Search for New Physics in Proton-Antiproton Collisions in Final States with Leptons and Jets

von

Oliver Kraff

Diplomarbeit in Physik

vorgelegt der

Fakultät für Mathematik, Informatik und Naturwissenschaften
der Rheinisch-Westfälischen Technischen Hochschule Aachen

im
November 2004

angefertigt im

III. Physikalischen Institut A

bei

Prof. Dr. Thomas Hebbeker

Contents

Chapter 1

Introduction

Since the Greek philosophers of the antique mankind asks one question: what is the origin of matter? The ultimate goal of all research in high energy physics is to answer this question by setting up a final theory of the fundamental particles and the forces between them. Although we do not know how far away we are from this goal our understanding increased especially since the development of quantum mechanics in the 1930s as well as with the progress in technology, making accelerator and detector programmes realisable. Today, our knowledge about the structure of matter is summarised in the Standard Model of particle physics (SM) which was built up through decades of intensive dialogue between theory and experiment. Leptons and quarks were discovered as the fundamental building blocks of matter. Carriers of the electromagnetic, weak and strong force are the massless photon, the massive W and Z bosons and eight massless gluons, respectively. Towards a grand unification of the fundamental forces, the electromagnetic and weak force are combined into the electroweak gauge field theory. The quantum chromodynamics gauge field theory has been confirmed as the theory of strong interactions.
The Standard Model has passed many stringent experimental tests with colliding electron-positron (e^+e^-), electron-proton (ep) and proton-antiproton ($p\bar{p}$) beams, giving complementary contributions. Some experiments are to a very high precision as the Z mass measurement at the Large Electron-Positron storage ring (LEP) at CERN[1] giving a Z mass of 91.1876 ± 0.0021 GeV [1]. Also QCD predictions have been thoroughly tested like measurements of the strong coupling α_s at LEP and at the Hadron-Elektron Ring Anlage (HERA) in Hamburg. The proton structure is being probed to the shortest accessible distances using ep collisions at HERA. The SM has succesfully predicted the τ neutrino, to name only one, discovered in July 2000 at the Fermi National Laboratory.
On the other hand, there are still many gaps in our understanding. We have no explanation for the wide range of masses of the leptons and quarks: from less than 1 eV for neutrinos to ≈ 178 GeV for the top quark [3]. Also the origin of the masses

---

[1] Centre Européenne pour la Recherche Nucléaire, Geneva, Switzerland

themselves is not yet understood. Another unsolved riddle is the excess of matter over antimatter in our universe. CP (charge conjugation, parity) violating processes are not understood at the level required to account for that question. Finally, the hierachy between the scale of electroweak physics and the scale of gravity is still unanswered and gravity is far away from being brought close to the other forces.

This analysis investigates data from the DØ experiment at the Tevatron collider at Fermilab near Chicago, USA. At the Tevatron protons and antiprotons collide with a centre of mass energy of $\sqrt{s} = 1.96\ TeV$. The multi-purpose DØ detector provides a precise measurement of leptons as well as hadronic showers, which are called jets in the following. Around 230 pb^{-1} data are analysed to search for deviations from the SM prediction in a model-independent way, considering final states with electrons and/or muons together with additional jets. Model-independent means that every detected event will be compared with the SM prediction without any prejudice. A deviation may be a hint for new physics. Furthermore, this analysis provides a broad understanding of the detector as well as the Monte Carlo simulations which represent the theoretical prediction.

The algorithms and programs used are a combined work of this analysis and [4], which investigates final states with leptons and missing transverse energy.

Chapter 2.2 gives an introduction to Standard Model physics, followed by a description of the Tevatron accelerator and the DØ detector in chapter 3. The datasets investigated and the selection criteria for the physics objects electrons, muons and jets are summarised in chapter 4, while chapter 5 gives an overview of the Monte Carlo simulations and explains the estimation of the QCD background. Systematic uncertainties are listed in chapter 6. Finally, chapter 7 presents the search for new physics which is performed in three steps. Frist, a simple comparison by total number of events is done. The next stage is a sophisticated search algorithm, described in chapter 8, which finds regions of largest deviation between data and SM prediction. Then, the probability of occurance of such a deviation is derived. A summary and an outlook are given in chapter 9.

Chapter 2

Theory

This chapter introduces the use of natural units and explains abbreviations used in this document. Also, a brief introduction to Standard Model physics and some of the possible extensions of this theory are given, followed by a description of the properties of proton-antiproton collisions and fragmentation.

2.1 Natural Units and Abbreviations

In particle physics it is common to measure quantities in units of 1 $GeV = 10^9\ eV$, where 1 electron volt (eV) is defined as the energy gained by a particle of charge $e = 1.602 \cdot 10^{-19}\ C$ traversing a potential difference of 1 Volt. This choice is motivated by the fact that the rest mass of the proton is around 1 GeV. The two fundamental constants in relativistic quantum mechanics are Planck's constant h and the velocity of light in vacuum c:

$$\begin{aligned} \hbar = \frac{h}{2\pi} &= 1.055 \cdot 10^{-34} Js \\ c &= 2.998 \cdot 10^{8} ms^{-1} \end{aligned} \tag{2.1}$$

In the system of natural units, these two constants are chosen as $\hbar = c = 1$ so that they need not be written explicitly in the formulae. Table 2.1 gives the conversion formulae between natural units and SI units.

For further simplification of the notation the following abbreviations are used

- e and the name electron both stand for e^+ (positron) or e^- (electron)
- muon and the abbreviation μ mean μ^+ or μ^-
- W stands for both types of the W boson: W^+ or W^-
- p and $\bar{p}$ denote the proton and antiproton, respectively
- p_T is the transverse momentum of a particle, i.e. the momentum perpendicular to the beam pipe (defined as z axis here) of an accelerator:

$$\vec{p} = (p_x, p_y, p_z) \Rightarrow |\vec{p}| \equiv p = \sqrt{p_x^2 + p_y^2 + p_z^2} = \sqrt{p_T^2 + p_z^2}$$

- cross sections are given in the unit barn: $1\ b = 1\ barn = 10^{-28}\ m^2$.

conversion factor	$\hbar = c = 1$ units	SI unit
$1\ kg = 5.61 \cdot 10^{26}\ GeV$	GeV	$GeV\ c^{-2}$
$1\ m = 5.07 \cdot 10^{15}\ GeV^{-1}$	GeV^{-1}	$\hbar\ c\ GeV^{-1}$
$1\ s = 1.52 \cdot 10^{24}\ GeV^{-1}$	GeV^{-1}	$\hbar\ GeV^{-1}$
$e = \sqrt{4\pi\alpha}$	-	$(\hbar c)^{-1/2}$

Table 2.1: The SI units for mass (kg), length (m), time (s) and elementary charge (e) in terms of natural units.

2.2 The Standard Model

According to our current understanding there are three distinct types of elementary particles: the first two are fermions (spin-$\frac{1}{2}$), called leptons and quarks, and the third are bosons (spin-1) which are called gauge bosons and are the force carriers. The word *gauge* refers to gauge theories which are based on the idea that the Lagrangian describing the theory must possess local symmetries as well as global symmetries. Leptons and quarks are the building blocks of matter. There are three generations of lepton and quark families (as well as of their corresponding antiparticles[1]) ordered by increasing mass (see table 2.2). The first generation consists of the leptons electron (e) and its corresponding electron neutrino (ν_e) and the two quarks up (u) and down (d). In the second generation there are the muon (μ) and muon neutrino (ν_μ) as well as the charm (c) and strange (s) quark. The third generation consits of the leptons

---

[1] Antiparticles have the same mass but opposite electrical charge, colour and third component of the weak iso-spin.

	I	II	III	Charge
Lepton	e	μ	τ	-1
	ν_e	ν_μ	ν_τ	0
Quark	u	c	t	+2/3
	d	s	b	-1/3

Table 2.2: The three generations of leptons and quarks.

tau (τ) and tau neutrino (ν_τ) and the quarks top (t) and bottom (b). Electron, muon and tau carry an electrical charge of one unit whereas the neutrinos are neutral. Neutrinos were long assumed to be massless but recent experiments[2] have shown that their masses or not zero. Quarks carry an electrical charge of $+\frac{2}{3}$ or $-\frac{1}{3}$. Leptons and quarks distinguish themselves in their interactions which will be described in the following.

The Standard Model is a quantum field theory describing all particles as fields. The interaction between the fundamental particles is explained by the exchange of twelve mediating gauge bosons. In principle, there are four different types of fundamental forces in nature, which can be ordered by increasing strength to: gravity, weak force, electromagnetic force and strong force (see table 2.3).

Although the gravitational force is the most experienced one in every day life it is not included in the SM as it is much weaker compared to the other three and too weak to be observable in any particle physics laboratory of our time. However, gravitational effects can not be neglected when considering masses or energies of the order of the Planck mass $m_{Pl} = \sqrt{\frac{1}{G}} = 1.22 \cdot 10^{19}\ GeV$ with G being the gravitational constant.
The electromagnetic force is mediated by the exchange of massless photons (γ) coupling to all electrically charged particles. The strength of this interaction is given by the fine structure constant $\alpha \approx \frac{1}{137} \approx 10^{-2}$ and has infinite range.
The weak force is responsible for e.g. the neutron decay $n \rightarrow p + e^- + \bar{\nu}_e$. The mediating particles are the charged W-bosons and the neutral Z-boson. These bosons are massive which corresponds to their interaction range of $\hbar/M_W c \approx 10^{-17}\ m$. The strength is given by the Fermi constant $G_F \approx 10^{-5}\ GeV^{-2}$.
The strong force binds quarks to hadrons. This interaction reflects the main difference between quarks and leptons as the latter do not interact strongly. Quarks, on the other hand, carry an additional charge called colour which comes in red (R), green (G) and blue (B). Antiquarks are assigned complementary anticolours $\bar{R}$, $\bar{G}$, $\bar{B}$. The carriers of the strong force are eight massless gluons which are themselves bicoloured objects of colour and anticolour. Unlike photons which do not carry electrical charges themselves,

---

[2] For example, the K2K Long-baseline Neutrino Oscillation Experiment [6]

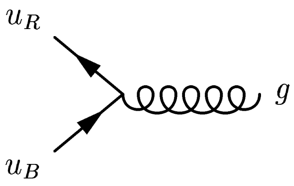
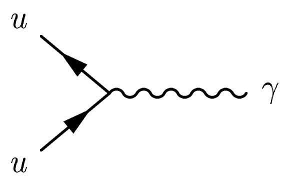
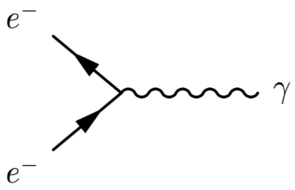
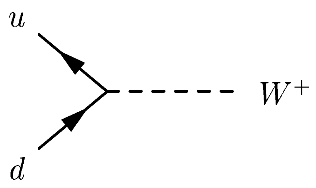
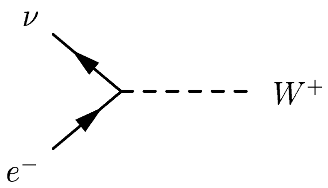
Force	Carrier	Quarks	Leptons
strong	8 gluons		
electromagnetic	photon		
weak	W, Z		

Table 2.3: The fundamental forces and their typical interactions. Gravity is not described here as it is not included in the SM.

gluons interact with each other which results in a running of the strong coupling constant α_s. The coupling depends on the energy scale μ as

$$\alpha_s(\mu) = \frac{\alpha_s(\mu_0)}{1 - \beta_0^s \alpha_s(\mu_0) ln\left(\frac{\mu^2}{\mu_0^2}\right)}$$

with $\alpha_s(\mu_0)$ being a value at a certain reference scale μ_0 and β_0^s being the β-function which is negative in QCD meaning that at low energies, i.e. large distances, the coupling and hence the strong force are very large[3]. This effect is called confinement and is the reason for the fact that free quarks cannot exist. On the other hand, at small distances, i.e. huge energy transfers, quarks behave almost like free particles; this is called asymptotic freedom. For their work on the theory of the strong force the physicists D. Gross, H. Politzer and F. Wilczek were honoured with the Nobel prize in physics of the year 2004.

The Noether theorem says that every invariance of the Lagrangian under a global symmetry transformation $U(\alpha) = e^{i\alpha}$, α being a phase constant, results in a conserved quantity:

invariance under translations ↔ conservation of momentum,
invariance under rotations ↔ conservation of angular momentum,
invariance under time displacements ↔ conservation of energy.

The previously introduced fundamental interactions are based on local gauge invariances requiring photons and gluons to be massless. A local gauge transformation is described by $U(\alpha) = e^{i\alpha(x)}$ with the phase α now being dependent on time and space $x_\mu = (t, \vec{r})$. In the Glashow-Salam-Weinberg model the electromagnetic and weak interaction are combined to the electroweak interaction within the SM, whose corresponding symmetry group is the $SU(2)_L \otimes U(1)_Y$. The subscripts L and Y denote left-handed iso-spin doublets and hyper charge, respectively. The left-handed fermions differ in the third component of the weak iso-spin $T_3 = \pm\frac{1}{2}$ which is reflected in their position in the iso-spin doublets. Right-handed fermions form iso-spin singlets $T_3 = 0$. Both right- and left-handed fermions carry a hyper charge Y defined by the Gell-Mann-Nishijima relation $Q = T_3 + Y/2$ with Q being the electric charge. Table 2.4 summarises these three quantum numbers for leptons and quarks. Hence, the $SU(2)_L \otimes U(1)_Y$ relates to rotations in weak iso-spin and in hyper charge space. For the strong force the $SU(3)_C$ is the corresponding symmetry group which is related to rotations in colour (C) space.

To explain the masses of the gauge bosons W and Z the theory introduces another scalar particle, the Higgs boson. The W and Z bosons acquire their masses through the coupling

---

[3] In QED the β-function is positive resulting in an effective coupling which increases with energy.

	Fermions			Quantum Numbers		
	1. Generation	2. Generation	3. Generation	Q	Y	T_3
Leptons	$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$	0 -1	-1 -1	$1/2$ $-1/2$
	e^-_R	μ^-_R	τ^-_R	-1	-2	0
	$\nu_{e,R}$	$\nu_{\mu,R}$	$\nu_{\tau,R}$	0	0	0
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	$2/3$ $-1/3$	$1/3$ $1/3$	$1/2$ $-1/2$
	u_R	c_R	t_R	$2/3$	$4/3$	0
	d_R	s_R	b_R	$-1/3$	$-2/3$	0

Table 2.4: Fermions and three of their quantum numbers: electric charge (Q), hyper charge (Y) and iso-spin (T_3). The weak iso-spin doublets are combined in brackets.

to the Higgs field. The Higgs mechanism also introduces the masses of the fermions through the so-called Yukawa coupling to the Higgs field. It should be emphasised that it does not explain the value of the masses. Up to now, the Higgs boson has not been observed. At LEP, direct searches lead to the conclusion that the Higgs boson must be heavier than 114.4 GeV (at 95% confidence level limit) [2]. Recent results like the improved measurement of the mass of the top quark at the Tevatron to a new world average of 178.0 ± 4.3 GeV increased the most likely Higgs mass to 117 GeV, just above the excluded mass range. The upper limit on the Higgs mass at 95% confidence level (CL) is raised from 219 GeV to 251 GeV [3]. Figure 2.1 shows these indirect Higgs limits.

2.3 Beyond Standard Model Physics

This section gives a brief overview of a few possible scenarios for physics beyond the Standard Model: technicolor [7], leptoquarks [8] and supersymmetry [9].

In 1979, Weinberg and Salam introduced technicolor to explain the electroweak symmetry breaking. According to this theory the electroweak symmetry is broken by the dynamics of a postulated QCD-like force and not through the existence of a Higgs particle. This new strong interaction of technifermions creates the technipion, which plays the role of the Higgs field. Just like the quarks in QCD, in technicolor a new spectrum of particles exists, called techniquarks, which transform according to the fundamental respresentation of the technicolor gauge group $SU(N_{TC})$ as well as the usual SM group. The original concept needed extensions and modifications to explain the existence of fermions and the heavy top quark. The latter is explained by an additional top-color.

Leptoquarks are predicted by many extensions of the SM which aim at explaining the

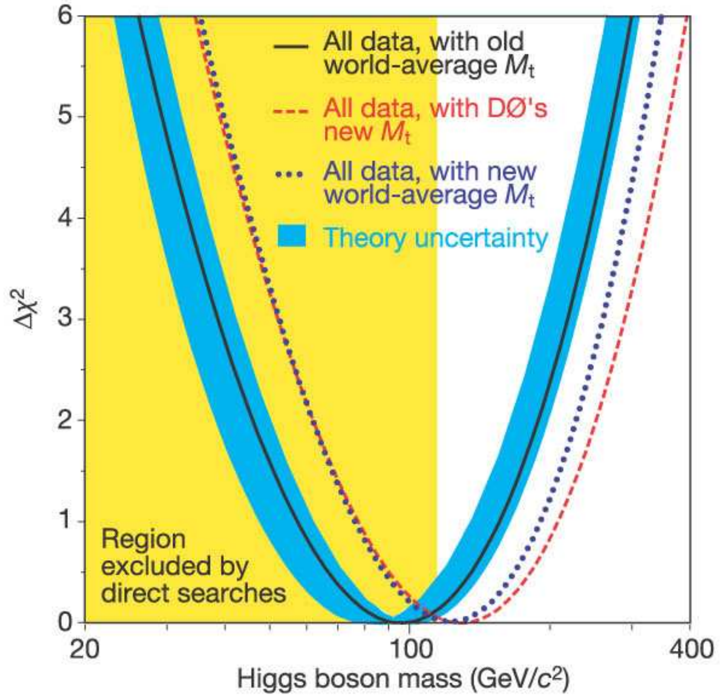


Figure 2.1: Current experimental constraints on the Higgs mass after a reanalysis of DØ Run I data, giving a new world average for the top mass. Shown is the χ^2 for a global fit to electroweak data as a function of the Higgs mass [3].

symmetry between quarks and leptons, e.g. Grand Unified Theories (GUT) or models where quarks and leptons are not elementary. These new spin-1 particles are a composition of quarks and leptons and carry the quantum numbers of a quark-lepton system. At hadron colliders, leptoquarks would be dominantly pair produced by gluon-gluon fusion and $q\bar{q}$ annihilation (figure 2.2).

In supersymmetry (SUSY) theories, an additional symmetry between fermions and bosons is predicted. To every SM particle a supersymmetric partner with a spin difference of $\frac{1}{2}$ is assigned. Each fermion of the SM has a scalar partner, i.e. a spin-0 boson, like:

SM particle		SUSY particle
electron e	$\leftrightarrow$	selectron $\tilde{e}$
neutrino ν	$\leftrightarrow$	sneutrino $\tilde{\nu}$
quark q	$\leftrightarrow$	squark $\tilde{q}$
	etc.	

The bosons of the SM have SUSY partners of spin-$\frac{1}{2}$. The partners of the neutral bosons

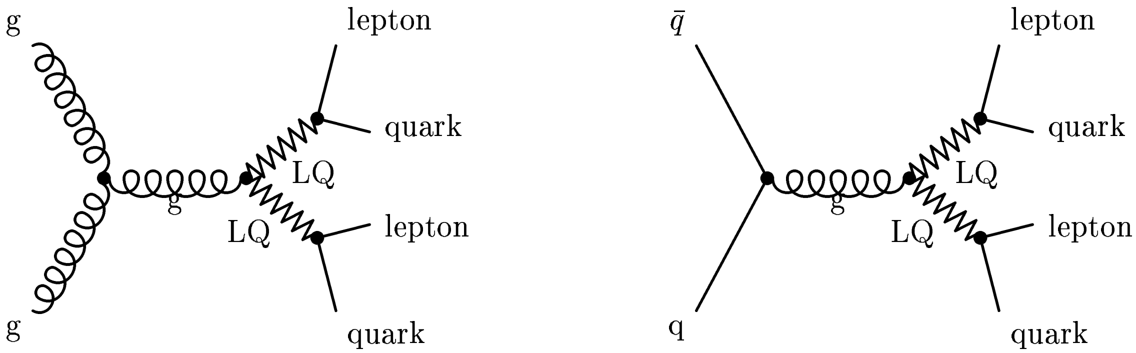


Figure 2.2: Pair production of leptoquarks (LQ) by gluon-gluon fusion and by quark-antiquark annihilation.

mix to form four neutral mass eigenstates, the neutralinos $\tilde{\chi}_i^0$ with $i = 1, ..., 4$ and the partners of the charged bosons mix to form two charginos $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_2^{\pm}$, where the index 1 indicates the lightest and 2(4) the heaviest of these new particles. A new quantum number called R-parity is introduced by giving SM particles the value $R = +1$ and SUSY particles $R = -1$. There are many SUSY models: minimal supersymmetric model (MSSM), minimal supergravity (mSUGRA), gauge mediated SUSY breaking (GMSB) and anomaly mediated SUSY breaking (AMSB). Each of these models postulates different mechanisms for the breaking of supersymmetry. In some models, the quantum number R is conserved resulting in a lightest stable SUSY particle (LSP) which is a good dark matter candidate in cosmology. Supersymmetry is at the root to unify gravity with the other fundamental interactions and explains the hierachy between the energy scales.

2.4 Properties of $p\bar{p}$ Collisions and Fragmentation

As previously described, proton and antiproton are composite objects of gluons and quarks. Due to quantum fluctuations of the gluons the three valence quarks of the proton "swim" in a diffuse sea of quark-antiquark pairs, known as "sea" quarks. These are mainly u, d and s quarks. Thus, the total momentum of the proton is not concentrated in a single parton, but split into many different fractions according to its constituents. Figure 2.3 shows the momentum fraction x of the partons of the proton, for a momentum transfer[4] of $Q^2 = 10\ GeV^2$. When a collision takes place essentially only one parton of the proton interacts with one parton of the antiproton. Only a part of the center of mass energy of the $p\bar{p}$ system is used in a collision.

---

[4] The momentum transfer Q is defined as the difference between the momenta of one of the particles before and after a collision/scattering.

Proton and antiproton are destroyed and those partons not contributing to a direct collision continue their flight under a small angle to the beam pipe and hadronise into jets. Hence, QCD is a very dominant process at the Tevatron.

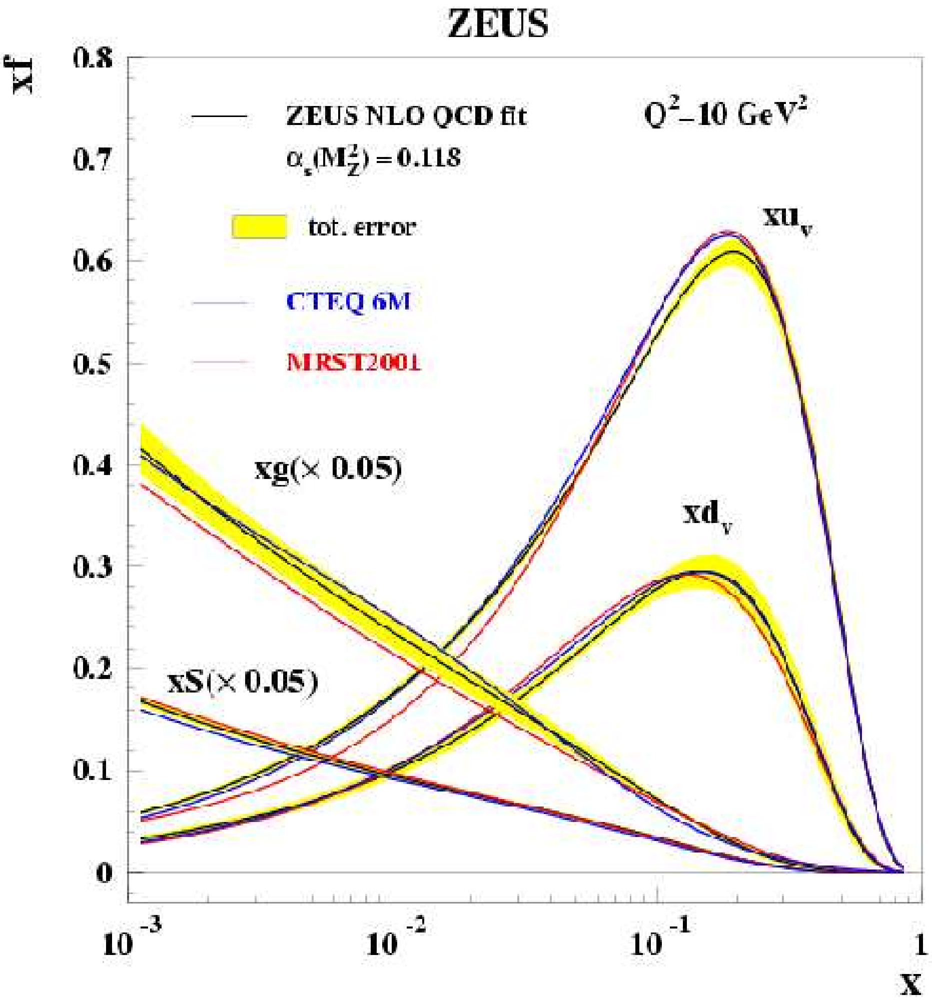


Figure 2.3: Measurements at ZEUS, DESY. Shown are the momentum distributions of the u- and d-valence quarks (xu_v, xd_v), the gluons (xg) and sea quarks (xS) as a function of the proton momentum fraction [5].

Fragmentation

Fragmentation, or sometimes called hadronisation, is a process of non-perturbative QCD and is yet not understood in all details. A common technique to probe quarks and antiquarks is the head on collision of very high energy electron and positron beams (see figure 2.4). The virtual photon, produced at rest by the e^+e^- collision can decay into a $q\bar{q}$ pair. Due to the conservation of energy and momentum the partons distance from each other. When the pair separates by distances of the order of 1 fm, α_s and hence the colour interaction becomes so strong that the increasing potential energy is sufficient to create another $q\bar{q}$ pair. This process takes place in many steps until all colour charges, i.e. quarks and gluons, are combined to colourless hadrons (mostly light π-mesons).

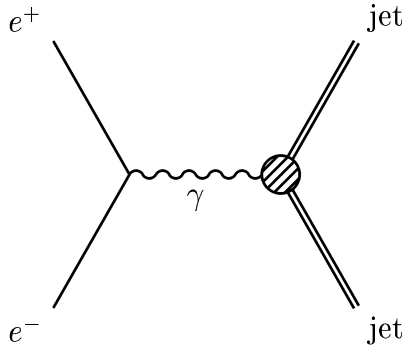


Figure 2.4: Electron-positron collision.

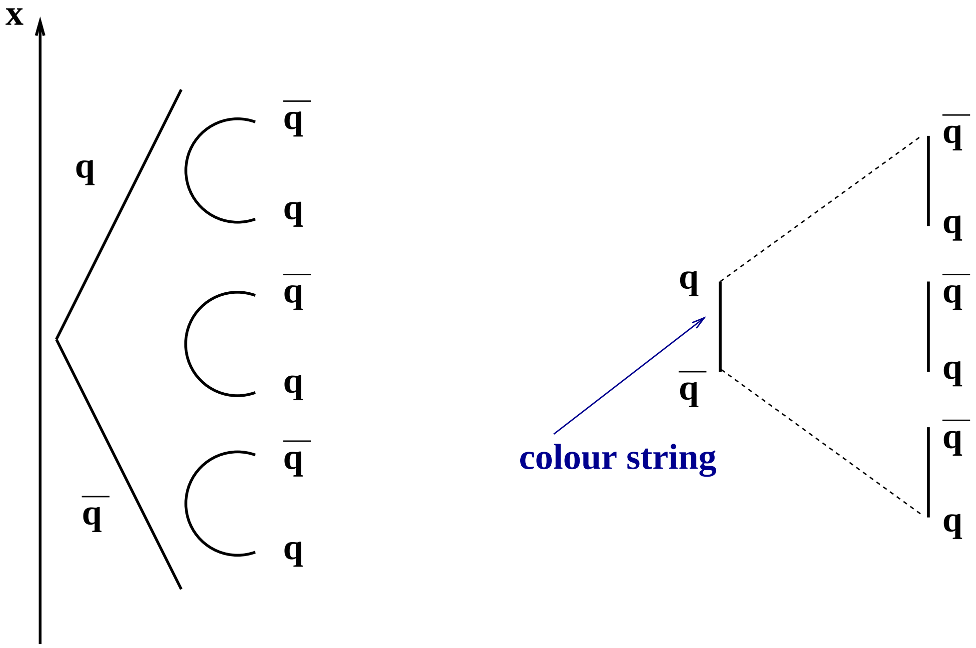


Figure 2.5: Drawing of the independent (left) and string (right) fragmentation.

There are mainly three models of fragmentation: independent and string fragmentation as well as the Herwig Cluster Model. The first two are explained in the following: In the model of independent fragmentation both quarks fragmentate independently into hadrons through the creation of $q\bar{q}$ pairs from the vacuum. The other model is the string fragmentation describing this process by the analogy of a rubber band representing the strong interaction. At a certain distance, the band disrupts and a new quark or antiquark is created at the band's end points. More and more $q\bar{q}$ pairs are produced until their kinetic energy has degraded into clusters of quarks and gluons, each of which has zero net colour. An illustration of both models is given in figure 2.5.

2.5 Strategy of this Search for New Physics

Unlike analyses searching for a certain signal in specific channels predicted by theoretical models, the approach here conists of a general search for deviations from the Standard Model prediction in final states with at least one electron (e) or at least one muon (μ) as well as additional jets (j). The analysis is based on a proper and unambiguous identification of these physics objects: e, μ and j. There is no additional criterion to select final state configurations. All events are classified into exclusive and inclusive event classes. An example for an exclusive event class is $1e\ 1\mu$ where exactly one electron and one muon satisfy the selection criteria but no jet. Inclusive means, that at least one physics object passed the selection, e.g. in $1\mu\ 2j\ +X$ at least one muon and at least 2 jets were found. The exclusive event classes provide a clear separation of the final states and an unambiguous statistical interpretation of deviations, whereas the inclusive event classes are mainly used to check that all exclusive classes are generated [5] and for the estimation of the QCD-background from data, which is described in detail in section 5.3. In a first step the total event numbers per event class in data and Monte Carlo samples (i.e. the Standard Model prediction) are compared. Then, the distributions of the scalar sum of transverse momenta $\sum p_T = \sum_i \sqrt{p_{x,i}^2 + p_{y,i}^2}$ are studied. This kinematical quantity is chosen since it is expected to have a large discovery potential [10]. New physics is assumed to appear at high energies, so a first sign may be visible as an excess or a deficit in one of these distributions (cf. chapter 7). Finally, a statistical algorithm is used to study the distributions systematically (cf. chapter 8). The algorithm locates the region of greatest deviation from the SM prediction within each distribution. By generating hypothetical data histograms, i.e. repeating the experiment many times, the probability of occurance of such a deviation is derived which in turn determines quantitatively the level of agreement between the SM prediction and the data.

---

[5] The total number of events in the exclusive 1e event class, for example, must be the same as the sum over all entries in each exclusive event class with electrons.

Chapter 3

The Tevatron and the DØ Experiment

The Tevatron and its two collider experiments, DØ and CDF, are located at the Fermi National Accelerator Laboratory (Fermilab) near Chicago, USA. Fermilab, named in honor of 1938 Nobel Prize winner Enrico Fermi, was founded in November 1967 and is operated by the U.S. Department of Energy. In 1977 the discovery of the bottom quark and in 1995 the first observation of the top quark as well as the verificaion of the existence of the tau neutrino in 2000, are the most important results of the research at Fermilab.

Figure 3.1 shows an aerial view of the 27.5 km^2 large site. The Tevatron is the upper ring, the DØ detector is marked with a blue dot, the CDF detector is marked in orange. During the first period of data taking in the so called Run I between 1991 and 1996 the Tevatron was operated at a centre of mass energy of $\sqrt{s} = 1.8\ TeV$. An integrated luminosity of approximately 130 fb^{-1} was recorded by each detector. The second phase of data taking, the so-called Run II, started in 2001 after a major upgrade of the accelerator and the two experiments. The physics motivation for the upgrade is to study low cross sections, hight-p_T processes such as top and W/Z and to search for new phenomena. The centre of mass energy has been increased to $\sqrt{s} = 1.96\ TeV$. During Run II, up to $4-9\ fb^{-1}$ data are expected until 2009 depending on the accelerator performance.

Subsequently the accelerator chain and the DØ detector are described.

Figure 3.1: Aerial view of the Fermilab and Tevatron with its two collider experiments DØ and CDF [14].

3.1 The Tevatron and its Accelerator Chain

The Tevatron was one of the first superconducting synchrotrons when it was build in 1983 and it is the worlds most energetic collider until the start of the Large Hadron Collider (LHC) at CERN in 2007. The main ring has a circumference of 6.4 km, where the protons (moving clockwise) collide with the antiprotons, moving in the opposite direction in the synchrotron. Due to the opposite charge of proton and antiproton, both particles can be accelerated in one ring only. A typical collision experiment takes place in three steps. First, the initial production and injection of the particles, followed by successive acceleration and finally the collison at defined interaction points, as DØ or CDF in this case. A schematic illustration is given in figure 3.2, more detailed information is provided in [12].

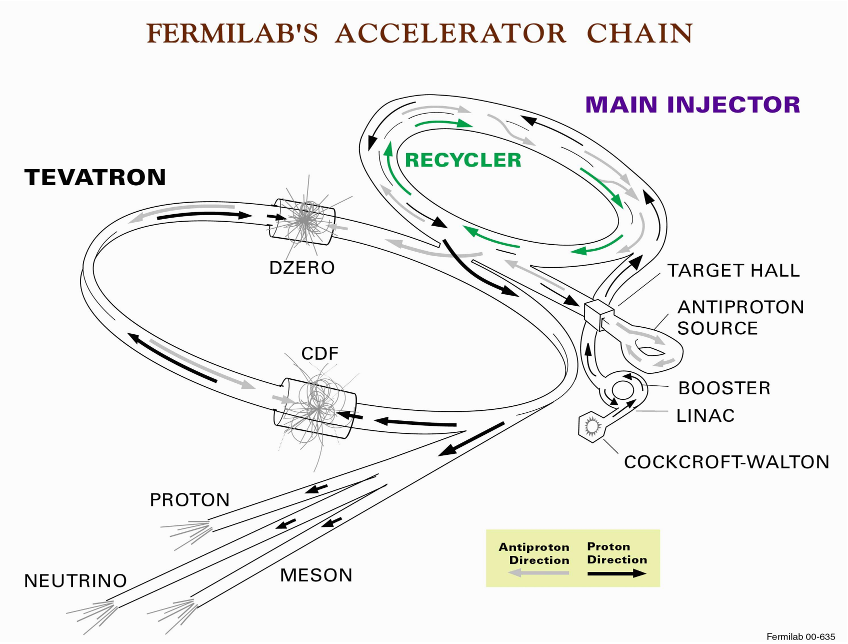


Figure 3.2: Schematic drawing of the Tevatron accelerator complex [14].

The initial production of the protons starts with hydrogen atoms, which are ionised in several steps to negatively charged ions (H^{-1}) and then pass through a 750 keV Cockcroft-Walton accelerator, where they are bunched and subsequently fed into a linear accelerator (LINAC) of approximately 150 m length. Being accelerated to 400 MeV by the LINAC, the hydrogen ions lose both electrons during their passage through thin graphite foil strips. This leaves only the protons that are injected into the next level, the Booster. The Booster is a synchrotron of 475 m length and accelerates the protons to $\approx$ 8 GeV, so that they can be fed into the Main Injector. The Main Injector is a synchrotron of 3 km circumference and delivers protons at an energy of 150 GeV to the Tevatron.

Some of the protons from the Main Injector are used for the production of antiprotons. Bunches with 120 GeV hit a nickel/copper target. On average, around 50,000 protons are necessary to produce one antiproton emerging with a required energy of $\approx$ 8 GeV. About 10^{11} antiprotons are produced within an hour. Before being injected into the Main Injector, the antiprotons are focused from the target particle shower, de-bunched, stochastically cooled and stored in the Accumulator. Protons and antiprotons move in opposite directions in the Main Injector and are finally injected into the Tevatron ring, where they are accelerated up to 980 GeV each, resulting in a centre of mass energy of $\sqrt{s} = 1.96\ TeV$.

To deflect the particles and to keep them focused on their circular path, a magnetic field of 4.2 T is necessary. This is achieved by superconducting dipole and quadrupole magnets. The cryogenic cooling system provides a temperature of 4.2 K, essential for superconductivity.

A proton bunch consists of $\mathcal{O}(10^{11})$ protons, an antiproton bunch consists of $\mathcal{O}(10^{10})$ antiprotons. Protons and antiprotons circulate in three trains of twelve bunches each, making 36×36 bunches in the ring at the same time. The trains are sperated by 7 ns, the bunches themselves by 396 µs. The ultimate goal for Run II is to reach an instantaneous luminosity of $2 \cdot 10^{32}\ cm^{-2}s^{-1}$. Due to the low efficiency in the number of antiprotons that are finally delivered to the Tevatron the presently achieved maxium luminosity is a factor of two smaller. In every bunch crossing there are 2.3 interactions on average.

3.2 The DØ Detector

The DØ detector was build in 1992 and has been upgraded for Run II. It is a multipurpose detector, covering an almost 4π solid angle and, like all modern detectors in high energy physics, it is composed of different layers each of which has its specific purpose (see figure 3.3). The first part is the tracking detector where vertices and tracks of charged particles are measured. This layer is followed by the electromagnetic calorimeter where electrons and photons deposite all of their energy while creating certain shower shapes. Hadronic jets on the other hand deposite almost all their energy in the hadronic calorimeter. The minimal ionising muons go straight through the calorimeter and give hits in the outermost muon detector. In order to identify particles, it is necessary to deflect them in a magentic field which gives conclusions on charge and momentum of a particle. The DØ detector consists of two magnets: a solenoid magnet providing a magnetic field of 2 T in the tracking system and an additional toroid magnet providing a magnetic field of 1.8 T in the muon system. Hence, a separate momentum measurement can be obtained in the muon system. In case of the DØ detector, the design of all systems is optimised to allow a precise measurement of particles emerging from the interaction point, to determine their properties as well as to reconstruct secondary vertices of b-jets, which are important for top-physics, for example. An extensive description of all components is given in [13] and concerning the upgrade for Run II in [15]. Figure 3.4 shows a side view of the detector which can be classified into three parts: the tracking system, the calorimeter and the muon system. The detector is 20 m long, around 11 m high and wide, and weighs approximately 5000 tons.

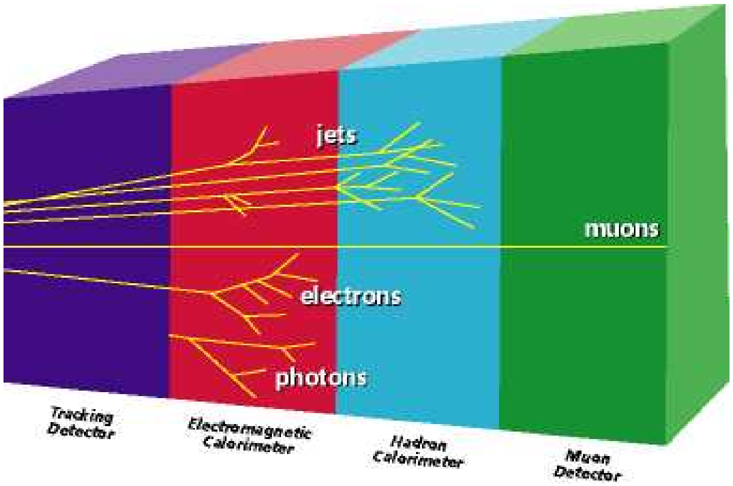


Figure 3.3: The principle construction of a modern detector in high energy physics [14].

In the following, the coordinate system is defined as right-handed. The z-axis coincides with the proton beam direction, the x-axis is pointing inwards to the middle of the Tevatron ring and the y-axis is pointing upwards. The polar and azimuthal angles are θ and ϕ, respectively, where θ is related to the pseudo-rapidity η:

$$\eta = -\ln(\tan\frac{\theta}{2}). \tag{3.1}$$

If the origin of the coordinate system is defined to be at the interaction point of an event, η is called *physics*-η, otherwise, if the coordinates are relative to the centre of the detector, η is called *detector*-η. In reality, the collider interaction point is spread over a distance of 25 cm along the z-axis.

3.2.1 The Tracking System

A major element of the upgrade was the replacement of the inner tracking systems. This was required because of the expected radiation damage[1] to those components during Run II as well as to improve the physics capabilities of the detector, e.g. precise

---

[1] In its current design, the Run II system resists a radiation dose of $\approx 1\ Mrad$, corresponding to an integrated luminosity of about 2 fb^{-1}. Another upgrade, Run IIb, will be necessary in the near future and is currently under review.

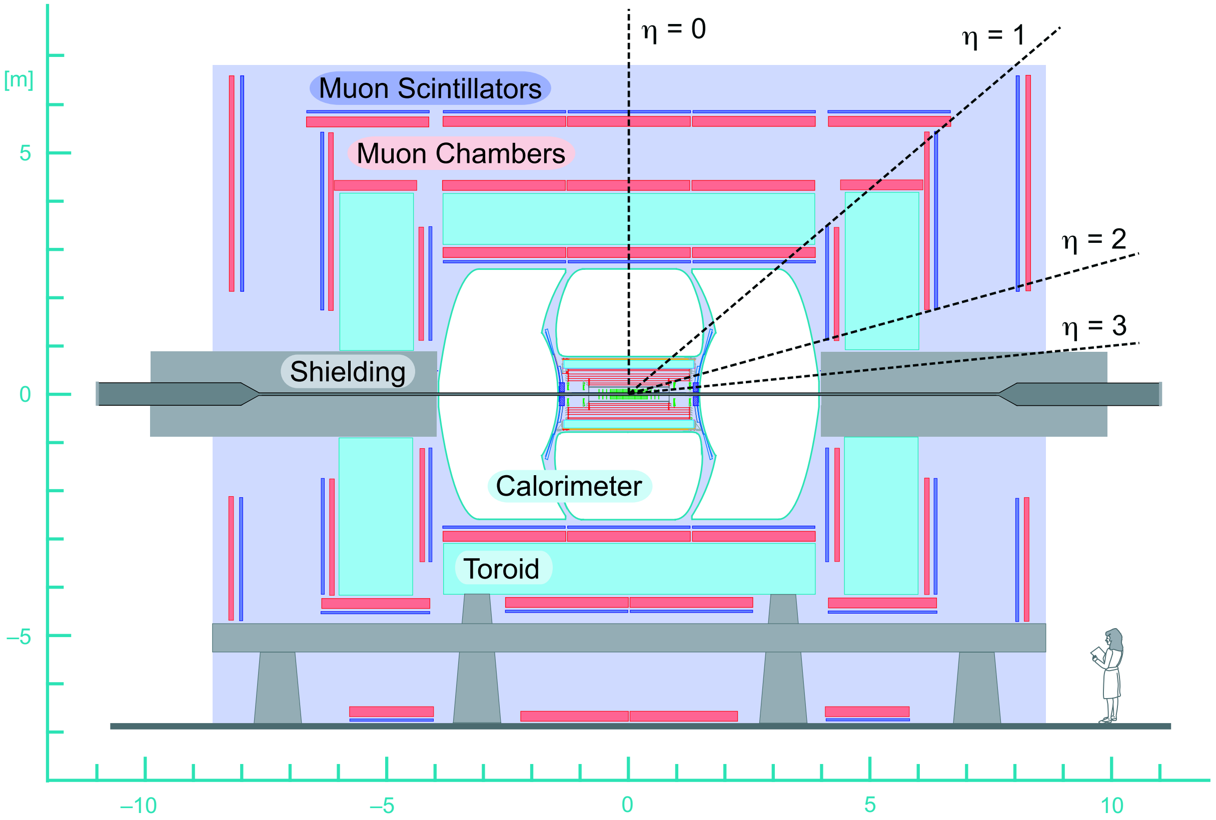


Figure 3.4: The DØ detector. It consists of a central tracking system, the calorimeter and the muon system [16].

tracking over a large range in pseudorapidity ($\eta \approx \pm 2$). The silicon microvertex tracker (SMT), nearest to the beam pipe and followed by the central fibre tracker (CFT) are both located within a 2 T superconducting solenoidal magnet. Magnet and cryostat are $\approx$ 0.9 radiation lengths thick [18]. SMT and CFT together are designed to provide a momentum resolution of $\Delta p_T/p_T^2 = 0.002\ GeV^{-1}$. A schematic view of the tracking system including the solenoid is given in figure 3.5. As one can see a precise tracking is limited in η by the dimensions of the CFT.

The **SMT** has 800,000 indiviual strips, around 793,000 read-out channels and is built of barrels and disks. There are six barrels in longitudinal direction each consisting of four detector layers, arranged axially around the beam pipe. Each barrel segment is 12 cm long and separated by 8 mm. The system is interspersed with twelve disks (F-disks) placed in the 8 mm gaps as well as at each end of the barrels. Two more disks of larger diameter (H-disks) at $z = \pm 110\ cm$ and $\pm 120\ cm$ complete the system. The structure is illustrated in figure 3.6. The spatial resolution perpendicular to the z-axis is $\approx$ 10 μm, while it is about 100 μm along the z-axis [17]. Depending on the position of the primary

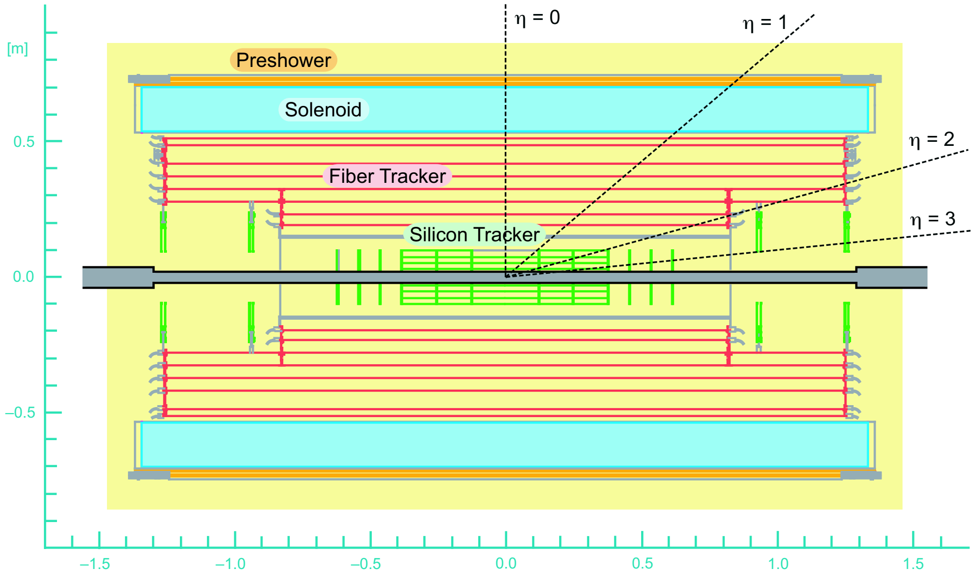


Figure 3.5: The DØ tracking system in a 2d illustration [16].

vertex and the direction of the particle, up to 20 hits can be registered by the SMT.

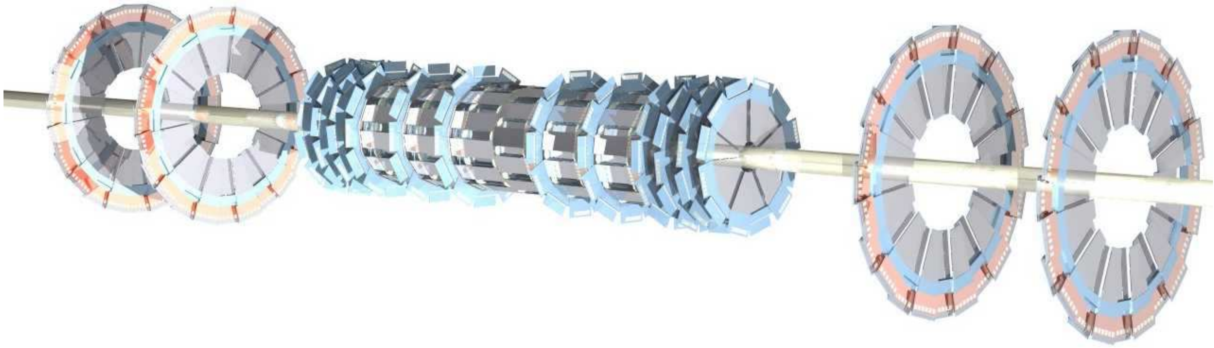


Figure 3.6: The silicon vertex detector with its arrangement of barrels and disks [16].

The **central fibre tracker** is the second part of the tracking system and encloses the SMT. A total of about 74,000 fibres, 0.8 mm in diameter, are mounted on eight concentric cylinders at different radii in the range of $19.5\ cm \leq r \leq 51.5\ cm$. Each cylinder is composed of two fiber doublets aligned parallel to the z-axis where one is an axial doublet for r and ϕ measurement and the other is a stereo doublet tilted by ±2° to provide stereo information of the tracks along z, resulting in an effective measurement of θ (η) and ϕ. Optical waveguides connect the fibers to visible light photon counters

(VLPC). These are small silicon devices which convert the light into electrical pulses. When operated at a temperature between 6 K and 15 K they have a quantum efficiency of over 80% and a gain of 20,000 to 50,000. The CFT covers the pseudo-rapidity range of $|\eta| < 2$ and has a spatial resolution per doublet of $\approx 100\ \mu m$. A maximum of 16 hits can be detected by the CFT.

3.2.2 The Central and the Forward Preshower Detectors

To improve the electron identification, triggering and to determine the energy loss in the solenoid two additional tracking detectors are located just outside the solenoid. Both are composed of a layer of lead absorber (1.1 X_0 thick for the central and 2 X_0 thick for the forward detector) sandwiched between several layers of scintillating strips which are arranged in axial and stereo ($\pm 23^\circ$) views. The central preshower detector (CPS) covers a pseudo-rapidity range of $|\eta| < 1.2$ while the forward preshower detectors (FPS) cover the range $1.4 < |\eta| < 2.5$ with one detector mounted on each of the end cap calorimeter (EC) cryostats (cf. section 3.2.3).

3.2.3 The Calorimeter

The liquid argon calorimeter is the second main system of the detector, located behind the tracking system and solenoid. It is basically unchanged for Run II. Only the read-out electronics have been upgraded to accomodate the higher rates and shorter bunch spacing [19]. In figure 3.7 a three dimensional view of the calorimeter system is given. A complete hermetic coverage up to $|\eta| \approx 4$ is achieved by three sections: one central calorimeter (CC) covering $|\eta| < 1.2$ and two forward/backward endcap calorimeters (EC) for $1.4 < |\eta| < 4$.

When electrons, photons and hadrons travel through the calorimeter modules they interact with the material and produce a cascade of many more particles of lower energy, called a shower. Many of these particles are electrically charged and cause ionisation in the argon. The ionisation, in turn, collects on metal plates and gives an electrical signal that can be read out. Since the amount of charge deposited in the calorimeter in a single shower and the depth to which the shower develops are proportional to the total energy of the particle that initiated the shower, the calorimeter can measure energies for both charged and neutral particles.

The main properties of argon are a very high number of electron-ion pairs that can be produced ($5 \cdot 10^4$ pairs per MeV compared to, for instance, a scintillator where 10^4

primary photons per MeV are produced) and its good energy resolution. Furthermore, it provides excellent homogeneity and hermiticity and it is radiation hard. On the other hand, this technique requires a cryostat for cooling, is difficult to repair and has a slow drift time of about 450 ns.

The read-out cells are composed of an alternating sandwich of signal boards and absorber material separated by a 2.3 mm liquid argon gap. The absorber layers mainly consist of depleted uranium due to its high density and compensating e/π response. Between absorber and signal board a voltage of 2 kV is applied.

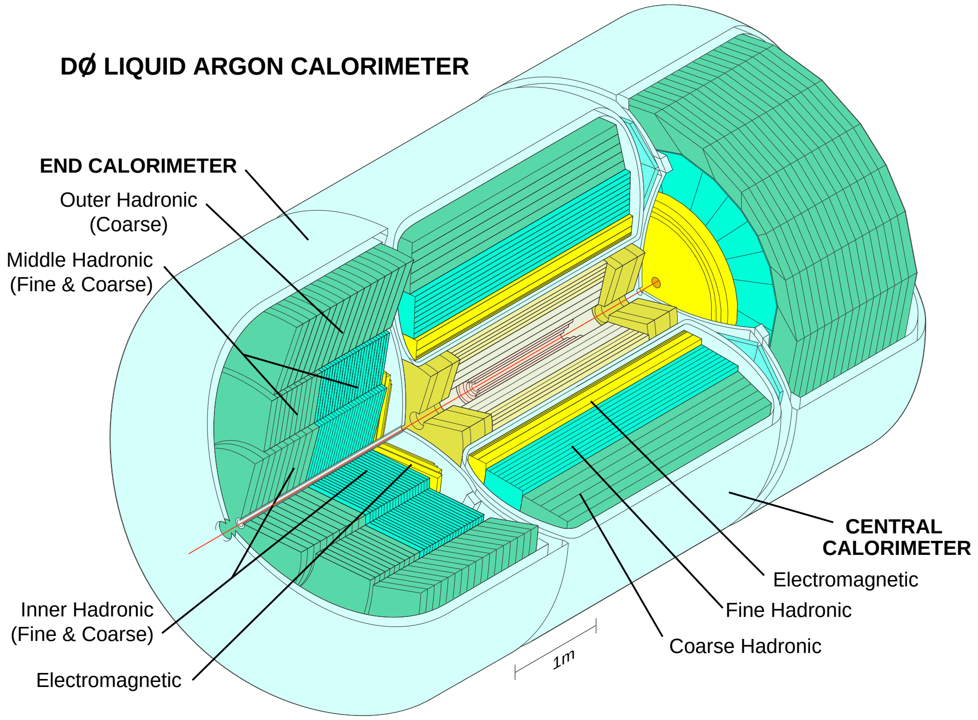


Figure 3.7: Outline of the DØ liquid argon calorimeter [16].

The read-out layers are arranged longitudinally into electromagnetic (EM), fine hadronic (FH) and coarse hadronic (CH) sections and are segmented into cells of 0.1 × 0.1 in $\Delta\eta \times \Delta\phi$. The third EM layer, though, is even finer segmented into 0.05×0.05 cells since this region corresponds to the electron shower maximum. The total radiation length for all four EM layers at $|\eta| = 0$ amounts to 20.6 with a thickness of 25 cm. The nuclear absorption lengths for the EM module amounts to 0.76 λ_A. Hadrons from jets deposite most of their energy in the FH and CH part, which are combined at $|\eta| = 0$ around 135 cm thick corresponding to 6.4 absorbtion lengths. Table 3.1 summarises information about radiation and absorption lengths in the different sections and layers of the calorimeter.

Layer	CC	EC
EM 1,2,3,4 d = 25 cm	X_0: 2, 2, 6.8, 9.8 3 mm Ur	X_0: 0.3, 3, 8, 9 4 mm Ur
FH 1,2,3,(4) d = 75 cm	λ_A: 1.3, 1.0, 0.9 6 mm Ur	λ_A: 1.3, 1.2, 1.2, 1.2 6 mm Ur
CH 1,(2,3) d = 60 cm	λ_A: 3 46.5 mm Cu	λ_A: 3, 3, 3 46.5 mm Fe

Table 3.1: Radiation length (X_0) for read-out layers in the electromagnetic (EM) and absorbtion length (λ_A) for the fine (FH) and the coarse (CH) hadronic sections of the central (CC) and end cap (EC) calorimeters as well as absorber material and total thickness of sections (d) at $|\eta| = 0$.

A calibration has been done so that measured pulses from the read-out cells can be related to physical energies. The energy resolutions for both the electromagnetic and hadronic section can be parameterised as

$$\left(\frac{\Delta E}{E}\right)^2 = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2} \tag{3.2}$$

where the terms have the following meaning:

- The systematic term C includes calibration errors.
- The statistical term S is due to sampling fluctuations
- The noise term N accounts for residual radioactivity from the uranium in the calorimeter.

Table 3.2 lists the values for the three terms obtained from test-beam measurements (for pions) [13] and recent studies from the DØ Electron and Photon Identification Group (for electrons) [20].

Particle	C	S	N
e	0.004	0.23 $\sqrt{GeV}$	0.20 GeV
π	0.032	0.41 $\sqrt{GeV}$	1.28 GeV

Table 3.2: Values for calorimeter resolution for electrons (e), obtained from Monte Carlo simulations, and pions (π), obtained from test beam measurements.

In most cases, the resolution of the hadronic calorimeter will be worse than the value measured for pions as it depends on the particle content of the hadronic shower. With an average fraction of electromagnetic response to hadronic response varying from 1.11 for 10 GeV to 1.04 for 150 GeV the calorimeter is close to being compensating.

Also it was found that the energy measurement in the EM calorimeter is less reliable in certain regions of the calorimeter. A particle is defined to be within the fiducial region, where the energy measurement is very good, if it is outside the transition regions, $\phi_{crack} \pm 0.02$, of the calorimeter modules and if it hits the third EM layer within $|z| < 115\ cm$. Since there are 32 calorimeter modules, the ϕ-cracks represent around 20% of the ϕ-space.

In the intermediate region $1.1 < |\eta| < 1.4$ the Intercryostat Detector (ICD) compensates the dead material and improves the measurement of the jet energy and the missing transverse energy (MET). It consists of a single layer of 384 scintillating tiles as active elements, mounted on both end cryostats. Wavelength-shifting fibres lead the detector signals to photo-tubes.

3.2.4 The Muon System

An outline of the DØ muon system is given in figure 3.8. This is the outmost part of the detector and it consists of two separate systems. The wide angle muon spectrometer (WAMUS) covers the region with $|\eta| < 1$ and the forward angle muon spectrometer (FAMUS), new for Run II, covers the region $1 < |\eta| < 2$. The upgrade for Run II also included a new shielding for the muon system to reduce background mainly from scattered proton and antiproton fragments.

The WAMUS consists of three layers (A, B and C) with a total of 94 proportional drift tubes (PDTs) built for Run I. The A layer lies between the calorimeter cryostat and the 1.8 Tesla muon toroid magnet. The B and C layers, separated by 1.5 m, are located outside the toroid. This design has been chosen to provide muon identification and a confirming momentum measurement independent of the central tracking. The A layer is made of 4 PDT decks in the top and, due to space constraints, of only three decks in the bottom part. The B and C layers have three decks each. Also, layers of scintillator are placed on top and upper sides of the system to help reject muons from cosmic rays. For Run II a new layer of scintillators, called $A\phi$ counters, was included between the A layer and the calorimeter. Those counters with a ϕ-segmentation of 4.5° are used for muon triggering, as well as for the rejection of out-of-time scattered particles and cosmic muons. The PDTs are rectangular aluminum tubes with 5.7 cm by 10 cm cells and provide a drift distance resolution of about 1 mm. The gas mixture used consists of 80% Ar, 10% CF_4 and 10% CH_4 which gives a maximum drift time of 450 ns. The

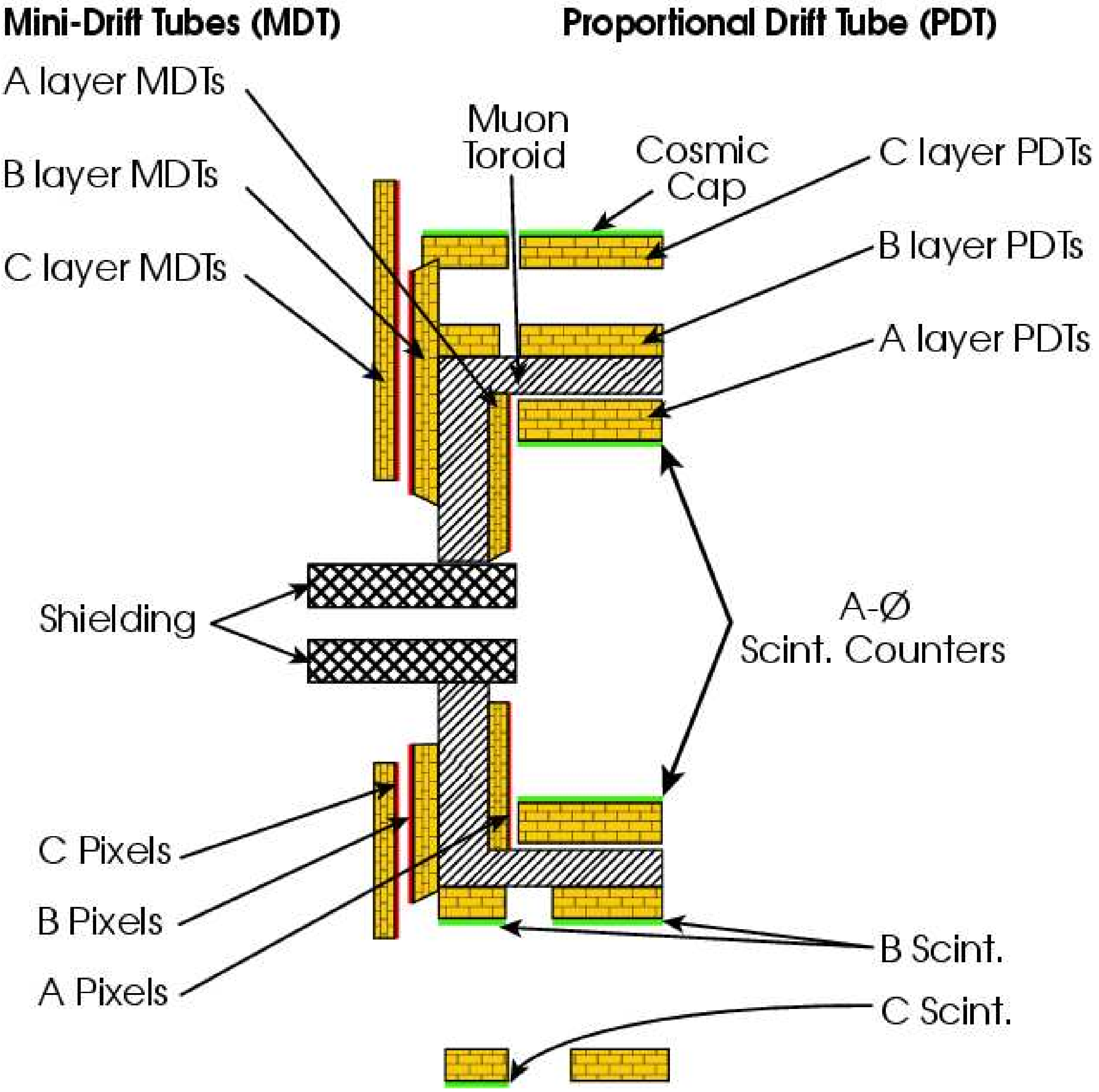


Figure 3.8: The DØ muon system. Shown are the central and forward region with its detector components [21].

scintillators have a time resolution of $\sim 5\ ns$.

The FAMUS has been completly replaced for Run II. Analogously to the central muon system, the forward system consists of three layers (A, B and C) with the A layer placed inside the toroid magnet and the B, C layers outside. The A layer is made of four decks of mini drift tubes (MDTs) while the B and C layer consist of three MDT decks each.

The MDTs are more radiation hard than the older PDTs, which showed clear aging effects in the forward region and thus needed to be replaced. They are composed of eight cells of 1 *cm* × 1 *cm* cross section and are made of aluminium extruded combs and plastic sleeves. The gas mixture is the same as for the WAMUS but due to the smaller size of the cells, the maximum drift time of the MDTs is only 60 *ns*. The distance resolution is of the order 0.7 *mm*. In addition to the drift tubes, each of the three layers also consists of scintillator pixel arrays which cover approximately the same η-region. These counters are segmented in ϕ of 4.5° and, different to the WAMUS, in η of 0.1. The typical size is 20 *cm* × 30 *cm* and they are primarily used for triggering.

The momentum resolution is different for muons in the central and in the forward region. Considering only $|\eta| < 1$ the resolution from the PDTs is $\sim 30\%$ for muons with $p_T = 100\ GeV/c$, which can be improved to $\sim 15\%$ when the muon track is matched to a track identified by the central tracking system. In general, the momentum resolution is better in the central tracking system than in the local muon system. The resolution is limited at low momentum by multiple scattering and is significantly worse at high momentum as the momentum is derived by measuring the sagitta of the track. The sagitta s of a curved track segment between two points (A,B) is defined as the maximum deviation of the curve from the straight line connecting these two points. If the strength of the magnetic field B (in Tesla) is assumed to be constant and perpendicular to the direction of the muon, the momentum p (in GeV) of the muon can be expressed as

$$p = 0.3\frac{BL^2}{8s} \quad , \tag{3.3}$$

where L is the length of the track between the two points A and B; L and s are given in metres. For example, a muon with $p = 500\ GeV$ will have a sagitta of 1 *mm* (for $L = 3\ m$ and $B = 1.8\ T$). Hence, the momentum resolution is limited by the resolution of the sagitta measurement.

3.2.5 The Trigger System and Luminosity Measurement

At the Tevatron, collisions occur at the interaction points every 396 *ns* which translates into a rate of 2.5 *MHz*. Since not all events are interesting from a physics point of view and data acquisition can only handle a rate of 50 *Hz*, a three-level trigger system is used for reduction. In order to accomodate the high luminosities of Run II the DØ upgrade included a significant enhancement of the trigger system as well [15]. In the following the three levels (L1, L2, L3) are explained. Figure 3.9 illustrates the operating mode of level 1 and 2.

The **level 1** trigger is a pure hardware trigger which requires simple criteria to be satisfied at detector level. This includes the central fibre tracker and muon scintillators,

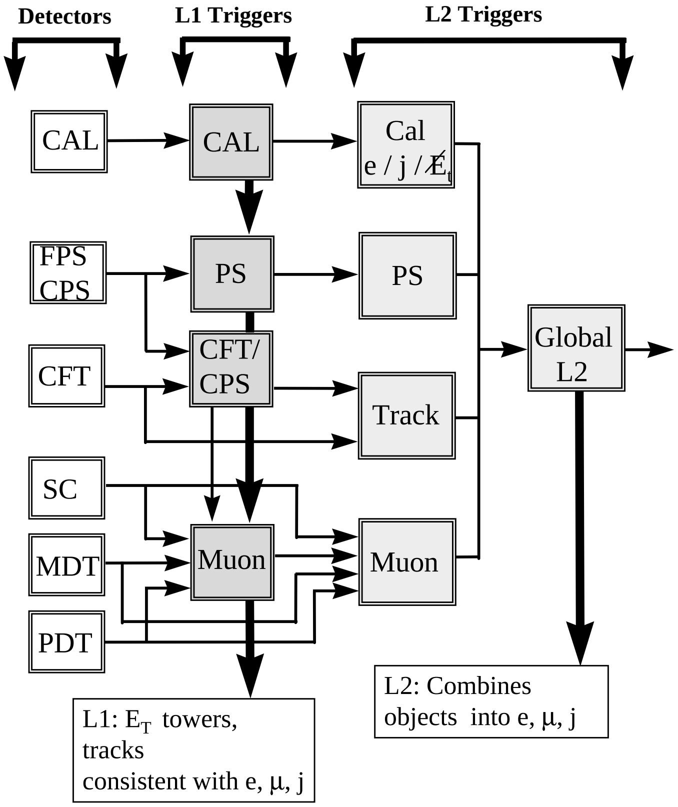


Figure 3.9: A simplified view of the first two levels of the trigger system [23].

information from preshower detectors and a threshold in calorimeter towers of size $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$. The level 1 triggers report their findings to the level 1 framework which has sufficient memory to ensure deadtimeless operation for even 7 MHz [2]. If the framework issues an accept the event is passed to level 2 just 4.2 μs after the beam crossing. The rate is reduced to 10 kHz.

---

[2] It was planned to go to shorter bunch crossing times of 132 ns in Run IIb. Both experiments CDF and DØ argued against this plan recently as the average number of interactions per bunch crossing would increase up to ten or more.

At **level 2** decisions are now taken by a mixture of hardware and software information. Data from each level 1 trigger is prepared in a preprocessor stage where a first reconstruction and identification of physics objects is done. Within 100 μs the input rate is further reduced by a factor of 10 to 1 kHz.

Level 3 uses software only, run on a computer farm. The algorithms access all event information and perform a full reconstruction. The output rate of filtered events is reduced to 50 Hz which is written to tape. Especially for level 3, the software is continuously improved resulting in new trigger lists. All information about triggers and corresponding runs can be found in [22].

Figure 3.10 shows the development of the interaction rate, i.e. luminosity, and the effects of the triggers during a store. The level 3 rate (red) is always just under 50 Hz as required. Level 1 and 2 are marked in black and blue, respectively, and need to fulfill certain accept rates which change from time to time. The luminosity (dashed pink line) is high at the beginning of a store and decreases exponentially with time. Hence another way to keep the writing rate at the maximum of 50 Hz is to introduce prescale factors. They work in addition to the trigger system and are set after level 1 [3]. A prescale of n transmits only every n^{th} triggered event to level 2. The peaks at level 1 in figure 3.10 indicate the setting of new prescales. With every new prescale a new run starts as well.

Luminosity

At DØ the luminosity is measured with monitors made of plastic scintillator arrays located in front of each end cap calorimeter cryostat, covering a range of $2.7 < |\eta| < 4.4$. They have a time-of-flight resolution of $\approx$ 200 ps. To discriminate between collisions (luminosity) and beam losses (halo) the time difference between signals from both luminosity monitors, north and south, is measured. Protons travel clockwise around the Tevatron, so that the rests of a proton (i.e. proton halo) will first hit the north luminosity monitors and then, around 9 ns later, are verfied by the south monitors. Antiprotons travel counter-clockwise, hitting first the south and then the north detector. On the other hand, particles produced in a $p\bar{p}$ collision should strike the north and south monitors at approximately the same time. Hence, the difference between both monitors gives a determination of the z-position of the interaction. The instantaneous luminosity is then calculated from the monitor countings (N) using elastic and inelastic cross sections ($\sigma_{p\bar{p}}$), the monitors acceptance (A) and efficiency (ϵ_{LM}):

---

[3] Prescales could be set after each trigger level in principle, though.

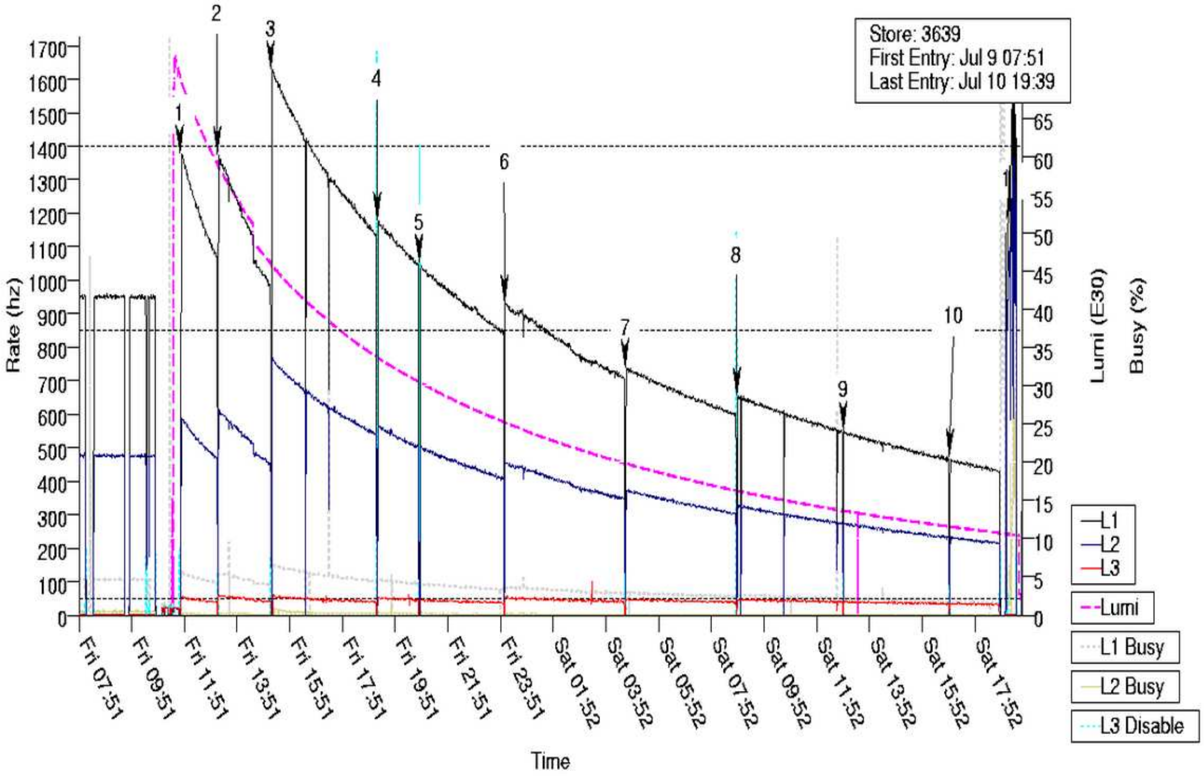


Figure 3.10: Luminosity and trigger rates during a store [24].

$$\mathcal{L}_{inst} = \frac{1}{\sigma_{p\bar{p}} \cdot A \cdot \epsilon_{LM}} \frac{dN}{dt} \tag{3.4}$$

Elastic processes are $p\bar{p} \rightarrow p\bar{p}$ while inelastic processes are divided into single diffractive $p\bar{p} \rightarrow pX$, double diffractive $p\bar{p} \rightarrow X_1 X_2$ and hard scattering $p\bar{p} \rightarrow X$.
The cross section $\sigma_{p\bar{p}}$ was measured by other experiments to be (60.7 ± 2.4) mb [25]. The luminosity is measured by dividing data-taking runs in 60 s blocks of integrated luminosity (LBN) with

$$\mathcal{L}_{int} = \int_{t_1}^{t_2} \mathcal{L}_{inst} \quad dt \quad . \tag{3.5}$$

The luminosity used for the analysis is finally corrected for:

- deadtime introduced by the trigger framework

- prescales
- loss of data due to data acquisition or level 3 errors and crashes
- loss of data due to fatal errors in reconstruction, streaming or in storage media

The estimated overall error on the DØ luminosity is 6.5% [27].

Data Management

Monitoring of data quality is done during data acquisition as well as in analyses. Results are summarized in an online database [28] marking every run, according to each detector system, as:

good	A run without problems, which can likely be used for physics analysis.
reasonable	A run with minor problems, which needs to be treated with care.
bad	This run should not be used due to mayor problems.
special	A run for experts only where special conditions have been applied.
unknown	This run has not been classified.

As explained, the events are written to tape with a rate of 50 Hz where each event needs about 250 kB of disk space. The events are then reconstructed with the latest update of reconstruction software and saved in the *thumbnail*-format (`.tmb`). This format reduces the data size by one magnitude. The user can access these data samples as well as Monte Carlo simulations via the data management system SAM (**S**equential **A**ccess via **M**etadata) [26]. Furthermore, physics groups provide preselected data samples and Monte Carlo simulations, which are available via SAM and from the local computer clusters at Fermilab.

Chapter 4

Data Selection and Object Classification

In this chapter the criteria for the reconstruction of the physics objects electrons, muons and jets are explained. Furthermore, the investigated datasets and the triggers used for this analysis are introduced.

4.1 Reconstruction

Almost all criteria for the physics objects are certified by the according object identification groups EM ID [29], Muon ID [30] and Jet ID [31]. Only for muons, further cuts are applied to reject misreconstructed muons and muons from cosmic rays. These cuts are not certified by the Muon ID group but recommended by other analysts within the collaboration [32].

4.1.1 Electrons

For electron identification the information from the central tracking detectors and the calorimeter is used. The electron interacts with the material of the electromagnetic calorimeter and deposits a large fraction of its energy in this part, creating a characteristic shower. Both, electrons and photons usually loose all their energy within the electromagnetic calorimeter. However, a small fraction may reach the hadron calorimeter.

First, an electromagnetic cluster is defined if the energy deposited within a cone of radius

$$R = \sqrt{\eta^2 + \phi^2} = 0.4 \tag{4.1}$$

exceeds a certain threshold. This analysis uses the *simple cone cluster algorithm* with the above mentioned size for data and Monte Carlo. The electron candidate is required to have an EM fraction *emf* of

$$emf = \frac{E_{0.4}^{ECal}}{E_{0.4}^{tot}} > 0.9 \quad , \tag{4.2}$$

with $E_{0.4}^{ECal}$ being the energy within a cluster of size $R = 0.4$ in the EM calorimeter and $E_{0.4}^{tot}$ is the energy in the complete calorimeter, i.e. EM plus hadron calorimeter.

The shower shape is also taken into account for electron reconstruction. To measure how consistent a certain shower shape is with the shape of a genuine electromagnetic shower (obtained from Monte Carlo simulations), a χ^2, which is also called H-matrix ($HMX7$), with seven correlated variables is attributed to the shower. These seven correlated variables are: the energy fractions of each of the four EM calorimeter layers, the total energy in the EM calorimeter, the vertex z-position and the transverse shower shape width in ϕ. The H-matrix is required to be

$$HMX7 < 20.0 \quad . \tag{4.3}$$

The last cut on the calorimeter information is the requirement of isolated clusters, meaning that most of the energy in a cluster around the candidate should be of electromagnetic origin:

$$iso = \frac{E_{0.4}^{tot} - E_{0.2}^{ECal}}{E_{0.2}^{ECAL}} = \frac{E_{0.4}^{tot}}{E_{0.2}^{ECAL}} - 1 < 0.15 \tag{4.4}$$

i.e. only 15% of the cluster energy are allowed to be deposited outside the electromagnetic part of the calorimeter. Here, a smaller size of the cluster in the EM calorimeter is taken ($R = 0.2$).

To suppress the overwhelming QCD background a further confirmation is needed. For the electron candidate an associated track is required. The track finding algorithms use information from the silicon microvertex (SMT) and central fibre tracker (CFT), requiring a certain number of hits[1], e.g. 8-16 for the CFT and at least 5 for the SMT, depending on the position. Finally, the χ^2-probability $P(\chi^2_{trackmatch})$ is calculated for the track to match the electron candidate found in the calorimeter cluster in ϕ, z and E_T/p_T:

$$P(\chi^2_{trackmatch}) > 0.01 \quad \text{with} \tag{4.5}$$

$$\chi^2_{trackmatch} = \left(\frac{\delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\delta z}{\sigma_z}\right)^2 + \left(\frac{E_T/p_T - 1}{\sigma_{E_T/p_T}}\right)^2 \quad . \tag{4.6}$$

---

[1] For the central region, a maximum of 16 hits in the CFT and 14 in the SMT is possible.

In equation 4.6 $\delta\phi$ denotes the difference between the ϕ coordinate of the third EM-calorimeter layer (EM3) and the ϕ coordinate of the extrapolated track in the EM3 layer. Similarly, δz is obtained. E_T is the transverse energy measured in the calorimeter and p_T is the transverse momentum of the track measured in the central tracking system. σ_ϕ, σ_z and σ_{E_T/p_T} are the resolutions of the three variables.
If a matching track is found, the electron's charge is derived from the curvature of the track, as well as the variables η and ϕ which are subsequently used to calculate the momentum:

$$\vec{P} = \begin{pmatrix} E \cdot \cos\phi \cdot \sin\theta \\ E \cdot \sin\phi \cdot \sin\theta \\ E \cdot \cos\theta \end{pmatrix} \quad , \tag{4.7}$$

where $\eta = -\ln(\tan\frac{\theta}{2})$ and E is the energy of the electron measured by the calorimeter.

Due to the fact that on average 2.3 interactions per collision take place, a very loose cut on the difference between the z_0 position of the electron and the z of the primary vertex is required

$$|z_0(e) - z_0(\text{vertex})| < 2 \text{ cm} \quad . \tag{4.8}$$

This ensures that the considered physics objects all come from the same vertex. The distribution of this difference in z is shown in figure 4.1.

Finally, the electron is required to be within a η-range of $|\eta| < 1.1$, i.e. taking into account the central calorimeter (CC) region only and especially excluding the intercryostat region. In the part of the analysis dealing with electron final states, the first electron must have a transverse momentum of $p_T > 30\ GeV$ and all others of $p_T > 15\ GeV$. On the other hand, if a muon is selected first, all electrons must satisfy $p_T > 15\ GeV$. This choice of p_T is taken due to the trigger efficiency (see section 4.2.1) and due to the assumption that new physics is expected in the high-p_T region.

4.1.2 Muons

Muons are reconstructed using information from two independent detector components: the muon system and the central tracking system. Each muon candidate in an event is required to satisfy certain quality criteria of both detector parts. A "local" track in the muon system is the basis of the muon identification. In this analysis, only muons with hits in all three layers of the muon system (A, B and C), called $nseg = 3$, and of at least "medium" quality are investigated. A muon is called medium if it has

- at least two hits in the A-layer drift tubes,

- at least one hit in the A-layer scintillator,
- at least two hits in the BC-layer drift tubes,
- at least one hit in the BC-layer scintillator.

Since the muon gives hits in two different detector systems there are two tracks found which can be associated with the muon: one from the local muon system plus one track in the central tracking system. The track matching algorithm first tries to match the track from the muon system to the central track if the muon track converges and vice versa otherwise. Since the resolution of the central tracking system is much better than the one of the muon system alone, only events with a matching central track are kept. Furthermore, the χ^2/dof of this global fit must be $\chi^2_{global}/4 < 10$.

In addition, the muons must meet two further requirements to ensure a good quality of the tracks. One criterion is a minimum number of hits in the silicon microvertex tracker (SMT) of at least three. The other one requires a $\chi^2_{track}/dof < 3.5$ for the fit of the central track. Here, the number of degrees of freedom depends on the number of hits registered by the tracking system. A reliable central track measurement is important since only the p_T measurement from the central track is used in this analysis. Muons

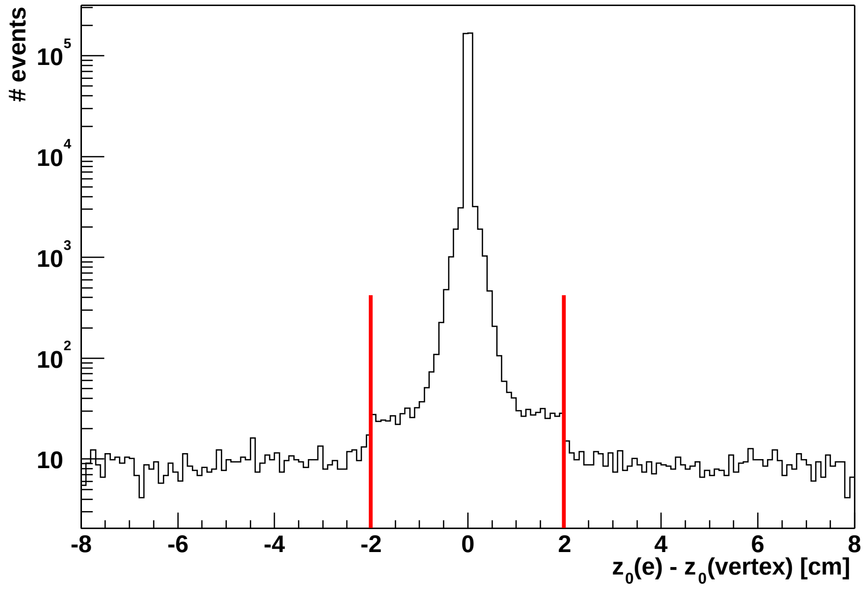


Figure 4.1: Distribution of the difference in z-position between electron (z_0(e)) and vertex ($z_0(vertex)$) from a $W \to e\nu$ Monte Carlo. Events within the two red lines are selected.

that fail these two cuts are poorly measured and tend to fake very high p_T muons. This cut has almost no influence in single boson production Monte Carlo samples but removes a lot of background in data. For example, muons from K decays tend to have tracks with a higher χ^2_{track}/dof. The remaining high p_T muons are believed to be genuine high p_T muons with a well measured track curvature q/p_T (q: electric charge) within the finite resolution of the central tracker. Figure 4.2 shows a two dimensional plot of χ^2_{track}/dof versus p_T of the muon in a $W \to \mu\nu$ Monte Carlo sample.

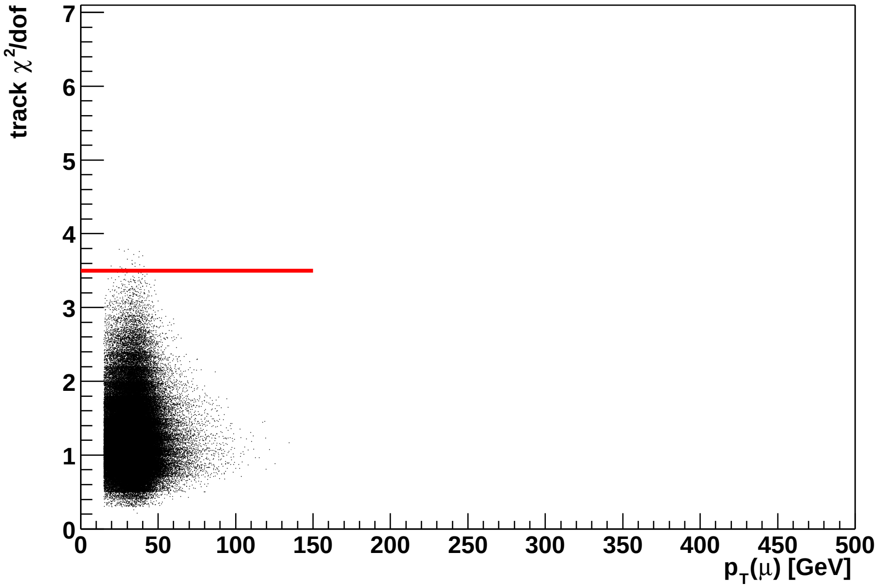


Figure 4.2: Distribution of the variable χ^2_{track}/dof versus p_T of the muon in a $W \to \mu\nu$ Monte Carlo sample. The red line indicates the cut value of 3.5.

Muons from cosmic rays are rejected by a cut on the scintillator counter times. For the A-layer scintillators, the times should be in the range of -10 $ns < t_A <$ 10 ns and for the BC layer, within -15 $ns < t_{BC} <$ 10 ns. The scintillator times are calibrated such that a muon from a proton-antiproton collision would reach them with $t = 0$ ns.

In order to get a clean sample of muons, it is important to use the isolation criteria to distinguish between muons coming from a leptonic W decay, for example, and muons from semileptonic decays in heavy quark jets. The latter tend to be non-isolated and have lower transverse momenta. The isolation criteria use information from the central tracker and the calorimeter. Muons satisfying the following two requirements are called isolated and are selected:

- The sum of the transverse momenta of all tracks, except the muon track itself, within a cone of radius $R = 0.5$ surrounding the muon is less than $0.6 \cdot p_T(\mu)$.

- The sum of the calorimeter transverse energies, considering electromagnetic and fine hadronic calorimeter layers, in a hollow cone with radius R between $0.1 < R < 0.4$ around the muon is less than $0.8 \cdot p_T(\mu)$. Since the muon is a minimal ionising particle, the lower bound for the size of the hollow cone is taken to be 0.1.

Finally, in analogy to the electron selection criteria, the z-position of the muon should not deviate more than 2 cm from the primary vertex z-position to ensure that all physics objects come from the same interaction point (cf. equation 4.8).

The range in η is restricted to $\eta < 1.4$ because of the geometrical acceptance of the central tracker. All muon transverse momenta are required to be above 15 GeV due to the trigger efficiency (cf. section 4.2.1).

4.1.3 Jets

Particle jets are reconstructed with information from the energy depositions in the DØ calorimeter cells using the *Run II Cone Algorithm* [33]. The cone algorithm proceeds in several steps: preclustering, clustering, addition of midpoints and merging or splitting.

First, *preclusters* are built by summing up the energy in all cells that share the same pseudo-rapidity and azimuthal angle. Subsequently they serve as seeds for the Run II Cone Algorithm. In this analysis the default value of $R_{cone} = 0.5$ for the cone size is taken. In the first stage, the algorithm defines so-called *proto-jets*, i.e. temporarily defined jets. Therefore, the algorithm loops over a list of p_T-ordered preclusters P_i and calculates the distance ΔR_{min} in $\eta \times \phi$ space between P_i and the closest proto-jet candidate $(PC)^2$. If it is close to the proto-jet, i.e. $\Delta R_{min} < \frac{1}{2} \cdot R_{cone}$, the algorithm goes to the next precluster in the list. Otherwise, P_i is considered as a seed for a proto-jet candidate PC. In the second stage (clustering), the algorithm tries to find a stable cone around the candidate in an iterative process. After forming a cone of radius R_{cone} around PC, all energy depositions within this cone are combined by adding their 4-momenta

$$p^{PC} = (E^{PC}, \vec{p}^{PC}) = \sum_i (E^i, p_x^i, p_y^i, p_z^i) \tag{4.9}$$

giving a new proto-jet candidate PC' which replaces PC. This process is then repeated unless a stopping condition, tested in the following order, is satisfied:

---

[2] If no proto-jet exists, P_i is directly considered as a seed for a proto-jet candidate PC.

1. If the transverse momentum of PC' is less than half the minimal p_T required for final jets, i.e. $p_T^{PC'} < \frac{1}{2} \cdot p_T^{min}$ with $p_T^{min} = 8\ GeV$, the candidate is discarded in order to keep only jets relevant for physics studies.

2. The proto-jet candidate is stable, i.e. $\Delta R(PC', PC) < 0.001$, and thus added to the list of proto-jets.

3. The number of iterations reaches a default limit of 50.

Furthermore, proto-jets that are close ($\Delta R < 0.5$) to an isolated electron with $p_T > 15\ GeV$ are discarded. Proto-jets are formed until the algorithm reaches the end of the list of preclusters.

So far, the algorithm is not infrared safe since soft radiation between two jets may cause a merging of them. To eliminate this sensitivity, the algorithm also searches for proto-jets in the vicinity of midpoints between any combination of two proto-jets satisfying

$$R_{cone} < \Delta R < 2 \cdot R_{cone} \quad . \tag{4.10}$$

Hence, a list of proto-jets from preclusters and midpoints is built, which is used in the final merging and splitting step. The proto-jets often share energy depositions, so the algorithm must avoid double counting and ensure that each energy deposition belongs to one jet at most. Two proto-jets are merged if the sum of p_T, which are shared, is greater than 50% of one of the proto-jet's total p_T. If the fraction of shared p_T is less than 50%, the proto-jets are splitted assigning each shared energy deposition to only one jet, namely the one to which the energy deposition is closest in $\eta \times \phi$ space. After this merging and splitting algorithm has been applied, the remaining objects are called jets and are used for physics analyses.

In order to remove different kind of fake jets, mainly due to noise in the calorimeter and electronics (called hot cells), the objects are required to pass the following *quality criteria*:

- The electromagnetic fraction emf must lie between 5% and 95% of the total energy deposited in the calorimeter, $0.05 < emf < 0.95$, to remove isolated electromagnetic particles.

- To remove jets clustered around noise cells in the coarse hadronic layer of the calorimeter a cut on the fraction of the jet energy deposited therein (CHF) is applied at $CHF < 0.4$.

- The ratio of the highest to the next-to-highest transverse energy cell, called $HotF$, must be smaller than 10. This prevents jets from being clustered of hot cells only.

- The number of calorimeter towers containing 90% of the jet energy, $n90$, is required to be greater than one to remove jets clustered around a single hot tower.
- Again, only the central region is considered requiring the jets to be within $|\eta| < 1.5$. The minimum transverse momentum for jets is $p_T > 30\ GeV$ in this analysis.

In data, an additional cut is applied to the energy of the jet seen by the level 1 trigger, which is almost not affected by noise, compared to the reconstructed transverse energy:

$$\frac{L1SET}{p_T^{reco} \cdot (1 - CHF)} > \begin{cases} 0.4 & \text{within} \quad |\eta| < 0.8 \quad \text{or} \quad |\eta| > 1.5 \\ 0.2 & \text{within} \quad 0.8 < |\eta| < 1.5 \end{cases} \tag{4.11}$$

The variable $L1SET$ is the scalar sum of the trigger towers transverse momenta and p_T^{reco} is the reconstructed transverse momentum, which is not corrected by the jet energy scale (see section 4.1.4). Since the level 1 trigger does not use information from the coarse hadronic section of the calorimeter, $L1SET$ has to be compared to the transverse momentum without this component. This was found to be the most powerful discriminant against calorimeter noise.

4.1.4 Jet Energy Scale Correction

The energy of the jets is corrected to account for non-linearities, dead material, noise and showering effects. The jet energy scale (JES) correction attempts to correct the jet energy for these effects in order to get back to the particle level energy E_{jet}^{PL} as it would have been before the particles interacted with the calorimeter. For this pupose the following expression is used:

$$E_{jet}^{PL} = \frac{E_{jet}^{CAL} - E_{offset}}{R \cdot S} \quad , \tag{4.12}$$

where E_{jet}^{CAL} is the measured jet energy by the calorimeter and E_{offset} is the energy offset due to the underlying event[3], energy pile-up[4], multiple interactions and noise from electronics and the uranium absorbers. R is the calorimeter response to a jet and S is the showering correction that compensates for the fraction of shower leakage outside the jet cone. The JES correction factor is given as the ratio of E_{jet}^{PL} to E_{jet}^{CAL}. Since Monte Carlo simulations do not simulate these effects completely, different JES correction are applied to data and Monte Carlo samples.

---

[3] The underlying event consists of the proton-antiproton remnants.

[4] Additional energy in the calorimeter from previous bunch crossings is called pile-up.

The jet energies are studied in γ+jet(s) events: The energy depositions of the photon in the calorimeter are similar to those from electrons and hence, can be calibrated using $Z \rightarrow ee$ events. This allows a very precise measurement of the photon-energy. Since the photon recoils from the jet, transverse momentum conservation then provides the correction of the jet energy. In this analysis the software package *JetCorr v5.3* [34] is used which provides corrections for data as well as for the Monte Carlo samples.

4.1.5 Missing Transverse Energy

Although not considered in this analysis directly, the missing transverse energy (MET) shall be explained for completeness. Neutrinos can not be detected directly, but their presence or that of any undetected object can be found from an energy imbalance in the transverse plane. The MET is calculated from the vector sum of the transverse energies of all cells with an energy above 0 MeV while excluding unclustered energy in the coarse hadronic section[5] except for those belonging to a jet. Further, the MET must be corrected for jet energy scale corrections applied to all jets, as well as for electron scale corrections applied to all electromagnetic objects, and finally for muons since they are minimal ionising particles which deposit little energy in the calorimeter.

4.2 Datasets

Data collected with the DØ detector during the period starting from April 2002 until April 2004 are investigated. For this period the total event count is around 640 million. All data are "skimmed" to provide smaller subsamples based on physics objects like electrons, muons and jets as well as on certain kinematical requirements which are relevant for the analysis. The Common Sample Group [35] at DØ, for instance, provides such skims for a broad range of analyses. For this particular search for new physics in final states with muons and electrons, two skims from the Common Samples Group are used: 1MUloose and EM1TRK. For the 1MUloose skim, events with at least one muon of loose type (cf. section 4.1.2) with $p_T > 8$ GeV are considered. Events with at least one electromagnetic object with $p_T > 8$ GeV and at least one track with $p_T > 5$ GeV pass the EM1TRK preselection criteria. These two skims consist of about 55 million events each (for the period mentioned above). They were reconstructed with the "DØ Reco" software versions p14.03.00/01/02, p14.05.00, p14.05.02 and p14.06.00 and saved in the

---

[5] The coarse hadronic layer has been partially removed from the MET calculation due to high levels of noise.

thumbnail-format (`.tmb`).

To be able to analyze the events with the ROOT-software [36], the thumbnails are first processed using the framework analysis package TopAnalyze [37], which converts the files to root-trees. At this stage, the size of the samples is further reduced by applying more preselection cuts, listed in table 4.1, in order to analyse the data samples much faster. All of them are softer than the final cuts described in the previous sections to keep a flexibility for control distributions. In the muon dataset, the muons must be of medium type and they must be detected by the A-layer of the muon system. In addition, the scintillator timing cuts for the cosmic veto are applied as well as the cut on the transverse momentum $p_T > 15\ GeV$. For the electron dataset, the preselection cuts are $|\eta| < 1.5$, a soft isolation cut $iso < 0.2$ as well as $p_T > 30\ GeV$ for the first electron and $p_T > 15\ GeV$ for all other electrons.

muon skim	electron skim
medium type	$\|\eta\| < 1.5$
$nseg \geq 1$	$iso < 0.2$
cosmic veto	$p_T > 30\ GeV$ for the first electron
$p_T > 15\ GeV$ for all muons	$p_T > 15\ GeV$ for all other electrons

Table 4.1: Soft preselection cuts to further reduce the sample sizes during the conversion from thumbnail to root-tree format.

Finally, 4,235,676 events are left in the muon data sample and 2,597,813 events in the electron data sample.

4.2.1 Triggers

The events are preselected by triggers. The triggers used in this analysis are inclusive single muon and EM triggers recommended by the Electroweak Physics Group for W/Z analyses [38]. The trigger definitions are summarised in so-called trigger lists. As the software is continuously improved there are changes in these trigger lists from time to time.

Muon Trigger

For the search in final states with at least one muon, the trigger `MU_W_L2M5_TRK10` is used for trigger list version 10.03 and below and the trigger `MUW_W_L2M3_TRK10` for trigger list versions 10.30 and higher. Runs between the list versions 10.03 and 10.30

were not physics runs and are not included in the sample.

At level 1, `MU_W_L2M5_TRK10` is based on the muon scintillators in $|\eta| < 1.5$ only and requires that the luminosity monitors are in coincidence. At level 2, there must be at least one muon found of medium quality and with $p_T > 5\ GeV$. At level 3, the trigger bit is set to true if one global muon track is found with $p_T > 10\ GeV$.
Although the prescales for this trigger are high (1-361), it was found to be the best choice among the available triggers for the according runrange. Besides, the trigger is used for only 13% of the whole muon dataset and, in particular, the MC samples are corrected for the differences in trigger and reconstruction efficiencies by a global scale factor (cf. section 5.2).

The other trigger, `MUW_W_L2M3_TRK10`, asks for a single muon with tight scintillator and loose drift tube requirements within $|\eta| < 1.5$ at level1. At level 2 at least one medium muon must be found with $p_T > 3\ GeV$. Again, the event is accepted at level 3 if one global muon track is found with $p_T > 10\ GeV$.
This trigger has a precale of 1-2.

Electron Trigger

For the part of the analysis dealing with electron final states the logical OR of different single electron triggers is used. For trigger list versions 11 and below, the OR of `EM_HI`, `EM_HI_SH`, `EM_HI_EMFR8`, `EM_HI_TR` and `EM_HI_SH_TR` is taken, while for trigger list version 12 and higher the OR of `E1_SHT20`, `E2_SHT20`, `E3_SHT20`, `E1_SH30`, `E2_SH30` and `E3_SH30` is considered.

For the first OR-combination, all triggers share the same level 1 requirement: there must be at least one EM calorimeter trigger tower with a transverse energy of 10 GeV. All of them do not have a restriction at level 2, but different criteria for level 3 which are listed in the following:

- `EM_HI`
  The transverse energy of an electron candidate must exceed 30 GeV in the calorimeter trigger towers and the candidate must be found within $|\eta| < 3$.
- `EM_HI_SH`
  The electron candidate has to deposit 20 GeV transverse energy in the calorimeter trigger towers with an electromagnetic shower shape and be found within $|\eta| < 3$.
- `EM_HI_EMFR8`
  The electron candidate has to deposit 40 GeV transverse energy in the calorimeter

trigger towers and meets the requirement $emf > 0.8$. Again it must be found within $|\eta| < 3$.

- `EM_HI_TR`
  A track with $p_T > 25\ GeV$ must be found by the central tracking system.
- `EM_HI_SH_TR`
  The electron candidate has to deposit 12 GeV transverse energy in the calorimeter trigger towers with an electromagnetic shower shape. Additionally, there must be a reconstructed track with $p_T > 12\ GeV$.

The prescales range from 1-50.

For the second OR-combination there are different requirements at level 1 and 3. At level 1, they require one EM calorimeter object with $E_T > 11\ GeV$ for `E1_xy`, two EM calorimeter objects with $E_T > 6\ GeV$ for `E2_xy` and for `E3_xy` there must be two EM calorimeter towers, one with $E_T > 3\ GeV$ and the other with $E_T > 9\ GeV$. The level 3 requirements are:

- `E1_SHT20`, `E2_SHT20` and `E3_SHT20`
  These triggers require an electron candidate satisfying tight shower shape requirements with $E_T > 20\ GeV$.
- `E1_SH30`, `E2_SH30` and `E3_SH30`
  The electron candidate has to deposit 30 GeV transverse energy in the calorimeter trigger towers and must satisfy loose shower requirements.

All these triggers share the same prescale of 1.

Hence, these triggers provide a preselection of electrons with $p_T > 30\ GeV$ with a high efficiency (cf. section 4.2.3).

4.2.2 Run Quality Criteria and Luminosity

The integrated luminosities for the datasets are calculated excluding the luminosity that corresponds to runs marked as *bad* for one of the detector subsystems: Muon, CFT- and SMT-system and the calorimeter. Those runs are summarised in a special bad run list. The corresponding jet and MET run-selections are also available, but they are not recommended since they entail a larger data loss (about 10% of the luminosity). Instead, a list of bad luminosity block numbers (LBN) according to the jet/MET information is created. These two lists are used by the program *runrange_luminosity* [40] which calculates the luminosity with respect to the chosen triggers.

For the muon dataset, the integrated luminosity amounts to 218.4 pb^{-1}, with 28.9 pb^{-1} according to the `MU_W_L2M5_TRK10` trigger and 189.5 pb^{-1} according to the `MUW_W_L2M3_TRK10` trigger.
For the electron dataset, the integrated luminosity has been calculated to 228.8 pb^{-1} with 128.8 pb^{-1} according to the OR of `EM_HI`, `EM_HI_SH`, `EM_HI_EMFR8`, `EM_HI_TR` and `EM_HI_SH_TR`, and 100.1 pb^{-1} according to the OR of `E1_SHT20`, `E2_SHT20`, `E3_SHT20`, `E1_SH30`, `E2_SH30` and `E3_SH30`.

4.2.3 Trigger Efficiencies

The efficiency of the trigger `MU_W_L2M5_TRK10` was found to be $\epsilon_{trigger1} = 0.60 \ \pm 0.02$ according to [48], whereas the efficiency of `MUW_W_L2M3_TRK10` is $\epsilon_{trigger2} = 0.92 \ \pm 0.03$, which has been obtained in the following way: First, an independent sample of adequate statistics is created from the muon skim by requiring one of the electron triggers to have fired. Then, the efficiency is calculated via:

$$\epsilon_{trigger2} = \frac{\text{number of selected muons where the trigger has fired}}{\text{total number of selected muons}} \quad . \tag{4.13}$$

Combining both efficiencies according to their corresponding integrated luminosities yields a total efficiency of

$$\epsilon_{total} = \frac{\mathcal{L}_1 \cdot \epsilon_1 + \mathcal{L}_2 \cdot \epsilon_2}{\mathcal{L}_1 + \mathcal{L}_2} = 0.88 \pm 0.03 \tag{4.14}$$

for both muon triggers.

The efficiencies of the electron triggers are 0.98 ± 0.01 for the first OR-combination and 0.98 ± 0.02 for the second [39], giving a total efficiency of $\epsilon_{total}^{electron} = 0.98 \ \pm 0.01$.

The obtained efficiencies are used in chapter 5.2 as a cross check for the derived MC correction factors as well as to determine systematic uncertainties of the scaling factors.

In both cases, the efficiencies are assumed to be independent of kinematical variables, e.g. p_T and η.

Chapter 5

Monte Carlo Simulations and QCD Background

In section 5.1 an overview of the Monte Carlo samples is given. Section 5.2 deals with the technique of fitting the sum of Monte Carlo to data in a certain region in order to account for different efficiencies. Finally, in section 5.3 the estimation of the QCD contributions from data is explained.

5.1 Monte Carlo Samples

The Standard Model prediction is given by Monte Carlo (MC) samples. Most of them (e.g. $W \to l\nu$, $Z/\gamma^* \to ll$ and $\Upsilon(1S) \to ll$, $\Upsilon(2S) \to ll$) are official DØ-MC and were generated either centrally at Fermilab or at remote production sites with the PYTHIA [41] generator, version 6.2, using CTEQ4L and CTEQ5L parton distribution functions [44].

Two $t\bar{t}$ Monte Carlo samples are taken from the Top Group [45], $t\bar{t} \to 2l + X$ and $t\bar{t} \to l + X$. They are generated with the Alpgen [46] generator version 1.2.1 with a top mass of 175 GeV.
The $WW \to X$, $WZ \to X$ and $ZZ \to X$ Monte Carlo samples are "private" MCs, generated by Ulla Blumenschein [42] of the New Phenomena Group [43] using the PYTHIA generator.

All these SM background Monte Carlo simulations are listed in table 5.1 together with their cross section (σ_{MC}) × branching ratio (BR) in leading order calculation (LO), next-to leading order (NLO) or next-to-next-to leading order (NNLO) as well as their equivalent MC-luminosity ($\mathcal{L}_{MC}$). The corresponding k-factors, defined as

$k = \sigma_{NNLO}/\sigma_{LO}$, are taken from [47] and are also listed in the table. It is ensured that the luminosities of the main background samples are significantly higher than those of the analysed data samples. The simulations of $Z/\gamma^* \to ll$ $(5 - 15\ GeV)$ and $\Upsilon \to ll$ are of lower statistics, though, but since they give only small or no contributions to the event classes this effect is assumed to be negligible.

To the Monte Carlo samples, additional scale factors and smearing of energy and momentum of the physics objects are applied, in order to reproduce the data at the known Z-peak. Details of the techniques and their systematic uncertainties are given in chapter 6.

For the QCD background to the inclusive lepton datasets, one strategy would be to generate a Monte Carlo sample to model this background as it is done with all other backgrounds. This technique was tried in the past, but there are several problems: The production mechanisms of, e.g. $b\bar{b}$ and $c\bar{c}$-processes, are difficult to simulate in MC because there are many different processes contributing to the production of heavy flavours. Finally, the large cross sections, for example $\approx 0.5\ mb$ for $b\bar{b}$ (compared to $\approx 3\ nb$ for $W \to l\nu$), complicate the production of sufficient Monte Carlo samples. A large scale production usable for many analyses would have to be coordinated, and this has not happened to date. Only samples with very low statistics exist. Hence, the common approach is to estimate this background from data which will be described in detail in the subsequent section 5.3.

5.2 Efficiency Corrections

In Monte Carlo simulations the detector is well modelled but still, there are slightly different reconstruction efficiencies ($\Delta\epsilon_{reco}$) compared to data. Besides, trigger efficiencies ($\epsilon_{trigger}$) are not simulated in the Monte Carlo simulations. To account for all different efficiencies a MC correction factor of $f = \Delta\epsilon_{reco} \cdot \epsilon_{trigger}$ is obtained from a χ^2-fit. The procedure is described in the following.

First, a region is selected in which the SM is well understood and the detector modelling is satisfying, so that new physics is not expected in that region. Then, a scale factor f is obtained by minimizing the $\chi^2(f)$:

$$\chi^2(f) = \sum_{i=\text{ first bin}}^{\text{last bin}} \frac{(Data_i - f \cdot MC_i)^2}{\sigma^2(f \cdot MC_i)} \tag{5.1}$$

where $Data_i$ is the number of data events in bin i, MC_i is the sum over all Monte Carlo events in bin i and $\sigma(f \cdot MC_i)$ is the statistical error of the scaled MC contribution. At

MC sample (M_{Z/γ^*} $[GeV]$)	$\sigma_{MC} \times BR$ $[pb]$	K_{NNLO}	$\mathcal{L}_{MC}$ $[pb^{-1}]$
$W \to e\nu$	2676.9 (NNLO)	1.39	832.8
$W \to \mu\nu$	2676.9 (NNLO)	1.39	824.2
$W \to \tau\nu$	2676.9 (NNLO)	1.39	648.4
$Z/\gamma^* \to \mu\mu$ (5-15)	4579.7 (NNLO)	1.29	47.9
$Z/\gamma^* \to \mu\mu$ (15-60)	450.4 (NNLO)	1.35	222.0
$Z/\gamma^* \to \mu\mu$ (60-130)	252.7 (NNLO)	1.39	1024.9
$Z/\gamma^* \to ee$ (5-15)	4579.7 (NNLO)	1.29	47.0
$Z/\gamma^* \to ee$ (15-60)	450.4 (NNLO)	1.35	222.0
$Z/\gamma^* \to ee$ (60-130)	252.7 (NNLO)	1.39	977.4
$Z/\gamma^* \to \tau\tau$ (5-15)	4579.7 (NNLO)	1.29	65.8
$Z/\gamma^* \to \tau\tau$ (15-60)	450.4 (NNLO)	1.35	485.8
$Z/\gamma^* \to \tau\tau$ (60-130)	252.7 (NNLO)	1.39	997.2
$WW \to X$	13.0 (NLO)		1076.9
$ZZ \to X$	1.56 (NLO)		9615.4
$WZ \to X$	3.96 (NLO)		4166.7
$t\bar{t} \to$ 1 ν 4j	2.97 (NLO)		15404.0
$t\bar{t} \to$ 2l ν 2j	0.70 (NLO)		13571.4
$\Upsilon(1S) \to \mu\mu$	190.4 (LO)		178.5
$\Upsilon(2S) \to \mu\mu$	183.8 (LO)		163.2
$\Upsilon(1S) \to ee$	184.82 (LO)		167.7
$\Upsilon(2S) \to ee$	201.35 (LO)		149.9

Table 5.1: Listed are all considered SM Monte Carlo samples.

this stage, the Monte Carlo samples are already normalised to the data luminosity ($\mathcal{L}_{data}$) using the weight $w = \mathcal{L}_{data}/\mathcal{L}_{MC}$. The correction factor f is derived at the minimum of the $\chi^2(f)$. Hence, f gives the best match between data and SM prediction in the selected region.

In this analysis, the selected region is the Z peak ($75\ GeV < p_T < 95\ GeV$) and the fit is done in the p_T-distributions of the exclusive event classes with two muons and two electrons, respectively, since the QCD contributions in these event classes are negligible:

$$f(2\mu) = 0.763 \pm 0.018 \text{ (stat.) with } \chi^2/\text{dof} = \tfrac{29.7}{19} = 1.56$$
$$f(2e) = 0.791 \pm 0.012 \text{ (stat.) with } \chi^2/\text{dof} = \tfrac{33.4}{19} = 1.76$$

Exemplarly, the graph of the fit in the event class with two electrons is given in figure 5.1.

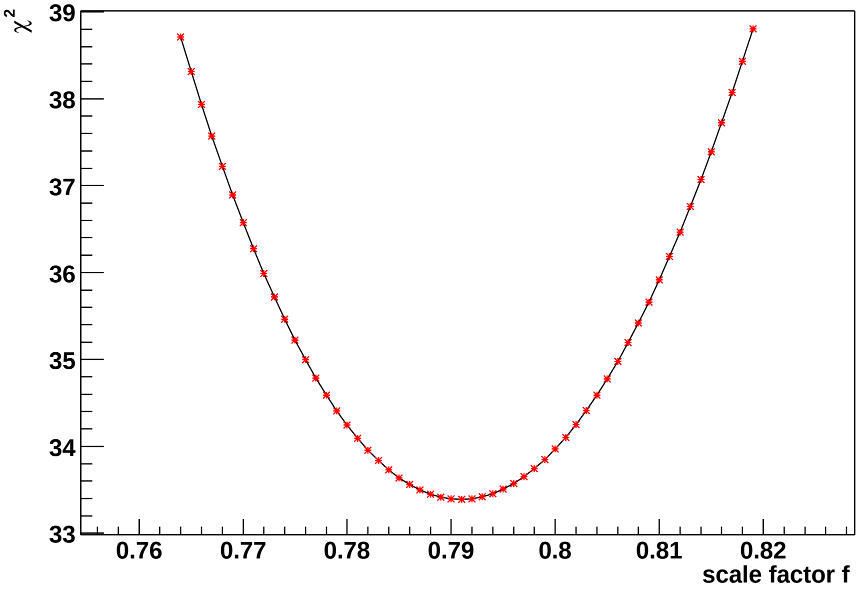


Figure 5.1: The χ^2-fit in the event class with two electrons. The MC correction factor f is taken at the minimum of the χ^2.

To derive the correction factor for the one object class, $f(1)$, from the factor of the class with 2 objects, $f(2)$, one has to consider the trigger efficiency which is given by

$$\epsilon_{trigger}(2) = \epsilon_{trigger}(1) + [1 - \epsilon_{trigger}(1)] \cdot \epsilon_{trigger}(1) = 2 \cdot \epsilon_{trigger}(1) - \epsilon_{trigger}(1)^2 \quad . \quad (5.2)$$

The first term, $\epsilon_{trigger}(1)$, is the efficiency of the trigger to select object 1. The second term arises from the possibility, that the trigger has not selected object 1, $1 - \epsilon_{trigger}(1)$, but selects object 2. Therefore, the efficiency is given by the product $[1-\epsilon_{trigger}(1)] \cdot \epsilon_{trigger}(1)$. The division of $f(2)$ and $\epsilon_{trigger}(2)$ gives the difference in reconstruction efficiencies for 2 objects, $\Delta\epsilon_{reco}(2) = f(2)/\epsilon_{trigger}(2)$ and the square root of $\Delta\epsilon_{reco}(2)$ is the difference in reconstruction efficiencies for 1 object. Hence, $f(1)$ is given by

$$f(1) = \sqrt{\frac{f(2)}{2 \cdot \epsilon_{trigger}(1) - \epsilon_{trigger}(1)^2}} \cdot \epsilon_{trigger}(1) = \sqrt{\frac{f(2) \cdot \epsilon_{trigger}(1)}{2 - \epsilon_{trigger}(1)}} \quad , \tag{5.3}$$

and applied to the correction factors $f(2\mu)$ and $f(2e)$ mentioned above, yields $f(1\mu) = 0.774$ and $f(1e) = 0.872$, respectively.

Now, these results are combined with those derived in [4] where the fit is obtained at the W peak in exclusive event classes with one muon and one electron, respectively. This gives an estimation of the systematic uncertainties of the MC correction factors by taking the difference between the value of $f(1)$, calculated in the explained way, and the value of the MC correction factor from [4]. In the same way, the systematic uncertainty for $f(2)$ is derived by the author of reference [4]. In table 5.2 the correction factors for event classes with electrons and muons are listed as well as their expected values by taking the product of $\epsilon_{trigger}$ from section 4.2.1 and $\Delta\epsilon_{reco}$. For the electron reconstruction, the efficiency correction $\Delta\epsilon_{reco}$, is taken from [29], while for muons the corresponding efficiency correction $\Delta\epsilon_{reco}$ is derived in the following way: A Z-sample is obtained from data where one object, the “tag object”, is required to satisfy all selection criteria. The remaining object, the so-called “probe object”, must again satisfy all criteria, but this time the criterion for which the efficiency is to be determined is not required explicitely. The efficiency is then defined as the number of probe objects satisfying the criterion in question divided by the total number of probe objects. In the same way, the efficiencies are determined from a $Z \to \mu\mu$ Monte Carlo sample, to finally derive the efficiency correction $\Delta\epsilon_{reco}$. As one can see from table 5.2, the correction factors derived are reproducable.

	correction factors from		$f = \Delta\epsilon_{reco} \cdot \epsilon_{trigger}$	$\Delta\epsilon_{reco}$	$\epsilon_{trigger}$
	Z peak	W peak	expected		
1e	0.872	0.894	0.85	0.870	0.980
1μ	0.774	0.749	0.766	0.876	0.874

Table 5.2: The MC correction factors for event classes with one electron (one muon), as well as their expected values. The value for the difference in electron reconstruction, $\Delta\epsilon_{reco}$, is taken from [29].

The final MC correction factors used for the rest of the analysis are, for event classes containing one lepton

$$\begin{aligned} f(1\mu) &= 0.749 \pm 0.003 \text{ (stat.) } \pm 0.025 \text{ (syst.)} \\ f(1e) &= 0.894 \pm 0.003 \text{ (stat.) } \pm 0.022 \text{ (syst.)} \end{aligned}$$

and for event classes containing two leptons

$$\begin{aligned} f(2\mu) &= 0.763 \pm 0.018 \text{ (stat.) } \pm 0.049 \text{ (syst.)} \\ f(2e) &= 0.791 \pm 0.012 \text{ (stat.) } \pm 0.041 \text{ (syst.)} \end{aligned}$$

In the figures 5.2 and 5.3 the invariant mass spectrum of two electrons and two muons, respectively, is given. These figures show clearly how data and MC agree at the Z peak and how well the scaling factors correct the MC samples.

5.3 Estimation of the QCD Background from Data

This section introduces the techniques of estimating the QCD background in event classes with electrons and muons. It consists of three steps that are common to both types of final states:

1. One of the cut criteria from sections 4.1.1 and 4.1.2, respectively, is inverted in order to describe the QCD background. Only one cut is modified while all other selection criteria remain unchanged. Hence, an independant QCD sample is extracted from data. Details are given in section 5.3.1 and section 5.3.2. This new cut combination is applied to data and to all MC samples.
2. The sum of all MC events (selected with the new cut combination) are subtracted bin by bin from data (selected with the new cut combination) in order to subtract the known Monte Carlo contributions from the QCD sample.
3. The QCD sample obtained (X_{QCD}) is scaled with a factor f_{QCD} determined with a χ^2-fit in the region of the inclusive event classes 1e + X or 1μ + X, where the QCD contribution is dominant. In particular, this is the low p_T region where no new physics are expected. Similarly to the determination of the MC correction factors in section 5.2, the factor f_{QCD} is derived by minimizing the $\chi^2(f_{QCD})$ of the fit:

$$\chi^2(f_{QCD}) = \sum_{i=\text{ first bin}}^{\text{last bin}} \frac{[(Data_i - f \cdot MC_i) - f_{QCD} \cdot X_{QCD,i}]^2}{\sigma^2(f_{QCD} \cdot X_{QCD,i})} \tag{5.4}$$

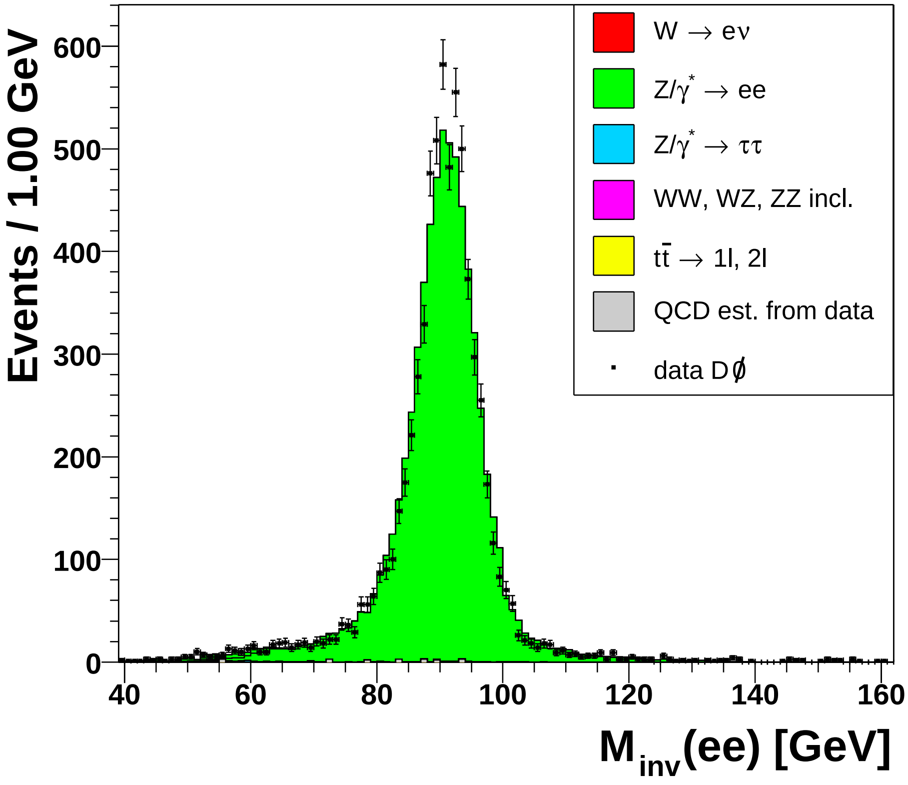


Figure 5.2: Invariant mass distribution of two electrons.

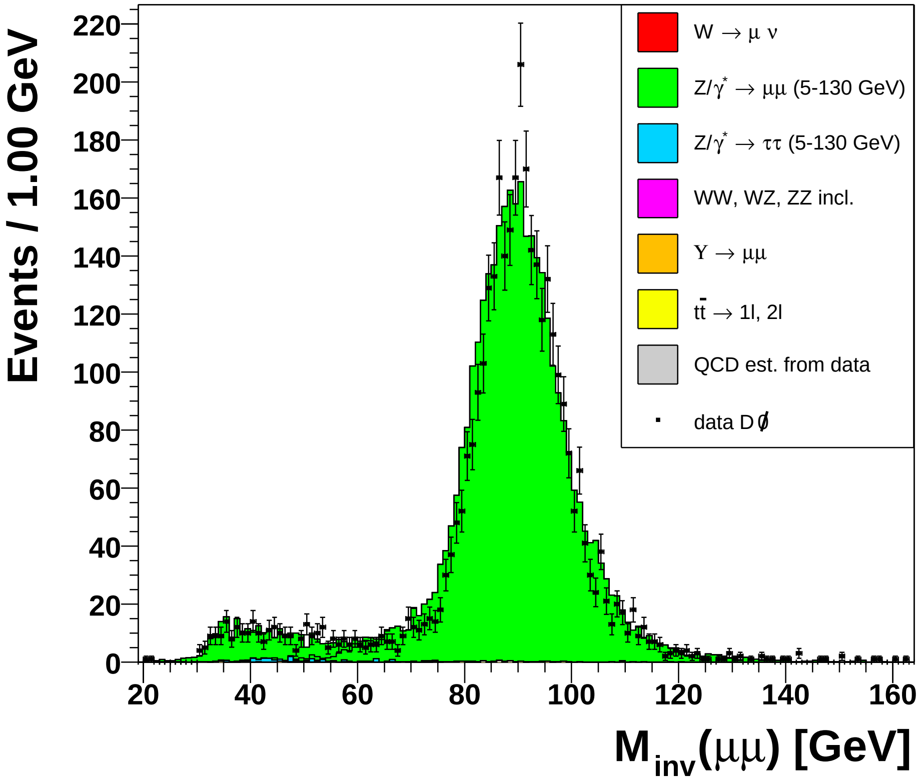


Figure 5.3: Invariant mass distribution of two muons.

where $Data_i$ is the number of data events in bin i and MC_i is the sum over all Monte Carlo events in bin i. The factor f is the MC correction factor from section 5.2. In this fit, only the statistical error of the scaled QCD sample $\sigma(f_{QCD} \cdot X_{QCD})$ is considered.
This factor f_{QCD}, derived at the minimum of the $\chi^2(f_{QCD})$, is then used to scale the QCD sample in all event classes.

5.3.1 QCD Background for Electron Final States

An example of the QCD contribution in event classes with electron final states are so-called "fake" electrons emerging from pion jets, where a π^0 decays into two gammas ($\pi^0 \rightarrow 2\gamma$ with branching ratio 98.9% [49]). The photons form an electromagnetic shower in the calorimeter. As they stem from a jet, which consists of charged and uncharged pions, a track may also be assigned to the fake electron. Besides, real electrons may be produced via pair production. An illustration is given in figure 5.4. It is assumed that the electromagnetic shower of these fake electrons fluctuate much more in width and depth than genuine electrons, so that the QCD background can be extracted from data by inverting the cut on the shower shape, i.e. $HMX7 > 20$ (cf. equation 4.3).

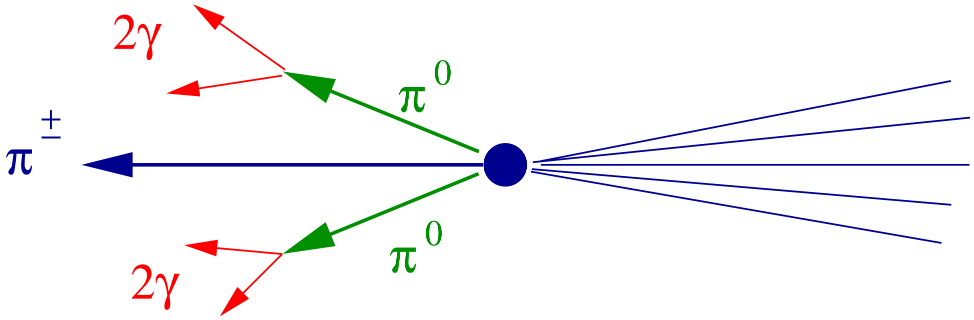


Figure 5.4: Fake electrons from pion jets.

In figure 5.5 the p_T-distribution of the inclusive event class 1e + X with all MC contributions but without the QCD sample is given. The fit to scale the QCD sample is done in this distribution in the range $p_T = 30 - 40\ GeV$ with a binning of $1\ GeV$:

$$f_{QCD} = 2.710 \pm 0.015 \text{ (stat.) } \pm 0.090 \text{ (syst.) with } \chi^2/\text{dof} = \tfrac{153}{9} = 17.0$$

The systematic uncertainty has been derived by varying the range of the fit to $p_T = 30 - 50\ GeV$. Although the fit has a high χ^2/dof, probably due to the fact that other errors than the statistical error of the QCD sample are neglected in the determination, the scale factor f_{QCD} gives a satisfying result. Figure 5.6 shows the p_T-distribution of the event class 1e + X with all MC contributions and the fitted QCD sample (shown as the foremost contribution). As desired, the QCD sample clearly fills out the part where it is dominant. Also, η- and ϕ-distributions are given in figures 5.7 and 5.8, respectively, where data and the sum of MC and QCD estimation agree well, too.

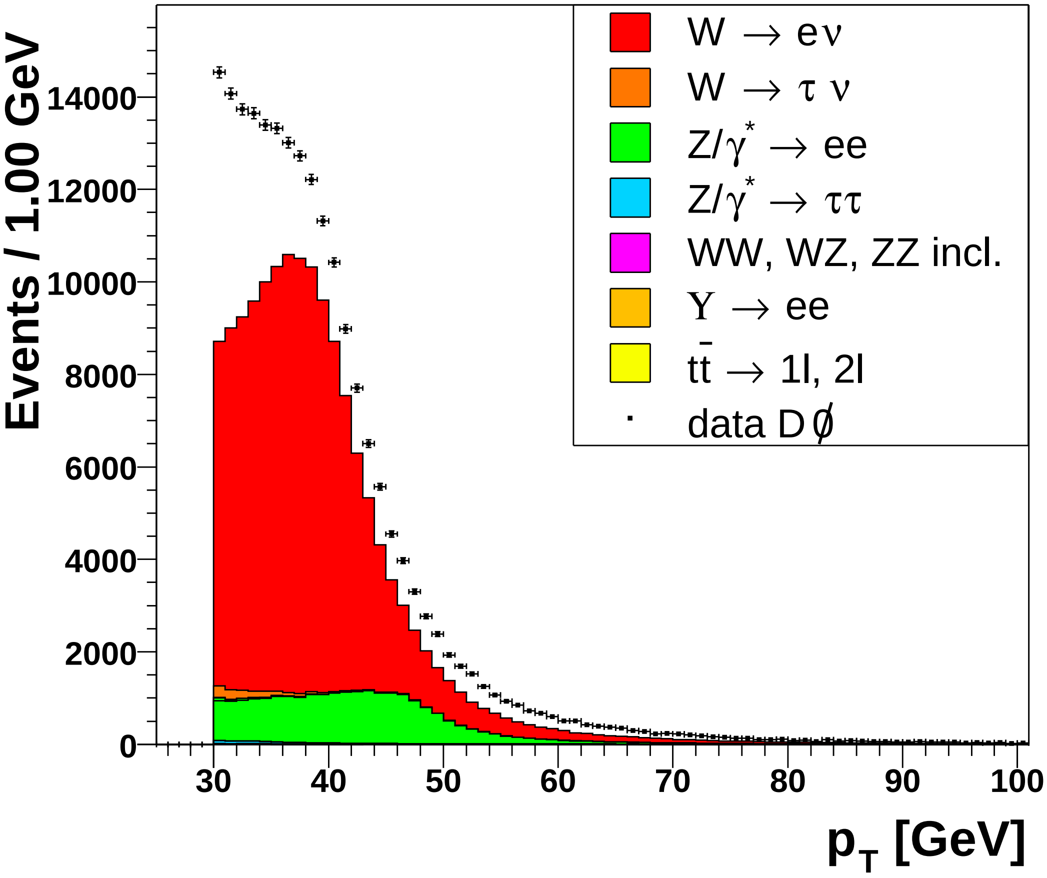


Figure 5.5: p_T-distribution of 1e + X showing the region where the QCD sample needs to be fitted into.

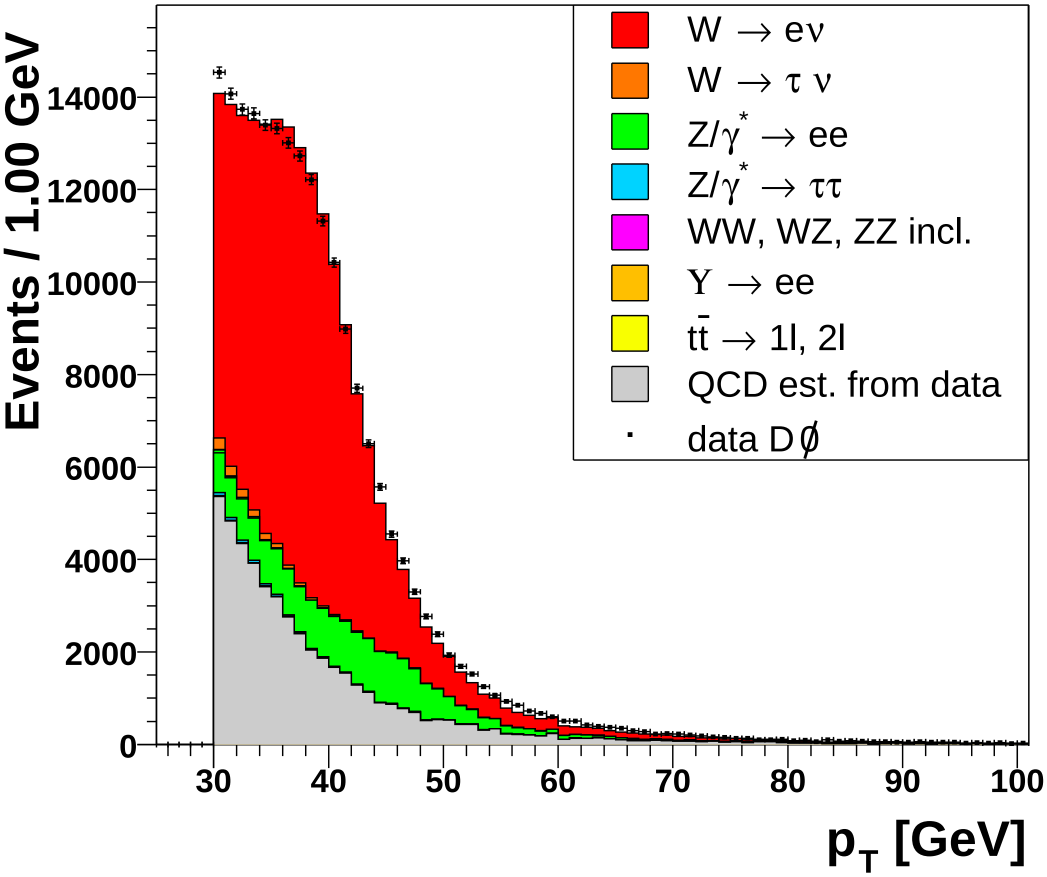


Figure 5.6: p_T-distribution of 1e + X with fitted QCD sample.

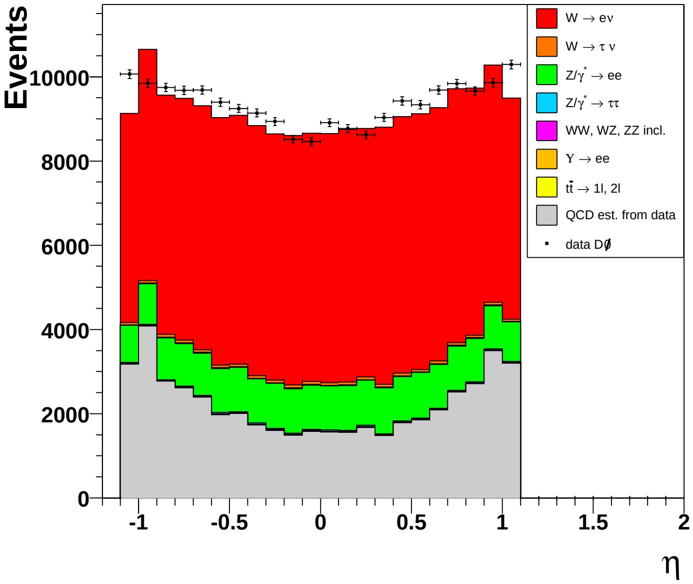


Figure 5.7: η-distribution of 1e + X with fitted QCD sample.

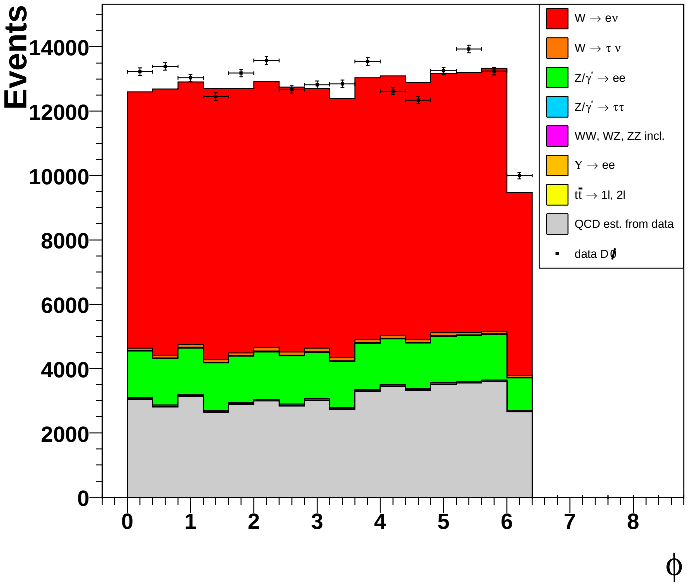


Figure 5.8: ϕ-distribution of 1e + X with fitted QCD sample.

5.3.2 QCD Background for Muon Final States

The QCD contribution in event classes with muons is expected to come mainly from semileptonic decays of $b\bar{b}$ and $c\bar{c}$ events (cf. figure 5.9) where a "fake" muon emerges from the jet. Hence, the isolation cut (cf. section 4.1.2) is modified to derive the QCD sample from data by requiring:

$$either\ 2.5\ GeV < E^{cal}_{cone} < 7\ GeV \text{ and } \sum_{tracks} p_T^{cone} < 7\ GeV$$

$$or\ 2.5\ GeV < \sum_{tracks} p_T^{cone} < 7\ GeV \text{ and } E^{cal}_{cone} < 7\ GeV,$$

where E^{cal}_{cone} is the sum of the calorimeter transverse energies, considering electromagnetic and fine hadron calorimeter layers, in a hollow cone with radius R between $0.1 < R < 0.4$ around the muon. The sum of the transverse momenta of all tracks except the muon track itself within a cone of radius $R = 0.5$ surrounding the muon is labelled $\sum_{\text{tracks}} p_T^{cone}$.

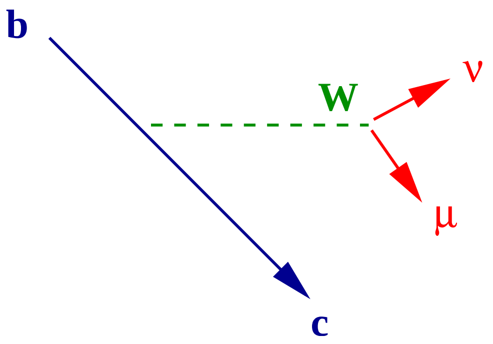


Figure 5.9: A muon from $b\bar{b}$ jets.

The fit of the QCD sample is done in the p_T-distribution of the inclusive event class with 1μ + X in the range $p_T = 15 - 25\ GeV$ with a $1\ GeV$ binning:

$$f_{QCD} = 0.105 \pm 0.001\ \text{(stat.)}\ \pm 0.021\ \text{(syst.)}\ \text{with}\ \chi^2/\text{dof} = \tfrac{60.5}{9} = 6.7$$

Figure 5.10 shows the p_T-distribution of the event class 1μ + X with all MC contributions but without the QCD sample. Since the fit range is limited to ten bins only ($p_T = 15 - 25\ GeV$) the systematic uncertainty of the fit can not be derived by varying this bin range. Instead, an estimation is done by varying the upper energy bounds of the modified isolation criterion by $\pm 1\ GeV$ to

$$either\ 2.5\ GeV < E^{cal}_{halo} < 6\ GeV\ (8\ GeV)\ \text{and}\ \sum_{tracks} p_T^{cone} < 6\ GeV\ (8\ GeV)$$

$$or\ 2.5\ GeV < \sum_{tracks} p_T^{cone} < 6\ GeV\ (8\ GeV)\ \text{and}\ E^{cal}_{halo} < 6\ GeV\ (8\ GeV),$$

giving a systematic uncertainty of the order of 20%.
Again, f_{qcd} is then used to scale the QCD sample in all event classes. Figure 5.11 shows the p_T-distribution of the event class 1μ + X with the included QCD sample (shown as the foremost contribution), which clearly fills out the part of the distribution where it is dominant. In figures 5.12 and 5.13 the η- and ϕ-distributions are shown. Hence, the results of this technique are satisfying.

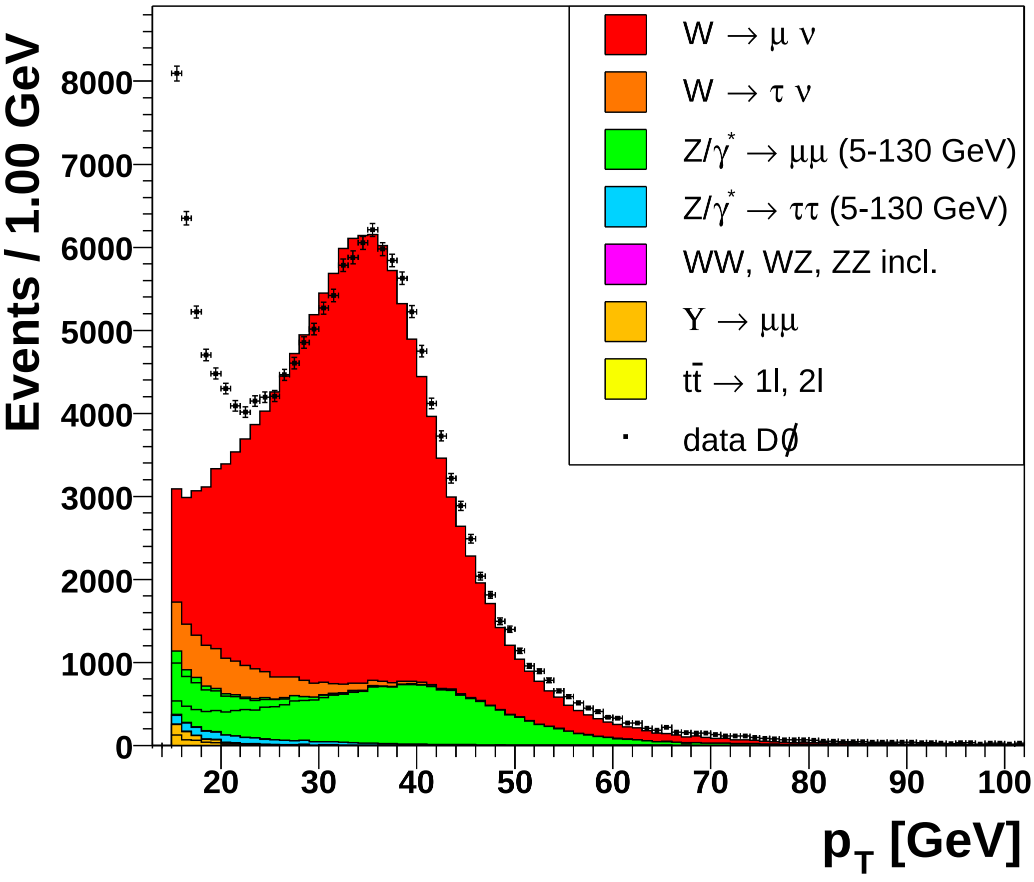


Figure 5.10: p_T-distribution of 1μ + X showing the region where the QCD sample needs to be fitted into.

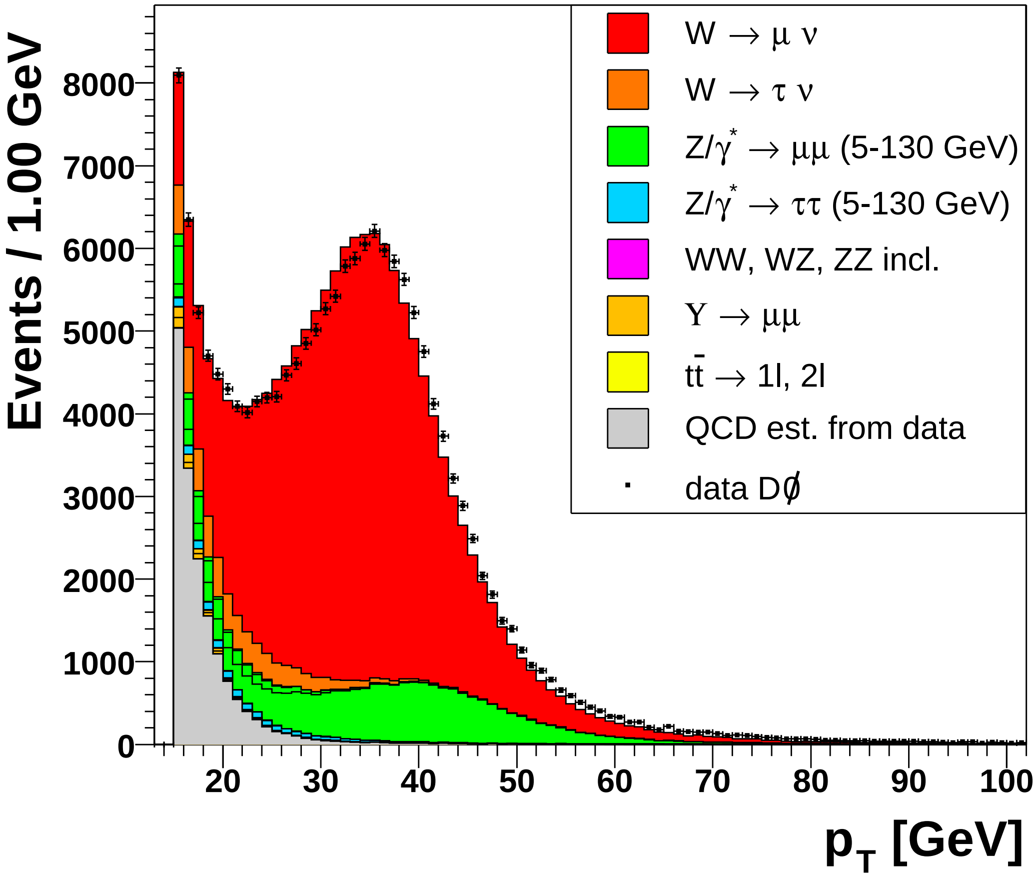


Figure 5.11: p_T-distribution of 1μ + X with fitted QCD sample.

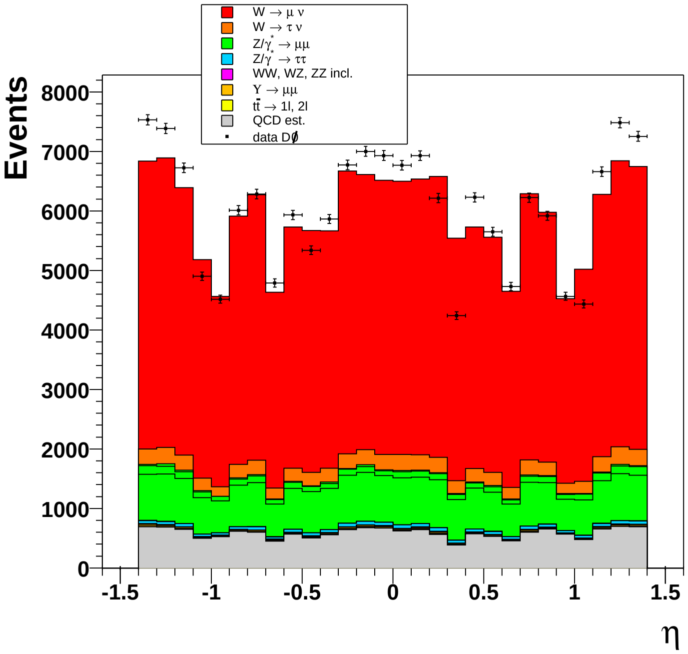


Figure 5.12: η-distribution of 1μ + X with fitted QCD sample.

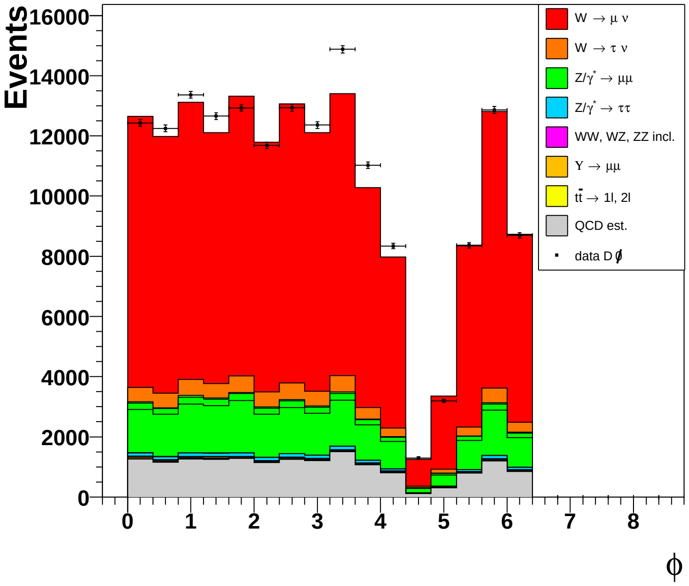


Figure 5.13: ϕ-distribution of 1μ + X with fitted QCD sample.

Chapter 6

Systematic Uncertainties

This chapter summarises the sources of systematic uncertainties which need to be accounted for when comparing data and Monte Carlo simulations. The systematic uncertainties of the Monte Carlo correction factors and of the QCD-background estimation from data have already been explained in the previous chapter.

6.1 Energy Scaling and Smearing

Electron Energy Scale and Resolution

By comparing the distribution of the dielectron invariant mass in Z candidate events it was found that the energy resolution in MC simulations is better than in data and that the position of the Z-peak in MC is slightly shifted from the data Z-peak. Hence, a tuning of the MC to data by applying an additional scale factor and an additional smearing of the electron energy is needed [50]. A parametrisation of the energy resolution is given in equation 3.2, section 3.2.3, which leads to the following expression to adjust the EM-scale and width

$$\widetilde{E} = E \cdot \alpha + E \cdot \mathrm{Gauss}(0, \sigma) \tag{6.1}$$

where α is the scale factor and additional smearing is provided by the Gaussian distribution with zero mean and standard devaition σ. The mean values and their uncertainties for the central calorimeter region are listed in table 6.1. Although the uncertainty in the scale factor is negligible its contribution is included in the search algorithm (cf. section 8).

Muon Momentum Scale and Resolution

Similar to the electron case, the difference in width and position of the $Z \to \mu\mu$ peak in data and MC can be corrected by applying a scale factor α and additional smearing in

	Central Calorimeter	
	in fiducial	not in fiducial
σ	0.042 ± 0.004	0.083 ± 0.006
α	1.007 ± 0.001	0.971 ± 0.012

Table 6.1: Values for additional smearing (σ) and EM energy scale factor α, taken from [50].

form of a Gaussian of width σ:

$$\frac{1}{\widetilde{p_T}} = \frac{1}{p_T \cdot \alpha} + \mathrm{Gauss}(0, \sigma) \quad . \tag{6.2}$$

In this analysis, α and σ are chosen as: $\alpha = 0.991 \pm 0.003$ and $\sigma = (2.47\,^{+0.01}_{-0.02}) \cdot 10^{-2}$. The scale factor uncertainty of 0.26% is negligible whereas the uncertainty of the Gaussian smearing is of the order of 6%.

Jet Energy Scale

The jet energy scale has already been discussed in section 4.1.4. The error of the correction factor varies with energy and pseudo-rapidity. For simplicity, an average value of 6% is assumed for the uncertainty, which is typical for the high p_T region [34].

6.2 Cross Sections of Monte Carlo Simulations

Due to uncertainties in the parton distribution functions (PDFs), variations in the renormalisation and factorisation scale as well as in the consideration of higher orders (using k-factors), the uncertainties of the production cross sections of all SM-processes must also be taken into account. Since the cross sections of W and Z are highly correlated the uncertainties are taken to be common to both processes. Also, the cross sections of the diboson production processes WW, WZ and ZZ are assumed to be correlated.
For W and Z, the uncertainty on the cross section is of the order of 3.5% [47].
By comparing the cross sections of WW, WZ and ZZ calculated with two different parton density functions, MRS98 and CTEQ5, an uncertainty of 4% was found in [51].
The uncertainty of the $t\bar{t}$ cross section of the order of 5% is taken from [52].
Since the two Upsilon-Monte Carlo samples contribute in the p_T-range where the χ^2-fits are done only and, in particular, which is excluded in the search algorithm, the cross section uncertainties are assumed to be negligible here.

Chapter 7

Search for Deviations via the Total Number of Events

Deviations may be visible in the total number of events. Hence, the first step of the search procedure consists in comparing the total number of events found in data and Monte Carlo simulations. The results are summarised in different tables; there are three tables each: for exclusive and inclusive event classes of the electron dataset the results are summarised in the tables 7.2 to 7.4 and 7.5 to 7.7, respectively as well as for exclusive and inclusive event classes of the muon dataset in the tables 7.8 to 7.10 and 7.11 to 7.13, respectively. The first column shows the event class, followed by the number of events found in the DØ dataset and the sum over all Monte Carlo samples and the QCD contribution. Subsequently the single contributions of each Monte Carlo sample are listed together with the statistical uncertainties. The contributions of the according QCD samples together with the contributions of the Upsilon-Monte Carlo samples are given in each third table. In the case that less than 5 MC events are found in an event class, i.e. the outcome of the measurement is a Poisson variable, the 68.3% confidence level (CL) is taken as an upper limit (denoted with "<" in the tables). The probability $\beta = 1 - 0.683 = 0.317$ is the probability to observe a value less than or equal to the one measured (N_{meas}) and is related to the upper limit b via

$$\beta = \sum_{n=0}^{N_{meas}} \frac{b^n}{n!} \cdot e^{-b} \quad . \tag{7.1}$$

For example, if the measurement results in $N_{meas} = 0$, equation 7.1 becomes

$$\beta = \sum_{n=0}^{0} \frac{b^n}{n!} e^{-b} = e^{-b} \;\Leftrightarrow\; b = -\ln\beta \tag{7.2}$$

giving an upper limit of $b = 1.15$. The upper limits for $N_{meas} = 0, \ldots, 4$ are summarised in table 7.1. Otherwise the event counts are assumed to follow a Gaussian distribution where the statistical error is $\sqrt{N_{meas}}$. The MC contributions are then scaled according to the integrated luminosity of the dataset and corrected for efficiencies by applying the MC correction factors from section 5.2. For the QCD sample the systematic uncertainty is given in the tables which is the dominant error.

N_{meas}	β in %	upper limit
0	31.7	1.15
1	31.7	2.36
2	31.7	3.52
3	31.7	4.65
4	31.7	5.77

Table 7.1: Poisson limits calculated for 68% CL.

Since some of the Monte Carlo samples are of low statistics and give contributions in one or two event classes only, one has to be careful when adding up all upper limits (in the case $N_{meas} = 0$) from these samples which would result in too large SM predictions. Hence, in certain event classes some of the Monte Carlo samples are classified to be of no contribution (n.c.). One criterion to label a sample n.c. is, if it does not contribute in a superset of the event sample under study. For example, if the contribution in 1e 3j is zero, the sample will be classified n.c. in 1e 4j if the contribution in this class is zero, too. This can be justified since, e.g. the radiation of each additional gluon is a process of higher order and the probability is directly proportional to the strong coupling constant α_s. Furthermore, a sample is classified n.c. if its signature is different from the event class, i.e. it is very unlikely that the sample $Z \to \tau\tau$ $(5-15\ GeV)$ produces 2 high energy muons of at least $p_T = 15\ GeV$ each as well as 2 jets with more than $p_T = 30\ GeV$ each.

A first look at the tables suggests that data and SM prediction agree quite well. In the event class 1e 3j (cf. tables 7.2 and 7.5) significantly more events are found in data than in Monte Carlo simulations. A reason for this 15% deviation may be that gluon radiation is not well modelled by the event generator PYTHIA.

Also, the event class 1e 1μ 3j (see tables 7.2 and 7.5) shows a deviation where one event is found in data and the only MC contribution is $t\bar{t} \to 2lX$ with 0.11 ± 0.04 events. The Figures 7.1 and 7.2 show this event in Lego view and in RZ, XY view, respectively. Energy depositions in the electromagnetic calorimeter are in red while depositions in the hadron calorimeter are in blue. Hits in the muon layers are in red (A), orange (B) and green (C). The selected muon is the one that hits all three layers of the muon system and has a transverse momentum of 30 GeV. The p_T of the electron is 136 GeV and the jets have transverse momenta of 83 GeV, 60 GeV and 46 GeV, respectively. Also, a fourth jet with $p_T = 34\ GeV$ is shown in the figures, which failed the selection criteria since it is very close to the beam pipe (at $\eta = 3$) and might be a beam remnant. The missing transverse energy (yellow) amounts to 82 GeV. However, this event is also observed in the $t\bar{t}$-analysis of reference [53], so that it is most likely a $t\bar{t} \rightarrow 1e\ 1\mu\ 3j$ event and no hint for new physics.

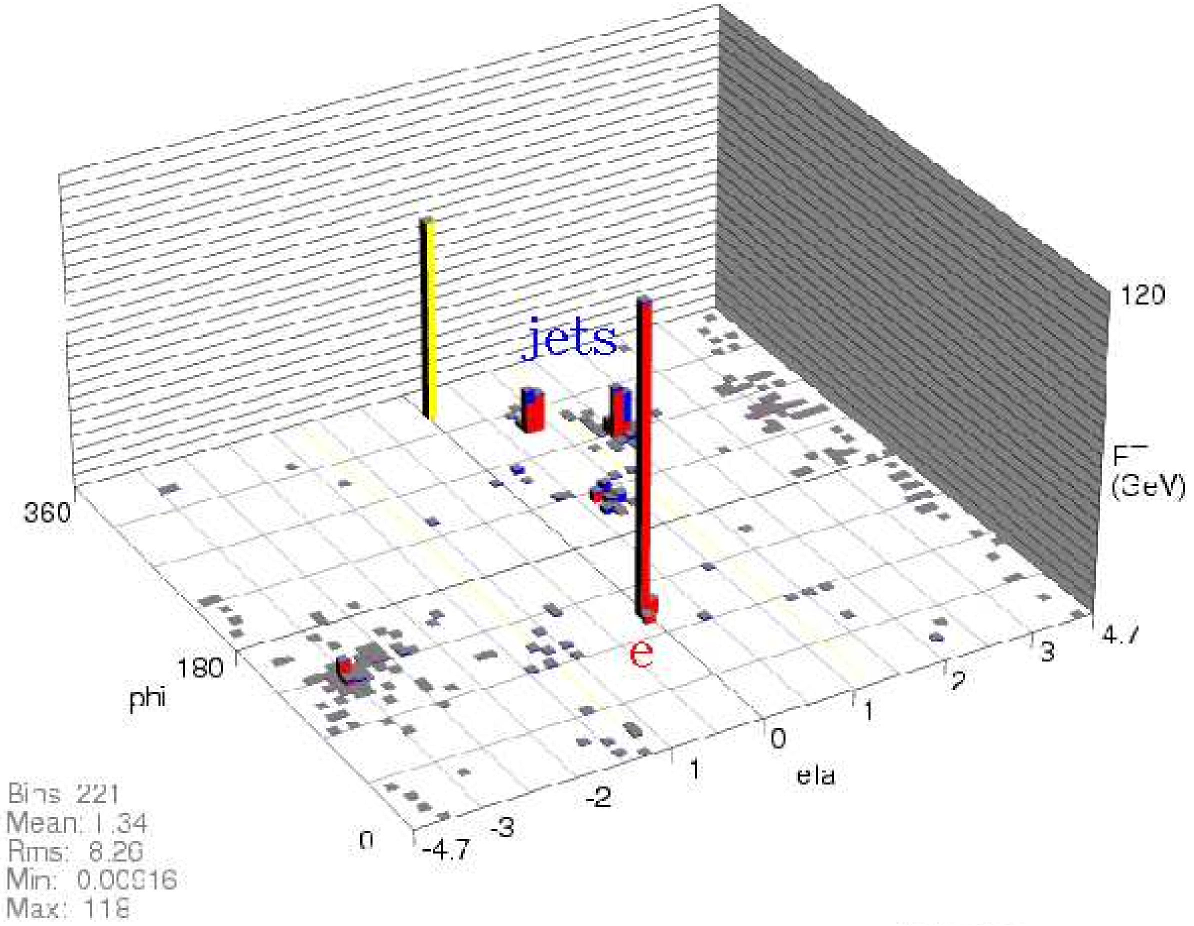


Figure 7.1: The event 1e 1μ 3j in Lego view, showing energy depositions in the calorimeter only.

For the event classes containing one muon and one jet (cf. tables 7.8 and 7.11), 20% more events are found in the total sum of all MC samples and QCD estimation than actually observed in data. The figure 7.3 shows the p_T sum of muon and jet including all Monte Carlo samples and the QCD estimation; in figure 7.4 the QCD contribution is excluded. These figures clearly show that SM prediction and data are not in good agreement. In particular, figure 7.4 shows that there is almost no space for any QCD contribution in this event class. On the other hand, for the event class 1e 1j, the QCD contribution is

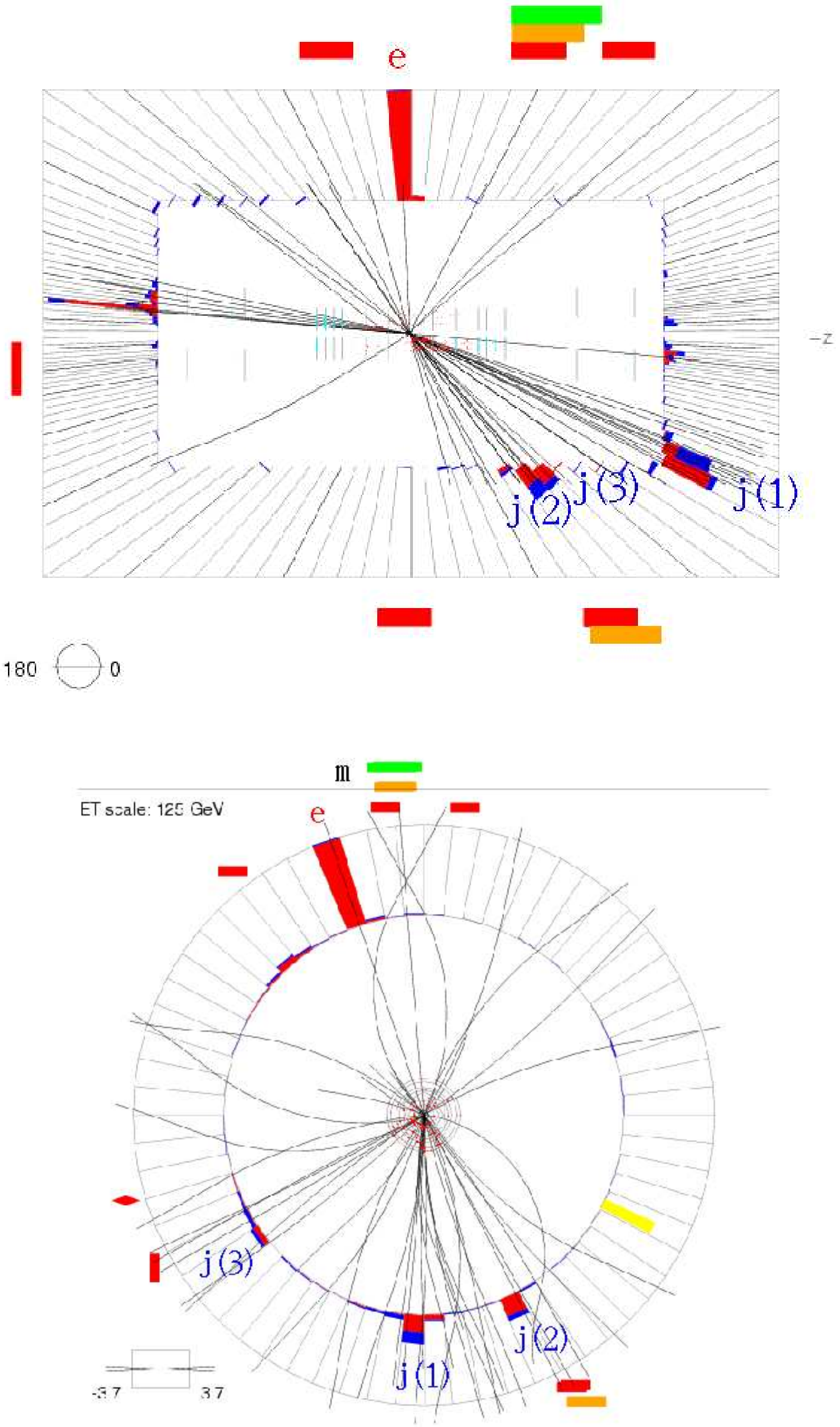


Figure 7.2: The event 1e 1μ 3j in RZ view (top) and XY view (bottom).

the dominant part and the SM prediction agrees well with data. A difference between these two cases is that jets which are close to electrons ($\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \leq 0.5$) are automatically removed from the sample by the reconstruction software TopAnalyze. Although the same cut on ΔR between muons and jets is tested the excess can not be reduced. The figure 7.5 shows the distribution of ΔR including the five most contributing Monte Carlo samples (W- and Z-processes) and the QCD contribution. The distribution clearly peaks at π, indicating that muons and jets are well separated in this event class. A small excess is observed at zero, but the main excess is found at π. Although the muons are isolated, according to the isolation criterion of section 4.1.2, figure 7.6 shows that still a lot of jets are spatially close to a muon. In this figure, the ratio of the p_T of isolated muons to the p_T of jets as a function of the jet p_T for the case that both are spatially close to each other ($\Delta R < 0.5$) is given. Also, altering the isolation criterion from the default p_T-dependency (cf. section 4.1.2) to an isolation criterion with fixed boundaries for the calorimeter energy in a hollow cone ($0.1 < R < 0.4$) around the muon and for the sum of p_T of all tracks excluding the muon track in a cone ($R < 0.5$) around the muon

$$E_{cone}^{cal} < 2.5 \ GeV \tag{7.3}$$

$$\sum_{tracks} p_T^{cone} < 2.5 \ GeV \tag{7.4}$$

does not change the picture. Further observables are studied, as in figure 7.7, where the p_T-distribution of the muon (of the inclusive 1μ 1j + X event class) is given. The shape of the $W \rightarrow \mu\nu$ contribution is not in good agreement with data and exceeds the data points in the range $30 - 40 \ GeV$. The equivalent distribution for the jet-p_T is illustrated in figure 7.8. Same results are found independently by the author of reference [4]. This disagreement occurs in the 1μ 1j event class only, and in lack of better methods, the applied selection criteria and QCD estimation are retained for this study.

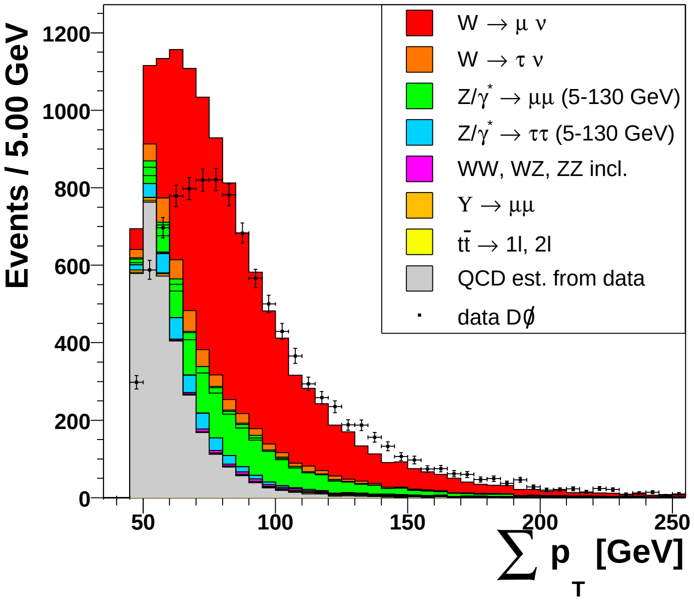


Figure 7.3: The inclusive event class 1μ 1j + X: sum of p_T-distributions including all background contributions.

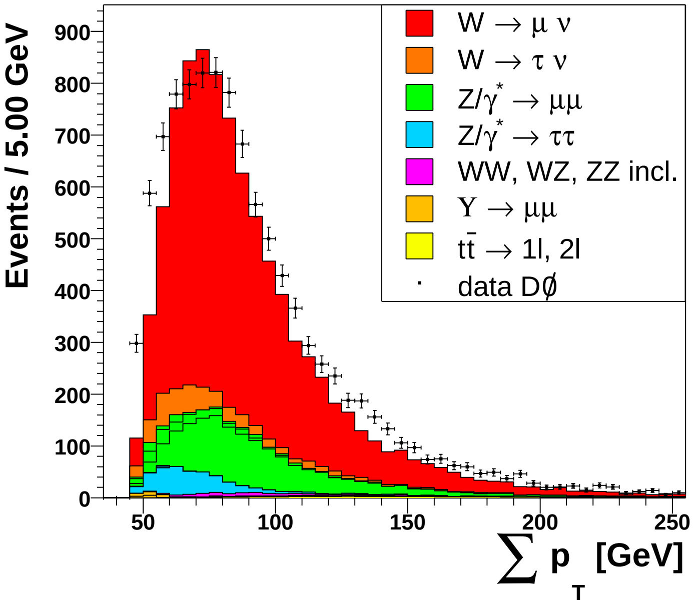


Figure 7.4: The inclusive event class 1μ 1j + X: sum of p_T-distributions without the QCD contributions.

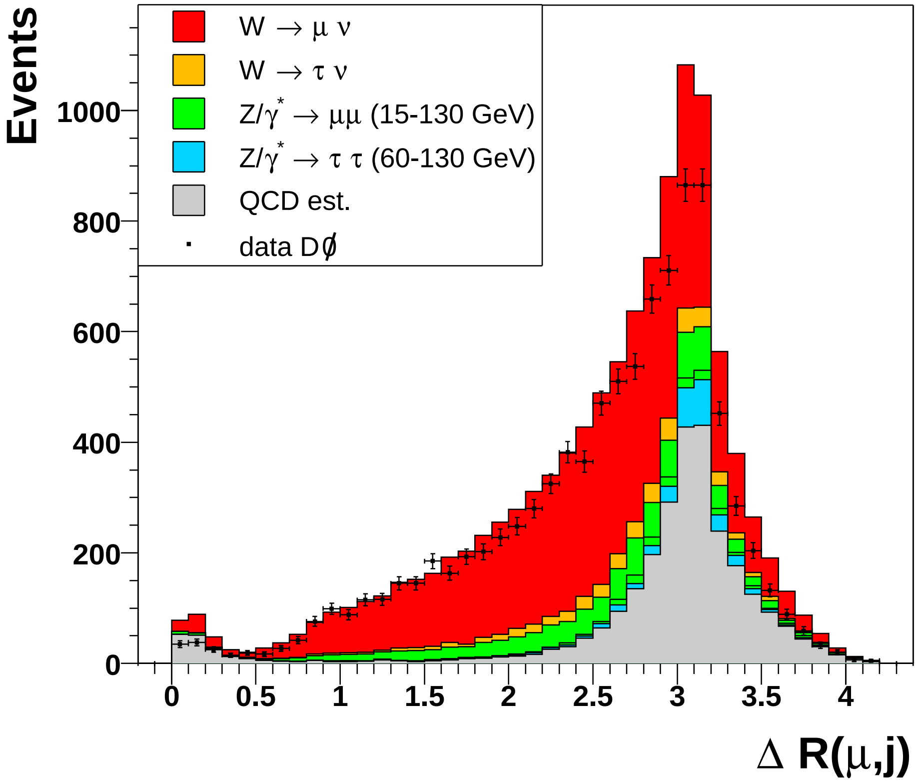


Figure 7.5: Distance ΔR between selected muon and jet of event class 1μ 1j.

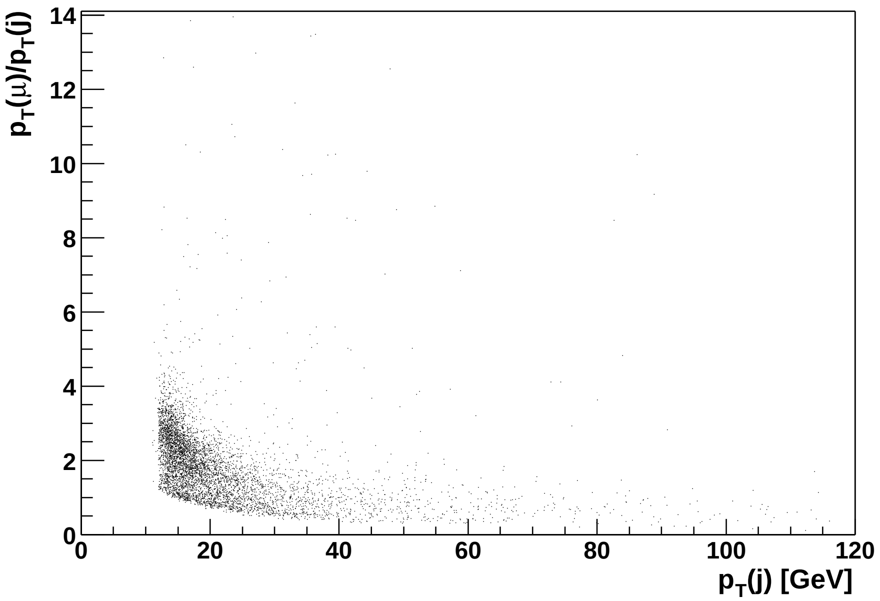


Figure 7.6: Ratio of muon p_T to jet p_T vs. jet p_T for events with jets spatially close to muons ($\Delta R < 0.5$) of event class $1\mu\ +\ X$.

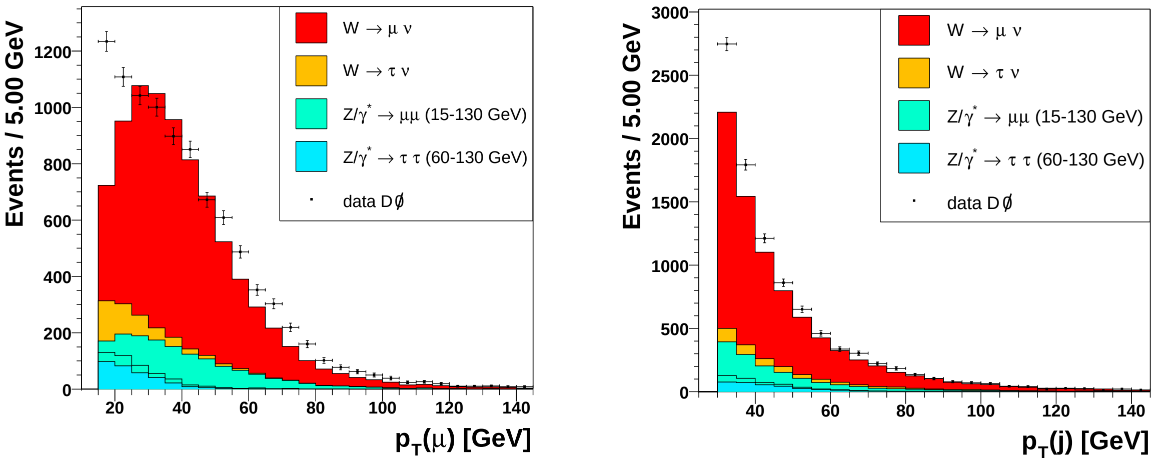


Figure 7.7: p_T of the muon of event class $1\mu\ 1j + X$.

Figure 7.8: p_T of the jet of event class $1\mu\ 1j + X$.

exclusive event classes of electron dataset: part 1							
Class	Data	sum of MC (+ stat. error)	$W \rightarrow e\nu$	$W \rightarrow \tau\nu$	$Z/\gamma^* \rightarrow ee$ (5 – 15 GeV)	$Z/\gamma^* \rightarrow ee$ (15 – 60 GeV)	$Z/\gamma^* \rightarrow ee$ (60 – 130 GeV)
1e	163902	159880 ±890	119930 ±170	1230 ±20	< 25	118 ±10	12210 ±50
2e	6444	6400 ±40	1.5 ±0.6	n.c.	< 4.4	56.2 ±6.8	6300 ±30
3e	1	$1.2\ ^{+1.6}_{-0.5}$	n.c.	n.c.	n.c.	< 0.9	0.8 ±0.4
1e 1j	32776	32740 ±730	7530 ±40	200 ±10	26.1 ±10.7	86.6 ±8.9	2760 ±20
1e 2j	2308	2025 ±50	421 ±10	13 ±2.0	< 5.0	< 4.3	151
2e 1j	420	$401\ ^{+11}_{-9}$	< 0.5	n.c.	n.c.	21.2 ±4.2	374 ±8
2e 2j	38	$23.5\ ^{+3.4}_{-1.9}$	< 0.2	n.c.	n.c.	< 0.9	24.4±1.9
1e 3j	207	$172\ ^{+6}_{-5}$	19.4 ±2.2	1.6 ±0.7	n.c.	n.c.	8.6 ±1.3
2e 3j	2	$0.2\ ^{+1.6}_{-0.2}$	n.c.	n.c.	n.c.	n.c.	< 1.1
1e 4j	21	$25.5\ ^{+2.4}_{-0.6}$	< 0.6	< 0.4	n.c.	n.c.	< 0.7
1e 5j	2	$1.0\ ^{+0.6}_{-0.1}$	< 0.3	n.c.	n.c.	n.c.	< 0.2
1e 1μ	25	$17.9\ ^{+2.1}_{-1.7}$	< 0.2	n.c.	n.c.	n.c.	< 0.2
1e 1μ 1j	2	$4.8\ ^{+1.6}_{-0.1}$	n.c.	n.c.	n.c.	n.c.	n.c.
1e 1μ 2j	5	$1.8\ ^{+0.6}_{-0.2}$	n.c.	n.c.	n.c.	n.c.	n.c.
1e 1μ 3j	1	$0.11\ ^{+0.04}_{-0.04}$	n.c.	n.c.	n.c.	n.c.	n.c.

Table 7.2: Data-MC comparison and first MC-contributions for exclusive event classes.

exclusive event classes of electron dataset: part 2							
Class	$Z/\gamma^* \to \tau\tau$ (15 – 60 GeV)	$Z/\gamma^* \to \tau\tau$ (60 – 130 GeV)	WW→X	ZZ→X	WZ→X	$t\bar{t}$ →2lX	$t\bar{t}$ →1lX
1e	2.5 ±1.0	313 ±8	66.5 ±3.6	13.9 ±0.8	2.9 ±0.2	1.6 ±0.2	0.9 ±0.1
2e	< 0.4	14.7 ±1.6	5.5 ±1.0	2.0 ±0.3	1.8 ±0.2	0.4 ±0.1	n.c.
3e	n.c.	< 0.2	< 0.2	0.2 ±0.1	0.14 ±0.05	< 0.01	n.c.
1e 1j	< 2.4	153 ±6	69.7 ±3.6	11.7 ±0.8	2.9 ±0.2	10.5 ±0.4	9.0 ±0.3
1e 2j	n.c.	12.3 ±1.6	30.4 ±2.4	6.7 ±0.6	2.0 ±0.2	12.4 ±0.4	28.9 ±0.6
2e 1j	n.c.	1.1 ±0.4	< 0.8	1.7	1.4	1.9	< 0.04
2e 2j	n.c.	< 0.2	< 0.2	0.7 ±0.2	0.6 ±0.1	1.7 ±0.2	n.c.
1e 3j	n.c.	< 1.0	2.1 ±0.6	0.5 ±0.2	0.2 ±0.1	2.5 ±0.2	32.7 ±0.7
2e 3j	n.c.	n.c.	< 0.2	< 0.2	< 0.1	0.13 ±0.04	< 0.03
1e 4j	n.c.	< 0.2	n.c.	< 0.1	n.c.	0.17 ±0.05	12.3 ±0.4
1e 5j	n.c.	n.c.	n.c.	n.c.	n.c.	< 0.02	1.1 ±0.1
1e 1μ	< 0.4	11.6 ±1.4	5.4 ±0.9	0.6 ±0.2	0.11 ±0.05	0.3 ±0.1	< 0.01
1e 1μ 1j	n.c.	< 0.8	< 0.6	< 0.1	< 0.1	2.1 ±0.2	n.c.
1e 1μ 2j	n.c.	< 0.2	< 0.2	n.c.	< 0.02	1.8 ±0.2	n.c.
1e 1μ 3j	n.c.	n.c.	n.c.	n.c.	n.c.	0.11 ±0.04	n.c.

Table 7.3: Second part of MC-contributions for exclusive event classes.

exclusive event classes of electron dataset: part 3			
Class	$\Upsilon(1S) \rightarrow ee$	$\Upsilon(2S) \rightarrow ee$	QCD
1e	6.1 ±2.7	8.2 ±3.4	25980 ±860
2e	< 1.2	< 1.4	13.3 ±0.4
3e	n.c.	n.c.	n.c.
1e 1j	n.c.	n.c.	21880 ±730
1e 2j	n.c.	n.c.	1350 ±50
2e 1j	n.c.	n.c.	n.c.
2e 2j	n.c.	n.c.	n.c.
1e 3j	n.c.	n.c.	105 ±4
2e 3j	n.c.	n.c.	n.c.
1e 4j	n.c.	n.c.	13.0 ±0.4
1e 5j	n.c.	n.c.	n.c.
1e 1μ	n.c.	n.c.	n.c.
1e 1μ 1j	n.c.	n.c.	2.7 ±0.1
1e 1μ 2j	n.c.	n.c.	n.c.
1e 1μ 3j	n.c.	n.c.	n.c.

Table 7.4: Rest of MC-contributions for exclusive event classes.

inclusive event classes of electron dataset: part 1

Class	Data	sum of MC (+ stat. error)	$W \to e\nu$	$W \to \tau\nu$	$Z/\gamma^* \to ee$ (5 – 15 GeV)	$Z/\gamma^* \to ee$ (15 – 60 GeV)	$Z/\gamma^* \to ee$ (60 – 130 GeV)
1e + X	206154	202630 ±1650	127900 ±180	1440 ±20	47.9 ±14.4	295 ±17	22700 ±70
2e + X	6905	6820 ±50	1.7 ±0.6	n.c.	< 9.1	77.4 ±7.9	6700 ±35
3e + X	1	$1.2\ ^{+1.7}_{-0.4}$	n.c.	n.c.	n.c.	< 0.9	0.8 ±0.4
1e 1j + X	35782	35460 ±790	7970 ±40	217 ±8	30.5 ±11.5	113 ±10	3360 ±30
2e 1j + X	460	$425\ ^{+11}_{-10}$	< 0.5	< 0.3	n.c.	21.2 ±4.2	395 ±9
1e 2j + X	2584	2250 ±60	441 ±10	14.5 ±2.1	< 5.0	< 4.3	184 ±6
2e 2j + X	40	$24.4\ ^{+3.3}_{-2.0}$	n.c.	n.c.	n.c.	< 0.9	21.1 ±2.0
1e 3j + X	233	$200\ ^{+8}_{-5}$	19.6 ±2.2	1.6 ±0.7	n.c.	n.c.	9.8 ±1.4
2e 3j + X	2	$0.2\ ^{+1.8}_{-0.1}$	n.c.	n.c.	n.c.	n.c.	< 1.1
1e 4j + X	23	$26.6\ ^{+2.9}_{-0.6}$	< 0.6	< 0.4	n.c	n.c.	< 0.2
1e 5j + X	2	$1.1\ ^{+0.7}_{-0.1}$	< 0.3	n.c.	n.c.	n.c.	< 0.2
1e 1μ + X	33	$23.2\ ^{+2.3}_{-1.8}$	n.c.	n.c.	n.c.	n.c.	< 0.2
1e 1μ 1j + X	8	$5.6\ ^{+0.6}_{-0.1}$	n.c.	n.c.	n.c.	n.c.	n.c.
1e 1μ 2j + X	6	$1.86\ ^{+0.7}_{-0.2}$	n.c.	n.c.	n.c.	n.c.	n.c.
1e 1μ 3j + X	1	$0.11\ ^{+0.04}_{-0.04}$	n.c.	n.c.	n.c.	n.c.	n.c.

Table 7.5: Data-MC comparison and first MC-contributions for inclusive event classes.

inclusive event classes of electron dataset: part 2

Class	$Z/\gamma^* \to \tau\tau$ (15 – 60 GeV)	$Z/\gamma^* \to \tau\tau$ (60 – 130 GeV)	WW→X	ZZ→X	WZ→X	$t\bar{t}$ →2lX	$t\bar{t}$ →1lX
1e + X	4.2 ±1.3	511 ±10	182 ±6	1.4 ±1.4	12.8 ±0.5	36.7 ±0.7	85.0 ±1.1
2e + X	< 0.4	15.8 ±1.7	6.1 ±1.0	0.5 ±0.5	4.1 ±0.3	4.1 ±0.2	0.09 ±0.03
3e + X	n.c.	< 0.2	< 0.2	0.1 ±0.1	0.2 ±0.1	< 0.01	n.c.
1e 1j + X	< 2.4	168 ±6	103 ±4	1.0 ±1.0	7.5 ±0.4	34.3 ±0.7	84.1 ±1.1
2e 1j + X	n.c.	1.1 ±0.4	< 0.8	0.3 ±0.3	2.1 ±0.2	3.7 ±0.2	0.09 ±0.03
1e 2j + X	< 0.5	12.9 ±1.6	32.5 ±2.5	< 0.6	3.0 ±0.3	19.3 ±0.5	75.1 ±1.0
2e 2j + X	n.c.	< 0.2	< 0.2	0.2 ±0.2	0.7 ±0.1	1.8 ±0.2	0.07 ±0.03
1e 3j + X	n.c.	< 1.0	2.1 ±0.6	0.2 ±0.2	0.3 ±0.1	3.0 ±0.2	46.1 ±0.8
2e 3j + X	n.c.	n.c.	< 0.2	< 0.3	< 0.1	0.13 ±0.04	< 0.04
1e 4j + X	n.c.	< 0.2	< 0.2	< 0.1	< 0.1	0.17 ±0.05	13.4 ±0.4
1e 5j + X	n.c.	n.c.	n.c.	< 0.1	n.c.	< 0.02	1.1 ±0.1
1e 1μ + X	< 0.4	12.1 ±1.5	5.7 ±1.0	0.2 ±0.2	0.2 ±0.1	4.2 ±0.2	< 0.1
1e 1μ 1j + X	n.c.	< 0.2	< 0.2	< 0.1	< 0.02	3.95 ±0.01	n.c.
1e 1μ 2j + X	n.c.	< 0.2	< 0.2	< 0.1	< 0.02	1.86 ±0.01	n.c.
1e 1μ 3j + X	n.c.	n.c.	n.c.	n.c.	n.c.	0.1 ±0.2	n.c.

Table 7.6: Second part of MC-contributions for inclusive event classes.

inclusive event classes of electron dataset: part 3			
Class	$\Upsilon(1S) \rightarrow ee$	$\Upsilon(2S) \rightarrow ee$	QCD
1e + X	7.3 ±3.0	13.7 ±4.3	49350 ±1640
2e + X	< 1.2	< 1.4	11.9 ±0.4
3e + X	n.c.	n.c.	n.c.
1e 1j + X	< 2.9	< 7.9	23350 ±780
2e 1j + X	n.c.	n.c.	n.c.
1e 2j + X	< 1.4	< 3.2	1460 ±50
2e 2j + X	n.c.	n.c.	n.c.
1e 3j + X	n.c.	< 1.6	118 ±4
2e 3j + X	n.c.	n.c.	n.c.
1e 4j + X	n.c.	n.c.	13.0 ±0.4
1e 5j + X	n.c.	n.c.	n.c.
1e 1μ + X	n.c.	n.c.	n.c.
1e 1μ 1j + X	n.c.	n.c.	1.7 ±0.1
1e 1μ 2j + X	n.c.	n.c.	n.c.
1e 1μ 3j + X	n.c.	n.c.	n.c.

Table 7.7: Rest of MC-contributions for inclusive event classes.

exclusive event classes of muon dataset: part 1

Class	Data	sum of MC (+ stat. error)	$W \to \mu\nu$	$W \to \tau\nu$	$Z/\gamma^* \to \mu\mu$ (5 – 15 GeV)	$Z/\gamma^* \to \mu\mu$ (15 – 60 GeV)	$Z/\gamma^* \to \mu\mu$ (60 – 130 GeV)
1μ	155050	150150 ±2670	110990 ±150	6080 ±40	465 ±40	2400 ±30	15040 ±50
2μ	3573	3740 ±40	< 0.5	n.c.	< 8.2	244 ±9	3460 ±20
1e 1μ	82	$77.4^{+5.3}_{-3.4}$	19.0 ±1.9	n.c.	n.c.	< 0.8	9.6 ±1.2
1e 2μ	1	$2.5^{+1.6}_{-0.6}$	n.c.	n.c.	n.c.	< 0.4	2.5 ±0.6
1μ 1j	9368	11300 ±570	6349 ±36	454 ±11	65 ±15	148 ±7	994 ±13
2μ 1j	194	222^{+7}_{-6}	n.c.	n.c.	n.c.	17.2 ±2.5	200.9 ±5.7
1μ 2j	892	769^{+58}_{-55}	332 ±8	27 ±3	n.c.	4.2 ±1.2	58.7 ±3.1
2μ 2j	19	$14.8^{+2.3}_{-1.5}$	n.c.	n.c.	n.c.	< 0.4	13.3 ±1.5
1μ 3j	84	$64.5^{+5.3}_{-4.1}$	14.3 ±1.7	1.3 ±0.6	n.c.	< 1.2	3.4 ±0.7
2μ 3j	2	$0.14^{+0.59}_{-0.04}$	n.c.	n.c.	n.c.	n.c.	< 0.4
1μ 4j	9	$11.9^{+1.78}_{-0.6}$	1.0 ±0.4	< 0.3	n.c.	n.c.	< 0.2
1μ 5j	1	$0.9^{+0.5}_{-0.1}$	< 0.2	n.c.	n.c.	n.c.	n.c.
1e 1μ 1j	5	$9.2^{+4.9}_{-1.0}$	2.4 ±0.7	n.c.	n.c.	< 3.7	0.9 ±0.4
1e 1μ 2j	5	$2.2^{+1.4}_{-0.2}$	< 0.6	n.c.	n.c.	n.c.	< 0.2
1e 1μ 3j	1	$0.16^{+0.04}_{-0.04}$	n.c.	n.c.	n.c.	n.c.	n.c.

Table 7.8: Data-MC comparison and first MC-contributions for exclusive event classes.

exclusive event classes of muon dataset: part 2								
Class	$Z/\gamma^* \to \tau\tau$ ($5-15$ GeV)	$Z/\gamma^* \to \tau\tau$ ($15-60$ GeV)	$Z/\gamma^* \to \tau\tau$ ($60-130$ GeV)	WW→X	ZZ→X	WZ→X	$t\bar{t}$ →2lX	$t\bar{t}$ →1lX
1μ	< 14.3	30.3 ±3.2	977 ±13	56.2 ±2.9	14.0 ±0.7	3.0 ±0.2	1.1 ±0.1	0.4 ±0.1
2μ	< 2.9	< 0.4	20.4 ±1.8	3.4 ±0.7	1.4 ±0.2	0.8 ±0.1	0.18 ±0.05	< 0.011
1e 1μ	n.c.	< 1.1	34.8 ±2.4	7.1 ±1.0	0.8 ±0.2	0.14 ±0.05	0.3 ±0.1	< 0.01
1e 2μ	n.c.	n.c.	< 0.2	< 0.2	< 0.2	< 0.02	< 0.01	n.c.
1μ 1j	n.c.	4.0 ±1.2	322 ±7	56.9 ±2.9	10.8 ±0.7	3.0 ±0.2	6.8 ±0.3	4.4 ±0.2
2μ 1j	n.c.	n.c.	1.5 ±0.5	< 0.4	0.8 ±0.2	0.7 ±0.1	0.7 ±0.1	n.c.
1μ 2j	n.c.	< 0.8	19.5 ±1.8	22.9 ±1.9	4.9 ±0.4	1.6 ±0.2	9.0 ±0.3	17.6 ±0.4
2μ 2j	n.c.	n.c.	< 0.2	< 0.2	0.2 ±0.1	0.3 ±0.1	1.0 ±0.1	n.c.
1μ 3j	n.c.	n.c.	1.0 ±0.4	0.9 ±0.4	0.3 ±0.1	0.2 ±0.1	2.0 ±0.2	23.5 ±0.5
2μ 3j	n.c.	n.c.	n.c.	n.c.	< 0.05	< 0.1	0.14 ±0.04	n.c.
1μ 4j	n.c.	n.c.	< 0.2	< 0.4	< 0.04	< 0.04	0.18 ±0.04	9.3 ±0.3
1μ 5j	n.c.	n.c.	n.c.	< 0.2	n.c.	n.c.	< 0.01	0.8 ±0.1
1e 1μ 1j	n.c.	n.c.	2.4 ±0.6	0.7 ±0.3	< 0.1	< 0.1	2.2 ±0.2	n.c.
1e 1μ 2j	n.c.	n.c.	< 0.2	< 0.2	n.c.	n.c.	2.0 ±0.2	n.c.
1e 1μ 3j	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	0.16 ±0.04	n.c.

Table 7.9: Second part of MC-contributions for exclusive event classes.

exclusive event classes of muon dataset: part 3			
Class	$\Upsilon(1S) \rightarrow \mu\mu$	$\Upsilon(2S) \rightarrow \mu\mu$	QCD
1μ	397 ±19	385 ±20	13310 ±2660
2μ	< 4.3	< 4.7	3.9 ±0.8
1e 1μ	n.c.	n.c.	2.0 ±0.4
1e 2μ	n.c.	n.c.	n.c.
1μ 1j	20.2 ±4.3	9.0 ±3.0	2850 ±570
2μ 1j	n.c.	n.c.	0.6 ±0.1
1μ 2j	< 1.1	< 1.2	272 ±54
2μ 2j	n.c.	n.c.	n.c.
1μ 3j	n.c.	n.c.	17.7 ±3.5
2μ 3j	n.c.	n.c.	n.c.
1μ 4j	n.c.	n.c.	1.4 ±0.3
1μ 5j	n.c.	n.c.	0.18 ±0.04
1e 1μ 1j	n.c.	n.c.	0.6 ±0.1
1e 1μ 2j	n.c.	n.c.	0.17 ±0.03
1e 1μ 3j	n.c.	n.c.	n.c.

Table 7.10: Rest of MC-contributions for exclusive event classes.

inclusive event classes of muon dataset: part 1							
Class	Data	sum of MC (+ stat. error)	$W \to \mu\nu$	$W \to \tau\nu$	$Z/\gamma^* \to \mu\mu$ (5 – 15 GeV)	$Z/\gamma^* \to \mu\mu$ (15 – 60 GeV)	$Z/\gamma^* \to \mu\mu$ (60 – 130 GeV)
1μ + X	169286	166680 ±3300	117710 ±150	6560 ±40	536 ±43	2810 ±30	19720 ±60
2μ + X	3789	3980 ±40	< 0.5	n.c.	< 12.2	261 ±10	3680 ±30
1e 1μ + X	94	$97.0\ ^{+7.8}_{-3.7}$	21.7 ±2.0	n.c.	n.c.	< 0.8	13.2 ±1.4
1e 2μ + X	1	$2.8\ ^{+1.6}_{-0.6}$	n.c.	n.c.	n.c.	< 0.4	2.8 ±0.6
1μ 1j + X	10580	12460 ±630	6700 ±40	480 ±11	68.3 ±15.3	169 ±8	1270 ±10
1μ 2j + X	1013	$893\ ^{+66}_{-59}$	348 ±8	27.7 ±2.6	< 3.9	4.9 ±1.3	75.5 ±3.5
2μ 1j + X	215	$239\ ^{+7}_{-6}$	n.c.	n.c.	n.c.	17.2 ±2.5	215 ±6
2μ 2j + X	21	$15.5\ ^{+2.3}_{-1.5}$	n.c.	n.c.	n.c.	< 0.4	13.7 ±1.5
1μ 3j + X	97	$90.0\ ^{+5.6}_{-4.4}$	15.3 ±1.7	1.3 ±0.6	n.c.	< 1.2	3.5 ±0.7
2μ 3j + X	2	$0.14\ ^{+0.59}_{-0.04}$	n.c.	n.c.	n.c.	n.c.	< 0.4
1μ 4j + X	10	$16.3\ ^{+1.7}_{-0.6}$	1.0 ±0.4	< 0.3	n.c.	n.c.	< 0.2
1μ 5j + X	1	$1.2\ ^{+0.3}_{-0.1}$	n.c.	n.c.	n.c.	n.c.	n.c.
1e 1μ 1j + X	11	$11.0\ ^{+1.4}_{-1.1}$	2.3 ±0.7	n.c.	n.c.	n.c.	1.0 ±0.4
1e 1μ 2j + X	6	$2.0\ ^{+1.9}_{-0.2}$	< 0.6	n.c.	n.c.	n.c.	< 0.3
1e 1μ 3j + X	1	$0.14\ ^{+0.04}_{-0.04}$	n.c.	n.c.	n.c.	n.c. n.c.	

Table 7.11: Data-MC comparison and first MC-contributions for inclusive event classes.

inclusive event classes of muon dataset: part 2

Class	$Z/\gamma^* \to \tau\tau$ (5 – 15 GeV)	$Z/\gamma^* \to \tau\tau$ (15 – 60 GeV)	$Z/\gamma^* \to \tau\tau$ (60 – 130 GeV)	WW→X	ZZ→X	WZ→X	$t\bar{t}$ →2lX	$t\bar{t}$ →1lX
1μ +X	12.4 ±5.6	35.3 ±3.5	1390 ±15	199 ±5	45.0 ±1.2	13.1 ±0.4	26.3 ±0.6	74.9 ±0.8
2μ +X	n.c.	< 0.4	21.9 ±1.9	4.7 ±0.7	3.4 ±0.3	2.4 ±0.2	2.0 ±0.2	< 0.03
1e 1μ +X	n.c.	< 1.1	41.0 ±2.5	11.4 ±1.0	1.8 ±0.2	0.3 ±0.1	4.7 ±0.2	0.09 ±0.02
1e 2μ +X	n.c.	n.c.	< 0.2	< 0.2	< 0.2	< 0.04	< 0.01	n.c.
1μ 1j +X	n.c.	4.4 ±1.2	347 ±8	109 ±4	22.8 ±0.8	7.8 ±0.3	24.7 ±0.5	74.4 ±0.8
1μ 2j +X	n.c.	< 0.4	20.7 ±1.8	32.0 ±1.9	7.1 ±0.4	2.9 ±0.2	14.7 ±0.4	68.5 ±0.4
2μ 1j +X	n.c.	n.c.	1.5 ±0.5	< 0.4	1.3 ±0.2	1.3 ±0.1	1.9 ±0.2	n.c.
2μ 2j +X	n.c.	n.c.	< 0.2	< 0.2	0.3 ±0.1	0.4 ±0.1	1.2 ±0.1	n.c.
1μ 3j +X	n.c.	n.c.	1.1 ±0.4	1.4 ±0.4	0.4 ±0.1	0.3 ±0.1	2.5 ±0.2	44.9 ±0.6
2μ 3j +X	n.c.	n.c.	n.c.	n.c.	< 0.05	< 0.1	0.15 ±0.04	n.c.
1μ 4j +X	n.c.	n.c.	< 0.2	< 0.4	n.c.	n.c.	0.19 ±0.05	13.5 ±0.3
1μ 5j +X	n.c.	n.c.	n.c.	< 0.2	n.c.	n.c.	< 0.01	1.0 ±0.1
1e1μ1j+X	n.c.	n.c.	2.1 ±0.6	1.0 ±0.3	< 0.2	< 0.1	3.7 ±0.2	0.09 ±0.02
1e1μ2j+X	n.c.	n.c.	< 0.4	< 0.2	< 0.04	< 0.02	1.8 ±0.2	< 0.1
1e1μ3j+X	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	0.14 ±0.04	n.c.

Table 7.12: Second part of MC-contributions for inclusive event classes.

inclusive event classes of muon dataset: part 3			
Class	$\Upsilon(1S) \to \mu\mu$	$\Upsilon(2S) \to \mu\mu$	QCD
1μ + X	561 ±20	531 ±20	16460 ±3290
2μ + X	< 5.4	< 5.9	4.5 ±0.9
1e 1μ + X	n.c.	< 2.2	2.8 ±0.6
1e 2μ + X	n.c.	n.c.	n.c.
1μ 1j + X	28.1 ±4.4	12.0 ±3.0	3140 ±630
1μ 2j + X	< 1.0	< 1.2	292 ±58
2μ 1j + X	n.c.	n.c.	0.6 ±0.1
2μ 2j + X	n.c.	n.c.	n.c.
1μ 3j + X	n.c.	n.c.	19.3 ±3.9
2μ 3j + X	n.c.	n.c.	n.c.
1μ 4j + X	n.c.	n.c.	1.6 ±0.3
1μ 5j + X	n.c.	n.c.	0.18 ±0.04
1e 1μ 1j + X	n.c.	n.c.	0.8 ±0.2
1e 1μ 2j + X	n.c.	n.c.	0.17 ±0.03
1e 1μ 3j + X	n.c.	n.c.	n.c.

Table 7.13: Rest of MC-contributions for inclusive event classes.

Chapter 8

Search for Deviations with the Algorithm

This chapter deals with the search algorithm which systematically investigates all event classes and determines the significance of a deviation found. This part has been inspired by [10] and [11].

8.1 The Search Algorithm

The algorithm is divided in two subroutines. In a first step, all possible **regions** in a given one-dimensional histogram are defined. A region consists of connected histogram bins where all bins are in p_T 5 GeV wide. In [10] a reduction of the bin size has been studied which turned out to have a negligible effect on the results. Figure 8.1 illustrates the definition of regions. Blue indicates the data distribution, whereas the sum of all Monte Carlo simulations is shown in red. Three regions are framed in green: bin 2-3, bin 2-4 and bin 10-12. The algorithm takes into account regions of any width and at any position and is thus sensitive to narrow resonances and single peaks as well as to signals spread over large regions. In each region, the number of observed events N_{data} is determined. Also, the number of expected events N_{SM}, predicted by the Standard Model and given as the sum over all MC contributions plus QCD, and the statistical error and systematical uncertainties of N_{SM} in the same region are calculated.

These values are then used in a **probability calculation** which evaluates the level of agreement between data and SM prediction in a distribution of a particular event class. A statistical estimator p is used to find the region of highest interest. This p-value is defined by the convolution of two probability density functions (PDF): a Poisson PDF to account for statistical errors and a Gaussian PDF with mean N_{SM} and width δN_{SM} to

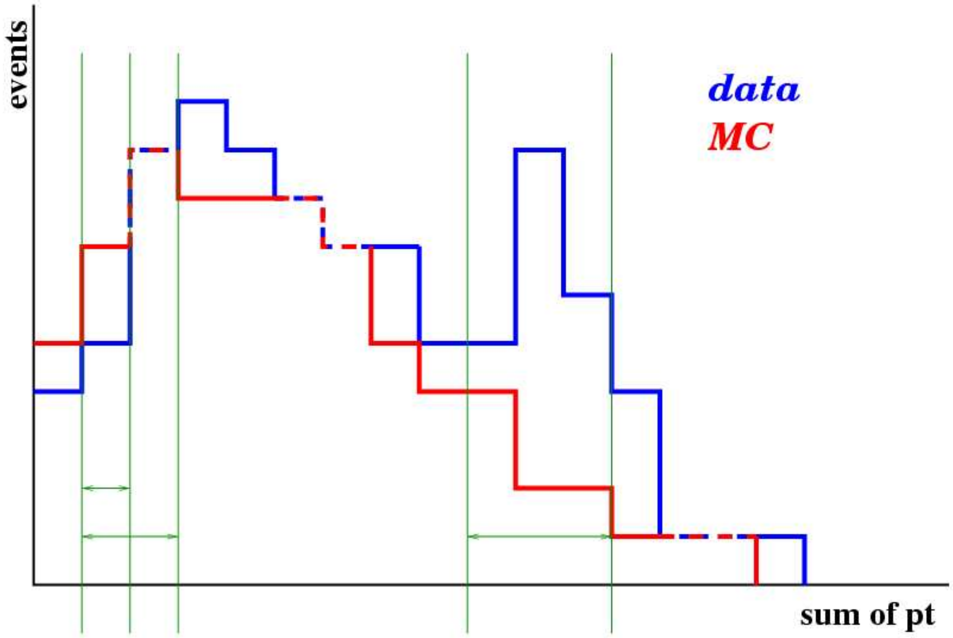


Figure 8.1: Definition of regions in a histogram.

include systematic uncertainties

$$p = \begin{cases} A \cdot \int\limits_0^\infty db\ \mathrm{Gauss}(\mathrm{b}, \mathrm{N_{SM}}, \delta \mathrm{N_{SM}}) \sum\limits_{\mathrm{i}=\mathrm{N_{data}}}^{\infty} \frac{\mathrm{e^{-b}b^i}}{\mathrm{i!}} & \text{if} \quad \mathrm{N_{SM}} \leq \mathrm{N_{data}} \\ A \cdot \int\limits_0^\infty db\ \mathrm{Gauss}(\mathrm{b}, \mathrm{N_{SM}}, \delta \mathrm{N_{SM}}) \sum\limits_{\mathrm{i}=0}^{\mathrm{N_{data}}} \frac{\mathrm{e^{-b}b^i}}{\mathrm{i!}} & \text{if} \quad \mathrm{N_{SM}} < \mathrm{N_{data}} \end{cases} \tag{8.1}$$

$$\text{where} \quad \mathrm{A} = \left[\int\limits_0^\infty \mathrm{db\ Gauss}(\mathrm{b}, \mathrm{N_{SM}}, \delta \mathrm{N_{SM}}) \sum_{\mathrm{i}=0}^{\infty} \frac{\mathrm{e^{-b}b^i}}{\mathrm{i!}} \right]^{-1} \tag{8.2}$$

ensures the normalisation to unity. Thus, for a region the probability is calculated, that the expected number of events N_{SM} fluctuates upwards ($N_{SM} \leq N_{data}$) or downwards ($N_{SM} > N_{data}$) to the number of events observed in data N_{data}. After determining p for each region, the algorithm arranges the p-values in decreasing order, i.e. greatest deviation. The **region of greatest deviation** is the one having the smallest p-value, p_{min}.

The second subroutine determines the **significance of a deviation per event class**. In order to ensure that the deviation found by the first subroutine is not a fluctuation, the experiment is repeated many times by producing hypothetical data histograms (*HDH*). For a hypothetical data histogram, the number of entries in every bin is generated according to the convolution of the probability density functions (equation 8.1). This new distribution is then compared with the original SM prediction. The algorithm is run to find the region of greatest deviation and to calculate the according p-value, $p_{\mathrm{min}}^{\mathrm{HDH}}$. One should realise, that again a region of greatest deviation can occur at any position and be of any width in the investigated distribution. Hence, the significance per event class is given by the probability $\widetilde{P}$ to find a p-value $p_{\mathrm{min}}^{\mathrm{HDH}}$ equal or smaller than the $p_{\mathrm{min}}^{\mathrm{data}}$-value obtained from the comparison between real data and SM prediction:

$$\widetilde{P} = \frac{\text{number of HDS with} \quad p_{min}^{HDH} \leq p_{min}^{data}}{\text{total number of HDS}} \quad . \tag{8.3}$$

If $\widetilde{P}$ is close to 1, data and SM prediction agree very well. On the other hand, in an event class with very small $\widetilde{P}$, a similar (or larger) deviation was not found in the *HDH* distributions. Thus, the observed deviation between real data and SM prediction is highly interesting and should be investigated further. For most event classes, 100 hypothetical data histograms are generated and compared with the SM prediction, whereas for some event classes with a very small p-value the number of *HDH* is increased.

The width of the Gaussian PDF is given by the **total uncertainty** of N_{SM}, $\delta_{N_{SM}}$, which is given by

$$\delta_{N_{SM}}^2 = \delta_{p_T}^2 + \delta_{smear}^2 + \delta_{corr}^2 + \delta_{xs}^2 + \delta_{QCD}^2 + \delta_{stat}^2 \quad . \tag{8.4}$$

Subsequently, each contribution is described:
The uncertainties of the energy and momentum scale, δ_{p_T}, are considered to be correlated bin by bin and for all Monte Carlo samples; they are determined in the following way: First, the absolute errors σ_{abs} of the sum of transverse momenta are added to every bin entry of the whole spectrum and of all MC samples, giving $MC(\sigma_{up})$. In a second run, the errors are subtracted from every bin entry, giving $MC(\sigma_{down})$. According to error propagation the absolute error σ_{abs} of the sum of transverse momenta is derived via

$$\sigma_{abs}^2 = \sum_{k=1}^{3} \sum_{i=0}^{N_k} [p_{T,i}(k) \cdot \sigma(k)]^2 \tag{8.5}$$

where the index k represents the physics objects electron, muon and jet and the index i indicates the particular number of one such object. The transverse momentum of the i^{th} physics object k is given by $p_{T,i}(k)$ while $\sigma(k)$ is the corresponding relative error. The mean values MC_{mean} are then subtracted from $MC(\sigma_{up})$ and $MC(\sigma_{down})$:

$$\sigma_{up}^2 = \left[\sum_{MC} \sum_{bin} MC(\sigma_{up}) - \sum_{MC} \sum_{bin} MC_{mean} \right]^2 \tag{8.6}$$

$$\sigma_{down}^2 = \left[\sum_{MC} \sum_{bin} MC(\sigma_{down}) - \sum_{MC} \sum_{bin} MC_{mean} \right]^2 \tag{8.7}$$

$$\tag{8.8}$$

This is done for each region (or new bin entry to be generated for a *HDH*), so that $\sum\limits_{bin}$ is the sum over all bins of the region (the bin) in question. The errors are symmetrised by taking the larger one.
For the uncertainty of smearing, δ_{smear}, only σ_{up} can be derived since the original smearing is implemented in the reconstruction software. Therefore, the uncertainty is estimated by a second Gaussian of mean one and width σ_{smear}. The error δ_{smear} is symmetrised by assuming that its size is the same for both directions, i.e. up and down.
Since the MC correction factors are global scaling factors, the systematic uncertainties are also statistically correlated bin by bin and for all Monte Carlo samples, i.e. the direction of the uncertainty is the same for each bin and each MC sample. This error is called δ_{corr}.
Due to the correlation between the cross sections of W and Z the uncertainties are applied in the same way. This means, both cross sections increase or both decrease. On the other hand, the errors of other Monte Carlo simulations may tend in opposite directions. The processes WW, ZZ and WZ are treated in the same way as explained for W and Z processes. The single errors are added in quadrature giving δ_{xs}.
The two remaining terms of equation 8.4 are the uncertainty of the QCD estimation δ_{QCD} and the statistical error δ_{stat}.

The results of the search algorithm are presented and discussed in the next section.

8.2 Search Results

All event classes are passed through the algorithm to systematically search the DØ data for signals of new physics. The presented analysis quantifies the overall agreement between DØ data and the SM prediction. The $p_{\mathrm{min}}^{\mathrm{data}}$-values and their corresponding significances $\widetilde{P}$ are summarised in table 8.1 for all exclusive event classes of the electron dataset. In table 8.2 the results for the inclusive event classes are given. For exclusive and inclusive event classes of the muon dataset, the values are shown in the tables 8.3 and 8.4, respectively. In addition, the regions of greatest deviation as well as their corresponding number of observed events N_{data} and the number of expected events N_{SM} together with their uncertainties $\sigma(N_{SM})$ are given in these table. In the figures 8.2 to 8.5 the regions of greatest deviation in the sum of p_T-distributions are presented for those event classes with very low $\widetilde{P}$ values. In order to demonstrate the reliability of the algorithm and due to the plentifulness of event classes, the sum of p_T-distributions of only a few event classes with high $\widetilde{P}$ values are presented. Subsequently, the results are discussed.

Event classes of the electron dataset:
In the analysis of the electron dataset, the most significant deviations are found in the

event classes containing one and two electrons. In the exclusive event class 1e no p-value equal or less than $p^{data}_{min} = 7.7 \cdot 10^{-7}$ can be found in 1000 hypothetical data histograms *HDH* (cf. table 8.1). In the selected p_T region between 60 *GeV* and 80 *GeV*, 1655 data events are found compared to an expectation of 1230 ± 80 events. The p_T-distribution is given in figure 8.2 (top row), where the region is shown in an enlarged view on the right side. For the exclusive event class 2e, out of 500 generated *HDH* no similar or larger deviation than the calculated p^{data}_{min}-value of $6.2 \cdot 10^{-9}$ is found. Here, an excess is observed in a region above the Z boson peak between 130 *GeV* and 180 *GeV* (see figure 8.2, lower left). In this region, 42 data events are found, while only 11.6 ± 2.1 events are predicted by the SM.
Similarly, for the inclusive event classes 1e + X and 2e + X, $\widetilde{P}$ values of zero are found (see table 8.2). For the event class 1e + X, 500 *HDH* are generated in order to determine the significance of the calculated p^{data}_{min}-value of $4.2 \cdot 10^{-13}$ for the p_T region between 85 *GeV* and 125 *GeV* (figure 8.3, top row), where 1146 data events are found compared to 850 ± 26 expected events. In the 2e + X event class, 100 generated *HDH* are analysed. The p^{data}_{min}-value of $7.2 \cdot 10^{-12}$ corresponds to the $\sum p_T$ region between 125 *GeV* and 230 *GeV* (figure 8.3, second row) where 106 events are found in data and 43 ± 4 are expected.
In the exclusive 1e 3j event class (table 8.1), where a deviation has already been observed in the search via the total number of events in chapter 7, a $\widetilde{P}$ value of zero is found, corresponding to 200 generated *HDH* and $p^{data}_{min} = 2.2 \cdot 10^{-3}$. The region of greatest deviation is $\sum p_T = 255 - 375$ *GeV* with 73 data events compared to 49.7 ± 2.6 events predicted by the SM (shown in figure 8.2, lower right).
A low $\widetilde{P}$ value of 0.01 is found in the exclusive event class 1e 1j. The region of greatest deviation is found by the algorithm between 235 *GeV* and 240 *GeV* where a deficit is observed in the $\sum p_T$-distribution. The distribution is shown in the second row of figure 8.2. Compared to a SM expectation of 25.5 ± 1.9 only 7 data events are measured.

In the remaining event classes of the electron dataset, a good agreement between data and SM prediction is found. As expected, a moderate $\widetilde{P}$ value of 0.2 is found for the event class 1e 1μ 3j, which strengthens the hypothesis of chapter 7, that the one data event observed indeed is a $t\bar{t} \to 1e\ 1\mu\ 3j$ event.
Although low p^{data}_{min}-values are observed for several exclusive event classes (1e 2j, 1e 4j and 1e 1μ 2j; see table 8.1) and inclusive event classes (1e 1j + X, 1e 4j + X and 1e 1μ + X; see table 8.2) the determination of their significances yields high $\widetilde{P}$ values, i.e. the deviations can be explained by statistical fluctuations. Exemplarily, the $\sum p_T$-distributions of the event classes 1e 2j and 1e 1j + X are given in figure 8.3.

Hence, the introduced search method is able to find deviations between data and SM prediction. Furthermore, it is shown that this method quantifies the significances of these deviations in a correct way. In event classes where only a small deviation is found, the test of significance returned high $\widetilde{P}$ values, as it is expected. On the other hand, in event classes with huge deviations in up to 1000 generated *HDH* no similar deviation can be found. These deviations are found in regions of high transverse momentum, which is also expected, since the low p_T region has already been thoroughly investigated by a number of experiments, showing no sign of new physics. At this stage, it can be concluded that the search algorithm provides a reliable and powerful tool for a general determination of the level of agreement between data and SM prediction.

A comparison with results from the model-independent search in final states with leptons and missing transverse energy [4] can not be performed, since no final results are available from this analysis to date. It is not known within the DØ collaboration or from the other collider experiment at the Tevatron, CDF, that the significant deviations found in this analysis have been observed by other analysts as well. For Run I, a similar search strategy has been applied to $\approx 100\ pb^{-1}$ data colected by the DØ experiment, but no evidence for new high p_T physics has been observed [54].
The deviations in the event classes containing electrons and jets may be explained by an underestimation of the uncertainty of the QCD estimation, which is the dominant contribution in these event classes. Similarly, the uncertainty of the jet energy scale correction (section 6.1) may be underestimated. As a third explanation, gluon radiation is not well modelled by the event generator PYTHIA.
For the dielectron event class it is remarkable, that large deviations are also found in the same $\sum p_T$ region of the dimuon event class (see below).
As a conclusion on the results of the electron dataset it can be said, that the observed deviations are statistically significant and should be further investigated in dedicated analyses. Also, an extension of the search algorithm to additional distributions (e.g. invariant mass spectra and angular distributions between the physics objects) should be considered to learn more about the deviations, since new physics and detector effects may appear in the very same regions and are difficult to distinguish from each other. Thus, no evidence for new physics can be claimed to date.

Event classes of the muon dataset:
The analysis of the muon dataset shows a huge deviation in the event classes 1μ, both exclusive and inclusive. For the exclusive event class a p_{min}^{data}-value of $1.2 \cdot 10^{-61}$ is found in the p_T region between 85 GeV and 200 GeV, where 536 data events are measured compared to a SM prediction of 157 ± 13 events. In 600 generated *HDH* no p-value less or equal to p_{min}^{data} is observed (cf. table 8.3, top row of figure 8.4). The inclusive event class 1μ + X also shows an excess in a region of very high transverse momentum, between 130 GeV and 450 GeV, where 401 events are found in the dataset compared to a prediction of 111 ± 7 (cf. table 8.3; figure 8.5, upper left). This corresponds to a p_{min}^{data}-value of $6 \cdot 10^{-66}$. A $\widetilde{P}$ value of zero is found in the analysis of 500 generated *HDH*. These deviations reveal that data and SM prediction are not in good agreement at high transverse momenta of the muons. Since the high p_T region was found to be problematic in this analysis (cf. section 4.1.2) as well as in other DØ analyses [32], these

deviations are probably due to a mismeasurement of the muon transverse momenta. Further investigations are needed.

Significant discrepancies are also found in the exclusive and inclusive event classes containing one muon and one jet. The event class 1μ 1j has already been discussed in chapter 7, where a 20% deficit in data was observed in the total number of events. The search algorithm selects the region of greatest deviation between 610 GeV and 1030 GeV in $\sum p_T$, which is shown in the second row of figure 8.4. Only 11 data events are measured while 65.8 ± 11.7 are expected. A p_{min}^{data}-value of $6.5 \cdot 10^{-6}$ is calculated for this region with a significance of $\widetilde{P} = 0/200$. For the inclusive event class 1μ 1j + X an excess of 3 data events compared to 0.03 ± 0.03 predicted by the SM is found in the $\sum p_T$ region between 775 GeV and 780 GeV, corresponding to $p_{min}^{data} = 1.7 \cdot 10^{-5}$ and $\widetilde{P} = \frac{0}{100} = 0$. A better understanding of the detector is necessary.

The inclusive event class 2μ + X shows an excess of 38 events in data compared to an expectation of 13.8 ± 2.4 events in the $\sum p_T$ region between 145 GeV and 210 GeV (see top row, right column of figure 8.5). The corresponding p_{min}^{data}-value amounts to $7.1 \cdot 10^{-6}$ with a significance of $\widetilde{P} = 0/500$. A low $\widetilde{P}$ value of 0.02 is also found in the exclusive event class 2μ for the $\sum p_T$ region between 145 GeV and 180 GeV (second row of figure 8.4). In data, 17 events are observed compared to 4.8 ± 1.2 predicted events; the p_{min}^{data}-value amounts to $1.6 \cdot 10^{-4}$. As introduced before, deviations in the same $\sum p_T$ region are found in the analysis of the electron dataset. Candidates may be heavy Z like resonances, decaying into two leptons. In [55], a search for heavy resonances in the dimuon channel with 100 pb^{-1} data, collected by the DØ detector in Run II, has been obtained. The authors of this analysis found no evidence for such a resonance and set a limit of 610 GeV on the mass of a Standard Model like Z'.

Furthermore, low $\widetilde{P}$ values are found in the exclusive and inclusive event classes 1μ $3j$. For both event classes, the regions of greatest deviation are between 375 GeV and 385 GeV in $\sum p_T$. The p_{min}^{data}-values are of the order of 10^{-3} and $\widetilde{P}$ values of 0.05 (for the exclusive event class) and 0.06 (for the inclusive event class) are found. The corresponding $\sum p_T$-distributions are given in figure 8.4 (third row) and in figure 8.5 (second row), respectively.

For the remaining event classes of the muon dataset moderate to high $\widetilde{P}$ values are found, showing no signals of new physics. Exemplarily, the $\sum p_T$-distributions for the event classes 1e 1μ + X and 1μ 2j + X are presented in 8.5. Although low p_{min}^{data}-values are found for these event classes (cf. table 8.4), the test of significance describes the deviations as fluctuations.

Similar to the interpretation of the results found in the electron dataset, further analyses of the deviations in event classes with muon final states are needed to distinguish between detector effects and a real signal of new physics. In these event classes, deviations due to mismeasurements and resolution effects are likely to occur, even more since the muon momentum resolution is limited by the resolution of the sagitta measurement (cf. section 3.2.4). Poorly measured muons tend to fake high p_T muons (cf. section 4.1.2), as it is observed in numerous DØ analyses.

In the analysis of the Run I data [54], no evidence for new physics has been found in muon final states either. It is difficult to point out a deviation to be most likely a signal of new physics. The results should motivate the collaboration to further investigate these event classes.

Results from algorithm for excl. event classes of electron dataset						
Class	$p_{\min}^{\text{data}}$	$\widetilde{P}$	region $[GeV]$	N_{data}	N_{SM}	$\sigma(N_{SM})$
1e	$7.7 \cdot 10^{-7}$	$\frac{0}{1000} = 0$	60-80	1655	1230	80
2e	$6.2 \cdot 10^{-9}$	$\frac{0}{500} = 0$	130-180	42	11.6	2.1
3e	0.29	$\frac{96}{100} = 0.96$	65-240	1	2.8	1.0
1e 1j	$4.2 \cdot 10^{-5}$	$\frac{2}{150} = 0.01$	235-240	7	25.5	1.9
2e 1j	0.018	$\frac{85}{100} = 0.85$	155-160	8	19.0	3.4
1e 2j	$8.7 \cdot 10^{-10}$	$\frac{100}{100} = 1$	200-265	374	257	8
2e 2j	0.011	$\frac{15}{100} = 0.15$	295-470	12	4.9	1.2
1e 3j	$2.2 \cdot 10^{-3}$	$\frac{0}{200} = 0$	255-375	73	49.7	2.6
2e 3j	0.39	$\frac{100}{100} = 1$	380-385	1	0.5	0.3
1e 4j	$2.8 \cdot 10^{-3}$	$\frac{33}{100} = 0.33$	300-315	0	6.6	1.12
1e 5j	0.049	$\frac{60}{100} = 0.60$	340-345	1	0.03	0.05
1e 1μ	0.024	$\frac{68}{100} = 0.68$	90-105	0	3.8	0.4
1e 1μ 1j	0.023	$\frac{34}{100} = 0.34$	95-485	1	5.8	0.5
1e 1μ 2j	$9.5 \cdot 10^{-3}$	$\frac{43}{100} = 0.43$	305-320	2	0.09	0.10
1e 1μ 3j	0.018	$\frac{20}{100} = 0.20$	350-355	1	0.01	0.01

Table 8.1: Regions of greatest deviation and their significances.

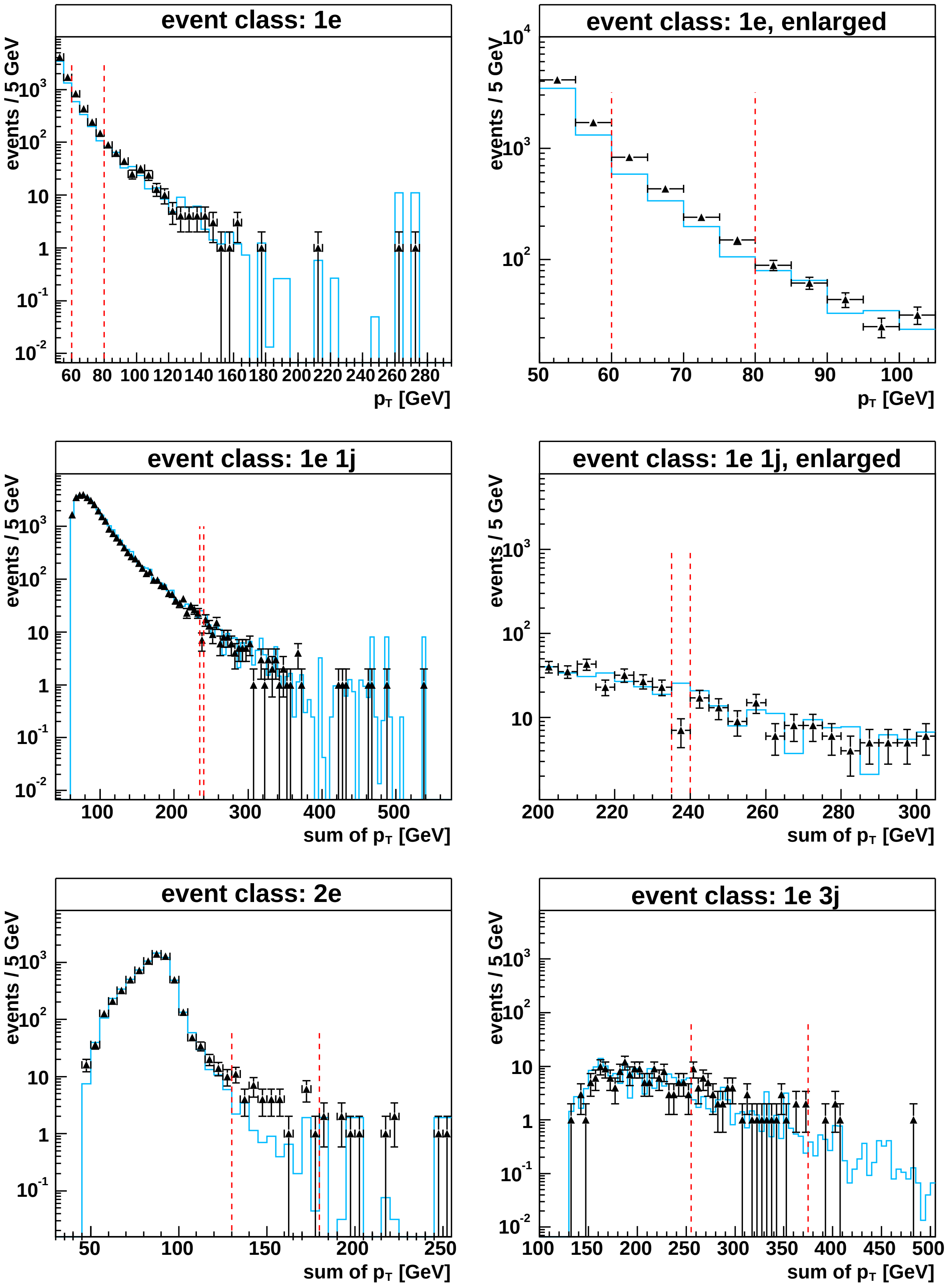


Figure 8.2: Selected event classes of the electron dataset. The black triangles indicate data points, the SM prediction is given by the blue line and the regions of greatest deviation are defined by the red dashed lines.

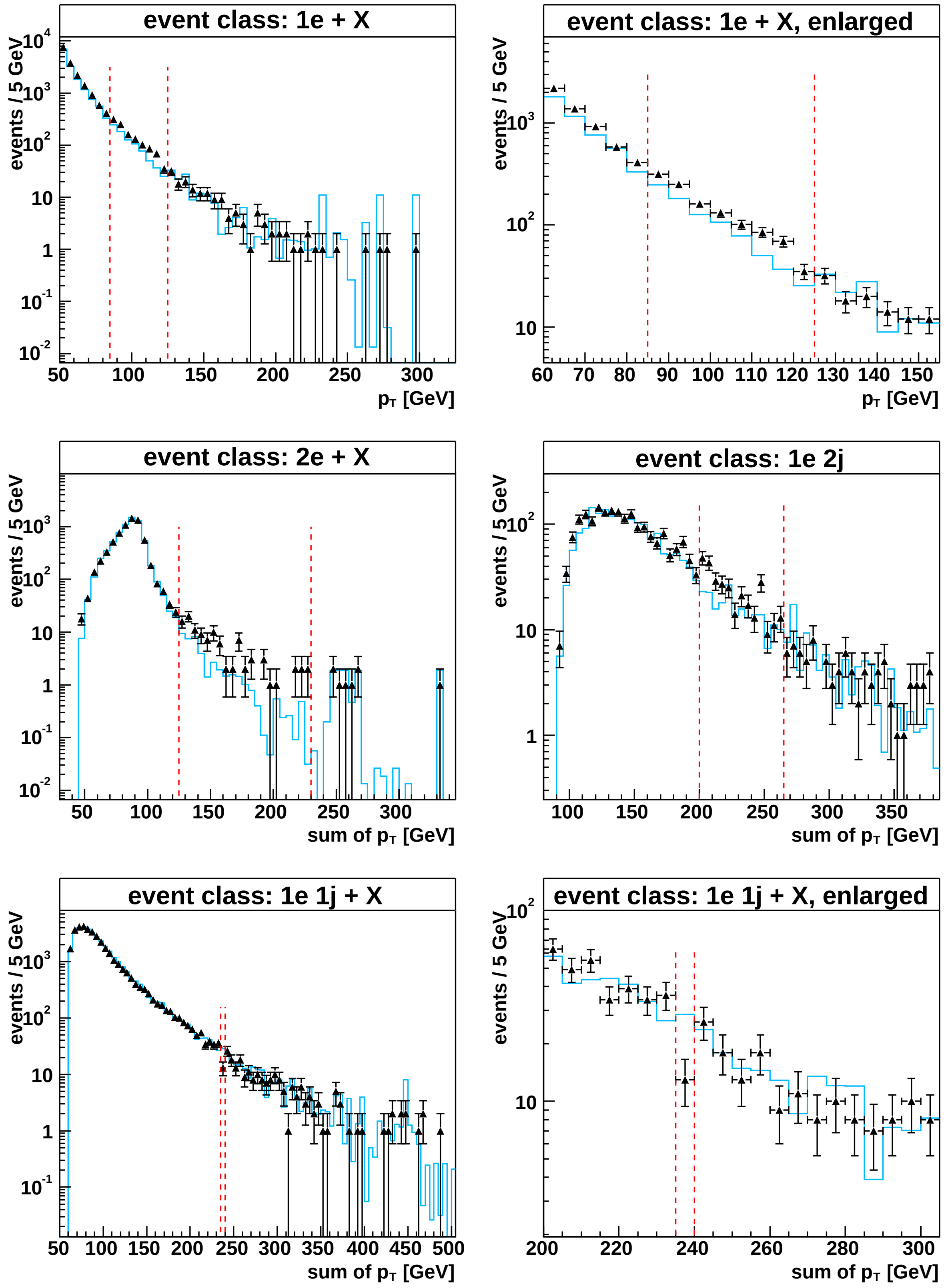


Figure 8.3: Selected event classes of the electron dataset. The black triangles indicate data points, the SM prediction is given by the blue line and the regions of greatest deviation are defined by the red dashed lines.

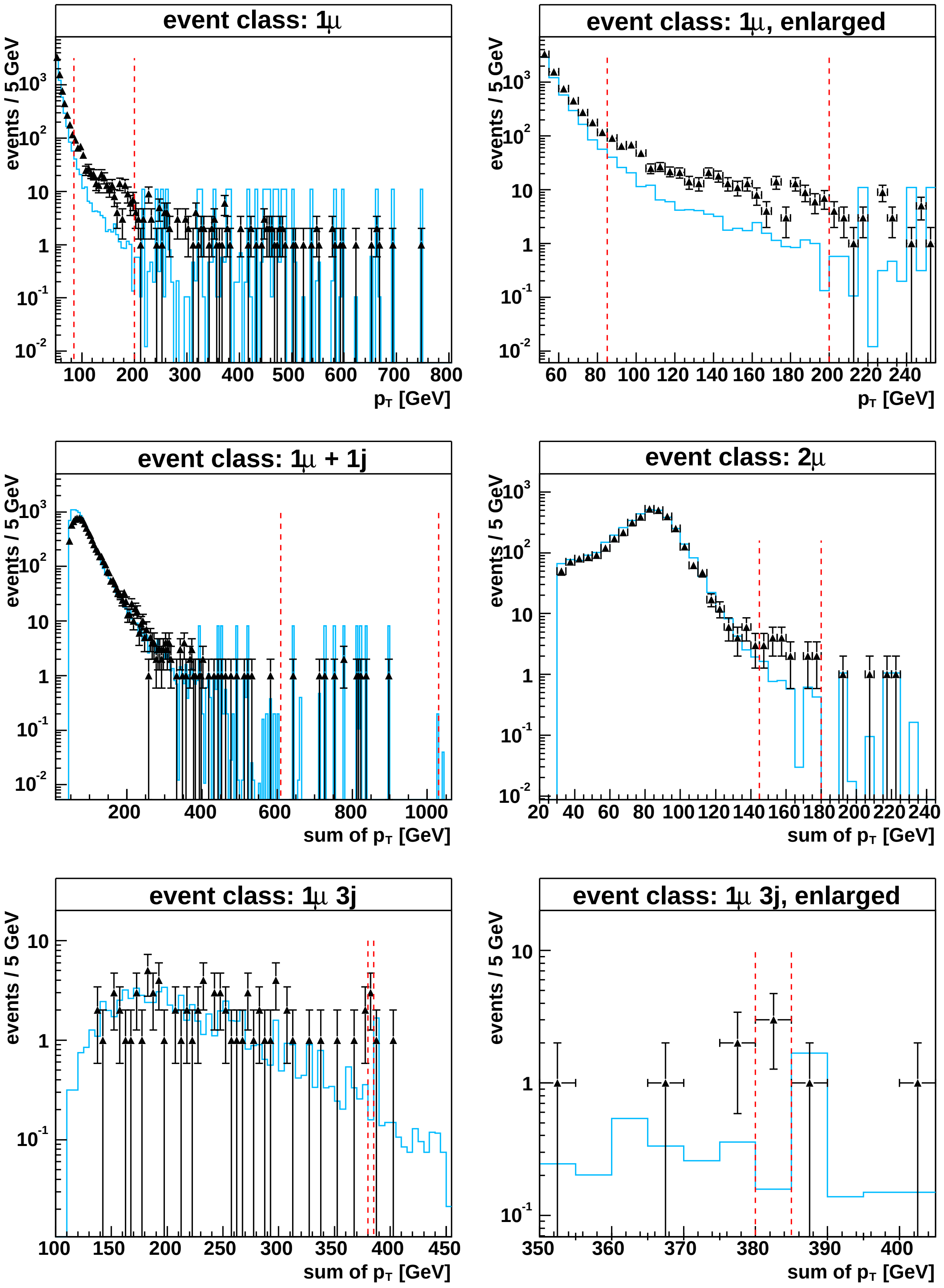


Figure 8.4: Selected event classes of the muon dataset. The black triangles indicate data points, the SM prediction is given by the blue line and the regions of greatest deviation are defined by the red dashed lines.

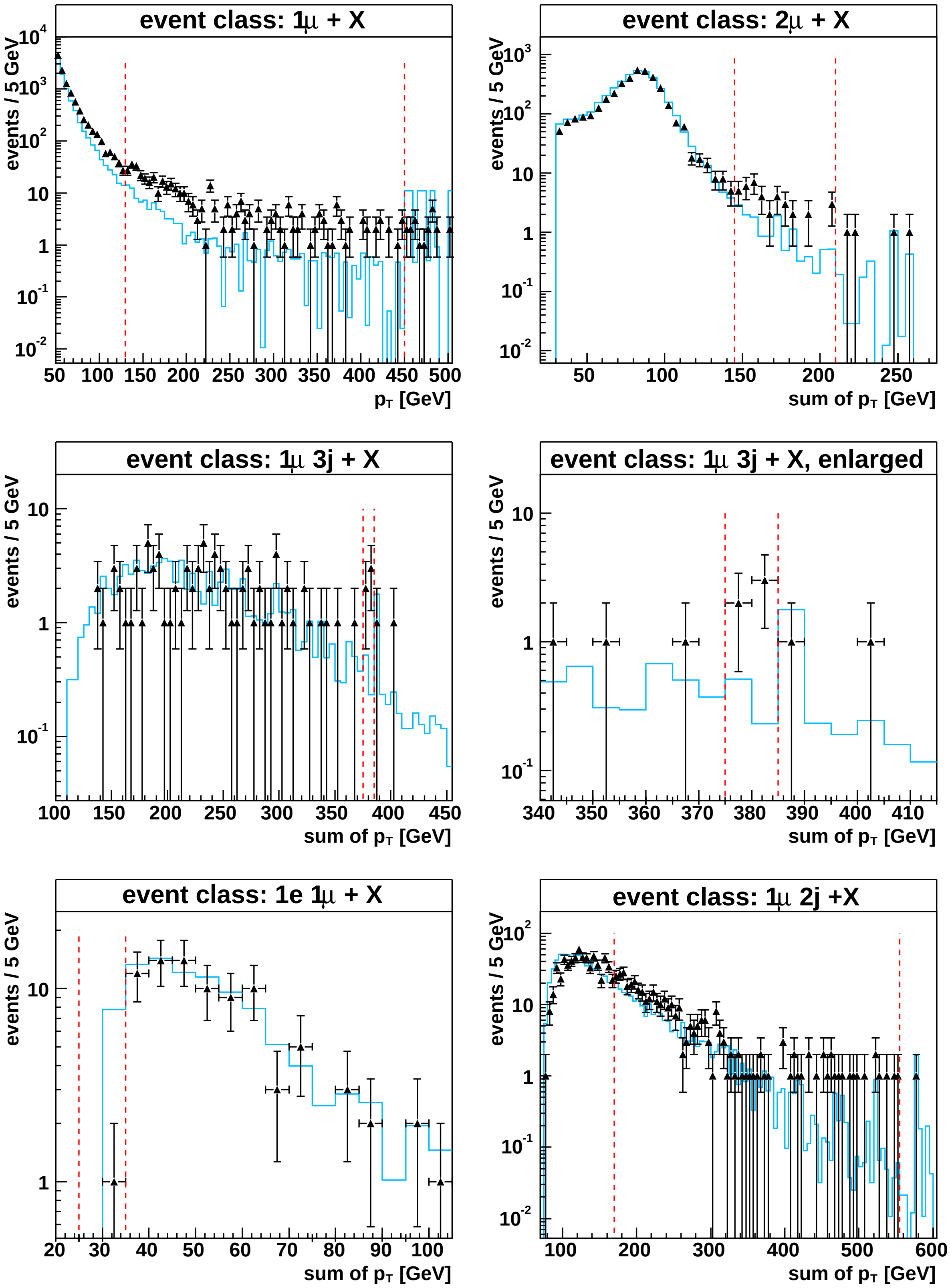


Figure 8.5: Selected event classes of the muon dataset. The black triangles indicate data points, the SM prediction is given by the blue line and the regions of greatest deviation are defined by the red dashed lines.

Results from algorithm for incl. event classes of electron dataset

Class	$p_{\min}^{\text{data}}$	$\widetilde{P}$	region $[GeV]$	N_{data}	N_{SM}	$\sigma(N_{SM})$
1e + X	$4.2 \cdot 10^{-13}$	$\frac{0}{500} = 0$	85-125	1146	850	26
2e + X	$7.2 \cdot 10^{-12}$	$\frac{0}{100} = 0$	125-230	106	43	4
3e + X	0.28	$\frac{96}{100} = 0.96$	10-240	1	2.8	1.0
1e 1j + X	$3.1 \cdot 10^{-3}$	$\frac{26}{100} = 0.26$	235-240	13	28.5	3.0
2e 1j + X	0.048	$\frac{100}{100} = 1$	155-160	11	20.2	3.3
1e 2j + X	0.024	$\frac{100}{100} = 1$	430-450	3	9.9	1.9
2e 2j + X	0.014	$\frac{53}{100} = 0.53$	355-470	6	1.3	0.9
1e 3j + X	0.039	$\frac{43}{100} = 0.43$	215-220	10	4.2	1.7
2e 3j + X	0.39	$\frac{100}{100} = 1$	250-270	1	0.5	0.3
1e 4j + X	$2.8 \cdot 10^{-3}$	$\frac{31}{100} = 0.31$	305-315	0	6.3	0.9
1e 5j + X	0.21	$\frac{53}{100} = 0.53$	340-390	2	0.03	0.9
1e 1μ + X	$2.2 \cdot 10^{-3}$	$\frac{35}{100} = 0.35$	130-150	4	0.5	0.1
1e 1μ 1j + X	0.016	$\frac{65}{100} = 0.65$	210-295	5	1.4	0.3
1e 1μ 2j + X	0.023	$\frac{15}{100} = 0.15$	250-390	5	1.1	0.8
1e 1μ 3j + X	0.018	$\frac{21}{100} = 0.21$	350-355	1	0.01	0.01

Table 8.2: Regions of greatest deviation and their significances.

Results from algorithm for excl. event classes of muon dataset						
Class	$p_{\text{min}}^{\text{data}}$	$\widetilde{P}$	region $[GeV]$	N_{data}	N_{SM}	$\sigma(N_{SM})$
1μ	$1.2 \cdot 10^{-61}$	$\frac{0}{600} = 0$	85-200	536	157	13
2μ	$1.6 \cdot 10^{-4}$	$\frac{11}{500} = 0.02$	145-180	17	4.8	1.2
1e 1μ	$9.9 \cdot 10^{-3}$	$\frac{36}{100} = 0.36$	10-35	1	7.6	1.5
1e 2μ	0.18	$\frac{81}{100} = 0.81$	0-155	1	3.6	1.2
1μ 1j	$6.5 \cdot 10^{-6}$	$\frac{0}{200} = 0$	610-1030	11	65.8	11.7
2μ 1j	0.011	$\frac{74}{100} = 0.74$	95-110	12	26.9	4.5
1μ 2j	$1.5 \cdot 10^{-9}$	$\frac{100}{100} = 1$	175-525	270	167	9
2μ 2j	0.049	$\frac{90}{100} = 0.9$	175-180	3	0.4	0.6
1μ 3j	$3.6 \cdot 10^{-3}$	$\frac{5}{100} = 0.05$	380-385	3	0.2	0.2
2μ 3j	0.062	$\frac{29}{100} = 0.29$	320-410	2	0.3	0.2
1μ 4j	0.045	$\frac{17}{100} = 0.17$	305-360	0	3.3	0.6
1μ 5j	0.073	$\frac{20}{100} = 0.2$	315-325	1	0.06	0.06
1e 1μ 1j	0.069	$\frac{91}{100} = 0.91$	75-570	4	9.1	1.3
1e 1μ 2j	0.020	$\frac{41}{100} = 0.41$	315-340	2	0.14	0.14
1e 1μ 3j	0.016	$\frac{14}{100} = 0.14$	350-355	1	0.01	0.01

Table 8.3: Regions of greatest deviation and their significances.

Results from algorithm for incl. event classes of muon dataset						
Class	$p_{\min}^{\text{data}}$	$\widetilde{P}$	region $[GeV]$	N_{data}	N_{SM}	$\sigma(N_{SM})$
1μ + X	$6 \cdot 10^{-66}$	$\frac{0}{500} = 0$	130-450	401	111	7
2μ + X	$7.1 \cdot 10^{-6}$	$\frac{0}{500} = 0$	145-210	38	13.8	2.4
1e 1μ + X	$8.6 \cdot 10^{-3}$	$\frac{52}{100} = 0.52$	25-35	1	7.8	1.5
1e 2μ + X	0.14	$\frac{73}{100} = 0.73$	40-155	1	3.9	1.2
1μ 1j + X	$1.7 \cdot 10^{-5}$	$\frac{0}{100} = 0$	775-780	3	0.03	0.03
2μ 1j + X	$7.9 \cdot 10^{-3}$	$\frac{71}{100} = 0.71$	95-110	12	27.5	4.5
1μ 2j + X	$7.9 \cdot 10^{-11}$	$\frac{100}{100} = 1$	170-555	369	234	12
2μ 2j + X	0.037	$\frac{89}{100} = 0.89$	375-380	1	0.03	0.03
1μ 3j + X	$5.9 \cdot 10^{-3}$	$\frac{6}{100} = 0.06$	375-385	5	0.74	0.5
2μ 3j + X	0.062	$\frac{34}{100} = 0.34$	320-410	2	0.3	0.2
1μ 4j + X	0.036	$\frac{15}{100} = 0.15$	250-285	6	2.3	0.4
1μ 5j + X	0.067	$\frac{15}{100} = 0.15$	320-330	1	0.06	0.05
1e 1μ 1j + X	0.016	$\frac{47}{100} = 0.47$	275-295	2	0.18	0.08
1e 1μ 2j + X	0.004	$\frac{10}{100} = 0.10$	305-340	3	0.2	0.2
1e 1μ 3j + X	0.016	$\frac{14}{100} = 0.14$	340-355	1	0.01	0.01

Table 8.4: Regions of greatest deviation and their significances.

Chapter 9

Summary

In this work, a model-independent search for signals of new physics in proton-antiproton collisions is presented. About 230 pb^{-1} data, collected with the DØ detector at the Tevatron between the years 2002 and 2004 are systematically investigated for deviations from the Standard Model (SM) prediction. This coherent analysis of all event topologies involving isolated electrons and muons, as well as additional jets is performed for the first time in DØ Run II. A similar approach was done in Run I [54], where no signs of new physics have been found.

The identification criteria for each physics object are based on certified recommendations and previous analyses performed on specific final states. For muons, additional requirements have been developed to ensure an unambiguous identification. All objects are measured in the central region of the detector. The transverse momenta p_T of electrons and jets are required to be above 30 GeV and for muons, to be above 15 GeV. According to these definitions, all events are divided into exclusive and inclusive event classes, depending on the number and type of the objects found in the final state.

The SM prediction is given by Monte Carlo simulations, which are created with the PYTHIA event generator. A global scaling method is developed to account for different efficiencies observed in data and Monte Carlo samples. For the multijet background, techniques are developed to estimate this contribution from the data. Both methods give satisfying results.

For all event classes, the total number of events and the distributions of the scalar sum of transverse momenta $\sum p_T$ are investigated. A good agreement between the data and the SM prediction is observed in most event classes.

Deviations are found in event classes containing a lepton and jets. A reason for the discrepancy in the event class 1e 3j may be that higher jet multiplicities are not well simulated by the event generator PYTHIA. Although the QCD background is expected to be the dominant contribution in the event class 1μ 1j, a deficit in the data is observed, even when considering only contributions from Monte Carlo samples. Further studies, for example of the muon isolation criterion, are needed.

In order to systematically search for signals of new physics, an algorithm is developed. The algorithm defines all possible regions in a one-dimensional histogram and locates the region of largest deviation between the data and SM prediction. In a test of significance, the probability of occurance of such a deviation is calculated. Therefore, hypothetical data histograms (HDH) are generated and searched for a similar or larger deviation. In this analysis, the $\sum p_T$-distributions of all event classes are investigated.

The results show that the algorithm is able to find deviations and that the quantification of the significances of these deviations are reliable.

The most significant deviation is found in the inclusive 1μ event class. In the p_T region between 130 GeV and 450 GeV, 401 events are found in data compared to a SM prediction of only 111 ± 7. In 500 generated HDH no similar or larger deviation is observed.

Furthermore, the dilepton event classes are of high interest since in both 2e- and 2μ-classes, large deviations are found in similar $\sum p_T$ regions above the Z boson peak. For example, in the exclusive 2e event class, 42 data events are measured in the $\sum p_T$ region between 130 GeV and 180 GeV, while 11.6 ± 2.1 are predicted by the SM.

In order to distinguish between detector effects, an incorrect description of the SM Monte Carlo simulations and real signals of new physics, further studies are necessary. The same algorithm is used in another DØ analysis [4] to search for deviations in final states with leptons and missing transverse energy. Results are expected with eagerness in the very near future.

In addition, up to $4 - 9\ fb^{-1}$ data are expected during Run II, which will help to clarify whether the deviations found and discussed in this work, are signs of new physics or statistical fluctuations.

The results of this analysis show, that the developed search method would also be ideally suited to effectively reveal interesting event topologies in the data of forthcoming experiments - for example at the Large Hadron Collider LHC.

Bibliography

[1] The LEP collaborations: Combination procedure for the precise determination of Z boson parameters from results of the LEP experiments
hep-ex/0101027 (2001)

[2] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration and The LEP Working Group for Higgs Boson Searches:
Search for the Standard Model Higgs boson at LEP
Phys. Lett. B 565 (2003) 61

[3] The DØ Collaboration: A precision measurement of the mass of the top quark
Nature, Vol. 429, 10 June 2004

[4] Philipp Biallass, diploma thesis: Model Independent Search for Deviations from the Standard Model in Final States with Missing Transverse Energy at the Tevatron
RWTH Aachen University, III. Physics Institute

[5] Public Access to ZEUS 2002 PDFs
http://www-spires.dur.ac.uk/hepdata/zeus2002.html

[6] The K2K Long-baseline Neutrino Oscillation Experiment
http://neutrino.kek.jp/

[7] Steven Weinberg: Implications of dynamical symmetry breaking
Phys. Rev. D13:974-996, 1976
Leonard Susskind: Dynamics of spontaneous symmetry breaking in the Weinberg-Salam theory
Phys. Rev., D20:2619, 1979

[8] Lecture "Angewandte Physik: Teilchenphysik mit höchstenergetischen Beschleunigern"
http://www.mppmu.mpg.de/english/Lectures_ws0304.html

[9] Stephen P. Martin: A Supersymmetry Primer
hep-ph/9709356 (1999)

[10] Martin Wessel: General Search for New Phenomena in ep Scattering at HERA (2004)
http://www-h1.desy.de/psfiles/theses/h1th-355.ps

[11] The DØ Collaboration, B. Abbott et al.:
Search for New Physics in e mu X Data at D0 Using Sleuth: A Quasi-Model-Independent Search Strategy for New Physics
hep-ex/0006011, Phys. ReV. D 62, 92004 (2000)

[12] RunII Handbook, http://www-bd.fnal.gov/runII/index.html

[13] The DØ Collaboration, S. Abachi et al.:
Nucl. Instr. and Methods, A338, 185 (1994)
FERMILAB-PUB-93/179-E

[14] Fermilab Visual Media Services Photo Database
http://vmsfmp2.fnal.gov/v1/VMSChoice.html

[15] The DØ Collaboration, S.Abachi et al.:
"The DØ Upgrade, The Detector and Its Physics", 1996

[16] DØ Presentations and Picture Gallery
http://www-d0.fnal.gov/Run2Physics/displays/presentations/

[17] DØ Silicon Tracker Technical Design Report
http://d0server1.fnal.gov/projects/silicon/www/tdr_final.ps

[18] Brzezniak et al.: Conceputanl Design of a 2T Superconducting Solenoid for the Fermilab DØ Detector Upgrade

DØ Note 2167 (1994)

[19] Calorimeter electronics Upgrade for Run II TDR, May 1998
http://www-d0.fnal.gov/~d0upgrad/calelec/intro/tdr/tdr17.pdf

[20] DØ Electron and Photon Identification Group: Certification Results from December 2002 (version 2.2.1)
http://www-d0.fnal.gov/phys_id/emid/d0_private/certification/main_v2_2_1.html

[21] Drawing of the DØ Muon System
http://www-d0.fnal.gov/~hagopian/muon_run2_detector.ps

[22] Global Trigger List Descriptions
http://www-d0.fnal.gov/trigger_meister/private/www/tl_desc/global.html

[23] DØ Run II Level 2 Trigger, General Information
http://www.pa.msu.edu/hep/d0/ftp/l2/overview/figures/configl1l2.eps

[24] Run II Stores and Physics Runs
http://www-d0.fnal.gov/runcoor/d0_private/operations/stores_runs.html

[25] Klimenko et al., Fermilab FN-741 (2003)

[26] SAM Data Handling at DØ
http://www-d0.fnal.gov/computing/sam/

[27] Edwards et al.: The Updated DØ Luminosity Determination
DØ Note 4328 (2004)

[28] Offline Run Quality Database Queries
http://d0db.fnal.gov/qualitygrabber/qualQueries.html

[29] The Electron and Photon Identification Group
http://www-d0.fnal.gov/phys_id/emid/d0_private/emid.html

[30] The Muon Identification Group
http://www-d0.fnal.gov/computing/algorithms/muon/muon_algo.html

[31] The Jet and Missing E_T Identification Group
http://www-d0.fnal.gov/ d0upgrad/d0_private/software/jetid/jetid.html

[32] Personal communication with other analysts in the DØ Muon ID mailing list
http://listserv.fnal.gov/archives/d0_muon_id.html

[33] E. Busato, B. Andrieu:
Jet Algorithms in the DØ Run II Software: Description and User's Guide
DØ Note 4457 (2004)

[34] Certified Jet Energy Scale
http://www-d0.fnal.gov/phys_id/jes/d0_private/certified/certified.html

[35] The Common Sample Group
http://www-d0.fnal.gov/Run2Physics/cs/index.html

[36] ROOT, An Object-Oriented Data Analysis Framework
http://root.cern.ch/

[37] Markus Klute, Lukas Phaf and Daniel Whiteson: TopAnalyze - A Framework Analyze Package For Top Group Analyses
DØ Note 4122 (2003)

[38] Inclusive single MU triggers for W/Z analyses:
http://www-d0.fnal.gov/Run2Physics/wz/d0-private/triggers/MUtriggers.html
Inclusive single EM triggers for W/Z analyses:
http://www-d0.fnal.gov/Run2Physics/wz/d0-private/triggers/EMtriggers.html

[39] D. Chapin et al.: Measurement of $Z \to ee$ and $W \to e\nu$ Production Cross Sections with $|eta| < 2.3$
DØ Note Number: 4403 (2004)

[40] Heidi Schellman, Marco Verzocchi
http://www-d0.fnal.gov/phys_id/luminosity/data_access/lm_access/doc/runrange_luminosity.html

[41] T. Sjöstrand, L. Lönnblad, S. Mrenna and P. Skands: PYTHIA 6.206
hep-ph/0108264 (2001)

[42] private communication with Ulla Blumenschein, Ph.D. student, University of Freiburg

[43] The DØ New Phenomena Group
http://www-d0.fnal.gov/Run2Physics/np/

[44] H. Plothow-Besch: PDFLIB,
http://www-d0.fnal.gov/computing/MonteCarlo/generator_tools/pdflib_v8.04_doc.ps

[45] The DØ Top Quark Group
http://www-d0.fnal.gov/Run2Physics/top/index.html

[46] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. Polosa:
ALPGEN, a generator for hard multiparton processes in hadronic collisions
hep-ph/0206293 (2002)

[47] Thomas Nunnemann: Cross-Sections for Drell-Yan, Z and W Production using Modern Parton Distribution Functions
DØ Note 4476 (2004)

[48] Alton, Askew, Maravin, Mattingly: Measurement of Wγ events in DØ Run II Data
DØ Note 4410 (2004)

[49] S. Eidelman, et al.: Review of Particle Physics
Physics Letters B 592, 1 (2004)

[50] S. Jain: A study of the scale and over-smearing parameters for MC electrons.
DØ Note 4402

[51] Campbell, Ellis: An update on vector boson pair production at hadron colliders
hep-ph/9905386 (1999)

[52] Campbell, Huston: PDF Error Discussion: the sequel to the sequel
http://www-d0.fnal.gov/computing/MonteCarlo/generator_tools/pdf_error4.pdf

[53] S. Anderson et al.: Measurement of the $t\bar{t}$ Production Cross-section at $\sqrt{s} = 1.96\,TeV$ in Dilepton Final States
DØ Note 4623 (2004)

[54] The DØ Collaboration, B. Abbott et.al:
A Quasi-Model-Independent Search for New High p_T Physis at DØ
hep-ex/0011071 (2000)

[55] R. Hooper, G. Landsberg: Search for Heavy Resonances in the Di-Muon Channel With 100 pb^{-1} of Run II Data
DØ Note 4229 (2003)

Chapter 10

Acknowledgement

First of all, I would like to express my gratitude to my parents who made it possible for me to study and who always supported me.

I would like to thank Prof. Dr. T. Hebbeker for giving me the possibility of writing this thesis. I am grateful for the opportunity of visiting the DØ experiment at Fermilab as well as for his patience and fruitful ideas, especially in discussions about the QCD background and statistics. I would like to thank Prof. Dr. M. Erdmann for agreeing to be my second referee.

I am grateful for the advice and support of my colleagues in the local DØ group: Christian Autermann, Jan Coenen, Dr. Steffen Kappler, Matthias Kirsch, Carsten Magaß, Dr. Arnd Meyer and Lotte Wilke. I am especially grateful to Daniela Käfer, who always found time and patience to help me, to give advice and for proof-reading every single chapter. This thesis would never have been accomplished without the great and fruitful atmosphere between Philipp Biallass and me. I would like to thank him for the harmonic collaboration during the year.
I would also like to thank the former members of the local DØ group, Thomas Stapelberg and Dr. Martin Wegner, who gave me good start into this work.

I owe special thanks to my brother Torsten for his support and to my girlfriend Simone for her loving care.