

Investigation of the High Mass Drell Yan Spectrum with ATLAS



Dissertation der Fakultät für Physik

der

Ludwig-Maximilians-Universität München

vorgelegt von

Thomas A. Müller

geboren in München

München, den 26. Juli 2010

Erstgutachter: Prof. Dr. Otmar Biebel
Zweitgutachter: Prof. Dr. Wolfgang Dünneberger
Datum der mündlichen Prüfung: 14. September 2010

*Die Würde des Menschen ist unantastbar.
Sie zu achten und zu schützen ist Verpflichtung aller staatlichen Gewalt. [1]*

Abstract

The Large Hadron Collider started data taking at the end of 2009 and an integrated luminosity of 1 fb^{-1} is hoped for by the end of 2011. A precise measurement of the high mass Drell Yan spectrum offers a good opportunity for a model independent search for new physics. The muon channel is well suited for this, due to the clean signature and the good muon identification in the Muon Spectrometer.

Previous studies at high dimuon masses neglected all background contributions [2]. This study investigated the impact of background on the Drell Yan spectrum and it was found that $t\bar{t}$ decays are the most important contribution.

Various selection cuts to suppress those background contributions were studied. A method to take systematic uncertainties into account, whilst optimising these selection cuts, has been developed. It was shown that two additional selection cuts based on b -tagging and Missing Transverse Energy (\cancel{E}_T) will reduce the overall uncertainty for a bin from 200 GeV to 300 GeV from 19.1% to 17.2% for an integrated luminosity of 50 pb^{-1} .

An important aspect of this analysis is to ensure that the efficiency for all selection cuts remains stable at very high dimuon masses of up to 1 TeV. This is not the case for the conventional \cancel{E}_T , so a derived variable has been introduced and tested.

Zusammenfassung

Der Large Hadron Collider hat Ende 2009 mit der Datennahme begonnen und man hofft auf eine integrierte Luminosität von 1 fb^{-1} bis Ende 2011. Eine präzise Messung des Drell Yan Spektrums bietet eine gute Möglichkeit für eine modellunabhängige Suche nach neuer Physik. Der Myon-Kanal ist für eine derartige Suche aufgrund der klaren Signatur und der guten Myon-Identifikation im Myonenspektrometer gut geeignet.

Bisherige Studien haben die Beiträge des Untergrunds vernachlässigt [2]. In der vorliegenden Studie wurde der Einfluss von Untergrund auf das Drell Yan Spektrum untersucht. Dabei stellte sich heraus, dass der Zerfall von $t\bar{t}$ den wichtigsten Beitrag zum Untergrund liefert.

Um den Untergrund zu reduzieren wurden verschiedene Auswahlkriterien studiert. Es wurde eine Methode zur Optimierung der Auswahlkriterien entwickelt, die auch den Einfluss von systematischen Unsicherheiten berücksichtigt. Durch die Einführung zweier zusätzlicher Auswahlkriterien, basierend auf b -Tagging und fehlender Transversalenergie (\cancel{E}_T), kann der Gesamtfehler im Bereich von 200 GeV bis 300 GeV bei einer integrierten Luminosität von 50 pb^{-1} von 19,1% auf 17,2% reduziert werden.

Ein wichtiger Aspekt dieser Analyse ist es, sicherzustellen, dass die Effizienz für alle Auswahlkriterien auch für sehr hohe Di-Myonen-Massen von bis zu 1 TeV gleichbleibend hoch ist. Für die herkömmliche Definition von \cancel{E}_T ist dies nicht der Fall. Daher wurde eine von \cancel{E}_T abgeleitete Variable eingeführt und getestet.

Contents

1	Introduction	13
2	Theoretical Background	15
2.1	The Standard Model	15
2.2	Theory of Proton-Proton Collisions	17
2.2.1	Parton Model	17
2.2.2	Factorisation	17
2.3	The Drell Yan Process	19
2.3.1	Leading Order cross section for $m_{\mu\mu} \ll m_Z$	19
2.3.2	Next-to-Leading Order corrections	20
2.3.3	Including the Z Boson	20
3	Experimental Setup	21
3.1	The Large Hadron Collider	21
3.1.1	Experiments at LHC	22
3.2	The Atlas Experiment	23
3.2.1	The ATLAS coordinate system	23
3.2.2	Inner Detector	24
3.2.3	Electromagnetic Calorimeter	25
3.2.4	Hadronic Calorimeter	26
3.2.5	Muon Spectrometer	26
3.2.6	Trigger	27
3.3	Physics Objects	28

3.3.1	Jets	28
3.3.2	Muons	29
3.3.3	Missing Transverse Energy	30
3.3.4	b -Tagging	30
3.4	Computing Model	33
3.4.1	ATHENA Framework	34
3.4.2	Monte Carlo Production	34
3.4.3	Grid Computing	35
4	High Mass Drell Yan Spectrum	37
4.1	Backgrounds	37
4.1.1	$t\bar{t}$	37
4.1.2	$b\bar{b}$	39
4.1.3	$Z \rightarrow \tau^+\tau^-$	39
4.1.4	$W \rightarrow \mu\nu$	40
4.2	Monte Carlo Generators	40
4.3	Samples	41
4.4	Existing Analyses	45
4.4.1	Z peak	45
4.4.2	Z'	46
4.4.3	Present Analysis	46
4.5	Systematic Uncertainties	47
5	Cuts	53
5.1	Quality criteria for cuts	53
5.2	Cut Optimisation	54
5.3	List of Cuts	56
5.3.1	Preselection	56
5.3.2	Quality Cuts	57
5.3.3	Isolation Cuts	58
5.3.4	b -Tagging	63

5.3.5	Missing Transverse Energy	71
5.4	Variables Unused in this Analysis	79
5.4.1	Sphericity	79
5.4.2	Jet Energy	81
5.4.3	HT	83
5.5	Significance of Cut Optimisation	83
5.6	Results	84
6	Conclusions and Outlook	88
A	Corrections	90
A.1	Bin-by-Bin Correction	90
A.2	Iterative Bin-by-Bin Correction	91
B	Definitions	96
B.1	Luminosity	96
B.2	Invariant dimuon mass	96
B.3	Sphericity	96
B.4	Transverse Sphericity	97
B.5	Resolution	97
C	Additional Studies	98
C.1	B-Tagging Algorithms	98
C.2	Differences between High Mass Drell Yan and Pythia Sample	99
D	Iterative Cut Optimisation	101
E	Datasets used	102

1 Introduction

Trying to understand the laws of nature has always been a motivation for scientific experiments. The questions “What are we made of?” and “What holds us together?” are the most fundamental of all. The best answer we have is the Standard Model of particle physics, which describes the constituents of matter and three of the four known forces. Developed nearly 50 years ago, its predictive power and the numerous confirmations through experiments make it one of the most successful theories of physics ever.

Nevertheless, there are still open questions unanswered by the Standard Model: The constituents of the Dark Matter, the asymmetry between matter and antimatter and the fine tuning problem are only three prominent examples.

To answer those questions and to find the last predicted particle of the Standard Model, a new experiment of hitherto unknown proportions and complexity has been devised. Roughly 10,000 physicists and engineers from nearly 100 countries have been working for many years to realise this project: The Large Hadron Collider (LHC) and the associated particle detectors ATLAS, CMS, ALICE and LHCb. With a design centre of mass energy of 14 TeV¹ it will push the high energy frontier by a factor of 7 compared to the previous world record of 1.96 TeV from the Tevatron.

LHC will enable ATLAS to measure many parameters of the Standard Model, and more importantly allows for various searches for new physics beyond the Standard Model.

One example for such a search is the measurement of the differential cross section of the Drell Yan process at high dilepton masses. This process is well understood in theory and deviations from the expected behaviour could be a first indication of new physics. A large variety of theoretical models predict such deviations, thus by performing a precise measurement of the Drell Yan spectrum a model independent search for new physics is possible.

Due to the clear signature the muon channel is well suited for such a measurement. However, various background contributions have to be considered, $t\bar{t}$ decaying into two muons being the most important one. This thesis focuses on reducing those backgrounds by implementing a cut-based analysis. Various systematic uncertainties are taken into account,

¹Throughout this thesis so-called *natural units* are used, i.e. $\hbar = 1$ and $c = 1$.

and the selection cuts are optimised to minimise the overall systematic and statistical uncertainty of the measured spectrum. Monte Carlo data with a centre-of-mass energy of 10 TeV was used for this study, since a long run at that energy was originally planned for 2010.

In chapter 2 an overview of the Standard Model and a brief introduction to the Drell Yan process is given. The leading order and next-to-leading order production mechanisms are explained, and the differential cross section is quantified. In chapter 3 the LHC and ATLAS experiment are described. Chapter 4 introduces the different background processes and explains how they are treated in this analysis. The systematic uncertainties are also handled in this chapter. In chapter 5 the cut optimisation is explained, all applied cuts are discussed in detail and the result of the cut optimisation is presented. In chapter 6 the results are summarized and discussed and an outlook for further analyses is given. In appendix A two methods to correct the measured spectrum for detector effects are compared.

2 Theoretical Background

2.1 The Standard Model

The Standard Model of particle physics (SM) is a gauge theory based on the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ symmetry. All known elementary particles and the fundamental interactions except gravity can be described by the SM. The particles can be divided into two groups: All matter is built from fermions with half-integer spin, while the forces between them are mediated by bosons with integer spin.

The color group $SU(3)_C$ is responsible for the strong interaction which is mediated by eight gluons. The Quantum-Chromodynamics (QCD) describes how those gluons hold the quarks together in order to form baryons and mesons. The unification of the electromagnetic and the weak force is described by the $SU(2)_L \otimes U(1)_Y$ symmetry. The corresponding gauge bosons are the massless photon (γ) and the massive W^\pm and Z^0 bosons. The electroweak interaction is responsible for flavour changing processes. In table 2.1 the gauge bosons of the SM are listed.

The fermions are divided into leptons and quarks with three generations each. In the quark sector the transition between generations is caused by the fact that the mass-eigenstates of the quarks are not identical to the weak-eigenstates. The Cabibbo-Kobayashi-Maskawa (CKM) matrix is used to assign each mass-eigenstate the corresponding flavour content. The SM contains no generation transition mechanism for the lepton sector. In table 2.2 the fermions of the SM are shown.

Boson	Mass [GeV]	Charge [e]	Interaction
γ	0	0	electromag.
W^\pm	80.398 ± 0.025	± 1	electroweak
Z	91.1876 ± 0.0021	0	electroweak
g	0	0	strong

Table 2.1: The gauge bosons of the Standard Model [3].

	1 st gen.		2 nd gen.		3 rd gen.	
	Flavour	Mass [MeV]	Flavour	Mass [MeV]	Flavour	Mass [MeV]
Leptons	e	0.510999	μ	105.658	τ	1776.84 ± 0.17
	ν_e	0	ν_μ	0	ν_τ	0
Quarks	u	1.5-3.3	c	$(1.27^{+0.07}_{-0.11}) \cdot 10^3$	t	$(171.3 \pm 1.6) \cdot 10^3$
	d	3.5-6.0	s	70-130	b	$(4.20^{+0.17}_{-0.07}) \cdot 10^3$

Table 2.2: Properties of the leptons and quarks of the Standard Model [3]. For a definition of quark masses see [3].

Even though the SM has been a great success over the last decades many open questions remain:

- Grand Unified Theories (GUTs) predict that the electroweak and the strong force can be unified, but the GUT scale seems to be many orders of magnitude higher than the scale at which electromagnetic and weak force are unified. This is known as the hierarchy problem and causes large corrections in theoretical calculations.
- All three forces should have the same strength at the GUT scale, but extrapolations suggest that this is not the case.
- The mass of the massive vector bosons can be explained by the Higgs mechanism, but the corresponding Higgs boson has not yet been found.
- The CP violation can be included in the SM by introducing a complex phase in the CKM matrix, but the resulting CP violation can not explain the asymmetry between matter and antimatter in our universe.
- The standard model has 19 free parameters which can not be predicted or explained.
- The newly discovered neutrino oscillation implicates that neutrinos do have a mass and allows transition between lepton generations. It also adds up to ten additional free parameters to the SM.

Various theoretical models expand the SM to solve some of those problems. Many of those expansions predict additional particles or otherwise observable distortions in the Drell Yan spectrum, thus motivating a precise measurement of the high mass Drell Yan spectrum.

2.2 Theory of Proton-Proton Collisions

As LHC is a proton-proton collider the theory of proton interactions will be briefly described here. In section 2.2.1 the hard scattering will be described using a ‘naive’ parton model while section 2.2.2 will discuss the remaining parts of the collision process.

2.2.1 Parton Model

In the naive parton model the proton is described by its constituents, the quarks and gluons. It is assumed that all processes can be described by a scattering in which only one of those partons participates. As the proton consists of the quarks and gluons i the proton structure function $F_2(x)$ can be composed from the distribution of the partons as:

$$F_2(x, Q^2) = \sum_i e_i^2 x f_i(x, Q^2) \quad (2.1)$$

where e_i is the charge of the parton and $f_i(x, Q^2)$ gives the probability of a parton of type i to carry a fraction x of the proton’s momentum for a process at the energy scale Q^2 , also called *parton density function* (PDF). An important result of experiment was the so called *Bjorken scaling* which describes the fact that, the proton structure functions are independent from the energy scale Q^2 , if gluon radiation is neglected: $F_2(x, Q^2) \rightarrow F_2(x)$. This implies that protons consist of pointlike partons and required the development of an asymptotically free theory, the QCD [4].

One of the assumptions in the parton model is that the partons carry a momentum fraction x of the proton’s momentum, while the transverse momentum k_T of the parton is small. In QCD partons can gain a significant k_T by emitting a gluon (as shown in fig. 2.2 (e-g)) and thus break the scaling. This behaviour is expected for all renormalizable gauge theories with pointlike fermion-vector boson interactions [4].

One problem of QCD is that the PDFs cannot be explained by theory, but have to be measured in experiments. The DGLAP [5] evolution equations can be used to translate PDFs to different Q^2 , so at least some extrapolation is possible.

2.2.2 Factorisation

The running coupling strength α_s of QCD is small for high Q^2 which allows the use of perturbation theory when describing the hard scattering. In hadronic initial and final states the soft gluon emissions play an important role and perturbation theory cannot be used. Therefore an important tool to describe hadronic interaction is the factorisation theorem [6], which basically shows that the process can be divided into an initial state described

by the PDFs and a hard scattering. The products of the hard scattering are then further split up as shown in fig. 2.1.

The emission of gluons in the initial and final state is approximated by the so called parton shower approach: The partons are split into pairs of partons with a energy distribution modeled by the Altarelli-Parisi splitting function [5]. This splitting is performed iteratively until the energy of the parton is below a certain cutoff value, which is usually in the order of 1 GeV.

At this point the hadronisation forms hadrons out of those partons, thus ensuring that no colored objects remain. This is achieved by creating additional $q\bar{q}$ pairs and combining them with the existing partons to color neutral hadrons. Two different approaches are commonly used [7]: The String Model and the Cluster Model. Both models are rather heuristic but can be tuned well to describe experimental data.

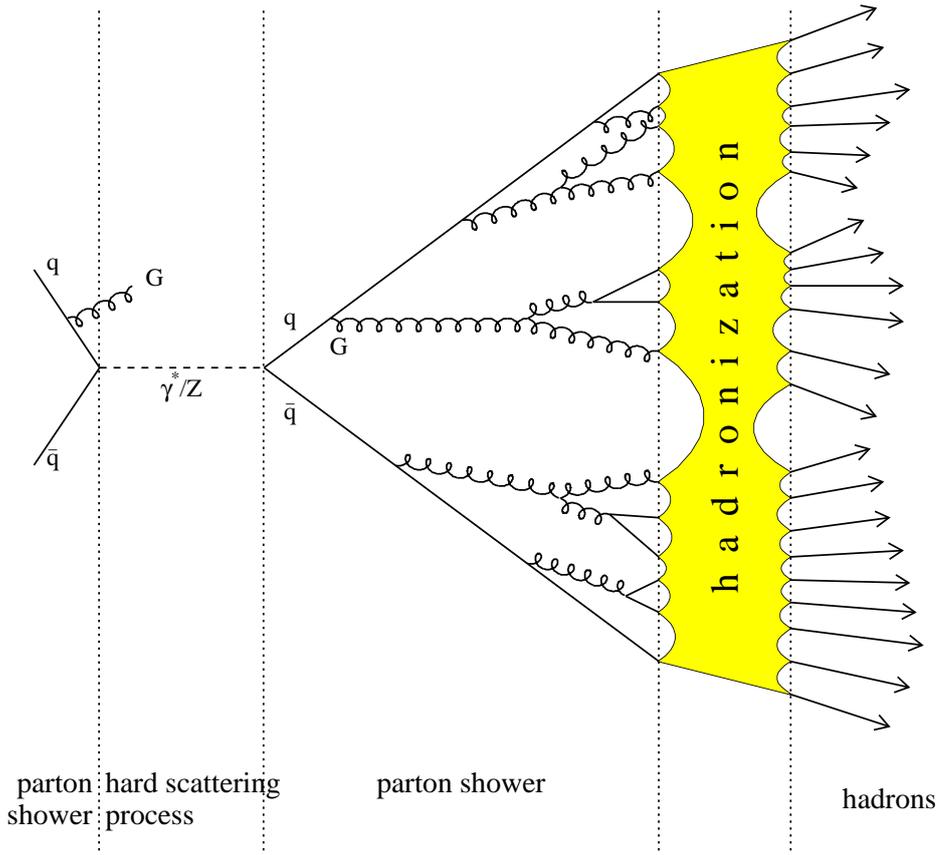


Figure 2.1: The Drell Yan process split up into the different stages [8].

2.3 The Drell Yan Process

In this section the cross section of the Drell Yan process will be discussed. First the leading order cross section at dimuon masses below the Z mass will be derived and the next-to-leading order corrections will be introduced. Finally the Z/γ^* interference will be included.

2.3.1 Leading Order cross section for $m_{\mu\mu} \ll m_Z$

Drell and Yan were the first to point out [9] that the cross section of lepton pair production in hadron-hadron collisions can be calculated from first principles using the parton model. The cross section of the process can be written as the sum over all quark-antiquark combinations and the integral over all momentum fractions x of the two partons:

$$\sigma = \sum_q \int_0^1 dx_1 dx_2 f_q(x_1) f_{\bar{q}}(x_2) \hat{\sigma}_{q\bar{q} \rightarrow l^+ l^-}. \quad (2.2)$$

The PDFs $f_q(x)$ and $f_{\bar{q}}(x)$ can be extracted from deep inelastic scattering. The cross section of the parton level subprocess $\hat{\sigma}_{q\bar{q} \rightarrow l^+ l^-}$ is known from the inverse process $l^+ l^- \rightarrow q\bar{q}$ which can also be calculated from QED:

$$\hat{\sigma}_{q\bar{q} \rightarrow l^+ l^-} = \frac{4\pi\alpha^2}{3\hat{s}} \frac{1}{N} Q_q^2, \quad (2.3)$$

where $1/N = 1/3$ is the colourfactor and $\hat{s} = (p_1 + p_2)^2$ is the squared invariant mass of the two partons. Due to the PDFs the collision energy \hat{s} of the partons is not fixed, so the differential lepton pair mass distribution is more appropriate to be studied:

$$\frac{d\hat{\sigma}}{dM^2} = \frac{4\pi\alpha^2}{3M^2 N} Q_q^2 \delta(\hat{s} - M^2) \quad (2.4)$$

If we now use $\hat{s} = (p_1 + p_2)^2 = x_1 x_2 s$ where \sqrt{s} is the centre of mass energy of the colliding hadrons and substitute eq. 2.4 into eq. 2.2 we get:

$$\frac{d\sigma}{dM^2} = \frac{4\pi\alpha^2}{3M^2 N} \int_0^1 dx_1 dx_2 \delta(x_1 x_2 s - M^2) \left[\sum_{q,\bar{q}} Q_q^2 f_q(x_1) f_{\bar{q}}(x_2) \right] \quad (2.5)$$

In the parton model the PDFs $f(x)$ are independent of M^2 and thus eq. 2.5 exhibits scaling in the variable $\tau = M^2/s$:

$$M^4 \frac{d\sigma}{dM^2} = \frac{4\pi\alpha^2}{3N} \tau \mathcal{F}(\tau). \quad (2.6)$$

2.3.2 Next-to-Leading Order corrections

The parton model approach used in the previous section is able to incorporate perturbative QCD corrections. The Factorisation theorems allows us to absorb all collinear singularities of real and virtual gluon corrections into renormalized PDFs. The cross section can then be written as

$$\sigma = \sum_q \int_0^1 dx_1 dx_2 f_q(x_1) f_{\bar{q}}(x_2) \times [\hat{\sigma}_0 + a \hat{\sigma}_1 + a^2 \hat{\sigma}_2 + \dots]_{q\bar{q} \rightarrow l^+ l^-}, \quad (2.7)$$

where $a = \alpha_s(M^2)/2\pi$ [4]. As mentioned before the parton model is limited by the assumption that the transverse momentum k_T of the partons is small, thus for properly describing the p_T distribution of the lepton pair a more complex approach has to be used (see for example [4]). In fig. 2.2 the leading and next-to-leading order QCD diagrams for the $q\bar{q}$ side of the Drell Yan process are shown.

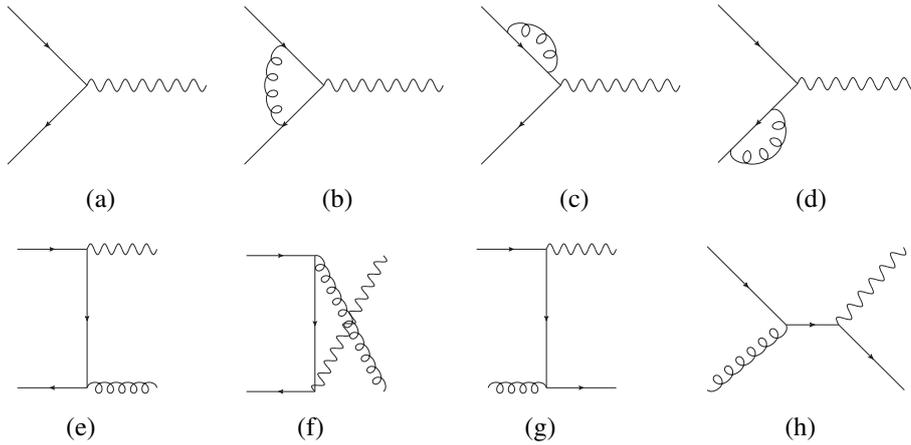


Figure 2.2: The leading (a) and next-to-leading-order (b-h) diagrams for the Drell-Yan process. (c-d) are virtual gluon corrections, (e-f) real gluon corrections and (g-h) are quark-gluon scattering processes.

2.3.3 Including the Z Boson

If the invariant mass of the produced dilepton is not small compared to the mass of the Z boson the process $q\bar{q} \rightarrow \gamma^* \rightarrow l^+ l^-$ has to be expanded to $q\bar{q} \rightarrow Z/\gamma^* \rightarrow l^+ l^-$. The Z boson production is well understood from QED and eq. 2.2 can be adapted to include the interference between γ^* and Z. The resulting equations are rather long and are presented for example in [4]. Even though to the Z boson has a small decay width of only 2.50 GeV the high mass Drell Yan spectrum is dominated by the $Z^*/\gamma^* \rightarrow l^+ l^-$ interference.

3 Experimental Setup

3.1 The Large Hadron Collider

The Large Hadron Collider is a proton proton collider with a design centre of mass energy of 14 TeV located at CERN near Geneva. It is built 100 m underground in a 27 km long ring tunnel, which was formerly used for the LEP experiment. LEP was an electron positron collider with a maximum centre of mass energy of up to 209 GeV, which was limited by synchrotron radiation. The well defined initial state of an e^+e^- collider is well suited for precision measurements, but the direct search for new physics is limited by its centre of mass energy. The more direct predecessor to the LHC is the Tevatron collider at Fermilab near Chicago, which is a $p\bar{p}$ collider with a centre of mass energy of 1.96 TeV.

Due to the hadronic environment Tevatron and LHC are quite similar in respect to detector requirements, analysis techniques and the underlying physics. At a centre of mass energy of 1.96 TeV quark antiquark annihilation is the dominating mechanism for many physics processes. Since sea quarks tend to have a lower momentum fraction than valence quarks, anti protons were used. At the higher energies of the LHC, gluon reactions start to become more important, which allows the use of proton proton collisions. This allows to reach higher luminosities¹, as the complex production of antiprotons can be omitted.

Hadron colliders, and especially LHC, are often designed as discovery machines, as they can access higher centre of mass energies than e^+e^- colliders. The downside of pp collisions is the unknown initial state and the large QCD background. Also only a fraction x of the momentum of the proton is carried by the interacting parton. Thus the energy available for the production of new particles will be different for each collision. A higher number of collision events will therefore extend the energy reach of the collider. With that in mind the LHC has been designed to deliver a very high luminosity of $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$.

In the beginning LHC will be operated with a lower luminosity, but once the design luminosity is reached 2808 [10] bunches with over 10^{11} protons each will be accelerated in each direction. They will collide every 25 ns at four different interaction points where the beams are focused to a radius of $16.7\mu\text{m}$. The 1232 dipole magnets needed to bend

¹See App. B.1 for a definition of luminosity

the beams will operate at 8.33 T and are cooled by super-liquid Helium at a temperature of 1.9°K. More than 400 magnets of higher multipole order are used to keep the beam focused and to squeeze it at the interaction points.

The centre of mass energy of LHC is limited by the magnetic field of the dipole magnets. During initial testing it turned out that the magnets can currently only be operated at a magnetic field equivalent to a centre of mass energy of 10 TeV. It was then decided to operate LHC at 10 TeV and make the necessary improvements to reach 14 TeV during the first shutdown [11]. During the first attempts to ramp up to 10 TeV a faulty electrical connection damaged the cooling system which led to a helium leak and subsequent mechanical damage to several magnets [12]. After time-consuming repairs it was decided to start the 2010 run with 7 TeV and later that year increase the energy to 10 TeV [13]. LHC started successful operations in 2009 [14] and the current plan is to operate for 18-24 month at 7 TeV. After that a long shutdown will allow the necessary upgrades to increase the centre of mass energy to 14 TeV [15].

3.1.1 Experiments at LHC

There are four big particle detectors installed at the four interactions zones of LHC shown in fig. 3.1, as well as two forward detectors:

ATLAS(**A Toroidal LHC AparatuS**) [16] and CMS(**Compact Muon Solenoid**) [17] are general purpose detectors. The focus of the ATLAS experiment is a precise Muon Spectrometer with as few material as possible in the path of the muons, while CMS put more focus on a precise tracking and electromagnetic calorimeter. Both ATLAS and CMS have a broad physics programme including the search for the higgs boson as well as searches for new physics like supersymmetry, black holes, extra dimensions and leptoquarks.

The ALICE (**A Large Ion Collider Experiment**) [18] is designed to study the quark gluon plasma, especially in lead-lead collisions. For this kind of collisions a much higher track multiplicity is expected, so ALICE is designed to handle much higher data rates than ATLAS or CMS.

The LHCb (**LHC beauty**) [19] experiment will study B-physics to gain a better understanding of CP-violation. Unlike the other three particle detectors LHCb is not symmetric, but is constructed as a single arm forward spectrometer.

The TOTEM (**TOTAL Elastic and diffractive cross section Measurement**) [20] experiment is situated near the CMS interaction point very close to the beamline. It aims at measuring elastic and diffractive cross sections of proton-proton collisions as well as the size of the proton. TOTEM will also be useful in determining the luminosity of LHC.

LHCf (**LHC forward**) [21] consists of two detectors located on either side of the ATLAS experiment about 140 meters from the interaction point at a zero degree collision angle.

LHCf will measure the energy and amount of neutral pions produced in the pp collisions of LHC. This will help to improve the understanding of cosmic showers produced by ultra-high-energy cosmic rays.

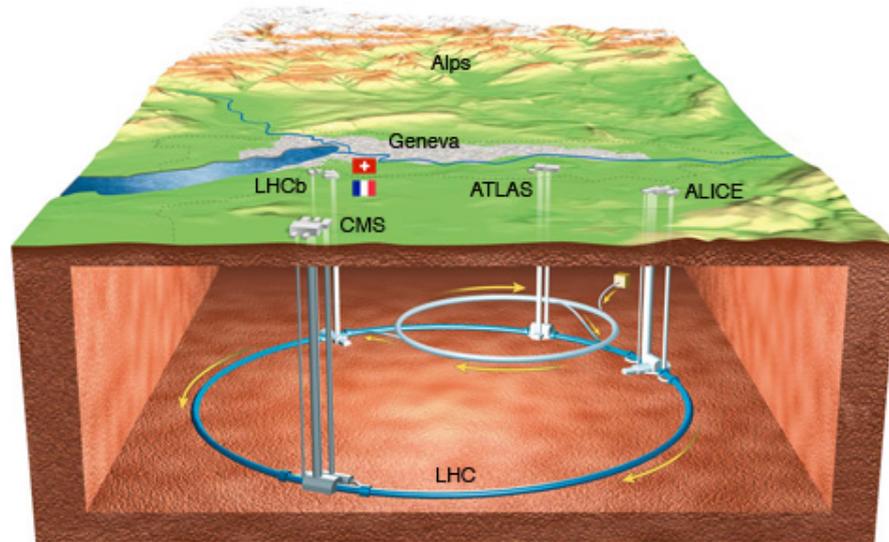


Figure 3.1: Overview of the LHC ring and the four experiments.

3.2 The Atlas Experiment

The ATLAS detector shown in fig. 3.2 is a classical omni-purpose detector for a collider experiment [22]. It has a symmetric cylindrical barrel layout with several subdetectors in a concentric arrangement around the beam axis of LHC. Starting from the inside the main components are the Inner Detector, the Electromagnetic and Hadronic Calorimeter and the large Muon Spectrometer. Each subdetector is divided in a barrel and two endcap regions. The Inner Detector is embedded in a solenoid magnet with a field strength of 2 Tesla. The Muon Spectrometer is inside a large toroidal magnet with non-uniform magnetic field. In total, ATLAS has a length of 44 m, a diameter of 25 m and a mass of about 7000 t.

3.2.1 The ATLAS coordinate system

The barrel shape of ATLAS suggests using a cylindrical coordinate system with its origin at the interaction point in the detector's centre. In Cartesian coordinates the x-axis points toward the centre of the LHC ring, the y-axis upwards and the z-axis is defined by the

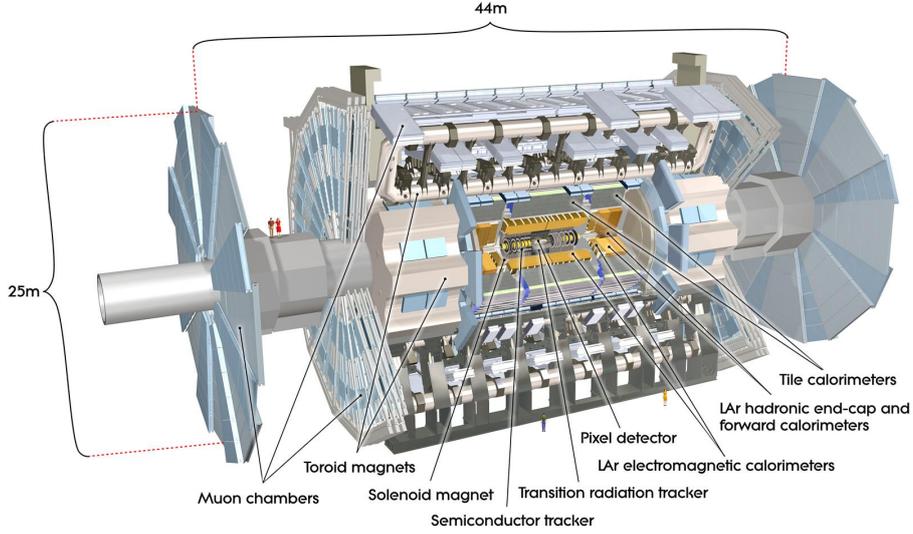


Figure 3.2: The ATLAS Detector.

beam direction. In cylindrical coordinates an azimuthal angle ϕ , a polar angle θ and a radius r are defined, where $\tan \phi = \frac{\sqrt{x^2+y^2}}{z}$. It is more convenient to use the pseudorapidity

$$\eta = -\log \left(\tan \frac{\theta}{2} \right) \quad (3.1)$$

which is in a good approximation additive under Lorentz boosts along the z -direction.

The spatial distance ΔR in the η - ϕ -plane is defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (3.2)$$

3.2.2 Inner Detector

The Inner Detector is placed a few centimetres from the beam axis and extends to a radius of 1.2 metres. It consists of the Silicon Pixel Detector, the Semiconductor and the Transition Radiation Tracker (only the first two are shown in fig. 3.3). All three subdetectors will supply several hits with a good spatial resolution for each track, allowing a precise reconstruction of the track. The bending of the tracks in the magnetic field will then allow to measure the momentum of the particles. The location of the Silicon Pixel Detector close to the interaction point allows for a good vertex resolution, while the Transition Radiation Tracker will allow a separation of electrons from pions. The geometric acceptance of the Inner Detector is limited to $|\eta| < 2.5$.

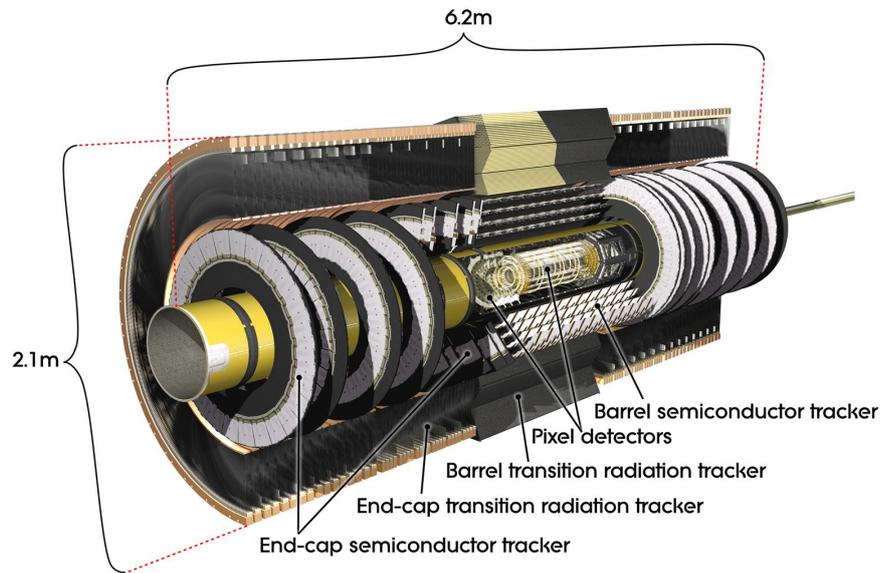


Figure 3.3: Central part of the ATLAS Inner Detector.

3.2.3 Electromagnetic Calorimeter

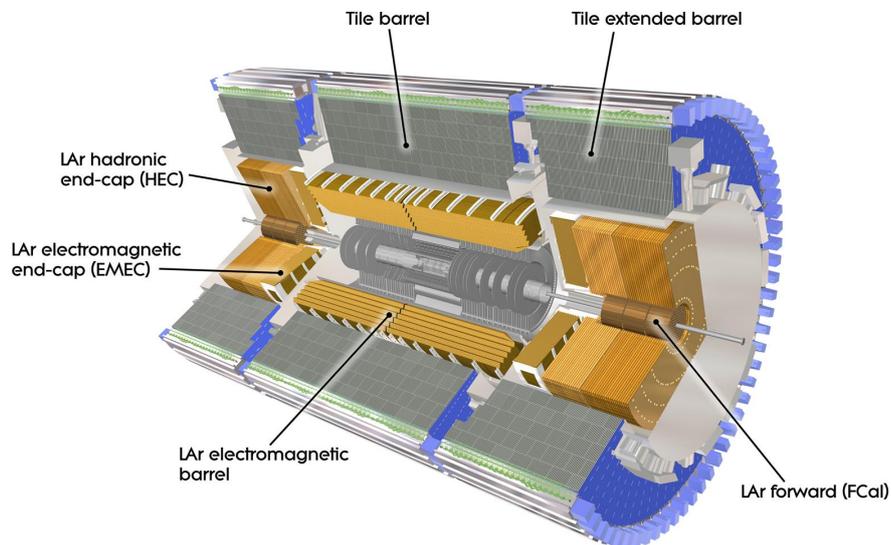


Figure 3.4: The ATLAS calorimeter system.

The Electromagnetic Calorimeter (ECAL) is part of the ATLAS calorimeter system shown

in fig. 3.4. It is designed to measure the energy of electromagnetically interacting particles, such as electrons, positrons and photons. These particles deposit most of their energy within the ECAL through showering. To discriminate between electrons and photons the Inner Detector is needed. The ECAL is an accordion-shaped lead and liquid argon sampling detector with a geometrical acceptance of $|\eta| < 3.2$.

The spatial granularity is $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ and the design energy resolution for electromagnetic showers is [23]

$$\frac{\Delta E}{E} \approx \frac{0.1}{\sqrt{E/[\text{GeV}]}} \oplus 0.01. \quad (3.3)$$

3.2.4 Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is placed around the ECAL. It is designed to measure the energy of charged and neutral hadrons. The barrel part of the HCAL is a sampling calorimeter with steel as the absorber and scintillating tiles as the active material. In the endcap and forward region a liquid argon sampling calorimeter is used. In order to prevent showers from leaking outside the calorimeter, the HCAL is several interaction lengths thick over the whole η region. The high geometrical acceptance of $|\eta| < 4.9$ allows a nearly hermetic measurement of all produced particles, and thus a good measurement of missing transverse energy is possible. The design resolution of the HCAL is [23]

$$\frac{\Delta E}{E} \approx \frac{0.5}{\sqrt{E/[\text{GeV}]}} \oplus 0.03. \quad (3.4)$$

3.2.5 Muon Spectrometer

Muons are the only detectable particles that will be able to reach the Muon Spectrometer, which is located outside the HCAL. Its goal is to identify muons, give a precise standalone momentum measurement and a trigger for high energetic muons. As shown in fig. 3.7 the barrel region of the Muon Spectrometer uses Monitored Drift Tubes (MDTs) for a precise measurement of η and Resistive Plate Chambers (RPCs) for measuring ϕ , trigger and timing information. A combination of MDTs, Cathode Strip Chambers (CSCs) and Thin Gap Chambers (TGCs) are used in the endcap region.

As mentioned before, the Muon Spectrometer uses a toroidal magnetic field and thus the muons will be bent in η . Three layers of precision chambers will measure the sagitta, which is defined in fig. 3.5. With a known field strength the sagitta can be used to determine p . As the magnetic field is not homogeneous the actual computation is complicated,

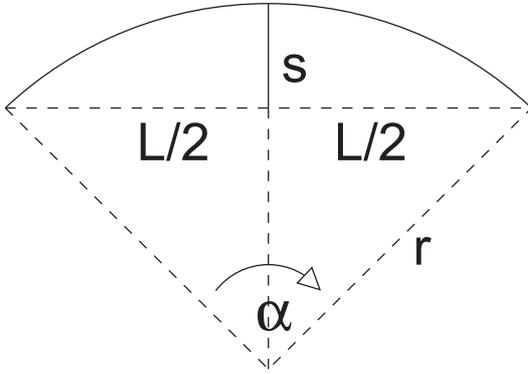


Figure 3.5: Definition of sagitta [24].

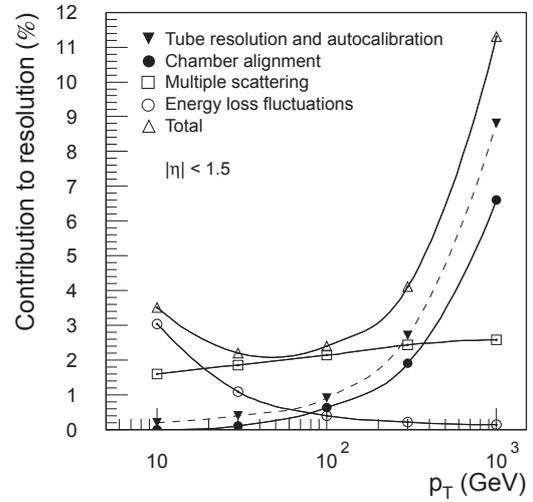


Figure 3.6: Expected contributions to the momentum resolution in the barrel region [24].

but it has to be noted, that p is proportional to $1/s$. The advantage of the torroidal design of the ATLAS Muon Spectrometer is the low material budget which keeps the impact of multiple scattering low, even for muons with very high momentum, as seen in fig. 3.6.

3.2.6 Trigger

At LHC's design luminosity the collision rate in ATLAS will be 40 MHz. The amount of data would be much too high to be completely stored to disk. It is thus necessary to make a fast selection of interesting events that will be recorded. The ATLAS trigger system consists of three different levels:

- The **Level One Trigger** is implemented in hardware and can only access basic information of the calorimeter and the Muon Spectrometer. The Level One Trigger specifies certain Regions of Interest (ROI). It has a target rate of 75 kHz.
- The **Level Two Trigger** is a software trigger, which can use the full data from the ROIs previously defined, including information from the Inner Detector. It will reduce the event rate to 1 kHz.
- The **Event Filter** has access to the full detector information, and can use more sophisticated algorithms than the Level Two Trigger. It is run on a computer farm near the ATLAS pit, and has a target rate of 100 Hz.

Only events that subsequently pass all cuts are recorded to disk.

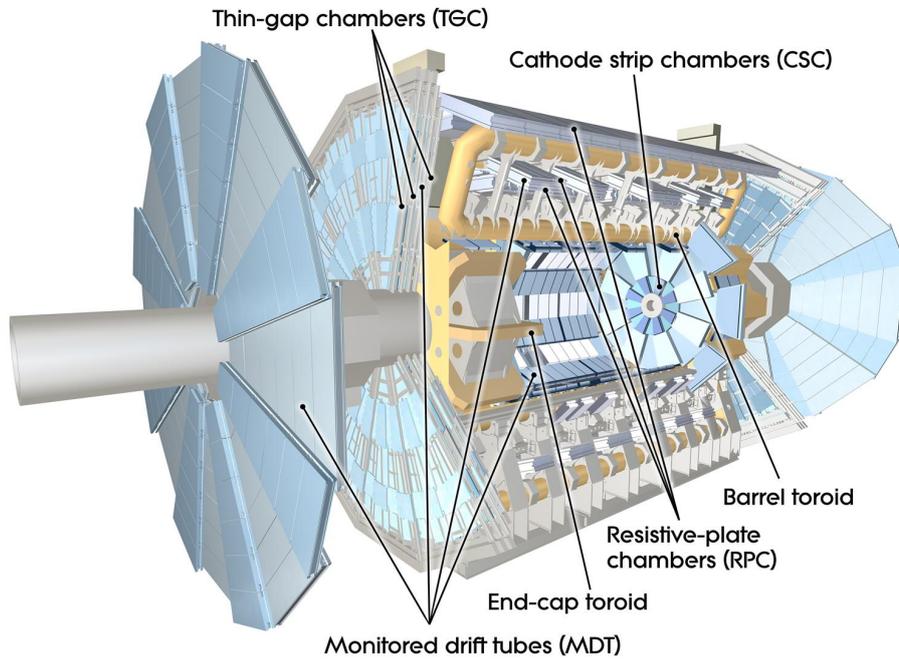


Figure 3.7: The ATLAS Muon Spectrometer.

3.3 Physics Objects

A large variety of algorithms has been developed to reconstruct usable physics objects like muons, jets or electrons from the recorded output of the various subdetectors of ATLAS. Also the construction of some deduced variables like missing transverse energy (E_T) and b -tagging are explained. Most objects (e.g. muons, jets and E_T) are reconstructed by competing algorithms, and only the ones that were used in this analysis are presented here.

3.3.1 Jets

One jet algorithm in ATLAS is the iterative seeded fixed-cone jet finder, which was already used at Tevatron [25]. It uses the calibrated calorimeter tower cells as input and creates individual objects from each tower. The object with the highest p_T above a threshold of 1 GeV is used as a seed. All objects with $\Delta R < 0.4$ are merged with the seed into a new object, thus defining a new cone. Now all objects inside the new cone are (re-)collected, and the direction of the cone is updated again. This process is repeated until the cone does not change anymore, and the resulting object defines the jet.

Now the next remaining object with a p_T above 1 GeV is used to seed another jet. This jet can share some of the objects with the previously constructed jets, thus creating some kind of overlap. In a last step, jets which share common objects are either merged or split, depending on energy of the shared objects.

This jet algorithm is not completely infrared safe, and some alternative jet-algorithms like kt [26] or anti-kt [27] are available for this purpose. In this analysis only jets with a p_T above 10 GeV are used.

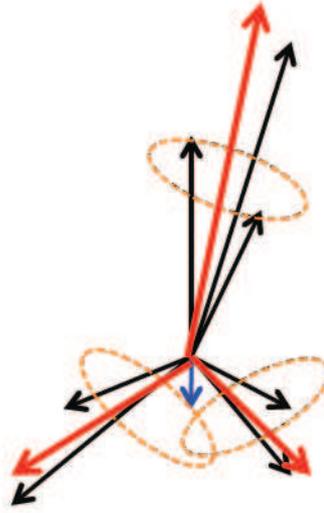


Figure 3.8: Three jets (red arrows) consisting of various calorimeter objects (black) [28].

3.3.2 Muons

The Staco algorithm [29] combines tracks from the Muon Spectrometer with ID tracks. Muonboy [30] is used to find the tracks inside the Muon Spectrometer and to extrapolate them to the Inner Detector, using a parametrisation for energy loss in the calorimeter. Staco then searches for tracks in the ID that can be matched with the extrapolated muon track. The quality of the match between the ID and Muonboy track, $\chi^2_{\text{match}}/N_{\text{DOF}}$, can later be used as a quality criterium for the muons. In some η regions the Muon Spectrometer will not be able to find tracks by itself. An additional tagger called MuTag [29] extrapolates ID tracks to the first layer of the Muon Spectrometer and can thus find tracks that would be lost otherwise. It is also possible to use information from the calorimeter, to identify muons that only have an ID track and no hit in the Muon Spectrometer at all. This is especially important for muons with low momentum and is done by the CaloMuonTag algorithm [31].

3.3.3 Missing Transverse Energy

Many physics studies in ATLAS use a selection cut on missing transverse energy, \cancel{E}_T . Especially SUSY searches often require a high \cancel{E}_T , while standard model processes like QCD events or $Z \rightarrow \mu^+ \mu^-$ will not produce real \cancel{E}_T . A good \cancel{E}_T resolution is also important for reconstructing the top-quark mass in the semileptonic decay channel. This analysis uses the so called Cell-based algorithm which uses mostly information from the calorimeter cells. For this, a noise suppression procedure is applied to all calorimeter cells, and an individual calibration is applied to each cell. This individual calibration takes into account if the energy in that cell was deposited by an electromagnetic shower, a hadronic shower or a muon. The vectorial sum of those energies gives the term $\cancel{E}_{x,y}^{\text{Calo}}$ in eq. 3.5.

Two additional corrections are applied: The energy $\cancel{E}_{x,y}^{\text{Cryo}}$ deposited in the walls of the cryostat located between the ECAL and the HCAL is estimated by comparing the energy in the outermost layer of the ECAL with the innermost layer of the HCAL. As muons will not deposit all of their energy in the calorimeter, the term $\cancel{E}_{x,y}^{\text{Muon}}$ corrects for the energy of the escaping muons. The energy deposition of the muon in the calorimeter is taken into account.

$$\cancel{E}_{x,y}^{\text{Final}} = \cancel{E}_{x,y}^{\text{Calo}} + \cancel{E}_{x,y}^{\text{Cryo}} + \cancel{E}_{x,y}^{\text{Muon}}. \quad (3.5)$$

3.3.4 b -Tagging

The ability to identify jets containing b -hadrons, so called b -tagging, is an important part of a general purpose detector like ATLAS. It is very helpful for a variety of physics analysis. b -tagging allows a selection of very pure $t\bar{t}$ samples, since top quarks dominantly decay into b quarks. It is helpfully in selecting Higgs decay channels including b -quarks, as well as various potential supersymmetric decay chains. Another benefit of b -tagging is the possibility to suppress $t\bar{t}$ and $b\bar{b}$ background, which will also be used in this analysis. Depending on the specific purpose, the b -tagging has to fulfill different requirements. The b -tagging efficiency is defined as the number of b -jets that are tagged as b -jets divided by the number of b -jets. This mis-tagging rate is defined as the number of light jets that are tagged as b -jets divided by the number of light jets.

To identify b -jets various properties of the b -quark can be used. First, due to the relatively high mass of the b -hadrons (> 5 GeV) their decays have some detectable characteristics. The decay products may have a large transverse momentum with respect to the jet axis, thus making it possible to identify them. The most important property is the long lifetime of b -hadrons, which is in the order of 1.5 ps. This allows b -hadrons to live long enough to travel on average about 3 mm before decaying. Thus a secondary decay vertex

is formed, which can be separated from the primary vertex. Different algorithms have been developed [2] to accomplish this, and a brief discussion will follow: First, different discriminating variables are discussed, and afterwards the likelihood ratio formalism is introduced to calculate cut variables from those discriminating variables.

Impact parameter tagging algorithms

The principle of b -tagging with algorithms based on impact parameter is best described in fig. 3.9. Every track is assigned a signed impact parameter with respect to the primary vertex. This sign is defined as

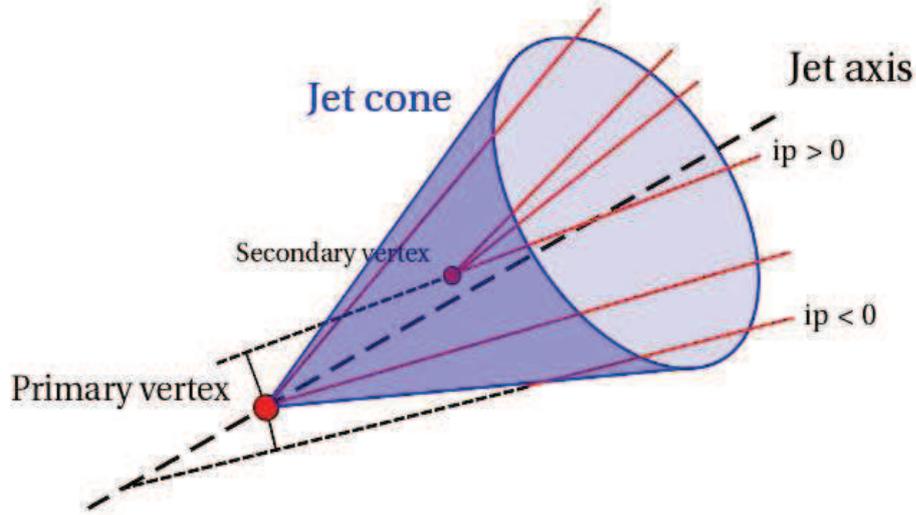


Figure 3.9: b -tagging principle [32].

$$\text{sign}(d_0) = (\vec{P}_j \times \vec{P}_t) \cdot (\vec{P}_t \times (\vec{X}_{pv} - \vec{X}_t)) \quad (3.6)$$

where \vec{P}_j and \vec{P}_t are the directions of the jet, and the track respectively, \vec{X}_t is the position of the point of closest approach of the track and \vec{X}_{pv} is the position of the primary vertex. Tracks originating from the primary vertex will have a random sign with equal distributions, while tracks that actually are from a secondary vertex will tend to have a positive sign. To give more weight to precisely measured tracks, the significance distribution d_0/σ_{d_0} is used as discriminating variable. There are three algorithms that use this method: IP1D uses the longitudinal impact parameter and IP2D uses the transverse impact parameter. IP3D uses correlations of both parameters by using a two-dimensional histogram.

Secondary vertex tagging algorithms

A more complex but also more rewarding approach is the reconstruction of the secondary vertex from the b decay. For this, all tracks with a certain distance to the primary vertex are used to build two-track pairs forming a good vertex. Vertices compatible with decays of short lived particles or material interactions are rejected. All remaining tracks are combined into a single vertex. Then tracks are removed until the χ^2 of the vertex fit is good. Three properties of the vertex are used as discriminating variables: The invariant mass of the reconstructed vertex, the ratio of the sum of the energies of the tracks contributing to the vertex to the sum of the energies of all tracks in the jet as well as the number of two tracks vertices. One algorithm using those variables is the so-called SV1 tagger, which uses a 2D-distribution of the first two variables and a 1D-distribution of the last.

Formalism of the likelihood ratio

The two algorithms only define a set of discriminating variables S_i . The likelihood ratio formalism is used to transform those variables into a b weight for each jet. A pre-defined, smoothed and normalized distribution of the discriminating variables S_i is used, one for the b -jet hypothesis, $b(S_i)$ and one for the light jet hypothesis, $u(S_i)$. In case of SV1 and IP3D these distributions can be two-dimensional. The ratio of the probabilities $b(S_i)/u(S_i)$ defines the weight of a vertex or track. To combine N_T tracks the logarithms of the individual track weights W_i are added:

$$W_{Jet} = \sum_{i=1}^{N_T} \ln W_i = \sum_{i=1}^{N_T} \ln \frac{b(S_i)}{u(S_i)} \quad (3.7)$$

In case of the secondary vertex tagger a weight of $\ln \frac{1-\epsilon_b^{SV}}{1-\epsilon_u^{SV}}$ is assigned if no vertex is found. The typical efficiency for b jets ϵ_b^{SV} and light jets ϵ_u^{SV} must be known a priori. The same is true for the distributions $b(S_i)$ and $u(S_i)$. This is also the biggest drawback of the likelihood ratio formalism: It requires a priori knowledge of the detector, which will not be available from first data. It is expected that a few hundred pb^{-1} will be needed to be able to use the more sophisticated algorithms like SV1. The advantage of the likelihood ratio is the possibility to easily combine different algorithms by simply adding their weight. A commonly used variable is IP3D+SV1.

Algorithms for early data

To have a reliable b -tagging for early data, some new algorithms have been implemented in ATLAS in the recent past [32]. The so called TrackCounting algorithm uses the signed

impact parameter significances that are used in the Impact parameter tagging algorithms. Instead of comparing all tracks with predefined distributions, the two tracks with the second and third highest impact parameter significances are compared to a fixed cut value. In fig. 3.10 one can see that TrackCounting is not as efficient as the IP2D algorithm. It can be shown however that TrackCounting is less sensitive to misalignment, and it does require less a priori knowledge of the detector performance. It is therefore a valid alternative for early data. Unfortunately, this algorithm is only available in Athena 15 and newer, while this analysis is based on version 14.2. Therefore only IP3D+SV1 and IP2D have been studied here.

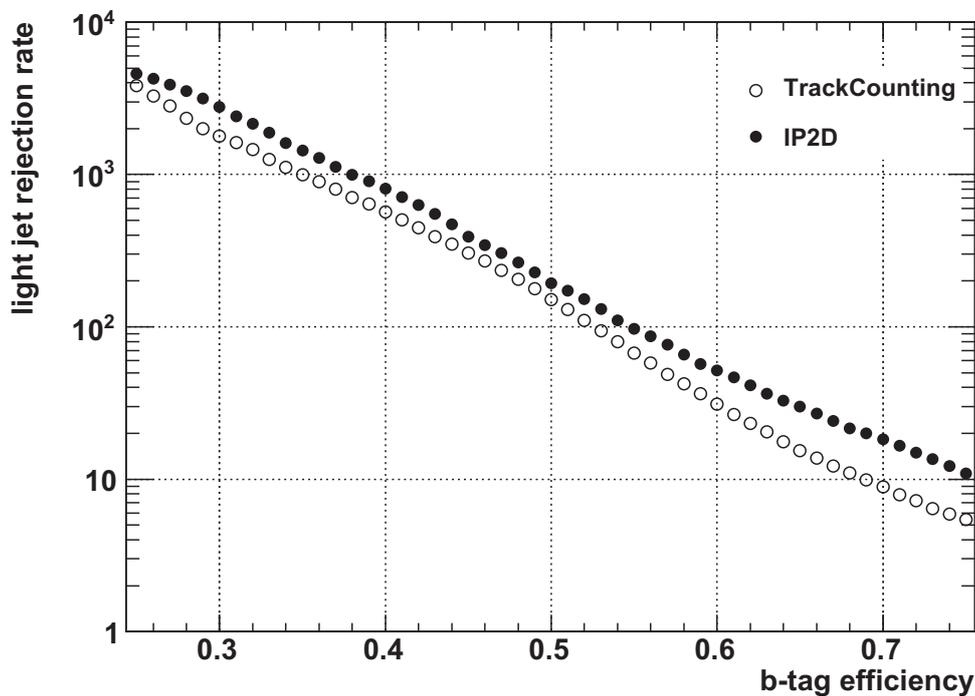


Figure 3.10: Comparison of Track Counting and IP2D [32].

3.4 Computing Model

An experiment as complex as ATLAS has also high requirements concerning analysis software and computing infrastructure. ATHENA, a multi purpose software framework has been developed to handle the various computing aspects of ATLAS. The object-orientated analysis framework ROOT [33] is used for storing and analysing data. To process the huge amounts of data the grid, a worldwide network of computing centres, is used.

3.4.1 ATHENA Framework

ATHENA [34] is an adaptation of the GAUDI [35] framework for high energy physics experiments. It is highly modular and consists of python scripts and C++ programs. The following tasks are performed inside the ATHENA framework:

- **Monte Carlo generation:** ATHENA controls the input options for various Monte Carlo programs and offers a common filter for the produced Monte Carlo events.
- **Simulation:** GEANT4 [36] is used to simulate the propagation and interaction of the final state particles through the detector.
- **Digitisation:** The interactions of the particles with the detector are translated into an electronic response.
- **Reconstruction:** All algorithms for reconstruction are implemented in ATHENA and can communicate with each other. Thus algorithms can use information provided by other algorithms while still being developed separately.
- **Calibration and alignment:** The calibration and alignment of the various subdetectors is also handled within ATHENA. The information is stored in a database, which can be accessed for reconstruction and analysis purposes within ATHENA.
- **Analysis:** For analyses with real data ATHENA can be used to utilize the full power of the available algorithms and to access additional detector information.

3.4.2 Monte Carlo Production

As the simulation of a single Monte Carlo event for ATLAS takes several minutes, it is necessary to organise the production of events centrally [37]. In fig. 3.11 the individual steps of the production are shown, in table 3.1 an overview of the used data formats is given. The event generation and simulation is performed by different working groups, and the resulting AODs can be used by all members of the collaboration. As the AODs are large in size analyses using AODs are performed on the grid and not locally. To allow a faster development HPTV [38] is used to create D3PDs which can then be used for analysing data locally.

Name	Size	Description
RAW	≈ 1.6 MB	Contains all recorded information from the subdetectors.
ESD	≈ 500 kB	Derived from RAW by applying offline reconstruction algorithms. It contains information about the different physics objects, and is mainly used for performance studies. Athena is needed to access the information in ESDs.
AOD	≈ 100 kB	Similar to ESDs, contains all information necessary for physics analysis
D3PD	≈ 10 kB	Derived from AODs by omitting less important information and keeping only the interesting events. Can be accessed from ROOT.

Table 3.1: Different data formats used in ATLAS.

3.4.3 Grid Computing

At the design trigger rate LHC will record more than 1 PB of data per year [39]. Processing those huge amounts of data and producing the previously mentioned large number of Monte Carlo samples needed is beyond the scope of a single computing centre. The world-wide LHC Computing Grid [40] provides the necessary computing power by operating more than 100,000 processors in over 170 computing centres in 34 countries. The network is structured hierarchically in different tiers:

- **TIER-0** is located at CERN. All datasets are stored there and a first processing of the data is performed. For further analysis, the data is distributed to other sites.
- **TIER-1** centres are responsible for different regions of the world and are used for further processing, data storage and distribution. The TIER-1 responsible for Germany is the GridKa [41] located in Karlsruhe.
- **TIER-2** centres are smaller and are used for larger physics analyses, Monte Carlo production and storing selected data samples. The local TIER-2 is the LRZ-LMU Tier2 [41].

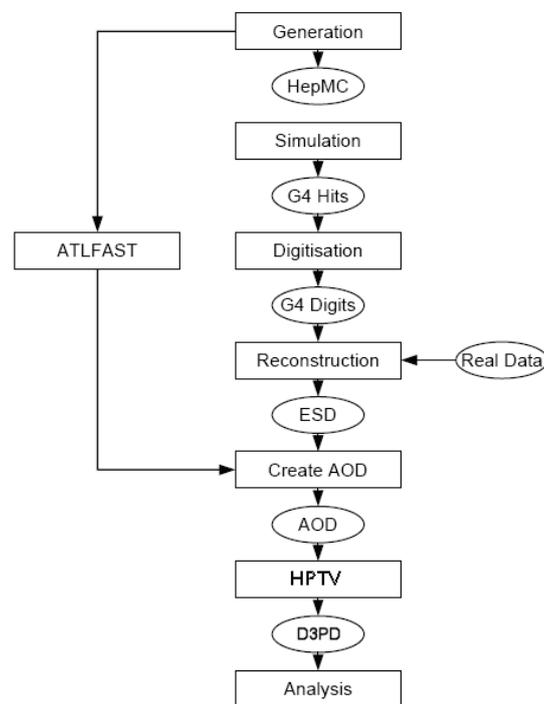


Figure 3.11: The full analysis chain for Monte-Carlo data and real data in the ATHENA framework [28].

4 High Mass Drell Yan Spectrum

In this chapter the necessary ingredients for a measurement of the high mass Drell Yan spectrum are discussed. First the contributing background processes are discussed. The Monte Carlo generators used to simulate those processes are presented, and an overview of the produced samples is given. Next the different approaches of the existing analyses are compared with this analysis and finally the systematic uncertainties are discussed.

4.1 Backgrounds

Even though the signal process $Z \rightarrow \mu^+ \mu^-$ has a clear signature of two isolated muons there are some standard model processes which will contribute some background. Those backgrounds are presented in the following section.

4.1.1 $t\bar{t}$

As $t\bar{t}$ is the most important background for this analysis, the production and decay mechanisms are described in the following.

$t\bar{t}$ Production

There are three important production mechanisms for top-quark pair production: quark anti-quark annihilation shown in fig. 4.1(a), the gluon fusion channel shown in fig. 4.1(b) and the dominant gluon t-channel shown in fig. 4.1(c). At the high centre of mass energy of LHC the gluon fusion channel will dominate over the quark anti-quark annihilation, which is more important at lower centre of mass energies.

The cross section for $t\bar{t}$ production is highly dependent on the centre of mass energy of the collider. Assuming a top mass of 172.5 GeV the cross section at near-NNLO precision is 886.28 pb for 14 TeV, 402.53 pb for 10 TeV and 160.79 pb at 7 TeV.

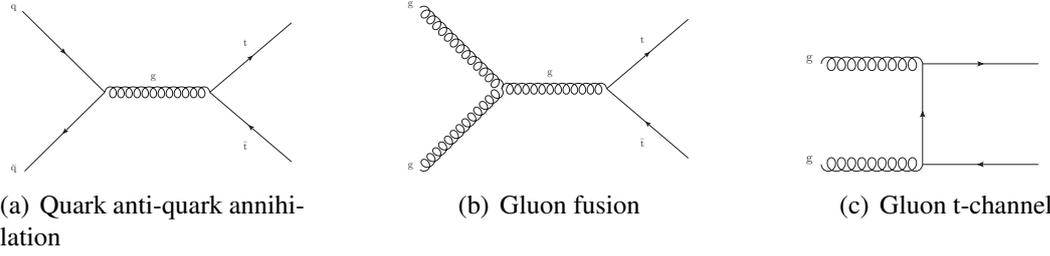


Figure 4.1: Two different LO production channels for $t\bar{t}$.

$t\bar{t}$ Decay

Due to the extremely short lifetime of ca. 10^{-24} s the top quark decays directly into a b -quark plus a W boson, without hadronisation. The W boson can then decay either leptonically or hadronically. The combination of the two W bosons leaves 16 different decay channels shown in fig. 4.3. The most important background to $Z \rightarrow \mu^+\mu^-$ stems from dileptonic $t\bar{t}$ decays¹, where both W -bosons decay into a muon. Due to the high mass of the top quark the W -bosons are real, and the resulting muons can have a high p_T and are well isolated. The dileptonic decay channel will produce two neutrinos and two b -jets, as can be seen in fig. 4.2. The two neutrinos will escape the ATLAS detector undetected and thus produce some \cancel{E}_T . The decaying b -quarks will hadronize and produce jets in the detector. Thus the jet activity for $t\bar{t}$ will typically be higher than for $Z \rightarrow \mu^+\mu^-$, and it might be possible to identify the jets from b -decays with the so called b -tagging.

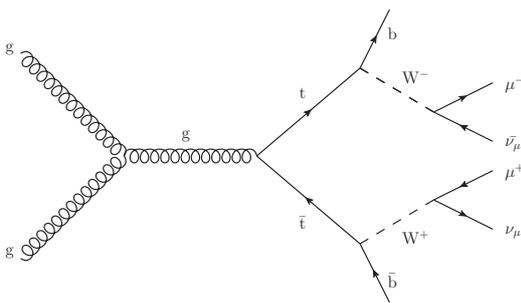


Figure 4.2: Leptonic $t\bar{t}$ decay into two muons.

$c\bar{s}$	electron+jets muon+jets tau+jets			all-hadronic			
$u\bar{d}$							
τ^-	e τ	$\mu\tau$	$\tau\tau$			tau+jets	
μ^-	e μ	$\mu\mu$	$\mu\tau$				
e^-	ee	e μ	e τ				
W decay	e^+	μ^+	τ^+	$u\bar{d}$	$c\bar{s}$		

Figure 4.3: Combination of $t\bar{t}$ decay channels [42].

¹There is also some background from the semileptonic channel, where one W -boson decays into a muon, and one muon is produced from a b quark. The muons from b -decays have a lower p_T and are not isolated, so this channel will be easier to remove by cuts on p_T and isolation.

4.1.2 $b\bar{b}$

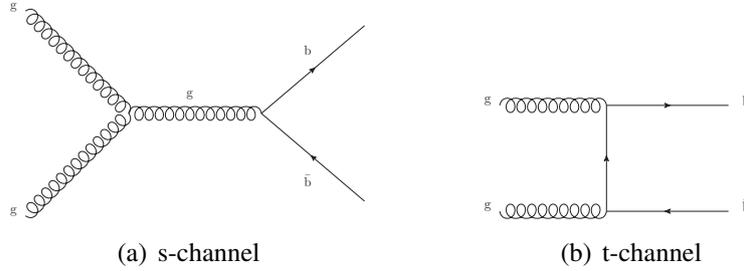


Figure 4.4: Two different LO production channels for $b\bar{b}$.

The production of $b\bar{b}$ pairs has a cross section which is some orders of magnitude higher than the $t\bar{t}$ production, but the impact on this analysis is nonetheless smaller for three reasons: b -quarks are mainly produced in forward direction, the squared branching ratio for $b \rightarrow \mu$ is low, and muons from b -decays are not expected to be isolated. The t-channel production of $b\bar{b}$ shown in fig. 4.4(b) is dominant compared to the s-channel production shown in fig. 4.4(a), and thus most b -quarks will be produced at small angles from the beam axis [43]. Since the mass of the b -quark is below that of the W boson, the W will be virtual, and thus the direction of the produced muons will be close to the original b -quark. This results in non-isolated muons with a high $|\eta|$ and a low p_T . The geometrical acceptance of the detector and a cut on the p_T of the muons will therefore lead to a strongly reduced acceptance of $b\bar{b}$ decays [44].

It has been shown [24] that muon production from QCD processes is completely dominated by $b\bar{b}$ production² for muons with a p_T greater than 10 GeV. Thus, the $b\bar{b}$ channel is studied representatively for all QCD processes.

4.1.3 $Z \rightarrow \tau^+\tau^-$

The Drell Yan process $Z \rightarrow \tau^+\tau^-$ has nearly the same cross section as the signal process, but due to the squared branching ratio of $\tau \rightarrow \mu\nu$ its impact on this analysis is limited. As it is an electroweak process, it is very well understood in theory and thus the impact on the high mass Drell Yan spectrum can be predicted quite well [3]. We will later see that this impact is rather small.

²Technically $t\bar{t}$ is also a QCD process, nonetheless it is treated separately.

4.1.4 $W \rightarrow \mu\nu$

The production of a W boson which decays into $\mu\nu$ has a ten times larger cross section than the signal, but a second muon has to be produced from a jet or faked by other detector effects. The probability for that is estimated to be in the order of 10^{-3} and can be determined with data [45].

4.2 Monte Carlo Generators

Monte Carlo generators are tools to stochastically simulate the physics processes that take place in collisions. By simulating large amounts of events, it is possible to translate the theoretical models into predictions for certain event variables. Thus Monte Carlo generators are an important tool to compare the theories with the measured data. Monte Carlo generators are also important to estimate the potential performance of the experiment and to test analysis methods before real data become available. The collision simulations include the structure of the proton, the hard scattering, the parton showering and the hadronization. There are different approaches on how to simulate each of the above steps, leading to a variety of different Monte Carlo generators with individual pros and cons. The event generators that have been used in this analysis are outlined in the following:

- **PYTHIA** [46] is a general purpose event generator which is often used in high energy physics. Its easy use, the large variety of supported physics processes and the sophisticated hadronisation scheme are the main advantages of PYTHIA. The hard scattering process uses leading order (LO) matrix elements and some higher order corrections are approximated with a parton shower approach.
- **PythiaB** [47, 48] is an ATLAS-specific modification of PYTHIA which can generate events containing b -quarks with higher efficiency than PYTHIA itself. It also allows to force certain decay-channels for b -hadrons.
- **HERWIG** [49, 50, 51] is another general purpose event generator quite similar to PYTHIA, but uses a different approach for parton shower and hadronisation.
- **ALPGEN** [52] is an event generator that is especially suited to describe events with large jet multiplicities. As it can only describe the hard scattering, the parton shower and hadronisation have to be performed by a different program, for example PYTHIA or HERWIG.
- **MC@NLO** [53, 54] includes full next-to-leading order calculations for the hard scattering process. Like ALPGEN it does not support parton shower and hadronisation of its own.

4.3 Samples

In order to study the properties of the signal and background processes various Monte Carlo samples have been used. As described in section 3.4.2 this analysis had to use the samples that were produced in the central ATLAS Monte Carlo generation. Due to the different needs of the various physics groups some consensus had to be reached for the type of produced samples. Therefore the available samples are not optimised to be used for a high mass Drell Yan study. In table 4.1 the samples used in this analysis are listed. All samples are generated with Athena version 14.2.0.2, simulated with Athena version 14.2.10.1 and reconstructed with Athena version 14.2.20.3. The only exception is the high mass Drell Yan sample, which is generated, simulated and reconstructed with Athena version 14.2.25. In order to reduce CPU and disk usage, all samples have been converted to DPDs locally.

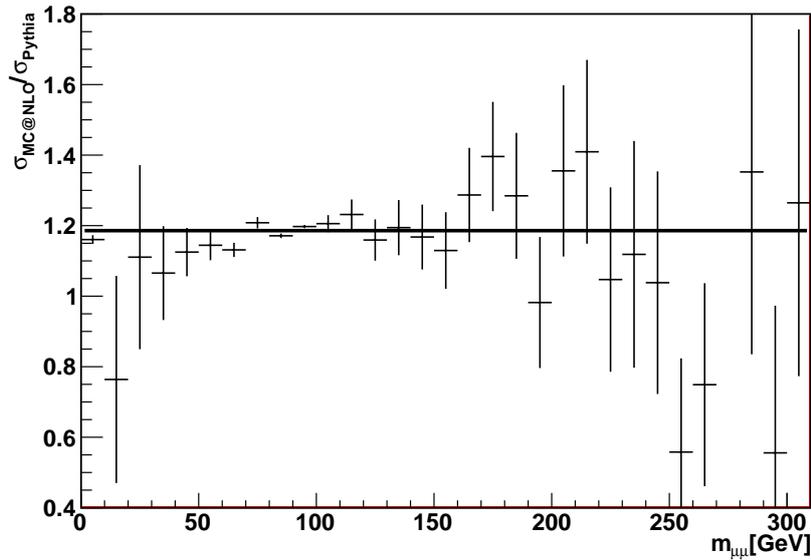


Figure 4.5: True dimuon mass distribution ratio of MC@NLO over Pythia.

Combination of Sliced Samples

It is quite common to produce some samples in different slices, e.g. with a different number of additional jets. This procedure is useful to study for example the properties of events with a high number of additional jets, even though the cross section would be too small to get a sufficient statistics with an inclusive Monte Carlo production. Each slice will have a different cross section, and thus the different subsamples have to be

Process	MC Generator	$\sigma_{NLO} \cdot BR[\text{pb}]$	MC Events	$\int \mathcal{L} [\text{pb}^{-1}]$
$Z \rightarrow \mu^+ \mu^-, m_{Z/\gamma^*} > 60 \text{ GeV}$	Pythia	1302	2.94 M	2258
$Z \rightarrow \mu^+ \mu^-, m_{Z/\gamma^*} > 60 \text{ GeV}$	MC@NLO	1303	189 k	145
$Z \rightarrow \mu^+ \mu^-, 60 \text{ GeV} < m_{Z/\gamma^*} < 200 \text{ GeV}$	Alpgen	1298	617 k	N.A. ⁵
$Z \rightarrow \mu^+ \mu^-, m_{Z/\gamma^*} > 200 \text{ GeV}$	Pythia	1.921	15 k	7808
$Z \rightarrow \mu^+ \mu^-, m_{Z/\gamma^*} > 75 \text{ GeV}$	Pythia	731	238 k	N.A. ⁵
$t\bar{t}$ decaying not fully hadronically	MC@NLO	216	487 k	2255
$Z \rightarrow \tau^+ \tau^-$	Pythia	1357	188 k	123.3
$W \rightarrow \mu \nu$	Alpgen	16149	568 k	N.A. ⁵
$W b \bar{b}$	Alpgen	17.86	45 k	N.A. ⁵
$b \bar{b} \rightarrow \mu_{15}$	PythiaB	88500	40 k	0.45
$b \bar{b} \rightarrow \mu_5 \mu_5$	PythiaB	600000	902 k	1.5

Table 4.1: List of used samples

scaled according to their integrated luminosity. After combining several subsamples to one combined sample the cross section is calculated as the sum of the cross sections of all subsamples. The calculation of an integrated luminosity for the combined sample is not possible anymore, as the scaling factors are different for each slice. See appendix E for a detailed list of used subsamples.

In order to combine different subsamples, it is important that no overlap between the samples can exist. For this reason the Pythia $Z \rightarrow \mu^+ \mu^-$ samples with $m_{Z/\gamma^*} > 60 \text{ GeV}$ and $m_{Z/\gamma^*} > 200 \text{ GeV}$ cannot be combined, because a double counting of all events with $m_{Z/\gamma^*} > 200 \text{ GeV}$ would occur. Unfortunately the cut on m_{Z/γ^*} is made on generator level, but m_{Z/γ^*} is not accessible in the DPD which is used in this analysis. Therefore this overlap cannot be removed afterwards.

$Z \rightarrow \mu^+ \mu^-$ Pythia Sample

The main signal sample used for this analysis is generated with Pythia and has a generator level cut on $m_{Z/\gamma^*} > 60 \text{ GeV}$ and at least one muon with $|\eta| < 2.8$. Three million events have been produced, resulting in an integrated luminosity of 2.3 fb^{-1} . This sample will from now on be referenced as Pythia sample.

$Z \rightarrow \mu^+ \mu^-$ MC@NLO Sample

In order to study effects of next-to-leading order (NLO) corrections a signal sample produced with MC@NLO has been studied as well. The events were passed to HERWIG for

⁵This sample is produced in several slices with different cross sections. Therefore an integrated luminosity is not applicable. See appendix E for a detailed list of subsamples used.

parton showering. Only 189 k events have been produced with this sample, corresponding to an integrated luminosity of 145 pb^{-1} . Thus the Pythia sample with the much higher statistics has been used for the main part of this analysis, while the MC@NLO sample was used to determine the impact of NLO corrections. As an effect of the NLO corrections 8% of the events in this sample have a negative weight which is taken into account in all distributions. This sample will from now on be referenced as MC@NLO sample.

$Z \rightarrow \mu^+ \mu^- + \text{jets}$ Alpgen Sample

To estimate the impact of high jet multiplicities a $Z \rightarrow \mu^+ \mu^-$ sample produced with Alpgen has been studied as well. It has six slices ranging from zero to five or more additional jets. Unfortunately there is a cut on generator level to require $60 \text{ GeV} < m_{Z/\gamma^*} < 200 \text{ GeV}$ which prevents this sample from being used in the main part of this analysis.

High Mass Drell Yan Sample

For studies of very high mass Drell Yan a sliced $Z \rightarrow \mu^+ \mu^-$ sample with $m_{Z/\gamma^*} > 75 \text{ GeV}$ and 11 slices up to 2 TeV has been produced with Pythia. There are some big differences between the high mass Drell Yan sample and the Pythia sample: Athena version 14.2.25 has been used to reconstruct this sample, and some changes have been made to the b -tagging algorithms in this version. This results in an incompatibility between the different samples, and thus the high mass Drell Yan sample can not be used to be compared with other samples. Also the Monte Carlo generation is done with different parameters, which results in the significantly smaller $\sigma_{NLO} \cdot \text{BR}$ (see appendix C.2 for details). Also the very low statistics around the Z -peak prevents this sample from being used as main signal sample.

However, some variables like \cancel{E}_T and isolation criteria have not changed significantly between those versions and therefore the high mass Drell Yan Sample can be used with some limitations. Especially for studying the stability of cuts at very high dimuon masses this sample will increase the available statistics. In appendix C.2 the difference of the two versions is studied in detail.

$t\bar{t}$ Sample

The $t\bar{t}$ sample has been produced with MC@NLO as well. A cut on generator level requires at least one of the W bosons to not decay completely hadronically. In this sample 13% of the events have a negative weight.

$b\bar{b}$ Sample

Two samples with $b\bar{b}$ production were produced using PythiaB. One sample with a cut on at least one muon with a p_T above 15 GeV, the other sample with a cut on at least two muons above 5 GeV. Both samples have an overlap, so it is not possible to combine them into one sample. The latter sample has a higher statistics as well as a higher impact on this study, so it is used from now on as $b\bar{b}$ sample. The samples were generated using leading order cross sections only, but since the uncertainty on the cross section is larger than the k-factor, this is not a problem.

$Z \rightarrow \tau^+\tau^-$ Sample

The $Z \rightarrow \tau^+\tau^-$ sample has been produced with Pythia and used the same generator level cut on $m_{Z/\gamma^*} > 75$ GeV, but no additional cuts on the τ decay products.

$W \rightarrow \mu\nu + \text{jets}$ Sample

The $W \rightarrow \mu\nu$ sample is produced with Alpgen and sliced into six subsamples from zero to 5 or more additional jets. The k-factor of 1.22 which was used here was calculated for 14 TeV [55]. Moreover the k-factor is the same for all subsamples, even though some studies [56] suggest, that each subsample should have its k-factor calculated independently. This will result in a systematic uncertainty of the $W \rightarrow \mu\nu$ cross section of about 20%.

$Wb\bar{b} + \text{jets}$ Sample

Like all other Alpgen samples this sample is sliced into subsamples with different numbers of jets, in this case from zero to three or more jets. There is a small overlap of 10% with the $W \rightarrow \mu\nu + \text{jets}$ sample, due to the W from $Wb\bar{b}$ decaying into a muon plus a neutrino. Both samples are used together in this analyses, and the cross section of the $Wb\bar{b}$ has been reduced by 10% for scaling to compensate for the overlap. This procedure will introduce some inaccuracy, but since the background from $Wb\bar{b}$ is much smaller than the one from $t\bar{t}$ this is acceptable.

Cross sections

In order to compare different samples they have to be scaled according to their integrated luminosities. The cross section given by the different Monte Carlo generators are not

directly comparable, since some of them are leading order and some are partially NLO. Thus a so called k-factor is defined as:

$$\text{k-factor} = \frac{\sigma_{NLO}}{\sigma_{LO}}, \quad (4.1)$$

and all samples are scaled to their NLO cross section. The cross sections in table 4.1 are already given at NLO using the appropriate k-factors. For the Pythia $Z \rightarrow \mu^+\mu^-$ sample no k-factor is available, thus the NLO cross section has to be determined separately. The NLO cross section of the MC@NLO sample is known, so the NLO cross section of the other signal samples can be derived from that. Due to slightly different kinematic cuts on generator level it is not clear whether the NLO cross section of the MC@NLO and the Pythia sample are identical. Thus the distribution of the true dimuon mass is used to determine the k-factor of the Pythia sample: Both samples are scaled with cross section over the number of events in the sample, and the true dimuon mass distribution from MC@NLO is divided by the one from Pythia. A constant is then fitted and the result of 1.186 ± 0.003 is used as a k-factor for the Pythia sample. This is shown in fig. 4.5. The same procedure is used for the $Z \rightarrow \mu^+\mu^-$ Alpgen sample, where the correction factor is 1.078 ± 0.003 .

4.4 Existing Analyses

The existing analyses for $Z \rightarrow \mu^+\mu^-$ have a different scope than the present analysis. They are either focused on the Z peak [24] or are model dependent searches for hypothetical Z' or other resonances at the TeV scale [2]. In the following the differences of the various approaches are discussed:

4.4.1 Z peak

In the dimuon mass region around the Z peak isolation criteria will be an efficient way to reduce background from the processes described before. As shown in fig. 4.6 the highly increased cross section at the Z peak together with the isolation cuts will provide a good signal to background ratio. This will allow a measurement of the cross section of $Z \rightarrow \mu^+\mu^-$ with an expected precision of

$$\frac{\Delta\sigma}{\sigma} \approx 0.006(\text{stat}) \pm 0.008(\text{sys})_{-0.008}^{+0.016}\text{pdf} \quad (4.2)$$

with an integrated luminosity of 50pb^{-1} , not including uncertainties in the luminosity measurements.

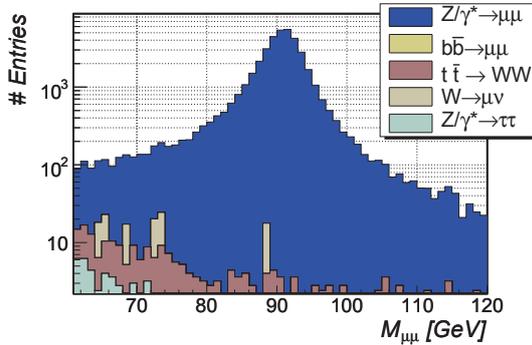


Figure 4.6: Expected reconstructed invariant mass spectrum $m_{\mu\mu}$ after isolation cuts [24].

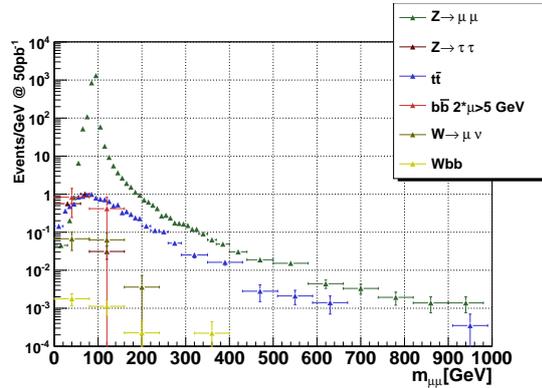


Figure 4.7: Expected dimuon mass spectrum with the cuts used for Z' searches.

4.4.2 Z'

The existing studies on the high mass tail of the Drell Yan spectrum are focused on searching new physics predicted by various theoretical models. The search for Z' is the most prominent one and will be discussed here briefly. The approaches for other model dependent searches are similar, and will not be discussed here.

The CDF experiment at Tevatron excluded a standard model like Z' for masses below 1 TeV [57] and thus the main focus of Z' searches is at masses above 1 TeV. Due to the resonance of the Z' the hypothetical signal will be much higher than the Drell Yan spectrum. Other background processes like $t\bar{t}$ will contribute less than 10% of the Drell Yan spectrum. It is then argued, that the Drell Yan process is the most important background for Z' searches, and all other backgrounds can be neglected. The typical cuts for a Z' search only require two muons with p_T above 30 GeV, opposite charge and some quality criteria. Since the main background is irreducible no further cuts like isolation are applied. In fig. 4.7 the expected high mass Drell Yan spectrum for a similar collection of cuts is shown. In the intermediate dimuon mass region above the Z peak but below one TeV the background from $t\bar{t}$ is indeed lower than the background contribution from Drell Yan.

4.4.3 Present Analysis

As described in chapter 1 a precise measurement of the Drell Yan process is a valid approach as well. Contrary to the Z' search the background contribution from $t\bar{t}$ and other processes can not be neglected anymore. It is thus necessary to study the impact of these background processes and if appropriate apply additional cuts to suppress them. As we

will see later, additional cuts on isolation, missing transverse energy and b -tagging will help to reduce the background while increasing the precision of the measurement.

4.5 Systematic Uncertainties

There are a variety of systematic effects that will affect the presented analysis. The selection cuts will be optimised to minimise the overall uncertainty of the measured Drell Yan cross section. To achieve this, all systematic effects that correlate with the selection cuts have to be studied. In order to understand the cut optimisation in the next chapter, an overview of all systematic uncertainties will be given in this chapter. Nearly all systematic uncertainties can be applied on an event per event basis, which will be described in this chapter. The details of how the resulting impact of each systematic uncertainty is calculated will be presented in section 5.2. How those systematic uncertainties are used to optimise the selection cuts, and how each cut will be affected by those uncertainties will be explained in chapter 5.

Jet Energy Scale (JES)

To correct for various effects within the calorimeter, the measured energy deposition has to be scaled in order to obtain the true energy of the jet. The scaling factors for that will be very imprecise in the beginning of data taking. With processes like Z+jets, γ +jets and QCD dijets it will be possible to determine those factors from data. A statistical precision of 1% can be reached with an integrated luminosity of 10pb^{-1} for 100 GeV jets. However, jet reconstruction is rather complicated and thus various physical effects will systematically limit the precision of the measured JES. This limit will be somewhere between 5 - 10% for 20 GeV jets, and will slowly decrease for higher jet energies [2].

Since the main focus of this study is not on jets, a rather pessimistic global uncertainty on the JES of 10% is assumed. This is applied to an event by simply increasing the energy of all jets within the event by 10%. This systematic uncertainty is considered to be symmetric, therefore a decrease in the JES of 10% has been studied as well. The effects are comparable to an increase of 10% and from now on, only an increase in the JES of 10% will be considered.

Jet Energy Resolution

The resolution of jets also depends on various effects inside the calorimeter. Gaps in the detector, varying detector responses and physics effects in the showering can influence the detector resolution. The Atlas Hadronic Calorimeter aims for a resolution of

$60\%/\sqrt{E[\text{GeV}]} \oplus 3\%$ but until now no methods to determine the resolution from data have been established.

Therefore a rather pessimistic gaussian smearing with a width of 10% is added³ to all jet energies in an event [28].

Both the uncertainties on the JES and on the jet energy resolution might affect the analysis in several ways. Obviously, a cut on the maximum jet energy of the first, second or third jet will be affected the most. There will also be a small impact on any isolation cuts, that require a minimum distance of a muon to the closest jet above a certain threshold. Mismeasured jets will also affect the \cancel{E}_T , but this effect will be discussed separately.

Muon Momentum Scale

Several factors will have an impact on the momentum scale of the muons: The limited knowledge of the magnetic field, the alignment of the muon spectrometer and the uncertainty in the energy loss of the muons are the most important ones. Using the known Z Boson mass and $Z \rightarrow \mu^+\mu^-$ events, it is possible to measure the muon momentum scale. With 100pb^{-1} a precision of 1% for muons within a range of $20\text{GeV} < p_T < 70\text{GeV}$ can be achieved [2]. It is in principle possible to extend the measurement of the muon momentum scale to higher momenta, but a higher luminosity is needed for that [2]. To reflect the fact that for this study muons with higher momenta are needed, the uncertainty on the muon momentum scale is considered to be 2%.

To apply this systematic uncertainty to an event, the momentum of all muons has been increased by 2%. The impact of a reduction by 2% has been studied as well and is of comparable size.

Muon Resolution

The muon resolution is affected by the same factors as the muon momentum scale. By looking at the shape of the Z peak it is possible to determine the muon resolution from data. With an integrated luminosity of 50pb^{-1} the muon momentum resolution can be measured with a relative precision of 6% [24]. Since this method is limited to the Z peak, a safety factor of 2 is applied to account for uncertainties due to extrapolation from the Z peak. Thus a relative precision of 12% is used in this analysis. A common approach to describe uncertainties on the muon resolution is to smear the measured muon momentum with a gaussian distribution with the width of the assumed precision. There are several disadvantages for this method: First, the muon momentum resolution is highly dependent on η and p_T . To properly describe those dependencies the resolution and the uncertainties

³In this context a smearing of 10% means a gaussian random number with $\mu = 0$ and $\sigma = 0.1$ is calculated for each jet and the jet energy is then multiplied by that number.

should be known as a function of η and p_T . To obtain such a two dimensional function and its uncertainties, a high statistics is needed. Second, the muon momentum resolution is not completely gaussian⁴. For a majority of muons, the resolution is described well by a gaussian smearing, but the p_T resolution distribution has rather wide tails. Those tails are not negligible in this analysis, since they are an important reason for fake \cancel{E}_T and can also cause a systematic shift in the Drell Yan spectrum.

Therefore another approach is used in this analysis: The momentum of every muon is changed as follows: The difference in p_T of the reconstructed muon and the matched true muon is multiplied by the uncertainty of the resolution measurement of 12% and then added vectorially to the reconstructed muon. The difference between the true muon and the reconstructed muon already takes η and p_T dependent effects into account. The tails of the resolution are also described to some extent in the detector simulation and are therefore taken into account with this method as well.

Both uncertainties concerning muons have negligible influence on the selection criteria of this analysis, so only the indirect contribution to the \cancel{E}_T described in the next section is taken into account.

Missing Transverse Energy \cancel{E}_T

As described in section 3.3.3 \cancel{E}_T is built on many different variables and therefore also collects the systematic uncertainties from those variables. To give an estimation on the uncertainties of \cancel{E}_T the following five contributions to \cancel{E}_T are considered:

- **Muon Resolution and Muon Scale** If for one of the reasons described above a muon is reconstructed with the wrong p_T this translates directly into a fake \cancel{E}_T component of the same size and opposite direction. Especially in the signal sample, which should in principle be free of \cancel{E}_T , this is the most important contribution to a wrongly measured \cancel{E}_T .

The impact of those uncertainties to the \cancel{E}_T is calculated as follows: For every muon in the event, the systematic uncertainty is applied as described before. The vectorial difference between the changed and the unchanged muon is then added vectorially to the \cancel{E}_T .

- **Jet Energy Scale and Jet Resolution** The same argument holds true for the p_T of jets, though this is a more important contribution in the $t\bar{t}$ background sample, which has a higher jet activity.

The impact of those uncertainties is determined in the same way as for the muons.

⁴In principle the fractional muon momentum resolution described in appendix B.5 should be gaussian, but it has wide tails as well.

- **Unclustered Energy Resolution** The unclustered energy in the calorimeter is also used for \cancel{E}_T reconstruction. It contains all energy depositions in the Calorimeters, that are not assigned to jets. Detector noise, underlying event and pileup will be the most important factors for this systematic uncertainty. It is expected to be rather small in absolute terms, so the impact will be more important for events with low \cancel{E}_T .

To obtain the unclustered energy contribution to \cancel{E}_T the p_T of all reconstructed muons and jets is subtracted vectorially from the \cancel{E}_T . The X and Y components of this unclustered energy are then both smeared with a gaussian with a width of 10% [28]. Afterwards the X and Y component, the reconstructed jets and muons are added vectorially again to obtain the systematically shifted \cancel{E}_T for that event.

Those five contributions are treated as independent systematic uncertainties. For the muon and jet contributions the same parameters as above are used.

B-Tagging

B-tagging efficiency and mis-tagging rates for jets will be determined with a relative uncertainty of 10% from real data. It is however not possible to apply this on an event per event basis. Also, the cut which will be studied later relies on a combination of more than one b-tag. A detailed description on how the systematic uncertainty for b-tagging is handled will be given in section 5.3.4.

Cross Sections

Another systematic uncertainty is the cross section for the most important background processes $t\bar{t}$ and $b\bar{b}$. For $t\bar{t}$ the cross section is known with a precision of 15%. The limiting factor at this energy is mainly the uncertainty of the parton distribution function. The isolation criteria for the muons will also affect the number of selected $t\bar{t}$ events. The uncertainties for those isolation criteria are not treated individually, but are included within the background cross section uncertainty. Thus, the uncertainty for the $t\bar{t}$ background is increased to 20% [24].

The theoretical precision of the cross section for $b\bar{b}$ is not as good as for $t\bar{t}$. An uncertainty of 50% is assumed for the $b\bar{b}$ background. This uncertainty also includes the uncertainty on the isolation criteria for $b\bar{b}$ events.

Due to the unprecise k-factors for $W \rightarrow \mu\nu$ an uncertainty of 20% is assumed for the $W \rightarrow \mu\nu$ cross section as well. It turned out, that due to the low impact of the $W \rightarrow \mu\nu$ background this uncertainty can be neglected.

Uncertainties on Monte Carlo Generators

Some selection variables might depend on NLO effects or on details of the parton showering. For that reasons different Monte Carlo generators should be used to study those effects. A MC@NLO sample for $Z \rightarrow \mu^+ \mu^-$ exists, but the statistics in the high mass Drell Yan spectrum is too low to be used for the numerical cut optimisation. There is also an Alpgen sample for Z+jets, but it only includes $60 \text{ GeV} < m_{\mu\mu} < 200 \text{ GeV}$.

To check for effects from different Monte Carlo generators the following approach will be used: The Pythia sample is compared with the one from MC@NLO for $m_{\mu\mu} > 60 \text{ GeV}$. It is then assumed, that the impact of NLO effects is comparable in the high mass Drell Yan spectrum. The uncertainty from NLO is approximated as half the difference of the selection efficiencies of Pythia and MC@NLO. For each selection cut the cut variable is compared between the Pythia and the MC@NLO sample to check for systematic shifts.

The description of events with high jet multiplicity differs a lot between Pythia and Alpgen. This can be seen in fig. 4.8 and is an expected effect. The systematic uncertainties would be overestimated if this difference was taken into account. Once real data is available, the jet multiplicity of Z+jet events should be studied, and only Monte Carlo generators that can be tuned to reproduce the jet multiplicity accurately should be used to calculate selection efficiencies.

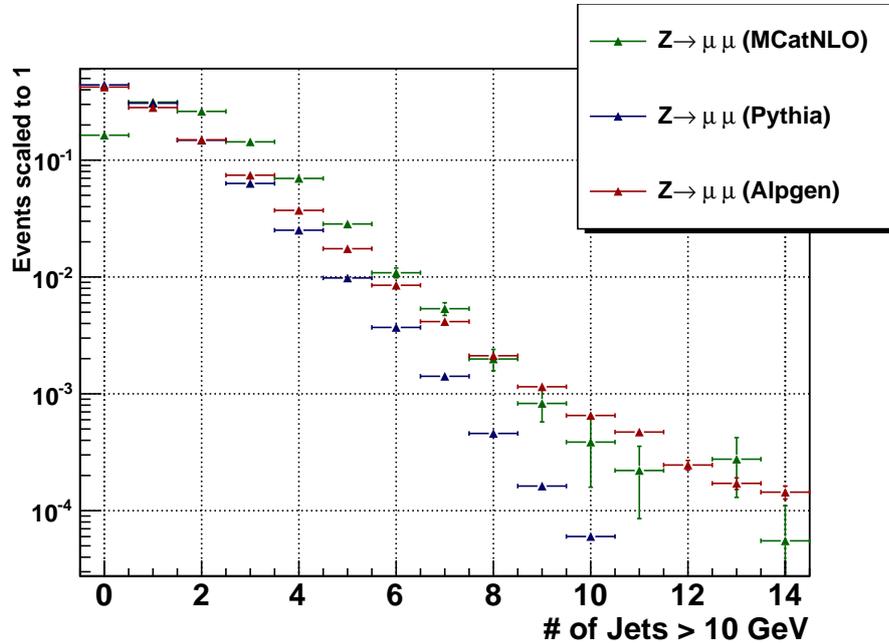


Figure 4.8: Jet Multiplicity for jets with $p_T > 10 \text{ GeV}$ for $60 \text{ GeV} < m_{\mu\mu} < 200 \text{ GeV}$.

Other uncertainties

Other systematic effects like misalignment of the Muon Spectrometer, cross sections for the signal process, luminosity and parton distribution functions have already been discussed in detail in other publications [2, 24]. Those uncertainties are either already included in the above uncertainties or do not correlate with the studied selection criteria, and therefore they are not relevant for optimising the selection cuts. It should however be noted, that those uncertainties should be taken into account if one wants to calculate discovery limits or to give an overall systematic uncertainty on the cross section measurement.

5 Cuts

In a physics wise perfect world, it would be possible to determine which physics processes happened in a given event. Since we unfortunately do live in an imperfect world, this is not the case. The closest thing we can achieve is making statistical assumptions on how likely it is that the given event contains our signal or our background.

The simplest method to do that is to develop a set of cuts on certain variables and require the event to fulfill all cuts in order to be treated as signal.

Section 5.1 describes what quality criteria a cut has to fulfill in order to be usable. In section 5.2 a method to optimise cuts is presented, and in section 5.3 all cuts used in this analysis are discussed in detail. In section 5.4 all cuts that have been studied but are not used are listed. The significance of the cut optimisation is briefly discussed in section 5.5 and the results are summarised in section 5.6.

5.1 Quality criteria for cuts

- **Reducing Background:** The most obvious criterium for a cut is the ability to reduce the background without removing too much of the signal. A good way to quantify this is the receiver operating characteristic (ROC) [58] which graphically plots the fraction of removed background against the fraction of removed signal¹. The ROC allows for a good comparison between different cuts in one single plot.
- **Stability to NLO:** NLO corrections can affect some of the variables used for cuts, especially if they are related to jets. By comparing the distribution of the cut variables between the Pythia and the MC@NLO sample the impact on the newly suggested cuts is estimated.

¹The receiver operating characteristic was developed during the second World War to analyse radar images, hence the name. Usually $1 - \text{fraction of removed signal}$ is used, which would give a mirror inverted version of the plots used in this thesis.

- **Stability to high dimuon masses:** As this analysis focuses on the high mass Drell Yan spectrum, it is important that the cut efficiency is stable even at very high dimuon masses. A cut which has an unaccounted for dependency on the dimuon mass would introduce a shift on the measured spectrum.
- **Stability to Systematic Errors:** All additional cuts will introduce new systematic uncertainties which have to be taken into account. The cut optimisation method described in the next section will minimise the impact of those uncertainties.

5.2 Cut Optimisation

The main focus of this thesis is the reduction of background through cuts, therefore a mathematical method to optimise cuts was developed, implemented and tested. The main purpose of measuring the Drell Yan spectrum at high dimuon masses is a model independent search for new physics, so the goal of all cuts can be summarized as:

Reduce the quadratic sum of syst. and stat. uncertainties of the measured cross section. (5.1)

The cross section is calculated as

$$\sigma = \frac{N_{selected} - B_{Selected}}{\epsilon_S \cdot L} \quad (5.2)$$

where $N_{selected}$ is the total number of selected events within the data and L is the integrated recorded luminosity. $B_{Selected}$ is the number of selected background events. ϵ_S is defined as the number of events with a reconstructed dimuon mass above 200 GeV divided by the number of selected events with a true dimuon mass above 200 GeV.

$B_{Selected}$ and ϵ_S have to be determined from Monte Carlo data.

As we will see later the exact value of the cuts will depend on the recorded luminosity and the resolution of various variables measured by the detector. Thus, the cut values will change with time, and only an overview can be given here.

To achieve the goals formulated in eq. 5.1 all systematic uncertainties listed in section 4.5 are taken into account, as well as all possible backgrounds described in section 4.1. The optimisation is performed on all cuts subsequently. For various reasons not all cuts are optimised with this method, as described in the description of the specific cuts.

The three main contributions to the uncertainty are:

- **Systematic Errors** For a given cut the impact of all systematics are considered separately.

The impact of a given systematic uncertainty is determined by performing the complete analysis once without any changes and once with a smearing to all the reconstructed particles according to the systematic uncertainty. The relative systematic error is then calculated from eq. 5.2 as

$$\frac{\Delta\sigma}{\sigma_{Systematic}} = \frac{\sigma - \tilde{\sigma}}{\sigma} = 1 - \left(\frac{B_{selected} + S_{selected} - \tilde{B}_{selected}}{\tilde{S}_{selected}} \right) \quad (5.3)$$

where \tilde{S}/\tilde{B} and S/B are the number of events from signal/background that survived all cuts with and without applying the systematic distortion respectively.

Since the focus of this study is the high mass tail of the Drell Yan spectrum only events with a dimuon mass above 200 GeV are taken into account.

- **Statistical Uncertainty** The relative statistical uncertainty is calculated as

$$\frac{\Delta\sigma}{\sigma_{Statistical}} = \frac{\sqrt{S_{selected} + B_{selected}}}{S_{selected}} \quad (5.4)$$

for a bin from 200 GeV to 300 GeV. This uncertainty introduces two opposing effects: The dependency on $S_{selected}$ will in general favour looser cuts, while the term $B_{selected}$ in the numerator will favour stronger cuts. This is especially important for low integrated luminosities, since the statistical uncertainty will dominate the overall uncertainty in this regime.

- **Systematic Uncertainty due to limited Monte Carlo Statistics** Another contribution to the total systematic uncertainty is the statistical uncertainty on the efficiencies $\epsilon_B = \frac{B_{Selected}}{B_{Total}}$ and ϵ_S due to limited Monte Carlo statistics. This uncertainty depends on the specific efficiencies ϵ_B and ϵ_S themselves, and therefore also on the value of the cut. Hence, it is also included here.

$$\frac{\Delta\sigma}{\sigma_{MonteCarlo}} = \frac{1}{S_{Total} \cdot \epsilon_S} \cdot \sqrt{S_{Total} \cdot \epsilon_S \cdot (1 - \epsilon_S) + B_{Total} \cdot \epsilon_B \cdot (1 - \epsilon_B)}, \quad (5.5)$$

where S_{Total} and B_{Total} are the number of produced Monte Carlo events for signal and background respectively.

It is assumed, that all statistical and systematic uncertainties are uncorrelated, which allows us to add the contribution of all contributing errors quadratically. Therefore the overall systematic uncertainty is

$$\frac{\Delta\sigma}{\sigma} = \sqrt{\left(\frac{\Delta\sigma}{\sigma_{Statistical}} \right)^2 + \sum_{System.} \left(\frac{\Delta\sigma}{\sigma_{Systematic}} \right)^2 + \left(\frac{\Delta\sigma}{\sigma_{MonteCarlo}} \right)^2} \quad (5.6)$$

All of those terms depend on the selected cut value. To optimise a given cut, the selection is performed with all other cuts enabled. For every possible cut value the variables $\epsilon_S, \epsilon_B, S_{Selected}$, and $B_{Selected}$ are calculated. With those variables we can calculate the overall uncertainty of the cross section from eq. 5.6 and plot it versus the cut value. The minimum in that histogram is the optimised cut value for that cut. If there is no minimum, the cut is not applied. The plots for all studied cuts will be discussed in detail in section 5.3.

This procedure is run consecutive for all cuts. As we will see later, only two additional cuts are applied, a cut on E_T and one using b -tagging. Those cuts are not completely uncorrelated, so the optimisation of the first might change, after the second is changed. Therefore the optimisation is iterated until the values do not change any more, which is described in detail in appendix D. The presented optimisation plots are shown after this iteration is performed, in order to illustrate the way how the cuts affect each other.

5.3 List of Cuts

5.3.1 Preselection

At least two muons

In order to identify a dimuon event, the most obvious requirement is at least two reconstructed muons. Both muons are required to have $p_T > 20\text{ GeV}$. Even though indispensable, this cut removes 58% of the simulated muon events². One limiting factor is the geometric acceptance of the ATLAS detector. As described in section 3.2 only muons with $|\eta| < 2.7$ can be reconstructed in the Muon Spectrometer. Another limiting factor is the reconstruction efficiency of about 95% for muons with $p_T > 20\text{ GeV}$.

Exactly two muons

In this analysis only events with exactly two muons with $p_T > 20\text{ GeV}$ are taken into account³. Additional muons can have various sources: Soft muons from jets, fake muons or muons from pileup events. All those muons have a rather low transverse momentum and thus only 0.02% of the simulated signal events are removed by this cut. Pileup events are completely uncorrelated to the dimuon mass of the event, and thus only the overall cross section could be affected, but not the shape of the spectrum. However, no simulated

²Only events with at least one muon with $|\eta| < 2.8$ are simulated. The percentage of all $Z \rightarrow \mu^+ \mu^-$ events that are removed would be even higher.

³In fact only muons that are reconstructed as *combined muons* (see Section 5.3.2) are counted for this cut.

pileup events are available for the software version used in this analysis, so this effect has not been studied in detail.

Opposite charge

The two muons are required to have opposite charge. This reduces the background from $t\bar{t}$ by 24%, while it has only very small effect on the signal. The 0.01% of events that are rejected by this cut have a mis-reconstructed charge of the muon, a so called charge flip. Charge flips only occur together with a massive mis-reconstruction of the muon p_T , so those can not be used in this analysis.

5.3.2 Quality Cuts

To ensure a well reconstructed dimuon mass the muons have to fulfill some quality criteria in order to be used. All of those criteria have to be met for both muons.

Combined Muons

As described in section 3.3.2 there are two subdetectors that can reconstruct muons: The Inner Detector and the Muon Spectrometer. It is in principle possible to use muons that are only reconstructed by the muon spectrometer. In this analysis only muons that have a matching track in both the Inner Detector and the Muon Spectrometer are used. Those muons are called combined muons. The advantages of using combined muons are the decreased rate of fake muons and the increased momentum resolution. The disadvantage is the slightly decreased efficiency for muon reconstruction.

Track Matching Quality ($\frac{\chi_{\text{match}}^2}{N_{\text{DOF}}}$)

An additional advantage of only using combined muons is the possibility to use the quality of the matching between Inner Detector and Muon Spectrometer as selection variable. For this the variable $\chi_{\text{match}}^2/N_{\text{DOF}}$ is used. If one of the two subdetectors mis-reconstructed the muon momentum by a large amount, the quality of the matching will decrease, and thus the $\chi_{\text{match}}^2/N_{\text{DOF}}$ will increase. The correlation between $\chi_{\text{match}}^2/N_{\text{DOF}}$ and the muon momentum resolution can be seen in fig. 5.1. The distribution of $\chi_{\text{match}}^2/N_{\text{DOF}}$ is compared for various backgrounds in fig. 5.2. As expected the quality of the matching for the muons is comparable between $t\bar{t}$ and $Z \rightarrow \mu^+\mu^-$.

A cut on $\chi_{\text{match}}^2/N_{\text{DOF}}$ of 5 is well established [2] and thus will be used here as well.

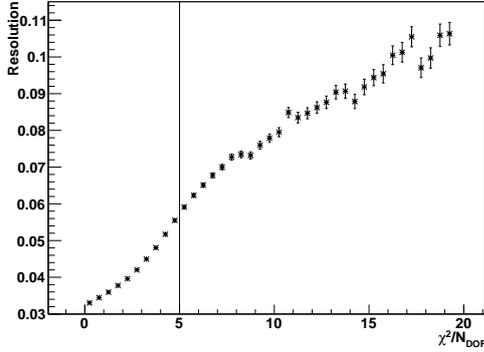


Figure 5.1: $\frac{\chi^2_{\text{match}}}{N_{\text{DOF}}}$ vs. the average resolution of muons in the signal sample.

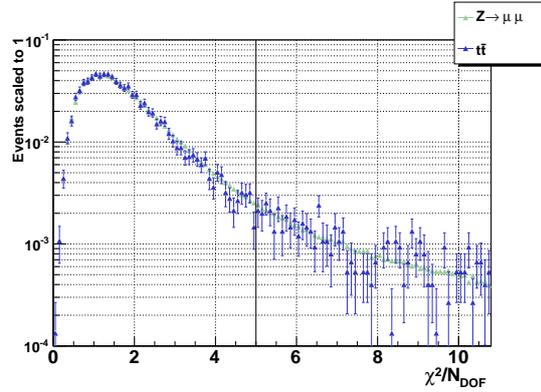


Figure 5.2: Distribution of the maximum $\frac{\chi^2_{\text{match}}}{N_{\text{DOF}}}$ for both muons.

Pseudorapidity constraint ($|\eta| < 2.5$)

The geometry of the Inner Detector allows a precise reconstruction of tracks with an $|\eta|$ of up to 2.5. It is well established within the ATLAS collaboration to only use muons with $|\eta| < 2.5$ for analyses that require high quality muons. Therefore the same cut is applied in this analysis.

5.3.3 Isolation Cuts

Muons from the leptonic Z boson decay appear isolated in the detector, while muons from QCD interactions are mostly produced within jets. At high dimuon masses QCD interactions do not contribute significantly to the background, and $t\bar{t}$ events become the most important background. Dileptonic $t\bar{t}$ events can produce two isolated muons as well, hence the isolation cuts are not very efficient in that regime. Therefore an optimisation as described in section 5.2 is not applicable. It is however important to reduce the background contribution also for dimuon masses around the Z peak, in order to measure a precise Drell Yan spectrum. Isolation criteria for $Z \rightarrow \mu^+\mu^-$ are studied thoroughly in previous publications [24, 2], and will be optimised with real data. The main focus of this analysis is to study additional cuts, thus no extensive optimisation for the isolation cuts has been done. The cuts variables and cut values are chosen to achieve a similar background suppression as previous studies, in order to get comparable results. The specifics of the high mass drell spectrum are discussed and taken into account.

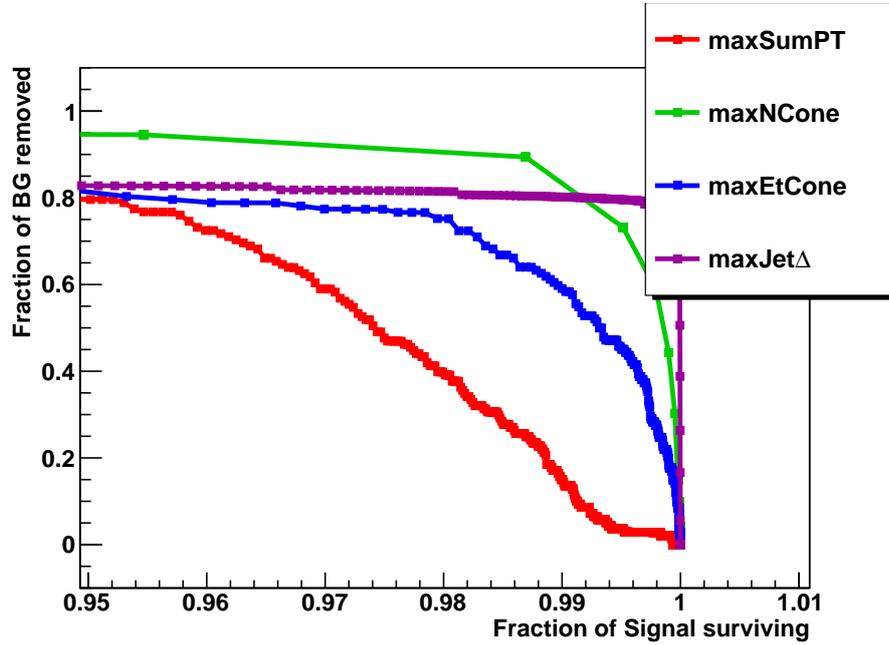


Figure 5.3: Comparison of various isolation methods.

Distance to closest jet ($maxJet\Delta R$)

One possibility is to measure the distance⁴ ΔR from a muon to its closest jet. If there is a jet in close proximity to the muon, it is possible that the muon was produced in a decay cascade within that jet. Since the preselection requires muons with a p_T of above 20 GeV, only jets with an energy above 20 GeV are taken into account. Each preselected event contains two muons, and the highest distance to the closest jet of both muons is used.

$$Jet\Delta R = \min(\Delta R(\mu, jet^n)) \quad (5.7)$$

$$maxJet\Delta R = \max(Jet\Delta R, Jet\Delta R) \quad (5.8)$$

If the event contains no jet, the value of $maxJet\Delta R$ is set to 9. In other words, only one muon has to be isolated with respect to the closest jet to fulfill this isolation requirement. Thus the impact of underlying event, an additional jet from Initial State Radiation, and other NLO effects can be kept to a minimum, because it is very unlikely, that an additional jet is produced close to each muon. In fig. 5.3 this variable is compared to some other isolation criteria, and it turns out to be the most efficient one. The cut value is set to 0.7, which removes 78% of the $b\bar{b}$ background if only preselection and quality cuts are applied, while only 0.03% of the signal are rejected. In fig. 5.4 the distribution of $maxJet\Delta R$ after

⁴ ΔR is a measure of distance commonly used in ATLAS. It is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

preselection and quality cuts is shown while the stability to high dimuon masses is shown in fig. 5.5.

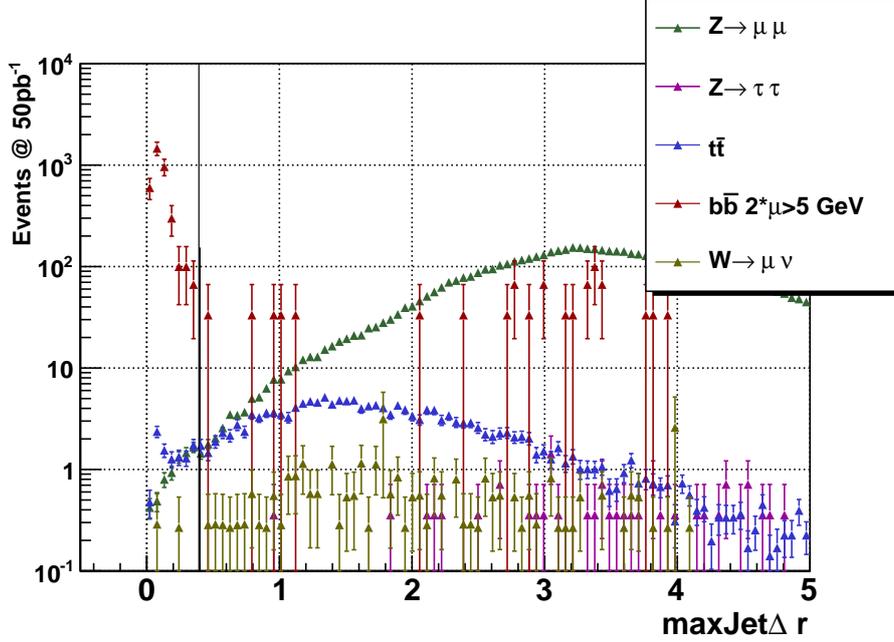


Figure 5.4: Distribution of $maxJet\Delta R$ after preselection and quality cuts.

Number of tracks close to the muon ($maxNCone$)

This variable counts all additional tracks within a cone of $\Delta R < 0.3$ around the muon:

$$NCone = \sum_{\Delta R < 0.3} N^{ID \text{ Tracks}} \quad (5.9)$$

Both muons are required to have less than 3 additional tracks within the cone, so a variable

$$maxNCone = max(NCone^{\mu_1}, NCone^{\mu_2}) \quad (5.10)$$

is defined. This cut is tighter than the cut on $maxJet\Delta R$ and removes 1.5% of the signal, but removes 93% of the $b\bar{b}$ and 40% of the $t\bar{t}$ background. The distribution of $maxNCone$ is shown in fig. 5.6, and the stability for high dimuon masses is shown in fig. 5.7.

Sum of p_T 's close to the muon ($maxSumPT$)

An alternative to counting the number of tracks close to the muon is to use the scalar sum of the p_T 's of reconstructed tracks in the Inner Detector with a ΔR to the muon below 0.3,

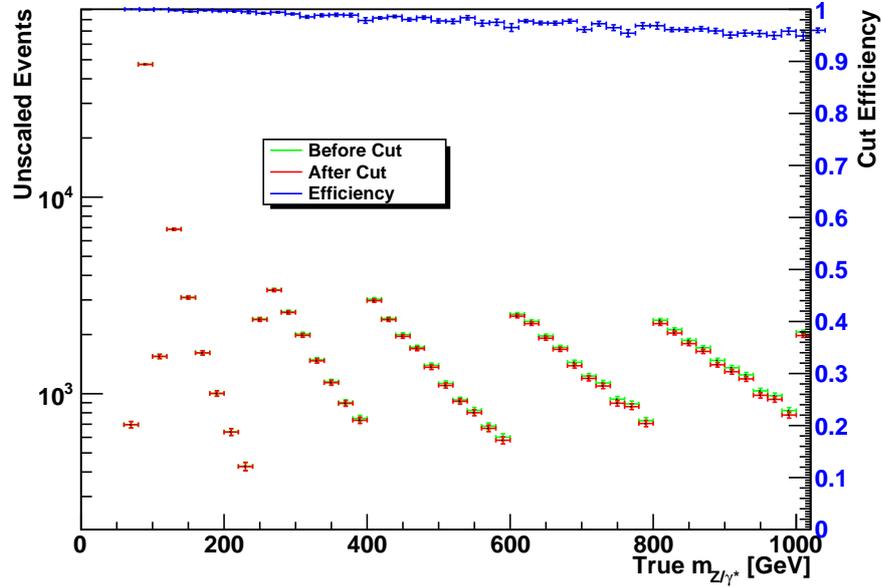


Figure 5.5: Cut efficiency for $\text{maxJet}\Delta R$ up to 1 TeV (blue) and its effect on the number of events selected by that cut (red).

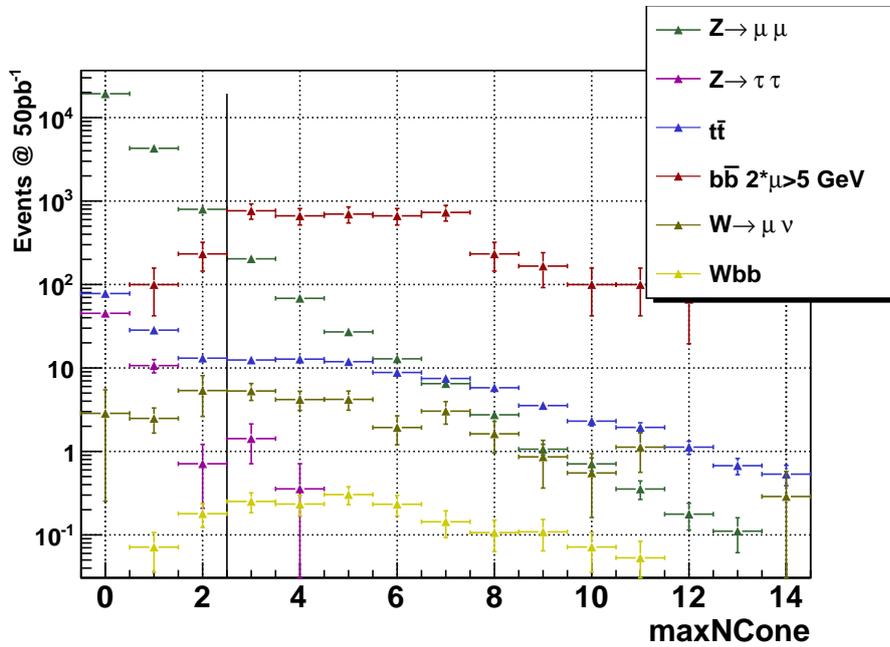


Figure 5.6: Distribution of maxNCone after preselection and quality cuts.

the so called *SumPT*.

$$SumPT = \sum_{\Delta R < 0.3} p_T^{\text{ID Tracks}} \quad (5.11)$$

This variable takes into account that tracks with higher momentum are a stronger indication of background processes, while tracks with very low p_T have a higher probability of being fake⁵ tracks.

$$maxSumPT = \max(SumPT^{\mu_1}, SumPT^{\mu_2}) \quad (5.12)$$

As for *maxNCone* both muons have to fulfill the cut, so *maxSumPT* is the actual cut variable. In fig. 5.3 *maxSumPT* is compared to the already described *NmaxCone*. Even though cutting on *maxSumPT* is better motivated physics wise, *maxNCone* turns out to be a more efficient selection criterion for this analysis. Once enough real data is available the performance of both isolation criteria should be compared. Until then the more efficient cut on *maxNCone* is used.

Calorimeter Isolation (*maxEtCone*)

A commonly used isolation criterion is the energy deposited in the calorimeters close to the extrapolated track of the muon.

$$EtCone = \sum_{\Delta R < 0.3} E^{\text{Calorimeter}} \quad (5.13)$$

$$maxEtCone = \max(EtCone^{\mu_1}, EtCone^{\mu_2}) \quad (5.14)$$

In fig. 5.3 a cut on *EtCone* of the least isolated muon (*maxEtCone*) is compared to the other isolation criteria. It turns out that a cut on *maxEtCone* is less effective than the other isolation cuts, and is therefore not used in this analysis. One reason for the weaker performance of *maxEtCone* is the fact, that even isolated muons will deposit energy in the calorimeter. There are two ways to estimate the energy deposition of the muon: The first one is to compare the reconstructed muon p_T in the Inner Detector with the reconstructed muon p_T in the Muon Spectrometer, and thus measure the energy deposition. The second method is to parametrize the deposited energy from the reconstructed p_T , η and ϕ of the muon. Especially for high p_T muons the deposited energy has a broad distribution with strong tails, thus those corrections are not very precise.

⁵Fake tracks in this context means tracks, that are produced by detector noise and have no corresponding truth particle

Isolation efficiencies

In order to measure a cross section, the isolation efficiencies must be known. It is in principle possible to determine isolation efficiencies for single muons with the so called 'tag and probe' method, which is described in [24]. However the isolation probability of the two muons is not uncorrelated⁶, and there is a slight correlation between the isolation efficiency and the dimuon mass⁷. Similar studies at the Z peak predict an uncertainty of $0.0012(stat) + 0.003(sys)$ for a tight muon isolation, which is considerably smaller than for example the statistical uncertainty for high mass Drell Yan measurements.

5.3.4 b -Tagging

The signal process $Z/\gamma^* \rightarrow \mu^+\mu^-$ should only contain very few b -jets, while the two most important backgrounds $t\bar{t}$ and $b\bar{b}$ will typically contain at least two b -jets.

Algorithms

As described in section 3.3.4 there are several algorithms available that quantify the likelihood whether a jet contains b -quarks or not. As shown in appendix C.1 the best algorithm for this cut is IP3D+SV1. The weights of all the jets within one event can be combined in various ways to obtain a cut variable:

- **Cut on the weight of the highest weighted jet** This method assumes that one b -jet is enough to identify an event as background. One advantage of that method is, that background events where the second b -jet is not tagged as b -jet, or not reconstructed at all are still rejected. The downside is, that the information about the second b -jet is not used at all.
- **Cut on two jets above a certain threshold** This method requires two jets with a b -weight above a certain threshold to reject the event. It uses the information from both jets, but requires both jets to be reconstructed and b -tagged.
- **Cut on the sum of the two highest weighted jets** This method is a combination of the first two methods. The b weight of the two highest weighted jets is summed up. If one jet has a high b -weight a second jet is not required, while two b -jets with a smaller b -weight are also enough to reject the event.

⁶Events with many reconstructed jets are more likely to produce tracks or jets close to the muon

⁷Events with a higher dimuon mass have a higher average number of jets

- **Cut on the sum of the two highest positive weighted jets above 10 GeV** This method is an improvement of the third method. It only takes jets into account that have a p_T above 10 GeV and a positive b -weight. Both requirements reduce the impact of additional low p_T jets, and thus the dependency on NLO corrections.

In fig. 5.8 all four methods are compared. As expected, the fourth method is the most efficient one. The drawback of method four is the unknown efficiency of a cut based on this method. The b -tagging performance group will measure and provide efficiencies and rejection rates for various b -tagging algorithms at different working points. The fourth method is a cut on a derived variable, and the efficiency will be hard to measure in real data. Thus the efficiency has to be determined from Monte Carlo, which is only possible, if the Monte Carlo is tuned well enough to describe the b -tagging efficiency and rejection rate with a reasonable uncertainty. Whether or not this is the case should be checked with real data.

The same difficulties arise while estimating the systematic uncertainties of the cut on b -tagging. For a b -tagging cut with an efficiency of 60% the b -tagging efficiency will be determined with an estimated relative uncertainty of 10% with 100pb^{-1} of data. The relative uncertainty on the rejection rate at the same efficiency will also be 10% [2]. For this analysis a relative uncertainty of 10% on the signal efficiency is used, as well as a relative uncertainty of 10% on the number of background events rejected by that cut. This assumes that the efficiency for a b -tagging cut which is a combination of two b -weights can be determined as precise as the b -tagging for one jet.

In summary, the fourth method is more sophisticated and will require further work once enough data are available. Yet it offers a higher background rejection and is better motivated physics wise. Therefore this analysis will use a cut on the sum of the two highest positive weighted jets above 10 GeV.

Fig. 5.9 shows the distribution of the b -tagging cut value as defined by method four.

Optimisation Results

In fig. 5.10 the optimisation described in section 5.2 is shown. The most important thing to look at in this plot is the overall uncertainty as calculated from eq. 5.6, which includes all statistical and systematic uncertainties. Several interesting effects and contributions can be discussed in this plot. For better visibility fig. 5.10(b) shows a magnification of the overall uncertainty and fig. 5.10(c) shows a magnification of the systematic uncertainties.

One contribution favouring a stronger cut, is the systematic uncertainty on the cross section of the $t\bar{t}$ background. The harder the cut on the b -tagging is, the lower the percentage of $t\bar{t}$ in the selected signal, and thus the effect of an over or underestimated $t\bar{t}$ cross section will decrease.

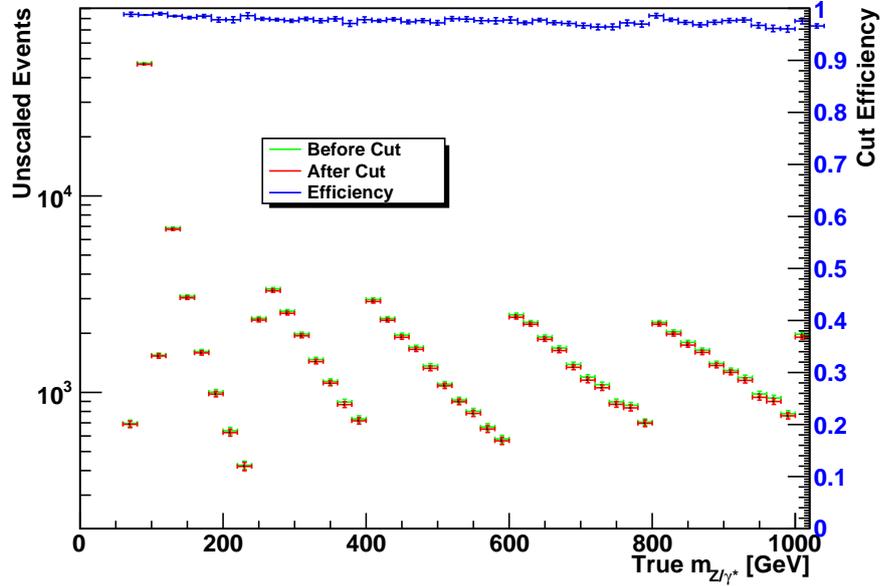


Figure 5.7: Cut efficiency for *maxNCone* up to 1 TeV (blue) and its effect on the number of events selected by that cut (red).

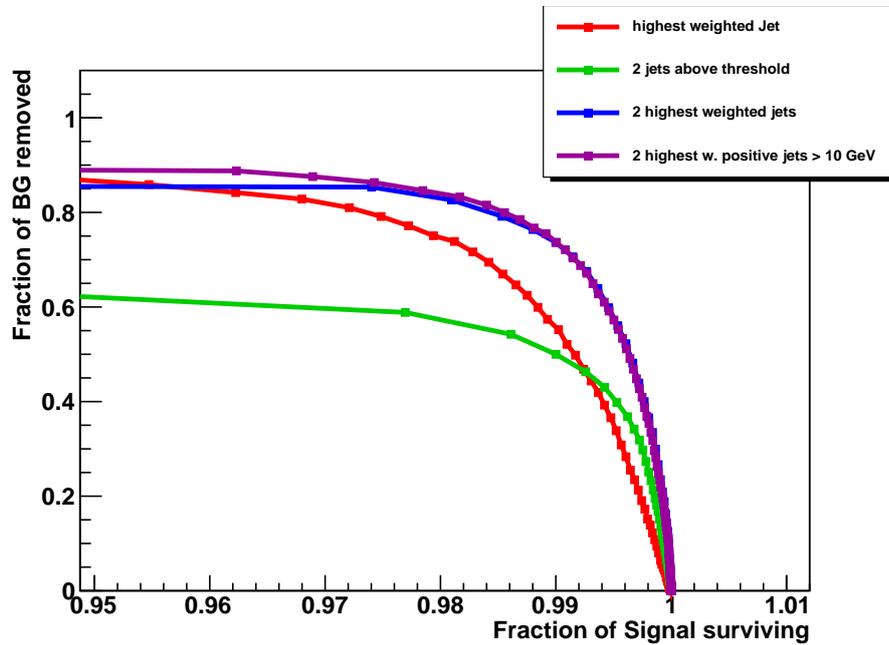


Figure 5.8: Comparison of the four methods for *b*-tagging. Only the signal and $t\bar{t}$ sample are included and preselection, quality and isolation cuts have been applied.

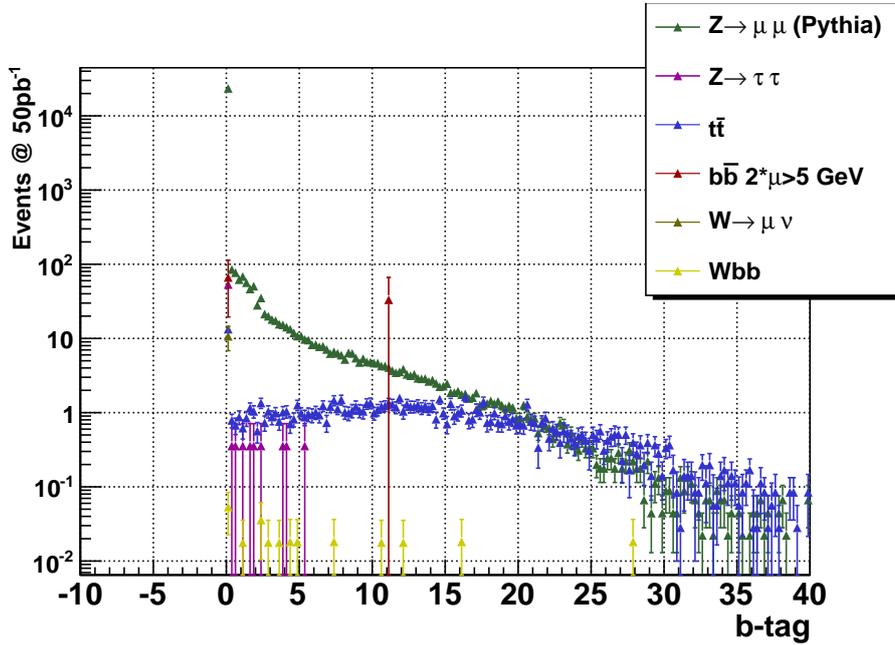


Figure 5.9: Distribution of b -Tagging weight.

A more indirect contribution is introduced by systematic uncertainties on variables affecting the cut on \cancel{E}_T which will be described later. The argument is similar to the one for the uncertainty on the $t\bar{t}$ cross section. A systematic shift affecting the \cancel{E}_T will cause background events to pass the cut on \cancel{E}_T while they are expected to be removed, and thus causing an underestimated background (or vice versa). If the signal to background ratio is improved by a harder cut on the b -weight, the impact of those uncertainties on \cancel{E}_T will be decreased as well. Therefore the uncertainty on the jet energy resolution and the uncertainty of the unclustered energy in the calorimeter will favour harder cuts on b -Tagging.

Two factors favouring looser cuts are the uncertainties that affect the cut on b -tagging itself, namely the uncertainty on mis-tagging and the uncertainty on the tagging efficiency. It can be seen from fig. 5.10, that the effect is much larger for the tagging efficiency. This is as expected, since only a few percent of the signal will be removed, so a 10% uncertainty on that percentage will create a very small error. For the b -tagging efficiency the effect is bigger, since up to 60% of the background can be removed with a very strong cut.

The dominating contribution for low luminosities is the statistical uncertainty though. To calculate the cross section of the signal process, the number of estimated background events has to be subtracted. Therefore the statistical uncertainty on the expected background is also part of the overall uncertainty, as seen in eq. 5.6. This statistical uncertainty gets reduced, if the background is reduced, therefore favouring harder cuts. If the

cut gets too hard, too many signal events are rejected, therefore increasing the statistical uncertainty. Thus, the statistical uncertainty has a minimum, which will dominate the overall uncertainty for low luminosities.

In fig. 5.10 a clear minimum for a cut on b -tagging can be seen for 50pb^{-1} . The overall uncertainty without the cut would be 17.81%, the global minimum is at a cut value of 3, resulting in an overall uncertainty of 17.22%. The minimum is rather broad, ranging from a cut value of 2 (17.24%) to 6.5 (17.23%). In order to minimise unaccounted systematic effects, it is best to choose the loosest possible cut, so a cut value of 6.5 is used. As mentioned before, the optimisation is in principle done iteratively, but since the systematic effect is dominating, the results do not change after the first iteration step.

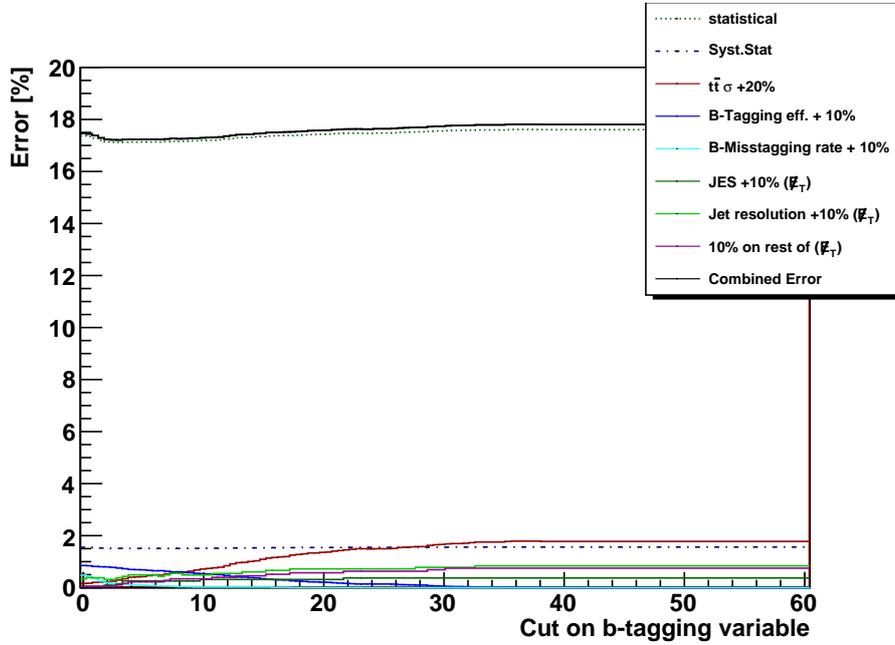
The optimisation plots for other luminosities are shown in fig. 5.11 and 5.12, the results are presented in table 5.1. However, those have to be taken with a grain of salt, as we can expect a decrease of systematic uncertainties with higher luminosities. Unfortunately it is already very hard to estimate reasonable values for expected uncertainties, but it is nearly impossible to predict how those systematic uncertainties will develop over the years. The main conclusion that can be drawn from those higher luminosities is the fact that even though the statistical uncertainty is the dominating effect for low luminosities it is not the only effect favouring a cut on b -tagging.

	50pb^{-1}	500pb^{-1}	5fb^{-1}
Uncertainty w/o cut	17.8%	6.2%	3.2%
Uncertainty with cut	17.2%	5.7%	2.5%
Cutvalue	6.5	3	3

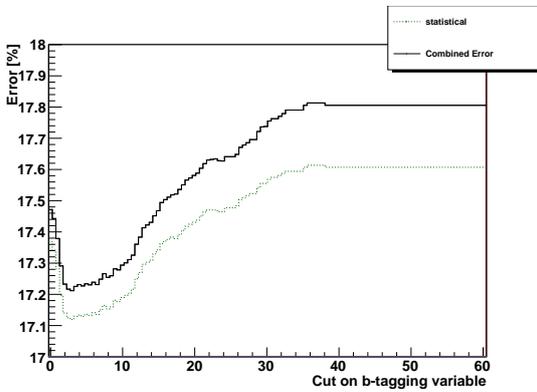
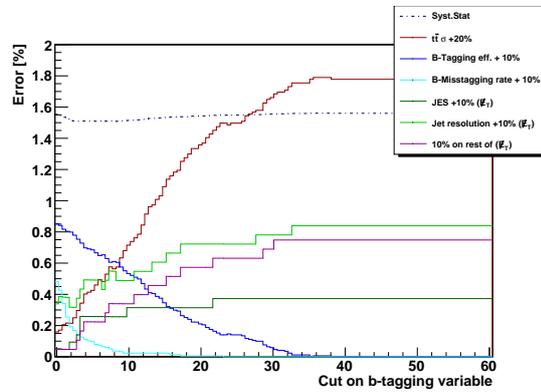
Table 5.1: Optimisation results for b -tagging with various luminosities after all other cuts have been applied.

NLO Corrections

As mentioned in section 4.5 there are effects from NLO corrections, that will affect the b -tagging. In fig. 5.13 the ratio Pythia to MC@NLO is shown, and only a small systematic shift is observable. This is most likely due to the higher jet multiplicity with MC@NLO and a slightly different jet composition and will result in small changes to the selection efficiencies. The uncertainty from that correction is estimated as half the difference in efficiency between the two different samples, which is 0.1%. This additional uncertainty is however smaller than the 0.6% uncertainty reduction gained by introducing this cut for 50pb^{-1} . With higher luminosities it will get more and more important to improve and tune the Monte Carlo generators to reproduce the basic physics observables like jet multiplicity.



(a)

(b) Overall and statistical Uncertainty for 50pb^{-1} .

(c) Impact of various systematic uncertainties

Figure 5.10: Optimisation plot for a cut on b -tagging at 50pb^{-1} (a-c). In (b) and (c) different parts are zoomed in for better visibility. The value of the applied cut is plotted on the x-axis, the resulting systematic error contributions and the combined overall uncertainty on the y-axis.

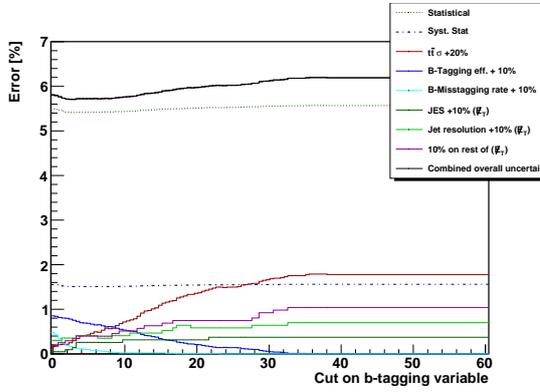


Figure 5.11: Optimisation plot for a cut on b -tagging with 500pb^{-1} .

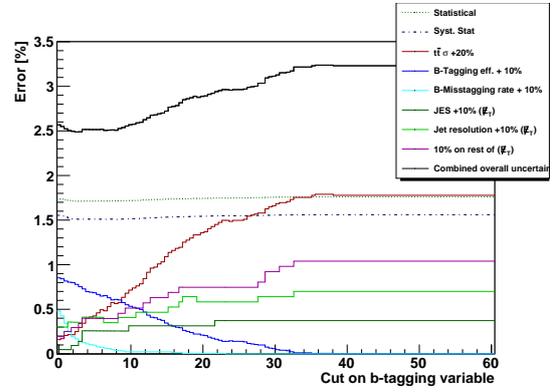


Figure 5.12: Optimisation plot for a cut on b -tagging with 5000pb^{-1} .

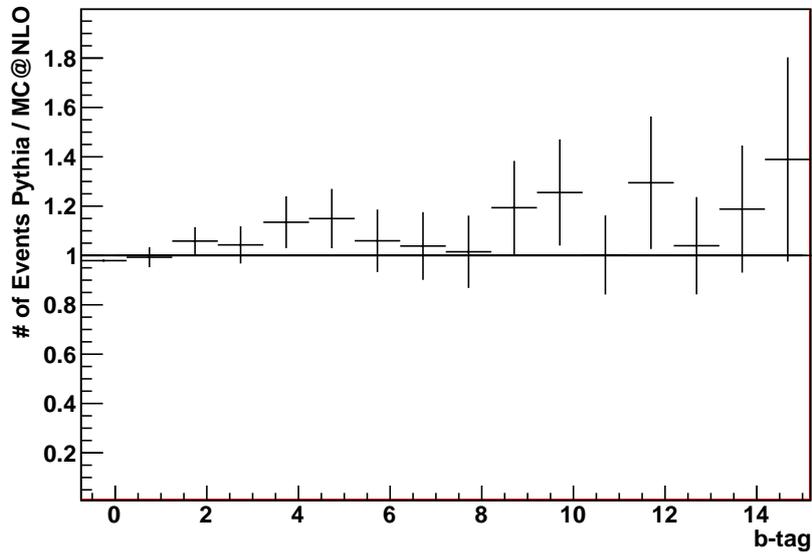


Figure 5.13: Number of events from *Pythia* divided by the number of events for *MC@NLO* for different cut values of the b -tagging cut.

Stability for high dimuon masses

As described in the beginning of this chapter, it is important for any cut to be stable even for very high dimuon masses. A cut on b -tagging will be stable up to 400 GeV, as can be seen in fig. 5.14. This is in principle not surprising, as events with higher dimuon masses are not expected to fake more b -jets than events with medium or low dimuon masses. Due to the limited statistics of the Pythia sample and the differences between the Pythia and the high mass Drell Yan sample described in appendix C.2 it is not possible to conclusively show the stability up to 1 TeV. However, tests with the high mass Drell Yan sample suggest, that the efficiency of a cut on b -tagging is not affected up to 1 TeV with Athena version 15.

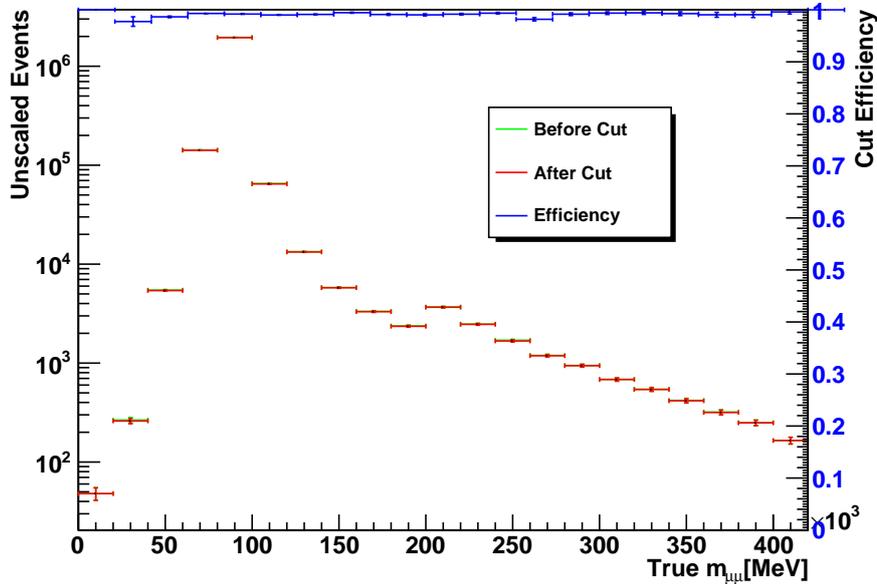


Figure 5.14: Cut efficiency for b -tagging cut up to 400 GeV (blue) and its effect on the number of events selected by that cut (red).

5.3.5 Missing Transverse Energy

As described in section 3.3.3 the \cancel{E}_T should reflect the vectorial sum of the p_T 's of all escaping particles. In the standard model the only particle which can not be detected inside ATLAS is the neutrino. A $Z \rightarrow \mu^+ \mu^-$ event is not expected to contain neutrinos, while the background processes $t\bar{t}$ and $b\bar{b}$ contain one neutrino per charged lepton, as described in section 4.1. Therefore a cut on \cancel{E}_T should provide a good separation between signal and background events. Unfortunately \cancel{E}_T is a very complicated variable with many sources of experimental errors, which are already described in section 3.3.3. Another problem is, that only the vectorial sum of the invisible particles momenta can be measured, therefore the missing energy of two or more neutrinos might cancel out.

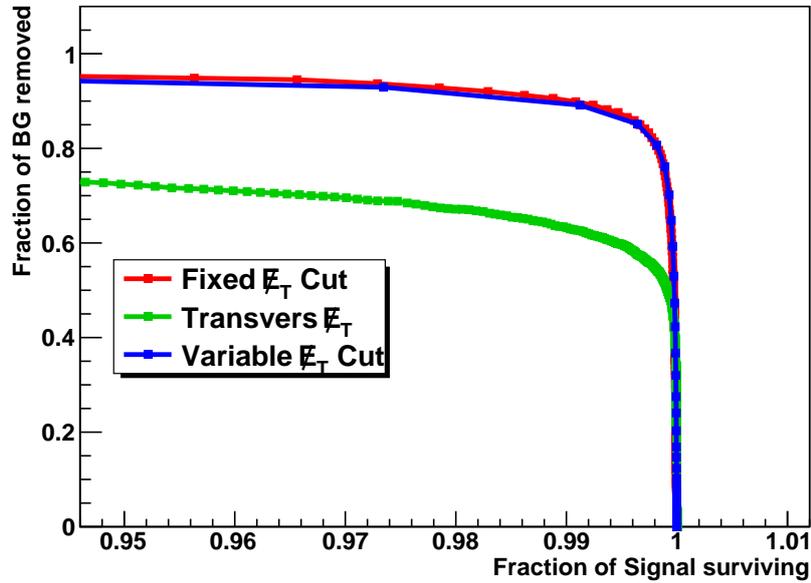


Figure 5.15: Comparison between three different \cancel{E}_T cuts. Only $Z \rightarrow \mu^+ \mu^-$ and $t\bar{t}$ are included since other backgrounds have only small impact and contribute low statistics for high dimuon masses. Preselection, Quality and Isolation cuts have been applied.

Cut on fixed Missing Transverse Energy

Nonetheless \cancel{E}_T is a variable worth looking at and it shows a promising efficiency in fig. 5.15. As a first iteration, a cut requiring $\cancel{E}_T < 50$ GeV will remove 65% of the $t\bar{t}$ background while only 0.06% of the signal are rejected. The drawback of a cut on a fixed value of \cancel{E}_T is revealed when plotting the efficiency of the cut against $m_{\mu\mu}$ as in fig. 5.19(a).

For low $m_{\mu\mu}$ the cut is well suited, but with rising $m_{\mu\mu}$ the efficiency drops significantly. The reason for this is the p_T dependency of the muon resolution shown in fig. 5.16. With higher $m_{\mu\mu}$ the p_T of the muon goes up as well, leading to a degraded muon resolution. A misreconstructed muon will lead to overestimated \cancel{E}_T , therefore wrongly rejecting events. This leads to two negative effects. First, the signal efficiency decreases for higher dimuon masses as seen in fig. 5.19(a), therefore increasing the statistical uncertainty in the high dimuon mass regions of the spectrum. Second, the muon momentum resolution cannot be measured with very high accuracy for high muon momenta, especially for the tails of the resolution distribution. A cut on a fixed \cancel{E}_T introduces a high dependency on the tails of the muon momentum resolution, creating a systematic uncertainty which is hard to control. Therefore the cut on \cancel{E}_T has to be refined.

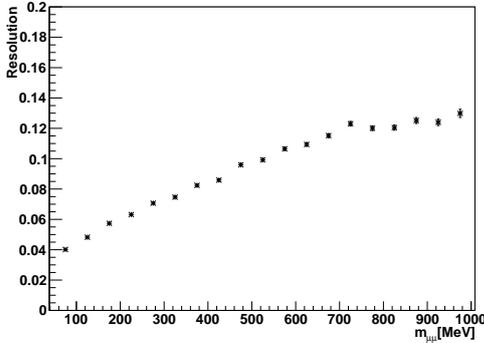


Figure 5.16: Muon Resolution vs. p_T of the muon for all muons in the high mass Drell Yan sample. The resolution is defined as the width of a gaussian fitted to the fractional muon momentum resolution (see appendix B.5 for further discussion). The error bars reflect the uncertainty on the fit.

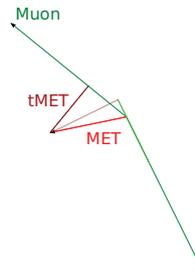


Figure 5.17: This shows the construction of $t\cancel{E}_T$ in the transversal plane. $t\cancel{E}_T$ is the \cancel{E}_T contribution perpendicular to a muon.

Cut on Transverse Missing Transverse Energy

As mentioned before, the main contribution to false Missing Transverse Energy stems from misreconstructed muons. This problem can be reduced by giving less weight to the \cancel{E}_T if it points in the same direction as one of the muons. If the \cancel{E}_T vector is perpendicular to the muons it gets a higher weight. This is achieved by the following definition of the transverse missing transverse energy $t\cancel{E}_T$:

$$t\cancel{E}_T = \cancel{E}_T \cdot \min(\sin(\angle_{\cancel{E}_T, \text{Muon}_i})) \quad (5.15)$$

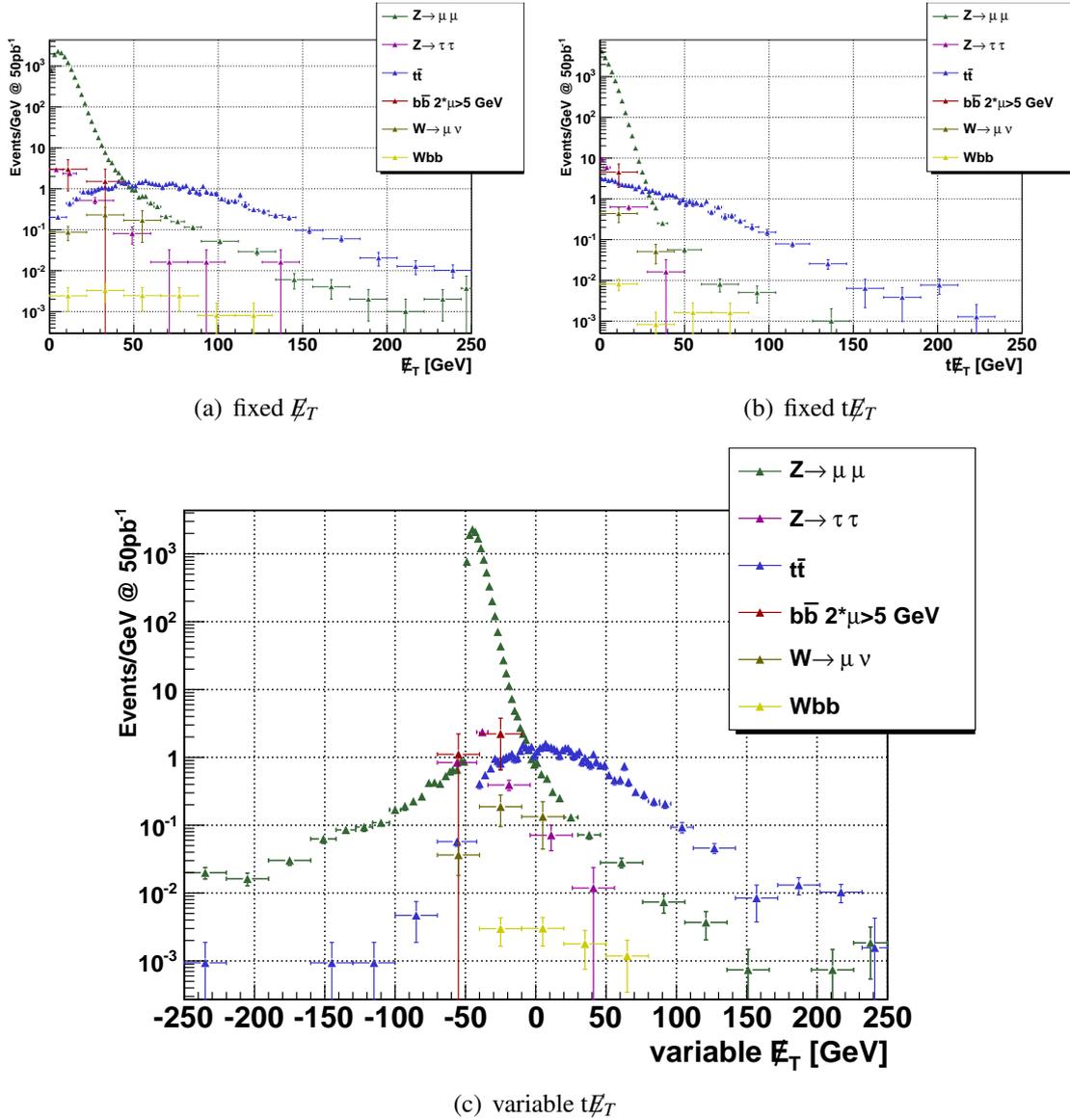


Figure 5.18: Distributions for various E_T cuts.

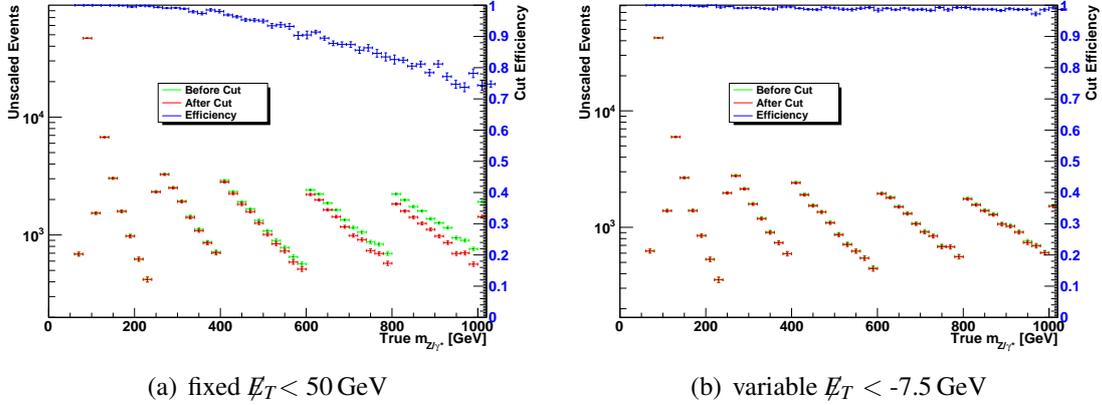


Figure 5.19: Cut efficiency for two \cancel{E}_T cuts up to 1 TeV (blue) and its effect on the number of events selected (red). As discussed in appendix C.2 the high mass Drell Yan sample can be used here.

where $Muon_i$ stands for every muon with p_T above 20 GeV. This definition is illustrated in fig. 5.17. Only \cancel{E}_T contributions perpendicular to a muon are taken into account, while \cancel{E}_T contributions parallel to a muon are discarded. As can be seen in fig. 5.18(b) the $t\cancel{E}_T$ distribution is narrower than the \cancel{E}_T distribution for the signal. Unfortunately the $t\bar{t}$ background is also shifted to lower values. The reason for that is, that in $t\bar{t}$ the muon and \cancel{E}_T can point in the same direction by chance and thus produce a very small $t\cancel{E}_T$. Looking at fig. 5.15 it becomes clear, that the $t\cancel{E}_T$ is not well suited to select signatures with low \cancel{E}_T . For signatures that have muons plus a high \cancel{E}_T this variable might be a good way to reduce the impact of muon momentum resolution.

Cut on variable MET

Since a fixed cut on \cancel{E}_T is not stable at high dimuon masses, and a cut on $t\cancel{E}_T$ is not very effective, a variable \cancel{E}_T cut has been developed. As described before, the muon resolution will reduce the signal efficiency of a cut on a fixed \cancel{E}_T . Therefore the cut is refined to allow a higher \cancel{E}_T for higher muon momenta. As can be seen in fig. 5.16 the relative muon resolution is increasing linearly with the transverse momentum, therefore the absolute resolution is proportional to p_T^2 . To also take \cancel{E}_T contributions from other sources into account at least 50 GeV are subtracted from the \cancel{E}_T . The cutvalue will be optimised as described in section 5.2

$$\text{variable } \cancel{E}_T = \cancel{E}_T - \max(50 \text{ GeV}, \sqrt{p_{T1}^4 + p_{T2}^4}) / 500 \text{ GeV} \quad (5.16)$$

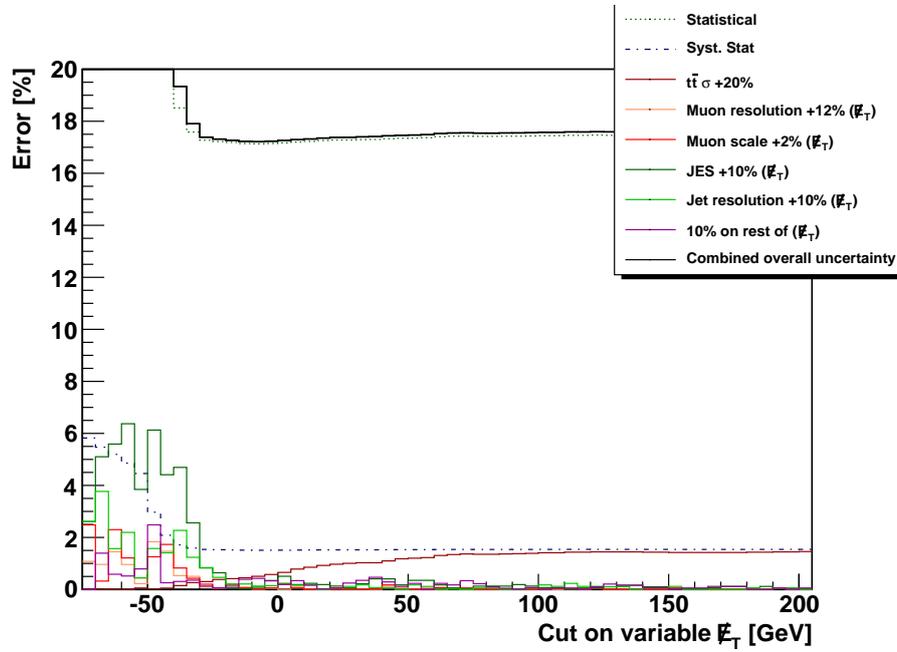
Other formulas have been tested as well, and eq. 5.16 turned out to be the most efficient one. The distribution of variable \cancel{E}_T differs only slightly from the E_T distribution, as can be seen in fig. 5.18. This is expected, since only for a minority of events the muons have a high enough momentum to subtract more than 50 GeV in eq. 5.16. According to fig. 5.19(b) a cut on variable \cancel{E}_T has a stable efficiency of above 98% for dimuon masses up to 1 TeV.

Optimisation Results

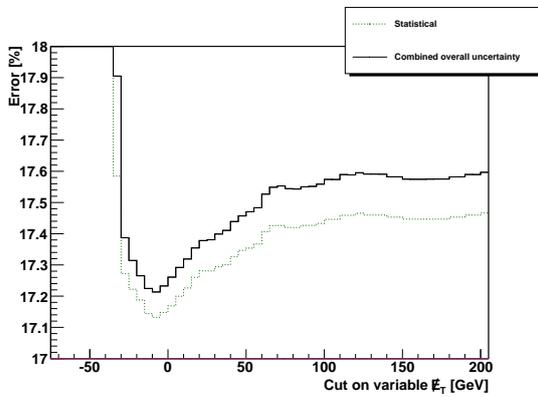
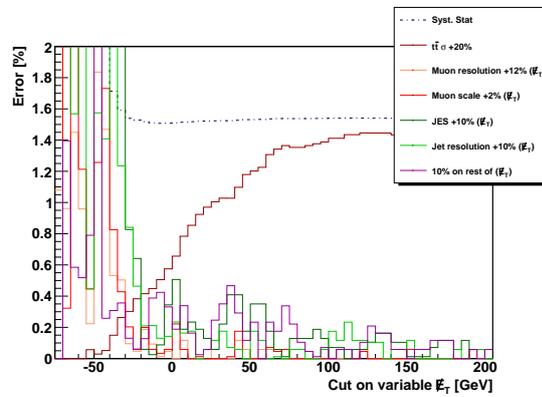
In fig. 5.20(a) the optimisation of a cut on variable \cancel{E}_T is shown for 50pb^{-1} . As before the most important line is the overall systematic uncertainty, which has a clearly visible minimum at -7.5 in fig 5.20(b). The different contributions are in principle the same as for the cut on b -tagging: Statistical uncertainties and the uncertainty on the background cross section favour a stronger cut, while the systematic uncertainties on \cancel{E}_T favour a looser cut. It can be seen, that \cancel{E}_T is affected by various systematic uncertainties, the most important ones being the uncertainty on the unclustered energy, the jet resolution and jet energy scale.

For cutvalues below -30GeV only a very small fraction of the signal is selected. This leads to very high statistical uncertainties in that area. Due to the very low number of events statistical fluctuations can create high peaks in the calculated systematic errors in those regions. Since a cutvalue below -30GeV rejects 97% of the signal, this region is not interesting for cut optimisation anyhow and is only shown for completeness. The spikes above -30GeV in fig. 5.20(c) can be understood by looking at a hypothetical event with 100GeV \cancel{E}_T . A systematic shift of e.g. the JES could bring the \cancel{E}_T of that event to 103GeV . Therefore, if the cut on \cancel{E}_T is below 100GeV or above 103GeV , the systematic shift would not have any effect, only in the range from 100GeV to 103GeV the cut would indeed increase the total systematic uncertainty. Those effects are artefacts of a limited Monte Carlo statistics, and could be reduced by higher statistics or broader binning. Since the spikes are low compared to the overall uncertainty, they do not have any significant influence on the result of the optimisation and can therefore be ignored.

For 50pb^{-1} the cut on variable \cancel{E}_T is set to -7.5GeV , which results in 71% of the remaining $t\bar{t}$ background being rejected by that cut, while only 0.08% of the signal are removed. The overall uncertainty is reduced from 17.6% to 17.2%. The optimisation plots for other luminosities are shown in fig. 5.21 and fig. 5.22. The resulting cutvalues and reductions in overall uncertainty are shown in table 5.2. It has to be noted that for higher integrated luminosities the results “with cut” are slightly different than the results for b -tagging in table 5.1. The reason for that is the fact, that the uncertainty on b -tagging is not calculated on an event by event basis and thus not included in the uncertainty for the variable \cancel{E}_T cut. The overall uncertainty given in table 5.1 contains this uncertainty and is therefore the more precise result.



(a)

(b) Overall and statistical Uncertainty for 50 pb^{-1} . The minimum at -7.5 is clearly visible

(c) Impact of various systematic uncertainties

Figure 5.20: Optimisation plot for a cut on variable E_T at 50 pb^{-1} (a-c). In (b) and (c) different parts are zoomed in for better visibility. The value of the applied cut is plotted on the x-axis, the resulting systematic error contributions and the combined overall uncertainty on the y-axis.

	50pb^{-1}	500pb^{-1}	5fb^{-1}
Uncertainty w/o cut	17.6%	5.8%	2.5%
Uncertainty with cut	17.2%	5.6%	2.3%
Cutvalue	-7.5	-7.5	-7.5

Table 5.2: Optimisation results for a cut on variable \cancel{E}_T with various integrated luminosities after all other cuts have been applied.

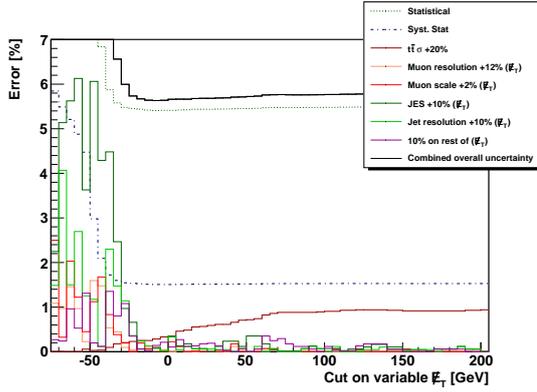


Figure 5.21: Optimisation plot for a cut on variable \cancel{E}_T at 500pb^{-1} .

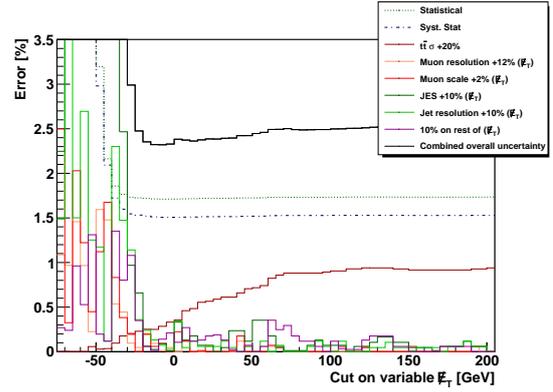


Figure 5.22: Optimisation plot for a cut on variable \cancel{E}_T at 5000pb^{-1} .

NLO Corrections

The \cancel{E}_T measurement is affected by jets within an event. The number of jets in $Z \rightarrow \mu^+\mu^-$ changes when next-to-leading order corrections are taken into account, thus a cut on \cancel{E}_T might have some dependency on NLO corrections. In fig. 5.23 the ratio of predictions from Pythia to those of MC@NLO are shown. No significant differences are observable in the region around the intended cut. A cutvalue of -7.5 GeV removes $(0.086 \pm 0.003)\%$ of the Pythia Events and $(0.079 \pm 0.01)\%$ of the MC@NLO events, so the efficiency of the selection cut does not differ significantly. Even if the small difference of 0.007% were significant, the impact would be negligible due to the very small number of events removed by a cut on variable \cancel{E}_T .

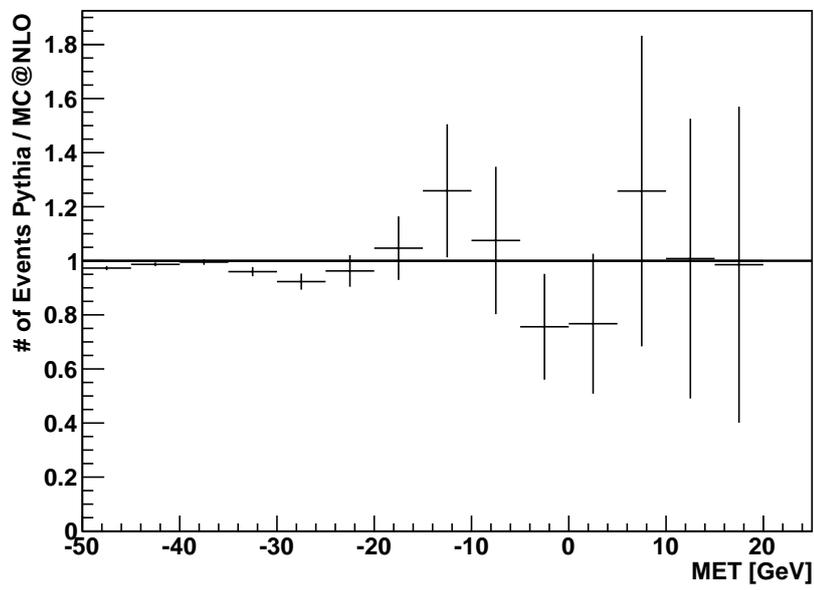


Figure 5.23: Number of events from Pythia divided by the number of events for MC@NLO for different cut values of the \cancel{E}_T cut.

5.4 Variables Unused in this Analysis

Various other selection variables have been studied as well. Three of them are presented in the following section. For the optimisation plots all other cuts have been applied already.

5.4.1 Sphericity

The sphericity S of an event describes how isotropic the shape of an event is. A perfectly isotropic event would have a sphericity of 1, while a completely linear event would have a sphericity of 0. Sphericity has been studied in a number of analyses at the Tevatron [59] and ATLAS[42] e.g. for $t\bar{t}$ selection. More recent analyses [28] use the transverse sphericity S_T which uses a projection on the plane perpendicular to the beam axis. For a mathematical definition of S and S_T see appendix B.3 and B.4. Fig. 5.24 shows a comparison between S and S_T and it turns out, that S_T is the more efficient variable for this analysis.

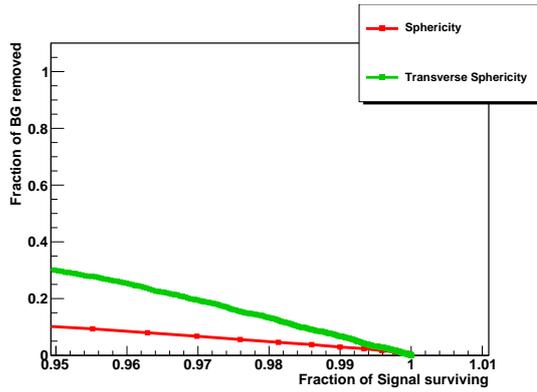


Figure 5.24: Comparison between cuts on S and S_T . Only the signal and $t\bar{t}$ sample are included and preselection, quality and isolation cuts have been applied.

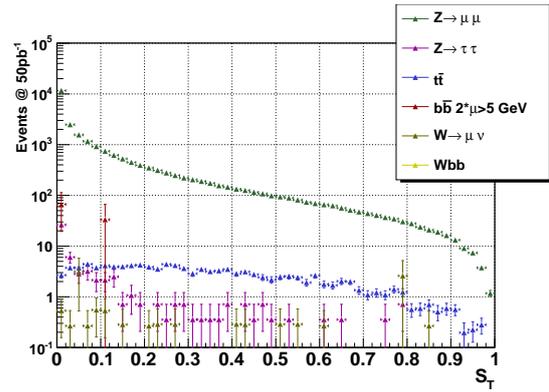


Figure 5.25: Distribution of transverse sphericity after isolation cuts have been applied.

$t\bar{t}$ events are expected to contain various decay products, and therefore the transverse sphericity should be higher than in $Z \rightarrow \mu^+\mu^-$ events. In fig. 5.25 the distribution of the transverse sphericity is shown. The signal has a high peak at 0, while the distribution of $t\bar{t}$ events is much broader with no clear peak.

Several additional systematic uncertainties that might be relevant for calculating the transverse sphericity of an event are not taken into account at the moment. Examples for those are the angular resolution of jets and muons, underlying event or next-to-leading order

effects. Even without those systematic uncertainties, the optimisation plots show no improvement in the overall uncertainty for a cut on transverse sphericity for 50pb^{-1} (see fig. 5.26(a)), 500pb^{-1} (see fig. 5.26(b)) or 5fb^{-1} (see fig. 5.26(c)). It is therefore safe to assume, that no cut on transverse sphericity should be applied.

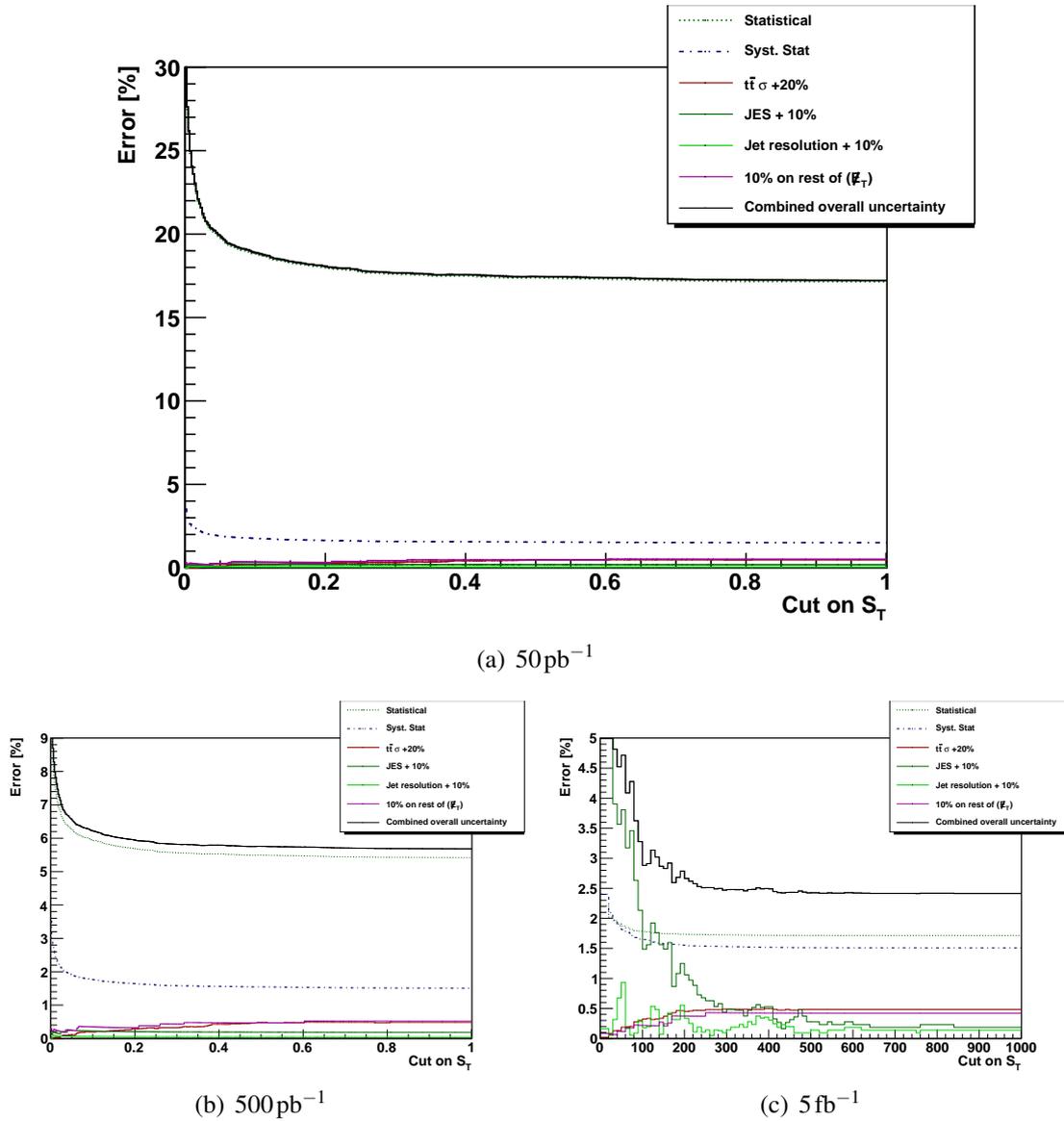


Figure 5.26: Optimisation plot for a cut S_T for different integrated luminosities. The value of the applied cut is plotted on the x-axis, the resulting systematic error contributions and the combined overall uncertainty on the y-axis.

5.4.2 Jet Energy

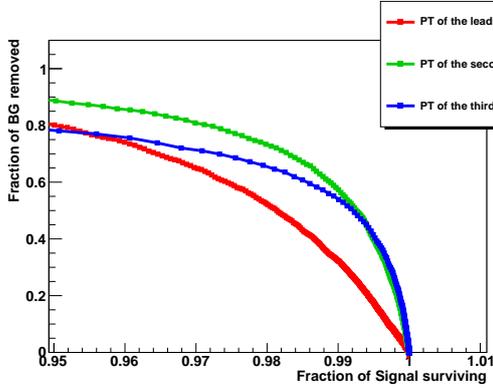


Figure 5.27: Comparison for various jet cuts. Only the signal and $t\bar{t}$ sample are included and preselection, quality and isolation cuts have been applied.

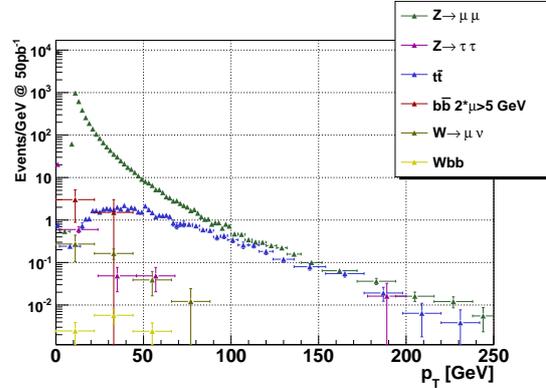


Figure 5.28: Distribution of the p_T of the second leading jet after isolation cuts have been applied.

As mentioned in section 4.1 the $t\bar{t}$ background is expected to have at least two jets per event, while $Z \rightarrow \mu^+\mu^-$ is not expected to contain jets at leading order. The NLO contributions described in chapter 2 and the underlying event are responsible for jets in the signal sample. In order to use a cut on jets, it is important to have a precise understanding of the properties of additional jets from NLO and underlying event. The current predictions of Monte Carlo generators will not be accurate enough to accomplish this, but it is in principle possible to measure the jet rates in $Z \rightarrow \mu^+\mu^-$ from data. Thus a cut on the p_T of the jet with the first, second and third highest p_T are studied here.

Fig. 5.27 shows that a cut on the jet with the second or third highest p_T are the most effective ones. Since dileptonic $t\bar{t}$ events are expected to contain two jets, a cut on the second leading jet has a better motivation and is thus studied here. A cut on the third leading jet has been studied in detail as well, but turned out to be less effective. In fig. 5.28 the p_T distribution of the second jet is shown for various samples. The $t\bar{t}$ background has a broader distribution with tails to much higher jet p_T 's. An optimisation as described in section 5.2 is shown in fig. 5.29. Neither for an integrated luminosity of 50pb^{-1} (fig. 5.29(a)) nor for luminosities up to 5fb^{-1} (fig. 5.29(b) and fig. 5.29(c)) a minimum is visible. Unlike most other cuts in this case the overall uncertainty is dominated by the systematic uncertainty on the jet energy scale. If at some point in the future the uncertainty on the jet energy scale can be reduced from the expected 10% to 3%, the overall uncertainty for 5fb^{-1} can be reduced from 2.38% to 2.37% by introducing a cut of 60 GeV on the second leading jet p_T . It has to be noted that no systematic uncertainty on jet production for the signal process has been assumed. Those uncertainties would make a cut on jet p_T less desirable. In conclusion a cut on jet p_T will not add any benefit

within the expected uncertainties. Even if the uncertainties on the jet energy scale will be lower than expected, the reduction in overall uncertainty of 0.01% is too low to justify the introduction of new systematic uncertainties.

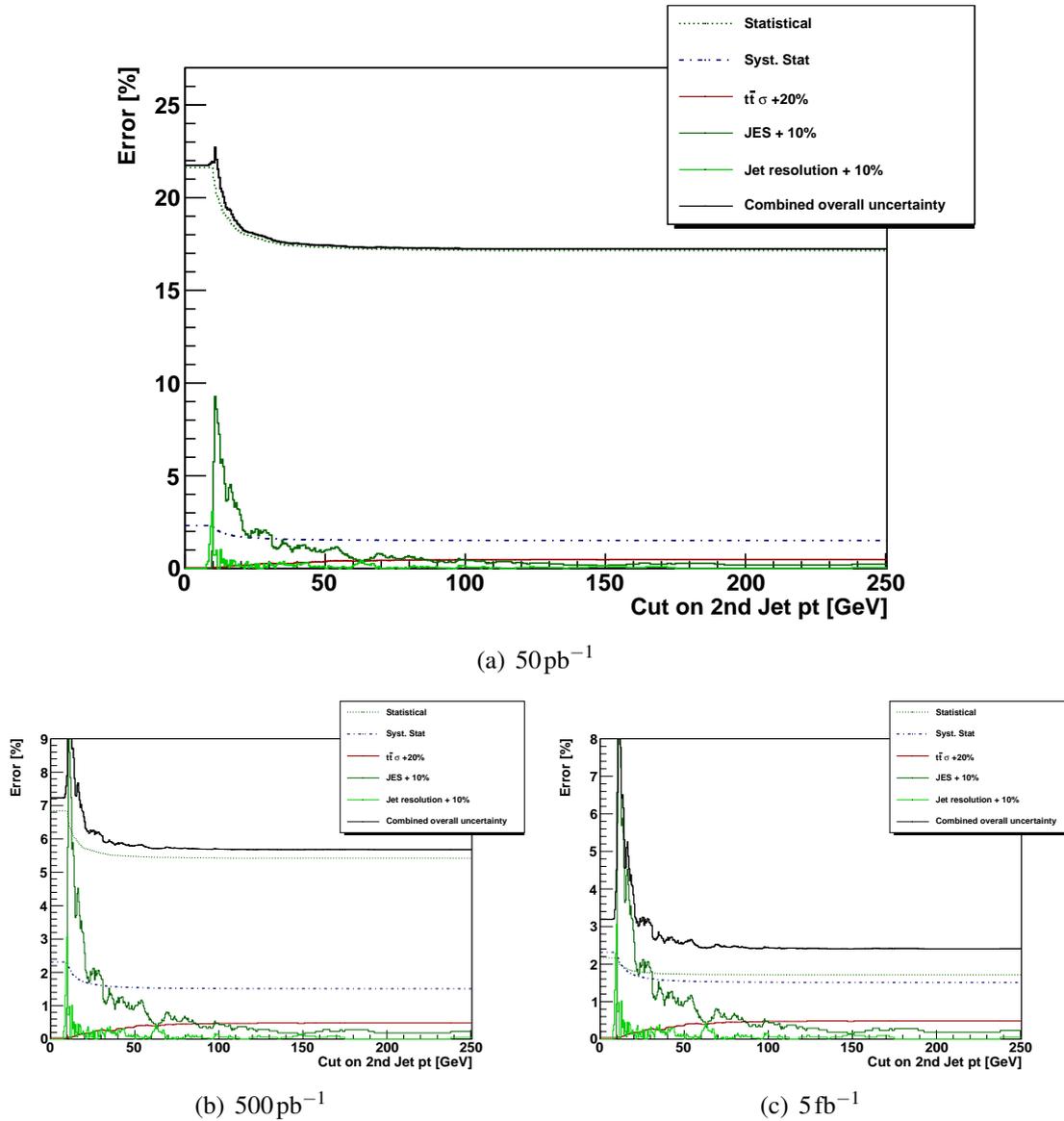


Figure 5.29: Optimisation plot for a cut on the p_T of the second leading jet for different integrated luminosities. The value of the applied cut is plotted on the x-axis, the resulting systematic error contributions and the combined overall uncertainty on the y-axis.

5.4.3 HT

As an alternative to cutting on the p_T of one jet, it is also possible to cut on the scalar sum of all jets above a certain threshold. Various values for that threshold have been tested and 20 GeV turned out to be the best choice:

$$HT = \sum_{jets, p_T > 20 \text{ GeV}} p_{T,jet} \quad (5.17)$$

This variable takes advantage of the fact that $t\bar{t}$ events have more *and* higher energetic jets than $Z \rightarrow \mu^+\mu^-$ events. In fig. 5.30 the distribution of HT is shown for various samples. Like a cut on the second leading jet p_T a cut on HT is highly dependent on the jet energy scale and the correct understanding of jet production in the signal channel. Therefore jet production has to be studied with real data and a Monte Carlo generator that describes the jet multiplicity well for $Z \rightarrow \mu^+\mu^-$ has to be used. However, the optimisation plot in fig. 5.31 for HT does not have a minimum, mostly due to the uncertainties on the jet energy scale and the impact on the statistical uncertainty. Even if the JES uncertainty would at some point get lower than anticipated now, the stability to high dimuon masses shown in fig. 5.32 would prevent the usage of a cut on HT . Some approaches similar to the variable \cancel{E}_T have been tested, but none of them solved the problem. In conclusion, there are various reasons against a cut on HT .

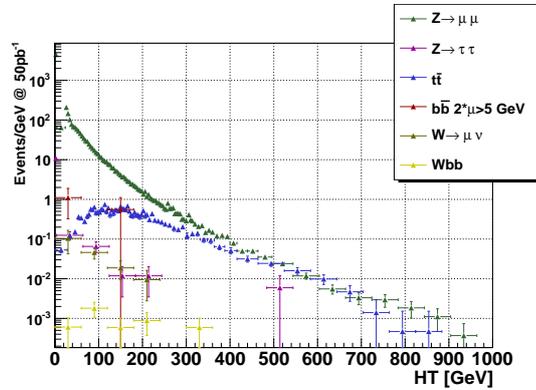


Figure 5.30: Distribution of HT for various samples after preselection and isolation cuts have been applied.

5.5 Significance of Cut Optimisation

To discuss the significance of the minima found in the optimisation plots three contributions are important:

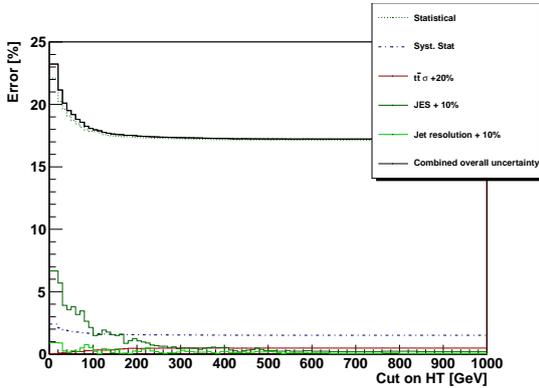


Figure 5.31: Optimisation plot for a cut HT for different integrated luminosities. The value of the applied cut is plotted on the x-axis, the resulting systematic error contributions and the combined overall uncertainty on the y-axis.

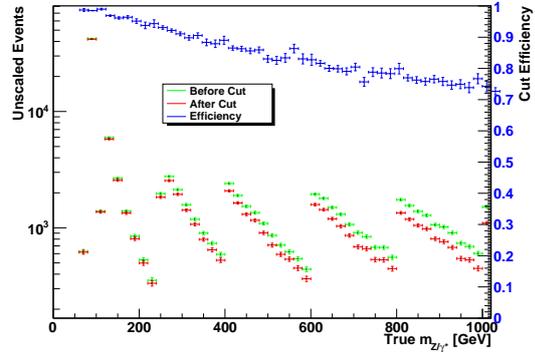


Figure 5.32: Cut efficiency for a hypothetical cut on $HT < 100$ GeV up to 1 TeV (blue) and its effect on the number of events selected by that cut (red).

The statistical uncertainty can be determined with a high accuracy due to the large integrated luminosity of the signal and the $t\bar{t}$ background. The uncertainty on the depth of the minimum of the statistical contribution can therefore be neglected.

The contribution of the individual systematic uncertainties is more complicated: Their sizes are supplied by the according working groups, but none of them gives an uncertainty on the uncertainty. Thus a worst case scenario is assumed, where all systematic uncertainties that favour looser cuts are multiplied by a factor of 2. For an integrated luminosity of 50pb^{-1} no change in the position of the minimum is observed. For higher integrated luminosities the size of the systematic uncertainties is going to be determined from data and will hopefully also include uncertainties on the uncertainties.

Some of the systematic uncertainties are applied to the events by using random gaussian numbers, thus the exact shape of the impact of those uncertainties might vary when the optimisation is performed again. Therefore the optimisation has been performed seven times with the same settings for each cut. The position of the minimum did not change, and the change of the depth of the minimum was negligible.

It is therefore safe to assume that the minima found in the optimisation plots are indeed significant for an integrated luminosity of 50pb^{-1} .

5.6 Results

In fig. 5.33 the reconstructed invariant dimuon mass is shown for different sets of cuts. As expected isolation cuts can significantly reduce the background contribution from $b\bar{b}$, but

$t\bar{t}$ will remain an important background even at high dimuon masses. The two additional cuts on b -tagging and variable \cancel{E}_T can reduce this contribution by more than a factor of 10.

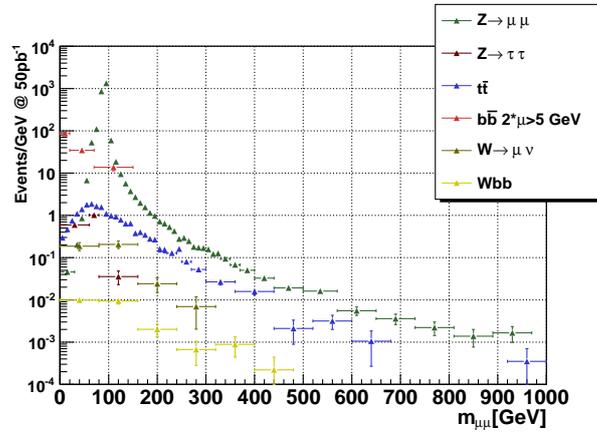
In table 5.4 the cut-flow normalised for an integrated luminosity of 50pb^{-1} is shown. The low statistics of the $b\bar{b}$ sample results in a large statistical uncertainty from Monte Carlo for that process. However, fig. 5.33(a) shows that dimuon events from $b\bar{b}$ have a low dimuon mass in most cases. Together with the large suppression factor of nearly 70 from selection cuts this will ensure that the background contribution from $b\bar{b}$ will be well below the one from $t\bar{t}$.

As discussed before the two additional cuts on variable \cancel{E}_T and b -tagging will not only improve the signal to background ratio but also reduce the overall uncertainty of the measured spectrum. For an integrated luminosity of 50pb^{-1} the overall uncertainty for a bin from 200 GeV to 300 GeV will be reduced from 19.1% to 17.2% by the two additional cuts. The impact of the two cuts for various integrated luminosities is shown in table 5.3. It has to be noted that only uncertainties that affect or are affected by the selection cuts are included in this calculation. Once theoretical uncertainties and uncertainties concerning the dimuon mass resolution are included the overall uncertainty will get higher. The existence and position of the minimum in the optimisation plot will not be affected.

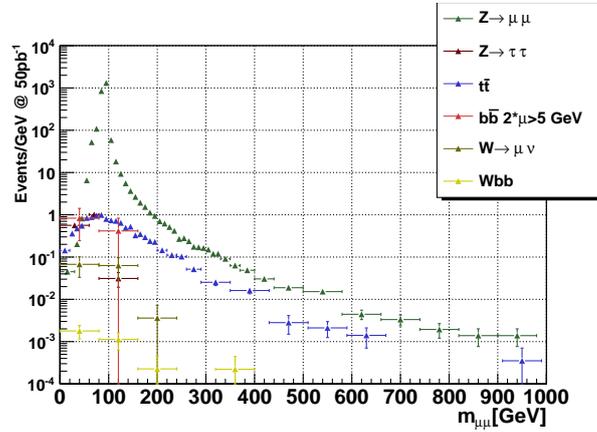
Due to the increased impact of the systematic uncertainties at higher integrated luminosities the results for 500pb^{-1} and 5fb^{-1} can only be seen as an indication, that those additional cuts will still be able to improve the precision. The actual numbers should be calculated again once the systematic uncertainties for those integrated luminosities are known.

	50pb^{-1}	500pb^{-1}	5fb^{-1}
Uncertainty with no additional cuts	19.1%	7.9%	5.6%
Uncertainty with only b -tagging cut	17.6%	5.8%	2.5%
Uncertainty with only \cancel{E}_T cut	17.8%	6.2%	3.2%
Uncertainty with both cuts	17.2%	5.7%	2.5%

Table 5.3: Optimisation results for b -tagging with various luminosities after all other cuts have been applied.



(a) After preselection and quality cuts



(b) After isolation cuts

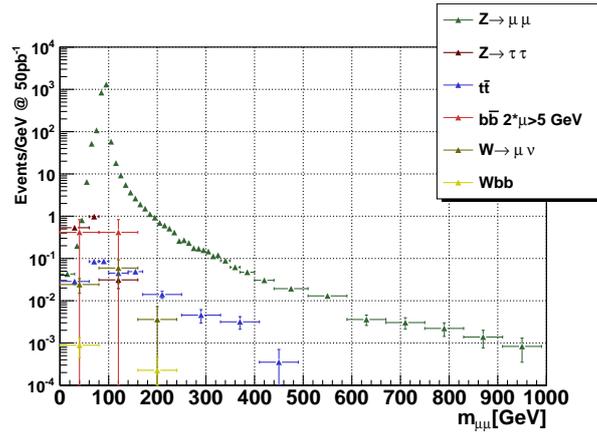
(c) After \cancel{E}_T and b -tagging cuts

Figure 5.33: The reconstructed invariant dimuon mass after different cuts. (a) has similar cuts as applied in Z' searches [2] and the cuts in (b) are similar to the ones used in [24] to measure the $Z \rightarrow \mu^+ \mu^-$ cross section. In (c) the result after applying the additional cuts suggested in this analysis is shown.

Sample Cut	$Z \rightarrow \mu^+\mu^-$	$Z \rightarrow \tau^+\tau^-$	$t\bar{t}$	$b\bar{b} \rightarrow \mu_5\mu_5$	$W \rightarrow \mu\nu$	$Wb\bar{b}$
Generated	65114.57	66881.38	10100	$30 \cdot 10^6$	807475.30	803.74
More than 2 Muons	27201.67	69.39	286.21	5784.16	78.82	4.373
Exactly 2 Muons	27195.49	69.03	277.49	5750.92	78.26	4.301
Opposite Charge	27192.87	68.68	210.65	5052.83	56.57	2.060
$\chi^2_{\text{match}}/N_{\text{DOF}} < 5$	24707.05	58.71	189.81	4554.20	33.81	1.845
$\text{maxJet}\Delta R < 0.7$	24699.48	58.71	179.18	964.03	33.26	1.703
$\text{maxNCone} < 3$	24380.05	56.58	118.98	99.73	10.70	0.269
$b\text{-tag} < 6.5$	24181.70	56.58	36.16	66.48	10.70	0.179
Variable $\cancel{E}_T < -7.5 \text{ GeV}$	24162.12	54.09	10.15	66.48	6.97	0.089
Statistical uncertainty from MC	± 23	± 4.2	± 0.53	± 47	± 1.7	± 0.040
Statistical uncertainty for 50 pb^{-1}	± 155	± 7.22	± 3.19	± 8.2	± 2.6	< 1

Table 5.4: Cut-flow scaled to an integrated luminosity of 50 pb^{-1} . The statistical uncertainty from Monte Carlo and the statistical uncertainty which is expected for 50 pb^{-1} are given as well. See the text for a discussion of the large uncertainty on $b\bar{b}$.

6 Conclusions and Outlook

The Large Hadron Collider started running at a centre of mass energy of 7 TeV in 2010, allowing the beginning of a new era in particle physics. The first scientific results have been published [60, 61, 62] and thousands of physicists are looking forward to getting sufficient integrated luminosity to perform the analyses they have prepared in the last years.

One of these analyses will be the precise measurement of the differential cross section of the high mass Drell Yan process. By comparing the theoretical predictions to the data a model independent search for new physics is possible.

In this thesis the background for the $Z \rightarrow \mu^+\mu^-$ channel is studied in detail using Monte Carlo data. The well established selection cuts for $Z \rightarrow \mu^+\mu^-$ at the Z-peak are adapted for high dimuon masses and two additional selection cuts are proposed, thus allowing the reduction of the background for dimuon masses above 200 GeV by a factor of 10. To reach the highest possible precision, a sophisticated method of optimising those cuts has been implemented: The systematic uncertainties and the statistical uncertainty are combined and the dependency on the selection cuts is determined. By scanning all possible cut values the overall uncertainty can be minimised for each selection cut. Thus a variety of different cuts have been studied and the following two proved to be able to reduce the overall uncertainty.

The b -tagging weight of a jet is basically an estimation of how likely it is that the jet contained a b quark, whilst the missing transverse energy (\cancel{E}_T) is an indication of neutrinos within the event. The signal process $Z \rightarrow \mu^+\mu^-$ contains neither b quarks nor neutrinos, whereas the most important background contribution $t\bar{t}$ typically contains two b -quarks and two neutrinos. By applying selection cuts on those two variables the overall uncertainty in the bin from 200 GeV to 300 GeV can be reduced from 19.1% to 17.2% for an integrated luminosity of 50pb^{-1} . The optimisation has been performed for integrated luminosities of 500pb^{-1} and 5fb^{-1} as well and the additional cuts will still be useful at higher integrated luminosities.

This result is contrary to the previous assumptions that background contributions from $t\bar{t}$ can be neglected and that the uncertainties for b -tagging and \cancel{E}_T will be too high to be useful in the early data taking phase. The increased systematic uncertainties are more than

compensated for by the reduction of the statistical uncertainty due to improved signal to background ratio. The suitability of all cuts to the high dimuon mass region was tested thoroughly, and the impact of next-to-leading order corrections has been estimated to be acceptable.

Once the first 50pb^{-1} have been recorded roughly 57 events with a dimuon masses above 200 GeV are expected. If only the established isolation cuts of [2] are applied, 11.3 background events are expected. With the additional cuts on b -tagging and \cancel{E}_T , this can be reduced to 1.1 expected background events where these two additional cuts do not introduce $m_{\mu\mu}$ dependent biases in the selection efficiency.

For a reliable measurement additional studies with real data are necessary. The systematic uncertainties have to be quantified and the reconstruction and isolation efficiencies must be determined from data. This will be possible with the first $50\text{-}100\text{pb}^{-1}$ paving the way for a model independent search for new physics by a precise measurement of the high mass Drell Yan spectrum.

A Corrections

In order to measure the differential cross section of $Z \rightarrow \mu^+\mu^-$ the data have to be corrected for two detector effects: Efficiency and resolution. An efficiency below 1 will arise, if not all of the $Z \rightarrow \mu^+\mu^-$ candidates are reconstructed. This can happen due to limited geometrical acceptance, detector inefficiencies or because the event is removed by some of the selection cuts. The limited resolution of the detector will cause some *bin-to-bin migration*, where an event is reconstructed with a wrong dimuon mass. To be able to compare the measured spectrum with the predicted differential cross section, one has to correct for both effects.

A.1 Bin-by-Bin Correction

A straight forward approach is the so called *bin-by-bin correction*, in which Monte Carlo data is used to determine a correction factor cf_i for each bin in the differential cross section distribution:

$$cf_i = \frac{N_i^{Truth}}{N_i^{Reco}} \quad (\text{A.1})$$

where i is the number of the bin, N_i^{Reco} is the number of events reconstructed to that bin and N_i^{Truth} is the number of events with a true dimuon mass in that bin. By construction this method also takes effects from bin-to-bin migration into account, as long as the migration is properly described in the detector simulation. There are two prerequisites for this method to work properly:

- **Correct detector simulation** To calculate the correction factors cf the detector simulation must be able to describe the efficiency and resolution of the detector. This can be tested with data driven methods like *Tag-and-Probe* [24].
- **Correct Monte Carlo** In order to properly describe the bin-to-bin migration the overall shape of the measured distribution has to be correct in the Monte Carlo. This is illustrated in fig. A.1

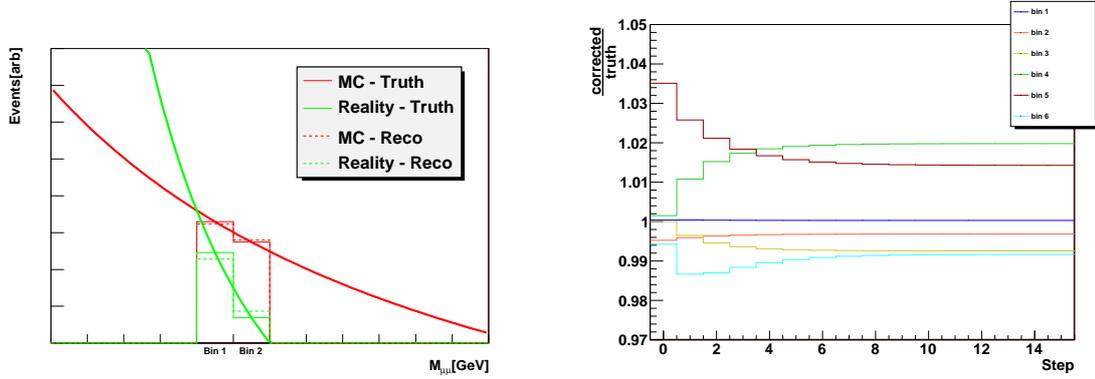


Figure A.1: If the Monte Carlo distribution differs from reality bin-to-bin migration will create errors in the bin-by-bin correction. In this example 10% of the events migrate from bin 1 to bin 2 and vice versa. For a low slope the migrations cancel out, while for a higher slope the net migration becomes larger.

Figure A.2: The iterative bin-by-bin method converges after a few steps. Step 0 equals the non-iterative method.

A.2 Iterative Bin-by-Bin Correction

To reduce the effect of mispredicted bin-to-bin migration an iterative bin-by-bin correction is used in some other analyses [63, 24], and has been implemented and tested for this analysis as well. The idea behind the iterative bin-by-bin correction is to repeatedly shift the reconstructed spectrum in such a way, that it will finally match the data. The correction factor cf_i^n for bin i in step n is obtained as:

$$cf_i^n = \frac{Data}{Reco_i^n}. \quad (\text{A.2})$$

It is then assumed, that if the same series of shifts is applied to the Monte Carlo truth it will return the true spectrum once the iteration converges:

$$Truth_i^{n+1} = cf_i \cdot Truth_i^n. \quad (\text{A.3})$$

In order to take bin-to-bin migration into account the reweighing of each event is applied according to its true dimuon mass, instead of the reconstructed dimuon mass:

$$Reco_i^{n+1} = \sum_j cf_j \cdot Reco_{i,j}^n, \quad (\text{A.4})$$

where $Reco_{i,j}$ is the number of events with a reconstructed dimuon mass in bin i , and a true dimuon mass in bin j . Once all cf_i^n are close to 1 the iteration has converged, and $Truth^n$ is the corrected distribution.

In order to test the implemented method various tests have been performed. Each test consisted of 50 runs with independent Monte Carlo and data samples. Due to the large CPU requirements of ATLAS full simulation this is not possible with fully simulated events, thus a so called *Toy Monte Carlo* has been produced. Each run was divided into two subsamples, one called Monte Carlo sample, the other one data sample. Both subsamples have a true $m_{\mu\mu}$ distribution generated from a parameterised Drell Yan Spectrum. For some samples a different parameterisation was used for data and Monte Carlo, in order to study the effect a wrongly predicted $m_{\mu\mu}$ distribution will have on the bin-by-bin correction. The shape of the Drell Yan spectrum was parameterised as $a \cdot e^{-bM^c}$, where M is the invariant dimuon mass in GeV [2]. Fits to fully simulated events between 200 GeV and 500 GeV resulted in a shape of

$$\frac{d\sigma}{dm_{\mu\mu}} = e^{-2.7 \cdot x^{0.3}}. \quad (\text{A.5})$$

In both subsamples a reconstructed distribution was created. To simulate the impact of the limited detector resolution a relative smearing of 5% was applied to the true dimuon mass. The Monte Carlo samples contained five million events each. The data samples contained 10,000 events each, which corresponds to an integrated luminosity of roughly 5 fb^{-1} .

Now the reconstructed and true $m_{\mu\mu}$ distribution of the Monte Carlo subsample and the reconstructed $m_{\mu\mu}$ distribution of the data sample have been used to obtain the true $m_{\mu\mu}$ distribution in the data sample. This result was then compared with the true $m_{\mu\mu}$ distribution already stored in the data sample to get an estimation of the error. This was done for the bin-by-bin correction and for the iterative bin-by-bin correction to compare the two methods.

In the first test, a bin width of 50 GeV was chosen, and the true $m_{\mu\mu}$ distributions for data and Monte Carlo are identical. In fig. A.2 the convergence of the iterative algorithm is shown. Fig. A.3 shows the comparison of the relative errors for both methods. On average, the error of the iterative method is larger than the one from the non-iterative method. A possible explanation for this is the low statistics for the highest bin in the data-sample. In the iterative method those errors can migrate to other bins, and thus degrade the precision there as well.

In a second test, a different shape for the true $m_{\mu\mu}$ in the data sample has been used. The parameter b was changed from 2.7 to 2.16. In fig. A.4 both methods are compared again. The changed slope introduces a systematic shift to higher values for the non-iterative method in fig. A.4(a). The iterative method can not completely compensate that

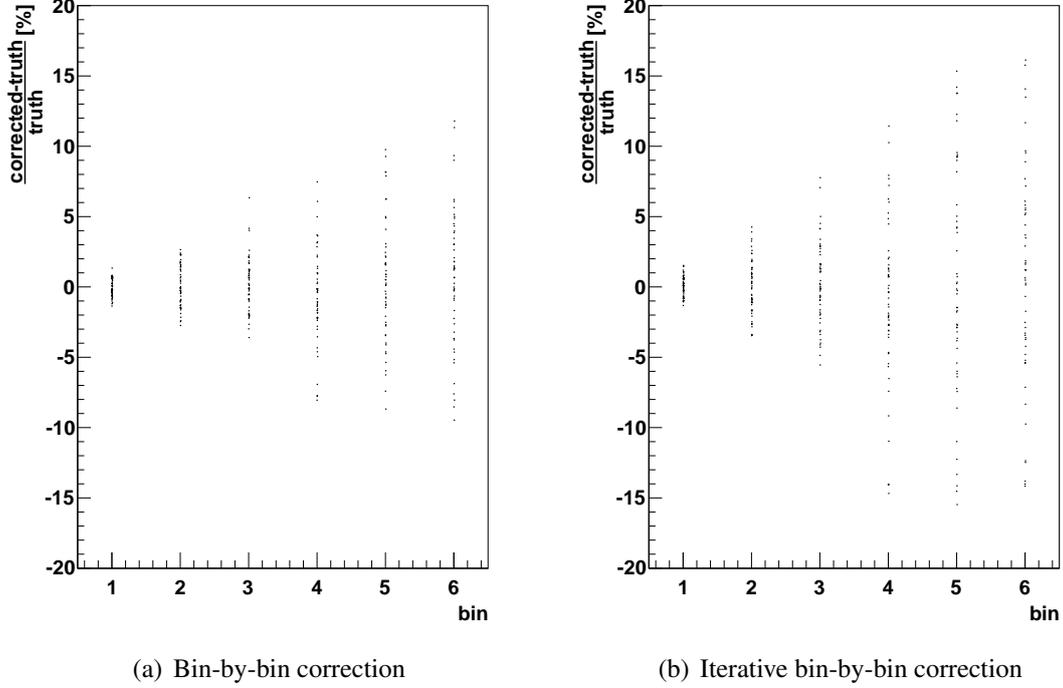


Figure A.3: Comparison between non-iterative and iterative bin-by-bin correction. Both subsamples used the same Drell Yan shape.

shift, but leads to an increased uncertainty for the higher bins. Thus the implemented iterative bin-by-bin correction is not suitable to correct the high mass Drell Yan spectrum. Possible reasons for this are the low statistics in the highest bins combined with some numerical instabilities. The other mentioned analyses used this method to study the p_T of the Z boson and the p_T of jets in Z +jets events. Both studies looked at regions with higher available statistics, while this analysis is focused on the tails of the $Z \rightarrow \mu^+\mu^-$ distribution.

In order to estimate the systematic uncertainty in bin i from a wrongly predicted $m_{\mu\mu}$ distribution in combination with bin-to-bin migration we first define the slope s_i as

$$s_i = \frac{\text{Content of bin } i}{\text{Content of bin } (i+1)} \quad (\text{A.6})$$

and then approximate that s_i does not change significantly between adjacent bins. The correction factor cf for a bin i with x entries is according to eq. A.2

$$cf_i = \frac{x}{x + m \cdot x \cdot s - m \cdot x + m \cdot \frac{x}{s} - m \cdot x}, \quad (\text{A.7})$$

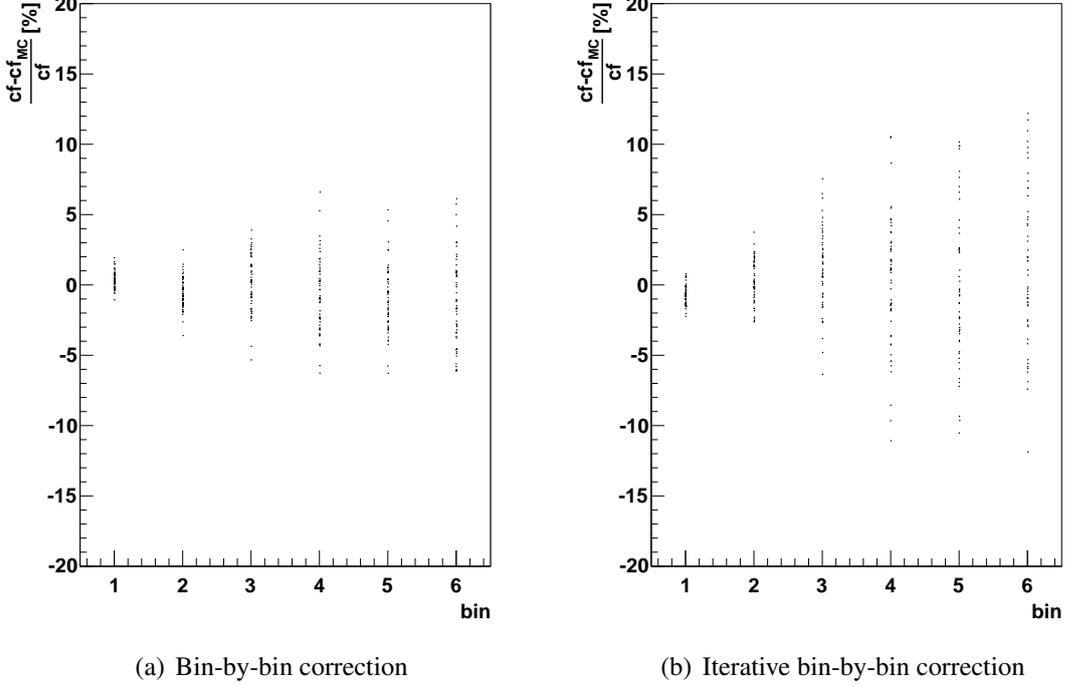


Figure A.4: Comparison between non-iterative and iterative bin-by-bin correction. Data and Monte Carlo sample use a different Drell Yan shape.

where the migration factor m is the percentage of events that will migrate into a neighbouring bin. In the denominator the terms $m \cdot x \cdot s$ and $m \cdot x$ reflect the migration from and to the bin $i - 1$, the terms $m \cdot \frac{x}{s}$ and $m \cdot x$ reflect the migration from and to the bin $i + 1$. If the slope s is predicted wrongly in the Monte Carlo by a factor $w = \frac{s_{MC}}{s}$ the resulting error in the correction factor will be

$$\frac{cf - cf_{MC}}{cf} = 1 - \frac{\frac{1}{m} + s - 2 + \frac{1}{s}}{\frac{1}{m} + w \cdot s - 2 + \frac{1}{w \cdot s}}. \quad (\text{A.8})$$

For a bin width of 50 GeV the full detector simulation gives a migration factor m in the order of 15% for a dimuon mass of 200 GeV. This migration factor is in fact neither symmetric nor constant with $m_{\mu\mu}$, but both approximations are reasonable for a rough estimation of the impact of bin-to-bin migration. From eq. A.5 we get a slope s of 2.5 for a bin at $m_{\mu\mu} = 200$ GeV. If we assume that in reality the parameter b in eq. A.5 is 2.4 instead of 2.7 we get a factor $w \approx 1.1$. According to eq. A.8 this would result in a systematic error of 3% on the bin around 200 GeV. A shift to the true Drell Yan spectrum of that size would change the ratio of events in a bin around 200 GeV to events in a bin

around 500 GeV by 60%, which would be easy to detect with high enough integrated luminosities.

The bin width of 50 GeV is only reasonable at integrated luminosities, where the statistics is high enough in the respective bin. For lower integrated luminosities a higher bin size would be used, which would in turn also reduce the bin-to-bin migration factor m . With $m = 0.05$ instead of 0.15 a shift of the parameter b from 2.7 to 2.4 would result in a systematic error of 1% in the bin around 200 GeV.

In conclusion, new physics in the form of a changed slope in the Drell Yan spectrum introduces a systematic error in the measured shape. This error is low compared to the change of the overall shape of the spectrum. If the measured Drell Yan spectrum is in agreement with the predictions, a non-iterative bin-by-bin correction is justified. If this is not the case, an improved iterative method or some other deconvolution technique has to be applied.

B Definitions

B.1 Luminosity

The luminosity \mathcal{L} is an important parameter of every accelerator, as the number of generated events per second is $\mathcal{L}\sigma_{event}$. Assuming a Gaussian beam distribution the luminosity can be calculated from the beam parameters as

$$\mathcal{L} = \frac{N_B^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F, \quad (\text{B.1})$$

where N_B is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalized transverse beam emittance, β^* the beta function at the collision point and F the geometric luminosity reduction factor due to the crossing angle at the Interaction Point [10].

B.2 Invariant dimuon mass

The invariant mass of two particles is defined as

$$m^2 = \left[\begin{pmatrix} E_1 \\ \vec{p}_1 \end{pmatrix} + \begin{pmatrix} E_1 \\ \vec{p}_1 \end{pmatrix} \right]^2 \quad (\text{B.2})$$

and can be interpreted as the total energy the particles have in their centre of mass system. In a two body decay the mass of the decaying particle is identical to the invariant mass of the two decay products.

B.3 Sphericity

To calculate the sphericity S of an event, the normalised momentum tensor M is defined as

$$M_{jk} = \frac{\sum_i p_j^i p_k^i}{\sum_i |\vec{p}^i|^2} \quad (\text{B.3})$$

where \vec{p}^i is the momentum vector of a reconstructed jet, electron or muon. j and k are Cartesian coordinates x, y and z . The three eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ define

$$S = \frac{3}{2}(\lambda_2 + \lambda_3) \quad (\text{B.4})$$

B.4 Transverse Sphericity

The transverse sphericity S_T is defined similarly to the sphericity S , using the same momentum tensor M from eq. B.3. The only difference is, that for the transverse sphericity the coordinates are only x and y . The two eigenvalues $\lambda_1 \geq \lambda_2$ define

$$S_T = \frac{2\lambda_1}{\lambda_1 + \lambda_2} \quad (\text{B.5})$$

B.5 Resolution

As mentioned in section 3.2.5 the Muon Spectrometer of ATLAS measures the sagitta s of transversing muons. The muon momentum p is then proportional to $1/s$. Assuming the sagitta has a gaussian uncertainty we define a fractional momentum resolution

$$\frac{\Delta p_f}{p} = \frac{1/p_{reco} - 1/p_{true}}{1/p_{true}} = \frac{p_{true} - p_{reco}}{p_{reco}} \quad (\text{B.6})$$

which has a gaussian distribution as well. For most physics purposes the resolution of the transverse momentum p_T is more important:

$$\frac{\Delta p_T}{p_T} = \frac{p_{Treco} - p_{Ttrue}}{p_{Ttrue}}. \quad (\text{B.7})$$

Since p_T is directly proportional to p and the angular resolution is negligible the relative resolution of p_T and p should be the same. The p_T resolution is not symmetric and has more accentuated tails than the fractional momentum resolution. For a resolution of 10% the difference in RMS between both definitions is only 4%, thus it is possible to get a good estimate of the p_T resolution by looking at the fractional momentum resolution. To study the impact of the muon resolution the asymmetry of the p_T resolution has to be taken into account.

C Additional Studies

C.1 B-Tagging Algorithms

There are several different algorithms available for b -tagging, the most commonly used being IP2D, SV1+IP3 and TrackCounting. As mentioned in section 3.3.4 TrackCounting is not available in the Athena version used in this analysis. In fig. C.1 the two other algorithms are compared, clearly showing that SV1+IP3D is more efficient than IP2D.

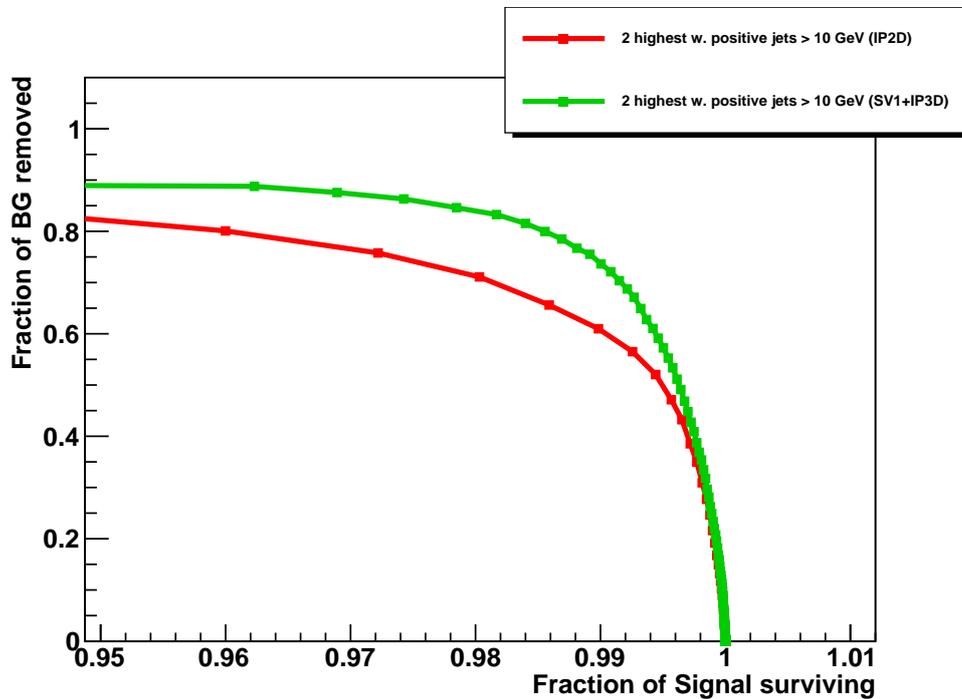


Figure C.1: Comparison of IP2D and SV1+IP3D

C.2 Differences between High Mass Drell Yan and Pythia Sample

The high mass Drell Yan and Pythia sample differ not only in the Athena version used for reconstruction, but also have different Monte Carlo generation options. For the Pythia sample, final state radiation is turned off, while it is turned on for the high mass Drell Yan sample. Moreover the Pythia sample requires at least one lepton with $|\eta| < 2.8$, while the high mass Drell Yan Sample requires two muons with $|\eta| < 2.8$ and $p_T > 5 \text{ GeV}$. According to the generation file those cuts have an efficiency of 0.86 for Pythia and 0.55 for the first slice of the high mass Drell Yan sample.

The first slice of the high mass Drell Yan sample contains approximately 40k events while all other slices contain around 20k events. Thus the integrated luminosity of the first slice is only around 50 pb^{-1} . The high relative weight of the events from the first slice combined with the large statistical uncertainty from the first slice will make it hard to compare those two samples for low dimuon masses.

To prevent both problems, the properties of the different variables are compared only for dimuon masses above 300 GeV, where the statistics is high enough and the different p_T cut has less impact.

Fig. C.2 shows, that it is safe to use the high mass Drell Yan sample to study the properties of \cancel{E}_T and isolation variables at very high dimuon masses. The differences in b -tagging prevent the high mass Drell Yan sample from being used for a cut on b -tagging.

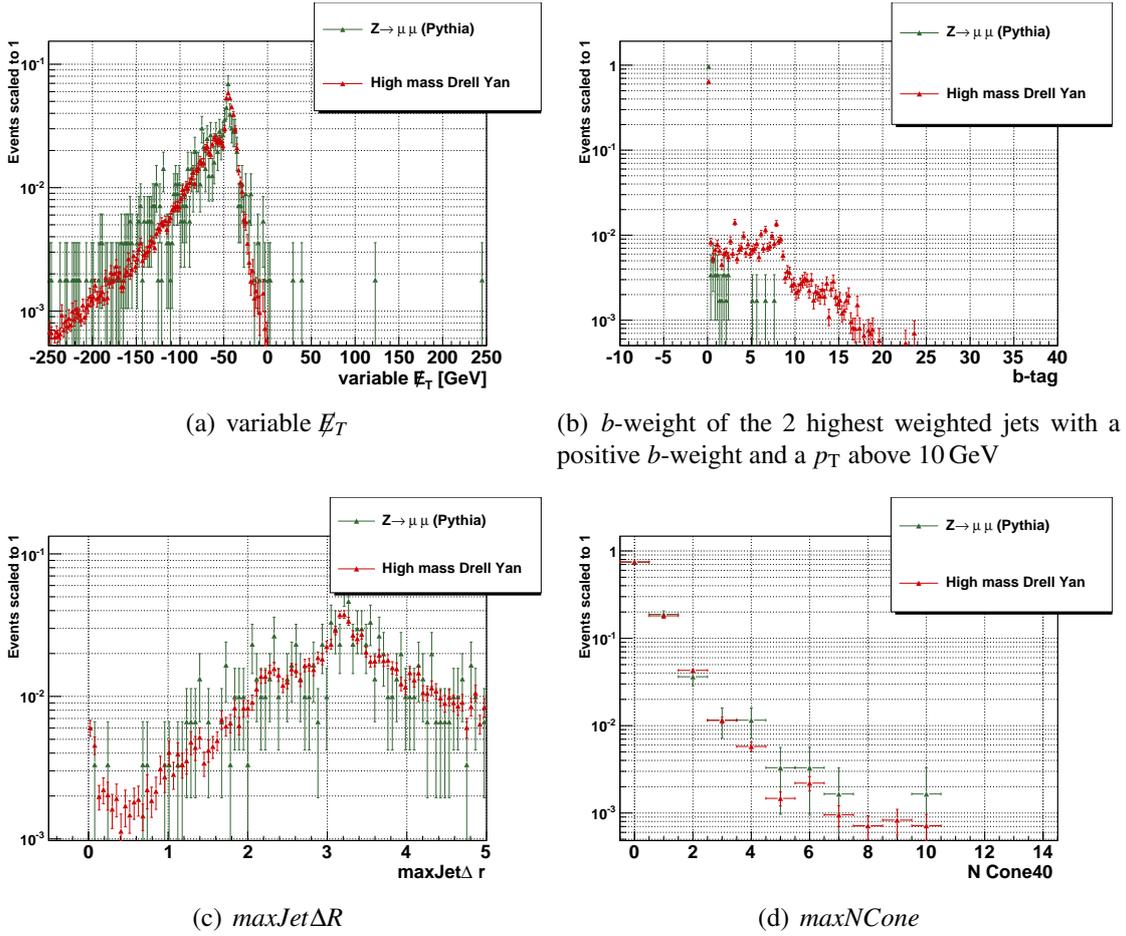


Figure C.2: Comparison of various cut variables between Pythia and high mass Drell Yan samples. For b -tagging the distribution differs significantly, while all other variables are in good agreement. All plots are for $m_{\mu\mu} > 300$ GeV and after preselection and quality cuts.

D Iterative Cut Optimisation

As mentioned before, the optimisation process is run iteratively until the cut value does not change anymore. In table D.1 this process is shown for the different integrated luminosities. The iteration is first performed for 50pb^{-1} and the result is used as a starting point for 500pb^{-1} and 5fb^{-1} .

Iteration	Integrated luminosity			
	Optimised cut	50pb^{-1}	500pb^{-1}	5000pb^{-1}
1	\cancel{E}_T b -tagging	-22.5 GeV 8	-7.5 GeV 3	-7.5 GeV 3
2	\cancel{E}_T b -tagging	-12.5 GeV 6.5	-7.5 GeV	-7.5 GeV
3	\cancel{E}_T b -tagging	-7.5 GeV 6.5		

Table D.1: Iterative optimisation of cut values for three different integrated luminosities.

E Datasets used

In table E.1 all samples are listed with the cross section times branching ratio given by the Monte Carlo generators.

Process	CSC ID	MC Generator	$\sigma \cdot \text{BR}[\text{pb}]$	MC Events
$Z \rightarrow \mu^+ \mu^-$, $m_{Z/\gamma^*} > 60 \text{ GeV}$	6051	Pythia	1098	2.94 M
$Z \rightarrow \mu^+ \mu^-$, $m_{Z/\gamma^*} > 60 \text{ GeV}$	6061	MC@NLO	1315	189 k
$Z \rightarrow \mu^+ \mu^-$, $60 \text{ GeV} < m_{Z/\gamma^*} < 200 \text{ GeV}$		Alpgen		
0 jets	7660	Alpgen	895	267 k
1 jet	7661	Alpgen	199	61.0 k
2 jets	7662	Alpgen	63	203 k
3 jets	7663	Alpgen	19	62.7 k
4 jets	7664	Alpgen	5.0	18.5 k
5+ jets	7665	Alpgen	1.4	5.44 k
$Z \rightarrow \mu^+ \mu^-$, $m_{Z/\gamma^*} > 200 \text{ GeV}$	5122	Pythia	1.6	15 k
$Z \rightarrow \mu^+ \mu^-$, $m_{Z/\gamma^*} > 75 \text{ GeV}$		Pythia		
75 GeV $< m_{Z/\gamma^*} < 120 \text{ GeV}$	5428	Pythia	1314	39.6 k
120 GeV $< m_{Z/\gamma^*} < 250 \text{ GeV}$	5429	Pythia	14.3	19.8 k
250 GeV $< m_{Z/\gamma^*} < 400 \text{ GeV}$	5430	Pythia	$721 \cdot 10^{-3}$	19.8 k
400 GeV $< m_{Z/\gamma^*} < 600 \text{ GeV}$	5431	Pythia	$125 \cdot 10^{-3}$	19.8 k
600 GeV $< m_{Z/\gamma^*} < 800 \text{ GeV}$	5432	Pythia	$22.6 \cdot 10^{-3}$	19.8 k
800 GeV $< m_{Z/\gamma^*} < 1000 \text{ GeV}$	5433	Pythia	$6.03 \cdot 10^{-3}$	19.9 k
1000 GeV $< m_{Z/\gamma^*} < 1250 \text{ GeV}$	5434	Pythia	$2.26 \cdot 10^{-3}$	19.9 k
1250 GeV $< m_{Z/\gamma^*} < 1500 \text{ GeV}$	5435	Pythia	$691 \cdot 10^{-6}$	19.9 k
1500 GeV $< m_{Z/\gamma^*} < 1750 \text{ GeV}$	5436	Pythia	$241 \cdot 10^{-6}$	19.9 k
1750 GeV $< m_{Z/\gamma^*} < 2000 \text{ GeV}$	5437	Pythia	$91.9 \cdot 10^{-6}$	19.8 k
$m_{Z/\gamma^*} > 2000 \text{ GeV}$	5438	Pythia	$66.5 \cdot 10^{-6}$	19.9 k
$t\bar{t}$ decaying not fully hadronically	5200	MC@NLO	203	487 k
$Z \rightarrow \tau^+ \tau^-$	6052	Pythia	1357	188 k
$W \rightarrow \mu \nu$		Alpgen		
0 jets	7690	Alpgen	10126	303 k
1 jets	7691	Alpgen	2156	50.7 k
2 jets	7692	Alpgen	682	156 k
3 jets	7693	Alpgen	203	42.7 k
4 jets	7694	Alpgen	57.0	11.9 k
5+ jets	7695	Alpgen	17.5	3.50 k
$W b\bar{b}$		Alpgen		
0 jets	6280	Alpgen	5.1	15.5 k
1 jets	6281	Alpgen	5.0	15.2 k
2 jets	6282	Alpgen	2.9	8.95 k
3+ jets	6283	Alpgen	1.6	5.00 k
$b\bar{b} \rightarrow \mu_{15}$	8405	PythiaB	88500	40 k
$b\bar{b} \rightarrow \mu_5 \mu_5$	8421	PythiaB	600000	902 k

Table E.1: List of all used samples and subsamples

Bibliography

- [1] Grundgesetz für die Bundesrepublik Deutschland. Artikel 1, Absatz 1
- [2] The ATLAS Collaboration. Expected Performance of the ATLAS Experiment. CERN-OPEN-2008-020
- [3] C. Amsler et al. (Particle Data Group). PL B667, 1 (2008) and 2009 partial update for the 2010 edition
- [4] R.K. Ellis, W.J. Stirling, and B.R. Webber. QCD and Collider Physics. *Cambridge University Press* (1996)
- [5] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys.*, B126:298 (1977)
- [6] J. Collins and D. Soper. The Theorems of Perturbative QCD *Annu. Rev. Nucl. Part. Sci.* 37 (1987), and references therein.
- [7] B. Webber. Fragmentation and hadronization. CERN Theory Division, <http://www.slac.stanford.edu/econf/C990809/docs/webber.pdf>, accessed on 2010-07-19
- [8] O. Biebel. Experimental tests of the strong interaction and its energy dependence in electron+positron annihilation, *Physics Reports* 340 (2001)
- [9] S.D. Drell and T.M. Yan, *Ann. Phys* 66 (1971)
- [10] O. Bruning (Ed.) et al. LHC design report. Vol. I: The LHC main ring. CERN-2004-003-V-1 (2004)
- [11] *CERN press release*, 2008-08-07: <http://press.web.cern.ch/press/PressReleases/Releases2008/PR06.08E.html>, accessed on 2010-07-11
- [12] *CERN press release*, 2008-09-19: <http://press.web.cern.ch/press/PressReleases/Releases2008/PR09.08E.html>, accessed on 2010-07-11

- [13] *CERN press release*, 2009-08-06: <http://press.web.cern.ch/press/PressReleases/Releases2009/PR13.09E.html>, accessed on 2010-07-11
- [14] *CERN press release*, 2009-11-23: <http://press.web.cern.ch/press/PressReleases/Releases2009/PR17.09E.html>, accessed on 2010-07-11
- [15] *CERN press release*, 2010-03-30: <http://press.web.cern.ch/press/PressReleases/Releases2010/PR07.10E.html>, accessed on 2010-07-11
- [16] ATLAS Collaboration. ATLAS:Detector and physics performance technical design report. Volume 1. CERN-LHCC-99-14 (2009)
- [17] M. Della Negra et al. CMS physics : Technical Design Report , v.1 : Detector performance and software. CERN-LHCC-2006-001 (2006)
- [18] P. Cortese et al. ALICE physics performance : Technical Design Report CERN-LHCC-2005-030 (2005)
- [19] LHCb Collaboration. LHCb reoptimized detector design and performance : Technical Design Report, CERN-LHCC-2003-030 (2003)
- [20] V. Berardi et al. Total cross-section, elastic scattering and diffraction dissociation at the Large Hadron Collider at CERN : TOTEM Technical Design Report CERN-LHCC-2004-002 (2004)
- [21] O. Adriani et al. OLHCf experiment : Technical Design Report CERN-LHCC-2006-004 (2006)
- [22] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider, *JINST* 3 (2008) S08003
- [23] ATLAS Collaboration. ATLAS calorimeter performance. CERN-LHCC-96-40 (1996)
- [24] M. Schott. Study of the Z Boson Production at the ATLAS Experiment with First Data. PhD thesis, Ludwig-Maximilians-Universität, Munich (2007)
- [25] G.C. Blazey et al. Run II jet physics,(2000) [hep-ex/0005012v2]
- [26] M. Cacciari, G. P. Salam and G. Soyez. The anti-kt jet clustering algorithm *JHEP* 04 (2008) 063
- [27] S. Catani, Y.L. Dokshitzer, M.H. Seymour and B.R. Webber. Longitudinally invariant K(t) clustering algorithms for hadron hadron collisions,*Nucl. Phys. B* 406 (1993) 187

-
- [28] C. Kummer. Search for supersymmetry events with two same-sign leptons. PhD thesis, Ludwig-Maximilians-Universität, Munich (2010)
- [29] S. Hassani et al. A Muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using (Muonboy, STACO, MuTag) reconstruction packages, *NIM A* 572, (2007) 77-79
- [30] S. Hassani et al. *NIM A* 572 (2007) 77-79
- [31] *Internal ATLAS Twiki* <https://twiki.cern.ch/twiki/bin/view/Atlas/CaloMuonContainer>, accessed on 2010-07-13
- [32] T. Goepfert. Counting signed impact parameter based b-tagging algorithm with ATLAS. *ATLAS internal note*, ATL-PHYS-INT-2009-064 (2009)
- [33] ROOT - An Object Oriented Data Analysis Framework, Rene Brun and Fons Rademakers, *NIM A* 389 (1997) 81-86, <http://root.cern.ch>, accessed on 2010-07-11
- [34] ATHENA, <http://twiki.cern.ch/twiki/bin/view/Atlas/AthenaFramework>, accessed on 2010-07-11
- [35] The Gaudi Collaboration: <http://proj-gaudi.web.cern.ch>, accessed on 2010-07-11
- [36] S. Agostinelli et al. Geant4: A simulation toolkit. Nuclear Instruments and Methods in *Physics Research A* 506 (2003) 250303
- [37] K. Assamagan et al. The ATLAS Monte Carlo Project. *ATLAS Internal Note* ATL-SOFT-INT-2010-002 (2010)
- [38] *Internal ATLAS Twiki* <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HighPtView>, accessed on 2010-07-13
- [39] The ATLAS Collaboration. ATLAS Computing TDR. Technical Design Report (2005)
- [40] Worldwide LHC Computing Grid: <http://lcg.web.cern.ch/LCG/public/overview.htm>, accessed on 2010-07-11
- [41] J. Kennedy et al *J. Phys.: Conf. Ser.* 219 072039 (2010)
- [42] R. Mameghani. Semi- and Dileptonic Top Pair Decays at the ATLAS Experiment PhD thesis, Ludwig-Maximilians-Universität, Munich (2008)
- [43] O. Biebel. *private communication* (2010).

- [44] R. V. Gvai, S. Gupta, P.L. McGaughey, E. Quack, P. V. Ruuskanen, R. Vogt, Xinnian Wang. Heavy Quark Production in pp Collisions. *Int.J.Mod.Phys A*10 (1995) 2999-3042 [hep-ph/9411438v1]
- [45] G. Wodden. Leptonic fake rates from jets at ATLAS, *ATLAS Communication*, ATLCOM-PHYS-2009-202 (2009)
- [46] Sjöstrand, S. Mrenna, and P. Skands. PYTHIA 6.4 physics and manual, *JHEP* 05 (2006) 026 [hep-ph/0603175]
- [47] M. Smizanska, S. Baranov, J. Hrivnc, and E. Kneringer. Overview of monte carlo simulations for atlas b-physics in the period 1996-1999, *ATLAS Note* ATL-PHYS-2000-025 (2000)
- [48] C. Anastopoulos et al., Physics analysis tools for beauty physics in atlas, *J. Phys.: Conf. Ser.* 119 (2008) 032003.
- [49] G. Marchesini, B. Webber, G. Abbiendi, I. Knowles, M. Seymour, and L. Stanco. Herwig: A monte carlo event generator for simulating hadron emission reactions with interfering gluons. version 5.1 - april 1991. *Comput. Phys. Commun.* 67 (1992) 465
- [50] G. Corcella et al. HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP* 01 (2001) 010, [hep-ph/0011363]
- [51] G. Corcella et al. Herwig 6.5 release note. CERN-TH/2002-270 (2005) [hep-ph/0210213v2]
- [52] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001 (2003) [hep-ph/0206293]
- [53] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP* 0206 (2002) 029 [hep-ph/0204244]
- [54] S. Frixione, P. Nason and B. R. Webber. Matching NLO QCD and parton showers in heavy flavour production. *JHEP* 0308 (2003) 007 [hep-ph/0305252]
- [55] *Internal ATLAS Twiki* <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopReferences10TeV>, accessed on 2010-07-11
- [56] J. Campbell, R.K. Ellis and D. Rainwater. Next-to-leading order QCD predictions for W+2j and Z+2j production at the CERN LHC. *Phys.Rev.D*68:094021(2003) [hep-ph/0308195v2]

-
- [57] The CDF Collaboration. Search for high mass resonances decaying to muon pairs, (2010) CDF/PHYS/EXO/PUBLIC/10165
- [58] Lena Kallin Westin. Receiver operating characteristic (ROC) analysis. Evaluating discriminance effects among decision support systems, Umea universitet ; UMINF-101.18
- [59] C. Clement, C. Gerber, G. Otero, M. A. Pleier, E. Shabalina, and J.-R. Vlimant. Measurement of the $t\bar{t}$ Production Cross-Section at $\sqrt{s} = 1.96$ TeV in the Electron+Jets Final State using a Topological Method. *DO note* 4662 (2005)
- [60] The ALICE Collaboration. First proton–proton collisions at the LHC as observed with the ALICE detector: measurement of the charged particle pseudorapidity density at $\sqrt{s} = 900$ GeV *Eur.Phys.J.* C65:111-125 (2010) [hep-ex/0911.5430v2]
- [61] The CMS Collaboration. Transverse momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 0.9$ and 2.36 TeV. *JHEP* 02 (2010) 041 [hep-ex/1002.0621v2]
- [62] The ATLAS Collaboration. Charged-particle multiplicities in pp interactions at $\sqrt{s} = 900$ GeV measured with the ATLAS detector at the LHC. *Phys Lett B* 688 (2010) Issue 1, 21-42 [hep-ex/1003.3124v2]
- [63] B. Tiller. Measurement of the Associated Production of Z Bosons and Jets in ppbar Collisions at the Tevatron. PhD thesis, Ludwig-Maximilians-Universität, Munich (2009)

Acknowledgements

Many people contributed to this work and I want to thank them at this place.

Above all I want to thank my “Doktorvater” Otmar Biebel for giving me the opportunity to write this thesis and for his continuous trust and encouragement. I also want to thank Dorothee Schaile for welcoming me at her institute.

Many thanks to my supervisor Raimund Ströhmer for his various inspirations and suggestions. With his amazing feeling for correlations he provided valuable feedback to this analysis.

I also want to thank my MPI-colleagues Sonja Leber, Jürgen Zimmer, Steffen Kaiser, Jörg von Loeben and Reinhard Sedlmeyer and the LMU-colleagues Felix Rauscher and Attila Varga for the support they gave me while staying at CERN and all the things they taught me while commissioning the ATLAS detector.

Marie-Helene Genest was a great help in providing the DPDs and helping to understand their contents.

Otmar Biebel and Julien de Graat deserve many thanks for carefully reviewing this dissertation.

I want to thank Stefanie Adomeit, Johannes Becker, Johannes Ebke, Albert Engl, Julien de Graat, Christian Kummer, Markus Lichtnecker and Benjamin Ruckert for all the interesting discussions about QCD and related topics and all other members of the Munich group for the good working climate.

Matthias Schott and Stefanie Adomeit were the best possible office roommates I could wish for.

I also want to thank my ex-girlfriend Sabine Öttl for supporting me during my numerous travels to CERN and my girlfriend Alexandra Wörner for being there while I wrote this thesis. And finally I want to thank my parents for enabling me to study physics and to do my PhD.

