ali and G. Kramer, “Jets and QCD: A Historical Review of the Discovery of the Quark and Gluon Jets and its Impact on QCD”, *Eur. Phys. J. H* **36** (2011) 245–326 [arXiv:1012.2288](https://arxiv.org/abs/1012.2288) [hep-ph].
- [83] M. Bowler “ $e^+ e^-$ Production of Heavy Quarks in the String Model”, *Z. Phys. C* **11** (1981) 169.
- [84] B. Andersson, G. Gustafson and B. Soderberg, “A General Model for Jet Fragmentation”, *Z. Phys. C* **20** (1983) 317.
- [85] T. Sjostrand “The Lund Monte Carlo for Jet Fragmentation and e^+e^- Physics: Jetset Version 6.2”, *Comput. Phys. Commun.* **39** (1986) 347–407.
- [86] B. Andersson, G. Gustafson and B. Soderberg, “A Probability Measure on Parton and String States”, *Nucl. Phys. B* **264** (1986) 29.
- [87] B. Andersson, G. Gustafson, G. Ingelman and T. Sjöstrand, “Parton fragmentation and string dynamics”, *Physics Reports* **97** (1983) 2Ü3 31 – 145.
- [88] “Simulation of QCD jets including soft gluon interference”, *Nuclear Physics B* **238** (1984) 1 1 – 29.
- [89] “A QCD model for jet fragmentation including soft gluon interference”, *Nuclear Physics B* **238** (1984) 3 492 – 528.
- [90] “Monte Carlo simulation of general hard processes with coherent QCD radiation”, *Nuclear Physics B* **310** (1988) 461.
- [91] G. Ballochi and R. Odorico, “The string effect and independent fragmentation models: Lore and facts”, *Nuclear Physics B* **345** (1990) 1 173 – 185.
- [92] http://www.imperialhep.blogspot.dk/2011_08_01_archive.html, retrieved: July 16, 2014.
- [93] G. Hanson *et. al.* “Evidence for Jet Structure in Hadron Production by e^+e^- Annihilation”, *Phys. Rev. Lett.* **35** (1975) 1609–1612.

- [94] G. Hanson *et. al.* “Hadron production by e^+e^- annihilation at centre-of-mass energies between 2.6 and 7.8 GeV ; 2, jet structure and related inclusive distributions”, *Phys. Rev. D* **26** (1981) LBL-13887. SLAC-PUB-2855 991–1012. 80 p.
- [95] CMS Collaboration “Search for Narrow Resonances using the Dijet Mass Spectrum with 19.6 fb^{-1} of pp collisions at $\sqrt{s} = 8\text{ TeV}$ ”, CMS PAS EXO-12-059 2013.
- [96] CMS Collaboration “Search for Black Holes in pp collisions at $\sqrt{s} = 7\text{ TeV}$ ”, CMS PAS EXO-11-021 2011.
- [97] M. Seymour “Jet shapes in hadron collisions: Higher orders, resummation and hadronization”, *Nucl. Phys. B* **513** (1998) 269–300 [arXiv:hep-ph/9707338](#) [[hep-ph](#)].
- [98] G. P. Salam “Towards Jetography”, [arXiv:0906.1833](#) [[hep-ph](#)].
- [99] G. F. Sterman and S. Weinberg, “Jets from Quantum Chromodynamics”, *Phys. Rev. Lett.* **39** (1977) 1436.
- [100] G. P. Salam and G. Soyez, “A practical Seedless Infrared-Safe Cone jet algorithm”, *JHEP* **0705** (2007) 086, [arXiv:0704.0292](#) [[hep-ph](#)].
- [101] JADE Collaboration “Experimental Studies on Multi-Jet Production in e^+e^- Annihilation at PETRA Energies”, *Z. Phys. C* **33** (1986) 23.
- [102] JADE Collaboration “Experimental Investigation of the Energy Dependence of the Strong Coupling Strength”, *Phys. Lett. B* **213** (1988) 235.
- [103] S. D. Ellis and D. E. Soper, “Successive Combination Jet Algorithm For Hadron Collisions”, *Phys. Rev. D* **48** (1993) 3160–3166 [arXiv:hep-ph/9305266](#) [[hep-ph](#)].
- [104] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock and B. Webber, “New clustering algorithm for multi-jet cross-sections in e^+e^- annihilation”, *Phys. Lett. B* **269** (1991) 432.
- [105] S. Catani, Y. L. Dokshitzer, M. Seymour, and B. Webber, “Longitudinally invariant $k_\perp$ clustering algorithms for hadron hadron collisions”, *Nucl. Phys. B* **406** (1993) 187.
- [106] M. Cacciari, G. P. Salam and G. Soyez, “The anti- k_t jet clustering algorithm”, *JHEP* **0804** (2008) 063, [arXiv:0802.1189](#) [[hep-ph](#)].
- [107] Y. Dokshitzer, G. Leder, S. Moretti and B. Webber, “Better Jet Clustering Algorithms”, *JHEP* **8** (1997) 001, [arXiv:9707323](#) [[hep-ph](#)].
- [108] M. Wobisch and T. Wengler, “Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering”, [arXiv:9907280](#) [[hep-ph](#)].
- [109] M. Rubin A New LHC Search Channel for a Light Higgs Boson and Associated QCD Calculations. PhD thesis Paris Universitat, Laboratoire de physique theorique et hautes energies 201.
- [110] L3 Collaboration “Inclusive Jet Production in Two-Photon Collisions at LEP”, *Phys. Lett. B* **602** (2004) 157–166 [arXiv:hep-ex/0410012](#) [[hep-ex](#)].
- [111] J. Bjorken “Asymptotic Sum Rules at Infinite Momentum”, *Phys. Rev.* **179** (1969) 1547–1553.
- [112] J. Stirling “Parton Distribution Functions and the LHC”, Presentation slides, Cambridge University <http://lhc.fuw.edu.pl/Stirling.pdf>.
- [113] V. N. Gribov and L. N. Lipatov, “Deep inelastic ep scattering in perturbation theory”, *Sov. J. Nucl. Phys.* **15** (1972) 4 438–450.

- [114] L. Lipatov “*The parton model and perturbation theory*”, *Sov. J. Nucl. Phys.* **20** (1975) 94–102.
- [115] G. Altarelli and G. Parisi, “*Asymptotic Freedom in Parton Language*”, *Nucl. Phys.* **B 126** (1977) 298.
- [116] Y. L. Dokshitzer “*Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics*”, *Sov. Phys. JETP* **46** (1977) 641–653.
- [117] A. Martin, R. Roberts, W. Stirling and R. Thorne, “*NNLO global parton analysis*”, *Phys. Lett.* **B 531** (2002) 216–224 [arXiv:hep-ph/0201127](#) [[hep-ph](#)].
- [118] R. D. Ball, V. Bertone *et. al.* “*Parton distributions with LHC data*”, *Nucl. Phys.* **B 867** (2012) 244–289 [arXiv:1207.1303](#) [[hep-ph](#)].
- [119] J. Pumplin, D. Stump, J. Huston, H. Lai, P. Nadolsky and W. Tung, “*New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*”, *JHEP* **0207** (2002) 012, [arXiv:hep-ph/0201195](#) [[hep-ph](#)].
- [120] J. Campbell, J. Huston and W. Stirling, “*Hard Interactions of Quarks and Gluons: a Primer for LHC Physics*”, *Rept.Prog.Phys.* **70** (2007) 89 [arXiv:hep-ph/0611148](#) [[hep-ph](#)].
- [121] J. Engelen and P. Kooijman, “*Deep inelastic scattering at HERA: A review of experimental results in the light of quantum chromodynamics*”, *Progress in Particle and Nuclear Physics* **41** (1998) 1–47.
- [122] B. Badelek “*Review of fixed target experimental results on the structure of the proton*”, *Acta Phys.Polon.* **B27** (1995) 1145–1173.
- [123] **H1 & ZEUS** Collaboration R. Placakyte “*Parton Distribution Functions*”, [arXiv:1111.5452](#) [[hep-ph](#)]. proceedings of XXXI Physics in Collision, Vancouver, BC Canada.
- [124] S. Albino, F. Anulli *et. al.* “*Parton fragmentation in the vacuum and in the medium*”, *Mini-Proceedings Workshop ECT**, Trento (2008) [arXiv:0804.2021](#) [[hep-ph](#)].
- [125] F. Arleo “*(medium-modified) Fragmentation Functions*”, *Eur.Phys.J.* **C61** (2009) 603–627 [arXiv:0810.1193](#) [[hep-ph](#)].
- [126] M. Cacciari and G. P. Salam, “*Pileup subtraction using jet areas*”, *Phys. Lett.* **B 659** (2008) 119–126 [arXiv:0707.1378](#) [[hep-ph](#)].
- [127] T. Sjostrand and M. van Zijl, “*A Multiple Interaction Model for the Event Structure in Hadron Collisions*”, *Phys.Rev.* **D 36** (1987) 2019.
- [128] J. Butterworth, J. Forshaw and M. Seymour, “*Multiparton Interactions in Photoproduction at HERA*”, *Z. Phys.* **C 72** (1996) 637–646 [arXiv:hep-ph/9601371](#) [[hep-ph](#)].
- [129] T. Sjostrand and P. Z. Skands, “*Transverse-momentum-ordered showers and interleaved multiple interactions*”, *Eur. Phys. J.* **C 39** (2005) 129–154 [arXiv:hep-ph/0408302](#) [[hep-ph](#)].
- [130] R. Corke and T. Sjöstrand, “*Multiparton Interactions and Rescattering*”, *JHEP* **1001** (2010) 035, [arXiv:0911.1909](#) [[hep-ph](#)].
- [131] M. Bähr, S. Gieseke and M. H. Seymour, “*Simulation of multiple partonic interactions in Herwig++*”, *JHEP* **0807** (2008) 076, [arXiv:0803.3633](#) [[hep-ph](#)].

- [132] Y. Cui, Z. Han and M. D. Schwartz, “*W-jet Tagging: Optimizing the Identification of Boosted Hadronically-Decaying W Bosons*”, *Phys. Rev. D* **83** (2011) 074023, [arXiv:1012.2077v2 \[hep-ph\]](#).
- [133] **ATLAS** Collaboration “*Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*”, [arXiv:1306.4945v1 \[hep-ex\]](#).
- [134] J. Thaler and L.-T. Wang, “*Strategies to Identify Boosted Tops*”, *JHEP* **07** (2008) 092, [arXiv:0806.0023 \[hep-ph\]](#).
- [135] L. G. Almeida, S. J. Lee, G. Perez, I. Sung and J. Virzi, “*Top Jets at the LHC*”, *Phys. Rev. D* **79** (2009) 074012, [arXiv:0810.0934 \[hep-ph\]](#).
- [136] **ATLAS** Collaboration “*Measurement of Jet Mass and Substructure for Inclusive Jets in $\sqrt{s} = 7$ TeV pp Collisions with the ATLAS Experiment*”, Tech. Rep. ATLAS-CONF-2011-073, CERN , Geneva, May, 2011.
- [137] G. D. Kribs, A. Martin, T. S. Roy and M. Spannowsky, “*Discovering the Higgs Boson in New Physics Events using Jet Substructure*”, *Phys. Rev. D* **81** (2010) 111501, [arXiv:0912.4731 \[hep-ph\]](#).
- [138] **OPAL** Collaboration “*Experimental properties of gluon and quark jets from a point source*”, *Eur. Phys. J. C* **11** (1999) 217, [arXiv:9903027 \[hep-ex\]](#).
- [139] A. Abdesselam, E. Kuutmann, U. Bitenc *et. al.* “*Boosted objects: a probe of beyond the standard model physics*”, *Eur. Phys. J. C* **71** (2011) 1661, [arXiv:1012.5412 \[hep-ph\]](#).
- [140] M. Dasgupta, A. Fregoso, S. Marzani and G. P. Salam, “*Towards an understanding of jet substructure*”, *JHEP* **09** (2013) 029, [arXiv:1307.0007 \[hep-ph\]](#).
- [141] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, “*Jet Substructure as a new Higgs search channel at the LHC*”, *Phys. Rev. Lett.* **100** (2008) 242001, [arXiv:0802.2470 \[hep-ph\]](#).
- [142] D. Krohn, J. Thaler and L.-T. Wang, “*Jet Trimming*”, *JHEP* **2** (2010) 84, [arXiv:0912.1342 \[hep-ph\]](#).
- [143] S. D. Ellis, C. K. Vermilion and J. R. Walsh, “*Techniques for improved heavy particle searches with jet substructure*”, *Phys. Rev. D* **80** (2009) 051501, [arXiv:0903.5081 \[hep-ph\]](#).
- [144] **ATLAS** Collaboration “*ATLAS Sensitivity to the Standard Model Higgs in the HW and HZ Channels at High Transverse Momenta*”, Tech. Rep. ATLAS-PHYS-PUB-2009-088, CERN , Geneva, August, 2009.
- [145] T. Plehn, G. P. Salam and M. Spannowsky, “*Fat Jets for a Light Higgs*”, *Phys. Rev. Lett.* **104** (2010) 111801, [arXiv:0910.5472 \[hep-ph\]](#).
- [146] G. D. Kribs, A. Martin, T. S. Roy and M. Spannowsky, “*Discovering Higgs Bosons of the MSSM using Jet Substructure*”, *Phys. Rev. D* **82** (2010) 095012, [arXiv:1006.1656 \[hep-ph\]](#).
- [147] **ATLAS** Collaboration “*Search for Dark Matter in Events with a Hadronically Decaying W or Z Boson and Missing Transverse Momentum in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*”, *Phys. Rev. Lett* **112** (2014) 041802, [arXiv:1309.4017 \[hep-ex\]](#).
- [148] J. R. Walsh and S. Zuberi, “*Factorization Constraints on Jet Substructure*”, [arXiv:1110.5333 \[hep-ph\]](#).

- [149] A. Altheimer *et. al.* “Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks”, *J. Phys.* **G 39** (2012) 063001, [arXiv:1201.0008 \[hep-ph\]](#).
- [150] D. Miller, A. Schwartzman and D. Su, “Pile-up jet energy scale corrections using the jet-vertex fraction method”, Tech. Rep. ATL-COM-2009-180, CERN, Geneva, September, 2009.
- [151] **ATLAS** Collaboration “Search for pair production of massive particles decaying into three quarks with the ATLAS detector in $\sqrt{s} = 7$ TeV *pp* collisions at the LHC”, *JHEP* **1212** (2012) 86, [arXiv:1210.4813v2 \[hep-ex\]](#).
- [152] S. D. Ellis, C. K. Vermilion and J. R. Walsh, “Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches”, *Phys. Rev.* **D 81** (2010) 094023, [arXiv:0912.0033 \[hep-ph\]](#).
- [153] **CMS** Collaboration “Search for heavy resonances in the *W/Z*-tagged dijet mass spectrum in *pp* collisions at 7 TeV”, *Phys. Lett.* **B 723** (2013) 280, [arXiv:1212.1910v2 \[hep-ex\]](#).
- [154] **CMS** Collaboration “Search for anomalous $t\bar{t}$ production in the highly-boosted all-hadronic final state”, *Phys. Lett.* **B 723** (2013) 280, [arXiv:1212.1910v2 \[hep-ex\]](#).
- [155] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. Seymour and R. Webber, “*HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*”, *JHEP* **0101** (2001) 010, [arXiv:hep-ph/0011363 \[hep-ph\]](#).
- [156] T. Sjöstrand “*PYTHIA 8 Status Report*”, [arXiv:0809.0303 \[hep-ph\]](#).
- [157] T. Sjöstrand, S. Mrenna and P. Skands, “A Brief Introduction to *PYTHIA 8.1*”, *Comput.Phys.Commun.* **178** (2008) 852–867 [arXiv:0710.3820 \[hep-ph\]](#).
- [158] A. Martin, W. Stirling, R. Thorne and G. Watt, “Parton distributions for the LHC”, *Eur. Phys. J.* **C 63** (2009) 189–285 [arXiv:0901.0002 \[hep-ph\]](#).
- [159] “Summary of ATLAS Pythia 8 tunes”, Tech. Rep. ATL-PHYS-PUB-2012-003 CERN Geneva Aug, 2012.
- [160] L. Tompkins “*Just the Beginning: The Post-Higgs Discovery LHC*”, Presentation slides March, 2014 http://www-physics.lbl.gov/seminars/old/tompkins_lbnl.pdf.
- [161] W. J. Waalewijn “Calculating the Charge of a Jet”, *Phys. Rev.* **D 86** (2012) 094030, [arXiv:1209.3019 \[hep-ph\]](#).
- [162] D. Krohn, M. D. Schwartz, T. Lin and W. J. Waalewijn, “Calculating the Charge of a Jet”, *Phys. Rev. Lett.* **110** (2013) 212001, [arXiv:1209.2421 \[hep-ph\]](#).
- [163] “Jet Charge Studies with the ATLAS Detector Using $\sqrt{s} = 8$ TeV Proton-Proton Collision Data”, Tech. Rep. ATLAS-CONF-2013-086 CERN Geneva Aug, 2013.
- [164] R. Field and R. Feynman, “A Parametrization of the Properties of Quark Jets”, *Nucl. Phys.* **B 136** (1978) 1.
- [165] **ALEPH** Collaboration “Measurement of triple gauge boson couplings at 172-GeV”, *Phys.Lett.* **B422** (1998) 369–383.
- [166] R. Astaloš “A study of *b*-jet charge structure”, in *Študentská vedecká konferencia FMFI UK* pp. 162–172, Bratislava, 2010.
- [167] J. Berge, D. Bogert, R. Endorf, R. Hanft, J. Malko *et. al.* “Quark Jets from anti-neutrino Interactions. 1. Net Charge and Factorization in the Quark Jets”, *Nucl.Phys.* **B184** (1981) 13.

- [168] ATLAS jet trigger performance 2012:
<https://twiki.cern.ch/twiki/bin/view/Atlas/JetTriggerPerformance2012>,
retrieved: July 02, 2014.
- [169] M. Bahr, S. Gieseke, M. Gigg *et al.* “*Herwig++ Physics and Manual*”, *Eur. Phys. J. C* **58** (2008) 639–707 [arXiv:0803.0883](https://arxiv.org/abs/0803.0883) [hep-ph].
- [170] N. Davidson, T. Przedzinski and Z. Was, “*PHOTOS Interface in C++: Technical and Physics Documentation*”, *CERN-PH-TH/2010-261* [arXiv:1011.0937](https://arxiv.org/abs/1011.0937) [hep-ph].
- [171] N. Davidson, G. Nanava, T. Przedzinski, E. Richter-Was and Z. Was, “*Universal Interface of TAUOLA Technical and Physics Documentation*”, *IFJPAN-IV-2009-10* [arXiv:1002.0543](https://arxiv.org/abs/1002.0543) [hep-ph].
- [172] S. Agostinelli *et al.* “*GEANT4: A Simulation toolkit*”, *Nucl. Instrum. Meth. A* **506** (2003) 250–303.
- [173] ATLAS Collaboration “*The ATLAS Simulation Infrastructure*”, *Eur. Phys. J. C* **70** (2010) 823–874 [arXiv:1005.4568](https://arxiv.org/abs/1005.4568) [physics.ins-det].
- [174] S. Albrand, J. Fulachier and F. Lambert, “*The ATLAS metadata interface*”, *J. Phys.: Conf. Ser.* **219** (2010) 042030.
- [175] ATLAS Collaboration “*Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*”, *Eur. Phys. J. C* **73** (2013) 3, [arXiv:1112.6426](https://arxiv.org/abs/1112.6426) [hep-ex].
- [176] Recommendations for jet cleaning for data 2011:
<https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets2011>,
retrieved: May 13, 2014.
- [177] JetEtMiss group on jet mass calibration:
<https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ApplyJetCalibration2012>,
retrieved: June 18, 2014.
- [178] ATLAS Collaboration “*Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment*”, *JHEP* **1205** (2012) 128, [arXiv:1203.4606](https://arxiv.org/abs/1203.4606) [hep-ex].
- [179] J. Thaler and K. Van Tilburg, “*Identifying Boosted Objects with N-subjettiness*”, *JHEP* **1103** (2011) 015, [arXiv:1011.2268](https://arxiv.org/abs/1011.2268) [hep-ph].
- [180] J. Thaler and K. Van Tilburg, “*Maximizing Boosted Top Identification by Minimizing N-subjettiness*”, *JHEP* **1202** (2012) 093, [arXiv:1108.2701](https://arxiv.org/abs/1108.2701) [hep-ph].
- [181] CMS Collaboration “*Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in pp collisions at $\sqrt{s} = 8$ TeV*”, CMS-EXO-12-024, CERN-PH-EP-2014-071 May, 2014.
- [182] I. Feige, M. D. Schwartz, I. W. Stewart and J. Thaler, “*Precision Jet Substructure from Boosted Event Shapes*”, *Phys. Rev. Lett.* **109** (2012) 092001, [arXiv:1204.3898](https://arxiv.org/abs/1204.3898) [hep-ph].
- [183] ATLAS Collaboration “*ATLAS measurements of the properties of jets for boosted particle searches*”, *Phys. Rev. D* **86** (2012) 072006, [arXiv:1206.5369](https://arxiv.org/abs/1206.5369) [hep-ex].
- [184] L. G. Almeida, S. J. Lee, G. Perez, G. Sterman, I. Sung and J. Virzi, “*Substructure of high- p_T Jets at the LHC*”, *Phys. Rev. D* **79** (2009) 074017, [arXiv:0807.0234](https://arxiv.org/abs/0807.0234) [hep-ph].
- [185] CDF Collaboration “*Study of Substructure of High Transverse Momentum Jets Produced in Proton-Antiproton Collisions at $\sqrt{s} = 1.96$ TeV*”, *Phys. Rev. D* **85** (2012) 091101, [arXiv:1106.5952](https://arxiv.org/abs/1106.5952) [hep-ex].

- [186] J. Gallicchio and M. D. Schwartz, “*Seeing in Color: Jet Superstructure*”, *Phys.Rev.Lett.* **105** (2010) 022001, [arXiv:1001.5027 \[hep-ph\]](#).
- [187] A. Hook, M. Jankowiak and J. G. Wacker, “*Jet Dipolarity: Top Tagging with Color Flow*”, *JHEP* **1204** (2012) 007, [arXiv:1102.1012 \[hep-ph\]](#).
- [188] D. W. Miller and A. Schwartzman, “*Additional substructure observables*”, Presentation slides, SLAC National Accelerator Laboratory and Stanford University May, 2011 <http://indico.cern.ch/event/141019/contribution/3/material/slides/0.pdf>.
- [189] C. Chen “*New approach to identifying boosted hadronically decaying particles using jet substructure in its center-of-mass frame*”, *Phys. Rev. D* **85** (2012) 034007, [arXiv:1112.2567 \[hep-ph\]](#).
- [190] H. R. Quinn and P. F. Harrison, *The BaBar physics book: physics at an asymmetric B factory*. SLAC Stanford, CA 1998.
- [191] S. Chekanov, C. Chen, J. Cochran *et. al.* “*First observation, cross section measurement and jet substructure properties of the inclusive production of boosted hadronically-decaying $W^\pm$ and Z^0 bosons from pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC.*”, ATL-COM-PHYS-2012-1817 2013.
- [192] J. D. Bjorken and S. J. Brodsky, “*Statistical Model for Electron-Positron Annihilation into Hadrons*”, *Phys. Rev. D* **1** (1970) 1416–1420.
- [193] I. Stroescu “*Event Shape Variables*”, Presentation slides, Kirchhoff-Institute for Physics, University of Heidelberg August, 2007 <http://www.kip.uni-heidelberg.de/ion/talks/EventShapeVariables.pdf>.
- [194] S. Brandt, C. Peyrou, R. Sosnowski and A. Wroblewski, “*The principal axis of jets - An attempt to analyse High-energy collisions as two-body processes*”, *Phys. Lett.* **12** (1964) 57–61.
- [195] E. Farhi “*Quantum Chromodynamics Test for Jets*”, *Phys. Rev. Lett.* **39** (1977) 1587–1588.
- [196] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne and H. Voss, “*TMVA: Toolkit for Multivariate Data Analysis*”, *PoS ACAT* (2007) 040, [arXiv:physics/0703039 \[physics.data-an\]](#).
- [197] F. J. Massey “*The Kolmogorov-Smirnov Test for Goodness of Fit*”, *Journal of the American Statistical Association* **46** (1951) 253.
- [198] **ATLAS Collaboration** Collaboration “*Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8$ TeV with ATLAS*”, Tech. Rep. CERN-PH-EP-2014-148. [arXiv:1407.4222](#) CERN Geneva Jul, 2014.
- [199] **ATLAS Collaboration** “*Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*”, *Eur.Phys.J.* **C73** (2013) 2518 [arXiv:1302.4393 \[hep-ex\]](#).
- [200] **CMS Collaboration** Collaboration “*Performance of quark/gluon discrimination in 8 TeV pp data*”, Tech. Rep. CMS-PAS-JME-13-002 CERN Geneva 2013.
- [201] **ATLAS Collaboration** “*Search for a Standard Model Higgs boson in the mass range 200-600 GeV in the $H \rightarrow ZZ \rightarrow llqq$ decay channel*”, *Phys. Lett. B* **717** (2012) 70–88 [arXiv:1206.2443 \[hep-ex\]](#).
- [202] **ATLAS Collaboration** Collaboration “*Search for the Higgs boson in the $H \rightarrow WW \rightarrow l\nu jj$ decay channel at $\sqrt{s} = 7$ TeV with the ATLAS detector*”, *Phys. Lett. B* **718** 391–410 [arXiv:1206.6074 \[hep-ex\]](#).

- [203] A. Djouadi “*Phenomenology of SM and SUSY Higgs bosons at the LHC*”, *Czech. J. Phys.* **55** (2005) B23–B44 [arXiv:hep-ph/0412238](#) [hep-ph].
- [204] S. Dittmaier, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka *et. al.* “*Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*”, [arXiv:1201.3084](#) [hep-ph].
- [205] **ATLAS Collaboration** “*Light-quark and Gluon Jets: Calorimeter Response, Jet Energy Scale Systematics and Properties*”, Tech. Rep. ATLAS-CONF-2012-138 CERN Geneva Sep, 2012.
- [206] **ATLAS Collaboration** “*Identification and Tagging of Double b -hadron jets with the ATLAS Detector*”, Tech. Rep. ATLAS-CONF-2012-100 CERN Geneva Jul, 2012.
- [207] E. Kajomovitz *et. al.* “*Searching for Diboson Resonances in fully hadronic final states*”, Exotics Plenary - May 15th 2014.
- [208] F. Ragusa “*An Introduction to Charged Particles Tracking*”, Presentation slides, Martignano, Italy June, 2006
http://www.physics.iitm.ac.in/sercehep2013/track2_Gagan_Mohanty.pdf.
- [209] L. Breiman, J. Friedman, C. J. Stone and R. Olshen, *Classification and Regression Trees*. Wadsworth Statistics/Probability. Chapman and Hall/CRC 1984. first edition.
- [210] S. Kotsiantis “*Decision trees: a recent overview*”, *Artificial Intelligence review* **39** (2013).
- [211] B. P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu and G. McGregor, “*Boosted Decision Trees as an Alternative to Artificial Neural Networks for Particle Identification*”, *Nucl. Instrum. Meth. A* **543** (2005) [arXiv:physics/0408124](#) [physics.data-an].
- [212] R.-H. Li and G. G. Belford, , eds. *Instability of Decision Tree Classification Algorithms* KDD’02 Proceedings of the eight ACM SIGKDD international conference on Knowledge discovery and data mining ACM 2002.
- [213] Y. Freund and R. E. Schapire, “*A Short Introduction to Boosting*”, in *In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence* pp. 1401–1406 Morgan Kaufmann 1999.
- [214] R. E. Schapire “*The Boosting Approach to Machine Learning: An Overview*”, 2002.